What Affects the Accuracy of Quasi-Monte Carlo Quadrature? ⋆

Fred J. Hickernell

Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong SAR, China, fred@hkbu.edu.hk, http://www.math.hkbu.edu.hk/~fred

Abstract. Quasi-Monte Carlo quadrature methods have been used for several decades. Their accuracy ranges from excellent to poor, depending on the problem. This article discusses how quasi-Monte Carlo quadrature error can be assessed, and what are the factors that influence it.

1 Introduction

Many practical problems arising in finance [PT95,PT96,CMO97], physics and engineering [MC95,Spa95,Kei96,PT97], and statistics [Gen92,Gen93,FW94] require the evaluation of an integral, which, after suitable transformation, may be written as

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

Integrals over one-dimensional intervals may be approximated by sophisticated algorithms, such as Gaussian quadrature, adaptive quadrature and Romberg extrapolation. However, integrals in higher dimensions are often simply approximated by the sample mean over some set P of N points,

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

This quadrature rule is the integral with respect to the measure defined by the **empirical distribution function** corresponding to P, i.e. $Q(f;P) = \int_{[0,1)^s} f(x)\, dF_P(x)$, where

$$F_P(x) = \frac{1}{N}\sum_{z\in P} 1_{\{x\ge z\}}.$$

Here $1_{\{\cdot\}}$ denotes the characteristic function, and $x \ge z$ means $x_j \ge z_j\ \forall j$. The quadrature error may then be written as the integral with respect to a

⋆ This research was partially supported by Hong Kong Research Grants Council grant RGC/97-98/47 and Hong Kong Baptist University grant FRG/96-97/II-67.

signed measure:

$$\text{(2)} \quad \mathrm{Err}(f;P) = I(f) - Q(f;P) = \int_{[0,1)^s} f(x)\ dx - \frac{1}{N}\sum_{z\in P} f(z)$$
$$= \int_{[0,1)^s} f(x)\ d[F_{\mathrm{unif}}(x) - F_P(x)],$$

where $F_{\mathrm{unif}}(x) = x_1 \cdots x_s$ is the **uniform distribution**.

The previous equation suggests that the quadrature error depends on the dimension, s, the integrand, f, and the set of points P. One would like to understand this dependence more precisely, so that one could predict the performance of quadrature rules, and perhaps design better quadrature rules. This article explores such issues.

A good quadrature rule should work well, not just for one integrand, but for a broad class of integrands. Sections 2–5 describe a general worst-case bound on the quadrature error. This bound is the product of the **discrepancy** of the set P, which measures the quality of the quadrature rule, and the **variation** of f, which measures the roughness of the integrand. An important tool in the worst-case analysis is the **reproducing kernel**, described in Sect. 4. The average-case error analysis in Sect. 6 also identifies the discrepancy as the measure of quality of the quadrature rule. Sect. 7 shows some similarities among different discrepancies, and Sect. 8 discusses how these discrepancies depend on various factors. Some concluding remarks and open questions are presented in Sect. 9.

Beginning at the end of Sect. 5 there are several **reality checks**. These are observations that the author believes to be important, and that may be contrary to what one is inclined to believe.

2 Traditional Quadrature Error Bounds

The best-known multidimensional quadrature error bound is the Koksma-Hlawka inequality [Hla61], [Nie92, Theorem 2.11], which takes the form

$$\text{(3)} \qquad |\mathrm{Err}(f;P)| \le D^*(P)V^*(f).$$

Here $D^*(P)$ denotes the **star discrepancy** of P, which is defined as

$$\text{(4)} \quad D^*(P) = \|F_{\mathrm{unif}}(x) - F_P(x)\|_\infty = \left\| x_1 \dots x_s - \frac{1}{N}\sum_{z\in P} 1_{\{x\ge z\}} \right\|_\infty .$$

The term $V^*(f)$ is the variation of f in the sense of Hardy and Krause. In order to define this variation in a compact form, let $S = \{1,\dots,s\}$ be the set of coordinate indices, and for any $u \subset S$, let $|u|$ denote the cardinality of u. Let x_u be the $|u|$-dimensional vector formed by the components of x

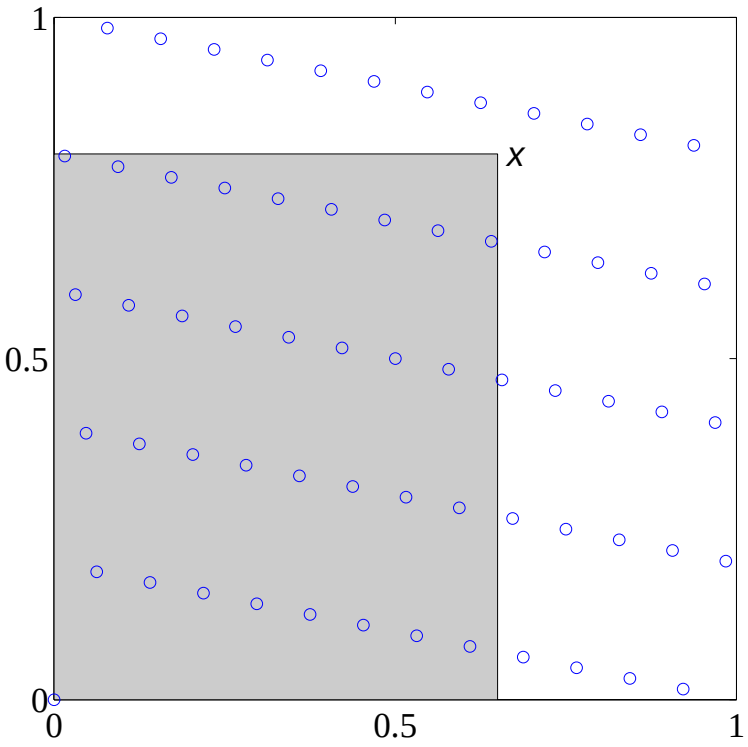

Fig. 1. For $x = (0.65, 0.8)$ the volume of the box $[0, x)$ is 0.52, and the proportion of points in the box is $34/64 = 0.53125$. The difference between these two is $R^*(x) = x_1 x_2 - \frac{1}{64} \sum_{z \in P} 1_{\{x \geq z\}} = -0.01125$.

that are indexed by the elements of u. Furthermore, let $[0,1)^u$ denote the $|u|$-dimensional unit cube that is the projection of $[0,1)^s$ into the coordinates indexed by the elements of u. Then the variation in the sense of Hardy and Krause may be written as

$$V^*(f) = \sum_{\emptyset \subset u \subseteq S} \int_{[0,1)^u} \left| \frac{\partial^{|u|}}{\partial x_u} f(x_u, 1) \right| \, dx_u. \tag{5}$$

Here and below, the notation $f(x_u, 1)$ means that all the components x_j with $j \notin u$ are set equal to one. Figure 1 illustrates the geometric meaning of the star discrepancy — the norm of the difference between the volume of the box $[0, x)$ and the proportion of the points inside the box. The definition of the star discrepancy above can be generalized by using the $\mathcal{L}_p$-norm instead of the $\mathcal{L}_\infty$-norm. Quadrature error bounds involving the $\mathcal{L}_p$-star discrepancy are derived in [Zar68,Sob69] and presented later in (30).

For integrating periodic functions lattice quadrature rules are favored because of their potentially high accuracy. Suppose that the integrand has an absolutely convergent Fourier series with Fourier coefficients, $\hat{f}(k)$:

$$f(x) = \sum_{k \in \mathbf{Z}^s} \hat{f}(k) e^{2\pi i k' x}, \qquad \hat{f}(k) = \int_{[0,1)^s} f(x) e^{-2\pi i k' x} \, dx. \tag{6}$$

Here $k'x$ denotes the dot product of the s-dimensional wavenumber vector k with x. An integration lattice, L, is a discrete subset of $\mathbf{R}^s$ that is also a superset of $\mathbf{Z}^s$, and that is closed under addition and subtraction [SJ94, Def. 2.2]. Lattice quadrature rules are rules of the form (1), where P is the node set of an integration lattice, i.e. $P = L \cap [0,1)^s$. The quadrature error for a lattice rule has the following bound:

$$|\mathrm{Err}(f;P)| \le \mathcal{P}_\alpha(L) \sup_{k \neq 0} [(\bar{k}_1 \cdots \bar{k}_s)^\alpha |\hat{f}(k)|], \tag{7}$$

where the quality of the lattice is measured by

$$\mathcal{P}_\alpha(L) = \sum_{0 \neq k \in L^\perp} \frac{1}{(\bar{k}_1 \cdots \bar{k}_s)^\alpha}, \quad \alpha > 1.$$

Here the dual lattice, $L^\perp$, is defined as the set of integer vectors whose dot product with any lattice point is an integer:

$$L^\perp = \{k \in \mathbf{Z}^s : k'z = 0 \pmod 1 \quad \forall z \in L\}, \tag{8}$$

and the over-bar notation is defined as

$$\bar{k}_j = \begin{cases} |k_j| & \text{for } k_j \neq 0, \\ 1 & \text{for } k_j = 0. \end{cases} \tag{9}$$

The number α is a measure of how smooth the integrands are assumed to be, with larger α implying greater smoothness. For a derivation of (7) and a detailed discussion of lattice rules see [Nie92, Chap. 5] and [SJ94].

The quadrature error bounds (3) and (7) appear to be quite different at first glance. The former applies to any quadrature rule of the form (1), while the latter applies only to lattice rules. The former measures the roughness of the integrand by the size of its mixed partial derivatives, while the latter measures roughness by the rate of decay of the Fourier coefficients.

On the other hand, both error bounds (3) and (7) are the products of two parts, one depending only on the quadrature rule or P, and the other depending only on the integrand, f. In fact these two quadrature error bounds, along with many others are members of a large family, which is described in the following section.

3 General Error Bounds

Formula (2) gives the quadrature error in terms of the integrand, f, and the difference between the uniform distribution and F_P, the empirical distribution of the points used for quadrature. To measure the size of the integrand, suppose that it lies in some Banach space $\mathcal{W}$ with norm $\|\cdot\|_{\mathcal{W}}$. Furthermore, define $f_\perp \in \mathcal{W}$ as the *non-constant* part of the integrand, i.e.

$$f_\perp = \operatorname{argmin}\{\|g\|_{\mathcal{W}} : g = f - c \in \mathcal{W},\ c \text{ is a constant} \in \mathcal{W}\}.$$

If $\mathcal{W}$ is a Hilbert space, then $f_\perp$ is the part of f that is orthogonal to constant functions. Since quadrature rule (1) is exact for constants, the quadrature error for f is the same as that for $f_\perp$, i.e. $\operatorname{Err}(f;P) = \operatorname{Err}(f_\perp;P)$. Let

$$V(f) = \|f_\perp\|_{\mathcal{W}} \tag{10}$$

be called the **variation** of f, i.e. the size of the non-constant part of f. It is shown in (30d) that the variation in the sense of Hardy and Krause in (5) is a special case of the variation just defined.

To determine the quality of the points used for quadrature, one may analyze the error formula, (2), from either a functional analysis or statistics viewpoint and obtain the same result. This section describes both approaches.

For a fixed quadrature rule, the quadrature error $\operatorname{Err}(\cdot;P)$ is a linear functional on the set of integrands $\mathcal{W}$. Suppose that $\operatorname{Err}(\cdot;P)$ is an element of $\mathcal{W}'$, the set of all *bounded* linear functionals on $\mathcal{W}$. Then the norm of this functional is defined as:

$$\|\operatorname{Err}(\cdot;P)\|_{\mathcal{W}'} = \sup_{\substack{f\in\mathcal{W}\\ f\neq 0}} \frac{|\operatorname{Err}(f;P)|}{\|f\|_{\mathcal{W}}}.$$

This implies a quadrature error bound of the form

$$|\operatorname{Err}(f;P)| \le D(P)V(f), \tag{11}$$

where the **discrepancy**, $D(P)$, is defined as the norm of the error functional

$$D(P) = \|\operatorname{Err}(\cdot;P)\|_{\mathcal{W}'}\,. \tag{12}$$

From a statistical point of view suppose that $\mathcal{M}$ is a Banach space of signed measures, with a norm

$$\|G(x)\|_{\mathcal{M}} = \sup_{\substack{f\in\mathcal{W}\\ f\neq 0}} \frac{\left|\int_{[0,1)^s} f(x)\ dG(x)\right|}{\|f\|_{\mathcal{W}}} \quad \forall G \in \mathcal{M}.$$

Assuming that $F_{\text{unif}} \in \mathcal{M}$ and that all point mass measures $F_{\{z\}}$ are in $\mathcal{M}$, this norm defines a **goodness-of-fit statistic**,

$$D(P) = \|F_{\text{unif}} - F_P\|_{\mathcal{M}}\,, \tag{13}$$

which measures how far F_P is from F_{unif}. Again one may conclude that (11) holds. For further discussion relating the discrepancy to goodness-of-fit statistics and robust experimental designs see [Hic99,YH99].

The quadrature error bounds arrived at by the functional analysis and statistics approaches are the same — only the interpretations are different. Furthermore, error bound (11) is tight. The following theorem summarizes these results.

Theorem 1. *Suppose that $\mathcal{W}$ is some Banach space of integrands with norm $\|\cdot\|_{\mathcal{W}}$, and that the variation or non-constant part of an integrand, $V(f)$, is defined as in* (10)*. Let the discrepancy, $D(P)$, be defined equivalently as the norm of the error functional,* (12)*, or as a goodness-of-fit statistic,* (13)*. If the discrepancy is finite, then error bound* (11) *holds, and the worst possible quadrature error for an integrand with unit variation is the discrepancy, i.e.*

$$\sup_{\substack{f\in\mathcal{W}\\ V(f)=1}} |\text{Err}(f;P)| = D(P).$$

The general error bound given in this theorem is shown in Sect. 5 to include the Koksma-Hlawka inequality, (3), and the traditional error bound for lattice rules, (7), as special cases. However, first there is a more fundamental question to be answered.

