A GENERALIZED DISCREPANCY AND QUADRATURE ERROR BOUND

FRED J. HICKERNELL

ABSTRACT. An error bound for multidimensional quadrature is derived that includes the Koksma-Hlawka inequality as a special case. This error bound takes the form of a product of two terms. One term, which depends only on the integrand, is defined as a generalized variation. The other term, which depends only on the quadrature rule, is defined as a generalized discrepancy. The generalized discrepancy is a figure of merit for quadrature rules and includes as special cases the $\mathcal{L}^p$-star discrepancy and P_α that arises in the study of lattice rules.

1. INTRODUCTION

The multidimensional integral

$$I(f) = \int_{C^s} f(x)\, dx$$

can be approximated by the sample mean,

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

where $C^s = [0,1)^s$ is the s-dimensional unit cube and P is some random or deterministic sample of N points in C^s. (P may have multiple copies of the same point.) Many quasi-random samples designed for quadrature have been discussed in the literature. These include (t, m, s)-nets [Nie92, Chapter 4] and integration lattices [SJ94].

Samples P whose points are uniformly scattered over C^s tend to give more accurate approximations to the integral. Some quadrature error bounds in the literature take the form

$$(1.2) \qquad |I(f) - Q(f)| \le D(P)V(f),$$

where $V(f)$ is a measure of the variation or fluctuation of the integrand, and $D(P)$ is a figure of merit for the quadrature rule, or equivalently, a measure of non-uniformity for the sample P. A popular figure of merit is the star discrepancy [Nie92, Chapter 2] and its generalization the $\mathcal{L}^p$-star discrepancy. For integration lattices and periodic integrands the quantity P_α is often used as a figure of merit

[SJ94, Chapter 4]. A generalization of P_α to arbitrary samples and non-periodic integrands is also possible [Hic96]. Asymptotic bounds on these figures of merit as N tends to infinity have been derived, but in practice it may be desirable to compute their actual values in order to compare different samples P. The question then arises as to which figure of merit to choose.

This article derives a family of quadrature error bounds of the form (1.2). The figure of merit, $D(P)$, which we call a *generalized $\mathcal{L}^p$-discrepancy*, includes the star discrepancy and P_α as special cases. Other figures of merit, which have certain advantages over the star discrepancy, are also derived. We first review the figures of merit mentioned above and then describe the key idea underlying our analysis, reproducing kernel Hilbert spaces.

Let $|\ |$ denote the number of points in a set. The empirical distribution function associated with the sample P can be written as $|P \cap [0, x)|/N$. The uniform distribution function on the unit cube is $\mathrm{Vol}([0, x))$, the volume of the rectangular solid $[0, x)$. The star discrepancy is defined as the maximum deviation between these two distributions [Nie92, Definition 2.1]:

$$(1.3) \qquad D^*(P) = \sup_{x \in C^s} \left| \frac{|P \cap [0, x)|}{N} - \mathrm{Vol}([0, x)) \right|.$$

The Koksma-Hlawka inequality [Nie92, Theorem 2.11] is

$$(1.4) \qquad |I(f) - Q(f)| \leq D^*(P)V(f),$$

where $V(f)$ is the variation of f on $\bar{C}^s = [0, 1]^s$ in the sense of Hardy and Krause. It is possible to generalize the star discrepancy by defining an $\mathcal{L}^p$-star discrepancy. Before doing so some useful notation is introduced.

Let $S = \{1, \ldots, s\}$ be the set of coordinate indices. For any $u \subseteq S$ let $|u|$ denote its cardinality and let $C^u = [0, 1)^u$ denote the $|u|$-dimensional unit cube involving the coordinates in u. This notation allows us to distinguish cubes of the same dimension in different coordinate directions. Furthermore, let x_u denote the vector containing the components of x whose indices are in u, and let $dx_u = \prod_{j \in u} dx_j$ denote the uniform measure on $C^u = [0, 1)^u$. This notation has been used in [Owe92, Hic96].

Let $\|\ \|_p$ denote the $\mathcal{L}^p$-norm of a function on C^u, that is

$$\|f\|_p = \left[\int_{C^u} |f|^p \, dx_u \right]^{1/p} \quad (1 \leq p < \infty), \quad \|f\|_\infty = \inf \left\{ \gamma : |f| \leq \gamma \text{ a.e.} \right\},$$

where "a.e." means almost everywhere in C^u. This notation is extended to the case of a vector of functions (f_u), where u is an index running over some or all of the subsets of S, and each f_u is a function on C^u:

$$\|(f_u)\|_p = \left[\sum_u \|f_u\|_p^p \right]^{1/p} = \left[\sum_u \int_{C^u} |f_u|^p \, dx_u \right]^{1/p} \quad (1 \leq p < \infty),$$

$$\|(f_u)\|_\infty = \max_u \|f_u\|_\infty = \max_u \inf \left\{ \gamma : |f_u| \leq \gamma \text{ a.e.} \right\}.$$

In the special case of a vector of constants $\|\ \|_p$ corresponds to the ℓ^p-norm.

The $\mathcal{L}^p$-star discrepancy is often defined as

$$\left\| \frac{|P \cap [0, x)|}{N} - \mathrm{Vol}([0, x)) \right\|_p,$$

a natural generalization of (1.3). However, if one wishes to generalize (1.4) to

$$(1.5) \qquad |I(f) - Q(f)| \le D_p^*(P) V_q(f) \quad (p^{-1} + q^{-1} = 1),$$

then the appropriate definition of $D_p^*(P)$ is

$$(1.6) \qquad D_p^*(P) = \left\| \left(\frac{|P_u \cap [0, x_u)|}{N} - \mathrm{Vol}([0, x_u)) \right)_{u \ne \emptyset} \right\|_p,$$

where P_u denotes the projection of the sample P into the cube C^u. This definition also corresponds to (1.3) in the case $p = \infty$. The $\mathcal{L}^p$-variation, $V_p(f)$, is a generalization of the variation in the sense of Hardy and Krause and can be written as

$$(1.7) \qquad V_p(f) = \left\| \left(\frac{\partial^{|u|} f}{\partial x_u} \bigg|_{x_{S-u} = (1, \dots, 1)} \right)_{u \ne \emptyset} \right\|_p.$$

Error bound (1.5) was derived by Zaremba [Zar68] for $p = 2$ and by Sobol' [Sob69, Chapter 8] in general. This paper provides an alternative derivation.

The figure of merit P_α arises in the study of good lattice point sets and their generalization, integration lattices. It can be written [SJ94, Equation (4.8)] as

$$(1.8) \qquad P_\alpha = -1 + \frac{1}{N} \sum_{z \in P} \prod_{j=1}^{s} \left[1 - \frac{(-4\pi^2)^{\alpha/2}}{\alpha!} B_\alpha(z_j) \right] \quad (\alpha \text{ even}),$$

where B_α denotes the Bernoulli polynomial of degree α [AS64, Chapter 23]. Let $\mathcal{F}(f)(k)$ denote the Fourier coefficients of the function f, that is,

$$\mathcal{F}(f)(k) = \int_{C^s} f(x) e^{-2\pi i k \bullet x} \, dx.$$

For periodic integrands belonging to the set

$$(1.9) \qquad \mathcal{E}^\alpha(K) = \left\{ f : |\mathcal{F}(f)(k)| \le \frac{K}{(\bar{k}_1 \dots \bar{k}_s)^\alpha} \right\}, \qquad \bar{k}_j = \max(|k_j|, 1),$$

the quadrature error for integration lattices satisfies [SJ94, Section 5.2]

$$(1.10) \qquad |I(f) - Q(f)| \le P_\alpha K,$$

which is also of the form (1.2).

For $\alpha = 2$ the author has extended this error bound to general samples P and integrands that are not necessarily periodic. For any positive constant β let

$$(1.11\text{a}) \qquad c^2 = -1 + \frac{1}{N^2} \sum_{z, z' \in P} \prod_{j=1}^{s} \left\{ 1 + \beta^2 \left[\frac{1}{2} B_2(\{z_j - z_j'\}) + B_1(z_j) B_1(z_j') \right] \right\},$$

$$(1.11\text{b}) \qquad \tilde{c}^2 = -1 + \frac{1}{N^2} \sum_{z, z' \in P} \prod_{j=1}^{s} \left[1 + \frac{\beta^2}{2} B_2(\{z_j - z_j'\}) \right],$$

where $\{\}$ denotes the fractional part of a real number or vector. For a certain broad class of integrands Hickernell [Hic96, Theorems 2.1 and 3.1] showed that

$$(1.12\text{a}) \qquad |I(f) - Q(f)| \le c \, \|f\| \quad \text{in general},$$

$$(1.12\text{b}) \qquad |I(f) - Q(f)| \le \tilde{c} \, \|f\| \quad \text{if } f \text{ is periodic}.$$

Here $\|\| \ \|\|$ is a norm whose definition involves $\mathcal{L}^2$-norms of the mixed partial derivatives of the function. Note that $\tilde{c}^2 = P_2$ for lattice rules when $\beta = 2\pi$.

The primary mathematical tool used in [Hic96] is reproducing kernel Hilbert spaces. It is also the main tool used here. Let $(X, \langle, \rangle)$ be some Hilbert space of real-valued functions on C^s, where $\langle, \rangle$ denotes the inner product. For any $x \in C^s$ let T_x denote the evaluation functional defined as

(1.13) $$T_x(f) = f(x) \quad \forall f \in X.$$

If T_x is bounded, then by the Riesz Representation Theorem there exists a reproducing kernel $\eta(\bullet, x) \in X$ such that

$$f(x) = T_x(f) = \langle \eta(\bullet, x), f \rangle \quad \forall f \in X.$$

Given the reproducing kernel it is straightforward to compute the representer ζ for any other bounded linear functional T:

$$T(f) = \langle \zeta, f \rangle \quad \forall f \in X, \quad \text{where } \zeta(x) = \langle \eta(\bullet, x), \zeta \rangle = T(\eta(\bullet, x)).$$

In particular, if the linear functional corresponding to quadrature error, $I - Q$, is bounded, then its representer can be found in terms of η:

$$(I - Q)(f) = \langle \xi, f \rangle \quad \forall f \in X, \quad \text{where } \xi(x) = (I - Q)(\eta(\bullet, x)).$$

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

(1.14) $$|I(f) - Q(f)| = |\langle \xi, f \rangle| \le \|\|\xi\|\| \ \|\|f\|\|,$$

where $\|\| \ \|\|$ is the norm induced by the inner product (not the $\mathcal{L}^p$-norm). Equality holds when f is a constant multiple of ξ. This means that ξ is the worst-case integrand. The quantity $\|\|\xi\|\|$ depends only on the points P and may be identified as $D(P)$ in (1.2), a figure of merit for P. Likewise $\|\|f\|\|$ can be identified as $V(f)$, a measure of the size or variation of the integrand. (In fact, the actual definition of $V(f)$ will be the norm of f less its constant part.)

