Weighted Geometric Discrepancies and Numerical Integration on Reproducing Kernel Hilbert Spaces

Michael Gnewuch

Institut für Informatik, Christian-Albrechts-Universität Kiel,
Christian-Albrechts-Platz 4,
24098 Kiel, Germany

December 19, 2010

Abstract

We extend the notion of L_2-B-discrepancy introduced in [E. Novak, H. Woźniakowski, L_2 discrepancy and multivariate integration, in: Analytic number theory. Essays in honour of Klaus Roth. W. W. L. Chen, W. T. Gowers, H. Halberstam, W. M. Schmidt, and R. C. Vaughan (Eds.), Cambridge University Press, Cambridge, 2009, 359 – 388] to what we want to call weighted geometric L_2-discrepancy. This extended notion allows us to consider weights to moderate the importance of different groups of variables, and additionally volume measures different from the Lebesgue measure as well as classes of test sets different from measurable subsets of Euclidean spaces.

We relate the weighted geometric L_2-discrepancy to numerical integration defined over weighted reproducing kernel Hilbert spaces and settle in this way an open problem posed by Novak and Woźniakowski.

Furthermore, we prove an upper bound for the numerical integration error for cubature formulas that use admissible sample points. The set of admissible sample points may actually be a subset of the integration domain of measure zero. We illustrate that particularly in infinite-dimensional numerical integration it is crucial to distinguish between the whole integration domain and the set of those sample points that actually can be used by algorithms.

1 Introduction

It is known that many notions of L_2-discrepancy are intimately related to multivariate or infinite-dimensional numerical integration over corresponding normed function spaces, see, e.g., [Zar68, Woź91, Hic98, SW98, HW01, NW01a, NW01b, NW09, DP10, NW10] and the related literature mentioned therein. In particular, Novak and Woźniakowski introduced in [NW09] (see also [NW10, Chapter 9]) the quite general notion of L_2-B-discrepancy. Here B refers to a function that maps elements t from some measurable Euclidean set D to measurable subsets $B(t)$ of $\mathbb{R}^d$. The L_2-B-discrepancy of a point set $\{t_1, \dots, t_n\}$ and real coefficients $a_1, \dots, a_n$ is then taken with respect to the class of test sets $\mathcal{B} = \{B(t) \,|\, t \in D\}$ and a probability density ρ on D,

$$\operatorname{disc}_2^B(\{t_j\}, \{a_j\}) = \left(\int_D \left(\operatorname{vol}(B(t)) - \sum_{j=1}^n a_j 1_{B(t)}(t_j) \right)^2 \rho(t) \, \mathrm{d}t \right)^{1/2},$$

where $1_{B(t)}$ is the characteristic function of the set $B(t)$ and $\operatorname{vol}(B(t))$ is the d-dimensional Lebesgue measure of $B(t)$, see also Section 7.1. Novak and Woźniakowski showed that the L_2-B-discrepancy corresponds to multivariate numerical integration over a Hilbert space with some reproducing kernel K_d related to the class of test sets $\mathcal{B}$ and the probability density ρ.

Their notion of L_2-B-discrepancy does not take into account the concept of weights to model the different importance of distinct subsets of coordinates, which is often helpful to overcome the curse of dimensionality. In the context of multivariate numerical integration such weights were probably first studied by Sloan and Woźniakowski in [SW98].

In their new book [NW10] Novak and Woźniakowski posed the open problem to extend the notion of L_2-B-discrepancy to include weights and to find relations of the new discrepancy notion to multivariate numerical integration over weighted reproducing kernel Hilbert spaces (cf. [NW10, Open Problem 35]).

In this paper we introduce the even more general definition of weighted geometric L_2-discrepancy[1], which allows not only to consider weights, but also admits measures that may differ from the Lebesgue measure on domains that are not necessarily measurable subsets of $\mathbb{R}^d$. Especially, it covers discrepancies related to infinite-dimensional numerical integration. We prove relations of this discrepancy notion to numerical integration over corresponding weighted reproducing kernel Hilbert spaces and thus, in particular, settle the open problem posed by Novak and Woźniakowski.

The paper is organized as follows: In Section 2 we introduce the setting we want to consider and state the general assumptions we want to make throughout the paper. In Section 3 we define the weighted geometric L_2-discrepancy and in Section 4 we introduce the numerical integration problems we want to study. We call the worst case error of integration by linear algorithms "weighted numerical discrepancy". With this notion the central question of Section 5 can be put as "Under which conditions do weighted geometric discrepancy and weighted numerical L_2-discrepancy coincide?". Of special interest is the situation, where the test sets which are used to determine the discrepancy and the measures on these classes of test sets exhibit a certain product structure, see Section 5.2. In Section 6 we prove an upper bound for the weighted geometric and the weighted

[1] The term "geometric discrepancy" has been used in the literature before, see, e.g., the title of the monograph [Mat10], but, as far as we can see, this term has never been defined in a rigorous way.

numerical L_2-discrepancy. Stated in the setting of numerical integration, we prove that there exist linear algorithms using n admissible sample points such that the integration error is smaller than a constant divided by $\sqrt{n}$. By refining the standard quasi-Monte Carlo averaging proof technique, we get this result also for sets of admissible sample points which may form a subset of measure zero of the actual integration domain. In Section 7 we discuss several examples.

2 General Assumptions

Let (M, Σ, μ) be a measure space. We assume M to be σ-finite, i.e., M can be written as a countable union of sets of finite measure.

Let I be a countable index set which may have finitely or infinitely many elements. For $\nu \in I$ let $(M_\nu, \Sigma_\nu, \mu_\nu)$ be a σ-finite measure space, which is related to the measure space (M, Σ, μ) in the following way: There exists a surjective measurable map $\Phi_\nu : M \to M_\nu$ such that μ_ν is the direct image of μ under Φ_ν, i.e., $\mu_\nu = \mu \circ \Phi_\nu^{-1}$. In particular, we have $\mu_\nu(M_\nu) = \mu(M)$.

Most important for us is the case where Φ_ν is some kind of projection and thus typically a non-injective function. Hence we understand Φ_ν^{-1} not as a function on M_ν, but as a function on the power set of M_ν – it maps each subset A of M_ν to its pre-image $\Phi_\nu^{-1}(A) := \{m \in M \mid \Phi_\nu(m) \in A\}$.

Let $\mathcal{B}_\nu$ be a subset of Σ_ν, consisting of sets of finite measure, endowed with a σ-algebra $\Sigma(\mathcal{B}_\nu)$ and a probability measure ω_ν. We put $\mathcal{B} := (\mathcal{B}_\nu)_{\nu \in I}$. We assume for all $\nu \in I$ that the function

$$\chi_\nu : \, M_\nu \times \mathcal{B}_\nu \to \{0, 1\} \, , \, (x_\nu, B_\nu) \mapsto 1_{B_\nu}(x_\nu) \tag{1}$$

is measurable with respect to the product σ-algebra $\Sigma_\nu \otimes \Sigma(\mathcal{B}_\nu)$ on $M_\nu \times \mathcal{B}_\nu$. Due to Tonelli's theorem the function

$$B_\nu \mapsto \mu_\nu(B_\nu) = \int_{M_\nu} 1_{B_\nu}(x_\nu) \, \mathrm{d}\mu_\nu(x_\nu)$$

is measurable with respect to $\Sigma(\mathcal{B}_\nu)$. Additionally, we require that

$$\int_{\mathcal{B}_\nu} \mu_\nu(B_\nu)^2 \, \mathrm{d}\omega_\nu(B_\nu) < \infty. \tag{2}$$

Let $\gamma := (\gamma_\nu)_{\nu \in I}$ be a family of non-negative weights, i.e., $\gamma_\nu \in [0, \infty)$ for all $\nu \in I$.

Furthermore, we consider a subset S of M which we want to call *set of admissible sample points*. For many discrepancies and numerical integration problems S will be equal to M. But for some numerical integration problems, in particular for infinite-dimensional integration as described in Sect. 7.4, S will be a proper subset or even a null set of M. With regard to such applications it is particularly important to distinguish between S and M in Sect. 6 and Theorem 6.1.

