*IMA Journal of Numerical Analysis* (2019) **00**, 1–24
doi:10.1093/imanum/drz017



# Richardson extrapolation allows truncation of higher-order digital nets and sequences

TAKASHI GODA
*School of Engineering, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*
goda@frcer.t.u-tokyo.ac.jp

[Received on 23 May 2018; revised on 9 January 2019]

We study numerical integration of smooth functions defined over the $s$-dimensional unit cube. A recent work by Dick *et al.* (2019, Richardson extrapolation of polynomial lattice rules. *SIAM J. Numer. Anal.*, **57**, 44–69) has introduced so-called extrapolated polynomial lattice rules, which achieve the almost optimal rate of convergence for numerical integration, and can be constructed by the fast component-by-component search algorithm with smaller computational costs as compared to interlaced polynomial lattice rules. In this paper we prove that, instead of polynomial lattice point sets, truncated higher-order digital nets and sequences can be used within the same algorithmic framework to explicitly construct good quadrature rules achieving the almost optimal rate of convergence. The major advantage of our new approach compared to original higher-order digital nets is that we can significantly reduce the precision of points, i.e., the number of digits necessary to describe each quadrature node. This finding has a practically useful implication when either the number of points or the smoothness parameter is so large that original higher-order digital nets require more than the available finite-precision floating-point representations.

*Keywords*: numerical integration; quasi-Monte Carlo; higher-order digital nets and sequences; Richardson extrapolation.

## 1. Introduction

In this paper we study numerical integration of multivariate functions defined over the $s$-dimensional unit cube. For an integrable function $f\colon [0,1)^s \to \mathbb{R}$ we denote the integral of $f$ by

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

We consider approximating $I(f)$ by a linear algorithm of the form

$$A_N(f) = \sum_{h=0}^{N-1} w_h f(\boldsymbol{x}_h),$$

where $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}$ and $w_0, \ldots, w_{N-1}$ denote the quadrature nodes and the quadrature weights, respectively. We call the algorithm $A_N$ a *quasi-Monte Carlo (QMC) rule* if the weights are given by $w_0 = \cdots = w_{N-1} = 1/N$.

For a Banach space $V$ with norm $\|\cdot\|_V$ the worst-case error of $A_N$ is defined by

$$e^{\mathrm{wor}}(A_N, V) := \sup_{\substack{f\in V\\ \|f\|_V\leqslant 1}} \left|I(f) - A_N(f)\right|.$$

Our aim is then to design a good quadrature rule $A_N$ such that $e^{\mathrm{wor}}(A_N, V)$ is made as small as possible, since for any function $f \in V$, we have

$$\left|I(f) - A_N(f)\right| \leqslant \|f\|_V \cdot e^{\mathrm{wor}}(A_N, V),$$

meaning that a single algorithm works well for all functions belonging to $V$. In this paper we are particularly interested in Banach spaces with dominating mixed smoothness $\alpha \in \mathbb{N}$, $\alpha \geqslant 2$, consisting of functions that have partial mixed derivatives up to order $\alpha$ in each variable (see Section 2.1 for more details). Such function spaces have been motivated by Dick *et al.* (2014) for the study of partial differential equations with random coefficients.

For function spaces of our interest QMC rules using *higher-order digital nets and sequences* as quadrature nodes are known to achieve the almost optimal rate of convergence of the worst-case error, which is $O(N^{-\alpha+\varepsilon})$ with arbitrarily small $\varepsilon > 0$. The concept and explicit construction of higher-order digital nets and sequences were originally introduced by Dick (2007, 2008) (see Section 2.2 for more details). Since then, on the one hand, further theoretical investigations on them have been made (see, e.g., Baldeaux *et al.*, 2011; Hinrichs *et al.*, 2016; Goda *et al.*, 2017, 2018). On the other hand, how to efficiently search for good quadrature node sets in a weighted function space setting as considered by Sloan & Woźniakowski (1998) has also attracted some interest (see, e.g., Baldeaux *et al.*, 2012; Goda, 2015; Goda & Dick, 2015; Gantner & Schwab, 2016; Goda *et al.*, 2016). In particular, so-called interlaced polynomial lattice rules originated by Goda & Dick (2015) and Goda (2015), which are based on the digit interlacing composition due to Dick (2007, 2008), have been applied in the context of partial differential equations with random coefficients (see, e.g., Dick *et al.*, 2014; Kuo & Nuyens, 2016).

Recently, a new alternative approach to interlaced polynomial lattice rules has been developed by Dick *et al.* (2019). Instead of searching for a single interlaced polynomial lattice point set their approach is to search for $\alpha$ classical polynomial lattice point sets with geometric spacing of $N$ first and then to apply Richardson extrapolation recursively to $\alpha$ numerical values $A_N(f)$. Such *extrapolated polynomial lattice rules* have been proved to achieve the almost optimal rate of convergence, and, moreover, the fast component-by-component algorithm can be used to find good rules with smaller computational costs as compared to interlaced polynomial lattice rules. A further advantage can be found in the fact that the fast QMC matrix-vector multiplication technique from Dick *et al.* (2015) applies to extrapolated polynomial lattice rules, whereas it is not straightforwardly applicable to interlaced ones.

