JOURNAL OF COMPLEXITY **14,** 1–33 (1998)
ARTICLE NO. CM970463

When Are Quasi-Monte Carlo Algorithms Efficient for High Dimensional Integrals?

Ian H. Sloan*

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

and

Henryk Woźniakowski†

Department of Computer Science, Columbia University, New York 10027 and Institute of Applied Mathematics and Mechanics, University of Warsaw, ul. Banacha 2, 02–097 Warsaw, Poland

Received April 1997

Recently, quasi-Monte Carlo algorithms have been successfully used for multivariate integration of high dimension d, and were significantly more efficient than Monte Carlo algorithms. The existing theory of the worst case error bounds of quasi-Monte Carlo algorithms does not explain this phenomenon. This paper presents a partial answer to why quasi-Monte Carlo algorithms can work well for arbitrarily large d. It is done by identifying classes of functions for which the effect of the dimension d is negligible. These are *weighted* classes in which the behavior in the successive dimensions is moderated by a sequence of weights. We prove that the minimal *worst case* error of quasi-Monte Carlo algorithms does not depend on the dimension d iff the sum of the weights is finite. We also prove that the minimal number of function values in the worst case setting needed to reduce the initial error by ε is bounded by $C\varepsilon^{-p}$, where the exponent $p \in [1, 2]$, and C depends exponentially on the sum of weights. Hence, the relatively small sum of the weights makes some quasi-Monte Carlo algorithms strongly tractable. We show in a nonconstructive way that many quasi-Monte Carlo algorithms are strongly tractable. Even random selection of sample points (done once for the whole weighted class of functions and then the worst case error is established for that particular selection, in contrast to Monte Carlo where random selection of sample points is carried out for a fixed function) leads to strong tractable quasi-Monte Carlo algorithms. In this case the minimal number of function values in the *worst case* setting is of order ε^{-p} with the exponent $p = 2$. The deterministic construction of strongly tractable quasi-Monte Carlo algorithms as well as the minimal exponent p is open.

*E-mail: i.sloan@unsw.edu.au.
†E-mail: henryk@cs.columbia.edu.

0885-064X/98 $25.00

1. INTRODUCTION

Monte Carlo algorithms are typically used for multivariate integration of high dimension d. The expected error of Monte Carlo algorithms that use n function values is of order $n^{-1/2}$. The rate of convergence, although not very fast, does *not* depend on the dimension d. The number of function values needed for Monte Carlo algorithms to reduce the initial error by ε is of order ε^{-2}.

Recently quasi-Monte Carlo algorithms have been successfully used for very large values of d, especially in financial applications; see [3, 4, 11, 12, 15, 16, 18, 20–23]. For example, calculations with $d = 360$ for Sobol points and generalized Faure points have been reported by Papageorgiou, Paskov, and Traub in [20, 22, 23] for collateralized mortgage obligations. The errors for these examples were observed to be independent of d and were of order n^{-1}. Hence, quasi-Monte Carlo algorithms win in two ways over Monte Carlo, in that we have both a better exponent of convergence and a better assurance of error.

The apparent success of these quasi-Monte Carlo calculations presents a major challenge to many computational theorists. The challenge is to explain why quasi-Monte Carlo algorithms are so efficient for high dimensions. This problem provides the spur for the present study.

In this paper, we do *not* explain why quasi-Monte Carlo algorithms are so efficient for finance problems. Instead, we identify classes of functions for which the worst case error estimates of some quasi-Monte Carlo algorithms essentially do not depend on the dimension d and are of order $n^{-1/p}$ with $p \in [1, 2]$. Hence, to explain the behavior of quasi-Monte Carlo algorithms for particular finance problems it would be sufficient to show that they belong to these weighted classes of functions and that $p = 1$. This work remains to be done.

We now present an informal derivation of classes of functions for which the effect of the dimension d is negligible. Our starting point is the classical Sobolev space of once differentiable functions with respect to each variable $t_1, t_2, \ldots, t_d$. For this class there exists a well-established error analysis of the quasi-Monte Carlo algorithms based on the Koksma–Hlawka inequality and discrepancy (see, for example [5, 17, 26, 27]), which we shall discuss in Section 3. While this analysis leads to an effective design principle of low-discrepancy sets and sequences, its theoretical usefulness appears to be restricted to moderate dimensions d—by general consent perhaps up to 12 but certainly not 360.

Every existing analysis that we are aware of assumes that the behavior with respect to each of the d variables is essentially the same. On the other hand, it has been pointed out (see, for example, [3]) that this is not a realistic assumption. The concept of effective dimension appears in a number of papers; see [3, 16, 22]. For example, it is claimed in [3] that in a specific financial calculation with $d = 360$ the effective dimension is only of the order of 30.

To allow the theory to better match the real needs, in this paper we assume that the components of $t = [t_1, \ldots, t_d]$ are ordered so that t_1 is the most important, etc. and then assume that the behavior in the successive dimensions is moderated by weights $\gamma_1, \gamma_2, \ldots, \gamma_d$, with $1 = \gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_d \geq 0$.

The precise way in which this is done will be made clear in Section 3. Here, we assume first that all γ_j are positive and explain the role of γ_j by means of the function

$$f(t_1, t_2, \ldots, t_d) = g(\gamma_1^{1/2} t_1, \gamma_2^{1/2} t_2, \ldots, \gamma_d^{1/2} t_d),$$

and f is to be integrated over some region. The choice of the weights γ_j should be such as to make the behavior of g with respect to each of the d variables essentially the same. As before, assume that partial derivatives $\partial^k f/(\partial t_{i_1} \partial t_{i_2} \cdots \partial t_{i_k})$ exist for $k \in [1, d]$ and all subscript choices satisfying $1 \leq i_1 < i_2 < \cdots < i_k \leq d$. Then

$$\begin{aligned} &\frac{\partial^k}{\partial t_{i_1} \partial t_{i_2} \cdots \partial t_{i_k}} f(t_1, t_2, \ldots, t_d) \\ &\quad = \gamma_{i_1}^{1/2} \gamma_{i_2}^{1/2} \cdots \gamma_{i_k}^{1/2} \frac{\partial^k}{\partial t_{i_1} \partial t_{i_2} \cdots \partial t_{i_k}} g(\gamma_1^{1/2} t_1, \gamma_2^{1/2} t_2, \ldots, \gamma_d^{1/2} t_d). \end{aligned}$$

If the behavior of all partial derivatives of g is more or less the same then the partial derivatives of f depend inversely on the products of $\gamma_{i_j}^{1/2}$. This explains how the weights affect the behavior of partial derivatives of f, and we will capture this in defining a *weighted* norm in Section 3.

The weights γ_j can also model the case when the function f is constant with respect to, say, $t_k, t_{k+1}, \ldots, t_d$. Then we set $\gamma_k = \gamma_{k+1} = \cdots = \gamma_d = 0$. In this way, the dimension d will be reduced to the dimension k. If f is "almost" constant with respect to $t_k, t_{k+1}, \ldots, t_d$ then small $\gamma_k, \gamma_{k+1}, \ldots, \gamma_d$ will have a similar effect of reducing the dimension.

For a weighted sequence $\gamma = \{\gamma_j\}$, we work in *weighted* classes of functions, and the error bound is given by the *weighted* Koksma–Hlawka inequality and is expressed in terms of a *weighted* discrepancy. In this setting, it makes sense to consider the quasi-Monte Carlo algorithms with arbitrarily large values of d, provided the weights γ_j approach zero sufficiently rapidly. In fact, we define the *limiting discrepancy* as the limit of the weighted discrepancy as d approaches infinity.

It turns out that the quality of the worst case errors of quasi-Monte Carlo algorithms in the weighted classes of functions depends on the sum

$$s(\gamma) := \sum_{j=1}^{\infty} \gamma_j. \tag{1}$$

More precisely, let $n = n_\gamma(\varepsilon, d)$ be the minimal number of function values necessary to reduce the initial error by a factor of ε for the d-dimensional case when we use quasi-Monte Carlo algorithms. We stress that $n_\gamma(\varepsilon, d)$ describes the behavior of a best quasi-Monte Carlo algorithm in the worst case setting.

Then we prove that $n_\gamma(\varepsilon, d)$ is independent of d and depends polynomially on $1/\varepsilon$ iff $s(\gamma)$ is finite. That is,

$$n_\gamma(\varepsilon, d) \leq C\varepsilon^{-p} \tag{2}$$

for some positive C and p independent of d and ε holds iff $s(\gamma) < \infty$. Clearly, p in (2) must be at least 1, since for $d = 1$ we have $n_\gamma(\varepsilon, 1) = \Theta(\varepsilon^{-1})$. We prove that $p \leq 2$. Because the bound (2) is independent of d, we say that the multivariate integration problem is *strongly quasi-Monte Carlo tractable in the worst case setting*, or briefly *strongly QMC-tractable*.

Hence, if $s(\gamma) < \infty$ then some quasi-Monte Carlo algorithms are superior to Monte Carlo algorithms. They are superior because we have a worst case assurance of the error rather than a stochastic one. However, it is not clear if we have a better bound on the minimal number of function values since we do not know whether $p < 2$. The problem of finding the exponent of strong tractability, which is defined as the minimal p in (2), is open.

Assume now that $s(\gamma) = \infty$. This obviously holds for the unweighted case $\gamma_j = 1\ \forall j$. Then the bound on $n_\gamma(\varepsilon, d)$ depends on how fast the partial sum $s_d(\gamma) = \sum_{j=1}^d \gamma_j$ approaches infinity. If the limit superior of $s_d(\gamma)/\ln d$ is finite and equal to a then $n_\gamma(\varepsilon, d)$ depends polynomially on d and $1/\varepsilon$. More precisely, we have

$$n_\gamma(\varepsilon, d) \leq C d^q \varepsilon^{-p}$$

for some positive C, q, and p independent of d and ε. Furthermore, $q \in [a/12, a/6]$, and $p \in [1, 2]$. We then say that the multivariate integration problem is *quasi-Monte Carlo tractable in the worst case setting*, or briefly, *QMC-tractable*.

Finally, if $s_d(\gamma)/\ln d$ is unbounded, then we show that $n_\gamma(\varepsilon, d)$ must increase faster than any polynomial in d. In this case, we say that the multivariate integration problem is *quasi-Monte Carlo intractable in the worst case setting*, or more briefly, *QMC-intractable*.

This shows that quasi-Monte Carlo algorithms are not tractable for the unweighted case, and the weighted sequence γ makes them tractable iff $s_d(\gamma)/\ln d$ is bounded.

We now stress the role of the partial sums $s_d(\gamma)$ independently of whether their limit is finite or infinite. We prove that $n_\gamma(\varepsilon, d)$ depends *exponentially* on $s_d(\gamma)$, namely we have

$$n_\gamma(\varepsilon, d) \geq (1 - \varepsilon^2) 1.055^{s_d(\gamma)}.$$

Observe that the last bound does not really address the dependence on ε. The essence of this bound is the dependence on d through the partial sum $s_d(\gamma)$. Although 1.055 is barely larger than 1, we do have an exponential dependence on

$s_d(\gamma)$. Hence, even when the limit of $s_d(\gamma)$ is finite and strong QMC-tractability holds, we are in trouble if $s_d(\gamma)$ is large. To guarantee a reasonable bound on $n_\gamma(\varepsilon, d)$ for all d we must also assume that the limit $s(\gamma)$ is relatively small.