4 Reproducing Kernel Hilbert Spaces

Theorem 1 is abstract because it gives no explicit, computable formula for the discrepancy. However, an explicit formula can easily be obtained if the space of integrands, $\mathcal{W}$, is a Hilbert space with a reproducing kernel $K(x,y)$. Because the Hilbert space is uniquely determined by its reproducing kernel, we denote the inner product and norm for this Hilbert space as $\langle\cdot,\cdot\rangle_K$ and $\|\cdot\|_K$, respectively. (See [Aro50,Sai88,Wah90] for more details about reproducing kernel Hilbert spaces.) A reproducing kernel is a function with the following property:

$$f(x) = \langle f, K(\cdot,x)\rangle_K \quad \forall x\in[0,1)^s,\ f\in\mathcal{W}, \tag{14}$$

that is, $K(\cdot,x)$ is the representer for the functional that evaluates a function at a point x. A reproducing kernel is symmetric in its arguments and positive definite:

$$K(x,y) = K(y,x) \quad \forall x,y\in[0,1)^s, \tag{15a}$$

$$\sum_{i,k} a^{(i)}a^{(k)}K(x^{(i)},x^{(k)}) \geq 0 \quad \forall a^{(i)}\in\mathbf{R},\ x^{(i)}\in[0,1)^s. \tag{15b}$$

Moreover, any function satisfying (15) is the reproducing kernel for some unique Hilbert space [Wah90, Chap. 1].

The reproducing kernel is related to $f_y(x)$, the function with unit norm that has the largest possible value at y, i.e.

$$f_y(x) = \text{argmax}\{f(y) : f \in \mathcal{W},\ \|f\|_K = 1\}.$$

This function defines the reproducing kernel as follows [YH99]:

$$K(x,y) = f_y(y)f_y(x) = f_x(y)f_x(x), \qquad f_y(x) = \frac{K(x,y)}{\sqrt{K(y,y)}}.$$

The variation of a function defined in (10), when $\mathcal{W}$ is a Hilbert space can be calculated in terms of the inner product. The "non-constant" part of f is either f itself, or the part of f that is orthogonal to the function 1:

$$f_\perp = \begin{cases} f & \text{if } 1 \notin \mathcal{W}, \\ f - 1\langle f, 1\rangle_K / \|1\|_K^2 & \text{if } 1 \in \mathcal{W}. \end{cases}$$

The variation of f is then the norm of the non-constant part of f:

$$V(f) = V(f;K) = \|f_\perp\|_K \tag{16}$$
$$= \begin{cases} \|f\|_K & \text{if } 1 \notin \mathcal{W}, \\ \left(\|f\|_K^2 - \langle f, 1\rangle_K^2 / \|1\|_K^2\right)^{1/2} & \text{if } 1 \in \mathcal{W}. \end{cases}$$

Suppose that the reproducing kernel satisfies the following integrability assumption:

$$\int_{[0,1)^s} K(x,x)\ dx < \infty. \tag{17}$$

This implies that the integration functional, $I(\cdot)$, is bounded because

$$\begin{aligned} |I(f)| &= \left| \int_{[0,1)^s} f(x)\ dx \right| \le \int_{[0,1)^s} |\langle f, K(\cdot,x)\rangle_K|\ dx \\ &\le \int_{[0,1)^s} \|f\|_K \|K(\cdot,x)\|_K\ dx = \|f\|_K \int_{[0,1)^s} \sqrt{K(x,x)}\ dx \\ &\le \|f\|_K \sqrt{\int_{[0,1)^s} K(x,x)\ dx}, \end{aligned}$$

where we have used the fact that the norm of the reproducing kernel is:

$$\|K(\cdot,x)\|_K = \sqrt{\langle K(\cdot,x), K(\cdot,x)\rangle_K} = \sqrt{K(x,x)} \ge 0.$$

The quadrature functional, $Q(\cdot;P)$, is bounded because

$$\begin{aligned} |Q(f;P)| &= \left| \frac{1}{N} \sum_{z\in P} f(z) \right| \le \frac{1}{N} \sum_{z\in P} |\langle f, K(\cdot,z)\rangle_K| \\ &\le \frac{1}{N} \sum_{z\in P} \|f\|_K \|K(\cdot,z)\|_K = \|f\|_K \left[\frac{1}{N} \sum_{z\in P} \sqrt{K(z,z)} \right]. \end{aligned}$$

Therefore, the error functional, $\mathrm{Err}(\cdot; P)$ is also bounded.

By the Riesz Representation Theorem there exists a $\xi \in \mathcal{W}$ which is the representer of the error functional, i.e. $\mathrm{Err}(f; P) = \langle \xi, f \rangle_K$ for all $f \in \mathcal{W}$. Furthermore, the reproducing kernel allows us to compute that representer explicitly:

$$\xi(x) = \langle \xi, K(\cdot, x) \rangle_K = \mathrm{Err}(K(\cdot, x); P).$$

The Cauchy-Schwarz inequality then implies that

$$|\mathrm{Err}(f; P)| = |\mathrm{Err}(f_\perp; P)| = |\langle \xi, f_\perp \rangle_K| \leq \|\xi\|_K \, \|f_\perp\|_K = D(P)V(f). \tag{18}$$

The discrepancy is the norm of the error functional by (12), or equivalently the norm of its representer, ξ. Again using of the reproducing kernel, we can obtain the following explicit formula for the discrepancy:

$$\begin{aligned} D(P) = D(P; K) &= \|\mathrm{Err}(\cdot; P)\|_{\mathcal{W}'} = \|\xi\|_K = \sqrt{\langle \xi, \xi \rangle_K} = \sqrt{\mathrm{Err}(\xi)} \\ &= \left\{ \int_{[0,1)^{2s}} K(x, y) \; d(F_{\mathrm{unif}} - F_P)(x) \; d(F_{\mathrm{unif}} - F_P)(y) \right\}^{1/2} \\ &= \left\{ \int_{[0,1)^{2s}} K(x, y) \; dx \; dy - \frac{2}{N} \sum_{z \in P} \int_{[0,1)^s} K(z, y) \; dy \right. \\ &\qquad \left. + \frac{1}{N^2} \sum_{z,t \in P} K(z, t) \right\}^{1/2}. \end{aligned} \tag{19}$$

Therefore, if the reproducing kernel is known, then the discrepancy may be computed in at most $\mathcal{O}(N^2)$ operations. For some cases, such as the $\mathcal{L}_2$-star discrepancy, there are more efficient algorithms [Hei96]. Since any symmetric, positive definite function, $K(x, y)$, is the reproducing kernel for some Hilbert space of integrands, one may bypass the description of $\mathcal{W}$ and the inner product, and jump straight to the computation of the discrepancy once $K(x, y)$ is chosen. The following theorem summarizes the results of this section:

Theorem 2. *Suppose that $\mathcal{W}$ is a Hilbert space with reproducing kernel $K(x, y)$ that satisfies condition* (17)*. Then the quadrature error is bounded as*

$$|\mathrm{Err}(f; P)| \leq D(P; K)V(f; K)$$

in terms of the discrepancy of P defined in (19) *and the variation of f defined in* (16)*. This bound is attained for the worst-case integrand, ξ, defined above.*

The reproducing kernel methods presented in this section only apply when the space of integrands is a Hilbert space. When the space of integrands, $\mathcal{W}$, described in the previous section is a Banach space, there is no known direct method for deriving a practical formula for the discrepancy. However, in many cases one can invoke a Hölder's type of inequality instead of the

Cauchy-Schwarz inequality in (18) to obtain quadrature error bounds for Banach spaces of integrands. The next section provides some examples of this approach.

5 Examples of Error Bounds and Discrepancies

There are a number of different error bounds and discrepancies that have been proposed in the literature. In this section we briefly summarize them. For further examples of reproducing kernel Hilbert spaces for which Theorem 2 can be applied see [Wah90].

5.1 Integrands Written as Fourier Series

Consider the case where the integrand, f, has an absolutely convergent Fourier series as in (6). Suppose that $\mathcal{W}_{\mathcal{F}}$ is a Hilbert space of integrands with the inner product and associated norm defined as follows:

$$\langle f, g \rangle_{\mathcal{F}} \equiv \sum_{k \in \mathbf{Z}^s} [w(k)]^2 \hat{f}(k) \hat{g}^*(k), \tag{20a}$$

$$\|f\|_{\mathcal{F}} = \left\{ \sum_{k \in \mathbf{Z}^s} [w(k)]^2 \left| \hat{f}(k) \right|^2 \right\}^{1/2} = \left\| \left(w(k) \hat{f}(k) \right)_{k \in \mathbf{Z}^s} \right\|_2, \tag{20b}$$

where * denotes the complex conjugate, and $\|\cdot\|_p$ is the ℓ_p-norm. The quantity $w(k)$ is some positive-valued weight function that satisfies the following summability condition:

$$\left\| \left(\frac{1}{w(k)} \right)_{k \in \mathbf{Z}^s} \right\|_2 = \left\{ \sum_{k \in \mathbf{Z}^s} \frac{1}{[w(k)]^2} \right\}^{1/2} < \infty. \tag{21}$$

The subscript $\mathcal{F}$ in (20) indicates that the inner product and norm are defined in terms of the Fourier coefficients. The space of integrands, $\mathcal{W}_{\mathcal{F}} \equiv \{f : \|f\|_{\mathcal{F}} < \infty\}$, is a Hilbert space of real-valued, periodic functions, whose Fourier coefficients decay to zero sufficiently fast. The reproducing kernel for this space is:

$$K_{\mathcal{F}}(x, y) = \sum_{k \in \mathbf{Z}^s} \frac{e^{2\pi i k'(x-y)}}{[w(k)]^2}. \tag{22}$$

Condition (21) insures that the series defining $K_{\mathcal{F}}(x, y)$ is absolutely convergent. The fact that $K_{\mathcal{F}}(x, y)$ is the reproducing kernel for the space $\mathcal{W}_{\mathcal{F}}$ can be checked by noting that $K_{\mathcal{F}}(\cdot, y) \in \mathcal{W}_{\mathcal{F}}$ for all y and that

$$\langle K_{\mathcal{F}}(\cdot, y), f \rangle_{\mathcal{F}} = \sum_{k \in \mathbf{Z}^s} [w(k)]^2 \frac{e^{-2\pi i k' y}}{[w(k)]^2} \hat{f}^*(k) = \sum_{k \in \mathbf{Z}^s} e^{2\pi i k' y} \hat{f}(k) = f(y).$$

Theorem 2 provides a quadrature error bound of the form $|\mathrm{Err}(f;P)| \le D_{\mathcal{F}}(P)V_{\mathcal{F}}(f)$, where the discrepancy given by (19) is:

$$D_{\mathcal{F}}(P) = D(P;K_{\mathcal{F}}) = \left\{ \sum_{\substack{k\in\mathbf{Z}^s \\ k\neq 0}} \frac{1}{[w(k)]^2} \left[\frac{1}{N^2} \sum_{z,t\in P} e^{2\pi i k'(z-t)} \right] \right\}^{1/2} \tag{23}$$
$$= \left\| \left(\frac{1}{w(k)} \left[\frac{1}{N} \sum_{z\in P} e^{2\pi i k' z} \right] \right)_{0\neq k\in\mathbf{Z}^s} \right\|_2 .$$

and the variation given by (16) is

(24)
$$V_{\mathcal{F}}(f) = V(f;K_{\mathcal{F}}) = \left\{ \sum_{\substack{k\in\mathbf{Z}^s \\ k\neq 0}} [w(k)]^2 \left|\hat{f}(k)\right|^2 \right\}^{1/2} = \left\| \left(w(k)\hat{f}(k) \right)_{0\neq k\in\mathbf{Z}^s} \right\|_2 .$$

Note that the constant part of the integrand, $\hat{f}(0)$, does not enter the definition of the variation because the quadrature rule integrates constants exactly. Likewise the constant term in the reproducing kernel does not enter the definition of the discrepancy.

The quadrature error for a particular integrand with an absolutely convergent Fourier series is simply the sum of the quadrature errors of each term:

$$\mathrm{Err}(f;P) = -\sum_{\substack{k\in\mathbf{Z}^s \\ k\neq 0}} \hat{f}(k) \left[\frac{1}{N} \sum_{z\in P} \mathrm{e}^{2\pi i k' z} \right],$$

where again the term corresponding to $k = 0$ does not enter because constants are integrated exactly. One can multiply and divide by $w(k)$ inside this sum. Then by applying Hölder's inequality one has a generalization of the error bound above:

$$|\mathrm{Err}(f;P)| \le D_{\mathcal{F},p}(P)V_{\mathcal{F},q}(f), \quad \frac{1}{p} + \frac{1}{q} = 1, \tag{25a}$$

$$D_{\mathcal{F},p}(P) = \left\| \left(\frac{1}{w(k)} \left[\frac{1}{N} \sum_{z\in P} e^{2\pi i k' z} \right] \right)_{0\neq k\in\mathbf{Z}^s} \right\|_p , \tag{25b}$$

$$V_{\mathcal{F},q}(f) = \left\| \left(w(k)\hat{f}(k) \right)_{0\neq k\in\mathbf{Z}^s} \right\|_q . \tag{25c}$$

In order to insure that the discrepancy is finite we assume that $\left\|(1/w(k))_{k\in\mathbf{Z}^s}\right\|_p$ is finite. The case of $p = q = 2$ corresponds to the error bound, discrepancy and variation derived above using the reproducing kernel $K_{\mathcal{F}}(x,y)$.

