A UNIVERSAL MEDIAN QUASI-MONTE CARLO INTEGRATION[*]

TAKASHI GODA[†], KOSUKE SUZUKI[‡], AND MAKOTO MATSUMOTO[§]

Abstract. We study quasi-Monte Carlo (QMC) integration over the multi-dimensional unit cube in several weighted function spaces with different smoothness classes. We consider approximating the integral of a function by the median of several integral estimates under independent and random choices of the underlying QMC point sets (either linearly scrambled digital nets or infinite-precision polynomial lattice point sets). Even though our approach does not require any information on the smoothness and weights of a target function space as an input, we can prove a probabilistic upper bound on the worst-case error for the respective weighted function space, where the failure probability converges to 0 exponentially fast as the number of estimates increases. Our obtained rates of convergence are nearly optimal for function spaces with finite smoothness, and we can attain a dimension-independent super-polynomial convergence for a class of infinitely differentiable functions. This implies that our median-based QMC rule is *universal* in the sense that it does not need to be adjusted to the smoothness and the weights of the function spaces and yet exhibits the nearly optimal rate of convergence. Numerical experiments support our theoretical results.

Key words. numerical integration, quasi-Monte Carlo, median, scrambled digital net, polynomial lattice point set, weighted function space, universality

MSC codes. 65D30, 65D32, 41A55, 42C10

1. Introduction. We study numerical integration for functions defined over the s-dimensional unit cube. For a Lebesgue integrable function $f : [0, 1]^s \to \mathbb{R}$, we denote its integral by

$$I_s(f) := \int_{[0,1]^s} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}.$$

In this paper, we consider integrands from Banach spaces that have continuous representatives. Quasi-Monte Carlo (QMC) methods approximate $I_s(f)$ by the equally-weighted mean of function evaluations over a carefully designed point set P:

$$Q_P(f) := \frac{1}{N} \sum_{\boldsymbol{x} \in P} f(\boldsymbol{x}),$$

where P is considered to be a multiset in the sense of combinatorics, i.e., a set that allows for repeated elements and takes into account the number of times each element appears, and $N := |P|$ denotes its cardinality as a multiset.

The history of QMC methods dates back to the work of Weyl [72] and has witnessed plenty of works in connection to discrepancy theory. One of the key findings is the Koksma-Hlawka inequality, which states that the integration error $|Q_P(f) - I_s(f)|$ is bounded above by the product of the total variation of f in the sense of Hardy and Krause and the star discrepancy of P [45, 12]. This error bound motivates us to study

  [†]School of Engineering, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan (goda@frcer.t.u-tokyo.ac.jp).

  [‡]Graduate School of Advanced Science and Engineering, Hiroshima University, 1-3-1 Kagamiyama, Higashi-Hiroshima 739-8526, Japan. Current address: Faculty of Pharmacy, Daiichi University of Pharmacy, 22-1 Tamagawa-machi, Minami-ku, Fukuoka 815-8511, Japan (k-suzuki@daiichi-cps.ac.jp).

  [§]Graduate School of Advanced Science and Engineering, Hiroshima University, 1-3-1 Kagamiyama, Higashi-Hiroshima 739-8526, Japan (m-mat@math.sci.hiroshima-u.ac.jp).

how to design a good point set such that the star discrepancy is small. In fact, there are many explicit constructions of low-discrepancy point sets and sequences provided in the literature, including those by Halton [26], Sobol' [62], Faure [17], Niederreiter [43], Tezuka [66] and Niederreiter and Xing [46]. If f has a bounded total variation, the error decays at the rate of either $(\log N)^{s-1}/N$ or $(\log N)^s/N$, depending on whether P is a low-discrepancy point set or the first N points of a low-discrepancy sequence, respectively.

More recently, one of the main approaches in QMC research is to consider a Banach space B with the norm $\|\cdot\|_B$, instead of looking only at functions with bounded variation, and study how to design a point set such that the worst-case error

$$e^{\mathrm{wor}}(Q_P; B) := \sup_{\substack{f \in B \\ \|f\|_B \leq 1}} |Q_P(f) - I_s(f)|$$

becomes small. As we have $|Q_P(f) - I_s(f)| \leq \|f\|_B \, e^{\mathrm{wor}}(Q_P; B)$ for any function $f \in B$, a single point set P with a small worst-case error bound should work well for all functions in B. Here it is important to note that, as long as we consider the worst-case for inputs f from the unit ball of B, which is a convex and symmetric set, optimal algorithms are *non-adaptive* and *linear* in the deterministic worst-case setting, as stated in [49, Theorem 4.7]. QMC methods are a special case of such non-adaptive, linear algorithms as they simply take the average of function evaluations. If we consider the worst-case over a non-convex or asymmetric input set $F \subset B$, though, then adaption may sometimes help [48].

As another key ingredient, Sloan and Woźniakowski used *weighted function spaces* to study how the worst-case error depends on the dimension s more precisely [61]; we also refer to earlier works by Hickernell [28, 29], who had already introduced weights in the definition of norms for function spaces. There the symbols $\boldsymbol{\gamma} = (\gamma_u)_{u \subseteq \{1,\ldots,s\}}$ with $\gamma_\emptyset = 1$ control the relative amount in the norm of the function to which each individual variable x_j or each subset of the variables $\boldsymbol{x}_u = (x_j)_{j \in u}$ contributes. Nowadays, these two ingredients have come together and there is a large body of work on *construction of good QMC point sets in weighted function spaces*. Typical examples include rank-1 lattice point sets in weighted Korobov spaces consisting only of smooth periodic functions [34, 50, 7] and (interlaced or extrapolated) polynomial lattice point sets in weighted Sobolev spaces with dominating mixed smoothness $\alpha \geq 2$ containing non-periodic smooth functions [11, 21, 9].

However, there are some shortcomings in this research direction. Firstly, to search for good point sets, the worst-case error or its upper bound must be computable in a reasonable time, which forces us to model the set of $2^s - 1$ weights γ_u by a smaller number of parameters. Secondly, although the information both on the smoothness and the weights is required as an input to define a weighted function space for construction, it is generally quite a hard problem to find appropriate smoothness and weights for a given application, except for some special cases such as partial differential equations with random coefficients [36, 11, 35]. Last, but not least, if good point sets are constructed in a weighted function space to which a target integrand does not belong, we may not have any theoretical guarantee that the error converges at the desired rate. Although these issues have been partly addressed in several works [5, 6, 16], none of them could completely eliminate the need to specify the weights.

A recent work by L'Ecuyer and the first named author of this paper has introduced a novel method that does not require any knowledge of the smoothness and weights at all but still achieves a desired convergence rate of the worst-case error with

a high probability [23]. Instead of trying to find good point sets, their approach goes as follows: select rank-1 lattice point sets $P_1, \ldots, P_r$ independently and randomly for an odd integer r, compute the integral estimates $Q_{P_1}(f), \ldots, Q_{P_r}(f)$ and take the median of them. Even with such a construction-free rule, a probabilistic worst-case error bound could be established for weighted Korobov spaces with any smoothness and general weights. Moreover, a similar result has been shown for high order polynomial lattice point sets in weighted Sobolev spaces with smoothness $\alpha \geq 2$, getting rid of the necessity that the integrand should be periodic. Although each of $Q_{P_1}, \ldots, Q_{P_r}$ is a linear non-adaptive algorithm, taking their median results in a non-linear algorithm. This non-linearity distinguishes it from the theoretically optimal deterministic integration methods. We justify the use of such a non-linear algorithm by observing that we are now working with randomized algorithms, where achieving a high confidence level is desirable because it is usually hard to check the actual worst-case error associated with every realization of randomized algorithms. Additionally, by utilizing the median trick, it becomes possible for the probability of "failure" in meeting a probabilistic error bound to decay exponentially fast toward 0 as the number of trials r increases, see [33, Proposition 1] and [23, Theorem 2.5] among some others. For the worst-case error analysis, Proposition 3.2 in this paper establishes a general median principle. Here we point out that taking the median of several randomized QMC estimates has been also studied quite recently in [20, 31, 56, 55].

In this paper, we push forward this novel research direction by extending some of the results shown in [23]. Our primary aim is to establish a *universality* of QMC-based integration methods in the sense that they do not need to be adjusted to the smoothness and the weights of non-periodic function spaces and yet exhibit the nearly optimal rate of convergence with high probability. To do so, we consider two special classes of QMC point sets, namely, linearly scrambled digital nets and (infinite-precision) polynomial lattice point sets, and cover three weighted function spaces with different smoothness classes simultaneously. Other than one function space which is exactly the same as the one studied in [23] for high order polynomial lattice point sets, we consider a weighted Sobolev space of first order at one end, which contains functions with absolutely integrable mixed first derivatives, and a weighted space of infinitely many times differentiable functions at the other end. As related works, we refer to [32, 70, 69] which study the universality of (randomized) Frolov cubature [19] for unweighted function spaces with finite smoothness, and also to [56] which shows a universality of one-dimensional scrambled nets in terms of smoothness of functions, covering both the finite and infinite smoothness classes. A part of the results in the latter paper has been extended to multi-dimensional scrambled nets quite recently in [55]. However, a universality in terms of the weights has been largely missing in the literature.

The rest of this paper is organized as follows. In the next section, we introduce linearly scrambled digital nets and polynomial lattice point sets and give some new results on their dual properties. Then, in section 3, we show a probabilistic worst-case error bound for each of three weighted function spaces and discuss its dependence on the dimension s. In section 4, we perform some numerical experiments to support our theoretical claims.

Notation. Throughout this paper, we denote the set of positive integers by $\mathbb{N}$ and write $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For $s \in \mathbb{N}$, we use the shorthand $1{:}s$ to denote the set $\{1, \ldots, s\}$. We always assume that b is a prime. We denote by $\mathbb{F}_b = \{0, 1, \ldots, b-1\}$ the field of b elements, i.e., the simplest construction of the prime field of order b by taking the integers modulo b, and by $\mathbb{F}_b^{n \times m}$ the set of $n \times m$ matrices over $\mathbb{F}_b$ for $m, n \in \mathbb{N}$. For

$k \in \mathbb{N}_0$ having the b-adic expansion $k = \kappa_0 + \kappa_1 b + \cdots$, where all but a finite number of κ_i's are 0, we write $k(x) = \kappa_0 + \kappa_1 x + \cdots \in \mathbb{F}_b[x]$ and $\vec{k} = (\kappa_0, \kappa_1, \ldots)^\top \in \mathbb{F}_b^{\mathbb{N}}$.

2. QMC point sets. In this paper we study two special classes of QMC point sets, linear scrambled digital (t, m, s)-nets and (infinite-precision) polynomial lattice point sets. Both of them can be regarded as a class of digital nets, respectively, which we define as follows.

DEFINITION 2.1 (digital net). *Let* $m \in \mathbb{N}$, $n \in \mathbb{N} \cup \{\infty\}$ *and* $C_1, \ldots, C_s \in \mathbb{F}_b^{n \times m}$. *For each integer* $0 \le k < b^m$, *we denote its b-adic expansion by* $k = \sum_{i=0}^{m-1} \kappa_i b^i$ *with* $\kappa_i \in \mathbb{F}_b$, *and obtain the k-th point* $\boldsymbol{x}_k \in [0, 1]^s$ *by*

(2.1) $$\boldsymbol{x}_k := (\psi_n(\boldsymbol{y}_{k,1}), \ldots, \psi_n(\boldsymbol{y}_{k,s})),$$

where we write $\boldsymbol{y}_{k,j} := C_j(\kappa_0, \ldots, \kappa_{m-1})^\top \in \mathbb{F}_b^n$ *for each j, and the map* $\psi_n \colon \mathbb{F}_b^n \to [0, 1]$ *is defined by*

$$\psi_n((y_1, \ldots, y_n)^\top) := \sum_{i=1}^{n} \frac{y_i}{b^i}.$$

The b^m-element point set $\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{b^m-1}\}$ *constructed this way is called a digital net over $\mathbb{F}_b$ with generating matrices* $C_1, \ldots, C_s$, *which we denote by* $P(C_1, \ldots, C_s)$.

In this section, we first introduce the definitions of linear scrambled digital (t, m, s)-nets and polynomial lattice point sets, respectively, and then give some new results on their dual properties.

2.1. Digital (t, m, s)-nets and linear scrambling. The t-value is the central measure of how uniformly distributed a digital net is, which comes from the concept of (t, m, s)-nets defined in terms of the elementary intervals as follows.

DEFINITION 2.2 (elementary interval and (t, m, s)-net). *A b-adic s-dimensional elementary interval denotes an axis-parallel, right half-open box of the form*

$$\prod_{j=1}^{s} \left[\frac{c_j}{b^{a_j}}, \frac{c_j + 1}{b^{a_j}} \right), \quad \text{with } a_j, c_j \in \mathbb{N}_0 \text{ and } 0 \le c_j < b^{a_j}.$$

For $m \in \mathbb{N}_0$, a b^m-element point set $P \subset [0, 1)^s$ *is called a (t, m, s)-net in base b if every elementary interval of volume b^{-m+t} contains exactly b^t points of P.*

It is obvious from the definition that a smaller t-value ensures a finer uniformity of points and $t = 0$ is the best possible. Now it is natural to call a point set P a *digital (t, m, s)-net in base b* if P is a digital net over $\mathbb{F}_b$ as well as a (t, m, s)-net in base b. Since any b^m-element point set or digital net over $\mathbb{F}_b$ is an (m, m, s)-net in base b, the concept of (t, m, s)-net is only useful when $t < m$ [12, Section 2.5].

The least t-value of a given digital net can be determined in the following way. We refer to [15, Theorem 4.52] for the proof.

LEMMA 2.3. *Let* $m, n \in \mathbb{N}$ *and* $C_1, \ldots, C_s \in \mathbb{F}_b^{n \times m}$. *The point set* $P(C_1, \ldots, C_s)$ *is a digital (t, m, s)-net in base b with t being the smallest non-negative integer such that, for any* $a_1, \ldots, a_s \in \mathbb{N}_0$ *with* $a_1 + \cdots + a_s = m - t$,

the first a_1 rows of C_1,
the first a_2 rows of C_2,
$\vdots$

and the first a_s rows of C_s

are linearly independent over $\mathbb{F}_b$.

Remark 2.4. If $n = \infty$, the t-value determined by Lemma 2.3 does not necessarily coincide with the one in the sense of Definition 2.2. This problem occurs because a digital net with generating matrices having infinitely many rows allows for two different representations of b-adic rationals, such as

$$x = \frac{1}{b} \quad \text{and} \quad x = \frac{b-1}{b^2} + \frac{b-1}{b^3} + \cdots.$$

In such cases, we always consider the t-value in the sense of Lemma 2.3.

Explicit constructions of digital (t, m, s)-nets with small t-value have been provided by Sobol' [62], Niederreiter [43] and Tezuka [66] as well as many others, see [45, Chapter 4] and [15, Chapter 8]. As a key ingredient for the subsequent analysis, let us mention the result from [71] on the t-value of the Niederreiter sequence in base b: for any $m, s \in \mathbb{N}$ and any non-empty subset $u \subseteq 1{:}s$, the projection of the first b^m points from the s-dimensional Niederreiter sequence onto the $|u|$-dimensional unit cube $[0, 1)^{|u|}$ is a digital $(t_u, m, |u|)$-net in base b with

$$(2.2) \qquad t_u \le \sum_{j \in u} \left(\log_b j + \log_b \log_b (j + b) + 1 \right).$$

As explained in [12, Example 2.20], the Niederreiter sequence replaces the sequence of primitive polynomials over $\mathbb{F}_2$, used in the Sobol' sequence construction, with a sequence of irreducible polynomials over $\mathbb{F}_b$ with prime power b. As primitive polynomials are always irreducible, this replacement leads to a smaller bound on the t-value.

We now introduce the definition of linearly scrambled digital nets (without shift).

DEFINITION 2.5 (linearly scrambled digital net). *For $w, n \in \mathbb{N} \cup \{\infty\}$, we define a set of non-singular lower triangular matrices over $\mathbb{F}_b$, denoted by $\mathcal{L}_{w,n}$, as*

$$\mathcal{L}_{w,n} := \left\{ L = (\ell_{i,j}) \in \mathbb{F}_b^{w \times n} \;\middle|\; \ell_{i,j} = 0 \text{ if } i < j \text{ and } \ell_{i,j} \neq 0 \text{ if } i = j \right\}.$$

For a digital net $P(C_1, \ldots, C_s)$ with $C_1, \ldots, C_s \in \mathbb{F}_b^{n \times m}$ and $L_1, \ldots, L_s \in \mathcal{L}_{w,n}$, the corresponding linearly scrambled digital net (with precision w) is a digital net $P(L_1 C_1, \ldots, L_s C_s)$. If each L_j is independently and randomly chosen from $\mathcal{L}_{w,n}$, we call $P(L_1 C_1, \ldots, L_s C_s)$ a randomly linearly scrambled digital net over $\mathbb{F}_b$. Here a random sampling from $\mathcal{L}_{w,n}$ is done by choosing each element $\ell_{i,j}$ independently and randomly under the above conditions.