In this paper, as a continuation of Dick *et al.* (2019), we push forward the idea of applying Richardson extrapolation to QMC rules for achieving a high order of convergence for multivariate numerical integration. In particular, we consider QMC rules using *truncated* higher-order digital nets or sequences as quadrature nodes, where truncation is done in the following way: we apply the following map $\mathrm{tr}_m\colon [0,1) \to [0,1)$ component-wise to each node $\boldsymbol{x}_h = (x_{h,1},\dots,x_{h,s}) \in [0,1)^s$ of higher-order digital nets with prime base $p$ and size $N = p^m$:

$$\mathrm{tr}_m\left(\sum_{i=1}^{\infty}\frac{\xi_i}{p^i}\right) = \sum_{i=1}^{m}\frac{\xi_i}{p^i} \quad \text{with } \xi_i \in \{0,1,\dots,p-1\}. \tag{1.1}$$

Then we prove that, by applying Richardson extrapolation recursively to QMC rules using such truncated higher-order digital nets or sequences with geometric spacing of $N$, the resulting linear algorithm to approximate $I(f)$ achieves the almost optimal rate of convergence.

Our finding has the following practically useful implication, especially when $p = 2$. The original digit interlacing composition approach to constructing higher-order digital nets with size $N = p^m$ requires $\alpha m$ digits in the $p$-adic expansion of each component of each node. Hence, the round-off error is inevitable when $\alpha m$ is larger than what is available via finite-precision floating-point representations (for instance, 23 and 52 for IEEE 754 single- and double-precision floating-point formats, respectively). Depending on the considered integrand the round-off error becomes comparable to the approximation error for numerical integration already when $m$ is of practical size, say $m \approx 20$. In such a situation the approximation error will remain more or less unchanged, even by increasing $m$. Since our extrapolation approach can reduce the necessary number of digits from $\alpha m$ to $m$, the round-off error problem will not happen until $m$ is large enough and, importantly, becomes independent of the smoothness parameter $\alpha$. Therefore, with the help of Richardson extrapolation, higher-order QMC rules become available for wider ranges of $N$ and $\alpha$ than before without suffering from the rounding problem.

The rest of this paper is organized as follows. After describing the necessary background and notation in Section 2, we propose an extrapolation-based quadrature rule using truncated higher-order digital nets or sequences, and prove the worst-case error bound of the proposed rule in Banach spaces with dominating mixed smoothness in Section 3. In the same section we further provide another possible, similar but different quadrature rule, together with its worst-case error bound. We conclude this paper with numerical experiments in Section 4.

## 2. Preliminaries

Throughout this paper we denote the set of positive integers by $\mathbb{N}$ and write $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For a prime $p$ let $\mathbb{F}_p$ be the finite field with $p$ elements, which is identified with the set of integers $\{0, 1, \ldots, p-1\} \subset \mathbb{Z}$ equipped with addition and multiplication modulo $p$. For an $s$-dimensional vector $\boldsymbol{x} = (x_1, \ldots, x_s)$ and a subset $u \subseteq \{1, \ldots, s\}$, we write $\boldsymbol{x}_u = (x_j)_{j \in u}$ and denote the cardinality and the complement of $u$ by $|u|$, and $-u := \{1, \ldots, s\} \setminus u$, respectively.

### 2.1 *Banach spaces with dominating mixed smoothness*

Following Dick *et al.* (2014) here we introduce the definition of function spaces, which we consider in this paper. Let $\alpha \in \mathbb{N}$, $\alpha \geqslant 2$ and $1 \leqslant q, r \leqslant \infty$ be real numbers. Further, let $\boldsymbol{\gamma} = (\gamma_u)_{u \subseteq \{1,\ldots,s\}}$ be a set of non-negative real numbers called weights, which has been introduced by Sloan & Woźniakowski (1998) to moderate the relative importance of different variables or groups of variables. In this paper we do not discuss the dependence of the worst-case error on the dimension, and just consider the weights for making consistent use of the notations of previous works.

