

## ON TRACTABILITY OF WEIGHTED INTEGRATION OVER BOUNDED AND UNBOUNDED REGIONS IN $\mathbb{R}^s$

FRED J. HICKERNELL, IAN H. SLOAN, AND GRZEGORZ W. WASILKOWSKI

**ABSTRACT.** We prove that for the space of functions with mixed first derivatives bounded in  $L_1$  norm, the weighted integration problem over bounded or unbounded regions is equivalent to the corresponding classical integration problem over the unit cube, provided that the integration domain and weight have product forms. This correspondence yields tractability of the general weighted integration problem.

### 1. INTRODUCTION

In recent years there has been great interest in the tractability of multiple integration in high dimensions, much of it stimulated by the apparent success of quasi–Monte Carlo methods applied to integrals from mathematical finance over hundreds or even thousands of dimensions; see, e.g., [8], [9] and [13] for more references. Most analysis has been carried out for the problem of integration over the  $s$ -dimensional unit cube in a reproducing kernel Hilbert space setting of functions whose mixed first derivatives are square integrable; see, e.g., [7], [12].

However, as pointed out in [11], there is a fundamental difficulty in applying the Hilbert space results to the integrals from mathematical finance. These integrals are typically with respect to probability densities over unbounded regions. The difficulty (discussed in more detail below) is that after mapping to the unit cube most problems of this kind yield integrands that do not belong to the Hilbert space: the derivatives are integrable, but not square integrable. In the present paper we study the tractability of the weighted integration problem, over both bounded and unbounded regions, in the Banach space of functions whose mixed first derivatives are in  $L_1$ . We assume that the weight function and the domain of integration have product forms (see below). We shall see that only in  $L_1$  (in contrast to  $L_p$  for  $p > 1$ ) is there a natural correspondence between the weighted integration problems over a general region and the unweighted integration problem over the unit cube.

We emphasise that in this paper we do not weight the various coordinate directions in the manner of [12]. It is known from the work of [2] that for the  $L_1$  case the unweighted integration problem over the unit cube is tractable, whereas this is believed to be not the case for  $p > 1$ , and is known [7] not to be so for  $p = 2$ . In this

---

Received by the editor May 27, 2002 and, in revised form, March 4, 2003.

2000 *Mathematics Subject Classification.* Primary 65D05, 65D30 65Y20, 62M20, 60G25.

*Key words and phrases.* Weighted integration, quasi–Monte Carlo methods, discrepancy, tractability.

paper we establish the tractability of the weighted integration problem over general regions for the  $L_1$  case by exploiting the above-mentioned connection between the weighted integration problem over general regions and the unweighted problem over the unit cube, and then extending the result of [2].

In more detail, in this paper we study the tractability in a worst-case setting of the weighted integration problem

$$I_\rho(f) = \int_D f(\mathbf{x})\rho(\mathbf{x}) d\mathbf{x},$$

where  $D$  is an  $s$ -dimensional box

$$(1) \quad D = \overline{(a_1, b_1)} \times \cdots \times \overline{(a_s, b_s)} \subseteq R^s,$$

and  $\rho(\mathbf{x})$  has the product form

$$(2) \quad \rho(\mathbf{x}) = \prod_{k=1}^s \rho_k(x_k)$$

for nonnegative functions  $\rho_k \in L_1((a_k, b_k))$ . The intervals  $(a_k, b_k)$  can be either bounded or unbounded. For simplicity, we shall assume

$$(3) \quad \int_{a_k}^{b_k} \rho_k(x) dx = 1.$$

Our analysis will be carried out in tensor product spaces of functions defined on  $D$ , with norm  $\|f\|_{p,s}$  given for  $p = 1$  by (9) below, and for general  $p > 1$  by (10). For the explanatory purposes of this introduction it is sufficient to consider just the 1-dimensional case, for which the norm in the case  $p = 1$  is

$$(4) \quad \|f\|_{1,1} := |f(c_1)| + \int_{a_1}^{b_1} |f'(x)| dx,$$

where  $c_1$  (the ‘‘anchor’’) is a fixed number in  $\overline{(a_1, b_1)}$ . In the 1-dimensional case the weighted integration problem, namely

$$I_{\rho_1}(f) := \int_{a_1}^{b_1} f(x)\rho_1(x) dx,$$

is mapped to an unweighted integration problem on the unit interval by the simple transformation

$$(5) \quad y = W_1(x) := \int_{a_1}^x \rho_1(z) dz.$$

Specifically, we find

$$I_{\rho_1}(f) = \int_0^1 F(y) dy,$$

where

$$(6) \quad F(y) = f(W_1^{-1}(y)).$$

At the same time this transformation converts the norm (4) to

$$\|f\|_{1,1} = |F(W_1(c_1))| + \int_0^1 |F'(y)| dy =: \|F\|_{1,1}.$$

The norm on the right is appropriate for the analysis of the unweighted integration problem on the unit interval.

To understand the special nature of the  $L_1$  case, suppose that instead we were to define the norm on the unit interval by

$$\|F\|_{p,1} := \left( |F(W_1(c_1))|^p + \int_0^1 |F'(y)|^p dy \right)^{1/p},$$

with  $p > 1$ . The connection between the unweighted problem on the unit interval and the original weighted integration problem is now more complicated. Indeed, under transformations (5) and (6) we find

$$\|F\|_{p,1} = \left( |f(c_1)|^p + \int_{a_1}^{b_1} |f'(x)|^p \frac{dx}{(\rho_1(x))^{p-1}} \right)^{1/p} =: \|f\|_{p,1}.$$

The norm on the right is made unattractive by the factor  $\rho_1^{p-1}$  in the denominator. This factor is especially restrictive if the region is unbounded, say if  $b_1 = \infty$ , since in that case a necessary condition for  $\|f\|_{p,1}$  to be finite is that  $f'(x) = o(\rho_1(x)^{(p-1)/p}/x^{1/p})$  as  $x \rightarrow \infty$ , a condition which in applications may seem quite unnatural. The difficulty disappears when  $p = 1$ , since in that case  $\rho_1^{p-1} = 1$ . In this paper, as in many others, we exploit the connection between worst-case error for  $f$  in the unit ball of our space and “discrepancy”. However, we need here a somewhat more general notion of discrepancy: we introduce in Section 3 the notion of “same-quadrant discrepancy”, of which the “ $L_\infty$  star discrepancy” is a special case. (Another special case is the “centered discrepancy” of [3].) There is a natural motivation for this generalization within the present work, arising from the fact that we allow  $D$  to be an unbounded region, say  $\mathbb{R}^s$ . In such a case the “anchor”  $\mathbf{c}$  in the  $s$ -dimensional generalization of (4) cannot be on the boundary of  $D$ , and hence after a smooth transformation cannot be on the boundary of the unit cube. The star discrepancy arises if in our language the anchor  $\mathbf{c}$  is chosen to be at one of the vertices of the unit cube. But that choice is too restrictive when unbounded regions are considered. By applying the present results to the classical unweighted integration problem over the unit cube, we obtain in Section 3 a simple proof of the famous Koksma-Hlawka inequality, and of a generalization when the anchor is allowed to move freely over the unit cube. In Section 4 we show that the weighted integration problem over a general region can be reduced to the unweighted problem over the unit cube. Tractability of the general problem is discussed in Section 5. Finally, some extensions are considered briefly in Section 6.