Remark 2.6. To be more precise on random sampling from $\mathcal{L}_{w,n}$ in Definition 2.5, let us denote the full index set by $U_{w,n} = \{(i, j) \mid 1 \le i \le w, 1 \le j \le n, i \le j\}$ for given $w, n \in \mathbb{N} \cup \{\infty\}$, and consider the Cartesian product

$$\mathcal{X}_{w,n} := \prod_{\substack{(i,j) \in U_{w,n} \\ i=j}} (\mathbb{F}_b \setminus \{0\}) \times \prod_{\substack{(i,j) \in U_{w,n} \\ i<j}} \mathbb{F}_b,$$

which is canonically isomorphic to $\mathcal{L}_{w,n}$. The Kolmogorov extension theorem ensures that, even if either w or n is infinite, there exists a measure ν on $\mathcal{X}_{w,n}$ induced by the uniform probability measures on $\mathbb{F}_b \setminus \{0\}$ and $\mathbb{F}_b$. Specifically, for every cylinder set

$$E = \prod_{\substack{(i,j) \in V \\ i=j}} F_{ii} \times \prod_{\substack{(i,j) \in V \\ i<j}} F_{ij} \times \prod_{\substack{(i,j) \in U_{w,n} \setminus V \\ i=j}} (\mathbb{F}_b \setminus \{0\}) \times \prod_{\substack{(i,j) \in U_{w,n} \setminus V \\ i<j}} \mathbb{F}_b,$$

with a finite index subset $V \subset U_{w,n}$, $F_{ii} \subseteq \mathbb{F}_b \setminus \{0\}$ and $F_{ij} \subseteq \mathbb{F}_b$, it holds that

$$\nu(E) = \prod_{\substack{(i,j) \in V \\ i=j}} \frac{|F_{ii}|}{b-1} \times \prod_{\substack{(i,j) \in V \\ i<j}} \frac{|F_{ij}|}{b}.$$

Subsequently, when computing the probability of an event or the expectation of functions with respect to $L_1, \ldots, L_s \in \mathcal{L}_{w,n}$, we consider the s-fold product of this cylinder set measure.

In what follows, we focus on the case where $w = \infty$, which facilitates our theoretical analysis. In practice, representing infinite digit expansions on computers is quite hard, but, even with finite expansions, the accompanying truncation error diminishes as w increases, as shown in [41, Lemma 2.1] and [55, Lemma 1 & Corollary 3].

Remark 2.7. A deterministic linear scrambling as in Definition 2.5 has been originally introduced by Tezuka [67] to generalize Faure sequences. A random linear scrambling together with a random digital shift has been studied by Matoušek [39] as a computationally more efficient alternative to the fully nested scrambling by Owen [51, 52, 53]. It is important that, no matter how we choose $L_1, \ldots, L_s \in \mathcal{L}_{w,n}$ with $w \geq m$, the t-value of a linearly scrambled net $P(L_1 C_1, \ldots, L_s C_s)$ remains the same as that of $P(C_1, \ldots, C_s)$, i.e., a linear scrambling preserves the t-value. In passing, we point out that further derandomization of random linear scrambling has been studied in [54, 68]. For example, i-binomial scrambling developed in [68] restricts the matrices in $\mathcal{L}_{w,n}$ to those of Toeplitz type. Although we will not go into the details, the results of this paper also apply to such cases.

2.2. Polynomial lattice point sets. As the other class of point sets we study in this paper, here we introduce polynomial lattice point sets.

DEFINITION 2.8 (polynomial lattice point set). *For $m, s \in \mathbb{N}$ and $w \in \mathbb{N} \cup \{\infty\}$ with $w \geq m$, let $p \in \mathbb{F}_b[x]$ with $\deg(p) = m$ and $\boldsymbol{g} = (g_1, \ldots, g_s) \in (\mathbb{F}_b[x])^s$ with $\deg(g_j) < m$. The polynomial lattice point set (of precision w) with modulus p and generating vector $\boldsymbol{g}$ is a b^m-element point set*

$$P(p, \boldsymbol{g}, w) = \left\{ \boldsymbol{x}_k := \nu_w \left(\frac{k(x)\boldsymbol{g}(x)}{p(x)} \right) \,\middle|\, 0 \leq k < b^m \right\},$$

where $\nu_w : \mathbb{F}_b((x^{-1})) \to [0,1]$ is applied component-wise to a vector and is given by

$$\nu_w \left(\sum_{i=c}^{\infty} \frac{a_i}{x^i} \right) = \sum_{i=\max(1,c)}^{w} \frac{a_i}{b^i}.$$

When $w = \infty$, we simply write $P(p, \boldsymbol{g})$ instead of $P(p, \boldsymbol{g}, \infty)$.

Polynomial lattice point sets have been originally introduced by Niederreiter [44] and their relation to linear shift-register sequences has been observed in [38]. We also refer to [37, 14, 57]. It is well-known that polynomial lattice point sets can be regarded as a special class of digital nets over $\mathbb{F}_b$ as follows. We refer to [15, Chapter 10] for the proof.

LEMMA 2.9. *For $m, s \in \mathbb{N}$ and $w \in \mathbb{N} \cup \{\infty\}$ with $w \geq m$, let $p \in \mathbb{F}_b[x]$ with $\deg(p) = m$ and $\boldsymbol{g} = (g_1, \ldots, g_s) \in (\mathbb{F}_b[x])^s$ with $\deg(g_j) < m$. The polynomial lattice point set $P(p, \boldsymbol{g}, w)$ is a digital net over $\mathbb{F}_b$ with generating matrices $C_1, \ldots, C_s \in$*

$\mathbb{F}_b^{w \times m}$, *where each matrix* $C_j = (c_{i,r}^{(j)})$ *is of Hankel type:*

$$c_{i,r}^{(j)} = u_{i+r-1}^{(j)},$$

and $u_1^{(j)}, u_2^{(j)}, \ldots$ *are the coefficients appearing in the infinite (periodic) Laurent series of the quotient*

$$\frac{g_j(x)}{p(x)} = \sum_{i=1}^{\infty} \frac{u_i^{(j)}}{x^i} \in \mathbb{F}_b((x^{-1})).$$

Remark 2.10. Polynomial lattice point sets of infinite precision $w = \infty$ have been studied in the randomized setting carefully in [7, Section 5]. We also refer to a relevant work [25] on digital nets with infinite digit expansions. In this paper, we focus only on the case $w = \infty$ in the theoretical analysis, but the result can be well-approximated by a large enough w. Note that, for $w = \infty$, QMC integration is given by

$$Q_{P(p,\boldsymbol{g})}(f) := \lim_{w \to \infty} Q_{P(p,\boldsymbol{g},w)}(f) = \frac{1}{b^m} \sum_{k=0}^{b^m-1} f(\boldsymbol{x}_k-),$$

where $f(\boldsymbol{x}-) = \lim_{\boldsymbol{y} \nearrow \boldsymbol{x}} f(\boldsymbol{y})$ denotes the component-wise left limit. Therefore, if f is not left continuous, we may have

$$Q_{P(p,\boldsymbol{g})}(f) \neq Q_{P(p,\boldsymbol{g},\infty)}(f) = \frac{1}{b^m} \sum_{k=0}^{b^m-1} f(\boldsymbol{x}_k),$$

whereas such a problem does not happen in this paper as we only deal with continuous functions.

In the subsequent analysis, we restrict the modulus p to be monic and irreducible. For $m \in \mathbb{N}$, we write

$$\mathbb{P}_m := \{p \in \mathbb{F}_b[x] \mid \deg(p) = m \text{ and } p \text{ is monic and irreducible}\},$$

and

$$\mathbb{G}_m := \{g \in \mathbb{F}_b[x] \mid g \neq 0 \text{ and } \deg(g) < m\}.$$

We shall always choose $p \in \mathbb{P}_m$ and $\boldsymbol{g} \in (\mathbb{G}_m)^s$. It is obvious that $|\mathbb{G}_m| = b^m - 1$ and also it is known from [58, Lemma 4] that $|\mathbb{P}_m| \geq b^m/(2m)$.

2.3. Duality and some new results. The error of QMC integration using digital nets is usually analyzed through the concept of dual nets, which are defined as follows.

DEFINITION 2.11 (Dual net). *Let* $m \in \mathbb{N}$, $n \in \mathbb{N} \cup \{\infty\}$ *and* $C_1, \ldots, C_s \in \mathbb{F}_b^{n \times m}$. *The dual net of the digital net* $P(C_1, \ldots, C_s)$ *is defined by*

$$P^\perp(C_1, \ldots, C_s) := \left\{ \boldsymbol{k} \in \mathbb{N}_0^s \ \middle| \ C_1^\top \operatorname{tr}_n(\vec{k}_1) + \cdots + C_s^\top \operatorname{tr}_n(\vec{k}_s) = \boldsymbol{0} \in \mathbb{F}_b^m \right\},$$

where we use the notation

$$\operatorname{tr}_n(\vec{k}) = (\kappa_0, \kappa_1, \ldots, \kappa_{n-1})^\top \in \mathbb{F}_b^n,$$

for $k = \kappa_0 + \kappa_1 b + \cdots \in \mathbb{N}_0$, *which stands for the truncated* b-*adic* n-*digit representation.*

In this subsection we give some new results on dual nets for linearly scrambled digital nets and polynomial lattice point sets, respectively.

2.3.1. Linear scrambled digital nets. We first describe an alternative way to Lemma 2.3 for determining the least t-value of a digital net through its dual net. To do so, we need to introduce the Niederreiter–Rosenbloom–Tsfasman (NRT) weight.

DEFINITION 2.12 (NRT weight). *For $k \in \mathbb{N}$, we denote its b-adic expansion by*

$$k = \kappa_1 b^{c_1 - 1} + \kappa_2 b^{c_2 - 1} + \cdots + \kappa_v b^{c_v - 1},$$

with $\kappa_1, \ldots, \kappa_v \in \{1, \ldots, b-1\}$ and $c_1 > c_2 > \cdots > c_v > 0$. The NRT weight $\mu_1 : \mathbb{N}_0 \to \mathbb{N}_0$ is defined by $\mu_1(k) = c_1$ and $\mu_1(0) = 0$, or equivalently by $\mu_1(k) = \lceil \log_b(k+1) \rceil$. Moreover, in the case of vectors in $\mathbb{N}_0^s$, we define

$$\mu_1(\boldsymbol{k}) = \sum_{j=1}^{s} \mu_1(k_j).$$

Now the t-value and the dual net of a digital net can be associated with each other in the following way. We refer to [15, Chapter 7] for the proof.

LEMMA 2.13. *Let $m \in \mathbb{N}$, $n \in \mathbb{N} \cup \{\infty\}$ and $C_1, \ldots, C_s \in \mathbb{F}_b^{n \times m}$. Define*

$$\mu_1(P^\perp(C_1, \ldots, C_s)) := \min_{\boldsymbol{k} \in P^\perp(C_1, \ldots, C_s) \setminus \{\boldsymbol{0}\}} \mu_1(\boldsymbol{k}).$$

Then the least t-value of the digital net $P(C_1, \ldots, C_s)$ equals

$$m - \mu_1(P^\perp(C_1, \ldots, C_s)) + 1.$$

In what follows, we give some new results on the dual net. Although we have already mentioned in Remark 2.7 that a linear scrambling preserves the t-value, we need a bit stronger result as follows.

LEMMA 2.14. *Let $k \in \mathbb{N}$ and let us write $c := \mu_1(k)$. Then, for any $L \in \mathcal{L}_{\infty,\infty}$, by letting $\ell \in \mathbb{N}$ such that $\vec{\ell} = L^\top \vec{k}$, we always have $\mu_1(\ell) = c$. Moreover, for any $k' \in \mathbb{N}$ satisfying $\mu_1(k') = c$ and any $n \geq c$, if L is randomly sampled from $\mathcal{L}_{\infty,n}$, we have*

$$\Pr\left[L^\top \vec{k} = \vec{k'} \right] = \frac{1}{(b-1)b^{c-1}}.$$

Proof. The first claim is trivial from the definition of $\mathcal{L}_{w,n}$. Noting that only the first c rows of L affect the result of the multiplication $L^\top \vec{k}$, and also that, for any choice of the first $c-1$ rows of L, there exists exactly one vector for the c-th row of L such that $L^\top \vec{k} = \vec{k'}$ holds, the second claim follows immediately from the fact that the number of possible choices for the c-th row is $(b-1)b^{c-1}$. $\square$

This lemma can be trivially extended to the multi-dimensional case as follows.

LEMMA 2.15. *For $\boldsymbol{k} = (k_1, \ldots, k_s) \in \mathbb{N}_0^s$, let us write $c_j := \mu_1(k_j)$ and $u := \{j \mid k_j \neq 0\}$. Then, for any j and $L_j \in \mathcal{L}_{\infty,\infty}$, by letting ℓ_j such that $\vec{\ell}_j = L_j^\top \vec{k}_j$, we have $\mu_1(\ell_j) = c_j$. Moreover, for any $\boldsymbol{k'} = (k'_1, \ldots, k'_s) \in \mathbb{N}_0^s$ with $\mu_1(k'_j) = c_j$ for all j and any $n \geq \max_j c_j$, if $L_1, \ldots, L_s$ are independently and randomly sampled from $\mathcal{L}_{\infty,n}$, we have*

$$\Pr\left[L_j^\top \vec{k}_j = \vec{k'}_j \text{ for all } j \right] = \left(\frac{b}{b-1} \right)^{|u|} b^{-(c_1 + \cdots + c_s)}.$$

The following result is a straightforward generalization of the recent result shown in [24, Lemma 3.4] by building up on [15, Corollary 13.8]. We also refer to [60, Lemma 2.2] for the previously known result.

LEMMA 2.16. *Let P be a digital net over $\mathbb{F}_b$ of size b^m. For a non-empty $u \subseteq 1{:}s$, let $P_u := \{\boldsymbol{x}_u \mid \boldsymbol{x} \in P\}$ be the projection of P onto the $|u|$-dimensional unit cube and assume that P_u is a digital $(t_u, m, |u|)$-net in base b. Then, for any $\boldsymbol{c}_u = (c_j)_{j \in u} \in \mathbb{N}^{|u|}$, by writing $|\boldsymbol{c}_u|_1 = \sum_{j \in u} c_j$, we have*

$$\left| \{\boldsymbol{k} = (k_1, \ldots, k_s) \in P^\perp \mid \mu_1(k_j) = c_j \text{ if } j \in u \text{ and } k_j = 0 \text{ otherwise} \} \right|$$
$$\leq \begin{cases} 0 & \text{if } |\boldsymbol{c}_u|_1 \leq m - t_u, \\ (b-1)^{|\boldsymbol{c}_u|_1 - (m - t_u)} & \text{if } m - t_u < |\boldsymbol{c}_u|_1 \leq m - t_u + |u|, \\ (b-1)^{|u|} b^{|\boldsymbol{c}_u|_1 - (m - t_u + |u|)} & \text{if } |\boldsymbol{c}_u|_1 > m - t_u + |u|. \end{cases}$$

In particular, we have

$$\left| \{\boldsymbol{k} = (k_1, \ldots, k_s) \in P^\perp \mid \mu_1(k_j) = c_j \text{ if } j \in u \text{ and } k_j = 0 \text{ otherwise} \} \right|$$
$$\leq \left(\frac{b-1}{b} \right) b^{|\boldsymbol{c}_u|_1 - (m - t_u)}.$$

We now prove what we need in the subsequent analysis. The result can be regarded as a generalization of the known result recently shown in [55, Lemma 3 and Corollary 1].

LEMMA 2.17. *For $m \in \mathbb{N}$, let $P = P(C_1, \ldots, C_s)$ be a digital net over $\mathbb{F}_b$ of size b^m with square generating matrices $C_1, \ldots, C_s \in \mathbb{F}_b^{m \times m}$, and assume that, for every non-empty $u \subseteq 1{:}s$, the projection P_u is a digital $(t_u, m, |u|)$-net in base b. Moreover, let $L_1, \ldots, L_s$ be independently and randomly sampled from $\mathcal{L}_{\infty, m}$. For any $\boldsymbol{k} = (k_1, \ldots, k_s) \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}$, by writing $u = \{j \mid k_j \neq 0\}$, we have*

$$\Pr\left[\boldsymbol{k} \in P^\perp(L_1 C_1, \ldots, L_s C_s) \right] \leq \left(\frac{b}{b-1} \right)^{|u|-1} b^{-m + t_u},$$

if $k_j < b^m$ holds for all j. Otherwise, assuming that each C_j is non-singular, we have

$$\Pr\left[\boldsymbol{k} \in P^\perp(L_1 C_1, \ldots, L_s C_s) \right] = \frac{1}{b^m}.$$

Proof. First let us consider the case $k_j < b^m$ for all j. From Definition 2.11, we have an equivalence

$$\boldsymbol{k} \in P^\perp(L_1 C_1, \ldots, L_s C_s)$$
$$\iff (L_1 C_1)^\top \vec{k}_1 + \cdots + (L_s C_s)^\top \vec{k}_s = \boldsymbol{0} \in \mathbb{F}_b^m$$
$$\iff \boldsymbol{k}' \in P^\perp(C_1, \ldots, C_s) \quad \text{with } k_j' \in \mathbb{N}_0 \text{ satisfying } \vec{k}_j' = L_j^\top \vec{k}_j.$$

It follows from Lemmas 2.15 and 2.16 that

$$\Pr\left[\boldsymbol{k} \in P^\perp(L_1 C_1, \ldots, L_s C_s) \right]$$
$$= \sum_{\boldsymbol{k}' \in P^\perp(C_1, \ldots, C_s)} \Pr\left[L_j^\top \vec{k}_j = \vec{k}_j' \text{ for all } j \right]$$
$$\leq \left(\frac{b}{b-1} \right)^{|u|} b^{-|\boldsymbol{c}_u|_1}$$
$$\quad \times \left| \{\boldsymbol{k}' \in P^\perp(C_1, \ldots, C_s) \mid \mu_1(k_j') = c_j \text{ if } j \in u \text{ and } k_j' = 0 \text{ otherwise} \} \right|$$

$$\leq \left(\frac{b}{b-1}\right)^{|u|} b^{-|\boldsymbol{c}_u|_1} \left(\frac{b-1}{b}\right) b^{|\boldsymbol{c}_u|_1 - (m-t_u)} = \left(\frac{b}{b-1}\right)^{|u|-1} b^{-(m-t_u)},$$

which proves the first claim.