Assume that a function $f\colon [0, 1)^s \to \mathbb{R}$ has partial mixed derivatives up to order $\alpha$ in each variable. We define the norm of $f$ by

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

with the obvious modification if either $q$ or $r$ is infinite. Here $(\boldsymbol{\tau}_{u\setminus v}, \boldsymbol{\alpha}_v, \mathbf{0})$ denotes the vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_s)$ such that

$$\beta_j = \begin{cases} \tau_j & \text{if } j \in u \setminus v, \\ \alpha & \text{if } j \in v, \\ 0 & \text{otherwise,} \end{cases}$$

and $f^{(\boldsymbol{\tau}_{u\setminus v}, \boldsymbol{\alpha}_v, \mathbf{0})}$ denotes the partial mixed derivative of order $(\boldsymbol{\tau}_{u\setminus v}, \boldsymbol{\alpha}_v, \mathbf{0})$ of $f$. If there exist subsets $u$ such that $\gamma_u = 0$, then we assume that the corresponding inner double sum is 0 and formally set $0/0 = 0$. Now we define the Banach space with dominating mixed smoothness $\alpha$ by

$$W_{s,\alpha,q,r} := \left\{ f\colon [0,1)^s \to \mathbb{R} : \|f\|_{s,\alpha,q,r} < \infty \right\}.$$

For $\tau \in \mathbb{N}$ we denote the Bernoulli polynomial of degree $\tau$ by $B_\tau\colon [0,1) \to \mathbb{R}$ and we put $b_\tau(\cdot) = B_\tau(\cdot)/\tau!$. With a slight abuse of notation we write $b_\tau = b_\tau(0)$. Further, we denote the one-periodic extension of the polynomial $b_\tau$ by $\tilde{b}_\tau\colon \mathbb{R} \to \mathbb{R}$. As shown below we have a point-wise representation for functions in $W_{s,\alpha,q,r}$.

Lemma 2.1 For $f \in W_{s,\alpha,q,r}$, we have

$$f(\boldsymbol{x}) = \sum_{u \subseteq \{1,\ldots,s\}} f_u(\boldsymbol{x}_u),$$

where each $f_u$ depends only on $\boldsymbol{x}_u$ and is given by

$$f_u(\boldsymbol{x}_u) = \sum_{v \subseteq u} (-1)^{(\alpha+1)|v|} \sum_{\boldsymbol{\tau}_{u\setminus v} \in \{1,\ldots,\alpha\}^{|u\setminus v|}} \prod_{j \in u\setminus v} b_{\tau_j}(x_j) \int_{[0,1)^s} f^{(\boldsymbol{\tau}_{u\setminus v}, \boldsymbol{\alpha}_v, \mathbf{0})}(\boldsymbol{y}) \prod_{j \in v} \tilde{b}_\alpha(y_j - x_j)\, \mathrm{d}\boldsymbol{y}.$$

Moreover, we have

$$\|f\|^r_{s,\alpha,q,r} = \sum_{u \subseteq \{1,\ldots,s\}} \|f_u\|^r_{s,\alpha,q,r}.$$

*Proof.* See the proof of Dick *et al.* (2014, Theorem 3.5). □

### 2.2 *Higher-order digital nets and sequences*

2.2.1 *Digital construction scheme.* We first introduce a class of point sets called digital nets, which are originally due to Niederreiter (1992).

Definition 2.2 (Digital nets). For a prime $p$ and $m, n \in \mathbb{N}$ let $C_1, \ldots, C_s \in \mathbb{F}_p^{n\times m}$. For $h \in \mathbb{N}_0$, $h < p^m$ we denote the $p$-adic expansion of $h$ by

$$h = \eta_0 + \eta_1 p + \cdots + \eta_{m-1} p^{m-1}.$$

Set $\boldsymbol{x}_h = (x_{h,1}, \ldots, x_{h,s}) \in [0, 1)^s$ where

$$x_{h,j} = \frac{\xi_{h,j,1}}{p} + \frac{\xi_{h,j,2}}{p^2} + \cdots + \frac{\xi_{h,j,n}}{p^n},$$

in which $\xi_{h,j,1}, \ldots, \xi_{h,j,n}$ are given by

$$(\xi_{h,j,1}, \ldots, \xi_{h,j,n}) = (\eta_0, \eta_1, \ldots, \eta_{m-1}) \cdot C_j^\top$$

for $1 \leqslant j \leqslant s$. Then the set of points $P_{m,n} = \{\boldsymbol{x}_h : 0 \leqslant h < p^m\}$ is called a digital net over $\mathbb{F}_p$ (with generating matrices $C_1, \ldots, C_s$).

It is easy to see from the definition that the parameter $m$ determines the total number of points, while $n$ does determine the precision of points.

Remark 2.3 Let us consider the case $n = \infty$. For each $C_j = (c_{k,l}^{(j)})_{k \in \mathbb{N}, 1 \leqslant l \leqslant m}$, if there exists a function $K_j \colon \{1, \ldots, m\} \to \mathbb{N}$ such that $c_{k,l}^{(j)} = 0$ whenever $k > K_j(l)$, the vector-matrix product appearing in the above definition gives $\xi_{h,j,i} = 0$ for all $i > \max_{1 \leqslant l \leqslant m} K_j(l) =: n_j$. Thus, each number $x_{h,j}$ is uniquely written in a finite $p$-adic expansion with the precision at most $n' = \max_{1 \leqslant j \leqslant s} n_j$. By identifying $C_1, \ldots, C_s$ with their upper $n' \times m$ submatrices Definition 2.2 still applies to such cases.

