Quasi-Monte Carlo for Integrands with Point Singularities at Unknown Locations

Art B. Owen

Stanford University, Stanford CA 94305, USA owen@stat.stanford.edu

Summary. This article considers quasi-Monte Carlo sampling for integrands having isolated point singularities. It is usual for such singular functions to be approached via importance sampling. Indeed one might expect that very uniform sampling, such as QMC uses, should be unhelpful in such problems, and the Koksma-Hlawka inequality seems to indicate as much. Perhaps surprisingly, we find that the expected errors in randomized QMC converge to zero at a faster rate than holds for Monte Carlo sampling, under growth conditions for which $2+\epsilon$ moments of the integrand are finite. The growth conditions do place constraints on certain partial derivatives of the integrand, but unlike importance sampling, they do not require knowledge of the locations of the singularities.

1 Introduction

The core problem in quasi-Monte Carlo (QMC) integration is to compute an estimate of

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

The estimates typically take the form

$$\hat{I} = \frac{1}{n}\sum_{i=1}^{n} f(x_i), \tag{2}$$

for carefully chosen x_i. In applications of QMC, some problem specific manipulations are usually made to render the problem and solution into the forms given by equations (1) and (2) respectively.

The classical Monte Carlo method (MC) takes the x_i to be independent random vectors with the $U(0,1)^d$ distribution. When $\int f(x)^2dx < \infty$, MC achieves a root mean square error (RMSE) that is $O(n^{-1/2})$. Hlawka (1961) proved the Koksma-Hlawka inequality: for Riemann integrable f, and $x_1,\ldots,x_n \in [0,1]^d$,

$$|\hat{I} - I| \le V_{\mathrm{HK}}(f) D_n^*(x_1, \dots, x_n), \tag{3}$$

where D_n^* is the star discrepancy of $x_1, \dots, x_n$, and V_{HK} denotes total variation in the sense of Hardy and Krause. Many constructions are known to satisfy $D_n^*(x_1, \dots, x_n) = O(n^{-1+\epsilon})$ for all $\epsilon > 0$. If $V_{\mathrm{HK}}(f) < \infty$ then Hlawka's theorem assures the asymptotic superiority of QMC over MC.

Now suppose that $V_{\mathrm{HK}}(f) = \infty$. Then (3) reduces to $|\hat{I} - I| \le \infty$, and so is completely uninformative. We cannot tell whether QMC is asymptotically better than MC, nor even whether the QMC error will converge to zero.

Infinite variation is not an obscure issue. It arises commonly in applications. The indicator function for a subset of $[0,1]^d$ for $d \ge 2$, has infinite variation, unless that subset has boundaries aligned to the axes of the unit cube. Similarly, functions formed piecewise such as $\max(g_1(x), g_2(x))$ commonly contain a cusp along the set where $g_1(x) = g_2(x)$ that gives them infinite variation, when $d \ge 3$. A detailed investigation of the variation of such functions is given in Owen (2005).

The integrands we consider here are unbounded. An unbounded integrand necessarily has infinite variation in the sense of Hardy and Krause. Unbounded integrands arise in some problems with Feynmann diagrams from physics, as in Aldins, Brodsky, Dufner & Kinoshita (1970). They can also arise from variable transformations such as importance sampling and mappings used to convert problems on infinite domains such as $[0, \infty)^d$ or $\mathbb{R}^d$ to the unit cube.

For some unbounded functions, QMC still attains the $O(n^{-1+\epsilon})$ rate of convergence. For others, the QMC error $|\hat{I} - I|$ diverges to infinity as $n \to \infty$, even as the star discrepancy satisfies $D_n^* = O(n^{-1+\epsilon})$. Owen (2006) includes examples of both cases.

Section 2 reviews the literature on QMC for unbounded functions. Most of that work considers singularities located on the boundary faces or corners of the unit cube, and in the exceptions we must know something about where the internal singularities are. The emphasis here is on functions that diverge to infinity at a finite number of interior points, whose locations are not necessarily known. Section 3 outlines some error bounds obtained from approximating f by another function $\widetilde{f}$ that has finite Hardy-Krause variation while remaining close to f. Section 4 presents the notation and some background material. Section 5 develops the error rates that are the main result of this paper. Section 6 ends with some more general comments on randomized QMC versus Monte Carlo.

2 Literature

Lehman (1955) may have been the first to investigate integration of general unbounded functions by equidistributed sequences. He considered periodic functions on $\mathbb{R}$ with $f(x+1) = f(x)$ and sampling at points $x_i = i\xi$ for irrational ξ. His functions can be singular either as $x \uparrow 1$ or as $x \downarrow 0$. His

Theorem 4 also allows finitely many internal singularities, though one has to check that the points $i\xi$ for $1 \le i \le n$ avoid the singularities suitably.

Sobol' (1973) made another large step in the development of quasi-Monte Carlo methods for singular functions. His paper has many results for the case $d = 1$, but we focus here on the multidimensional results. Sobol' (1973) considers functions $f(x)$ that are singular as x approaches the origin of $[0, 1]^d$, such as products of negative powers of the components of x. These functions are in fact singular as x approaches the entire 'lower' boundary of $[0, 1]^d$. He shows that quasi-Monte Carlo points which avoid a hyperbolic region near that boundary lead to consistent integral estimates, that is $|\hat{I} - I| \to 0$ as $n \to \infty$.