Observe that for the unweighted case, $\gamma_j \equiv 1$, $n_\gamma(\varepsilon, d)$ depends exponentially on d. For $d = 360$ we have

$$n_\gamma(\varepsilon,\ 360) \geq (1 - \varepsilon^2)\, 2.78 * 10^8.$$

Hence, $n_\gamma(\varepsilon, 360)$ is really huge and it is impossible to guarantee a small error for the unweighted case for large d. This indicates that the success of quasi-Monte Carlo algorithms for finance applications cannot be explained on the grounds of the worst case error for the classical (unweighted) Sobolev space. This may indicate that finance problems belong to more restricted spaces of functions, and one may hope that they belong to a weighted class with a finite and relatively small $s(\gamma)$.

The theoretical approach in this paper is considerably more general than has been indicated so far. The results rest on general results for multivariate integration in reproducing kernel Hilbert spaces. The general analysis for reproducing kernel Hilbert spaces is developed in Section 6 and then applied to deduce necessary and sufficient conditions for strong QMC- and QMC-tractability in spaces associated with the weighted Koksma–Hlawka inequality.

Our proof technique is based on averaging arguments. Hence, even if $s(\gamma)$ is finite, our arguments do not allow us to explicitly construct any good choices for the sample points. This averaging procedure is presented in Section 5, where we introduce a new notion of tractability. We call it *tractability for average sample points*. In this case, we take n-tuples of sample points which are independent and uniformly distributed over the d-dimensional unit cube. Then we determine the worst case error of the quasi-Monte Carlo algorithm that uses these sample points. Finally, we define the average error by averaging the worst case errors of the n-tuples in the L_2 norm. Tractability for average sample points is then defined as before in terms of the behavior of the average error. We stress this is not the same as in the Monte Carlo algorithms. For the Monte Carlo algorithms the average is taken for a *fixed* function, not for the worst case error as in our case.

Surprisingly enough, tractability for average sample points holds under the same conditions on the weighted sequence γ as before. In particular, finiteness of the sum $s(\gamma)$ is a necessary and sufficient condition for strong QMC-tractability for average sample points. It therefore follows under the same condition of finiteness of $s(\gamma)$ that there are many sample point sets which lead to strongly tractable quasi-Monte Carlo algorithms. In Section 5 we indicate how such sample points can be found computationally.

The minimal number of function values needed for tractability for average sample points has a sharp bound $C\, d^q\, \varepsilon^{-p}$ for some positive C and $q = a/6$, $p = 2$, where, as before, a is the limit superior of $s_d(\gamma)/\ln d$. From the mean value theorem we conclude that the minimal number $n_\gamma(\varepsilon, d)$ of function values has the same bound. It would be interesting to improve the bound on $n_\gamma(\varepsilon, d)$ and, in particular, to check whether we can set $p = 1$ in (2).

We wish to stress that good quality of quasi-Monte Carlo algorithms for high d is *not* restricted to only weighted classes of functions. In a recent paper of Papageorgiou and Traub [21] it was empirically observed that the effect of dimension d is negligible for an isotropic class of functions where integrands depend on a norm of the vector. The good quality of quasi-Monte Carlo algorithms in this isotropic class cannot be explained by the analysis of the weighted classes. In general, there are probably many different classes of functions for which quasi-Monte Carlo algorithms behave successfully for high d. Of course, it would be interesting to identify all such classes. It would enable us to better understand the essence of quasi-Monte Carlo algorithms.

We end this introduction by stating one more open problem. This problem is to estimate the weighted sequence γ for some practically important applications. Here, natural candidates are finance problems. We believe (see also [3, 4]) that many finance problems may be defined in terms of *path* integrals which are infinite-dimensional integrals with respect to the Wiener measure; see [29] where tractability of path integration is studied. An approximation of a path integral is a d-dimensional integral, and the error of such an approximation tends to zero as d approaches infinity. This explains why arbitrarily large d can be met in computational practice. It seems plausible that the weights γ_j should be related to the eigenvalues of the covariance operator of the Wiener measure. If so, γ_j should be proportional to j^{-2}. Then the series $s(\gamma)$ is indeed convergent and we get strong tractability. This would partially explain why some finance problems can be solved by quasi-Monte Carlo algorithms so efficiently even for huge d.

2. TRACTABILITY

We deal with multivariate integration

$$I_d(f) = \int_{[0,1]^d} f(t)\,dt \tag{3}$$

for functions defined over the d-dimensional unit cube $[0, 1]^d$ which belong to a normed space F_d. The norm in F_d is denoted by $\|\cdot\|_d$. In most cases, we will assume that F_d is a Hilbert space. Here $d \geq 1$; we are mainly interested in large d.

As mentioned in the introduction, we restrict our analysis to quasi-Monte Carlo algorithms, since they are often used in computational practice for high-dimensional integration. A quasi-Monte Carlo algorithm $Q_{n,d}$ is of the form

$$Q_{n,d}(f) = \frac{1}{n}\sum_{i=1}^{n} f(t_i). \tag{4}$$

Here, the sample points t_i are deterministic, belong to $[0, 1]^d$, and may depend on n and d, as well as on the space F_d. We stress that the weights of quasi-Monte Carlo algorithms are by definition all equal to $1/n$.

We define the (worst case) error of $Q_{n,d}$ by its worst case performance over the unit ball of F_d,

$$e(Q_{n,d}) = \sup_{f \in F_d,\, \|f\|_d \le 1} |I_d(f) - Q_{n,d}(f)|. \tag{5}$$

For $n = 0$ we formally set $Q_{0,d} = 0$, and

$$e(Q_{0,d}) = \sup_{f \in F_d,\, \|f\|_d \le 1} |I_d(f)| = \|I_d\|$$

is the initial error. This is the a priori error in multivariate integration without sampling the function.

We would like to reduce the initial error by a factor of ε, where $\varepsilon \in [0, 1)$. That is, we are looking for the smallest $n = n_{\min}(\varepsilon, d, \{Q_{n,d}\})$ for which [1]

$$e(Q_{n,d}) \le \varepsilon e(Q_{0,d}).$$

We are ready to define what we mean by various notions of tractability. A general discussion of tractability can be found in [19, 28, 30, 31]. We say that a family $\{Q_{n,d}\}$ of quasi-Monte Carlo algorithms is *tractable* [2] iff there exist nonnegative C, q, and p such that

$$n_{\min}(\varepsilon, d, \{Q_{n,d}\}) \le C d^q \varepsilon^{-p} \quad \forall d = 1, 2, \dots; \forall \varepsilon \in (0, 1). \tag{6}$$

Tractability means that we need a number of function evaluations at most a polynomial in d and ε^{-1} to approximate multivariate integration to within $\varepsilon\|I_d\|$. The smallest q and p (or the infima of q and p) are called the d-exponent and the ε-exponent of tractability for $\{Q_{n,d}\}$.

We say that a family $\{Q_{n,d}\}$ of quasi-Monte Carlo algorithms is *strongly tractable* if (6) holds with $q = 0$. In this case, the number of samples is independent of d and depends polynomially on ε^{-1}.

Of course, if $\{Q_{n,d}\}$ is strongly tractable then the d-exponent is zero. However, the converse is, in general, not true. That is, it may happen that the d-exponent

[1] In many papers $n_{\min}(\varepsilon, d, \{Q_{n,d}\})$ is defined as the minimal n for which the condition $e(Q_{n,d}) \le \varepsilon$ holds. Our condition $e(Q_{n,d}) \le \varepsilon e(Q_{0,d})$ can be viewed as the normalization of the functional I_d. That is, for $I'_d := \|I_d\|^{-1} I_d$ we have $\|I'_d\| = 1$ and the two conditions coincide.

[2] Alternatively, as in [28], such algorithms may be called polynomial time algorithms.

of $\{Q_{n,d}\}$ is zero and $\{Q_{n,d}\}$ is *not* strongly tractable. Indeed, assume that $n_{\min}(\varepsilon, d, \{Q_{n,d}\})$ is of order, say, $(\ln d)^{1/2}\varepsilon^{-1}$. Then (6) holds with $p = 1$ and any positive q, and therefore, the d-exponent is zero and the ε-exponent is 1. Still, we cannot set $q = 0$ in (6), and therefore, $\{Q_{n,d}\}$ is *not* strongly tractable. In Section 3 we will see that such a case can indeed happen.

We say that multivariate integration in the space F_d is *QMC-tractable* (or *strongly QMC-tractable*) iff there exists a family of quasi-Monte Carlo algorithms $\{Q_{n,d}\}$ which is tractable (or strongly tractable). The infima of the d- and ε-exponents of tractability for $\{Q_{n,d}\}$ are called the d- and ε-exponents of QMC-tractability for multivariate integration in the space F_d, or for short the d- and ε-exponents.

If such a family does not exist we say that multivariate integration is *QMC-intractable* (or *strongly QMC-intractable*) in the space F_d. The lack of QMC-tractability means that a polynomial number of arbitrary samples is not enough to approximate multivariate integration by a quasi-Monte Carlo algorithm to within $\varepsilon\|I_d\|$. We stress that intractability of multivariate integration in this paper is defined in terms of quasi-Monte Carlo algorithms. It may happen that the use of other algorithms may break intractability. Since we consider quasi-Monte Carlo with arbitrary sample points, it would mean that the equal weights of size $1/n$ for quasi-Monte Carlo algorithms are causing the trouble. This is known to happen for some (rather esoteric) spaces F_d, as explained in Remark 2 of Section 6.

The main purpose of this paper is to explore for which spaces F_d we have tractability and strong tractability. In particular, we show that strong QMC-tractability holds in some weighted spaces, whereas in simple tensor product[3] spaces with nonnegative reproducing kernels QMC-tractability holds only for trivial cases.

3. WEIGHTED KOKSMA–HLAWKA INEQUALITY AND WEIGHTED DISCREPANCY

We first recall the classical error formula for multivariate integration derived by Zaremba in 1968; see [17, 32]. We use the standard notation as in many papers. Let $D = \{1, 2, \ldots, d\}$ be the set of coordinate indices. For any $u \subset D$ we denote by $|u|$ its cardinality. Obviously, we have 2^dsuch subsets. For the vector $x \in [0, 1]^d$, let x_u denote the vector from $[0, 1]^{|u|}$ containing the components of x whose indices are in u, and let $dx_u = \prod_{j\in u} dx_j$. By $(x_u, 1)$ we mean the vector x from $[0, 1]^d$, with all components whose indices are not in u replaced by 1.

[3]By a simple tensor product space F_d we mean a space that is a tensor product of d copies of the same space F_1, that is, $F_d = F_1 \otimes F_1 \otimes \cdots \otimes F_1$.

Consider the quasi-Monte Carlo algorithm $Q_{n,d}$ that uses sample points $\{t_i\}$. For the sample points $\{t_i\}$, define the discrepancy function as

$$\operatorname{disc}(x) = x_1 x_2 \cdots x_d - \frac{|\{i\colon\ t_i \in [0,\, x)\}|}{n}.$$

Here, $[0, x) = [0, x_1) \times [0, x_2) \times \cdots \times [0, x_d)$.