It is straightforward to extend the definition of digital nets to digital sequences that are infinite sequences of points in $[0, 1)^s$.

Definition 2.4 (Digital sequences). For a prime $p$ let $C_1, \ldots, C_s \in \mathbb{F}_p^{\mathbb{N} \times \mathbb{N}}$. For each $C_j = (c_{k,l}^{(j)})_{k,l \in \mathbb{N}}$ assume that there exists a function $K_j \colon \mathbb{N} \to \mathbb{N}$ such that $c_{k,l}^{(j)} = 0$ if $k > K_j(l)$. For $h \in \mathbb{N}_0$ we denote the $p$-adic expansion of $h$ by

$$h = \eta_0 + \eta_1 p + \cdots,$$

where all but a finite number of $\eta_i$s are 0. Set $\boldsymbol{x}_h = (x_{h,1}, \ldots, x_{h,s}) \in [0, 1)^s$, where

$$x_{h,j} = \frac{\xi_{h,j,1}}{p} + \frac{\xi_{h,j,2}}{p^2} + \cdots,$$

in which $\xi_{h,j,1}, \xi_{h,j,2}, \ldots$ are given by

$$(\xi_{h,j,1}, \xi_{h,j,2}, \ldots) = (\eta_0, \eta_1, \ldots) \cdot C_j^\top$$

for $1 \leqslant j \leqslant s$. Then the sequence of points $\mathscr{S} = \{\boldsymbol{x}_h : h \in \mathbb{N}_0\}$ is called a digital sequence over $\mathbb{F}_p$ (with generating matrices $C_1, \ldots, C_s$).

As mentioned in Remark 2.3 the existence of functions $K_j$ in this definition is assumed to ensure that every number $x_{h,j}$ is uniquely written in a finite $p$-adic expansion.

2.2.2 *Dual nets.* Next we introduce the concept of dual nets and also the weight function due to Dick (2008), which generalizes the original weight function introduced independently by Niederreiter (1986) and Rosenbloom & Tsfasman (1997). Thereafter, we give the definition of higher-order digital nets and sequences.

DEFINITION 2.5 (Dual nets). For a prime $p$ and $m, n \in \mathbb{N}$ let $P_{m,n}$ be a digital net over $\mathbb{F}_p$ with generating matrices $C_1, \ldots, C_s \in \mathbb{F}_p^{n\times m}$. The dual net of $P_{m,n}$, denoted by $P_{m,n}^{\perp}$, is defined by

$$P_{m,n}^{\perp} := \left\{ \boldsymbol{k} = (k_1, \ldots, k_s) \in \mathbb{N}_0^s : C_1^{\top} \nu_n(k_1) \oplus \cdots \oplus C_s^{\top} \nu_n(k_s) = \boldsymbol{0} \in \mathbb{F}_p^m \right\},$$

where we write $\nu_n(k) = (\kappa_0, \ldots, \kappa_{n-1})^{\top}$ for $k \in \mathbb{N}_0$ whose $p$-adic expansion is given by $k = \kappa_0 + \kappa_1 p + \cdots$, where all but a finite number of $\kappa_i$'s are 0.

REMARK 2.6 Again, even for the case $n = \infty$, as long as there exists a function $K_j \colon \{1, \ldots, m\} \to \mathbb{N}$ such that $c_{k,l}^{(j)} = 0$ whenever $k > K_j(l)$ for each $C_j = (c_{k,l}^{(j)})_{k,l\in\mathbb{N}}$, Definition 2.5 still applies.

DEFINITION 2.7 (Weight function). Let $\alpha \in \mathbb{N}$. We denote the $p$-adic expansion of $k \in \mathbb{N}$ by

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

with $\kappa_1, \ldots, \kappa_v \in \{1, \ldots, p-1\}$ and $c_1 > c_2 > \cdots > c_v > 0$. Then we define the weight function $\mu_\alpha \colon \mathbb{N}_0 \to \mathbb{N}_0$ by

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

and $\mu_\alpha(0) = 0$. In case of vectors in $\mathbb{N}_0^s$ we define

$$\mu_\alpha(k_1, \ldots, k_s) := \sum_{j=1}^{s} \mu_\alpha(k_j).$$

Now we are ready to introduce higher-order digital nets and sequences.