3 Weighted Geometric L_2-Discrepancy

For $\nu \in I$ we define the *local (geometric) discrepancy function* of a multi-set of points $\{t_{1,\nu}, \ldots, t_{n,\nu}\}$ in M_ν for a multi-set of real coefficients $\{a_1, \ldots, a_n\}$ and a test set $B_\nu \in \mathcal{B}_\nu$ by

$$\operatorname{disc}(B_\nu, \{t_{j,\nu}\}, \{a_j\}) := \mu_\nu(B_\nu) - \sum_{j=1}^{n} a_j 1_{B_\nu}(t_{j,\nu}), \tag{3}$$

and the *weighted geometric L_2-discrepancy* for a multi-set $\{t_1, \ldots, t_n\}$ in M with respect to $\mathcal{B} = (\mathcal{B}_\nu)_{\nu\in I}$ and $\gamma = (\gamma_\nu)_{\nu\in I}$ by

$$\operatorname{disc}^{\mathcal{B}}_{2,\gamma}(\{t_j\}, \{a_j\}) := \left(\sum_{\nu\in I} \gamma_\nu \int_{\mathcal{B}_\nu} \operatorname{disc}(B_\nu, \{\Phi_\nu(t_j)\}, \{a_j\})^2 \, \mathrm{d}\omega_\nu(B_\nu) \right)^{1/2}. \tag{4}$$

We suppress the attribute “weighted” if all weights except of one are equal to zero. We deduce from (3)

$$\begin{aligned} &\operatorname{disc}^{\mathcal{B}}_{2,\gamma}(\{t_j\}, \{a_j\}) \\ =&\left(\sum_{\nu\in I} \gamma_\nu \left[\int_{\mathcal{B}_\nu} \mu_\nu(B_\nu)^2 \, \mathrm{d}\omega_\nu(B_\nu) - 2 \sum_{j=1}^{n} a_j \int_{\mathcal{B}_\nu} \mu_\nu(B_\nu) 1_{B_\nu}(\Phi_\nu(t_j)) \, \mathrm{d}\omega_\nu(B_\nu) \right.\right. \\ &\left.\left. + \sum_{i,j=1}^{n} a_i a_j \int_{\mathcal{B}_\nu} 1_{B_\nu}(\Phi_\nu(t_i)) 1_{B_\nu}(\Phi_\nu(t_j)) \, \mathrm{d}\omega_\nu(B_\nu) \right] \right)^{1/2}. \end{aligned} \tag{5}$$

We are mostly interested in the situation where $\operatorname{disc}^{\mathcal{B}}_{2,\gamma}(\{t_j\}, \{a_j\})$ is finite for any choice of $\{t_j\}$. Due to (5) and (2) this is always satisfied for finite I, and, if the weights γ decay rapidly enough, also for infinite I, see the examples in Section 7. If, e.g., $\mu(M)$ is finite, then it is sufficient that $\sum_{\nu\in I} \gamma_\nu < \infty$.

Let us define the *nth S-minimal weighted geometric L_2-discrepancy* $\operatorname{disc}^{\mathcal{B}}_{2,\gamma}(n, S)$ by

$$\operatorname{disc}^{\mathcal{B}}_{2,\gamma}(n, S) := \inf\{\operatorname{disc}^{\mathcal{B}}_{2,\gamma}(\{t_j\}, \{a_j\}) \,|\, t_1, \ldots, t_n \in S, a_1, \ldots, a_n \in \mathbb{R}\}.$$

4 Integration on Weighted Reproducing Kernel Hilbert Spaces

Let $(\widetilde{K}_\nu)_{\nu\in I}$ be a family of reproducing kernels $\widetilde{K}_\nu : M_\nu \times M_\nu \to \mathbb{R}$. That is, for each $\nu \in I$ the function $\widetilde{K}_\nu$ is symmetric

$$\widetilde{K}_\nu(x_\nu, y_\nu) = \widetilde{K}_\nu(y_\nu, x_\nu) \quad \text{for all } x_\nu, y_\nu \in M_\nu$$

and positive semi-definite

$$\sum_{i,j=1}^{n} \widetilde{K}_\nu(x_i, x_j)\xi_i\xi_j \geq 0 \quad \text{for all } n \in \mathbb{N},\, x_1, \ldots, x_n \in M_\nu,\, \xi_1, \ldots, \xi_n \in \mathbb{R}.$$

In general, we denote the reproducing kernel Hilbert space of a reproducing kernel K by $H(K)$ and its scalar product by $\langle \cdot\,, \cdot \rangle_{H(K)}$. Our standard reference for the theory of reproducing kernel Hilbert spaces and their kernels is [Aro50]. We assume that $\widetilde{K}_\nu$ is measurable on $M \times M$ for all $\nu \in I$. For each $\nu \in I$ the function K_ν, defined by

$$K_\nu(x, y) := \widetilde{K}_\nu(\Phi_\nu(x), \Phi_\nu(y)) \quad \text{for all } x, y \in M,$$

inherits from $\widetilde{K}_\nu$ the properties of symmetry and of positive semi-definiteness, and is therefore a reproducing kernel on $M \times M$. Furthermore, K_ν is measurable on $M \times M$. Let us assume that

$$\sum_{\nu \in I} \gamma_\nu K_\nu(x, x) < \infty \quad \text{for all } x \in M, \tag{6}$$

which, of course, is trivially satisfied if I is a finite set. Since

$$|K_\nu(x, y)|^2 \leq K_\nu(x, x) K_\nu(y, y) \quad \text{for all } x, y \in M,$$

the function K_γ defined by

$$K_\gamma(x, y) := \sum_{\nu \in I} \gamma_\nu K_\nu(x, y) \quad \text{for all } x, y \in M, \tag{7}$$

is well-defined. K_γ is a measurable map and a reproducing kernel on $M \times M$, see [Aro50, Sect. I.9, Thm.II]. The corresponding Hilbert space $H(K_\gamma)$ can be described as follows: If we assume for convenience that $I = \mathbb{N}$ and $\gamma_\nu > 0$ for all $\nu \in I$, we may define for $n \in \mathbb{N}$ the Hilbert space $F_n = \sum_{\nu=1}^n H(K_\nu)$ with the norm

$$\|f\|_n^2 := \min \sum_{\nu=1}^n \gamma_\nu^{-1} \|f_\nu\|_{H(K_\nu)}^2,$$

where the minimum is taken over all decompositions $f = \sum_{\nu=1}^n f_\nu$, $f_\nu \in H(K_\nu)$. Put $F_0 := \cup_{n \in \mathbb{N}} F_n$, endowed with the norm $\|f\|_0 = \lim_{n \to \infty} \|f\|_n$. (The limit exists, since we have for $n \geq m$ and $f \in F_m$ that $\|f\|_n \leq \|f\|_m$.) Now $f_0^* : M \to \mathbb{R}$ is in $H(K_\gamma)$ if and only if there exists a Cauchy sequence $(f_0^{(n)})_{n \in \mathbb{N}}$ in F_0 with

$$f_0^*(x) := \lim_{n \to \infty} f_0^{(n)}(x) \quad \text{for all } x \in M. \tag{8}$$

The norm of f_0^* in $H(K_\gamma)$ is then given by

$$\|f_0^*\|_{H(K_\gamma)} = \min \lim_{n \to \infty} \|f_o^{(n)}\|_0,$$

where the minimum is taken over all Cauchy sequences $(f_0^{(n)})_{n \in \mathbb{N}}$ in F_0 that satisfy (8).

Recall that due to the reproducing kernel properties we have

$$K_\gamma(\cdot, y) \in H(K_\gamma) \quad \text{for all } y \in M$$

and

$$f(x) = \langle f, K_\gamma(\cdot, x) \rangle_{H(K_\gamma)} \quad \text{for all } f \in H(K_\gamma),\ x \in M$$

(and the same holds, of course, if we substitute all γs by any fixed $\nu \in I$).

Lemma 4.1. *For all $x \in M$ and all $\nu \in I$ we have $K_\nu(\cdot, x) \in H(K_\gamma)$. Furthermore,*

$$K_\gamma(\cdot, x) = \sum_{\nu \in I} \gamma_\nu K_\nu(\cdot, x), \tag{9}$$

where the sum converges unconditionally to $K_\gamma(\cdot, x)$ in $H(K_\gamma)$.

The lemma follows again from [Aro50, Sect. I.9, Thm.II].

We assume that $H(K_\gamma)$ consists of integrable functions with respect to μ and that the integral

$$\mathcal{I}(f) = \int_M f(x) \, \mathrm{d}\mu(x)$$

is a bounded linear functional on $H(K_\gamma)$, i.e, that the function

$$h_\gamma := \int_M K_\gamma(x, \cdot) \, \mathrm{d}\mu(x) \in H(K_\gamma). \tag{10}$$

Note that

$$\mathcal{I}(f) = \langle f, h_\gamma \rangle_{H(K_\gamma)} \quad \text{for all } f \in H(K_\gamma);$$

the function h_γ is called the representer of $\mathcal{I}$ in $H(K_\gamma)$.