Assume that the function f belongs to the Sobolev space $W_2^{(1,1,\ldots,1)}([0, 1]^d)$.[4] Then Zaremba's identity states that

$$I_d(f) - Q_{n,d}(f) = \sum_{\emptyset \neq u \subset D} (-1)^{|u|} \int_{[0,1]^{|u|}} \operatorname{disc}(x_u,\, 1)\, \frac{\partial^{|u|}}{\partial x_u} f(x_u,\, 1)\, dx_u. \quad (7)$$

We want to consider functions whose dependence on successive variables is increasingly limited. Intuitively, we would like to assume that the jth variable x_j is the jth most important and that the partial derivative of f with respect to x_j is bounded by some nonnegative parameter γ_j. More precisely, suppose that we are given a sequence $\gamma = \{\gamma_j\}$ such that

$$\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_j \geq \cdots \geq 0.$$

We will normalize the sequence γ by assuming that $\gamma_1 = 1$. Then we define $\gamma_\emptyset = 1$ and for nonempty $u \subset D$,

$$\gamma_u = \prod_{j \in u} \gamma_j. \quad (8)$$

If all γ_j are positive, we can rewrite (7) by multiplying and dividing by $\gamma_u^{1/2}$,

$$\begin{aligned} &I_d(f) - Q_{n,d}(f) \\ &\quad = \sum_{\emptyset \neq u \subset D} (-1)^{|u|} \gamma_u^{1/2} \int_{[0,1]^{|u|}} \operatorname{disc}(x_u,\, 1)\gamma_u^{-1/2}\, \frac{\partial^{|u|}}{\partial x_u} f(x_u,\, 1)\, dx_u. \end{aligned} \quad (9)$$

Applying the Cauchy–Schwarz inequality for integrals and sums, we obtain

$$|I_d(f) - Q_{n,d}(f)| \leq \operatorname{disc}_\gamma(\{t_i\}) \|f\|_{d,\gamma}, \quad (10)$$

where

$$\operatorname{disc}_\gamma(\{t_i\}) = \left(\sum_{\emptyset \neq u \subset D} \gamma_u \int_{[0,1]^{|u|}} \operatorname{disc}^2(x_u,\, 1)\, dx_u \right)^{1/2} \quad (11)$$

and

[4] This space is the tensor product $W_2^1([0, 1]) \otimes \cdots \otimes W_2^1([0, 1])$, d times, where $W_2^1([0, 1])$ is the space of scalar absolutely continuous functions whose first derivatives belong to $L_2([0, 1])$.

$$\| f \|_{d,\gamma} = \left(\sum_{u \subset D} \gamma_u^{-1} \int_{[0,1]^{|u|}} \left| \frac{\partial^{|u|}}{\partial x_u} f(x_u, 1) \right|^2 dx_u \right)^{1/2}. \tag{12}$$

A few words of comments are in order. For $\gamma_j = 1$, $j \geq 1$, (10) is the L_2 version of the classical Koksma–Hlawka inequality. That is why we call (10) for arbitrary γ_j a *weighted Koksma–Hlawka inequality*.

We will call $\mathrm{disc}_\gamma(\{t_i\})$ *the weighted discrepancy* (or the weighted L_2 discrepancy) of the sample points $\{t_i\}$ of $Q_{n,d}$. Here, the word "weighted" refers to the sequence γ. As we shall see, for some families $\{Q_{n,d}\}$ of quasi-Monte Carlo algorithms and some sequences γ we will be able to bound the discrepancy $\mathrm{disc}_\gamma(\{t_i\})$ by a function of n which goes to zero polynomially in $1/n$. The classical discrepancy is obtained for $\gamma_j \equiv 1$.

The square of the weighted discrepancy is defined as the sum of $2^d - 1$ terms. It is therefore surprising that the weighted discrepancy can be exactly computed in time proportional to $n^2 d$, as recently shown by Joe in [10]. For the case $\gamma_j = 1$ for all j this was proved earlier by Hickernell in [9]; see also the work of Heinrich and Frank [6, 8] for fast evaluation of discrepancy, as well as a recent paper of Matoušek [14].

Remark 1: *Limiting discrepancy.* The concept of weights allows us to define the discrepancy for dimension d tending to infinity. This can be done as follows. Consider points $t_i^{(\infty)} = [t_{i,1}, t_{i,2}, \ldots] \in [0, 1]^\infty$ and their d-dimensional projections $t_i^{(d)} = [t_{i,1}, t_{i,2}, \ldots, t_{i,d}]$ for $i = 1, 2, \ldots, n$. Then the discrepancy $\mathrm{disc}_\gamma(\{t_i^{(d)}\})$ is a nondecreasing function of d. This follows from the fact that if d is replaced by $d + 1$, the sum in (11) has the same terms as for d when $u \subset \{1, 2, \ldots, d\}$ and it has extra nonnegative terms corresponding to $\{u, d + 1\}$ with $u \subset \{1, 2, \ldots, d\}$. Hence, the limit of the discrepancy with respect to d exists, although it can be infinite. The limit,

$$\mathrm{disc}_\gamma(\{t_i^{(\infty)}\}) = \lim_{d\to\infty} \mathrm{disc}_\gamma(\{t_i^{(d)}\}), \tag{13}$$

will be called the *limiting discrepancy*. It is natural to ask for which sample points $\{t_i^{(\infty)}\}$ the limiting discrepancy is finite. It turns out that this does not depend on the sample points $\{t_i^{(\infty)}\}$ but only on the sequence γ. Specifically, for any choice of sample points $\{t_i^{(\infty)}\}$, we find

$$\mathrm{disc}_\gamma(\{t_i^{(\infty)}\}) < \infty \quad \text{iff} \sum_{j=1}^\infty \gamma_j < \infty. \tag{14}$$

A proof of (14) is given in the Appendix; part of the argument uses results obtained later in the paper.

We now discuss (12). Observe that in defining $\| f \|_{d,\gamma}$ we included $u = \emptyset$; that is, we added one more term $|f(1)|^2$ in the sum. This was done in order to make $\| \cdot \|_{d,\gamma}$ a norm. In this section we shall be analyzing quasi-Monte Carlo algorithms for the Sobolev space $W_2^{(1,1,\dots,1)}([0,1]^d)$ equipped with the norm $\| \cdot \|_{d,\gamma}$. Formally, we set

$$F_{d,\gamma} = \{ f \in W_2^{(1,1,\dots,1)}([0,1]^d) \colon \; \| f \|_{d,\gamma} < \infty \}. \tag{15}$$

We now explain the role of the sequence γ in the space $F_{d,\gamma}$. Since the error of quasi-Monte Carlo algorithms is defined here over the unit ball of $F_{d,\gamma}$, small γ_u means that the L_2-norm of the partial derivative $\partial^{|u|} f / \partial x_u$ must also be small. In fact, we can even permit that the γ_j are zero beyond some index. In that case in (12) we adopt the convention that $0/0 = 0$; hence, $\gamma_j = 0$ implies that the functions must be constant with respect to x_j. In the extreme case when all $\gamma_j = 0$ for $j \geq 2$ this means that we permit only dependence on x_1. It is clear that the unit ball of $F_{d,\gamma}$ shrinks for small γ_j, and so it makes multivariate integration easier. It is natural to ask what are the minimal conditions on the sequence γ to guarantee tractability and strong tractability of families of quasi-Monte Carlo algorithms. These conditions are presented in the next section.

The weighted discrepancy plays an important role in the error analysis for the space $F_{\gamma,d}$. Note first that the weighted Koksma–Hlawka inequality (10) tells us that $\mathrm{disc}_\gamma(\{t_i\})$ is an upper bound on $e(Q_{n,d})$, the worst-case error in $F_{\gamma,d}$ for the quasi-Monte Carlo algorithm $Q_{n,d}$ employing the points $t_1, \dots, t_n$. The following theorem tells us that $\mathrm{disc}_\gamma(\{t_i\})$ is, not only an upper bound on the error, it *is* the error! This is not surprising since, as is well known, the use of the Cauchy–Schwarz inequality to obtain (10) is sharp. For completeness, and as a warmup, we provide a short proof.

THEOREM 1. *If $Q_{n,d}$ is a quasi-Monte Carlo algorithm employing the points $t_1, \dots, t_n$, and if $e(Q_{n,d})$ is the worst-case error for $Q_{n,d}$ in the space $F_{d,\gamma}$, then*

$$e(Q_{n,d}) = \mathrm{disc}_\gamma(\{t_i\}).$$

Proof. Given the Koksma–Hlawka inequality (10), to prove the theorem it is sufficient to demonstrate one nontrivial function $f \in F_{\gamma,d}$ for which the Koksma–Hlawka inequality becomes an equality. Such a function is

$$f(x) := \prod_{j=1}^{d} \left(1 + \frac{1}{2} \gamma_j (1 - x_j^2) \right) - \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{d} (1 + \gamma_j \min(1 - x_j,\, 1 - t_{i,j})),$$

where $t_i = [t_{i,1}, t_{i,2}, \dots, t_{i,d}]$. Note that $f(1) = 0$.

Indeed, for nonempty $u \subset D = \{1, 2, \ldots, d\}$, we have for this function f and for almost all $x \in [0, 1]^d$,

$$\frac{\partial^{|u|}}{\partial x_u} f(x_u, 1) = (-1)^{|u|}\gamma_u \prod_{j \in u} x_j - \frac{(-1)^{|u|}\gamma_u}{n} |\{i: (t_i)_u \in [0, x_u)\}|$$
$$= (-1)^{|u|}\gamma_u \operatorname{disc}(x_u, 1).$$

Hence, the identity (9) for this particular function f becomes

$$I_d(f) - Q_{n,d}(f) = \sum_{\emptyset \neq u \subset D} \int_{[0,1]^{|u|}} \left(\gamma_u^{-1/2} \frac{\partial^{|u|}}{\partial x_u} f(x_u, 1)\right)^2 dx_u$$
$$= \operatorname{disc}_\gamma(\{t_i\}) \| f \|_{d,\gamma}.$$

This completes the proof. ■

We end this section by pointing out that weighted discrepancy can also be defined and studied in the L_p norm, $p \in [1, \infty]$. Indeed, let $1/p + 1/q = 1$. We apply Hölder's inequality for integrals and sums to (9) and obtain

$$|I_d(f) - Q_{n,d}(f)| \leq \operatorname{disc}_{\gamma,p}(\{t_i\}) \| f \|_{d,\gamma,q},$$

where

$$\operatorname{disc}_{\gamma,p}(\{t_i\}) = \left(\sum_{\emptyset \neq u \subset D} \gamma_u^{p/2} \int_{[0,1]^{|u|}} \operatorname{disc}^p(x_u, 1)\, dx_u\right)^{1/p}$$

and

$$\| f \|_{d,\gamma,q} = \left(\sum_{u \subset D} \gamma_u^{-q/2} \int_{[0,1]^{|u|}} \left|\frac{\partial^{|u|}}{\partial x_u} f(x_u, 1)\right|^q dx_u\right)^{1/q}.$$

Obviously, for $p = \infty$, we have

$$\operatorname{disc}_{\gamma,\infty}(\{t_i\}) = \sup_{x \in [0,1]^d} \max_{\emptyset \neq u \subset D} \gamma_u^{1/2} |\operatorname{disc}(x_u, 1)|.$$

The weighted discrepancy $\operatorname{disc}_{\gamma,p}(\{t_i\})$ is a nondecreasing function of d, and the limiting discrepancy is always well defined. It can be shown in the same way as in Section A.1 that the limiting discrepancy is finite iff $(\sum_{j=1}^{\infty} \gamma_j^{p/2})^{1/p} < \infty$.

To find bounds on $\operatorname{disc}_{\gamma,p}(\{t_i\})$ one needs to have bounds on the L_p norm of $\operatorname{disc}(x_u, 1)$ for all nonempty u. Such bounds are difficult to obtain, especially for large d and $p \neq 2$. In any case, we anticipate that, as in the case $p = 2$, one can find quasi-Monte Carlo algorithms for which $\operatorname{disc}_{\gamma,p}(\{t_i\})$ is uniformly bounded in d and goes to zero polynomially in n^{-1} as long as the series $(\sum_{j=1}^{\infty} \gamma_j^{p/2})^{1/p} < \infty$ is convergent.

