

Low Discrepancy Sequences in High Dimensions: How Well Are Their Projections Distributed?

Xiaoqun Wang^{1,2} and Ian H. Sloan²

¹*Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China*

²*School of Mathematics, University of New South Wales, Sydney 2052, Australia*

Abstract Quasi-Monte Carlo (QMC) methods have been successfully used to compute high-dimensional integrals arising in many applications, especially in finance. To understand the success and the potential limitation of QMC, this paper focuses on quality measures of point sets in high dimensions. We introduce the order- ℓ , superposition and truncation discrepancies, which measure the quality of selected projections of a point set on lower-dimensional spaces. These measures are more informative than the classical ones. We study their relationships with the integration errors and study the tractability issues. We present efficient algorithms to compute these discrepancies and perform computational investigations to compare the performance of the Sobol' nets with that of the sets of Latin hypercube sampling and random points. Numerical results show that in high dimensions the superiority of the Sobol' nets mainly derives from the one-dimensional projections and the projections associated with the earlier dimensions; for order-2 and higher-order projections all these point sets have similar behavior (on the average). In weighted cases with fast decaying weights, the Sobol' nets have a better performance than the other two point sets. The investigation enables us to better understand the properties of QMC and throws new light on when and why QMC can have a better (or no better) performance than Monte Carlo for multivariate integration in high dimensions.

Keywords: quasi-Monte Carlo methods, low discrepancy sequences, discrepancy, multivariate integration.

2000 Mathematics Subject Classification: 11K38, 65C05, 65D30.

Email addresses: xwang@math.tsinghua.edu.cn, i.sloan@unsw.edu.au.

1 Introduction

Many practical problems can be transformed into the computation of multivariate integrals

$$I_s(f) = \int_{[0,1]^s} f(\mathbf{x}) d\mathbf{x}, \quad \mathbf{x} = (x_1, \dots, x_s).$$

The dimension s can be very large. The classical methods based on product rules are not suitable for large s because of the curse of dimensionality. Monte Carlo (MC) or quasi-Monte Carlo (QMC) estimates of $I_s(f)$ take the form

$$Q_{n,s}(f) = \frac{1}{n} \sum_{i=1}^n f(\mathbf{x}_i), \quad \mathbf{x}_i \in [0, 1]^s.$$

MC methods are based on random samples and have a convergence order of $O(n^{-1/2})$ for square integrable functions, independently of s , while QMC methods use deterministic points having better uniformity properties.

Our purpose in this paper is to study the low-dimensional projections of low discrepancy sequences in high dimensions, aiming at answering the question of when and why QMC can have a better (or no better) performance than MC. We introduce new quality measures of a point set and explain their importance for understanding the success of QMC in some applications. Note that some projection properties of low discrepancy sequences were studied in [7, 15, 23].

A classical QMC error bound is the Koksma-Hlawka inequality (see [16]):

$$|I_s(f) - Q_{n,s}(f)| \leq D_{\infty,*}(P) V_{\text{HK}}(f),$$

where $D_{\infty,*}(P)$ is the L_∞ -star discrepancy and $V_{\text{HK}}(f)$ is the variation in the sense of Hardy and Krause. Traditionally, a sequence is called a *low discrepancy sequence* if the L_∞ -star discrepancy of the first n points satisfies

$$D_{\infty,*}(P) \leq c(s) \frac{(\log n)^s}{n}.$$

There are several known low discrepancy sequences, such as Halton [9], Sobol' [28], Faure [8], Niederreiter [16] and Niederreiter-Xing [17]. A QMC algorithm based on a low discrepancy sequence has a deterministic error bound $O(n^{-1}(\log n)^s)$ for functions of bounded variation, which is asymptotically better than that of MC. However, the QMC error bound increases with n until n equals approximately e^s . For large s , the Koksma-Hlawka bound does not imply any advantage of QMC for practical values of n .

Empirical investigations have shown that QMC is significantly more efficient than MC for high-dimensional integrals in some applications, especially in finance (see [1, 3, 18, 22]). The question of how to explain the success of QMC has been extensively investigated. Caflisch *et al* [3], Owen [20] and Wang and Fang [31] studied the *effective dimensions*. Sloan and Woźniakowski [27] introduced *weighted function spaces* and showed the existence of QMC algorithms for which the curse of dimensionality is not present under some conditions on the weights. Good shifted lattice rules which achieve strong tractability error

bounds were constructed in Sloan *et al* [25]. Hickernell and Wang [12] and Wang [30] showed that Halton, Sobol' and Niederreiter sequences achieve error bounds with the optimal convergence order $O(n^{-1+\delta})$ (for arbitrary $\delta > 0$) independently of the dimension under appropriate conditions. Sloan *et al* [26] showed that the convergence rate of QMC becomes essentially independent of the dimension for functions for which the “interactions” of many variables together are small enough.

The efficiency of an algorithm depends on both the point set and the integrand. The success of QMC for some high-dimensional problems indicates that there must be some inherent properties of low discrepancy point sets and/or there must be some special features in the integrands. About the integrands, it is known that many high-dimensional problems in finance are of low effective dimension in the superposition sense, i.e., the functions can be well approximated by sums of low-dimensional functions (but the effective dimension in the truncation sense can be large), see [3, 32]. Moreover, if dimension reduction techniques are used, such as the Brownian bridge [14] and the principal component analysis [1], the integrands are likely to be of low effective dimension in the truncation sense [31]. For functions of small effective dimension (in the superposition or truncation sense), the integration errors are determined by the discrepancy of the low-order projections or the projections of the initial coordinates. Note that most results on effective dimensions are empirical and most discussions on the relation of the effective dimension with the QMC errors are of qualitative nature.

This paper focuses on the quality of QMC point sets in high dimensions. In deciding which of two point sets is the better, the usual preference is for the one with smaller discrepancy (such as the classical L_∞ -star or L_2 -star discrepancy, see [16]). Computational investigations of the classical discrepancy confirmed the superiority of low discrepancy point sets over random point sets only in low dimensions (say $s \leq 12$), but failed to demonstrate any superiority of QMC in high dimensions (see [15]). Indeed, in high dimensions the classical discrepancies of low discrepancy point sets are found to behave initially like $O(n^{-1/2})$, the same as for random points, while an order close to $O(n^{-1})$ can only be observed for huge n , with the transition value of n appearing to grow exponentially with the dimension s . Thus the classical discrepancies are not good enough for measuring the quality of point sets in high dimensions. Another popular empirical method to study the uniformity of QMC point sets is to plot their two-dimensional projections [15]. However, there are too many projections for a point set in high dimensions, and one can almost always find some good and some bad projections. This indicates the need of new methodology to study the uniformity of point sets.

This paper is organized as follows. In Section 2, after an overview of the classical discrepancies, we introduce new quality measures: the order- ℓ , superposition and truncation discrepancies. Their generalizations and tractability issues are studied in Section 3. Efficient algorithms for computing these discrepancies are presented in Section 4, whereas computational investigations are performed in Section 5. The investigations throw new light on the efficiency and potential limitations of QMC in high dimensions, and indicate when and why QMC may have a better or no better performance than MC and indicate ways of how to enhance the superiority of QMC.

2 Superposition and truncation discrepancies

2.1 The classical discrepancies

We start with the classical star discrepancy. Let $P := \{\mathbf{x}_i = (x_{i,1}, \dots, x_{i,s}) : i = 1, \dots, n\}$ be a set of n points in $[0, 1]^s$. For a vector $\mathbf{y} = (y_1, \dots, y_s) \in [0, 1]^s$, let $[\mathbf{0}, \mathbf{y}] = [0, y_1] \times \dots \times [0, y_s]$. Define the *local discrepancy* as

$$\text{disc}(\mathbf{y}) := y_1 \cdots y_s - \frac{1}{n} \sum_{i=1}^n \mathcal{X}_{[\mathbf{0}, \mathbf{y}]}(\mathbf{x}_i), \quad (1)$$

where $\mathcal{X}_{[\mathbf{0}, \mathbf{y}]}$ is the characteristic function of $[\mathbf{0}, \mathbf{y}]$. The classical L_∞ -star discrepancy is the L_∞ -norm of the local discrepancy function:

$$D_{\infty,*}(P) := \sup_{\mathbf{y} \in [0,1]^s} |\text{disc}(\mathbf{y})|.$$

The *classical L_2 -star discrepancy* is the L_2 -norm of the local discrepancy function

$$D_*(P) := \left(\int_{[0,1]^s} \text{disc}^2(\mathbf{y}) d\mathbf{y} \right)^{1/2}. \quad (2)$$

An important property of the classical L_2 -star discrepancy is its relation to the average-case error of the algorithms $Q_{n,s}(f)$, as shown in [34]:

$$D_*^2(\{\mathbf{1} - \mathbf{x}_i\}) = \int_F [I_s(f) - Q_{n,s}(f)]^2 \mu(df),$$

where $F = C([0, 1]^s)$ is the class of continuous functions defined over $[0, 1]^s$, and μ is the classical Wiener sheet measure.

For a point set $P := \{\mathbf{x}_i = (x_{i,1}, \dots, x_{i,s}) : i = 1, \dots, n\}$ in $[0, 1]^s$, an explicit formula is available for $D_*(P)$ (see [33]):

$$D_*^2(P) = \left(\frac{1}{3}\right)^s - \frac{2}{n} \sum_{i=1}^n \prod_{j=1}^s \left(\frac{1}{2} - \frac{x_{i,j}^2}{2}\right) + \frac{1}{n^2} \sum_{i=1}^n \sum_{k=1}^n \prod_{j=1}^s [1 - \max(x_{i,j}, x_{k,j})]. \quad (3)$$

Many computational investigations were based on this. A faster algorithm was presented in [10]. The classical L_2 -star discrepancy fails to properly discriminate among different (low discrepancy or random) point sets in high dimensions. One important reason is that it ignores the lower-order projections.

In order to define new notions of discrepancies, we introduce some notation. Let $\mathcal{A} = \{1, \dots, s\}$ and let $1 : \ell = \{1, \dots, \ell\}$. For any subset $\mathbf{u} \subseteq \mathcal{A}$, let $|\mathbf{u}|$ denotes its cardinality (or order). For $\mathbf{x} \in [0, 1]^s$, let $\mathbf{x}_{\mathbf{u}}$ be the $|\mathbf{u}|$ -dimensional vector containing x_j for $j \in \mathbf{u}$. By $(\mathbf{x}_{\mathbf{u}}, \mathbf{1})$ we mean the s -dimensional vector with the same components as \mathbf{x} for indices in \mathbf{u} and the rest of the components replaced by 1.