From Lemma 4.1 follows for all $y \in M$ that $K_\nu(\cdot, y)$ is integrable with respect to μ and

$$\begin{aligned} \int_M K_\gamma(x, y) \, \mathrm{d}\mu(x) &= \langle h_\gamma, K_\gamma(\cdot, y) \rangle_{H(K_\gamma)} = \sum_{\nu \in I} \gamma_\nu \langle h_\gamma, K_\nu(\cdot, y) \rangle_{H(K_\gamma)} \\ &= \sum_{\nu \in I} \gamma_\nu \int_M K_\nu(x, y) \, \mathrm{d}\mu(x). \end{aligned} \tag{11}$$

Furthermore, $h_\gamma \in H(K_\gamma)$ implies that h_γ is integrable with respect to μ and

$$\|h_\gamma\|^2_{H(K_\gamma)} = \int_M \left(\int_M K_\gamma(x, y) \, \mathrm{d}\mu(x) \right) \mathrm{d}\mu(y) < \infty. \tag{12}$$

Notice that $\|h_\gamma\|_{H(K_\gamma)}$ is the operator norm of $\mathcal{I}$. Since we are only interested in non-trivial integration problems, we assume $\|h_\gamma\|_{H(K_\gamma)} > 0$. Furthermore, we assume that the kernel functions

$$K_\gamma \text{ and } K_\nu, \nu \in I, \text{ are integrable on } M \times M \tag{13}$$

and

$$\int_M \int_M K_\gamma(x, y) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y) = \sum_{\nu \in I} \gamma_\nu \int_M \int_M K_\nu(x, y) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y). \tag{14}$$

In the important case where for each $\nu \in I$ the kernel K_ν takes only non-negative values (see Sect. 5), (12) and Tonelli's theorem already imply the integrability of K_γ on $M \times M$ which in turn, together with the dominated convergence theorem, ensures the integrability of the K_νs and (14). For convenience, we want to call weights γ that ensure that all assumptions made above are satisfied *admissible weights*.

Let Q_n be a linear algorithm given by

$$Q_n(f) = \sum_{j=1}^{n} a_j f(t_j) \quad \text{with } t_1, \dots, t_n \in S \text{ and } a_1, \dots, a_n \in \mathbb{R}. \tag{15}$$

Then

$$\mathcal{I}(f) - Q_n(f) = \langle f, h_{\gamma,n} \rangle_{H(K_\gamma)} \quad \text{for all } f \in H(K_\gamma),$$

where

$$h_{\gamma,n} := h_\gamma - \sum_{j=1}^{n} a_j K_\gamma(\cdot, t_j).$$

If we want to approximate the functional $\mathcal{I}$ by the linear algorithm Q_n, then the *worst case error* of the approximation taken over the norm unit ball of $H(K_\gamma)$ is given by

$$e^{\mathrm{wor}}(Q_n, H(K_\gamma)) = \sup_{\|f\|_{H(K_\gamma)} \le 1} |\mathcal{I}(f) - Q_n(f)| = \|h_{\gamma,n}\|_{H(K_\gamma)}. \tag{16}$$

In the case of finite-dimensional integration of functions defined on $[0,1]^d$ whose mixed first partial derivatives are square integrable, the quantity $\|h_{\gamma,n}\|_{H(K_\gamma)}$ was called generalized L_2-discrepancy in [Hic98]. In the case of infinite-dimensional integration of functions defined on $[0,1]^{\mathbb{N}}$ it was simply called L_2-discrepancy in [HW01]. To distinguish it clearly from the weighted geometric L_2-discrepancy defined in (4), we prefer to call $e^{\mathrm{wor}}(Q_n, H(K_\gamma)) = \|h_{\gamma,n}\|_{H(K_\gamma)}$ the *weighted numerical L_2-discrepancy* of the linear algorithm Q_n (or of the corresponding multi-sets $\{t_1, \dots, t_n\}$ of sample points and $\{a_1, \dots, a_n\}$ of coefficients). As in the case of the weighted geometric L_2-discrepancy, we drop the attribute "weighted" if all weights γ_ν except of one are equal to zero.

We obtain

$$\begin{aligned} & e^{\mathrm{wor}}(Q_n, H(K_\gamma))^2 \\ =& \|h_\gamma\|^2_{H(K_\gamma)} - 2\sum_{j=1}^{n} a_j \langle h_\gamma, K_\gamma(\cdot, t_j)\rangle_{H(K_\gamma)} + \sum_{i,j=1}^{n} a_i a_j \langle K_\gamma(\cdot, t_i), K_\gamma(\cdot, t_j)\rangle_{H(K_\gamma)} \\ =& \int_M \int_M K_\gamma(x,y)\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) - 2\sum_{j=1}^{n} a_j \int_M K_\gamma(x, t_j)\,\mathrm{d}\mu(x) + \sum_{i,j=1}^{n} a_i a_j K_\gamma(t_i, t_j). \end{aligned}$$

Thus we have

$$\begin{aligned} & e^{\mathrm{wor}}(Q_n, H(K_\gamma))^2 \\ =& \sum_{\nu \in I} \gamma_\nu \left[\int_M \int_M K_\nu(x,y)\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) - 2\sum_{j=1}^{n} a_j \int_M K_\nu(x, t_j)\,\mathrm{d}\mu(x) + \sum_{i,j=1}^{n} a_i a_j K_\nu(t_i, t_j) \right], \end{aligned} \tag{17}$$

where in the case of infinite I the identity follows from (11) and (14).

Let us also define the *nth S-minimal worst case error* $e^{\mathrm{wor}}(n, S, H(K_\gamma))$ by

$$e^{\mathrm{wor}}(n, S, H(K_\gamma)) = \inf\{ e^{\mathrm{wor}}(Q_n, H(K_\gamma)) \,|\, Q_n \text{ as in (15)} \}.$$

5 Relation between Weighted Numerical Integration and Weighted Geometric L_2-Discrepancy

We are interested in the question when do weighted numerical L_2-discrepancy and weighted geometric L_2-discrepancy coincide, that is, under which conditions does the identity

$$e^{\mathrm{wor}}(Q_n, H(K_\gamma)) = \mathrm{disc}^{\mathcal{B}}_{2,\gamma}(\{t_j\}, \{a_j\}) \tag{18}$$

hold?

5.1 The General Case

Let us first assume we have

$$K_\nu(x,y) = \int_{\mathcal{B}_\nu} 1_{B_\nu}(\Phi_\nu(x)) 1_{B_\nu}(\Phi_\nu(y)) \,\mathrm{d}\omega_\nu(B_\nu) \quad \text{for all } x, y \in M \text{ and all } \nu \in I. \tag{19}$$

The function K_ν defined by (19) is measurable on $M \times M$ due to (1), the measurability of Φ_ν, and Tonelli's theorem. It is indeed a reproducing kernel, since it is obviously symmetric and also positive semi-definite: Let $n \in \mathbb{N}$, $x_1, \ldots, x_n \in M$, and $a_1, \ldots, a_n \in \mathbb{R}$. Then

$$\sum_{i,j=1}^{n} K_\nu(x_i, x_j) a_i a_j = \int_{\mathcal{B}_\nu} \left(\sum_{i=1}^{n} a_i 1_{B_\nu}(\Phi_\nu(x_i)) \right)^2 \mathrm{d}\omega_\nu(B_\nu) \geq 0.$$

We have to assume (6), which is now, e.g., satisfied if $\sum_{\nu \in I} \gamma_\nu < \infty$. Furthermore, we assume that $H(K_\gamma)$ consists of μ-integrable functions and that integration is a bounded linear functional on $H(K_\gamma)$, i.e., that (10) holds. Then, due to the fact that the K_νs are non-negative, condition (13) and (14) are also satisfied.