4. CONDITIONS FOR TRACTABILITY IN $F_{d,\gamma}$

It is not difficult to show, and we shall do so formally in Section 6, that the initial error in multivariate integration in the space $F_{d,\gamma}$ is

$$e(Q_{0,d}) = \|I_d\| = \prod_{j=1}^{d} \left(1 + \frac{1}{3}\gamma_j\right)^{1/2}. \tag{16}$$

Observe from (16) that the initial error, equaling the norm of multivariate integration, is bounded as a function of d iff $\sum_{j=1}^{\infty} \gamma_j$ is finite. This may indicate that the assumption $\sum_{j=1}^{\infty} \gamma_j < \infty$ is quite natural. As we already mentioned, this assumption guarantees that the limiting discrepancy is finite, and as we shall see below, this assumption is also necessary and sufficient for strong QMC-tractability. For completeness, we shall also analyze what happens if this assumption does not hold. In this case, the situation will be seen to depend on how fast $\sum_{j=1}^{d} \gamma_j$ goes to infinity. If it goes no faster than $\ln d$ we still have QMC-tractability; if it goes faster than $\ln d$ then multivariate integration is not QMC-tractable.

THEOREM 2. (i) *Multivariate integration in the space $F_{d,\gamma}$ is strongly QMC-tractable iff*

$$\sum_{j=1}^{\infty} \gamma_j < \infty. \tag{17}$$

If (17) *holds, the ε-exponent belongs to* $[1, 2]$.

(ii) *Multivariate integration in the space $F_{d,\gamma}$ is QMC-tractable iff*

$$a := \limsup_{d\to\infty} \frac{\sum_{j=1}^{d} \gamma_j}{\ln d} < \infty. \tag{18}$$

If a is finite then the d-exponent belongs to $[a/12, a/6]$, *and the ε-exponent belongs to* $[1, 2]$.

(iii) *Let $n_\gamma(\varepsilon, d)$ be the minimal number of sample points needed to reduce the initial error by a factor of ε by a quasi-Monte Carlo algorithm. Then*

$$n_\gamma(\varepsilon, d) \le \left\lceil \frac{\exp\left(\frac{1}{6}\sum_{j=1}^{d} \gamma_j\right) - 1}{\varepsilon^2} \right\rceil = \left\lceil \frac{(1.1836\ldots)^{\sum_{j=1}^{d} \gamma_j} - 1}{\varepsilon^2} \right\rceil \tag{19}$$

and

$$n_\gamma(\varepsilon, d) \ge (1-\varepsilon^2)\exp\left(2c\sum_{j=1}^{d} \gamma_j\right), \tag{20}$$

where

$$c = \min_{x,\, y \in [0,\, 1]} \frac{1}{y} \left(1 - \frac{1 + y(1-x^2)/2}{\sqrt{(1+y(1-x))(1+y/3)}} \right) \approx 0.027.$$

Hence, since $\exp(2 * 0.027) = 1.055\ldots$, *we have*

$$n_\gamma(\varepsilon,\, d) \geq (1-\varepsilon^2)\, 1.055^{\sum_{j=1}^d \gamma_j}$$

and $n_\gamma(\varepsilon,\, d)$ *depends exponentially on* $\sum_{j=1}^d \gamma_j$.

The proof of Theorem 2 is presented in the Appendix; see Section A.3. It will follow from a more general analysis of quasi-Monte Carlo algorithms in Hilbert spaces with reproducing kernels. It also relies on an averaging argument presented in the next section.

Theorem 2 presents necessary and sufficient conditions on the sequence γ to guarantee strong QMC-tractability and QMC-tractability. From classical results for the L_2 discrepancy it is known (see [7, 24]), that for $\gamma_j \equiv 1$ the function $n_\gamma(\varepsilon,\, d)$ is asymptotically (as ε tends to zero) equal to $c_d \varepsilon^{-1}(\ln \varepsilon^{-1})^{(d-1)/2}$ for some positive number c_d. Thus, even for $d = 1$ the minimal number of sample points is of order ε^{-1}, and therefore the ε-exponent must be at least 1. The weights γ_j which satisfy the assumption (17) or (18) in Theorem 2 cancel the effect of $c_d(\ln \varepsilon^{-1})^{(d-1)/2}$ at the possible expense of increasing ε^{-1} to ε^{-2}.

Observe that for the ε- and d-exponents of QMC-tractability we only know bounds, and these bounds differ by a factor of 2. The bounds for the d-exponent depend on a given in Theorem 2(ii). Clearly, the d-exponent goes to infinity with a. It is easy to show that a can take any nonnegative value. For example, define $\gamma_j = 1/(2j\sqrt{\ln j}\,)$ for $j \geq 2$. Then $\sum_{j=1}^d \gamma_j = (\ln d)^{1/2}(1 + o(1))$ as $d \to \infty$, and hence, $a = 0$. Although the d-exponent is zero, multivariate integration is not strongly QMC-tractable since the series $\sum_{j=1}^\infty \gamma_j$ is not convergent and (17) is not satisfied. This is an example which we mentioned in Section 2. If we want to obtain a positive a in (ii) of Theorem 2 we may define $\gamma_j = 1$ for $j \leq \lceil a \rceil$ and $\gamma_{j+1} = a \ln(1 + 1/j)$ for $j \geq \lceil a \rceil$. Then the sequence γ is nonincreasing and $\sum_{j=1}^d \gamma_j = a \ln d(1 + o(1))$ as $d \to \infty$.

The last point (iii) of Theorem 2 states of the role of the sum $s_d(\gamma) = \sum_{j=1}^d \gamma_j$. As already mentioned in the Introduction, the minimal number of sample points depends exponentially on $s_d(\gamma)$, so we are in trouble if $s_d(\gamma)$ is large. Observe that the bound in (iii) does not really address the dependence on ε. This will be done later.

5. TRACTABILITY FOR AVERAGE SAMPLE POINTS

In this section we define a new kind of tractability. To motivate this concept, suppose we want to show that multivariate integration in a certain space F_d

is QMC-tractable (or strongly QMC-tractable). This means showing that there exists a quasi-Monte Carlo algorithm which is tractable (or strongly tractable). However, it does not necessarily mean that we must know how to construct such an algorithm. In fact, this is the situation in Theorem 2 for the space $F_{d,\gamma}$, where we state necessary and sufficient conditions for QMC-tractability and strong QMC-tractability without having an explicit construction of tractable and strongly tractable quasi-Monte Carlo algorithms. Our proof of Theorem 2 rests on showing, under appropriate conditions, that even a uniformly random selection of sample points leads on the average to tractable or strongly tractable quasi-Monte Carlo algorithms.

To define precisely the notion of tractability for average sample points, we proceed as follows. We take random sample points t_i which are independent and uniformly distributed over $[0, 1]^d$. For each such resulting n-tuple of sample points $t_1, t_2, \ldots, t_n$ we take the corresponding quasi-Monte Carlo algorithm $Q_{n,d} = Q_{n,d}(t_1, t_2, \ldots, t_n)$ given by (4) and determine the *worst case* error $e(Q_{n,d}(t_1, t_2, \ldots, t_d))$ defined by (5). We then average over these n-tuples by taking the L_2-norm of these errors,

$$e_n^{\text{avg}}(F_d) = \left(\int_{[0,1]^{dn}} e(Q_{n,d}(t_1, t_2, \ldots, t_n))^2 \, dt_1 \, dt_2 \cdots dt_n\right)^{1/2}. \tag{21}$$

We say that multivariate integration in the space F_d is *QMC-tractable for average sample points* iff there exist nonnegative C, q, and p such that

$$\min\{n\colon\ e_n^{\text{avg}}(F_d) \le \varepsilon e(Q_{0,d})\} \le C d^q \varepsilon^{-p} \qquad \forall d = 1, 2, \ldots;\ \forall \varepsilon \in (0, 1). \tag{22}$$

If $q = 0$ we say that multivariate integration in the space F_d is *strongly QMC-tractable for average sample points*. The infima of q and p for which (22) holds are called the d and *ε-exponents* of tractability for average sample points.

We stress that the use of random sample points in the concept of tractability for average sample points is different from the use of random sample points in the Monte Carlo algorithms. In the Monte Carlo algorithms the average is taken for a *fixed* function f, not for the worst case error as we did in (21).

Obviously, tractability for average sample points implies tractability, but the converse need not be true. The d- and ε-exponents for tractability are not greater than the corresponding d- and ε-exponents for tractability for average sample points. Furthermore, QMC-tractability implies the existence of at least one quasi-Monte Carlo algorithm that is tractable, whereas QMC-tractability for average sample points implies that there are many quasi-Monte Carlo algorithms that are tractable. Indeed, Chebyshev's inequality yields

$$\lambda(\{(t_1, t_2, \ldots, t_n)\colon\ e(Q_{n,d}(t_1, t_2, \ldots, t_n)) \le m e_n^{\text{avg}}(F_d)\}) \ge 1 - m^{-2}.$$

Here, λ is the Lebesgue measure and m is an arbitrary positive number. For instance, take $m = 10$ and assume that (22) holds. Then the measure of quasi-Monte Carlo algorithms that are tractable (and for which (22) holds with C replaced by $10^p C$) is at least 0.99.

For a number of cases we will be able to show tractability or strong tractability for average sample points. Although there are then many quasi-Monte Carlo algorithms which are tractable or strongly tractable, it is not clear how to construct such sample points. One possibility is as follows. Assume that we may compute the error $e(Q_{n,d}(t_1, t_2, \ldots, t_n))$ in time proportional to $n^2 d$. If both n and d are not too large then computation of the error is feasible. Now we select $n = n(\varepsilon, d)$ such that the error of random quasi-Monte Carlo algorithms is at most, say $\varepsilon/2$. Due to tractability, such n should not be too large for reasonable d and ε. Next we choose uniformly random t_i; see [2, 13] for information on how this can be done computationally. For the n-tuple $t_1, t_2, \ldots, t_n$ we check whether $e(Q_{n,d}(t_1, t_2, \ldots, t_n)) \le \varepsilon$. If so we are done. If not we repeat the selection of sample points $\{t_i\}$. We will need a relatively small number of such selections, since the average error is at most $\varepsilon/2$. In Section 6 we will see that, indeed, the error $e(Q_{n,d}(t_1, t_2, \ldots, t_n))$ can be computed for a number of cases.

We now present the main theorem on tractability and strong tractability for average sample points for the space $F_{d,\gamma}$.

THEOREM 3. *The concepts of tractability and tractability for average sample points are equivalent in the space $F_{d,\gamma}$. More precisely,*

(i) *multivariate integration in the space $F_{d,\gamma}$ is strongly QMC-tractable iff multivariate integration in the space $F_{d,\gamma}$ is strongly QMC-tractable for average sample points.*

(ii) *multivariate integration in the space $F_{d,\gamma}$ is QMC-tractable iff multivariate integration in the space $F_{d,\gamma}$ is QMC-tractable for average sample points.*

Necessary and sufficient conditions for the two cases are given by (17) *and* (18), *respectively. If we have QMC-tractability then the d- and ε-exponents for tractability for average sample points are $a/6$ and,* 2, *respectively, where a is given in* (18).

This is a restatement of Corollary 10 in Section 6.3, which proves strong QMC-tractability (or QMC-tractability) for average sample points under the condition (17) (or (18)) of Theorem 2. (Corollary 10 is also used in the proof of Theorem 2 to prove the sufficiency of the condition (17) or (18).)

6. TRACTABILITY IN HILBERT SPACES

In this section we study quasi-Monte Carlo algorithms in certain Hilbert spaces F_d. We present lower bounds on the error of quasi-Monte Carlo algorithms and compute the error of quasi-Monte Carlo algorithms for average sample points. From the former we deduce necessary conditions on strong QMC-tractability and QMC-tractability of multivariate integration in the space F_d. From the latter we deduce conditions on tractability for average sample points. The results will be illustrated for the space $F_{d,\gamma}$.