DEFINITION 2.8 (Higher-order digital nets). Let $\alpha \in \mathbb{N}$. For a prime $p$ and $m, n \in \mathbb{N}$ let $P_{m,n}$ be a digital net over $\mathbb{F}_p$. We call $P_{m,n}$ an *order $\alpha$ digital $(t, m, s)$-net* over $\mathbb{F}_p$ if there exists an integer $0 \leqslant t \leqslant \alpha m$ such that the following holds:

$$\mu_\alpha(P_{m,n}^{\perp}) := \min_{\boldsymbol{k} \in P_{m,n}^{\perp} \setminus \{\boldsymbol{0}\}} \mu_\alpha(\boldsymbol{k}) > \alpha m - t.$$

REMARK 2.9 It follows from Definition 2.5 that $(p^n, 0, \ldots, 0) \in P_{m,n}^{\perp}$, which gives

$$\mu_\alpha(P_{m,n}^{\perp}) \leqslant \mu_\alpha(p^n, 0, \ldots, 0) = n + 1.$$

Thus, in order for $P_{m,n}$ to be an order $\alpha$ digital $(t, m, s)$-net, it is necessary to have $n \geqslant \alpha m - t$. Together with Remark 2.12 below, this means that the precision $n$ should scale linearly with $\alpha$ and $m$.

DEFINITION 2.10 (Higher-order digital sequences). Let $\alpha \in \mathbb{N}$. For a prime $p$ let $\mathscr{S}$ be a digital sequence over $\mathbb{F}_p$. We call $\mathscr{S}$ an *order $\alpha$ digital $(t,s)$-sequence* over $\mathbb{F}_p$ if there exists $t \in \mathbb{N}_0$ such that the first $p^m$ points of $\mathscr{S}$ are an order $\alpha$ digital $(t,m,s)$-net over $\mathbb{F}_p$ when $\alpha m > t$.

2.2.3 *Explicit constructions.* It is important to note that higher-order digital nets and sequences can be constructed explicitly. In fact, many explicit constructions of order 1 digital $(t,s)$-sequences with small $t$-values for arbitrary $s$ have been known already. Among them are those by Sobol' (1967), Faure (1982), Niederreiter (1988), Tezuka (1993); Niederreiter & Xing (2001). Some of them hold the property on functions $K_j$ in Definition 2.4 such that $K_j(l) \leqslant l$ for all $j, l \in \mathbb{N}$. This means that the first $p^m$ points of such digital sequences are an order 1 digital $(t,m,s)$-net over $\mathbb{F}_p$ with the precision $n \leqslant m$. We refer to Dick & Pillichshammer (2010, Chapter 8) for more information on these special constructions.

Moreover, the digit interlacing composition due to Dick (2007, 2008) enables us to construct order $\alpha$ digital $(t,m,s)$-nets and $(t,s)$-sequences in the following way. For $\alpha \in \mathbb{N}$, $\alpha \geqslant 2$ let us consider a generic point $\boldsymbol{x} = (x_1,\ldots,x_\alpha) \in [0,1)^\alpha$. We denote the $p$-adic expansion of each $x_j$ by

$$x_j = \frac{\xi_{j,1}}{p} + \frac{\xi_{j,2}}{p^2} + \cdots,$$

which is understood to be unique in the sense that infinitely many of the $\xi_{j,i}$s are different from $p-1$. Then we define the map $\mathscr{D}_\alpha : [0,1)^\alpha \to [0,1)$ by

$$\mathscr{D}_\alpha(x_1,\ldots,x_\alpha) := \sum_{i=1}^{\infty}\sum_{j=1}^{\alpha} \frac{\xi_{j,i}}{p^{\alpha(i-1)+j}}.$$

We extend the map $\mathscr{D}_\alpha$ to vectors by setting

$$\begin{aligned} &\mathscr{D}_\alpha : [0,1)^{\alpha s} \to [0,1)^s, \\ &(x_1,\ldots,x_{\alpha s}) \mapsto \big(\mathscr{D}_\alpha(x_1,\ldots,x_\alpha),\ldots,\mathscr{D}_\alpha(x_{\alpha(s-1)+1},\ldots,x_{\alpha s})\big), \end{aligned}$$

i.e., $\mathscr{D}_\alpha$ is applied to nonoverlapping consecutive $\alpha$ components of $(x_1,\ldots,x_{\alpha s})$. Using this digit interlacing composition $\mathscr{D}_\alpha$ we can construct higher-order digital nets and sequences explicitly as follows.

LEMMA 2.11 Let $\alpha \in \mathbb{N}$, $\alpha \geqslant 2$ and $p$ be a prime.

1. For $m \in \mathbb{N}$ let $P_{m,m}$ be an order 1 digital $(t,m,\alpha s)$-net over $\mathbb{F}_p$. Then

$$\mathscr{D}_\alpha(P_{m,m}) := \big\{\mathscr{D}_\alpha(\boldsymbol{x}) : \boldsymbol{x} \in P_{m,m}\big\} \subset [0,1)^s$$

is an order $\alpha$ digital $(t',m,s)$-net over $\mathbb{F}_p$ with

$$t' = \alpha \min\left\{m, t + \left\lfloor \frac{s(\alpha-1)}{2} \right\rfloor\right\}.$$