Under these assumptions (19) implies that identity (18) holds independently of the choice of the finite sequences $\{t_j\}$, $\{a_j\}$, and the admissible weights $\gamma = (\gamma_\nu)_{\nu \in I}$. Indeed, due to our assumptions $\mu_\nu = \mu \circ \Phi_\nu^{-1}$ and the measurability of χ_ν defined in (1), and to the theorem of Fubini and Tonelli,

$$\begin{aligned}
\int_M \int_M K_\nu(x,y) \,\mathrm{d}\mu(x) \,\mathrm{d}\mu(y) &= \int_M \int_M \int_{\mathcal{B}_\nu} 1_{B_\nu}(\Phi_\nu(x)) 1_{B_\nu}(\Phi_\nu(y)) \,\mathrm{d}\omega_\nu(B_\nu) \,\mathrm{d}\mu(x) \,\mathrm{d}\mu(y) \\
&= \int_{\mathcal{B}_\nu} \left(\int_M 1_{B_\nu}(\Phi_\nu(x)) \,\mathrm{d}\mu(x) \right)^2 \mathrm{d}\omega_\nu(B_\nu) \\
&= \int_{\mathcal{B}_\nu} \left(\int_{M_\nu} 1_{B_\nu}(\xi_\nu) \,\mathrm{d}\mu_\nu(\xi_\nu) \right)^2 \mathrm{d}\omega_\nu(B_\nu) \\
&= \int_{\mathcal{B}_\nu} \mu_\nu(B_\nu)^2 \,\mathrm{d}\omega_\nu(B_\nu).
\end{aligned}$$

Furthermore,

$$\begin{aligned}\int_M K_\nu(x,t_j)\,\mathrm{d}\mu(x) &= \int_{\mathcal{B}_\nu}\left(\int_M 1_{B_\nu}(\Phi_\nu(x))\,\mathrm{d}\mu(x)\right)1_{B_\nu}(\Phi_\nu(t_j))\,\mathrm{d}\omega_\nu(B_\nu)\\ &= \int_{\mathcal{B}_\nu}\mu_\nu(B_\nu)1_{B_\nu}(\Phi_\nu(t_j))\,\mathrm{d}\omega_\nu(B_\nu).\end{aligned}$$

Hence identity (18) follows from identity (5) and (17).

A comparison of (5) and (17) reveals that condition (19) is not only sufficient, but also necessary for (18) to hold for all choices of $\{t_j\}$, $\{a_j\}$, and γ. It is even necessary if we restrict ourselves to the case $n=2$, arbitrary $a_1,a_2>0$, $t_1,t_2\in M$, and admissible positive weights γ. This is easily verified by first varying the positive weights γ, which shows that for each $\nu\in I$ the corresponding summands in (5) and (17) have to be equal, and then, for fixed ν, t_1, and t_2, varying the coefficients a_1 and a_2.

Theorem 5.1. *Let $\gamma=(\gamma_\nu)_{\nu\in I}$ be a sequence of weights, and assume that (6) holds. Let K_γ be the reproducing kernel defined by equation (7). Furthermore, assume that $H(K_\gamma)$ consists of μ-integrable functions and that (10), (13), and (14) hold.*

If additionally condition (19) is satisfied, then the identity

$$e^{\mathrm{wor}}(Q_n,H(K_\gamma))=\mathrm{disc}^{\mathcal{B}}_{2,\gamma}(\{t_j\},\{a_j\}) \tag{20}$$

holds for all linear algorithms $Q_n(f)=\sum_{j=1}^n a_jf(t_j)$, $a_1,\dots,a_n\in\mathbb{R}$, $t_1,\dots,t_n\in S$. Consequently, we have

$$e^{\mathrm{wor}}(n,S,H(K_\gamma))=\mathrm{disc}^{\mathcal{B}}_{2,\gamma}(n,S).$$

Condition (19) is also necessary for (20) to hold for all choices of sample points $\{t_j\}$, coefficients $\{a_j\}$, and admissible weights γ.

Corollary 5.2. *Let the assumptions from Theorem 5.1 hold. If additionally (19) holds, we have the following* generalized Zaremba inequality

$$\left|\int_M f(x)\,\mathrm{d}\mu(x)-\sum_{i=1}^n a_if(t_i)\right|\le \mathrm{disc}^{\mathcal{B}}_{2,\gamma}(\{t_j\},\{a_i\})\|f\|_{H(K_\gamma)}$$

for all $f\in H(K_\gamma)$, $t_1,\dots,t_n\in S$, and $a_1,\dots,a_n\in\mathbb{R}$.

5.2 The Product Structure Case

Here we want to study a situation where condition (19) can be simplified reasonably. Let us assume that there exists a set $\widetilde{M}$, and a class $\widetilde{\mathcal{B}}$ of subsets of $\widetilde{M}$, endowed with a σ-algebra $\Sigma(\widetilde{\mathcal{B}})$ and a probability measure $\widetilde{\omega}$ such that the following holds:

Assumption 1. *For each $\nu\in I$ exists a number $n(\nu)\in\mathbb{N}$ such that*

(*i*) *M_ν is the $n(\nu)$-fold Cartesian product of $\widetilde{M}$, i.e., $M_\nu=\prod_{i=1}^{n(\nu)}\widetilde{M}$,* *(ii)* *each $B_\nu \in \mathcal{B}_\nu$ is an $n(\nu)$-fold Cartesian product of sets in $\widetilde{\mathcal{B}}$, i.e.,*

$$\mathcal{B}_\nu = \times_{i=1}^{n(\nu)} \widetilde{\mathcal{B}} := \left\{ \prod_{i=1}^{n(\nu)} B_i \,\middle|\, B_i \in \widetilde{\mathcal{B}} \right\},$$

(iii) *the σ-algebra $\Sigma(\mathcal{B}_\nu)$ on $\mathcal{B}_\nu$ is the $n(\nu)$-fold product σ-algebra of $\Sigma(\widetilde{\mathcal{B}})$, i.e., $\Sigma(\mathcal{B}_\nu) = \otimes_{i=1}^{n(\nu)} \Sigma(\widetilde{\mathcal{B}})$,*

(iv) *the measure ω_ν on $\Sigma(\mathcal{B}_\nu)$ is the $n(\nu)$-fold product measure of $\widetilde{\omega}$, i.e., $\omega_\nu = \otimes_{i=1}^{n(\nu)} \widetilde{\omega}$.*

(Formally, the product σ-algebra $\otimes_{i=1}^{n(\nu)} \Sigma(\widetilde{B})$ is defined on the $n(\nu)$-fold Cartesian product $\prod_{i=1}^{n(\nu)} \widetilde{\mathcal{B}}$, but as a measure space we simply identify $\times_{i=1}^{n(\nu)} \widetilde{\mathcal{B}}$ with $\prod_{i=1}^{n(\nu)} \widetilde{\mathcal{B}}$. As long as, e.g., $\emptyset \notin \widetilde{\mathcal{B}}$, we have the canonical bijection $\prod_{i=1}^{n(\nu)} \widetilde{\mathcal{B}} \to \times_{i=1}^{n(\nu)} \widetilde{\mathcal{B}}$, $(B_1, \ldots, B_{n(\nu)}) \mapsto \prod_{i=1}^{n(\nu)} B_i$; note that the empty set is irrelevant for discrepancy questions, since it always leads to the trivial local discrepancy zero.) For $j = 1, \ldots, n(\nu)$ let $\Phi_{\nu,j} : M \to \widetilde{M}$ denote the jth component function of Φ_ν, that is $\Phi_\nu = (\Phi_{\nu,1}, \ldots, \Phi_{\nu,n(\nu)})$. Furthermore, Assumption 1 and (1) ensure that $B \mapsto 1_B(r)$ is a measurable map on $\widetilde{\mathcal{B}}$ for all $r \in \widetilde{M}$.

Under Assumption 1 condition (19) reads

$$K_\nu(x,y) = \prod_{i=1}^{n(\nu)} \int_{\widetilde{\mathcal{B}}} 1_B(\Phi_{\nu,i}(x)) 1_B(\Phi_{\nu,i}(y)) \, \mathrm{d}\widetilde{\omega}(B) \quad \text{for all } x, y \in M \text{ and all } \nu \in I.$$

Thus, defining the reproducing kernel $\widetilde{K}$ on $\widetilde{M} \times \widetilde{M}$ by

$$\widetilde{K}(r,s) = \int_{\widetilde{\mathcal{B}}} 1_B(r) 1_B(s) \, \mathrm{d}\widetilde{\omega}(B) \quad \text{for all } r, s \in \widetilde{M}, \tag{21}$$

we get

$$K_\nu(x,y) = \prod_{i=1}^{n(\nu)} \widetilde{K}(\Phi_{\nu,i}(x), \Phi_{\nu,i}(y)) \quad \text{for all } x, y \in M \text{ and all } \nu \in I. \tag{22}$$

On the other hand, it is easily seen that under the assumption that (22) holds for some function $\widetilde{K} : \widetilde{M} \times \widetilde{M} \to \mathbb{R}$, the conditions (19) and (21) are equivalent (apart from the fact that in the case where all $n(\nu)$ are even, we have the additional freedom to multiply $\widetilde{K}$ in (21) by a factor -1).