6.1. *Hilbert Spaces with Reproducing Kernels*

We assume that F_d is a Hilbert space of functions defined over $[0, 1]^d$. Its inner product is denoted by $\langle \cdot, \cdot \rangle_d$, and obviously $\| f \|_d = \langle f, f \rangle_d^{1/2}$. We always assume that F_d is a subset of $L_2([0, 1]^d)$. Since quasi-Monte Carlo algorithms use function values, we need to assume that the linear functional $f(t)$ is continuous for arbitrary $t \in [0, 1]^d$. This is equivalent (see [1]) to the statement that F_d has a reproducing kernel K_d, which is a function defined over $[0, 1]^d \times [0, 1]^d$ such that $K_d(\cdot, t) \in F_d$ for all $t \in [0, 1]^d$ and

$$f(t) = \langle f, K_d(\cdot, t) \rangle_d \quad \forall f \in F_d;\ \forall t \in [0, 1]^d.$$

The reproducing kernel K_d has a number of algebraic properties. For example, the matrix $(K_d(t_i, t_j))$ for any choice of the sample points t_j, $j = 1, 2, \ldots, n$, is symmetric and nonnegative definite, since $K_d(t_i, t_j) = \langle K_d(\cdot, t_j), K_d(\cdot, t_i) \rangle_d$, and hence,

$$\sum_{i,j=1}^{n} K_d(t_j, t_i) a_j a_i = \left\| \sum_{i=1}^{n} K_d(\cdot, t_i) a_i \right\|_d^2 \geq 0 \quad \forall a_i \in \mathbb{R}.$$

The diagonal elements $K_d(t, t) = \| K_d(\cdot, t) \|_d^2$ are nonnegative, and we have

$$| f(t) | \leq \| f \|_d \sqrt{K_d(t, t)} \quad \forall f \in F_d;\ \forall t \in [0, 1]^d.$$

Obviously, $L_2([0, 1]^d)$ is not a Hilbert space with a reproducing kernel, since the functional $f(t)$ is not even well defined for $L_2([0, 1]^d)$. Hence, F_d is a proper subset of $L_2([0, 1]^d)$.

EXAMPLE. The space $F_{d,\gamma}$. Take $F_d = F_{d,\gamma}$ from (15). Noting the norm $\| \cdot \|_{d,\gamma}$ defined by (12), it is apparent that the space $F_{d,\gamma}$ is a Hilbert space with the inner product

$$\langle f, g \rangle_{d,\gamma} = \sum_{u \subset D} \gamma_u^{-1} \int_{[0,1]^{|u|}} \frac{\partial^{|u|}}{\partial x_u} f(x_u, 1) \frac{\partial^{|u|}}{\partial x_u} g(x_u, 1)\, dx_u. \tag{23}$$

It is easy to check that $F_{d,\gamma}$ has the reproducing kernel

$$K_{d,\gamma}(x, t) = \prod_{j=1}^{d} (1 + \gamma_j \min(1 - x_j, 1 - t_j)). \tag{24}$$

For completeness, we provide a short proof of this fact in the Appendix, Section A.2. Here we only remark that $K_{d,\gamma}(x, t)$ is always nonnegative (in fact, $K_{d,\gamma}(x, t) \geq 1$). It turns out that for nonnegative reproducing kernels we can find a necessary condition for QMC-tractability and strong QMC-tractability of multivariate integration.

We return to a general Hilbert space F_d with a reproducing kernel K_d. Since F_d is a subset of $L_2([0, 1]^d)$, the function $K_d(\cdot, t)$ belongs to $L_2([0, 1]^d)$ for arbitrary $t \in [0, 1]^d$. We now additionally assume that $K_d \in L_1([0, 1]^d \times [0, 1]^d)$. Consider now the multivariate integration functional I_d given by (3). It is easy to see that

$$I_d(f) = \langle f, h_d \rangle_d \quad \text{with } h_d(x) = \int_{[0,1]^d} K_d(x, t)\, dt, \tag{25}$$

so that h_d is the representer of multivariate integration, and, by use of the reproducing kernel property,

$$\begin{aligned} e(Q_{0,d}) = \|I_d\| = \|h_d\|_d &= \left(\int_{[0,1]^{2d}} K_d(x, t)\, dx\, dt \right)^{1/2} \\ &= \left(\int_{[0,1]^d} h_d(x)\, dx \right)^{1/2}. \end{aligned} \tag{26}$$

In passing, note that we have

$$\int_{[0,1]^{2d}} K_d(x, t)\, dx\, dt = \int_{[0,1]^d} h_d(x)\, dx \geq 0. \tag{27}$$

Due to our assumption on K_d, the norm of h_d is finite and I_d is a linear continuous functional. In general, it may happen that $\|I_d\| = 0$. Then multivariate integration is trivial, and since $e(Q_{0,d}) = 0$ it is obviously strongly QMC-tracable. To avoid this case, we will assume that $\|I_d\|$ is positive for all d, in which case the inequality in (27) is strict.

Clearly, for any quasi-Monte Carlo algorithm $Q_{n,d}$ we have for the error

$$I_d(f) - Q_{n,d}(f) = \langle f, g_d \rangle_d,$$

where the representer g_d of the error is

$$g_d = h_d - \frac{1}{n} \sum_{i=1}^{n} K_d(\cdot, t_i).$$

Hence, the error of $Q_{n,d}$ can be expressed as

$$e(Q_{n,d}) = \sup_{f \in F_d,\, \|f\|_d \le 1} |I_d(f) - Q_{n,d}(f)| = \|g_d\|_d$$
$$= \left\| h_d - \frac{1}{n} \sum_{i=1}^{n} K_d(\cdot,\, t_i) \right\|_d .$$

Since the norm $\|\cdot\|_d$ is $\langle \cdot, \cdot \rangle_d^{1/2}$, and since $\langle h_d,\, K_d(\cdot,\, t_i)\rangle_d = h_d(t_i)$ and $\langle K_d(\cdot,\, t_i),\, K_d(\cdot,\, t_j)\rangle_d = K_d(t_i,\, t_j)$, we obtain

$$e(Q_{n,d})^2 = \|h_d\|_d^2 - \frac{2}{n} \sum_{i=1}^{n} h_d(t_i) + \frac{1}{n^2} \sum_{i,\,j=1}^{n} K_d(t_i,\, t_j), \tag{28}$$

with h_d and $\|h_d\|_d$ given by (25) and (26).

EXAMPLE. The space $F_{d,\gamma}$ (continued). For the space $F_{d,\gamma}$, the reproducing kernel $K_{d,\gamma}$ is given by (24). Thus multivariate integration has the representer function $h_d = h_{d,\gamma}$ of the form

$$h_{d,\gamma}(x) = \int_{[0,\,1]^d} K_{d,\gamma}(x,\, t)\, dt = \prod_{j=1}^{d} \left(1 + \frac{1}{2}\gamma_j (1 - x_j^2)\right)$$

and

$$e(Q_{0,d})^2 = \|I_d\|^2 = \|h_{d,\gamma}\|_{d,\gamma}^2 = \int_{[0,\,1]^{2d}} K_{d,\gamma}(x,\, t)\, dx\, dt$$
$$= \prod_{j=1}^{d} \left(1 + \frac{1}{3}\gamma_j\right). \tag{29}$$

The error of $Q_{n,d}$ depends on the representer

$$g_{d,\gamma}(x) := h_{d,\gamma}(x) - \frac{1}{n} \sum_{i=1}^{n} K_{d,\gamma}(x,\, t_i),$$

which is now equal to

$$g_{d,\gamma}(x) = \prod_{j=1}^{d} \left(1 + \frac{1}{2}\gamma_j (1 - x_j^2)\right) - \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{d} (1 + \gamma_j \min(1 - x_j,\, 1 - t_{i,\,j})),$$

where $t_i = [t_{i,1},\, t_{i,2},\, \ldots,\, t_{i,d}]$.

By Theorem 1 we have

$$e(Q_{n,d} = \|g_{d,\gamma}\|_{d,\gamma} = \mathrm{disc}_\gamma(\{t_i\}).$$

6.2. *Lower Bounds*

We now present lower bounds on the error of quasi-Monte Carlo algorithms for a class of reproducing kernel Hilbert space F_d. These lower bounds will be useful for deriving necessary conditions for strong QMC-tractability and QMC-tractability of multivariate integration.

For a general Hilbert space F_d with reproducing kernel K_d, let

$$\kappa_d = \sup_{x\in[0,1]^d} \frac{|h_d(x)|}{\sqrt{K_d(x,x)}}\,\frac{1}{\|h_d\|_d}, \tag{30}$$

with the convention that 0/0 = 0. In passing, we note that it is shown in [19] that the minimal error of the quadrature formula that uses only one function value is $e(Q_{0,d})\sqrt{1-\kappa_d^2}$. Clearly, $\kappa_d \le 1$ since

$$|h_d(x)| = |\langle h_d, K_d(\cdot,x)\rangle_d| \le \|h_d\|_d\sqrt{K_d(x,x)}.$$

We are ready to prove the following lemma.

LEMMA 4. *If the reproducing kernel K_d of the space F_d is nonnegative then for an arbitrary quasi-Monte Carlo algorithm $Q_{n,d}$ we have*

$$e(Q_{n,d})^2 \ge e(Q_{0,d})^2(1-n\kappa_d^2).$$

Hence, $e(Q_{n,d}) \le \varepsilon e(Q_{0,d})$ implies that

$$n \ge (1-\varepsilon^2)\kappa_d^{-2}.$$

Proof. We make use of the identity (28). From the definition of κ_d we have $|h_d(t_i)| \le \kappa_d\|h_d\|_d\sqrt{K_d(t_i,t_i)}$. Thus by the Cauchy–Schwarz inequality

$$\left|\frac{1}{n}\sum_{i=1}^n h_d(t_i)\right| \le \kappa_d\|h_d\|_d\,\frac{1}{n}\sum_{i=1}^n\sqrt{K_d(t_i,t_i)} \le \kappa_d\|h_d\|_d\beta,$$

with $\beta = \sqrt{1/n\sum_{i=1}^n K_d(t_i,t_i)}$.

Since by assumption $K_d(t_i,t_j) \ge 0$, we can estimate the double sum $\sum_{i,j=1}^n K_d(t_i,t_j)$ from below by the sum of the diagonal elements $\sum_{i=1}^n K_d(t_i,t_i) = n\beta^2$. Thus it follows from (28) that

$$e(Q_{n,d}^2) \ge \|h_d\|_d^2 - 2\kappa_d\|h_d\|_d\beta + \frac{1}{n}\beta^2.$$

We minimize the right-hand side with respect to β. The minimum is achieved for $\beta = n\kappa_d \|h_d\|_d$, yielding

$$e(Q_{n,\,d})^2 \geq \|h_d\|_d^2(1 - n\kappa_d^2) = e(Q_{0,\,d})^2(1 - n\kappa_d^2),$$

as claimed. ∎

Lemma 4 presents a lower bound on the error of any quasi-Monte Carlo algorithm in a space F_d with nonnegative kernel. We do not know if the assumption on nonnegativity of the kernel is essential. The essence of the lower bound is that for small κ_d the number n of sample points must be large, since otherwise the error $e(Q_{n,\,d})$ is close to the initial error $e(Q_{0,\,d})$.

From Lemma 4 it is easy to deduce necessary conditions on strong QMC-tractability and QMC-tractability of multivariate integration in terms of the behavior of κ_d as a function of d.

THEOREM 5. *Let the space F_d have a nonnegative reproducing kernel K_d, and let $\|I_d\| > 0$ for all d:*

(i) *If*

$$\liminf_{d\to\infty} \kappa_d = 0,$$

then multivariate integration in the space F_d is not strongly QMC-tractable.

