Lattice Rules: How Well Do They Measure Up?

Fred J. Hickernell [1]

ABSTRACT Lattice rules are numerical methods for evaluating multi-dimensional integrals. The formula for an integration lattice is relatively simple and straightforward to program. A popular quadrature error bound for lattice rules involves the term $\mathcal{P}_\alpha(L)$, a measure of quality of the integration lattice L. Asymptotic upper bounds on the decay rate of $\mathcal{P}_\alpha(L)$ in terms of the size of the lattice are known. However, the original definition of $\mathcal{P}_\alpha(L)$ makes sense only for lattices. This makes it difficult to compare lattice rules with other quasi-Monte Carlo quadrature rules. Furthermore, the existing upper bounds on $\mathcal{P}_\alpha(L)$ and some recent tractability results give a pessimistic view of lattice rules for high dimensional quadrature. This chapter demonstrates the relationship between $\mathcal{P}_\alpha(L)$ and other measures of quality for quasi-random sets, such as the star and unanchored discrepancies. Moreover, it is shown how by generalizing the definition of $\mathcal{P}_\alpha(L)$ one may obtain a more optimistic view of lattice rules in high dimensions.

1 Introduction

A simple, but often effective, way to approximate an integral over the s-dimensional unit cube is to take the average of the integrand over some set P of N points. Monte Carlo methods choose P randomly and typically obtain an error of $\mathrm{O}(N^{-1/2})$. Quasi-Monte Carlo methods attempt to decrease the error by choosing P in a deterministic (or quasi-random) way so that the points are more uniformly spread over the integration domain.

One popular family of such quasi-random points is **integration lattices**, and the resulting quadrature formulae are called **lattice rules**. These lattice rules may be thought of as s-dimensional generalizations of the rectangle rule for one-dimensional quadrature. Lattice rules go back at least as far as Korobov [Kor59], but their general formulation and analysis began with the work of Sloan and his collaborators [Slo85, SK87]. Several excellent monographs review the large body of research on lattice rules [HW81, Nie92, SJ94].

[1] Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong, fred@hkbu.edu.hk, http://www.math.hkbu.edu.hk/~fred. This research was supported by an HKBU FRG grant 96-97/II-67.

The purpose of this chapter is to highlight some recent developments in the study of lattice rules. As the title suggests, the focus is on how well integration lattices perform for quadrature. The answer depends on how one measures performance.

The next section reviews some basic properties of lattice rules, including a popular quadrature error bound involving $\mathcal{P}_\alpha(L)$, a measure of the quality of an integration lattice. Section 3 shows this error bound to be a special case of quite general worst-case and average-case error analyses. Several other examples included in this general framework are given in Section 4, such as the Koksma-Hlawka inequality involving the star discrepancy. In Section 5 the notion of a shift-invariant discrepancy is introduced, and it is seen that a weighted form of $\mathcal{P}_\alpha(L)$ is closely connected to many interesting discrepancies.

Upper and lower bounds on $\mathcal{P}_\alpha(L)$ and related discrepancies are discussed in Section 6, and the discrepancies of integration lattices and nets are compared in Section 7. In Section 8 we look again at how the choice of weights in the definition of $\mathcal{P}_\alpha(L)$ can affect its size, in particular for high dimensions. The choice of weights both reflects and determines whether one takes an optimistic or pessimistic view of lattice rules for large s. Concluding remarks are given in the final section.

Several places in this chapter we quote results in somewhat more general form than they appear in the literature. Rather than proving the generalizations here, we refer the reader to the original proofs, which can be suitably modified without much difficulty.

Although this chapter touches on several recent developments in the study of lattice rules, the choice of topics is somewhat subjective and definitely not comprehensive. Some important topics not covered in this chapter include recent work on finding lattice rules with low Zaremba merit or trigonometric degree [CS96, Lan96, LS97], and canonical forms of lattice rules [Lan95, LJ96].

2 Some Basic Properties of Lattice Rules

The s-dimensional integral that we wish to evaluate has the form:

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

where $C^s = [0,1)^s$ is the unit cube. A general quasi-Monte Carlo quadrature rule to approximate this integral takes the average of the integrand values over some set $P \subset C^s$ with N points:

$$Q(f) \equiv Q(f;P) \equiv \frac{1}{N}\sum_{z\in P} f(z). \tag{2.1}$$

Throughout this chapter P is allowed to have multiple copies of a point (see [Nie92, p. 14]). In this rule the weight given each integrand value is $1/N$. So, the weight is the same for integrand value and the sum of all the weights is one, which implies that the rule is exact for constant integrands. Quadrature rules with unequal weights and which may not be exact for constant integrands are considered in Section 6.3.

In the one-dimensional case the choice

$$P = \left\{0, \frac{1}{N}, \ldots, \frac{N-1}{N}\right\},$$

yields the left rectangle rule. There is more than one way to generalize this to a two-dimensional rule. The most straightforward way might be a grid,

$$P = \left\{\left(\frac{i_1}{\sqrt{N}}, \frac{i_2}{\sqrt{N}}\right) : i_1, i_2 = 0, \ldots, \sqrt{N}-1\right\},$$

but another possibility is

$$P = \left\{\left(\frac{i}{N}, \left\{\frac{ih}{N}\right\}\right) : i = 0, \ldots, N-1\right\},$$

where h is some integer, and $\{x\} = x - \lfloor x \rfloor$ denotes the fractional part of x. Figures 2.1 and 2.2 show both of these sets P for the case $N = 64$ and $h = 51$. These P are the node sets of integration lattices as defined below:

Definition 2.1.

a. An s-dimensional **integration lattice**, L, is a discrete subset of $\mathbf{R}^s$ that is closed under addition and subtraction and which contains the integer vectors $\mathbf{Z}^s$ as a subset.

b. A **shifted lattice** with shift $\Delta \in \mathbf{R}^s$ is the set $L+\Delta \equiv \{z+\Delta : z \in L\}$ for some lattice L.

c. The **node set** for a shifted integration lattice, $L + \Delta$, is the set of points in the lattice that fall inside the unit cube, that is, $P = (L + \Delta) \cap C^s$.

d. The **dual lattice** of a lattice L is denoted $L^\perp$ and is defined as $\{k \in \mathbf{R}^s : k \cdot z \in \mathbf{Z} \text{ for all } z \in L\}$. (Here and below $\cdot$ denotes the dot product of two vectors.)

e. A **rank-1 lattice** is a lattice whose node set may be expressed as

$$P = \{\{ih/N\} : i = 0, \ldots, N-1\}, \tag{2.2}$$

for some generating vector $h \in \mathbf{Z}^s$.

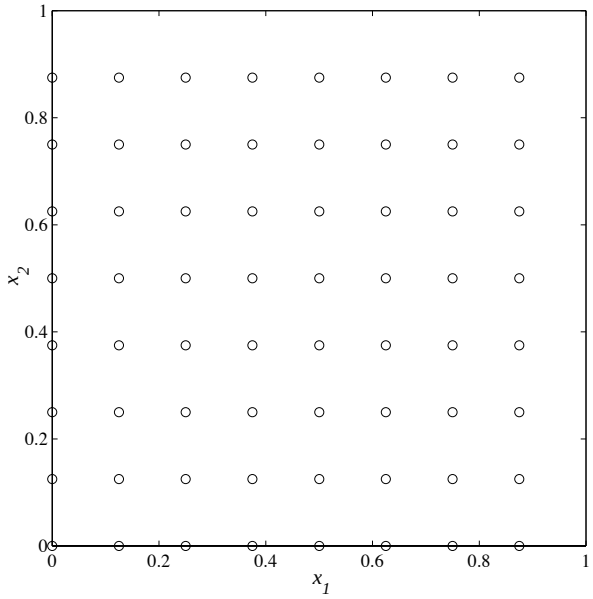

FIGURE 2.1. A grid of 64 points.

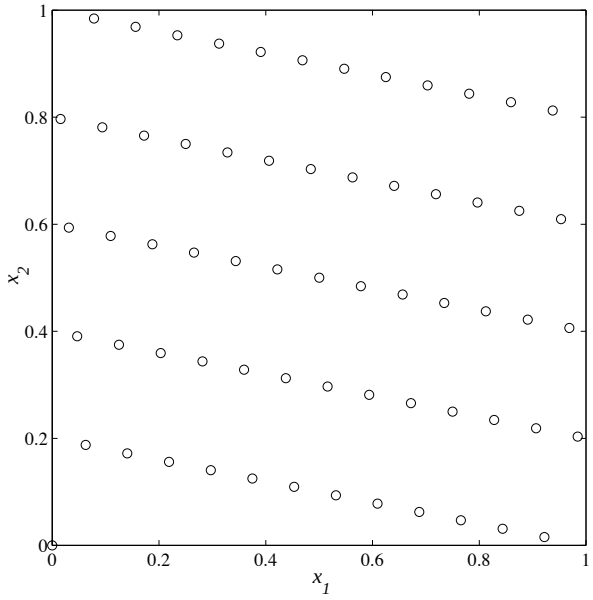

FIGURE 2.2. A rank-1 lattice with 64 points.

Although an integration lattice contains an infinite number of points, its node set, P, which defines the lattice quadrature rule through (2.1), is always finite. Below the notation $Q(f; L+\Delta)$ is sometimes used to denote a quadrature rule based on a (shifted) lattice, that is, $Q(f; L+\Delta) \equiv Q(f; P)$, where P is the node set for $L + \Delta$. The one-dimensional set

$$P = \left\{ \frac{1}{2N}, \frac{3}{2N}, \ldots, \frac{2N-1}{2N} \right\}$$

is the node set of the shifted lattice (with shift $\Delta = \frac{1}{2N}$) and gives the midpoint rule.

The left rectangle rule gives low accuracy in general, but it gives high accuracy if the integrand is periodic and has a sufficient degree of smoothness. The same general principle is true for lattice rules. Therefore, one often uses Fourier analysis to investigate the error of lattice rules. Suppose that the integrand, f, has an absolutely convergent Fourier series,

$$f(x) = \sum_{k \in \mathbf{Z}^s} \hat{f}(k) e^{2\pi i k \cdot x}, \tag{2.3a}$$

where the Fourier coefficients, $\hat{f}(k)$ are defined by

$$\hat{f}(k) \equiv \mathcal{F}\{f\}(k) \equiv \int_{C^s} f(x) e^{-2\pi i k \cdot x} \, dx. \tag{2.3b}$$

The quadrature error for an absolutely convergent Fourier series is the sum of the error in each term. Since constants are integrated exactly, this implies that

$$\begin{aligned} I(f) - Q(f; P) &= -{\sum_{k \in \mathbf{Z}^s}}' \hat{f}(k) Q(e^{2\pi i k \cdot x}; P) \\ &= -{\sum_{k \in \mathbf{Z}^s}}' \left\{ \hat{f}(k) \left[\frac{1}{N} \sum_{z \in P} e^{2\pi i k \cdot z} \right] \right\}, \end{aligned} \tag{2.4}$$

where the notation $\sum'$ means that the zero term is to be omitted from the sum.

An intriguing feature of lattice rules is that they integrate the term $e^{2\pi i k \cdot x}$ with either no error or 100% error, depending on whether k is in the dual lattice:

Lemma 2.2. *[SJ94, Lemma 2.7] If L is an integration lattice, then*

$$Q\left(e^{2\pi i k \cdot x}; L\right) = 1_{\{k \in L^\perp\}} = \begin{cases} 1, & k \in L^\perp, \\ 0, & k \notin L^\perp. \end{cases}$$

Here and below, $1_{\{\cdot\}}$ denotes the indicator function. This lemma allows one to write the error of a shifted lattice rule in terms of the Fourier coefficients of the integrand evaluated on the dual lattice:

Theorem 2.3. *[SJ94, Theorem 2.10] The error of a shifted lattice rule is:*

$$I(f) - Q(f; L + \Delta) = -\sum_{k \in L^\perp}{}' \hat{f}(k) e^{2\pi i k \cdot \Delta}.$$

From this theorem one can easily derive a lattice rule error bound. First, we introduce the notation

$$\bar{k}_j = \begin{cases} |k_j|, & k_j \neq 0, \\ 1, & k_j = 0. \end{cases} \tag{2.5}$$

Note that his definition is somewhat different than the usual one, $\bar{k}_j = \max(|k_j|, 1)$, however, the two definitions are equivalent if k_j is an integer. An error bound for lattice rules follows by applying Hölder's inequality to Theorem 2.3:

Theorem 2.4. *[SJ94, Section 5.2] Suppose that the Fourier coefficients of the integrand decay fast enough so that:*

$$\left\| (\bar{k}_1 \cdots \bar{k}_s)^\alpha \hat{f}(k) \right\|_\infty < \infty \text{ for some } \alpha > 1,$$

where $\|\cdot\|_p$ denotes the ℓ_p-norm. Then the error of a shifted lattice rule is bounded as follows:

$$|I(f) - Q(f; L + \Delta)| \leq \mathcal{P}_\alpha(L) \left\| 1_{\{k \neq 0\}} (\bar{k}_1 \cdots \bar{k}_s)^\alpha \hat{f}(k) \right\|_\infty, \tag{2.6}$$

where $\mathcal{P}_\alpha(L)$ is defined as:

$$\mathcal{P}_\alpha(L) \equiv \left\| \frac{1_{\{k \neq 0, k \in L^\perp\}}}{(\bar{k}_1 \cdots \bar{k}_s)^\alpha} \right\|_1 = \sum_{k \in L^\perp}{}' \frac{1}{(\bar{k}_1 \cdots \bar{k}_s)^\alpha}. \tag{2.7}$$

This theorem and the quantity $\mathcal{P}_\alpha(L)$ are the springboard for much of the analysis and discussion in the remainder of this chapter. To motivate what follows a couple of remarks are in order.

Remark 2.5. The error bound above is a product of *two* parts. One part measures the quality of the lattice. The other part measures the roughness of the integrand. The following sections investigate other error bounds that have this same general form.

Remark 2.6. Error bound (2.6) and $\mathcal{P}_\alpha(L)$ are specific to *lattice* rules. It is both desirable and possible to generalize these so that they apply to all quasi-Monte Carlo quadrature rules of the form (2.1). This will be done in Sections 3 and 4.

The quantity $\mathcal{P}_\alpha(L)$ may be written as the lattice rule error for the particular integrand $\xi_\alpha(x)$ defined as follows:

$$\xi_\alpha(x) = -\sum_{k \in \mathbf{Z}^s} \frac{e^{2\pi i k \cdot x}}{(\bar{k}_1 \cdots \bar{k}_s)^\alpha}. \tag{2.8}$$

By Theorem 2.3 it follows that $\mathcal{P}_\alpha(L)$ is the quadrature error obtained when integrating ξ_α using an unshifted lattice L:

$$\mathcal{P}_\alpha(L) = I(\xi_\alpha) - Q(\xi_\alpha; L).$$

Remark 2.7. The error bound in Theorem 2.4 is tight, equality holding when the integrand is a constant multiple of ξ_α.

When α is a positive even integer one may write ξ_α and $\mathcal{P}_\alpha(L)$ in terms of Bernoulli polynomials. The definitions and some important properties of these polynomials are given in [AS64, Chapter 23]. The first few Bernoulli polynomials are:

$$B_0(x) = 1, \qquad B_1(x) = x - \frac{1}{2}, \qquad B_2(x) = x^2 - x + \frac{1}{6},$$
$$B_3(x) = x^3 - \frac{3}{2}x^2 + \frac{1}{2}x, \qquad B_4(x) = x^4 - 2x^2 + x - \frac{1}{30}.$$

The Fourier expansions of the Bernoulli polynomials of even degree are:

$$B_\alpha(x) = \frac{-(-1)^{\alpha/2}\alpha!}{(2\pi)^\alpha} \sum_{k \neq 0} \frac{e^{2\pi i k x}}{|k|^\alpha}, \quad 0 \leq x \leq 1 \tag{2.9}$$

which implies that for even integer $\alpha \geq 2$:

$$\xi_\alpha(x) = -\prod_{j=1}^{s} \left[1 - \frac{(-1)^{\alpha/2}(2\pi)^\alpha}{\alpha!} B_\alpha(x_j)\right], \quad x \in C^s,$$

$$\mathcal{P}_\alpha(L) = -1 + \frac{1}{N} \sum_{z \in P} \prod_{j=1}^{s} \left[1 - \frac{(-1)^{\alpha/2}(2\pi)^\alpha}{\alpha!} B_\alpha(z_j)\right],$$

where P is the node set of the lattice L.

Remark 2.8. The advantage of this formula for $\mathcal{P}_\alpha(L)$ is that it involves only a sum over the N points in the node set as opposed to the general formula (2.7) that requires a sum over an infinite number of points in the dual lattice.

3 A General Approach to Worst-Case and Average-Case Error Analysis

The previous section provides a formula for the error of lattice rules (Theorem 2.3) and an error bound for lattice rules (Theorem 2.4). In this section a general framework is presented that extends these results to general quasi-Monte Carlo quadrature rules and other spaces of integrands. Worst-case error analysis is discussed first, as this more closely parallels the development of the previous section. An average-case error analysis is also presented. For related presentations of worst-case and average-case error analyses see [Rit95, FH96, Hic98].

3.1 Worst-Case Quadrature Error for Reproducing Kernel Hilbert Spaces

In the worst-case error analysis the integrands are assumed to lie in a Hilbert space, $\mathcal{W}$, of real-valued functions on C^s. This Hilbert space is assumed to have a reproducing kernel $K(x,y)$ and an inner product $\langle \cdot, \cdot \rangle_K$. The definition of a reproducing kernel is a function on $C^s \times C^s$ such that $K(\cdot, y) \in \mathcal{W}$ for all $y \in C^s$, and

$$f(y) = \langle K(\cdot, y), f \rangle_K \text{ for all } f \in \mathcal{W} \text{ and } y \in C^s.$$

For a more thorough discussion of reproducing kernels see [Sai88, Wah90]. It can be shown that any reproducing kernel is symmetric in its arguments and positive definite:

$$K(x,y) = K(y,x) \quad \text{for all } x, y \in C^s \tag{3.1a}$$

$$\sum_{i,k=1}^{N} a^{(i)} a^{(k)} K(x^{(i)}, x^{(k)}) \geq 0 \quad \text{for all } a^{(i)} \in \mathbf{R},\ x^{(i)} \in C^s. \tag{3.1b}$$

Conversely, any function $K(x,y)$ satisfying these two conditions uniquely defines a reproducing kernel Hilbert space.

The advantage of assuming the integrands to be in a reproducing kernel Hilbert space is that the quadrature error is then a bounded linear functional on that space. Furthermore, its representer, ξ_K, may be easily found in terms of the reproducing kernel. By the definition of the reproducing kernel, $\xi_K(x) = \langle K(\cdot, x), \xi_K \rangle_K$, and by the definition of representer, $\langle K(\cdot, x), \xi_K \rangle_K = I(K(\cdot, x)) - Q(K(\cdot, x); P)$, so

$$\xi_K(x) = I(K(\cdot, x)) - Q(K(\cdot, x); P) \tag{3.2}$$

(The dependence of ξ_K on the point set P defining the quadrature rule is suppressed for ease of notation). The Cauchy-Schwarz inequality then yields a worst-case error bound:

$$|I(f) - Q(f;P)| = |\langle \xi_K, f \rangle_K| \leq \|\xi_K\|_K \, \|f\|_K .$$

This error bound is tight since equality holds when f is a multiple of ξ_K, the worst-case integrand. The discrepancy is defined as the norm of the worst-case integrand, which is equivalent to the square root of the quadrature error for the worst-case integrand. The discrepancy may be written in terms of integrals and sums involving the reproducing kernel:

$$\begin{aligned} D(P;K) \equiv \|\xi_K\|_K &= \langle \xi_K, \xi_K \rangle_K^{1/2} = [I(\xi_K) - Q(\xi_K; P)]^{1/2} \\ &= \left\{ \int_{C^s \times C^s} K(x,y) \, dx \, dy - \frac{2}{N} \sum_{z \in P} \int_{C^s} K(z,y) \, dy \right. \end{aligned} \tag{3.3}$$

$$+\frac{1}{N^2}\sum_{z,z'\in P} K(z,z')\Bigg\}^{1/2}.$$

Since quadrature rule (2.1) is exact for constants, one may replace the norm of the integrand by the variation of the integrand, which is defined as the norm of the nonconstant part, that is, the part which is orthogonal to the function 1.

$$V(f;K) \equiv \|f_\perp\|_K, \quad \text{where } f_\perp \equiv f - \frac{1\,\langle f,1\rangle_K}{\langle 1,1\rangle_K}. \tag{3.4}$$

Note that the discrepancy is the variation of the worst-case integrand. The following theorem gives the quadrature error bound as derived above.