## 2. BASIC DEFINITIONS

In this section, we briefly present some definitions and basic facts concerning the worst case setting. For more detailed discussion we refer to, e.g., [6], [14]. We study the approximation of multivariate integrals of the form

$$(7) \quad I_\rho(f) = \int_D f(\mathbf{x}) \rho(\mathbf{x}) d\mathbf{x},$$

where  $f$  is from a Banach space  $\mathcal{F}$  of functions that will be specified later. Here  $D$  is the  $s$ -dimensional box (1). The intervals can be unbounded; this is why we

write  $\overline{(a_i, b_i)}$  instead of  $[a_i, b_i]$ . The *weight* function  $\rho$  is of the tensor product form (2). For simplicity, we shall assume that each  $\rho_k$  is a probability density on  $\overline{(a_k, b_k)}$  (i.e., a nonnegative, Lebesgue integrable function such that (3) holds), although, as explained in Section 6, for the tractability result it is sufficient to assume that the integral of  $\rho_k$  is finite. Classical examples of suitable weights  $\rho$  are provided by

- the uniform distribution with  $D = [0, 1]^s$  and  $\rho(\mathbf{x}) = 1$ ,
- exponential distribution with  $D = \mathbb{R}_+^s$  and

$$\rho(\mathbf{x}) = \prod_{k=1}^s \exp(-x_k/\lambda_k)/\lambda_k,$$

- Gaussian distribution with  $D = \mathbb{R}^s$  and

$$\rho(\mathbf{x}) = \prod_{k=1}^s \exp(-x_k^2/(2\lambda_k))/(\sqrt{2\pi\lambda_k}).$$

Without loss of generality, we can restrict the analysis to *linear algorithms* (or *cubatures*) of the form

$$\mathcal{A}(f) := \sum_{i=1}^n a_i f(\mathbf{x}^i)$$

that use  $n$  function values at points  $\mathbf{x}^i$ . The worst-case error of the algorithm  $\mathcal{A}$  is defined as

$$\text{error}(\mathcal{A}, \mathcal{F}) := \sup_{\|f\| \leq 1} |I_\rho(f) - \mathcal{A}(f)|,$$

where  $\|f\|$  is the norm of  $f$  in the space  $\mathcal{F}$ . Since the error equals the operator norm  $\|I_\rho - \mathcal{A}\|$ , we have

$$|I_\rho(f) - \mathcal{A}(f)| \leq \text{error}(\mathcal{A}, \mathcal{F}) \|f\|, \quad \forall f \in \mathcal{F}.$$

The number  $n$  of function values used by an algorithm is called the *cardinality* of  $\mathcal{A}$ , and is denoted by  $\text{card}(\mathcal{A})$ . Since it is a natural measure of the cost, we are interested in the smallest errors that can be achieved with  $n$  optimally chosen function evaluations,

$$\text{error}(n, \mathcal{F}) := \inf\{\text{error}(\mathcal{A}, \mathcal{F}) : \mathcal{A} \text{ such that } \text{card}(\mathcal{A}) = n\}.$$

Such minimal errors are called  *$n$ th minimal errors*. To provide a standard to compare with, let

$$(8) \quad \text{error}(0, \mathcal{F}) := \|I_\rho\|$$

be the initial error, i.e., the error of the zero algorithm. If the ratio of  $n$ th minimal error and the initial error depends polynomially on  $s$  and  $n^{-1}$ , then we say that the corresponding problem is tractable.

More specifically, consider now a sequence of weighted integration problems, each defined for a space of functions of  $s$  variables,  $s = 1, 2, \dots$ . To stress the dependence on  $s$ , we will write sometimes  $\text{error}(n, \mathcal{F}, s)$  instead of  $\text{error}(n, \mathcal{F})$ . We say that the corresponding *problem* (or more formally *a sequence of problems*) is *tractable* if there exist positive constants  $C, q_1$ , and  $q_2$  such that

$$\frac{\text{error}(n, \mathcal{F}, s)}{\text{error}(0, \mathcal{F}, s)} \leq C s^{q_1} n^{-q_2} \quad \forall n, s.$$

## 3. FUNCTION SPACES, ERRORS, AND DISCREPANCIES

Of course, tractability (or lack of it) depends very much on the choice of  $\mathcal{F}$ . In this paper, we consider the following space denoted by  $\mathcal{F}_{1,s}$ . Let  $\mathcal{H}_k$  be the space of absolutely continuous functions on  $\overline{(a_k, b_k)}$  whose first derivatives are in  $L_1((a_k, b_k))$ . Recall that  $L_1((a_k, b_k))$  is the space of functions whose absolute value is integrable over the interval  $(a_k, b_k)$ . Let  $\mathcal{H}^s = \bigotimes_{k=1}^s \mathcal{H}_k$  be the (algebraic) tensor product space, i.e., the space of all linear combinations of functions  $f$  of the product form

$$f : D \rightarrow \mathbb{R} \quad \text{and} \quad f(\mathbf{x}) = \prod_{k=1}^s h_k(x_k) \quad \text{with} \quad h_k \in \mathcal{H}_k.$$

The space  $\mathcal{F}_{1,s}$  is the completion of  $\mathcal{H}^s$  with respect to the norm

$$(9) \quad \|f\|_{1,s} := |f(\mathbf{c})| + \sum_{U \neq \emptyset} \|f'_U\|_{L_1(D_U)}.$$

Here  $\mathbf{c} = [c_1, \dots, c_s]$ , called an *anchor*, is a point from  $D$ . The summation is with respect to subsets  $U$  of  $\{1, \dots, s\}$ , and

$$f'_U(\mathbf{x}_U) := \frac{\partial^{|U|}}{\prod_{k \in U} \partial x_k} f(\mathbf{x}_U, \mathbf{c}),$$

where  $(\mathbf{x}_U, \mathbf{c})$  denotes the  $s$ -dimensional vector whose  $k$ th component is  $x_k$  if  $k \in U$ , and is  $c_k$  if  $k \notin U$ . By  $\mathbf{x}_U$  we mean the  $|U|$ -dimensional vector obtained from  $\mathbf{x}$  by removing all components  $x_k$  with  $k \notin U$ . This means that  $\mathbf{x}_U \in D_U = \prod_{k \in U} (a_k, b_k)$  and  $f'_U$  is a function defined on  $D_U$ . To simplify the notation, we shall also write  $f'_\emptyset$  and  $\|f'_\emptyset\|_{L_p}$  to denote  $f(\mathbf{c})$  and  $|f(\mathbf{c})|$ , respectively, and we often drop  $D_U$  by writing  $\|\cdot\|_{L_p}$  instead of  $\|\cdot\|_{L_p(D_U)}$ . This allows a more concise formula for the norm,

$$\|f\|_{1,s} = \sum_U \|f'_U\|_{L_1}.$$

The essence of  $\mathcal{F}_{1,s}$  is that it is a tensor product space consisting of functions whose mixed first partial derivatives are bounded in  $L_1$  norm. Natural extensions are spaces  $\mathcal{F}_{p,s}$  of functions with derivatives bounded in  $L_p$  norms for  $p \in (1, \infty]$ . That is, the norm in  $\mathcal{F}_{p,s}$  is given by

$$(10) \quad \|f\|_{p,s} := \left( \sum_U \|f'_U\|_{L_p}^p \right)^{1/p} \quad \text{for} \quad p < \infty$$

and

$$\|f\|_{\infty,s} := \max_U \|f'_U\|_{L_\infty} \quad \text{for} \quad p = \infty.$$

The following representation of  $f$  will play an important role in our analysis. Let

$$M_k(x, t) := \begin{cases} 1 & \text{if } c_k < t < x, \\ -1 & \text{if } x < t < c_k, \\ 0 & \text{otherwise,} \end{cases}$$

$$M(\mathbf{x}, \mathbf{t}) := \prod_{k=1}^s M_k(x_k, t_k),$$

and

$$M_U(\mathbf{x}_U, \mathbf{t}_U) := \prod_{k \in U} M_k(x_k, t_k),$$

with the convention that  $M_\emptyset \equiv 1$ . We will refer to  $M_k, M,$  and  $M_U$  as *kernels*.

