Quasi-Monte Carlo for an Integrand with a Singularity along a Diagonal in the Square

Kinjal Basu and Art B. Owen

Abstract Quasi-Monte Carlo methods are designed for integrands of bounded variation, and this excludes singular integrands. Several methods are known for integrands that become singular on the boundary of the unit cube $[0,1]^d$ or at isolated possibly unknown points within $[0,1]^d$. Here we consider functions on the square $[0,1]^2$ that may become singular as the point approaches the diagonal line $x_1 = x_2$, and we study three quadrature methods. The first method splits the square into two triangles separated by a region around the line of singularity, and applies recently developed triangle QMC rules to the two triangular parts. For functions with a singularity 'no worse than $|x_1 - x_2|^{-A}$ is' for $0 < A < 1$ that method yields an error of $O((\log(n)/n)^{(1-A)/2})$. We also consider methods extending the integrand into a region containing the singularity and show that method will not improve upon using two triangles. Finally, we consider transforming the integrand to have a more QMC-friendly singularity along the boundary of the square. This then leads to error rates of $O(n^{-1+\varepsilon+A})$ when combined with some corner-avoiding Halton points or with randomized QMC but it requires some stronger assumptions on the original singular integrand.

1 Introduction

Quasi-Monte Carlo (QMC) integration is designed for integrands of bounded variation in the sense of Hardy and Krause (BVHK). Such integrands must necessarily be bounded. Singular integrands cannot be BVHK; they cannot even be Riemann

Kinjal Basu
LinkedIn Inc.
e-mail: kbasu@linkedin.com

Art B. Owen
Stanford University
e-mail: owen@stanford.edu

integrable. It is known since [6] and [3] that for any integrand f on $[0,1]^d$ that is not Riemann integrable, there exists a sequence $\boldsymbol{x}_i \in [0,1]^d$ for which the star discrepancy $D_n^*(\boldsymbol{x}_1,\dots,\boldsymbol{x}_n) \to 0$ as $n \to \infty$ while $(1/n)\sum_{i=1}^n f(\boldsymbol{x}_i)$ fails to converge to $\int_{[0,1]^d} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$.

We are interested in problems where the singularity arises along a manifold in $[0,1]^d$. For motivation, see the engineering applications by Mishra and Gupta in [10] and several other papers. Apart from a few remarks, we focus solely on the problem where there is a singularity along the line $x_1 = x_2$ in $[0,1]^2$.

It is possible for QMC integration to succeed on unbounded integrands. Sobol' [15] noticed this when colleagues used his methods on such problems. He explained it in terms of QMC points that avoid a hyperbolic region around the lower boundary of the unit cube where the integrands became singular. Klinger [9] shows that Halton points and some digital nets avoid a cubical region around the origin. Halton points (after the zero'th) avoid hyperbolic regions around the boundary faces of the unit cube at a rate suitable to get error bounds for QMC [13]. Certain Kronecker sequences avoid hyperbolic regions around the boundary of the cube [8]. In all of these examples, avoiding the singularity should be understood as using points that approach it, but not too quickly, as the number n of function evaluations increases.

For plain Monte Carlo, the location of the singularity is not important. One only needs to consider the first two moments of the integrand. Because QMC exploits mild smoothness of the integrand, the nature of the singularity matters. Reference [14] considers randomized QMC (RQMC) methods for integrands with point singularities at unknown locations. In RQMC, the integrand is evaluated at points that, individually, are uniformly distributed on $[0,1]^d$ and this already implies a singularity avoidance property via the Borel-Cantelli lemma. If $\int f(\boldsymbol{x})^2\mathrm{d}\boldsymbol{x} < \infty$ then scrambled nets yield an unbiased estimate of $\mu = \int f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$ with RMSE $o(n^{-1/2})$ [11].

The analyses in [13] and [14] employ an extension $\tilde{f}$ of f from a set $K = K_n \subset [0,1]^d$ to $[0,1]^d$. The extension satisfies $\tilde{f}(\boldsymbol{x}) = f(\boldsymbol{x})$ for $\boldsymbol{x} \in K$. Now the quadrature error is

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^n f(\boldsymbol{x}_i) - \int_{[0,1]^d} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x} &= \frac{1}{n}\sum_{i=1}^n f(\boldsymbol{x}_i) - \frac{1}{n}\sum_{i=1}^n \tilde{f}(\boldsymbol{x}_i)\\ &\quad + \frac{1}{n}\sum_{i=1}^n \tilde{f}(\boldsymbol{x}_i) - \int_{[0,1]^d} \tilde{f}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\\ &\quad + \int_{[0,1]^d} \tilde{f}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} - \int_{[0,1]^d} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}.\end{aligned}$$