Theorem 3.1. *Suppose that $\mathcal{W}$ is a reproducing kernel Hilbert space of real-valued functions on C^s with reproducing kernel $K(x,y)$, inner product $\langle\cdot,\cdot\rangle_K$, and norm $\|\cdot\|_K$. Then the error for the quadrature rule defined in* (2.1) *has the following upper bound:*

$$|I(f) - Q(f;P)| \le D(P;K)V(f;K),$$

in terms of the discrepancy of P and the variation of f as defined above. This bound is attained for the worst-case integrand, ξ_K, defined above.

Remark 3.2. In this theorem the set P is an arbitrary subset of the unit cube and need not come from a lattice. Therefore the error bound applies to any quasi-Monte Carlo quadrature rule. However, the special case where P is the node set of a lattice will be discussed below.

Remark 3.3. The discrepancy is uniquely determined by any $K(x,y)$ satisfying (3.1). Moreover, the choice of K also uniquely determines the space of integrands $\mathcal{W}$ and the accompanying inner product (see [Wah90]).

Remark 3.4. The formula for the discrepancy in (3.3) requires only $O(N^2)$ operations to evaluate. In fact, for the case of the $\mathcal{L}_2$-star discrepancy, which is a special case of (3.3) (see Section 4.5), the number of operations required can be reduced to $O(N[\log N]^s)$ [Hei96].

Remark 3.5. The worst-case integrand is the quadrature error obtained upon integrating the reproducing kernel.

This theorem provides many worst-case error bounds, each arising from a different choice of the reproducing kernel $K(x,y)$. To illustrate this theorem consider an inner product motivated by Theorem 2.3 with the associated

norm:

$$\langle f, g\rangle_{\mathcal{F},\alpha} \equiv \sum_{k\in\mathbf{Z}^s} (\bar{k}_1\cdots\bar{k}_s)^\alpha \hat{f}(k)\hat{g}^*(k), \tag{3.5}$$

$$\|f\|_{\mathcal{F},\alpha} = \left\{\sum_{k\in\mathbf{Z}^s} (\bar{k}_1\cdots\bar{k}_s)^\alpha |\hat{f}(k)|^2\right\}^{1/2},$$

where * denotes the complex conjugate. The subscript $\mathcal{F}$ denotes the fact that this inner product is defined in terms of the Fourier coefficients of a function. The space of integrands is a Hilbert space of real-valued, periodic functions, whose Fourier coefficients decay to zero sufficiently fast:

$$\mathcal{W}_{\mathcal{F},\alpha} \equiv \{f : \|f\|_{\mathcal{F},\alpha} < \infty\}, \quad \alpha > 1.$$

The reproducing kernel for this space is:

$$K_{\mathcal{F},\alpha}(x,y) = \sum_{k\in\mathbf{Z}^s} \frac{e^{2\pi i k\cdot(x-y)}}{(\bar{k}_1\cdots\bar{k}_s)^\alpha}. \tag{3.6}$$

This can be verified by noting that $K_{\mathcal{F},\alpha}(\cdot,y) \in \mathcal{W}_{\mathcal{F},\alpha}$ for all y and that

$$\langle K(\cdot,y), f\rangle_{\mathcal{F},\alpha} = \sum_{k\in\mathbf{Z}^s} (\bar{k}_1\cdots\bar{k}_s)^\alpha \frac{e^{-2\pi i k\cdot y}}{(\bar{k}_1\cdots\bar{k}_s)^\alpha}\hat{f}^*(k)$$
$$= \sum_{k\in\mathbf{Z}^s} e^{2\pi i k\cdot y}\hat{f}(k) = f(y).$$

According to Theorem 3.1 the quadrature error bound is

$$|I(f) - Q(f;P)| \le D_{\mathcal{F},\alpha}(P) V_{\mathcal{F},\alpha}(f), \tag{3.7}$$

where the worst-case integrand, (3.2), the discrepancy, (3.3), and the variation of the integrand, (3.4) are:

$$\xi_{\mathcal{F},\alpha}(x) = I(K(\cdot,x)) - Q(K(\cdot,x);P) \tag{3.8}$$
$$= -\frac{1}{N}\sum_{z\in P}{\sum_{k\in\mathbf{Z}^s}}' \frac{e^{2\pi i k\cdot(z-x)}}{(\bar{k}_1\cdots\bar{k}_s)^\alpha},$$

$$D_{\mathcal{F},\alpha}(P) = \left\{\frac{1}{N^2}\sum_{z,z'\in P}{\sum_{k\in\mathbf{Z}^s}}' \frac{e^{2\pi i k\cdot(z-z')}}{(\bar{k}_1\cdots\bar{k}_s)^\alpha}\right\}^{1/2}, \tag{3.9}$$

$$V_{\mathcal{F},\alpha}(f) \equiv \|f_\perp\|_{\mathcal{F},\alpha} = \left\{{\sum_{k\in\mathbf{Z}^s}}' (\bar{k}_1\cdots\bar{k}_s)^\alpha |\hat{f}(k)|^2\right\}^{1/2}. \tag{3.10}$$

Since $\xi_{\mathcal{F},\alpha}$ is also the representer of the quadrature error functional, the quadrature error for the integrand f is $\langle f, \xi_{\mathcal{F},\alpha}\rangle_{\mathcal{F},\alpha}$, which is just the right-hand side of (2.4).

For positive, even integers α the formulas for the reproducing kernel, the worst-case integrand, and the discrepancy may be written in terms of Bernoulli functions. Using the Fourier expansion (2.9), one obtains for all even integers $\alpha \geq 2$:

$$K_{\mathcal{F},\alpha}(x,y) = \prod_{j=1}^{s}\left[1 - \frac{(-1)^{\alpha/2}(2\pi)^{\alpha}}{\alpha!}B_{\alpha}(\{x_j - y_j\})\right],$$

$$\xi_{\mathcal{F},\alpha}(x) = 1 - \frac{1}{N}\sum_{z\in P}\prod_{j=1}^{s}\left[1 - \frac{(-1)^{\alpha/2}(2\pi)^{\alpha}}{\alpha!}B_{\alpha}(\{x_j - z_j\})\right],$$

(3.11)
$$D_{\mathcal{F},\alpha}(P) = \left\{-1 + \frac{1}{N^2}\sum_{z,z'\in P}\prod_{j=1}^{s}\left[1 - \frac{(-1)^{\alpha/2}(2\pi)^{\alpha}}{\alpha!}B_{\alpha}(\{z_j - z'_j\})\right]\right\}^{1/2}.$$

Remark 3.6. When computing $D_{\mathcal{F},\alpha}(P)$ in practice it is advisable to choose α to be an even integer. Formula (3.11) requires $\mathrm{O}(N^2)$ operations, whereas the number of operations required by formula (3.9) is $\mathrm{O}(N^2)$ times the work required to sum an infinite series.

If P is the node set of a shifted lattice, $L + \Delta$, then the worst-case integrand and the discrepancy may be written in terms of the dual lattice, $L^{\perp}$, by using Lemma 2.2 as follows:

$$\xi_{\mathcal{F},\alpha}(x) = -{\sum_{k\in L^{\perp}}}' \frac{e^{2\pi i k\cdot(x-\Delta)}}{(\bar{k}_1\cdots\bar{k}_s)^{\alpha}},$$

$$D_{\mathcal{F},\alpha}(P) = \left\{{\sum_{k\in L^{\perp}}}' \frac{1}{(\bar{k}_1\cdots\bar{k}_s)^{\alpha}}\right\}^{1/2} = \sqrt{\mathcal{P}_{\alpha}(L)},$$

and when α is a positive even integer:

$$\xi_{\mathcal{F},\alpha}(x) = -\prod_{j=1}^{s}\left[1 - \frac{(-1)^{\alpha/2}(2\pi)^{\alpha}}{\alpha!}B_{\alpha}(\{x_j - \Delta_j\})\right],$$

$$D_{\mathcal{F},\alpha}(P) = \left\{-1 + \frac{1}{N}\sum_{z\in P}\prod_{j=1}^{s}\left[1 - \frac{(-1)^{\alpha/2}(2\pi)^{\alpha}}{\alpha!}B_{\alpha}(\{z_j - \Delta_j\})\right]\right\}^{1/2}.$$

Remark 3.7. The number of operations required to compute $D_{\mathcal{F},\alpha}(P)$ is only $\mathrm{O}(N)$ when α is an even integer and P is the node set of a lattice. This is substantially faster than the case when P is arbitrary (see Remark 3.6).

3.2 A More General Worst-Case Quadrature Error Analysis

The error bound in (3.7) that came from Theorem 3.1 may be obtained by applying the Cauchy-Schwarz inequality to Theorem 2.3. This *does not* correspond to the lattice rule error bound in Theorem 2.4, which is derived by applying Hölder's inequality. However, one can extend the idea behind Theorem 3.1 to cover Theorem 2.4.

Suppose that the space of integrands, $\mathcal{W}$, is a Banach space of real-valued functions on the unit cube C^s with norm $\|\cdot\|_{\mathcal{W}}$. Instead of an inner product, $\langle\cdot,\cdot\rangle_K$, suppose, that one has a bilinear, non-negative definite function, $\langle\cdot,\cdot\rangle$ from $\mathcal{W}\times\mathcal{W}'$ into $\mathbf{R}$, where $\mathcal{W}'$ is another Banach space of real-valued functions on the unit cube C^s. The norm for $\mathcal{W}'$ is assumed to be related to $\|\cdot\|_{\mathcal{W}}$ and $\langle\cdot,\cdot\rangle$ in the following way:

$$\|f\|_{\mathcal{W}}\,\|g\|_{\mathcal{W}'} = \sup_{f\in\mathcal{W}}|\langle f,g\rangle| = \sup_{g\in\mathcal{W}'}|\langle f,g\rangle|. \tag{3.12}$$

Thus, for each $g\in\mathcal{W}'$, the function $\langle\cdot,g\rangle$ is a bounded linear functional on $\mathcal{W}$, and $\mathcal{W}'$ is isomorphic to the dual space of $\mathcal{W}$. Conversely, for each $f\in\mathcal{W}$, the function $\langle f,\cdot\rangle$ is a bounded linear functional on $\mathcal{W}'$. Suppose that the evaluation functional is bounded on $\mathcal{W}'$, so that there exists a real-valued function K on $C^s\times C^s$ such that

$$K(\cdot,x)\in\mathcal{W} \quad \text{for all } x\in C^s, \tag{3.13a}$$

$$g(x) = \langle K(\cdot,x),g\rangle \quad \text{for all } x\in C^s, g\in\mathcal{W}'. \tag{3.13b}$$

Then the quadrature error may be written in terms of the representer for the quadrature error, ξ_K, which in turn is found from the kernel as in (3.2) above:

$$\begin{aligned} I(f)-Q(f;P) &= \langle f,\xi_K\rangle\,, \\ \xi_K(x) &= \langle K(\cdot,x),\xi_K\rangle = I(K(\cdot,x))-Q(K(\cdot,x);P). \end{aligned} \tag{3.14}$$

Applying condition (3.12) then yields the desired error bound.

Theorem 3.8. *Suppose that $\mathcal{W}$ and $\mathcal{W}'$ are two Banach spaces of real-valued functions on the unit cube C^s, as described above, and that $\langle\cdot,\cdot\rangle$ is a bilinear, non-negative definite function from $\mathcal{W}\times\mathcal{W}'$ into $\mathbf{R}$ satisfying* (3.12). *Suppose that both $\mathcal{W}$ and $\mathcal{W}'$ contain the function 1, and that a kernel $K(\cdot,\cdot)$ exists that satisfies conditions* (3.13). *Then the error for the quadrature rule defined in* (2.1) *has the following upper bound:*

$$|I(f)-Q(f;P)| = |\langle f,\xi_K\rangle| \le D(P;\mathcal{W}')V(f;\mathcal{W}).$$

The discrepancy of P is defined as the norm of the function ξ_K defined in (3.14)*, and the variation of f is defined as the norm of its non-constant part:*

$$D(P;\mathcal{W}') \equiv \|\xi_K\|_{\mathcal{W}'},$$
$$V(f;\mathcal{W}) \equiv \|f_\perp\|_{\mathcal{W}}, \qquad f_\perp \equiv f - \frac{1\,\langle f,1\rangle}{\langle 1,1\rangle}.$$

Remark 3.9. Theorem 3.8 contains Theorem 3.1 as a special case when $\mathcal{W} = \mathcal{W}'$ is a reproducing kernel Hilbert space with reproducing kernel $K(x,y)$, and the bilinear function $\langle\cdot,\cdot\rangle$ is $\langle\cdot,\cdot\rangle_K$.

Remark 3.10. For Theorem 3.1 one only needs to know the reproducing kernel to compute the discrepancy. In contrast Theorem 3.8 seems more difficult to apply because one must specify two Banach spaces with their norms and $\langle\cdot,\cdot\rangle$. However, if one starts with a reproducing kernel Hilbert space satisfying the hypotheses of Theorem 3.1, one is often able to use $\langle\cdot,\cdot\rangle_K$ as the $\langle\cdot,\cdot\rangle$ in Theorem 3.8 along with the same $K(\cdot,\cdot)$. The definitions of the $\mathcal{W}$ and $\mathcal{W}'$ often suggest themselves in the same way as an $\mathcal{L}_p$ space generalizes an $\mathcal{L}_2$ space.

Theorem 2.4 can now be derived as a special case of Theorem 3.8. The appropriate bilinear function $\langle\cdot,\cdot\rangle$ is just $\langle\cdot,\cdot\rangle_{\mathcal{F},\gamma}$ as defined in (3.5). The relevant Banach spaces and norms are defined in terms of the ℓ_p-norm. For any non-negative number α, define the norm

$$\|f\|_{\mathcal{F},\alpha,p} \equiv \left\| (\bar{k}_1 \cdots \bar{k}_s)^\alpha \hat{f}(k) \right\|_p, \qquad 1 \le p \le \infty,$$

and the Banach space $\mathcal{W}_{\mathcal{F},\alpha,p} = \{f : \|f\|_{\mathcal{F},\alpha,p} < \infty\}$. If the space of integrands $\mathcal{W}$ in Theorem 3.8 is $\mathcal{W}_{\mathcal{F},\alpha,q}$, then the corresponding $\mathcal{W}'$ is $\mathcal{W}_{\mathcal{F},\gamma-\alpha,p}$, where $\frac{1}{p} + \frac{1}{q} = 1$. Condition (3.12) is satisfied by applying Hölder's inequality to $\langle\cdot,\cdot\rangle_{\mathcal{F},\gamma}$. The appropriate kernel in Theorem 3.8 is $K_{\mathcal{F},\gamma}$ as defined above in (3.6), so ξ_K in Theorem 3.8 is just $\xi_{\mathcal{F},\gamma}$ as defined in (3.8). The quadrature error bound is

$$\begin{aligned} |I(f) - Q(f;P)| &= |\langle f, \xi_{\mathcal{F},\gamma}\rangle_{\mathcal{F},\gamma}| \\ &= \left| \sum_{k\in\mathbf{Z}^s} \left\{ \hat{f}(k) \left[\frac{1}{N} \sum_{z\in P} e^{2\pi i k\cdot z} \right] \right\} \right| \\ &\le D_{\mathcal{F},\alpha,p}(P) V_{\mathcal{F},\alpha,q}(f) \quad \forall f \in \mathcal{W}_{\mathcal{F},\alpha,q}, \end{aligned} \tag{3.15}$$

where the corresponding discrepancy and variation are:

$$\begin{aligned} D_{\mathcal{F},\alpha,p}(P) &\equiv \|\xi_{\mathcal{F},\gamma}\|_{\mathcal{F},\gamma-\alpha,p} \\ &= \left\| 1_{\{k\ne 0\}} (\bar{k}_1 \cdots \bar{k}_s)^{-\alpha} \frac{1}{N} \sum_{z\in P} e^{2\pi i k\cdot z} \right\|_p, \end{aligned} \tag{3.16}$$

$$V_{\mathcal{F},\alpha,q}(f) \equiv \left\| 1_{\{k\neq 0\}} (\bar{k}_1 \cdots \bar{k}_s)^{\alpha} \hat{f}(k) \right\|_q, \quad 1 \le p \le \infty, \quad \frac{1}{p} + \frac{1}{q} = 1. \tag{3.17}$$

For $p < \infty$ there is a worst-case integrand, $\xi_{\mathcal{F},\alpha,p}$, that is, an integrand for which the above error bound is attained. Its Fourier coefficients satisfy the condition:

$$\hat{\xi}_{\mathcal{F},\alpha,p}(k) \left[\frac{1}{N} \sum_{z \in P} e^{2\pi i k \cdot z} \right] = 1_{\{k\neq 0\}} (\bar{k}_1 \cdots \bar{k}_s)^{-\alpha p} \left| \frac{1}{N} \sum_{z \in P} e^{2\pi i k \cdot z} \right|^p, \tag{3.18}$$

Remark 3.11. Error bound (3.15) may be obtained in a more straightforward manner by applying Hölder's inequality directly to (2.4). However, Theorem 3.8 also has many other interesting cases, some of which are explored in the next section.

Remark 3.12. The computational complexity of calculating the discrepancy, (3.16), is high unless one chooses $p = 2$. In this case one obtains $D_{\mathcal{F},\alpha,2} = D_{\mathcal{F},2\alpha}(P)$ (see (3.9)).

In the error bound above there should be some condition relating α and p to insure that the discrepancy is well-defined and that the space of integrands is a subset of absolutely convergent Fourier series. To derive this condition note that for $1 \le r \le p$, and $\alpha, \gamma \ge 0$ it follows by Hölder's inequality that

$$\begin{aligned} \|f\|_{\mathcal{F},\gamma,r} &= \left\| (\bar{k}_1 \cdots \bar{k}_s)^{\gamma} \hat{f}(k) \right\|_r \\ &\le \left\| (\bar{k}_1 \cdots \bar{k}_s)^{-\alpha} \right\|_{pr/(p-r)} \left\| (\bar{k}_1 \cdots \bar{k}_s)^{\alpha+\gamma} \hat{f}(k) \right\|_p \\ &= \left[1 + 2\zeta\left(\frac{\alpha p r}{p-r} \right) \right]^{s(p-r)/(rp)} \|f\|_{\mathcal{F},\alpha+\gamma,p}, \end{aligned}$$

provided that $\alpha p r/(p-r) > 1$, where ζ is the Riemann zeta function. Therefore,

$$\mathcal{W}_{\mathcal{F},\alpha+\gamma,p} \subseteq \mathcal{W}_{\mathcal{F},\gamma,r} \text{ for } \alpha > \frac{1}{r} - \frac{1}{p}.$$

By choosing $r = 1$ and $\gamma = 0$, it follows that $\mathcal{W}_{\mathcal{F},\alpha,p}$ is a subset of absolutely convergent Fourier series when $\alpha > 1 - \frac{1}{p}$. To insure that the two spaces $\mathcal{W}_{\mathcal{F},\alpha,q}$ and $\mathcal{W}_{\mathcal{F},\gamma-\alpha,p}$ contain only absolutely convergent Fourier series, the following constraint is made on α:

$$\frac{1}{p} = 1 - \frac{1}{q} < \alpha < \gamma - 1 + \frac{1}{p}. \tag{3.19}$$

Since error bound (3.15) does not depend explicitly on γ, one may always choose γ large enough so that the upper bound on α is satisfied.