For any nonempty subset \mathbf{u} of \mathcal{A} , let $P_{\mathbf{u}} = \{(\mathbf{x}_i)_{\mathbf{u}} : i = 1, \dots, n\}$. According to (2), the square of the classical L_2 -star discrepancy of $P_{\mathbf{u}}$ is

$$D_*^2(P_{\mathbf{u}}) := \int_{[0,1]^{|\mathbf{u}|}} \text{disc}^2(\mathbf{x}_{\mathbf{u}}) d\mathbf{x}_{\mathbf{u}}, \quad (4)$$

where $\text{disc}(\mathbf{x}_{\mathbf{u}})$ is defined as in (1). Suppose now that we are given a sequence of non-negative weights $\{\gamma_{\mathbf{u}} : \mathbf{u} \subseteq \mathcal{A}\}$ with $\gamma_{\emptyset} = 1$. The squared (fully) *weighted L_2 -star discrepancy* is defined as (see [11, 27])

$$D^2(P; K_s^*) := \sum_{\emptyset \neq \mathbf{u} \subseteq \mathcal{A}} \gamma_{\mathbf{u}} D_*^2(P_{\mathbf{u}}). \quad (5)$$

(The reason for the label K_s^* will be clear soon.)

The purpose of introducing weights is to allow a greater or smaller emphasis to various terms in (5). Two kinds of weights will be of particular interest. One is the *product weights*

$$\gamma_{\emptyset} = 1, \quad \gamma_{\mathbf{u}} := \prod_{j \in \mathbf{u}} \gamma_j \quad \text{for } \emptyset \neq \mathbf{u} \subseteq \mathcal{A}; \quad (6)$$

the other is the *order-dependent weights*

$$\gamma_{\emptyset} = 1, \quad \gamma_{\mathbf{u}} := \Gamma_{(\ell)} \quad \text{if } |\mathbf{u}| = \ell, \quad (7)$$

where $\gamma_1, \dots, \gamma_s$ and $\Gamma_{(1)}, \dots, \Gamma_{(s)}$ are non-negative numbers. The product weights (6), introduced in [27], are the simplest to handle. Their drawback is that they are not sufficiently flexible: the “interaction” among x_j with $j \in \mathbf{u}$ is automatically determined by the product $\prod_{j \in \mathbf{u}} \gamma_j$. In the order-dependent case, introduced in [6], the weights depend on \mathbf{u} only through the order $|\mathbf{u}|$. This choice is more appropriate if all variables are (nearly) equally important. The order-dependent weights are of the product form iff $\Gamma_{(\ell)} = \beta^{\ell}$ for some β .

For the product weights (6), an explicit formula is available for the weighted L_2 -star discrepancy [27]. For the order-dependent weights (7), a formula will be given in Section 4.

The weighted L_2 -star discrepancy involves *all* projections. The relative importance of different projections depends strongly on the weights. It would seem better to separate the projections and to examine the uniformity of each projection separately. However, there are $2^s - 1$ projections, a huge number for large s . Such considerations lead us to define new discrepancies, which measure the uniformity over selected aggregations of projections. We introduce new discrepancies based on the geometric interpretation and then generalize them via the reproducing kernel approach.

2.2 The order- ℓ , superposition and truncation discrepancies

The quality of the projections of low discrepancy point sets varies significantly with the order $|\mathbf{u}|$ (this is made mathematically precise in Section 3). It is desirable to distinguish different orders of projections.

Definition 1. For $\ell = 1, \dots, s$ the order- ℓ (weighted L_2 -star) discrepancy of the point set P , denoted by $\mathcal{D}_{(\ell)}(P)$, is defined by

$$\mathcal{D}_{(\ell)}^2(P) := \sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}| = \ell} \gamma_{\mathbf{u}} D_*^2(P_{\mathbf{u}}), \quad (8)$$

where $D_*(P_{\mathbf{u}})$ is the classical L_2 -star discrepancy of $P_{\mathbf{u}}$.

The order- ℓ discrepancy $\mathcal{D}_{(\ell)}(P)$ is a measure of the uniformity of the order- ℓ projections all taken together. The classical L_2 -star discrepancy in (2) is actually the order- s discrepancy with $\gamma_{\mathcal{A}} = 1$, which only measures the uniformity of the *highest* order projection (this exposes a deficiency of the classical L_2 -star discrepancy). Moreover, from (5) and (8) it follows that $D^2(P; K_s^*) = \sum_{\ell=1}^s \mathcal{D}_{(\ell)}^2(P)$.

For many low discrepancy point sets in dimension s , each projection of order- ℓ is an ℓ -dimensional low discrepancy point set (see Theorem 3 below). That result shows that for two important low discrepancy point sets and for a fixed small ℓ , the order- ℓ discrepancy has asymptotically better behavior than a random point set. However, this asymptotically better behavior may not appear for practical values of n . Indeed, as mentioned in [15], even order-2 projections of common low discrepancy point sets may have bad distribution properties. Later we will show in a numerical experiment that the Sobol' net may have larger order-2 discrepancy than a random point set does, if s is large and n is relatively small. (In the experiment $s = 64$ and $n \leq 2^{12}$.)

Definition 2. For $\ell = 1, \dots, s$ the superposition (weighted L_2 -star) discrepancies of P , denoted by $\mathcal{SD}_{\ell}(P)$, are defined by

$$\mathcal{SD}_{\ell}^2(P) = \sum_{k=1}^{\ell} \mathcal{D}_{(k)}^2(P) = \sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}| \leq \ell} \gamma_{\mathbf{u}} D_*^2(P_{\mathbf{u}}).$$

The superposition discrepancies $\mathcal{SD}_{\ell}(P)$ is an aggregate measure of the uniformity of all projections with order $|\mathbf{u}|$ from 1 up to ℓ . Clearly, we have $\mathcal{SD}_s^2(P) = D(P; K_s^*)$.

Definition 3. For $\ell = 1, \dots, s$ the truncation (weighted L_2 -star) discrepancies of P , denoted by $\mathcal{TD}_{1:\ell}(P)$, are defined by

$$\mathcal{TD}_{1:\ell}^2(P) := \sum_{\emptyset \neq \mathbf{u} \subseteq 1:\ell} \gamma_{\mathbf{u}} D_*^2(P_{\mathbf{u}}).$$

The truncation discrepancy $\mathcal{TD}_{1:\ell}(P)$ is the weighted L_2 -star discrepancy of $P_{1:\ell}$, the projection of P on the first ℓ dimensions; moreover, $\mathcal{TD}_{1:s}(P)$ is just the weighted L_2 -star discrepancy $D(P; K_s^*)$ in (5), i.e., $\mathcal{TD}_{1:s}(P) = D(P; K_s^*)$.

2.3 Relations to QMC integration error

We now make use of the theory of reproducing kernel Hilbert spaces (RKHS). For general properties of RKHS, we refer the reader to [2]. If $H(K)$ is a RKHS of functions defined on $[0, 1]^s$ with the reproducing kernel $K(\mathbf{x}, \mathbf{y})$ and norm $\|\cdot\|_{H(K)}$, then

$$|I_s(f) - Q_{n,s}(f)| \leq D(P; K) \|f\|_{H(K)}, \quad \forall f \in H(K),$$

where

$$D(P; K) := \sup\{|I_s(f) - Q_{n,s}(f)| : f \in H(K), \|f\|_{H(K)} \leq 1\}, \quad (9)$$

is called the *discrepancy* of $P = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ with respect to $K(\mathbf{x}, \mathbf{y})$, or the *worst-case error* of the algorithm $Q_{n,s}(f)$ in the unit ball of $H(K)$. Since $H(K)$ is a RKHS, it can be shown that (see [11, 27]):

$$D^2(P; K) = \int_{[0,1]^{2s}} K(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y} - \frac{2}{n} \sum_{i=1}^n \int_{[0,1]^s} K(\mathbf{x}, \mathbf{x}_i) d\mathbf{x} + \frac{1}{n^2} \sum_{i,k=1}^n K(\mathbf{x}_i, \mathbf{x}_k). \quad (10)$$

Now we consider a specific reproducing kernel

$$K_s^*(\mathbf{x}, \mathbf{y}) = \sum_{\mathbf{u} \subseteq \mathcal{A}} \gamma_{\mathbf{u}} K_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}, \mathbf{y}_{\mathbf{u}}), \quad (11)$$

where $K_{\emptyset} = 1, \gamma_{\emptyset} = 1$ and

$$K_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}, \mathbf{y}_{\mathbf{u}}) = \prod_{j \in \mathbf{u}} \min(1 - x_j, 1 - y_j) \quad \text{for } \mathbf{u} \neq \emptyset.$$

It is known (see [11, 27]) that for this kernel the discrepancy (9) is just the weighted L_2 -star discrepancy $D(P; K_s^*)$ defined by (5) (this is the reason for the notation $D(P; K_s^*)$ in (5)). Thus $D(P, K_s^*)$ in (5) may be interpreted as the worst-case error in the RKHS $H(K_s^*)$ in which the square of the norm is known to be

$$\|f\|_{H(K_s^*)}^2 = \sum_{\mathbf{u} \subseteq \mathcal{A}} \gamma_{\mathbf{u}}^{-1} \int_{[0,1]^{|\mathbf{u}|}} \left(\frac{\partial^{|\mathbf{u}|}}{\partial \mathbf{x}_{\mathbf{u}}} f(\mathbf{x}_{\mathbf{u}}, \mathbf{1}) \right)^2 d\mathbf{x}_{\mathbf{u}}.$$

It is shown in [12] that the RKHS $H(K_s^*)$ can be written as the direct sum of the orthogonal spaces $H(K_{\mathbf{u}})$ (the RKHS with the kernel $K_{\mathbf{u}}$):

$$H(K_s^*) = \bigoplus_{\mathbf{u} \subseteq \mathcal{A}} H(K_{\mathbf{u}}), \quad (12)$$

and every function $f \in H(K_s^*)$ has a unique decomposition

$$f(\mathbf{x}) = \sum_{\mathbf{u} \subseteq \mathcal{A}} f_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) \quad \text{with } f_{\mathbf{u}} \in H(K_{\mathbf{u}}).$$

The squared norm of $f_{\mathbf{u}}$ in $H(K_{\mathbf{u}})$ can be expressed as (see [12])

$$\|f_{\mathbf{u}}\|_{H(K_{\mathbf{u}})}^2 = \int_{[0,1]^{|\mathbf{u}|}} \left(\frac{\partial^{|\mathbf{u}|}}{\partial \mathbf{x}_{\mathbf{u}}} f(\mathbf{x}_{\mathbf{u}}, \mathbf{1}) \right)^2 d\mathbf{x}_{\mathbf{u}}. \quad (13)$$

The norm of $f_{\mathbf{u}}$ in $H(K_{\mathbf{u}})$ is related to its norm in $H(K_s^*)$ by $\|f_{\mathbf{u}}\|_{H(K_{\mathbf{u}})}^2 = \gamma_{\mathbf{u}} \|f_{\mathbf{u}}\|_{H(K_s^*)}^2$.