If all of the points satisfy $\boldsymbol{x}_i \in K$, then the first term drops out and we find that

$$\left|\frac{1}{n}\sum_{i=1}^n f(\boldsymbol{x}_i) - \int_{[0,1]^d} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right| \le \left|\frac{1}{n}\sum_{i=1}^n \tilde{f}(\boldsymbol{x}_i) - \int_{[0,1]^d} \tilde{f}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right| + \int_{-K} |\tilde{f}(\boldsymbol{x}) - f(\boldsymbol{x})|\mathrm{d}\boldsymbol{x},$$

where $-K = [0,1]^d \setminus K$. The extension used in [13] and [14] is due to Sobol' [15]. It is particularly well suited to a Koksma-Hlawka bound for the first term above as $\tilde{f}$ has low variation.

In our case, we can isolate the singularity in the set $\{\boldsymbol{x} \mid |x_1 - x_2| < \varepsilon\}$. A set $K \subset [0,1]^d$ is Sobol'-extensible to $[0,1]^d$ with anchor $\boldsymbol{c}$ if for every $\boldsymbol{x} \in K$ the rectangle $\prod_{j=1}^d [\min(x_j, c_j), \max(x_j, c_j)] \subset K$. In our case, the set $\{\boldsymbol{x} \mid |x_1 - x_2| \geq \varepsilon\}$ in which f is bounded is not Sobol' extensible. The extension $\tilde{f}$ used in [13] and [14] cannot be defined for this problem.

Section 2 presents a strategy of avoiding a region near the singularity and integrating over two triangular regions using the method from [1]. The error is then a sum of two quadrature errors and one truncation error. We consider functions where the singularity is not more severe than that in $|x_1 - x_2|^{-A}$ where $0 < A < 1$. Section 3 shows that the truncation error in this approach is $O(\varepsilon^{-A})$ and the quadrature error is $O(\varepsilon^{-A-1}\log(n)/n)$ using the points from [1] and a Koksma-Hlawka bound from [4]. The result is that we can attain a much better quadrature error bound of $O((\log(n)/n)^{(1-A)/2})$. Section 4 shows that an approach based on finding an extension $\tilde{f}$ of f would not yield a better rate for this problem. Section 5 transforms the problem so that each triangular region becomes the image of a unit square, with the singularity now on the boundary of the square. The singularity may be too severe for QMC. However, with an additional assumption on the nature of the singularity it is possible to attain a quadrature error of $O(n^{-1+\varepsilon+A})$. Section 6 summarizes the findings and relates them to QMC-friendliness as discussed by several authors, including Ian Sloan in his work with Xiaoqun Wang.

2 Background

In the context of a Festschrift for Ian Sloan, we presume that the reader is familiar with quasi-Monte Carlo, discrepancy and variation. Modern approaches to QMC and discrepancy are covered in [7]. See [12] for an outline of variation for QMC including variation in the senses of Vitali and of Hardy and Krause.

We will use a notion of functions that are singular but not too badly singular.

Definition 1. The function f defined on $[0,1]^2$ has a diagonal singularity no worse than $|x_1 - x_2|^{-A}$ for $0 < A < 1$, if

$$\begin{aligned} |f(\boldsymbol{x})| &\leq B|x_1 - x_2|^{-A} \\ \left|\frac{\partial f(\boldsymbol{x})}{\partial x_j}\right| &\leq B|x_1 - x_2|^{-A-1}, \quad j \in \{1,2\}, \quad \text{and} \\ \left|\frac{\partial^2 f(\boldsymbol{x})}{\partial x_j \partial x_k}\right| &\leq B|x_1 - x_2|^{-A-2}, \quad j,k \in \{1,2\} \end{aligned} \tag{1}$$

all hold for some $B < \infty$.

We take $A > 0$ in order to allow a singularity and $A < 1$ because f must be integrable. Smaller values of A describe easier cases to handle. The value of A to use for a given integrand may be evident from its analytical form. If $A < 1/2$ then f^2 is integrable. Definition 1 is modeled on some previous notions:

Definition 2. The function $f(\boldsymbol{x})$ defined on $[0,1]^d$ has a lower edge singularity no worse than $\prod_{j=1}^d x_j^{-A_j}$, for constants $0 < A_j < 1$, if

$$|\partial^u f(\boldsymbol{x})| \le B \prod_{j=1}^d x_j^{-A_j - 1_{j\in u}},$$

holds for some $B < \infty$ and all $u \subseteq \{1,2,\dots,d\}$.