If P is the node set of a shifted lattice $L+\Delta$, then the discrepancy above can be written in terms of $\mathcal{P}_\alpha(L)$. By Lemma 2.2 it follows that:

$$\begin{aligned} D_{\mathcal{F},\alpha,p}(P) &= \left\| 1_{\{k\neq 0, k\in L^\perp\}} (\bar{k}_1 \cdots \bar{k}_s)^{-\alpha} \right\|_p \\ &= \left\{ {\sum_{k\in L^\perp}}' (\bar{k}_1 \cdots \bar{k}_s)^{-p\alpha} \right\}^{1/p} = [\mathcal{P}_{p\alpha}(L)]^{1/p}, \quad 1 \le p < \infty, \\ D_{\mathcal{F},\alpha,\infty}(P) &= \left\| 1_{\{k\neq 0, k\in L^\perp\}} (\bar{k}_1 \cdots \bar{k}_s)^{-\alpha} \right\|_\infty = [\rho(L)]^{-\alpha}, \end{aligned}$$

where $\rho(L)$ is the Zaremba figure of merit for lattice rules [Nie92, Definition 5.31] defined as

$$\rho(L) = \min_{\substack{k\in L^\perp \\ k\neq 0}} \bar{k}_1 \cdots \bar{k}_s.$$

In fact one may consider $\rho(L)$ to be the limit of $[\mathcal{P}_p(L)]^{-1/p}$ as p tends to infinity. Error bound (3.15) when applied to lattice rules may be written as:

$$|I(f)-Q(f;L+\Delta)| \le \begin{cases} [\mathcal{P}_{p\alpha}(L)]^{1/p} \left\| 1_{\{k\neq 0\}} (\bar{k}_1 \cdots \bar{k}_s)^\alpha \hat{f}(k) \right\|_q, & 1 \le p < \infty, \\ & \frac{1}{p} + \frac{1}{q} = 1, \\ [\rho(L)]^{-\alpha} \left\| 1_{\{k\neq 0\}} (\bar{k}_1 \cdots \bar{k}_s)^\alpha \hat{f}(k) \right\|_1, & p = \infty. \end{cases}$$

The case $p = 1$, $q = \infty$ corresponds to Theorem 2.4.

Remark 3.13. For lattice rules the worst-case integrand $\xi_{\mathcal{F},\alpha,p}$, defined in (3.18), is just $\xi_{\mathcal{F},p\alpha}$, as defined in (3.8). Furthermore, $\xi_{\mathcal{F},\alpha,p}$ is similar to $\xi_{p\alpha}$ defined in (2.8).

Remark 3.14. According to restriction (3.19), α must be positive. However, $p = \infty$ and $\alpha = 0$ is allowable since $\mathcal{W}_{\mathcal{F},0,1}$ and $\mathcal{W}_{\mathcal{F},\gamma,\infty}$, $(\gamma > 1)$ contain only absolutely convergent Fourier series. In this case the discrepancy of the node set of a shifted lattice is one, no matter what the lattice is. So, for any lattice L, one may construct an integrand $f \in \mathcal{W}_{\mathcal{F},0,1}$ such that the quadrature error is the size of variation of the integrand, $V_{\mathcal{F},0,1}(f) = {\sum_{k\in\mathbf{Z}^s}}' |\hat{f}(k)|$. The construction is straightforward. Since lattice rules integrate Fourier coefficients with $k \notin L^\perp$ exactly, but give 100% error for Fourier coefficients with $k \in L^\perp$, one chooses f to have nonzero Fourier coefficients only for $k \in L^\perp$.

3.3 Average-Case Quadrature Error Analysis

The error bounds given in Theorems 2.4, 3.1, and 3.8 give an upper bound on the quadrature error. Another approach is to study the **average or**

mean error over some space of random integrands. In this average-case analysis the same discrepancy as defined above serves as a measure of the quality of the quadrature rule.

Theorem 3.15. *Suppose that $\mathcal{A}$ is a space of random functions such that*

$$E_{f\in\mathcal{A}}[f(x)] = 0, \quad E_{f\in\mathcal{A}}[f(x)f(y)] = K(x,y) \quad \forall x, y \in C^s,$$

for some given covariance kernel, K. Then the mean square quadrature error for quadrature rule (2.1) *is the discrepancy as defined in* (3.3)*:*

$$E_{f\in\mathcal{A}}\{[I(f) - Q(f;P)]^2\} = [D(P;K)]^2.$$

Proof. Proofs of this theorem or special cases are given in [Woź91, MC94, Rit95, HH98] and elsewhere. The conclusion follows by interchanging the order of the expectation and the integration or summation:

$$\begin{aligned}
&E_{f\in\mathcal{A}}\{[I(f) - Q(f;P)]^2\}\\
&= E_{f\in\mathcal{A}}\left[\int_{C^s\times C^s} f(x)f(y)\,dx\,dy\right.\\
&\qquad\left. -\frac{2}{N}\sum_{z\in P}\int_{C^s} f(z)f(y)\,dy + \frac{1}{N^2}\sum_{z,z'\in P} f(z)f(z')\right]\\
&= \int_{C^s\times C^s} K(x,y)\,dx\,dy - \frac{2}{N}\sum_{z\in P}\int_{C^s} K(z,y)\,dy + \frac{1}{N^2}\sum_{z,z'\in P} K(z,z')\\
&= [D(P)]^2.
\end{aligned}$$

□

Remark 3.16. It may seem odd that the same discrepancy acts as a measure of both average-case and worst-case quadrature error. The reason why this is possible is that the space of integrands for the average-case analysis, $\mathcal{A}$, is larger than the space of integrands for the worst-case analysis, $\mathcal{W}$.

Remark 3.17. The double integral over the reproducing/covariance kernel that appears in the definition of the discrepancy in (3.3) has interesting interpretations in both the worst-case and average-case error analyses. In the worst-case analysis the discrepancy is the norm of the linear functional corresponding to the quadrature error, i.e. $\|I - Q\|_K = D(P;K)$. The norm of the integration functional itself is:

$$\|I\|_K^2 = \int_{C^s\times C^s} K(x,y)\,dx\,dy.$$

This corresponds to the expected value of the square of the integral in the average-case analysis:

$$E_{f\in\mathcal{A}}\{[I(f)]^2\} = \int_{C^s\times C^s} K(x,y)\,dx\,dy.$$

The expected value of the variance of the function can also be written in terms of the covariance kernel:

$$E_{f\in\mathcal{A}}\{\mathrm{Var}(f)\} = E_{f\in\mathcal{A}}\{I(f^2) - [I(f)]^2\} = \int_{C^s} K(x,x)\,dx - \int_{C^s\times C^s} K(x,y)\,dx\,dy. \tag{3.20}$$

Because quasi-Monte Carlo methods, such as lattice rules, are intended to be more accurate than simple Monte Carlo, one might wonder how large the discrepancy is for a simple uniform random sample, P. Elementary calculations give the following result.

Theorem 3.18. *The mean square of the discrepancy defined in* (3.3) *for simple, uniform random samples, P is*

$$E_P[D(P)]^2 = \frac{1}{N}\left\{\int_{C^s} K(x,x)\,dx - \int_{C^s\times C^s} K(x,y)\,dx\,dy\right\} = \frac{1}{N}E_{f\in\mathcal{A}}\{\mathrm{Var}(f)\}.$$

This theorem implies that the discrepancy of a simple random sample is typically of $\mathrm{O}(N^{-1/2})$. For quasi-random points, such as lattices, it is expected to be much smaller.

4 Examples of Other Discrepancies

The previous section provided only one example of the error bounds and discrepancies that may be obtained from Theorems 3.1, 3.8, and 3.15. This section gives some more examples of error bounds and discrepancies that can be derived using these theorems. These examples include the popular star discrepancy and the related unanchored discrepancy.

4.1 The ANOVA Decomposition

The first example, however, is a generalization of the one given in the previous section. One observation made by several researchers is that the nominal dimension of an integration domain, s, may not be as important as some effective dimension of the integrand. Consider a possible integrand on C^4:

$$f(x_1,x_2,x_3,x_4) = 5x_1\cos(3\pi x_2) - 4e^{x_3}. \tag{4.1}$$

What is the effective dimension of this integrand? Only x_1, x_2 and x_3 appear, so one could argue that the effective dimension is three. On the other

hand f is a sum of two terms, each of which depends on only two variables, so one could also argue that the dimension is two. The analysis of variance (ANOVA) decomposition provides a way to look more closely at which variables and combinations of variables are important.

First, some notation must be introduced. Let $S = \{1, 2, \ldots, s\}$. For any set of coordinate indices $u \subseteq S$, let $|u|$ denote the number of elements in u. Let x_u denote the $|u|$-vector containing the components of x indexed by u, and let C^u denote the $|u|$-dimensional cube which is the projection of C^s into the coordinates indexed by u. Any function $f \in \mathcal{L}_2(C^s)$ may be written as the sum of its ANOVA effects, f_u, defined recursively as follows [ES81, Owe92, Hic96b]:

$$f(x) = \sum_{u \in S} f_u(x), \text{ where } f_u(x) = \int_{C^{S-u}} f(x)\, dx_{S-u} - \sum_{v \subset u} f_v(x). \tag{4.2}$$

(By convention $\int_{C^\emptyset} f\, dx_\emptyset = f(x)$.) Every function has a total of 2^s ANOVA effects. The ANOVA effect f_u is the part of the function depending only on the x_j with $j \in u$. The constant $f_\emptyset$ is the average value or integral of the function:

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

The main effect along the j^{th} axis is $f_{\{j\}}$, while $f_{\{j,m\}}$ is the interaction along the axes j and m. From the definition it can be shown that

$$\int_0^1 f_u(x_u)\, dx_j = 0 \text{ for any } j \in u. \tag{4.3}$$

The order of an effect f_u is $|u|$. The function f defined in (4.1) has ANOVA effects:

$$f_\emptyset = 4(1-e), \quad f_{\{2\}} = \frac{5}{2}\cos(3\pi x_2),$$
$$f_{\{3\}} = 4\left(e - 1 - e^{x_3}\right), \quad f_{\{1,2\}} = 5\left(x_1 - \frac{1}{2}\right)\cos(3\pi x_2),$$
$$f_{\{1\}} = f_{\{4\}} = f_{\{1,3\}} = f_{\{1,4\}} = f_{\{2,3\}} = f_{\{2,4\}} = f_{\{3,4\}} = 0,$$
$$f_{\{1,2,3\}} = f_{\{1,2,4\}} = f_{\{1,3,4\}} = f_{\{2,3,4\}} = f_{\{1,2,3,4\}} = 0.$$

Remark 4.1. One would expect that it is easier to approximate the integrals of low order ANOVA effects because they are essentially functions on low-dimensional cubes. This point is developed further in Section 8.

The notation $\|\cdot\|_p$ may correspond to either the $\mathcal{L}_p$-norm of a function or the ℓ_p-norm of a sequence, depending on the context. This notation is extended to the case of a vector of ANOVA effects (f_u) or their Fourier coefficients $(\hat{f}_u)$, where u is an index running over some or all of the subsets

of S, each f_u is a function on C^u, and each $\hat{f}_u$ is a function on $\mathbf{Z}^u$:

$$\|(f_u)\|_p = \left[\sum_u \|f_u\|_p^p\right]^{1/p}, \quad \|(\hat{f}_u)\|_p = \left[\sum_u \|\hat{f}_u\|_p^p\right]^{1/p}, \quad 1 \le p < \infty,$$
$$\|(f_u)\|_\infty = \max_u \|f_u\|_\infty, \quad \|(\hat{f}_u)\|_\infty = \max_u \|\hat{f}_u\|_\infty.$$

In the case where (f_u) is a vector of constants $\|(f_u)\|_p$ corresponds to the ℓ_p-norm, and when (f_u) is a single function $\|(f_u)\|_p$ corresponds to the $\mathcal{L}_p$-norm.

Now we take another look at the quadrature error, as given in (3.15), by writing it in terms of the ANOVA effects of the integrand and of $\xi_{\mathcal{F},\alpha}$. Let $\hat{f}_u(k_u)$ denote the Fourier coefficients of the ANOVA effect $f_u(x_u)$. From (4.2) and (4.3) one can show that

$$\hat{f}_u(k_u) = \begin{cases} \hat{f}(k_u) & \text{if } k_j \neq 0 \text{ for all } j \in u, \\ 0 & \textit{otherwise.} \end{cases}$$

This implies that $\langle f_u, g_v \rangle_{\mathcal{F},\alpha} = 0$ for any ANOVA effects f_u and g_v where $u \neq v$. Thus, error bound (3.15) may be written as:

$$\begin{aligned} (4.4) \qquad |I(f) - Q(f;P)| &= \left| \sum_{\emptyset \subset u \subseteq S} [I(f_u) - Q(f_u;P)] \right| \\ &= \left| \sum_{\emptyset \subset u \subseteq S} \langle f_u, \xi_{\mathcal{F},\alpha,u} \rangle_{\mathcal{F},\alpha} \right| \\ &\le \sum_{\emptyset \subset u \subseteq S} D_{\mathcal{F},\alpha,p,u}(P) V_{\mathcal{F},\alpha,q,u}(f), \quad \frac{1}{p} + \frac{1}{q} = 1, \end{aligned}$$

where

$$\begin{aligned} (4.5) \qquad D_{\mathcal{F},\alpha,p,u}(P) &\equiv D_{\mathcal{F},\alpha,p,u}(P_u) \equiv \|\xi_{\mathcal{F},\alpha,u}\|_{\mathcal{F},\gamma-\alpha,p} \\ &= \left\| \left[\prod_{j \in u} 1_{\{k_j \neq 0\}} \bar{k}_j^{-\alpha} \right] \frac{1}{N} \sum_{z \in P} e^{2\pi i k_u \cdot z_u} \right\|_p, \\ (4.6) \qquad V_{\mathcal{F},\alpha,q,u}(f) &\equiv V_{\mathcal{F},\alpha,q,u}(f_u) \equiv \|f_u\|_{\mathcal{F},\alpha,q} \\ &= \left\| \left[\prod_{j \in u} \bar{k}_j^{\alpha} \right] \hat{f}_u(k_u) \right\|_q, \end{aligned}$$

and P_u denotes the projection of P into C^u.

Remark 4.2. Consider the space of integrands:

$$\mathcal{W}_{\mathcal{F},\alpha,q,u} \equiv \left\{ f \in \mathcal{W}_{\mathcal{F},\alpha,q} : \int_0^1 f \, dx_j = 0 \text{ for all } j \in u \right\}.$$

This is a subspace of $\mathcal{W}_{\mathcal{F},\alpha,q}$, and in fact, is the space consisting of all ANOVA effects, f_u, of all $f \in \mathcal{W}_{\mathcal{F},\alpha,q}$. The quadrature error bound for integrands in $\mathcal{W}_{\mathcal{F},\alpha,q,u}$ involves the quantities defined in (4.5) and (4.6), specifically:

$$|I(f) - Q(f;P)| \le D_{\mathcal{F},\alpha,p,u}(P) V_{\mathcal{F},\alpha,q,u}(f) \quad \text{for all } f \in \mathcal{W}_{\mathcal{F},\alpha,q,u}.$$

Remark 4.3. Another viewpoint is that the $D_{\mathcal{F},\alpha,p,u}$ are pieces of the discrepancy $D_{\mathcal{F},\alpha,p,}$ and the $V_{\mathcal{F},\alpha,q,u}$ are pieces of the variation $V_{\mathcal{F},\alpha,q}$, because

$$D_{\mathcal{F},\alpha,p}(P) = \left\| (D_{\mathcal{F},\alpha,p,u}(P_u))_{u \ne \emptyset} \right\|_p, \tag{4.7a}$$

$$V_{\mathcal{F},\alpha,q}(f) = \left\| (V_{\mathcal{F},\alpha,q,u}(f_u))_{u \ne \emptyset} \right\|_q. \tag{4.7b}$$

By applying Hölder's inequality to the last line of (4.4) one then arrives back at (3.15)

When the quadrature error bound is written as (4.4), one can identify how each ANOVA effect contributes to the quadrature error. For the integrand f in (4.1), the only nonzero ANOVA effects, f_u, are for $u = \emptyset, \{2\}, \{3\}$ and $\{1,2\}$. Therefore, the variation terms $V_{\mathcal{F},\alpha,q,u}(f_u)$ are zero for all other $u \subseteq \{1,2,3,4\}$, and the only necessary discrepancy terms in the quadrature error bound are $D_{\mathcal{F},\alpha,p,u}(P_u)$ for $u = \{2\}, \{3\}$ and $\{1,2\}$. If the set P is evenly distributed along the unit interval in the second and third coordinates and in the unit square in the first and second coordinates, then the quadrature rule based on P will be accurate for this particular integrand f.

4.2 A Generalization of $\mathcal{P}_\alpha(L)$ with Weights

One may generalize the discrepancy and variation as written in (4.7) by introducing positive weights, β_u, $\emptyset \subset u \subseteq S$, specifically,

$$\begin{aligned} D_{\mathcal{F},\alpha,p}(P) &= \left\| (\beta_u D_{\mathcal{F},\alpha,p,u}(P_u))_{u \ne \emptyset} \right\|_p \\ &= \left\| \left(\beta_u \left[\prod_{j \in u} 1_{\{k_j \ne 0\}} |k_j|^{-\alpha} \right] \frac{1}{N} \sum_{z \in P} e^{2\pi i k_u \cdot z_u} \right)_{u \ne \emptyset} \right\|_p, \end{aligned} \tag{4.8a}$$

$$V_{\mathcal{F},\alpha,q}(f) = \left\| (\beta_u^{-1} V_{\mathcal{F},\alpha,q,u}(f_u))_{u \ne \emptyset} \right\|_q. \tag{4.8b}$$

The dependence of the discrepancy and variation on the weights is suppressed for ease of notation. Under these modified definitions one still has the quadrature error bound:

$$|I(f) - Q(f;P)| \le D_{\mathcal{F},\alpha,p}(P) V_{\mathcal{F},\alpha,q}(f). \tag{4.8c}$$

Weighted discrepancies were proposed by [Hic96b, Hic98], with $\beta_u = \beta^{|u|}$, and by [SW97] for the star discrepancy, with $\beta_u = \prod_{j\in u} \beta_j$.

It is natural to ask how best to choose the weights. To make the error bound as tight as possible, that is,

$$\sum_{\emptyset \subset u \subseteq S} D_{\mathcal{F},\alpha,p,u}(P_u) V_{\mathcal{F},\alpha,q,u}(f_u) = D_{\mathcal{F},\alpha,p}(P) V_{\mathcal{F},\alpha,q}(f),$$

one should choose the β_u such that

$$\beta_u [D_{\mathcal{F},\alpha,p,u}(P_u)]^{1/q} [V_{\mathcal{F},\alpha,q,u}(f_u)]^{-1/p} = \text{constant} \quad \forall \emptyset \subset u \subseteq S.$$

In practice this condition is difficult to satisfy exactly since it requires detailed knowledge about the integrand. However, this condition suggests a useful qualitative principle.

Remark 4.4. A set of points with low discrepancy $D_{\mathcal{F},\alpha,p}(P)$ must have low values for its components $D_{\mathcal{F},\alpha,p,u}(P_u)$ (see (4.8)), but the relative importance of each component depends on the weight β_u assigned to it. If one expects the variation of the ANOVA effect f_u to be relatively large, then one should make β_u relatively large according to the condition above. This gives heavier weight to the component of the discrepancy $D_{\mathcal{F},\alpha,p,u}(P_u)$, which implies that a low-discrepancy set must do better in integrating the ANOVA effect f_u.