If P is the node set of an integration lattice, then [SJ94, Lemma 2.7]

$$\frac{1}{N}\sum_{z\in P} e^{2\pi i k' z} = 1_{\{k\in L^\perp\}},$$

where the dual lattice, $L^\perp$, is defined in (8). Thus, for lattices the definition of discrepancy above may be simplified to

$$D_{\mathcal{F},p}(P) = \left\| \left(\frac{1}{w(k)} \right)_{0\neq k\in L^\perp} \right\|_p .$$

Certain explicit choices of $w(k)$ have appeared in the literature. For example, one may choose

$$w(k) = \left[\overline{(\beta_1^{-1}k_1)} \cdots \overline{(\beta_s^{-1}k_s)} \right]^\alpha, \quad \alpha > 0, \tag{26}$$

where the over-bar notation is defined in (9), and $\beta_1, \ldots, \beta_s$ are arbitrary positive weights. Suppose that $\beta_1 = \cdots = \beta_s = 1$ and P is the node set of a lattice. Then for $p = 1$ the error bound in (25) corresponds to the traditional error bound for lattice rules, (7), and $D_{\mathcal{F},1}(P) = P_\alpha(L)$. For $1 \le p < \infty$ the discrepancy is $D_{\mathcal{F},p}(P) = [P_{\alpha p}(L)]^{1/p}$, and for $p = \infty$, the discrepancy is $D_{\mathcal{F},\infty}(P) = [\rho(L)]^{-\alpha}$, where $\rho(L)$ is the **Zaremba figure of merit** for lattice rules [Nie92, Definition 5.31].

The more general case of P not a lattice is considered in [Hic96b], and the inclusion of unequal weights β_j is discussed in [Hic98b]. If the weight function $w(k)$ takes the form (26) for *positive integer* α, then one has an error bound of the form:

$$|\mathrm{Err}(f;P)| \le D_{\mathcal{F}}(P) V_{\mathcal{F}}(f), \quad \forall f \in \mathcal{W}_{\mathcal{F},\alpha}, \tag{27a}$$

where the integrands lie in the Hilbert space of periodic functions whose α-order mixed partial derivatives are square integrable:

$$\mathcal{W}_{\mathcal{F},\alpha} = \left\{ f : \frac{\partial^{\alpha|u|} f}{\partial x_u^\alpha} \in \mathcal{L}_2([0,1)^s) \text{ and } \int_0^1 \frac{\partial^{\gamma|u|} f}{\partial x_u^\gamma}\, dx_j = 0 \right. \\ \left. \text{for all } \gamma = 1, \ldots, \alpha,\ j \in u,\ \emptyset \subset u \subseteq S \right\}. \tag{27b}$$

The variation defined in (24) may be written in terms of the α-order mixed partial derivatives of the integrand:

$$V_{\mathcal{F}}(f) = \left\{ \sum_{\emptyset\subset u\subseteq S} \left[\left(\prod_{j\in u} \frac{1}{4\pi^2\beta_j^2} \right)^\alpha \left\| \int_{[0,1)^{S-u}} \frac{\partial^{\alpha|u|} f}{\partial x_u^\alpha}\, dx_{S-u} \right\|_2^2 \right] \right\}^{1/2}. \tag{27c}$$

The discrepancy defined in (23) is

$$D_{\mathcal{F}}(P) = \left\{ -1 + \frac{1}{N^2} \sum_{z,t \in P} \prod_{j=1}^{s} \left[1 - \frac{(-4\pi^2 \beta_j^2)^\alpha}{(2\alpha)!} B_{2\alpha}(\{z_j - t_j\}) \right] \right\}^{1/2}, \tag{27d}$$

for general sets P, where $B_{2\alpha}$ denotes the Bernoulli polynomial of degree 2α [AS64, Chap. 23]. The notation $\{x\}$ means the fractional part of the scalar or vector x, i.e. $\{x\} = x \mod 1$. When P is the node set of an integration lattice, the double sum in this discrepancy can be simplified to a single sum:

$$D_{\mathcal{F}}(P) = \left\{ -1 + \frac{1}{N} \sum_{z \in P} \prod_{j=1}^{s} \left[1 - \frac{(-4\pi^2 \beta_j^2)^\alpha}{(2\alpha)!} B_{2\alpha}(z_j) \right] \right\}^{1/2}. \tag{27e}$$

Another choice for $w(k)$ is a weighted ℓ_2-norm of the vector k to some power:

$$w(k) = \left[\left(\frac{k_1}{\beta_1} \right)^2 + \cdots + \left(\frac{k_s}{\beta_s} \right)^2 \right]^\alpha, \quad \alpha > 0,$$

again for arbitrary positive weights $\beta_1, \ldots, \beta_s$. When these weights are unity, P is the node set of a lattice, and $p = \infty$, then

$$D_{\mathcal{F},\infty}(P) = \max_{0 \neq k \in L^\perp} \|k\|_2^{-\alpha} = \left\{ \min_{0 \neq k \in L^\perp} \|k\|_2 \right\}^{-\alpha}.$$

This discrepancy is equivalent to the **spectral test**, which is used to measure the quality of linear congruential pseudo-random number generators. The spectral test has been used to generate lattices for quasi-Monte Carlo quadrature by L'Écuyer and his collaborators [EHL99,L'É99,LL98,LL99].

5.2 Periodic Integrands

Another family of error bounds for periodic integrands is derived in [Hic98a]. This is a generalization of (27) and is derived without reference to Fourier series.

For ease of presentation we extend the notation for the ℓ_p- and $\mathcal{L}_p$-norms. Let (f_u) denote a vector of functions indexed by $u \subseteq S$. The range of u is typically $\emptyset \subseteq u \subseteq S$ or $\emptyset \subset u \subseteq S$ and will be clear from the context. Then the norm $\|(f_u)\|_p$ is defined as follows:

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

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 scalar function, $\|(f_u)\|_p$ corresponds to the $\mathcal{L}_p$-norm.

Let $\mu(t)$ be any periodic function of one variable whose α-order derivative, $\mu^{(\alpha)}(t)$, is essentially bounded on the interval $[0,1)$ with $\int_0^1 \mu^{(\gamma)}(t)\,dt = 0$ for $\gamma = 0, 1, \ldots, \alpha$. For ease of notation let

$$M = \int_0^1 [\mu^{(\alpha)}(t)]^2\,dt. \tag{28a}$$

Define a reproducing kernel, which is a product of one-dimensional kernels, of the form:

$$K_{\mathrm{per},\alpha}(x,y) = \prod_{j=1}^{s} \Big\{ 1 + \tilde{\beta}_j^{2\alpha} \left[M + \mu(x_j) + \mu(y_j) - \frac{(-1)^\alpha}{(2\alpha)!} B_{2\alpha}(\{x_j - y_j\}) \right] \Big\}, \tag{28b}$$

where the $\tilde{\beta}_j$ are arbitrary positive weights. Theorem 2 provides a worst-case error bound. This can be generalized to an error bound of the form given in Theorem 1 by applying Hölder's inequality instead of the Cauchy-Schwarz inequality [Hic98a]:

$$|\mathrm{Err}(f;P)| \le D_{\mathrm{per},p}(P) V_{\mathrm{per},q}(f) \quad \text{for } f \in \mathcal{W}_{\mathrm{per},q},\ \frac{1}{p} + \frac{1}{q} = 1, \tag{28c}$$

$$\mathcal{W}_{\mathrm{per},q} = \left\{ f : \frac{\partial^{\alpha|u|} f}{\partial x_u^\alpha} \in \mathcal{L}_q([0,1)^s) \text{ and } \int_0^1 \frac{\partial^{\gamma|u|} f}{\partial x_u^\gamma}\,dx_j = 0 \text{ for all } \gamma = 1,\ldots,\alpha,\ j \in u,\ \emptyset \subset u \subseteq S \right\}, \tag{28d}$$

$$V_{\mathrm{per},q}(f) = \left\| \left(\left[\prod_{j\in u} \tilde{\beta}_j^{-\alpha} \right] \times \int_{[0,1)^{S-u}} \left\{ \prod_{j \in S-u} \left[1 - \mu^{(\alpha)}(x_j) \frac{\partial^\alpha}{\partial x_j^\alpha} \right] \right\} \frac{\partial^{\alpha|u|} f}{\partial x_u^\alpha}\,dx_{S-u} \right)_{\emptyset \subset u \subseteq S} \right\|_q, \tag{28e}$$

$$D_{\mathrm{per},p}(f) = \left\| \left(\left[\prod_{j\in u} \tilde{\beta}_j^{\alpha} \right] \left\{ \prod_{j\in u} \mu^{(\alpha)}(x_j) - \frac{1}{N} \sum_{z\in P} \prod_{j\in u} \left[\mu^{(\alpha)}(x_j) - \frac{(-1)^\alpha}{\alpha!} B_\alpha(\{x_j - z_j\}) \right] \right\} \right)_{\emptyset \subset u \subseteq S} \right\|_p. \tag{28f}$$

For $p = 2$, the formula for the discrepancy is a special case of (19) and may be written as follows:

$$\text{(28g)} \quad D_{\text{per},2}(P) = \left\{ \prod_{j=1}^{s} \left[1 + \tilde{\beta}_j^{2\alpha} M\right] - \frac{2}{N} \sum_{z \in P} \prod_{j=1}^{s} \left[1 + \tilde{\beta}_j^{2\alpha} \left\{M + \mu(z_j)\right\}\right] \right.$$
$$\left. + \frac{1}{N^2} \sum_{z,t \in P} \prod_{j=1}^{s} \left[1 + \tilde{\beta}_j^{2\alpha} \left\{M + \mu(z_j) + \mu(t_j) - \frac{(-1)^\alpha}{(2\alpha)!} B_{2\alpha}(\{z_j - t_j\})\right\}\right] \right\}^{1/2}.$$

The case of $p = 2$ and $\mu = 0$ is equivalent to the error bound (27), and $D_{\mathcal{F}}(P) = D_{\text{per},2}(P)$ in this case.

5.3 Non-Periodic Integrands

The error bounds described in the previous subsections are for periodic integrands. However, many integrands are not periodic. The family of error bounds described in the previous subsection has an analog for non-periodic integrands when $\alpha = 1$, which includes the Koksma-Hlawka inequality, (3), as a special case.

Let $\mu(t)$ now be any function of one variable whose first order derivative, $\mu'(t)$, is essentially bounded on the interval $[0, 1)$ and for which $\int_0^1 \mu(t)\, dt = 0$. Again, for ease of notation let

$$\text{(29a)} \qquad M = \int_0^1 [\mu'(t)]^2 \, dt.$$

Define a reproducing kernel which is a product of one-dimensional kernels of the form:

$$\text{(29b)} \quad K_{\text{non}}(x, y) = \prod_{j=1}^{s} \left\{1 + \hat{\beta}_j^2 \left[M + \mu(x_j) + \mu(y_j) \right.\right.$$
$$\left.\left. + \frac{1}{2} B_2(\{x_j - y_j\}) + B_1(x_j) B_1(y_j)\right]\right\},$$

where the $\hat{\beta}_j$ are arbitrary positive weights. Theorem 2 provides a worst-case error bound. This can be generalized by applying Hölder's inequality instead of the Cauchy-Schwarz inequality to obtain a family of error bounds [Hic98a]:

$$\text{(29c)} \qquad |\text{Err}(f; P)| \le D_{\text{non},p}(P) V_{\text{non},q}(f) \quad \text{for } f \in \mathcal{W}_{\text{non},q}, \ \frac{1}{p} + \frac{1}{q} = 1,$$

$$\text{(29d)} \qquad \mathcal{W}_{\text{non},q} = \left\{ f : \frac{\partial^{|u|} f}{\partial x_u} \in \mathcal{L}_q([0,1)^s), \quad \emptyset \subset u \subseteq S \right\},$$

$$\text{(29e)}\quad V_{\text{non},q}(f) = \left\| \left(\left[\prod_{j\in u} \hat{\beta}_j^{-1} \right] \times \int_{[0,1)^{S-u}} \left\{ \prod_{j\in S-u} \left[1 - \mu'(x_j) \frac{\partial}{\partial x_j} \right] \right\} \frac{\partial^{|u|} f}{\partial x_u}\, dx_{S-u} \right)_{\emptyset\subset u\subseteq S} \right\|_q ,$$

$$\text{(29f)}\quad D_{\text{non},p}(f) = \left\| \left(\left[\prod_{j\in u} \hat{\beta}_j \right] \times \left\{ \prod_{j\in u} \mu'(x_j) - \frac{1}{N} \sum_{z\in P} \prod_{j\in u} \left[\mu'(x_j) + x_j - 1_{\{x_j \geq z_j\}} \right] \right\} \right)_{\emptyset\subset u\subseteq S} \right\|_p .$$

For $p = 2$, the formula for the discrepancy is a special case of (19) and may be written as follows:

$$\begin{aligned}\text{(29g)}\quad D_{\text{non},2}(P) = \Bigg\{ &\prod_{j=1}^{s} \left[1 + \hat{\beta}_j^2 M \right] - \frac{2}{N} \sum_{z\in P} \prod_{j=1}^{s} \left[1 + \hat{\beta}_j^2 \left(M + \mu(z_j) \right) \right] \\ &+ \frac{1}{N^2} \sum_{z,t\in P} \prod_{j=1}^{s} \Big[1 + \hat{\beta}_j^2 \{ M + \mu(z_j) + \mu(t_j) + \\ &\qquad \frac{1}{2} B_2(\{z_j - t_j\}) + B_1(z_j) B_1(t_j) \Big\} \Big] \Bigg\}^{1/2},\end{aligned}$$

5.4 Geometric Interpretations of Discrepancies

Star Discrepancy. The discrepancies (28f) and (29f) presented in the previous subsections involve the arbitrary function $\mu(t)$. For certain choices of $\mu(t)$ the discrepancy has a nice geometric interpretation. For example, the star discrepancy in (4) is simply the $\mathcal{L}_\infty$-norm of

$$\text{(30a)}\quad R^*(x) = \int_{[0,1)^s} 1_{[0,x]}(y)\; d(F_{\text{unif}} - F_P)(y) = \text{Vol}([0,x]) - \frac{|P \cap [0,x]|}{N},$$

which is the difference in the measure of the region $[0, x]$ according to the uniform distribution and the empirical distribution for P. In other words, $R^*(x)$ is the difference of the volume of the region $[0, x]$ and the proportion of points in P that are also in $[0, x]$. Here the notation $|A|$ means the number of points in the set A. Figure 1 illustrates the geometric meaning of $R^*(x)$. An $\mathcal{L}_p$-star discrepancy may be defined in terms of the $\mathcal{L}_p$-norm of $R^*(x)$.

