STRONG TRACTABILITY OF MULTIVARIATE INTEGRATION USING QUASI–MONTE CARLO ALGORITHMS

XIAOQUN WANG

ABSTRACT. We study quasi–Monte Carlo algorithms based on low discrepancy sequences for multivariate integration. We consider the problem of how the minimal number of function evaluations needed to reduce the worst-case error from its initial error by a factor of $\varepsilon$ depends on $\varepsilon^{-1}$ and the dimension $s$. Strong tractability means that it does not depend on $s$ and is bounded by a polynomial in $\varepsilon^{-1}$. The least possible value of the power of $\varepsilon^{-1}$ is called the $\varepsilon$-exponent of strong tractability. Sloan and Woźniakowski established a necessary and sufficient condition of strong tractability in weighted Sobolev spaces, and showed that the $\varepsilon$-exponent of strong tractability is between 1 and 2. However, their proof is not constructive.

In this paper we prove in a constructive way that multivariate integration in some weighted Sobolev spaces is strongly tractable with $\varepsilon$-exponent equal to 1, which is the best possible value under a stronger assumption than Sloan and Woźniakowski’s assumption. We show that quasi–Monte Carlo algorithms using Niederreiter’s $(t, s)$-sequences and Sobol sequences achieve the optimal convergence order $O(N^{-1+\delta})$ for any $\delta > 0$ independent of the dimension with a worst case deterministic guarantee (where $N$ is the number of function evaluations). This implies that strong tractability with the best $\varepsilon$-exponent can be achieved in appropriate weighted Sobolev spaces by using Niederreiter’s $(t, s)$-sequences and Sobol sequences.

1. INTRODUCTION

Multivariate integration is a standard field of application of Monte Carlo (MC) and quasi–Monte Carlo (QMC) methods. Consider the integral of $f(x)$ over the $s$-dimensional unit cube

$$I_s(f) = \int_{[0,1]^s} f(x) dx.$$

Dimensions in the hundreds or even thousands occurs in many applications, in particular, in physics and in mathematical finance [4, 16, 21]. For large $s$, the typical approach for approximating $I_s(f)$ is to use MC or QMC algorithms:

$$Q_{N,s}(f) = \frac{1}{N} \sum_{n=0}^{N-1} f(x_n).$$

Received by the editor March 7, 2001 and, in revised form, August 16, 2001.

2000 Mathematics Subject Classification. Primary 65C05, 65D30, 68Q25.

Key words and phrases. Quasi–Monte Carlo methods, low discrepancy sequences, tractability, strong tractability, multidimensional integration.

Supported by the NSF of China Grants 79970120 and 10001021.
MC methods use independent random samples from the uniform distribution on 
\([0, 1]^s\) and have asymptotic convergence rate \(O(N^{-1/2})\) independent of the dimension. In QMC methods, the points are chosen in some deterministic way. The classical QMC error bound is the following Koksma-Hlawka inequality \([13]\):

\[
|I_s(f) - Q_N,s(f)| \leq V^*(f)D^*_N(\{x_n\}),
\]

where \(D^*_N(\{x_n\})\) is the star discrepancy and \(V^*(f)\) is the variation of \(f\) in the sense of Hardy and Krause. Several methods for constructing sequences with star discrepancy (for the first \(N\) points of the sequence)

\[
D^*_N \leq C(s)\log s N + O\left(\frac{\log N}{N}\right)
\]

have been proposed by Halton \([6]\), Sobol \([19]\), Faure \([5]\), Niederreiter \([12, 13]\) and Niederreiter and Xing \([14]\). Sequences for which the star discrepancy satisfies the bound \((2)\) are called low discrepancy sequences. Hence, the asymptotic convergence rate of QMC algorithms based on low discrepancy sequences can be \(O(N^{-1} \log^2 N)\) for infinite sequences or \(O(N^{-1} \log^{s-1} N)\) for finite point sets, which is asymptotically much better than the order of MC for fixed dimension \(s\). But when \(s\) is large, the factor \((\log N)^s\) becomes huge, and it takes impractically large samples before the asymptotic is relevant. We cannot afford to sample the function so many times. However, empirical tests on integrals from finance have shown that QMC consistently beats MC by wide margins \([16, 21]\).

It is a challenging problem for theory to understand the apparent success of QMC algorithms for high and very high dimensional integrals. This problem has been extensively studied in recent years (see \([21]\) for references) and is related to the concepts of tractability and strong tractability of multivariate integration.

Tractability means that we can reduce the initial error by a factor of \(\varepsilon\) by using a number of function values which is polynomial in the dimension \(s\) and \(\varepsilon^{-1}\), whereas strong tractability means that we have only a polynomial in \(\varepsilon^{-1}\) that is independent of \(s\). The least possible value (or the infimum) of the power of \(\varepsilon^{-1}\) is called the \(\varepsilon\)-exponent of strong tractability. These concepts will be explained in detail later.

Tractability of multivariate integration has been studied in a number of papers \([9, 15, 17, 18, 24]\). Sloan and Woźniakowski \([17]\) established a necessary and sufficient condition of strong tractability in weighted Sobolev spaces and showed that the \(\varepsilon\)-exponent of strong tractability is between 1 and 2. However, their proof is not constructive. That is, we only known about the existence of such strongly polynomial-time QMC algorithms, but do not know how to find such algorithms. Therefore, it is an interesting problem to construct QMC algorithms that achieve an error bound appropriate to strong tractability.

In this paper we consider weighted classes of spaces where the weights \(\gamma_i\) moderate the behavior of functions with respect to the successive variables. The main results of this paper state that QMC algorithms based on Niederreiter’s \((t, s)\)-sequences and Sobol sequences are strongly tractable for some weighted Sobolev spaces with \(\varepsilon\)-exponent equal to 1, which is the best possible value if the weights \(\gamma_i\) go to zero sufficiently fast. In Section 2, we give background on tractability and on weighted Sobolev spaces. In Section 3, we show that QMC algorithms based on Niederreiter’s \((t, s)\)-sequences and Sobol sequences achieve an error bound independent of the dimension \(s\) and of order \(O(N^{-1+\delta})\) for any \(\delta > 0\) with a worst
case deterministic guarantee. Some open problems are presented in the concluding section.