Definition 3. The function $f(\boldsymbol{x})$ defined on $[0,1]^d$ has a point singularity no worse than $\|\boldsymbol{x}-\boldsymbol{z}\|^{-A}$, for $\boldsymbol{z} \in [0,1]^d$, if

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

holds for some $B < \infty$ and all $u \subseteq \{1,2,\dots,d\}$.

Definition 2 is one of several conditions in [13] for singularities that arise as $\boldsymbol{x}$ approaches the boundary of the unit cube. Definition 3 is used in [14] for isolated point singularities. Definition 1 is more stringent than Definitions 2 and 3 are, because it imposes a constraint on partial derivatives taken twice with respect to x_1 or x_2.

To estimate $\mu = \int_{[0,1]^2} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$ we will sample points $\boldsymbol{x}_i \in [0,1]^2$. The points we use will avoid a region near the singularity by sampling only within

$$S_\varepsilon = \bigl\{\boldsymbol{x} \in [0,1]^2 \bigm| |x_1 - x_2| \ge \varepsilon\bigr\}$$

where $0 < \varepsilon < 1$. The set S_ε is the union of two disjoint triangles:

$$T_\varepsilon^u = \bigl\{\boldsymbol{x} \in [0,1]^2 \bigm| x_2 \ge x_1 + \varepsilon\bigr\}, \quad \text{and}$$
$$T_\varepsilon^d = \bigl\{\boldsymbol{x} \in [0,1]^2 \bigm| x_2 \le x_1 - \varepsilon\bigr\}.$$

We let $-S_\varepsilon$ denote the set $[0,1]^2 \setminus S_\varepsilon$. As remarked in the introduction, the set $T_u \cup T_d$ is not Sobol' extensible to $[0,1]^2$.

We will choose points $\boldsymbol{x}_{i,u} \in T_\varepsilon^u$ for $i = 1,\dots,n$ and estimate $\mu_{\varepsilon,u} = \int_{T_\varepsilon^u} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$ by

$$\hat{\mu}_{\varepsilon,u} = \frac{\mathrm{vol}(T_\varepsilon^u)}{n}\sum_{i=1}^n f(\boldsymbol{x}_{i,u}).$$

Using a similar estimate for T_ε^d we arrive at our estimate of μ,

$$\hat{\mu}_\varepsilon = \hat{\mu}_{\varepsilon,u} + \hat{\mu}_{\varepsilon,d}.$$

Our error then consists of two quadrature errors and a truncation error and it satisfies the bound

$$|\hat{\mu}_\varepsilon - \mu| \le \left|\hat{\mu}_{\varepsilon,u} - \int_{T_\varepsilon^u} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right| + \left|\hat{\mu}_{\varepsilon,d} - \int_{T_\varepsilon^d} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right| + \left|\int_{-S_\varepsilon} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right|. \quad (2)$$

3 Error bounds

We show in Proposition 1 below that the truncation error bound $|\int_{-S_\varepsilon} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}|$ is $O(\varepsilon^{1-A})$ as $\varepsilon \to 0$. We will use the construction from [1] and the Koksma-Hlawka inequality from [4] to provide an upper bound for the integration error over T_ε^u. That bound grows as $\varepsilon \to 0$ and so to trade them off we will tune the way ε depends on n.

Proposition 1. *Under the regularity conditions* (1)*,*

$$\left|\int_{-S_\varepsilon} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right| \le \frac{2B\varepsilon^{1-A}}{1-A}.$$

Proof. We take the absolute value inside the integral and obtain

$$\int_{-S_\varepsilon} |f(\boldsymbol{x})|\mathrm{d}\boldsymbol{x} \le \int_{-S_\varepsilon} B|x_1-x_2|^{-A}\mathrm{d}\boldsymbol{x} \le B\int_0^1 2\int_0^\varepsilon x_2^{-A}\mathrm{d}x_2\mathrm{d}x_1$$

from which the conclusion follows. □

Next we turn to the quadrature errors over T_ε^u. Of course, T_ε^d is similar. The Koksma-Hlawka bound in [4] has

$$|\hat{\mu}_{\varepsilon,u} - \mu_{\varepsilon,u}| \le D^*_{T_\varepsilon^u}(\boldsymbol{x}_{1,u},\dots,\boldsymbol{x}_{n,u})V_{T_\varepsilon^u}(f)$$

where $D^*_{T_{u,\varepsilon}}$ and $V_{T_\varepsilon^u}$ are measures of discrepancy and variation suited to the triangle. Basu and Owen [1] provide a construction in which $D^*_{T_\varepsilon^u} = O(\log(n)/n)$, the best possible rate.

