THE ASYMPTOTIC EFFICIENCY
OF RANDOMIZED NETS FOR QUADRATURE

FRED J. HICKERNELL AND HEE SUN HONG

ABSTRACT. An $\mathcal{L}_2$-type discrepancy arises in the average- and worst-case error analyses for multidimensional quadrature rules. This discrepancy is uniquely defined by $K(x, y)$, which serves as the covariance kernel for the space of random functions in the average-case analysis and a reproducing kernel for the space of functions in the worst-case analysis. This article investigates the asymptotic order of the root mean square discrepancy for randomized $(0, m, s)$-nets in base b. For moderately smooth $K(x, y)$ the discrepancy is $O(N^{-1}[\log(N)]^{(s-1)/2})$, and for $K(x, y)$ with greater smoothness the discrepancy is $O(N^{-3/2}[\log(N)]^{(s-1)/2})$, where $N = b^m$ is the number of points in the net. Numerical experiments indicate that the (t, m, s)-nets of Faure, Niederreiter and Sobol' do not necessarily attain the higher order of decay for sufficiently smooth kernels. However, Niederreiter nets may attain the higher order for kernels corresponding to spaces of periodic functions.

1. INTRODUCTION

Multidimensional integrals over the s-dimensional unit cube $C^s = [0, 1)^s$ may be approximated by the sample mean of the integrand evaluated on a set P with N points:

$$(1.1) \qquad Q(f) \equiv \frac{1}{N} \sum_{z \in P} f(z).$$

Here, P is allowed to have multiple copies of the same point [Nie92, p. 14]. The quadrature error,

$$\text{Err}(f) \equiv \int_{C^s} f(x) \, dx - \frac{1}{N} \sum_{z \in P} f(z),$$

depends on the quality of the quadrature rule Q and the roughness of the integrand f.

Average-case and worst-case quadrature error analyses are equivalent in the following sense (see [Wah90], [Rit95] and below):

$$(1.2) \qquad E_{f \in F}[\text{Err}(f)]^2 = [D(P)]^2 = \sup_{\substack{f \in X \\ V(f)=1}} [\text{Err}(f)]^2.$$

Here E denotes the expectation, and F is a space of random functions with zero mean at every point and with covariance

$$(1.3) \qquad K(x,y) \equiv E_{f \in F}[f(x)f(y)] \quad \forall (x,y) \in C^s \times C^s.$$

The function K also serves as a reproducing kernel for the Hilbert space $(X, \langle, \rangle)$. The variation of the integrand is defined as

$$(1.4) \qquad V(f) \equiv \left[\langle f, f \rangle - \langle 1, f \rangle^2 / \langle 1, 1 \rangle \right]^{1/2},$$

where 1 denotes the unit function. The *discrepancy*, $D(P)$, serves as a measure of uniformity of the set P and is defined as

$$(1.5) \quad D(P) \equiv \left\{ \int_{C^s \times C^s} K(x,y)\, dx\, dy - \frac{2}{N} \sum_{z \in P} \int_{C^s} K(z,y)\, dy \right.$$
$$\left. + \frac{1}{N^2} \sum_{z,z' \in P} K(z,z') \right\}^{1/2}.$$

This discrepancy may be thought of as an $\mathcal{L}_2$-type discrepancy because it is a generalization of the $\mathcal{L}_2$-star discrepancy. It does not include the popular $\mathcal{L}_\infty$-star discrepancy [Nie92, Definition 2.1], however, it does include other $\mathcal{L}_2$-type discrepancies derived in [Hic98].

Since the discrepancy depends only on the quadrature rule and not on the integrand, it is a suitable figure of merit for quadrature rules, or equivalently sets P. There has been much research on generating low discrepancy sets. For moderate N one may compare different sets by actually computing their discrepancies [MC94], [Hic95], [Hic96a], and for large N one may find the asymptotic order of the discrepancies for certain kinds of sets [HW81], [Nie92], [SJ94]. One popular family of low discrepancy sets is the (t, m, s)-nets in base b. The purpose of this article is to present some new results for the discrepancy defined in (1.5) for $(0, m, s)$-nets that have been randomized according to [Owe95], [Owe97a], [Owe97b]. A summary of the main results are as follows:

    i. For reproducing kernels with *moderate* smoothness the root mean square discrepancy is $O(N^{-1}[\log(N)]^{(s-1)/2})$ for large N (Theorem 5.1).

    ii. For reproducing kernels with *greater* smoothness the root mean square discrepancy is $O(N^{-3/2}[\log(N)]^{(s-1)/2})$ for large N (Theorem 5.1).

    iii. Numerical evaluation of the discrepancy for Faure, Niederreiter and Sobol' (t, m, s)-nets show that they do not, in general, have $O(N^{-3/2}[\log(N)]^{(s-1)/2})$ discrepancy for sufficiently smooth reproducing kernels (Figure 2). However, the Niederreiter nets appear to have $O(N^{-3/2}[\log(N)]^{(s-1)/2})$ discrepancy for a smooth reproducing kernel corresponding to a space of *periodic* functions (Figures 3–6).

The next section provides some background on reproducing kernels and the average- and worst-case error analyses mentioned above. Section 3 defines (t, m, s)-nets and describes the randomization proposed by Art Owen. The main results of this article are derived in Sections 4 and 5. Some examples are discussed in the last section.

2. Reproducing Kernels and Error Analysis

For the sake of completeness we outline the derivation of (1.2). For the average-case analysis it is assumed that the functions in F have zero mean and covariance $K(x, y)$ as given in (1.3). The left hand side of (1.2) is derived by interchanging the order of the expectation and the integration or summation:

$$E_{f \in F}[\text{Err}(f)]^2 = E_{f \in F} \left[\int_{C^s} f(x)\, dx - \frac{1}{N} \sum_{z \in P} f(z) \right]^2$$

$$= E_{f \in F} \left[\int_{C^s \times C^s} f(x)f(y)\, dx\, dy \right.$$

$$\left. - \frac{2}{N} \sum_{z \in P} \int_{C^s} f(z)f(y)\, dy + \frac{1}{N^2} \sum_{z, z' \in P} f(z)f(z') \right]$$

$$= \int_{C^s \times C^s} K(x, y)\, dx\, dy - \frac{2}{N} \sum_{z \in P} \int_{C^s} K(z, y)\, dy + \frac{1}{N^2} \sum_{z, z' \in P} K(z, z')$$

$$= [D(P)]^2.$$

For the worst-case error analysis the function $K(x, y)$ is assumed to be a reproducing kernel of a Hilbert space $(X, \langle, \rangle)$, that is,

$$f(y) = \langle K(\bullet, y), f \rangle \quad \forall f \in X.$$

(See [Sai88], [Wah90] for a fuller discussion of reproducing kernels.) This definition implies that the functional that evaluates f at a point must be bounded, something which is not true for $\mathcal{L}_2(C^s)$, but is true for spaces of sufficiently smooth functions. It can be shown that any reproducing kernel K is symmetric in its arguments,

$$K(x, y) = K(y, x),$$

and positive-definite,

$$\sum_{i, k=1}^{N} a^{(i)} a^{(k)} K(x^{(i)}, x^{(k)}) \geq 0 \quad \forall a^{(i)} \in \mathbf{R}, x^{(i)} \in C^s.$$

Furthermore, for any positive-definite function K there corresponds a unique Hilbert space $(X, \langle, \rangle)$ for which it is the reproducing kernel [Wah90, Section 1.1].

The worst-case analysis relies on the Riesz Representation Theorem. Since X has a reproducing kernel, the quadrature error, $\text{Err}(f)$, is a bounded linear functional, and so

$$\text{Err}(f) = \langle \xi, f \rangle \quad \forall f \in X, \quad \text{where } \xi(x) = \text{Err}(K(\bullet, x)).$$

Since all quadrature rules of the form (1.1) are exact for constants, $0 = \text{Err}(f) = \langle \xi, 1 \rangle$. This implies that $\langle \xi, f \rangle = \langle \xi, f_\perp \rangle$, where $f_\perp$ is defined as the part of f which is orthogonal to the function 1:

$$f_\perp \equiv f - \langle 1, f \rangle / \langle 1, 1 \rangle.$$

The Cauchy-Schwarz inequality then implies the following error bound:

(2.1) $$|\text{Err}(f)| = |\langle \xi, f \rangle| = |\langle \xi, f_\perp \rangle| \leq \langle \xi, \xi \rangle^{1/2} \langle f_\perp, f_\perp \rangle^{1/2}.$$

The variation, $V(f)$, may be identified as $\langle f_\perp, f_\perp \rangle^{1/2}$, and

(2.2) $\langle \xi, \xi \rangle = \mathrm{Err}(\xi(x))$

$$= \int_{C^s} \xi(x)\, dx - \frac{1}{N} \sum_{z \in P} \xi(x)$$

$$= \int_{C^s \times C^s} K(x,y)\, dy\, dx - \frac{2}{N} \sum_{z \in P} \int_{C^s} K(z,y)\, dy + \frac{1}{N^2} \sum_{z,z' \in P} K(z,z')$$

$$= [D(P)]^2.$$

Thus, (2.1) is equivalent to the right hand side of (1.2). Note that this error bound is attained when f is a constant multiple of ξ. In other words, ξ is a worst-case integrand.

It is important to note that in general a random function in F will have infinite variation with probability one (see [Wah90, Section 1.1]). This means that F, the space of functions arising in the average-case error analysis, is much larger than X, the space for which the worst-case error bound is derived.

The $\mathcal{L}_2$-star discrepancy is an important example of the discrepancy. Let $S = \{1, \ldots, s\}$ be the set of coordinate indices. For any $u \subseteq S$ let $|u|$ denote the number of points in u. Let $C^u = [0,1)^u$ denote the $|u|$-dimensional unit cube involving the coordinates in u. This notation allows us to distinguish spaces of the same dimension in different coordinate directions. Let P_u denote the projection of P into C^u. The quantity

$$\frac{|P_u \cap [0, x_u)|}{N} - \mathrm{Vol}([0, x_u))$$

is the difference between the empirical distribution associated with the sample P_u and the uniform distribution on the unit cube C^u. The $\mathcal{L}_2$-star discrepancy is based on the $\mathcal{L}_2$-norm, $\| \ \|_2$, of this difference [Hic98]:

(2.3) $\quad [D_2^*(P)]^2 \equiv \sum_{\emptyset \subset u \subseteq S} \left\| \frac{|P_u \cap [0, x_u)|}{N} - \mathrm{Vol}([0, x_u)) \right\|_2^2,$

(2.4) $\qquad = \left(\frac{4}{3} \right)^s - \frac{2}{N} \sum_{z \in P} \prod_{j=1}^s \left(\frac{3 - z_j^2}{2} \right) + \frac{1}{N^2} \sum_{z,z' \in P} \prod_{j=1}^s \left[2 - \max(z_j, z_j') \right].$

The worst-case error bound in (1.2) involving the $\mathcal{L}_2$-star discrepancy was derived by Zaremba [Zar68]. The variation of the function in this case is

$$[V(f)]^2 = \sum_{\emptyset \subset u \subseteq S} \left\| \frac{\partial^{|u|} f}{\partial x_u} \bigg|_{x_{S-u} = (1, \ldots, 1)} \right\|_2^2.$$

The average-case error analysis in (1.2) for the $\mathcal{L}_2$-star discrepancy was obtained by [Woź91], [MC94]. The space of random functions, F, corresponds to Brownian sheets. In the context outlined above, the $\mathcal{L}_2$-star discrepancy is based on the reproducing kernel

(2.5) $\qquad K(x,y) = \prod_{j=1}^s \left[2 - \frac{x_j + y_j + |x_j - y_j|}{2} \right].$

This reproducing kernel is a special case of reproducing kernels that are products of a 1-dimensional reproducing kernel $\tilde{K}$:

$$(2.6) \qquad K(x,y) = \prod_{j=1}^{s} \tilde{K}(x_j, y_j).$$

The discrepancy defined in (1.5) for product reproducing kernels can be written as

$$(2.7) \quad D(P) \equiv \left\{ \left[\int_0^1 \int_0^1 \tilde{K}(x_1, y_1) \ dx_1 \ dy_1 \right]^s - \frac{2}{N} \sum_{z \in P} \prod_{j=1}^{s} \int_0^1 \tilde{K}(z_j, y_j) \ dy_j \right.$$

$$\left. + \frac{1}{N^2} \sum_{z, z' \in P} \prod_{j=1}^{s} \tilde{K}(z_j, z'_j) \right\}^{1/2}.$$

Some specific examples of product reproducing kernels are considered in this article.