This reproducing kernel Hilbert space approach is quite general. It will be shown that for different choices of $(X, \langle, \rangle)$ one may obtain the error bounds and figures of merit summarized above ((1.4)-(1.12)) as well as some new ones. To facilitate the exposition we first present the analysis for dimension $s = 1$ in the next section. Section 3 considers the arbitrary s-dimensional problem and contains the main results — quadrature error bounds in terms of a generalized $\mathcal{L}^p$-discrepancy and variation. Periodic integrands are treated in Section 4. In Section 5 several specific examples of the generalized discrepancy and variation are given and their relative merits are discussed.

2. One-dimensional case

For $s = 1$ the key arguments and results are similar to that for multidimensional quadrature $(s > 1)$. However, the technical details are simpler.

Bernoulli polynomials arise in the derivation of the Euler-Maclaurin Summation Formula [AS64, Equation 23.1.30]. They also facilitate the derivation of the theory presented here. Bernoulli polynomials are defined⋅by the following generating function:

$$\frac{te^{xt}}{e^t - 1} = \sum_{n=0}^{\infty} B_n(x) \frac{t^n}{n!}.$$

The first few are

$$B_0(x) = 1, \qquad {}^{\prime}B_1(x) = x - 1/2, \qquad B_2(x) = x^2 - x + 1/6.$$

Some relevant properties of Bernoulli polynomials quoted in [AS64, Chapter 23] are given in the following lemma. Of particular importance is (2.1) below, which relates $f(y)$ to an integral involving the derivatives of f and Bernoulli polynomials.

Lemma 2.1. *For non-negative integers n Bernoulli polynomials have the following properties:*

$$B_n(1-x) = (-1)^n B_n(x) \quad (n \geq 0), \qquad B_n(0) = B_n(1) \quad (n \neq 1),$$

$$\frac{dB_n(x)}{dx} = nB_{n-1}(x) \quad (n > 0), \qquad \int_0^1 B_n(x) \, dx = 0 \quad (n > 0).$$

For any fixed y, $\{x - y\}$ as a function of x has jump discontinuities at $\{x\} = \{y\}$ and has slope 1 everywhere else. Furthermore,

$$B_n(\{x - y\}) \text{ is continuous for } n \neq 1,$$

$$\frac{\partial}{\partial x} B_n(\{x - y\}) = nB_{n-1}(\{x - y\}) \quad (n > 1).$$

If f is any function whose n^{th} derivative is Lebesgue integrable on $[0, 1)$, and F is an anti-derivative of f, then for all $y \in [0, 1)$

$$(2.1) \qquad \frac{(-1)^n}{n!} \int_0^1 B_n(\{x - y\}) \frac{d^n f}{dx^n} \, dx = \sum_{i=0}^{n} \frac{B_i(y)}{i!} \left. \frac{d^i F}{dx^i} \right|_0^1 - f(y) \quad (n > 0),$$

where $|_0^1$ represents the change in a function from 0 to 1.

Reproducing kernels require that the space of integrands have enough regularity to insure that T_x, as defined in (1.13), is bounded. Integrability is not enough, but requiring the first derivative of the integrand to be integrable is sufficient. For any p, $1 \leq p \leq \infty$, let

$$(2.2) \qquad X_p \equiv \left\{ f : \frac{df}{dx} \in \mathcal{L}^p([0, 1)) \right\}.$$

The sets X_p are subsets of absolutely continuous functions, and $X_q \subset X_p$ for $q > p$. For example, the function f defined as

$$f(x) = \text{sign}(x - 1/2)|x - 1/2|^{1/q}, \qquad \frac{df}{dx} = \frac{1}{q}|x - 1/2|^{-1+1/q},$$

is continuous on $[0, 1)$ for all $q \geq 0$, and $f \in X_p$ for all $p^{-1} + q^{-1} > 1$.

One can define an inner product and reproducing kernel on X_2. Suppose η has the following form:

$$(2.3) \qquad \eta(x, y) = M + \beta^2 \left[\mu(x) + \mu(y) + \frac{1}{2} B_2(\{x - y\}) + B_1(x)B_1(y) \right],$$

for some constant M (to be specified below), some arbitrary positive constant β and some arbitrary function μ satisfying

$$(2.4) \qquad \mu \in X_\infty, \qquad \int_0^1 \mu(x)dx = 0.$$

For any $y \in [0, 1)$, $\eta(\bullet, y) \in X_2$, and for any function $f \in X_2$,

$$\frac{\partial \eta}{\partial x} = \beta^2 \left[\frac{d\mu}{dx} + B_1(\{x - y\}) + B_0(x)B_1(y) \right],$$

$$\beta^{-2} \int_0^1 \frac{\partial \eta}{\partial x} \frac{df}{dx} \, dx = f(y) - \int_0^1 \left(f - \frac{d\mu}{dx} \frac{df}{dx} \right) \, dx,$$

by (2.1). The second term in the above expression can be identified as some linear functional of f:

$$(2.5) \qquad\qquad L(f) \equiv \int_0^1 \left(f - \frac{d\mu}{dx} \frac{df}{dx} \right) \, dx.$$

The inner product on X and the induced norm are defined as

$$(2.6) \quad \langle f, g \rangle = L(f)L(g) + \beta^{-2} \int_0^1 \frac{df}{dx} \frac{dg}{dx} \, dx, \quad \|f\|_2 = \left\| \left(L(f), \beta^{-1} \frac{df}{dx} \right) \right\|_2.$$

The η defined above is then the reproducing kernel, that is,

$$(2.7) \qquad\qquad f(x) = \langle \eta(\bullet, x), f \rangle \quad \forall f \in X_2,$$

provided that M satisfies

$$(2.8) \qquad 1 = L(\eta(\bullet, y)) = \int_0^1 \left(\eta - \frac{d\mu}{dx} \frac{\partial \eta}{\partial x} \right) \, dx = M - \beta^2 \int_0^1 \left(\frac{d\mu}{dx} \right)^2 \, dx.$$

Thus, the constant M is determined in terms of β and μ:

$$(2.9) \qquad\qquad M = 1 + \beta^2 \int_0^1 \left(\frac{d\mu}{dx} \right)^2 \, dx.$$

Note that $L(\eta(\bullet, y)) = 1$, implying that the constant 1 is the representer for the linear functional L:

$$L(f) = \langle 1, f \rangle \quad \forall f \in X_2.$$

Having defined the Hilbert space $(X, \langle, \rangle)$ and found its reproducing kernel it is straightforward to compute the worst-case integrand ξ and its norm:

$$(2.10a) \qquad \xi(x) = I(\eta(\bullet, x)) - Q(\eta(\bullet, x))$$

$$= -\frac{\beta^2}{N} \sum_{z \in P} \left[\mu(z) + \frac{1}{2} B_2(\{x - z\}) + B_1(x)B_1(z) \right],$$

$$(2.10b) \qquad \frac{d\xi}{dx} = -\frac{\beta^2}{N} \sum_{z \in P} [B_1(\{x - z\}) + B_1(z)]$$

$$= -\frac{\beta^2}{N} \sum_{z \in P} [x - 1_{x>z}] = \beta^2 \left[-x + \frac{1}{N} \sum_{z \in P} 1_{x>z} \right],$$

where $1_{x>z}$ is the indicator function. Since 1 is the representer for L, and quadrature rule (1.1) is exact for constants, it follows that

$$L(\xi) = \langle \xi, 1 \rangle = (I - Q)(1) = 0.$$

Finally the norm of ξ follows by straightforward calculations:

$$(2.11) \quad \|\|\xi\|\|_2^2 = \langle \xi, \xi \rangle = \left\|\left(L(\xi), \beta^{-1}\frac{d\xi}{dx}\right)\right\|_2^2 = \left\|\beta^{-1}\frac{d\xi}{dx}\right\|_2^2$$

$$= \frac{\beta^2}{N^2} \sum_{z,z' \in P} \int_0^1 [B_1(\{x-z\}) + B_1(z)][B_1(\{x-z'\}) + B_1(z')] \ dx$$

$$= \frac{\beta^2}{N^2} \sum_{z,z' \in P} \left[\frac{1}{2}B_2(\{z-z'\}) + B_1(z)B_1(z')\right]$$

$$= \beta^2 \left[\frac{1}{12N^2} + \frac{1}{N}\sum_{i=1}^{N}\left(z_{(i)} - \frac{2i-1}{2N}\right)^2\right],$$

where $z_{(1)} \le \cdots \le z_{(N)}$ are the ordered values of P, that is, the order statistics.

Quadrature error bound (1.14) is

$$(2.12) \qquad |I(f) - Q(f)| = |\langle \xi, f \rangle| = \beta^{-2}\left|\int_0^1 \frac{d\xi}{dx}\frac{df}{dx} \ dx\right| \le D_2(P)V_2(f),$$

where

$$D_2(P) \equiv \|\|\xi\|\|_2 = \left\|\beta^{-1}\frac{d\xi}{dx}\right\|_2, \qquad V_2(f) \equiv \|\|f - L(f)\|\|_2 = \left\|\beta^{-1}\frac{df}{dx}\right\|_2.$$

The quantity $D_2(P)$ is called the generalized $\mathcal{L}^2$-discrepancy since it is the $\mathcal{L}^2$-norm of $d\xi/dx$. In one dimension $D_2(P)$ is independent of the choice of μ, and β appears only as a constant multiple. However, this is not true for $s > 1$ as shall be seen in the next section. From (2.11), it follows that the minimum value of $D_2(P)$ is $\beta/(\sqrt{12}N)$, which occurs when $P = \{(2i-1)/2N : i = 1, \ldots, N\}$. This P corresponds to the midpoint rule, which has $O(N^{-2})$ error when the integrand is twice continuously differentiable [DR84, p. 53], but in this case has only $O(N^{-1})$ error under a weaker assumption on the integrand. The quantity $V_2(f)$ is defined as the generalized $\mathcal{L}^2$-variation of the f. The variation of a function does not change if an arbitrary constant is added to it. Likewise, adding a constant to a function leaves the absolute quadrature error unchanged.