Brandolini et al. [4, p. 46] provide a bound for $V_{T_\varepsilon^u}$, the variation on the simplex as specialized to the triangle. To translate their bound into our setting, we introduce the notation $f_{rs} = \partial^{r+s}f/\partial^r x_1\partial^s x_2$. Specializing their bound to the domain T_ε^u we find that the variation is

$$\begin{aligned}O\Big(&|f(0,1)| + |f(0,\varepsilon)| + |f(1-\varepsilon,1)|\\ &+\int_\varepsilon^1 |f(0,x_2)|\mathrm{d}x_2 + \int_0^{1-\varepsilon}|f(x_1,1)|\mathrm{d}x_1 + \int_0^{1-\varepsilon}|f(x_1,x_1+\varepsilon)|\mathrm{d}x_1\\ &+\int_\varepsilon^1 |f_{01}(0,x_2)|\mathrm{d}x_2 + \int_0^{1-\varepsilon}|f_{10}(x_1,1)|\mathrm{d}x_1\\ &+\int_0^{1-\varepsilon}|f_{10}(x_1,x_1+\varepsilon)|\mathrm{d}x_1 + \int_0^{1-\varepsilon}|f_{01}(x_1,x_1+\varepsilon)|\mathrm{d}x_1\\ &+\int_{T_\varepsilon^u}|f(\boldsymbol{x})| + |f_{01}(\boldsymbol{x})| + |f_{10}(\boldsymbol{x})| + |f_{20}(\boldsymbol{x})| + |f_{02}(\boldsymbol{x})| + |f_{11}(\boldsymbol{x})|\mathrm{d}\boldsymbol{x}\Big)\end{aligned} \tag{3}$$

as $\varepsilon \to 0$. The implied constant in (3) includes their unknown constant C_2, the reciprocals of edge lengths of T_ε^u, the reciprocal of the area of T_ε^u, some small integers and some factors involving $\sqrt{2}(1-\varepsilon)$, the length of the hypotenuse of T_ε^u.

Proposition 2. *Let f satisfy the regularity condition* (1). *Then the trapezoidal variation of f over T^u_ε satisfies*

$$V_{T^u_\varepsilon}(f) = O(\varepsilon^{-A-1})$$

as $\varepsilon \to 0$.

Proof. Under condition (1),

$$|f(0,1)| + \int_\varepsilon^1 |f(0,x_2)|\mathrm{d}x_2 + \int_0^{1-\varepsilon} |f(x_1,1)|\mathrm{d}x_1 + \int_{T^u_\varepsilon} |f(\boldsymbol{x})| = O(1).$$

Next

$$|f(0,\varepsilon)| + |f(1-\varepsilon,1)| + \int_0^{1-\varepsilon} |f(x_1,x_1+\varepsilon)|\mathrm{d}x_1 = O(\varepsilon^{-A})$$

and

$$\int_\varepsilon^1 |f_{01}(0,x_2)|\mathrm{d}x_2 + \int_0^{1-\varepsilon} |f_{10}(x_1,1)|\mathrm{d}x_1 = O(\varepsilon^{-A})$$

as well. Continuing through the terms, we find that

$$\int_0^{1-\varepsilon} |f_{10}(x_1,x_1+\varepsilon)|\mathrm{d}x_1 + \int_0^{1-\varepsilon} |f_{01}(x_1,x_1+\varepsilon)|\mathrm{d}x_1 = O(\varepsilon^{-A-1}).$$

The remaining terms are integrals of absolute partial derivatives of f over T^u_ε. They are dominated by integrals of second derivatives and those terms obey the bound

$$\int_0^{1-\varepsilon}\int_{x_1+\varepsilon}^1 B_2|x_1-x_2|^{-A-2}\mathrm{d}x_2\mathrm{d}x_1 = O(\varepsilon^{-A-1}).$$

□

Theorem 1. *Under the regularity conditions* (1)*, we may choose $\varepsilon \propto \sqrt{\log(n)/n}$ and get*

$$|\hat{\mu}-\mu| = O\Big(\Big(\frac{\log(n)}{n}\Big)^{(1-A)/2}\Big). \tag{4}$$

Proof. From Propositions 1 and 2 we get

$$|\hat{\mu}-\mu| = O\Big(\varepsilon^{1-A} + \frac{\log(n)}{n}\varepsilon^{-1-A}\Big).$$

Taking ε to be a positive multiple of $\sqrt{\log(n)/n}$ yields the result. □

The choice of $\varepsilon \propto \sqrt{\log(n)/n}$ optimizes the upper bound in (4).