Suppose that $\mu(t) = 1/6 - t^2/2$. Then the discrepancy defined in (29f) becomes the $\mathcal{L}_p$-star discrepancy, which may be written in terms of the $\mathcal{L}_p$-norm of $R^*(x)$ [Hic98a]:

$$D_p^*(P) = \left\| \left(\left[\prod_{j \in u} \hat{\beta}_j \right] R^*(x_u, 1) \right)_{\emptyset \subset u \subseteq S} \right\|_p . \tag{30b}$$

Here, $R^*(x_u, 1)$ means that all the components x_j with $j \notin u$ are set equal to one in (30a). For $p = \infty$ and all $\hat{\beta}_j = 1$ this discrepancy is the same as the star discrepancy defined in (4). For $p < \infty$, this discrepancy involves $\|R^*(x_u, 1)\|_p$, which is the traditional $\mathcal{L}_p$-star discrepancy of the projection of P into the $|u|$-dimensional unit cube $[0, 1)^u$ for $\emptyset \subset u \subseteq S$. The $\mathcal{L}_2$-star discrepancy takes the form

$$D_2^*(P) = \left\{ \prod_{j=1}^{s} \left(1 + \frac{\hat{\beta}_j^2}{3} \right) - \frac{2}{N} \sum_{z \in P} \prod_{j=1}^{s} \left[1 + \frac{\hat{\beta}_j^2 (1 - z_j^2)}{2} \right] + \frac{1}{N^2} \sum_{z,t \in P} \prod_{j=1}^{s} \left(1 + \hat{\beta}_j^2 [1 - \max(z_j, t_j)] \right) \right\}^{1/2} , \tag{30c}$$

which is similar to, though not the same as, the formula derived by Warnock [War72]. The variation defined in (29e) for the above choice of $\mu(t)$ is

$$V_q^*(f) = \left\| \left(\left[\prod_{j \in u} \hat{\beta}_j^{-1} \right] \frac{\partial^{|u|}}{\partial x_u} f(x_u, 1) \right)_{\emptyset \subset u \subseteq S} \right\|_q . \tag{30d}$$

The star discrepancy and variation given above are generalizations of the traditional definitions (4) and (5) presented in Sect. 2.

Although the function $R^*(x)$ defining the $\mathcal{L}_p$-star discrepancy is simple, it favors a certain corner of the unit cube. Thus, the $\mathcal{L}_p$-star discrepancy lacks the following property:

Definition 3. A discrepancy, $D(P)$, is said to be **reflection invariant** if and only if $D(P) = D(P^{(1)}) = \cdots = D(P^{(s)})$, where $P^{(j)}$ is obtained by reflecting the points in P about the plane $x_j = 1/2$ passing through the center of the cube, that is

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

The discrepancies (28f) and (29f) for certain other choices of $\mu(t)$ are reflection invariant. Moreover, they can also be interpreted geometrically as the $\mathcal{L}_p$-norm of a function defined analogously to $R^*(x)$, where the box $[0, x]$ is replaced by another subset of the unit cube. Several examples of these geometric interpretations are provided below.

Centered Discrepancy. For any two points $x, y \in [0,1]^s$ let $J(x,y)$ denote the box containing the points between x and y, i.e.

$$J(x,y) = \{(t_1,\ldots,t_s) : \min(x_j,y_j) \le t_j \le \max(x_j,y_j)\ \forall j = 1,\ldots,s\}.$$

The set of vertices of the unit cube is $\{0,1\}^s$, and $b = (1/2,\ldots,1/2)$ is the center of the cube $[0,1)^s$. There are 2^s cubes of the form $J(a,b)$ for $a \in \{0,1\}^s$.

Let $a_x \in \{0,1\}^s$ be defined as the vertex of $[0,1)^s$ closest to x, that is, $x \in J(a_x,b)$. (This definition is ambiguous when any $x_j = 1/2$, but since such cases lie in a set of measure zero, they can be ignored.)

$$R^c(x) = \mathrm{Vol}(J(a_x,x)) - \frac{|P \cap J(a_x,x)|}{N}, \tag{31a}$$

and let $R^c_u(x_u)$ be the analog of $R^c(x)$ for the cube $[0,1)^u$, and for P_u, the projection of P into the cube $[0,1)^u$. Figure 2 illustrates the geometric meaning of $R^c(x)$. The $\mathcal{L}_p$-centered discrepancy is obtained by choosing $\mu(t) = -B_2(\{t - 1/2\})/2$ and can be written in terms of $R^c_u(x_u)$ [Hic98a]:

$$D^c_p(f) = \left\| \left(\left[\prod_{j\in u} \hat\beta_j \right] R^c_u(x_u) \right)_{\emptyset \subset u \subseteq S} \right\|_p . \tag{31b}$$

The $\mathcal{L}_2$-centered discrepancy takes the form

$$\begin{aligned} D^c_2(P) = \Bigg\{ & \prod_{j=1}^s \left(1 + \frac{\hat\beta_j^2}{12}\right) \\ & - \frac{2}{N} \sum_{z\in P} \prod_{j=1}^s \left[1 + \frac{\hat\beta_j^2}{2}\left(\left|z_j - \frac{1}{2}\right| - \left|z_j - \frac{1}{2}\right|^2\right)\right] \\ & + \frac{1}{N^2} \sum_{z,t\in P} \prod_{j=1}^s \left[1 + \frac{\hat\beta_j^2}{2}\left(\left|z_j - \frac{1}{2}\right| + \left|t_j - \frac{1}{2}\right| - |z_j - t_j|\right)\right] \Bigg\}^{1/2} . \end{aligned} \tag{31c}$$

The variation corresponding to this choice of $\mu(t)$ is

$$V^c_q(f) = \left\| \left(\left[\prod_{j\in u} \hat\beta_j^{-1} \right] \frac{\partial^{|u|}}{\partial x_u} f(x_u, 1/2) \right)_{\emptyset \subset u \subseteq S} \right\|_q . \tag{31d}$$

Symmetric Discrepancy. For any $x \in [0,1)^s$ there are 2^s boxes of the form $J(a,x)$ for $a \in \{0,1\}^s$. Let $J_e(x)$ denote the union of the *even* boxes $J(a,x)$,

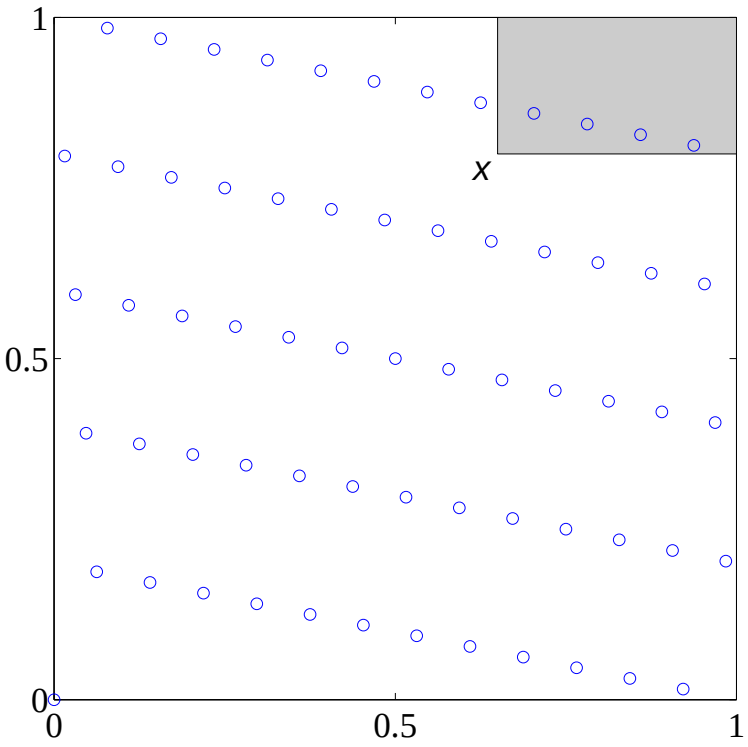

Fig. 2. For $x = (0.65, 0.8)$ the volume of the box $J(a_x, x)$ is 0.07, and the proportion of points in this box is $4/64 = 0.0625$. The difference between these two is $R^c(x) = \mathrm{Vol}\left(J(a_x, x)\right)) - \frac{1}{64}|P \cap J(a_x, x)| = 0.0075$.

that is, those for which the sum of the vertex coordinates, $a_1 + \cdots + a_s$, is even, and let $J_o(x)$ denote the union of the *odd* boxes $J(a, x)$. Let

$$R^s(x) = \mathrm{Vol}(J_e(x)) - \frac{|P \cap J_e(x)|}{N} = -\mathrm{Vol}(J_o(x)) + \frac{|P \cap J_o(x)|}{N}. \tag{32a}$$

The $\mathcal{L}_p$-symmetric discrepancy can be written in terms of $R^s(x)$, and is obtained by choosing $\mu(t) = -B_2(t)/2$ in (29f) [War95,Hic98a]:

$$D_p^s(f) = 2 \left\| \left(\left[\prod_{j \in u} \frac{\hat{\beta}_j}{2} \right] R^s(x_u, 1) \right)_{\emptyset \subset u \subseteq S} \right\|_p . \tag{32b}$$

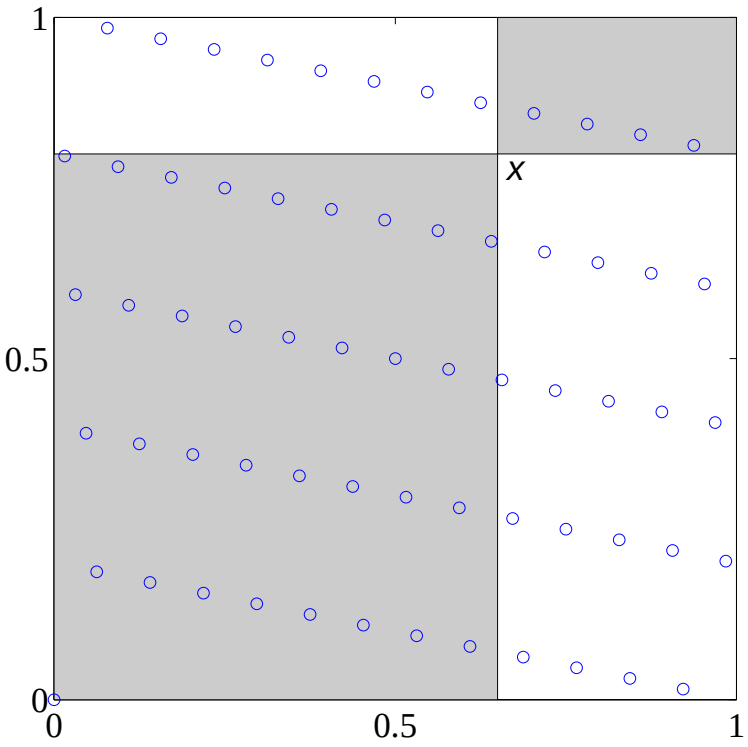

Fig. 3. For $x = (0.65, 0.8)$ the total volume of the even rectangles is 0.59, and the proportion of points in these rectangles is $38/64 = 0.59375$. The difference between these two is $R^s(x) = \mathrm{Vol}\,(J_e(x)) - \frac{1}{64}|P \cap J_e(x)| = -0.00375$.

The $\mathcal{L}_2$-symmetric discrepancy takes the form

$$D_2^s(P) = \left\{ \prod_{j=1}^{s} \left(1 + \frac{\hat{\beta}_j^2}{12}\right) - \frac{2}{N} \sum_{z \in P} \prod_{j=1}^{s} \left[1 + \frac{\hat{\beta}_j^2}{2} z_j (1 - z_j)\right] + \frac{1}{N^2} \sum_{z,t \in P} \prod_{j=1}^{s} \left[1 + \frac{\hat{\beta}_j^2}{4} (1 - 2|z_j - t_j|)\right] \right\}^{1/2}. \tag{32c}$$

The variation corresponding to this choice of $\mu(t)$ is

$$V_q^s(f) = \left\| \left(\left[\prod_{j \in u} \frac{2}{\hat{\beta}_j} \right] 2^{-s} \sum_{a \in \{0,1\}^{S-u}} \frac{\partial^{|u|}}{\partial x_u} f(x_u, a) \right)_{\emptyset \subset u \subseteq S} \right\|_q. \tag{32d}$$

$\mathcal{L}_2$-Unanchored Discrepancy. One may generalize the star discrepancy in a geometric way by replacing the box $[0, x]$ appearing in (30a) by the box $[x', x]$. That is, let

$$R^u(x', x) = \mathrm{Vol}([x', x]) - \frac{|P \cap [x', x]|}{N}. \tag{33a}$$

Figure 4 illustrates the geometric meaning of $R^u(x', x)$, which is used to define the $\mathcal{L}_p$-unanchored discrepancy (see [Nie92, Definition 2.2] and [MC94]). In contrast to the previous discrepancies, the $\mathcal{L}_p$-unanchored discrepancy corresponds to one of the general discrepancies in Sects. 5.2 and 5.3 only for $p = 2$. Whereas the star discrepancy is a special case of the analysis for non-periodic integrands in Sect. 5.3, the $\mathcal{L}_2$-unanchored discrepancy is a special case of the analysis for *periodic* integrands in Sect. 5.2. Choosing $\mu = -\frac{1}{2}B_2(t)$ and $\alpha = 1$ in (28g) gives

$$\begin{aligned} D_2^u(P) = \Bigg\| \Bigg(\left[\prod_{j \in u} \hat{\beta}_j \right] & \\ \times \left\{ \int_{[0,1)^{2u}, x'_u \le x_u} [R^u((x'_u, 0), (x_u, 1))]^2 \, dx'_u \, dx_u \right\}^{1/2} \Bigg) & \Bigg\|_{\emptyset \subset u \subseteq S} \Bigg\|_2 \\ = \Bigg\{ \prod_{j=1}^{s} \left(1 + \frac{\hat{\beta}_j^2}{12} \right) - \frac{2}{N} \sum_{z \in P} \prod_{j=1}^{s} & \left[1 + \frac{\hat{\beta}_j^2 z_j (1 - z_j)}{2} \right] \\ + \frac{1}{N^2} \sum_{z,t \in P} \prod_{j=1}^{s} & \Big(1 + \hat{\beta}_j^2 [\min(z_j, t_j) - z_j t_j] \Big) \Bigg\}^{1/2}. \end{aligned} \tag{33b}$$