(ii) *If*

$$\liminf_{d\to\infty} d^k\kappa_d = 0 \quad \forall\, k \geq 0,$$

then multivariate integration in the space F_d is not QMC-tractable.

Proof. Suppose that multivariate integration in the space F_d is strongly QMC-tractable (or QMC-tractable). Then there exists a family of quasi-Monte Carlo algorithms $\{Q_{n,\,d}\}$ which is strongly tractable (or tractable). This means that for an arbitrary ε, $\varepsilon \in (0, 1)$, we need to perform $n = n(\varepsilon, d)$ function values to guarantee that $e(Q_{n,\,d}) \leq \varepsilon e(Q_{0,\,d})$ and $n(\varepsilon,\, d) \leq Cd^q\varepsilon^{-p}$ for some positive C, q, and p independent of d and ε, and with $q = 0$ in the case of strong QMC-tractability. Since Lemma 4 is valid for any quasi-Monte Carlo algorithm, we apply it to $\{Q_{n,\,d}\}$ and we have

$$(1 - \varepsilon^2)\kappa_d^{-2} \leq n(\varepsilon,\, d) \leq Cd^q\varepsilon^{-p}.$$

We fix $\varepsilon \in (0, 1)$. If (i) is satisfied then the left-hand side is unbounded as $d \to \infty$. This implies that q cannot be equal to zero, which contradicts strong tractability. If (ii) is satisfied then we have

$$(1 - \varepsilon^2)\kappa_d^{-2}d^{-q} \leq n(\varepsilon,\, d)d^{-q} \leq C\varepsilon^{-p}.$$

Once more, the left-hand side is unbounded, whereas the right-hand side is bounded. This contradiction proves the lack of tractability. ■

We now check the assumptions of Theorem 5 for simple tensor product spaces $F_d = F_1 \otimes F_1 \otimes \cdots \otimes F_1$. Here F_1 is a Hilbert space with the reproducing kernel K_1. Then the reproducing kernel K_d of F_d is given by

$$K_d(x, y) = \prod_{j=1}^{d} K_1(x_j, y_j).$$

In this case, $\|h_d\|_d = \|h_1\|_1^d$, and $\kappa_d = \kappa_1^d$. Obviously $\kappa_1 \le 1$. As shown in [19], $\kappa_1 = 1$ means that we can solve multivariate integration with arbitrarily small ε by just using only one function value. This means that multivariate problem is trivial. To omit trivial cases we need to assume that $\|I_1\| > 0$ and $\kappa_1 < 1$. Then $d^k \kappa_d = d^k \kappa_1^d$ goes to zero for all k. Due to Theorem 5 we have the following corollary.

COROLLARY 6. *If the space F_1 has a nonnegative reproducing kernel K_1, and $\|I_1\| > 0$ and $\kappa_1 < 1$, then multivariate integration in the simple tensor product space $F_d = F_1 \otimes F_1 \otimes \cdots \otimes F_1$ is not QMC-tractable. In this case, any quasi-Monte Carlo algorithm must use exponentially many sample points n,*

$$n \ge (1 - \varepsilon^2)\kappa_1^{-2d},$$

to guarantee that $e(Q_{n,d}) \le \varepsilon e(Q_{0,d})$.

Observe that the space $F_{d,\gamma}$ is a tensor product space, $F_{d,\gamma} = F_{1,\gamma_1} \otimes F_{1,\gamma_2} \otimes \cdots \otimes F_{1,\gamma_d}$ with F_{1,γ_j} being a Hilbert space with the reproducing kernel K_{1,γ_j}. The space $F_{d,\gamma}$ is a simple tensor product space iff $\gamma_j = 1$ for all $j \ge 1$. In this case, we do not have QMC-tractability.

Remark 2: *Tractability for Arbitrary Algorithms.* We stress that the lack of QMC-tractability of multivariate integration does not necessarily mean that multivariate integration is intractable. It may be that the use of algorithms which are not quasi-Monte Carlo is very effective, allowing the reduction of the initial error by a factor of ε to be achieved by a polynomial in d of function values. More precisely (see [19], Theorem 5, and Remark 3), for any integer $p > 1$ there exist Hilbert spaces F_1 for which $n = \binom{d+p-1}{p-1} = \Theta(d^{p-1})$ function values are enough to compute exactly the multivariate integral $I_d(f)$ for all $f \in F_d = F_1 \otimes F_1 \otimes \cdots \otimes F_1$. We add that the spaces F_1 for which we know this to hold are rather esoteric.

An interesting problem is to investigate for which spaces the use of quasi-Monte Carlo algorithms is not restrictive. In particular, for which spaces are QMC-tractability and tractability of multivariate integration equivalent.

We now check the assumptions of Theorem 5 for the space $F_{d,\gamma}$ for arbitrary sequences γ. We translate the conditions on κ_d in terms of the behavior of the sequence γ.

COROLLARY 7.

(i) *If*

$$\sum_{j=1}^{\infty} \gamma_j = \infty$$

then multivariate integration in the space $F_{d,\gamma}$ is not strongly QMC-tractable.

(ii) *If*

$$\limsup_{d\to\infty} \frac{\sum_{j=1}^{d} \gamma_j}{\ln d} = \infty$$

then multivariate integration in the space $F_{d,\gamma}$ is not QMC-tractable.

Proof. For the space $F_{d,\gamma}$ we have $\kappa_d = \kappa_{d,\gamma}$ with

$$\begin{aligned}\kappa_{d,\gamma} &= \max_{x\in[0,1]^d} \prod_{j=1}^{d} \frac{1+\frac{1}{2}\gamma_j(1-x_j^2)}{\sqrt{(1+\gamma_j(1-x_j))\left(1+\frac{1}{3}\gamma_j\right)}}\\ &= \prod_{j=1}^{d} \max_{x\in[0,1]} \frac{1+\frac{1}{2}\gamma_j(1-x^2)}{\sqrt{(1+\gamma_j(1-x))\left(1+\frac{1}{3}\gamma_j\right)}}. \end{aligned} \tag{31}$$

Suppose first that γ_j do not tend to zero. Since they are nonincreasing, this means that there exists a positive γ_0 such that $\gamma_j \geq \gamma_0 > 0$ for all $j \geq 1$. It is easy to show that there exists a positive number $\rho = \rho(\gamma_0)$ such that

$$\rho := \max_{y\in[\gamma_0,1]} \max_{x\in[0,1]} \frac{1+\frac{1}{2}y(1-x^2)}{\sqrt{(1+y(1-x))\left(1+\frac{1}{3}y\right)}} < 1.$$

Indeed, if we define

$$F_x(y) := \frac{1+\frac{1}{2}y(1-x^2)}{\sqrt{(1+y(1-x))\left(1+\frac{1}{3}y\right)}}, \quad x, y \in [0,1],$$

it is enough to observe that

$$F_x(0) = 1, \quad x \in [0,1]; \quad F_x'(y) < 0, \quad x, y \in [0,1].$$

Then $\kappa_{d,\gamma} \leq \rho^d$, and $d^k \kappa_{d,\gamma}$ goes to zero for all k. Hence, we do not have QMC-tractability.

Assume then that $\lim_j \gamma_j = 0$. For small y it is easy to show that

$$\max_{x\in[0,\,1]} F_x(y) = 1 - \frac{1}{24}\, y + O(y^2).$$

Hence, $\kappa_{d,\,\gamma} = \Theta(q_{d,\,\gamma})$, where

$$\begin{aligned} q_{d,\,\gamma} &= \prod_{j=1}^{d}\left(1 - \frac{1}{24}\,\gamma_j\right) = \exp\left(\sum_{j=1}^{d} \ln(1-\gamma_j/24)\right) \\ &= \Theta\left(\exp\left(-\frac{1}{24}\sum_{j=1}^{d}\gamma_j\right)\right). \end{aligned} \tag{32}$$

If $\sum_{j=1}^{\infty}\gamma_j = \infty$ then we have $\lim_{d\to\infty}\kappa_{d,\,\gamma} = 0$ and the lack of strong QMC-tractability. If $\limsup_{d\to\infty}\sum_{j=1}^{d}\gamma_j/\ln d = \infty$ then

$$\ln d^k \kappa_{d,\,\gamma} = k\,\ln d - \frac{1}{24}\sum_{j=1}^{d}\gamma_j + O(1),$$

which tend to $-\infty$ for some subsequence of d_i. Hence, $\liminf_d d^k\kappa_{d,\,\gamma} = 0$ for all k, and we do not have QMC-tractability. ∎

6.3. *Error for Average Sample Points*

We now compute the error of quasi-Monte Carlo algorithms for average sample points. From this we derive conditions on strong QMC-tractability and QMC-tractability for average sample points. These conditions will also be sufficient conditions for strong QMC-tractability and QMC-tractability of multivariate integration.

For a general Hilbert space F_d with reproducing kernel K_d we assume that $\|I_d\| > 0$. From (26) this means that $\int_{[0,\,1]^{2d}} K_d(x,\,t)\,dx\,dt$ is positive. Let

$$\rho_d = \frac{\displaystyle\int_{[0,\,1]^d} K_d(x,\,x)\,dx}{\displaystyle\int_{[0,\,1]^{2d}} K_d(x,\,t)\,dx\,dt}. \tag{33}$$

We now show that

$$\rho_d \ge \kappa_d^{-2} \ge 1, \tag{34}$$

where κ_d is given by (30). Indeed, the definition of κ_d and (26), together with the second part of (25), yield

$$\int_{[0,1]^d} K_d(x, t)\,dt \le \kappa_d K_d(x, x)^{1/2} \left(\int_{[0,1]^{2d}} K_d(x, t)\,dx\,dt \right)^{1/2} \quad \forall x \in [0, 1]^d \tag{35}$$

and, on integrating over x,

$$\int_{[0,1]^{2d}} K_d(x, t)\,dx\,dt \le \kappa_d \int_{[0,1]^d} K_d(x, x)^{1/2}\,dx \left(\int_{[0,1]^{2d}} K_d(x, t)\,dx\,dt \right)^{1/2}. \tag{36}$$

In general, the left-hand side of (35) may be negative for some values of x, but by (27) the left-hand side of (36) is necessarily positive. Thus, on squaring and applying the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}\int_{[0,1]^{2d}} K_d(x, t)\,dx\,dt &\le \kappa_d^2 \left(\int_{[0,1]^d} K_d(x, x)^{1/2}\,dx \right)^2 \\ &\le \kappa_d^2 \int_{[0,1]^d} K_d(x, x)\,dx.\end{aligned}$$

This yields $\rho_d \ge \kappa_d^{-2}$. Since $\kappa_d \le 1$, (34) is shown.

We are ready to compute the error $e_n^{\text{avg}}(F_d)$ for average sample points given by (21).