Wahba [Wah90, Section 10.2] and others have used the product reproducing kernel with

$$(2.8a) \qquad \tilde{K}(x_1, y_1) = -\frac{(-\beta^2)^\gamma}{(2\gamma)!} B_{2\gamma}(\{x_1 - y_1\}) + \sum_{i=0}^{\gamma} \frac{\beta^{2i}}{(i!)^2} B_i(x_1) B_i(y_1),$$

where β is an arbitrary positive constant, the $B_i(x)$ are Bernoulli polynomials [AS64, Chapter 23], and the notation $\{x\}$ means the fractional part of a scalar or vector x. The positive integer γ indicates the degree of smoothness of the reproducing kernel and the corresponding Hilbert space X. In one dimension the Hilbert space X with this reproducing kernel $\tilde{K}$ consists of functions whose γ order derivative is square integrable over $[0, 1)$. In s dimensions the Hilbert space X with the corresponding product reproducing kernel consists of functions whose mixed partial derivatives of order up to γ in each coordinate are square integrable over C^s. The case $\gamma = 1$ appears in [Hic96b]. The discrepancy corresponding to the product reproducing kernel based on (2.8a) is

$$D(P) \equiv \left\{ -1 + \frac{1}{N^2} \sum_{z, z' \in P} \prod_{j=1}^{s} \left[-\frac{(-\beta^2)^\gamma}{(2\gamma)!} B_{2\gamma}(\{z_j - z'_j\}) \right. \right.$$

$$(2.8b)$$

$$\left. \left. + \sum_{i=0}^{\gamma} \frac{\beta^{2i}}{(i!)^2} B_i(z_j) B_i(z'_j) \right] \right\}^{1/2}.$$

The first author [Hic96b], [Hic98] has considered the reproducing kernel

$$(2.9a) \qquad \tilde{K}(x_1, y_1) = -\frac{(-\beta^2)^\gamma}{(2\gamma)!} B_{2\gamma}(\{x_1 - y_1\}) + 1,$$

where the meanings of β and γ are the same as above. The Hilbert space corresponding to the product reproducing kernel based on (2.9a) consists of *periodic functions* with the same smoothness condition as for (2.8a). The discrepancy corresponding to (2.9a) is

$$(2.9b) \qquad D(P) \equiv \left\{ -1 + \frac{1}{N^2} \sum_{z, z' \in P} \prod_{j=1}^{s} \left[-\frac{(-\beta^2)^\gamma}{(2\gamma)!} B_{2\gamma}(\{z_j - z'_j\}) + 1 \right] \right\}^{1/2}.$$

The discrepancies based on the above reproducing kernels will be considered in Sections 4-6.

3. RANDOMIZED (t, m, s)-NETS

To define a (t, m, s)-net consider the *base b* representation of a point $z = (z_1, \ldots, z_s) \in C^s$:

$$z = (0.z_{11}z_{21}z_{31}\ldots, \ 0.z_{12}z_{22}z_{32}\ldots, \ \ldots\ldots, \ 0.z_{1s}z_{2s}z_{3s}\ldots),$$

where the b-nary digits z_{ij} range from 0 to $b - 1$. Let $\mathbf{Z}_+^s$ denote the space of s-dimensional non-negative integer vectors. For any $k = (k_1, \ldots, k_s) \in \mathbf{Z}_+^s$, let $\sigma(k) = k_1 + \cdots + k_s$. There are $b^{k_1} \cdots b^{k_s} = b^{\sigma(k)}$ different ways to choose the first k b-nary digits of a point z:

(3.1) $z_{11}, \ldots, z_{k_1 1}, z_{12}, \ldots, z_{k_2 2}, \ldots, z_{1s}, \ldots, z_{k_s s}.$

(If $k_j = 0$, then no digit z_{ij} is being specified.) A (t, m, s)-net contains at least one point with every possible choice of the first k digits, provided that $\sigma(k)$ is small enough.

Definition 3.1. Let t and m be non-negative integers with $0 \leq t \leq m$, let s be a positive integer, and let b be a positive integer greater than one. A (t, m, s)-net in base b is a set P containing $N = b^m$ points in C^s. For any possible choice of the first k b-nary digits (3.1) there exist $b^{m-\sigma(k)}$ points in P with these digits, provided that $\sigma(k) \leq m - t$.

Owen [Owe95] proposed randomizing the digits of the points in a given (t, m, s)-net P_0 to obtain a new (t, m, s)-net P. For every digit index $i = 1, 2, \ldots$, every coordinate index $j = 1, \ldots, s$ and every $\zeta \in P_0$ one obtains a random digit z_{ij} that contributes to a random point $z \in P$. The randomized net P satisfies the following two assumptions:

Assumption 3.2.   a. For any $z \in P$ each digit z_{ij} is uniformly distributed on the set $\{0, \ldots, b - 1\}$.
  b. For any two points $z, z' \in P$ the random vectors $(z_1, z_1'), \ldots, (z_s, z_s')$ are *mutually independent*.

Assumption 3.3. For any two points $\zeta, \zeta' \in P_0$ let $z, z' \in P$ be the corresponding points in the randomized net. Suppose that ζ_j and ζ_j' share the same first k_j digits, but that their $k_j + 1^{\text{st}}$ digits are different. Then
  a. $z_{ij} = z_{ij}'$ for $i = 1, \ldots, k_j$,
  b. the random vector $(z_{k_j+1,j}, z_{k_j+1,j}')$ is uniformly distributed on the set $\{(n, n') : n \neq n'; \ n, n' = 0, \ldots, b - 1\}$, and
  c. $z_{k_j+2,j}, z_{k_j+3,j}, \ldots, z_{k_j+2,j}', z_{k_j+3,j}', \ldots$ are *mutually independent*.

Assumption 3.2 implies that the marginal probability distribution of any $z = (z_1, \ldots, z_s) \in P$ is uniform on C^s and that the $z_1, \ldots, z_s$ are mutually independent. Assumption 3.3 maintains the correlation between different points in P that is necessary for retaining its net properties and thus a low discrepancy. For a simple random sample Assumption 3.2 is also satisfied, but instead of Assumption 3.3 one has $z_{1j}, z_{2j}, \ldots, z_{1j}', z_{2j}', \ldots$ mutually independent for $z \neq z'$.

The aim of this article is to analyze the mean square discrepancy for randomized nets, i.e. $E\{[D(P)]^2\}$, where E is understood here to be the expectation over random sets P. This section concludes with a lemma that simplifies formula (1.5) for

any P satisfying Assumption 3.2 and a formula for the mean square discrepancy of a simple random sample.

Lemma 3.4. *If P is a random set satisfying Assumption 3.2, then*

$$(3.2) \qquad E\{[D(P)]^2\} = \frac{1}{N^2} \sum_{z,z' \in P} E[K(z,z')] - \int_{C^s \times C^s} K(x,y) \, dy \, dx.$$

Proof. By Assumption 3.2 each $z \in P$ is uniformly distributed on C^s, so

$$E\left[\frac{2}{N} \sum_{z \in P} \int_{C^s} K(z,y) \, dy\right] = 2 \int_{C^s \times C^s} K(x,y) \, dy \, dx.$$

Substituting this into (1.5) gives (3.2). $\qquad \square$

For a simple random sample

$$E[K(z,z')] = \begin{cases} \int_{C^s} K(x,x) \, dx, & z = z', \\ \int_{C^s \times C^s} K(x,y) \, dy \, dx, & z \neq z', \end{cases}$$

which implies the following theorem:

Theorem 3.5. *If P is a simple random sample, then*

$$(3.3) \qquad E\{[D(P)]^2\} = \frac{1}{N} \left\{ \int_{C^s} K(x,x) \, dx - \int_{C^s \times C^s} K(x,y) \, dy \, dx \right\}.$$

This formula serves as a benchmark for other (presumably superior) quasi-random sets. Since the mean square discrepancy is $O(N^{-1})$, the discrepancy itself is typically $O(N^{-1/2})$ for a simple random sample. The variance of a function f may be defined as $\mathrm{Var}(f) \equiv \int_{C^s} f^2 \, dx - \left(\int_{C^s} f \, dx\right)^2$. The term in braces in (3.3) is $E_{f \in F}[\mathrm{Var}(f)]$.

4. THE DISCREPANCY FOR GENERAL REPRODUCING KERNELS

The calculation of the mean square discrepancy of a randomized net is aided by decomposing the reproducing kernel into a sum of components $K_{u,v}$ for all $(u,v) \subset S \times S$. For this purpose the sets (u,v) are partially ordered so that $(u',v') < (u,v)$ means that $u' \subseteq u$, $v' \subseteq v$ and $(u',v') \neq (u,v)$. The components $K_{u,v}$ are defined in Definition 4.1, and some of their properties are given in Lemma 4.2. The two terms in (3.2) are shown to depend only on the components $K_{u,u}$ in Lemma 4.3. An expression for the mean square discrepancy as a double sum is given in Theorem 4.4.