4 Extension based approaches

Another approach to this problem is to construct a function $\tilde{f}$ where $\tilde{f}(\boldsymbol{x}) = f(\boldsymbol{x})$ for $\boldsymbol{x} \in S_\varepsilon$ and apply QMC to $\tilde{f}$. The function $\tilde{f}$ can smoothly bridge the gap between T^u_ε and T^d_ε. With such a function, the quadrature error satisfies

$$\left|\frac{1}{n}\sum_{i=1}^n \tilde{f}(\boldsymbol{x}_i) - \int_{[0,1]^2} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right| \le D^*_n(\boldsymbol{x}_1,\dots,\boldsymbol{x}_n)V_{\mathrm{H}K}(\tilde{f}) + \int_{-S_\varepsilon} |f(\boldsymbol{x}) - \tilde{f}(\boldsymbol{x})|\mathrm{d}\boldsymbol{x} \quad (5)$$

where $V_{\mathrm{H}K}$ is total variation in the sense of Hardy and Krause.

Our regularity condition (1) allows for f to take the value ε^{-A} along the line $x_2 = x_1 - \varepsilon$ and to take the value $-\varepsilon^{-A}$ along $x_2 = x_1 + \varepsilon$. By placing squares of side 2ε along the main diagonal we then find that the Vitali variation of an extension $\tilde{f}$ is at least $\lfloor(2\varepsilon)^{-1}\rfloor 2\varepsilon^{-A} \sim \varepsilon^{-1-A}$. Therefore the Hardy-Krause variation of $\tilde{f}$ grows at least this quickly for some of the functions f that satisfy (1). More generally, for singular functions along a linear manifold M within $[0,1]^d$, and no worse than $\mathrm{dist}(\boldsymbol{x},M)^{-A}$, an extension over the region within ε of M could have a variation lower bound growing as fast as $\varepsilon^{-(d-1)-A}$.

This result is much less favorable than the one for isolated point singularities [14]. For integrands on $[0,1]^d$ no worse than $\|\boldsymbol{x}-\boldsymbol{z}\|^{-A}$, where $\boldsymbol{z} \in [0,1]^d$, Sobol's low variation extension yields a function $\tilde{f}$ that agrees with f for $\|\boldsymbol{x}-\boldsymbol{z}\| \ge \varepsilon > 0$ having $V_{\mathrm{H}K}(\tilde{f}) = O(\varepsilon^{-A})$. Here we see that no extension can have such low variation for this type of singularity.

Owen [13] considers functions with singularities along the lower boundary of $[0,1]^d$ that are no worse than $\prod_{j=1}^d x_j^{-A_j}$. Sobol's extension from the region where $\prod_j x_j \ge \varepsilon$ has variation $O(\varepsilon^{-\max A_j})$ when the A_j are distinct (otherwise logarithmic factors enter). So that problem with singularities along the boundary also has a more accurate extension than can be obtained for singularities along the diagonal.

No extension $\tilde{f}$ from S_ε to $[0,1]^2$ can yield a bound (5) with a better rate than $O((\log n/n)^{(1-A)/2})$. To show this we first clarify one of the rules we impose on extensions. When we extend f from $\boldsymbol{x} \in S$ to values of $\boldsymbol{x} \notin S$ we do not allow the construction of $\tilde{f}$ to depend on $f(\boldsymbol{x})$ for $\boldsymbol{x} \notin S$. That is, we cannot peek outside the set we are extending from. Some such rule must be necessary or we could trivially get 0 error from an extension based on an oracle that uses the value of μ to define $\tilde{f}$. With our rule, any two functions f_1 and f_2 with $f_1(\boldsymbol{x}) = f_2(\boldsymbol{x})$ on S_ε have the same extension $\tilde{f}$. From the triangle inequality,

$$\max_{j=1,2}\left(\int_{-S_\varepsilon} |\tilde{f}(\boldsymbol{x}) - f_j(\boldsymbol{x})|\mathrm{d}\boldsymbol{x}\right) \ge \frac{1}{2}\int_{-S_\varepsilon} |f_1(\boldsymbol{x}) - f_2(\boldsymbol{x})|\mathrm{d}\boldsymbol{x}.$$