For the weighted form of the discrepancy $D_{\mathcal{F},\alpha,p}(P)$ a convenient special case of β_u is

$$\beta_u = \beta_0 \prod_{j\in u} \beta_j^{\alpha}, \tag{4.9}$$

for some positive $\beta_0, \beta_1, \ldots, \beta_s$. Weights of this form are called **product-type weights**. For this type of weights $D_{\mathcal{F},\alpha,p}(P)$ becomes

$$D_{\mathcal{F},\alpha,p}(P) = \beta_0 \left\| \left[\overline{(\beta_1^{-1} k_1)} \cdots \overline{(\beta_s^{-1} k_s)}\right]^{-\alpha} \frac{1}{N} \sum_{z\in P} e^{2\pi i k\cdot z} \right\|_p, \tag{4.10}$$

where the definition of the notation $\bar{k}$ is given in (2.5).

The discrepancy, $D_{\mathcal{F},\alpha}(P)$ defined in (3.9) with product-type weights

now takes the form:

$$D_{\mathcal{F},\alpha}(P) = \beta_0 \left\{ \sum_{k\in\mathbf{Z}^s}{}' \left(\left[\overline{(\beta_1^{-1}k_1)} \cdots \overline{(\beta_s^{-1}k_s)} \right]^{-\alpha} \times \frac{1}{N^2} \sum_{z,z'\in P} e^{2\pi i k\cdot(z-z')} \right) \right\}^{1/2}. \tag{4.11}$$

The corresponding reproducing/covariance kernel originally defined in (3.6) becomes:

$$K_{\mathcal{F},\alpha}(x,y) = \beta_0^2 \sum_{k\in\mathbf{Z}^s} \left(e^{2\pi i k\cdot(x-y)} \left[\overline{(\beta_1^{-1}k_1)} \cdots \overline{(\beta_s^{-1}k_s)} \right]^{-\alpha} \right). \tag{4.12}$$

The relationship $D_{\mathcal{F},\alpha,p}(P) = D_{\mathcal{F},2\alpha}(P)$ continues to hold for these weighted discrepancies.

If $p = 2$ and α is a positive integer, then the above discrepancies can be written in terms of Bernoulli polynomials as was done in (3.11):

$$\begin{aligned} D_{\mathcal{F},\alpha,2}(P) &= D_{\mathcal{F},2\alpha}(P) \\ &= \tilde{\beta}_0 \left\{ -1 + \frac{1}{N^2} \sum_{z,z'\in P} \prod_{j=1}^{s} \left[1 - \frac{(-\tilde{\beta}_j^2)^\alpha}{(2\alpha)!} B_{2\alpha}(\{z_j - z_j'\}) \right] \right\}^{1/2}, \end{aligned} \tag{4.13}$$

where

$$\beta_0 = \tilde{\beta}_0, \quad \beta_j = \frac{\tilde{\beta}_j}{2\pi}, \ j = 1, \ldots, s. \tag{4.14}$$

The weighted form of $\mathcal{P}_\alpha(L)$ for the above product-type weights is defined as

$$\mathcal{P}_\alpha(L) = \beta_0 \sum_{k\in L^\perp}{}' \left[\overline{(\beta_1^{-1}k_1)} \cdots \overline{(\beta_s^{-1}k_s)} \right]^{-\alpha}.$$

This may also be written in terms of Bernoulli polynomials as

$$\mathcal{P}_{2\alpha}(L) = \tilde{\beta}_0 \left\{ -1 + \frac{1}{N} \sum_{z\in P} \prod_{j=1}^{s} \left[1 - \frac{(-\tilde{\beta}_j^2)^\alpha}{(2\alpha)!} B_{2\alpha}(z_j) \right] \right\}^{1/2}, \tag{4.15}$$

for integers α, where P is the node set of the lattice L. The relationship between the weighted form of $\mathcal{P}_\alpha(L)$ and the weighted forms of the discrepancies $D_{\mathcal{F},\alpha,p}(P)$ and $D_{\mathcal{F},\alpha}(P)$ for lattices is

$$\mathcal{P}_\alpha(L) = \beta_0^{-1}[D_{\mathcal{F},\alpha}(P)]^2 = \beta_0^{1-p}[D_{\mathcal{F},\alpha/p,p}(P)]^p, \tag{4.16}$$

for any $\alpha > 1$ and any $1 \le p < +\infty$.

Remark 4.5. There is always an implicit, if not explicit, choice of weights for the discrepancies $D_{\mathcal{F},\alpha,p}(P)$, $D_{\mathcal{F},\alpha}(P)$ and $\mathcal{P}_\alpha(L)$ and the other weighted discrepancies introduced below. If one prefers the simplicity of the original definition of $\mathcal{P}_\alpha(L)$ in (2.7), then one would choose $\beta_u = 1$. However, if one prefers the formulas for discrepancy in terms of Bernoulli polynomials (see (3.11) and (4.13)) to be elegant, then one would choose the product-type weights $\beta_u = (2\pi)^{-\alpha|u|}$. It is difficult to argue for one "natural" choice of weights. It is better to realize that regardless which choice is made, the relative values of the weights β_u reflect one's assumptions about the relative sizes of the variations of the ANOVA effects f_u of the integrands one expects to encounter as discussed in Remark 4.4.

The space $\mathcal{W}_{\mathcal{F},\alpha,p}$ consists of functions, f, with finite variation, $V_{\mathcal{F},\alpha,p}(f)$. If one writes the weights in (4.8) as $\beta_u = (2\pi)^{-\alpha|u|}\tilde{\beta}_u$, then the definition of the variation may be written as:

$$\begin{aligned} V_{\mathcal{F},\alpha,p}(f) &= \left\| \left(\tilde{\beta}_u^{-1} \left\{ \prod_{j\in u} \left[2\pi \bar{k}_j\right]^\alpha \right\} \hat{f}_u(k) \right)_{\emptyset\subset u\subseteq S} \right\|_p \\ &= \left\| \left(\tilde{\beta}_u^{-1} \mathcal{F}\left\{ \frac{\partial^{\alpha|u|} f_u}{\partial x_u^\alpha} \right\}(k_u) \right)_{\emptyset\subset u\subseteq S} \right\|_p , \end{aligned}$$

for integer α. Here $\mathcal{F}$ denotes the Fourier transform as defined in (2.3). If $p = 2$ this norm may also be written without referring to the Fourier coefficients:

$$V_{\mathcal{F},\alpha,2}(f) = \left\{ \sum_{\emptyset\subset u\subseteq S} \tilde{\beta}_u^{-2} \left\| \frac{\partial^{\alpha|u|} f_u}{\partial x_u^\alpha} \right\|_2^2 \right\}^{1/2} .$$

Therefore, $\mathcal{W}_{\mathcal{F},\alpha,2}$ consists of all periodic integrands where the α-order mixed partial derivatives of the ANOVA effects are square integrable.

4.3 The Periodic Bernoulli Discrepancy — Another Generalization of $\mathcal{P}_\alpha(L)$

A family of error bounds and discrepancies that are related to $\mathcal{P}_\alpha$ is given in [Hic96b, Hic98]. These may be derived from Theorem 3.8 by using the ANOVA decomposition. Let $\mathcal{W}_{\mathrm{per},\alpha,p}$ be a space of periodic functions whose first α mixed partial derivatives are in $\mathcal{L}_p(C^s)$, that is, let

$$\mathcal{W}_{\mathrm{per},\alpha,p} \equiv \left\{ f : \frac{\partial^{\alpha|u|} f}{\partial x_u^\alpha} \in \mathcal{L}_p(C^u) \text{ and } \int_0^1 \frac{\partial^{\gamma|u|} f}{\partial x_u^\gamma}\, dx_j = 0 \right. \\ \left. \text{for all } \gamma \le \alpha,\ j \in u,\ \emptyset \subset u \subseteq S \right\}. \tag{4.17a}$$

Then one has the quadrature error bound:

$$|I(f) - Q(f;P)| \le \sum_{\emptyset \subset u \subseteq S} D_{\mathrm{per},\alpha,p,u}(P) V_{\mathrm{per},\alpha,q,u}(f)$$
$$\le D_{\mathrm{per},\alpha,p}(P) V_{\mathrm{per},\alpha,q}(f), \quad \forall f \in \mathcal{W}_{\mathrm{per},\alpha,p}, \ \frac{1}{p} + \frac{1}{q} = 1,$$

where the discrepancy and variation are defined as:

$$V_{\mathrm{per},\alpha,p,u}(f) \equiv \left\| \frac{\partial^{\alpha|u|} f_u}{\partial x_u^{\alpha}} \right\|_p, \tag{4.17b}$$

$$V_{\mathrm{per},\alpha,p}(f) \equiv \left\| \left(\tilde{\beta}_u^{-1} V_{\mathrm{per},\alpha,p,u}(f) \right)_{\emptyset \subset u \subseteq S} \right\|_p, \tag{4.17c}$$

$$D_{\mathrm{per},\alpha,p,u}(P) \equiv \left\| \frac{1}{N} \sum_{z \in P} \prod_{j \in u} \left[\frac{(-1)^{\alpha+1}}{\alpha!} B_{\alpha}(\{x_j - z_j\}) \right] \right\|_p, \tag{4.17d}$$

$$D_{\mathrm{per},\alpha,p}(P) \equiv \left\| \left(\tilde{\beta}_u D_{\mathrm{per},\alpha,p,u}(P) \right)_{\emptyset \subset u \subseteq S} \right\|_p. \tag{4.17e}$$

Because the definition of the discrepancy involves Bernoulli polynomials, it can be called a periodic Bernoulli discrepancy. The error bound, discrepancy and variation are equivalent to those in (4.8) for $p = q = 2$, but they differ for other values of p and q. For $p = q = 2$ and for product-type weights the discrepancy is the same as (4.13).

4.4 The Non-Periodic Bernoulli Discrepancy

The discrepancies derived so far appear in error bounds for *periodic* integrands. Although lattice rules are particularly suited for such integrands, they can perform satisfactorily for non-periodic integrands. Extensions of the periodic Bernoulli discrepancy and the corresponding error bounds to spaces of non-periodic integrands are derived by the author [Hic96b, Hic98, HH98] on the basis of Theorems 3.1, 3.8, and 3.15. These make use of the ANOVA decomposition introduced above. They are introduced briefly below.

As a generalization to (4.17) for $\alpha = 1$, let

$$\mathcal{W}_{\mathrm{non},1,p} \equiv \left\{ f : \frac{\partial^{|u|} f}{\partial x_u} \in \mathcal{L}_p(C^u) \ \forall u \subseteq S \right\},$$
$$V_{\mathrm{non},1,p,u}(f) \equiv \left\| \frac{\partial^{|u|} f_u}{\partial x_u} \right\|_p,$$
$$V_{\mathrm{non},1,p}(f) \equiv \left\| \left(\hat{\beta}_u^{-1} V_{\mathrm{non},1,p,u}(f) \right)_{\emptyset \subset u \subseteq S} \right\|_p,$$

$$D_{\mathrm{non},1,p,u}(P) \equiv \left\| \frac{1}{N} \sum_{z \in P} \prod_{j \in u} [x_j - 1_{\{x_j > z_j\}}] \right\|_p ,$$

$$D_{\mathrm{non},1,p}(P) \equiv \left\| \left(\hat{\beta}_u D_{\mathrm{non},1,p,u}(P) \right)_{\emptyset \subset u \subseteq S} \right\|_p .$$

Then one has the quadrature error bound:

$$\begin{aligned} |I(f) - Q(f;P)| &\leq \sum_{\emptyset \subset u \subseteq S} D_{\mathrm{non},1,p,u}(P) V_{\mathrm{non},1,q,u}(f) \\ &\leq D_{\mathrm{non},1,p}(P) V_{\mathrm{non},1,q}(f) \quad \forall f \in \mathcal{W}_{\mathrm{non},1,q}, \ \frac{1}{p} + \frac{1}{q} = 1. \end{aligned}$$

For integers $\alpha > 1$ the generalization for all p is tedious, but the generalization for $p = 2$ with product-type weights is manageable. The reproducing/covariance kernel and corresponding $\mathcal{L}_2$-non-periodic Bernoulli discrepancy are:

$$K_{\mathrm{non},\alpha}(x,y) = \hat{\beta}_0^2 \prod_{j=1}^{s} \left[-\frac{(-\hat{\beta}_j^2)^\alpha}{(2\alpha)!} B_{2\alpha}(\{x_j - y_j\}) + \sum_{i=0}^{\alpha} \frac{\hat{\beta}_j^{2i}}{(i!)^2} B_i(x_j) B_i(y_j) \right],$$

$$D_{\mathrm{non},\alpha}(P) \equiv \hat{\beta}_0 \left\{ -1 + \frac{1}{N^2} \sum_{z,z' \in P} \prod_{j=1}^{s} \left[-\frac{(-\hat{\beta}_j^2)^\alpha}{(2\alpha)!} B_{2\alpha}(\{z_j - z'_j\}) + \sum_{i=0}^{\alpha} \frac{\hat{\beta}_j^{2i}}{(i!)^2} B_i(z_j) B_i(z'_j) \right] \right\}^{1/2}.$$

This reproducing/covariance kernel has been used by Wahba [Wah90, Section 10.2] and others.

4.5 The Star Discrepancy

The Koksma-Hlawka inequality [Hla61], [Nie92, Theorem 2.11] is perhaps the most well-known and popular multidimensional quadrature error bound. The original inequality is the case $p = \infty$ below, but it has been extended by [Zar68] (for $p = 2$), [Sob69, Chapter 8] (for $1 \leq p \leq \infty$) and [SW97] (by introducing weights). The following error bound may be derived by using

Theorem 3.8 (see [Hic98]):

$$\begin{aligned}|I(f) - Q(f;P)| &\le \sum_{\emptyset \subset u \subseteq S} D^*_{p,u}(P) V^*_{q,u}(f) \\ &\le D^*_p(P) V^*_q(f) \quad \forall f \in \mathcal{W}_{\text{non},1,q},\ \frac{1}{p} + \frac{1}{q} = 1,\end{aligned}$$

where,

$$\begin{aligned}V^*_{p,u}(f) &\equiv \left\| \left. \frac{\partial^{|u|} f}{\partial x_u} \right|_{x_{S-u} = (1,\ldots,1)} \right\|_p, \\ V^*_p(f) &\equiv \left\| \left(\hat{\beta}_u^{-1} V^*_{p,u}(f) \right)_{u \ne \emptyset} \right\|_p, \\ D^*_{p,u}(P) &\equiv \left\| \prod_{j \in u} x_j - \frac{1}{N} \sum_{z \in P} 1_{\{x_u > z_u\}} \right\|_p, \\ D^*_p(P) &\equiv \left\| \left(\hat{\beta}_u D^*_{p,u}(P) \right)_{u \ne \emptyset} \right\|_p.\end{aligned}$$

Here $x_u > z_u$ means that $x_j > z_j$ for all $j \in u$. The variation $V_1(f)$ with unit weights corresponds to the variation in the sense of Hardy and Krause.

The piece of the discrepancy $D^*_{p,u}(P)$ is the $\mathcal{L}_p$-norm of the volume of the rectangular solid $[0, x_u)$ minus the proportion of points in P that are also in $[0, x_u)$. (See Figure 4.1 for a graphical representation of the star discrepancy.) In fact, $D^*_{p,S}(P)$ is often defined as the $\mathcal{L}_p$-star discrepancy. However, such a definition has the disadvantage that the corresponding error bound,

$$|I(f) - Q(f;P)| \le D^*_{p,S}(P) V^*_{q,S}(f), \quad \frac{1}{p} + \frac{1}{q} = 1,$$

applies only to integrands which vanish for any $x_j = 1$. For unit weights $D^*_{p,S}(P) < D^*_p(P)$ in general, but $D^*_{\infty,S}(P) = D^*_\infty(P)$.

The $\mathcal{L}_2$-star discrepancy is a special case of Theorems 3.1 and 3.15. The discrepancy and the corresponding kernel on which it is based are given below for the case of product-type weights:

$$\begin{aligned}K^*(x,y) &= \hat{\beta}_0^2 \prod_{j=1}^s \left\{ 1 + \hat{\beta}_j^2 [1 - \max(x_j, y_j)] \right\}, \\ D^*_2(P) &= \hat{\beta}_0 \left\{ \prod_{j=1}^s \left(1 + \frac{\hat{\beta}_j^2}{3} \right) - \frac{2}{N} \sum_{z \in P} \prod_{j=1}^s \left[1 + \frac{\hat{\beta}_j^2 (1 - z_j^2)}{2} \right] \right. \\ &\qquad \left. + \frac{1}{N^2} \sum_{z, z' \in P} \prod_{j=1}^s \left(1 + \hat{\beta}_j^2 [1 - \max(z_j, z'_j)] \right) \right\}^{1/2}.\end{aligned}$$

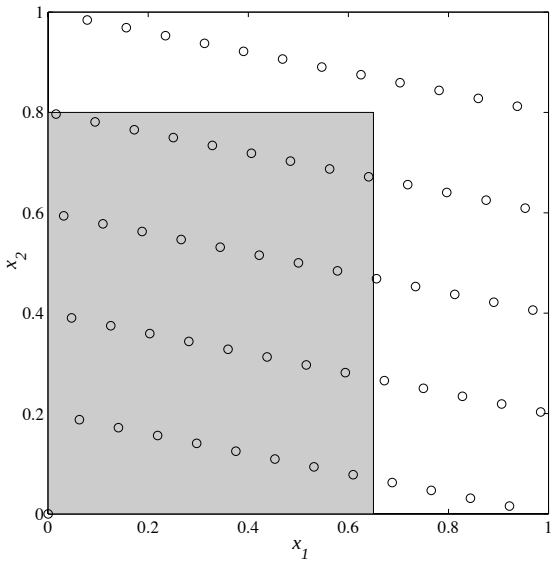

FIGURE 4.1. For $x_{\{1,2\}} = (0.65, 0.8)$ the volume of the box is 0.52, and the proportion of points in the box is $34/64 = 0.53125$. The difference between these two is $x_1 x_2 - \frac{1}{64} \sum_{z \in P} 1_{\{x_{\{1,2\}} > z_{\{1,2\}}\}} = -0.01125$.

4.6 The Unanchored Discrepancy

Although the star discrepancy enjoys widespread popularity as a figure of merit for quasi-random points, it lacks the following property:

Definition 4.6. A discrepancy, $D(P)$, is said to be **reflection invariant** if and only if $D(P) = D(P')$, where P' is obtained by reflecting the points in P about any plane $x_j = 1/2$ passing through the center of the cube, that is

$$P' = \{(z_1, \ldots, z_{j-1}, 1 - z_j, z_{j+1}, \ldots, z_s) : (z_1, \ldots, z_s) \in P\}.$$

The star discrepancy is not reflection invariant because it counts points in a box *anchored at the origin*. One modification of the star discrepancy that is reflection invariant is the unanchored discrepancy. This discrepancy, defined in [Nie92, Definition 2.2] and [MC94], counts points in an unanchored box $[x'_u, x_u) \subseteq C^u$ with $x'_u < x_u$ (see Figure 4.2). In the formula below $|A|$ is the number of points in a set A. The corresponding error bound for the $\mathcal{L}_2$-unanchored discrepancy can be found using Theorem 3.1, and is given as follows:

$$\begin{aligned} |I(f) - Q(f;P)| &\le \sum_{\emptyset \subset u \subseteq S} D_{\mathrm{un},u}(P) V_{\mathrm{un},u}(f) \\ &\le D_{\mathrm{un}}(P) V_{\mathrm{un}}(f) \quad \forall f \in \mathcal{W}_{\mathrm{per},1,2}, \end{aligned}$$

where,

$$
\begin{aligned}
V_{\mathrm{un},u}(f) &\equiv \left\| \left. \frac{\partial^{|u|} f}{\partial x_u} \right|_{x_{S-u}=(1,\ldots,1)} \right\|_2, \\
V_{\mathrm{un}}(f) &\equiv \left\| \left(\hat{\beta}_u^{-1} V_{\mathrm{un},u}(f) \right)_{u \neq \emptyset} \right\|_2, \\
D_{\mathrm{un},u}(P) &\equiv \left\{ \int_{C^{2s}, x' < x} \left[\prod_{j \in u} (x_j - x_j') - \frac{|P_u \cap [x_u', x_u)|}{N} \right]^2 dx' dx \right\}^{1/2}, \\
D_{\mathrm{un}}(P) &\equiv \left\| \left(\hat{\beta}_u D_{\mathrm{un},u}(P) \right)_{u \neq \emptyset} \right\|_2.
\end{aligned}
$$

For product-type weights the discrepancy and its kernel are:

$$
\begin{aligned}
K_{\mathrm{un}}(x,y) &= \hat{\beta}_0^2 \prod_{j=1}^{s} \left\{ 1 + \hat{\beta}_j^2 [\min(x_j, y_j) - x_j y_j] \right\}, \\
D_{\mathrm{un}}(P) &= \hat{\beta}_0 \left\{ \prod_{j=1}^{s} \left(1 + \frac{\hat{\beta}_j^2}{12} \right) - \frac{2}{N} \sum_{z \in P} \prod_{j=1}^{s} \left[1 + \frac{\hat{\beta}_j^2 z_j (1 - z_j)}{2} \right] \right. \\
&\qquad \left. + \frac{1}{N^2} \sum_{z, z' \in P} \prod_{j=1}^{s} \left(1 + \hat{\beta}_j^2 [\min(z_j, z_j') - z_j z_j'] \right) \right\}^{1/2}.
\end{aligned}
$$

Remark 4.7. As is the case for the star discrepancy, the unanchored discrepancy is and ought to be defined by looking at boxes not only in C^s, but in all possible lower dimensional unit cubes $C^u, \emptyset \subset u \subseteq S$. Otherwise, the corresponding quadrature error analyses apply only to integrands that vanish on all faces of the cube C^s.