2. Background on tractability

2.1. Tractability and strong tractability. We start by defining tractability and strong tractability more precisely (see [17, 23]). Let \( F \) be a normed space of integrable functions defined on \([0, 1]^s\); the norm in \( F \) is denoted by \( \| \cdot \|_F \). Define the worst case error of the algorithm \( Q_{N,s}(f) \) by its worst case performance over the unit ball of \( F \),

\[
e(Q_{N,s}(f), F) = \sup_{f \in F, \|f\|_F \leq 1} |I_s(f) - Q_{N,s}(f)|.
\]

For \( N = 0 \), we formally set \( Q_{0,s}(f) = 0 \), and

\[
e(0, F) = \sup_{f \in F, \|f\|_F \leq 1} |I_s(f)| = \|I_s\|
\]

is the initial error in multivariate integration without sampling the function. We would like to reduce the initial error by a factor of \( \varepsilon \), where \( \varepsilon \in (0, 1) \). We are looking for the smallest \( N = N(\varepsilon, F) \) for which there exists an algorithm \( Q_{N,s}(f) \) such that \( e(Q_{N,s}(f), F) \leq \varepsilon e(0, F) \). That is,

\[
N(\varepsilon, F) = \min\{N : \exists Q_{N,s}(f) \text{ such that } e(Q_{N,s}(f), F) \leq \varepsilon e(0, F)\}.
\]

We are interested in how fast \( N(\varepsilon, F) \) grows with \( \varepsilon^{-1} \) and \( s \).

We say that multivariate integration in a space \( F \) is tractable if there exist nonnegative constants \( C, p \) and \( q \) such that

\[
(3) \quad N(\varepsilon, F) \leq C\varepsilon^{-p}s^q, \forall \varepsilon \in (0, 1) \text{ and } \forall s \geq 1.
\]

The infima of the numbers \( p \) and \( q \) for which \( (3) \) holds are called the \( \varepsilon \)-exponent and \( s \)-exponent of tractability. If the inequality \( (3) \) holds with \( q = 0 \), then we say that multivariate integration is strongly tractable in \( F \), and the infimum of \( p \) is called the \( \varepsilon \)-exponent of strong tractability and is denoted by \( p^* \). In this case, the number of samples is independent of \( s \) and depends polynomially on \( \varepsilon^{-1} \).

Tractability and strong tractability of multivariate integration are obviously dependent on the class of functions. Multivariate integration is intractable for many natural spaces (see [15]). For instance, a classical result of Bakhvalov [2] states that the complexity of integration over the \( s \)-dimensional unit cube of \( r \) (\( r > 0 \)) times continuously differentiable functions is \( O(\varepsilon^{-s/r}) \). It depends exponentially on the dimension \( s \). Note that the complexity of integration is the minimal worst case cost of computing an approximation with error at most \( \varepsilon \), and is directly related to the minimal number \( N(\varepsilon, F) \) defined above. It is natural to ask for which spaces of functions the problem of multivariate integration is tractable or strongly tractable. With suitable choices of spaces \( F \) and by introducing weights that model the behavior of successive variables, we can sometimes break intractability or even achieve strong tractability.

2.2. Weighted Sobolev spaces. Let us introduce some notation. Let \( x = (x^{(1)}, \ldots, x^{(s)}) \) and \( D = \{1, 2, \ldots, s\} \). For any subset \( u \subseteq D \), let \( |u| \) denotes its cardinality. For the vector \( x \in [0, 1]^s \), let \( x^{(u)} \) be the \(|u|\)-dimensional vector containing the components of \( x \) with superscripts in \( u \), and let \( dx^{(u)} = \prod_{i \in u} dx^{(i)} \).
Furthermore, let \([0,1)^n\) denote the \(|u|\)-dimensional unit cube involving the components in \(u\). By \((x^{(i)},1)\) we mean the vector from \([0,1)^s\) with the same components as \(x\) for superscripts in \(u\) and the rest of the components being replaced by 1.

We consider functions whose dependence on successive variables is increasingly limited. To quantify this, following Sloan and Woźniakowski [17], take a sequence \(\{\gamma_i\}\) such that

\[\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_i \geq \cdots > 0,\]

and let \(\gamma_0 = 1\) and \(\gamma_u = \prod_{i \in u} \gamma_i\) for any nonempty subset \(u \subseteq D\). We associate the first variable \(x^{(1)}\) to \(\gamma_1\), the second variable \(x^{(2)}\) to \(\gamma_2\), the variables \(x^{(u)}\) to \(\gamma_u\) and so on. Consider the reproducing kernel Hilbert space \(H(K_{s,\gamma})\) with the following reproducing kernel:

\[K_{s,\gamma}(x,y) = \prod_{i=1}^s \left[ 1 + \gamma_i \min(1-x^{(i)},1-y^{(i)}) \right].\]

The theory of reproducing kernel Hilbert spaces can be found in [1]. The inner product in the space \(H(K_{s,\gamma})\) is given by

\[
\langle f,g \rangle_{H(K_{s,\gamma})} = \sum_{u \subseteq D} \gamma_u^{-1} \int_{[0,1)^s} \frac{\partial^{|u|}}{\partial x_u} f(x^{(u)},1) \frac{\partial^{|u|}}{\partial x_u} g(x^{(u)},1) dx^{(u)}
\]

for \(f, g \in H(K_{s,\gamma})\). The induced norm is

\[||f||_{H(K_{s,\gamma})} = \left\{ \sum_{u \subseteq D} \gamma_u^{-1} \int_{[0,1]^s} \left| \frac{\partial^{|u|}}{\partial x_u} f(x^{(u)},1) \right|^2 dx^{(u)} \right\}^{1/2}.\]

Obviously, \(H(K_{s,\gamma})\) is a tensor product space \(H(K_{s,\gamma}) = \bigotimes_{i=1}^s H_i\), where the \(H_i\) are 1-dimensional Hilbert spaces of functions with the reproducing kernels

\[K_{1,\gamma_i}(x,y) = 1 + \gamma_i \min(1-x,1-y).\]

If \(\gamma_i = 1\) for all \(i\), we obtain the classical (tensor product) Sobolev space. If \(\gamma_i\) is decreasing with \(i\), then we obtain weighted Sobolev spaces \(H(K_{s,\gamma})\) consisting of functions with diminishing dependence on the \(i\)-th variable. Such spaces of functions provide models for problems with decreasing importance of successive variables.