Now let

$$f_1(\boldsymbol{x}) = \begin{cases} -|x_1 - x_2|^{-A}, & x_2 - x_1 > 0 \\ |x_1 - x_2|^{-A}, & x_2 - x_1 < 0, \end{cases}$$

and

$$f_2(\boldsymbol{x}) = \begin{cases} |x_1 - x_2|^{-A}, & x_2 - x_1 > 0 \\ \phi(x_2 - x_1), & 0 > x_2 - x_1 \geq -\varepsilon \\ |x_1 - x_2|^{-A}, & -\varepsilon > x_2 - x_1, \end{cases}$$

for a quadratic polynomial ϕ with $\phi(-\varepsilon) = \varepsilon^{-A}$, $\phi'(-\varepsilon) = -A\varepsilon^{-A-1}$, and $\phi''(-\varepsilon) = A(A+1)\varepsilon^{-A-2}$. Both f_1 and f_2 satisfy (1) and $\int_{-S_\varepsilon} |f_1(\boldsymbol{x}) - f_2(\boldsymbol{x})| \mathrm{d}\boldsymbol{x}$ is larger than a constant times ε^{1-A}. That is the same rate as the truncation error from Proposition 1 and the quadrature error from this approach also attains the same rate as the error in Proposition 2. As a result, we conclude that even if we could construct the best extension $\tilde{f}$, it would not lead to a bound with a better rate than the one in Theorem 1.

5 Transformation

Here we consider applying a change of variable to move the singularity from the diagonal to an edge of the unit square. We focus on integrating $f(\boldsymbol{x})$ over $T^u = \{(x_1,x_2) \in [0,1]^2 \mid 0 \leq x_1 \leq x_2 \leq 1\}$ for f with a singularity no worse than $|x_1 - x_2|^{-A}$. The same strategy and same convergence rate hold on $T^d = \{(x_1,x_2) \in [0,1]^2 \mid 0 \leq x_2 \leq x_1 \leq 1\}$. Using a standard change of variable we have

$$\int_{T^u} f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \frac{1}{2} \int_0^1 \int_0^1 f((1-u_1)\sqrt{u_2}, \sqrt{u_2}) \mathrm{d}\boldsymbol{u},$$

which we then write as

$$\frac{1}{2} \int_{[0,1]^2} g(\boldsymbol{u}) \mathrm{d}\boldsymbol{u}, \quad \text{for } g(\boldsymbol{u}) = f((1-u_1)\sqrt{u_2}, \sqrt{u_2}).$$

That is $g(\boldsymbol{u}) = f(\tau(\boldsymbol{u}))$ for a transformation $\tau : [0,1]^2 \to T_u \subset [0,1]^2$ given by $\tau_1(\boldsymbol{u}) = (1-u_1)\sqrt{u_2}$ and $\tau_2(\boldsymbol{u}) = \sqrt{u_2}$.

The archetypal function with diagonal singularity satisfying Definition 1 is $f(\boldsymbol{x}) = |x_1 - x_2|^{-A}$. The corresponding function g for this f is

$$g(\boldsymbol{u}) = |\tau_1(\boldsymbol{u}) - \tau_2(\boldsymbol{u})|^{-A} = u_1^{-A} u_2^{-A/2}.$$

We see that the change of variable has produced an integrand with a singularity no worse than $u_1^{-A} u_2^{-A/2}$ according to Definition 2. Taking $\boldsymbol{u}_i$ to be the Halton points leads to a quadrature error at rate $O(n^{-1+\varepsilon+A})$ for any $\varepsilon > 0$, because Halton points (after the zeroth one) avoid the origin at a suitable rate [13, Corollary 5.6]. For this integrand g, randomized quasi-Monte Carlo points for will attain the mean error rate $\mathbb{E}(|\hat{\mu} - \mu|) = O(n^{-1+\varepsilon+A})$ as shown in Theorem 5.7 of [13].

We initially thought that the conversion from a diagonal singularity to a lower edge singularity no worse than $u_1^{-A}u_2^{-A/2}$ would follow for other functions satisfying Definition 1. Unfortunately, that is not necessarily the case.