Let us move on to the second case in which there exists at least one index j^* such that $k_{j^*} \geq b^m$. Because of the assumption that all C_j are non-singular, we have an equivalence

$$\boldsymbol{k} \in P^{\perp}(L_1 C_1, \ldots, L_s C_s) \iff (L_1 C_1)^{\top} \vec{k}_1 + \cdots + (L_s C_s)^{\top} \vec{k}_s = \boldsymbol{0} \in \mathbb{F}_b^m$$

$$\iff L_{j^*}^{\top} \vec{k}_{j^*} = - \sum_{j \neq j^*} (C_{j^*}^{\top})^{-1} (L_j C_j)^{\top} \vec{k}_j.$$

Since $k_{j^*} \geq b^m$, there exists a unique integer $v > m$ such that $\vec{k}_{j^*,v} \neq 0$ and $\vec{k}_{j^*,v+1} = \vec{k}_{j^*,v+2} = \cdots = 0$. Then we see that only the first v rows of L_{j^*} affect the result of the multiplication $L_{j^*}^{\top} \vec{k}_{j^*}$, and that, for any choice of the first $v-1$ rows of L_{j^*}, $L_{j^*}^{\top} \vec{k}_{j^*}$ is distributed uniformly on the set $\mathbb{F}_b^m$ when the v-th row of L_{j^*} is distributed uniformly on the set $\mathbb{F}_b^m$. Hence we have

$$\Pr \left[L_{j^*}^{\top} \vec{k}_{j^*} = - \sum_{j \neq j^*} (C_{j^*}^{\top})^{-1} (L_j C_j)^{\top} \vec{k}_j \right] = \frac{1}{b^m},$$

which proves the second claim. □

Remark 2.18. We note that, for the second claim of Lemma 2.17, we have assumed that each C_j is non-singular. This implies that any one-dimensional projection of $P(C_1, \ldots, C_s)$ is a $(0, m, 1)$-net in base b. Such a good property holds for Sobol' sequences and Niederreiter sequences, where a proper reordering of the rows of each matrix C_j is necessary for the latter [18]. Here reordering the rows does not change the t-value of nets shown in (2.2).

2.3.2. Polynomial lattice point sets. Here we show a result corresponding to polynomial lattice point sets. First, let us recall that the dual net of a polynomial lattice point set is expressed in a specific way as follows. We refer to [15, Lemma 10.6] for the proof and also to [7, Section 5.1] for a relevant result on the character property.

LEMMA 2.19. *For $m, s \in \mathbb{N}$ and $w \in \mathbb{N} \cup \{\infty\}$ with $w \geq m$, let $p \in \mathbb{F}_b[x]$ with $\deg(p) = m$ and $\boldsymbol{g} \in (\mathbb{F}_b[x])^s$ with $\deg(g_j) < m$. The dual net of the polynomial lattice point set $P(p, \boldsymbol{g}, w)$ is given by*

$$P^{\perp}(p, \boldsymbol{g}, w) = \{\boldsymbol{k} \in \mathbb{N}_0^s \mid \mathrm{tr}_w(\boldsymbol{k}(x)) \cdot \boldsymbol{g}(x) \equiv 0 \pmod{p(x)}\},$$

where, extending the notation of Definition 2.11, we write

$$tr_w(k(x)) = \kappa_0 + \kappa_1 x + \cdots + \kappa_{w-1} x^{w-1},$$

for $k(x) = \kappa_0 + \kappa_1 x + \cdots \in \mathbb{F}_b[x]$, which stands for the truncated polynomial over $\mathbb{F}_b$ of degree less than w and is applied component-wise to a vector. For $w = \infty$ we simply have $tr_w(k(x)) = k(x)$ for any $k \in \mathbb{N}_0$ and write $P^{\perp}(p, \boldsymbol{g})$ instead of $P^{\perp}(p, \boldsymbol{g}, \infty)$.

Now the following result is what we need in the subsequent analysis.

LEMMA 2.20. *For $m, s \in \mathbb{N}$, let p and $\boldsymbol{g}$ be independently and randomly sampled from $\mathbb{P}_m$ and $(\mathbb{G}_m)^s$, respectively. For any $\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}$, it holds that*

$$\Pr\left[\boldsymbol{k} \in P^{\perp}(p, \boldsymbol{g})\right] \leq \frac{3\mu_1(\boldsymbol{k})}{b^m - 1}.$$

Moreover, if $k_j < b^m$ holds for all j, we have

$$\Pr\left[\boldsymbol{k} \in P^{\perp}(p, \boldsymbol{g})\right] \leq \frac{1}{b^m - 1}.$$

Proof. Let χ_A be the indicator function of an event A. It follows from Lemma 2.19 that

$$\Pr\left[\boldsymbol{k} \in P^{\perp}(p, \boldsymbol{g})\right] := \frac{1}{|\mathbb{P}_m|} \sum_{p \in \mathbb{P}_m} \frac{1}{|\mathbb{G}_m|^s} \sum_{\boldsymbol{g} \in \mathbb{G}_m^s} \chi_{\boldsymbol{k} \cdot \boldsymbol{g} \equiv 0 \pmod{p}},$$

where, with abuse of notation, $\boldsymbol{k}$ appearing on the right-hand side denotes the associated vector of polynomials in $\mathbb{F}_b[x]$. If the modulus p is a divisor of all the components of $\boldsymbol{k}$, on the one hand, the condition $\boldsymbol{k} \in P^{\perp}(p, \boldsymbol{g})$ holds for any choice of $\boldsymbol{g} \in \mathbb{G}_m^s$, which implies that

$$\frac{1}{|\mathbb{G}_m|^s} \sum_{\boldsymbol{g} \in \mathbb{G}_m^s} \chi_{\boldsymbol{k} \cdot \boldsymbol{g} \equiv 0 \pmod{p}} = 1.$$

Assume, on the other hand, that there exists a non-empty subset $u \subseteq 1{:}s$ such that k_j is not divisible by p if and only if $j \in u$. Then the condition $\boldsymbol{k} \in P^{\perp}(p, \boldsymbol{g})$ is equivalent to $\boldsymbol{k}_u \cdot \boldsymbol{g}_u = \sum_{j \in u} k_j g_j \equiv 0 \pmod{p}$. Here the cardinality of u must be larger than 1, since, if $u = \{j\}$, the equivalent condition $k_j g_j \equiv 0 \pmod{p}$ contradicts the fact that p is irreducible and $g_j \in \mathbb{G}_m$. Thus, defining $\ell = \max_{j \in u} j$, we have $\ell \geq 2$, and the above condition $\sum_{j \in u} k_j g_j \equiv 0 \pmod{p}$ is further equivalent to

$$k_\ell g_\ell \equiv -\boldsymbol{k}_{u \setminus \{\ell\}} \cdot \boldsymbol{g}_{u \setminus \{\ell\}} \equiv -\boldsymbol{k}_{1:\ell-1} \cdot \boldsymbol{g}_{1:\ell-1} \pmod{p}.$$

If $\boldsymbol{k}_{1:\ell-1} \cdot \boldsymbol{g}_{1:\ell-1} \equiv 0 \pmod{p}$ holds, no $g_\ell \in \mathbb{G}_m$ satisfies this equation. Otherwise, there exists exactly one $g_\ell \in \mathbb{G}_m$ which solves the equation. This leads to

$$\frac{1}{|\mathbb{G}_m|^s} \sum_{\boldsymbol{g} \in \mathbb{G}_m^s} \chi_{\boldsymbol{k} \cdot \boldsymbol{g} \equiv 0 \pmod{p}} \leq \frac{1}{|\mathbb{G}_m|}.$$

Now, since any polynomial $k \in \mathbb{F}_b[x]$ has at most $\lfloor \deg(k)/m \rfloor$ prime divisors of a fixed degree m, by using $|\mathbb{G}_m| = b^m - 1$ and $|\mathbb{P}_m| \geq b^m/(2m)$, we obtain

$$\begin{aligned}
\Pr\left[\boldsymbol{k} \in P^{\perp}(p, \boldsymbol{g})\right] &\leq \frac{1}{|\mathbb{P}_m|} \sum_{\substack{p \in \mathbb{P}_m \\ p | \boldsymbol{k}}} 1 + \frac{1}{|\mathbb{P}_m|} \sum_{\substack{p \in \mathbb{P}_m \\ p \nmid \boldsymbol{k}}} \frac{1}{|\mathbb{G}_m|} \\
&\leq \frac{1}{|\mathbb{P}_m|} \frac{\max_{1 \leq j \leq s} \deg(k_j)}{m} + \frac{1}{|\mathbb{G}_m|} \\
&\leq \frac{2 \max_{1 \leq j \leq s} \mu_1(k_j)}{b^m} + \frac{1}{b^m - 1} \leq \frac{3\mu_1(\boldsymbol{k})}{b^m - 1},
\end{aligned}$$

which proves the first statement of this lemma. The second statement follows immediately from the fact that there is no prime divisor of a fixed degree m for any polynomial $k \in \mathbb{F}_b[x]$ with $\deg(k) < m$. $\square$

3. Median QMC integration. Here we provide a basic framework for what we call *median QMC integration*. Let S be a collection of point sets in $[0,1]^s$. In the remainder of the paper, S will be either a set of linearly scrambled digital nets

$$(3.1) \qquad \{P(L_1 C_1, \ldots, L_s C_s) \mid L_1, \ldots, L_s \in \mathcal{L}_{\infty,n}\},$$

for some good generating matrices $C_1, \ldots, C_s \in \mathbb{F}_b^{n \times m}$, or a set of polynomial lattices

$$(3.2) \qquad \{P(p, \boldsymbol{g}) \mid p \in \mathbb{P}_m, \boldsymbol{g} \in \mathbb{G}_m^s\}.$$

For an odd integer r, we draw point sets $P_1, \ldots, P_r$ independently and randomly from the set S, and we approximate $I_s(f)$ by the median-of-means

$$M_r(f) := \operatorname{median}(Q_{P_1}(f), \ldots, Q_{P_r}(f)).$$

Because of random selection for $P_1, \ldots, P_r$, our estimate $M_r(f)$ is a random variable, which means that our median QMC integration is a randomized quadrature algorithm; for every realization of $P_1, \ldots, P_r$, the resulting estimator has a certain worst-case error. In what follows, we are interested in a lower bound on the probability of a small worst-case error (less than or equal to ϵ) with respect to independently and randomly chosen $P_1, \ldots, P_r$. A study of our median QMC integration in terms of different randomized error criteria is left open for future research. We refer to [33] as a relevant work in this research direction, which studies a probabilistic error guarantee of the form

$$\sup_{\substack{f \in B \\ \|f\|_B \leq 1}} \Pr\left[|A_n(f) - I_s(f)| > \epsilon\right] \leq \delta,$$

for a randomized algorithm A_n that uses n function values.

Remark 3.1. Although we do not go into the details, the results shown in this section also apply to the collections of randomly digitally shifted point sets

$$\{P(L_1 C_1, \ldots, L_s C_s) \oplus \Delta \mid L_1, \ldots, L_s \in \mathcal{L}_{\infty,n}, \Delta \sim U([0,1)^s)\},$$

and

$$\{P(p, \boldsymbol{g}) \oplus \Delta \mid p \in \mathbb{P}_m, \boldsymbol{g} \in \mathbb{G}_m^s, \Delta \sim (U[0,1)^s)\},$$

respectively. Here we write $P \oplus \Delta = \{\boldsymbol{x} \oplus \Delta \mid \boldsymbol{x} \in P\}$, and $\oplus$ denotes the digit-wise addition modulo b, that is, for $x = \xi_1/b + \xi_2/b^2 + \cdots$ and $y = \eta_1/b + \eta_2/b^2 + \cdots$, we define

$$x \oplus y = \frac{\zeta_1}{b} + \frac{\zeta_2}{b^2} + \cdots, \quad \text{with } \zeta_i \equiv \xi_i + \eta_i \pmod{b},$$

and apply component-wise to vectors. As mentioned in Remark 2.7, the former set above is nothing but the set of randomly scrambled digital nets in the sense of Matoušek [39]. Applying a random digital shift makes every individual estimator unbiased, whereas it is biased without a random digital shift. However, taking the median of r repetitions makes the resulting final estimator biased regardless of whether a random digital shift is applied or not.

3.1. Basic result. Let B be a Banach space over $[0,1]^s$ with the norm $\|\cdot\|_B$. If a realization of $P_1, \ldots, P_r \in S$ is given, our median QMC integration M_r is deterministic and the worst-case error is defined by

$$e^{\operatorname{wor}}(M_r; B) := \sup_{\substack{f \in B \\ \|f\|_B \leq 1}} |M_r(f) - I_s(f)|.$$

Then the question we are interested in is: *What are the possible pairs (ϵ, δ) that satisfy*

$$\Pr\left[e^{\mathrm{wor}}(M_r; B) \leq \epsilon\right] \geq 1 - \delta,$$

if the probability is taken with respect to independently and randomly chosen $P_1, \ldots, P_r$?
Regarding this question, we can establish the following "meta" result.

PROPOSITION 3.2. *Assume that there exists a pair (ϵ, δ) such that $\delta \in (0, 1/4)$ and*

(3.3)
$$\Pr\left[e^{\mathrm{wor}}(Q_P; B) \leq \epsilon\right] \geq 1 - \delta$$

holds if P is randomly chosen from S. Then, if $P_1, \ldots, P_r$ are independently and randomly chosen from S, we have

$$\Pr\left[e^{\mathrm{wor}}(M_r; B) \leq \epsilon\right] \geq 1 - 2^{r-1}\delta^{(r+1)/2}.$$

Proof. For each realization of $P_1, \ldots, P_r$, the corresponding worst-case error is bounded by

$$
\begin{aligned}
e^{\mathrm{wor}}(M_r; B) &= \sup_{\substack{f \in B \\ \|f\|_B \leq 1}} \left|\mathrm{median}\left(Q_{P_1}(f) - I_s(f), \ldots, Q_{P_r}(f) - I_s(f)\right)\right| \\
&\leq \sup_{\substack{f \in B \\ \|f\|_B \leq 1}} \mathrm{median}\left(|Q_{P_1}(f) - I_s(f)|, \ldots, |Q_{P_r}(f) - I_s(f)|\right) \\
&\leq \sup_{\substack{f \in B \\ \|f\|_B \leq 1}} \mathrm{median}\left(\|f\|_B\, e^{\mathrm{wor}}(Q_{P_1}; B), \ldots, \|f\|_B\, e^{\mathrm{wor}}(Q_{P_r}; B)\right) \\
&= \mathrm{median}\left(e^{\mathrm{wor}}(Q_{P_1}; B), \ldots, e^{\mathrm{wor}}(Q_{P_r}; B)\right),
\end{aligned}
$$

where the first inequality follows from Jensen's inequality for medians, see [42] and [23, Lemma 2.6].

Following an argument similar to that in [47, Proposition 2.1], the median of $e^{\mathrm{wor}}(Q_{P_1}; B), \ldots, e^{\mathrm{wor}}(Q_{P_r}; B)$ is larger than ϵ if and only if the number of point sets that satisfy $e^{\mathrm{wor}}(Q_{P_i}; B) > \epsilon$ is greater than or equal to $(r+1)/2$. Considering the Bernoulli scheme in which $e^{\mathrm{wor}}(Q_{P_i}; B) \leq \epsilon$ is interpreted as the "success" in the i-th trial, we obtain

$$
\begin{aligned}
&\Pr\left[\mathrm{median}\left(e^{\mathrm{wor}}(Q_{P_1}; B), \ldots, e^{\mathrm{wor}}(Q_{P_r}; B)\right) > \epsilon\right] \\
&= \sum_{i=(r+1)/2}^{r} \binom{r}{i} \left(\Pr\left[e^{\mathrm{wor}}(Q_P; B) > \epsilon\right]\right)^i \left(1 - \Pr\left[e^{\mathrm{wor}}(Q_P; B) > \epsilon\right]\right)^{r-i} \\
&\leq \sum_{i=(r+1)/2}^{r} \binom{r}{i} \delta^i \leq \delta^{(r+1)/2} \sum_{i=(r+1)/2}^{r} \binom{r}{i} = 2^{r-1}\delta^{(r+1)/2}.
\end{aligned}
$$

By combining this with the earlier bound on $e^{\mathrm{wor}}(M_r; B)$, we complete the proof. □

As already pointed out in [23, Remark 2.7], since we assume $0 < \delta < 1/4$, the probability given in Proposition 3.2 converges to 1 exponentially fast as r increases. Therefore it is clear that, with the help of taking the median, increasing r amplifies the probability of success and what we need to do is to prove that most of the point sets in S are "good" in the sense of (3.3). To prove such an existence, probably the

most straightforward way is to apply Markov inequality which ensures that ϵ can be given by

$$(3.4) \qquad \epsilon = \frac{1}{\delta} \mathbb{E}_{P \in S} \left[e^{\mathrm{wor}}(Q_P; B) \right],$$

so that it suffices to evaluate the expectation of the worst-case error with respect to $P \in S$. However, it is generally hard to show a rate of convergence better than N^{-1} by (3.4). In such situations, together with Markov inequality, we shall use Jensen's inequality

$$(3.5) \qquad \phi\left(\sum_n a_n\right) \leq \sum_n \phi(a_n),$$

which holds for any sequence of non-negative real numbers $(a_n)_n$ and concave function $\phi : [0, \infty) \to [0, \infty)$, see [10, Section 2.3]. Then ϵ can be given by

$$(3.6) \qquad \epsilon = \inf_{\phi \in \Phi} \phi^{-1}\left(\frac{1}{\delta} \mathbb{E}_{P \in S}\left[\phi(e^{\mathrm{wor}}(Q_P; B))\right]\right),$$

for a family Φ of concave functions ϕ.