The truncation discrepancy $\mathcal{TD}_{1:\ell}(P)$ is the ℓ -dimensional version of the weighted L_2 -star discrepancy $D(P; K_s^*)$, thus it is the worst-case error in the RKHS $H(K_{\ell}^*)$ with the kernel $K_{\ell}^*(\mathbf{x}_{1:\ell}, \mathbf{y}_{1:\ell})$, i.e.,

$$\mathcal{TD}_{1:\ell}(P) = D(P_{1:\ell}, K_{\ell}^*). \quad (14)$$

Note that, based on the orthogonal decomposition (12), for $f \in H(K_{\ell}^*)$ its norm in $H(K_{\ell}^*)$ is the same as that in $H(K_s^*)$ because of the orthogonality (see [12]).

It can be verified by using (10) and (3) that the order- ℓ discrepancy $\mathcal{D}_{(\ell)}(P)$ is the worst-case error for the RKHS with the kernel

$$K_{(\ell)}(\mathbf{x}, \mathbf{y}) = \sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}|=\ell} \gamma_{\mathbf{u}} K_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}, \mathbf{y}_{\mathbf{u}}),$$

i.e.,

$$\mathcal{D}_{(\ell)}(P) = D(P, K_{(\ell)}). \quad (15)$$

Similarly, for the superposition discrepancy we have

$$\mathcal{SD}_{\ell}(P) = D\left(P, \sum_{k=1}^{\ell} K_{(k)}\right). \quad (16)$$

Moreover, the order- ℓ , superposition and truncation discrepancies appear as pieces in the general integration error bound (9) for $f \in H(K_s^*)$. From the Hlawka-Zaremba identity

$$I_s(f) - Q_{n,s}(f) = \sum_{\emptyset \neq \mathbf{u} \subseteq \mathcal{A}} (-1)^{|\mathbf{u}|} \int_{[0,1]^{|\mathbf{u}|}} \text{disc}(\mathbf{x}_{\mathbf{u}}, \mathbf{1}) \frac{\partial^{|\mathbf{u}|}}{\partial \mathbf{x}_{\mathbf{u}}} f(\mathbf{x}_{\mathbf{u}}, \mathbf{1}) d\mathbf{x}_{\mathbf{u}},$$

together with the Cauchy-Schwarz inequality, (4) and (13), it follows that

$$|I_s(f) - Q_{n,s}(f)| \leq \sum_{\emptyset \neq \mathbf{u} \subseteq \mathcal{A}} D_*(P_{\mathbf{u}}) \|f_{\mathbf{u}}\|_{H(K_s^*)}. \quad (17)$$

Based on this, on applying the Cauchy-Schwarz inequality once again (after multiplying and dividing by $\gamma_{\mathbf{u}}^{1/2}$), we obtain

$$|I_s(f) - Q_{n,s}(f)| \leq \sum_{\ell=1}^s \mathcal{D}_{(\ell)}(P) \mathcal{V}_{(\ell)}(f), \quad (18)$$

where

$$\mathcal{V}_{(\ell)}^2(f) = \sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}|=\ell} \gamma_{\mathbf{u}}^{-1} \|f_{\mathbf{u}}\|_{H(K_{\mathbf{u}})}^2.$$

In terms of the superposition discrepancy, from (18) we have the error bound (for $\ell = 1, \dots, s$)

$$|I_s(f) - Q_{n,s}(f)| \leq \mathcal{SD}_{\ell}(P) \left(\sum_{k=1}^{\ell} \mathcal{V}_{(k)}^2(f) \right)^{1/2} + \sum_{k=\ell+1}^s \mathcal{D}_{(k)}(P) \mathcal{V}_{(k)}(f). \quad (19)$$

Similarly, from (17) we have, for $\ell = 1, \dots, s$,

$$|I_s(f) - Q_{n,s}(f)| \leq \mathcal{TD}_{1:\ell}(P) \mathcal{V}_{1:\ell}(f) + \sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}| \geq 1:\ell} D_*(P_{\mathbf{u}}) \|f_{\mathbf{u}}\|_{H(K_{\mathbf{u}})}, \quad (20)$$

where

$$\mathcal{V}_{1:\ell}^2(f) = \sum_{\emptyset \neq \mathbf{u} \subseteq 1:\ell} \gamma_{\mathbf{u}}^{-1} \|f_{\mathbf{u}}\|_{H(K_{\mathbf{u}})}^2.$$

We summarize the results in the following.

Theorem 1 *The order- ℓ , superposition and truncation discrepancies of P are the worst-case errors in the RKHS with the kernels $K_{(\ell)}(\mathbf{x}, \mathbf{y})$, $\sum_{k=1}^{\ell} K_{(k)}(\mathbf{x}, \mathbf{y})$ and $K_{\ell}^*(\mathbf{x}_{1:\ell}, \mathbf{y}_{1:\ell})$, respectively, i.e.,*

$$\mathcal{D}_{(\ell)}(P) = D(P, K_{(\ell)}) = \sup\{|I_s(f) - Q_{n,s}(f)| : f \in H(K_{(\ell)}), \|f\|_{H(K_s^*)} \leq 1\},$$

$$\mathcal{SD}_{\ell}(P) = D\left(P, \sum_{k=1}^{\ell} K_{(k)}\right) = \sup\{|I_s(f) - Q_{n,s}(f)| : f \in H\left(\sum_{k=1}^{\ell} K_{(k)}\right), \|f\|_{H(K_s^*)} \leq 1\},$$

$$\mathcal{TD}_{1:\ell}(P) = D(P_{1:\ell}, K_{\ell}^*) = \sup\{|I_s(f) - Q_{n,s}(f)| : f \in H(K_{\ell}^*), \|f\|_{H(K_s^*)} \leq 1\}.$$

Moreover, for $f \in H(K_s^*)$ the QMC error can be bounded in terms of order- ℓ , superposition or truncation discrepancy as in (18) or (19) or (20).

3 Generalizations and tractability issues

3.1 Generalizations via reproducing kernel approach

The star discrepancy is based on rectangles with one vertex at the origin. Correspondingly, the reproducing kernel K_s^* has an “anchor” at the opposite vertex $(1, \dots, 1)$. There are other ways of defining discrepancies (see [11]). Consider the RKHS $H(K_s)$ with the kernel

$$K_s(\mathbf{x}, \mathbf{y}) = \sum_{\mathbf{u} \subseteq \mathcal{A}} \gamma_{\mathbf{u}} K_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}, \mathbf{y}_{\mathbf{u}}), \quad (21)$$

where $K_{\emptyset} = 1, \gamma_{\emptyset} = 1$ and

$$K_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}, \mathbf{y}_{\mathbf{u}}) = \prod_{j \in \mathbf{u}} \eta_j(x_j, y_j) \quad \text{for } \mathbf{u} \neq \emptyset. \quad (22)$$

We are interested in the following two choices for the function $\eta_j(x)$:

- Choice (a):

$$\eta_j(x, y) = \begin{cases} \min(|x - a_j|, |y - a_j|), & \text{for } (x - a_j)(y - a_j) > 0, \\ 0, & \text{otherwise,} \end{cases} \quad (23)$$

where $(a_1, \dots, a_s) \in [0, 1]^s$ is a given point (the anchor). If $(a_1, \dots, a_s) = (1, \dots, 1)$, then $\eta_j(x, y) = \min(1-x, 1-y)$, which corresponds to star discrepancy. If $(a_1, \dots, a_s) = (1/2, \dots, 1/2)$, then one obtains the centered discrepancy, see [11].

- Choice (b):

$$\eta_j(x, y) = \frac{1}{2} B_2(\{x - y\}) + (x - \frac{1}{2})(y - \frac{1}{2}), \quad (24)$$

where $B_2(x) = x^2 - x + 1/6$ is the Bernoulli polynomial of degree 2 and $\{x\}$ means the fractional part of x .

We shall call the corresponding RKHS an *anchored Sobolev space* or *unanchored Sobolev space*, respectively. According to [12], in both cases the RKHS $H(K_s)$ can be written as the direct sum of $H(K_u)$:

$$H(K_s) = \bigoplus_{u \subseteq \mathcal{A}} H(K_u),$$

where $H(K_u)$ is the RKHS with the kernel K_u ; moreover, an arbitrary function $f \in H(K_s)$ has a unique *projection decomposition*:

$$f(\mathbf{x}) = \sum_{u \subseteq \mathcal{A}} f_u(\mathbf{x}_u) \quad \text{with} \quad f_u \in H(K_u), \quad (25)$$

and $f_u(\mathbf{x}_u) = (f, K_u(\cdot, \mathbf{x}_u))_{K_u}$. Note that for choice (a) the projection decomposition of a function is in general different from its ANOVA decomposition, on which some concepts of effective dimension are based [3]. For choice (b) these two decompositions coincide [5, 11].

Define

$$K_{(\ell)}(\mathbf{x}, \mathbf{y}) := \sum_{u \subseteq \mathcal{A}, |u|=\ell} \gamma_u K_u(\mathbf{x}_u, \mathbf{y}_u), \quad (26)$$

and

$$K_{\ell}(\mathbf{x}_{1:\ell}, \mathbf{y}_{1:\ell}) := \sum_{u \subseteq 1:\ell} \gamma_u K_u(\mathbf{x}_u, \mathbf{y}_u). \quad (27)$$

Motivated by the relations of the order- ℓ , superposition and truncation discrepancies with the worst-case errors for the special kernel $K_s^*(\mathbf{x}, \mathbf{y})$, we introduce the following generalizations of (14), (15) and (16).

Definition 4. *The order- ℓ , superposition and truncation discrepancies with respect to the general kernel $K_s(\mathbf{x}, \mathbf{y})$ are defined by*

$$\begin{aligned} \mathcal{D}_{(\ell)}(P) &:= D(P; K_{(\ell)}), \\ \mathcal{SD}_{\ell}(P) &:= D\left(P; \sum_{k=1}^{\ell} K_{(k)}\right), \\ \mathcal{TD}_{1:\ell}(P) &:= D(P_{1:\ell}; K_{\ell}), \end{aligned}$$

respectively, where $D(P; K)$ denotes the discrepancy of the point set P with respect to the kernel K , see (10).

As in Section 2, relations exist among the new discrepancies. The superposition and truncation discrepancies can be written in function of the order- ℓ discrepancies as

$$\mathcal{SD}_{\ell}^2(P) = \sum_{k=1}^{\ell} \mathcal{D}_{(k)}^2(P), \quad (28)$$

and

$$\mathcal{TD}_{1:s}^2(P) = \sum_{\ell=1}^s \mathcal{D}_{(\ell)}^2(P) = \mathcal{SD}_s^2(P) = D^2(P; K_s).$$

Each of the discrepancies can be interpreted as a worst-case error as in (9). Relations similar to (18) - (20) can be established for the general order- ℓ , superposition and truncation discrepancies with respect to the general kernels. We omit the details.