To define the generalized $\mathcal{L}^p$-discrepancy and $\mathcal{L}^p$-variation it is necessary to define norms, $\|\| \ \|\|_p$, for X_p. For any p, $1 \le p \le \infty$, let

$$(2.13) \qquad\qquad \|\|f\|\|_p = \left\|\left(L(f), \beta^{-1}\frac{df}{dx}\right)\right\|_p.$$

Given these norms it is possible to define the $\mathcal{L}^p$-discrepancy and $\mathcal{L}^p$-variation as a natural extension of the $\mathcal{L}^2$-discrepancy and $\mathcal{L}^2$-variation because $d\xi/dx$ is piecewise continuous.

Definition 2.2. For any $f \in X_p$ the generalized $\mathcal{L}^p$-variation is

$$(2.14) \qquad\qquad V_p(f) \equiv \|\|f - L(f)\|\|_p = \left\|\beta^{-1}\frac{df}{dx}\right\|_p.$$

For any finite sample $P \subset [0,1)$ the generalized $\mathcal{L}^p$-discrepancy is defined as

$$(2.15) \qquad\qquad D_p(P) \equiv V_p(\xi) = \|\|\xi\|\|_p = \left\|\beta^{-1}\frac{d\xi}{dx}\right\|_p.$$

Hölder's inequality for the $\mathcal{L}^2$-inner product and $\mathcal{L}^p$-norms implies a Hölder's inequality for the inner product defined in (2.6) and the norms defined in (2.13), that is,

$$|\langle f, g \rangle| \leq \|\|f\|\|_p \, \|\|g\|\|_q \quad \forall f \in X_p, \ g \in X_q \ (1 \leq p, q \leq \infty, \ p^{-1} + q^{-1} = 1).$$

Norms, variations and discrepancies of different orders $p \leq q$ obey the following relationships:

$$\|\|f\|\|_p \leq \|\|f\|\|_q, \quad V_p(f) \leq V_q(f), \quad D_p(P) \leq D_q(P) \quad \forall f \in X_q, \ \forall P.$$

To derive quadrature error bounds in terms of the $\mathcal{L}^p$-discrepancy and $\mathcal{L}^q$-variation we note that for any fixed x the reproducing kernel $\eta(\bullet, x)$ has sufficient smoothness to lie not only in X_2 but also in X_∞. Thus, equation (2.7) holds for all $f \in X_1$. Likewise $\xi \in X_\infty$, so $I(f) - Q(f) = \langle \xi, f \rangle$ for all $f \in X_1$. Applying Hölder's inequality yields a quadrature error bound that includes (2.12) as a special case:

$$(2.16) \quad |I(f) - Q(f)| = |\langle \xi, f \rangle| = \beta^{-2} \left| \int_0^1 \frac{d\xi}{dx} \frac{df}{dx} \, dx \right|$$

$$\leq D_p(P) \, V_q(f) \quad \forall f \in X_q \ (1 \leq q \leq \infty, \ p^{-1} + q^{-1} = 1).$$

The worst-case integrand, $\xi_q \in X_q$, is that for which the error bound is attained. This occurs when

$$\left| \int_0^1 \frac{d\xi}{dx} \frac{d\xi_q}{dx} \, dx \right| = \left\| \frac{d\xi}{dx} \right\|_p \ \left\| \frac{d\xi_q}{dx} \right\|_q$$

in (2.16), that is,

$$(2.17) \qquad \xi_q(y) = \int_0^y \text{sign}\left(\frac{d\xi}{dx}\right) \left|\frac{d\xi}{dx}\right|^{1/(q-1)} \, dx \quad (1 < q \leq \infty).$$

Of course, adding an arbitrary constant to ξ_q or multiplying ξ_q by an arbitrary constant gives another worst-case integrand.

A worst-case integrand ξ_1 does not exist. For it to exist $d\xi_1/dx$ would need to be a Dirac delta function centered at the point where $|d\xi/dx|$ attains its maximum value, but such a ξ_1 does not have sufficient regularity to lie in X_1. However, for any $\epsilon > 0$ one can find a "nearly" worst-case integrand $\xi_{1,\epsilon} \in X_1$ such that

$$|I(\xi_{1,\epsilon}) - Q(\xi_{1,\epsilon})| > [D_\infty(P) - \epsilon]V_1(\xi_{1,\epsilon}).$$

This is done by constructing a $d\xi_{1,\epsilon}/dx$ that is only non-zero near the point where $|d\xi/dx|$ attains its maximum.

So far, a family of error bounds and figures of merit have been derived in the one-dimensional case assuming a mild amount of regularity on the integrand. The following theorem summarizes these results:

Theorem 2.3. *For any $f \in X_q$ $(1 \leq q \leq \infty)$ and for $s = 1$ quadrature rule (1.1) has an error*

$$|I(f) - Q(f)| \leq D_p(P)V_q(f) \quad (p^{-1} + q^{-1} = 1).$$

Equality holds for the worst-case integrand ξ_q given in (2.17) for $1 < q \leq \infty$.

3. Multidimensional case

This section generalizes the ideas of the previous one to $s > 1$, resulting in a multidimensional version of Theorem 2.3. The integrands are now functions of $x = (x_1, \ldots, x_s)$ and the notation introduced following (1.4) is employed.

The multidimensional generalization of X_p, the space of integrands defined in (2.2), is a space of functions whose mixed partial derivatives are all integrable:

$$(3.1) \qquad X_p \equiv X_p(C^s) \equiv \left\{ f : \frac{\partial^{|u|} f}{\partial x_u} \in \mathcal{L}^p(C^u) \; \forall u \subseteq S \right\}.$$

Note that $X_p(C^s) \subset X_p(C^t)$ for any $s < t$. In the previous section any $f \in X_p([0,1))$ was decomposed into two parts: a constant, $L(f)$, and a nonconstant part, $f - L(f)$. The variation of the integrand, $V_p(f)$, depended only on the second part. This decomposition is now generalized to 2^s components for $s \geq 1$. Let L_j denote the operator L defined in (2.5) acting on the jth coordinate, let $L_u = \prod_{j \in u} L_j$, and let $L_\emptyset$ be defined as the identity.

Definition 3.1. For any function $f \in X_p(C^s)$ iteratively define its components, f_u as follows:

$$(3.2) \qquad f_u = L_{S-u}(f) - \sum_{v \subset u} f_v \quad (u \subseteq S).$$

Lemma 3.2. *The components of $f \in X_p(C^s)$ defined in Definition 3.1 satisfy the following properties:*

$$(3.3a) \qquad L_j f_u = \begin{cases} 0 & \text{for } j \in u, \\ f_u & \text{for } j \notin u, \end{cases}$$

$$(3.3b) \qquad f_u \in X_{p*}(C^u) \equiv \{ f \in X_p(C^u) : L_j(f) = 0 \; \forall j \in u \},$$

and

$$(3.3c) \qquad f = \sum_{u \subseteq S} f_u,$$

Conversely, if $\hat{f}_u \in X_{p*}(C^u)$ are arbitrary functions, and $f = \sum_{u \subseteq S} \hat{f}_u$, then the components of f as defined in Definition 3.1 are, in fact, the $\hat{f}_u$.

Proof. Note from the definition of L in (2.5) that $L(1) = 1$, thus it follows that $L_j f_u = f_u$ for $j \notin u$. The proof of $L_j f_u = 0$ for $j \in u$ proceeds by induction. Note that it holds vacuously for $u = \emptyset$. Now for any $u \subseteq S$ suppose that $L_j f_v = 0$ if $j \in v$ for all $v \subset u$. For any particular $j \in u$ let $w = u - \{j\} \subset u$. It follows that

$$L_j f_u = L_j \left[L_{S-u}(f) - \sum_{v \subset u} f_v \right] = L_{S-w}(f) - \sum_{v \subset u} L_j(f_v)$$

$$= L_{S-w}(f) - \sum_{v \subseteq w} f_v = 0.$$

Equation (3.3b) follows from (3.3a), and equation (3.3c) follows from the definition (3.2) for the case $u = S$.

Induction is also used to to prove the second part of the lemma. Since $L_j \hat{f}_u = 0$ for all $j \in u$, it follows that $f_\emptyset = L_S(f) = \hat{f}_\emptyset$. Next for any given u it is assumed

that $f_v = \hat{f}_v$ for all $v \subset u$. Applying the definition of f_u and (3.3a) completes the proof:

$$f_u = L_{S-u}\left[\sum_{w \subseteq S} \hat{f}_w\right] - \sum_{v \subset u} f_v = \sum_{w \subseteq u} \hat{f}_w - \sum_{v \subset u} f_v = \hat{f}_u. \quad \square$$

The component f_u is the part of the function that depends only on x_u. It belongs to $X_{p*}(C^u)$, a subspace of $X_p(C^u)$ whose elements satisfy the condition $L_j(f) = 0 \ \forall j \in u$. The order of a component is $|u|$. The constant term $f_\emptyset$ is integrated exactly by quadrature rule (1.1), however, the other components are subject to quadrature error. If $\mu = 0$, then the f_u are ANOVA effects, used in earlier analysis of quadrature error by Owen [Owe92, Owe95] and the author [Hic95, Hic96]. They showed how certain quadrature rules integrate components of a certain order better than others.