Definition 4.1. For any $u, v \subset S$ the u, v component of the reproducing kernel K, denoted $K_{u,v}$, is a function on $C^u \times C^v$ recursively defined as follows:

$$2^{|(u \cup v) - (u \cap v)|} K_{u,v}(x_u, y_v) = \int_{C^{S - (u \cup v)}} K \Big|_{\substack{y_{S-v} = x_{S-v} \\ x_{v-u} = y_{v-u}}} dx_{S-(u \cup v)}$$
$$- \sum_{(u',v') < (u,v)} 2^{|(u' \cup v') - (u \cap v)|} K_{u',v'}(x_{u'}, y_{v'}).$$

The decomposition of product reproducing kernels (2.6) is relatively simple. Define

$$(4.1a) \qquad M \equiv \int_0^1 \tilde{K}(x_1, x_1) \, dx_1, \qquad \mu(x_1) \equiv (\tilde{K}(x_1, x_1) - M)/2,$$

$$(4.1b) \qquad \nu(x_1, y_1) \equiv \tilde{K}(x_1, y_1) - [\tilde{K}(x_1, x_1) + \tilde{K}(y_1, y_1)]/2.$$

The one-dimensional reproducing kernel can be written as

$$\tilde{K}(x_1, y_1) = M + \mu(x_1) + \mu(y_1) + \nu(x_1, y_1),$$

and the components of the s-dimensional product reproducing kernel are

$$(4.2) \qquad K_{u,v} = M^{s-|u \cup v|} \prod_{j \in u-v} \mu(x_j) \prod_{j' \in v-u} \mu(y_{j'}) \prod_{j'' \in u \cap v} \nu(x_{j''}, y_{j''}).$$

It follows from the definitions of μ and ν that $\int_0^1 \mu(x_1) \, dx_1 = \nu(x_1, x_1) = 0$. Therefore, for product kernels

$$K_{u,v}|_{y_j = x_j} = 0 \quad \forall j \in u \cap v,$$

$$\int_0^1 K_{u,v} \, dx_j = 0 \quad \forall j \in (u \cup v) - (u \cap v).$$

In fact, these properties hold not only for product kernels, but in general.

Lemma 4.2. *The components of the reproducing kernel have the following properties:*

$$(4.3a) \qquad K = \sum_{u,v \subseteq S} K_{u,v},$$

$$(4.3b) \qquad K_{u,v}|_{y_j = x_j} = 0 \quad \forall j \in u \cap v,$$

$$(4.3c) \qquad \int_0^1 K_{u,v} \, dx_j = 0 \quad \forall j \in (u \cup v) - (u \cap v).$$

Proof. The proof of (4.3a) follows from the definition by setting $u = v = S$. The proofs of (4.3b) and (4.3c) proceed by induction. Both properties hold vacuously for $u = v = \emptyset$. By assuming that they hold for all $(u'', v'') < (u, v)$ it is possible to show that they hold for (u, v) itself.

Let j be some coordinate index in $u \cap v$. The definition of a component implies that

$$2^{|(u \cup v) - (u \cap v)|} K_{u,v}\Big|_{y_j = x_j} = \int_{C^{S-(u \cup v)}} K\Big|_{\substack{y_{S-v} = x_{S-v} \\ x_{v-u} = y_{v-u} \\ y_j = x_j}} dx_{S-(u \cup v)}$$

$$- \sum_{(u', v') < (u, v)} 2^{|(u' \cup v') - (u \cap v)|} K_{u', v'}\Big|_{y_j = x_j}.$$

In the sum over (u', v') those terms with $j \in u' \cap v'$ vanish by the induction assumption, leaving only those terms for which j is in either u' only or in v' only or in neither. Let $\hat{u} = u - j$ and $\hat{v} = v - j$. Then

$$|(u' \cup v') - (u \cap v)|$$
$$= \begin{cases} |(u' \cup v') - (u \cap \hat{v})| - 1 = |(u' \cup v') - (\hat{u} \cap v)| - 1 & \text{if } j \in u' \cup v', \\ |(u' \cup v') - (u \cap \hat{v})| = |(u' \cup v') - (\hat{u} \cap v)| & \text{if } j \notin u' \cup v'. \end{cases}$$

This allows us to rewrite the sum over (u', v') in the right hand side above:

$$
\begin{aligned}
2^{|(u \cup v) - (u \cap v)|} K_{u,v}\Big|_{y_j = x_j} &= \int_{C^{S-(u \cup v)}} K\Big|_{\substack{y_{S-\hat{v}} = x_{S-\hat{v}} \\ x_{v-u} = y_{v-u}}} dx_{S-(u \cup v)} \\
&\quad - \frac{1}{2} \sum_{(u',v') \leq (\hat{u},v)} 2^{|(u' \cup v') - (\hat{u} \cap v)|} K_{u',\hat{v}'}\Big|_{y_j = x_j} \\
&\quad - \frac{1}{2} \sum_{(u',v') \leq (u,\hat{v})} 2^{|(u' \cup v') - (u \cap \hat{v})|} K_{u',v'}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{2}\Bigg\{ -2^{|(\hat{u} \cup v) - (\hat{u} \cap v)|} K_{\hat{u},v} + \int_{C^{S-(\hat{u} \cup v)}} K\Big|_{\substack{y_{S-v} = x_{S-v} \\ x_{v-u} = y_{v-u}}} dx_{S-(\hat{u} \cup v)} \\
&\qquad\qquad\qquad\qquad\qquad - \sum_{(u',v') < (\hat{u},v)} 2^{|(u' \cup v') - (\hat{u} \cap v)|} K_{u',v'} \Bigg\}\Bigg|_{y_j = x_j} \\
&\quad + \frac{1}{2}\Bigg\{ -2^{|(u \cup \hat{v}) - (u \cap \hat{v})|} K_{u,\hat{v}} + \int_{C^{S-(u \cup \hat{v})}} K\Big|_{\substack{y_{S-\hat{v}} = x_{S-\hat{v}} \\ x_{\hat{v}-u} = y_{\hat{v}-u}}} dx_{S-(u \cup \hat{v})} \\
&\qquad\qquad\qquad\qquad\qquad - \sum_{(u',v') < (u,\hat{v})} 2^{|(u' \cup v') - (u \cap \hat{v})|} K_{u',v'} \Bigg\}.
\end{aligned}
$$

Each set of braces is zero by Definition 4.1, thus completing the proof of (4.3b).

To prove (4.3c) one may assume without loss of generality that $j \in u - v$. Again let $\hat{u} = u - j$ and note that $v - u = v - \hat{u}$. The integral over $K_{u,v}$ with respect to the jth coordinate is

$$
\begin{aligned}
2^{|(u \cup v) - (u \cap v)|} \int_0^1 K_{u,v}\, dx_j &= \int_{C^{S-(\hat{u} \cup v)}} K\Big|_{\substack{y_{S-v} = x_{S-v} \\ x_{v-\hat{u}} = y_{v-\hat{u}}}} dx_{S-(\hat{u} \cup v)} \\
&\quad - \sum_{(u',v') < (u,v)} 2^{|(u' \cup v') - (u \cap v)|} \int_0^1 K_{u',v'}\, dx_j.
\end{aligned}
$$

The integrals of $K_{u',v'}$ on the right hand side vanish if $j \in u'$ by the induction assumption. If $j \notin u'$, that is, $u' \subseteq \hat{u}$, then the integral is the same as the function itself. Therefore,

$$
\begin{aligned}
2^{|(u \cup v) - (u \cap v)|} \int_0^1 K_{u,v}\, dx_j &= \int_{C^{S-(\hat{u} \cup v)}} K\Big|_{\substack{y_{S-v} = x_{S-v} \\ x_{v-\hat{u}} = y_{v-\hat{u}}}} dx_{S-(\hat{u} \cup v)} \\
&\quad - \sum_{(u',v') \leq (\hat{u},v)} 2^{|(u' \cup v') - (\hat{u} \cap v)|} K_{u',v'}.
\end{aligned}
$$

The right-hand side vanishes by the definition of $K_{\hat{u},v}$. $\qquad\square$

Substituting the decomposition of the reproducing kernel into the two terms in formula (3.2) for mean square discrepancy of a randomized set under Assumption

3.2 yields

$$\frac{1}{N^2} \sum_{z,z' \in P} E[K(z, z')] = \sum_{u,v \subseteq S} \frac{1}{N^2} \sum_{z,z' \in P} E[K_{u,v}(z_u, z'_v)],$$

$$\int_{C^s \times C^s} K(x, y) \, dy \, dx = \sum_{u,v \subseteq S} \int_{C^u \times C^v} K_{u,v}(x_u, y_v) \, dx_u \, dy_v.$$

Using Lemma 4.2, one sees that those terms involving $K_{u,v}$ for $v \neq u$ vanish. One may suppose without loss of generality that there is some coordinate j in u but not in v. The expectation $E[K_{u,v}(z_u, z'_v)]$ depends on z_j, which is independent of z_{u-j} and z'_v. Since z_j is uniformly distributed on $[0, 1)$, $E[K_{u,v}(z_u, z'_v)]$ vanishes by (4.3c). The integral of $K_{u,v}$ over x_j also vanishes by (4.3c). This gives the following lemma:

Lemma 4.3. *If P is a random set satisfying Assumptions 3.2, then*

(4.4a)
$$\frac{1}{N^2} \sum_{z,z' \in P} E[K(z, z')] = \sum_{u \subseteq S} \frac{1}{N^2} \sum_{z,z' \in P} E[K_{u,u}(z_u, z'_u)],$$

(4.4b)
$$\int_{C^s \times C^s} K(x, y) \, dy \, dx = \sum_{u \subseteq S} \int_{C^u \times C^u} K_{u,u}(x_u, y_u) \, dx_u \, dy_u.$$

To further simplify (4.4) one must analyze in what subset of $C^u \times C^u$ a pair of points $(z_u, z'_u) \in P_u \times P_u$ may fall. Suppose the coordinates z_j and z'_j share exactly the first k_j base b digits but no more. If this is true for all j in some $u \subseteq S$, one may say that the points z_u and z'_u share exactly the same k_u coordinates, where k_u is in $\mathbf{Z}_+^u$, the set of $|u|$-dimensional non-negative integer vectors. Define the following conditional expectations:

(4.5a) $H(k_u) \equiv E[K_{u,u}(z_u, z'_u) | z_u \text{ and } z'_u \text{ share exactly } k_u \text{ digits}],$

(4.5b) $H_{l,\tau} \equiv E[K_{u,u}(z_u, z'_u) | z_u \text{ and } z'_u \text{ share exactly } k_u \text{ digits}$
                    for some k_u with $\sigma(k_u) = \tau$ and $|u| = l$].