Remark 4.8. The worst-case and average-case error analyses for the unanchored discrepancy apply to *periodic* integrands. This follows from the fact that the reproducing/covariance kernel satisfies $K_{\mathrm{un}}(x,y)|_{x_j=0} = K_{\mathrm{un}}(x,y)|_{x_j=1}$ for all $j \in S$.

4.7 The Wrap-Around Discrepancy

A variation on the unanchored discrepancy that is also reflection-invariant arises by counting points in a "box" $[x_u', x_u)$, allowing for wrap-around

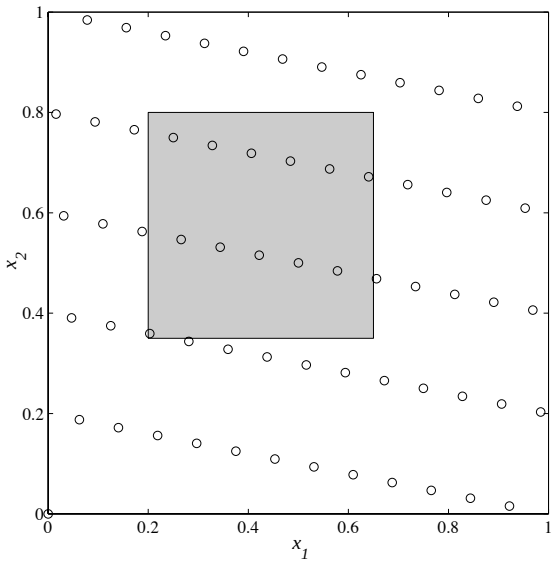

FIGURE 4.2. For $x'_{\{1,2\}} = (0.2, 0.35)$ and $x_{\{1,2\}} = (0.65, 0.8)$ the volume of the box is 0.2025, and the proportion of points in the box is $12/64 = 0.1875$. The difference between these two is $(x_1 - x'_1)(x_2 - x'_2) - \frac{1}{64}|P_{\{1,2\}} \cap [x'_{\{1,2\}}, x_{\{1,2\}})| = 0.015$.

when $x'_j > x_j$ (see Figure 4.3). Specifically, let

$$J_{\text{wrap}}(x'_j, x_j) = \begin{cases} [x'_j, x_j) & x'_j < x_j \\ [0, x_j) \cup [x'_j, 1) & x_j \le x'_j, \end{cases}$$

$$J_{\text{wrap}}(x'_u, x_u) = \bigotimes_{j \in u} J_{\text{wrap}}(x'_j, x_j).$$

The volume of $J_{\text{wrap}}(x'_u, x_u)$ is $\prod_{j \in u} \{x_j - x'_j\}$. The $\mathcal{L}_2$-wrap-around discrepancy is then defined analogously to the unanchored discrepancy. Assuming product-type weights the resulting discrepancy is:

$$D_{\text{wrap},u}(P) \equiv \left\{ \int_{C^{2s}} \left[\prod_{j \in u} \{x_j - x'_j\} - \frac{|P_u \cap J_{\text{wrap}}(x'_u, x_u)|}{N} \right]^2 dx' dx \right\}^{1/2},$$

$$D_{\text{wrap}}(P) \equiv \hat{\beta}_0 \left\| \left(\left[\prod_{j \in u} \hat{\beta}_j \right] D_{\text{wrap},u}(P) \right)_{u \neq \emptyset} \right\|_2$$

$$= \hat{\beta}_0 \left\{ -\prod_{j=1}^{s} \left(1 + \frac{\hat{\beta}_j^2}{3} \right) \right.$$

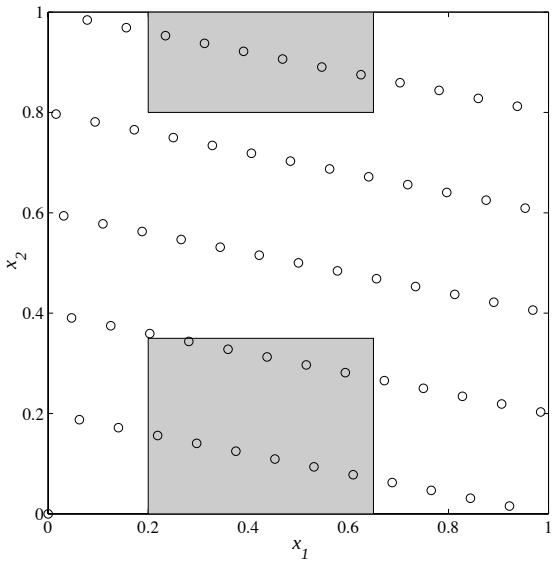

FIGURE 4.3. For $x'_{\{1,2\}} = (0.2, 0.8)$ and $x_{\{1,2\}} = (0.65, 0.35)$ the volume of the wrapped-around box is 0.2475, and the proportion of points in the box is $17/64 = 0.265625$. The difference between these two is $\{x_1 - x'_1\}\{x_2 - x'_2\} - \frac{1}{64}|P_{\{1,2\}} \cap J_{\text{wrap}}(x'_{\{1,2\}}, x_{\{1,2\}})| = -0.018125$.

$$+\frac{1}{N^2}\sum_{z,z'\in P}\prod_{j=1}^{s}\left(1+\hat{\beta}_j^2\left[\frac{1}{3}+B_2(\{z_j-z'_j\})\right]\right)\Bigg\}^{1/2},$$

$$K_{\text{wrap}}(x,y) = \hat{\beta}_0^2\prod_{j=1}^{s}\left\{1+\hat{\beta}_j^2\left[\frac{1}{3}+B_2(\{x_j-y_j\})\right]\right\}.$$

Remark 4.9. This discrepancy is the same as the weighted definition of $D_{\mathcal{F},1,2}(P) = D_{\mathcal{F},2}(P)$ in (4.13) if one chooses

$$\tilde{\beta}_0^2 = \hat{\beta}_0^2\prod_{j=1^s}\left(1+\frac{\hat{\beta}_j^2}{3}\right), \qquad \tilde{\beta}_j^2 = \frac{2\hat{\beta}_j^2}{1+\hat{\beta}_j^2/3}.$$

Recall, too, from (4.16) that $D_{\mathcal{F},2}(P)$ is equivalent to $\sqrt{\mathcal{P}_2(L)}$ if P is the node set of a lattice. Thus, the wrap-around discrepancy provides a geometric interpretation for a weighted version of $\mathcal{P}_2(L)$.

4.8 The Symmetric Discrepancy

Another reflection invariant discrepancy is the symmetric discrepancy [Hic98, Section 5.3]. Let $\{0,1\}^s$ denote the vertices of the unit cube C^s. Any point

x_u, divides the cube C^u into $2^{|u|}$ rectangular solids, each being defined as the region between a vertex $a_u \in \{0,1\}^u$ and x_u (see Figure 4.4). Let $J_e(x_u)$ be the union of the "even" rectangular solids, that is those touching a vertex a_u where $\sum_{j\in u} a_j$ is an even number. The symmetric discrepancy is defined by comparing the proportion of points in $P_u \cap J_e(x_u)$, to the volume of $J_e(x_u)$. (An equivalent definition would be obtained by looking at odd rather than even solids.) The associated quadrature error bound is:

$$\begin{aligned}|I(f) - Q(f;P)| &\le \sum_{\emptyset\subset u\subseteq S} D_{\mathrm{sym},p,u}(P) V_{\mathrm{sym},q,u}(f) \\ &\le D_{\mathrm{sym},p}(P) V_{\mathrm{sym},q}(f) \quad \forall f \in \mathcal{W}_{\mathrm{non},1,q},\ \frac{1}{p}+\frac{1}{q}=1,\end{aligned}$$

where,

$$\begin{aligned}V_{\mathrm{sym},p,u}(f) &\equiv \left\| \frac{1}{2^s} \sum_{a_{S-u}\in\{0,1\}^{S-u}} \left.\frac{\partial^{|u|} f}{\partial x_u}\right|_{x_{S-u}=a_{S-u}} \right\|_p, \\ V_{\mathrm{sym},p}(f) &\equiv \left\| \left(\hat\beta_u^{-1} V_{\mathrm{sym},p,u}(f)\right)_{u\neq\emptyset} \right\|_p, \\ D_{\mathrm{sym},p,u}(P) &\equiv 2 \left\| \mathrm{Vol}\left(J_e(x_u)\right) - \frac{|P_u \cap J_e(x_u)|}{N} \right\|_p, \\ D_{\mathrm{sym},p}(P) &\equiv \left\| \left(\hat\beta_u D_{\mathrm{sym},p,u}(P)\right)_{u\neq\emptyset} \right\|_p.\end{aligned}$$

For unit weights $\hat\beta_u$ it was shown in [Hic98] that $D_{\mathrm{sym},\infty}(P) = D_{\mathrm{sym},\infty,S}(P)$, as is the case for the star discrepancy.

The $\mathcal{L}_2$-symmetric discrepancy and its defining kernel are given below for the case of product-type weights:

$$\begin{aligned}K_{\mathrm{sym}}(x,y) &= \hat\beta_0^2 \prod_{j=1}^{s} [1 + \hat\beta_j^2 (1 - |x_j - y_j|)], \\ D_{\mathrm{sym},2}(P) &= \hat\beta_0 \left\{ \prod_{j=1}^{s} \left(1 + \frac{\hat\beta_j^2}{3}\right) - \frac{2}{N} \sum_{z\in P} \prod_{j=1}^{s} [1 + \hat\beta_j^2 (2z_j - 2z_j^2)] \right. \\ &\qquad \left. + \frac{1}{N^2} \sum_{z,z'\in P} \prod_{j=1}^{s} [1 + \hat\beta_j^2 (1 - |z_j - z_j'|)] \right\}^{1/2}.\end{aligned}$$

5 Shift-Invariant Kernels and Discrepancies

The previous section provides several examples of discrepancies for measuring the quality of quasi-random point sets and the associated quadrature

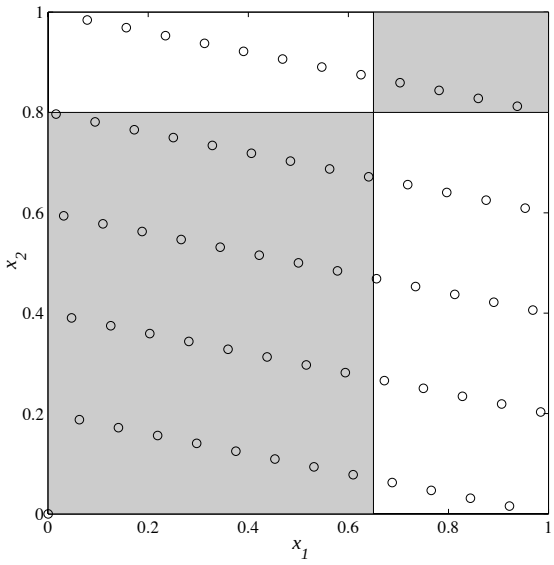

FIGURE 4.4. For $x_{\{1,2\}} = (0.65, 0.8)$ the total volume of the even rectangles is 0.59, and the proportion of points in these rectangles is $38/64 = 0.59375$. The difference between these two is $\text{Vol}\left(J_e(x_{\{1,2\}})\right) - \frac{1}{64}|P_{\{1,2\}} \cap J_e(x_{\{1,2\}})| = -0.00375$.

rules. This presents the difficulty of knowing which discrepancy to use in practice. This section shows how nearly all of the discrepancies described in the previous section are in some way related to $D_{\mathcal{F},\alpha}(P)$, which corresponds to $\mathcal{P}_\alpha(L)$ for lattices by (4.16).

The discrepancy $D_{\mathcal{F},\alpha}(P)$, defined in (3.9) and (4.11), is based on the a kernel whose value remains the same when an arbitrary shift (modulo 1) is added to both of the arguments simultaneously, that is,

$$K(\{x+\Delta\},\{y+\Delta\}) = K(x,y) \quad \text{for all } x,y,\Delta \in C^s.$$

This **shift-invariance** condition may be written equivalently as

$$K(x,y) = K(\{x-y\},0) \quad \text{for all } x,y \in C^s. \tag{5.1}$$

A discrepancy based on a shift-invariant kernel is called a **shift-invariant discrepancy**. The formula for such a discrepancy is simpler than the original definition (3.3), and further simplifications can be made for the case of lattices.

Theorem 5.1. *If a kernel, K, satisfies the shift-invariance condition* (5.1)*, then the discrepancy as defined in* (3.3) *may be written as:*

$$D(P;K) = \left\{ -\int_{C^s} K(x,0)\, dx + \frac{1}{N^2} \sum_{z,z' \in P} K(\{z-z'\},0) \right\}^{1/2}.$$

Furthermore, if P is a node set of a shifted lattice $L+\Delta$, then this formula for the discrepancy may be further simplified as:

$$D(P;K) = \left\{ -\int_{C^s} K(x,0)\ dx + \frac{1}{N}\sum_{z\in P} K(\{z-\Delta\},0) \right\}^{1/2}. \tag{5.2}$$

Proof. The first part of the theorem follows directly by substituting condition (5.1) into (3.3) and performing a change of integration variable. The second part of the theorem follows from the first part by noting that the difference of any two node points of the shifted lattice, $\{z-z'\}$, must be some node point of the unshifted lattice. □

Remark 5.2. This theorem shows that the shift-invariant discrepancy of a set P is unchanged when one adds a shift Δ (modulo 1) to every point.

Remark 5.3. Because (5.2) contains only a single sum, it requires only $O(N)$ operations to evaluate, as opposed to the general discrepancy as defined by (3.3), which requires $O(N^2)$ operations to evaluate.

It might seem that a shift-invariant discrepancy is a rather special case. However, suppose one considers random shifts of a set P. Then according to the theorem below the mean square discrepancy defined by an arbitrary kernel, is the square discrepancy of P defined by a suitable shift-invariant kernel. The proof of this theorem follows from the definition of discrepancy and the shift-invariance condition.

Theorem 5.4. *Given an arbitrary kernel, K, define an associated shift-invariant kernel, K_{shift}, as follows:*

$$K_{shift}(x,y) = \int_{C^s} K(\{x+\Delta\},\{y+\Delta\})\ d\Delta. \tag{5.3}$$

For a given finite set of points $P \subset C^s$, and any $\Delta \in C^s$, let $P+\Delta \equiv \{\{z+\Delta\} : z \in P\}$ denote a shifted copy of P. If Δ is a random variable uniformly distributed on C^s, then

$$E_\Delta\{[D(P+\Delta;K)]^2\} = [D(P;K_{shift})]^2.$$

The discrepancy based on K_{shift} is called the associated shift-invariant discrepancy.

This theorem implies that $D(P;K_{\text{shift}})$ gives an upper bound on the value of $D(P+\Delta;K)$ that is possible by a good choice of Δ. Many different reproducing kernels have the same associated shift-invariant kernel, as is demonstrated below. Thus, knowing about the shift-invariant discrepancy can provide useful information about many different discrepancies. When one is interested in lattice rules, there is the added benefit that the computational complexity of evaluating a shift-invariant discrepancy reduces from $O(N^2)$ to $O(N)$.

Theorem 5.5. *The kernels defining the $\mathcal{L}_2$-non-periodic Bernoulli ($\alpha = 1$), $\mathcal{L}_2$-star, $\mathcal{L}_2$-unanchored, and $\mathcal{L}_2$-symmetric discrepancies with product-type weights all have $K_{\mathcal{F},\alpha}$ as their associated shift-invariant kernel, assuming the following choices of weights:*

$$\begin{aligned}
\mathcal{L}_2\text{-non-periodic Bernoulli } (\alpha = 1)\text{:} \quad & \tilde\beta_0^2 = \hat\beta_0^2, \quad \tilde\beta_j^2 = 2\hat\beta_j^2,\\
\mathcal{L}_2\text{-star:} \quad & \tilde\beta_0^2 = \hat\beta_0^2 \prod_{j=1}^{s} \left(1 + \frac{\hat\beta_j^2}{3}\right), \quad \tilde\beta_j^2 = \frac{2\hat\beta_j^2}{1 + \hat\beta_j^2/3},\\
\mathcal{L}_2\text{-unanchored:} \quad & \tilde\beta_0^2 = \hat\beta_0^2 \prod_{j=1}^{s} \left(1 + \frac{\hat\beta_j^2}{12}\right), \quad \tilde\beta_j^2 = \frac{\hat\beta_j^2}{1 + \hat\beta_j^2/12},\\
\mathcal{L}_2\text{-symmetric:} \quad & \tilde\beta_0^2 = \hat\beta_0^2 \prod_{j=1}^{s} \left(1 + \frac{2\hat\beta_j^2}{3}\right), \quad \tilde\beta_j^2 = \frac{4\hat\beta_j^2}{1 + 2\hat\beta_j^2/3}.
\end{aligned}$$

Recall that β_j is related to $\tilde\beta_j$ by (4.14).

Remark 5.6. The quantities in Remark 3.17 can be written in terms of the associated shift-invariant kernel defined in (5.3) as follows:

$$\begin{aligned}
\|I\|_K^2 = E_{f\in\mathcal{A}}\{[I(f)]^2\} &= \int_{C^s} K_{\text{shift}}(x,0)\ dx,\\
E_{f\in\mathcal{A}}\{\text{Var}(f)\} &= K_{\text{shift}}(0,0) - \int_{C^s} K_{\text{shift}}(x,0)\ dx.
\end{aligned}$$

The implication is that different reproducing/covariance kernels that share the same associated shift-invariant kernel have the same $\|I\|_K$, etc.

6 Discrepancy Bounds

Sections 3 and 4 have presented several worst-case error bounds for quasi-Monte Carlo quadrature. Each bound is a product of a suitably defined discrepancy, $D(P)$, which depends only on the set P defining the quadrature rule, and the variation of the integrand, $V(f)$ which is a measure of the integrand's roughness. The average-case quadrature error is also a suitably defined discrepancy, $D(P)$. Since a smaller discrepancy tends to imply greater accuracy, one would like to know how small the discrepancy can be. This chapter presents lower and upper bounds.