The inner product $\langle, \rangle$ on $X_2(C^s)$ and the norms $\| \ \|_p$ on $X_p(C^s)$ are defined as generalizations of (2.6) and (2.13) in terms of the decomposition in Lemma 3.2:

(3.4)

$$\langle f, g \rangle = \sum_{u \subseteq S} \beta^{-2|u|} \int_{C^u} \frac{\partial^{|u|} f_u}{\partial x_u} \frac{\partial^{|u|} g_u}{\partial x_u} \, dx_u, \quad \| f \|_p = \left\| \left(\beta^{-|u|} \frac{\partial^{|u|} f_u}{\partial x_u} \right)_{u \subseteq S} \right\|_p.$$

Note from Definition 3.1 that

(3.5)
$$\frac{\partial^{|u|} f_u}{\partial x_u} = L_{S-u}\left(\frac{\partial^{|u|} f}{\partial x_u}\right).$$

This fact is used in Section 5 in deriving formulas for specific cases of the generalized variation and discrepancy. As in the one-dimensional case we have a Hölder's inequality,

$$|\langle f, g \rangle| \leq \| f \|_p \| g \|_q \quad \forall f \in X_p, \ g \in X_q \ (1 \leq p, q \leq \infty, \ p^{-1} + q^{-1} = 1).$$

Under this inner product the reproducing kernel for $X_2(C^s)$ is simply the product of one-dimensional reproducing kernels (2.3).

Lemma 3.3. *Define*

$$\eta(x, y) = \prod_{j=1}^{s} \eta_1(x_j, y_j),$$

where

$$\eta_1(x_1, y_1) = M + \beta^2\left[\mu(x_1) + \mu(y_1) + \frac{1}{2}B_2(\{x_1 - y_1\}) + B_1(x_1)B_1(y_1)\right].$$

It follows that the components of η are

$$\eta_u(x, y) = \eta_u(x_u, y_u) = \prod_{j \in u}[\eta_1(x_j, y_j) - 1],$$

and that

(3.6a) $$f_u(y_u) = \beta^{-2|u|} \int_{C^u} \frac{\partial^{|u|} f_u}{\partial x_u} \frac{\partial^{|u|} \eta_u(\bullet, y_u)}{\partial x_u} \, dx_u \quad \forall y_u \in C^u,$$

(3.6b) $$f(y) = \langle f, \eta(\bullet, y) \rangle \quad \forall y \in C^s.$$

That is, η is the reproducing kernel for $X_p(C^s)$.

Proof. Recall from (2.8) that $L(\eta_1(\bullet, y_1)) = 1$. Using induction one can derive the components η_u.

Induction is also used to prove (3.6a). This is true for $u = \emptyset$. Now suppose it is true for all $v \subset u$. For any $j \in u$ let $v = u - j$. For any fixed x_j the function $\partial f_u / \partial x_j$ is in $X_p(C^v)$. Thus,

$$\beta^{-2|u|} \int_{C^u} \frac{\partial^{|u|} f_u}{\partial x_u} \frac{\partial^{|u|} \eta_u(\bullet, y_u)}{\partial x_u} \, dx_u$$

$$= \beta^{-2} \int_0^1 \left[\beta^{-2|v|} \int_{C^v} \frac{\partial^{|v|}}{\partial x_v} \left(\frac{\partial f_u}{\partial x_j} \right) \frac{\partial^{|v|} \eta_u(\bullet, y_v)}{\partial x_v} \, dx_v \right] \frac{\partial \eta_j(\bullet, y_j)}{\partial x_j} \, dx_j$$

$$= \beta^{-2} \int_0^1 \left. \frac{\partial f_u}{\partial x_j} \right|_{x_v = y_v} \frac{\partial \eta_j(\bullet, y_j)}{\partial x_j} \, dx_j = f(y_u).$$

Equation (3.6b) follows from (3.4) and (3.6a). $\square$

With the reproducing kernel we can now compute the worst-case integrand for $X_2(C^s)$ as was done in (2.10). The components of the worst-case integrand can be calculated as well.

$$\xi(x) = (I - Q)(\eta(\bullet, x)) = \prod_{j=1}^s [M + \beta^2 \mu(x_j)]$$

$$- \frac{1}{N} \sum_{z \in P} \prod_{j=1}^s \left\{ M + \beta^2 \left[\mu(x_j) + \mu(z_j) + \frac{1}{2} B_2(\{x_j - z_j\}) + B_1(x_j) B_1(z_j) \right] \right\},$$

$$\xi_u(x_u) = \prod_{j \in u} [M - 1 + \beta^2 \mu(x_j)]$$

$$- \frac{1}{N} \sum_{z \in P} \prod_{j \in u} \left\{ M - 1 + \beta^2 \left[\mu(x_j) + \mu(z_j) \right. \right.$$

$$\left. \left. + \frac{1}{2} B_2(\{x_j - z_j\}) + B_1(x_j) B_1(z_j) \right] \right\},$$

$$\frac{\partial^{|u|} \xi_u}{\partial x_u} = \beta^{2|u|} \left\{ \prod_{j \in u} \mu'(x_j) - \frac{1}{N} \sum_{z \in P} \prod_{j \in u} [\mu'(x_j) + B_1(\{x_j - z_j\}) + B_1(z_j)] \right\}$$

$$= \beta^{2|u|} \left\{ \prod_{j \in u} \mu'(x_j) - \frac{1}{N} \sum_{z \in P} \prod_{j \in u} [\mu'(x_j) + x_j - 1_{x_j > z_j}] \right\},$$

where μ' is the derivative of μ, and $1_{x_j > z_j}$ is the indicator function. The worst-case integrand, ξ, is used to generalize Definition 2.2 for the variation and discrepancy.

Definition 3.4. For any $f \in X_p(C^s)$ let

(3.7a) $$V_{p,u}(f) \equiv \left\| \frac{\partial^{|u|} f_u}{\partial x_u} \right\|_p.$$

The generalized $\mathcal{L}^p$-variation is

(3.7b) $\quad V_p(f) \equiv \||f - L_S(f)\||_p = \left\| \left(\beta^{-|u|} V_{p,u}(f) \right)_{u \neq \emptyset} \right\|_p = \left\| \left(\beta^{-|u|} \dfrac{\partial^{|u|} f_u}{\partial x_u} \right)_{u \neq \emptyset} \right\|_p .$

For any finite set $P \subset C^s$ let

(3.8a) $\qquad D_{p,u}(P) \equiv V_{p,u}(\beta^{-2|u|} \xi_u) = \left\| \beta^{-2|u|} \dfrac{\partial^{|u|} \xi_u}{\partial x_u} \right\|_p$

$$= \left\| \prod_{j \in u} \mu'(x_j) - \frac{1}{N} \sum_{z \in P} \prod_{j \in u} [\mu'(x_j) + x_j - 1_{x_j > z_j}] \right\|_p$$

The generalized $\mathcal{L}^p$-discrepancy is defined as

(3.8b) $\quad D_p(P) \equiv V_p(\xi) = \||\xi\||_p$

$$= \left\| \left(\beta^{|u|} D_{p,u}(P) \right)_{u \neq \emptyset} \right\|_p = \left\| \left(\beta^{-|u|} \frac{\partial^{|u|} \xi_u}{\partial x_u} \right)_{u \neq \emptyset} \right\|_p .$$

$$= \left\| \left(\beta^{|u|} \left\{ \prod_{j \in u} \mu'(x_j) - \frac{1}{N} \sum_{z \in P} \prod_{j \in u} [\mu'(x_j) + x_j - 1_{x_j > z_j}] \right\} \right)_{u \neq \emptyset} \right\|_p$$

In contrast to the one-dimensional case the generalized $\mathcal{L}^p$-discrepancy depends intrinsically on the choices of β and μ for $s > 1$. The consequences of these choices are discussed below and in Section 5. Whereas the integrals defining the generalized $\mathcal{L}^p$-discrepancy are generally intractable, the case $p = 2$ can be reduced to a double sum:

(3.9a) $\quad [D_{2,u}(P)]^2 \equiv \beta^{-2|u|} \||\xi\||_2^2 = \beta^{-2|u|} \langle \xi_u, \xi_u \rangle = \beta^{-2|u|} (I - Q)(\xi_u)$

$$= \bar{M}^{|u|} - \frac{2}{N} \sum_{z \in P} \prod_{j \in u} [\bar{M} + \mu(z_j)]$$

$$+ \frac{1}{N^2} \sum_{z,z' \in P} \prod_{j \in u} \left[\bar{M} + \mu(z_j) + \mu(z'_j) + \frac{1}{2} B_2(\{z_j - z'_j\}) + B_1(z_j) B_1(z'_j) \right],$$

(3.9b) $\qquad [D_2(P)]^2 \equiv \||\xi\||_2^2 = \langle \xi, \xi \rangle = (I - Q)(\xi)$

$$= M^s - \frac{2}{N} \sum_{z \in P} \prod_{j=1}^{s} [M + \beta^2 \mu(z_j)]$$

$$+ \frac{1}{N^2} \sum_{z,z' \in P} \prod_{j=1}^{s} \Bigg\{ M$$

$$+ \beta^2 \left[\mu(z_j) + \mu(z'_j) + \frac{1}{2} B_2(\{z_j - z'_j\}) + B_1(z_j) B_1(z'_j) \right] \Bigg\},$$

where

$$\bar{M} = (M - 1)/\beta^2 = \int_0^1 \left(\frac{d\mu}{dx} \right)^2 dx.$$

Error bounds for multidimensional quadrature are analogous to those for univariate quadrature and involve the generalized discrepancy and variation. Note that $\xi_\emptyset = 0$. For any $f \in X_q(C^s)$

$$(3.10) \quad |I(f) - Q(f)| = |\langle \xi, f \rangle| = \left| \sum_{u \neq \emptyset} \beta^{-2|u|} \int_{C^u} \frac{\partial^{|u|} \xi_u}{\partial x_u} \frac{\partial^{|u|} f_u}{\partial x_u} \, dx_u \right|$$

$$\leq \sum_{u \neq \emptyset} D_{p,u}(P) V_{q,u}(f) \leq D_p(P) V_q(f) \quad (p^{-1} + q^{-1} = 1),$$

by Hölder's inequality. There is no contribution from the constant $f_\emptyset$ in (3.10) because quadrature rule (1.1) integrates constants exactly. Recall that the integrand f is a sum of its components f_u. The term

$$\beta^{-2|u|} \int_{C^u} \frac{\partial^{|u|} \xi_u}{\partial x_u} \frac{\partial^{|u|} f_u}{\partial x_u} \, dx_u,$$

that appears above, is the quadrature error for f_u. Furthermore, $|I(f_u) - Q(f_u)| \leq D_{p,u}(P) V_{q,u}(f)$. The error bound

$$\sum_{u \neq \emptyset} D_{p,u}(P) V_{q,u}(f)$$

in (3.10) is independent of β since $D_{p,u}(P)$ and $V_{p,u}(f)$ are themselves independent of β. The quantity $D_{p,u}(P)$ measures how accurately the quadrature rule integrates functions in $X_{p*}(C^u)$, that is, functions depending on the coordinates whose indices are in u only. On the other hand, both $D_p(P)$ and $V_q(f)$ do depend on β (see (3.7) and (3.8)). They are weighted averages of the $D_{p,u}(P)$ and $V_{p,u}(f)$, respectively, where the weight is a power of β. Choosing a larger value of β gives a higher weight to the $D_{p,u}(P)$ with larger $|u|$ and so implies a preference for quadrature rules that accurately integrate the higher order components of the integrand. Conversely, choosing a smaller value of β implies a preference for quadrature rules that accurately integrate the lower order components of the integrand. By referring to the definition of the variation in (3.7) one may see that β^{-1} plays the role of a length scale.