If z_u and z'_u share exactly k_u digits, then (z_u, z'_u) is uniformly distributed on $\Omega_{k_u} \subset C^u \times C^u$, the Cartesian product of certain $\Omega_{k_j} \subset C^{\{j\}} \times C^{\{j\}}$, that is,

(4.6a)
$$\Omega_{k_u} \equiv \bigotimes_{j \in u} \Omega_{k_j},$$

(4.6b)
$$\Omega_{k_j} \equiv \bigcup_{i=0}^{b^{k_j}} [ib^{-k_j}, (i+1)b^{-k_j})^2 - \bigcup_{i=0}^{b^{k_j+1}} [ib^{-k_j-1}, (i+1)b^{-k_j-1})^2.$$

The volume of Ω_{k_j} is $b^{-k_j}(1 - b^{-1})$, so the volume of Ω_{k_u} is the product of the volumes of the corresponding Ω_{k_j}:

(4.7) $\text{Vol}(\Omega_{k_u}) = \text{Vol}_{|u|, \sigma(k_u)}, \text{ where } \text{Vol}_{l,\tau} \equiv b^{-\tau}(1 - b^{-1})^l.$

This notation now allows one to compute the conditional expectation defined in (4.5a):

(4.8a)
$$H(k_u) = \frac{1}{\text{Vol}_{|u|, \sigma(k_u)}} \int_{\Omega_{k_u}} K_{u,u}(x_u, y_u) \, dx_u \, dy_u.$$

The conditional expectation $H_{l,\tau}$ is the average of the $H(k_u)$ with $|u| = l$ and $\sigma(k_u) = \tau$. There are $\binom{s}{l}$ subsets u of S with cardinality l, and for any u there are $\binom{l+\tau-1}{l-1}$ different k_u with $\sigma(k_u) = \tau$, so

$$(4.8b) \qquad H_{l,\tau} = \frac{1}{\binom{s}{l}\binom{l+\tau-1}{l-1}} \sum_{|u|=l} \sum_{\sigma(k_u)=\tau} H(k_u).$$

It is known [Hic96a] that for a randomized $(0, m, s)$-net there are $\lambda_{|u|,\sigma(k_u)}$ points z' that share the same k_u digits as z, where

$$(4.9) \qquad \frac{\lambda_{l,\tau}}{N} \equiv b^{-\tau}(1 - b^{-1})^l R(l, m - \tau) \quad (= 0 \text{ for } m \le \tau),$$

and $R(l, \tau)$ is defined as the partial binomial sum:

$$(4.10) \qquad R(l, \tau) \equiv (1 - b^{-1})^{1-l} \sum_{r=0}^{\tau-1} \binom{l-1}{r}(-b)^{-r} = \begin{cases} 0, & \tau \le 0, \\ 1, & \tau \ge l. \end{cases}$$

Furthermore, when $z = z'$, then $K_{u,u} = 0$ for $u \neq \emptyset$ by (4.3b). Therefore, (4.4a) may be written as

$$(4.11) \qquad \sum_{u \subseteq S} \frac{1}{N^2} \sum_{z,z' \in P} E[K_{u,u}(z_u, z'_u)]$$

$$= K_{\emptyset,\emptyset} + \sum_{\emptyset \subset u \subseteq S} \sum_{k_u} [\text{Proportion of } (z, z') \in P \times P$$

$$\text{such that } z_u \text{ and } z'_u \text{ share exactly } k_u \text{ digits}]$$

$$\times E[K_{u,u}(z_u, z'_u)|z_u \text{ and } z'_u \text{ share exactly } k_u \text{ digits}]$$

$$= K_{\emptyset,\emptyset} + \sum_{l=1}^{s} \sum_{\tau=0}^{+\infty} \sum_{|u|=l} \sum_{\sigma(k_u)=\tau} \frac{\lambda_{|u|,|k_u|}}{N} H(k_u)$$

$$= K_{\emptyset,\emptyset} + \sum_{l=1}^{s} \sum_{\tau=0}^{+\infty} \binom{s}{l}\binom{l+\tau-1}{l-1}(1 - b^{-1})^l b^{-\tau} R(l, m - \tau) H_{l,\tau}.$$

The term (4.4b) is even simpler to express. Note that $C^u \times C^u$ is the union of Ω_{k_u} over all $k_u \in \mathbf{Z}_+^u$ plus a set of measure zero (the diagonal $\{(x_u, x_u) : x_u \in C^u\}$). Therefore, by (4.7) and (4.8a),

$$(4.12) \qquad \sum_{u \subseteq S} \int_{C^u \times C^u} K_{u,u}(x_u, y_u) \, dx_u \, dy_u$$

$$= K_{\emptyset,\emptyset} + \sum_{\emptyset \subset u \subseteq S} \sum_{k_u} \text{Vol}(\Omega_{k_u}) H(k_u)$$

$$= K_{\emptyset,\emptyset} + \sum_{l=1}^{s} \sum_{\tau=0}^{+\infty} \binom{s}{l}\binom{l+\tau-1}{l-1}(1 - b^{-1})^l b^{-\tau} H_{l,\tau}.$$

Combining (4.11) and (4.12) with Lemmas 3.4 and 4.3 yields the following formula for the mean square discrepancy of a randomized net:

$$(4.13) \quad E\{[D(P)]^2\} = \sum_{l=1}^{s} \sum_{\tau=0}^{+\infty} \binom{s}{l}\binom{l+\tau-1}{l-1}(1 - b^{-1})^l b^{-\tau} [R(l, m - \tau) - 1] H_{l,\tau}$$

Not only is $R(l, \tau)$ is a partial binomial sum, but $R(l, \tau) - 1$ is also a partial binomial sum. Let

$$(4.14) \qquad \tilde{R}(l, \tau) \equiv (1 - b)^{1-l} \sum_{r=0}^{\tau-1} \binom{l-1}{r} (-b)^r = \begin{cases} 0, & \tau \leq 0, \\ 1, & \tau \geq l. \end{cases}$$

By rearranging the order of summation we get

$$R(l, \tau) - 1 = (1 - b^{-1})^{1-l} \left[\sum_{r=l-\tau}^{l-1} \binom{l-1}{r} (-b)^{r-l+1} - \sum_{r=0}^{l-1} \binom{l-1}{r} (-b)^{r-l+1} \right]$$

$$= -(1 - b)^{1-l} \sum_{r=0}^{l-\tau-1} \binom{l-1}{r} (-b)^r$$

$$= -\tilde{R}(l, l - \tau).$$

Replacing $R(l, m - \tau)$ by $1 - \tilde{R}(l, \tau + l - m)$ in (4.13) gives

$$E\{[D(P)]^2\} = -\sum_{l=1}^{s} \sum_{\tau=0}^{+\infty} \binom{s}{l} \binom{l+\tau-1}{l-1} (1 - b^{-1})^l b^{-\tau} \tilde{R}(l, \tau + l - m) H_{l,\tau}.$$

Replacing τ by $\tau + m - l$ and noticing that $\tilde{R}(l, \tau)$ is only nonzero for $\tau \geq 1$ yields the following theorem:

Theorem 4.4. *Let $H_{l,\tau}$ be the conditional expectation given by (4.8) and $\tilde{R}(l, \tau)$ be the partial binomial sum defined in (4.14). The mean square discrepancy of a randomized $(0, m, s)$-net P satisfying Assumptions 3.2 and 3.3 is*

$$(4.15) \quad E\{[D(P)]^2\}$$

$$= -\sum_{l=1}^{s} \sum_{\tau=\max(1,l-m)}^{+\infty} \binom{s}{l} \binom{m+\tau-1}{l-1} (b-1)^l b^{-m-\tau} \tilde{R}(l, \tau) H_{l,\tau+m-l}.$$

The advantage of this formula is that the large N (or m) asymptotic behavior of the mean square discrepancy is seen to depend on how fast the $H_{l,\tau}$ tend to zero as τ tends to infinity. In fact, the values of $H_{l,\tau}$ for $\tau \leq m - s$ do not even enter into the formula for the mean square discrepancy. For large τ, the $H_{l,\tau}$ depend only on the values of the reproducing kernel $K(x, y)$ close to the planes $x_u = y_u$.

To apply Theorem 4.4 one must know more about the $H_{l,\tau}$ as given by (4.8). For product reproducing kernels the components $K_{u,u}$, as given by (4.2), are:

$$(4.16) \qquad K_{u,u} = M^{s-|u|} \prod_{j \in u} \nu(x_j, y_j).$$

Let h_{k_1} be the average value of $\nu(x_1, y_1)$ over Ω_{k_1} for any non-negative integer k_1:

$$h_{k_1} \equiv \frac{1}{\mathrm{Vol}_{1,k_1}} \int_{\Omega_{k_1}} \nu(x_1, y_1) \, dx_1 \, dy_1.$$

Then by (4.8)

$$H_{l,\tau} = \frac{M^{s-l}}{\binom{l+\tau-1}{l-1}} \sum_{k_1 + \cdots + k_l = \tau} h_{k_1} \cdots h_{k_l}.$$

For some cases this can be further simplified.

For example, consider the product kernel defined by (2.8a) with $\gamma = 1$. By (4.1) the M and ν are

$$M = 1 + \frac{\beta^2}{6}, \qquad \nu(x_1, y_1) = -\frac{\beta^2}{2}|x_1 - y_1|.$$

Some straightforward algebra (see Lemma 5.2 and the proof of Theorem 5.1 below) gives

$$h_{k_1} = -\left[\frac{\beta^2(1 + b^{-1})}{6}\right]b^{-k_1}, \qquad H_{l,\tau} = \left(\frac{6 + \beta^2}{6}\right)^s \left[\frac{-\beta^2(1 + b^{-1})}{(6 + \beta^2)}\right]^l b^{-\tau}.$$

Therefore, (4.15) becomes

$$(4.17) \quad E\{[D(P)]^2\} = -\left(\frac{6 + \beta^2}{6}\right)^s \sum_{l=1}^{s} \binom{s}{l} \left[\frac{-\beta^2(b^2 - 1)}{(6 + \beta^2)}\right]^l$$

$$\times \sum_{\tau = \max(1, l-m)}^{+\infty} \binom{m + \tau - 1}{l - 1} \tilde{R}(l, \tau) b^{-2m - 2\tau}.$$

For large numbers of points, i.e. large m, the binomial coefficient $\binom{m+\tau-1}{l-1}$ is asymptotic to $\frac{m^{l-1}}{(l-1)!}$. Making this substitution the inner sum in the equation above can be simplified by reversing the order of summation:

$$\sum_{\tau = \max(1, l-m)}^{+\infty} \binom{m + \tau - 1}{l - 1} \tilde{R}(l, \tau) b^{-2m - 2\tau}$$

$$\sim b^{-2m} \frac{m^{l-1}}{(l-1)!}(1 - b)^{1-l} \sum_{\tau=1}^{+\infty} \sum_{r=0}^{\tau-1} \binom{l-1}{r}(-b)^r b^{-2\tau}$$

$$= b^{-2m} \frac{m^{l-1}}{(l-1)!}(1 - b)^{1-l} \sum_{r=0}^{+\infty} \binom{l-1}{r}(-b)^r \sum_{\tau=r+1}^{+\infty} b^{-2\tau}$$

$$= b^{-2m} \frac{m^{l-1}}{(l-1)!}(1 - b)^{1-l} \sum_{r=0}^{+\infty} \binom{l-1}{r}(-b)^r \frac{b^{-2r}}{b^2 - 1}$$

$$= b^{-2m} \frac{m^{l-1}}{(l-1)!} \frac{(-b)^{-l+1}}{b^2 - 1}.$$

For large m the most significant term in the outer sum of (4.17) is $l = s$, because of the factor m^{l-1} above. Therefore, the asymptotic form of the mean square discrepancy is

$$E\{[D(P)]^2\} \sim \frac{\beta^2(b - b^{-1})^{s-1}}{6^s(s - 1)!}b^{-2m}m^{s-1} = \frac{\beta^2(b - b^{-1})^{s-1}}{6^s(s - 1)!}N^{-2}[\log_b(N)]^{s-1}.$$

This formula was obtained previously by [Hic96a].

5. THE DISCREPANCY FOR SMOOTH REPRODUCING KERNELS

For general choices of the reproducing kernel it is not possible to obtain a precise asymptotic form of the mean square discrepancy as in the example above. However, by assuming sufficient smoothness of the reproducing kernel it is possible to use Theorem 4.4 to compute the asymptotic order for general reproducing kernels.

Such a result is given in the the following theorem, which is the main result of this article.