2. Let $\mathscr{S}$ be an order 1 digital $(t, \alpha s)$-sequence over $\mathbb{F}_p$. Then

$$\mathscr{D}_\alpha(\mathscr{S}) := \left\{\mathscr{D}_\alpha(\boldsymbol{x}) : \boldsymbol{x} \in \mathscr{S}\right\} \subset [0,1)^s$$

is an order $\alpha$ digital $(t', s)$-sequence over $\mathbb{F}_p$ with

$$t' = \alpha t + \frac{s\alpha(\alpha-1)}{2}.$$

*Proof.* See Baldeaux *et al.* (2011, Corollary 3.4) and Dick (2008, Theorems 4.11 and 4.12) for the proofs of the first and second items, respectively. □

REMARK 2.12 Let $\mathscr{S}$ be an order 1 digital $(t, \alpha s)$-sequence over $\mathbb{F}_p$ with generating matrices $C_1, \ldots, C_{\alpha s} \in \mathbb{F}_p^{\mathbb{N}\times\mathbb{N}}$. We denote the $l$th row of $C_j$ by $\boldsymbol{c}_l^{(j)}$. Then $\mathscr{D}_\alpha(\mathscr{S})$ is a digital sequence over $\mathbb{F}_p$ with generating matrices $D_1, \ldots, D_s \in \mathbb{F}_p^{\mathbb{N}\times\mathbb{N}}$, where each $D_j$ whose $l$th row is denoted by $\boldsymbol{d}_l^{(j)}$ is given by

$$\boldsymbol{d}_{\alpha(l-1)+h}^{(j)} = \boldsymbol{c}_l^{(\alpha(j-1)+h)}$$

for $l \geqslant 1$ and $1 \leqslant h \leqslant \alpha$. If each $C_j = (c_{k,l}^{(j)})$ satisfies $c_{k,l}^{(j)} = 0$ whenever $k > l$, i.e., if $K_j(l) \leqslant l$ holds, each $D_j = (d_{k,l}^{(j)})$ satisfies $d_{k,l}^{(j)} = 0$ whenever $k > \alpha l$. This means, the first $p^m$ points of $\mathscr{D}_\alpha(\mathscr{S})$ are an order $\alpha$ digital $(t', m, s)$-net over $\mathbb{F}_p$ with the precision $n \leqslant \alpha m$.

Since several explicit constructions of order 1 digital $(t, \alpha s)$-sequences, including those of Sobol' (1967) and Tezuka (1993), fulfil the condition $K_j(l) \leqslant l$, we assume that the precision of the first $p^m$ points of an order $\alpha$ digital $(t, s)$-sequence over $\mathbb{F}_p$ is at most $\alpha m$ in the rest of this paper.

### 2.3 *Walsh functions*

Finally, in this section, we recall the definition of Walsh functions, which play a central role in the quadrature error analysis of QMC rules using (higher-order) digital nets and sequences.

DEFINITION 2.13 (Walsh functions). For a prime $p$ we write $\omega_p = \exp(2\pi\sqrt{-1}/p)$. For $k \in \mathbb{N}_0$ whose $p$-adic expansion is given by $k = \kappa_0 + \kappa_1 p + \cdots$, where all but a finite number of $\kappa_i$s are 0, the $k$th Walsh function $\mathrm{wal}_k \colon [0,1) \to \{1, \omega_p, \ldots, \omega_p^{p-1}\}$ is defined by

$$\mathrm{wal}_k(x) := \omega_p^{\kappa_0\xi_1 + \kappa_1\xi_2 + \cdots},$$

where we denote the $p$-adic expansion of $x \in [0,1)$ by $x = \xi_1/p + \xi_2/p^2 + \cdots$, which is understood to be unique in the sense that infinitely many of the $\xi_i$s are different from $p-1$.

In the multivariate case, for $\boldsymbol{k} = (k_1, \ldots, k_s) \in \mathbb{N}_0^s$ and $\boldsymbol{x} = (x_1, \ldots, x_s) \in [0,1)^s$, the $\boldsymbol{k}$th Walsh function is defined by

$$\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}) := \prod_{j=1}^{s} \mathrm{wal}_{k_j}(x_j).$$

LEMMA 2.14 For $k \in \mathbb{N}_0$ and $n \in \mathbb{N}$, we have

$$\frac{1}{p^n}\sum_{h=0}^{p^n-1} \mathrm{wal}_k\left(\frac{h}{p^n}\right) = \begin{cases} 1 & \text{if } p^n \text{ divides } k, \\ 0 & \text{otherwise.} \end{cases}$$

*Proof.* Write $k = k' + p^n l$ for $0 \leqslant k' < p^n$ and $l \geqslant 0$. From the definition of Walsh functions we see that $\mathrm{wal}_k(h/p^n) = \mathrm{wal}_{k'}(h/p^n)$ for any $0 \leqslant h < p^n$. Thus, it suffices to prove the result for the case $0 \leqslant k < p^n$. Actually, the result for $k = 0$ is trivial and the proof for $1 \leqslant k < p^n$ can be found in Dick & Pillichshammer (2010, Lemma A.8). □