**Proposition 1.** *For every  $f \in \mathcal{F}_{p,s}$  and every  $\mathbf{x} \in D,$*

$$(11) \quad f(\mathbf{x}) = \sum_U \int_{D_U} f'_U(\mathbf{t}_U) M_U(\mathbf{x}_U, \mathbf{t}_U) d\mathbf{t}_U.$$

As an illustration, consider  $s = 2$ . In this case the subsets  $U$  of  $\{1, 2\}$  are  $\emptyset, \{1\}, \{2\}$  and  $\{1, 2\}$ , and the representation (11) reduces to

$$\begin{aligned} f(x_1, x_2) &= f(c_1, c_2) + \int_{c_1}^{x_1} \frac{\partial}{\partial t_1} f(t_1, c_2) dt_1 \\ &\quad + \int_{c_2}^{x_2} \frac{\partial}{\partial t_2} f(c_1, t_2) dt_2 + \int_{c_1}^{x_1} \int_{c_2}^{x_2} \frac{\partial^2}{\partial t_1 \partial t_2} f(t_1, t_2) dt_2 dt_1 \\ &= f(c_1, c_2) + [f(x_1, c_2) - f(c_1, c_2)] + [f(c_1, x_2) - f(c_1, c_2)] \\ &\quad + [f(x_1, x_2) - f(c_1, x_2) - f(x_1, c_2) + f(c_1, c_2)]. \end{aligned}$$

The representation (11) has been used, at least implicitly, in a number of papers; see, e.g., [3]. For completeness, we present a short proof.

*Proof of Proposition 1.* Of course, (11) holds for  $s = 1$ . Suppose therefore that it holds for  $s$  and consider a function  $f$  of  $s + 1$  variables. Let  $R$  denote the resulting right-hand side of (11), i.e.,

$$R(\mathbf{x}) = \sum_U \int_{D_U} f'_U(\mathbf{t}_U) M_U(\mathbf{x}_U, \mathbf{t}_U) d\mathbf{t}_U,$$

where the summation now is over all subsets of  $\{1, \dots, s + 1\}$ . This sum can be expressed as a sum  $R_1$  over only those subsets  $U$  that do not contain  $s + 1$ , plus a sum  $R_2$  over every  $U$  containing  $s + 1$ . For the sum  $R_2$  write  $U = V \cup \{s + 1\}$ , where  $V \subseteq \{1, \dots, s\}$ . Then  $R_2$  can be rewritten as

$$\begin{aligned} R_2(\mathbf{x}) &= \sum_V \int_{D_V} \int_{a_{s+1}}^{b_{s+1}} \frac{\partial}{\partial t_{s+1}} f'_V(\mathbf{t}_V, t_{s+1}) M_{s+1}(x_{s+1}, t_{s+1}) dt_{s+1} M_V(\mathbf{x}_V, \mathbf{t}_V) d\mathbf{t}_V \\ &= \sum_V \int_{D_V} f'_V(\mathbf{t}_V, x_{s+1}) M_V(\mathbf{x}_V, \mathbf{t}_V) d\mathbf{t}_V \\ &\quad - \sum_V \int_{D_V} f'_V(\mathbf{t}_V, c_{s+1}) M_V(\mathbf{x}_V, \mathbf{t}_V) d\mathbf{t}_V, \end{aligned}$$

with the summations now over all subsets of  $\{1, \dots, s\}$ . Of course, the second sum above equals  $R_1$ . Hence

$$R(\mathbf{x}) = R_1(\mathbf{x}) + R_2(\mathbf{x}) = \sum_V \int_{D_V} f'_V(\mathbf{t}_V, x_{s+1}) M_V(\mathbf{x}_V, \mathbf{t}_V) d\mathbf{t}_V.$$

Since  $f = f(\cdot, x_{s+1})$  is a function of  $s$  variables only, it follows from the inductive hypothesis that  $R(\mathbf{x}) = f(x_1, \dots, x_s, x_{s+1})$ , so that (11) holds with  $s$  replaced by  $s + 1$ , thus completing the proof by induction.  $\square$

The representation of  $f$  given by (11) allows the error of the algorithm  $\mathcal{A}$  applied to  $f$  to be expressed in terms of the error of  $\mathcal{A}$  applied to the kernel  $M_U(\cdot, \mathbf{t}_U)$ . Indeed, due to linearity of  $\mathcal{A}$  and  $I_\rho$ , we have from (11) that

$$\mathcal{A}(f) = \mathcal{A}\left(\sum_U \int_{D_U} f'_U(\mathbf{t}_U) M_U(\cdot, \mathbf{t}_U) d\mathbf{t}_U\right) = \sum_U \int_{D_U} f'_U(\mathbf{t}_U) \mathcal{A}(M_U(\cdot, \mathbf{t}_U)) d\mathbf{t}_U,$$

and similarly

$$I_\rho(f) = \sum_U \int_{D_U} f'_U(\mathbf{t}_U) I_\rho(M_U(\cdot, \mathbf{t}_U)) d\mathbf{t}_U.$$

This leads to the following generalization of Zaremba's formula (see [15]):

$$(12) \quad I_\rho(f) - \mathcal{A}(f) = \sum_U \int_{D_U} f'_U(\mathbf{t}_U) h_U(\mathbf{t}_U) d\mathbf{t}_U,$$

where

$$(13) \quad h_U(\mathbf{t}_U) := I_\rho(M_U(\cdot, \mathbf{t}_U)) - \mathcal{A}(M_U(\cdot, \mathbf{t}_U)).$$

In this way, Proposition 1 leads to the following form of the worst case error of  $\mathcal{A}$ ; see also [3] where  $D = [0, 1]^s$  was considered.

**Theorem 1.** *The error of any linear algorithm  $\mathcal{A}$  is*

$$(14) \quad \text{error}(\mathcal{A}, \mathcal{F}_{1,s}) = \max_U \sup_{\mathbf{t} \in D} |I_\rho(M_U(\cdot, \mathbf{t}_U)) - \mathcal{A}(M_U(\cdot, \mathbf{t}_U))|.$$

*Proof.* By Hölder's inequality applied to (12),

$$|I_\rho(f) - \mathcal{A}(f)| \leq \sum_U \|f'_U\|_{L_1} \|h_U\|_{L_\infty} \leq \|f\|_{1,s} \max_U \|h_U\|_{L_\infty}.$$

This proves that the error of  $\mathcal{A}$  is bounded from above by the right-hand side of (14). We now prove the opposite inequality.

Given  $\varepsilon > 0$ , there exist  $Z \subseteq \{1, \dots, s\}$  and  $\mathbf{t}^* \in \text{Int}(D)$  such that  $h_Z$  is continuous at  $\mathbf{t}_Z^*$ , and

$$|h_Z(\mathbf{t}_Z^*)| \geq \max_U \|h_U\|_{L_\infty} - \varepsilon.$$

If  $Z = \emptyset$ , the worst case error is attained at the constant function  $\tilde{f} \equiv 1$ . Consider now  $Z \neq \emptyset$ . For  $\delta \simeq 0$ , define

$$v_{k,\delta}(x) = \begin{cases} (2\delta)^{-1} & \text{if } |x - t_k^*| \leq \delta, \\ 0 & \text{otherwise.} \end{cases}$$

Consider now  $g_Z \in L_1(D_Z)$  defined by

$$g_Z(\mathbf{x}_Z) = \text{sign}(h_Z(\mathbf{t}_Z^*)) \prod_{k \in Z} v_{k,\delta}(x_k)$$

and

$$\tilde{f}(\mathbf{y}) = \int_{D_Z} g_Z(\mathbf{t}) M_Z(\mathbf{y}_Z, \mathbf{t}_Z) d\mathbf{t}_Z.$$

For example, if  $Z = \{1, 2\}$ , then  $\tilde{f}(y_1, y_2) = \int_{c_1}^{y_1} \int_{c_2}^{y_2} g_Z(t_1, t_2) dt_2 dt_1$ . Note that  $\tilde{f}(\mathbf{y})$  is independent of  $y_k$  if  $k \notin Z$ , and  $\tilde{f}(\mathbf{y}) = 0$  if  $y_k = c_k$  for some  $k \in Z$ . From this it follows that  $\tilde{f}'_U \equiv 0$  if  $Z \neq U$ , and  $\tilde{f}'_Z = g_Z$ . Hence

$$\|\tilde{f}\|_{1,s} = \|g_Z\|_{L_1(D_Z)} = \prod_{k \in Z} \int_{a_k}^{b_k} v_{k,\delta}(x) dx = 1$$

for small enough  $\delta$ , and by (12),

$$I_\rho(\tilde{f}) - \mathcal{A}(\tilde{f}) = \int_{D_Z} g_Z(\mathbf{t}_Z) h_Z(\mathbf{t}_Z) d\mathbf{t}_Z = \int_{D_Z} \prod_{k \in Z} v_{k,\delta}(t_k) |h_Z(\mathbf{t}_Z)| d\mathbf{t}_Z.$$