The corresponding variation is

$$V_2^u(f) = \left\{ \sum_{\emptyset \subset u \subseteq S} \left[\prod_{j \in u} \hat{\beta}_j^{-2} \right] \left\| \frac{\partial^{|u|}}{\partial x_u} f(x_u, 1) \right\|_2^2 \right\}^{1/2}. \tag{33c}$$

$\mathcal{L}_2$-Wrap-Around Discrepancy. Another discrepancy that is similar to the unanchored discrepancy is defined by counting points in a "box" $[x'_u, x_u]$,

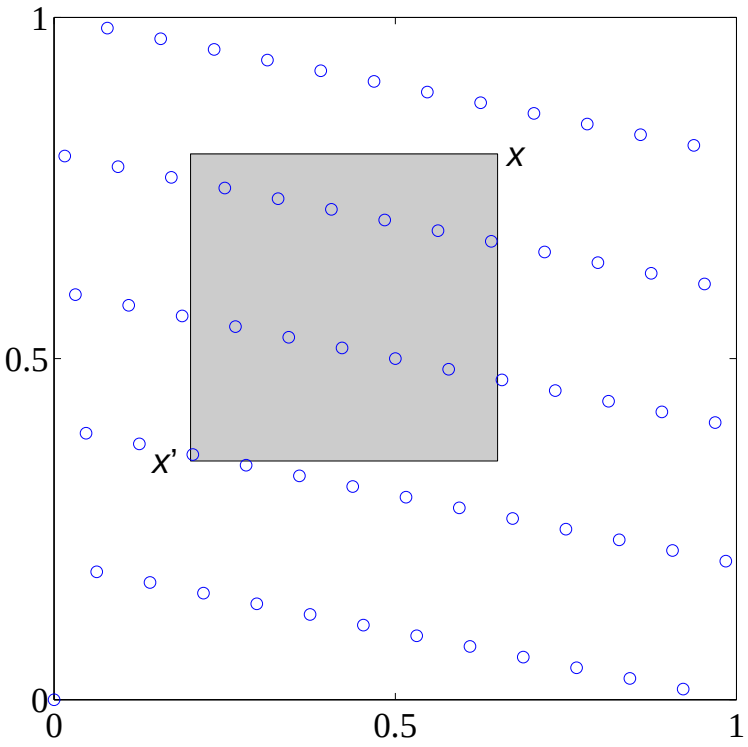

Fig. 4. For $x' = (0.2, 0.35)$ and $x = (0.65, 0.8)$ the volume of the box is 0.2025, and the proportion of points in the box is $12/64 = 0.1875$. The difference between these two is $R^u(x', x) = (x_1 - x'_1)(x_2 - x'_2) - \frac{1}{64}|P \cap [x', x]| = 0.015$.

allowing for wrap-around when $x'_j > x_j$ (see Fig. 5). Specifically, let

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

$$J_w(x', x) = \bigotimes_{j=1}^{s} J_w(x'_j, x_j),$$

$$R^w(x', x) = \mathrm{Vol}(J_w(x', x)) - \frac{|P \cap J_w(x', x)|}{N} \tag{34a}$$

The volume of $J_w(x', x)$ is $\prod_{j=1}^{s} \{x_j - x'_j\}$. The $\mathcal{L}_2$-wrap-around discrepancy is then defined analogously to the unanchored discrepancy, but with $\mu(t) = 0$

in (28g):

$$\text{(34b)} \qquad D_2^w(P) \equiv \left\| \left(\left[\prod_{j \in u} \hat{\beta}_j \right] \left\{ \int_{[0,1)^{2s}} \left[R^w(x', x) \right]^2 \, dx' dx \right\}^{1/2} \right)_{u \neq \emptyset} \right\|_2$$

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

The corresponding variation is

$$\text{(34c)} \quad V_2^w(f) = \left\{ \sum_{\emptyset \subset u \subseteq S} \left[\prod_{k \in S-u} \left(1 + \frac{\hat{\beta}_k^2}{3} \right) \right] \left[\prod_{j \in u} \hat{\beta}_j^{-2} \right] \left\| \int_{[0,1)^{S-u}} \frac{\partial^{|u|} f}{\partial x_u} \, dx_{S-u} \right\|_2^2 \right\}^{1/2}.$$

The star discrepancy has a nice geometric interpretation, but so do several other discrepancies introduced in this section. However, the star discrepancy has certain disadvantages. It does not have the reflection invariance property of Definition 3. The $\mathcal{L}_\infty$-star discrepancy, which is the most popular of the $\mathcal{L}_p$-star discrepancies, is more difficult to compute than the $\mathcal{L}_2$-star discrepancy. Moreover, because the Koksma-Hlawka inequality requires only moderate smoothness of the integrand, the decay of the discrepancy with respect to N is rather slow. As seen in Sect. 8, discrepancies based on smoother kernels can decay significantly faster with respect to N. To summarize the previous three sections, there are many other useful error bounds available other than the traditional ones of Sect. 2.

Reality Check 1. *The Koksma-Hlawka inequality* (3) *and the traditional error bound for lattice rules* (7) *are popular error bounds for quasi-Monte Carlo quadrature that have dominated the literature.* However*, they are just two cases of a large family of error bounds, and may not be the best in all situations.*

6 Average-Case Error Analysis

The error analysis in Sects. 3 and 4 is a worst-case analysis, i.e. a tight upper bound is given on the worst possible error for integrands in a space $\mathcal{W}$. It is also possible to analyze the quadrature error assuming that the integrand is a **random function** drawn from some sample space $\mathcal{A}$. The quadrature error

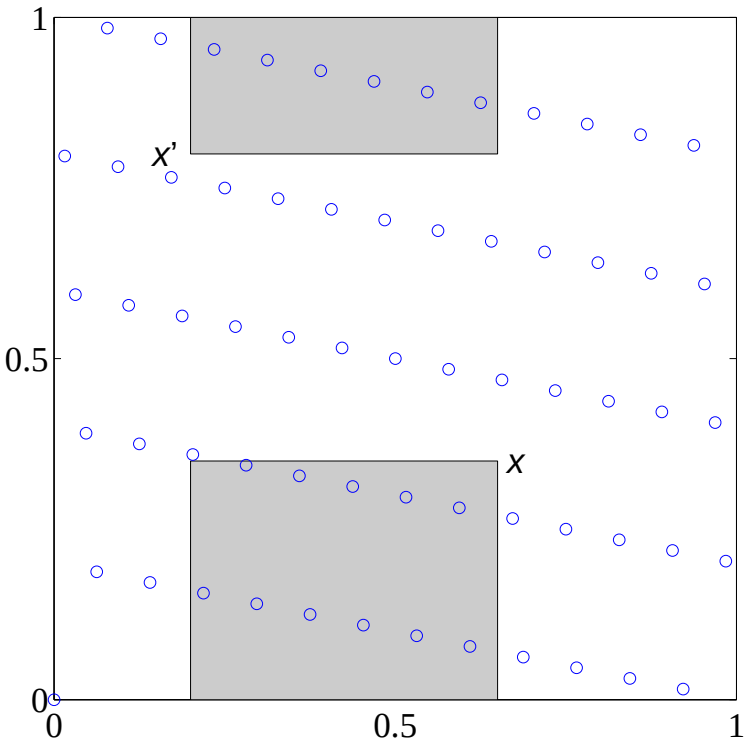

Fig. 5. For $x' = (0.2, 0.8)$ and $x = (0.65, 0.35)$ the volume of the wrapped-around box is 0.2475, and the proportion of points in the box is $17/64 = 0.265625$. The difference between these two is $R^w(x', x) = \{x_1 - x'_1\}\{x_2 - x'_2\} - \frac{1}{64}|P \cap J_w(x', x)| = -0.018125$.

analysis only requires that we know the point-wise mean and covariance of the random integrand. Suppose that

$$E_{f\in\mathcal{A}}[f(x)] = 0, \qquad E_{f\in\mathcal{A}}[f(x)f(y)] = K(x, y), \tag{35}$$

where $K(x, y)$ now plays the role of a covariance kernel. The root mean square quadrature error for a random integrand is given in the following theorem. The proof proceeds by interchanging the order of integration and the expectation (see, e.g. [Rit96,HH99]).

Theorem 4. *Suppose that $\mathcal{A}$ is a sample space of random functions satisfying* (35) *for some covariance kernel, K. Then the root mean square quadrature error for quadrature rule* (1) *is the discrepancy as defined in* (19)*:*

$$\sqrt{E_{f\in\mathcal{A}}\{[\mathrm{Err}(f; P)]^2\}} = D(P; K).$$

The worst-case and average-case analyses both yield a discrepancy as the measure of the quality of P, or equivalently of the quadrature rule. If the reproducing kernel for the worst-case analysis is the same as the covariance kernel for the average-case analysis, then the two discrepancies are the same, but $\mathcal{W}$ is a much smaller space than $\mathcal{A}$. In fact, $\mathcal{W}$, is a subset of $\mathcal{A}$ with measure zero. There are some important quantities involving the integrals of the reproducing or covariance kernel that have interpretations in both the worst-case and average-case error analyses. If $\|I\|_K$ denotes the norm of the integration functional on a reproducing kernel Hilbert space, then

$$\|I\|_K^2 = E_{f\in\mathcal{A}}\{[I(f)]^2\} = \int_{[0,1)^{2s}} K(x,y)\ dx\ dy.$$

Moreover, if the variance of a function is defined as $\mathrm{Var}(f) = I(f^2) - [I(f)]^2$, it can be shown that

$$\sup_{\substack{f\in\mathcal{W}\\ \|f\|_K=1}} \{\mathrm{Var}(f)\} \geq E_{f\in\mathcal{A}}\{\mathrm{Var}(f)\}$$
$$= \int_{[0,1)^s} K(x,x)\ dx - \int_{[0,1)^{2s}} K(x,y)\ dx\ dy.$$

7 The Similarity of Reproducing Kernels and Discrepancies

As illustrated in Sects. 4 and 5 there are many different possible choices of reproducing kernels, $K(x,y)$, and for each there exists a discrepancy. This might make it difficult to know which kernel to choose. In this section we show that the mean square discrepancies for certain random point sets are rather independent of the details of the kernel function.

The Monte Carlo and quasi-Monte Carlo quadrature literatures provide examples of random transformations applied to quasi-Monte Carlo points. For example, simple Monte Carlo, can be thought of as giving each point in P a random shift (mod 1) *independently* of every other point. The Cranley-Patterson shifts [CP76] give each point in P the *same* shift (mod 1). Owen [Owe95,Owe97a,Owe97b,Owe98,Owe99] describes a random scrambling of points that is especially useful for nets.

Let us consider the root mean square discrepancies of random sets of points, where the discrepancies are derived from reproducing kernels according to (19). For simple Monte Carlo sampling the root mean square discrepancy is given by the following theorem [Hic95,Hic96b], whose proof follows in a straightforward manner from the definition of the discrepancy in (19):

Theorem 5. *Let P_{MC} be a simple Monte Carlo sample, i.e. a set of N independently distributed random points chosen according to the uniform distribution over $[0,1)^s$. The root mean square of the discrepancy defined in* (19)

for this random sample is

$$\sqrt{E[D(P_{MC};K)]^2}$$
$$= \frac{1}{N^{1/2}} \left\{ \int_{[0,1)^s} K(x,x)\ dx - \int_{[0,1)^{2s}} K(x,y)\ dx\ dy \right\}^{1/2}$$
$$= \frac{1}{N^{1/2}} \sqrt{E_{f\in\mathcal{A}}\{\mathrm{Var}(f)\}}.$$

Note that the root mean square discrepancy is independent of the detailed form of the kernel K, but depends only on the difference of two integrals. Thus, the quality of a simple random sample for quadrature does not depend on the detailed structure of the corresponding spaces of integrands $\mathcal{W}$ and $\mathcal{A}$.