LEMMA 2.15 For a prime $p$ and $m, n \in \mathbb{N}$ let $P_{m,n} = \{\boldsymbol{x}_h : 0 \leqslant h < p^m\}$ be a digital net over $\mathbb{F}_p$. For $\boldsymbol{k} \in \mathbb{N}_0^s$, we have

$$\frac{1}{p^m}\sum_{h=0}^{p^m-1} \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}_h) = \begin{cases} 1 & \text{if } \boldsymbol{k} \in P_{m,n}^{\perp}, \\ 0 & \text{otherwise.} \end{cases}$$

*Proof.* See Dick & Pillichshammer (2010, Lemma 4.75) for the proof. □

As shown in Dick & Pillichshammer (2010, Theorem A.11) the system $\{\mathrm{wal}_{\boldsymbol{k}} : \boldsymbol{k} \in \mathbb{N}_0^s\}$ is a complete orthonormal system in $L_2([0,1)^s)$. Therefore, we can define the Walsh series of $f \in L_2([0,1)^s)$

$$\sum_{\boldsymbol{k}\in\mathbb{N}_0^s} \hat{f}(\boldsymbol{k})\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}),$$

where $\hat{f}(\boldsymbol{k})$ is the $\boldsymbol{k}$th Walsh coefficient of $f$:

$$\hat{f}(\boldsymbol{k}) := \int_{[0,1)^s} f(\boldsymbol{x})\overline{\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x}.$$

It is easy to see that $I(f) = \hat{f}(\boldsymbol{0})$.

For smooth functions $f \in W_{s,\alpha,q,r}$ the above Walsh series converges to $f$ point-wise absolutely, and, moreover, the following bounds on the Walsh coefficients are known.

LEMMA 2.16 Let $u$ be a subset of $\{1,\ldots,s\}$ and $\boldsymbol{k}_u \in \mathbb{N}^{|u|}$. The $(\boldsymbol{k}_u, \boldsymbol{0})$th Walsh coefficient of $f \in W_{s,\alpha,q,r}$ is bounded by

$$|\hat{f}(\boldsymbol{k}_u, \boldsymbol{0})| \leqslant \gamma_u \|f_u\|_{s,\alpha,q,r} C_\alpha^{|u|} p^{-\mu_\alpha(\boldsymbol{k}_u)},$$

where $f_u$ is given as in Lemma 2.1 and

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

*Proof.* See Dick (2009, Theorem 15) and Dick *et al.* (2014, Theorem 3.5) for the proof. □

## 3. Extrapolation of truncated higher-order digital nets and sequences

### 3.1 *Euler–Maclaurin formula*

Before providing our extrapolation-based quadrature rules here we show some necessary results as preparation. In what follows, for $l \in \mathbb{N}$ and $\boldsymbol{k} = (k_1, \ldots, k_s) \in \mathbb{N}_0^s$, we write $l \mid \boldsymbol{k}$ if $l$ divides $k_j$ for all $1 \leqslant j \leqslant s$ and $l \nmid \boldsymbol{k}$ if there exists at least one component $k_j$ that is not divided by $l$. Further, we write $\boldsymbol{k} < l$ (resp. $\boldsymbol{k} > l$) if $k_j < l$ (resp. $k_j > l$) holds for all $1 \leqslant j \leqslant s$.

Lemma 3.1 For a prime $p$ and $m, n \in \mathbb{N}$ let $P_{m,n}$ be a digital net over $\mathbb{F}_p$ with generating matrices $C_1, \ldots, C_s \in \mathbb{F}_p^{n\times m}$. Then we have

$$\left\{\boldsymbol{k} \in \mathbb{N}_0^s : p^n \mid \boldsymbol{k}\right\} \subseteq P_{m,n}^{\perp},$$

and

$$P_{m,n}^{\perp} \setminus \left\{\boldsymbol{k} \in \mathbb{N}_0^s : p^n \mid \boldsymbol{k}\right\} = \left\{\boldsymbol{k} + p^n \boldsymbol{l} : \boldsymbol{k} \in P_{m,n}^{\perp}, \boldsymbol{0} \neq \boldsymbol{k} < p^n, \boldsymbol{l} \in \mathbb{N}_0^s\right\}.$$

*Proof.* The first statement is trivial, since for $\boldsymbol{k} \in \mathbb{N}_0^s$ such that $p^n \mid \boldsymbol{k}$, we have $\nu_n(k_1) = \cdots = \nu_n(k_s) = (0, \ldots, 0)^\top$ that gives

$$C_1^\top \nu_n(k_1) \oplus \cdots \oplus C_s^\top \nu_n(k_s) = \boldsymbol{0}.$$

Hence, $P_{m,n}^{\perp}$ always contains such $\boldsymbol{k}$ as elements.