3.2 The tractability issues

As benchmarks, we give the expected values for the new discrepancies for random points. For a given kernel of the general form (21), let

$$B_j := \int_{[0,1]^2} \eta_j(x, y) dx dy; \quad A_j := \int_{[0,1]} \eta_j(x, x) dx. \quad (29)$$

For choice **(a)** above we have

$$B_j = a_j^2 - a_j + 1/3, \quad A_j = a_j^2 - a_j + 1/2, \quad (30)$$

whereas for choice **(b)**

$$B_j = 0, \quad A_j = 1/6. \quad (31)$$

If P is a set of random points, the expected value of the squared order- ℓ discrepancy is

$$\mathbb{E}[\mathcal{D}_{(\ell)}^2(P)] = \frac{1}{n} \left(\int_{[0,1]^s} K_{(\ell)}(\mathbf{x}, \mathbf{x}) d\mathbf{x} - \int_{[0,1]^{2s}} K_{(\ell)}(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y} \right).$$

By direct calculation, from (21) and (22) we have

$$\mathbb{E}[\mathcal{D}_{(\ell)}^2(P)] = \frac{1}{n} \left(\sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}|=\ell} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} A_j - \sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}|=\ell} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} B_j \right). \quad (32)$$

Based on (28), we have

$$\mathbb{E}[\mathcal{SD}_{\ell}^2(P)] = \sum_{k=1}^{\ell} \mathbb{E}[\mathcal{D}_{(k)}^2(P)] = \frac{1}{n} \left(\sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}| \leq \ell} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} A_j - \sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}| \leq \ell} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} B_j \right).$$

Similarly, the expected value of the squared truncation discrepancy is

$$\mathbb{E}[\mathcal{TD}_{1:\ell}^2(P)] = \frac{1}{n} \left(\sum_{\mathbf{u} \subseteq 1:\ell} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} A_j - \sum_{\mathbf{u} \subseteq 1:\ell} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} B_j \right). \quad (33)$$

For the product weights (6), this can be written as

$$\mathbb{E}[\mathcal{TD}_{1:\ell}^2(P)] = \frac{1}{n} \left(\prod_{j=1}^{\ell} (1 + \gamma_j A_j) - \prod_{j=1}^{\ell} (1 + \gamma_j B_j) \right).$$

The root mean square expected order- ℓ , superposition and truncation discrepancies behave like $O(n^{-1/2})$.

We briefly study the tractability issues. Since the superposition discrepancy corresponds to the situation of finite-order weights in [26], all results there apply to superposition discrepancy. At the same time, the order- ℓ discrepancy corresponds to finite-order weights with

$\gamma_{\mathbf{u}} = 0$ for $|\mathbf{u}| \neq \ell$. We start with the anchored space and study the bounds on the *normalized order- ℓ discrepancy*. By the normalized order- ℓ discrepancy, we mean $\mathcal{D}_{(\ell)}(P)/\mathcal{D}_{(\ell)}(\emptyset)$, where

$$\mathcal{D}_{(\ell)}^2(\emptyset) = \int_{[0,1]^{2s}} K_{(\ell)}(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y} = \sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}|=\ell} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} B_j$$

is the squared initial order- ℓ discrepancy. (Of course this is interesting only if $B_j \neq 0$ for all j , thus choice **(b)** is excluded from the present discussion.) Given $\varepsilon > 0$, we look for the smallest $n = n(\varepsilon, s, \ell)$ for which there exists a point set P such that $\mathcal{D}_{(\ell)}(P) \leq \varepsilon \mathcal{D}_{(\ell)}(\emptyset)$. That is,

$$n(\varepsilon, s, \ell) = \min\{n : \exists P, |P| = n, \text{ such that } \mathcal{D}_{(\ell)}(P) \leq \varepsilon \mathcal{D}_{(\ell)}(\emptyset)\}.$$

We ask how $n(\varepsilon, s, \ell)$ behaves with respect to ε^{-1} , s and ℓ . The following result is related to the results for finite-order weights [26], but can be proved in a simpler way.

Theorem 2 *Consider the order- ℓ discrepancy with respect to an anchored Sobolev space.*

(i) *For arbitrary weights $\gamma_{\mathbf{u}}$ and $n \geq 1$, there exists a point set P such that*

$$\frac{\mathcal{D}_{(\ell)}^2(P)}{\mathcal{D}_{(\ell)}^2(\emptyset)} \leq \frac{1}{n} (3^\ell - 1).$$

Hence,

$$n(\varepsilon, s, \ell) \leq \left\lceil \frac{3^\ell - 1}{\varepsilon^2} \right\rceil.$$

(ii) *For arbitrary weights $\gamma_{\mathbf{u}}$ and $n \geq 1$, and for an arbitrary point set P ,*

$$\frac{\mathcal{D}_{(\ell)}^2(P)}{\mathcal{D}_{(\ell)}^2(\emptyset)} \geq \left[1 - \left(\frac{8}{9}\right)^\ell n \right].$$

Hence,

$$n(\varepsilon, s, \ell) \geq (1 - \varepsilon^2) \left(\frac{9}{8}\right)^\ell.$$

Proof. For the first part, note that for the anchored Sobolev space, A_j and B_j are given by (30). It follows that for an arbitrary anchor $\mathbf{a} = (a_1, \dots, a_s) \in [0, 1]^s$ we have

$$\min B_j = 1/12, \quad A_j = B_j + 1/6. \quad (34)$$

Note that from (32) the expected value of the squared order- ℓ discrepancy can be written as

$$\mathbb{E}[\mathcal{D}_{(\ell)}^2(P)] = \frac{\rho_{s,\ell} - 1}{n} \mathcal{D}_{(\ell)}^2(\emptyset), \quad (35)$$

where

$$\rho_{s,\ell} = \frac{\sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}|=\ell} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} A_j}{\sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}|=\ell} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} B_j}.$$

From (34), it follows that

$$\begin{aligned}\rho_{s,\ell} &= \frac{\sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}|=\ell} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} B_j \left(1 + \frac{1}{6B_j}\right)}{\sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}|=\ell} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} B_j} \\ &\leq \max_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}|=\ell} \prod_{j \in \mathbf{u}} \left(1 + \frac{1}{6B_j}\right) \leq 3^\ell,\end{aligned}$$

where we used the fact that $\min B_j = 1/12$. Based on (35), there exists a point set P , such that

$$\mathcal{D}_{(\ell)}^2(P)/\mathcal{D}_{(\ell)}^2(\emptyset) \leq \mathbb{E}[\mathcal{D}_{(\ell)}^2(P)]/\mathcal{D}_{(\ell)}^2(\emptyset) \leq \frac{1}{n} (3^\ell - 1).$$

Now we prove the second part. Since the kernel $K_{(\ell)}(\mathbf{x}, \mathbf{y})$ in the anchored case is always non-negative, we may use Lemma 4 of [27], which for the special case of the kernel $K_{(\ell)}$ states that

$$\mathcal{D}_{(\ell)}^2(P) \geq (1 - n \Lambda^2) \mathcal{D}_{(\ell)}^2(\emptyset), \quad (36)$$

where

$$\Lambda^2 = \max_{\mathbf{x} \in [0,1]^s} \frac{h_{(\ell)}^2(\mathbf{x})}{\mathcal{D}_{(\ell)}^2(\emptyset) K_{(\ell)}(\mathbf{x}, \mathbf{x})}, \quad h_{(\ell)}(\mathbf{x}) = \int_{[0,1]^s} K_{(\ell)}(\mathbf{x}, \mathbf{y}) d\mathbf{y}. \quad (37)$$

In the same way as in Theorem 6 of [26], we may prove that

$$\Lambda^2 \leq \frac{\sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}|=\ell} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} W_j}{\sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}|=\ell} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} B_j},$$

where

$$W_j = \max \left(\frac{8a_j^3}{27}, \frac{8(1-a_j)^3}{27} \right).$$

For $a_j \in [1/2, 1]$, we have

$$\frac{W_j}{B_j} = \frac{8a_j^3}{27(a_j^2 - a_j + 1/3)} \in \left[\frac{4}{9}, \frac{8}{9} \right], \quad j = 1, \dots, s.$$

Thus

$$\Lambda^2 \leq \frac{\sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}|=\ell} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} B_j \prod_{j \in \mathbf{u}} \frac{W_j}{B_j}}{\sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}|=\ell} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} B_j} \leq \left(\frac{8}{9} \right)^\ell.$$

The second result in Theorem 2 now follows from (36). ■

The upper bound on $n(\varepsilon, s, \ell)$ limits how good a well chosen point set can be and the lower bound tells how hard the problem is. We observe that $n(\varepsilon, s, \ell)$ can be bounded above independently of the dimension s and the weights, so the normalized order- ℓ discrepancy is strongly tractable for any fixed ℓ and any weights. However, the number $n(\varepsilon, s, \ell)$ depends *exponentially* on ℓ , thus for large ℓ we may have trouble.

It is important to know how the order- ℓ discrepancies of the Sobol' sequence and the Niederreiter sequence behave with respect to s, ℓ and n . The following results can be deduced from the results in [26].

Theorem 3 *Let P_{Sobol} and P_{Nied} be the point sets of the first n points of the s -dimensional Sobol' sequence and the Niederreiter sequence in base b , respectively. For order- ℓ discrepancy with respect to the anchored Sobolev space (associated with the choice **(a)** of the RKHS), we have for arbitrary weights γ_u that*

$$\mathcal{D}_{(\ell)}(P_{\text{Sobol}}) \leq C_1 n^{-1} s^\ell [\log_2(s+1) \log_2 \log_2(s+3) \log_2(2n)]^\ell \mathcal{D}_{(\ell)}(\emptyset),$$

and

$$\mathcal{D}_{(\ell)}(P_{\text{Nied}}) \leq C_2 n^{-1} s^\ell [\log_2(s+b) \log_2(nb)]^\ell \mathcal{D}_{(\ell)}(\emptyset),$$

where C_1 and C_2 are constants independent of s and n .

Sometimes we may be interested in the *absolute* error criterion in which we want to guarantee that the order- ℓ discrepancy $\mathcal{D}_{(\ell)}(P)$ is at most ε . Let $N(\varepsilon, s, \ell)$ be defined as

$$N(\varepsilon, s, \ell) := \min\{n : \exists P, |P| = n, \text{ such that } \mathcal{D}_{(\ell)}(P) \leq \varepsilon\}.$$

We assume that $\gamma_u \leq \Gamma^*$ for some constant Γ^* .

Unlike the normalized order- ℓ discrepancy, the upper bound on the absolute order- ℓ discrepancy depends on s and the weights. For the anchored Sobolev space, the squared initial order- ℓ discrepancy is of the order s^ℓ . Therefore, based on Theorem 2, we have $D_{(\ell)}(P) = O(n^{-1/2} s^{\ell/2})$ for a good choice of P , implying an exponential dependence on ℓ , but a polynomial dependence on s . For the Sobol' or Niederreiter nets, based on Theorem 3 we have $D_{(\ell)}(P) = O(n^{-1} s^{3\ell/2})$, ignoring the logarithmic factors.