Note that (22) implies that the reproducing kernel Hilbert space $H(\widetilde{K}_\nu)$ is of tensor product structure. More precisely, we have that $H(\widetilde{K}_\nu)$ is equal to $\otimes_{i=1}^{n(\nu)} H(\widetilde{K})$, the complete $n(\nu)$-fold tensor product Hilbert space of $H(\widetilde{K})$, see, e.g., [Aro50, Sect. I.8].

Theorem 5.3. *Let the assumptions of Theorem 5.1 hold, and let Assumption 1 be satisfied.*

(*i*) *Condition (19) implies for all* $\nu \in I$ *that the reproducing kernel* K_ν *is of product structure (22) with* $\widetilde{K}$ *as in (21), and the reproducing kernel Hilbert space* $H(\widetilde{K}_\nu)$ *is the complete* $n(\nu)$*-fold tensor product Hilbert space of* $H(\widetilde{K})$*.*

(*ii*) *Let condition (22) hold. Then condition (19) is equivalent to condition (21). In particular, condition (21) is sufficient and necessary to ensure for all linear algorithms* $Q_n(f) = \sum_{j=1}^n a_j f(t_j)$, $a_1, \ldots, a_n \in \mathbb{R}$, $t_1, \ldots, t_n \in S$, *and all admissible weights* γ *that*

$$e^{\mathrm{wor}}(Q_n, H(K_\gamma)) = \mathrm{disc}^{\mathcal{B}}_{2,\gamma}(\{t_j\}, \{a_j\}).$$

(If all ν *are even, this holds only modulo the restriction that we have the additional freedom to multiply* $\widetilde{K}$ *in (21) by* -1*.)*

Notice that for Theorem 5.3 it is completely irrelevant whether the measure μ on M, or the measures μ_ν on M_ν, $\nu \in I$, have product structure, see also the example given in Subsection 7.2.

6 An Upper Bound for the Integration Error

Let us assume that condition (19) holds. Furthermore, we assume that (M, Σ, μ) is a finite measure space, i.e., $\mu(M) < \infty$, and that $\sum_{\nu \in I} \gamma_\nu < \infty$. The set $S \subseteq M$ of admissible sample points should be measurable.

If additionally $\mu(M \backslash S) = 0$, then we can prove an upper bound on $e^{\mathrm{wor}}(n, S, H(K_\gamma))$ by averaging over all properly normalized quasi-Monte Carlo algorithms that use admissible sample points. Now, in some applications, we may not have $\mu(M \setminus S) = 0$. Actually, in infinite-dimensional integration under realistic assumptions we have rather $\mu(S) = 0$, see the example in Subsection 7.4. That is why we require the following weaker conditions:

There exists a sequence $(\nu_m)_{m \in \mathbb{N}}$ in I which satisfies

$$\mu_{\nu_m}(M_{\nu_m} \setminus \Phi_{\nu_m}(S)) = 0 \quad \text{for all } m \in \mathbb{N}, \tag{23}$$

and additionally, we find for all $\nu \in I$ an $m_0 \in \mathbb{N}$ such that for all $m \geq m_0$ there exists a measurable map

$$\Psi_{m,\nu} : M_{\nu_m} \to M_\nu \quad \text{with} \quad \Psi_{m,\nu} \circ \Phi_{\nu_m} = \Phi_\nu. \tag{24}$$

(Indeed, these conditions hold if $\mu(M \setminus S) = 0$, since we may formally extend I by some index $\kappa \notin I$, define $(M_\kappa, \Sigma_\kappa, \mu_\kappa) := (M, \Sigma, \mu)$ and put $\gamma_\kappa := 0$ and $\nu_m := \kappa$ for all $m \in \mathbb{N}$, and $\Psi_{m,\nu} := \Phi_\nu$ and $\Phi_{\nu_m} := \mathrm{Id}_M$ for all $m, \nu \in \mathbb{N}$.)

If for ν_m and $\nu \in I$ condition (24) holds, we write $\nu \preceq \nu_m$. Note that this relation implies $\mu_\nu = \mu_{\nu_m} \circ \Psi^{-1}_{m,\nu}$, i.e., μ_ν is the direct image of μ_{ν_m} under $\Psi_{m,\nu}$. Recall that (19) implies $K_\nu(x,y) = \widetilde{K}_\nu(\Phi_\nu(x), \Phi_\nu(y)) \in [0,1]$ for all $x, y \in M$.

From (17) we get for all $m \in \mathbb{N}$ and all linear algorithms of the form

$$Q_n(f) = \frac{\mu(M)}{n} \sum_{j=1}^n f(t_j)\,, \quad t_1, \ldots, t_n \in S\,, \tag{25}$$

the estimate

$$e^{\mathrm{wor}}(Q_n, H(K_\gamma))^2 \le f_m(\Phi_{\nu_m}(t_1), \dots, \Phi_{\nu_m}(t_n)) + 2\mu(M)^2 \sum_{\nu \not\preceq \nu_m} \gamma_\nu \,,$$

where

$$\begin{aligned} f_m(\tau_1, \dots, \tau_n) = \sum_{\nu \preceq \nu_m} \gamma_\nu \Big[\int_{M_\nu} \int_{M_\nu} \widetilde{K}_\nu(x_\nu, y_\nu) \, \mathrm{d}\mu_\nu(x_\nu) \, \mathrm{d}\mu_\nu(y_\nu) \\ - \frac{2\mu(M)}{n} \sum_{j=1}^n \int_{M_\nu} \widetilde{K}_\nu(x_\nu, \Psi_{m,\nu}(\tau_j)) \, \mathrm{d}\mu_\nu(x_\nu) + \frac{\mu(M)^2}{n^2} \sum_{i,j=1}^n \widetilde{K}_\nu(\Psi_{m,\nu}(\tau_i), \Psi_{m,\nu}(\tau_j)) \Big] \end{aligned}$$

for $\tau_1, \dots, \tau_n \in M_{\nu_m}$. For any m we can average for fixed n over $f_m(\tau_1, \dots, \tau_n)$, $\tau_1, \dots, \tau_n \in \Phi_{\nu_m}(S)$. Due to (23) we get

$$\begin{aligned} &\frac{1}{\mu_{\nu_m}(\Phi_{\nu_m}(S))^n} \int_{(\Phi_{\nu_m}(S))^n} f_m(\tau_1, \dots, \tau_n) \, \mathrm{d}\mu_{\nu_m}(\tau_1) \dots \, \mathrm{d}\mu_{\nu_m}(\tau_n) \\ =& \frac{1}{\mu_{\nu_m}(M_{\nu_m})^n} \int_{M_{\nu_m}^n} f_m(\tau_1, \dots, \tau_n) \, \mathrm{d}\mu_{\nu_m}(\tau_1) \dots \, \mathrm{d}\mu_{\nu_m}(\tau_n) \\ =& \sum_{\nu \preceq \nu_m} \gamma_\nu \Big[\int_{M_\nu} \int_{M_\nu} \widetilde{K}_\nu(x_\nu, y_\nu) \, \mathrm{d}\mu_\nu(x_\nu) \, \mathrm{d}\mu_\nu(y_\nu) \\ &- \frac{2}{n} \sum_{j=1}^n \int_{M_{\nu_m}} \int_{M_\nu} \widetilde{K}_\nu(x_\nu, \Psi_{m,\nu}(\tau_j)) \, \mathrm{d}\mu_\nu(x_\nu) \, \mathrm{d}\mu_{\nu_m}(\tau_j) \\ &+ \frac{\mu(M)}{n^2} \sum_{i=1}^n \int_{M_{\nu_m}} \widetilde{K}_\nu(\Psi_{m,\nu}(\tau_i), \Psi_{m,\nu}(\tau_i)) \, \mathrm{d}\mu_{\nu_m}(\tau_i) \\ &+ \frac{1}{n^2} \sum_{i \neq j}^n \int_{M_{\nu_m}} \int_{M_{\nu_m}} \widetilde{K}_\nu(\Psi_{m,\nu}(\tau_i), \Psi_{m,\nu}(\tau_j)) \, \mathrm{d}\mu_{\nu_m}(\tau_i) \, \mathrm{d}\mu_{\nu_m}(\tau_j) \Big] \\ =& \frac{1}{n} \sum_{\nu \preceq \nu_m} \gamma_\nu \left[\mu(M) \int_{M_\nu} \widetilde{K}_\nu(x_\nu, x_\nu) \, \mathrm{d}\mu_\nu(x_\nu) - \int_{M_\nu} \int_{M_\nu} \widetilde{K}_\nu(x_\nu, y_\nu) \, \mathrm{d}\mu_\nu(x_\nu) \, \mathrm{d}\mu_\nu(y_\nu) \right]. \end{aligned}$$