LEMMA 8. *Let* $\|I_d\| > 0$. *Then*

$$e_n^{\text{avg}}(F_d) = \frac{\sqrt{\rho_d - 1}}{\sqrt{n}}\, e(Q_{0,d}).$$

Proof. Take an arbitrary quasi-Monte Carlo algorithm $Q_{n,d}$. Let t_i be sample points used by $Q_{n,d}$. To stress the role of the sample points, denote $Q_{n,d} = Q_{n,d}(\{t_i\})$. We now rewrite the error of $Q_{n,d}(\{t_i\})$ given by (28),

$$e(Q_{n,d}(\{t_i\}))^2 = \|h_d\|_d^2 - \frac{2}{n}\sum_{i=1}^n h_d(t_i) + \frac{1}{n^2}\sum_{i,j=1}^n K_d(t_i, t_j).$$

As in (21), we integrate this identity over uniformly distributed t_i. Keeping in mind the formulae for h_d and $\|h_d\|_d$ given by (25) and (26), and separating the diagonal and off-diagonal terms of the last sum, we obtain

$$\begin{aligned}e_n^{\text{avg}}(F_d)^2 &= \|h_d\|_d^2 - 2\|h_d\|_d^2 + \frac{1}{n^2}\left(n\int_{[0,1]^d} K_d(x, x)\,dx + (n^2 - n)\|h_d\|_d^2 \right) \\ &= \frac{1}{n}\left(\int_{[0,1]^d} K_d(x, x)\,dx - \int_{[0,1]^{2d}} K_d(x, t)\,dx\,dt \right) \\ &= \frac{\rho_d - 1}{n}\,\|h_d\|_d^2,\end{aligned}$$

which completes the proof, since $\|h_d\| = e(Q_{0,d})$. ■

From Lemma 8 we can immediately deduce conditions for QMC-tractability and strong QMC-tractability for average sample points in terms of the behavior of ρ_d as a function of d.

THEOREM 9. *Let* $\|I_d\| > 0$ *for all* d.

(i) *Multivariate integration in the space* F_d *is strongly QMC-tractable for average sample points iff*

$$\sup_{d=1,2,\dots} \rho_d < \infty.$$

If so, the ε*-exponent is* 2.

(ii) *Multivariate integration in the space* F_d *is QMC-tractable for average sample points iff there exists a nonnegative* k *for which*

$$\sup_{d=1,2,\dots} d^{-k} \rho_d < \infty. \tag{37}$$

If so, the d*-exponent is the infimum of* k *satisfying* (37), *and the* ε*-exponent is* 2.

Proof. From Lemma 8 we conclude that

$$\min\{n\colon\; e_n^{\mathrm{avg}}(F_d) \le \varepsilon e(Q_{0,d})\} = \left\lceil \frac{\rho_d - 1}{\varepsilon^2} \right\rceil. \tag{38}$$

Hence, we get strong QMC-tractability for average sample points iff $\sup_d \rho_d$ is bounded. We get QMC-tractability for average sample points iff $\sup_d d^{-k}\rho_d$ is bounded for some k. In this case the d-exponent is the infimum of such k. In both cases the ε-exponent is 2. ■

We stress that the conditions of Theorem 9 are sufficient conditions for QMC-tractability and QMC-tractability of multivariate integration in the space F_d.

As in the previous subsection, we check now the assumptions of Theorem 9 for simple tensor product spaces $F_d = F_1 \otimes F_1 \otimes \cdots \otimes F_1$. Clearly, we now have $\rho_d = \rho_1^d$. Since $\rho_1 \ge 1$, the only case for which (i) or (ii) of Theorem 9 holds is $\rho_1 = 1$. However, $\rho_1 = 1$ and (34) imply that $\kappa_1 = \kappa_d = 1$. As already remarked, this means that multivariate integration is trivial, since one function value is enough to solve the problem for an arbitrary positive ε. Excluding this case, we have $\rho_1 > 1$ and the assumptions of Theorem 9 do *not* hold for simple tensor product spaces. Hence, we do not have even QMC-tractability for average sample points in a simple tensor product space F_d.

We now check the assumptions of Theorem 9 for the space $F_{d,\gamma}$. We translate the conditions on ρ_d in terms of the behavior of the sequence γ.

COROLLARY 10.

(i) *Multivariate integration in the space* $F_{d,\gamma}$ *is strongly QMC-tractable for average sample points iff*

$$\sum_{j=1}^{\infty} \gamma_j < \infty.$$

If so, the ε*-exponent is* 2.

(ii) *Multivariate integration in the space* $F_{d,\gamma}$ *is QMC-tractable for average sample points iff*

$$a := \limsup_{d\to\infty} \frac{\sum_{j=1}^{d} \gamma_j}{\ln d} < \infty.$$

If so, the d*-exponent is* $a/6$ *and the* ε*-exponent is* 2.

Proof. For the space $F_{d,\gamma}$ we have

$$\rho_d = \prod_{j=1}^{d} \frac{1+\frac{1}{2}\gamma_j}{1+\frac{1}{3}\gamma_j} = \exp\left(\sum_{j=1}^{d} \ln \frac{1+\frac{1}{2}\gamma_j}{1+\frac{1}{3}\gamma_j}\right) \le \exp\left(\frac{1}{6}\sum_{j=1}^{d}\gamma_j\right)$$
$$= d^{1/6 \sum_{j=1}^{d}\gamma_j/\ln d}. \tag{39}$$

Hence, finiteness of the sum $\sum_{j=1}^{\infty}\gamma_j$ implies that ρ_d is bounded and we get strong QMC-tractability for average sample points. Similarly, a finite a implies that for any positive δ the sequence $d^{-a/6-\delta}\rho_d$ is bounded. Thus we get QMC-tractability for average sample points with the d-exponent no greater than $a/6$.

On the other hand, strong tractability for average sample points implies that ρ_d is bounded. Then $\sum_j \gamma_j$ must be finite. In turn, tractability for average sample points implies that $d^{-k}\rho_d$ is bounded for some k. Then $\lim_j \gamma_j = 0$ since otherwise ρ_d goes exponentially fast to infinity with d. For $\lim_j \gamma_j = 0$ we have

$$\rho_d = \Theta\left(\exp\left(\frac{1}{6}\sum_{j=1}^{d}\gamma_j\right)\right)$$

and

$$d^{-k}\rho_d = \Theta\big(d^{-k} d^{1/6\sum_{j=1}^{d}\gamma_j/\ln d}\big).$$

This implies $\limsup_d \sum_{j=1}^{d}\gamma_j/\ln d \le 6k$ and $a/6 \le k$. Since k can be arbitrarily close to the d-exponent, $a/6$ is no greater than the d-exponent. Hence, $a/6$ is the d-exponent. This completes the proof. ■

We end this section by a remark on the classical Monte Carlo algorithm.

Remark 3: *Tractability of the Classical Monte Carlo Algorithm.* Although in this paper we study only quasi-Monte Carlo algorithms, as a byproduct of our analysis we can also check when the classical Monte Carlo algorithm is tractable. Monte Carlo is a randomized algorithm and its error $e_{n,d}(f)$ for each function f is defined as the expected L_2-error with respect to random sample points. It is well known that

$$e_{n,d}(f) = \frac{1}{\sqrt{n}} V(f),$$

where $V(f)$ is the variance of f and is given by

$$V(f) = \left(I_d(f^2) - I_d^2(f)\right)^{1/2}.$$

Hence, the classical Monte Carlo algorithm is strongly tractable or tractable if we can find a uniform bound on the variances of f from the class F_d with $\|f\|_d \le 1$.

Clearly, $e_{n,d}(f)$ is no greater than $e_n^{\text{avg}}(F_d)$ from (21) since the latter is defined for a worst f from the unit ball of F_d for each quadrature point set. Thus,

$$e_{n,d}(f) \le e_n^{\text{avg}}(F_d).$$

From Lemma 8 we conclude that

$$\sup_{f \in F_d,\, \|f\|_d \le 1} e_{n,d}(f) \le \frac{\sqrt{\rho_d - 1}}{\sqrt{n}} \|I_d\|.$$

Hence, strong tractability and tractability of the classical Monte Carlo algorithm hold under the same conditions on ρ_d as in Theorem 9.

APPENDIX

In this appendix we provide proofs of a number of the results which were stated or used in the paper.

A.1. *Limiting Discrepancy*

We prove (14). Assume that $\sum_{j=1}^{\infty} \gamma_j = \infty$. Then the proof of Corollary 7 states that $\lim_d \kappa_{d,\gamma} = 0$. For any sample points $t_i^{(\infty)}$, Theorem 1 and Lemma 4 yield

$$\operatorname{disc}_\gamma(\{t_i^{(d)}\}) \ge e(Q_{0,d})(1 + o(1)), \quad \text{as } d \to \infty.$$

Since $e(Q_{0,d}) = \|h_{d,\gamma}\|_d = \prod_{j=1}^d (1+\gamma_j/3)^{1/2}$ goes to infinity with d, we conclude that

$$\operatorname{disc}_\gamma(\{t_i^{(\infty)}\}) = \lim_{d\to\infty} \operatorname{disc}_\gamma(\{t_i^{(d)}\}) = \infty.$$

On the other hand, assume that $\sum_{j=1}^\infty \gamma_j < \infty$. From the definition of the discrepancy function disc we have $\operatorname{disc}(x) \in [-1, 1]$ for all $x \in [0, 1]^d$, from which follows

$$\operatorname{disc}_\gamma(\{t_i^{(d)}\})^2 \le \sum_{\emptyset\neq u\subset D} \gamma_u.$$

We now prove by induction on d that

$$\sum_{\emptyset\neq u\subset D} \gamma_u = \sum_{j=1}^d \gamma_j \prod_{k=j+1}^d (1+\gamma_k) \le \left(\sum_{j=1}^d \gamma_j\right) \exp\left(\sum_{j=1}^d \gamma_j\right). \tag{40}$$

Let $A_d = \sum_{\emptyset\neq u\subset D} \gamma_u$. For $d = 1$ we have $A_1 = \gamma_1$. Consider now A_d for $d \ge 2$. We have two kinds of terms in the sum which defines A_d. The first kind corresponds to nonempty u's which are subsets of $\{1, 2, \ldots, d-1\}$. The sum of such terms is A_{d-1}. The second kind of term corresponds to $\{u, d\}$. Here u can either be empty or u is a nonempty subset of $\{1, 2, \ldots, d-1\}$. For $u = \emptyset$ we have the term γ_d, and the sum over $\{u, d\}$ for nonempty subsets u of $\{1, 2, \ldots, d-1\}$ is $\gamma_d A_{d-1}$. Hence,

$$A_d = (1+\gamma_d)A_{d-1} + \gamma_d.$$

Using the inductive assumption for A_{d-1} we get the first equality in (40). The inequality in (40) follows from the fact that

$$\prod_{k=j+1}^d (1+\gamma_k) \le \prod_{k=1}^d (1+\gamma_k) = \exp\left(\sum_{k=1}^d \ln(1+\gamma_k)\right).$$

Since $\ln(1+\gamma_k) \le \gamma_k$, we have (40).

Hence, $\sum_{j=1}^\infty \gamma_j < \infty$ implies that $\sum_{\emptyset\neq u\subset D} \gamma_u$ is bounded, and the limit of $\operatorname{disc}_\gamma(\{t_i^{(d)}\})$ is finite, as claimed.

A.2. *Reproducing the Kernel of $F_{d,\gamma}$*

We show that $K_{d,\gamma}$ given by (24) is a reproducing kernel of the Hilbert space $F_{d,\gamma}$ given by (15). We need to show that for any $f \in F_{d,\gamma}$ and any $t \in [0, 1]^d$ we have

$$\langle f, K_{d,\gamma}(\cdot, t)\rangle_{d,\gamma} = f(t). \tag{41}$$

First of all observe from (24) that

$$K_{d,\gamma}(\mathbf{1},\, t) = 1 \quad \forall\, t \in [0,\, 1]^d,$$

$$K_{d,\gamma}((x_u,\, 1),\, t) = \prod_{j\in u} (1 + \gamma_j\, \min(1 - x_j,\, 1 - t_j)) \quad \forall\, x,\, t \in [0,\, 1]^d;\ u \subset D.$$

Here, $\mathbf{1} = [1, 1, \ldots, 1]$. Observe also that the derivative of $\min(1 - x,\, 1 - t)$ with respect to x is zero for $x < t$, does not exist for $x = t$, and is -1 for $x > t$.