Remark 6.1. An easy way to make the discrepancy smaller is to redefine it by multiplying it by some positive constant $a < 1$. In the worst-case analysis one must then redefine the variation by multiplying it by a^{-1}. Thus the error bound, which is the product of the discrepancy and the variation, remains the same for any particular integrand f. In the average-case

analysis redefining the discrepancy in this way is equivalent to multiplying the covariance kernel by a. Thus, one is assuming that the integrands are not as rough as before. One way to avoid the ambiguities raised by an arbitrary constant multiple is to define the discrepancy with the following normalization:

$$\|I\|_K^2 = E_{f\in\mathcal{A}}\{[I(f)]^2\} = \int_{C^s\times C^s} K(x,y)\; dx\; dy = 1.$$

For the discrepancy $D_{\mathcal{F},\alpha}(P)$ this is equivalent to choosing $\beta_0 = 1$.

Our focus in this and the following sections is on the weighted version of the discrepancy $D_{\mathcal{F},\alpha,p}(P)$. The reasons are two-fold. For lattice rules this discrepancy can be written in terms of the familiar $\mathcal{P}_\alpha(L)$ (see (4.16)). Also, $D_{\mathcal{F},\alpha,2}(P) = D_{\mathcal{F},2\alpha}(P)$ is related to several of the other discrepancies introduced in Section 4 via Theorem 5.5.

Before looking at lattices we consider the discrepancies of simple random samples used in Monte Carlo quadrature. For product-type weights Theorem 3.18 can be applied to obtain a benchmark value for such samples.

Theorem 6.2. *The mean square of the discrepancy $D_{\mathcal{F},\alpha}(P)$ with product-type weights for simple uniform random samples, P, is:*

$$E_P\{[D_{\mathcal{F},\alpha}(P)]^2\} = \frac{\beta_0^2}{N}\left\{-1 + \prod_{j=1}^{s}[1 + 2\beta_j^\alpha\zeta(\alpha)]\right\}.$$

The mean square discrepancy, which is comparable to $\mathcal{P}_\alpha(L)$, is plotted in Figures 6.1–6.4 for different values of the dimension, s, and the weights, β_j. Although the magnitude of this discrepancy is smaller when β_j and/or s are smaller, the asymptotic order in N is the same for all cases, as is clear from the formula above.

6.1 Upper Bounds for $\mathcal{P}_\alpha(L)$

Upper bounds on $\mathcal{P}_\alpha(L)$ for rank-1 lattice rules are given by several authors (see [Nie92, SJ94]). Here we consider the bounds given by Disney and Sloan [DS91] for prime numbers of points, N, because of the relative simplicity of their proofs. Extending their results to the case of product-type weights we have two lemmas. The first is proved by computing the average $\mathcal{P}_\alpha(L)$ over a family of rank-1 lattices. The second lemma is proved by invoking Jensen's inequality.

Lemma 6.3. *[SJ94, Proposition 4.6 and Theorem 4.8] If N is a prime number, then there exists a generating vector $h \in \mathbf{Z}^s$ such that the rank-1*

lattice with node set $P = \{\{ih/N\} : i = 0, \dots, N-1\}$ *has discrepancy:*

$$\begin{aligned}
\mathcal{P}_\alpha(L) &= \beta_0^{-1}[D_{\mathcal{F},\alpha}(P)]^2 = \beta_0^{1-p}[D_{\mathcal{F},\alpha/p,p}(P)]^p \\
&\le \beta_0 \left\{ -1 + \frac{1}{N} \prod_{j=1}^{s} [1 + 2\beta_j^\alpha \zeta(\alpha)] \right. \\
&\qquad \left. + \frac{N-1}{N} \prod_{j=1}^{s} \left[1 - \frac{(1 - N^{1-\alpha}) 2\beta_j^\alpha \zeta(\alpha)}{N-1} \right] \right\} \\
&\le \frac{\beta_0}{N} \prod_{j=1}^{s} [1 + 2\beta_j^\alpha \zeta(\alpha)] \quad \textit{for } N > 1 + 2\zeta(\alpha) \max_j \beta_j.
\end{aligned}$$

Lemma 6.4. *[SJ94, Proposition 4.7] If* $N \ge 2$*, then for any lattice:*

$$\mathcal{P}_\alpha(L) \le \beta_0^{1-\alpha/\gamma} [\mathcal{P}_\gamma(L)]^{\alpha/\gamma}.$$

As noted by Disney and Sloan the first result by itself is not very strong because it indicates that asymptotic order of the discrepancy $D_{\mathcal{F},\alpha}(P)$ for node sets of lattices is the same as that for simple random samples, that is, $\mathrm{O}(N^{-1/2})$. However, one may combine the previous two lemmas as follows:

Theorem 6.5. *If* N *is a prime number, then there exists a generating vector* $h \in \mathbf{Z}^s$ *such that the rank-1 lattice with node set* $P = \{\{ih/N\} : i = 0, \dots, N-1\}$ *has discrepancy:*

$$\begin{aligned}
\mathcal{P}_\alpha(L) &= \beta_0^{-1}[D_{\mathcal{F},\alpha}(P)]^2 = \beta_0^{1-p}[D_{\mathcal{F},\alpha/p,p}(P)]^p \\
&\le \beta_0 \min_{1<\gamma\le\alpha} \left\{ -1 + \frac{1}{N} \prod_{j=1}^{s} [1 + 2\beta_j^\gamma \zeta(\gamma)] \right. \\
&\qquad \left. + \frac{N-1}{N} \prod_{j=1}^{s} \left[1 - \frac{(1 - N^{1-\gamma}) 2\beta_j^\gamma \zeta(\gamma)}{N-1} \right] \right\}^{\alpha/\gamma}.
\end{aligned}$$

Disney and Sloan used this theorem to show that there exist rank-1 lattices with $\mathcal{P}_\alpha(L) = \mathrm{O}(N^{-\alpha}[\log N]^{\alpha s})$ for prime N. However, the value of N at which one sees the nearly $N^{-\alpha}$ decay may be quite large, even for moderate dimensions.

Figures 6.1–6.4 show the upper bound on $\mathcal{P}_2(L)$ from Theorem 6.5 for dimensions 2 and 13 and two cases of product-type weights. One case is $\beta_0 = \beta_1 = \cdots = \beta_s = 1$, which corresponds to the original (or unweighted) $\mathcal{P}_2(L)$. The second case is $\beta_0 = 1, \beta_1 = \cdots = \beta_s = \sqrt{3/(8\pi^2)}$. These weights correspond to setting $\hat\beta_1 = \cdots = \hat\beta_s = 1$ for the $\mathcal{L}_2$-star discrepancy, and then computing the corresponding β_j according to Theorem 5.5. Although

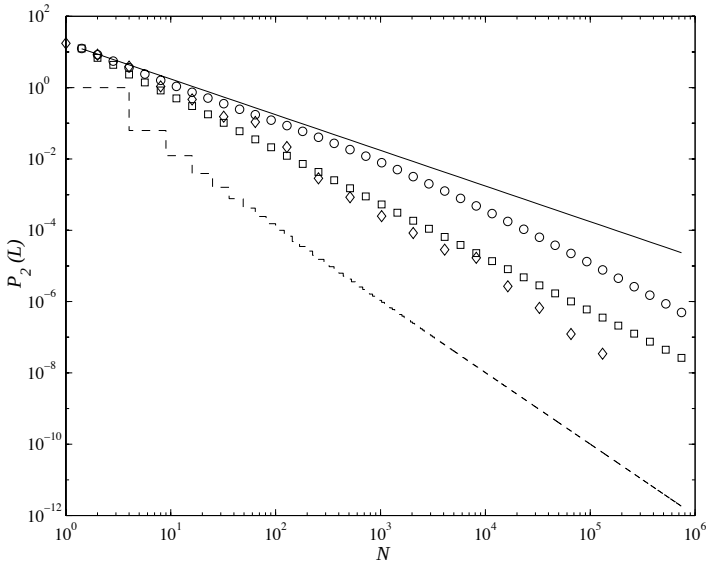

FIGURE 6.1. The mean square of $D_{\mathcal{F},2}(P)$ for simple random samples by Theorem 6.2 (solid); the upper bound on $\mathcal{P}_2(L)$ for rank-1 lattices from Theorem 6.5 ($\circ$); the upper bound on $\mathcal{P}_2(L)$ for copy rules using Theorem 6.7 ($\square$) ; the lower bound on $D_{\mathcal{F},2,1}(P)$ from Theorem 6.6 (dashed); $\mathcal{P}_2(L)$ for the lattice sequence described in Section 7.2 ($\diamond$); for $s = 2$, $\beta_0 = \beta_1 = \cdots = \beta_s = 1$.

Theorem 6.5 applies strictly to prime numbers N, this restriction is ignored for convenience in these figures

The upper bounds on $\mathcal{P}_2(L)$ vary substantially with β_j and s. Although the upper bounds are smaller than the corresponding discrepancy of a simple random sample, the difference is small for higher dimensions for the range of N plotted here. Also, the eventual asymptotic decay rate of nearly $O(N^{-2})$ is not evident for moderate N.

6.2 A Lower Bound on $D_{\mathcal{F},\alpha,1}(P)$

In contrast to the upper bounds on $\mathcal{P}_\alpha(L)$ above, Sloan and Woźniakowski [SW97] have given some lower bounds on the discrepancy $D_{\mathcal{F},\alpha,1}(P)$ for any P and for $N < 2^s$. Here we generalize their argument for the case of a weighted discrepancy and for other values of N.

Theorem 6.6. *For any set $P \subset C^s$ with N points there is the following lower bound on the weighted discrepancy:*

$$D_{\mathcal{F},\alpha,1}(P) \geq \min_{\emptyset \subset u \subseteq S} (\beta_u \lfloor N^{1/s} \rfloor^{-\alpha|u|}).$$

Proof. Suppose that $r = \lfloor N^{1/s} \rfloor$, which implies that $N < (r+1)^s$. For any

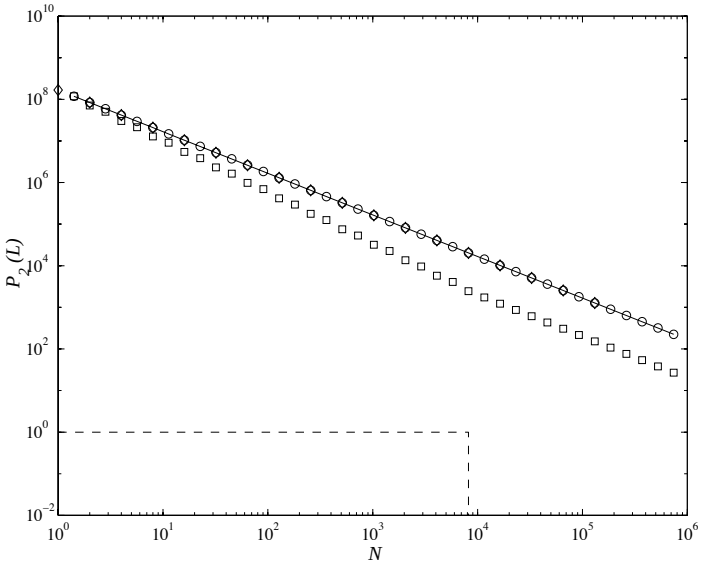

FIGURE 6.2. The same as Figure 6.1, except that $s = 13$.

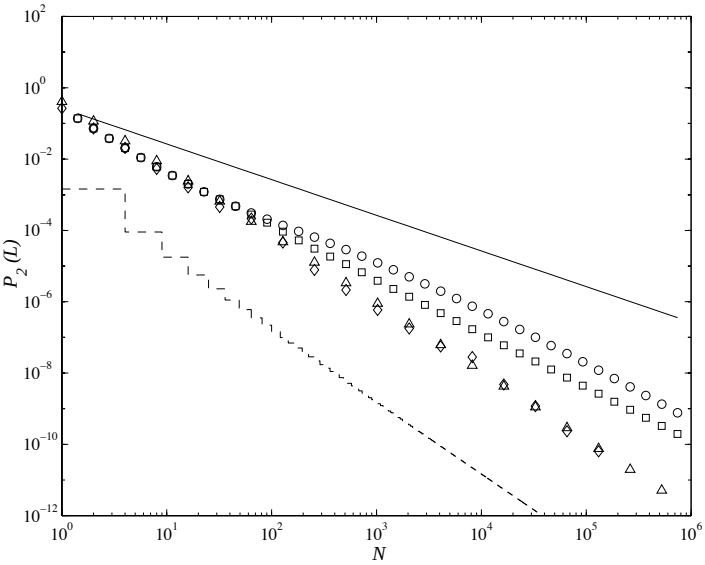

FIGURE 6.3. The same as Figure 6.1, except that $\beta_1 = \cdots = \beta_s = \sqrt{3/(8\pi^2)}$. Also, includes the mean square $\mathcal{L}_2$-star discrepancy of a randomized $(0, m, s)$-net in base s according to (7.1) ($\triangle$).

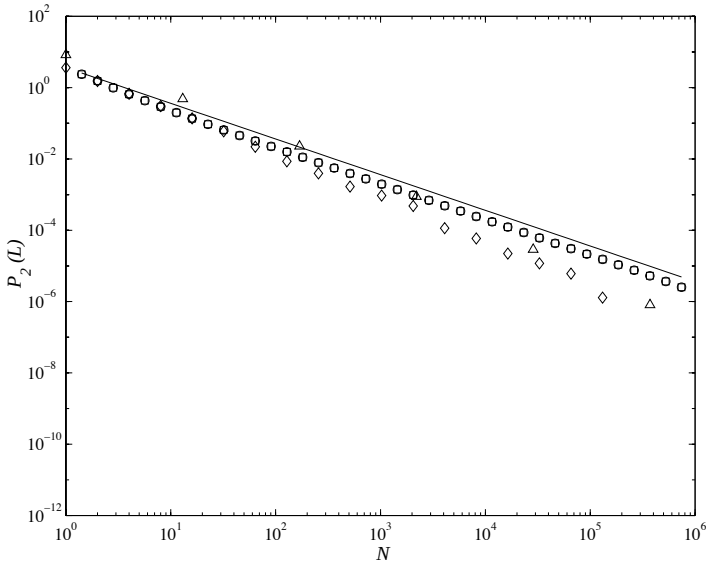

FIGURE 6.4. The same as Figure 6.3, except that $s = 13$.

set $P \subset C^s$ with N points, the system of N linear equations,

$$\sum_{k \in \{0,1,\ldots,r\}^s} a_k e^{2\pi i k \cdot z} = 0 \quad \text{for all } z \in P,$$

has a nontrivial solution for the $(r+1)^s$ coefficients a_k. This nontrivial solution may be scaled so that $|a_k| \le a_{k^*} = 1$, $\forall k \in \{0,1,\ldots,r\}^s$ for some particular k^*. Now form the integrand

$$f(x) = \sum_{k \in \{0,1,\ldots,r\}^s} a_k e^{2\pi i (k-k^*)\cdot x}.$$

Because of the above conditions on the a_k the quadrature rule applied to f gives 0, but the integral of f is 1. So,

$$1 = |I(f) - Q(f;P)| \le D_{\mathcal{F},\alpha,1}(P) V_{\mathcal{F},\alpha,\infty}(f).$$

A lower bound on the variation of f implies an upper bound on the discrepancy:

$$\begin{aligned} D_{\mathcal{F},\alpha,1}(P) \ge [V_{\mathcal{F},\alpha,\infty}(f)]^{-1} &= \left\| (\beta_u^{-1} V_{\mathcal{F},\alpha,\infty,u}(f_u))_{u \ne \emptyset} \right\|_\infty^{-1} \\ &= \left\| \left(\beta_u^{-1} \left[\prod_{j \in u} (\overline{k_j - k_j^*})^\alpha \right] \hat{f}_u(k_u) \right)_{\emptyset \subset u \subseteq S} \right\|_\infty^{-1}. \end{aligned}$$

For $k, k^* \in \{0, 1, \ldots, r\}^s$ it follows that $\overline{k_j - k_j^*} \le r^\alpha$. Moreover, $|\hat{f}_u(k_u)| \le \max_k |a_k| = 1$. Substituting these bounds into the above formula completes the proof. □

Sloan and Woźniakowski's version of Theorem 6.6 is the case $\beta_u = 1$ and $N < 2^s$, which implies that $D_{\mathcal{F},\alpha,1}(P) \ge 1$. Sloan and Woźniakowski thus argued that it is safer to assume that the integral is zero unless one cannot afford at least 2^s integrand evaluations. This means that at least 10^3 points are needed for $s = 10$ and at least 10^6 points are needed for $s = 20$.

However, if one chooses product-type weights of the form (4.9), with $\beta_0 = 1$, $\beta_1 = \cdots = \beta_s = \beta < 1$, then the lower bound on the discrepancy becomes $(\beta \lfloor N^{1/s} \rfloor)^{-\alpha s}$, which is $\beta^{-\alpha s}$ for $N < 2^s$. This lower bound vanishes as α or s approach infinity. Therefore, the pessimistic result of Sloan and Woźniakowski [SW97] is strongly dependent on the choice of weights in the definition of the discrepancy. As mentioned in Remarks 4.4 and 4.5 the choice of weights reflects one's assumptions about the sizes of the lower and higher dimensional parts of the integrand.

Figures 6.1–6.4 show the lower bound in Theorem 6.6. In most of the cases considered the lower bound is not close to the known upper bound, except for small N and s. In some of the figures the lower bound is too small to appear within the plotting window. This suggests the need for improved upper and/or lower bounds.

6.3 Quadrature Rules with Different Weights

Even if one interprets the lower bound on the discrepancy in Theorem 6.6 pessimistically, there is a better alternative to choosing the quadrature rule $Q(f) = 0$ for $N < 2^s$. Consider a quadrature rule that is the *weighted average* of the values of the integrand, that is,

$$Q(f; P, \{w^{(i)}\}) = \sum_{i=1}^{N} w^{(i)} f(z^{(i)}).$$

(The choice $w^{(i)} = N^{-1}$ is the original quadrature rule.) The worst-case error and average-case analyses in Theorems 3.1 and 3.15 can be modified to accommodate this new quadrature rule. If the weights do not sum up to 1, that is, the new rule is not exact for constants, then the variation of the integrand must be defined as its norm, that is, $V(f; K) = \|f\|_K$. The definition of the discrepancy is modified as follows:

$$D(P, \{w^{(i)}\}; K) = \left\{ \int_{C^s \times C^s} K(x, y)\, dx\, dy - 2 \sum_{i=1}^{N} w^{(i)} \int_{C^s} K(z^{(i)}, y)\, dy + \sum_{i,i'=1}^{N} w^{(i)} w^{(i')} K(z^{(i)}, z^{(i')}) \right\}^{1/2}.$$

This formula is a quadratic function of the weights. Therefore, the weights which minimize the discrepancy satisfy the following system of N linear equations for the N unknown $w^{(i)}$:

$$-\int_{C^s} K(z^{(i)}, y)\ dy + \sum_{i'=1}^{N} w^{(i')} K(z^{(i)}, z^{(i')}) = 0, \quad i = 1, \ldots, N.$$

This system of equations may be too large to solve in practice. However, in the case where P is the node set of a lattice, and the kernel is shift-invariant, it can be shown that the optimal weights are uniform, i.e. $w^{(1)} = \cdots = w^{(N)}$. Even in cases where the optimal weights are not uniform, one may easily solve for the best uniform weights:

$$w^{(1)} = \cdots = w^{(N)} = \frac{w_{\text{opt}}}{N} = \left[\sum_{z \in P} \int_{C^s} K(z, y)\ dy\right] \left[\sum_{z, z' \in P} K(z, z')\right]^{-1}.$$

In general, this weight satisfies $0 < w_{\text{opt}} < 1$. For small N the weight w_{opt} is close to 0, which corresponds to a cautious approach in estimating the integral with too few function evaluations. As N tends to infinity, w_{opt} tends to 1, which corresponds to the original quadrature rule, (2.1). The discrepancy for the best uniform weights is:

$$D(P, w_{\text{opt}} N^{-1}; K) = \left\{ \int_{C^s \times C^s} K(x, y)\ dx\ dy \right.$$
$$\left. - \left[\sum_{z \in P} \int_{C^s} K(z, y)\ dy\right]^2 \left[\sum_{z, z' \in P} K(z, z')\right]^{-1} \right\}^{1/2}.$$

For shift-invariant kernels, the discrepancy for the best uniform weights can be written in terms of the original discrepancy with each integrand value weighted by N^{-1}:

$$D(P, w_{\text{opt}} N^{-1}; K) = \left\{ [D(P; K)]^{-2} + \|I\|_K^{-2} \right\}^{-1/2}.$$

Therefore, by choosing optimal uniform weights one obtains a smaller discrepancy than either $D(P; K) = D(P, N^{-1}; K)$ or $\|I\|_K = D(P, 0; K)$. The former case corresponds to the original quadrature rule, and the latter case corresponds to pessimistically choosing $Q(f) = 0$.