For an arbitrary kernel K one may define an associated **Monte Carlo-invariant** kernel K_{MC} as follows:

$$K_{\mathrm{MC}}(x,y) = \begin{cases} \int_{[0,1)^{2s}} K(z,t)\ dz\ dt, & x \neq y, \\ \int_{[0,1)^s} K(z,z)\ dz, & x = y. \end{cases}$$

The kernel K_{MC} takes a constant value on the diagonal plane $x = y$, and another constant value off the diagonal plane. Thus, the discrepancy (19) defined by the kernel K_{MC} for any set of points depends only on how many of the points are duplicates. In particular, for all possible ordered pairs of points $(t,z) \in P \times P$, suppose that n of these have $t = z$. Obviously, $N \leq n \leq N^2$, with $n = N$ if the points in P are distinct, and $n > N$ if P contains the same point more than once. Then the discrepancy of P defined by the kernel K_{MC} is

$$D(P;K_{\mathrm{MC}}) = \frac{n^{1/2}}{N} \left\{ \int_{[0,1)^s} K(x,x)\ dx - \int_{[0,1)^{2s}} K(x,y)\ dx\ dy \right\}^{1/2}.$$

The reason we choose to call K_{MC} Monte Carlo-invariant is that the corresponding discrepancy of a point set remains the same (with probability one) even when each point is given an arbitrary shift modulo one:

Corollary 6. *For any point set $P = \{z^{(1)}, \ldots, z^{(N)}\}$, let $\mathcal{T}_{MC}$ be a transformation that gives each point in P a random shift (mod 1), i.e.*

$$\mathcal{T}_{MC}(P) = \{\{z^{(1)} + \Delta^{(1)}\}, \ldots, \{z^{(N)} + \Delta^{(N)}\}\},$$

where the $\Delta^{(i)}$ are random variables uniformly distributed on $[0,1)^s$. For any $i,l = 1,\ldots,N$ assume that $\Delta^{(i)} = \Delta^{(l)}$ if $z^{(i)} = z^{(l)}$, and assume that $\Delta^{(i)}$ and $\Delta^{(l)}$ are uncorrelated otherwise. Then for any Monte Carlo-invariant kernel defined above

$$D(\mathcal{T}_{MC}(P);K_{MC}) = D(P;K_{MC}) \text{ with probability one.}$$

Furthermore, if P is a set of distinct *points, then $\mathcal{T}_{MC}(P)$ is a simple Monte Carlo sample, and*

$$\sqrt{E[D(P_{MC};K)]^2} = \sqrt{E[D(P;K_{MC})]^2}.$$

There are two points to note about this corollary. Firstly, for certain (Monte Carlo-invariant) kernels, any (Monte Carlo) randomization of the points leaves the discrepancy unchanged. Secondly, the root mean square discrepancy for an arbitrary kernel of an arbitrary set that has undergone (Monte Carlo) randomization, is equivalent to the discrepancy for a particular (Monte Carlo-invariant) kernel of an arbitrary set. These two ideas are extended to the case of Cranley-Patterson shifts in Theorem 7 below. Extension to the case of Owen's scramblings Owen [Owe95] is work in progress.

Consider the three main kernels introduced in Sect. 5, namely (22), (28b), and (29b). The Monte Carlo-invariant kernels associated to these three kernels are

$$K_{\mathrm{MC},\mathcal{F}}(x,y) = \frac{1}{[w(0)]^2} + 1_{\{x=y\}} \sum_{0\neq k\in\mathbf{Z}^s} \frac{1}{[w(k)]^2}, \tag{36a}$$

$$K_{\mathrm{MC,per},\alpha}(x,y) = \begin{cases} \prod_{j=1}^s \left\{1+\tilde{\beta}_j^{2\alpha} M\right\}, & x\neq y, \\ \prod_{j=1}^s \left\{1+\tilde{\beta}_j^{2\alpha}\left[M - \frac{(-1)^\alpha}{(2\alpha)!}B_{2\alpha}(0)\right]\right\}, & x=y, \end{cases} \tag{36b}$$

$$K_{\mathrm{MC,non}}(x,y) = \begin{cases} \prod_{j=1}^s \left\{1+\hat{\beta}_j^{2} M\right\}, & x\neq y, \\ \prod_{j=1}^s \left\{1+\hat{\beta}_j^{2}\left[M + \frac{1}{6}\right]\right\}, & x=y, \end{cases}, \tag{36c}$$

respectively.

Computing the discrepancy defined by any of the kernels defined in (36) gives an answer that is *dependent on the dimension*. This may seem puzzling, since Monte Carlo is known to be a dimension-independent quadrature method. The reason for the dimension dependence is that the kernels as defined do not keep the variance of the integrands bounded as the dimension increases. Specifically, if $\mathcal{A}$ is a sample space of random integrands with covariance kernel K given by (22), (28b), and (29b), then $E_{f\in\mathcal{A}}\{\mathrm{Var}(f)\}$ is dimension-dependent. However, this can be rectified by defining a suitable normalized kernel. For any kernel K, let

$$\tilde{K}(x,y) = K(x,y)\left[\int_{[0,1)^s} K(z,z)\,dz - \int_{[0,1)^{2s}} K(z,t)\,dz\,dt\right]^{-1}. \tag{37}$$

This scaling insures that if $\tilde{\mathcal{A}}$ is a sample space of random integrands with covariance kernel $\tilde{K}$, then $E_{f\in\tilde{\mathcal{A}}}\{\mathrm{Var}(f)\} = 1$ and $\sqrt{E[D(P_{\mathrm{MC}};K)]^2} = N^{-1/2}$, a dimension-independent result. The effect of this normalization is to multiply the original discrepancy by a constant that depends on s, but not on N. The variation is multiplied by the inverse of this constant, which leaves the worst-case error bound unchanged.

Reality Check 2. *The error of Monte Carlo quadrature may be dimension-independent,* but only if *the variance of the integrand is independent of dimension.*

Now, consider the Cranley-Patterson shifts. Analogous to Corollary 6 we have a result [Hic98b] that the root mean square discrepancy under such shifts is just the discrepancy according to a shift-invariant kernel of the original set P.

Theorem 7. *For any set P, let $\mathcal{T}_{sh}(P) = \{\{z + \Delta\} : z \in P\}$ denote a shifted copy of P, where Δ is a random variable distributed uniformly on $[0,1)^s$. Then*

$$\sqrt{E_{\mathcal{T}_{sh}}\{[D(\mathcal{T}_{sh}(P); K)]^2\}} = D(P; K_{sh}),$$

where the **shift-invariant** *kernel is defined as:*

$$K_{sh}(x,y) = \int_{[0,1)^s} K(\{x+z\},\{y+z\})\, dz = \int_{[0,1)^s} K(\{x-y+z\}, z)\, dz.$$

Again, consider the three main kernels introduced earlier, namely (22), (28b), and (29b). Their shift-invariant counterparts are:

$$K_{\mathrm{sh},\mathcal{F}}(x,y) = K_{\mathcal{F}}(x,y) = \sum_{k\in\mathbf{Z}^s} \frac{e^{2\pi i k'(x-y)}}{[w(k)]^2},$$

$$K_{\mathrm{sh,per},\alpha}(x,y) = \prod_{j=1}^{s} \left\{1 + \tilde{\beta}_j^{2\alpha}\left[M - \frac{(-1)^\alpha}{(2\alpha)!} B_{2\alpha}(\{x_j - y_j\})\right]\right\},$$

$$K_{\mathrm{sh,non}}(x,y) = \prod_{j=1}^{s} \left\{1 + \hat{\beta}_j^2\,[M + B_2(\{x_j - y_j\})]\right\}.$$

Note that $K_{\mathcal{F}}(x,y)$ is itself shift-invariant, and, in fact, any shift-invariant kernel takes the form of $K_{\mathcal{F}}(x,y)$. This is because a general shift-invariant kernel, $K_{\mathrm{sh}}(x,y)$ must be a function of $\{x-y\}$. Thus, it can be written as a Fourier series $K_{\mathrm{sh}}(x,y) = \sum_{k\in\mathbf{Z}^s} \hat{K_{\mathrm{sh}}}(k) e^{2\pi i k'(x-y)}$. The Fourier coefficients, $\hat{K_{\mathrm{sh}}}(k)$, are non-negative because the kernel is positive definite. So, the general shift-invariant kernel, $K_{\mathrm{sh}}(x,y)$, is the same as $K_{\mathcal{F}}(x,y)$ with $\hat{K_{\mathrm{sh}}}(k) = 1/[w(k)]^2$.

The kernel $K_{\mathrm{sh,per},\alpha}(x,y)$ is a special case of $K_{\mathcal{F}}(x,y)$ where $w(k)$ is chosen according to (26) and α is a positive integer. The kernel $K_{\mathrm{sh,non}}(x,y)$ is a special case of $K_{\mathrm{sh,per},\alpha}(x,y)$ with $\alpha = 1$. Neither of the shift-invariant kernels $K_{\mathrm{sh,non}}(x,y)$ nor $K_{\mathrm{sh,per},\alpha}(x,y)$ depends on the arbitrary function $\mu(t)$ appearing in the definitions of $K_{\mathrm{non}}(x,y)$ and $K_{\mathrm{per},\alpha}(x,y)$. The shift-invariant kernels for all the kernels introduced in Sect. 5.4 are essentially the same. Thus, if a set of points has a low $\mathcal{L}_2$-wrap-around discrepancy, then, on average, random shifts of that set will have low $\mathcal{L}_2$-star, $\mathcal{L}_2$-unanchored, $\mathcal{L}_2$-symmetric, and $\mathcal{L}_2$-centered discrepancies.

8 The Dependence of the Discrepancy on the Number of Points and the Dimension

Having described different discrepancies we now turn to the question of how the discrepancy depends on N, s and other factors. The literature contains asymptotic upper bounds on the attainable star discrepancy of the form $c(s,b)N^{-1}(\log N)^{s-1}$ [Nie92] for certain sets. Similar bounds exist for $P_\alpha(L)$, the traditional figure of merit for integration lattices [Nie92,SJ94]. However, such bounds are not useful for large dimensions because the power of $\log N$ becomes quite substantial. In this section we focus on examples where the actual discrepancy can be calculated in a reasonable number of operations.

8.1 Dissecting an Error Bound

In order to understand what affects the size of the discrepancy let us look more closely at a particular error bound derived in Sect. 5. Choosing $\alpha = 1$, $\mu = 0$, $\tilde{\beta}_1 = \cdots = \tilde{\beta}_s = \tilde{\beta}$, and quadrature by lattice rules, error bound (28) becomes:

$$|\mathrm{Err}(f;P)| \le D_{\mathrm{per},2}(P) V_{\mathrm{per},2}(f) \quad \text{for } f \in \mathcal{W}_{\mathrm{per},2}, \tag{38a}$$

$$\mathcal{W}_{\mathrm{per},2} = \left\{ f : \frac{\partial^{|u|} f}{\partial x_u} \in \mathcal{L}_2([0,1)^s) \text{ and } \int_0^1 \frac{\partial^{|u|} f}{\partial x_u}\, dx_j = 0 \right. \\ \left. \text{for all } j \in u,\ \emptyset \subset u \subseteq S \right\}, \tag{38b}$$

$$D_{\mathrm{per},2}(P) = \left\{ -1 + \frac{1}{N} \sum_{z \in P} \prod_{j=1}^{s} \left[1 + \frac{\tilde{\beta}^2}{2} B_2(z_j) \right] \right\}^{1/2}, \tag{38c}$$

$$V_{\mathrm{per},2}(f) = \left\{ \sum_{\emptyset \subset u \subseteq S} \tilde{\beta}^{-2|u|} \left\| \int_{[0,1)^{S-u}} \frac{\partial^{|u|} f}{\partial x_u}\, dx_{S-u} \right\|_2^2 \right\}^{1/2}. \tag{38d}$$

A closer look at this error bound shows that it can be derived in several steps [Hic98a]. First, one writes the error as the inner product of the worst-case integrand, ξ, with $f_\perp$. After expanding the inner product as a sum over u, the absolute value is taken inside the sum. Next, the Cauchy-Schwarz inequality is applied to each integral in the sum. The u^{th} term in this sum is multiplied and divided by $\tilde{\beta}^{|u|}$. Finally, the Cauchy-Schwarz inequality is

applied to the sum to obtain the product $D_{\mathrm{per},2}(P)V_{\mathrm{per},2}(f)$:

$$
\begin{aligned}
(39)\quad & |\mathrm{Err}(f;P)| = |\langle \xi, f_\perp \rangle_K| \\
&= \left| \sum_{\emptyset \subset u \subseteq S} \int_{[0,1)^u} \left\{ \left[\frac{-1}{N} \sum_{z \in P} \prod_{j \in u} B_1(\{x_j - z_j\}) \right] \int_{[0,1)^{S-u}} \frac{\partial^{|u|} f}{\partial x_u} \, dx_{S-u} \right\} dx_u \right| \\
&\le \sum_{\emptyset \subset u \subseteq S} \left| \int_{[0,1)^u} \left\{ \left[\frac{-1}{N} \sum_{z \in P} \prod_{j \in u} B_1(\{x_j - z_j\}) \right] \int_{[0,1)^{S-u}} \frac{\partial^{|u|} f}{\partial x_u} \, dx_{S-u} \right\} dx_u \right| \\
&\le \sum_{\emptyset \subset u \subseteq S} \left\{ \frac{1}{N} \sum_{z \in P} \prod_{j \in u} \left[\frac{1}{2} B_2(z_j) \right] \right\}^{1/2} \left\| \int_{[0,1)^{S-u}} \frac{\partial^{|u|} f}{\partial x_u} \, dx_{S-u} \right\|_2 \\
&= D_{\mathrm{per},2}(P) V_{\mathrm{per},2}(f).
\end{aligned}
$$

Note that other quadrature error bounds derived in Sects. 5.2 and 5.3 can also be dissected in this manner. See [Hic98a] for details.

8.2 A Quadrature Example

One might wonder how tight this quadrature error bound is in practice. As an indication we consider the following example:

$$
(40)\qquad f(x) = \left(\frac{3.02}{2.53} \right)^s \prod_{j=1}^{s} (1 + x_j - x_j^{0.51}).
$$

The constant in front has been chosen so that the integral of this function over the unit cube is one. This integrand has period interval $[0,1)^s$ and sufficient smoothness to apply (38). Since the integrand is a product of relatively simple functions, the computation of the variation is straightforward.

The integral of $f(x)$ is approximated by a lattice quadrature rule. The rank-1 integration lattice used for this quadrature rule has the following node set:

$$
(41a)\qquad P = \{\{\phi_2(i)h\} : i = 0, \ldots, 2^m - 1\},
$$

$$
(41b)\qquad h = (1, 89715, 89715^2, \ldots, 89715^{s-1}),
$$

where $\phi_2(i)$ is the base 2 Van der Corput sequence. The generating vector h for the lattice is of Korobov type for convenience. Although the above formula may look unusual, it has the advantage of being the node set of a lattice with 2^m points for *any* positive integer m. Lattices of this type have been studied in [HH97].