3.2. Universality. In order to demonstrate the universality of our median QMC integration, we prove upper bounds on the expected worst-case error (in the sense of (3.4) or (3.6)) in function spaces with different smoothness classes, which constitutes the main contribution of this paper. We consider three weighted function spaces with different smoothness classes as our target Banach space. By following the seminal work by Sloan and Woźniakowski [61], we model the relative importance of a subset of variables $\boldsymbol{x}_u = (x_j)_{j \in u}$ by introducing a set of weight parameters $\boldsymbol{\gamma} = (\gamma_u)_{u \subseteq 1:s}$ with $\gamma_u \in [0, 1]$ for the first two spaces with finite smoothness. For the third one, we follow [64, 8] and introduce a sequence $u_1 \geq u_2 \geq \cdots > 0$ to define a weighted space of infinitely many times differentiable functions. For each function space, we derive the possible values of the quantity ϵ for a given $\delta \in (0, 1)$, and then derive sufficient conditions on the weights for the obtained values to be bounded independently of the dimension. At the end of each sub-subsection, we include brief comments for comparison on what is known in the literature.

3.2.1. Weighted Sobolev space of first order. The first function space, studied for instance in [61, 13], requires the minimum differentiability for integrands among the three spaces we consider.

DEFINITION 3.3 (weighted Sobolev space of first order). *For a set of weights $\boldsymbol{\gamma}$, the weighted Sobolev space of first order, denoted by $\mathcal{F}_{s,1,\boldsymbol{\gamma}}^{\mathrm{sob}}$, is a Banach space with the norm*

$$\|f\|_{s,1,\boldsymbol{\gamma}}^{\mathrm{sob}} := \sum_{u \subseteq 1:s} \gamma_u^{-1} \int_{[0,1]^{|u|}} \left| \frac{\partial^{|u|}}{\partial \boldsymbol{x}_u} f(\boldsymbol{x}_u, \mathbf{1}) \right| \, \mathrm{d}\boldsymbol{x}_u,$$

where $(\boldsymbol{x}_u, \mathbf{1})$ denotes the vector $\boldsymbol{y} \in [0, 1]^s$ such that $y_j = x_j$ if $j \in u$ and $y_j = 1$ otherwise. For any subset $u \subseteq 1{:}s$ with $\gamma_u = 0$, we assume that the corresponding integral over $[0, 1]^{|u|}$ equals 0 and we set $0/0 = 0$.

It is well known that, for any point set P, the worst-case error in $\mathcal{F}_{s,1,\boldsymbol{\gamma}}^{\mathrm{sob}}$ is bounded above by the weighted star discrepancy:

$$e^{\mathrm{wor}}(Q_P; \mathcal{F}_{s,1,\boldsymbol{\gamma}}^{\mathrm{sob}}) \leq D_{\boldsymbol{\gamma}}^*(P) := \sup_{\boldsymbol{y} \in [0,1]^s} \max_{\emptyset \neq u \subseteq 1:s} \gamma_u \left| \Delta_P(\boldsymbol{y}_u, \mathbf{1}) \right|,$$

where Δ_P is the local discrepancy function defined by

$$\Delta_P(\boldsymbol{y}) = \frac{1}{|P|} \sum_{\boldsymbol{x} \in P} \mathbf{1}_{\boldsymbol{x} \in [\mathbf{0}, \boldsymbol{y})} - \prod_{j=1}^{s} y_j,$$

with $[\mathbf{0}, \boldsymbol{y}) = [0, y_1) \times \cdots \times [0, y_s)$ being the anchored axis-parallel box, see [61, Section 3]. According to [15, Corollary 10.16], if P is a digital net over $\mathbb{F}_b$ with square generating matrices $C_1, \ldots, C_s \in \mathbb{F}_b^{m \times m}$, the weighted star discrepancy is bounded above by

$$D_{\boldsymbol{\gamma}}^*(P) \leq \sum_{\emptyset \neq u \subseteq 1:s} \gamma_u \left[1 - \left(1 - \frac{1}{b^m} \right)^{|u|} + \sum_{\substack{\boldsymbol{k}_u \in P_{u,0}^\perp \\ \forall j \in u, k_j < b^m}} \prod_{j \in u} \tilde{r}_b(k_j) \right],$$

where $\tilde{r}_b : \mathbb{N}_0 \to \mathbb{R}_{\geq 0}$ is defined by

$$\tilde{r}_b(k) = \begin{cases} 1 & \text{if } k = 0, \\ b^{-a}(\sin(\pi \kappa_{a-1}/b))^{-2} & \text{if } k = \kappa_0 + \kappa_1 b + \cdots + \kappa_{a-1} b^{a-1} \text{ with } \kappa_{a-1} \neq 0, \end{cases}$$

and, for a non-empty subset $u \subseteq 1:s$, we write

$$P_{u,0}^\perp = \left\{ \boldsymbol{k}_u \in \mathbb{N}_0^{|u|} \setminus \{\mathbf{0}\} \,\middle|\, (\boldsymbol{k}_u, \mathbf{0}) \in P^\perp \right\}.$$

First let S be given by (3.1) with good square generating matrices. The following result holds for this first function space.

THEOREM 3.4. *For a set of weights* $\boldsymbol{\gamma}$, *let* $B = \mathcal{F}_{s,1,\boldsymbol{\gamma}}^{\text{sob}}$. *For* $m \in \mathbb{N}$, *let* S *be a set of linear scrambled digital nets* (3.1) *with* $C_1, \ldots, C_s \in \mathbb{F}_b^{m \times m}$ *all non-singular and satisfying* (2.2). *In Proposition 3.2, the assumption* (3.3) *holds for any* $\delta \in (0,1)$ *and*

$$\epsilon = \frac{3}{\delta b^m} \sum_{\emptyset \neq u \subseteq 1:s} \gamma_u \prod_{j \in u} \left[1 + m \frac{b(b+1)}{3} j \log_b(j+b) \right].$$

Proof. We use (3.4) to prove the theorem. In this proof, we denote the left upper $m \times m$ submatrices of $L_1 C_1, \ldots, L_s C_s$ by $(L_1 C_1)^{(m)}, \ldots, (L_s C_s)^{(m)}$, respectively. It follows from the linearity of expectation that

$$\mathbb{E}\left[e^{\text{wor}}(Q_{P(L_1 C_1, \ldots, L_s C_s)}; \mathcal{F}_{s,1,\boldsymbol{\gamma}}^{\text{sob}}) \right]$$

$$\leq \mathbb{E}\left[D_{\boldsymbol{\gamma}}^*(P(L_1 C_1, \ldots, L_s C_s)) \right]$$

$$= \mathbb{E}\left[D_{\boldsymbol{\gamma}}^*(P((L_1 C_1)^{(m)}, \ldots, (L_s C_s)^{(m)})) \right]$$

$$\quad + \mathbb{E}\left[D_{\boldsymbol{\gamma}}^*(P(L_1 C_1, \ldots, L_s C_s)) - D_{\boldsymbol{\gamma}}^*(P((L_1 C_1)^{(m)}, \ldots, (L_s C_s)^{(m)})) \right]$$

$$\leq \sum_{\emptyset \neq u \subseteq 1:s} \gamma_u \mathbb{E}\left[1 - \left(1 - \frac{1}{b^m} \right)^{|u|} + \sum_{\substack{\boldsymbol{k}_u \in P_{u,0}^\perp((L_1 C_1)^{(m)}, \ldots, (L_s C_s)^{(m)}) \\ \forall j \in u, k_j < b^m}} \prod_{j \in u} \tilde{r}_b(k_j) \right]$$

$$(3.7) \qquad + \mathbb{E}\left[D_{\boldsymbol{\gamma}}^*(P(L_1 C_1, \ldots, L_s C_s)) - D_{\boldsymbol{\gamma}}^*(P((L_1 C_1)^{(m)}, \ldots, (L_s C_s)^{(m)})) \right].$$

Regarding the expectations for the first term on the rightmost side above, since we have

$$P^\perp((L_1 C_1)^{(m)}, \ldots, (L_s C_s)^{(m)}) \cap \{0, \ldots, b^m - 1\}^s$$
$$= P^\perp(L_1 C_1, \ldots, L_s C_s) \cap \{0, \ldots, b^m - 1\}^s,$$

we can apply the first statement of Lemma 2.17 to obtain

$$\mathbb{E}\left[\sum_{\substack{\boldsymbol{k}_u \in P_{u,0}^\perp((L_1 C_1)^{(m)}, \ldots, (L_s C_s)^{(m)}) \\ \forall j \in u, k_j < b^m}} \prod_{j \in u} \tilde{r}_b(k_j) \right]$$

$$= \sum_{\substack{\boldsymbol{k}_u \in \mathbb{N}_0^{|u|} \setminus \{\boldsymbol{0}\} \\ \forall j \in u, k_j < b^m}} \Pr\left[(\boldsymbol{k}_u, \boldsymbol{0}) \in P^\perp(L_1 C_1, \ldots, L_s C_s)\right] \prod_{j \in u} \tilde{r}_b(k_j)$$

$$= \sum_{\emptyset \neq v \subseteq u} \sum_{\substack{\boldsymbol{k}_v \in \mathbb{N}^{|v|} \\ \forall j \in v, k_j < b^m}} \Pr\left[(\boldsymbol{k}_v, \boldsymbol{0}) \in P^\perp(L_1 C_1, \ldots, L_s C_s)\right] \prod_{j \in v} \tilde{r}_b(k_j)$$

$$\leq \sum_{\emptyset \neq v \subseteq u} \left(\frac{b}{b-1}\right)^{|v|-1} b^{-m+t_v} \sum_{\substack{\boldsymbol{k}_v \in \mathbb{N}^{|v|} \\ \forall j \in v, k_j < b^m}} \prod_{j \in v} \tilde{r}_b(k_j)$$

$$= \sum_{\emptyset \neq v \subseteq u} \left(\frac{b}{b-1}\right)^{|v|-1} b^{-m+t_v} \left(m \frac{b^2-1}{3b}\right)^{|v|} \leq \frac{1}{b^m} \sum_{\emptyset \neq v \subseteq u} b^{t_v} \left(m \frac{b+1}{3}\right)^{|v|},$$

where we used the result on a sum of $\tilde{r}_b(k)$, shown in [13, Lemma 2.2], for the last equality. By using a bound (2.2) on the t-value of the projected point set, we have

$$\mathbb{E}\left[D_{\boldsymbol{\gamma}}^*(P((L_1 C_1)^{(m)}, \ldots, (L_s C_s)^{(m)})) \right]$$

$$\leq \sum_{\emptyset \neq u \subseteq 1:s} \gamma_u \left[1 - \left(1 - \frac{1}{b^m}\right)^{|u|} + \frac{1}{b^m} \sum_{\emptyset \neq v \subseteq u} b^{t_v} \left(m \frac{b+1}{3}\right)^{|v|} \right]$$

$$\leq \sum_{\emptyset \neq u \subseteq 1:s} \gamma_u \left[1 - \frac{1}{b^m} - \left(1 - \frac{1}{b^m}\right)^{|u|} + \frac{1}{b^m} \sum_{v \subseteq u} \left(m \frac{b(b+1)}{3}\right)^{|v|} \prod_{j \in v} j \log_b(j+b) \right]$$

$$= \sum_{\emptyset \neq u \subseteq 1:s} \gamma_u \left[1 - \frac{1}{b^m} - \left(1 - \frac{1}{b^m}\right)^{|u|} + \frac{1}{b^m} \prod_{j \in u} \left[1 + m \frac{b(b+1)}{3} j \log_b(j+b)\right] \right]$$

$$\leq \frac{1}{b^m} \sum_{\emptyset \neq u \subseteq 1:s} \gamma_u \left[|u| - 1 + \prod_{j \in u} \left[1 + m \frac{b(b+1)}{3} j \log_b(j+b)\right] \right]$$

$$\leq \frac{2}{b^m} \sum_{\emptyset \neq u \subseteq 1:s} \gamma_u \prod_{j \in u} \left[1 + m \frac{b(b+1)}{3} j \log_b(j+b)\right],$$

where the third inequality follows from the fact that $1 - \frac{1}{b^m} - \left(1 - \frac{1}{b^m}\right)^{|u|} \leq \frac{|u|-1}{b^m}$ holds for any non-empty $u \subseteq 1:d$, which itself can be proven by an induction on the

cardinality $|u|$, and the last inequality is obtained by noticing that

$$|u| - 1 \le 2^{|u|} = \prod_{j \in u}(1 + 1) \le \prod_{j \in u}\left[1 + m\frac{b(b+1)}{3}j\log_b(j+b)\right].$$

For the second term of (3.7), by denoting the k-th points of $P(L_1C_1, \ldots, L_sC_s)$ and $P((L_1C_1)^{(m)}, \ldots, (L_sC_s)^{(m)})$ by $\boldsymbol{x}_k$ and $\tilde{\boldsymbol{x}}_k$, respectively, we have

$$\tilde{x}_{k,j} - x_{k,j} \le \frac{1}{b^m},$$

for all $1 \le j \le s$, so that, by applying [15, Theorem 3.15], we obtain

$$\mathbb{E}\left[D^*_{\boldsymbol{\gamma}}(P(L_1C_1, \ldots, L_sC_s)) - D^*_{\boldsymbol{\gamma}}(P((L_1C_1)^{(m)}, \ldots, (L_sC_s)^{(m)}))\right]$$
$$\le \frac{1}{b^m}\sum_{\emptyset \ne u \subseteq 1:s}\gamma_u|u| \le \frac{1}{b^m}\sum_{\emptyset \ne u \subseteq 1:s}\gamma_u\prod_{j \in u}\left[1 + m\frac{b(b+1)}{3}j\log_b(j+b)\right].$$

Now the result of the theorem follows from (3.4). □

Using [30, Lemma 3], we can derive a condition under which the worst-case error is bounded independently of the dimension s in the case of product weights, i.e., the case where, for any subset $u \subseteq 1:s$, the weight γ_u is given by $\prod_{j \in u}\gamma_j$ with a one-dimensional sequence $\gamma_1, \gamma_2, \ldots$.

COROLLARY 3.5. *For product weights satisfying $\sum_{j=1}^{\infty}\gamma_j j\log_b(j+b) < \infty$, the probabilistic worst-case error bound in Theorem 3.4 is bounded independently of s by*

$$C_{(\gamma_1,\gamma_2,\ldots),\lambda}b^{-(1-\lambda)m},$$

where $\lambda > 0$ is arbitrarily small, and $C_{(\gamma_1,\gamma_2,\ldots),\lambda}$ is a positive constant that satisfies $\lim_{\lambda \to 0^+}C_{(\gamma_1,\gamma_2,\ldots),\lambda} = \infty$.

Next let S be given by (3.2), for which the associated result is as follows. By replacing the first statement of Lemma 2.17 with the second statement of Lemma 2.20, the proof is almost identical to that of Theorem 3.4, so we omit it.

THEOREM 3.6. *For a set of weights $\boldsymbol{\gamma}$, let $B = \mathcal{F}^{\text{sob}}_{s,1,\boldsymbol{\gamma}}$. For $m \in \mathbb{N}$, let S be a set of polynomial lattices (3.2). In Proposition 3.2, the assumption (3.3) holds for any $\delta \in (0,1)$ and*

$$\epsilon = \frac{3}{\delta(b^m - 1)}\sum_{\emptyset \ne u \subseteq 1:s}\gamma_u\left(1 + m\frac{b^2 - 1}{3}\right)^{|u|}.$$

Following an argument similar to that used in [15, Corollaries 5.45 & 10.30], we can derive a weaker condition than that presented above in Corollary 3.5 under which the worst-case error is bounded independently of the dimension s.