Due to (19) we have

$$\int_{M_\nu} \widetilde{K}_\nu(x_\nu, x_\nu) \, \mathrm{d}\mu_\nu(x_\nu) \le \mu_\nu(M_\nu) = \mu(M).$$

For given $n \in \mathbb{N}$ we may choose $m = m(n) \in \mathbb{N}$ such that

$$2\mu(M)^2 \sum_{\nu \not\preceq \nu_m} \gamma_\nu \le \frac{1}{n} \sum_{\nu \preceq \nu_m} \gamma_\nu \int_{M_\nu} \int_{M_\nu} \widetilde{K}_\nu(x_\nu, y_\nu) \, \mathrm{d}\mu_\nu(x_\nu) \, \mathrm{d}\mu_\nu(y_\nu).$$

(Recall that the sum on the right hand side converges to $\|h_\gamma\|^2_{H(K_\gamma)} > 0$ for $m \to \infty$, see (12), (14) and the following comment. Furthermore, we assumed that the weights $(\gamma_\nu)_{\nu \in I}$ are summable.)

From this follows that there exists at least one normalized quasi-Monte Carlo algorithm Q_n that uses n admissible sample points with

$$e^{\mathrm{wor}}(Q_n, H(K_\gamma)) \le \frac{\mu(M)\sqrt{\sum_{\nu\in I}\gamma_\nu}}{\sqrt{n}}.$$

Altogether we have proved the following theorem.

Theorem 6.1. *Assume that $\sum_{\nu\in I}\gamma_\nu < \infty$, $\mu(M) < \infty$, and that the set S of admissible sample points is a measurable subset of M. Assume that (19) holds and let the weighted reproducing kernel K_γ be defined by equation (7). Assume furthermore that (10) holds.*

If $\mu(M \setminus S) = 0$ or if the weaker conditions (23) and (24) hold, then there exists a normalized quasi-Monte Carlo algorithm Q_n as in (25) such that

$$e^{\mathrm{wor}}(n, S, H(K_\gamma)) \le e^{\mathrm{wor}}(Q_n, H(K_\gamma)) \le \frac{\mu(M)\sqrt{\sum_{\nu\in I}\gamma_\nu}}{\sqrt{n}}, \tag{26}$$

or equivalently, there exists points $t_1, \ldots, t_n \in S$ and coefficients $a_1 = \ldots = a_n = \mu(M)/n$ such that

$$\mathrm{disc}^{\mathcal{B}}_{2,\gamma}(n, S) \le \mathrm{disc}^{\mathcal{B}}_{2,\gamma}(\{t_j\}, \{a_j\}) \le \frac{\mu(M)\sqrt{\sum_{\nu\in I}\gamma_\nu}}{\sqrt{n}}.$$

Remark 6.2. In Theorem 6.1 we actually did not need condition (19) to prove the estimate (26), but only the weaker condition that K_ν takes only values in $[0, 1]$ for all $\nu \in I$. In general, it is sufficient to get a (properly scaled) version of estimate (26) if all the K_νs are non-negative and uniformly bounded.

7 Examples

Here we want to discuss some special cases of the quite general notion of weighted geometric L_2-discrepancy from Section 3 and relate them to numerical integration on corresponding reproducing kernel Hilbert spaces.

7.1 L_2-B-Discrepancy

We start with the *L_2-B-discrepancy* as defined in [NW09], see also [NW10]. This discrepancy fits in our more general definition if we make the following choices: Let M be a measurable subset of $\mathbb{R}^d$, Σ the Borel σ-algebra, and μ the d-dimensional Lebesgue measure restricted to M. Furthermore, let $I = \{1\}$, $\gamma_1 = 1$, and let $\Phi_1 : M \to M$ be the identity mapping. Let $\mathcal{B}_1 = \mathcal{B}$ be a class of measurable subsets of M with $\cup_{B\in\mathcal{B}} B = M$. For a given positive integer $\tau(d)$ let $D \subseteq \mathbb{R}^{\tau(d)}$ be Borel measurable and $\rho : D \to [0, \infty)$ a probability density.

Let $B : D \to \mathcal{B}$, $x \mapsto B(x)$ be a parametrization such that the mapping $(t, x) \mapsto 1_{B(x)}(t)$ is measurable on $M \times D$ with respect to the product σ-algebra. (The last important measurability condition was actually forgotten in [NW09], but is added in the more recent and more comprehensive exposition in [NW10, Chapter 9].)

Formally, we endow $\mathcal{B}$ with the σ-algebra

$$\Sigma(\mathcal{B}) = \{A \subseteq \mathcal{B} \,|\, B^{-1}(A) \quad \text{Borel measurable}\}.$$

Let the probability measure ω on $\mathcal{B}$ be induced by the probability measure $\rho(x)\,\mathrm{d}x$, where $\mathrm{d}x$ is the $\tau(d)$-dimensional Lebesgue measure, that is,

$$\omega(A) = \int_{B^{-1}(A)} \rho(x)\,\mathrm{d}x \quad \text{for all } A \in \Sigma(\mathcal{B}).$$

For these special choices the weighted geometric L_2-discrepancy defined in (4) is nothing but the L_2-B-discrepancy

$$\begin{aligned}\operatorname{disc}_2^B(\{t_i\},\{a_j\}) &= \left(\int_{\mathcal{B}} \left(\operatorname{vol}(A) - \sum_{j=1}^n a_j 1_A(t_j)\right)^2 \mathrm{d}\omega(A)\right)^{1/2} \\ &= \left(\int_D \left(\operatorname{vol}(B(x)) - \sum_{j=1}^n a_j 1_{B(x)}(t_j)\right)^2 \rho(x)\,\mathrm{d}x\right)^{1/2}\end{aligned}$$

defined in [NW09]. In this situation Theorem 5.1 and Theorem 6.1 (under the additional assumption $S = M$) were already proved in [NW09]. If K_d^B denotes the reproducing kernel corresponding to disc_2^B, then condition (19) becomes

$$K_d^B(y,z) = \int_D 1_{B(x)}(y) 1_{B(x)}(z)\rho(x)\,\mathrm{d}x \quad \text{for all } y,z \in M.$$

More concrete examples for L_2-B-discrepancies as, e.g., the *centered discrepancy* [Hic98], the *quadrant discrepancy* [HSW04, NW09], the *extreme discrepancy* [MC94] or the *periodic ball discrepancy* [CT09] are discussed in [NW09, NW10]. That is why we confine ourselves in the rest of this section to present examples of (weighted) geometric L_2-discrepancies which are not covered by the notion of L_2-B-discrepancy.

7.2 G-Discrepancy

The d-dimensional *L_2-G-discrepancy* or *L_2-G-star discrepancy* is defined as the L_2-B-discrepancy in the special case where $M = D = [0,1]^d$, the mapping B is given by $B(x) = [0,x)$ (where $[0,x) = [0,x_1) \times \cdots \times [0,x_d)$ for a vector $x = (x_1,\ldots,x_d)$), and $\rho \equiv 1$, except that $\mu = \mu_G$ is in general not the d-dimensional Lebesgue measure, but some probability measure given by a distribution function G via $\mu([0,x)) = G(x)$ for all $x \in [0,1]^d$. Thus

$$\operatorname{disc}_2^G(\{t_i\},\{a_j\}) = \left(\int_{[0,1]^d} \left(G(x) - \sum_{j=1}^n a_j 1_{[0,x)}(t_j)\right)^2 \mathrm{d}x\right)^{1/2}.$$

The reproducing kernel K_d^G of the corresponding Hilbert space of d-variate functions is given by

$$K_d^G(y,z) = \int_{[0,1]^d} 1_{[0,x)}(y) 1_{[0,x)}(z)\,\mathrm{d}x = \prod_{j=1}^d \int_0^1 1_{[0,\xi)}(y_j) 1_{[0,\xi)}(z_j)\,\mathrm{d}\xi = \prod_{j=1}^d (1-\max\{y_j,z_j\})$$

and does actually not depend on G. Using the short hand $\widetilde{K}(\xi,\eta) = 1-\max\{\xi,\eta\}$, we see that

$$K_d^G(y,z) = \prod_{j=1}^d \widetilde{K}(y_j,z_j),$$

i.e., condition (22) is satisfied (and condition (21), too).