Since the right-hand side of the above equation tends to  $|h_Z(\mathbf{t}_Z^*)|$  as  $\delta$  tends to zero, this proves that  $I_\rho(\tilde{f}) - \mathcal{A}(\tilde{f}) \geq \max_U \|h_U\|_{L_\infty} - 2\varepsilon$  if  $\delta$  is sufficiently small. Since  $\varepsilon$  can be arbitrarily close to 0, this completes the proof.  $\square$

For illustration, we apply Theorem 1 to the *classical case*, i.e.,  $\rho \equiv 1$  and  $D = [0, 1]^s$ , choosing two different anchors, namely  $\mathbf{c} = \mathbf{1}$  and  $\mathbf{c}$  inside the cube  $D$ . In both cases, we shall take for  $\mathcal{A}$  a *quasi-Monte Carlo* algorithm,

$$\mathcal{A}(f) = \mathcal{QMC}(f) = n^{-1} \sum_{i=1}^n f(\mathbf{x}^i),$$

with  $\mathbf{x}^i \in D$ .

When  $\mathbf{c} = \mathbf{1}$  we have  $M_k(x, t) = -1$  if  $x < t < 1$ , and  $M_k(x, t) = 0$  otherwise. Hence

$$I_1(M_U(\cdot, \mathbf{t}_U)) = (-1)^{|U|} \prod_{k \in U} t_k = (-1)^{|U|} \text{Vol}([\mathbf{0}, \mathbf{t}_U]),$$

and

$$\mathcal{QMC}(M_U(\cdot, \mathbf{t}_U)) = (-1)^{|U|} n^{-1} \sum_{i=1}^n \mathbf{1}_{[\mathbf{0}, \mathbf{t}_U]}(\mathbf{x}_U^i),$$

where  $\mathbf{1}_A$  denotes the characteristic (indicator) function of  $A \subseteq D$ , and  $\text{Vol}(A)$  denotes the volume of  $A$ . Thus from (13)

$$(15) \quad \|h_U\|_{L_\infty} = \sup_{\mathbf{t}_U} \left| \text{Vol}([\mathbf{0}, \mathbf{t}_U]) - n^{-1} \sum_{i=1}^n \mathbf{1}_{[\mathbf{0}, \mathbf{t}_U]}(\mathbf{x}_U^i) \right|,$$

which is the classical  $L_\infty$  star discrepancy for the  $|U|$ -dimensional unit cube  $D_U$ . Thus Theorem 1, when applied to  $\mathcal{QMC}$  methods for the classical case with  $\mathbf{c} = \mathbf{1}$ , yields the famous Koksma-Hlawka inequality:

$$|I_1(f) - \mathcal{QMC}(f)| \leq \|f\|_{1,s} \mathcal{D}_\infty^*(\mathbf{x}^1, \dots, \mathbf{x}^n),$$

where  $\mathcal{D}_\infty^*$  is the  $L_\infty$  star discrepancy of the point-set  $\{\mathbf{x}^1, \dots, \mathbf{x}^n\}$ . That is,

$$\mathcal{D}_\infty^*(\mathbf{x}^1, \dots, \mathbf{x}^n) := \sup_{\mathbf{t} \in [0,1]^s} \left| \text{Vol}([\mathbf{0}, \mathbf{t}]) - n^{-1} \sum_{i=1}^n \mathbf{1}_{[\mathbf{0}, \mathbf{t}]}(\mathbf{x}^i) \right|;$$

see, e.g., [1], [5]. Note that the maximum over  $U$  no longer appears, because for  $\mathbf{c} = \mathbf{1}$  the largest value of  $\|h_U\|_{L_\infty}$  in (15) occurs at  $U = \{1, \dots, s\}$ . Theorem 1 also states that this inequality is attained arbitrarily closely for some functions from  $\mathcal{F}_{1,s}$ . We next consider the case in which the anchor  $\mathbf{c} = [c_1, \dots, c_s]$  is in the interior  $(0, 1)^s$  of the cube; for example  $c_1 = \dots = c_s = 1/2$ . In the right-hand side of (14)



we may assume that  $t_k \neq c_k$  for  $k = 1, \dots, s$ , since  $M_U(\cdot, \mathbf{t}_U) = 0$  if  $t_k = c_k$  for some  $k \in U$ . The set of all points  $\mathbf{t} \in [0, 1]^s$  which satisfy this condition is partitioned into  $2^s$  quadrants, according to whether  $t_k < c_k$  or  $t_k > c_k$  for  $k = 1, \dots, s$ . Given a  $\mathbf{t}$  in one of these quadrants, let

$$B(\mathbf{t}) = B(\mathbf{t}; \mathbf{c}) \subseteq [0, 1]^s$$

denote the box with one corner at  $\mathbf{t}$  and the opposite corner given by the unique vertex of  $D$  that lies in the same quadrant as  $\mathbf{t}$ . This box is to be a product of half open intervals, with the open end of the  $k$ th interval at  $t_k$ . For example, for  $s = 2$  and  $c_1 = c_2 = 1/2$ , if  $\mathbf{t} = (1/3, 2/3)$ , then

$$B(\mathbf{t}; \mathbf{c}) = [0, 1/3) \times (2/3, 1].$$

With this notation, it is easily seen that

$$I_1(M_U(\cdot, \mathbf{t}_U)) = (-1)^{|U_-|} \text{Vol}(B_U(\mathbf{t}_U)),$$

where  $B_U(\mathbf{t}_U)$  is the projection of  $B(\mathbf{t})$  onto  $D_U$ , and  $U_-$  is the subset of  $U$  containing those indices  $k$  for which  $t_k < c_k$ . Similarly,

$$\mathcal{A}(M_U(\cdot, \mathbf{t}_U)) = \frac{(-1)^{|U_-|}}{n} \sum_{i=1}^n \mathbf{1}_{B_U(\mathbf{t}_U)}(\mathbf{x}_U^i).$$

Thus the expression (14) is the same as that for the classical case  $\mathbf{c} = \mathbf{1}$ , except that the box  $\prod_k [0, t_k]$  is replaced by  $B(\mathbf{t})$ . Another difference is that for interior  $\mathbf{c}$ , the maximum with respect to  $U$  is not necessarily attained by  $U = \{1, \dots, s\}$ . In other words, Theorem 1 shows that the error of  $\mathcal{A}$  is the  $L_\infty$  same-quadrant discrepancy with anchor at  $\mathbf{c}$ , given by

$$(16) \quad \mathcal{D}_{\infty, \mathbf{c}}(\mathbf{x}^1, \dots, \mathbf{x}^n) := \sup_{\mathbf{t} \in D} \max_{U \neq \emptyset} \left| \text{Vol}(B_U(\mathbf{t}_U; \mathbf{c}_U)) - n^{-1} \sum_{i=1}^n \mathbf{1}_{B_U(\mathbf{t}_U; \mathbf{c}_U)}(\mathbf{x}^i) \right|.$$

For completeness of the definition (16), we define  $B_U(\mathbf{t}_U; \mathbf{c}_U) = \emptyset$  if  $t_k = c_k$  for some  $k \in U$ . With this understanding, the  $L_\infty$  star discrepancy  $\mathcal{D}_\infty^*(\mathbf{x}^1, \dots, \mathbf{x}^n)$  is the special case of (16) with  $\mathbf{c} = \mathbf{1}$ . When  $\mathbf{c}$  is at the center of the cube,  $\mathcal{D}_{\infty, \mathbf{c}}$  is also called the  $L_\infty$  centered discrepancy; see, e.g., [3], [7]. The following proposition sums up the classical case.