Figure 6 displays the true absolute quadrature error using lattice rule (41) and the error bound (38) for $s = 5$ and for various choices of the parameters $\tilde{\beta}$. This figure shows that the error bound is pessimistic, in some cases extremely

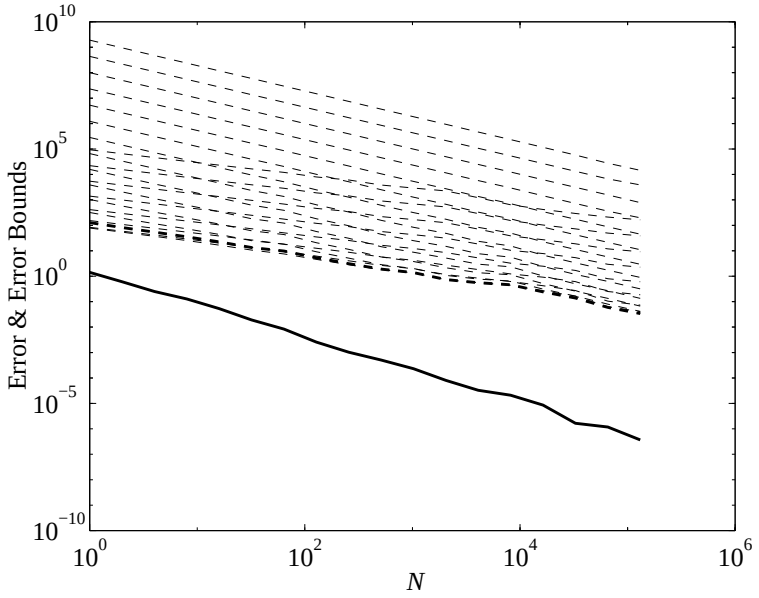

Fig. 6. Actual absolute quadrature error (solid) for integrand (40) using lattice rule (41) and corresponding error bounds (38) (dashed) for $s = 5$, and various choices of $\tilde{\beta}$, $(2\pi)^{-2} \le \tilde{\beta} \le (2\pi)^2$.

so. For smaller values of $\tilde{\beta}$ the slope of the error bound is steeper, and for larger values of $\tilde{\beta}$ it becomes more shallow. As the value of $\tilde{\beta}$ increases the error bound curves move downward for awhile, but eventually move upward.

This behavior can be explained by looking at the intermediate results in (39). Since one sacrifices equality at several steps in (39), it is not surprising that the quadrature error bound is not tight for this integrand. The choice of $\tilde{\beta}$ effects the tightness of the error bound in the last step of (39). In the next to last step of (39) the error is bounded by a sum of terms, each comprised of two factors:

$$\left\{ \frac{1}{N} \sum_{z \in P} \prod_{j \in u} \left[\frac{1}{2} B_2(z_j) \right] \right\}^{1/2}, \tag{42}$$

$$\left\| \int_{[0,1)^{S-u}} \frac{\partial^{|u|} f}{\partial x_u} \, dx_{S-u} \right\|_2 . \tag{43}$$

The first factor depends on the projection of P into the unit cube $[0,1)^u$, while the second factor depends on how much the integrand varies with respect to the variables $x_j,\ j \in u$. The discrepancy is a weighted average of the factors

(42) with weights $\tilde{\beta}^{2|u|}$, and the variation is a weighted average of the factors (43) with weights $\tilde{\beta}^{-2|u|}$.

Suppose that the integrand is approximately a sum of functions of only one or two coordinates at a time, i.e. the terms in (43) with $|u|$ small are bigger than those terms with $|u|$ large. Then, error bound in the last step of (39) is tightest for a *relatively small* value of $\tilde{\beta}$. In this case, the definition of the discrepancy gives heavier weight to the terms in (42) with $|u|$ small, and thus it emphasizes the uniformity of the low dimensional projections of P. If the terms in (43) with $|u|$ large are bigger than those with $|u|$ small, then the opposite occurs. Thus, the choice of $\tilde{\beta}$ that gives the tightest error bounds indicates the relative importance of the low and high dimensional pieces of the integrand.

One could obtain an even tighter bound than (38) by defining the square discrepancy and the square variation as weighted averages of the terms (42) and (43), respectively, with weights $\tilde{\beta}_u^2$ and $\tilde{\beta}_u^{-2}$, respectively. The disadvantage of this approach is that one now has $2^s - 1$ parameters $\tilde{\beta}_u$, rather than just one parameter $\tilde{\beta}$. Sloan, Woźniakowski, and others have explored different choices of weights [SW97,Hic98b,HW98,Woź99]. However, in practice, one can rarely make the best choice of $\tilde{\beta}$ or $\tilde{\beta}_u$ for a particular integrand f. This is because the best choice depends on knowing something about the sizes of the terms (43), which are even more difficult to calculate than the unknown integral of f itself.

Equation (18) shows in a general way that the quadrature error only depends on the component of the integrand that is parallel to the worst-case integrand, ξ. For this reason the error bounds presented in this article may be quite conservative and, therefore, not directly useful for error estimation. The best suggestions for quasi-Monte Carlo error estimation seem to be based on replications via Cranley-Patterson shifts [CP76] or scrambling of nets [Owe95]. Performing r replications based on low discrepancy points in an rs-dimensional unit cube has recently been suggested by Halton and Warnock. Although error bounds may not be too useful for error estimation, they do seem to be useful for comparing different quadrature rules, since the general trend of the quadrature error with N and s often follows the discrepancy.

Reality Check 3. *Error bounds* do not*, in general, tell you about the quadrature error of your particular problem. Rather, they tell you about the overall suitability of a set P for quadrature of a given type of integrands.*

The next two subsections briefly present the discrepancies of some actual point sets. Other numerical investigations of discrepancies are given in [MC94,Hic95,KW97]. In order to make meaningful comparisons across different dimensions we normalize the reproducing kernel according to (37) so that the root mean square discrepancy of a simple random sample is $N^{-1/2}$. This then is the target which we expect low discrepancy sets to beat.

8.3 A Rank-1 Integration Lattice

Consider the rank-1 integration lattice in (41) above. The discrepancy is chosen to be of the form (38c), which is also equivalent to (27e) with $\alpha = 1$. Including the normalization (37) the discrepancy becomes:

$$[D(P)]^2 = \frac{-1 + \frac{1}{N}\sum_{z\in P}\prod_{j=1}^{s}\left[1 + \frac{\tilde{\beta}^2}{2}B_2(z_j)\right]}{-1 + \prod_{j=1}^{s}\left[1 + \frac{\tilde{\beta}^2}{2}B_2(0)\right]}, \tag{44}$$

This discrepancy is a weighted version of the figure of merit for lattice rules $\mathcal{P}_2(L)$.

Figures 7 and 8 plot the discrepancy of lattice (41) and compare it with $N^{-1/2}$, the root mean square discrepancy of a simple random sample. For a lattice consisting of only $N = 1$ point in any dimension, s, the discrepancy is 1, the same as for the simple random sample of one point. For dimension $s = 1$ the discrepancy of the lattice is N^{-1}, which is much better than the discrepancy of the random sample. As the dimension increases the discrepancy of the lattice approaches that of the simple random sample. However, the discrepancy of the lattice is almost always no worse than that of a simple random sample. To understand how this happens, note that (44) may also be written as a sum of two terms as follows:

$$[D(P)]^2 = \frac{1}{N} + \frac{\sum_{\emptyset\subset u\subseteq S}\left(\frac{\tilde{\beta}^2}{2}\right)^{|u|}\left\{\frac{1}{N}\sum_{0\neq z\in P}\left[\prod_{j\in u}B_2(z_j)\right]\right\}}{\sum_{\emptyset\subset u\subseteq S}\left[\frac{\tilde{\beta}^2}{2}B_2(0)\right]^{|u|}}.$$

The more negative the second term is, the better the lattice is in comparison to a simple random sample. The term $\frac{1}{N}\sum_{0\neq z\in P}\left[\prod_{j\in u}B_2(z_j)\right]$ depends only on the projection of P into the $|u|$-dimensional unit cube $[0,1)^u$. This term becomes less negative, or even positive, when $|u|$ is large. This is why as the dimension increases the discrepancy of the lattice approaches that of the simple random sample.

Comparing Figs. 7 and 8, the discrepancies are higher when the parameter $\tilde{\beta}$ is larger. This is because the larger $\tilde{\beta}$ places a heavier emphasis on the higher dimensional terms, that is, those $\frac{1}{N}\sum_{0\neq z\in P}\left[\prod_{j\in u}B_2(z_j)\right]$ with larger $|u|$. Thus, the discrepancy depends not simply on the nominal dimension of the problem, s, but on how heavily the high dimensional components of the discrepancy are weighted compared to the low dimensional components. The size of $\tilde{\beta}$ is also indicative of how large the high dimensional components of the integrand are assumed to be, as compared to the low dimensional components, in defining the space of integrands $\mathcal{W}$ or $\mathcal{A}$. Therefore, some authors have proposed definitions of an effective dimension of the problem [CMO97,Hic98b].

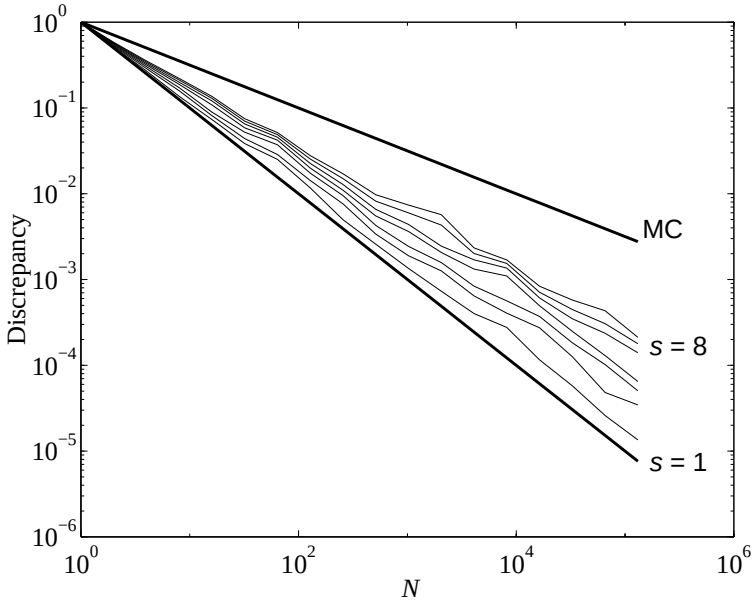

Fig. 7. Discrepancy (44) of the lattice (41) for $N = 1, 2, 4, \ldots, 2^{17}$ points, dimensions $s = 1, \ldots, 8$, and $\tilde{\beta} = 1$. The root mean square discrepancy of a simple random sample is shown for comparison.

Reality Check 4. *Although the error of quasi-Monte Carlo quadrature is somewhat dimension-dependent, this does not necessarily imply that it is inferior to Monte Carlo quadrature. For integrating integrands with low* **effective dimension** *quasi-Monte Carlo is often more accurate than Monte Carlo. When the effective dimension of the integrand is high, quasi-Monte Carlo is often no worse (or not much worse) than Monte Carlo.*

8.4 Randomized $(0, m, s)$-Nets

The discrepancy in (44) requires only $\mathcal{O}(N)$ operations to evaluate because the set P is a lattice. On the other hand, evaluation of the $\mathcal{L}_2$-star discrepancy requires $\mathcal{O}(N^2)$ operations according to a naive algorithm and $\mathcal{O}(N[\log N]^s)$ operations according to the algorithm of Heinrich [Hei96]. However, if one scrambles an $(0, m, s)$-net in base b, as recommended by Owen [Owe95,Owe97a,Owe97b], then the mean square $\mathcal{L}_2$-star discrepancy and other similar discrepancies can be computed in closed form [Hic96a,Hic98b].

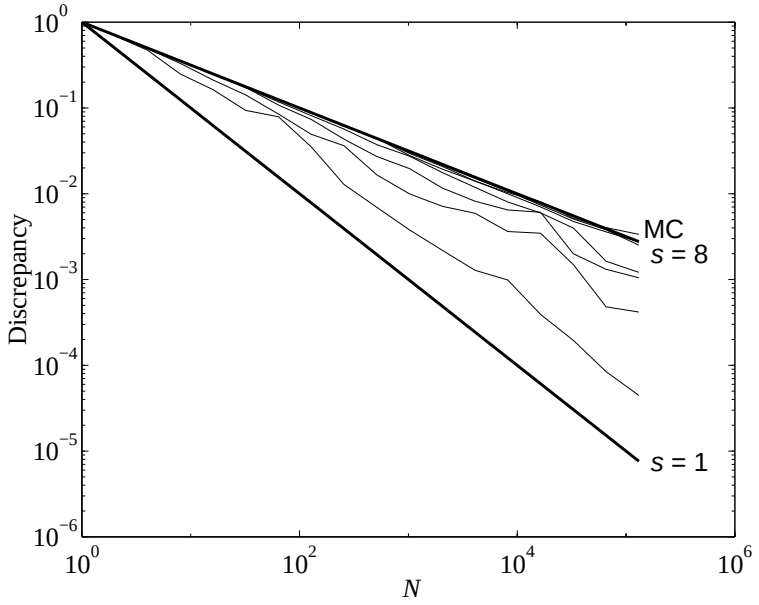

Fig. 8. Discrepancy (44) of the lattice (41) for $N = 1, 2, 4, \ldots, 2^{17}$ points, dimensions $s = 1, \ldots, 8$, and $\tilde{\beta} = 2\pi$. The root mean square discrepancy of a simple random sample is shown for comparison.

The following sum is the mean square $\mathcal{L}_2$-star discrepancy, (30c), with $\hat{\beta}_j = 1$:

$$\left(\frac{3}{2}\right)^s b^{-2m} \sum_{l=0}^{s-1} \sum_{k=0}^{l} \sum_{j=0}^{l-k} \left[\binom{s}{l+1} \binom{m}{k} \binom{l}{j} \left(\frac{1}{9}\right)^{l+1} \times b^{-l} (b^2-1)^k (-b-1)^j \right]. \tag{45}$$

Although it is ugly in appearance, the above formula requires only $\mathcal{O}(s^2 + s \log N)$ operations to evaluate.