Theorem 5.1. *Let* $\tilde{R}_{\max}(l)$ *be an upper bound on* $|\tilde{R}(l,\tau)|$ *for all* τ. *Suppose that* P *is a randomized* $(0,m,s)$*-net satisfying Assumptions 3.2 and 3.3. If the reproducing kernel satisfies a moderate smoothness condition:*

$$(5.1a) \qquad \left\| \frac{\partial^{|u|} K_{u,u}}{\partial x_{u-v} \partial y_v} \right\|_\infty \leq A_1(|u|) \quad \forall v \subseteq u \subseteq S,$$

for some finite $A_1(l)$, $l = 1, \ldots, s$, *then the mean square discrepancy for large* N *is asymptotically*

$$(5.1b) \qquad E\{[D(P)]^2\} \lesssim A_1(s)\tilde{R}_{\max}(s) \frac{(b^2-1)^{s-1}}{3^s(s-1)!} N^{-2}[\log_b(N)]^{s-1}.$$

If the reproducing kernel has even greater smoothness:

$$(5.2a) \qquad \left\| \frac{\partial^{|2u|} K_{u,u}}{\partial x_{u-v} \partial x_{u-w} \partial y_v \partial y_w} \right\|_\infty \leq A_2(|u|) \quad \forall v, w \subseteq u \subseteq S,$$

for some finite $A_2(l)$, $l = 1, \ldots, s$, *then the mean square discrepancy for large* N *is asymptotically*

$$(5.2b) \qquad E\{[D(P)]^2\} \lesssim A_2(s)\tilde{R}_{\max}(s) \frac{(b^3-1)^{s-1}}{6^s(s-1)!} N^{-3}[\log_b(N)]^{s-1}$$

The proof of this theorem is based on two lemmas. The first is a technical one whose proof involves straightforward integration. The second involves Taylor expansions of the the components of the reproducing kernel $K_{u,u}$ about the diagonal $x_u = y_u$.

Lemma 5.2. *For any constant* $p > -1$

$$\frac{1}{\mathrm{Vol}(\Omega_{k_u})} \int_{\Omega_{k_u}} \prod_{j \in u} |x_j - y_j|^p \, dx_u \, dy_u \leq \left[\frac{2(1-b^{-p-1})}{(p+1)(p+2)(1-b^{-1})} \right]^{|u|} b^{-p\sigma(k_u)}.$$

Proof. Because the domain of integration as defined in (4.6) is a Cartesian product of one-dimensional domains, the integral may be simplified as follows:

$$\int_{\Omega_{k_u}} \prod_{j \in u} |x_j - y_j|^p \, dx_u \, dy_u = \prod_{j \in u} \int_{\Omega_{k_j}} |x_j - y_j|^p \, dx_j \, dy_j$$

$$= \prod_{j \in u} \left\{ b^{k_j} \left[\int_0^{b^{-k_j}} \int_0^{b^{-k_j}} |x_j - y_j|^p \, dx_j \, dy_j \right. \right.$$

$$\left. \left. - b \int_0^{b^{-k_j-1}} \int_0^{b^{-k_j-1}} |x_j - y_j|^p \, dx_j \, dy_j \right] \right\}$$

$$= \prod_{j \in u} \left\{ b^{k_j} \left[\frac{2b^{-k_j(p+2)}}{(p+1)(p+2)} - b\frac{2b^{-(k_j+1)(p+2)}}{(p+1)(p+2)} \right] \right\}$$

$$= \prod_{j \in u} \left\{ \frac{2(1-b^{-p-1})}{(p+1)(p+2)} b^{-k_j(p+1)} \right\}$$

$$= \left[\frac{2(1-b^{-p-1})}{(p+1)(p+2)} \right]^{|u|} b^{-(p+1)\sigma(k_u)}$$

Dividing the term on the right hand side by the volume of Ω_{k_u} completes the proof. $\square$

Before stating and proving the next lemma it is convenient to introduce new coordinates $(\hat{x}_j, \hat{y}_j)$ defined as follows:

$$\hat{x}_j = x_j + y_j, \quad \hat{y}_j = x_j - y_j$$
$$x_j = (\hat{x}_j + \hat{y}_j)/2, \quad y_j = (\hat{x}_j - \hat{y}_j)/2.$$

The plane $x_j = y_j$ corresponds to $\hat{y}_j = 0$. Furthermore, let

$$\hat{K}_{u,u}(\hat{x}_u, \hat{y}_u) = K_{u,u}((\hat{x}_j + \hat{y}_j)/2, (\hat{x}_j - \hat{y}_j)/2).$$

Below we consider derivatives of $\hat{K}_{u,u}$ with respect to its second variable, $\hat{y}_u$. According to the above definitions

$$\frac{\partial}{\partial \hat{y}_j} = \frac{1}{2}\left(\frac{\partial}{\partial x_j} - \frac{\partial}{\partial y_j}\right).$$

Therefore, the mixed partial derivatives of $\hat{K}_{u,u}$ with respect to its second variable may be written as follows:

$$(5.3a) \qquad \frac{\partial^{|u|}\hat{K}_{u,u}}{\partial \hat{y}_u} = \frac{1}{2^{|u|}}\sum_{\emptyset \subseteq v \subseteq u}(-1)^{|v|}\frac{\partial^{|u|}K_{u,u}}{\partial x_{u-v}\partial y_v},$$

$$(5.3b) \qquad \frac{\partial^{|2u|}\hat{K}_{u,u}}{\partial \hat{y}_u^2} = \frac{1}{2^{2|u|}}\sum_{\emptyset \subseteq v \subseteq u}\sum_{\emptyset \subseteq w \subseteq u}(-1)^{|v|+|w|}\frac{\partial^{2|u|}K_{u,u}}{\partial x_{u-v}\partial x_{u-w}\partial y_v \partial y_w},$$

Conditions (5.1a) and (5.2a) guarantee the boundedness of these partial derivatives of $\hat{K}_{u,u}$ with respect to $\hat{y}_u$.

Lemma 5.3. *If the reproducing kernel K satisfies smoothness condition* (5.1a), *then*

$$(5.4) \qquad \hat{K}_{u,u}(\hat{x}_u, \hat{y}_u) = \int_0^{\hat{y}_u}\frac{\partial^{|u|}\hat{K}_{u,u}(\hat{x}_u, \hat{\eta}_u)}{\partial \hat{\eta}_u}\, d\hat{\eta}_u.$$

If the reproducing kernel K satisfies smoothness condition (5.2a), *then*

$$(5.5) \quad \hat{K}_{u,u}(\hat{x}_u, \hat{y}_u) = \sum_{\emptyset \subset v \subseteq u}(-1)^{|v|+1}\frac{\partial^{|v|}\hat{K}_{u,u}(\hat{x}_u, \hat{\eta}_u)}{\partial \hat{\eta}_v}\Bigg|_{\substack{\hat{\eta}_v=0 \\ \hat{\eta}_{u-v}=\hat{y}_{u-v}}}\hat{y}_v$$
$$+ \int_0^{\hat{y}_u}\int_0^{\hat{\eta}_u}\frac{\partial^{2|u|}\hat{K}_{u,u}(\hat{x}_u, \hat{\zeta}_u)}{\partial \hat{\zeta}_u^2}\, d\hat{\zeta}_u d\hat{\eta}_u.$$

Proof. Assuming sufficient smoothness of the reproducing kernel, one may expand $\hat{K}_{u,u}(\hat{x}_u, \hat{y}_u)$ in a Taylor approximation about $\hat{y}_u = 0$. This is derived by integrating the partial derivative of $\hat{K}_{u,u}$ with respect to its second variable over the rectangular solid from 0 to y_u:

$$\int_0^{\hat{y}_u}\frac{\partial^{|u|}\hat{K}_{u,u}(\hat{x}_u, \hat{\eta}_u)}{\partial \hat{\eta}_u}\, d\hat{\eta}_u = \sum_{\emptyset \subseteq v \subseteq u}(-1)^{|v|}\hat{K}_{u,u}(\hat{x}_u, \hat{y}_u)\Big|_{\hat{y}_v=0}.$$

By (4.3b) the terms on the right vanish for all $v \neq \emptyset$, which implies (5.4).

The second part of the lemma is proved similarly, except that it involves an iterated integral of a higher order partial derivative of $\hat{K}_{u,u}$:

$$\int_0^{\hat{y}_u} \int_0^{\hat{\eta}_u} \frac{\partial^{2|u|} \hat{K}_{u,u}(\hat{x}_u, \hat{\zeta}_u)}{\partial \hat{\zeta}_u^2} \, d\hat{\zeta}_u d\hat{\eta}_u$$

$$= \sum_{\emptyset \subseteq w \subseteq v \subseteq u} (-1)^{|v|} \frac{\partial^{|w|} \hat{K}_{u,u}(\hat{x}_u, \hat{\eta}_u)}{\partial \hat{\eta}_w}\bigg|_{\substack{\hat{\eta}_v = 0 \\ \hat{\eta}_{u-v} = \hat{y}_{u-v}}} \hat{y}_w.$$

Again by (4.3b) the terms on the right with $w \subset v$ vanish, which implies (5.5). $\square$

Proof of Theorem 5.1. The results of the previous two lemmas are now combined. If the reproducing kernel satisfies smoothness condition (5.1a), then it follows by (5.4) and (5.3a) that

$$|K_{u,u}(x_u, y_u)| = |\hat{K}_{u,u}(x_u + y_u, x_u - y_u)| \leq A_1(s) \prod_{j \in u} |x_j - y_j|.$$

Substituting this inequality into (4.8) and applying Lemma 5.2 for the case $p = 1$ yields upper bounds on $|H(k_u)|$ and $|H_{l,\tau}|$:

$$|H(k_u)| \leq A_1(s) \left[\frac{(1 - b^{-2})}{3(1 - b^{-1})} \right]^{|u|} b^{-\sigma(k_u)},$$

$$|H_{l,\tau}| \leq A_1(l) \left[\frac{(1 - b^{-2})}{3(1 - b^{-1})} \right]^l b^{-\tau}.$$

Formula (4.15) in Theorem 4.4 contains an infinite sum over τ. For large m the binomial coefficient $\binom{m+\tau-1}{l-1}$ in this sum is asymptotic to $\frac{m^{l-1}}{(l-1)!}$ as was noted earlier. Therefore, the inner sum has the following asymptotic upper bound for large m:

$$\left| \sum_{\tau=1}^{+\infty} \binom{m+\tau-1}{l-1} \tilde{R}(l,\tau) b^{-\tau} H_{l,\tau+m-l} \right|$$

$$\lesssim \frac{m^{l-1}}{(l-1)!} \tilde{R}_{\max}(l) \sum_{\tau=1}^{+\infty} b^{-\tau} |H_{s,\tau+m-s}|$$

$$\leq \frac{m^{l-1}}{(l-1)!} \tilde{R}_{\max}(l) A_1(l) \left[\frac{(1-b^{-2})}{3(1-b^{-1})} \right]^l \sum_{\tau=1}^{+\infty} b^{-m+l-2\tau}$$

$$= \frac{m^{l-1}}{(l-1)!} \tilde{R}_{\max}(l) A_1(l) \frac{(b^2-1)^{l-1}}{[3(b-1)]^l} b^{-m}.$$

In the outer sum over l in (4.15) the most significant term is $l = s$, due to the presence of the m^{l-1} above. This term, after substituting the above bound on the inner sum, is just (5.1b).