**Proposition 2.** *Let  $D = [0, 1]^s$  and  $\rho(\mathbf{x}) \equiv 1$ . The error of the quasi-Monte Carlo algorithm with the points  $\mathbf{x}^1, \dots, \mathbf{x}^n$  is*

$$\text{error}(\text{QMC}, \mathcal{F}_{1,s}) = \mathcal{D}_{\infty, \mathbf{c}}(\mathbf{x}^1, \dots, \mathbf{x}^n).$$

#### 4. REDUCTION TO CLASSICAL PROBLEM

Recall that by a classical problem we mean the integration problem over  $[0, 1]^s$  with constant weight function  $\rho \equiv 1$ . In this section, we show that the general problem over  $D$  with a general  $\rho$  can be reduced to a classical problem.

Given  $\rho$ , define the transformations

$$W_k(x) = \int_{a_k}^x \rho_k(z) dz, \quad k = 1, \dots, s.$$

Then  $W_k : \overline{(a_k, b_k)} \rightarrow [0, 1]$ . Define also

$$(17) \quad W(\mathbf{x}) = [W_1(x_1), \dots, W_s(x_s)] \quad \text{and} \quad \mathbf{d} = W(\mathbf{c}).$$

**Theorem 2.** *i) The error of any algorithm  $\mathcal{A}(f) = \sum_{i=1}^n a_i f(\mathbf{x}^i)$  for the integration problem (7) satisfies*

$$(18) \quad \text{error}(\mathcal{A}, \mathcal{F}_{1,s}) = \text{error}(\tilde{\mathcal{A}}, \tilde{\mathcal{F}}_{1,s}),$$

where  $\tilde{\mathcal{A}}(g) = \sum_{i=1}^n a_i g(W(\mathbf{x}^i))$  is an algorithm approximating  $I_1(g) = \int_{[0,1]^s} g(\mathbf{x}) \, d\mathbf{x}$  and  $\tilde{\mathcal{F}}_{1,s}$  is defined in the same way as  $\mathcal{F}_{1,s}$  but with  $D = [0, 1]^s$  and anchor  $\mathbf{d} = W(\mathbf{c})$ .

*ii) In particular, if  $\mathcal{A} = \mathcal{QMC}$  is a quasi-Monte Carlo algorithm, then*

$$(19) \quad \text{error}(\mathcal{QMC}, \mathcal{F}_{1,s}) = \mathcal{D}_{\infty, \mathbf{d}}(W(\mathbf{x}^1), \dots, W(\mathbf{x}^n)).$$

*Proof.* Due to Theorem 1, the error of  $\mathcal{A}$  is given by

$$\text{error}(\mathcal{A}, \mathcal{F}_{1,s}) = \max_{U \neq \emptyset} \sup_{\mathbf{t} \in D} |h_U(\mathbf{t})|$$

with

$$h_U(\mathbf{t}) = I_\rho(M_U(\cdot, \mathbf{t}_U)) - \sum_{i=1}^n a_i M_U(\mathbf{x}_U^i, \mathbf{t}_U).$$

Note that for  $t_k < c_k$ , or equivalently  $W_k(t_k) < d_k$ ,

$$\int_{a_k}^{b_k} M_k(x, t_k) \rho_k(x) \, dx = -W_k(t_k) = -\text{Vol}([0, W_k(t_k)])$$

and

$$M_k(x_k^i, t_k) = -\mathbf{1}_{[a_k, t_k]}(x_k^i) = -\mathbf{1}_{[0, W_k(t_k)]}(W_k(x_k^i)),$$

with the last equality holding because of the monotonicity of  $W_k$  and  $0 = W_k(a_k)$ .

On the other hand, for  $t_k > c_k$ , or equivalently  $W_k(t_k) > d_k$ ,

$$\int_{a_k}^{b_k} M_k(x, t_k) \rho_k(x) \, dx = 1 - W_k(t_k) = \text{Vol}([W_k(t_k), 1])$$

and

$$M_k(x_k^i, t_k) = \mathbf{1}_{(t_k, b_k]}(x_k^i) = \mathbf{1}_{(W_k(t_k), 1]}(W_k(x_k^i)).$$

This implies that

$$h_U(\mathbf{t}) = (-1)^{U^-} \left( \text{Vol}(B_U(\mathbf{y}_U; \mathbf{d}_U)) - \sum_{i=1}^n a_i \mathbf{1}_{B_U(\mathbf{y}_U, \mathbf{d}_U)}(W_U(\mathbf{x}_U^i)) \right),$$

where  $\mathbf{y}_U = W_U(\mathbf{t}_U)$ , from which (18) follows by a second application of Theorem 1.

The last part of the theorem then follows from Proposition 1.  $\square$

Theorem 2 implies that weighted integration over (bounded or unbounded)  $D$  has  $n$ th minimal error equal to the  $n$ th minimal error of the classical problem with the anchor  $\mathbf{d} = W(\mathbf{c})$ .

Since the classical case has been extensively studied, Theorem 2 provides (at least conceptually) a technique for studying tractability of and deriving algorithms for weighted integration. Indeed, suppose that we have a point-set  $\{\mathbf{y}^1, \dots, \mathbf{y}^n\}$

inside  $[0, 1]^s$  with low  $L_\infty$  same-quadrant discrepancy. Then for  $\|f\|_{1,s} \leq 1$ , the algorithm

$$\mathcal{QMC}(f) = n^{-1} \sum_{i=1}^n f(W^{-1}(\mathbf{y}^i))$$

approximates the weighted integral  $I_\rho(f)$  with error not exceeding the discrepancy of  $\{\mathbf{y}^1, \dots, \mathbf{y}^n\}$ .

## 5. TRACTABILITY OF THE WEIGHTED INTEGRATION PROBLEM

From (8) the initial error of the weighted integration problem in the space  $\mathcal{F}_{1,s}$  equals 1,

$$\text{error}(0, \mathcal{F}_{1,s}) = 1.$$

Thus according to Theorem 2 ii), to prove tractability it is enough to show that there exist  $n$  points  $\mathbf{z}^1, \dots, \mathbf{z}^n$  in  $(0, 1)^s$  whose  $L_\infty$  same-quadrant discrepancy anchored at  $\mathbf{d} \in D$  is bounded by a polynomial in  $s$  and  $n^{-1}$ . The first result in the following theorem is then an easy application of a result from [2]. The second is shown using proof ideas from [2]; however it is more elaborate and yields a sharper upper bound than the corresponding upper bound from Theorem 1 in [2].

**Theorem 3.** *The weighted integration problem (7) for a general region  $D$  of the form (1) and a general weight of the form (2) is tractable. The  $n$ th minimal error satisfies*

$$(20) \quad \text{error}(n, \mathcal{F}_{1,s}) \leq C \sqrt{s/n}$$

for an unknown constant  $C$  independent of  $s$  and  $n$ , and also

$$(21) \quad \text{error}(n, \mathcal{F}_{1,s}) \leq \left( \frac{s}{nc} \ln \left( 3 + \sqrt{sn/c \ln 2} \right) + \frac{\ln 2}{nc} \right)^{1/2}$$

with

$$c = 3 \ln 3 - 2 \simeq 1.2958.$$

*Proof.* Let  $\mathbf{d}$  be as in Theorem 2, and let  $\mathbf{z}^1, \dots, \mathbf{z}^n \in (0, 1)^s$ . Although the boxes  $B_U(\mathbf{t}_U; \mathbf{d}_U)$  in the definition of  $\mathcal{D}_{\infty, \mathbf{d}}$  are of dimension  $|U|$ , we can replace them by boxes of dimension  $s$ , since

$$\left| \text{Vol}(B_U(\mathbf{t}_U; \mathbf{d}_U)) - n^{-1} \sum_{i=1}^n \mathbf{1}_{B_U(\mathbf{t}_U; \mathbf{d}_U)}(\mathbf{z}_U^i) \right| = \left| \text{Vol}([\mathbf{g}, \mathbf{h}]) - n^{-1} \sum_{i=1}^n \mathbf{1}_{[\mathbf{g}, \mathbf{h}]}(\mathbf{z}^i) \right|,$$

where  $\mathbf{g}, \mathbf{h} \in [0, 1]^s$  with the  $k$ th edge of  $[\mathbf{g}, \mathbf{h}]$  being the same as that of  $B_U(\mathbf{t}_U; \mathbf{d}_U)$  if  $k \in U$ , and  $[0, 1)$  otherwise. Hence, the corresponding discrepancy is bounded from above by the corresponding *unanchored discrepancy*

$$\begin{aligned} \mathcal{D}_{\infty, \mathbf{d}}(\mathbf{z}^1, \dots, \mathbf{z}^n) &\leq \mathcal{D}_\infty(\mathbf{z}^1, \dots, \mathbf{z}^n) \\ &:= \sup_{\mathbf{g}, \mathbf{h} \in [0, 1]^s} \left| \text{Vol}([\mathbf{g}, \mathbf{h}]) - n^{-1} \sum_{i=1}^n \mathbf{1}_{[\mathbf{g}, \mathbf{h}]}(\mathbf{z}^i) \right|. \end{aligned}$$