After applying the normalization in (37), this discrepancy is compared to the $N^{-1/2}$ discrepancy of a random sample in Fig. 9. This graph is similar to those in Figs. 7 and 8, with the discrepancy of a net approaching that of a simple random sample as the dimension increases. An asymptotic analysis of (45) shows that as $N \to \infty$, the root mean square $\mathcal{L}_2$-star discrepancy is $\mathcal{O}(N^{-1}[\log N]^{(s-1)/2})$. This asymptotic order has been attained in an explicit construction of symmetrized $(0, m, s)$-nets by Larcher [Lar99]. It is known to be the best asymptotic order possible for the $\mathcal{L}_2$-star discrepancy of any set [Rot54]. It is also the asymptotic order of the root mean square quadrature error using randomized $(0, m, s)$-nets derived by Owen [Owe97a].

The arguments leading to (45) have been extended to general discrepancies based on reproducing kernels, i.e. discrepancies of the form (19). It has

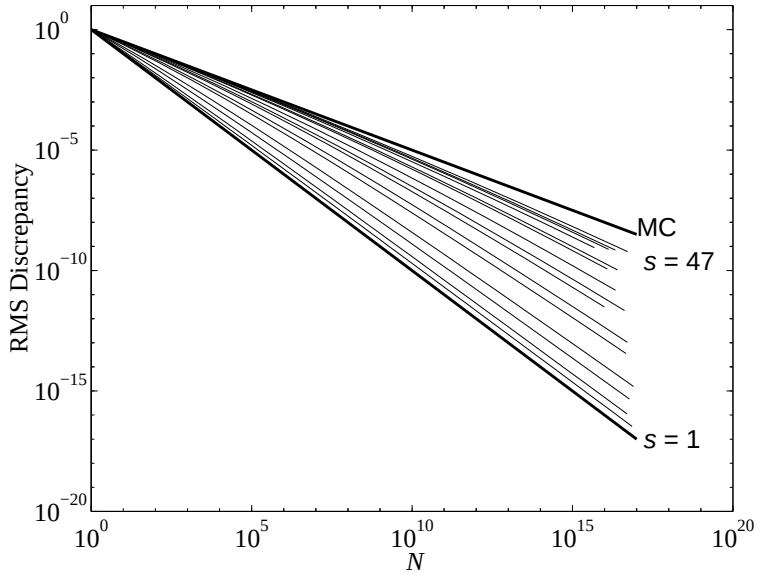

Fig. 9. Root mean square $\mathcal{L}_2$-star discrepancy (45) for randomly scrambled $(0, m, s)$-nets in base s for $N = 1, s, s^2, \ldots$ points and prime dimensions $s = 1, 2, 3, 5, 7, \ldots, 47$. Normalization (37) has been applied.

been found that if the reproducing kernel is *smoother* than the one used to define the $\mathcal{L}_2$-star discrepancy, then the root mean square discrepancy of a scrambled $(0, m, s)$-net is $\mathcal{O}(N^{-1.5}[\log N]^{(s-1)/2})$ [HH99]. Thus, quasi-Monte Carlo methods are able to integrate smoother integrands more accurately than the $\mathcal{O}(N^{-1+\epsilon})$ error one expects from the Koksma-Hlawka inequality. Owen [Owe97b] found this same asymptotic result for smooth integrands using a different approach.

Reality Check 5. *The asymptotic order of quasi-Monte Carlo quadrature error is often quoted as* N^{-1} *times some power of* $\log N$*, with effort being expended to reduce the power of the* $\log N$*. However, quasi-Monte Carlo quadrature error can be as low as* $\mathcal{O}(N^{-1.5+\epsilon})$ *when integrands are a bit smoother. Thus, it is possible to improve the power of* N*, not only the power of* $\log N$*.*

9 Conclusion and Open Questions

Many quasi-Monte Carlo quadrature error bounds exist — all taking the form of a discrepancy times a variation. The discrepancy, which measures the quality of the set of points defining the quadrature rule, depends on the following factors:

– the number of points and how they are chosen,
– the dimension of the problem, and
– the choice of the reproducing kernel, or equivalently, the definition of the normed space of integrands.

At present we have some understanding of what affects the size of discrepancy, and thus we know something about the size of quasi-Monte Carlo quadrature error. However, our understanding is far from complete.

The discrepancy of a good set typically decays like $\mathcal{O}(N^{-1+\epsilon})$ which is significantly faster than the $\mathcal{O}(N^{-1/2})$ decay for a simple random sample. However, for small N the decay rate may look more like $\mathcal{O}(N^{-1/2})$, as is seen in Sect. 8 and also in [MC94,Hic95]. As the dimension of the problem increases, the value of N where the decay switches from $\mathcal{O}(N^{-1/2})$ to $\mathcal{O}(N^{-1+\epsilon})$ increases.

This dimension of the problem is not strictly s. The definition of the reproducing kernel, or equivalently, the normed space of integrands, implies the relative importance of low dimensional effects to high dimensional effects. If the low dimensional effects are given heavier weight, the effective dimension of the problem may be much less than the nominal dimension. In the discrepancies introduced here, the choices of β_j, $\tilde{\beta}_j$, or $\hat{\beta}_j$ affect the effective dimension of the problem. Unfortunately, these parameters cannot be optimally chosen by sampling the integrand. Rather, their values reflect the a priori beliefs of the user.

The smoothness of the integrands, as reflected in the smoothness of the reproducing kernel, also affects the discrepancy. As mentioned in Sect. 8 one may obtain an $\mathcal{O}(N^{-1.5+\epsilon})$ decay in the discrepancy for sufficiently smooth kernels using good $(0, m, s)$-nets. It has been known for some time that periodic integrands with sufficient smoothness can be integrated very accurately by using lattice rules. Although many integrands are not periodic to begin with, there are transformations one can use to periodize them. However, these transformations may greatly increase the relevant variation of the integrand, thereby negating the benefits of periodicity.

Randomizing sets of low discrepancy points using Cranley-Patterson shifts or Owen's scrambling averages out some of the differences among discrepancies defined by different reproducing kernels. This is an extension of the principle in Monte Carlo methods that randomly choosing the sample points removes the bias in your estimator. The challenge in constructing good randomizations for quasi-Monte Carlo methods is to preserve the low discrepancy of the point set while randomizing in a non-trivial way.

In conclusion we present some open problems of personal interest in the study of quasi-Monte Carlo quadrature, several of which are the topics of ongoing research. If you, the reader, have some answers or partial answers, please contact us.

– There is a wealth of results about the existence of sets and sequences with low star discrepancy and the existence of lattices with small $\mathcal{P}_\alpha(L)$.

Which of these results extend to other discrepancies, such as those presented in Sects. 3–5?

- Are there computationally simple formulas for the minimum, average or maximum discrepancies of (t, s)-sequences, digital sequences and lattices, such as that given in (45) for $(0, m, s)$-nets?
- As mentioned in Sect. 8 scrambled $(0, m, s)$-nets give a root mean square discrepancy of $\mathcal{O}(N^{-1.5+\epsilon})$ if the reproducing kernel is sufficiently smooth. However, numerical computation of the discrepancies of known (t, m, s)-nets in [HH99] show only an $\mathcal{O}(N^{-1+\epsilon})$ decay. The midpoint rule for one-dimensional quadrature is an example of using a $(0, m, 1)$-net to obtain $\mathcal{O}(N^{-2})$ quadrature error for sufficiently smooth integrands. Can one explicitly construct (t, m, s)-nets or (t, s)-sequences that give $\mathcal{O}(N^{-1.5+\epsilon})$ or $\mathcal{O}(N^{-2+\epsilon})$ quadrature error for sufficiently smooth integrands?
- For (t, m, s)-nets or (t, s)-sequences in base b the scalar quality factor t becomes less informative for large dimensions, s, that occur, for example, in finance applications. It would be useful to have a vector quality factor $T = (t_u)_{\emptyset \subset u \subseteq S}$, where t_u measures the quality of the projection of the net or sequence into the unit cube $[0, 1)^u$. For some applications one may wish to insure that the points are distributed as well as possible in the first several dimensions, which means one would want to minimize $t_{\{1,\ldots,s'\}}$ for moderate values of s'. For other applications one might wish for all low-dimensional projections of points to be well distributed, which means that one would want to minimize t_u for all $|u|$ small. Can constructions of nets and sequences be designed to achieve such objectives?
- The quasi-Monte Carlo literature contains both upper and lower bounds on the size of the smallest attainable discrepancy. Some recent results are given in [SW97,Hic98b,HW98,Woź99]. Can these bounds be improved to narrow the gap?
- As mentioned in Sect. 8 the effective dimension of the integrand, and not necessarily s, determines how much better quasi-Monte Carlo methods perform as compared to Monte Carlo methods. However, this connection is qualitative. Is there a good definition of effective dimension that will make this connection more precise?

Acknowledgments

I would like to thank Professors Harald Niederreiter and Jerome Spanier for organizing a wonderful MCQMC98 in Claremont and for inviting me to speak. A great deal of gratitude is also due to my professors at nearby Pomona College for launching my research career while I was an undergraduate student. Many thanks to Regina H. S. Hong, Xiaoqun Wang, Rongxian Yue, and an anonymous referee for their valuable suggestions for improving this manuscript.

References

[Aro50] M. Aronszajn, *Theory of reproducing kernels*, Trans. Amer. Math. Soc. **68** (1950), 337–404.

[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.

[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.

[CP76] R. Cranley and T. N. L. Patterson, *Randomization of number theoretic methods for multiple integration*, SIAM J. Numer. Anal. **13** (1976), 904–914.

[EHL99] K. Entacher, P. Hellekalek, and P. L'Écuyer, *Quasi-Monte Carlo node sets from linear congruential generators*, In Niederreiter and Spanier [NS99].

[FW94] K. T. Fang and Y. Wang, *Number-theoretic methods in statistics*, Chapman and Hall, New York, 1994.

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

[Gen93] A. Genz, *Comparison of methods for the computation of multivariate normal probabilities*, Computing Science and Statistics **25** (1993), 400–405.

[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.

[HH99] F. J. Hickernell and H. S. Hong, *The asymptotic efficiency of randomized nets for quadrature*, Math. Comp. **68** (1999), 767–791.

[Hic95] F. J. Hickernell, *A comparison of random and quasirandom points for multidimensional quadrature*, In Niederreiter and Shiue [NS95], 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.

[Hic98a] F. J. Hickernell, *A generalized discrepancy and quadrature error bound*, Math. Comp. **67** (1998), 299–322.

[Hic98b] F. J. Hickernell, *Lattice rules: How well do they measure up?*, Random and Quasi-Random Point Sets (P. Hellekalek and G. Larcher, eds.), Lecture Notes in Statistics, vol. 138, Springer-Verlag, New York, 1998, pp. 109–166.

[Hic99] F. J. Hickernell, *Goodness-of-fit statistics, discrepancies and robust designs*, Statist. Probab. Lett. (1999), to appear.

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

[HW98] F. J. Hickernell and H. Woźniakowski, *Integration and approximation in arbitrary dimensions*, 1998, submitted for publication.

[Kei96] B. D. Keister, *Multidimensional quadrature algorithms*, Computers in Physics **10** (1996), 119–122.

[KW97] L. Kocis and W. J. Whiten, *Computational investigations of low-discrepancy sequences*, ACM Trans. Math. Software **23** (1997), 266–294.

[Lar99] G. Larcher, *Point sets with minimal L_2-discrepancy*, 1999, preprint.

[L'É99] P. L'Écuyer, *Tables of linear congruential generators of different sizes and good lattice structure*, Math. Comp. **68** (1999), 249–260.

[LL98] C. Lemieux and P. L'Écuyer, *Efficiency improvement by lattice rules for pricing asian options*, Proc. 1998 Winter Simulation Conference, IEEE Press, 1998, pp. 579–586.

[LL99] C. Lemieux and P. L'Écuyer, *A comparison of Monte Carlo, lattice rules and other low-discrepancy point sets*, In Niederreiter and Spanier [NS99].

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

[MC95] W. J. Morokoff and R. E. Caflisch, *Quasi-Monte Carlo integration*, J. Comput. Phys. **122** (1995), 218–230.

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

[NS95] H. Niederreiter and P. J.-S. Shiue (eds.), *Monte Carlo and quasi-Monte Carlo methods in scientific computing*, Lecture Notes in Statistics, vol. 106, Springer-Verlag, New York, 1995.

[NS99] H. Niederreiter and J. Spanier (eds.), *Monte Carlo and quasi-Monte Carlo methods 1998*, Springer-Verlag, Berlin, 1999.

[Owe95] A. B. Owen, *Randomly permuted (t, m, s)-nets and (t, s)-sequences*, In Niederreiter and Shiue [NS95], 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.

[Owe98] A. B. Owen, *Latin supercube sampling for very high dimensional simulations*, ACM Trans. Model. Comput. Simul. **8** (1998), 71–102.

[Owe99] A. B. Owen, *Monte Carlo, quasi-Monte carlo, and randomized quasi-Monte Carlo*, In Niederreiter and Spanier [NS99].

[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.

[PT97] A. Papageorgiou and J. F. Traub, *Faster evaluation of multidimensional integrals*, Computers in Physics **11** (1997), 574–578.

[Rit96] K. Ritter, *Average case analysis of numerical problems*, Habilitationsschrift, Universität Erlangen-Nürnberg, Erlangen, Germany, 1996.

[Rot54] K. F. Roth, *On irregularities of distribution*, Mathematika **1** (1954), 73–79.

[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.

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

[Spa95] J. Spanier, *Quasi-Monte Carlo methods for particle transport problems*, In Niederreiter and Shiue [NS95], pp. 121–148.

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

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

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

[War95] T. T. Warnock, *Computational investigations of low discrepancy point sets II*, In Niederreiter and Shiue [NS95], pp. 354–361.

[Woź99] H. Woźniakowski, *Efficiency of quasi-Monte Carlo algorithms for high dimensions*, In Niederreiter and Spanier [NS99].

[YH99] R. X. Yue and F. J. Hickernell, *Robust designs for fitting linear models with misspecification*, Statist. Sinica (1999), to appear.

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