COROLLARY 3.7. *For product weights satisfying $\sum_{j=1}^{\infty}\gamma_j < \infty$, the probabilistic worst-case error bound given in Theorem 3.6 is bounded independently of s with a form similar to that in Corollary 3.5.*

For the unweighted Sobolev space with the slightly different norm

$$\left(\sum_{u \subseteq 1:s}\int_{[0,1]^{|u|}}\left|\frac{\partial^{|u|}}{\partial\boldsymbol{x}_u}f(\boldsymbol{x}_u,\boldsymbol{0})\right|^2\mathrm{d}\boldsymbol{x}_u\right)^{1/2},$$

a lower bound on the worst-case error of order $(\log N)^{(s-1)/2}/N$ is proven, for instance, in [10, Theorem 2.5], which also applies to $\mathcal{F}^{\mathrm{sob}}_{s,1,\boldsymbol{\gamma}}$ with leading weight $\gamma_{\{1,\ldots,s\}} \neq 0$. Recall that the number of points in Theorems 3.4 and 3.6 is $N = b^m$. Then, we see that, with fixed values of δ and b, our obtained probabilistic worst-case error bounds are of order $(\log N)^s/N$, which is nearly optimal. Regarding the dimension-independence of the weighted star discrepancy, it was proven in [13] that the sufficient condition $\sum_{j=1}^{\infty} \gamma_j < \infty$ holds for polynomial lattices with finite precision constructed by component-by-component algorithms. In addition, a non-constructive existence result was shown in [1] that a weaker condition $\sum_{j=1}^{\infty} \exp(-c\gamma_j^{-2}) < \infty$ is sufficient for dimension-independence, but with a slower decay rate in terms of N.

3.2.2. Weighted Sobolev space of high order. The second function space allows for higher order differentiability than the first space, which has been introduced in the context of partial differential equations with random coefficients [11].

DEFINITION 3.8 (weighted Sobolev space of high order). *For $\alpha \in \mathbb{N}$, $\alpha \geq 2$, $1 \leq q \leq \infty$ and a set of weights $\boldsymbol{\gamma}$, the weighted Sobolev space of order α, denoted by $\mathcal{F}^{\mathrm{sob}}_{s,\alpha,\boldsymbol{\gamma},q}$, is a Banach space with the norm*

$$\|f\|^{\mathrm{sob}}_{s,\alpha,\boldsymbol{\gamma},q} :=$$

$$\sup_{u\subseteq 1:s} \gamma_u^{-1} \left(\sum_{v\subseteq u} \sum_{\boldsymbol{\tau}_{u\setminus v}\in\{1,\ldots,\alpha\}^{|u\setminus v|}} \int_{[0,1)^{|v|}} \left| \int_{[0,1)^{s-|v|}} f^{(\boldsymbol{\tau}_{u\setminus v},\boldsymbol{\alpha}_v,\boldsymbol{0})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_{-v} \right|^q \mathrm{d}\boldsymbol{x}_v \right)^{1/q},$$

where $(\boldsymbol{\tau}_{u\setminus v},\boldsymbol{\alpha}_v,\boldsymbol{0})$ denotes the vector $\boldsymbol{h} \in \mathbb{N}_0^s$ such that $h_j = \tau_j$ if $j \in u \setminus v$, $h_j = \alpha$ if $j \in v$, and $h_j = 0$ otherwise, and $f^{(\boldsymbol{\tau}_{u\setminus v},\boldsymbol{\alpha}_v,\boldsymbol{0})}$ denotes the mixed derivative of order $(\boldsymbol{\tau}_{u\setminus v},\boldsymbol{\alpha}_v,\boldsymbol{0})$ of f. For any subset $u \subseteq 1:s$ with $\gamma_u = 0$, we assume that all the corresponding integrals over $[0,1]^{|v|}$ with $v \subseteq u$ equal 0 and we set $0/0 = 0$.

For a digital net P, the worst-case error in $\mathcal{F}^{\mathrm{sob}}_{s,\alpha,\boldsymbol{\gamma},q}$ has been analyzed through Walsh analysis for smooth functions [3, 4], where the Dick weight, a generalization of the NRT weight, plays a central role.

DEFINITION 3.9 (Dick weight). *Let $\alpha \in \mathbb{N}$, $\alpha \geq 2$. For $k \in \mathbb{N}$, we denote its b-adic expansion by*

$$k = \kappa_1 b^{c_1-1} + \kappa_2 b^{c_2-1} + \cdots + \kappa_v b^{c_v-1},$$

with $\kappa_1,\ldots,\kappa_v \in \{1,\ldots,b-1\}$ and $c_1 > c_2 > \cdots > c_v > 0$. The Dick weight $\mu_\alpha : \mathbb{N}_0 \to \mathbb{N}_0$ is defined by $\mu_\alpha(0) = 0$ and

$$\mu_\alpha(k) = \sum_{i=1}^{\min(\alpha,v)} c_i.$$

Moreover, in the case of vectors in $\mathbb{N}_0^s$, we define

$$\mu_\alpha(\boldsymbol{k}) = \sum_{j=1}^{s} \mu_\alpha(k_j).$$

According to [11, Theorem 3.5], for any digital net P, the worst-case error in $\mathcal{F}^{\mathrm{sob}}_{s,\alpha,\boldsymbol{\gamma},q}$ is bounded independently of the parameter q by

$$(3.8) \qquad e^{\mathrm{wor}}(Q_P; \mathcal{F}^{\mathrm{sob}}_{s,\alpha,\boldsymbol{\gamma},q}) \leq \sum_{\emptyset \neq u \subseteq 1:s} \gamma_u C_\alpha^{|u|} \sum_{\boldsymbol{k}_u \in P_u^\perp} b^{-\mu_\alpha(\boldsymbol{k}_u)},$$

where

$$C_\alpha = \left(1 + \frac{1}{b} + \frac{1}{b(b+1)}\right)^{\alpha-2} \left(3 + \frac{2}{b} + \frac{2b+1}{b-1}\right) \max\left(\frac{2}{(2\sin\frac{\pi}{b})^\alpha} \max_{1\leq\tau<\alpha} \frac{1}{(2\sin\frac{\pi}{b})^\tau}\right),$$

and, for a non-empty subset $u \subseteq 1{:}s$, we write

$$P_u^\perp = \left\{ \boldsymbol{k}_u \in \mathbb{N}^{|u|} \,\Big|\, (\boldsymbol{k}_u, \boldsymbol{0}) \in P^\perp \right\}.$$

Here we point out a slight difference from $P_{u,0}^\perp$, which has been introduced and used in section 3.2.1.

First let S be given by (3.1) with good square generating matrices. The following result holds for the second function space.

THEOREM 3.10. *For* $\alpha \in \mathbb{N}$, $\alpha \geq 2$, $1 \leq q \leq \infty$ *and a set of weights* $\boldsymbol{\gamma}$, *let* $B = \mathcal{F}_{s,\alpha,\gamma,q}^{\mathrm{sob}}$. *For* $m \in \mathbb{N}$, *let* S *be a set of linearly scrambled digital nets* (3.1) *with* $C_1, \ldots, C_s \in \mathbb{F}_b^{m \times m}$ *all non-singular and satisfying* (2.2). *In Proposition* 3.2, *the assumption* (3.3) *holds for any* $\delta \in (0,1)$ *and*

$$\epsilon = \inf_{\lambda \in (1/\alpha, 1]} \left(\frac{1}{\delta b^m} \sum_{\emptyset \neq u \subseteq 1{:}s} \gamma_u^\lambda \left(\frac{b^2}{b-1}\right)^{|u|} C_\alpha^{\lambda|u|} A_{\alpha,\lambda}^{|u|} \prod_{j \in u} j \log_b(j+b)\right)^{1/\lambda},$$

with

$$A_{\alpha,\lambda} := \sum_{k=1}^\infty b^{-\lambda\mu_\alpha(k)} = \sum_{\tau=1}^{\alpha-1} \prod_{i=1}^\tau \frac{b-1}{b^{\lambda i} - 1} + \frac{b^{\lambda\alpha}-1}{b^{\lambda\alpha}-b} \prod_{i=1}^\alpha \frac{b-1}{b^{\lambda i} - 1}.$$

Proof. We use (3.6) to prove the theorem. For a family of concave functions, let $\phi(x) = x^\lambda$ with $0 < \lambda \leq 1$. By applying Jensen's inequality (3.5), the linearity of expectation and Lemma 2.17, we have

$$\mathbb{E}\left[\phi\left(e^{\mathrm{wor}}(Q_{P(L_1 C_1, \ldots, L_s C_s)}; \mathcal{F}_{s,\alpha,\gamma,q}^{\mathrm{sob}})\right)\right]$$

$$\leq \mathbb{E}\left[\sum_{\emptyset \neq u \subseteq 1{:}s} \gamma_u^\lambda C_\alpha^{\lambda|u|} \sum_{\boldsymbol{k}_u \in P_u^\perp(L_1 C_1, \ldots, L_s C_s)} b^{-\lambda\mu_\alpha(\boldsymbol{k}_u)}\right]$$

$$= \sum_{\emptyset \neq u \subseteq 1{:}s} \gamma_u^\lambda C_\alpha^{\lambda|u|} \mathbb{E}\left[\sum_{\boldsymbol{k}_u \in P_u^\perp(L_1 C_1, \ldots, L_s C_s)} b^{-\lambda\mu_\alpha(\boldsymbol{k}_u)}\right]$$

$$= \sum_{\emptyset \neq u \subseteq 1{:}s} \gamma_u^\lambda C_\alpha^{\lambda|u|} \sum_{\boldsymbol{k}_u \in \mathbb{N}^{|u|}} b^{-\lambda\mu_\alpha(\boldsymbol{k}_u)} \Pr\left[\boldsymbol{k}_u \in P_u^\perp(L_1 C_1, \ldots, L_s C_s)\right]$$

$$\leq \frac{1}{b^m} \sum_{\emptyset \neq u \subseteq 1{:}s} \gamma_u^\lambda C_\alpha^{\lambda|u|} \left[\left(\frac{b}{b-1}\right)^{|u|-1} b^{t_u} \sum_{\substack{\boldsymbol{k}_u \in \mathbb{N}^{|u|} \\ \forall j \in u, k_j < b^m}} b^{-\lambda\mu_\alpha(\boldsymbol{k}_u)} + \sum_{\substack{\boldsymbol{k}_u \in \mathbb{N}^{|u|} \\ \exists j \in u, k_j \geq b^m}} b^{-\lambda\mu_\alpha(\boldsymbol{k}_u)}\right]$$

$$\leq \frac{1}{b^m} \sum_{\emptyset \neq u \subseteq 1{:}s} \gamma_u^\lambda C_\alpha^{\lambda|u|} \left(\frac{b}{b-1}\right)^{|u|} b^{t_u} \sum_{\boldsymbol{k}_u \in \mathbb{N}^{|u|}} b^{-\lambda\mu_\alpha(\boldsymbol{k}_u)}. \qquad \blacksquare$$

Here we know from [22, Lemma 7] that the inner sum over $\boldsymbol{k}_u$ is finite for any $\lambda \in (1/\alpha, 1]$ and equals $A_{\alpha,\lambda}^{|u|}$. By using a bound (2.2) on the t-value of the projected point set, we have

$$\mathbb{E}\left[\phi\left(e^{\mathrm{wor}}(Q_{P(L_1 C_1, \ldots, L_s C_s)}; \mathcal{F}_{s,\alpha,\gamma,q}^{\mathrm{sob}})\right)\right]$$

$$\leq \frac{1}{b^m} \sum_{\emptyset \neq u \subseteq 1:s} \gamma_u^\lambda \left(\frac{b^2}{b-1} \right)^{|u|} C_\alpha^{\lambda|u|} A_{\alpha,\lambda}^{|u|} \prod_{j \in u} j \log_b(j+b).$$

Applying the inverse map ϕ^{-1} in (3.6), the proof is completed. □

In the case of product weights, by using the elementary inequality $1 + x \leq \exp(x)$, the probabilistic worst-case error bound shown in Theorem 3.10 is bounded by

$$\epsilon \leq \inf_{\lambda \in (1/\alpha, 1]} \left(\frac{1}{\delta b^m} \exp \left(\left(\frac{b^2}{b-1} \right) C_\alpha^\lambda A_{\alpha,\lambda} \sum_{j=1}^s \gamma_j^\lambda j \log_b(j+b) \right) \right)^{1/\lambda},$$

which leads to the following corollary.

COROLLARY 3.11. *If there exists $\lambda' \in (1/\alpha, 1]$ such that the product weights satisfy $\sum_{j=1}^\infty \gamma_j^{\lambda'} j \log_b(j+b) < \infty$, the probabilistic worst-case error given in Theorem 3.10 is bounded independently of s and decays at the rate of $O(b^{-m/\lambda'})$.*

Again a similar result holds for the set (3.2). Although the proof is also similar to that of Theorem 3.10, we give its sketch for the sake of completeness.

THEOREM 3.12. *For $\alpha \in \mathbb{N}$, $\alpha \geq 2$, $1 \leq q \leq \infty$ and a set of weights $\boldsymbol{\gamma}$, let $B = \mathcal{F}_{s,\alpha,\boldsymbol{\gamma},q}^{\mathrm{sob}}$. For $m \in \mathbb{N}$, let S be a set of polynomial lattices (3.2). In Proposition 3.2, the assumption (3.3) holds for any $\delta \in (0,1)$ and*

$$\epsilon = \inf_{\substack{\lambda \in (1/\alpha, 1] \\ \tau \in (0, \lambda - 1/\alpha)}} \left(\frac{3}{(b^m-1)\delta \tau e \log b} \sum_{\emptyset \neq u \subseteq 1:s} \gamma_u^\lambda C_\alpha^{\lambda|u|} A_{\alpha,\lambda-\tau}^{|u|} \right)^{1/\lambda},$$

where the constant $A_{\alpha,\lambda-\tau}$ is the same as in Theorem 3.10.

Proof. Again let $\phi(x) = x^\lambda$ with $0 < \lambda \leq 1$ in (3.6). By applying Jensen's inequality (3.5), the linearity of expectation and Lemma 2.20, we have

$$\mathbb{E}\left[\phi\left(e^{\mathrm{wor}}(Q_{P(p,\boldsymbol{g})}; \mathcal{F}_{s,\alpha,\boldsymbol{\gamma},q}^{\mathrm{sob}})\right)\right] \leq \sum_{\emptyset \neq u \subseteq 1:s} \gamma_u^\lambda C_\alpha^{\lambda|u|} \sum_{\boldsymbol{k}_u \in \mathbb{N}^{|u|}} b^{-\lambda \mu_\alpha(\boldsymbol{k}_u)} \Pr\left[\boldsymbol{k}_u \in P_u^\perp(p,\boldsymbol{g})\right]$$

$$\leq \frac{3}{b^m - 1} \sum_{\emptyset \neq u \subseteq 1:s} \gamma_u^\lambda C_\alpha^{\lambda|u|} \sum_{\boldsymbol{k}_u \in \mathbb{N}^{|u|}} b^{-\lambda \mu_\alpha(\boldsymbol{k}_u)} \mu_1(\boldsymbol{k})$$

$$\leq \frac{3}{(b^m-1)\tau e \log b} \sum_{\emptyset \neq u \subseteq 1:s} \gamma_u^\lambda C_\alpha^{\lambda|u|} \sum_{\boldsymbol{k}_u \in \mathbb{N}^{|u|}} b^{-(\lambda-\tau)\mu_\alpha(\boldsymbol{k}_u)}$$

$$= \frac{3}{(b^m-1)\tau e \log b} \sum_{\emptyset \neq u \subseteq 1:s} \gamma_u^\lambda C_\alpha^{\lambda|u|} A_{\alpha,\lambda-\tau}^{|u|},$$

if $1/\alpha < \lambda \leq 1$ and $0 < \tau < \lambda - 1/\alpha$. Here, from the second to the third line, we used $\mu_1(\boldsymbol{k}) \leq \mu_\alpha(\boldsymbol{k})$ (cf. Definitions 2.12 and 3.9) and then applied the elementary inequality $x \leq b^{\tau x}/(\tau e \log b)$, which holds for any $\tau \in (0,1)$ and $x > 0$, to the case $x = \mu_\alpha(\boldsymbol{k})$. Applying the inverse map ϕ^{-1} in (3.6), the proof is completed. □

We note that a quite similar result has been shown recently in [23, Theorem 3.7] for higher order polynomial lattices, i.e., polynomial lattices with a modulus of larger degree $n \geq m$. This subtle difference between point sets requires including the additional parameter τ in the present analysis, especially to deal with $\mu_1(\boldsymbol{k})$ appearing in the first statement of Lemma 2.20.

COROLLARY 3.13. *If there exists $\lambda' \in (1/\alpha, 1]$ such that the product weights satisfy $\sum_{j=1}^{\infty} \gamma_j^{\lambda'} < \infty$, the probabilistic worst-case error given in Theorem 3.12 is bounded independently of s and decays at the rate of $O(b^{-m/\lambda'})$.*

Again, this sufficient condition compares favorably with that in Corollary 3.11.

A lower error estimate by Sharygin [59], originally for unweighted function spaces, applies to the weighed spaces $\mathcal{F}_{s,\alpha,\gamma,q}^{\mathrm{sob}}$, and the convergence rate of the worst-case error cannot be better than $N^{-\alpha}$. As the number of points is $N = b^m$ in Theorems 3.10 and 3.12, with fixed values of δ and b, our obtained probabilistic worst-case error bounds are of order $N^{-1/\lambda}$ for $\lambda \in (1/\alpha, 1]$, which is almost optimal as $\lambda \searrow 1/\alpha$ but with growing constants. Regarding the dimension-independence of the worst-case error, the sufficient condition same as that in Corollary 3.13 can be derived from [11, Theorem 3.10] for interlaced polynomial lattices. This condition was also shown to hold for extrapolated polynomial lattices in [9, Corollary 3.7].