Therefore, the first upper bound of the theorem follows from Theorem 2 ii) and the fact that there are points  $\mathbf{z}^1, \dots, \mathbf{z}^n \in (0, 1)^s$  for which  $\mathcal{D}_\infty(\mathbf{z}^1, \dots, \mathbf{z}^n) \leq C\sqrt{s/n}$ ; see, Theorem 4 in [2] and the discussion that follows it.

We now show the second result. For a positive  $\delta$ , define  $m = m(\delta, s) = \lceil s/\delta \rceil$ . Let  $\Gamma_m$  be the equidistant grid on  $[0, 1]^s$  with spacing  $1/m$  in each coordinate direction, and let  $\Gamma_{m,U}$  be its corresponding projection on  $[0, 1]^U$ . Of course, the cardinalities of these grids are

$$|\Gamma_m| = (m + 1)^s \quad \text{and} \quad |\Gamma_{m,U}| = (m + 1)^{|U|}.$$

As in [2] it can be shown that for any nonempty  $U$  and any points  $\mathbf{t}^i \in [0, 1]^s$ ,

$$\begin{aligned} & \sup_{\mathbf{y} \in [0,1]^s} \left| \text{Vol}(B_U(\mathbf{y}_U; \mathbf{d}_U)) - n^{-1} \sum_{i=1}^n \mathbf{1}_{B_U(\mathbf{y}_U; \mathbf{d}_U)}(\mathbf{t}_U^i) \right| \\ & \leq \delta + \sup_{\mathbf{y} \in \Gamma_{m,U}} \left| \text{Vol}(B_U(\mathbf{y}; \mathbf{d}_U)) - n^{-1} \sum_{i=1}^n \mathbf{1}_{B_U(\mathbf{y}_U; \mathbf{d}_U)}(\mathbf{t}_U^i) \right|. \end{aligned}$$

Now let the  $\mathbf{t}^i$  be independent random variables, uniformly distributed in  $[0, 1]^s$ . Given  $U$  and  $\mathbf{y} \in \Gamma_{m,U}$ , define random variables

$$\zeta_{\mathbf{y},U}^{(i)} = \text{Vol}(B_U(\mathbf{y}; \mathbf{d}_U)) - \mathbf{1}_{B_U(\mathbf{y}; \mathbf{d}_U)}(\mathbf{t}_U^i), \quad i = 1, \dots, n.$$

Of course, the expression in the last supremum equals the absolute value of  $\sum_{i=1}^n \zeta_{\mathbf{y},U}^{(i)}/n$ . We recall Bennett’s inequality (see, e.g., [10]), which states the following. Let  $Y_1, \dots, Y_n$  be independent random variables with zero means and bounded ranges  $|Y_i| \leq M$ . Let  $\sigma_i^2$  be the variance of  $Y_i$  and let  $V$  be any number such that  $V \geq \sum_{i=1}^n \sigma_i^2$ . Then for each  $\eta > 0$ ,

$$\text{Prob} \{ |Y_1 + \dots + Y_n| \geq \eta \} \leq 2 \exp \left( -\frac{\eta^2}{2V} B(M\eta/V) \right),$$

where  $B(t) = 2t^{-2} ((1+t) \ln(1+t) - t)$ . We shall use this inequality with  $Y_i = \zeta_{\mathbf{y},U}^{(i)}$ . Indeed,  $\zeta_{\mathbf{y},U}^{(i)}$  has a zero mean and its values are between  $-1$  and  $1$ . Denoting its variance by  $\sigma^2$  and taking  $\eta = n\delta$  and  $V = n\sigma^2$ , we get

(22)

$$\text{Prob} \left\{ \left| n^{-1} \sum_{i=1}^n \zeta_{\mathbf{y},U}^{(i)} \right| \geq \delta \right\} \leq 2 \exp \left( -n\sigma^2 ((1 + \delta/\sigma^2) \ln(1 + \delta/\sigma^2) - \delta/\sigma^2) \right).$$

It is easy to check that  $\sigma^2 = \text{Vol}(B_U(\mathbf{y}; \mathbf{d}_U))(1 - \text{Vol}(B_U(\mathbf{y}; \mathbf{d}_U))) \leq 1/4$ . Moreover, the right-hand side of (22) increases with  $\sigma^2$ . Hence

$$\text{Prob} \left\{ \left| n^{-1} \sum_i \zeta_{\mathbf{y},U}^{(i)} \right| \geq \delta \right\} \leq 2 \exp \left( -n4^{-1}((1 + 4\delta) \ln(1 + 4\delta) - 4\delta) \right).$$

Now assume that  $\delta \leq 1/2$ . Then we claim

$$(23) \quad (1 + 4\delta) \ln(1 + 4\delta) - 4\delta \geq 4c\delta^2$$

with  $c = 3 \ln 3 - 2$ . To show this, consider

$$f(x) = ((1 + x) \ln(1 + x) - x)/x^2 \quad \text{for} \quad x = 4\delta \in (0, 2].$$

To prove (23), it is enough to show that

$$(24) \quad \inf_{x \in (0,2]} f(x) \geq c/4.$$

Note that

$$f'(x) = g(x)/x^3 \quad \text{with} \quad g(x) = 2x - (2+x)\ln(1+x).$$

Moreover,

$$g'(x) = \frac{x}{1+x} - \ln(1+x) \quad \text{and} \quad g''(x) = \frac{-x}{(1+x)^2}.$$

Since  $g''(x)$  is negative,  $g'$  decreases and  $g'(x) \leq g'(0) = 0$ , implying in turn  $g(x) \leq g(0) = 0$ . Hence  $f$  also decreases and

$$f(x) \geq f(2) = \frac{3 \ln 3 - 2}{4} = \frac{c}{4},$$

as claimed in (24). Using (23), we get

$$\text{Prob} \left\{ \left| n^{-1} \sum_{i=1}^n \zeta_{\mathbf{y}, U}^{(i)} \right| \leq \delta \right\} \leq 2 \exp(-nc\delta^2).$$

Since the number of all possible pairs  $(\mathbf{y}, U)$  satisfies

$$\sum_{k=1}^s \binom{s}{k} (m+1)^k < (m+2)^s,$$

we have that

$$\begin{aligned} & \text{Prob} \{ \mathcal{D}_{\infty, \mathbf{d}}(\mathbf{t}^1, \dots, \mathbf{t}^n) \leq 2\delta \} \\ & > \text{Prob} \left\{ \max_U \max_{\mathbf{y} \in \Gamma_{m, U}} \left| \text{Vol}(B_U(\mathbf{y}; \mathbf{d}_U)) - n^{-1} \sum_{i=1}^n \mathbf{1}_{B_U(\mathbf{y}; \mathbf{d}_U)}(\mathbf{t}_U^i) \right| \leq \delta \right\} \\ & \geq 1 - 2(m+2)^s \exp(-nc\delta^2). \end{aligned}$$

The latter expression is nonnegative if  $\delta \geq \delta_0$ , where  $\delta_0 = \delta_0(n, s)$  is given by

$$\delta_0^2 = (nc)^{-1} (s \ln(\lceil s/\delta_0 \rceil + 2) + \ln(2)).$$

Of course,  $\delta_0^2 \geq (nc)^{-1}(s+1)\ln 2$ . Using this to estimate  $s/\delta_0$  on the right-hand side of the above equation, we conclude that there exist points  $\mathbf{t}^i$  with

$$\mathcal{D}_{\infty, \mathbf{d}}(\mathbf{t}^1, \dots, \mathbf{t}^n) \leq 2\delta$$

when

$$\delta^2 \geq \left( s \ln \left( 3 + \sqrt{sn c / \ln 2} \right) + \ln 2 \right) / (nc).$$

This completes the proof.  $\square$

We end this section with the following remark.