2.3. Related work. The decay of the weights \(\gamma_i\) determines whether tractability or strong tractability holds. Sloan and Woźniakowski [17] proved that strong tractability holds iff

\[\sum_{i=1}^\infty \gamma_i < \infty.\]

Furthermore, they showed that under the assumption (5) the strong \(\varepsilon\)-exponent \(p^*\) belongs to [1,2]. But their result is not constructive. Their arguments involve averaging over all choices of the quadrature points in \(Q_{N,s}(f)\). In Hickernell and Woźniakowski [9], it is proved that \(\sum_{i=1}^\infty \gamma_i^{1/2} < \infty\) yields the best possible exponent \(p^* = 1\). Again this result is not constructive. There is a constructive proof that \(p^* = 1\) in Wasilkowski and Woźniakowski [22] under the more restrictive assumption that \(\sum_{i=1}^\infty \gamma_i^{1/3} < \infty\). But this proof is based on weighted tensor product algorithms, which are generalizations of Smolyak’s construction and do not use low discrepancy
3. Strong tractability of multivariate integration

Observe that for known low discrepancy sequences their star discrepancies have the bound (2), where the constant $C(s)$ and the implied constant in the Landau symbol depend on the dimension $s$ and may increase superexponentially with $s$ for some known low discrepancy sequences [13]. Such a bound is enough for obtaining the convergence rate for fixed dimension, but do not say anything about tractability.

If we want to investigate tractability or strong tractability, we need to consider QMC algorithms with arbitrarily large dimension $s$ and need to know how the discrepancies of the lower dimensional projections on $[0,1]^u$ of the underlying low discrepancy sequences depend on the dimension $s$ and on the indices $i$ in the subset $u \subseteq D$.

3.1. QMC algorithms using Niederreiter’s $(t, s)$-sequence. The most powerful known methods for the construction of low discrepancy point sets and sequences are based on the theory of $(t, m, s)$-nets and $(t, s)$-sequences. We give a brief introduction to these concepts. We refer to Niederreiter [13] for details.

An elementary interval in base $b$ ($b \geq 2$) is an interval $E$ in $[0,1)^s$ of the form

$$E = \prod_{i=1}^{s} \left[ \frac{a_i}{b^{d_i}}, \frac{a_i + 1}{b^{d_i}} \right]$$

with integers $a_i, d_i \geq 0$ and $0 \leq a_i < b^{d_i}$ for $1 \leq i \leq s$.

Let $0 \leq t \leq m$ be integers. A $(t, m, s)$-net in base $b$ is a point set of $b^m$ points in $[0,1)^s$ such that every elementary interval $E$ in base $b$ of volume $b^{-m}$ contains exactly $b^t$ points.

Let $t \geq 0$ be an integer. A sequence $x_0, x_1, \ldots$ of points in $[0,1)^s$ is a $(t, s)$-sequence in base $b$ if for all integers $k \geq 0$ and $m > t$, the point set $\{x_n | kb^m \leq n < (k+1)b^m \}$ is a $(t, m, s)$-net in base $b$.

The parameter $t$ measures the regularity of the $(t, s)$-sequence. A smaller value of $t$ means stronger regularity properties of the sequence. Detailed information on upper bounds for the star discrepancy of $(t, s)$-sequences is provided in [13].

In the following investigation of strong tractability, we use the most interesting special case for the construction of $(t, s)$-sequences, due to Niederreiter [12, 13]. We choose the base $b$ to be a prime power $q$. We list all monic irreducible polynomials over the finite field $F_q$ in a sequence according to nondecreasing degree (polynomials with the same degree can be listed in any order) and let $p_1, p_2, \ldots, p_s$ be the first $s$ monic irreducible polynomials. The degree of $p_i$ is denoted by $\deg(p_i)$. Put $\deg(p_i) = e_i$ for $1 \leq i \leq s$.

The generation of Niederreiter’s $(t, s)$-sequence in base $q$ can be briefly described as follows [12, 13]. For $1 \leq i \leq s$ and integers $j \geq 1$ and $0 \leq k < e_i$, consider the
Laurent series expansions

\[
\frac{x^k}{p_i(x)^j} = \sum_{r=0}^{\infty} a^{(i)}(j, k, r)x^{-r-1},
\]

and define the collection of coefficients \(c^{(i)}_{jr}\) by

\[
c^{(i)}_{jr} = a^{(i)}(Q + 1, k, r) \in F_q, \quad \text{for } 1 \leq i \leq s, j \geq 1, r \geq 0,
\]

where the integers \(Q\) and \(k\) satisfy \(j - 1 = Qe_i + k\) with \(0 \leq k < e_i\).

To generate the \(n\)-th point of an \(s\)-dimensional Niederreiter sequence, write \(n\) in the digit expansion in base \(q\), i.e.,

\[
n = \sum_{r=0}^{\infty} a_r(n)q^r, \quad n = 0, 1, \ldots
\]

where \(a_r(n) \in \{0, 1, \ldots, q - 1\}\) for \(r \geq 0\). Note that \(a_r(n) = 0\) for all sufficiently large \(r\). Put

\[
x^{(i)}_n = \sum_{j=1}^{\infty} y^{(i)}_{nj} q^{-j}, \quad \text{for } n \geq 0 \quad \text{and } 1 \leq i \leq s,
\]

with

\[
y^{(i)}_{nj} = \sum_{r=0}^{\infty} c^{(i)}_{jr} a_r(n), \quad \text{for } n \geq 0, 1 \leq i \leq s, \quad \text{and } j \geq 1.
\]

Note that the sum over \(r\) is a finite sum. The \(s\)-dimensional Niederreiter sequence (denoted by \(\mathcal{P}\)) is defined by

\[
x_n = \left( x^{(1)}_n, \ldots, x^{(s)}_n \right), \quad n = 0, 1, \ldots
\]