The proof for the case where the reproducing kernel satisfies smoothness condition (5.2a) is similar, but more tedious. The right-hand side of (5.5) contains two parts. The first is a sum of terms that are linear in $\hat{y}_v = x_v - y_v$. These terms do not contribute to $H(k_u)$, since the integration domain Ω_{k_u} in (4.8a) is symmetric about the planes $x_j = y_j$. Only the second part of the right-hand side of (5.5)

contributes to the upper bound on $|H(k_u)|$. By Lemma 5.2 for the case $p = 2$ it follows that

$$|H(k_u)| \leq A_2(s) \left[\frac{(1 - b^{-3})}{6(1 - b^{-1})} \right]^{|u|} b^{-2\sigma(k_u)}.$$

The asymptotic bound on the mean square discrepancy then follows as for the previous smoothness condition:

$$|H_{l,\tau}| \leq A_2(l) \left[\frac{(1 - b^{-3})}{6(1 - b^{-1})} \right]^{l} b^{-2\tau},$$

$$\left| \sum_{\tau=1}^{+\infty} \binom{m + \tau - 1}{l - 1} \tilde{R}(l,\tau) b^{-\tau} H_{l,\tau+m-l} \right| \lesssim A_2(l) \frac{(b^3 - 1)^{l-1}}{[6(b-1)]^l} b^{-2m}.$$

Substituting this bound into (4.15) and taking the term $l = s$ gives (5.2b). $\square$

For product reproducing kernels the smoothness conditions in Theorem 5.1 may be written in terms of M and ν defined in (4.1). Recall that $\nu(x_1, y_1) = \nu(y_1, x_1)$. If

$$(5.6) \qquad \bar{\nu}_1 \equiv \left\| \frac{\partial \nu(x_1, y_1)}{\partial x_1} \right\|_\infty = \left\| \frac{\partial \nu(x_1, y_1)}{\partial y_1} \right\|_\infty,$$

then by (4.16)

$$\left\| \frac{\partial^{|u|} K_{u,u}}{\partial x_{u-v} \partial y_v} \right\|_\infty = M^{s-|u|} \prod_{j \in u-v} \left\| \frac{\partial \nu(x_j, y_j)}{\partial x_j} \right\|_\infty \prod_{j' \in v} \left\| \frac{\partial \nu(x_{j'}, y_{j'})}{\partial y_{j'}} \right\|_\infty = M^{s-|u|} \bar{\nu}_1^{|u|}.$$

Smoothness condition (5.1a) is satisfied by taking $A_1(l) = M^{s-l} \bar{\nu}_1^l$. If

$$(5.7) \qquad \bar{\nu}_2 \equiv \max \left(\left\| \frac{\partial^2 \nu(x_1, y_1)}{\partial x_1^2} \right\|_\infty, \left\| \frac{\partial^2 \nu(x_1, y_1)}{\partial x_1 \partial y_1} \right\|_\infty, \left\| \frac{\partial^2 \nu(x_1, y_1)}{\partial y_1^2} \right\|_\infty \right),$$

then by (4.16)

$$\left\| \frac{\partial^{|2u|} K_{u,u}}{\partial x_{u-v} \partial x_{u-w} \partial y_v \partial y_w} \right\|_\infty = M^{s-|u|} \prod_{j \in u-(v \cup w)} \left\| \frac{\partial^2 \nu(x_j, y_j)}{\partial x_j^2} \right\|_\infty$$

$$\times \prod_{j' \in (v \cup w)-(v \cap w)} \left\| \frac{\partial^2 \nu(x_{j'}, y_{j'})}{\partial x_{j'} \partial y_{j'}} \right\|_\infty \prod_{j'' \in v \cap w} \left\| \frac{\partial^2 \nu(x_{j''}, y_{j''})}{\partial y_{j''}^2} \right\|_\infty = M^{s-|u|} \bar{\nu}_2^{|u|}.$$

Smoothness condition (5.2a) is satisfied by taking $A_2(l) = M^{s-l} \bar{\nu}_2^l$.

For the product kernels defined by (2.8a) and (2.9a) the only discontinuities in the derivatives of $\nu(x_1, y_1)$ come from the term $B_{2\gamma}(\{x_1 - y_1\})$. For $\gamma = 1$ these product reproducing kernels satisfy (5.6), and thus (5.1a), so the root mean square discrepancies (2.8b) and (2.9b) are $O(N^{-1}[\log_b(N)]^{(s-1)/2})$, by Theorem 5.1. For $\gamma \geq 2$ the product reproducing kernels defined by (2.8a) and (2.9a) satisfy (5.7), and thus (5.2a), so the corresponding root mean square discrepancies are $O(N^{-3/2}[\log_b(N)]^{(s-1)/2})$, by Theorem 5.1.

For any $(0, m, s)$-net the base b must be no smaller than $s - 1$ [Nie92, Corollary 4.21]. This allows one to derive simple lower and upper bounds on $\tilde{R}(l, \tau)$. Suppose that l is odd. The terms in the definition of $\tilde{R}(l, \tau)$ in (4.14) are alternating in

sign and monotonically increasing in magnitude with the summation index, since $l \leq s \leq b + 1$. Therefore,

$$0 \leq \tilde{R}(l,1) \leq \tilde{R}(l,3) \leq \cdots \leq \tilde{R}(l,l) = 1,$$

and

$$0 \geq \tilde{R}(l,2) \geq \tilde{R}(l,4) \geq \cdots \geq \tilde{R}(l,l-1) = 1 - (1 - b^{-1})^{1-l}$$
$$\geq 1 - (1 - b^{-1})^{-b} \geq -3 \quad \text{for } b \geq 2.$$

A similar argument gives $-3 \leq \tilde{R}(l,\tau) \leq 1$ for even l as well. Therefore, one may choose $\tilde{R}_{\max}(l)$ in Theorem 5.1 to be 3.

6. Discussion and conclusion

The bounds on the $\mathcal{L}_\infty$-star discrepancy for (t,m,s)-nets [Nie92, Theorem 4.10] are of the form

$$D_\infty^*(P) \leq C(s,b) b^t \frac{(\log N)^{s-1}}{N},$$

that is, the value of t does not effect the order of the $\mathcal{L}_\infty$-star discrepancy, only the leading constant. Although the results proven in this article are only for $(0,m,s)$-nets, the author conjectures that the orders of the root mean square discrepancy for randomized nets will be the same even if t is nonzero. Of course, this remains to be proven.

An interesting question, then, is whether the discrepancies of known (t,m,s)-nets attain the same asymptotic order as the root mean square discrepancy of randomized $(0,m,s)$-nets. Although no mathematical proof has been found, some computational evidence is presented here. Figures 1–6 show three different discrepancies for three different kinds of (t,m,s)-nets. The nets are the first b^m points of:

  i. the $(0,s)$-sequence of Faure [Fau82] with b chosen to be the smallest prime number $\geq s$,
 ii. the (t,s)-sequence in base 2 of Niederreiter [BFN92], and
iii. the (t,s)-sequence in base 2 of Sobol' [BF88].

The initial point for all of these sequences is the origin.

The first discrepancy considered is (2.8b) with $\beta = 1$ and $\gamma = 1$. Only smoothness condition (5.1a) is satisfied since $\gamma = 1$. The root mean square discrepancy for randomized $(0,m,s)$-nets is $O(N^{-1+\epsilon})$, and Figure 1 shows that the three $(t,m,5)$-nets chosen also have this same asymptotic order.

The second discrepancy considered is (2.8b) with $\beta = 1$ and $\gamma = 2$. Now smoothness condition (5.2a) is satisfied, so the root mean square discrepancy for randomized $(0,m,s)$-nets is $O(N^{-3/2+\epsilon})$. However, Figure 2 shows that the discrepancies of the three $(t,m,5)$-nets are only $O(N^{-1+\epsilon})$. Thus, these three popular nets have *worse than average* discrepancy, in this case. This means that these three nets are not able to attain higher quadrature accuracy than $O(N^{-1+\epsilon})$, even if the integrand has sufficient smoothness.

To understand why consider a linear integrand $f(x_1) = x_1$, and a one-dimensional net with $N = b^m$ points, $\{0, b^{-m}, \ldots, 1 - b^{-m}\}$. Quadrature using this net corresponds to the left rectangle rule, which can only attain an error of $O(N^{-1})$, since
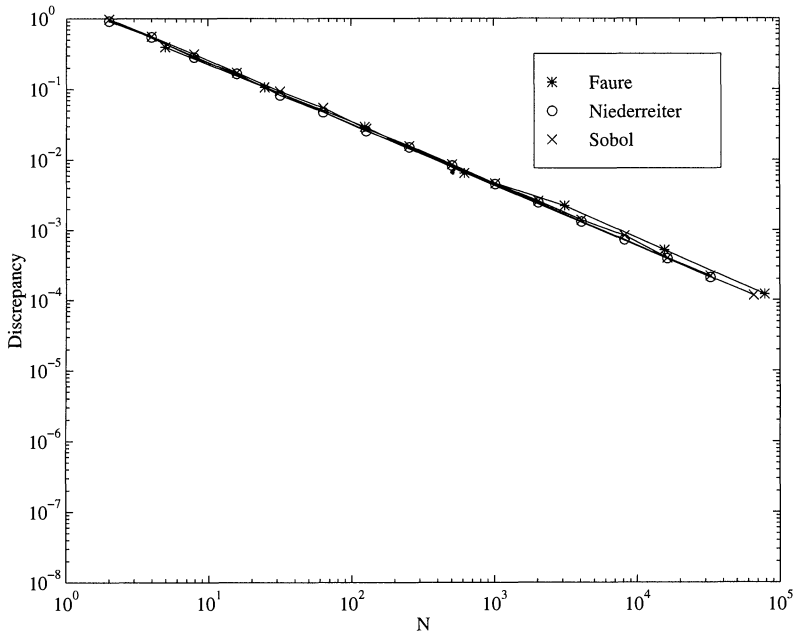
FIGURE 1. The discrepancy defined in (2.8b) with $\beta = 1$, $\gamma = 1$ as a function of N for three different $(t, m, 5)$-nets
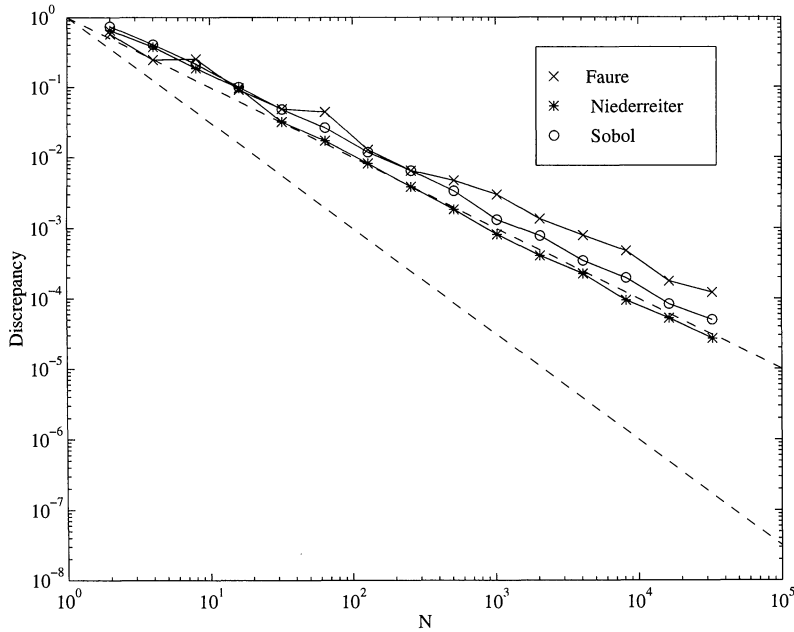