Sobol's work has been extended recently. Like Sobol', de Doncker & Guan (2003) consider power law singularities around the origin, and they find some benefit from extrapolation schemes. Klinger (1997*b*) shows that Halton points as well as certain digital nets avoid cubical regions around the origin, making them suited to integrands with point singularities at $(0, \dots, 0)$. In his dissertation, Klinger (1997*a*, Theorem 3.17) gives conditions under which the points of a Kronecker sequence, a d dimensional version of the points that Lehman (1955) studied, avoid a hyperbolic region around the origin. Hartinger, Kainhofer & Tichy (2004) show how to replace the uniform density implicit in (1) with a more general bounded probability density function, and they show how to generate low discrepancy points with respect to that density. In that paper the sample points have every component bounded away from zero, effectively avoiding an L-shaped region around the origin. Owen (2006) extends Sobol's work to get rates of convergence, and to consider functions singular around any or all of the 2^d corners of $[0, 1]^d$. The rate can be as good as $O(n^{-1+\epsilon})$ if the singular function obeys a strict enough growth rate. Both hyperbolic and L-shaped regions are considered. Owen (2006) also shows how the Halton sequences avoid the boundaries of the unit cube as do randomized QMC points. Hartinger & Kainhofer (2005) consider non-uniform densities, with singularities arising at any corner and with hyperbolic or L-shaped avoidance regions, and obtain the same rates as in Owen (2006).

As motivation, Sobol' (1973) makes the observation that his quadrature methods were being used in practice with success on singular integrands, despite the fact that the supporting theory is limited to bounded integrands. Much of the recent work is motivated in part by the observation that good results are obtained on certain Asian option valuation problems, despite the unboundedness of the corresponding integrands.

3 Use of Low Variation Approximations

Suppose that f has infinite variation, but is close to another function $\widetilde{f}$ that has finite variation and integral $\widetilde{I}$. Then a three epsilon argument gives

$$|I - \hat{I}| \le |I - \widetilde{I}| + D_n^*(x_1, \dots, x_n) V_{\text{HK}}(\widetilde{f}) + \frac{1}{n} \sum_{i=1}^{n} |f(x_i) - \widetilde{f}(x_i)|. \quad (4)$$

We want the sum on the right side of (4) to converge to zero as n increases. To achieve this we will let $\widetilde{f}$ change as n increases, and bound all three terms by a common rate in n. Because $D_n^* \to 0$ we can let $V_{\text{HK}}(\widetilde{f})$ increase with n in order to obtain a reduction in the first and third terms.

In principle, the infimum of (4) taken over functions $\widetilde{f}$ is also an upper bound on the error. In practice, we need strategies for constructing specific $\widetilde{f}$ that we can study. One strategy is that of "avoiding the singularity". We take a set K on which f is bounded, and arrange to sample only within K. The ideal set K might be a level set $\{x \in [0,1]^d \mid |f(x)| \le \eta\}$. When such a set is hard to work with, K might be geometrically chosen to avoid the origin or some other place where the singularity arises.

Suppose that x_i are constructed to belong to a subset K of $[0,1]^d$, and that f is not singular in K. The function $\widetilde{f}$ is an extension of f from $K \subseteq [0,1]^d$ to $[0,1]^d$ if $\widetilde{f}(x) = f(x)$ whenever $x \in K$. Then, for deterministic points $x_i \in K$ we obtain

$$|I - \hat{I}| \le |I - \widetilde{I}| + D_n^*(x_1, \dots, x_n) V_{\text{HK}}(\widetilde{f}). \quad (5)$$

In a closely related strategy, suppose that x_i are randomized QMC points, where each x_i individually has the $U[0,1]^d$ distribution and for any $\epsilon > 0$ there is $D_\epsilon < \infty$ such that $D_n^*(x_1, \dots, x_n) \le D_\epsilon n^{-1+\epsilon}$ always holds. Then taking expectations in (4) we get

$$E(|I - \hat{I}|) \le 2 \int_{[0,1]^d} |\widetilde{f}(x) - f(x)| dx + D_\epsilon n^{-1+\epsilon} V_{\text{HK}}(\widetilde{f}). \quad (6)$$

Random points generally have some positive probability of landing outside K, and as a result, the last term in (4) contributes to the first term in (6).

Both strategies require a function $\widetilde{f}$ that is close to f and has a small variation. Ilya M. Sobol' invented such a low variation extension. It was used in Sobol' (1973) but Sobol' did not publish it. An account of it appears in Owen (2005) and also in the next section.

Our present purpose is to study integration of functions with point singularities at possibly unknown locations. It is then hard to construct QMC points that avoid the singularity. Accordingly, the randomized QMC strategy is the one that we choose to employ here.

4 Notation

A point $x \in \mathbb{R}^d$ is written as $x = (x^1, \dots, x^d)$. For $a, b \in \mathbb{R}^d$ with $a^j \le b^j$, the hyperrectangle $[a, b]$ is the Cartesian product $\prod_{j=1}^d [a^j, b^j]$. Most of our work will be with the unit hyperrectangle defined via $a^j = 0$ and $b^j = 1$. The unit hyperrectangle is denoted $[0,1]^d$. We need more general hyperrectangles $[a,b]$, particularly for subsets of $[0,1]^d$.

For any points $a, b \in \mathbb{R}^d$ the rectangular hull of a and b is

$$\mathrm{rect}[a,b] = \prod_{j=1}^{d} [\min(a^j, b^j), \max(a^j, b^j)],$$

a kind of d dimensional bounding box determined by points a and b.