*Remark 1.* Recall that the maximum over the subsets  $U$  occurs in the definition (16) of the same-quadrant discrepancy. Consider now a discrepancy with  $U$  fixed and equal to  $\{1, \dots, s\}$ . That is, let

$$\mathcal{D}_{\infty, \mathbf{d}, s}(\mathbf{x}^1, \dots, \mathbf{x}^n) := \sup_{\mathbf{y} \in [0, 1]^s} \left| \text{Vol}(B(\mathbf{y}; \mathbf{d})) - n^{-1} \sum_{i=1}^n \mathbf{1}_{B(\mathbf{y}; \mathbf{d})}(\mathbf{x}^i) \right|.$$

As can easily be shown, this discrepancy is equal to the error of the quasi-Monte Carlo algorithm for the subspace of  $\mathcal{F}_{1, s}$  consisting of functions that satisfy the zero conditions at the anchor, i.e.,  $f_U \equiv 0$  for any proper subset  $U$  of  $\{1, \dots, s\}$ .

Suppose also that the anchor  $\mathbf{d}$  is inside the cube  $D = [0, 1]^s$  and denote by  $\gamma$  its distance from the boundary of  $D$ , measured in  $\|\cdot\|_\infty$ -norm. That is,

$$\gamma = \gamma(\mathbf{d}) := \max_k \max\{d_k, 1 - d_k\} < 1.$$

Observe that now  $\text{Vol}(B(\mathbf{y}; \mathbf{d})) \leq \gamma^s$  and

$$\sigma^2 = \text{Vol}(B(\mathbf{y}; \mathbf{d}))(1 - \text{Vol}(B(\mathbf{y}; \mathbf{d}))) \leq \gamma^s$$

for any point  $\mathbf{y}$ . Hence, repeating the proof of Theorem 3, one can show that there are points  $\mathbf{x}^i$  for which

$$(25) \quad \mathcal{D}_{\infty, \mathbf{d}, s}(\mathbf{x}^1, \dots, \mathbf{x}^n) \leq ((nc_s)^{-1} (\ln(2 + s\sqrt{nc_s}) + s^{-1} \ln 2))^{1/2}$$

with

$$c_s = \ln \gamma^{-1} - \ln(2e)/s.$$

Consider now a sequence of problems for increasing  $s$  with anchors  $\mathbf{d}_s$  such that

$$\gamma_\infty := \sup_s \gamma(\mathbf{d}_s) < 1;$$

e.g.,  $\gamma_\infty = 1/2$  for  $\mathbf{d}_s = (1/2, \dots, 1/2)$ . Then

$$c_s \geq c_\infty = \ln(1/\gamma_\infty)(1 + o(1)),$$

and

$$\begin{aligned} & \inf_{\{\mathbf{x}^i\}} \mathcal{D}_{\infty, \mathbf{d}, s}(\mathbf{x}^1, \dots, \mathbf{x}^n) \\ & \leq ((n \ln(1/\gamma_\infty))^{-1} (\ln(s) + 1/2 \ln(n \ln(\gamma_\infty))))^{1/2} \cdot (1 + o(1)) \quad \text{as } s \rightarrow \infty. \end{aligned}$$

Since for large  $s$  this is much smaller than the bounds given in (20) and (21), it seems unlikely that the maximum over  $U$  is achieved at  $U = \{1, \dots, s\}$  when the anchor  $\mathbf{d}$  is in the interior with  $\gamma_\infty < 1$ .

## 6. EXTENSIONS

We conclude the paper by discussing a number of possible generalizations.

**6.1. More general weight  $\rho$ .** The tractability results of this and the preceding sections hold for weights  $\rho$  of the product form (2) that are not probability densities, as long as  $\rho \geq 0$  and

$$\alpha_k := \int_{a_k}^{b_k} \rho_k(x) dx < \infty, \quad k = 1, 2, \dots$$

Let  $\alpha = \prod_{k=1}^s \alpha_k$ . For any linear algorithm  $\mathcal{A}$ ,

$$I_\rho(f) - \mathcal{A}(f) = \alpha \cdot (I_{\hat{\rho}}(f) - \alpha^{-1} \mathcal{A}(f)),$$

where now  $\hat{\rho} = \rho/\alpha$  is a probability density. Hence, the error bounds from previous theorems should be multiplied by  $\alpha$ . However, since the initial error now equals  $\alpha$ , this does not change the tractability property.

**6.2. More general class of functions.** The assumption that  $\|f\|_{1,s}$  is finite might be too strong for a number of applications with unbounded domain  $D$ . Indeed, it requires that the derivatives  $f'_U(\mathbf{x}_U)$  converge to zero sufficiently fast with  $\|\mathbf{x}\| \rightarrow \infty$ . In this subsection, we show how it can be weakened by considering tractability with respect to a larger class of functions. Suppose that  $\|f\nu\|_{1,s} < \infty$  for a positive function  $\nu$ . We assume that

$$(26) \quad \nu(\mathbf{x}) = \prod_{k=1}^s \nu_k(x_k)$$

and that

$$(27) \quad \beta := \int_D \frac{\rho(\mathbf{x})}{\nu(\mathbf{x})} d\mathbf{x} < \infty.$$

Replacing  $\rho$  by

$$\rho_\nu(\mathbf{x}) := \frac{\rho(\mathbf{x})}{\beta\nu(\mathbf{x})},$$

we conclude from Theorem 2 that the integral

$$I_\rho(f) = \beta I_{\rho_\nu}(f\nu)$$

can be approximated by

$$\mathcal{A}(f) = \beta \mathcal{QMC}(f\nu),$$

with the error bounded by  $\beta \|f\nu\|_{1,s} \mathcal{D}_{\infty,d}(W(\mathbf{x}^1), \dots, W(\mathbf{x}^n))$ . Since the initial error is also multiplied by  $\beta$ , the problem with bounded  $\|f\nu\|_{1,s}$  remains tractable.

To illustrate this, consider  $D = \mathbb{R}^s$ ,

$$\rho(\mathbf{x}) = \prod_{i=1}^s (2\sigma_i)^{-1/2} \exp\left(-\sum_{k=1}^s x_k^2/(2\sigma_k)\right),$$

and

$$\nu(\mathbf{x}) = \prod_{i=1}^s (2\delta_i)^{-1/2} \exp\left(-\sum_{k=1}^s x_k^2/(2\delta_k)\right)$$

with  $\sigma_i < \delta_i$ . Then the corresponding problem of approximating  $I_\rho(f)$  is tractable even though the space of functions with bounded  $\|f\nu\|_{1,s}$  is relatively very large, certainly much larger than the original  $\mathcal{F}_{1,s}$ .