Let f be defined on $[0,1]^2$ with a diagonal singularity no worse than $|x_1-x_2|^{-A}$ for $0<A<1$. First,

$$|g(\boldsymbol{u})| = |f((1-u_1)\sqrt{u_2},\sqrt{u_2})| \le B|u_1u_2^{1/2}|^{-A}$$

which fits Definition 2. Similarly,

$$g_{10}(\boldsymbol{u}) = f_{10}(\tau_1(\boldsymbol{u}),\tau_2(\boldsymbol{u}))\frac{\partial\tau_1(\boldsymbol{u})}{\partial u_1} = O(|\tau_1-\tau_2|^{-A-1})\times u_2^{1/2} = O(u_1^{-A-1}u_2^{-A/2})$$

which also fits Definition 2. However,

$$\begin{aligned} g_{01}(\boldsymbol{u}) &= f_{10}(\tau(\boldsymbol{u}))\frac{\partial\tau_1(\boldsymbol{u})}{\partial u_2} + f_{01}(\tau(\boldsymbol{u}))\frac{\partial\tau_2(\boldsymbol{u})}{\partial u_2} \\ &= \big(f_{10}(\tau(\boldsymbol{u}))+f_{01}(\tau(\boldsymbol{u}))\big)\frac{1}{2}u_2^{-1/2} - f_{10}(\tau(\boldsymbol{u}))\frac{1}{2}u_1u_2^{-1/2}. \end{aligned} \tag{6}$$

Now f_{10} and f_{01} appearing in (6) are both $O(u_1^{-A-1}u_2^{-A/2-1/2})$. Therefore the two terms there are $O(u_1^{-A-1}u_2^{-A/2-1})$ and $O(u_1^{-A}u_2^{-A/2-1})$ respectively. The first term is too large by a factor of u_1^{-1} to suit Definition 2. We would need $(f_{01}+f_{10})(\tau(\boldsymbol{u}))$ to be only $O(u_1^{-A}u_2^{-A/2-1/2})$. Definition 1 is also not strong enough for g_{11} to be $O(u_1^{-A-1}u_2^{-A/2-1})$ as it would need to be under Definition 2. That definition yields only $O(u_1^{-A-2}u_2^{-A/2-1})$ without stronger assumptions. Theorem 2 below gives a sufficient condition where f is a modulated version of $|x_1-x_2|^{-A}$.

Theorem 2. *Let $f(\boldsymbol{x}) = |x_1-x_2|^{-A}h(\boldsymbol{x})$ for $\boldsymbol{x}\in[0,1]^2$ and $0<A<1$ where h and its first two derivatives are bounded. Then $g(\boldsymbol{u}) = f((1-u_1)\sqrt{u_2},\sqrt{u_2})$ satisfies Definition 2 with $A_1 = A$ and $A_2 = A/2$.*

Proof. We begin with

$$g(\boldsymbol{u}) = u_1^{-A}u_2^{-A/2}h((1-u_1)u_2^{1/2},u_2^{1/2}) = O(u_1^{-1}u_2^{-A/2})$$

by boundedness of h. Next because u_1 is not in the second argument to h,

$$\begin{aligned} g_{10}(\boldsymbol{u}) &= -Au_1^{-A-1}u_2^{-A/2}h(\tau(\boldsymbol{u})) + u_1^{-A}u_2^{-A/2}h_{10}(\tau(\boldsymbol{u}))\partial\tau_1(\boldsymbol{u})/\partial u_1 \\ &= -Au_1^{-A-1}u_2^{-A/2}h(\tau(\boldsymbol{u})) - u_1^{-A}u_2^{-A/2+1/2}h_{10}(\tau(\boldsymbol{u})) \\ &= O(u_1^{-A-1}u_2^{-A/2}) \end{aligned}$$

as required. Similarly,

$$
\begin{aligned}
g_{01}(\boldsymbol{u}) &= -(A/2)u_1^{-A}u_2^{-A/2-1}h(\tau(\boldsymbol{u}))\\
&\quad + u_1^{-A}u_2^{-A/2}\big(h_{10}(\tau(\boldsymbol{u}))(1-u_1)+h_{01}(\tau(\boldsymbol{u}))\big)(1/2)u_2^{-1/2}\\
&= O(u_1^{-A}u_2^{-A/2-1})
\end{aligned}
$$

as required. Finally $g_{11}(\boldsymbol{u})$ equals

$$
\begin{aligned}
&(A^2/2)u_1^{-A-1}u_2^{-A/2-1}h(\tau(\boldsymbol{u}))\\
&\quad -(A/2)u_1^{-A}u_2^{-A/2-1}h_{10}(\tau(\boldsymbol{u}))(-u_2^{1/2})\\
&\quad -(A/2)u_1^{-A-1}u_2^{-A/2-1/2}\big(h_{10}(\tau(\boldsymbol{u}))(1-u_1)+h_{01}(\tau(\boldsymbol{u}))\big)\\
&\quad +(u_1^{-A}u_2^{-A/2-1/2}/2)\big(-h_{10}(\tau(\boldsymbol{u}))+(1-u_1)h_{20}(\tau(\boldsymbol{u}))(-u_2^{1/2})+h_{11}(\tau(\boldsymbol{u}))(-u_2^{1/2})\big)\\
&= O(u_1^{-A-1}u_2^{-A/2-1})
\end{aligned}
$$