For $u \subseteq 1:d \equiv \{1,\ldots,d\}$ we write $|u|$ for the cardinality of u, and $-u$ for the complement $\{1 \le j \le d \mid j \notin u\}$ of u. The set difference is $v - u = \{j \in v \mid j \notin u\}$. For disjoint sets u and v their union may be written $u + v$. The symbol x^u denotes the $|u|$-tuple consisting of all x^j for $j \in u$. The domain of x^u is written $[0^u, 1^u]$ or $[0,1]^u$.

For $x, y \in [0,1]^d$ the expression $x^u{:}y^{-u}$ denotes the point $z \in [0,1]^d$ with $z^j = x^j$ for $j \in u$ and $z^j = y^j$ for $j \notin u$. Similar more general expressions are used. For example, with disjoint sets $u, v, w \subseteq 1{:}d$, the expression $x^u{:}y^v{:}z^w$ designates a point in $[0,1]^{u+v+w}$.

When x^{-u} is held fixed then f can be used to define a function of x^u over $[0,1]^u$. This function is denoted by $f(x^u; x^{-u})$, with the argument x^u appearing before the semi-colon, and the parameter x^{-u} after it. The value of $f(x^u; x^{-u})$ is $f(x^u{:}x^{-u})$.

We let $\partial^u f$ denote the mixed partial derivative of f taken once with respect to each x^j for $j \in u$. By convention $\partial^\emptyset f = f$.

The variation of f over the hyperrectangle $[a,b]$ in the sense of Vitali, denoted by $V_{[a,b]}(f)$ is defined in Niederreiter (1992). We will use some properties of it, as surveyed in Owen (2005). Suppose that $\partial^{1:d} f$ exists. Then, a classical result of Fréchet (1910) is that

$$V_{[a,b]}(f) \le \int_{[a,b]} |\partial^{1:d} f(x)| dx, \tag{7}$$

with equality when $\partial^{1:d} f$ is continuous. The variation of f in the sense of Hardy and Krause is

$$V_{\mathrm{HK}}(f) = \sum_{v \ne \emptyset} V_{[0^v, 1^v]} f(x^v; 1^{-v}). \tag{8}$$

Sobol's extension begins with an extendable region. Such a region K, must be accompanied by an anchor point $c \in K$ such that $z \in K$ implies $\mathrm{rect}[c,z] \subseteq K$. Figure 1 illustrates some extendable regions.

The function f is extendable if $\partial^{1:d} f$ exists at every point of K. A more general definition of extendable function is available if $x \in K$ implies $x^j = c^j$ for some $1 \le j \le d$ (Owen 2005), but we do not need it here.

Sobol's extension is defined as follows. First, the function f can be written as

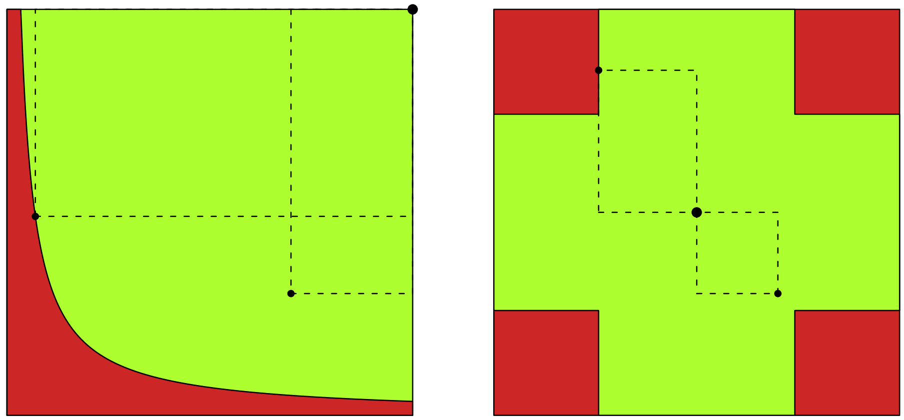

Fig. 1. This figure shows some extendable regions. In the left panel, the region K is above a hyperbolic branch, and the anchor point is at $(1,1)$. In the right panel, the region K excludes four small corner squares, and the anchor point is at $(1/2,1/2)$. For each panel one point is selected interior to K and one is selected on the boundary of K. The dashed lines indicate the rectangular hulls joining each selected point to its anchor.

$$f(x) = f(c) + \sum_{u\neq\emptyset}\int_{[c^u,x^u]} \partial^u f(y^u\!:\!c^{-u})dy^u. \tag{9}$$

In integrals like (9), the region of integration is not a proper hyperrectangle if some $c^j > x^j$. We take $\int_{[c^u,x^u]} g(y^u)dy^u = \pm\int_{\mathrm{rect}[c^u,x^u]} g(y^u)dy^u$ where the sign is negative if and only if $c^j > x^j$ holds for an odd number of $j \in u$. Sobol's extension is then

$$\widetilde{f}(x) = f(c) + \sum_{u\neq\emptyset}\int_{[c^u,x^u]} \partial^u f(y^u\!:\!c^{-u})1_{y^u:c^{-u}\in K}dy^u. \tag{10}$$

For $x \in K$ the factor $1_{y^u:c^{-u}\in K}$ in (10) always equals 1, so that $\widetilde{f}(x) = f(x)$ for $x \in K$, and then the term "extension" is appropriate. Owen (2005) shows that $V_{[a,b]}(\widetilde{f}) \le \int_K |\partial^{1:d} f(x)|dx = \int_K |\partial^{1:d}\widetilde{f}(x)|dx$. In view of Fréchet's result (7), Sobol's extension has as low a variation as one could expect.