3.2.3. Weighted space of infinitely smooth functions. The last, third function space contains only infinitely many times differentiable functions, for which it can be expected from [64, 8, 56, 55] that the worst-case error decays faster than any polynomial convergence.

DEFINITION 3.14 (weighted space of infinitely smooth functions). *For a sequence $\boldsymbol{u} = (u_1, u_2, \ldots)$ with $u_1 \geq u_2 \geq \cdots > 0$, the weighted space of infinitely smooth functions, denoted by $\mathcal{F}_{s,\infty,\boldsymbol{u}}$, is a Banach space consisting of $f \in C^{\infty}([0,1]^s)$ with the norm*

$$\|f\|_{\mathcal{F}_{s,\infty,\boldsymbol{u}}} := \sup_{(\alpha_1,\ldots,\alpha_s) \in \mathbb{N}_0^s} \prod_{j=1}^{s} u_j^{-\alpha_j} \int_{[0,1]^s} \left| f^{(\alpha_1,\ldots,\alpha_s)}(\boldsymbol{x}) \right| \, \mathrm{d}\boldsymbol{x}.$$

Similarly to the function space studied in section 3.2.2, the worst-case error in $\mathcal{F}_{s,\infty,\boldsymbol{u}}$ has been analyzed for digital nets though Walsh analysis for infinitely smooth functions [65, 73]. There the modified Dick weight with $\alpha = \infty$ plays a crucial role.

DEFINITION 3.15 (modified infinite Dick weight). *Let a be a real number. For $k \in \mathbb{N}$, we denote its b-adic expansion by*

$$k = \kappa_1 b^{c_1 - 1} + \kappa_2 b^{c_2 - 1} + \cdots + \kappa_v b^{c_v - 1},$$

with $\kappa_1, \ldots, \kappa_v \in \{1, \ldots, b-1\}$ and $c_1 > c_2 > \cdots > c_v > 0$. The modified infinite Dick weight $\mu_{\infty,a} : \mathbb{N}_0 \to \mathbb{R}$ is defined by $\mu_{\infty,a}(0) = 0$ and

$$\mu_{\infty,a}(k) = \sum_{i=1}^{v} (c_i + a) = \mu_{\infty}(k) + av.$$

Moreover, in the case of vectors in $\mathbb{N}_0^s$, for a sequence $\boldsymbol{a} = (a_1, \ldots, a_s) \in \mathbb{R}^s$, we define

$$\mu_{\infty,\boldsymbol{a}}(\boldsymbol{k}) = \sum_{j=1}^{s} \mu_{\infty,a_j}(k_j).$$

According to [64, Lemma 6.5], the worst-case error for any digital net P in $\mathcal{F}_{s,\infty,\boldsymbol{u}}$ is bounded above by

$$(3.9) \qquad e^{\mathrm{wor}}(Q_P; \mathcal{F}_{s,\infty,\boldsymbol{u}}) \leq \sum_{\boldsymbol{k} \in P^{\perp} \setminus \{\boldsymbol{0}\}} b^{-\mu_{\infty,\boldsymbol{a}}(\boldsymbol{k})},$$

where $\boldsymbol{a} = (a_1, a_2, \ldots)$ is a sequence defined by $a_j = -\log_b(K_b m_b^{-1} u_j)$, with $m_b = 2\sin(\pi/b)$,

$$M_b = \begin{cases} 2 & \text{if } b \text{ is even,} \\ 2\sin((b+1)\pi/2b) & \text{if } b \text{ is odd,} \end{cases}$$

and

$$K_b = \begin{cases} 2 & \text{if } b = 2, \\ M_b + bm_b/(b - M_b) & \text{if } b \geq 3. \end{cases}$$

First let S be given by (3.1) with good square generating matrices. Unfortunately, our average argument in the sense of either (3.4) or (3.6) prevents us from capturing a precise convergence behavior (cf. [64, Sections 6.2 & 6.3] and [55, Sections 4 & 5]). The following result only shows that the probabilistic worst-case error decays faster than any polynomial convergence of a fixed degree.

THEOREM 3.16. *For a sequence $\boldsymbol{u} = (u_1, u_2, \ldots)$ with $u_1 \geq u_2 \geq \cdots > 0$, let $B = \mathcal{F}_{s,\infty,\boldsymbol{u}}$. For $m \in \mathbb{N}$, let S be a set of linearly scrambled digital nets (3.1) with $C_1, \ldots, C_s \in \mathbb{F}_b^{m \times m}$ all non-singular and satisfying (2.2). In Proposition 3.2, the assumption (3.3) holds for any $\check{\delta} \in (0, 1)$ and*

$$\epsilon = \inf_{\lambda \in (0,1]} \left(\frac{1}{\check{\delta} b^m} \left[-1 + \prod_{j=1}^{s} \left(1 + \left(\frac{b}{b-1} \right) A_{\infty,\lambda} b^{-\lambda a_j} j \log_b(j+b) \right) \right] \right)^{1/\lambda},$$

where $a_j = -\log_b(K_b m_b^{-1} u_j)$, and $A_{\infty,\lambda}$ is finite for any $\lambda \in (0,1)$ and equals

$$A_{\infty,\lambda} := \sum_{k=1}^{\infty} b^{-\lambda \mu_\infty(k)} = \prod_{\ell=1}^{\infty} \left[1 + \frac{b-1}{b^{\lambda \ell}} \right] - 1.$$

Note that the constant $A_{\infty,\lambda}$ equals $\lim_{\alpha \to \infty} A_{\alpha,\lambda}$, where $A_{\alpha,\lambda}$ is as in Theorem 3.10.

Proof. Let us consider (3.6) with the choice $\phi(x) = x^\lambda$ for $0 < \lambda \leq 1$. Following an argument similar to the proof of Theorem 3.10, we have

$$\mathbb{E} \left[\phi \left(e^{\mathrm{wor}}(Q_{P(L_1 C_1, \ldots, L_s C_s)}; \mathcal{F}_{s,\infty,\boldsymbol{u}}) \right) \right]$$
$$\leq \sum_{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}} \Pr \left[\boldsymbol{k} \in P^\perp(L_1 C_1, \ldots, L_s C_s) \right] b^{-\lambda \mu_{\infty,\boldsymbol{a}}(\boldsymbol{k})}$$
$$\leq \frac{1}{b^m} \sum_{\emptyset \neq u \subseteq 1:s} \left(\frac{b}{b-1} \right)^{|u|} b^{t_u} \sum_{\boldsymbol{k}_u \in \mathbb{N}^{|u|}} b^{-\lambda \mu_{\infty,\boldsymbol{a}}(\boldsymbol{k}_u)}$$
$$\leq \frac{1}{b^m} \sum_{\emptyset \neq u \subseteq 1:s} \left(\frac{b}{b-1} \right)^{|u|} b^{t_u} b^{-\lambda |\boldsymbol{a}_u|_1} \sum_{\boldsymbol{k}_u \in \mathbb{N}^{|u|}} b^{-\lambda \mu_\infty(\boldsymbol{k}_u)}.$$

Here, in a way similar to [8, Proposition 3], we can prove that the inner sum over $\boldsymbol{k}_u$ equals $A_{\infty,\lambda}^{|u|}$. Finally, by using a bound (2.2) on the t-value of the projected point set, we have

$$\mathbb{E} \left[\phi \left(e^{\mathrm{wor}}(Q_{P(L_1 C_1, \ldots, L_s C_s)}; \mathcal{F}_{s,\infty,\boldsymbol{u}}) \right) \right]$$
$$\leq \frac{1}{b^m} \sum_{\emptyset \neq u \subseteq 1:s} \left(\frac{b}{b-1} \right)^{|u|} A_{\infty,\lambda}^{|u|} \prod_{j \in u} b^{-\lambda a_j} j \log_b(j+b)$$

$$= \frac{1}{b^m} \left[-1 + \prod_{j=1}^{s} \left(1 + \left(\frac{b}{b-1} \right) A_{\infty,\lambda} b^{-\lambda a_j} j \log_b(j+b) \right) \right].$$

Applying the inverse map ϕ^{-1} in (3.6) completes the proof.   □

Recalling that $a_j = -\log_b(K_b m_b^{-1} u_j)$, the following result holds:

COROLLARY 3.17. *If the weight sequence $\boldsymbol{u}$ satisfies $\sum_{j=1}^{\infty} u_j^\lambda j \log_b(j+b) < \infty$ for any $\lambda \in (0,1]$, the probabilistic worst-case error ϵ given in Theorem 3.16 is bounded independently of the dimension s and decays faster than any polynomial convergence of a fixed degree.*

For the set (3.2) it is possible to evaluate a convergence behavior more precisely under some conditions on the weights $\boldsymbol{u}$. To do so, as in [8], let us consider a family of concave functions of the form $\phi(x) = b^{-(-\log_b x)^\lambda}$ with $0 < \lambda \le 1$, which maps b^{-x} to b^{-x^λ}. In order to make ϕ concave and unbounded monotonically increasing over the support $[0, \infty)$, let

$$\phi(x) = \begin{cases} 2^{-(-\log_2 x)^\lambda} & \text{if } 0 < x \le \tilde{x}_\lambda, \\ \frac{\lambda(\log_2 \tilde{x}_\lambda)^{\lambda-1}}{e\tilde{x}_\lambda}(x - \tilde{x}_\lambda) + \frac{1}{e} & \text{otherwise}, \end{cases}$$

with $\tilde{x}_\lambda = 2^{-(\log 2)^{1/\lambda}}$ for $b = 2$, and

$$\phi(x) = \begin{cases} b^{-(-\log_b x)^\lambda} & \text{if } 0 < x \le 1/b, \\ \lambda(x - \frac{1}{b}) + \frac{1}{b} & \text{otherwise}, \end{cases}$$

for $b \ge 3$. We set $\phi(0) = 0$ for any base $b \ge 2$.[1] In what follows, we focus on the case $b \ge 3$ and assume $a_1 \ge 0$, i.e., $u_1 \le m_b/K_b$. Note that a similar result without the assumption $a_1 \ge 0$ can be proven in the same way, including the case $b = 2$ as in [8, 64]. Then, for any $\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}$, we have $\mu_{\infty,\boldsymbol{a}}(\boldsymbol{k}) \ge \mu_1(\boldsymbol{k}) \ge 1$ and $b^{-\mu_{\infty,\boldsymbol{a}}(\boldsymbol{k})} \le 1/b$. Now we show the following result.

THEOREM 3.18. *For a sequence $\boldsymbol{u} = (u_1, u_2, \ldots)$ with $m_b/K_b \ge u_1 \ge u_2 \ge \cdots > 0$, let $B = \mathcal{F}_{s,\infty,\boldsymbol{u}}$ and $a_j = -\log_b(K_b m_b^{-1} u_j)$. For $m \in \mathbb{N}$, let S be a set of polynomial lattices (3.2) with $b \ge 3$ and let ϕ be given as above. In Proposition 3.2, the assumption (3.3) holds for any $\delta \in (0,1)$ and the following value of ϵ.*

1. *(unweighted case) If there exists a constant $a \ge 0$ such that $a_1 = a_2 = \cdots = a$,*

$$\epsilon = \inf_{\substack{\lambda \in (1/2,1) \\ \tau \in (0, \min(\lambda, 1/(\log b)))}} \phi^{-1} \left(\frac{3C_{s,\lambda,\tau}}{(b^m - 1)\delta(\tau e \log b)^{1/\tau}} \right),$$

   where

$$C_{s,\lambda,\tau} = \sum_{i=1}^{\infty} \exp\left(2\sqrt{s(b-1)(i+1)} - (i^\lambda - i^\tau) \log b \right).$$

2. *(weighted case) If there exist constants $a, q > 0$ such that $a_j \ge a(j-1)^q$,*

$$\epsilon = \inf_{\substack{\lambda \in ((q+1)/(2q+1),1) \\ \tau \in (0, \min(\lambda, 1/(\log b)))}} \phi^{-1} \left(\frac{3C_{a,q,\lambda,\tau}}{(b^m - 1)\delta(\tau e \log b)^{1/\tau}} \right),$$

---

[1] In [8], there are some typos in the definition of ϕ, which we correct in this paper.

where

$$C_{a,q,\lambda,\tau} = \sum_{i=1}^{\infty} \exp\left(A_{a,q}(i+1)^{(q+1)/(2q+1)} - (i^{\lambda} - i^{\tau})\log b\right)$$

with $A_{a,q} = 1 + (b-1)\left(1 + \frac{\Gamma(1/q)}{qa^{1/q}}\right)$ *and* Γ *being the Gamma function.*

Proof. Similarly to the proof of Theorem 3.12, by using the elementary inequality $x \leq b^{x^{\tau}}/(\tau e \log b)^{1/\tau}$ which holds for any $\tau \in (0, 1/(\log b))$ and $x \geq 1$ instead of $x \leq b^{\tau x}/(\tau e \log b)$, we have

$$\mathbb{E}\left[\phi\left(e^{\mathrm{wor}}(Q_{P(p,\boldsymbol{g})}; \mathcal{F}_{s,\infty,\boldsymbol{u}})\right)\right] \leq \sum_{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}} \phi\left(b^{-\mu_{\infty,\boldsymbol{a}}(\boldsymbol{k})}\right) \Pr\left[\boldsymbol{k} \in P^{\perp}(p,\boldsymbol{g})\right]$$

$$\leq \frac{3}{b^m - 1} \sum_{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}} b^{-(\mu_{\infty,\boldsymbol{a}}(\boldsymbol{k}))^{\lambda}} \mu_1(\boldsymbol{k})$$

$$(3.10) \qquad \leq \frac{3}{(b^m - 1)(\tau e \log b)^{1/\tau}} \sum_{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}} b^{-(\mu_{\infty,\boldsymbol{a}}(\boldsymbol{k}))^{\lambda} + (\mu_{\infty,\boldsymbol{a}}(\boldsymbol{k}))^{\tau}}.$$

For $i \in \mathbb{N}$, let us define

$$\mathrm{vol}_{\boldsymbol{a}}(i) := \left|\{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\} \mid i \leq \mu_{\infty,\boldsymbol{a}}(\boldsymbol{k}) < i+1\}\right|.$$

Then, for any $x \in (0, 1)$, it holds that

$$\mathrm{vol}_{\boldsymbol{a}}(i) = \sum_{\substack{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\} \\ i \leq \mu_{\infty,\boldsymbol{a}}(\boldsymbol{k}) < i+1}} 1 \leq \sum_{\substack{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\} \\ \mu_{\infty,\boldsymbol{a}}(\boldsymbol{k}) < i+1}} 1 \leq \sum_{\substack{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\} \\ \mu_{\infty,\boldsymbol{a}}(\boldsymbol{k}) < i+1}} x^{\mu_{\infty,\boldsymbol{a}}(\boldsymbol{k}) - (i+1)}$$

$$\leq \sum_{\boldsymbol{k} \in \mathbb{N}_0^s} x^{\mu_{\infty,\boldsymbol{a}}(\boldsymbol{k}) - (i+1)} = \frac{1}{x^{i+1}} \prod_{j=1}^{s} \prod_{\ell=1}^{\infty} \left(1 + (b-1)x^{\ell + a_j}\right),$$

where the last equality follows from [64, Eq. (4)]. Taking the natural logarithm of the left-most and right-most sides and using the elementary inequality $\log(1 + x) \leq x$ for $x > 0$, we have

$$\log \mathrm{vol}_{\boldsymbol{a}}(i) \leq -(i+1)\log x + \sum_{j=1}^{s} \sum_{\ell=1}^{\infty} \log\left(1 + (b-1)x^{\ell + a_j}\right)$$

$$\leq -(i+1)\log x + (b-1) \sum_{j=1}^{s} x^{a_j} \sum_{\ell=1}^{\infty} x^{\ell}$$

$$= -(i+1)\log x + \frac{(b-1)x}{1-x} \sum_{j=1}^{s} x^{a_j} \leq -(i+1)\log x - \frac{b-1}{\log x} \sum_{j=1}^{s} x^{a_j}.$$

In what follows, we discuss two different cases separately.

1. (unweighted case) If $a_1 = a_2 = \cdots = a \geq 0$, by choosing

$$x = 1/\exp(\sqrt{s(b-1)/(i+1)}),$$

we obtain

$$\mathrm{vol}_{\boldsymbol{a}}(i) \leq \exp\left(2\sqrt{s(b-1)(i+1)}\right).$$

For any $0 < \tau < \min(\lambda, 1/(\log b))$, the sum over $\boldsymbol{k}$ of (3.10) is bounded by

$$\sum_{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}} b^{-(\mu_{\infty, \boldsymbol{a}}(\boldsymbol{k}))^\lambda + (\mu_{\infty, \boldsymbol{a}}(\boldsymbol{k}))^\tau}$$

$$= \sum_{i=1}^\infty \sum_{\substack{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\} \\ i \leq \mu_{\infty, \boldsymbol{a}}(\boldsymbol{k}) < i+1}} b^{-(\mu_{\infty, \boldsymbol{a}}(\boldsymbol{k}))^\lambda + (\mu_{\infty, \boldsymbol{a}}(\boldsymbol{k}))^\tau} \leq \sum_{i=1}^\infty b^{-i^\lambda + i^\tau} \operatorname{vol}_{\boldsymbol{a}}(i)$$

$$\leq \sum_{i=1}^\infty \exp\left(2\sqrt{s(b-1)(i+1)} - (i^\lambda - i^\tau)\log b\right) = C_{s, \lambda, \tau}.$$

This infinite sum converges when $\lambda > 1/2$ and $0 < \tau < \min(\lambda, 1/(\log b))$.