Here we give a form where the bijections \(\psi_r\) and \(\eta_{i,j}\) from Niederreiter [12, 13] are set equal to the identity. This choice is relatively practical for implementation purposes. According to Niederreiter [12, 13] the above construction yields a \((t_D, s)\)-sequence in base \(q\) with the quality parameter

\[
t_D = \sum_{i=1}^{s} \lceil \deg(p_i) - 1 \rceil.
\]

Note that the polynomials \(p_1, p_2, \ldots, p_s\) are not uniquely determined, but the quality parameter \(t_D\) is well defined. The quality parameter \(t_D\) measures the uniformity of the sequence \(\mathcal{P}\) and should be as small as possible. It is obvious that the use of the first \(s\) monic irreducible polynomials minimizes the quality parameter. The monic irreducible polynomials of a specific degree should be exhausted before moving to a higher degree. It has been proved that \(t_D\) is in the order of \(O(s \log s)\) (see [13]). The program to generate Niederreiter sequence is given in [3].

The Niederreiter sequence in base \(q\) constructed above has the following telescopic property: in order to obtain a sequence in dimension \(s\) \((s > 1)\), it suffices to add the last component \(x^{(s)}_n\) to the term of the \((s - 1)\)-dimensional sequence \((x^{(1)}_n, \ldots, x^{(s-1)}_n)\), i.e., with the previous \(s - 1\) components kept unchanged. Thus, the Niederreiter sequence can be constructed in arbitrarily large dimension, dimension by dimension. Furthermore, the \(i\)-th component of the Niederreiter sequence
depends on the $i$-th monic irreducible polynomial; this property allows us to investigate the quality of any lower dimensional projections of the Niederreiter sequence $\mathcal{P}$.

Let $u$ be a nonempty subset of $D$ and $\mathcal{P}_u$ be the projection of $\mathcal{P}$ on the $|u|$-dimensional unit cube $[0,1)^u$. Then $\mathcal{P}_u$ is a $|u|$-dimensional Niederreiter sequence in base $q$ based on the monic irreducible polynomials $p_i$ with indices $i$ in $u$. Thus, it is a $(t_u, |u|)$-sequence in base $q$ with the quality parameter

$$t_u = \sum_{i \in u} [\deg(p_i) - 1],$$

where the sum is taken over all the indices $i$ in $u$. The following lemma shows how to bound the star discrepancy of the sequence $\mathcal{P}_u$ explicitly in terms of the degrees of the monic irreducible polynomials $p_i$ with $i$ in the subset $u$.

**Lemma 1.** Let $\mathcal{P}_u$ be the projection on $[0,1)^u$ of the $s$-dimensional Niederreiter sequence $\mathcal{P}$ in prime power base $q$. The star discrepancy of the first $N$ terms of $\mathcal{P}_u$ satisfies

$$ND_N^*(\mathcal{P}_u) \leq \prod_{i \in u} \left[q^{\deg(p_i) \log_q(qN)}\right],$$

for all $N \geq 1$ and any nonempty set $u \subseteq D$.

**Proof.** We consider the case for base $q \geq 3$. Since $\mathcal{P}_u$ is a $(t_u, |u|)$-sequence in base $q$, for $N \geq q^{t_u}$ we have the following explicit star discrepancy bound (see [13], Theorem 4.12):

$$ND_N^*(\mathcal{P}_u) \leq \frac{q - 1}{2} q^{|u|} \sum_{i=1}^{\lfloor \log_q N \rfloor} \binom{|u| - 1}{i - 1} \binom{k + 1 - t_u}{i} \left[\frac{q}{2}\right]^{i-1}$$

$$+ \frac{1}{2} q^{|u|} \sum_{i=0}^{\lfloor \log_q N \rfloor} \binom{|u| - 1}{i} \binom{k + 1 - t_u}{i} + (k - t_u) \left[\frac{q}{2}\right]^{i}$$

$$=: S_1 + S_2$$

where $k$ is the largest integer with $q^k \leq N$, i.e., $k = \lfloor \log_q N \rfloor$, and $t_u$ is the quality parameter of $\mathcal{P}_u$ defined by (7).

Note that for $1 \leq i \leq |u|$ we have an upper bound for the binomial coefficients:

$$\binom{k + 1 - t_u}{i} \leq \frac{(k + 1)^i}{i!} \leq (k + 1)^{|u|}.$$
Thus for the first term $S_1$ we have

$$S_1 \leq \frac{q-1}{2} q^{t_u} \sum_{i=1}^{[u]} \left( \frac{|u| - 1}{i - 1} \right) (k + 1)^{|u|} \left\lfloor \frac{q}{2} \right\rfloor^{i-1}$$

$$\leq \frac{q-1}{2} q^{t_u} \log_q N + 1 |u| \sum_{i=1}^{[u]} \left( \frac{|u| - 1}{i - 1} \right) \left\lfloor \frac{q}{2} \right\rfloor^{i-1}$$

$$= \frac{q-1}{2} q^{t_u} [\log_q (qN)]^{|u|} \left( \left\lfloor \frac{q}{2} \right\rfloor - 1 \right)^{|u| - 1}$$

$$\leq \frac{q-1}{2} q^{t_u + |u| - 1} [\log_q (qN)]^{|u|}$$

$$= \frac{q-1}{2q} \prod_{i \in u} \left( q^{\deg(p_i)} \log_q (qN) \right).$$

Similarly, for the second term $S_2$ we have

$$S_2 \leq \frac{1}{q} \prod_{i \in u} \left( q^{\deg(p_i)} \log_q (qN) \right).$$

Therefore, for $N \geq q^{t_u}$ we have

$$ND_N^*(P_u) \leq S_1 + S_2 \leq \frac{q+1}{2q} \prod_{i \in u} \left( q^{\deg(p_i)} \log_q (qN) \right) < \prod_{i \in u} \left( q^{\deg(p_i)} \log_q (qN) \right).$$

For $1 \leq N < q^{t_u}$, this bound is also valid. Since the star discrepancy of any point set is between 0 and 1, it follows that $0 \leq D_N^*(P_u) \leq 1$ and

$$ND_N^*(P_u) \leq N \leq q^{t_u} = \prod_{i \in u} \left( q^{\deg(p_i) - 1} \right) < \prod_{i \in u} \left( q^{\deg(p_i) \log_q (qN)} \right).$$

The case for base $q = 2$ can be proved analogously using the star discrepancy bound for even base (see Theorem 4.13 in [13]). This completes the proof. 