To compute variation, we will need derivatives of $\widetilde{f}(x)$. First notice that the $f(c)$ term in (10) corresponds to $u = \emptyset$ in a natural convention. Also $\widetilde{f}(x)$ depends on x only through the role x plays in limits of integration. Accordingly, for $w \subseteq 1\!:\!d$,

$$\partial^w\widetilde{f}(x) = \sum_{u\supseteq w}\int_{[c^{u-w},x^{u-w}]} \partial^u f(y^{u-w}\!:\!x^w\!:\!c^{-u})1_{y^{u-w}:x^w:c^{-u}\in K}dy^{u-w}. \tag{11}$$

In particular

$$\partial^{1:d}\widetilde{f}(x) = \begin{cases} \partial^{1:d} f(x), & x \in K \\ 0, & \text{else.} \end{cases}$$

5 Integrable Point Singularities

Suppose that the function f has a finite number L of integrable singularities at distinct points $z_1, \ldots, z_L \in (0,1)^d$. Without loss of generality, we may suppose that $L = 1$ because integration errors sum. Let the function f have a point singularity at $z \in (0,1)^d$. We do not assume that the position of z is known.

We will suppose that f is such that $|f(x)| \le \|x-z\|_p^{-A}$ where $0 < A < d$, and $1 \le p < \infty$. The lower limit on A provides the singularity, while the upper limit ensures that the singularity is integrable:

$$\begin{aligned} \int_{[0,1]^d} \|x-z\|_p^{-A} dx &\le S_{d,p} \int_0^{\|(1,\ldots,1)\|_p} y^{-A+d-1} dy \\ &\le \frac{S_{d,p}}{d-A}\Bigl(d^{(d-A)/p} - 0^{d-A}\Bigr) = \frac{S_{d,p} d^{(d-A)/p}}{d-A}, \end{aligned}$$

where $S_{d,p}$ is the $d-1$ dimensional volume of the set $\{x \in \mathbb{R}^d \mid \|x\|_p = 1\}$.

If $A < d/2$, then f^2 is integrable. In this case Monte Carlo sampling has root mean squared error $O(n^{-1/2})$.

5.1 Growth conditions

We need a notion of a singularity "no worse than" $\|x-z\|_p^{-A}$, and to obtain it we impose growth conditions on the partial derivatives of f. If $f = \|x-z\|_p^{-A}$ for $1 < p < \infty$, then for $x \ne z$,

$$\frac{\partial f}{\partial x^j} = -A\|x-z\|_p^{-A-p}|x^j - z^j|^{p-1}\operatorname{sign}(x^j - z^j)$$

and generally for $u \subseteq \{1,\ldots,d\}$,

$$\partial^u f = \Biggl(\prod_{k=0}^{|u|-1}(-A-kp)\Biggr)\|x-z\|_p^{-A-|u|p}\prod_{j\in u}|x^j - z^j|^{p-1}\operatorname{sign}(x^j - z^j). \quad (12)$$

An easy Lagrange multiplier argument gives, for non-empty u,

$$\prod_{j\in u}|x^j - z^j| \le |u|^{-|u|/p}\|x^u - z^u\|_p^{|u|} \le |u|^{-|u|/p}\|x-z\|_p^{|u|}. \quad (13)$$

Now applying (13) to (12) and absorbing the constants, we obtain the growth condition

$$|\partial^u f(x)| \le B\|x-z\|_p^{-A-|u|} \tag{14}$$

required to hold for all $u \subseteq \{1,\dots,d\}$, all $x \ne z$, some $0 < A < d$, some $B < \infty$, and some $1 \le p < \infty$. For $p = 1$ the upper bound $\|x-z\|_1^{-A}$ for f is not differentiable, and hence does not itself satisfy the growth conditions. Yet, growth condition (14) for $p = 1$ may still be of use, when a specific integrand of interest can be shown to satisfy it.

5.2 Extension from $\|x-z\|_p \ge \eta$

To avoid the singularity in $\|x-z\|_p^{-A}$, x must satisfy $\|x-z\|_p \ge \eta$ for some $\eta > 0$. But if $z \in (0,1)^d$, then the level set $K = \{x \mid \|x-z\|_p \ge \eta\}$, for small η, is not an extendable region: there is no place to put the anchor c so that $\mathrm{rect}[c, z+\eta(1,0,\dots,0)]$ and $\mathrm{rect}[c, z-\eta(1,0,\dots,0)]$ are both subsets of K.

We will make 2^d extensions, one from every orthant of $[0,1]^d$, as defined with respect to an origin at z. These orthants are, for $u \subseteq \{1,\dots,d\}$, Cartesian products of the form

$$\overline{C}(u) = \overline{C}(z,u) \equiv \prod_{j\in u}[0,z^j]\prod_{j\notin u}[z^j,1]. \tag{15}$$

The set $\overline{C}(u)$ may also be written as $[0^u\!:\!z^{-u}, z^u\!:\!1^{-u}]$ or $[0,z]^u \times [z,1]^{-u}$.

For small positive η and $z \in (0,1)^d$, the region $K_u = K_u(\eta) \equiv \overline{C}(u) \cap \{\|x-z\|_p \ge \eta\}$ is extendable to $\overline{C}(u)$ with respect to the anchor at $c_u = 0^u\!:\!1^{-u}$. Figure 2 illustrates how $[0,1]^d$ can be split into orthants, for $d = 2$. We let $\widetilde{f}_u(x)$ be the Sobol' extension of f from K_u to $\overline{C}(u)$ with anchor c_u.