as required. □

6 Discussion

We find that for an integrand with a singularity 'no worse than $|x_1-x_2|^{-A}$' along the line $x_1 = x_2$ we can get a QMC estimate with error $O((\log(n)/n)^{(1-A)/2})$ by splitting the square into two triangles and ignoring a region in between them. The same method applies to singularities along the other diagonal of $[0,1]^2$. Moreover, the result extends to singularities along other lines intersecting the square. One can partition the square into rectangles, of which one has the singularity along the diagonal while the others have no singularity, and then integrate f over each of those rectangles.

That result does not directly extend to singularities along a linear manifold in $[0,1]^d$ for $d \geq 3$. The reason is that the QMC result for integration in the triangle from [1] has not been extended to the simplex. In a personal communcation, Dimitry Bilyk told us that such an extension would imply a counterexample to the Littlewood conjecture, which is widely believed to be true. Basu and Owen [2] present some algorithms for RQMC over simplices, but they come without a Koksma-Hlawka bound that would be required for limiting arguments using sequences of simplices.

The rate $O((\log(n)/n)^{(1-A)/2})$ is a bit disappointing. We do much better by transforming the problem to place the singularity along the boundary of a square region, for then we can attain $O(n^{-1+\varepsilon+A})$, under a stronger assumption that f is our prototypical singular function $|x_1-x_2|^{-A}$ possibly modulated by a function h with bounded second derivatives on $[0,1]^2$. As a result we find that there is something to be gained by engineering QMC-friendly singularities in much the same way that benefits of QMC-friendly discontinuities have been found valuable by Wang and Sloan [16].

Acknowledgements This work was supported by the US National Science Foundation under grants DMS-1407397 and DMS-1521145. We thank two anonymous reviewers for helpful comments.

References

1. Basu, K., Owen, A.B.: Low discrepancy constructions in the triangle. SIAM Journal on Numerical Analysis **53**(2), 743–761 (2015)
2. Basu, K., Owen, A.B.: Scrambled geometric net integration over general product spaces. Foundations of Computational Mathematics pp. 1–30 (2015)
3. Binder, C.: Über einen Satz von de Bruijn und Post. Österreichische Akademie der Wissenschaften Mathematisch-Naturwissenschaftliche Klasse. Sitzungsberichte. Abteilung II **179**, 233–251 (1970)
4. Brandolini, L., Colzani, L., Gigante, G., Travaglini, G.: A Koksma–Hlawka inequality for simplices. In: Trends in Harmonic Analysis, pp. 33–46. Springer (2013)
5. Brandolini, L., Colzani, L., Gigante, G., Travaglini, G.: On the Koksma–Hlawka inequality. Journal of Complexity **29**(2), 158–172 (2013)
6. de Bruijn, N.G., Post, K.A.: A remark on uniformly distributed sequences and Riemann integrability. Indagationes Mathematicae **30**, 149–150 (1968)
7. Dick, J., Pillichshammer, F.: Digital Sequences, Discrepancy and Quasi-Monte Carlo Integration. Cambridge University Press, Cambridge (2010)
8. Klinger, B.: Discrepancy of point sequences and numerical integration. Ph.D. thesis, Technische Universität Graz (1997)
9. Klinger, B.: Numerical integration of singular integrands using low-discrepancy sequences. Computing **59**, 223–236 (1997)
10. Mishra, M., Gupta, N.: Application of quasi Monte Carlo integration technique in EM scattering from finite cylinders. Progress In Electromagnetics Research Letters **9**, 109–118 (2009)
11. Owen, A.B.: Monte Carlo variance of scrambled equidistribution quadrature. SIAM Journal of Numerical Analysis **34**(5), 1884–1910 (1997)
12. Owen, A.B.: 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 (2005)
13. Owen, A.B.: Halton sequences avoid the origin. SIAM Review **48**, 487–583 (2006)
14. Owen, A.B.: Quasi-Monte Carlo for integrands with point singularities at unknown locations. In: Monte Carlo and Quasi-Monte Carlo Methods 2004, pp. 403–417. Springer (2006)
15. Sobol', I.M.: Calculation of improper integrals using uniformly distributed sequences. Soviet Mathematics Doklady **14**(3), 734–738 (1973)
16. Wang, X., Sloan, I.H.: Quasi-Monte Carlo methods in financial engineering: An equivalence principle and dimension reduction. Operations Research **59**(1), 80–95 (2011)