**6.3. Space  $\mathcal{F}_{p,s}$ .** The error analysis performed in this paper for the space  $\mathcal{F}_{1,s}$  has an analog for  $\mathcal{F}_{p,s}$  for  $1 < p \leq \infty$  that we summarize here. However, there are some wrinkles. In the proof of Theorem 1 one may take a general form of Hölder's inequality to obtain

$$(28) \quad \text{error}(\mathcal{A}, \mathcal{F}_{p,s}) = \left\{ \sum_U \int_{D_U} |I_\rho(M_U(\cdot, \mathbf{t}_U)) - \mathcal{A}(M_U(\cdot, \mathbf{t}_U))|^q d\mathbf{t}_U \right\}^{1/q},$$

where  $p^{-1} + q^{-1} = 1$ . By the proof of Theorem 2 it is known that

$$\begin{aligned} & \left\{ \sum_U \int_{D_U} |I_\rho(M_U(\cdot, \mathbf{t}_U)) - \mathcal{A}(M_U(\cdot, \mathbf{t}_U))|^q dt_U \right\}^{1/q} \\ &= \left\{ \sum_U \int_{D_U} |h_U(\mathbf{t})|^q dt_U \right\}^{1/q} \\ &= \left\{ \sum_U \int_{[0,1]^U} \left| \text{Vol}(B_U(\mathbf{y}_U; \mathbf{d}_U)) \right. \right. \\ & \quad \left. \left. - n^{-1} \sum_{i=1}^n \mathbf{1}_{B_U(\mathbf{y}_U; \mathbf{d}_U)}(W_U(\mathbf{x}_U^i)) \right|^q \frac{d\mathbf{y}_U}{\rho_U(W^{-1}(\mathbf{y}_U))} \right\}^{1/q} \end{aligned}$$

The quantity above looks like the  $L_q$  star discrepancy except for the factor  $1/\rho_U(W^{-1}(\mathbf{y}_U))$ .

Going back to the definition of the Banach space of integrands, one may define, instead of  $\mathcal{F}_{p,s}$ , the spaces  $\tilde{\mathcal{F}}_{p,s}$  with norms given by

$$(29a) \quad \|f\|_{p,s} := \left( \sum_U \|f'_U/\rho_U^{1/q}\|_{L_p}^p \right)^{1/p} \quad \text{for } p < \infty$$

and

$$(29b) \quad \|f\|_{\infty,s} := \max_U \|f'_U/\rho_U\|_{L_\infty} \quad \text{for } p = \infty.$$

Note that  $1/q = 0$  for  $p = 1$  and, hence  $\|\cdot\|_{1,s}$  and  $\|\cdot\|_{1,s}$  coincide. For this space of functions one has

$$\begin{aligned} \text{error}(\mathcal{A}, \tilde{\mathcal{F}}_{p,s}) &= \left\{ \sum_U \int_{D_U} |I_\rho(M_U(\cdot, \mathbf{t}_U)) - \mathcal{A}(M_U(\cdot, \mathbf{t}_U))|^q \rho_U(t_U) dt_U \right\}^{1/q} \\ &= \left\{ \sum_U \int_{[0,1]^U} \left| \text{Vol}(B_U(\mathbf{y}_U; \mathbf{d}_U)) - n^{-1} \sum_{i=1}^n \mathbf{1}_{B_U(\mathbf{y}_U; \mathbf{d}_U)}(W_U(\mathbf{x}_U^i)) \right|^q d\mathbf{y}_U \right\}^{1/q} \\ &= \mathcal{D}_{q,\mathbf{d}}(W(\mathbf{x}^1), \dots, W(\mathbf{x}^n)). \end{aligned}$$

So, the worst-case quadrature error for the space  $\tilde{\mathcal{F}}_{p,s}$  is the  $L_q$  same-quadrant discrepancy with anchor at  $\mathbf{d} = W(\mathbf{c})$ .

For unbounded domains one always has  $\rho_k(x_k) \rightarrow 0$  as  $x_k \rightarrow \infty$ , which implies that the derivatives of  $f$  must decay quite rapidly to zero. Thus, the spaces  $\tilde{\mathcal{F}}_{p,s}$  are typically much smaller than  $\mathcal{F}_{p,s}$ , except for the case  $p = 1$  where they coincide. On the other hand, one may also use the device of the previous subsection to obtain the error bound

$$|I_\rho(f) - \mathcal{A}(f)| \leq \beta \cdot \|f\nu\|_{p,s} \cdot \mathcal{D}_{q,\mathbf{d}}(W(\mathbf{x}^1), \dots, W(\mathbf{x}^n)).$$

By choosing  $\nu$  decaying to zero fast enough (but not too fast to make  $\beta$  infinite), one may enlarge the space of integrands for which one has a finite error bound.

Even for the classical problem tractability requires some weighting of the coordinates for  $p > 1$ ; see [7]. For general regions and weighted integrals, this problem will be taken up in another paper [4].



**Acknowledgments.** We thank S. Heinrich and H. Woźniakowski for valuable comments. The research was partially supported by the Hong Kong Research Grants Council grant HKBU/2030/99P, the Hong Kong Baptist University Faculty Research Grant FRG/00-01/II-62, the Australian Research Council, and the National Science Foundation under Grant CCR-0095709. The research was done when the third author visited the Hong Kong Baptist University and the University of New South Wales, and he thanks his hosts for their hospitality.

## REFERENCES

- [1] M. Drmota and R. F. Tichy, *Sequences, Discrepancies and Applications*, Lecture Notes in Math. **1651**, Springer, Berlin, 1997. MR **98j**:11057
- [2] S. Heinrich, E. Novak, G. W. Wasilkowski, and H. Woźniakowski, The inverse of the star-discrepancy depends linearly on the dimension, *Acta Arithmetica* **XCVI.3**, pp. 279-302, 2001. MR **2002b**:11103
- [3] F. H. Hickernell, A generalized discrepancy and quadrature error bound, *Math. Comp.* **67**, pp. 299-322, 1998. MR **98c**:65032
- [4] F. H. Hickernell, I. H. Sloan, and G. W. Wasilkowski, On tractability of integration for certain Banach spaces of functions, Monte Carlo and Quasi-Monte Carlo 2003 (H. Niederreiter, ed.), Springer 2003 (to appear).
- [5] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia, 1992. MR **93h**:65008
- [6] E. Novak, *Deterministic and Stochastic Error Bounds in Numerical Analysis*, Lecture Notes in Mathematics **1349**, Springer, 1988. MR **90a**:65004
- [7] E. Novak and H. Woźniakowski, Intractability results for integration and discrepancy, *J. Complexity* **17**, pp. 388-441, 2001. MR **2002f**:65204
- [8] S. H. Paskov, New methodologies for valuing derivatives, *Mathematics of Derivative Securities* (S. Pliska, M. Dempster eds.), pp. 545-582, Cambridge University Press, 1997.
- [9] S. H. Paskov and J. F. Traub, Faster valuation of financial securities, *J. Portfolio Management* **22**, pp. 113-120, 1995.
- [10] D. Pollard, *Convergence of Stochastic Processes*, 1984, Springer Verlag. MR **86i**:60074
- [11] I. H. Sloan, QMC integration—beating intractability by weighting the coordinate directions, *Monte Carlo and Quasi-Monte Carlo Methods 2000* (K.-T. Fang, F.J. Hickernell, H. Niederreiter, eds.), pp. 103-123, Springer 2002.
- [12] I. H. Sloan and H. Woźniakowski, When are quasi-Monte Carlo algorithms efficient for high dimensional integrals?, *J. Complexity* **14**, pp. 1-33 (1998). MR **99d**:65384
- [13] J. F. Traub and A. G. Werschulz, *Complexity and Information*, Cambridge University Press, 1998. MR **2000m**:65170
- [14] J. F. Traub, G. W. Wasilkowski, and H. Woźniakowski, *Information-Base Complexity*, Academic Press, New York, 1988. MR **90f**:68085
- [15] S. K. Zaremba, Some applications of multivariate integration by parts, *Ann. Pol. Math.* **21**, pp. 95-96, 1968. MR **38**:4034

DEPARTMENT OF MATHEMATICS, HONG KONG BAPTIST UNIVERSITY, KOWLOON TONG, HONG KONG

*E-mail address:* fred@math.hkbu.edu.hk

SCHOOL OF MATHEMATICS, UNIVERSITY OF NEW SOUTH WALES, SYDNEY 2052, AUSTRALIA

*E-mail address:* sloan@maths.unsw.edu.au

DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF KENTUCKY, 773 ANDERSON HALL, LEXINGTON, KENTUCKY 40506-0046

*E-mail address:* greg@cs.uky.edu