Next we consider the unanchored Sobolev space. For the unanchored Sobolev space we have $B_j = 0$ and $A_j = 1/6$ for $j = 1, \dots, s$. The initial order- ℓ discrepancy is $\mathcal{D}_{(\ell)}(\emptyset) = 0$. There is no sense to consider the normalized order- ℓ discrepancy, so we consider the absolute order- ℓ discrepancy. Based on (32), there exists a point set P such that $\mathcal{D}_{(\ell)}(P) = O(n^{-1/2} s^{\ell/2})$. For the Sobol' nets and the Niederreiter nets, it follows from [26] that their order- ℓ discrepancies for the unanchored Sobolev space satisfy the same bounds as in Theorem 3, but with the factor $\mathcal{D}_{(\ell)}(\emptyset)$ replaced by 1. Therefore, their order- ℓ discrepancies are of order $O(n^{-1} s^\ell)$, ignoring the logarithmic factors.

Therefore, for the Sobol' nets and the Niederreiter nets, their order- ℓ discrepancies achieve a faster asymptotic convergence rate than random points do, but their upper bounds have polynomial dependence on s . It is unknown where the asymptotic regime begins. One purpose of this paper is to investigate, for practical n and large s , whether and when the order- ℓ discrepancies of QMC point sets are smaller than these for random point sets. This is done empirically in Section 5.

4 Efficient computational algorithms

We consider the general kernel given by (21) and (22). For general weights, we face a problem to compute the order- ℓ , superposition and truncation discrepancies. For example, the computation of $\mathcal{D}_{(\ell)}(P)$ based on the formula (8) is equivalent to the computation of the classical L_2 -star discrepancy for each projection \mathbf{u} with $|\mathbf{u}| = \ell$ separately. Since P

has $\binom{s}{\ell}$ projections P_u with order $|u| = \ell$, and the calculation of $D_*(P_u)$ requires $O(\ell n^2)$ operations, a direct computation of $\mathcal{D}_{(\ell)}(P)$ would require $O(\binom{s}{\ell} \ell n^2)$ operations. Consequently, computing all $\mathcal{D}_{(\ell)}(P)$ for $\ell = 1, \dots, s$, requires $O(s 2^{s-1} n^2)$ operations, which depends exponentially on s , and thus is an impossible task even for moderate size of s .

Below we show that for the product weights (6) and the order-dependent weights (7), it is unnecessary to deal with each projection separately and efficient algorithms for the truncation and order- ℓ discrepancies exist. Computing the superposition discrepancy is based on the formula (28).

Theorem 4 Consider the reproducing kernel (21), (22). Put $B_j := \int_{[0,1]^2} \eta_j(x, y) dx dy$ and $B_{ij} := \int_{[0,1]} \eta_j(x, x_{i,j}) dx$. Let $P = \{\mathbf{x}_i = (x_{i,1}, \dots, x_{i,s}), i = 1, \dots, n\} \subset [0, 1]^s$. We have
(i) For the product weights (6) and for $\ell = 1, \dots, s$,

$$\mathcal{T}\mathcal{D}_{1:\ell}^2(P) = \prod_{j=1}^{\ell} (1 + B_j \gamma_j) - \frac{2}{n} \sum_{i=1}^n \prod_{j=1}^{\ell} (1 + B_{ij} \gamma_j) + \frac{1}{n^2} \sum_{i,k=1}^n \prod_{j=1}^{\ell} [1 + \gamma_j \eta_j(x_{i,j}, x_{k,j})],$$

$$\mathcal{D}_{(\ell)}^2(P) = \sum_{u \subseteq \mathcal{A}, |u|=\ell} \prod_{j \in u} (B_j \gamma_j) - \frac{2}{n} \sum_{i=1}^n \sum_{u \subseteq \mathcal{A}, |u|=\ell} \prod_{j \in u} (B_{ij} \gamma_j) + \frac{1}{n^2} \sum_{i,k=1}^n \sum_{u \subseteq \mathcal{A}, |u|=\ell} \prod_{j \in u} [\gamma_j \eta_j(x_{i,j}, x_{k,j})].$$

(ii) For the order-dependent weights (7) and for $\ell = 1, \dots, s$,

$$\mathcal{D}_{(\ell)}^2(P) = \Gamma_{(\ell)} \left(\binom{s}{\ell} \prod_{j \in u} B_j - \frac{2}{n} \sum_{i=1}^n \sum_{u \subseteq \mathcal{A}, |u|=\ell} \prod_{j \in u} B_{i,j} + \frac{1}{n^2} \sum_{i,k=1}^n \sum_{u \subseteq \mathcal{A}, |u|=\ell} \prod_{j \in u} \eta_j(x_{i,j}, x_{k,j}) \right), \quad (38)$$

$$\mathcal{T}\mathcal{D}_{1:\ell}^2(P) = \sum_{k=1}^{\ell} \mathcal{D}_{(k)}^2(P_{1:\ell}).$$

Proof If the weights have the product form (6), then the reproducing kernel (27) can be written as the product

$$K_{\ell}(\mathbf{x}_{1:\ell}, \mathbf{y}_{1:\ell}) = \prod_{j=1}^{\ell} (1 + \gamma_j \eta_j(x_j, y_j)).$$

The formula for the truncation discrepancy in (i) follows immediately from the formula (10). The formulas for the order- ℓ discrepancies in both (i) and (ii) are also based on the formula (10) and on the explicit form of the kernel (26). ■

Remark 1. The order- ℓ and superposition discrepancies with different order-dependent weights are related to each other. For order-dependent weights (7), it is clear from (38) that

$$\mathcal{D}_{(\ell)}(P) = \Gamma_{(\ell)}^{1/2} \mathcal{D}_{(\ell)}(P; \mathbf{1}), \quad \ell = 1, \dots, s,$$

where we temporarily use $\mathcal{D}_{(\ell)}(P; \mathbf{1})$ to denote the order- ℓ discrepancy with the specific order-dependent weights: $\Gamma_{(1)} = \dots = \Gamma_{(s)} = 1$. In terms of $\mathcal{D}_{(\ell)}(P; \mathbf{1})$, the square superposition discrepancy corresponding to general weights $\Gamma_{(1)}, \dots, \Gamma_{(s)}$ can be expressed

as

$$\mathcal{SD}_\ell^2(P) = \sum_{k=1}^{\ell} \mathcal{D}_{(k)}^2(P) = \sum_{k=1}^{\ell} \Gamma_{(k)} \mathcal{D}_{(k)}^2(P; \mathbf{1}), \quad \ell = 1, \dots, s.$$

Thus the order- ℓ and superposition discrepancies with arbitrary order-dependent weights can be easily obtained from $\mathcal{D}_{(k)}(P; \mathbf{1})$ for $k = 1, \dots, \ell$.

For product weights or order-dependent weights, the formulas for the order- ℓ discrepancy and the formula (32) involve quantities of the form

$$\sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}|=\ell} \prod_{j \in \mathbf{u}} C_j, \quad \text{for some } C_1, C_2, \dots, C_s. \quad (39)$$

It is crucial to be able to compute such quantities efficiently. Define

$$T(m, \ell, \{C_j\}_{j=1}^s) := T(m, \ell) := \sum_{\mathbf{u} \subseteq 1:m, |\mathbf{u}|=\ell} \prod_{j \in \mathbf{u}} C_j, \quad (40)$$

for $m = 1, \dots, s$ and $\ell = 1, \dots, m$. We can view T as an $s \times s$ lower triangular matrix. Obviously, $T(m, 1) = \sum_{j=1}^m C_j$ and $T(m, m) = \prod_{j=1}^m C_j$ for $m = 1, \dots, s$. The elements of the last row of T are $T(s, \ell) = \sum_{\mathbf{u} \subseteq \mathcal{A}, |\mathbf{u}|=\ell} \prod_{j \in \mathbf{u}} C_j$ for $\ell = 1, \dots, s$, which are the quantities in (39). The following lemma gives a recursive relation for $T(m, \ell)$.

Lemma 5 *Let $T(m, \ell)$ be defined by (40), then*

$$T(m, \ell) = T(m-1, \ell) + C_m T(m-1, \ell-1), \quad \text{for } m \geq 3, \ell \geq 2. \quad (41)$$

Proof From the definition of $T(m, \ell)$, it follows that

$$\begin{aligned} T(m, \ell) &= \sum_{\mathbf{u} \subseteq 1:m, |\mathbf{u}|=\ell} \prod_{j \in \mathbf{u}} C_j \\ &= \sum_{\mathbf{u} \subseteq 1:(m-1), |\mathbf{u}|=\ell} \prod_{j \in \mathbf{u}} C_j + \sum_{m \in \mathbf{u} \subseteq 1:m, |\mathbf{u}|=\ell} \prod_{j \in \mathbf{u}} C_j \\ &= T(m-1, \ell) + C_m T(m-1, \ell-1). \quad \blacksquare \end{aligned}$$

Based on the recursive relation (41), we can easily compute the elements of the last row of T : $T(s, 1), T(s, 2), \dots, T(s, s)$. Computing all these element requires $O(s^2)$ operations.

Lemma 5 makes an efficient computation of the order- ℓ discrepancy possible. For example, the order- ℓ discrepancy with order-dependent weights in Theorem 4 can be written as

$$\begin{aligned} \mathcal{D}_{(\ell)}^2(P) &= \Gamma_{(\ell)} \left[T(s, \ell, \{B_j\}_{j=1}^s) - \frac{2}{n} \sum_{i=1}^n T(s, \ell, \{B_{ij}\}_{j=1}^s) \right. \\ &\quad \left. + \frac{1}{n^2} \sum_{i,k=1}^n T(s, \ell, \{\eta_j(x_{i,j}, x_{k,j})\}_{j=1}^s) \right]. \end{aligned}$$

The number of operations needed for computing all the elements $\mathcal{D}_{(\ell)}(P)$, $\ell = 1, \dots, s$, is $O(s^2 n^2)$. By using symmetry the computing cost can be cut in half.

5 Computational investigations

We investigate empirically the order- ℓ , superposition and truncation discrepancies of the Sobol' nets and compare them with the mean values for the sets of random points and of Latin hypercube sampling (LHS). The LHS will be introduced below. Note that in the generation of the Sobol' nets, we choose the direction numbers such that certain additional uniformity properties are satisfied [29].

We consider the reproducing kernel (21) with η_j given by choice (b), see (24). We prefer this kernel, since in this case the projection decomposition (25) and the ANOVA decomposition of a function are identical. The quantities involved in the formulas of the order- ℓ and truncation discrepancies are given in (31) and $B_{ij} = \int_{[0,1]} \eta_j(x, x_{i,j}) dx = 0$. The computation is based on Theorem 4. The mean values for random point sets are given in (32) and (33). The mean values for LHS are given below.