K_d^G is the kernel of the Sobolev space anchored in 1, which is, e.g., described in [NW09, NW10].

This example underlines that the choice of the measure $\mu = \mu_G$ on M effects the form of the discrepancy disc_2^G, but not the kernel K_d^G or the corresponding reproducing kernel Hilbert space $H(K_d^G)$ (but obviously the integration problem $\mathcal{I}(f) = \int_M f(x)\,\mathrm{d}\mu_G(x)$ we want to solve).

Seemingly, the L_2-G-discrepancy has not been studied so far, in contrast to the (L_∞)-G- or G-star discrepancy

$$\mathrm{disc}_\infty^G(\{t_i\},\{a_j\}) = \sup_{x\in[0,1]^d} \left| G(x) - \sum_{i=1}^n a_i 1_{[0,x)}(t_i) \right| ,$$

which has applications in quasi-Monte Carlo importance sampling, see, e.g., [Ökt99]. Further results on the G-star discrepancy can, e.g., be found in [GR09].

7.3 Weighted L_2-Star Discrepancy

Let $d \in \mathbb{N}$, and denote the set $\{1,\dots,d\}$ by $[d]$. For a family of weights $\gamma = \{\gamma_u\}_{u\subseteq[d]}$ the *weighted L_2-star discrepancy* of a multi-set $\{t_1,\dots,t_n\}$ in $[0,1]^d$ and coefficients $a_1,\dots,a_n$ in $\mathbb{R}$ is defined as

$$\mathrm{disc}_{2,\gamma}^*(\{t_i\},\{a_j\}) = \left(\sum_{u\subseteq[d]} \gamma_u \int_{[0,1]^{|u|}} \left(\prod_{j\in u} x_j - \sum_{k=1}^n a_j \prod_{j\in u} 1_{[0,x_j)}(t_{k,j}) \right)^2 \mathrm{d}x_u \right)^{1/2} .$$

To get from our definition of the weighted geometric L_2-discrepancy the special case of the weighted L_2-star discrepancy (which is sometimes also called *weighted L_2-discrepancy anchored at* 0), we just have to make the following choices:

Let $M = [0,1]^d$, Σ the Borel σ-algebra on $[0,1]^d$, and μ the restriction of the d-dimensional Lebesgue measure to $[0,1]^d$. Let $I = \{u \,|\, u \subseteq [d]\}$. Let $M_u = [0,1]^{|u|}$, where $|u|$ denotes the cardinality of the set u, and let Σ_u be the Borel σ-algebra on $[0,1]^{|u|}$. Furthermore, let

$$\Phi_u : [0,1]^d \to [0,1]^{|u|}\,,\; x = (x_i)_{i=1}^d \mapsto (x_\nu)_{\nu\in u}.$$

Then $\mu_u = \mu \circ \Phi_u^{-1}$ is nothing but the restriction of the $|u|$-dimensional Lebesgue measure to $[0,1]^{|u|}$. Furthermore, let $\mathcal{B}_u = \{[0,\xi_u) \,|\, \xi_u \in [0,1]^{|u|}\}$. As a measure space we identify $(\mathcal{B}_u, \Sigma(\mathcal{B}_u), \omega_u)$ via the mapping $\iota : [0,1]^{|u|} \to \mathcal{B}_u,\ \xi_u \mapsto [0,\xi_u)$ with the measure space (M_u, Σ_u, μ_u). (Note that for $|u| > 1$ the map ι is not injective, since $\iota(\xi) = \emptyset$ for all $\xi \in \{y \in [0,1]^{|u|} \,|\, \exists i : y_i = 0\}$; but this is irrelevant for our purpose, since the latter set has zero $|u|$-dimensional Lebesgue measure.)

Clearly, for each $u \subseteq [d]$ the function

$$\chi_u : [0,1]^{2|u|} \to \{0,1\}\,,\ (x_u, y_u) \mapsto 1_{[0,y_u)}(x_u)$$

is measurable, and we have

$$\int_{[0,1]^{|u|}} \mu_u([0,y_u))^2 \,\mathrm{d}\mu_u(y_u) = 3^{-|u|} < \infty.$$

Condition (19) reads now as follows:

$$\begin{aligned} K_u(x,y) &= \int_{\mathcal{B}_u} 1_{B_u}(\Phi_u(x)) 1_{B_u}(\Phi_u(y)) \,\mathrm{d}\omega_u(B_u) \\ &= \int_{[0,1]^{|u|}} 1_{[0,\xi_u)}(\Phi_u(x)) 1_{[0,\xi_u)}(\Phi_u(y)) \,\mathrm{d}\xi_u \\ &= \prod_{j\in u} \int_0^1 1_{[0,\xi)}(x_j) 1_{[0,\xi)}(y_j) \,\mathrm{d}\xi \\ &= \prod_{j\in u} (1 - \max\{x_j, y_j\}). \end{aligned}$$

This leads us to the weighted reproducing kernel

$$K_\gamma(x,y) = \sum_{u\subseteq[d]} \gamma_u K_u(x,y) = \sum_{u\subseteq[d]} \gamma_u \prod_{j\in u} (1 - \max\{x_j, y_j\}).$$

The resulting Hilbert space is the weighted Sobolev space with mixed partial derivatives of order 1 anchored at 1, and is, e.g., discussed in detail in [NW09, NW10]. In that situation identity (20) and Theorem 6.1, under the assumption $S = M$, were proved in [SW98] for product weights. For general weights the corresponding results can be found in [NW09].

Due to the product structure of the sets $M_u = [0,1]^{|u|}$, of the classes of test sets

$$\mathcal{B}_u = \left\{ \prod_{j\in u} [0, x_j) \,\middle|\, \forall j \in u :\ x_j \in [0,1] \right\},$$

of the σ-algebras Σ_u, of the measures $\omega_u = \mathrm{d}\xi_u = \otimes_{j\in u} \mathrm{d}\xi$, and of the kernels

$$K_u(x,y) = \prod_{j\in u} \widetilde{K}(x_j, y_j), \quad \text{with} \quad \widetilde{K}(\xi,\eta) = 1 - \max\{\xi, \eta\},$$

condition (19) is equivalent to

$$\widetilde{K}(r,s) = \int_0^1 1_{[0,t)}(r) 1_{[0,t)}(s)\,\mathrm{d}t \quad \forall r,s \in [0,1], \tag{27}$$

as described in Theorem 5.3.

7.4 Infinite-Dimensional Integration and Limiting Discrepancy

Quite recently, there have been several papers on deterministic infinite-dimensional numerical integration on weighted reproducing or quasi-reproducing Hilbert spaces, see [KSWW10, NHMR10, Gne10, PW10]. An earlier paper dealing with infinite-dimensional integration and discrepancy is [HW01]. We want to discuss the setting studied in these papers.

Let $I = \{u \subset \mathbb{N} \,|\, |u| < \infty\}$. We consider here the setting described in [KSWW10] in Sect. 5 "Generalization":

Assume that there exists a Borel measurable set $\widetilde{M} \subseteq \mathbb{R}$, a point $a \in \widetilde{M}$, and a reproducing kernel $\widetilde{K} : \widetilde{M} \times \widetilde{M} \to \mathbb{R}$ with $\widetilde{K}(a,a) = 0$. The last condition implies $f(a) = 0$ for all $f \in H(\widetilde{K})$. Assume further that the corresponding Hilbert space $H(\widetilde{K})$ is separable and define

$$\widetilde{K}_u(x_u, y_u) = \prod_{j \in u} \widetilde{K}(x_j, y_j) \quad \text{for } u \in I \text{ and } x_u, y_u \in M_u = \widetilde{M}^{|u|}. \tag{28}$$

Each $f_u \in H(\widetilde{K}_u)$ is a function defined on $\widetilde{M}^{|u|}$ which satisfies $f_u(x_u) = 0$ if at least one component of x_u is a. With

$$\Phi_u : M = \widetilde{M}^{\mathbb{N}} \to M_u = \widetilde{M}^{|u|}\,,\ (x_j)_{j\in\mathbb{N}} \mapsto (x_j)_{j \in u},$$

let us write $K_u(x,y) = \widetilde{K}_u(\Phi_u(x), \Phi_u(y))$ for all $x, y \in \widetilde{M}^{\mathbb{N}}$. We define

$$\mathcal{H}_\gamma = \left\{ \sum_{u \in I} f_u \,\middle|\, f_u \in H(K_u)\,,\ \sum_{u \in I} \gamma_u^{-1} \|f_u\|^2_{H(K_u)} < \infty \right\}$$

for a sequence of weights $\gamma = (\gamma_u)_{u \in I}$. Under the assumption (6) $\mathcal{H}_\gamma$ is a reproducing kernel Hilbert space with norm

$$\|f\|_{H(K_\gamma)} := \left(\sum_{u \in I} \gamma_u^{-1} \|f_u\|^2_{H(K_u)} \right)^{1/2} \quad \text{if} \quad f = \sum_{u \in I} f_u \quad \text{with } f_u \in H(K_u),$$

and reproducing kernel K_γ defined by (7), i.e., $\mathcal{H}_\gamma = H(K_\gamma)$. Then $\mathcal{H}_\gamma = \oplus_{u \in I} H(K_u)$ with orthogonal spaces $H(K_u)$.