The bound in Lemma 1 involves the degrees of monic irreducible polynomials. In the study of tractability of multivariate integration using the Niederreiter sequence $\mathcal{P}$, the dimension $s$ can be arbitrarily large; thus the number of monic irreducible polynomials $p_i$ needed and their degrees can also be arbitrarily large. By estimating the degree of the $i$-th monic irreducible polynomial $p_i$, a further bound on the star discrepancy can be obtained.

Let $J_q(n)$ be the number of monic irreducible polynomials over the finite field $F_q$ of degree $\leq n$, with $J_q(0) = 0$. Let $I_q(n)$ be the number of monic irreducible polynomials over $F_q$ of degree $n$. For any prime power $q$, we have the relation [13]

$$J_q(n) \geq \frac{1}{n} q^n, \text{ for all } n \geq 1. \tag{8}$$

This can be shown by using the following explicit formula for the number $I_q(n)$:

$$I_q(n) = \frac{1}{n} \sum_{d | n} \mu(n/d) q^d, \text{ for all } n \geq 1,$$

where $\mu$ is the M"obius function (see [10]).

The next lemma establishes the relationship of the degree of the $i$-th monic irreducible polynomial $p_i(x)$ with its index $i$. A similar method was used in [13] to estimate the upper bound for the quality parameter $t_D$ defined in (6).
Lemma 2. For any prime power $q$, the degree of the $i$-th monic irreducible polynomial $p_i(x)$ over the finite field $F_q$ can be bounded by

\[ \deg(p_i) \leq \log_q i + \log_q \log_q(i + q) + 2, \quad \text{for } i = 1, 2, \ldots. \] (9)

Proof. The first $q$ monic irreducible polynomials are linear polynomials $p_i(x) = x - b_i$ for $1 \leq i \leq q$, where $b_1, \ldots, b_q$ are distinct elements of $F_q$. Thus for $1 \leq i \leq q$, we have \( \deg(p_i) = 1 \). The bound (9) holds in this case.

For $i > q$, put $k = \lceil \log_q i + \log_q \log_q(i + q) \rceil + 2$. Then

\[ \log_q i + \log_q \log_q(i + 1) < k \leq \log_q i + \log_q \log_q i + 2. \]

If either $q = 2, y \geq 4$, or $q \geq 3, y > 1$, then

\[ (q - 1)y \geq \log_q y + 2. \]

With $y = \log_q i$, we obtain that if either $q = 2, i \geq 16$, or $q \geq 3, i > q$, then

\[ q \log_q i \geq \log_q i + \log_q \log_q i + 2 \geq k. \]

In these cases, by taking the logarithm on both the left-hand and right-hand side of the above inequality, we have

\[ \log_q \log_q i + 1 \geq \log_q k. \]

Thus

\[ k > \log_q i + \log_q \log_q i + 1 \geq \log_q i + \log_q k. \]

It follows that

\[ \frac{1}{k} q^k > i. \]

According to (8), we immediately obtain

\[ J_q(k) \geq \frac{1}{k} q^k > i. \]

This means that the number of monic irreducible polynomials over $F_q$ of degree $\leq k$ exceeds the number $i$. Thus the degree of the $i$-th monic irreducible polynomial is not larger than $k$, i.e.,

\[ \deg(p_i) \leq k \leq \log_q i + \log_q \log_q i + 2 < \log_q i + \log_q \log_q(i + q) + 2. \]

For the remaining case where $q = 2, 3 \leq i \leq 15$, the bound (9) can be checked directly by using tables of monic irreducible polynomials (see, for instance, [10]). \(
\)

Combining Lemmas 1 and 2, we obtain the upper bound on the star discrepancy of $P_u$ in terms of $i$ with these indices $i$ in the subset $u$.

Lemma 3. The star discrepancy of the first $N$ terms of $P_u$ satisfies

\[ D_N(P_u) \leq \frac{1}{N} \prod_{i \in u} [C_1 i \log(i + q) \log(qN)], \]

for all $N \geq 1$ and any nonempty set $u \subseteq D$, where $C_1$ is a positive constant independent of the subset $u$ and the dimension $s$. 
Proof. From Lemma 2 we have for all $i \geq 1$
\[ \deg(p_i) \leq \log_q i + \log_q (i + q) + 2. \]
Thus
\[ q^{\deg(p_i)} \leq q^2 i \log_q (i + q). \]
It follows from Lemma 1 that
\[ \mathcal{N}D_N^*(P_u) \leq \prod_{i \in u} \left[ q^2 i \log_q (i + q) \log(qN) \right] = \prod_{i \in u} [C_1 i \log(i + q) \log(qN)], \]
for all $N \geq 1$, where $C_1 = q^2 / (\log q)^2$ is independent of $u$ and $s$.

Now we are ready to study the tractability of multivariate integration using the Niederreiter sequence $P$. Consider the weighted Sobolev space $H(K_{s, \gamma})$ defined in the previous section (see (4)). In order to study the tractability in $H(K_{s, \gamma})$, it suffices to consider the corresponding integration error over the unit ball of $H(K_{s, \gamma})$.