5.1 The discrepancies of LHS

LHS [13] provides a good set of points for integration. A LHS point set of size n , $P_{\text{LHS}} = \{(x_{i,1}, \dots, x_{i,s}) : i = 1, \dots, n\}$, is generated as

$$x_{i,j} = \frac{\pi_j(i) - U_{ij}}{n}, \quad i = 1, \dots, n; \quad j = 1, \dots, s, \quad (42)$$

where π_j are random permutations of $\{1, \dots, n\}$, each uniformly distributed over all $n!$ possible permutations, and U_{ij} are uniform $[0, 1]$ random variables; moreover, the permutations and the uniform random variates are mutually independent. Note that LHS stratifies each individual dimension, but imposes no higher dimensional stratification. We now derive formulas for the mean square truncation and order- ℓ discrepancies of LHS.

Theorem 6 *Let P_{LHS} be a LHS set of size n . Consider the reproducing kernel (21) with $\eta_j(x, y)$ defined in (24).*

(i) *For the product weights (6) and for $\ell = 1, \dots, s$,*

$$\mathbb{E}[\mathcal{T}\mathcal{D}_{1:\ell}^2(P_{\text{LHS}})] = -1 + \frac{1}{n} \prod_{j=1}^{\ell} \left(1 + \frac{\gamma_j}{6}\right) + \left(1 - \frac{1}{n}\right) \prod_{j=1}^{\ell} \left[1 + \gamma_j \left(\frac{1-n}{12n^2}\right)\right].$$

$$\mathbb{E}[\mathcal{D}_{(\ell)}^2(P_{\text{LHS}})] = \frac{1}{n} \sum_{u \subseteq \mathcal{A}, |u|=\ell} \prod_{j \in u} \left(\frac{\gamma_j}{6}\right) + \left(1 - \frac{1}{n}\right) \sum_{u \subseteq \mathcal{A}, |u|=\ell} \prod_{j \in u} \left(\frac{(1-n)\gamma_j}{12n^2}\right).$$

(ii) *For the order-dependent weights (7) and for $\ell = 1, \dots, s$,*

$$\mathbb{E}[\mathcal{D}_{(\ell)}^2(P_{\text{LHS}})] = \Gamma_{(\ell)} \left[\frac{1}{n6^\ell} + \left(1 - \frac{1}{n}\right) \left(\frac{1-n}{12n^2}\right)^\ell \right] \binom{s}{\ell},$$

$$\mathbb{E}[\mathcal{T}\mathcal{D}_{1:\ell}^2(P_{\text{LHS}})] = \sum_{k=1}^{\ell} \mathbb{E}[\mathcal{D}_{(k)}^2(P_{\text{LHS}, 1:\ell})].$$

Proof We prove only the first formula. For the product weights (6), from Theorem 4 we have, using (24), (31) and $B_{ij} = 0$ for all i, j ,

$$\mathcal{TD}_{1:\ell}^2(P_{\text{LHS}}) = -1 + \frac{1}{n^2} \sum_{i,k=1}^n \prod_{j=1}^{\ell} \left[1 + \gamma_j \left(\frac{1}{2} B_2(\{x_{i,j} - x_{k,j}\}) + (x_{i,j} - \frac{1}{2})(x_{k,j} - \frac{1}{2}) \right) \right],$$

where the random variables $x_{i,j}$ are defined by (42). On separating the $i = k$ terms and the $i \neq k$ terms, we obtain

$$\begin{aligned} \mathcal{TD}_{1:\ell}^2(P_{\text{LHS}}) &= -1 + \frac{1}{n^2} \sum_{i=1}^n \prod_{j=1}^{\ell} \left[1 + \gamma_j \left(\frac{1}{2} B_2(0) + (x_{i,j} - \frac{1}{2})^2 \right) \right] \\ &\quad + \frac{1}{n^2} \sum_{i \neq k} \prod_{j=1}^{\ell} \left[1 + \gamma_j \left(\frac{1}{2} B_2(\{x_{i,j} - x_{k,j}\}) + (x_{i,j} - \frac{1}{2})(x_{k,j} - \frac{1}{2}) \right) \right]. \end{aligned}$$

By taking expectations and by direct computation, we have

$$\mathbb{E}[(x_{i,j} - \frac{1}{2})^2] = \frac{1}{12}; \quad \mathbb{E}(x_{i,j} - \frac{1}{2}) = 0;$$

$$\mathbb{E}[B_2(\{x_{i,j} - x_{k,j}\})] = \frac{1}{6n^2} - \frac{1}{6n} \quad \text{for } i \neq k,$$

and the result follows. ■

5.2 Numerical comparisons

We consider two groups of weights:

- Group (A) (order-dependent weights):
 - (A1) $\Gamma_{(1)} = \dots = \Gamma_{(s)} = 1$ (or equivalently, $\gamma_1 = \dots = \gamma_s = 1$);
 - (A2) $\Gamma_{(1)} = 1, \Gamma_{(2)} = \dots = \Gamma_{(s)} = 1/100$;
 - (A3) $\Gamma_{(1)} = 1, \Gamma_{(2)} = \dots = \Gamma_{(s)} = 1/10000$.
- Group (B) (product weights): $\gamma_j = 1/2^{j-1}, j = 1, \dots, s$.

Note that for order-dependent weights, the order- ℓ and superposition discrepancies for any weights $\Gamma_{(\ell)}$ can be deduced from those for the basic case (A1) as mentioned in Remark 1; moreover, from equation (38) in Theorem 4 the relative performance of the order- ℓ discrepancies of any two point sets does not depend on the weights $\Gamma_{(\ell)}$. So we will focus on the order- ℓ discrepancy for the basic case (A1).

The computational results for $s = 64$ are given in the appendix. The convergence rates for the order- ℓ , superposition and truncation discrepancies of the Sobol' nets, i.e., the value a in an expression of the form $O(n^{-a})$, estimated from linear regression on the empirical data with $n = 2^8, 2^{10}, 2^{12}$, are also given (the convergence rates give sufficient information on the behavior of the Sobol' nets with the increasing of n). We observe the following:

- In cases (A1) and (B) (see Tables 1 and 2), the order-1 discrepancies of the Sobol' nets are much smaller than those of LHS and random point sets (this is not surprising, see Remark 2 below). But for $\ell = 2, 3, \dots$, the advantage for Sobol' nets may no longer exist for relatively small n ; in fact, in case (A1) the order- ℓ (even the order-2) discrepancies for all three kind of point sets are almost the same, which implies that the high-order (including order-2) projections of the Sobol' nets are no more uniform than LHS and random point sets on the average. This is somewhat surprising. In case (B), the superiority of order- ℓ discrepancy of the Sobol' nets is preserved up to at least order 4. (This is because of the often observed property of the Sobol' nets that the low-order projections involving earlier dimensions are much better than these involving later dimensions, and this superiority for the earlier dimensions is allowed to show itself because the later dimensions are associated with smaller weights.) Thus the advantage of Sobol' nets in high dimensions is mainly provided by the order-1 projections and by projections of the earlier dimension.
- The performance of the superposition discrepancy (see Tables 3 and 4) depends on the weights. For order-dependent weights, the ratio $\Gamma_{(\ell)}/\Gamma_{(1)}$ ($\ell > 1$) determines the possible superiority of low discrepancy point sets. Only when $\Gamma_{(\ell)}$ ($\ell > 1$) is much smaller than $\Gamma_{(1)}$ is the superposition discrepancy of the Sobol' nets smaller than that of LHS or random point set. Indeed, for the largest weights (A1) the superposition discrepancy even for $\ell = 2$ is largest for Sobol' nets, whereas for the weights (A3), where the effect of order-2 projections is much smaller, the superposition discrepancy for Sobol' nets is much smaller than for the other two. For the weights (B), the Sobol' nets always behave better or much better than LHS and random point sets do. Again this reflects the fact that any superiority of the order- ℓ projections of the Sobol' nets for $\ell > 1$ is confined to the earlier dimensions.
- The performance of truncation discrepancy (see Tables 5 and 6) also depends on the weights. For the basic case (A1) the truncation discrepancy $\mathcal{D}_{1:\ell}(P)$ for the Sobol' nets can be smaller or much smaller than that for LHS or random point set only if ℓ is relatively small (say $\ell \leq 10$). For fast decaying weights (B), the truncation discrepancy for Sobol' nets behaves better or much better than for LHS or random point set in all dimensions (this is because the contribution from the high-order projections is negligible due to the small weights γ_u). This is consistent with the fact that the truncation discrepancy of Sobol' nets can achieve $O(n^{-1+\delta})$ convergence for arbitrary $\delta > 0$ in certain weighted Sobolev spaces uniformly in ℓ (see [30]).

Remark 2. The $O(n^{-1})$ convergence of the order-1 discrepancy for the Sobol' nets is not surprising. In fact, a simple formula can be derived for $\mathcal{D}_{(1)}(P_{\text{Sobol}})$, where P_{Sobol} is the s -dimensional Sobol' net with $n = 2^m$. For each $j = 1, \dots, s$, the projection of P_{Sobol} on j -th dimension coincides with the set $P_{\{1\}} = \{x_i = i/n : i = 0, \dots, n-1\}$, whose discrepancy with respect to the kernel $\eta(x, y) = \frac{1}{2}B_2(\{x - y\}) + (x - 1/2)(y - 1/2)$ can be calculated

directly

$$D(P_{\{1\}}, \eta) = \frac{1}{n^2} \sum_{i,k=1}^n \eta(x_i, x_k) = \frac{\sqrt{3}}{3n}.$$

Since the order-1 discrepancy $\mathcal{D}_{(1)}(P_{\text{Sobol}})$ is the discrepancy of P_{Sobol} with respect to the kernel $K_{(1)}(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^s \gamma_{\{j\}} \eta(x_j, y_j)$, with $\gamma_{\{j\}} = 1$ we have

$$\mathcal{D}_{(1)}^2(P_{\text{Sobol}}) = \frac{s}{3n^2}.$$

So $\mathcal{D}_{(1)}(P_{\text{Sobol}})$ has a $O(n^{-1})$ convergence and depends linearly on \sqrt{s} (this result is slightly stronger than that of Theorem 3 for $\ell = 1$).

For a random point set P_{rand} or a LHS set P_{LHS} , from (32) and from Theorem 6 we have

$$\mathbb{E}(\mathcal{D}_{(1)}^2(P_{\text{rand}})) = \frac{s}{6n}, \quad \mathbb{E}(\mathcal{D}_{(1)}^2(P_{\text{LHS}})) = \frac{s(n^2 + 2n - 1)}{12n^3}.$$

Both have the same weak dependence on s , but a worse dependence on n .