The orthants in (15) overlap, and so we cannot simply take $\widetilde{f}$ to be $\widetilde{f}_u(x)$ for $x \in \overline{C}(u)$. The sets

$$C(u) = C(z,u) \equiv \prod_{j\in u}[0,z^j)\prod_{j\notin u}[z^j,1], \tag{16}$$

form a partition of $[0,1]^d$. We may now define

$$\widetilde{f}(x) = \widetilde{f}_u(x), \quad \text{for} \quad x \in C(z,u), \tag{17}$$

and then $\widetilde{f}(x)$ is well defined on $[0,1]^d$.

Lemma 1. *If f satisfies (14) with $z \in (0,1)^d$, then for $\widetilde{f}_u$ as defined above,*

$$\|\widetilde{f}_u(x) - f(x)\|_p \le B_u\|x-z\|_p^{-A} \tag{18}$$

holds for all $x \in \overline{C}(u)$ and some $B_u < \infty$, and

$$\|\widetilde{f}(x) - f(x)\|_p \le \widetilde{B}\|x-z\|_p^{-A} \tag{19}$$

for all $x \in [0,1]^d$ and some $\widetilde{B} < \infty$.

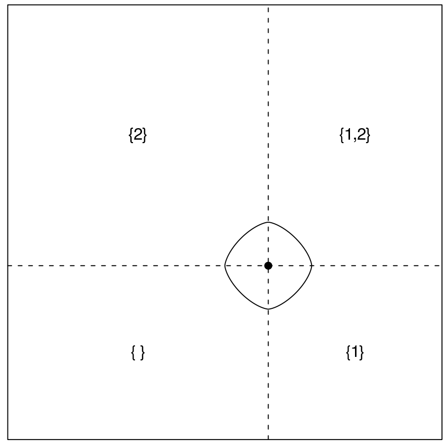

Fig. 2. This figure illustrates the orthants used to extend a function towards a point singularity. The solid point at $z = (0.6, 0.4)$, represents the site of the singularity. It is surrounded by a curve $\|x - z\|_p = \eta$, where $p = 3/2$ and $\eta = 1/10$. The orthants $\overline{C}(u)$ are rectangles with one corner at z and another at a corner c_u of $[0,1]^2$. They are labeled by the set u, taking values $\{\}$, $\{1\}$, $\{2\}$, and $\{1,2\}$. The part K_u of $\overline{C}(u)$ outside of the curve around z is a region extendable to $\overline{C}(u)$ with respect to an anchor at c_u.

Proof: Equation (19) follows from (18) by taking $\widetilde{B} = \max_u B_u$, so we only need to show (18). For $x \in \overline{C}(u)$, subtracting representation (10) from (9) term by term gives

$$
\begin{aligned}
|f(x) - \widetilde{f}(x)| &\le B \sum_{v \ne \emptyset} \int_{[c_u^v, x^v]} 1_{y^v : c_u^{-v} \notin K_v} \|y^v : c_u^{-v} - z\|_p^{-A-|v|} dy^v \\
&\le B \sum_{v \ne \emptyset} \int_{[c_u^v, x^v]} \|y^v : c_u^{-v} - z\|_p^{-A-|v|} dy^v \\
&\le B \sum_{v \ne \emptyset} \int_{\|x^v : c_u^{-v} - z\|_p}^{\|c_u - z\|_p} \alpha^{-A-|v|} S_{p,|v|} \alpha^{|v|-1} d\alpha \\
&\le \frac{B}{A} \sum_{v \ne \emptyset} S_{p,|v|} \|x^v : c_u^{-v} - z\|_p^{-A}.
\end{aligned}
$$

For $v \ne \{1, \dots, d\}$, we have

$$
\|x^v : c_u^{-v} - z\|_p^{-A} \le \min_{j \notin v} |z^j - c_u^j|^{-A} \le \min_{1 \le j \le d} |z^j - c_u^j|^{-A}.
$$

Thus $|f(x) - \widetilde{f}(x)| \le \bigl(B S_{p,d} \|x - z\|_p^{-A} + B'\bigr)/A$ where

$$\begin{aligned} B' &= B \min_{1\le j\le d} |z^j - c^j|^{-A} \sum_{0<|v|<d} \binom{d}{|v|} S_{p,|v|} \\ &\le \left(\frac{B \min_{1\le j\le d} |z^j - c^j|^{-A} \sum_{0<|v|<d} \binom{d}{|v|} S_{p,|v|}}{\inf_{y\in\overline{C}(u)} \|y-z\|_p^{-A}} \right) \|x-z\|_p^{-A}, \end{aligned}$$

so that a finite B_u exists with $|f(x) - \widetilde{f}(x)| \le B_u \|x-z\|_p^{-A}$ for $x \in \overline{C}(u)$. $\square$

Lemma 1 will allow us to find an upper bound to the first term in the expected absolute quadrature error (6). To handle the second term, we need to consider the variation of $\widetilde{f}$, in the sense of Hardy and Krause:

$$\begin{aligned} V_{\mathrm{HK}}(\widetilde{f}) &= \sum_{v\ne\emptyset} V_{[0^v,1^v]}(\widetilde{f}(x^v;1^{-v})) \\ &= \sum_{v\ne\emptyset} \sum_{u\subseteq v} V_{[0,z]^u\times[z,1]^{v-u}}(\widetilde{f}(x^v;1^{-v})). \end{aligned} \tag{20}$$

The second equality in (20) comes from an additive property of Vitali variation stated in Young (1913) and repeated in Owen (2005). Apart from boundary points, all of $[0,z]^u \times [z,1]^{v-u}$ is inside $C(u)$. But we cannot simply replace $\widetilde{f}$ by $\widetilde{f}_u$ in (20). The reason is that $\widetilde{f}$ is ordinarily discontinuous along the boundaries between orthants and such a discontinuity contributes to the Vitali variation of $\widetilde{f}$.