Let now ρ be a probability density on $\widetilde{M}$ and $\widetilde{\mu}(s) = \rho(s)\,\mathrm{d}s$. Let μ be the infinite-dimensional product probability measure $\otimes_{n \in \mathbb{N}} \widetilde{\mu}$.

As in Section 4, we consider the integral

$$\mathcal{I} = \int_M f(x)\,\mathrm{d}\mu(x).$$

By requiring that

$$A_0 = \left(\int_{\widetilde{M}}\int_{\widetilde{M}} \widetilde{K}(r,s)\rho(r)\rho(s)\,\mathrm{d}r\,\mathrm{d}s\right)^{1/2} < \infty$$

and

$$A_{0,\gamma} = \left(\sum_{u\in I}\gamma_u A_0^{2|u|}\right)^{1/2} < \infty\,,$$

it was ensured in [KSWW10] that (10) holds, i.e., that $\mathcal{I}$ is a bounded linear functional on $\mathcal{H}_\gamma = H(K_\gamma)$, and its operator norm is given by $A_{0,\gamma}$, see (12).

The set of admissible sample points is given by

$$S = \{x \in M \,|\, x_j = a \text{ for all but finitely many } j \in \mathbb{N}\}.$$

Notice that S is actually a set of measure zero, i.e., $\mu(S) = 0$. But, with $u_d = [d]$, we have that the sequence $(u_d)_{d\in\mathbb{N}}$ satisfies the conditions (23) and (24) if we choose for all $u \subseteq u_d$

$$\Psi_{d,u} : M_{u_d} \to M_u\,,\ (x_j)_{j\in[d]} \mapsto (x_j)_{j\in u},$$

hence in this setting the relation "$\preceq$" is the inclusion relation.

If there exists a set system $\widetilde{\mathcal{B}}$ of measurable subsets of $\widetilde{M}$, a σ-algebra $\Sigma(\widetilde{\mathcal{B}})$ on $\widetilde{\mathcal{B}}$, and a probability measure $\widetilde{\omega}$ on $(\widetilde{\mathcal{B}}, \Sigma(\widetilde{\mathcal{B}}))$ such that condition (21) holds, then, due to Theorem 5.3 and Theorem 6.1, we know that for any $n \in \mathbb{N}$ there exists a normalized quasi-Monte Carlo algorithm Q_n of the form (25) such that

$$e^{\mathrm{wor}}(Q_n, H(K_\gamma)) \le \frac{\sqrt{\sum_{\nu\in I}\gamma_\nu}}{\sqrt{n}}. \tag{29}$$

A similar estimate was proved in [HW01] in the case where $M = [0,1]^{\mathbb{N}}$. There the assumption that condition (21) holds was weakened to $\sup_{r,s\in\widetilde{M}} |\widetilde{K}(r,s)| < \infty$, see also Remark 6.2. But the authors assumed that the set of admissible sample points is the whole set M. In that case we are allowed to use sample points with infinitely many components different from the nominal value a, an assumption which will in practice usually not be realizable.

Let us have a look at the special case where $\widetilde{M} = [0,1]$ and $\widetilde{K}(r,s) = \prod_{j\in u}(1 - \max\{r,s\})$. From the previous subsection we know that

$$\widetilde{K}(r,s) = \int_0^1 1_{[0,x)}(r)1_{[0,x)}(s)\,\mathrm{d}x,$$

i.e., condition (21) holds. Formally, the discrepancy corresponding to the integration problem is the weighted L_2-star discrepancy with $d = \infty$, i.e.,

$$\mathrm{disc}^*_{2,\gamma}(\{t_j\},\{a_j\}) = \left(\sum_{u \in I} \gamma_u \int_{[0,1]^{|u|}} \left(\prod_{j \in u} x_j - \sum_{k=1}^{n} a_j \prod_{j \in u} 1_{[0,x_j)}(t_{k,j}) \right)^2 \mathrm{d}x \right)^{1/2} .$$

In the case of product weights this discrepancy was baptized "limiting discrepancy" in [SW98]. Here, we have the estimate (29).

For further bounds on the worst-case error of infinite-dimensional integration we refer the reader to the articles [Gne10, HW01, KSWW10, NHMR10, PW10] and the literature mentioned therein.

Acknowledgment

A major part of this work was done while the author was a research fellow at the Department of Computer Science of Columbia University in the City of New York. He thanks Henryk Woźniakowski for motivating and helpful discussions during that time.

The author also would like to thank Stefan Heinrich and Klaus Ritter for lively discussions during and after the talk he gave on this topic at the Technical University Kaiserslautern.

Furthermore, he acknowledges support from the German Science Foundation DFG under grant GN 91/3-1 and GN 91/4-1.

References

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

[CT09] W. W. L. Chen, G. Travaglini. Deterministic and probabilistic discrepancies. Ark. Mat. 47 (2009), 273–293.

[DP10] J. Dick, F. Pillichshammer. *Digital Nets and Sequences*. Cambridge University Press, Cambridge, 2010.

[Gne10] M. Gnewuch. *Infinite-dimensional integration on weighted Hilbert spaces*, Preprint cucs-016-10, Department of Computer Science, Columbia University in the City of New York, 2010.

[GR09] M. Gnewuch, A. V. Roşca. *On G-discrepancy and mixed Monte Carlo and quasi-Monte Carlo Sequences*. Acta Univ. Apulensis Math. Inform. 18 (2009).

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

[HSW04] F. J. Hickernell, I. H. Sloan, G. W. Wasilkowski. On tractability of weighted integration over bounded and unbounded regions in $\mathbb{R}^s$. Math. Comput. 73 (2004), 1885–1902.

[HW01] F. J. Hickernell, X. Wang. *The error bounds and tractability of quasi-Monte Carlo algorithms in infinite dimensions*. Math. Comp. 71 (2001), 1641–1661.

[KSWW10] F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, H. Woźniakowski. *Liberating the dimension.* J. Complexity (2010).

[Mat10] J. Matoušek. Geometric Discrepancy. Springer-Verlag Berlin Heidelberg, revised second printing, 2010.

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

[NHMR10] B. Niu, F. J. Hickernell, T. Müller-Gronbach, Klaus Ritter. *Deterministic multi-level algorithms for infinite-dimensional integration on* $\mathbb{R}^{\mathbb{N}}$. J. Complexity (2010).

[NW01a] E. Novak, H. Woźniakowski. *Intractability results for integration and discrepancy.* J. Complexity 17 (2001), 388–441.

[NW01b] E. Novak, H. Woźniakowski. *When are integration and discrepancy tractable?* In: Foundations of Computational Mathematics, R. A. DeVore, A. Iserles, E. Süli (Eds.), Cambridge University Press, 2001, 211–266.

[NW09] E. Novak, H. Woźniakowski. *L_2 discrepancy and multivariate integration.* In: Analytic number theory. Essays in honour of Klaus Roth. W. W. L. Chen, W. T. Gowers, H. Halberstam, W. M. Schmidt, and R. C. Vaughan (Eds.), Cambridge Univ. Press, Cambridge, 2009, 359–388,.

[NW10] E. Novak, H. Woźniakowski. *Tractability of multivariate problems. Vol. 2: Standard information.* EMS Tracts in Mathematics, European Mathematical Society (EMS), to appear.

[Ökt99] G. Ökten. *Error reduction techniques in quasi-Monte Carlo integration.* Math. Comput. Modelling 30 (1999), 61–69.

[PW10] L. Plaskota, G. W. Wasilkowski. *Tractability of infinite-dimensional integration in the worst case and randomized setting.* Preprint 2010.

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

[Woź91] H. Woźniakowski. *Average case complexity of multivariate integration.* Bull. Amer. Math. Soc. (N. S.) 24 (1991), 185–191.

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