Similar formulas can be derived for the weights (B):

$$\mathcal{D}_{(1)}^2(P_{\text{Sobol}}) = \frac{G}{3n^2}, \quad \mathbb{E}[\mathcal{D}_{(1)}^2(P_{\text{rand}})] = \frac{G}{6n}, \quad \mathbb{E}[\mathcal{D}_{(1)}^2(P_{\text{LHS}})] = \frac{G(n^2 + 2n - 1)}{12n^3}.$$

where $G = 2 - 2^{1-s}$. All are uniformly bounded in s .

Remark 3. The behavior of the order- ℓ discrepancy in the most general case (A1) gives an indication of the circumstances for which QMC can be more efficient than MC for functions with a large number of *equally important variables*. In such cases, the relative importance of the first-order part with respect to the higher-order part is an important indicator of QMC performance. If the function is dominated by the first-order part, then good results can be expected from using QMC. Consider an example (with large s , say $s = 64$)

$$f(\mathbf{x}) = \prod_{j=1}^s (1 + \Gamma_s(x_j - 1/2)),$$

where $\Gamma_s > 0$ is a parameter. Then

$$f_{(1)}(\mathbf{x}) = \Gamma_s \sum_{j=1}^s (x_j - 1/2), \quad f_{(2)}(\mathbf{x}) = \Gamma_s^2 \sum_{i < j} (x_i - 1/2)(x_j - 1/2).$$

are the sums of the first-order and second-order ANOVA terms of f , respectively. Simple tests show that the Sobol' nets work much better than MC for f and $f_{(1)} + f_{(2)}$ if Γ_s is small (since the first-order terms dominate the functions), but they work no better for $f_{(2)}$ and $f - f_{(1)}$, which include no first-order terms. This observation indicates the importance of the additive structure of the integrands in explaining the success of QMC when there are many variables of nearly equal importance.

6 Conclusions

Classical discrepancies are not good enough for measuring the quality of point sets in high dimensions. We therefore introduced more refined quality measures by focusing on selected aggregations of projections. Their connections with integration errors were studied, and bounds on the new discrepancies were established. In particular, we showed that the Sobol' sequence and the Niederreiter sequence have order- ℓ discrepancies of order $O(s^\ell n^{-1}(\log n)^\ell)$ if logarithms are ignored. Efficient algorithms for computing the new discrepancies were presented (faster algorithms are under development such that we could easily deal with much larger number of points. One possibility is to use the algorithm in [10]). Computational investigations were performed in dimension $s = 64$, with the results that in comparison with the LHS and random point sets, the Sobol' nets have perfect order-1 projections and better order-2 projections for the earlier components (as can be expected), but have no better order-2 and higher-order projections on the average for moderate values of n (this is somewhat surprising and pessimistic). It is unknown how large n needs to be to ensure that the higher-order projections of Sobol' nets are better than those of random point sets. More studies are needed to fully understand the behavior of the higher-order projections of various QMC point sets. Due to the advantages of the order-1 projections and the projections of the earlier dimensions, the Sobol' nets can have much smaller superposition and truncation discrepancies for fast decaying weights (this is fortunate, since many practical problems are indeed weighted, with the first few variables more important than later ones).

The investigation indicates the possible advantages and potential limitations of QMC in high dimensions. The answer to the question of whether QMC is superior to MC and how well QMC could work for high-dimensional integrals depends strongly on the structure of the function. If a function is dominated by its first-order part, or if the earlier variables are much more important than others, then good results can be expected from using QMC. However, if higher-order ANOVA terms play an important role, then the superiority of QMC is not guaranteed for relatively small n . It is important to construct QMC point sets in high dimensions, such that some higher-order discrepancies (at least the order-2 discrepancies) are much smaller than those of random point sets. The new quality measures proposed in this paper are useful in constructing better QMC point sets. It is possible that some randomization techniques, such as the scrambling [19] and the digital shift (see [4]), may improve the order-2 and higher-order projections. The new quality measures provide suitable means to assess the possible improvement. One way would be to compute the expected order- ℓ or superposition discrepancies for scrambling or digital shift.

According to [32], many financial problems have low effective dimensions, since they are controlled by some hidden weights which limit the contribution of higher-order ANOVA terms. Quite often, an ANOVA expansion up to the second-order can provide a quite satisfactory approximation to the functions, with the first-order part playing the major role (this might be called the QMC-friendly property). Interestingly, the QMC-friendly property can often be enhanced by dimension reduction techniques, so that the truncation dimension is reduced and the importance of the first-order part is increased, thereby further enhancing the superiority of QMC methods.

Acknowledgments

The authors would like to thank the referees for their valuable comments and suggestions. The support of the National Science Foundation of China and of the Australian Research Council under its Centre of Excellence Program are gratefully acknowledged.

References

- [1] P. Acworth, M. Broadie and P. Glasserman, A comparison of some Monte Carlo and quasi-Monte Carlo techniques for option pricing. In Monte Carlo and Quasi-Monte Carlo Methods 1996, H. Niederreiter, P. Hellekalek, G. Larcher and P. Zinterhof (eds.), 1-18, Springer-Verlag, New York, 1998.
- [2] N. Aronszajn, Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, **68** (1950), 337-404.
- [3] R. E. Caflisch, W. Morokoff and A. Owen, Valuation of Mortgage backed securities using Brownian bridges to reduce effective dimension. *J. Comp. Finance*, **1** (1997), 27-46.
- [4] J. Dick and F. Pillichshammer, On the mean square weighted L_2 discrepancy of randomized digital (t, m, s) -nets over Z_2 . *Acta Arith.*, **117** (2005), 371-403.
- [5] J. Dick, I. H. Sloan, X. Wang and H. Woźniakowski, Liberating the weights. *J. Complexity*, **20** (2004), 593-623.
- [6] J. Dick, I.H. Sloan, X. Wang and H. Woźniakowski, *Good lattice rules in weighted Korobov spaces with general weights*, *Numer. Math.*, **103** (2006), 63-97.
- [7] P. L'Ecuyer and C. Lemieux, Variance reduction via lattice rules, *Management Science*, **46** (2000), 1214-1235.
- [8] H. Faure, Discrepance de suites associées à un système de numération (en dimension s), *Acta Arith.*, **41** (1982), 337-351.
- [9] J. H. Halton, On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numer. Math.*, **2** (1960), 84-90.
- [10] S. Heinrich, Efficient algorithms for computing the L_2 -discrepancy *Math. Comp.*, **65** (1996), 1621-1633.
- [11] F. J. Hickernell, A generalized discrepancy and quadrature error bound. *Math. Comp.*, **67** (1998), 299-322.
- [12] F. J. Hickernell and X. Wang, The error bounds and tractability of quasi-Monte Carlo algorithms in infinite dimension. *Math. Comp.*, **71** (2002), 1641-1661.
- [13] M. D. McKay, W. J. Conover and R. J. Beckman, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, **21** (1979), 239-245.
- [14] B. Moskowitz and R. E. Caflisch, Smoothness and dimension reduction in quasi-Monte Carlo methods. *Math. Comput. Modelling*, **23** (1996), 37-54.
- [15] W. J. Morokoff and R. E. Caflisch, Quasi-random sequences and their discrepancies. *SIAM J. Sci. Comput.*, **15** (1994), 1251-1279.

- [16] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods. SIAM, Philadelphia, 1992.
- [17] H. Niederreiter, C. P. Xing, Low discrepancy sequences and global functional fields with many rational places, *Finite Fields Appl.*, **2** (1996), 241-273.
- [18] S. Ninomiya and S. Tezuka, Toward real time pricing of complex financial derivatives. *Applied Mathematical Finance*, **3** (1996), 1-20.
- [19] A. B. Owen, Randomly permuted (t, m, s) -nets and (t, s) -sequences. In Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, H. Niederreiter and Shiue (eds.), p.299-317, Springer-Verlag, 1995.
- [20] A. B. Owen, The dimension distribution, and quadrature test functions. *Statistica Sinica*, **13** (2003), 1-17.
- [21] A. Papageorgiou and J. F. Traub, Faster valuation of multi-dimensional integrals, *Computers in Physics*, Nov./Dec. (1997), 474-578.
- [22] S. H. Paskov and J. F. Traub, Faster valuation of financial derivatives. *J. Portfolio Management*, **22** (1995), 113-120.
- [23] W. Ch. Schmid, Projections of digital nets and sequences, *Math. Comput. Simulation*, **55** (2001), 239-247
- [24] I. H. Sloan and S. Joe, Lattice methods for Multiple Integration. Oxford University Press, Oxford, 1994.
- [25] I. H. Sloan, F. Y. Kuo and S. Joe, On the step-by-step construction of quasi-Monte Carlo integration rules that achieves strong tractability error bounds in weighted Sobolev spaces. *Math. Comp.*, **71** (2002), 1609-1640.
- [26] I. H. Sloan, X. Wang and H. Woźniakowski. Finite-order weights imply tractability of multivariate integration. *J. Complexity*, **20** (2004), 46-74.
- [27] I. H. Sloan and H. Woźniakowski, When are quasi-Monte Carlo algorithms efficient for high dimensional integrals? *J. Complexity*, **14** (1998), 1-33.
- [28] I. M. Sobol', On the distribution of points in a cube and the approximate evaluation of integrals. *Zh. Vychisl. Mat. i Mat. Fiz.*, **7** (1967), 784-802.
- [29] I. M. Sobol', Uniformly distributed sequences with additional uniform properties. *Zh. Vychisl. Mat. i Mat. Fiz.*, **16** (1976), 1332-1337.
- [30] X. Wang, Strong tractability of multivariate integration using quasi-Monte Carlo algorithms. *Math. Comp.*, **72** (2003), 823-838.
- [31] X. Wang and K. T. Fang, Effective dimensions and quasi-Monte Carlo integration, *J. Complexity*, **19** (2003), 101-124.
- [32] X. Wang and I. H. Sloan, Why are high-dimensional finance problems often of low effective dimension? *SIAM J. Sci. Comput.*, **27** (2005), 159-183.
- [33] T. T. Warnock, Computational investigations of low discrepancy point sets. In Application of Number Theory to Numerical Analysis, S. K. Zaremba (ed.). Academic Press, New York, 319-343, 1972.
- [34] H. Woźniakowski, Average case of complexity of multivariate integration, *Bull. Amer. Math. Soc.*, **24** (1991), 185-194.