For $d = 2$ and $x \in \overline{C}(\emptyset)$, we find that

$$\widetilde{f}(x) = \begin{cases} \widetilde{f}_\emptyset(x), & x \in [0,z^1)\times[0,z^2) \\ \widetilde{f}_{\{1\}}(x), & x \in \{z^1\}\times[0,z^2) \\ \widetilde{f}_{\{2\}}(x), & x \in [0,z^1)\times\{z^2\} \\ \widetilde{f}_{\{1,2\}}(x), & x \in \{z^1\}\times\{z^2\}. \end{cases} \tag{21}$$

For general d and general u, the function $\widetilde{f}(x)$ on $\overline{C}(u)$ takes on values from each of the $2^{d-|u|}$ functions $\widetilde{f}_v$ with $v \supseteq u$.

The small example in (21) can be written as a sum of inclusion-exclusions

$$\begin{aligned} \widetilde{f}(x) &= \widetilde{f}_\emptyset(x) + \Big(\widetilde{f}_{\{1\}}(x) - \widetilde{f}_\emptyset(x)\Big) 1_{x^1=z^1} + \Big(\widetilde{f}_{\{2\}}(x) - \widetilde{f}_\emptyset(x)\Big) 1_{x^2=z^2} \\ &\quad + \Big(\widetilde{f}_{\{1,2\}}(x) - \widetilde{f}_{\{1\}}(x) - \widetilde{f}_{\{2\}}(x) + \widetilde{f}_\emptyset(x)\Big) 1_{x^{\{1,2\}}=z^{\{1,2\}}}, \quad x \in \overline{C}(\emptyset), \end{aligned}$$

and more generally, for $x \in \overline{C}(u)$,

$$\widetilde{f}(x) = \sum_{w\subseteq -u} \widetilde{g}_{u,w}(x), \quad \text{where,} \tag{22}$$

$$\widetilde{g}_{u,w}(x) = 1_{x^w=z^w} \sum_{t\subseteq w} (-1)^{|w-t|} \widetilde{f}_{u\cup t}(x), \tag{23}$$

where by convention, $1_{x^\emptyset=z^\emptyset}=1$. Therefore

$$V_{\mathrm{HK}}(\widetilde{f}) \le \sum_{v\ne\emptyset}\sum_{u\subseteq v}\sum_{w\subseteq -u} V_{[0,z]^u\times[z,1]^{v-u}}(\widetilde{g}_{u,w}(x^v;1^{-v})) \tag{24}$$

because the Vitali variation of a sum of functions is no larger than the sum of the variations of the functions. Because z is an interior point, no $z^j=1$. If w intersects $-v$, then $\widetilde{g}_{u,w}(x^v;1^{-v})=0$ for all $x^v\in[0,z]^u\times[z,1]^{v-u}$. The sum over w in (24) can therefore be reduced to one over $v-u$.

Many of the functions $\widetilde{g}_{u,w}$ vanish except on one "face" of the hyperrectangle on which they are defined. Variation of such functions is covered by the next lemma.

Lemma 2. *Let $a,b\in\mathbb{R}^d$ with $a\le b$ componentwise. Let $u,v,w\subseteq\{1,\dots,d\}$ be disjoint with $u\cup v\cup w=\{1,\dots,d\}$. Suppose that $f(x)$ is defined on $[a,b]$ with $f(x)=0$ unless $x^u=a^u$ and $x^v=b^v$. Then*

$$V_{[a,b]}(f)=\begin{cases}0, & \mathit{Vol}([a,b])=0\\ |f(a^u\!:\!b^v)|, & \mathit{Vol}([a,b])>0,\, w=\emptyset\\ V_{[a^w,b^w]}(f(x^w;a^u\!:\!b^v)), & \mathit{Vol}([a,b])>0,\, w\ne\emptyset.\end{cases} \tag{25}$$

Proof: This follows from Propositions 18 and 19 of Owen (2005). □

Lemma 3. *Let f satisfy the growth conditions (14) for some $z\in(0,1)^d$. Let $\widetilde{f}(x)$ be the extension of $f(x)$ from $K=\{x\in[0,1]^d\mid \|x-z\|_p\ge\eta\}$ given by equation (17). Then as $\eta\to 0$,*

$$V_{HK}(\widetilde{f})=O(\eta^{-A}). \tag{26}$$

Proof: In view of Lemma 2, we may write

$$\begin{aligned}V_{\mathrm{HK}}(\widetilde{f}) &\le \sum_{v\ne\emptyset}\sum_{u\subseteq v}\sum_{w\subseteq v-u}\sum_{t\subseteq w} V_{[0,z]^u\times[z,1]^{v-u}}\Big(\widetilde{f}_{u+t}(x^v;1^{-v})1_{x^t=z^t}\Big)\\ &= \sum_{v\ne\emptyset}\sum_{u\subseteq v}\sum_{w\subseteq v-u}\sum_{t\subseteq w} V_{[0,z]^u\times[z,1]^{(v-t)-u}}\Big(\widetilde{f}_{u+t}(x^{v-t};z^t\!:\!1^{-v})\Big)\\ &\le \sum_{v\ne\emptyset}\sum_{u\subseteq v}\sum_{w\subseteq v-u}\sum_{t\subseteq w}\int_{[0,z]^u\times[z,1]^{(v-t)-u}} |\partial^{v-t}\widetilde{f}_{u+t}(x^{v-t};z^t\!:\!1^{-v})|\,dx^{v-t}.\end{aligned}$$