If one desires figures of merit for P that are independent of β one can compute the $D_{p,u}(P)$ for all $u \neq \emptyset$. However, the number of these $D_{p,u}(P)$ is $2^s - 1$, which grows exponentially as s increases. An alternative is to compute

$$D_{p,r}(P) \equiv \left\| (D_{p,u}(P))_{|u|=r} \right\|_p \quad (1 \leq r \leq s).$$

The quantity $D_{p,r}(P)$, which is also independent of β, measures how well the quadrature rule integrates r-dimensional components. Also, note that

$$D_p(P) = \left\| (\beta^r D_{p,r}(P))_{1 \leq r \leq s} \right\|_p.$$

Hickernell [Hic95] compared the relative merits of several different quasi-random and random samples using quantities similar to $D_{2,r}(P)$. It was found that some s-dimensional samples are good for integrating the lower-dimensional effects, while others were better for integrating the higher-dimensional effects. For example, a rectangular grid has relatively small $D_{2,s}(P)$, but a relatively large $D_{2,1}(P)$, compared to other samples.

Using Lemma 3.2 it is possible to construct the worst-case integrand, ξ_q. Each of its components, $\xi_{q,u}$, is defined according to the conditions:

(3.11a) $$\frac{\partial^{|u|}\xi_{q,u}}{\partial x_u} = \beta^{(q-2)|u|/(q-1)} \operatorname{sign}\left(\frac{\partial^{|u|}\xi_u}{\partial x_u}\right)\left|\frac{\partial^{|u|}\xi_u}{\partial x_u}\right|^{1/(q-1)}$$

(3.11b) $$L_v(\xi_{q,u}) = 0 \quad (\emptyset \subset v \subseteq u).$$

These conditions insure that the upper bound in each line of (3.10) is attained.

Theorem 3.5. *For any $f \in X_q(C^s)$ $(1 \le q \le \infty)$ quadrature rule (1.1) has an error*

$$|I(f) - Q(f)| = |\langle \xi, f \rangle| \le D_p(P)V_q(f) \quad (p^{-1} + q^{-1} = 1).$$

Equality holds for the worst-case integrand ξ_q given by (3.11) for $1 < q \le \infty$.

4. PERIODIC CASE

The error bounds in the previous two sections make no assumptions about the periodicity of the integrand, as do error bounds (1.10) and (1.12b). To consider periodic integrands define

$$\tilde{X}_{p,t}(C^s) \equiv \left\{ f : \frac{\partial^{k|u|}f}{\partial x_u^k} \in \mathcal{L}^p(C^s) \text{ and } \int_0^1 \frac{\partial^{k|u|}f}{\partial x_u^k}\, dx_j = 0 \right.$$

$$\left. \forall k \le t, \ \forall j \in u, \ \forall u \subseteq S \right\}.$$

The integer t denotes the degree of periodicity. Note that $\tilde{X}_{p,1}(C^s)$ is a subset of $X_p(C^s)$.

The inner products and norms for $X_p(C^s)$ are now modified to apply to $\tilde{X}_{p,t}(C^s)$. First we modify the definition of the operator L previously given in (2.5). For any $\mu \in \tilde{X}_{\infty,t}([0,1))$ with $\int_0^1 \mu\, dx = 0$ let

(4.1) $$L(f) \equiv \int_0^1 \left(f - \frac{d^t\mu}{dx^t}\frac{d^t f}{dx^t} \right) dx, \qquad M = 1 + \beta^{2t}\int_0^1 \left(\frac{d^t\mu}{dx^t}\right)^2 dx.$$

The components of a function in $\tilde{X}_{p,t}(C^s)$ are defined as in Definition 3.1, but using the new definition of L. The analogous result to Lemma 3.2 for functions in $\tilde{X}_{p,t}(C^s)$ holds. The inner product on $\tilde{X}_{2,t}(C^s)$ and the norms on $\tilde{X}_{p,t}(C^s)$ are defined as follows:

$$\langle f, g \rangle_t = \sum_{u \subseteq S} \beta^{-2t|u|} \int_{C^u} \frac{\partial^{t|u|}f_u}{\partial x_u^t}\frac{\partial^{t|u|}g_u}{\partial x_u^t}\, dx_u,$$

$$\|f\|_{p,t} = \left\| \left(\beta^{-t|u|}\frac{\partial^{t|u|}f_u}{\partial x_u^t} \right)_{u \subseteq S} \right\|_p.$$

Here and below we suppress the dependence of some quantities on t for ease of notation.

The reproducing kernel for $\tilde{X}_{2,t}(C^s)$ is denoted $\tilde{\eta}$ and takes the form:

$$\tilde{\eta}(x,y) = \prod_{j=1}^s \tilde{\eta}_1(x_j, y_j),$$

where

$$\tilde{\eta}_1(x_1, y_1) = M + \beta^{2t} \left[\mu(x_1) + \mu(y_1) - \frac{(-1)^t}{(2t)!} B_{2t}(\{x_1 - y_1\}) \right].$$

Because $\tilde{\eta}_1(x_1, y_1)$ must be periodic it cannot contain the term $B_1(x_1)B_1(y_1)$, which appears in the definition of η_1. Since $\tilde{X}_{p,t}(C^s)$ contains only periodic functions, it follows by (2.1) that

$$L(\tilde{\eta}_1(\bullet, y_1)) = 1 \quad \forall y_1, \qquad f(x) = \langle \tilde{\eta}(\bullet, x), f \rangle \quad \forall f \in \tilde{X}_{1,t}(C^s).$$

The worst-case integrand in $\tilde{X}_{2,t}(C^s)$ is

$$(4.2a) \quad \tilde{\xi}(x) = (I - Q)(\tilde{\eta}(\bullet, x)) = \prod_{j=1}^{s} [M + \beta^{2t}\mu(x_j)]$$

$$- \frac{1}{N} \sum_{z \in P} \prod_{j=1}^{s} \left\{ M + \beta^{2t} \left[\mu(x_j) + \mu(z_j) - \frac{(-1)^t}{(2t)!} B_{2t}(\{x_j - z_j\}) \right] \right\},$$

$$(4.2b) \quad \frac{\partial^{t|u|} \tilde{\xi}_u}{\partial x_u^t} = \beta^{2t|u|} \left\{ \prod_{j \in u} \mu^{(t)}(x_j) \right.$$

$$\left. \cdot - \frac{1}{N} \sum_{z \in P} \prod_{j \in u} \left[\mu^{(t)}(x_j) - \frac{(-1)^t}{t!} B_t(\{x_j - z_j\}) \right] \right\},$$

where $\mu^{(t)}$ is the tth derivative of μ. Given $\tilde{\xi}$ it is now possible to generalize Definition 3.4.

Definition 4.1. For any $f \in \tilde{X}_{p,t}(C^s)$ the generalized *periodic* $\mathcal{L}^{p,t}$-variation is

$$\tilde{V}_{p,t}(f) \equiv \| f - L_S(f) \|_p = \left\| \left(\beta^{-t|u|} \frac{\partial^{t|u|} f_u}{\partial x_u^t} \right)_{u \neq \emptyset} \right\|_p.$$

For any finite set $P \subset C^s$ the generalized *periodic* $\mathcal{L}^{p,t}$-discrepancy is defined as

$$(4.3a) \quad \tilde{D}_{p,t}(P) \equiv V_{p,t}(\tilde{\xi}) = \left\| \| \tilde{\xi} \| \right\|_{p,t}$$

$$= \left\| \left(\beta^{t|u|} \left\{ \prod_{j \in u} \mu^{(t)}(x_j) \right. \right. \right.$$

$$\left. \left. \left. - \frac{1}{N} \sum_{z \in P} \prod_{j \in u} \left[\mu^{(t)}(x_j) - \frac{(-1)^t}{t!} B_t(\{x_j - z_j\}) \right] \right\} \right)_{u \neq \emptyset} \right\|_p,$$

$$(4.3b) \quad [\tilde{D}_{2,t}(P)]^2 = M^s - \frac{2}{N} \sum_{z \in P} \prod_{j=1}^{s} [M + \beta^{2t}\mu(z_j)]$$

$$+ \frac{1}{N^2} \sum_{z,z' \in P} \prod_{j=1}^{s} \left\{ M + \beta^{2t} \left[\mu(z_j) + \mu(z_j') - \frac{(-1)^t}{(2t)!} B_{2t}(\{z_j - z_j'\}) \right] \right\}.$$

For $s = 1$ the formula for $\tilde{D}_{2,1}(P)$ can be simplified even further:

$$[\tilde{D}_{2,1}(P)]^2 = \beta^2 \left\{ \frac{1}{12N^2} + \frac{1}{N} \sum_{i=1}^{N} \left[z_{(i)} - \frac{i}{N} - \frac{1}{N} \sum_{j=1}^{N} \left(z_{(j)} - \frac{j}{N} \right) \right]^2 \right\},$$

where the $z_{(i)}$ are again the ordered values of P. From (2.11) it follows that $\tilde{D}_{2,1}(P)$ and $D_2(P)$ have the same minimum value. For $\tilde{D}_{2,1}(P)$ this is obtained for $P = \{(i - \nu)/N : i = 1, \ldots, N\}$ for all $0 < \nu \le 1$, whereas for $D_2(P)$ it is obtained only for $\nu = 1/2$. In fact, shifting any set of points leaves its periodic $\mathcal{L}^p$-discrepancy unchanged, that is, for any $\nu \in C^s$

$$\tilde{D}_{p,t}(\{\{z + \nu\} : z \in P\}) = \tilde{D}_{p,t}(P).$$

Quadrature for periodic functions has an error bound in terms of the periodic discrepancy and variation. The following theorem is a straightforward generalization of Theorem 3.5.