Let
\[ B_{s, \gamma} = \{ f \in H(K_{s, \gamma}) : ||f||_{H(K_{s, \gamma})} \leq 1 \} \]
be the unit ball of $H(K_{s, \gamma})$. According to Sloan and Woźniakowski [17], the worst case error of a QMC algorithm $Q_{N,s}(f)$ over $B_{s, \gamma}$ is equal to the weighted version of the generalized $L_2$-star discrepancy of the point set consisting of the quadrature points. That is,
\begin{equation}
(10) \quad e(Q_{N,s}(f), H(K_{s, \gamma})) = \sup_{f \in B_{s, \gamma}} |I_s(f) - Q_{N,s}(f)| = \mathcal{D}_N^*(P, \gamma),
\end{equation}
where
\begin{equation}
(11) \quad \mathcal{D}_N^*(P, \gamma) = \left\{ \sum_{\emptyset \neq u \subseteq D} \gamma_u \int_{[0,1]^u} \left[ \text{Disc} \left( (x^{(u)}, 1); P \right) \right]^2 dx^{(u)} \right\}^{1/2}
\end{equation}
is the weighted version of the generalized $L_2$-star discrepancy of the first $N$ points of the Niederreiter sequence $P$ and Disc$(x; P)$ is defined by
\[ \text{Disc} (x; P) = \left| \frac{\{n : x_n \in [0, x]\}}{N} - x^{(1)} x^{(2)} \ldots x^{(s)} \right|. \]

Note that the square of the weighted version of the generalized $L_2$-star discrepancy is the weighted sum of $\int_{[0,1]^u} \left[ \text{Disc} \left( (x^{(u)}, 1); P \right) \right]^2 dx^{(u)}$, which is the square of the traditional $L_2$-star discrepancy of $P_u$. So the weighted version of the generalized $L_2$-star discrepancy of a point set captures the uniformity of its lower dimensional projections ([7]).

From (10) we see that the problem of tractability becomes one of determining how small the weighted generalized $L_2$-star discrepancy can be made using $N$ sample points.

**Theorem 4.** Assume that the QMC algorithm $Q_{N,s}(f)$ employs the first $N$ points of Niederreiter’s $(t,s)$-sequence $P$ in prime power base $q$. If the weights $\gamma_i$ satisfy
\begin{equation}
(12) \quad \sum_{i=1}^{\infty} \gamma_i^{1/2} i \log i < \infty,
\end{equation}
then for any \( \delta > 0 \), there exists a constant \( C_* \), independent of \( s \), such that
\[
\sup_{f \in B_{s, \gamma}} |I_s(f) - Q_N(f)| = D_N^s(\mathcal{P}, \gamma) \leq C_* N^{-1+\delta}, \text{ for all } N \geq 1.
\]
Hence, multivariate integration in the weighted Sobolev space \( H(K_{s, \gamma}) \) is strongly tractable, and the QMC algorithm using Niederreiter’s \((t, s)\)-sequence \( \mathcal{P} \) achieves the strong tractability with the optimal \( \varepsilon \)-exponent \( p^* = 1 \).

Proof. Note that the (traditional) \( L_2 \)-star discrepancy of a point set is no larger than its (\( L_\infty \)-) star discrepancy [11]. Thus
\[
\int_{[0,1]^s} \left[ \text{Disc} \left( (x(u), 1); \mathcal{P} \right) \right]^2 \, dx(u) \leq \left[ D_N^s(\mathcal{P}_u) \right]^2.
\]

From Lemma 3 we have a bound for the star discrepancy of \( \mathcal{P}_u \):
\[
D_N^s(\mathcal{P}_u) \leq \frac{1}{N} \prod_{i \in u} [C_i \log(i + q) \log(qN)], \text{ for all } N \geq 1.
\]

Based on these and the definition of the weighted version of generalized \( L_2 \)-star discrepancy (11), it follows that
\[
[D_N^s(\mathcal{P}, \gamma)]^2 \leq \frac{1}{N^2} \sum_{\emptyset \neq u \subseteq \mathcal{D}} \gamma_u \prod_{i \in u} [C_i \log(i + q) \log(qN)]^2, \text{ for all } N \geq 1.
\]

For any given \( \delta > 0 \), under the assumption (12) there exists an integer \( \kappa \) such that
\[
\sum_{i = \kappa + 1}^{\infty} \gamma_i^{1/2} i \log(i + q) < \frac{\delta}{C_1}.
\]

Now define a constant
\[
A = \min \left\{ 1, \delta \left( C_1 \sum_{i = 1}^{\kappa} \gamma_i^{1/2} i \log(i + q) \right)^{-1} \right\}
\]
and a sequence
\[
w_i = \begin{cases} A^2 \gamma_i, & 1 \leq i \leq \kappa, \\ \gamma_i, & i > \kappa. \end{cases}
\]

Note that the constant \( A \) and the sequence \( \{w_i\} \) just defined are independent of the dimension \( s \). Because only the first \( \kappa \) weights among \( \{\gamma_i\} \) may have been changed and \( A \leq 1 \), for any nonempty subset \( u \subseteq \mathcal{D} \) it follows that
\[
w_u := \prod_{i \in u} w_i \geq A^{2\kappa} \gamma_u.
\]

Thus
\[
\sum_{i = 1}^{\infty} w_i^{1/2} i \log(i + q) = A \sum_{i = 1}^{\kappa} \gamma_i^{1/2} i \log(i + q) + \sum_{i = \kappa + 1}^{\infty} \gamma_i^{1/2} i \log(i + q) < \frac{2\delta}{C_1}.
\]
Therefore, from (13) and (14) we have for all \( N \geq 1 \)
\[
|D_N^*(P, \gamma)|^2 \leq \frac{1}{N^2} \sum_{\emptyset \neq u \subseteq D} \gamma_u \prod_{i \in u} [C_i i \log(i + q) \log(qN)]^2
\]
\[
\leq \frac{1}{N^2 A^{2s}} \sum_{\emptyset \neq u \subseteq D} w_u \prod_{i \in u} [C_i i \log(i + q) \log(qN)]^2
\]
\[
< \frac{1}{N^2 A^{2s}} \prod_{i=1}^s \left[ 1 + w_i [C_i i \log(i + q) \log(qN)]^2 \right]
\]
\[
= \frac{1}{N^2 A^{2s}} \exp \left( \sum_{i=1}^s \log \left( 1 + w_i [C_i i \log(i + q) \log(qN)]^2 \right) \right)
\]
\[
\leq \frac{1}{N^2 A^{2s}} \exp \left( \sum_{i=1}^s w_i^{1/2} [C_i i \log(i + q) \log(qN)] \right)
\]
\[
< \frac{1}{N^2 A^{2s}} \exp \left( \sum_{i=1}^\infty w_i^{1/2} [C_i i \log(i + q) \log(qN)] \right)
\]
\[
= \frac{1}{N^2 A^{2s}} \exp \left( C_1 \log(qN) \sum_{i=1}^\infty w_i^{1/2} i \log(i + q) \right)
\]
\[
< \frac{1}{N^2 A^{2s}} \exp (2\delta \log(qN))
\]
\[
= \frac{q^{2\delta}}{N^{2s}} N^{2\delta}
\]
\[
= C_2^* N^{-2+2\delta}
\]
uniformly for \( s \geq 1 \), where \( C_2^* = q^\delta / A^\epsilon \). In passing we have used the fact that \( \log(1 + x^2) \leq x \) for \( x \geq 0 \), and the inequality (15).