Applying (11) and (14)

$$
\begin{aligned}
&|\partial^{v-t}\widetilde{f}_{u+t}(x^{v-t};z^t\!:\!1^{-v})|\\
&\le \sum_{s\supseteq v-t}\int_{[c_{u+t}^{s-(v-t)},x^{s-(v-t)}]}|\partial^s f(y^{s-(v-t)}\!:\!x^{v-t}\!:\!c_{u+t}^{-s})|1_{y^{s-(v-t)}:x^{v-t}:c_{u+t}^{-s}\in K}\,dy^{s-(v-t)}\\
&\le B\sum_{s\supseteq v-t}\int_{[c_{u+t}^{s-(v-t)},x^{s-(v-t)}]}\|y^{s-(v-t)}\!:\!x^{v-t}\!:\!c_{u+t}^{-s}-z\|_p^{-A-|s|}1_{y^{s-(v-t)}:x^{v-t}:c_{u+t}^{-s}\in K}\,dy^{s-(v-t)}\\
&\le B\sum_{s\supseteq v-t}S_{p,|s|}\int_{\|x^s:c_{u+t}^{-s}-z\|_p}^{\|x^{v-t}:c_{u+t}^{-(v-t)}-z\|_p}\alpha^{-A-|s|+|s|-1}1_{\alpha\ge\eta}\,d\alpha\\
&\le \frac{B}{A}\sum_{s\supseteq v-t}S_{p,|s|}\max\Bigl(\eta,\|x^s:c_{u+t}^{-s}-z\|_p\Bigr)^{-A}\\
&=O(\eta^{-A}).
\end{aligned}
$$

The estimate $O(\eta^{-A})$ propagates through the integral and four summations in $V_{\mathrm{HK}}(\widetilde{f})$, to give $V_{\mathrm{HK}}(\widetilde{f})=O(\eta^{-A})$. $\square$.

Theorem 1. *Let f be a Lebesgue measurable function with integral I, and a singularity at $z\in(0,1)^d$ for which the growth conditions (14) hold. Let $x_1,\dots,x_n$ be randomized quasi-Monte Carlo points: each $x_i\sim U[0,1]^d$, and for any $\epsilon>0$ there exists $D_\epsilon<\infty$ with $\Pr(D_n^*(x_1,\dots,x_n)\le D_\epsilon n^{-1+\epsilon})=1$ for all $n\ge1$. Let $\hat I=(1/n)\sum_{i=1}^n f(x_i)$. Then*

$$E(|\hat I-I|)=O(n^{(-1+\epsilon)(d-A)/d}).$$

Proof: For small $\eta>0$, define $\widetilde{f}$ by extending f from $\{x\in[0,1]^d\mid \|x-z\|_p\ge\eta\}$ as described above. Then

$$E(|\hat I-I|)\le 2\int|\widetilde{f}(x)-f(x)|dx+E(D_n^*(x_1,\dots,x_n))\|\widetilde{f}\|_{\mathrm{HK}}.$$

Combining the hypothesis of the Theorem with Lemma 3 shows that the second term is $O(n^{-1+\epsilon}\eta^{-A})$. From Lemma 1,

$$
\begin{aligned}
\int_{[0,1]^d}|\widetilde{f}(x)-f(x)|dx &\le \widetilde{B}\int_{[0,1]^d}\|x-z\|_p^{-A}1_{\|x-z\|_p\le\eta}dx\\
&\le BS_{p,d}\int_0^\eta\alpha^{-A+d-1}d\alpha\\
&=O(\eta^{d-A}).
\end{aligned}
$$

Taking η proportional to $n^{(-1+\epsilon)/d}$ gives $E(|\hat I-I|)=O(n^{(-1+\epsilon)(d-A)/d})$. $\square$

Suppose that the integrand f has L singularities at points $z_\ell\in(0,1)^d$, for $\ell=1,\dots,L$. Under the following mild conditions on f, the worst of L growth conditions will determine the error rate. First, there are L hyperrectangles $[a_\ell, b_\ell] \subset (0,1)^d$ with $z_\ell \in (a_\ell, b_\ell)$ and $[a_\ell, b_\ell] \cap [a_{\widetilde{\ell}}, b_{\widetilde{\ell}}] = \emptyset$ for $\ell \neq \widetilde{\ell}$. Suppose that $|\partial^u f(x)| \le B_\ell \|x - z_\ell\|_{p_\ell}^{-A_\ell}$ holds for $x \in [a_\ell, b_\ell] - \{z_\ell\}$ and $B_\ell < \infty$, and $0 < A_\ell < d$, and $1 \le p_\ell < \infty$. Suppose also that $|\partial^{1:d} f(x)|$ is bounded on $[0,1]^d - \cup_{\ell=1}^L (a_\ell, b_\ell)$. We may take $\widetilde{f}$ to equal a low variation extension within each of the $[a_\ell, b_\ell]$ and to equal f outside of them. Then for the randomized QMC points in Theorem 1 we get

$$E(|\hat{I} - I|) = O(n^{(-1+\epsilon)(d - \max_{1 \le \ell \le L} A)/d}).$$

The next result (whose proof is immediate) shows that randomized QMC is asymptotically superior to independent Monte Carlo sampling in this setting, when $\int |f(x)|^{2+\epsilon} dx < \infty$ for some $\epsilon > 0$.