FIGURE 2. The discrepancy defined in (2.8b) with $\beta = 1$, $\gamma = 2$ as a function of N for three different $(t, m, 5)$-nets. The dashed reference lines are plots of N^{-1} and $N^{-1.5}$
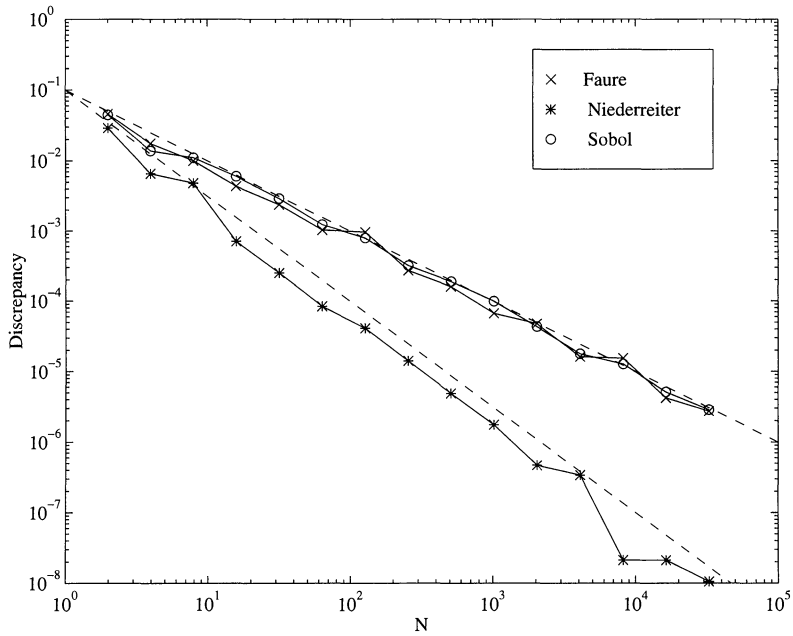
FIGURE 3. The discrepancy defined in (2.9b) with $\beta = 1$, $\gamma = 2$ as a function of N for three different $(t, m, 3)$-nets. The dashed reference lines are plots of $0.1N^{-1}$ and $0.1N^{-1.5}$
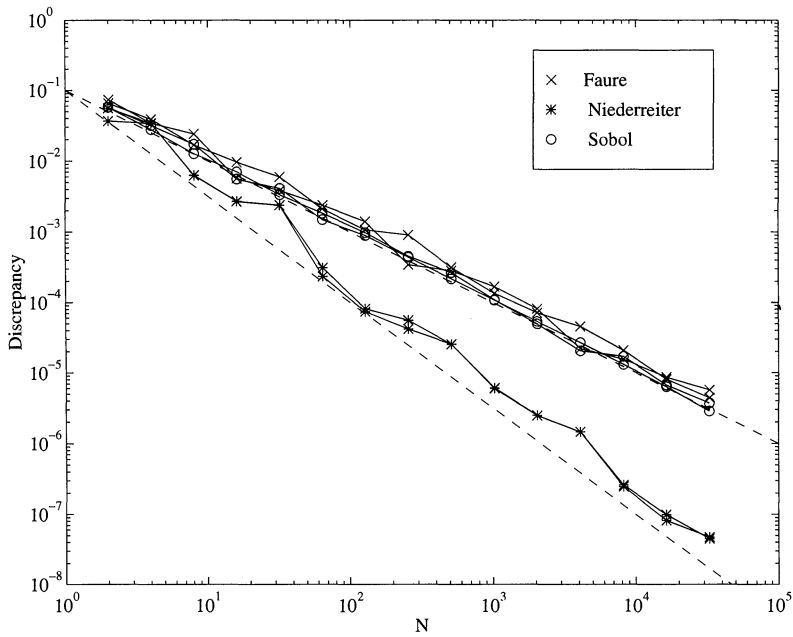

FIGURE 4. The discrepancy defined in (2.9b) with $\beta = 1$, $\gamma = 2$ as a function of N for three different $(t, m, 5)$-nets. The dashed reference lines are plots of $0.1N^{-1}$ and $0.1N^{-1.5}$

the points have a leftward bias. On the other hand, quadrature using Owen's randomization of this net attains an error of $O(N^{-3/2})$. The Faure, Niederreiter and Sobol' nets share the same problem as the one-dimensional net.

The third discrepancy considered is (2.9b) with $\beta = 1$ and $\gamma = 2$. Again the root mean square discrepancy for randomized $(0, m, s)$-nets is $O(N^{-3/2+\epsilon})$. For dimensions $s = 3, 5, 8, 10$ the discrepancies of the Faure and Sobol' nets are still only $O(N^{-1+\epsilon})$, as shown in Figures 3 through 6. However, the discrepancies of the Niederreiter nets decay faster than $O(N^{-1})$. For smaller dimensions (Figures 3 and 4) the decay rate appears to be $O(N^{-3/2+\epsilon})$. For the higher dimensions (Figures 5 and 6) the decay rate is not as good. This may be because for larger s the factor $[\log(N)]^{(s-1)/2}$ remains significant for the range of N values shown in the plots.

For $s = 5$ Figure 4 displays two curves for each type of net corresponding to taking different starting points in the $(t, 5)$-sequence. One curve takes the first 5^m points of the Faure sequence for $m = 1, \ldots, 7$. Another curve takes the $5^7 + 1^{\text{st}}$ to $5^{m+7^{\text{th}}}$ points of the Faure sequence for $m = 1, \ldots, 7$. The curves for the Niederreiter and Sobol' nets are obtained in a similar manner. The choice of starting point does not seem to affect the asymptotic order of the this discrepancy.

Because the discrepancy as plotted in Figure 4 is for a space of *periodic* integrands, the example of a linear integrand mentioned for the previous case does not apply. The Niederreiter nets are able to integrate periodic functions more accurately than non-periodic ones, and more accurately than Faure and Sobol' nets, especially in lower dimensions. Therefore, when using Niederreiter nets for quadrature it is suggested that one periodize the integrand, that is, use the quadrature rule

$$Q_{\text{per}}(f) \equiv \frac{1}{N} \sum_{z \in P} f(|2z_1 - 1|, \ldots, |2z_s - 1|)$$

instead of the $Q(f)$ defined in (1.1).

To demonstrate this suggestion with a practical example, consider the problem of computing multivariate normal probabilities:

$$(6.1) \qquad I(a, b, \Sigma) = \frac{1}{\sqrt{|\Sigma|(2\pi)^s}} \int_{a_1}^{b_1} \cdots \int_{a_s}^{b_s} e^{-\frac{1}{2}\theta' \Sigma^{-1} \theta} \, d\theta,$$

where a and b are known s-dimensional vectors that define the interval of integration, and Σ is a given $s \times s$ positive definite covariance matrix. Alan Genz [Gen92] has proposed a transformation of variables that results in $I(a, b, \Sigma)$ being written as an integral over an $s - 1$-dimensional unit cube. We consider the specific case of $s = 6$,

$$(6.2a) \qquad a = (-\infty, \ldots, -\infty), \qquad b = (0, 0.5, 1, 1.5, 2, 2.5, 3)$$

$$(6.2b) \qquad \Sigma = (\sigma_{ij}), \text{ where } \sigma_{ij} = \begin{cases} 1, & i = j, \\ 0.5, & i \neq j. \end{cases}$$

Because the off-diagonal elements of the covariance matrix are all equal, the multivariate normal probability may be reduced to a one-dimensional integral [Ton90], which may be evaluated by standard numerical techniques to provide the "exact" answer of 0.6135764963. Figure 7 shows how the absolute error decays when applying the quadrature rules Q and Q_{per} with Niederreiter nets to the integral (6.1) after Genz's transformation. The error of Q_{per} has a faster rate of decay.
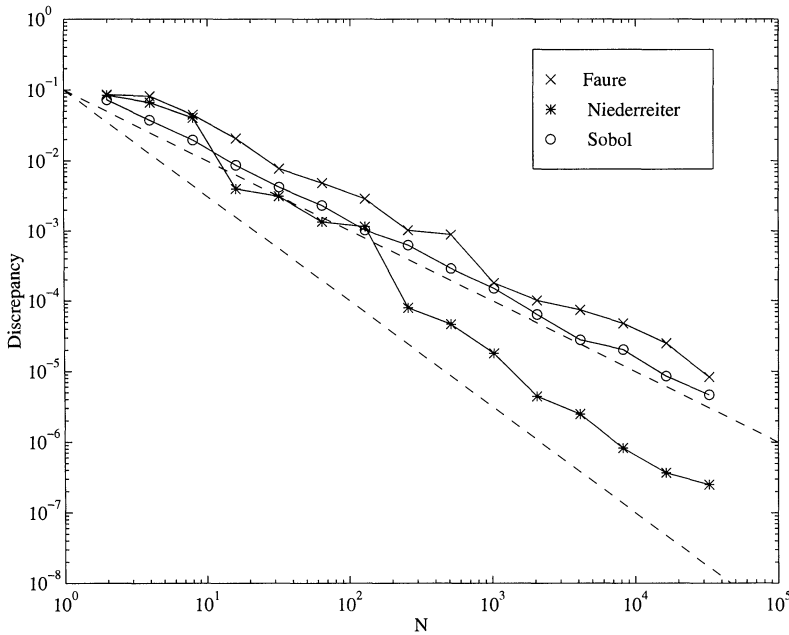
FIGURE 5. The discrepancy defined in (2.9b) with $\beta = 1$, $\gamma = 2$ as a function of N for three different $(t, m, 8)$-nets. The dashed reference lines are plots of $0.1N^{-1}$ and $0.1N^{-1.5}$

The $\mathcal{L}_2$-star discrepancy defined in (2.3) gives the average-case quadrature error for Brownian sheets. Because the corresponding reproducing kernel satisfies smoothness condition (5.1a), the $\mathcal{L}_2$-star discrepancy is $O(N^{-1}[\log(N)]^{(s-1)/2})$ for $(0, m, s)$-nets, a fact previously shown in [Hic96a]. For Brownian motion in Lévy's sense the covariance (reproducing kernel) is

$$K(x, y) = \frac{\|x\|_2 + \|y\|_2 - \|x - y\|_2}{2},$$

and the corresponding discrepancy is known to be at best $O(N^{-\frac{1}{2} - \frac{1}{2s}})$ for any set [Was93]. The reason that this discrepancy has larger order than the $\mathcal{L}_2$-star discrepancy for $s > 1$ is that the reproducing kernel is not sufficiently smooth. Consider the case $s = 2$. The components $K_{u,u}$ of this reproducing kernel are

$$K_{\emptyset,\emptyset} = \frac{\sqrt{2} + \log(1 + \sqrt{2})}{3},$$

$$K_{\{1\},\{1\}}(x_1, y_1) = -\frac{|x_1 - y_1|}{2}, \qquad K_{\{2\},\{2\}}(x_2, y_2) = -\frac{|x_2 - y_2|}{2},$$

$$K_{\{1,2\},\{1,2\}}(x_1, x_2, y_1, y_2)$$

$$= \frac{1}{4}\left\{\sqrt{x_1^2 + x_2^2} - \sqrt{x_2^2 + y_1^2} - \sqrt{x_1^2 + y_2^2} + \sqrt{y_1^2 + y_2^2}\right\}$$