2. (weighted case) If there exist $a, q > 0$ such that $a_j \geq a(j-1)^q$ for all j, the sum of x^{a_j} is bounded by

$$\sum_{j=1}^s x^{a_j} \leq \sum_{j=1}^s x^{a(j-1)^q} \leq 1 + \int_0^{s-1} x^{at^q}\, dt \leq 1 + \int_0^\infty x^{at^q}\, dt$$

$$= 1 + \frac{\Gamma(1/q)}{qa^{1/q}(-\log x)^{1/q}},$$

where we used the result from [64, Lemma 6.11] for the last equality. By choosing $x = 1/\exp((i+1)^{-q/(2q+1)})$, we obtain

$$\operatorname{vol}_{\boldsymbol{a}}(i) \leq \exp\left(A_{a, q}(i+1)^{(q+1)/(2q+1)}\right),$$

with

$$A_{a, q} = 1 + (b-1)\left(1 + \frac{\Gamma(1/q)}{qa^{1/q}}\right).$$

For any $0 < \tau < \min(\lambda, 1/(\log b))$, the sum over $\boldsymbol{k}$ of (3.10) is bounded by

$$\sum_{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}} b^{-(\mu_{\infty, \boldsymbol{a}}(\boldsymbol{k}))^\lambda + (\mu_{\infty, \boldsymbol{a}}(\boldsymbol{k}))^\tau}$$

$$\leq \sum_{i=1}^\infty \exp\left(A_{a, q}(i+1)^{(q+1)/(2q+1)} - (i^\lambda - i^\tau)\log b\right) = C_{a, q, \lambda, \tau},$$

which converges when $\lambda > (q+1)/(2q+1)$ and $0 < \tau < \min(\lambda, 1/(\log b))$. The result of the theorem follows from (3.6). $\qquad\square$

For the unweighted case, the constant $C_{s, \lambda, \tau}$ depends exponentially on s. For m being large enough such that there exist λ', τ' that satisfy

$$\frac{3C_{s, \lambda', \tau'}}{(b^m - 1)\delta(\tau' e \log b)^{1/\tau'}} \leq \frac{1}{b},$$

we have

$$\epsilon = b^{-\left(\log_b\left((b^m - 1)\delta(\tau' e \log b)^{1/\tau'} / (3C_{s, \lambda', \tau'})\right)\right)^{1/\lambda'}}.$$

This implies that the probabilistic worst-case error asymptotically decays faster than any polynomial convergence. However, due to the exponential dependence of the constant $C_{s, \lambda', \delta'}$ on s, such an asymptotic regime can be only achieved for exponentially large b^m in terms of s. On the contrary, the following holds for the weighted case:

COROLLARY 3.19. *If there exist $a, q > 0$ such that $u_j \leq (m_b/K_b) \exp(-a(j-1)^q)$ for all j, the probabilistic worst-case error ϵ is independent of the dimension s, as given in the second item of Theorem 3.18, and decays super-polynomially as*

$$\epsilon = b^{-\left(\log_b\left((b^m-1)\delta(\tau e \log b)^{1/\tau}\big/(3C_{a,q,\lambda,\tau})\right)\right)^{1/\lambda}},$$

for any $(q+1)/(2q+1) < \lambda \leq 1$ and $0 < \tau < \min(\lambda, 1/(\log b))$.

In the literature, a computable inversion formula for the worst-case error bound (3.9) has been studied in the unweighted cases in [41, 63]. This formula was used in [27] to optimize the linear scrambling of digital nets, and Theorem 3.16 provides partial theoretical support for such optimization. In the weighted cases, a dimension-independent super-polynomial convergence has been established under a slightly milder condition $\liminf_{j\to\infty} \log(u_j^{-1})/j^q > 0$ for some $q > 0$ in [64] by a non-constructive existence proof and in [8] for interlaced polynomial lattices with a carefully chosen interlacing factor. It should be noted that a lower bound on the worst-case error in $\mathcal{F}_{s,\infty,\boldsymbol{u}}$ is currently unknown. Instead, a lower bound in the Walsh space, into which $\mathcal{F}_{s,\infty,\boldsymbol{u}}$ is embedded, is shown in [64, Section 5], and the super-polynomial convergence rates obtained in Theorem 3.18 closely match that result.

4. Numerical experiments. We conclude this paper with numerical experiments. Although our theoretical results in section 3 are obtained for the infinite-precision setting $w = \infty$, we use finite-precision versions with $b = 2$:

$$\{P(L_1 C_1, \ldots, L_s C_s) \mid L_1, \ldots, L_s \in \mathcal{L}_{w=52,m}\},$$

with the generating matrices of Sobol' points $C_1, \ldots, C_s \in \mathbb{F}_2^{m \times m}$, and

$$\{P(p, \boldsymbol{g}, w = 52) \mid p \in \mathbb{P}_m, \boldsymbol{g} \in \mathbb{G}_m^s\},$$

instead of (3.1) and (3.2) for the set S, respectively, throughout this section. The number $w = 52$ not only seems large enough to approximate those results well but also enables us to implement reasonably in the double-precision floating-point format. Moreover, we fix the number of random draws at $r = 15$, whose choice is justified by a rule of thumb given in the following remark.

Remark 4.1. To reach a certain confidence level $1 - \theta$ with small $\theta < 1/4$, we use Proposition 3.2 to obtain $2^{r-1} \delta^{(r+1)/2} \leq \theta$. Let us set $\delta = (4\theta)^{2/r}/4$ to satisfy this inequality. Now, given a total budget of $N = rb^m$ function evaluations, consider maximizing the product $\delta b^m = N(4\theta)^{2/r}/(4r)$ with respect to r. Since the product δb^m commonly appears in the denominator of ϵ in the theorems presented in the previous section, it is a natural quantity to optimize. This maximization leads to $r = -2\log(4\theta)$, which corresponds to $\delta = 1/(4e)$. Using this rule of thumb, we obtain $\theta = 1/(4e^{r/2})$ for a given r. Therefore, choosing $r = 15$ ensures that the worst-case error is stabilized below a guaranteed probability of failure $\theta < 1/7000$. By changing the value of r, one can adjust to different confidence levels.

The numerical results for a series of four different experiments are reported below.

Example 4.2. We begin by examining the distribution of worst-case errors for randomly chosen point sets, here focusing only on the set of polynomial lattices in the weighted Sobolev space of high order $\mathcal{F}_{s,\alpha,\boldsymbol{\gamma},q}^{\mathrm{sob}}$. Note that the worst-case error is bounded as shown in (3.8). For the case of product weights, it follows from [2,

Lemma 1 & Corollary 1] that this bound can be rewritten as

$$\mathcal{S}_{\alpha,\boldsymbol{\gamma}}(P) = -1 + \frac{1}{|P|} \sum_{\boldsymbol{x} \in P} \prod_{j=1}^{s} \left[1 + \gamma_j C_\alpha \omega_\alpha(x_j) \right],$$

with $\omega_\alpha : [0,1) \to \mathbb{R}$ having a concise expression if $b = 2$ and $\alpha = 2$ or 3. We consider the cases $\alpha = 2$ and $\alpha = 3$, and set $m = 8$, $s = 5$, and $\gamma_j = C_\alpha^{-1} j^{-\alpha-1}$ for $j = 1, \ldots, s$. For each value of α, we generate 10^4 polynomial lattices independently and randomly from the set and compute $\mathcal{S}_{\alpha,\boldsymbol{\gamma}}$ for each polynomial lattice. The left panels of Figure 1 display histograms of the resulting values of $\log_2 \mathcal{S}_{\alpha,\boldsymbol{\gamma}}$. For both values of α, the obtained distributions exhibit a right-skewed shape, indicating that while the worst-case error is small for the majority of polynomial lattices, there is a small fraction of "bad" polynomial lattices with significantly larger worst-case errors.

We now evaluate how taking the median can help stabilize the worst-case error experimentally. From the proof of Proposition 3.2, we can see that the worst-case error for the median-of-means with point sets $P_1, \ldots, P_r$ is bounded above by the median of the worst-case errors for each individual point set. Justified by this, with $r = 15$, we generate 10^4 random sets of r polynomial lattices and compute the median of the worst-case errors associated with each set. The results are presented in the right panels of Figure 1, where we plot the histograms of all the realizations of

$$\log_2 \operatorname{median}\left(\mathcal{S}_{\alpha,\boldsymbol{\gamma}}(P_1), \ldots, \mathcal{S}_{\alpha,\boldsymbol{\gamma}}(P_r)\right) = \operatorname{median}\left(\log_2 \mathcal{S}_{\alpha,\boldsymbol{\gamma}}(P_1), \ldots, \log_2 \mathcal{S}_{\alpha,\boldsymbol{\gamma}}(P_r)\right).$$

We observe that the resulting distributions are more concentrated around small error values, indicating that taking the median successfully mitigates the influence of the "bad" polynomial lattices, exhibiting the effectiveness of the median-of-means approach in stabilizing the worst-case error.

Example 4.3. The rest of the examples is devoted to illustrating how our median QMC integration works for some concrete test functions. We compare our novel two median QMC rules with the deterministic QMC rule using Sobol' point sets as one of the "gold standard" methods. For this second experiment, we look at the simplest, one-dimensional integration problem. As we showed a universality for the three different function spaces in section 3, we prepare one test function from the respective function space:

$$f_1(x) = \sqrt{x},$$
$$f_2(x) = x^2 \left(\log x + \frac{1}{3} \right),$$
$$f_3(x) = x \exp(x).$$

For any $\gamma_1 > 0$, it is easy to check that $f_1 \in \mathcal{F}_{1,1,\gamma_1}^{\mathrm{sob}}$ but $f_1 \notin \mathcal{F}_{1,2,\gamma_1,q}^{\mathrm{sob}}$ for any $1 \leq q \leq \infty$. Similarly, we can check that $f_2 \in \mathcal{F}_{1,2,\gamma_1,q}^{\mathrm{sob}}$ for any finite $q \geq 1$ but $f_2 \notin \mathcal{F}_{1,3,\gamma_1,1}^{\mathrm{sob}}$, and also that $f_3 \in \mathcal{F}_{1,\infty,u_1}$ for any $u_1 > 0$. Note that, for the case $s = 1$, the first 2^m points of a Sobol' sequence is just the equi-spaced points $\{i/2^m \mid 0 \leq i < 2^m\}$.

The results for the three one-dimensional test functions are shown in Figure 2. As can be seen from the left panel, Sobol' points cannot exploit the smoothness of integrands so that the error decays at the rate of N^{-1} for all the test functions. On the contrary, our median QMC rules universally exploit the smoothness of integrands and exhibit a significantly better convergence behavior. In particular, for the function
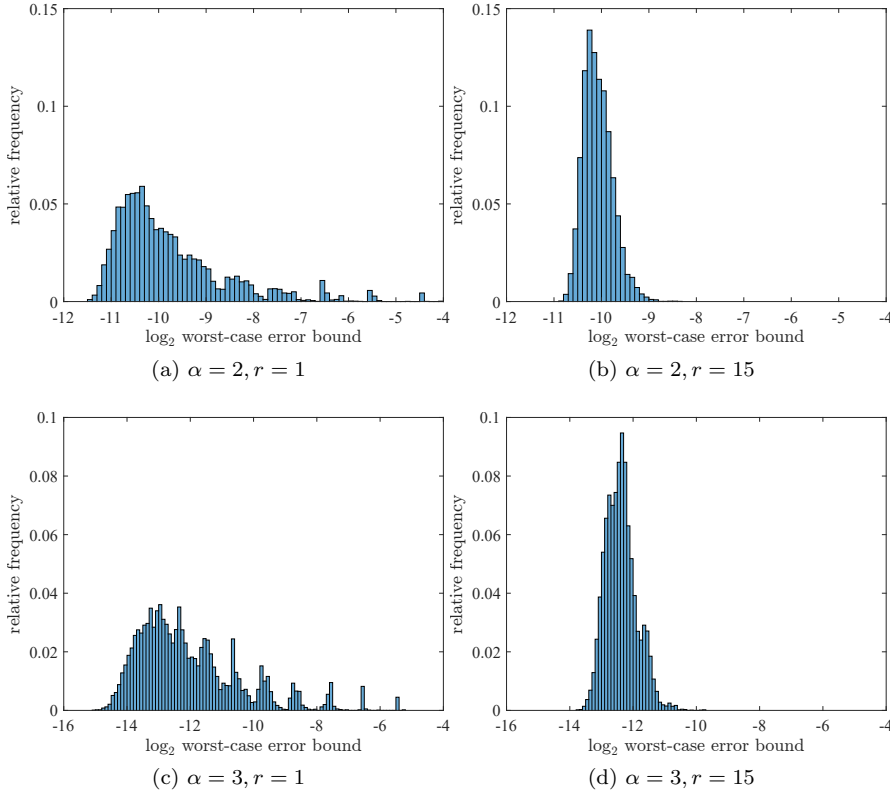
FIG. 1. *Histograms of $\log_2 \mathcal{S}_{\alpha,\gamma}$ for randomly generated polynomial lattices (left panels) and those of the median of $\log_2 \mathcal{S}_{\alpha,\gamma}$ for randomly generated sets of $r = 15$ polynomial lattices (right panels). The upper panels correspond to the case $\alpha = 2$, while the lower panels correspond to $\alpha = 3$.*

f_3, the rate of convergence looks accelerating until the error drops down to 2^{-52}, which agrees with a super-polynomial convergence proven in Theorem 3.18. It is interesting to see that the two median QMC rules perform almost equally for all the test functions.

Example 4.4. Let us move on to high-dimensional integration problems. Let $s = 20$ and consider the following test function

$$f_4(\boldsymbol{x}) = \prod_{j=1}^{s} \left[1 + \frac{1}{\exp(\lceil c \rceil j)} \left(x_j^c - \frac{1}{1+c} \right) \right],$$

with a parameter $c > 0$. When c is not an integer, the $\lceil c \rceil$-th derivative of the function $x \mapsto x^c$ is not absolutely continuous but is in $L_q([0,1])$ for any $1 \leq q < 1/(\lceil c \rceil - c)$, which means that $f_4 \in \mathcal{F}_{s,1,\gamma}^{\mathrm{sob}}$ if $0 < c < 1$ and $f_4 \in \mathcal{F}_{s,\lceil c \rceil,\gamma,q}^{\mathrm{sob}}$ for any $1 \leq q < 1/(\lceil c \rceil - c)$ if $c > 1$ is not an integer. This way, $\lceil c \rceil$ corresponds to the smoothness of this function. The factor $1/\exp(\lceil c \rceil j)$ is multiplied to model different relative importance for the individual variables.

The results for f_4 with $c = 0.5, 1.5, 2.5$ are shown in Figure 3. Here again, the left panel clearly depicts that Sobol' points cannot exploit the smoothness of integrands and the error decays at the rate of N^{-1} irrespective of the value of c. Both of the two
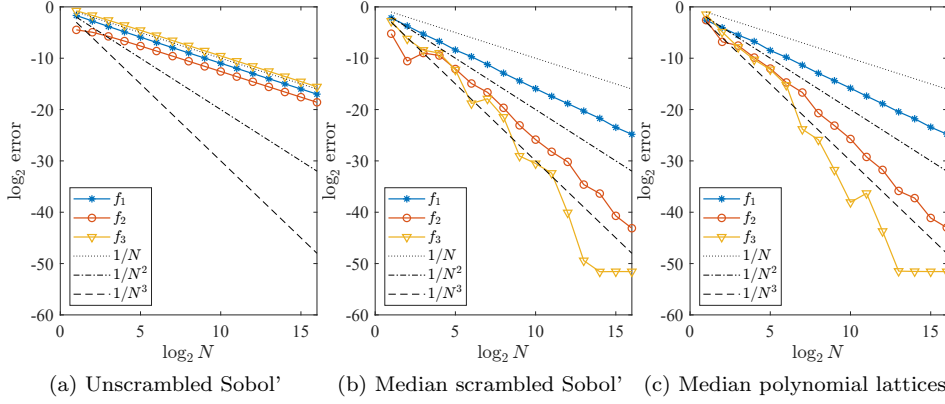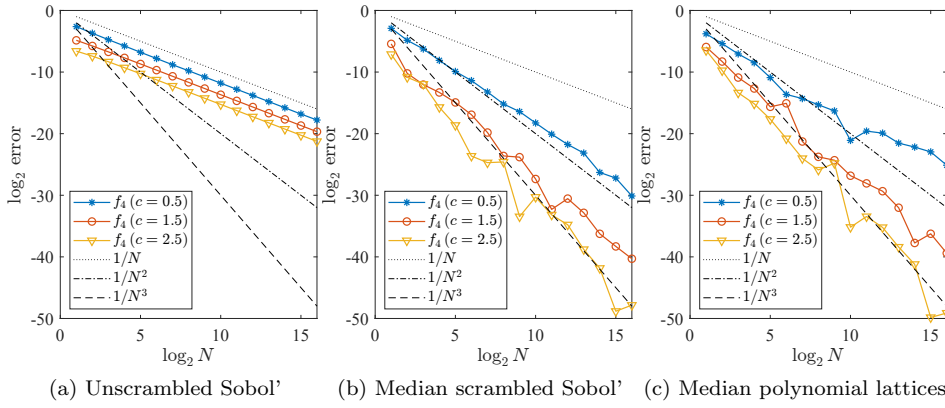
FIG. 2. *Comparison of the one-dimensional integration error by QMC rule using Sobol' points (left) and two median QMC rules using linearly scrambled Sobol' points (middle) and randomly chosen polynomial lattice point sets (right) for the test functions f_1 (blue), f_2 (orange) and f_3 (yellow), respectively.*
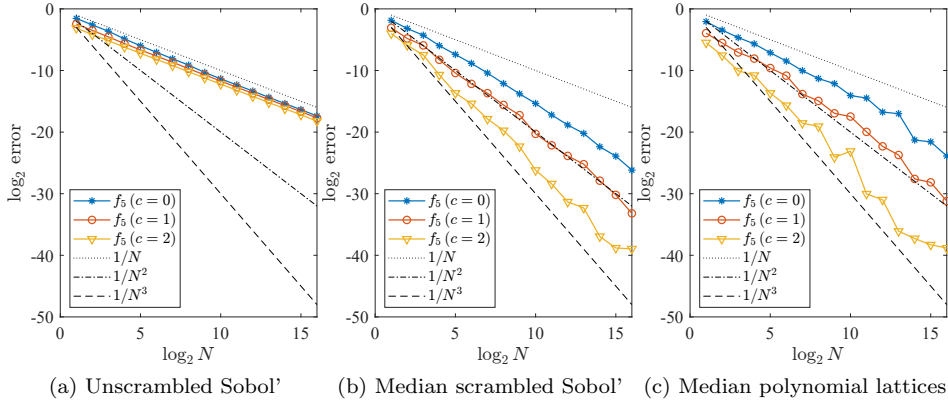

FIG. 3. *Comparison of the 20-dimensional integration error by QMC rule using Sobol' points (left) and two median QMC rules using linearly scrambled Sobol' points (middle) and randomly chosen polynomial lattice point sets (right) for the finitely smooth function f_4 with $c = 0.5$ (blue), $c = 1.5$ (orange) and $c = 2.5$ (yellow), respectively.*

median QMC rules compare quite favorably even for such a high-dimensional setting and both achieve an error decay nearly of order $N^{-\lceil c \rceil}$. Although taking the median has a stabilizing effect, a slight fluctuation in convergence is observed due to the fact that we randomly drew $r = 15$ point sets only once for each value of $m = \log_2 N$. Here again, the difference between the two median QMC rules is not significant.