Note that the initial error in multivariate integration in the space \( H(K_s, \gamma) \) is
\[
\epsilon(0, H(K_s, \gamma)) = ||I_s|| = \prod_{j=1}^s \left( 1 + \frac{1}{3} \gamma_j \right)^{1/2},
\]
which is uniformly bounded in \( s \) under the assumption (12). From this fact and the bound for the weighted discrepancy above, we can assert that the \( \epsilon \)-exponent of strong tractability is 1. This concludes the proof. \( \Box \)

3.2. QMC algorithms using Sobol sequences. In this subsection we study tractability using Sobol sequences in a similar way. The generation of Sobol sequence can be briefly described as follows (see [19]). We list all primitive polynomials over the field \( F_2 \) in a sequence arranged in order of nondecreasing degrees. Let \( P_1, P_2, \ldots, P_s \) be the first \( s \) primitive polynomials. Each component of an \( s \)-dimensional Sobol sequence is based on one primitive polynomial. Let
\[
P_i(x) = x^d + a_1 x^{d-1} + \cdots + a_{d-1} x + 1
\]
be the \( i \)-th primitive polynomial of degree \( d = \deg(P_i) \), where each \( a_j \) is 0 or 1. We use its coefficients to define a sequence \( \{m_j\} \) by the recurrence
\[
m_j = 2a_1 m_{j-1} \oplus 2^2 a_2 m_{j-2} \oplus \cdots \oplus 2^{d-1} a_{d-1} m_{j-d+1} \oplus 2^d m_{j-d} \oplus m_{j-d},
\]
for \( j = d + 1, d + 2, \ldots \), where \( \oplus \) is the bit-by-bit exclusive-or operator. The initial values of \( m_1, m_2, \ldots, m_d \) can be chosen freely provided that each \( m_j \) is odd and \( m_j < 2^j \). Then we define the direction numbers by

\[
v_j = m_j/2^j, \quad j = 1, 2, \ldots.
\]

The \( i \)-th component of the \( s \)-dimensional Sobol sequence (denoted by \( S \)) is defined by

\[
x_n^{(i)} = c_1 v_1 \oplus c_2 v_2 \oplus c_3 v_3 \oplus \cdots, \quad n = 0, 1, \ldots,
\]

where \( \cdots c_3 c_2 c_1 \) is the binary representation of \( n \). The \( n \)-th Sobol point is defined by \( x_n = (x_n^{(1)}, \ldots, x_n^{(s)}) \). The \( s \)-dimensional Sobol sequence \( S \) is a \((T_D, s)\)-sequence in base 2 with the quality parameter (see [19])

\[
T_D = \sum_{i=1}^{s} (\deg(P_i) - 1).
\]

Note that the Sobol sequence is different from the Niederreiter sequence in base 2, since the former is based on primitive polynomials, whereas the latter is based on irreducible polynomials. A comparison of their quality parameters (defined in (6)) is given in [13] and shows that \( T_D = t_D \) for \( 1 \leq s \leq 7 \), and \( T_D > t_D \) for all \( s \geq 8 \). Thus the Sobol sequence is worse theoretically when the dimension exceeds seven, and so the tractability of multivariate integration using the Sobol sequence is not a direct consequence of the tractability using the Niederreiter sequence.

The Sobol sequence also has the telescopic property. For a nonempty subset \( u \) of \( D \), let \( S_u \) be the projection of the \( s \)-dimensional Sobol sequence \( S \) on \([0, 1]^s \). Then \( S_u \) is a \(|u|\)-dimensional Sobol sequence based on the primitive polynomials \( P_i \) with indices \( i \) in \( u \). Therefore, \( S_u \) is a \((T_u, |u|)\)-sequence in base 2 with the quality parameter

\[
(16) \quad T_u = \sum_{i \in u} (\deg(P_i) - 1).
\]

The star discrepancy of \( S_u \) can be bounded in terms of \( i \) with these \( i \) in \( u \). The following lemma is parallel to Lemma 3 for Niederreiter sequences.

**Lemma 5.** The star discrepancy of the first \( N \) points of \( S_u \) satisfies

\[
D_N^*(S_u) \leq \frac{1}{N} \prod_{i \in u} \left[ C_2 i \log_2(i + 1) \log_2(i + 3) \log(2N) \right],
\]

for all \( N \geq 1 \) and \( \emptyset \neq u \subseteq D \), where \( C_2 \) is a positive constant independent of the subset \( u \) and the dimension \( s \).

**Proof.** Since \( S_u \) is a \(|u|\)-dimensional Sobol sequence based on the primitive polynomials \( P_i \) with \( i \) in \( u \), for \( N \geq 2^{|u|−1+T_u} \) we have (see [19])

\[
(17) \quad ND_N^*(S_u) \leq 2^T_u \sum_{i=0}^{|u|−1} \binom{k - T_u + 1}{i + 1} + 2^{T_u} - 1,
\]