Corollary 1. *Suppose that the conditions of Theorem 1 hold with $A < d/2$. Then $E(|\hat{I} - I|) = o(n^{-1/2})$.*

6 Discussion

It is surprising to find that even for functions with unknown point singularities, that randomized QMC has superior asymptotics to ordinary Monte Carlo sampling, which after all, includes importance sampling. One might have thought that spreading points uniformly through space would be wasteful for such functions. The key measure of strength for our growth conditions is the scalar A. When $A < d/2$, then Monte Carlo has the $O(n^{-1/2})$ rate and randomized QMC has expected absolute error $O(n^{(d-A)/d}) = o(n^{-1/2})$.

In fairness, we should note that the QMC approach does make some stronger assumptions than Monte Carlo. The randomized QMC approach requires conditions on mixed partial derivatives of f that are unnecessary in MC.

Perhaps a bigger loophole is that the Koksma-Hlawka bound can be extremely conservative. For large d it is certainly plausible that the superior rate for randomized QMC will be slow to set in, and that independent Monte Carlo with importance sampling will beat randomized QMC in practical samples sizes. Perhaps the best approach is to employ an importance sampling strategy in conjunction with randomized QMC points. The importance sampling strategy may well leave an integrand with singularities. Then, the results here show that singular integrands need not be a serious drawback for randomized QMC.

Randomized QMC has one important practical advantage compared to importance sampling. We do not need to know the locations of the singularities.

It would be interesting to know how generally randomized QMC is asymptotically superior to Monte Carlo for singular functions. From the Borel-Cantelli theorem (Chung 1974) we can find that randomized QMC points will avoid high level sets, in that $\max_{1 \le i \le n} |f(x_i)| > n$ will only happen finitely often with probability one, if $\int |f(x)|dx < \infty$. It is not however clear that one can always find a good extension $\widetilde{f}$ for f, from $\{x \in [0,1]^d \mid |f(x)| \le n\}$ to $[0,1]^d$.

Acknowledgments

I thank Harald Niederreiter and Denis Talay for organizing MC2QMC2004 and for editing this proceedings volume. I also thank Harald Niederreiter and one anonymous reviewer for their careful scrutiny of this paper. Further thanks are due to Reinhold Kainhofer for discussions on QMC for unbounded integrands, Bernhard Klinger for providing a copy of his dissertation, and again to Harald Niederreiter for telling me about Lehman's results. This work was supported by grant DMS-0306612 from the U.S. National Science Foundation.

References

Aldins, J., Brodsky, S. J., Dufner, A. J. & Kinoshita, T. (1970), 'Moments of the muon and electron', *Physical Review D* **1**(8), 2378–2395.

Chung, K.-L. (1974), *A Course in Probability Theory*, 2nd edn, Academic Press, New York.

de Doncker, E. & Guan, Y. (2003), 'Error bounds for the integration of singular functions using equidistributed sequences', *Journal of Complexity* **19**(3), 259–271.

Fréchet, M. (1910), 'Extension au cas des intégrales multiples d'une définition de l'intégrale due à Stieltjes', *Nouvelles Annales de Mathématiques* **10**, 241–256.

Hartinger, J. & Kainhofer, R. F. (2005), Non-uniform low-discrepancy sequence generation and integration of singular integrands, *in* H. Niederreiter & D. Talay, eds, 'Proceedings of MC2QMC2004, Juan-Les-Pins France, June 2004', Springer-Verlag, Berlin.

Hartinger, J., Kainhofer, R. F. & Tichy, R. F. (2004), 'Quasi-Monte Carlo algorithms for unbounded, weighted integration problems', *Journal of Complexity* **20**(5), 654–668.

Hlawka, E. (1961), 'Funktionen von beschränkter Variation in der Theorie der Gleichverteilung', *Annali di Matematica Pura ed Applicata* **54**, 325–333.

Klinger, B. (1997*a*), Discrepancy of Point Sequences and Numerical Integration, PhD thesis, Technische Universität Graz.

Klinger, B. (1997*b*), 'Numerical integration of singular integrands using low-discrepancy sequences', *Computing* **59**, 223–236.

Lehman, R. S. (1955), 'Approximation of improper integrals by sums over multiples of irrational numbers', *Pacific Journal of Mathematics* **5**, 93–102.

Niederreiter, H. (1992), *Random Number Generation and Quasi-Monte Carlo Methods*, S.I.A.M., Philadelphia, PA.

Owen, A. B. (2005), Multidimensional variation for quasi-Monte Carlo, *in* J. Fan & G. Li, eds, 'International Conference on Statistics in honour of Professor Kai-Tai Fang's 65th birthday'.

Owen, A. B. (2006), 'Halton sequences avoid the origin', *SIAM Review* **48**. To appear.

Sobol', I. M. (1973), 'Calculation of improper integrals using uniformly distributed sequences', *Soviet Math Dokl* **14**(3), 734–738.

Young, W. H. (1913), *Proceedings of the London Mathematical Society (2)* **11**, 142.