Theorem 4.2. *For any $f \in \tilde{X}_{q,t}(C^s)$ ($1 \le q \le \infty$) quadrature rule (1.1) has an error*

$$|I(f) - Q(f)| = |\langle \tilde{\xi}, f \rangle| \le \tilde{D}_{p,t}(P)\tilde{V}_{q,t}(f) \quad (p^{-1} + q^{-1} = 1).$$

Equality holds for the worst-case integrand

$$\frac{\partial^{t|u|}\tilde{\xi}_{q,u}}{\partial x_u^t} = \beta^{(q-2)|u|/(q-1)} \operatorname{sign}\left(\frac{\partial^{t|u|}\tilde{\xi}_u}{\partial x_u^t} \right) \left| \frac{\partial^{t|u|}\tilde{\xi}_u}{\partial x_u^t} \right|^{1/(q-1)}$$

$$L_v(\tilde{\xi}_{q,u}) = 0 \quad (\emptyset \subset v \subseteq u),$$

for $1 < q \le \infty$.

The space of periodic integrands $\tilde{X}_{p,1}(C^s)$ is a subspace of $X_p(C^s)$ since they share the same inner product (assuming the same β and same $\mu \in \tilde{X}_{\infty,1}([0,1))$). For any $f \in \tilde{X}_{p,1}(C^s)$ it follows that $\tilde{V}_{p,1}(f) = V_p(f)$, and for any sample P it follows that $\tilde{D}_{p,1}(P) \le D_p(P)$. The latter assertion follows by applying both Theorems 3.5 and 4.2 to the worst-case integrand $\tilde{\xi}_q$:

$$\tilde{D}_{p,1}(P)\tilde{V}_{q,1}(\tilde{\xi}_q) = |(I - Q)(\tilde{\xi}_q)| \le D_p(P)V_q(\tilde{\xi}_q) = D_p(P)\tilde{V}_{q,1}(\tilde{\xi}_q),$$

for $1 \le p < \infty$. This can be extended to the case $p = \infty$ by noting that for a fixed P the discrepancy is a continuous function of p and taking the limit as $p \to \infty$.

5. Choices of β and μ

In the previous section it was seen that the error bounds depend on the choices of β and μ. In this section we show how different choices allow us to recover the error bounds and figures of merit in (1.4) and (1.12). It will also be shown that in some sense different values of β and μ give equivalent error bounds.

Recall from (3.8) that the generalized $\mathcal{L}^\infty$-discrepancy is

$$D_\infty(P) = \left\| \left(\beta^{-|u|}\frac{\partial^{|u|}\xi_u}{\partial x_u} \right)_{u \ne \emptyset} \right\|_\infty.$$

For certain choices of β this formula for the generalized $\mathcal{L}^\infty$-discrepancy can be simplified in that one need· not consider all u, but only $u = S$. Taking the limit as x_{S-u} tends to the origin gives

$$\beta^{-s} \left| \frac{\partial^s \xi_S}{\partial x_S} \right|_{x_{S-u} \downarrow (0,\ldots,0)}$$

$$= \beta^s |\mu'(0)|^{s-|u|} \left| \prod_{j \in u} \mu'(x_j) - \frac{1}{N} \sum_{z \in P} \prod_{j \in u} [\mu'(x_j) + x_j - 1_{x_j > z_j}] \right|$$

$$= |\beta \mu'(0)|^{s-|u|} \beta^{-|u|} \left| \frac{\partial^{|u|} \xi_u}{\partial x_u} \right|,$$

assuming that μ' has limiting values as x tends to 0. Thus, the essential supremum of $|\beta^{-s} \partial^s \xi_S / \partial x_S|$ will be no smaller than that of $|\beta^{-|u|} \partial^{|u|} \xi_u / \partial x_u|$ if $|\beta \mu'(0)| = 1$. Likewise,

$$\beta^{-s} \left| \frac{\partial^s \xi_S}{\partial x_S} \right|_{x_{S-u} \uparrow (1,\ldots,1)} = |\beta \mu'(1)|^{s-|u|} \beta^{-|u|} \left| \frac{\partial^{|u|} \xi_u}{\partial x_u} \right|.$$

Similar results hold for limits of $\partial^{ts} \tilde{\xi}_S / \partial x_S^t$. This leads to the following theorem.

Theorem 5.1. *If either* $\lim_{x \downarrow 0} \beta |\mu'(x)|$ *or* $\lim_{x \uparrow 1} \beta |\mu'(x)|$ *is 1, then*

$$D_\infty(P) = \left\| \prod_{j=1}^s \mu'(x_j) - \frac{1}{N} \sum_{z \in P} \prod_{j=1}^s [\mu'(x_j) + x_j - 1_{x_j > z_j}] \right\|_\infty.$$

For any positive integer t *if* $\lim_{x \downarrow 0} \beta^t |\mu^{(t)}(x)|$ *or* $\lim_{x \uparrow 1} \beta^t |\mu^{(t)}(x)|$ *is 1, then*

$$\tilde{D}_{\infty,t}(P) = \left\| \prod_{j=1}^s \mu^{(t)}(x_j) - \frac{1}{N} \sum_{z \in P} \prod_{j=1}^s \left[\mu^{(t)}(x_j) - \frac{(-1)^t}{t!} B_t(\{x_j - z_j\}) \right] \right\|_\infty.$$

In the following subsections we consider four specific choices of μ and β and give the corresponding formulas for the generalized discrepancy and variation. These examples yield the star discrepancy (1.6) and P_α (1.8) as well as two new discrepancies.

5.1. Star discrepancy. To recover the $\mathcal{L}^p$-star discrepancy defined in (1.6) choose

$$\mu(x) = \frac{1}{6} - \frac{x^2}{2}, \qquad \mu'(x) = -x, \qquad \beta^{-1} = |\mu'(1)| = 1, \qquad M = \frac{4}{3}.$$

This implies that $L_u(f) = f|_{x_u = (1,\ldots,1)}$, and that the variation is defined as follows:

$$V_p(f) = \left\| \left(\frac{\partial^{|u|} f}{\partial x_u} \bigg|_{x_{S-u} = (1,\ldots,1)} \right)_{u \neq \emptyset} \right\|_p.$$

The formulas for the discrepancy from (3.8) and (3.9) are

$$\frac{\partial^{|u|} \xi_u}{\partial x_u} = \prod_{j \in u} -x_j - \frac{1}{N} \sum_{z \in P} \prod_{j \in u} (-1_{x_j > z_j}) = (-1)^{|u|} \left\{ \text{Vol}([0, x_u)) - \frac{|P_u \cap [0, x_u)|}{N} \right\},$$

(5.1a)
$$D_p(P) = \left\| \left(\frac{|P_u \cap [0, x_u)|}{N} - \mathrm{Vol}([0, x_u)) \right)_{u \neq \emptyset} \right\|_p,$$

(5.1b)
$$D_\infty(P) = \left\| \frac{|P \cap [0, x)|}{N} - \mathrm{Vol}([0, x)) \right\|_\infty,$$

(5.1c)

$$[D_2(P)]^2 = \left(\frac{4}{3} \right)^s - \frac{2}{N} \sum_{z \in P} \prod_{j=1}^s \left(\frac{3 - z_j^2}{2} \right) + \frac{1}{N^2} \sum_{z, z' \in P} \prod_{j=1}^s [2 - \max(z_j, z_j')],$$

where we have applied Theorem 5.1. The error bound in Theorem 3.5 for this choice of μ and β is the generalization of the Koksma-Hlawka inequality (1.5).

As remarked in the introduction the $\mathcal{L}^p$-star discrepancy is often defined as

(5.2)
$$D_{p,S}(P) = \left\| \frac{|P \cap [0, x)|}{N} - \mathrm{Vol}([0, x)) \right\|_p,$$

and this figure of merit has been used to compare different quasirandom samples (e.g. [MC94]). It is true that for the choice of μ and β above $D_\infty(P) = D_{\infty,S}(P)$ because of Theorem 5.1. However, for all other values of p

$$D_p(P) = \left\{ \sum_{u \neq \emptyset} [D_{p,u}(P)]^p \right\}^{1/p} > D_{p,S}(P).$$

Specifically, we may compare the above formula for $D_2(P)$ with the following formula given by [War72]:

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

The terms inside each of the three products in the formula for $[D_{2,S}(P)]^2$ are one less than the corresponding terms in the formula for $[D_2(P)]^2$. Because $D_2(P)$ appears in a quadrature error (Theorem 3.5), whereas $D_{2,S}(P)$ does not, we prefer the former as a figure of merit. Moreover, for a uniform simple random sample, P,

$$E\{[D_2(P)]^2\} = \left[\left(\frac{3}{2} \right)^s - \left(\frac{4}{3} \right)^s \right] \Big/ N = \left(\frac{4}{3} \right)^s \left[\left(\frac{9}{8} \right)^s - 1 \right] \Big/ N,$$

$$E\{[D_{2,S}(P)]^2\} = \left[\left(\frac{1}{2} \right)^s - \left(\frac{1}{3} \right)^s \right] \Big/ N,$$

(see [MC94]). As a figure of merit $E\{[D_{2,S}(P)]^2\}$ becomes smaller for a random sample simply by increasing the dimension. This is undesirable.