where \( k \) is the largest integer with \( 2^k \leq N \), i.e., \( k = \lfloor \log_2 N \rfloor \), and \( T_u \) is the quality parameter defined by (16).
Note that for $0 \leq i \leq |u| - 1$ the binomial coefficients have the upper bound
\[
\binom{k - T_u + 1}{i + 1} \leq (k + 1)^{|u|} \leq \left\lfloor \log_2 N + 1 \right\rfloor^{|u|} = \left\lfloor \log_2 (2N) \right\rfloor^{|u|}.
\]
Therefore, from (17) and from the formula (16) for the quality parameter $T_u$, it follows that for $N \geq 2^{|u| - 1 + T_u}$,
\[
ND_N^* (S_u) \leq 2T_u \sum_{i=0}^{|u|-1} \left\lfloor \log_2 (2N) \right\rfloor^{|u|} + 2^{T_u} - 1 < 2^{T_u} (|u| + 1) \left\lfloor \log_2 (2N) \right\rfloor^{|u|} \leq 2^{T_u + |u|} \left\lfloor \log_2 (2N) \right\rfloor^{|u|} = \prod_{i \in u} \left[ 2^{\deg(P_i)} \log_2 (2N) \right].
\]
This bound is also valid for $1 \leq N < 2^{|u| - 1 + T_u}$, since $0 \leq D_N^* (S_u) \leq 1$, and thus
\[
ND_N^* (S_u) \leq N < 2^{|u| - 1 + T_u} = \frac{1}{2} \prod_{i \in u} \left[ 2^{\deg(P_i)} \right] < \prod_{i \in u} \left[ 2^{\deg(P_i)} \log_2 (2N) \right].
\]
Note that the bound for $ND_N^* (S_u)$ can also be obtained by using the star discrepancy bound for a $(t, s)$-sequence in even base (take the base to be 2, see Theorem 4.13 in [13]).

The bound above for $ND_N^* (S_u)$ involves the degrees of the primitive polynomials. According to Sobol [20] (with a modification), the degree $\deg(P_i)$ can be bounded by
\[
\deg(P_i) \leq \log_2 i + \log_2 \log_2 (i + 1) + \log_2 \log_2 \log_2 (i + 3) + C_3.
\]
for all integers $i \geq 1$, where $C_3$ is a constant independent of $i$ and $s$. Note that this bound is a little worse than the bound for the degrees of monic irreducible polynomials, see Lemma 2. Thus
\[
2^{\deg(P_i)} \leq 2^{C_3 i \log_2 (i + 1) \log_2 \log_2 (i + 3)}.
\]
Therefore, for all $N \geq 1$
\[
D_N^* (S_u) \leq \frac{1}{N} \prod_{i \in u} \left[ C_2 i \log_2 (i + 1) \log_2 \log_2 (i + 3) \log (2N) \right],
\]
where $C_2 = 2^{C_3} / \log 2$ is independent of $u$ and $s$. This completes the proof. \hfill \Box

The next theorem is based on Lemma 5 and can be proved analogously to the proof of Theorem 4.

**Theorem 6.** Assume that the QMC algorithm $Q_{N,s}(f)$ uses the first $N$ points of the $s$-dimensional Sobol sequence $S$. If the weights $\gamma_i$ satisfy
\[
\sum_{i=2}^{\infty} \gamma_i^{1/2} i \log i \log \log i < \infty,
\]
then for any $\delta > 0$, there exists a constant $C^*$, independent of $s$, such that
\[
\sup_{f \in B_{s,\gamma}} |I_s(f) - Q_{N,s}(f)| = D_N^* (S, \gamma) \leq C^* N^{-1+\delta}, \quad \text{for all } N \geq 1.
\]
Hence, multivariate integration in $H(K_s, \gamma)$ is strongly tractable, and the QMC algorithm based on the Sobol sequence satisfies strong tractability with the best possible value of the $\varepsilon$-exponent $p^* = 1$.

Remark. The approach used in Theorems 4 and 6 for Niederreiter and Sobol sequences can be generalized. Indeed, for any sequence $P$, if it has the telescopic property and if there exist a nonnegative sequence $\beta_1, \beta_2, \ldots$ and a constant $d$ such that for any nonempty set $u \subseteq D$ and dimension $s \geq 1$, the star discrepancy of the projections of $P$ on $[0, 1)^u$ allows the multiplicative upper bound

$$D_N^*(P_u) \leq \frac{1}{N} \prod_{i \in u} [\beta_i \log(dN)]$$

then under the condition

$$\sum_{i=1}^{\infty} \gamma_i^{1/2} \beta_i < \infty,$$

we can prove that the QMC algorithm using the sequence $P$ achieves strong tractability with the optimal $\varepsilon$-exponent $p^* = 1$. Note that both the Niederreiter sequences $P$ and the Sobol sequences $S$ investigated in this paper have these properties, as shown in Lemmas 3 and 5. The Halton sequences also have these properties (see [8]). But the Faure sequences [5], i.e., $(0, s)$-sequences in prime base $\geq s$, have no such properties. Faure sequences have no telescopic property.

4. Open problems

In conclusion we present several open problems related to the topic of this paper.

It is open whether QMC algorithms based on Niederreiter or Sobol sequences can achieve a strong tractability error bound when the weights $\gamma_i$ only satisfy $\sum_{i=1}^{\infty} \gamma_i < \infty$, but do not satisfy the stronger assumptions in Theorems 4 and 6 (we are inclined to believe that they do). If they do, what is the exact $\varepsilon$-exponent then?

Another related question is: do we have similar strong tractability results for other low discrepancy sequences or for some kind of good lattice rules? Niederreiter-Xing sequences [14] yield significant improvement on all earlier constructions. Do QMC algorithms using Niederreiter-Xing sequences achieve strong tractability in some weighted Sobolev spaces of functions?

As mentioned earlier, there is a nonconstructive proof of strong $\varepsilon$-exponent $p^* = 1$ under $\sum_{i=1}^{\infty} \gamma_i^{1/2} < \infty$. Our conditions in Theorems 4 and 6 are slightly stronger than this condition. It is open whether the condition $\sum_{i=1}^{\infty} \gamma_i^{1/2} < \infty$ is necessary for $p^* = 1$. Can we construct low discrepancy sequences such that $p^* = 1$ under this condition? Do the known low discrepancy sequences satisfy this?

Extend the strong tractability results to other function spaces. Characterize function spaces for which QMC algorithms achieve strong tractability. Construct explicit QMC algorithms (low discrepancy sequences or lattice rules) for which strong tractability is achieved.

Acknowledgments

The author gratefully acknowledges the valuable comments and suggestions from F. J. Hickernell, E. Novak, I. H. Sloan, H. Woźniakowski, and anonymous referees.
REFERENCES


Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China
E-mail address: xwang@math.tsinghua.edu.cn