Appendix: Computational results

1. Comparisons of the order- ℓ discrepancies

n	Point set	$\mathcal{D}_{(1)}$	$\mathcal{D}_{(2)}$	$\mathcal{D}_{(4)}$	$\mathcal{D}_{(8)}$	$\mathcal{D}_{(16)}$	$\mathcal{D}_{(32)}$	$\mathcal{D}_{(64)}$
2^8	Random	2.04e-1	4.68e-1	1.38e00	3.21e00	8.22e-1	3.00e-5	7.85e-27
	LHS	1.45e-1	4.68e-1	1.38e00	3.21e00	8.22e-1	3.00e-5	7.85e-27
	Sobol	1.80e-2	5.51e-1	1.42e00	3.19e00	8.11e-1	2.82e-5	4.84e-27
2^{10}	Random	1.02e-1	2.34e-1	6.92e-1	1.60e00	4.11e-1	1.50e-5	3.93e-27
	LHS	7.22e-2	2.34e-1	6.92e-1	1.60e00	4.11e-1	1.50e-5	3.93e-27
	Sobol	4.51e-3	3.05e-1	7.13e-1	1.61e00	4.12e-1	1.56e-5	4.60e-27
2^{12}	Random	5.10e-2	1.17e-1	3.46e-1	8.02e-1	2.06e-1	7.50e-6	1.96e-27
	LHS	3.61e-2	1.17e-1	3.46e-1	8.02e-1	2.06e-1	7.50e-6	1.96e-27
	Sobol	1.13e-3	1.44e-1	3.57e-1	8.07e-1	2.10e-1	8.41e-6	2.41e-27
a	Sobol	1.00	0.48	0.50	0.50	0.49	0.49	0.25

Table 1. The order- ℓ discrepancies of Sobol', Latin hypercube sampling and random point sets in dimension $s = 64$ for the basic case (A1): $\Gamma_{(1)} = \dots = \Gamma_{(s)} = 1$. The last row gives the convergence orders for the Sobol' nets, estimated from linear regression on the empirical data with $n = 2^8, 2^{10}, 2^{12}$.

n	Point Set	$\mathcal{D}_{(1)}$	$\mathcal{D}_{(2)}$	$\mathcal{D}_{(4)}$	$\mathcal{D}_{(8)}$	$\mathcal{D}_{(16)}$	$\mathcal{D}_{(32)}$
2^8	Random	3.61e-2	1.20e-2	3.91e-4	5.47e-9	6.00e-26	9.11e-89
	LHS	2.56e-2	1.20e-2	3.91e-4	5.47e-9	6.00e-26	9.11e-89
	Sobol	3.18e-3	3.95e-3	3.90e-4	5.12e-9	5.02e-26	9.45e-89
2^{10}	Random	1.80e-2	6.01e-3	1.96e-4	2.73e-9	3.00e-26	4.56e-89
	LHS	1.28e-2	6.01e-3	1.96e-4	2.73e-9	3.00e-26	4.56e-89
	Sobol	7.97e-4	1.15e-3	1.67e-4	2.63e-9	2.88e-26	6.12e-89
2^{12}	Random	9.02e-3	3.07e-3	9.78e-5	1.37e-9	1.50e-26	2.28e-89
	LHS	6.38e-3	3.01e-3	9.78e-5	1.37e-9	1.50e-26	2.28e-89
	Sobol	1.99e-4	5.74e-4	6.03e-5	1.33e-9	1.59e-26	2.71e-89
a	Sobol	1.00	0.70	0.67	0.49	0.42	0.45

Table 2. The same as Table 1, but for the weights (B): $\gamma_j = 1/2^{j-1}, j = 1, \dots, s$.

2. Comparisons of the superposition discrepancies

Weights	Point Set	\mathcal{SD}_1	\mathcal{SD}_2	\mathcal{SD}_4	\mathcal{SD}_8	\mathcal{SD}_{16}	\mathcal{SD}_{32}	\mathcal{SD}_{64}
(A1)	Random	1.02e-1	2.55e-1	8.56e-1	2.83e00	4.32e00	4.34e00	4.34e00
	LHS	7.22e-2	2.45e-1	8.53e-1	2.83e00	4.32e00	4.34e00	4.34e00
	Sobol	4.51e-3	3.05e-1	8.78e-1	2.83e00	4.32e00	4.33e00	4.33e00
		(1.00)	(0.48)	(0.48)	(0.49)	(0.49)	(0.49)	(0.49)
(A2)	Random	1.02e-1	1.05e-1	1.33e-1	3.00e-1	4.44e-1	4.45e-1	4.45e-1
	LHS	7.22e-2	7.59e-2	1.12e-1	2.92e-1	4.38e-1	4.40e-1	4.40e-1
	Sobol	4.51e-3	3.08e-2	8.79e-2	2.83e-1	4.32e-1	4.36e-1	4.36e-1
		(1.00)	(0.48)	(0.49)	(0.49)	(0.49)	(0.49)	(0.49)
(A3)	Random	1.02e-1	1.02e-1	1.02e-1	1.06e-1	1.11e-1	1.11e-1	1.11e-1
	LHS	7.22e-2	7.23e-2	7.27e-2	7.76e-2	8.42e-2	8.42e-2	8.42e-2
	Sobol	4.51e-3	5.44e-3	9.87e-3	2.87e-2	4.34e-2	4.36e-2	4.36e-2
		(1.00)	(0.71)	(0.58)	(0.51)	(0.50)	(0.50)	(0.50)

Table 3. The superposition discrepancies in dimension $s = 64$ with $n = 2^{10}$ for the order-dependent weights of Group (A): (A1) $\Gamma_{(1)} = \dots = \Gamma_{(s)} = 1$; (A2) $\Gamma_{(1)} = 1, \Gamma_{(2)} = \dots = \Gamma_{(s)} = 1/100$; (A3) $\Gamma_{(1)} = 1, \Gamma_{(2)} = \dots = \Gamma_{(s)} = 1/10000$. The convergence orders for Sobol' nets (estimated from linear regression on the empirical data with $n = 2^8, 2^{10}, 2^{12}$) are given in the parentheses for each case. The superposition discrepancies for $n = 2^8$ and 2^{12} are omitted.

n	Point Set	\mathcal{SD}_1	\mathcal{SD}_2	\mathcal{SD}_4	\mathcal{SD}_8	\mathcal{SD}_{16}	\mathcal{SD}_{32}	\mathcal{SD}_{64}
2^8	Random	3.61e-2	3.80e-2	3.81e-2	3.81e-2	3.81e-2	3.81e-2	3.81e-2
	LHS	2.56e-2	2.83e-2	2.84e-2	2.84e-2	2.84e-2	2.84e-2	2.84e-2
	Sobol	3.18e-3	5.07e-3	5.38e-3	5.38e-3	5.38e-3	5.38e-3	5.38e-3
2^{10}	Random	1.80e-2	1.90e-2	1.91e-2	1.91e-2	1.91e-2	1.91e-2	1.91e-2
	LHS	1.28e-2	1.41e-2	1.41e-2	1.41e-2	1.41e-2	1.41e-2	1.41e-2
	Sobol	7.97e-4	1.40e-3	1.52e-3	1.52e-3	1.52e-3	1.52e-3	1.52e-3
2^{12}	Random	9.02e-3	9.51e-3	9.53e-3	9.53e-3	9.53e-3	9.53e-3	9.53e-3
	LHS	6.38e-3	7.05e-3	7.08e-3	7.08e-3	7.08e-3	7.08e-3	7.08e-3
	Sobol	1.99e-4	6.08e-4	6.40e-4	6.40e-4	6.40e-4	6.40e-4	6.40e-4
a	Sobol	1.00	0.77	0.77	0.77	0.77	0.77	0.77

Table 4. The superposition discrepancies in dimension $s = 64$ for the product weights (B). The convergence orders for Sobol' nets (estimated from linear regression on the empirical data with $n = 2^8, 2^{10}, 2^{12}$) are also given.

3. Comparisons of the truncation discrepancies

n	Point Set	$\mathcal{TD}_{\{1\}}$	$\mathcal{TD}_{1:2}$	$\mathcal{TD}_{1:4}$	$\mathcal{TD}_{1:8}$	$\mathcal{TD}_{1:16}$	$\mathcal{TD}_{1:32}$	$\mathcal{TD}_{1:64}$
2^8	Random	2.55e-2	3.76e-2	5.77e-2	9.75e-2	2.05e-1	7.34e-1	8.67e00
	LHS	1.81e-2	2.77e-2	4.52e-2	8.32e-2	1.92e-1	7.27e-1	8.67e00
	Sobol	2.25e-3	4.34e-3	1.10e-2	3.80e-2	1.60e-1	7.24e-1	8.61e00
2^{10}	Random	1.28e-2	1.88e-2	2.89e-2	4.87e-2	1.02e-1	3.67e-1	4.34e00
	LHS	9.03e-3	1.38e-2	2.25e-2	4.15e-2	9.61e-2	3.63e-1	4.34e00
	Sobol	5.63e-4	1.09e-3	3.08e-3	1.40e-2	9.86e-2	3.70e-1	4.33e00
2^{12}	Random	6.38e-3	9.39e-3	1.44e-2	2.44e-2	5.13e-2	1.83e-1	2.16e00
	LHS	4.51e-3	6.89e-3	1.13e-2	2.08e-2	4.80e-2	1.82e-1	2.16e00
	Sobol	1.41e-4	2.90e-4	9.12e-4	4.90e-3	4.21e-2	2.15e-1	2.19e00
a	Sobol	1.00	0.98	0.90	0.74	0.48	0.44	0.49

Table 5. The truncation discrepancy and the convergence order for the case (A1).

n	Point Set	$\mathcal{TD}_{\{1\}}$	$\mathcal{TD}_{1:2}$	$\mathcal{TD}_{1:4}$	$\mathcal{TD}_{1:8}$	$\mathcal{TD}_{1:16}$	$\mathcal{TD}_{1:32}$	$\mathcal{TD}_{1:64}$
2^8	Random	2.55e-2	3.21e-2	3.67e-2	3.80e-2	3.81e-2	3.81e-2	3.81e-2
	LHS	1.81e-2	2.34e-2	2.72e-2	2.83e-2	2.84e-2	2.84e-2	2.84e-2
	Sobol	2.25e-3	3.46e-3	4.73e-3	5.31e-3	5.38e-3	5.38e-3	5.38e-3
2^{10}	Random	1.28e-2	1.61e-2	1.83e-2	1.90e-2	1.91e-2	1.91e-2	1.91e-2
	LHS	9.03e-3	1.17e-2	1.36e-2	1.41e-2	1.42e-2	1.42e-2	1.42e-2
	Sobol	5.63e-4	8.69e-4	1.24e-3	1.42e-3	1.52e-3	1.52e-3	1.52e-3
2^{12}	Random	6.38e-3	8.02e-3	9.16e-3	9.51e-3	9.53e-3	9.53e-3	9.53e-3
	LHS	4.51e-3	5.82e-3	6.77e-3	7.07e-3	7.08e-3	7.08e-3	7.08e-3
	Sobol	1.41e-4	2.28e-4	3.35e-4	3.91e-4	6.40e-4	6.40e-4	6.40e-4
a	Sobol	1.00	0.98	0.95	0.94	0.77	0.77	0.77

Table 6. The same as Table 5, but for the weights (B).