5.2. **A centered discrepancy.** The discrepancy in the previous section is anchored to the origin because the interval $[0, x)$ appears in its definition. Also, there is reference to $(1, \ldots, 1)$, since it appears in the formula for the variation. Now we define a discrepancy and variation that refer to the center of the cube, $(1/2, \ldots, 1/2)$. This discrepancy and variation are invariant under reflections of P about any plane $x_j = 1/2$.

It is useful to introduce some notation describing the vertices of the cube. Let A^s denote the set of 2^s vertices of the cube C^s, that is,

$$A^s \equiv \{a \in \mathbf{R}^s : a_j = 0, 1 \; \forall j\}.$$

The planes passing through any point x parallel to the faces of the cube C^s can be thought of as dividing it into 2^s intervals, each containing the points between x and a vertex a:

$$J(a, x) = \{y \in C^s : \min(a_j, x_j) \leq y_j < \max(a_j, x_j) \; \forall j\}.$$

The set of $2^{|u|}$ vertices of the cube C^u is denoted A^u, which is the projection of A^s into $\mathbf{R}^u$. The projections of the intervals $J(a, x)$ into $\mathbf{R}^u$ are denoted $J(a_u, x_u)$. Also, for any vertex $a_u \in A^u$, let

$$\sigma(a_u) \equiv \sum_{j \in u} a_j \pmod{2}.$$

In this way one can differentiate even $(\sigma(a_u) = 0)$ and odd $(\sigma(a_u) = 1)$ vertices.

The centered discrepancy comes from the following choices of μ and β:

$$\mu(x) = -\frac{1}{2} B_2(\{x - 1/2\}), \qquad \mu'(x) = -x + 1_{x > 1/2}, \qquad \beta^{-1} = 1, \qquad M = \frac{13}{12},$$

which implies that $L_u(f) = f|_{x_u = (1/2, \ldots, 1/2)}$, and

$$V_p(f) = \left\| \left(\frac{\partial^{|u|} f}{\partial x_u} \bigg|_{x_{S-u} = (1/2, \ldots, 1/2)} \right)_{u \neq \emptyset} \right\|_p.$$

Let $a(x)$ denote the unique vertex of C^s which is closest to x, that is, $a(x)$ is the unique vertex such that $x \in J(a(x), (1/2, \ldots, 1/2))$. Then

$$\frac{\partial^{|u|} \xi_u}{\partial x_u} = \prod_{j \in u} (-x_j + 1_{x_j > 1/2}) - \frac{1}{N} \sum_{z \in P} \prod_{j \in u} (1_{x_j > 1/2} - 1_{x_j > z_j})$$

$$= (-1)^{|u| + \sigma(a(x)_u)} \left\{ \text{Vol}(J(a_u(x), x_u)) - \frac{|P_u \cap J(a_u(x), x_u)|}{N} \right\}.$$

The formula for the discrepancy can be written as

$$D_p(P) = \left\| \left(\frac{|P_u \cap J(a_u(x), x_u)|}{N} - \text{Vol}(J(a_u(x), x_u)) \right)_{u \neq \emptyset} \right\|_p,$$

$$[D_2(P)]^2 = \left(\frac{13}{12} \right)^s - \frac{2}{N} \sum_{z \in P} \prod_{j=1}^s \left(1 + \frac{1}{2}|z_j - 1/2| - \frac{1}{2}|z_j - 1/2|^2 \right)$$

$$+ \frac{1}{N^2} \sum_{z, z' \in P} \prod_{j=1}^s \left[1 + \frac{1}{2}|z_j - 1/2| + \frac{1}{2}|z_j' - 1/2| - \frac{1}{2}|z_j - z_j'| \right].$$

This formula is similar to that for the star discrepancy (5.1), except the origin is replaced by $a(x)$, the vertex of the cube closest to the point x. In contrast to the star discrepancy, replacing x_j by $1 - x_j$ for any j leaves the centered discrepancy and variation unchanged. Unfortunately, Theorem 5.1 cannot be invoked to simplify the formula for $D_\infty(P)$ because $\mu'(0) = \mu'(1) = 0$. The next discrepancy has both the desired invariance and a simplified formula for $D_\infty(P)$.

5.3. **A symmetric discrepancy.** In the previous two cases $L_u f$ and the variation referred to the function value at a single point. However, below we choose μ so that $L_u f$ is an average of the function values over the vertices of the cube. Let

$$\mu(x) = -\frac{1}{2} B_2(x), \qquad \mu'(x) = -x + 1/2,$$

$$\beta^{-1} = |\mu'(0)| = |\mu'(1)| = 1/2, \qquad M = \frac{4}{3}.$$

The notation for the vertices introduced above is used. Furthermore, let $J_e(x_u)$ and $J_o(x_u)$ be the unions of the even and odd intervals, respectively, that is

$$J_e(x_u) = \bigcup_{\sigma(a_u)=0} J(a_u, x_u), \qquad J_o(x_u) = \bigcup_{\sigma(a_u)=1} J(a_u, x_u).$$

One may show that

$$\mathrm{Vol}\,(J_e(x_u)) = \frac{1}{2} + \frac{1}{2} \prod_{j \in u}(2x_j - 1), \qquad \mathrm{Vol}\,(J_o(x_u)) = \frac{1}{2} - \frac{1}{2} \prod_{j \in u}(2x_j - 1),$$

which will be used below.

The linear operator L_u is defined as an average of the function values over the vertices in A^u:

$$L_u(f) = \frac{1}{2^{|u|}} \sum_{a_u \in A^u} f|_{x_u = a_u},$$

and the variation is defined as

$$V_p(f) = \frac{1}{2^s} \left\| \left(\sum_{a_{S-u} \in A^{S-u}} \frac{\partial^{|u|} f}{\partial x_u} \bigg|_{x_{S-u} = a_{S-u}} \right)_{u \neq \emptyset} \right\|.$$

The symmetric discrepancy is defined as follows:

$$\frac{\partial^{|u|} \xi_u}{\partial x_u} = 2^{2|u|} \left\{ \prod_{j \in u}(1/2 - x_j) - \frac{1}{N} \sum_{z \in P} \prod_{j \in u}(1/2 - 1_{x_j > z_j}) \right\}$$

$$= 2^{|u|} \left\{ \prod_{j \in u}(1 - 2x_j) - \frac{1}{N} \sum_{z \in P} \prod_{j \in u} \mathrm{sign}(z_j - x_j) \right\}$$

$$= (-2)^{|u|} \left\{ 2\mathrm{Vol}\,(J_e(x_u)) - 1 \right.$$

$$\left. - \frac{1}{N} \left[|P_u \cap J_e(x_u)| - |P_u \cap J_o(x_u)| \right] \right\}$$

$$= 2(-2)^{|u|} \left\{ \mathrm{Vol}\,(J_e(x_u)) - \frac{|P_u \cap J_e(x_u)|}{N} \right\}$$

$$= -2(-2)^{|u|} \left\{ \mathrm{Vol}\,(J_o(x_u)) - \frac{|P_u \cap J_o(x_u)|}{N} \right\},$$

$$D_p(P) = 2 \left\| \left(\frac{|P_u \cap J_e(x_u)|}{N} - \text{Vol}\,(J_e(x_u)) \right)_{u \neq \emptyset} \right\|_p,$$

$$D_\infty(P) = 2 \left\| \frac{|P \cap J_e(x)|}{N} - \text{Vol}\,(J_e(x)) \right\|_\infty,$$

$$[D_2(P)]^2 = \left(\frac{4}{3} \right)^s - \frac{2}{N} \sum_{z \in P} \prod_{j=1}^{s} \left(1 + 2z_j - 2z_j^2 \right) + \frac{2^s}{N^2} \sum_{z,z' \in P} \prod_{j=1}^{s} (1 - |z_j - z_j'|),$$

where we have applied Theorem 5.1. Comparing this formula to that for the star discrepancy, the single rectangular solid $[0,x)$ appearing in (5.1) has been replaced by a union of rectangular solids $J_e(x)$.

5.4. P_α and its generalizations. As the final example let $\mu(x) = 0$ so $M = 1$. For arbitrary β the error bound coefficients c and $\tilde{c}$ in (1.11) correspond to $D_2(P)$ and $\tilde{D}_{2,1}(P)$, respectively. The error bounds (1.12) are the same as those in Theorems 3.5 and 4.2. For $\beta = 2\pi$ it follows from (1.8) and (4.3a) that

$$P_{2t} = [\tilde{D}_{2,t}(P)]^2$$

for any positive integer t.

As one can see from the examples above the discrepancy and the variation, as well as the operator L and the norms depend inherently on β and μ. Although the choice of β and μ is important, it can be shown that different choices are equivalent in a certain sense to be made precise below. For $i = 1, 2$ let β_i and μ_i denote two different choices of β and μ, respectively. Furthermore let the superscript (i) denote the dependence of other pertinent quantities on this choice. The following theorem gives the relationship between the different discrepancies and variations defined by β_i and μ_i.