6.4 Copy Rules

The monograph of Sloan and Joe [SJ94] and the papers cited therein recommend copy rules for two reasons: i) they provide a way to estimate

quadrature error, and ii) they tend to give smaller $\mathcal{P}_\alpha(L)$ for the same number of points than rank-1 lattice rules. Here we generalize some of the results in [SJ94] on copy rules and comment on the assertion that they give smaller $\mathcal{P}_\alpha(L)$ values.

Define a rectangular grid with spacing n_j^{-1} in the j^{th} coordinate as follows:

$$G = \left\{ \left(\frac{m_1}{n_1}, \ldots, \frac{m_s}{n_s} \right) : m_j = 0, \ldots, n_j - 1 \right\},$$

for some integers $n_1, \ldots, n_s$. For any set $P \in C^s$ (P is not necessarily the node set of a lattice) let $P + G \in C^s$ denote $n_1 \cdots n_s$ shifted copies of P, specifically,

$$P + G = \{ \{ z + z' \} : z \in P, z' \in G \}.$$

Let $D_{\mathcal{F},\alpha,p}(P + G; \beta_0, \ldots, \beta_s)$ denote the weighted version of the discrepancy of this set as defined in (4.10), where the dependence on β_i is given explicitly. This discrepancy can be computed in terms of the discrepancy of a set related to P as follows:

Theorem 6.7. *[Hic96b, Lemma 4.1] For any set $P \in C^s$ and a grid G as defined above,*

$$D_{\mathcal{F},\alpha,p}(P + G; \beta_0, \ldots, \beta_s) = D_{\mathcal{F},\alpha,p}(\tilde{P}; \beta_0, n^{-1}\beta_1, \ldots, n^{-s}\beta_s), \tag{6.1}$$

where $\tilde{P} = \{ \{ (n_1 z_1, \ldots, n_s z_s) \} : z \in P \}$.

Proof. According to (4.10) the discrepancy of $P + G$ involves the following term, which can be rewritten using some straightforward algebraic manipulations:

$$\begin{aligned}
\frac{1}{n_1 \cdots n_s N} \sum_{z \in P+G} e^{2\pi i k \cdot z} &= \frac{1}{n_1 \cdots n_s N} \sum_{z \in P} e^{2\pi i k \cdot z} \sum_{z' \in G} e^{2\pi i k \cdot z'} \\
&= \frac{1}{n_1 \cdots n_s N} \sum_{z \in P} e^{2\pi i k \cdot z} \prod_{j=1}^{s} \sum_{m_j = 0}^{n_j} e^{2\pi i k_j m_j / n_j} \\
&= \frac{1}{n_1 \cdots n_s N} \sum_{z \in P} e^{2\pi i k \cdot z} \prod_{j=1}^{s} n_j 1_{\{k_j \in n_j \mathbf{Z}\}} \\
&= \prod_{j=1}^{s} 1_{\{k_j \in n_j \mathbf{Z}\}} \frac{1}{N} \sum_{z \in P} e^{2\pi i k \cdot z}.
\end{aligned}$$

Since one need only consider those k_j which are integer multiples of n_j, one can replace the vector k by $(\tilde{k}_1 n_1, \ldots, \tilde{k}_s n_s)$. Making this substitution into the formula for the discrepancy of $P + G$ completes the proof. □

The essence of this theorem is that copy rules correspond to reducing the sizes of the weights β_j. Since the discrepancy $D_{\mathcal{F},\alpha,p}(P; \beta_0, \ldots, \beta_s)$ is

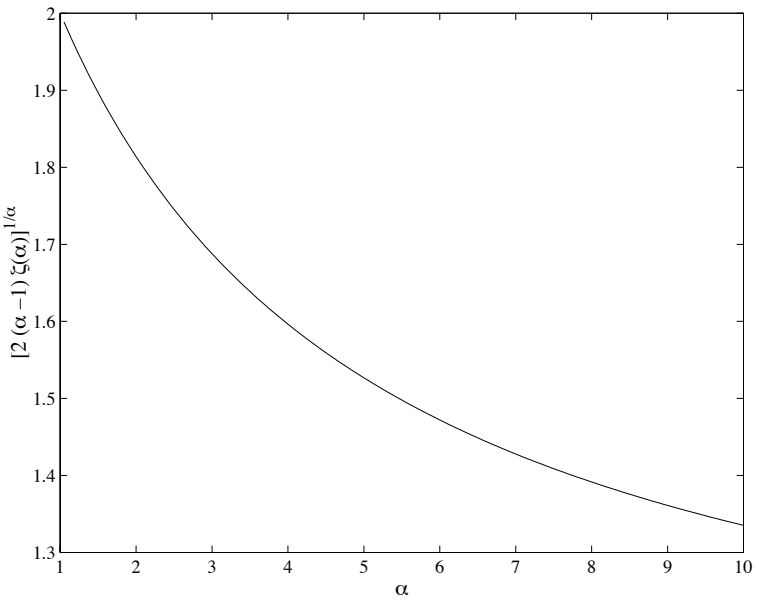

FIGURE 6.5. The optimal n_j for $\beta_j = 1$.

monotonically increasing in each β_j for any set P, this result would appear to favor copy rules. However, one must realize that the cost of reducing the β_j is that $P + G$ in the left-hand side of (6.1) contains $n_1 \cdots n_s$ times as many points as the $\tilde{P}$ in the right-hand side. It is not immediately obvious whether or not some other set, P', with the same number of points as $P+G$ might have an even lower discrepancy value.

For lattice rules Sloan and Joe [SJ94] look at the simpler, though not as tight, upper bound $\mathcal{P}_\alpha(L)$ based on Lemma 6.3 and Theorem 6.7 for a copy lattice rule with N points:

$$\frac{\beta_0}{N} \prod_{j=1}^{s} [n_j + 2n_j^{1-\alpha} \beta_j^\alpha \zeta(\alpha)].$$

The choice of n_j that minimizes this bound can be found using simple calculus, and corresponds to:

$$n_j = \beta_j [2(\alpha - 1)\zeta(\alpha)]^{1/\alpha}.$$

The quantity multiplying β_j is plotted in Figure 6.5. Recognizing that in practice n_j must be an integer it is easy to see from the Figure 6.5 that $n_j = 2$ is a good choice for $\beta_j = 1$. In fact, Sloan and Joe showed that if $\beta_j = 1$, then $n_j = 2$ is better than $n_j = 1$ (no copies) for all $\alpha > 1$. However, for small enough values of β_j the optimum is $n_j = 1$ (no copies).

Figures 6.1–6.4 show the upper bound on $\mathcal{P}_2(L)$ found by combining Theorems 6.5 and 6.7 and searching for copy lattice rules where $n_1 = \cdots = n_j = n_{j+1} + 1 = \cdots = n_s + 1$. Typical values of the total number of points, N, are chosen, where $N > n_1 \cdots n_s$, but N is not constrained to be an integer multiple of $n_1 \cdots n_s$, for convenience. As N increases, so do the n_j. While this might support the assertion that copy rules are preferable, the author believes that it is due to the fact that the upper bounds being minimized are not tight. An obvious disadvantage of copy rules is that while increasing the number of points by a factor $n_1 \cdots n_s$, you only get n_j times as many new values in the j^{th} coordinate direction.

7 Discrepancies of Integration Lattices and Nets

Rather than relying on upper bounds one may compute the discrepancies of specific integration lattices and other quasi-random sets and compare their values. This has been done by [MC94, Hic95] and others. This section considers the discrepancies of two kinds of quasi-random points. For a look at the discrepancies corresponding to Smolyak quadrature rules see [FH96].

7.1 The Expected Discrepancy of Randomized $(0, m, s)$-Nets

Besides integration lattices, (t, m, s)-nets and (t, s)-sequences in base b are popular choices of quasi-random points. This family includes the sequences of Sobol' [Sob67], Faure [Fau82] and Niederreiter [BFN92]. For a review of nets and sequences see [Nie92] and elsewhere in this volume.

Although computing a discrepancy of the form (3.3) for a particular (t, m, s)-net requires $\mathrm{O}(sN^2)$ operations (or somewhat less when the algorithm in [Hei96] can be used), the root mean square discrepancy over a randomized $(0, m, s)$-nets has a computationally simple formula (see [Hic96a]). Here the nets are assumed to be randomized according to Owen's procedure [Owe95, Owe97a, Owe97b]. For the $\mathcal{L}_2$-star discrepancy defined in Section 4.5 with weights $\hat{\beta}_1 = \cdots = \hat{\beta}_s = \hat{\beta}$ the mean square discrepancy is:

$$E_P\{[D_2^*(P)]^2\} = \hat{\beta}_0^2 \left(1 + \frac{\hat{\beta}^2}{2}\right)^s b^{-2m} \times \sum_{l=0}^{s-1} \sum_{k=0}^{l} \sum_{j=0}^{l-k} \left[\binom{s}{l+1} \binom{m}{k} \binom{l}{j} \left(\frac{\hat{\beta}^2}{6 + 3\hat{\beta}^2}\right)^{l+1} b^{-l} (b^2 - 1)^k (-b - 1)^j \right]. \tag{7.1}$$

Figures 6.3–6.4 show plots of the mean square $\mathcal{L}_2$-star discrepancy for $(0, m, s)$-nets with $\hat{\beta} = 1$. The corresponding shift-invariant kernel according to Theorem 5.5 is $K_{\mathcal{F},\alpha}$ with weights $\beta_1 = \cdots = \beta_s = \sqrt{3/(8\pi^2)}$, so the mean square $\mathcal{L}_2$-star discrepancy is comparable to $\mathcal{P}_\alpha(L)$ with these

weights. The $(0, m, s)$-nets have significantly smaller discrepancy than the upper bounds for lattice rules, especially as N increases.

Remark 7.1. Formula (7.1) requires only $O(s \max(s, \log N))$ operations to evaluate.

7.2 Infinite Sequences of Embedded Lattices

One advantage of (t, s)-sequences over integration lattices is that the number of points required, N, need not be chosen in advance. If according to some quadrature error estimate the original N is too small, one may take the next $N' - N$ points of the (t, s)-sequence to get a total of N' points without discarding the original N points. However, this is typically not the case for integration lattices. Since the generating vector is a function of the number of points, an N' point lattice cannot normally be obtained by adding $N' - N$ points to an N point lattice.

The author and his collaborator have proposed a method for overcoming this disadvantage of lattice rules by defining an infinite lattice sequence that is analogous to a (t, s)-sequence [HH97]. Let $\phi_b(i)$ denote the Van der Corput sequence, that is, if $\cdots i_3 i_2 i_1$ is the base-b representation of the non-negative integer i with digits $i_1, i_2, \ldots$, then

$$\phi_b(i) = 0.i_1 i_2 i_3 \ldots \quad \text{(base } b\text{)}.$$

For example, the first several terms of the base-2 Van der Corput sequence are:

$$0, \frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \frac{5}{8}, \ldots$$

A rank-1 infinite lattice sequence is defined by replacing i/N in (2.2) by $\phi_b(i)$, that is,

$$P_\infty = \{\{\phi_b(i)h\} : i = 0, 1, \ldots\}.$$

Each component of the generating vector is now considered to be an infinite string of digits,

$$h = (\ldots h_{12} h_{11},\ \ldots h_{22} h_{21},\ \ldots,\ \ldots h_{s2} h_{s1}) \quad \text{(base } b\text{)}.$$

The first b^m points of P_∞ are the node set of a lattice for any non-negative integer m. Moreover, the succeeding runs of b^m points are shifted copies of the first b^m points. Furthermore, only the first m digits of each component of h are needed to generate the first $N = b^m$ points of P_∞.

Remark 7.2. Although one may in theory use the first N points from P_∞ for quadrature for any value of N, the lattice property holds only if N is an integer power of the base, b. An analogous property holds for (t, s)-sequences, that is, the (t, m, s)-net is the first b^m points of the (t, s)-sequence. For both lattices and nets one expects the discrepancy to be smaller when choosing $N = b^m$ points.

Finding good generating vectors for lattices of the above form is an area of ongoing research. Some promising results have been obtained by using optimization methods to find h which minimize $\mathcal{P}_\alpha(L)$. To make the optimization problem tractable one may restrict the class of possible generating vectors, for example, requiring h to be of Korobov-type, that is,

$$h = (1, h_2, h_2^2, h_2^3, \ldots, h_2^{s-1}).$$

The values of $P_\alpha(L)$ for the first $N = 1, 2, 4, \ldots, 2^{17}$ points of an infinite lattice sequence of this type are shown in Figures 6.1–6.4. The generator is

$$h_2 = 89715 = 10101111001110011 \text{ (base 2)}$$

The infinite lattice has similar discrepancy values to the $(0, m, s)$-nets.

8 Tractability of High Dimensional Quadrature

The discussion so far has assumed the dimension s to be some fixed, finite quantity. It is also interesting to explore what happens if s becomes arbitrarily large.

8.1 Quadrature in Arbitrarily High Dimensions

Quasi-Monte Carlo methods have proven to be effective in evaluating some integrals arising in the valuation of financial securities (see [PT95, NT96, PT96, CMO97] and elsewhere in this volume). In such problems the dimension, s, may be in the hundreds or thousands. The asymptotic order of the discrepancy for a fixed s and letting N tend to infinity typically contains $\log N$ to some power that increases with s. Such expressions for the asymptotic order of the quadrature error are quite pessimistic for arbitrarily large s. This has motivated some scholars to try other approaches to understand quadrature error in arbitrarily high dimensions.

The theorems in Section 6 can be used to determine the behavior of the discrepancy when s tends to infinity. For Monte Carlo methods Theorem 6.2 says that the mean square discrepancy is uniformly bounded in s, if and only if

$$\lim_{s \to +\infty} \prod_{j=1}^{s} [1 + 2\beta_j^\alpha \zeta(\alpha)] < \infty.$$

By taking logarithms of the product, this condition can be shown to be equivalent to

$$\sum_{j=1}^{+\infty} \beta_j^\alpha < \infty. \tag{8.1}$$

The upper bound on $\mathcal{P}_\alpha(L)$ is also uniformly bounded in s if and only if condition (8.1) holds. A comparable condition is obtained by [SW98] for the $\mathcal{L}_2$-star discrepancy. The condition that the β_j tend to zero fast enough means that the integrand varies less and less with respect to x_j as j tends to infinity.

8.2 The Effective Dimension of an Integrand

Figures 6.1–6.4 show the dependence of the discrepancy or $\mathcal{P}_\alpha(L)$ on the dimension, s, and the weights, β_j. One can imagine integrands, such as that given in (4.1), whose effective dimension is smaller than the nominal dimension s. The ANOVA decomposition introduced in Section 4.1 provides a way of defining the effective dimension of an integrand.

The effective dimension of a particular integrand has been defined by Caflisch, Morokoff and Owen [CMO97]. It can be shown from the definition in (4.2) that the variance of a function is the sum of the variances of its ANOVA effects:

$$\mathrm{Var}(f) = \sum_{\emptyset \subset u \subseteq S} \mathrm{Var}(f_u).$$

(Note that $\mathrm{Var}(f_\emptyset) = 0$ because $f_\emptyset$ is a constant.) The dimension of an integrand may be defined in terms of the dimension of the ANOVA effects that contain some proportion of the variance of the integrand.

Definition 8.1. An integrand f is said to be proportion p_t of **truncated dimension** s_t if and only if

$$\sum_{u \subseteq \{1,\ldots,s_t\}} \mathrm{Var}(f_u) = p_t \,\mathrm{Var}(f),$$

where $0 \le p_t \le 1$. An integrand f is said to be proportion p_s of **superposition dimension** s_s if and only if

$$\sum_{|u| \le s_s} \mathrm{Var}(f_u) = p_s \,\mathrm{Var}(f),$$

where $0 \le p_s \le 1$.

The example in (4.1) is a relatively simple function, so its dimension can be computed easily. This function has 100% truncation dimension 3 and 100% superposition dimension 2.

An alternative to computing the dimension of a single integrand is to compute the average dimension of a family of integrands. For the spaces of random functions, $\mathcal{A}$, introduced in Section 3.3, recall formula (3.20) for the mean value of the variance of a function. An analogous definition to the one above is:

Definition 8.2. The integrands in a space of random functions, $\mathcal{A}$, are said to be proportion p_t of **truncated dimension** s_t if and only if

$$\sum_{u \subseteq \{1,\ldots,s_t\}} E_{f\in\mathcal{A}}[\mathrm{Var}(f_u)] = p_t[E_{f\in\mathcal{A}}\,\mathrm{Var}(f)].$$

The integrands in a space of random functions, $\mathcal{A}$, are said to be proportion p_s of **superposition dimension** s_s if and only if

$$\sum_{|u|\le s_s} E_{f\in\mathcal{A}}[\mathrm{Var}(f_u)] = p_s E_{f\in\mathcal{A}}[\mathrm{Var}(f)].$$

Remark 8.3. In the previous two definitions one may either fix the proportions, p_s and p_t, and then solve for s_s and s_t, or vice versa. In [CMO97] the choice $p_s = p_t = 0.95$ is recommended.

Remark 8.4. If $p_s = p_t$, then $s_s \le s_t$, and if $s_s = s_t$, then $p_s \ge p_t$.

Computing the truncated or superposition dimension of integrands in a space $\mathcal{A}$ requires formulae for $E_{f\in\mathcal{A}}[\mathrm{Var}(f_u)]$ in terms of the covariance kernel. From the definition of the ANOVA decomposition in (4.2) it follows that:

$$\begin{aligned}\sum_{(u',v')\subseteq(u,v)} E_{f\in\mathcal{A}}[f_{u'}(x_{u'})f_{v'}(y_{v'})] &= E_{f\in\mathcal{A}}\left[\sum_{u'\subseteq u} f_{u'}(x_{u'}) \sum_{v'\subseteq v} f_{v'}(y_{v'})\right]\\ &= E_{f\in\mathcal{A}}\left[\int_{C^{S-u}} f_u(x_u)\; dx_{S-u} \int_{C^{S-v}} f_v(y_v)\; dy_{S-v}\right]\\ &= \int_{C^{S-u}\times C^{S-v}} K(x,y)\; dx_{S-u}\; dy_{S-v}.\end{aligned}$$

The mean over $\mathcal{A}$ of $\mathrm{Var}(f_u)$ is obtained by setting $v = u$ and $y_v = x_u$ in the above equation and integrating with respect x_u over C^u. Furthermore, one may use the associated shift-invariant kernel, as defined in (5.3), to simplify the resulting formula:

$$\begin{aligned}(8.2)\quad & E_{f\in\mathcal{A}}[f_\emptyset^2] + \sum_{\emptyset\subset u'\subseteq u} E_{f\in\mathcal{A}}[\mathrm{Var}(f_{u'})]\\ &= \int_{C^s\times C^{S-u}} K(x,y)\big|_{y_u=x_u}\; dx\; dy_{S-u}\\ &= \int_{C^{S-u}} K_{\mathrm{shift}}(x,0)\big|_{x_u=0}\; dx_{S-u}.\end{aligned}$$

Further manipulation of this equation leads to the conclusion that:

$$(8.3)\qquad E_{f\in\mathcal{A}}[\mathrm{Var}(f_u)] = \sum_{v\subseteq u}(-1)^{|u-v|}\int_{C^{S-v}} K_{\mathrm{shift}}(x,0)\big|_{x_v=0}\; dx_{S-v}.$$

Using this formula and some counting arguments one may arrive at the following formulae for the truncated and superposition dimensions of functions in $\mathcal{A}$.