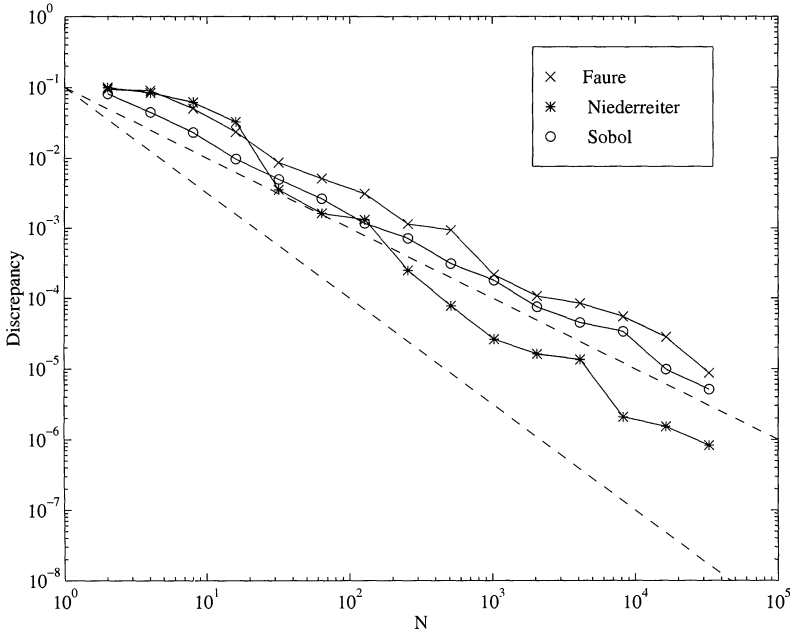$$+ \frac{1}{2}\left\{|x_1 - y_1| + |x_2 - y_2| - \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}\right\}.$$

FIGURE 6. The discrepancy defined in (2.9b) with $\beta = 1$, $\gamma = 2$ as a function of N for three different $(t, m, 10)$-nets. The dashed reference lines are plots of $0.1N^{-1}$ and $0.1N^{-1.5}$
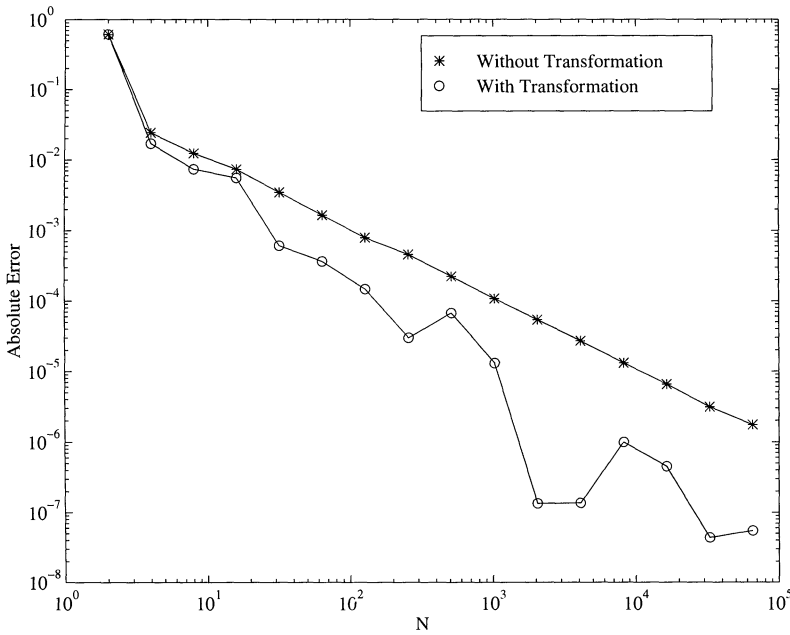

FIGURE 7. The absolute error for computing a multivariate normal probability using Niederreiter nets with and without a periodizing transformation of the integrand

The last term of component $K_{\{1,2\},\{1,2\}}$ does not satisfy condition (5.1a), so an $O(N^{-1}[\log(N)]^{(s-1)/2})$ discrepancy cannot be attained.

Theorem 5.1 and (1.2) imply that for randomized $(0, m, s)$-nets

$$E\left\{\sup_{V(f)=1}[\mathrm{Err}(f)]^2\right\} = E\{[D_2(P)]^2\} = O(N^{-p}[\log(N)]^{s-1}) \text{ as } N \to \infty,$$

where the variation of the integrand, $V(f)$, is defined in (1.4), and $p = 2$ or 3 depending on the smoothness of the integrands. The smoothness of the integrands (roughly that of the reproducing kernel) required for $p = 3$ is greater than that assumed in [Owe97b], where it is shown that, for a fixed integrand f,

$$E\left\{[\mathrm{Err}(f)]^2\right\} = O(N^{-3}[\log(N)]^{s-1}) \text{ as } N \to \infty$$

for randomized $(0, m, s)$-nets. The reason is that in the latter case the integrand is fixed in advance, whereas in the former case the integrand is chosen pessimistically after picking a specific randomized net. On the other hand, Theorem 5.1 and (1.2) also imply

$$E\left\{E_{f\in F}[\mathrm{Err}(f)]^2\right\} = E\{[D_2(P)]^2\} = O(N^{-p}[\log(N)]^{s-1}) \text{ as } N \to \infty,$$

where again $p = 2$ or 3 depending on the smoothness of the integrands. The smoothness of the integrands in F required for $p = 3$ is less than that required for the worst-case result and is similar to that assumed in [Owe97b].

Theorem 5.1 demonstrates that the asymptotic order of the discrepancy depends on the smoothness of the reproducing kernels. It appears that $O(N^{-3/2+\epsilon})$ is the best possible order attainable by assuming only sufficient smoothness of the reproducing kernel. However, if one considers spaces of Walsh series or Haar functions, it is likely that the order of the corresponding discrepancy may be improved further. This is because it has been shown by other means that (t, m, s)-nets accurately integrate these type of functions [LSW94], [LT94], [Ent96].

REFERENCES

[AS64]   M. Abramowitz and I. A. Stegun (eds.), *Handbook of mathematical functions with formulas, graphs and mathematical tables*, U. S. Government Printing Office, Washington, DC, 1964. MR **29**:4914

[BF88]   P. Bratley and B. L. Fox, *Algorithm 659: Implementing Sobol's quasirandom sequence generator*, ACM Trans. Math. Softw. **14** (1988), 88–100.

[BFN92] P. Bratley, B. L. Fox, and H. Niederreiter, *Implementation and tests of low-discrepancy sequences*, ACM Trans. Model. Comput. Simul. **2** (1992), 195–213.

[Ent96] K. Entacher, *Generalized Haar function systems in the theory of uniform distribution modulo one*, Ph.D. thesis, University of Salzburg, 1996.

[Fau82] H. Faure, *Discrépance de suites associées à un système de numération (en dimension s)*, Acta Arith. **41** (1982), 337–351. MR **84m**:10050

[Gen92] A. Genz, *Numerical computation of multivariate normal probabilities*, J. Comput. Graph. Statist. **1** (1992), 141–150.

[Hic95] F. J. Hickernell, *A comparison of random and quasirandom points for multidimensional quadrature*, Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing (H. Niederreiter and P. J.-S. Shiue, eds.), Lecture Notes in Statistics, vol. 106, Springer-Verlag, New York, 1995, pp. 213–227. MR **97j**:65002

[Hic96a] F. J. Hickernell, *The mean square discrepancy of randomized nets*, ACM Trans. Model. Comput. Simul. **6** (1996), 274–296.

[Hic96b] F. J. Hickernell, *Quadrature error bounds with applications to lattice rules*, SIAM J. Numer. Anal. **33** (1996), 1995–2016. MR **97m**:65050

[Hic98] F. J. Hickernell, *A generalized discrepancy and quadrature error bound*, Math. Comp. **67** (1998), 299–322. MR **98c**:65032

[HW81] L. K. Hua and Y. Wang, *Applications of number theory to numerical analysis*, Springer-Verlag, Berlin, and Science Press, Beijing, 1981. MR **83g**:10034

[LSW94] G. Larcher, W. Ch. Schmid, and R. Wolf, *Representation of functions as Walsh series to different bases and an application to the numerical integration of high-dimensional Walsh series.*, Math. Comp. **63** (1994), 701–716. MR **95a**:65045

[LT94] G. Larcher and C. Traunfellner, *On the numerical integration of Walsh series by number-theoretic methods*, Math. Comp. **63** (1994), 277–291. MR **94j**:65030

[MC94] W. J. Morokoff and R. E. Caflisch, *Quasi-random sequences and their discrepancies*, SIAM J. Sci. Comput. **15** (1994), 1251–1279. MR **95e**:65009

[Nie92] H. Niederreiter, *Random number generation and quasi-Monte Carlo methods*, SIAM, Philadelphia, 1992. MR **93h**:65008

[Owe95] A. B. Owen, *Randomly permuted (t, m, s)-nets and (t, s)-sequences*, Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing (H. Niederreiter and P. J.-S. Shiue, eds.), Lecture Notes in Statistics, vol. 106, Springer-Verlag, New York, 1995, pp. 299–317. MR **97k**:65013

[Owe97a] A. B. Owen, *Monte Carlo variance of scrambled equidistribution quadrature*, SIAM J. Numer. Anal. **34** (1997), 1884–1910. CMP 98:01

[Owe97b] A. B. Owen, *Scrambled net variance for integrals of smooth functions*, Ann. Stat. **25** (1997), 1541–1562. CMP 97:16

[Rit95] K. Ritter, *Average case analysis of numerical problems*, Ph.D. thesis, Universität Erlangen-Nürnberg, Erlangen, Germany, 1995.

[Sai88] S. Saitoh, *Theory of reproducing kernels and its applications*, Longman Scientific & Technical, Essex, England, 1988. MR **90f**:46045

[SJ94] I. H. Sloan and S. Joe, *Lattice methods for multiple integration*, Oxford University Press, Oxford, 1994. MR **98a**:65026

[Ton90] Y. L. Tong, *The multivariate normal distribution*, Springer-Verlag, New York, 1990. MR **91g**:60021

[Wah90] G. Wahba, *Spline models for observational data*, SIAM, Philadelphia, 1990. MR **91g**:62028

[Was93] G. W. Wasilkowski, *Integration and approximation of multivariate functions: Average case complexity with isotropic Wiener measure*, Bull. Amer. Math. Soc. **28** (1993), 308–314. MR **93i**:65136

[Woź91] H. Woźniakowski, *Average case complexity of multivariate integration*, Bull. Amer. Math. Soc. **24** (1991), 185–194. MR **91i**:65224

[Zar68] S. K. Zaremba, *Some applications of multidimensional integration by parts*, Ann. Polon. Math. **21** (1968), 85–96. MR **38**:4034

DEPARTMENT OF MATHEMATICS, HONG KONG BAPTIST UNIVERSITY, KOWLOON TONG, HONG KONG

E-mail address: fred@hkbu.edu.hk, http://www.math.hkbu.edu.hk/~fred