Theorem 5.2. *For any positive β_1 and β_2 and any μ_1 and μ_2 satisfying condition (2.4) let*

$$\Delta_p = \|\mu_2' - \mu_1'\|_p, \qquad d_p = \|(1, \Delta_p)\|_p,$$

$$K(p, \beta_1, \beta_2, \mu_1, \mu_2) = d_q^s \max_{k=1,s} \left\| \beta_2(1, (d_q \beta_1)^{-1}) \right\|_p^k, \quad p^{-1} + q^{-1} = 1.$$

For all finite samples P and all $f \in X_p$ the generalized $\mathcal{L}^p$-variations and discrepancies associated with the β_i, μ_i $(i = 1, 2)$ are related as follows:

$$V_p^{(1)}(f) \leq K(p, \beta_1, \beta_2, \mu_1, \mu_2) V_p^{(2)}(f),$$

$$D_p^{(1)}(P) \leq K(q, \beta_2, \beta_1, \mu_2, \mu_1) D_p^{(2)}(P),$$

where $p^{-1} + q^{-1} = 1$. Furthermore, if the $\mu_i \in \tilde{X}_{\infty,t}([0,1))$, then for all $f \in \tilde{X}_{p,t}$ the generalized periodic $\mathcal{L}^p$-variations and discrepancies associated with the β_i, μ_i $(i = 1, 2)$ are related as follows:

$$\tilde{V}_{p,t}^{(1)}(f) \leq K(p, \beta_1^t, \beta_2^t, \mu_1^{(t-1)}, \mu_2^{(t-1)}) \tilde{V}_{p,t}^{(2)}(f),$$

$$\tilde{D}_{p,t}^{(1)}(P) \leq K(q, \beta_2^t, \beta_1^t, \mu_2^{(t-1)}, \mu_1^{(t-1)}) \tilde{D}_{p,t}^{(2)}(P).$$

Proof. Let $L^{(i)}$ denote the operator L associated with μ_i, and let $f_u^{(i)}$ denote the component of f associated with μ_i. It follows that for any $u \subseteq S$ the mixed partial

derivative of the component $f_u^{(1)}$ can be written in terms of mixed partial derivatives of the components $f_v^{(2)}$ for $v \supseteq u$ by (3.5):

$$\frac{\partial^{|u|} f_u^{(1)}}{\partial x_u} = L_{S-u}^{(1)} \frac{\partial^{|u|} f}{\partial x_u} = \sum_{u \subseteq v \subseteq S} (L_{v-u}^{(1)} - L_{v-u}^{(2)}) L_{S-v}^{(2)} \frac{\partial^{|u|} f}{\partial x_u}$$

$$= \sum_{u \subseteq v \subseteq S} \int_{C^{v-u}} \left[\prod_{j \in v-u} \mu_2'(x_j) - \mu_1'(x_j) \right] \frac{\partial^{|v|} f_v^{(2)}}{\partial x_v} \, dx_{v-u}.$$

For $p^{-1} + q^{-1} = 1$ one may apply Hölder's inequality to the sum of integrals over C^{v-u}. This yields

$$\left| \frac{\partial^{|u|} f_u^{(1)}}{\partial x_u} \right| \leq \left\| \left(\prod_{j \in v-u} \mu_2'(x_j) - \mu_1'(x_j) \right)_{u \subseteq v \subseteq S} \right\|_q \left\| \left(\frac{\partial^{|v|} f_v^{(2)}}{\partial x_v} \right)_{u \subseteq v \subseteq S} \right\|_p .$$

$$\leq d_q^{s-|u|} \left\| \left(\frac{\partial^{|v|} f_v^{(2)}}{\partial x_v} \right)_{u \subseteq v \subseteq S} \right\|_p ,$$

where d_q was defined in the statement of this theorem. Substituting this inequality into the formula for the variation leads to the desired bound:

$$V_p^{(1)}(f) = \left\| \left(\beta_1^{-|u|} \frac{\partial^{|u|} f_u^{(1)}}{\partial x_u} \right)_{u \neq \emptyset} \right\|_p$$

$$\leq \left\| \left(\beta_1^{-|u|} d_q^{s-|u|} \left\| \left(\frac{\partial^{|v|} f_v^{(2)}}{\partial x_v} \right)_{u \subseteq v \subseteq S} \right\|_p \right)_{u \neq \emptyset} \right\|_p$$

$$= d_q^s \left\| \left(\left[\sum_{\emptyset \subset u \subseteq v} \beta_1^{-|u|} d_q^{-|u|} \right] \frac{\partial^{|v|} f_v^{(2)}}{\partial x_v} \right)_{v \neq \emptyset} \right\|_p$$

$$\leq d_q^s \left\| \left([1 + \beta_1^{-1} d_q^{-1}]^{|v|} \frac{\partial^{|v|} f_v^{(2)}}{\partial x_v} \right)_{v \neq \emptyset} \right\|_p$$

$$\leq d_q^s \max_{1 \leq k \leq s} [1 + \beta_1^{-1} d_q^{-1}]^k \beta_2^k \left\| \left(\beta_2^{-|v|} \frac{\partial^{|v|} f_v^{(2)}}{\partial x_v} \right)_{v \neq \emptyset} \right\|_p$$

$$= K(p, \beta_1, \beta_2, \mu_1, \mu_2) V_p^{(2)}(f).$$

The inequalities for the discrepancy follow by considering the inequalities for the variation and $\xi_q^{(1)}$, the worst-case integrand (3.11) under β_1 and μ_1. For $1 \leq p < \infty$ and $p^{-1} + q^{-1} = 1$ it follows that

$$D_p^{(1)}(P) V_q^{(1)}(\xi_q^{(1)}) = |I(\xi_q^{(1)}) - Q(\xi_q^{(1)})| \leq D_p^{(2)}(P) V_q^{(2)}(\xi_q^{(1)})$$

$$\leq D_p^{(2)}(P) K(q, \beta_2, \beta_1, \mu_2, \mu_1) V_q^{(1)}(\xi_q^{(1)}).$$

There is no worst-case integrand for $q = 1$ ($p = \infty$). The inequality in this case can be obtained by observing that the discrepancy is a continuous function of p

and taking the limit as p tends to infinity. The proof for the periodic case follows similarly.   □

The inequalities obtained above are not the tightest possible. For example, $K(p, \beta_1, \beta_1, \mu_1, \mu_1) > 1$. However, they are among the simplest the author has found.

Theorem 5.2 gives an equivalence among different discrepancies, that is, a low discrepancy sample for one particular choice of μ and β cannot have too high a discrepancy for another choice of μ and β. However, this does not imply that the choice of μ and β are unimportant. The better sample under one type of discrepancy may not be the better sample under a different type.

6. DISCUSSION AND CONCLUSION

A number of figures of merit for quasi-random points appear in the literature for quadrature and the related field of experimental design. For an extensive list see [FW94, Chapters 1 and 5] and [FH95]. One might prefer a figure of merit, $D(P)$, to have the following desirable qualities:

1. $D(P)$ should arise from an error bound (for quadrature, function approximation, or some other relevant application).
2. Projections of P into a lower-dimensional space should not increase $D(P)$.
3. $D(P)$ should be easy to compute.
4. $D(P)$ should have an intuitive interpretation.
5. $D(P)$ should be invariant under certain transformations of P, such as reflections about the plane $x_j = 1/2$ and permutations of the coordinates.

The generalized $\mathcal{L}^p$-discrepancy derived here satisfies the first two criteria. However, the traditional $\mathcal{L}^p$-star discrepancy (5.2) does not satisfy them unless $p = \infty$. Thus, we prefer the definition of star discrepancy given by (5.1).

The easiest generalized $\mathcal{L}^p$-discrepancy to compute is the case $p = 2$, since it involves at worst a double sum (see (3.9) and (4.3b)) that requires $O(N^2)$ operations to evaluate. In fact Heinrich [Hei96] gives an algorithm for the $\mathcal{L}^2$-star discrepancy that requires only $O(N(\log N)^s)$ operations. For $\mu = 0$ the generalized periodic $\mathcal{L}^2$-discrepancy can be reduced to a single sum for lattice rules, thus requiring only $O(N)$ operations. Some might prefer to use the generalized $\mathcal{L}^\infty$-discrepancy as a figure of merit because it makes the weakest assumption on the integrands. In this case choosing β and μ so that Theorem 5.1 can be applied will simplify the calculation.

The star, centered and symmetric discrepancies can all be interpreted in terms of the relative proportion of points lying in subsets of the integration domain. For P_α and its generalizations we have not yet found a geometric interpretation, however, the decomposition of the the integrand f into a sum of f_u corresponds to the analysis of variance (ANOVA) decomposition, which is very popular in statistics.

It has been shown [Woź91, MC94] that the traditional $\mathcal{L}^2$-star discrepancy gives the mean square quadrature error for a Brownian sheet which is zero on all faces of the cube $x_j = 1$. If one relaxes this condition so that the values of the integrand on these faces are themselves generalized Brownian sheets, then mean square quadrature error is the $\mathcal{L}^2$-star discrepancy defined here (Caflisch, private communication and [MC94]).

The star discrepancy is not invariant under reflections of P about planes $x_j = 1/2$. The other three figures of merit in Section 5 do have this quality. All figures of merit considered in this paper are invariant to permutations of the coordinates.

For lattice integration rules, the speed at which one can compute P_α makes it a very attractive figure of merit. However, for general quadrature rules the symmetric discrepancy has much to commend it. Although not backed by the weight of history, it satisfies all the above criteria whereas the star discrepancy does not.

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

[DR84]   P. J. Davis and P. Rabinowitz, *Methods of numerical integration*, Academic Press, Orlando, Florida, 1984. MR **86d**:65004

[FH95]   K. T. Fang and F. J. Hickernell, *The uniform design and its applications*, Bulletin of the International Statistical Institute, 50$^{\text{th}}$ Session, Book 1 (Beijing), 1995, pp. 333–349.

[FW94]   K. T. Fang and Y. Wang, *Number theoretic methods in statistics*, Chapman and Hall, New York, 1994. MR **95g**:65189

[Hei96] S. Heinrich, *Efficient algorithms for computing the L_2 discrepancy*, Math. Comp. **65** (1996), 1621–1633. MR **97a**:65024

[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.

[Hic96] F. J. Hickernell, *Quadrature error bounds with applications to lattice rules*, SIAM J. Numer. Anal. **33** (1996), 1995–2016. CMP 97:02

[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

[Owe92] A. B. Owen, *Orthogonal arrays for computer experiments, integration and visualization*, Statist. Sinica **2** (1992), 439–452. MR **93h**:62135

[Owe95] A. B. Owen, *Equidistributed Latin hypercube samples*, 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.

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

[Sob69] I. M. Sobol', *Multidimensional quadrature formulas and Haar functions (in Russian)*, Izdat. "Nauka", Moscow, 1969. MR **54**:10952

[War72] T. T. Warnock, *Computational investigations of low discrepancy point sets*, Applications of Number Theory to Numerical Analysis (S. K. Zaremba, ed.), Academic Press, New York, 1972, pp. 319–343. MR **50**:3526

[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

URL: http://www.math.hkbu.edu.hk/~fred