Theorem 8.5. *Consider a space of random functions, $\mathcal{A}$, with covariance kernel, K, and associated shift-invariant kernel, K_{shift}. Let $S_t = \{1, \ldots, s_t\}$. The truncated dimension, s_t, and the superposition dimension, s_s, with proportions p_t and p_s for the space $\mathcal{A}$ satisfy the following equations:*

$$\int_{C^{S-S_t}} K_{shift}(x,0)|_{x_{S_t}=0} \; dx_{S-S_t} = p_t K_{shift}(0,0) + (1-p_t) \int_{C^s} K_{shift}(x,0) \; dx,$$

$$\sum_{|u|\le s_s} \left[(-1)^{s_s-|u|} \binom{s-1-|u|}{s_s-|u|} \int_{C^{S-u}} K_{shift}(x,0)|_{x_u=0} \; dx_{S-u} \right] = p_s K_{shift}(0,0) + (1-p_s) \int_{C^s} K_{shift}(x,0) \; dx.$$

Remark 8.6. A function space $\mathcal{A}$ may be said to contain purely s dimensional integrands if the proportions p_t and p_s are zero for s_t and s_s less than s. This occurs if and only if $\int_0^1 K_{\text{shift}}(x_j,0) \; dx_j = 0$, for all $j = 1, \ldots, s$. An example of such a kernel would be the following weighted version of $K_{\mathcal{F},\alpha}$:

$$K_{\mathcal{F},\alpha}(x,y) = \beta_S \sum_{k_j \neq 0} \frac{e^{2\pi i k \cdot (x-y)}}{(\bar{k}_1 \cdots \bar{k}_s)^\alpha}.$$

Many interesting covariance kernels are of product form, that is,

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

In this case the associated shift-invariant kernel is also of product form:

$$K_{\text{shift}}(x,0) = \prod_{j=1}^{s} \tilde{K}_{\text{shift},j}(x_j,0).$$

Let $a_j = \int_0^1 \tilde{K}_{\text{shift},j}(x_j,0) \; dx_j$, and $b_j = \tilde{K}_{\text{shift},j}(0,0)$. Then some of the relevant formulae above may be written in terms of these a_j and b_j. Equation (8.2) becomes:

$$E_{f\in\mathcal{A}}[f_\emptyset^2] + \sum_{\emptyset \subset u' \subseteq u} E_{f\in\mathcal{A}}[\text{Var}(f_{u'})] = \prod_{j\in u} b_j \prod_{j'\in S-u} a_{j'}.$$

Thus, the formula for the truncated dimension is:

$$(b_1 \cdots b_{s_t})(a_{s_t+1} \cdots a_s) = p_t(b_1 \cdots b_s) + (1-p_t)(a_1 \cdots a_s).$$

The formula for the superposition dimension is best computed using a recursion relation. Let

$$A(s,s_s) = E_{f\in\mathcal{A}}[f_\emptyset^2] + \sum_{0<|u|\le s_s} E_{f\in\mathcal{A}}[\mathrm{Var}(f_u)],$$

so that the superposition dimension satisfies the formula:

$$A(s,s_s) = p_s(b_1 \cdots b_s) - (1-p_s)(a_1 \cdots a_s).$$

Because (8.3) can be written as

$$E_{f\in\mathcal{A}}[\mathrm{Var}(f_u)] = \prod_{j\in u}(b_j - a_j) \prod_{j'\in S-u} a_{j'},$$

it follows that $A(s,s_s)$ satisfies the following recursion relationship:

$$\begin{aligned} A(t,0) &= a_1 \cdots a_t, \quad t = 1,2,\ldots, \\ A(1,1) &= A(1,2) = \cdots = b_1, \\ A(t,s_s) &= A(t-1,s_s)a_s + A(t-1,s_s-1)(b_s - a_s). \end{aligned}$$

For the case of product-type weights, (4.9), the covariance kernel $K_{\mathcal{F},\alpha}$ is a shift-invariant, product-type kernel with

$$a_j = 1, \quad b_j = 1 + 2\beta_j^\alpha \zeta(\alpha), \quad j = 1,\ldots,s.$$

The values of p_t and p_s are plotted in Figures 8.1–8.6 for a range of s_t, s_s and s, and for three different types of weights. For $\beta_1 = \cdots = \beta_s = 1$ the integrands are only about 25% of truncation dimension $s-1$. The superposition dimension, s_s, is also close to the nominal dimension, s, for p_s in the range 10-90%. For the other two cases where the weights are smaller, the truncation dimension takes on most of the range of 1 through s for p_t in the range 10-90%, and the superposition dimension is only 1 through 3 for p_s in the range 10-90%. Therefore, the effective truncation and superposition dimensions of the integrands depend strongly on the choices of the weights, β_j.

9 Discussion and Conclusion

How well do lattice rules measure up? The answer depends on the measuring stick. In this chapter several discrepancies have been introduced, but

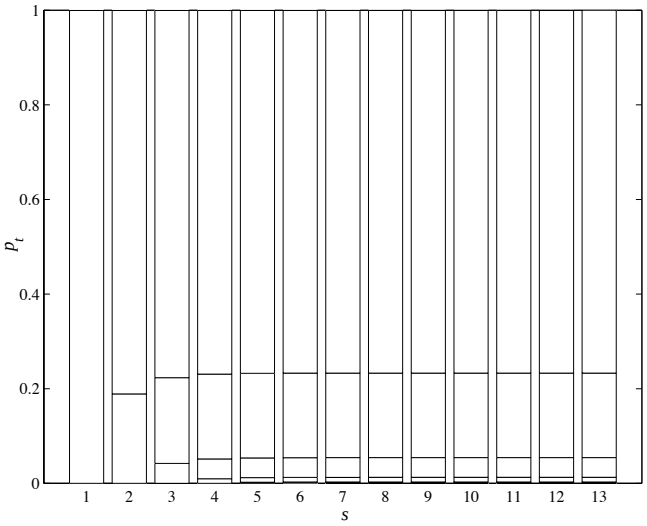

FIGURE 8.1. The proportions p_t for truncation dimensions $s_t = 1, \ldots, s$ and $s = 1, \ldots, 13$ for the space of random functions with covariance kernel $K_{\mathcal{F},\alpha}$ and $\beta_1 = \cdots = \beta_s = 1$.

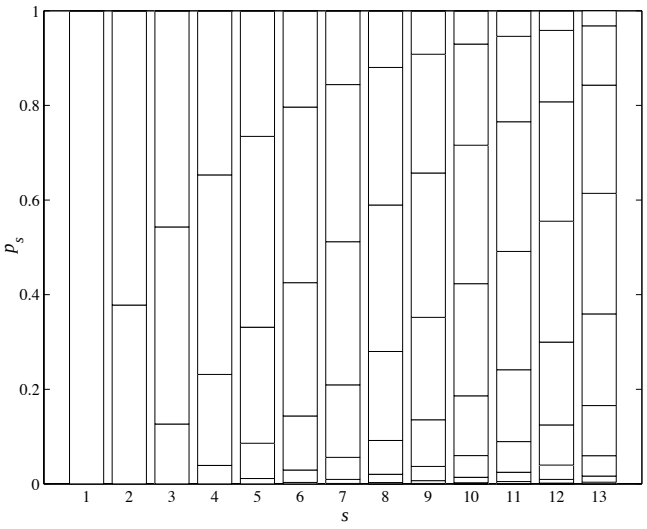

FIGURE 8.2. The same as Figure 8.1 but for the superposition dimension.

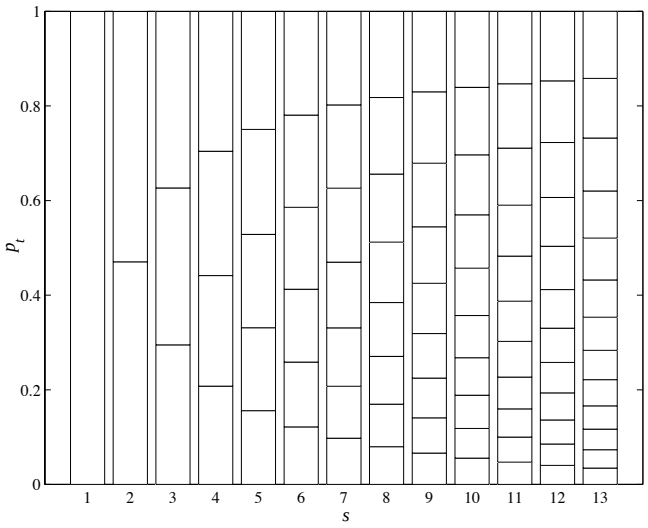

FIGURE 8.3. The same as Figure 8.1 but for $\beta_1 = \cdots = \beta_s = \sqrt{3/(8\pi^2)}$.

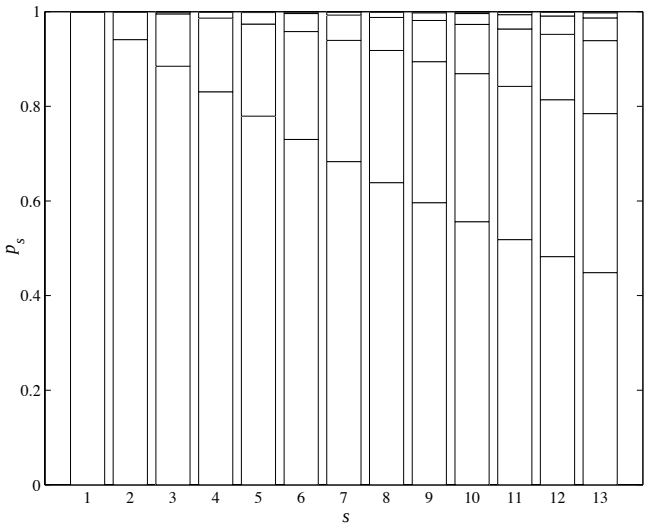

FIGURE 8.4. The same as Figure 8.3 but for the superposition dimension.

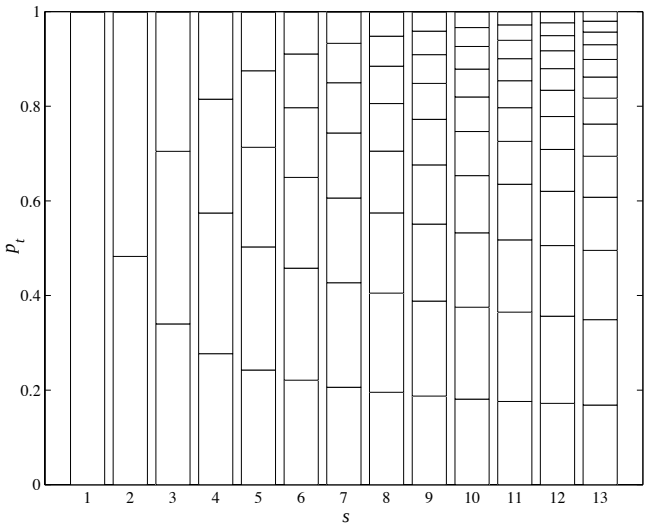

FIGURE 8.5. The same as Figure 8.1 but for $\beta_j = j^{-1},\ j = 1, \ldots, s$.

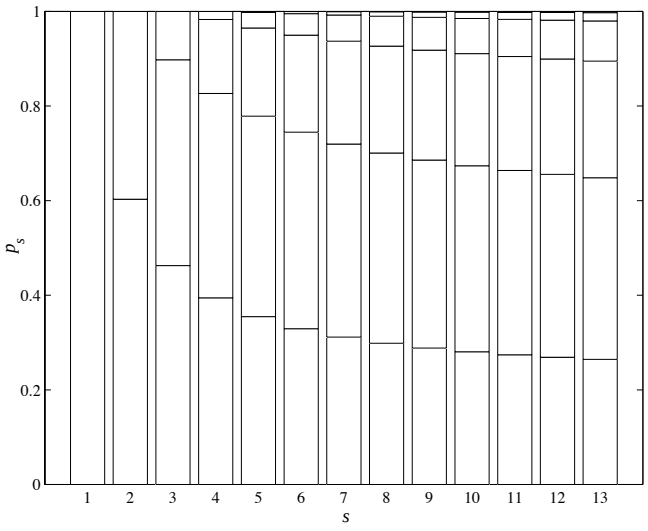

FIGURE 8.6. The same as Figure 8.5 but for the superposition dimension.

most of them are related to a weighted form of $\mathcal{P}_\alpha(L)$, either as a direct generalization or through the associated shift-invariant discrepancy (Theorem 5.5). In fact, the shift-invariant discrepancy associated with the star and unanchored discrepancies is just a weighted form of $\mathcal{P}_2(L)$. Furthermore, for product-type weights and even, positive integer α, the computation of $\mathcal{P}_\alpha(L)$ can be done in only $\mathrm{O}(sN)$ operations.

Therefore, for several reasons a weighted form of $\mathcal{P}_2(L)$ seems to be the best single method for measuring the quality of an integration lattice. However, the choice of weights, β_j, is important. The value and rate of decay of $\mathcal{P}_2(L)$ with N depend on the weights. The choice of the best generating vector for a lattice depends on the choice of weights. The desirability of copy rules depends on whether the weights are large or small.

The average effective dimension of integrands one expects to encounter also depends on the sizes of the weights. In fact, one may write $\mathcal{P}_2(L)$ in (4.15) without explicit reference to the nominal dimension, s, as follows:

$$\mathcal{P}_2(L) = \left\{ -1 + \frac{1}{N} \sum_{z \in P} \prod_{j=1}^{+\infty} \left[1 + \frac{\tilde{\beta}_j^2}{2} B_2(z_j) \right] \right\}^{1/2},$$

with the condition on the weights, (8.1), that

$$\sum_{j=1}^{+\infty} \tilde{\beta}_j^2 < \infty.$$

This form allows for the nominal dimension to be infinite provided that the β_j tend to zero fast enough. Alternatively, the nominal dimension will be some finite s, if, $0 = \tilde{\beta}_{s+1} = \tilde{\beta}_{s+2} = \cdots$. Whether or not the nominal dimension is finite, the truncation or superposition dimensions are likely to be more important in determining the tractability of lattice rules.

Again we stress that the weights in the definition of the discrepancy are not additional parameters that can be used to improve the performance of a particular quadrature rule. Rather, the choice of weights reflects the user's assumptions about the type of integrands that one can expect to encounter as discussed in Remarks 4.4 and 4.5.

Because of the importance of the weights it is advisable that future theoretical investigations of lattice rules allow for different choices of weights. This might mean that one could arrive at different conclusions based on one's choice of weights. For example, Theorem 6.6 gives a pessimistic view of the tractability of quadrature for one choice of weights, but the situation is much better for a different choice of weights.

There seems to be much room for improvement in determining what values of $\mathcal{P}_\alpha(L)$ are possible. There is a sizable gap between the lower and upper bounds that begs to be closed. Moreover, the "average" $(0, m, s)$-net has a much smaller discrepancy than the "average" rank-1 lattice. However,

a few specific lattices considered here have $\mathcal{P}_\alpha(L)$ substantially better than the average and have similar discrepancy values to $(0, m, s)$-nets. Perhaps, there is room to improve the upper bound on $\mathcal{P}_\alpha(L)$ by a more selective average that ignores bad lattices.

Although the discrepancies of lattices are similar to those of nets, the nets still have one practical advantage over lattices. Whereas there are several known explicit constructions of infinite (t, s)-sequences, whose first N points have low discrepancy, the situation is much less developed for lattices. Section 7.2 provides a way of constructing infinite lattice sequences. What is lacking is an explicit construction of a good generating vector.

Acknowledgments

I am grateful to many experts in quasi-Monte Carlo methods and lattice rules for their valuable comments and hearty encouragement, which have greatly benefitted my research. I would especially like to thank Professors Kai-Tai Fang and Yuan Wang for introducing me to integration lattices, and Professor Ian H. Sloan, who gave me helpful advice as I was writing my first paper on lattice rules.

10 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.

[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.

[CMO97] R. E. Caflisch, W. Morokoff, and A. Owen, *Valuation of mortgage backed securities using Brownian bridges to reduce effective dimension*, J. Comput. Finance **1** (1997), 27–46.

[CS96] R. Cools and I. H. Sloan, *Minimal cubature formulae of trigonometric degree*, Math. Comp. **65** (1996), 1583–1600.

[DS91] S. Disney and I. H. Sloan, *Error bounds for the method of good lattice points*, Math. Comp. **56** (1991), 257–266.

[ES81] B. Efron and C. Stein, *The jackknife estimate of variance*, Ann. Stat. **9** (1981), 586–596.

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

[FH96] K. Frank and S. Heinrich, *Computing discrepancies of Smolyak quadrature rules*, J. Complexity **12** (1996), 287–314.

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

[HH97] F. J. Hickernell and H. S. Hong, *Computing multivariate normal probabilities using rank-1 lattice sequences*, Proceedings of the Workshop on Scientific Computing (Hong Kong) (G. H. Golub, S. H. Lui, F. T. Luk, and R. J. Plemmons, eds.), Springer-Verlag, Singapore, 1997, pp. 209–215.

[HH98] F. J. Hickernell and H. S. Hong, *The asymptotic efficiency of randomized nets for quadrature*, Math. Comp. **67** (1998), to appear.

[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.

[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, corrected printing of Sections 3-6 in ibid., **34** (1997), 853–866.

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

[Hla61] E. Hlawka, *Funktionen von beschränkter Variation in der Theorie der Gleichverteilung*, Ann. Mat. Pura Appl. **54** (1961), 325–333.

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

[Kor59] N. M. Korobov, *The approximate computation of multiple integrals*, Dokl. Adad. Nauk. SSR **124** (1959), 1207–1210, (Russian).

[Lan95] T. N. Langtry, *The determination of canonical forms for lattice quadrature rules*, J. Comput. Appl. Math. **59** (1995), 129–143.

[Lan96] T. N. Langtry, *An application of Diophantine approximation to the construction of rank-1 lattice quadrature rules*, Math. Comp. **65** (1996), 1635–1662.

[LJ96] J. N. Lyness and S. Joe, *Triangular canonical forms for lattice rules of prime-power order*, Math. Comp. **65** (1996), 65–178.

[LS97] J. N. Lyness and I. H. Sloan, *Cubature rules of prescribed merit*, SIAM J. Numer. Anal. **34** (1997), 586–602.

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

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

[NT96] S. Ninomiya and S. Tezuka, *Toward real-time pricing of complex financial derivatives*, Appl. Math. Finance **3** (1996), 1–20.

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

[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.

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

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

[PT95] S. Paskov and J. Traub, *Faster valuation of financial derivatives*, J. Portfolio Management **22** (1995), 113–120.

[PT96] A. Papageorgiou and J. F. Traub, *Beating monte carlo*, Risk **9** (1996), no. 6, 63–65.

[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.

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

[SK87] I. H. Sloan and P. J. Kachoyan, *Lattice methods for multiple integration: Theory, error analysis and examples*, SIAM J. Numer. Anal. **24** (1987), 116–128.

[Slo85] I. H. Sloan, *Lattice methods for multiple integration*, J. Comput. Appl. Math. **12 & 13** (1985), 131–143.

[Sob67] I. M. Sobol', *The distribution of points in a cube and the approximate evaluation of integrals*, U.S.S.R. Comput. Math. and Math. Phys. **7** (1967), 86–112.

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

[SW97] I. H. Sloan and H. Woźniakowski, *An intractability result for multiple integration*, Math. Comp. **66** (1997), 1119–1124.

[SW98] I. H. Sloan and H. Woźniakowski, *When are quasi-Monte Carlo algorithms efficient for high dimensional integrals*, J. Complexity **14** (1998), 1–33.

[Wah90] G. Wahba, *Spline models for observational data*, SIAM, Philadelphia, 1990.

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

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