Example 4.5. Finally, let us consider the multivariate test function with infinite smoothness

$$f_5(\boldsymbol{x}) = \exp\left(-\sum_{j=1}^{s} \frac{x_j}{2^{j^c}}\right),$$

with a parameter $c \geq 0$. It is easy to see that $f_5 \in \mathcal{F}_{s,\infty,\boldsymbol{u}}$ with $u_j = 1/2^{j^c}$. Since we have $a_j = -\log_2(u_j) = j^c$, the case $c = 0$ is classified into the unweighted case

(a) Unscrambled Sobol'   (b) Median scrambled Sobol'   (c) Median polynomial lattices

FIG. 4. *Comparison of the 5-dimensional integration error by QMC rule using Sobol' points (left) and two median QMC rules using linearly scrambled Sobol' points (middle) and randomly chosen polynomial lattice point sets (right) for the infinitely smooth function f_5 with $c = 0$ (blue), $c = 1$ (orange) and $c = 2$ (yellow), respectively.*

with $a = 1$ in Theorem 3.18, while the case $c > 0$ is into the weighted case with $a = 1$ and $q = c$. In passing, it was shown in [40] that the worst-case error for a digital net over $\mathbb{F}_2$ in $\mathcal{F}_{s,\infty,\boldsymbol{u}}$ is bounded from above and below (up to constants) by the absolute integration error for the integrand f_5 with $1/2^{j^c}$ replaced by u_j. Therefore, the experimental results presented below on this function f_5 are expected to correspond to the behavior of the worst-case error.

The results for f_5 with $s = 5$ and $c = 0, 1, 2$ are shown in Figure 4. Regarding each of the two median QMC rules, although it is not obvious from the figure whether the error decays super-polynomially within the performed range of N, the error decays at a rate between N^{-1} and N^{-2} even for the worst case $c = 0$, which is better than that for naive QMC rule using Sobol' points. As c increases, the convergence behavior for small N gets improved and we can observe a higher order convergence than the case $c = 0$.

REFERENCES

[1] C. AISTLEITNER, *Tractability results for the weighted star-discrepancy*, Journal of Complexity, 30 (2014), pp. 381–391.
[2] J. BALDEAUX, J. DICK, G. LEOBACHER, D. NUYENS, AND F. PILLICHSHAMMER, *Efficient calculation of the worst-case error and (fast) component-by-component construction of higher order polynomial lattice rules*, Numerical Algorithms, 59 (2012), pp. 403–431.
[3] J. DICK, *Walsh spaces containing smooth functions and quasi–Monte Carlo rules of arbitrary high order*, SIAM Journal on Numerical Analysis, 46 (2008), pp. 1519–1553.
[4] J. DICK, *The decay of the Walsh coefficients of smooth functions*, Bulletin of the Australian Mathematical Society, 80 (2009), pp. 430–453.
[5] J. DICK, *Random weights, robust lattice rules and the geometry of the cbcrc algorithm*, Numerische Mathematik, 122 (2012), pp. 443–467.
[6] J. DICK AND T. GODA, *Stability of lattice rules and polynomial lattice rules constructed by the component-by-component algorithm*, Journal of Computational and Applied Mathematics, 382 (2021), p. 113062.

[7] J. Dick, T. Goda, and K. Suzuki, *Component-by-component construction of randomized rank-1 lattice rules achieving almost the optimal randomized error rate*, Mathematics of Computation, 91 (2022), pp. 2771–2801.

[8] J. Dick, T. Goda, K. Suzuki, and T. Yoshiki, *Construction of interlaced polynomial lattice rules for infinitely differentiable functions*, Numerische Mathematik, 137 (2017), pp. 257–288.

[9] J. Dick, T. Goda, and T. Yoshiki, *Richardson extrapolation of polynomial lattice rules*, SIAM Journal on Numerical Analysis, 57 (2019), pp. 44–69.

[10] J. Dick, A. Hinrichs, and F. Pillichshammer, *Proof techniques in quasi-Monte Carlo theory*, Journal of Complexity, 31 (2015), pp. 327–371.

[11] J. Dick, F. Y. Kuo, Q. T. L. Gia, D. Nuyens, and C. Schwab, *Higher order QMC Petrov–Galerkin discretization for affine parametric operator equations with random field inputs*, SIAM Journal on Numerical Analysis, 52 (2014), pp. 2676–2702.

[12] J. Dick, F. Y. Kuo, and I. H. Sloan, *High-dimensional integration: the quasi-Monte Carlo way*, Acta Numerica, 22 (2013), pp. 133–288.

[13] J. Dick, G. Leobacher, and F. Pillichshammer, *Construction algorithms for digital nets with low weighted star discrepancy*, SIAM Journal on Numerical Analysis, 43 (2005), pp. 76–95.

[14] J. Dick and F. Pillichshammer, *Strong tractability of multivariate integration of arbitrary high order using digitally shifted polynomial lattice rules*, Journal of Complexity, 23 (2007), pp. 436–453.

[15] J. Dick and F. Pillichshammer, *Digital Nets and Sequences: Discrepancy Theory and Quasi–Monte Carlo Integration*, Cambridge University Press, Cambridge, 2010.

[16] A. Ebert, P. Kritzer, D. Nuyens, and O. Osisiogu, *Digit-by-digit and component-by-component constructions of lattice rules for periodic functions with unknown smoothness*, Journal of Complexity, 66 (2021), p. 101555.

[17] H. Faure, *Discrépance de suites associées à un système de numération (en dimension s)*, Acta Arithmetica, 41 (1982), pp. 337–351.

[18] H. Faure and C. Lemieux, *Irreducible Sobol' sequences in prime power bases*, Acta Arithmetica, 173 (2016), pp. 59–80.

[19] K. K. Frolov, *Upper bounds for the errors of quadrature formulae on classes of functions*, Doklady Akademii Nauk SSSR, 231 (1976), pp. 818–821.

[20] E. Gobet, M. Lerasle, and D. Métivier, *Mean estimation for Randomized Quasi Monte Carlo method*, HAL preprint `hal-03631879`, (2022).

[21] T. Goda, *Good interlaced polynomial lattice rules for numerical integration in weighted Walsh spaces*, Journal of Computational and Applied Mathematics, 285 (2015), pp. 279–294.

[22] T. Goda, *Quasi-Monte Carlo integration using digital nets with antithetics*, Journal of Computational and Applied Mathematics, 304 (2016), pp. 26–42.

[23] T. Goda and P. L'Ecuyer, *Construction-free median quasi-Monte Carlo rules for function spaces with unspecified smoothness and general weights*, SIAM Journal on Scientific Computing, 44 (2022), pp. A2765–A2788.

[24] T. Goda and K. Suzuki, *Improved bounds on the gain coefficients for digital nets in prime power base*, Journal of Complexiy, 76 (2023), p. 101722.

[25] T. Goda, K. Suzuki, and T. Yoshiki, *Digital nets with infinite digit expansions and construction of folded digital nets for quasi-Monte Carlo integration*, Journal of Complexity, 33 (2016), pp. 30–54.

[26] J. H. Halton, *On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals*, Numerische Mathematik, 2 (1960), pp. 84–90.

[27] S. Harase, *Quasi-Monte Carlo point sets with small t-values and WAFOM*, Applied Mathematics and Computation, 254 (2015), pp. 318–326.

[28] F. J. Hickernell, *Quadrature error bounds with applications to lattice rules*, SIAM Journal on Numerical Analysis, 33 (1996), pp. 1995–2016.

[29] F. J. Hickernell, *A generalized discrepancy and quadrature error bound*, Mathematics of Computation, 67 (1998), pp. 299–322.

[30] F. J. Hickernell and H. Niederreiter, *The existence of goodextensible rank-1 lattices*, Journal of Complexity, 19 (2003), pp. 286–300.

[31] J. Hofstadler and D. Rudolf, *Consistency of randomized integration methods*, Journal of Complexity, 76 (2023), p. 101740.

[32] D. Krieg and E. Novak, *A universal algorithm for multivariate integration*, Foundations of Computational Mathematics, 17 (2017), pp. 895–916.

[33] R. J. Kunsch and D. Rudolf, *Optimal confidence for Monte Carlo integration of smooth functions*, Advances in Computational Mathematics, 45 (2019), pp. 3095–3122.

[34] F. Y. Kuo, *Component-by-component constructions achieve the optimal rate of convergence for multivariate integration in weighted Korobov and Sobolev spaces*, Journal of Complexity, 19 (2003), pp. 301–320.

[35] F. Y. Kuo and D. Nuyens, *Application of quasi-Monte Carlo methods to elliptic PDEs with random diffusion coefficients: a survey of analysis and implementation*, Foundations of Computational Mathematics, 16 (2016), pp. 1631–1696.

[36] F. Y. Kuo, C. Schwab, and I. H. Sloan, *Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients*, SIAM Journal on Numerical Analysis, 50 (2012), pp. 3351–3374.

[37] C. Lemieux and P. L'Ecuyer, *Randomized polynomial lattice rules for multivariate integration and simulation*, SIAM Journal on Scientific Computing, 24 (2003), pp. 1768–1789.

[38] P. L'Ecuyer and C. Lemieux, *Quasi-Monte Carlo via linear shift-register sequences*, in Proceedings of the 1999 Winter Simulation Conference, P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans, eds., 1999, pp. 632–639.

[39] J. Matoušek, *On the L_2-discrepancy for anchored boxes*, Journal of Complexity, 14 (1998), pp. 527–556.

[40] M. Matsumoto, R. Ohori, and T. Yoshiki, *Approximation of Quasi-Monte Carlo worst case error in weighted spaces of infinitely times smooth functions*, Journal of Computational and Applied Mathematics, 330 (2018), pp. 155–164.

[41] M. Matsumoto, M. Saito, and K. Matoba, *A computable figure of merit for quasi-Monte Carlo point sets*, Mathematics of Computation, 83 (2014), pp. 1233–1250.

[42] M. Merkle, *Jensen's inequality for medians*, Statistics & Probability Letters, 71 (2005), pp. 277–281.

[43] H. Niederreiter, *Low-discrepancy and low-dispersion sequences*, Journal of Number Theory, 30 (1988), pp. 51–70.

[44] H. Niederreiter, *Low-discrepancy point sets obtained by digital constructions over finite fields*, Czechoslovak Mathematical Journal, 42 (1992), pp. 143–166.

[45] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, Society for Industrial and Applied Mathematics, Philadelphia, 1992.

[46] H. Niederreiter and C. Xing, *Rational Points on Curves over Finite Fields: Theory and Applications*, Cambridge University Press, Cambridge, 2001.

[47] W. Niemiro and P. Pokarowski, *Fixed precision MCMC estimation by median of products of averages*, Journal of Applied Probability, 46 (2009), pp. 309–329.

[48] E. Novak, *On the power of adaption*, Journal of Complexity, 12 (1996), pp. 199–237.

[49] E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems, Volume I: Linear Information*, EMS Press, Zürich, 2008.

[50] D. Nuyens and R. Cools, *Fast algorithms for component-by-component construction of rank-1 lattice rules in shift-invariant reproducing kernel Hilbert spaces*, Mathematics of Computation, 75 (2006), pp. 903–920.

[51] A. B. Owen, *Randomly permuted (t, m, s)-nets and (t, s)-sequences*, in Monte Carlo and quasi-Monte Carlo methods in scientific computing, H. Niederreiter and P. J.-S. Shiue, eds., New York, 1995, Springer, pp. 299–317.

[52] A. B. Owen, *Monte Carlo variance of scrambled net quadrature*, SIAM Journal on Numerical Analysis, 34 (1997), pp. 1884–1910.

[53] A. B. Owen, *Scrambled net variance for integrals of smooth functions*, The Annals of Statistics, 25 (1997), pp. 1541–1562.

[54] A. B. Owen, *Variance with alternative scramblings of digital nets*, ACM Transactions on Modeling and Computer Simulation, 13 (2003), pp. 363–378.

[55] Z. Pan and A. B. Owen, *Super-polynomial accuracy of multidimensional randomized nets using the median-of-means*, arXiv preprint `arXiv:2208.05078`, (2022).

[56] Z. Pan and A. B. Owen, *Super-polynomial accuracy of one dimensional randomized nets using the median-of-means*, Mathematics of Computation, 92 (2023), pp. 805–837.

[57] F. Pillichshammer, *Polynomial lattice point sets*, in Monte Carlo and Quasi-Monte Carlo Methods 2010, L. Plaskota and H. Woźniakowski, eds., Berlin, 2012, Springer-Verlag, pp. 189–210.

[58] P. Pollack, *Irreducible polynomials with several prescribed coefficients*, Finite Fields and Their Applications, 22 (2013), pp. 70–78.

[59] I. F. Sharygin, *A lower estimate for the error of quadrature formulae for certain classes of functions*, Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki, 3 (1963), pp. 370–376.

[60] M. M. Skriganov, *Harmonic analysis on totally disconnected groups and irregularities of point distributions*, Journal für die reine und angewandte Mathematik, 600 (2006), pp. 25–49.

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

[62] I. M. Sobol', *On the distribution of points in a cube and the approximate evaluation of integrals*, Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki, 7 (1967), pp. 784–802.

[63] K. Suzuki, *WAFOM over abelian groups for quasi-Monte Carlo point sets*, Hiroshima Mathematical Journal, 45 (2015), pp. 341–364.

[64] K. Suzuki, *Super-polynomial convergence and tractability of multivariate integration for infinitely times differentiable functions*, Journal of Complexity, 39 (2017), pp. 51–68.

[65] K. Suzuki and T. Yoshiki, *Formulas for the Walsh coefficients of smooth functions and their application to bounds on the Walsh coefficients*, Journal of Approximation Theory, 205 (2016), pp. 1–24.

[66] S. Tezuka, *Polynomial arithmetic analogue of Halton sequences*, ACM Transactions on Modeling and Computer Simulation, 3 (1993), pp. 99–107.

[67] S. Tezuka, *A generalization of Faure sequences and its efficient implementation*, Research Report IBM, RT0105 (1994), pp. 1–10.

[68] S. Tezuka and H. Faure, *I-binomial scrambling of digital nets and sequences*, Journal of Complexity, 19 (2003), pp. 744–757.

[69] M. Ullrich, *A Monte Carlo method for integration of multivariate smooth functions*, SIAM Journal on Numerical Analysis, 55 (2017), pp. 1188–1200.

[70] M. Ullrich and T. Ullrich, *The role of Frolov's cubature formula for functions with bounded mixed derivative*, SIAM Journal on Numerical Analysis, 54 (2016), pp. 969–993.

[71] X. Wang, *Strong tractability of multivariate integration using quasi-Monte Carlo algorithms*, Mathematics of Computation, 72 (2003), pp. 823–838.

[72] H. Weyl, *Über die Gleichverteilung von Zahlen mod. Eins*, Mathematische Annalen, 77 (1916), pp. 313–352.

[73] T. Yoshiki, *Bounds on Walsh coefficients by dyadic difference and a new Koksma-Hlawka type inequality for Quasi-Monte Carlo integration*, Hiroshima Mathematical Journal, 47 (2017), pp. 155–179.