We prove (41) by induction on d. For $d = 1$ we have, from (23),

$$\begin{aligned}\langle f,\, K_{1,\gamma}(\cdot,\, t)\rangle_{1,\gamma} &= f(1)K_{1,\gamma}(1,\, t) + \gamma_1^{-1}\gamma_1 \int_0^1 f'(x)\,(\min(1 - x,\, 1 - t))'\, dx\\ &= f(1) - \int_t^1 f'(x)\, dx = f(t),\end{aligned}$$

as claimed.

Suppose that (41) holds for d, and check it for $d + 1$. Let $t = [\tau,\, t_{d+1}]$ with $\tau \in [0, 1]^d$. The sets u are now subsets of $\{1, 2, \ldots, d + 1\}$. We can first consider the sets u which are subsets of $D = \{1, 2, \ldots, d\}$ and then the sets $\{u, d + 1\}$ with u being once more a subset of D. Hence, we have

$$\langle f,\, K_{d+1,\gamma}(\cdot,\, t)\rangle_{d+1,\gamma} = T_1 + T_2,$$

where

$$\begin{aligned}T_1 &= \sum_{u\subset D} \gamma_u^{-1} \int_{[0,\,1]^{|u|}} \frac{\partial^{|u|}}{\partial x_u} f(x_u,\, 1)\, \frac{\partial^{|u|}}{\partial x_u} K_{d+1,\gamma}(((x_u,\, 1),\, 1),\, t)\, dx_u,\\ T_2 &= \sum_{u\subset D} a_u \int_{[0,\,1]^{|u|}} \frac{\partial^{|u|}}{\partial x_u} \bigg(\int_0^1 \frac{\partial}{\partial x_{d+1}} f((x_u,\, 1),\, x_{d+1})\\ &\quad \times \frac{\partial}{\partial x_{d+1}} \min(1 - x_{d+1},\, 1 - t_{d+1})\, dx_{d+1} \bigg)\\ &\quad \times \frac{\partial^{|u|}}{\partial x_u} K_{d,\gamma}((x_u,\, 1),\, \tau)\, dx_u,\end{aligned}$$

where $a_u = \gamma_{d+1}/(\gamma_u\gamma_{d+1}) = \gamma_u^{-1}$. Note that in T_1 the function $f(x_u,\, 1)$ is a function of at most d variables and that $K_{d+1,\gamma}(((x_u,\, 1),\, 1),\, t) = K_{d,\gamma}((x_u,\, 1),\, \tau)$. Therefore, we can apply the induction assumption to T_1 and claim that $T_1 = f(\tau,\, 1)$.

We now work on T_2. By simple integration with respect to x_{d+1} we get

$$\begin{aligned}T_2 = -\sum_{u\subset\{1,2,\ldots,d\}} \gamma_u^{-1} \int_{[0,\,1]^{|u|}} \frac{\partial^{|u|}}{\partial x_u} (f((x_u,\, 1),\, 1) - f((x_u,\, 1),\, t_{d+1}))\\ \times \frac{\partial^{|u|}}{\partial x_u} K_{d,\gamma}((x_u,\, 1),\, \tau)\, dx_u.\end{aligned}$$

Once more by induction, we conclude that $T_2 = -f(\tau,\ 1) + f(\tau,\ t_{d+1})$. Hence,

$$T_1 + T_2 = f(\tau,\ t_{d+1}) = f(t),$$

as claimed.

A.3. *Proof of Theorem* 2

Assume that multivariate integration in the space $F_{d,\gamma}$ is strongly QMC-tractable. Then (i) of Corollary 7 yields that $\sum_{j=1}^{\infty} \gamma_j < \infty$. Corollary 10 in turn states that the ε-exponent of tractability for average sample points is 2. Hence, the ε-exponent is at most 2. As already mentioned, the ε-exponent must be at least 1. Hence, the ε-exponent of multivariate integration belongs to [1, 2] as claimed.

Assume that multivariate integration is QMC-tractable in the space $F_{d,\gamma}$. Then (ii) of Corollary 7 yields that $a = \limsup_d \sum_{j=1}^{d} \gamma_j / \ln d$ is finite. Again Corollary 10 implies that the ε-exponent belongs to [1, 2]. We now estimate the d-exponent. Clearly, from Lemma 4 we conclude that for any quasi-Monte Carlo algorithm we have

$$n_{\min}(\varepsilon,\ d,\ \{Q_{n,d}\}) \geq (1 - \varepsilon^2)\kappa_{d,\gamma}^{-2}$$

with $\kappa_{d,\gamma}$ given by (31). From (32) we obtain

$$\kappa_{d,\gamma}^{-2} = \Theta\left(e^{1/12 \sum_{j=1}^{d} \gamma_j}\right) = \Theta\left(d^{1/12 \sum_{j=1}^{d} \gamma_j / \ln d}\right).$$

This proves that the d-exponent must be at least $a/12$. It is at most $a/6$ due to Corollary 10.

Assume now that $\sum_{j=1}^{\infty} \gamma_j < \infty$. Then Corollary 10 yields strong QMC-tractability for average sample points in the space $F_{d,\gamma}$, and hence strong QMC-tractability with the ε-exponent in [1, 2].

If $a = \limsup_d \sum_{j=1}^{d} \gamma_j / \ln d$ is finite then Corollary 10 yields QMC-tractability in the space $F_{d,\gamma}$. The bounds on the ε- and d-exponents are obtained as before.

To prove (iii), note that (19) follows from (38) and (39). Then Lemma 4 yields

$$n_\gamma(\varepsilon,\ d) \geq (1 - \varepsilon^2)\kappa_{d,\gamma}^{-2}.$$

From the definition of c in (iii) we get

$$\frac{1 + y(1 - x^2)/2}{\sqrt{(1 + y(1 - x))(1 + y/3)}} \leq 1 - cy \quad \forall\, x,\ y \in [0,\ 1].$$

Using this in (31) and the fact that $\ln(1 + x) \le x$ for $x > -1$, we obtain

$$\kappa_{d,\gamma} \le \prod_{j=1}^{d} (1 - c\gamma_j) = \exp\left(\sum_{j=1}^{d} \ln(1 - c\gamma_j)\right) \le \exp\left(-c \sum_{j=1}^{d} \gamma_j\right).$$

This yields (20). The value of c can be computed numerically and it is approximately equal to 0.027. This completes the proof. ■

ACKNOWLEDGMENTS

Both authors gratefully acknowledge the support of the Australian Research Council. The second of us (H.W.) was supported in part by the National Science Foundation. This work was done while the second author was visiting the University of New South Wales. We are grateful to E. Novak, J. F. Traub, G. W. Wasilkowski, and A. G. Werschulz for valuable comments on our paper.

REFERENCES

1. Aronszajn, N. (1950), Theory of reproducing kernels, *Trans. Amer. Math. Soc.* **68,** 337–404.
2. Box, G. E. P., and Muller, M. E. (1958), A note on the generation of random normal deviates, *Ann. Math. Statist.* **29,** 610–611.
3. Caflisch, R. E., and Morokoff, W. J. (1996), Quasi-Monte Carlo computation of a finance problem, *in* “Workshop on Quasi-Monte Carlo Methods and Their Applications,” Hong Kong Baptist University, (K. T. Fang and F. J. Hinkernell, Eds.), pp. 15–30.
4. Caflisch, R. E., Morokoff, W. J., and Owen, A. B. (1996), “Valuation of Mortgage Backed Securities Using Brownian Bridges to Reduce Effective Dimension,” UCLA CAM Report 96-16.
5. Drmota, M., and Tichy, R. (1997), Sequences, discrepancies, and applications, Lecture Notes in Math., Vol. 1651, Springer-Verlag, Berlin.
6. Frank, K., and Heinrich, S. (1996), Computing discrepancy of Smolyak quadrature rules, *J. Complexity* **12,** 287–314.
7. Frolov, K. K. (1980), Upper bounds on discrepancy in L_p, $2 \le p \le \infty$, *Dokl. Akad. Nauk SSSR* **252,** 805–807.
8. Heinrich, S. (to appear), Efficient algorithms for computing the L_2 discrepancy, *Math. Comp.*
9. Hickernell, F. J. (to appear), A generalized discrepancy and quadrature error bound.
10. Joe, S. (1997), Formulas for the computation of the weighted L_2 discrepancy, in preparation.
11. Joy, C., Boyle, P. P., and Tan, K. S. (1995), Quasi-Monte Carlo methods in numerical finance, working paper, University of Waterloo.
12. Keister, B. D. (1996), Multidimensional quadrature algorithms, *Comput. in Phys.* **10,** 119–122.
13. Knuth, D. E. (1981), “The Art of Computer Programming, Vol. 2, Semi-numerical Algorithms,” 2nd ed., Addison–Wesley, Reading, MA.
14. Matoušek, J. (1997), On the L_2-discrepancy for anchored boxes, Dept. of Applied Math. Report, Charles University, Praha, Czech Republic.
15. Morokoff, W., and Caflisch, R. E. (1995), Quasi-Monte Carlo integration, *J. Comput. Phys.* **122,** 218–230.

16. Moskowitz, B., and Caflisch, R. E. (1996), Smoothness and dimension reduction in quasi-Monte Carlo methods, *J. Math. Comp. Modeling* **23,** 37–54.

17. Niederreiter, H. (1992), “Random Number Generation and Quasi-Monte Carlo Methods,” CBMS-NSF, SIAM, Philadelphia.

18. Ninomiya, S., and Tezuka, S. (1996), Toward real-time pricing of complex financial derivatives, *Appl. Math. Finance* **3,** 1–20.

19. Novak, E., Sloan, I. H., and Woźniakowski, H. (1997), Tractability of tensor product linear operators, *J. Complexity.*

20. Papageorgiou, A., and Traub, J. F. (1996), Beating Monte Carlo, *Risk* **9**(6), 63–65.

21. Papageorgiou, A., and Traub, J. F. (1997), Faster evaluation of multidimensional integrals, *Comput. in Phys.*

22. Paskov, S. H. (1997), New methodologies for valuing derivatives, *in* “Mathematics of Derivative Securities” (S. Pliska and M. Dempster, Eds.), Isaac Newton Institute, Cambridge University Press, Cambridge.

23. Paskov, S. H., and Traub, J. F. (1995), Faster valuation of financial derivatives, *J. Portfolio Manage.* **22,** 113–120.

24. Roth, K. F. (1954), On irregularities of distribution, *Mathematika* **1,** 73–79.

25. Roth, K. F. (1980), On irregularities of distribution IV, *Acta Arith.* **37,** 67–75.

26. Sloan, I. H., and Joe, S. (1994), “Lattice Methods for Multiple Integration,” Clarendon Press, Oxford.

27. Tezuka, S. (1995), “Uniform Random Numbers: Theory and Practice,” Kluwer Academic, Dordrecht.

28. Wasilkowski, G. W., and Woźniakowski, H. (1995), Explicit cost bounds of algorithms for multivariate tensor product problems, *J. Complexity* **11,** 1–56.

29. Wasilkowski, G. W., and Woźniakowski, H. (1996), On tractability of path integration, *J. Math. Phys.* **37**(4), 2071–2088.

30. Woźniakowski, H. (1994), Tractability and strong tractability of linear multivariate problems, *J. Complexity* **10,** 96–128.

31. Woźniakowski, H. (1993), Tractability and strong tractability of multivariate tensor product problems, *in* “Proceedings of 2nd Gauss Symposium, München, 1993.” [*J. Comput. and Inform.* **4,** 1–19, 1994]

32. Zaremba, S. K. (1968), Some applications of multidimensional integration by parts, *Ann. Polon. Math.* **21,** 85–96.