Let us move on to the proof of the second statement. For $k, l \in \mathbb{N}_0$ with $k < p^n$, we have $\nu_n(k+p^n l) = \nu_n(k)$. This means that, for $\boldsymbol{k}, \boldsymbol{l} \in \mathbb{N}_0^s$ with $\boldsymbol{k} < p^n$, whether $P_{m,n}^{\perp}$ contains $\boldsymbol{k} + p^n \boldsymbol{l}$ as an element does not depend on $\boldsymbol{l}$, so that $\boldsymbol{k} + p^n \boldsymbol{l} \in P_{m,n}^{\perp}$ if and only if $\boldsymbol{k} \in P_{m,n}^{\perp}$. Therefore, we have

$$\begin{aligned} P_{m,n}^{\perp} &= \{\boldsymbol{k} + p^n \boldsymbol{l} : \boldsymbol{k} \in P_{m,n}^{\perp}, \boldsymbol{k} < p^n, \boldsymbol{l} \in \mathbb{N}_0^s\} \\ &= \{\boldsymbol{k} + p^n \boldsymbol{l} : \boldsymbol{k} \in P_{m,n}^{\perp}, \boldsymbol{0} \neq \boldsymbol{k} < p^n, \boldsymbol{l} \in \mathbb{N}_0^s\} \cup \{p^n \boldsymbol{l} : \boldsymbol{l} \in \mathbb{N}_0^s\}, \end{aligned}$$

where the last equality follows by separating the cases $\boldsymbol{k} \neq \boldsymbol{0}$ and $\boldsymbol{k} = \boldsymbol{0}$. Since the two sets on the right-most side above are disjoint, the result follows. □

Corollary 3.2 For a prime $p$ and $m, n \in \mathbb{N}$ let $P_{m,n} = \{\boldsymbol{x}_h : 0 \leqslant h < p^m\}$ be a digital net over $\mathbb{F}_p$ with generating matrices $C_1, \ldots, C_s \in \mathbb{F}_p^{n\times m}$. For $f \in W_{s,\alpha,q,r}$, we have

$$\frac{1}{p^m} \sum_{h=0}^{p^m-1} f(\boldsymbol{x}_h) = I(f) + \sum_{\boldsymbol{l} \in \mathbb{N}_0^s} \sum_{\substack{\boldsymbol{k} \in P_{m,n}^{\perp} \\ \boldsymbol{0} \neq \boldsymbol{k} < p^n}} \hat{f}(\boldsymbol{k} + p^n \boldsymbol{l}) + \sum_{\tau=1}^{\alpha-1} \frac{c_\tau(f)}{p^{\tau n}} + R_{s,\alpha,n}, \tag{3.1}$$

where $c_\tau(f)$ depends only on $f$ and $\tau$ and $R_{s,\alpha,n} \in O(p^{-\alpha n})$.

*Proof.* Using the Walsh series of $f$, Lemma 2.15 and Lemma 3.1, we have

$$
\begin{aligned}
\frac{1}{p^m}\sum_{h=0}^{p^m-1} f(\boldsymbol{x}_h) &= \frac{1}{p^m}\sum_{h=0}^{p^m-1}\sum_{\boldsymbol{k}\in\mathbb{N}_0^s}\hat{f}(\boldsymbol{k})\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}_h) = \sum_{\boldsymbol{k}\in\mathbb{N}_0^s}\hat{f}(\boldsymbol{k})\frac{1}{p^m}\sum_{h=0}^{p^m-1}\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}_h)\\
&= \sum_{\boldsymbol{k}\in P_{m,n}^{\perp}}\hat{f}(\boldsymbol{k}) = I(f) + \sum_{\boldsymbol{k}\in P_{m,n}^{\perp}\setminus\{\boldsymbol{0}\}}\hat{f}(\boldsymbol{k})\\
&= I(f) + \sum_{\boldsymbol{l}\in\mathbb{N}_0^s}\sum_{\substack{\boldsymbol{k}\in P_{m,n}^{\perp}\\ \boldsymbol{0}\neq\boldsymbol{k}<p^n}}\hat{f}(\boldsymbol{k}+p^n\boldsymbol{l}) + \sum_{\substack{\boldsymbol{k}\in\mathbb{N}_0^s\setminus\{\boldsymbol{0}\}\\ p^n|\boldsymbol{k}}}\hat{f}(\boldsymbol{k}). \qquad (3.2)
\end{aligned}
$$

Now we write

$$
P_n^* := \left\{\left(\frac{h_1}{p^n},\ldots,\frac{h_s}{p^n}\right) : 0 \leqslant h_1,\ldots,h_s < p^n\right\},
$$

which is called a *regular grid*. Using Lemma 2.14 for $\boldsymbol{k}\in\mathbb{N}_0^s$, we have

$$
\frac{1}{p^{ns}}\sum_{h_1,\ldots,h_s=0}^{p^n-1}\mathrm{wal}_{\boldsymbol{k}}\left(\frac{h_1}{p^n},\ldots,\frac{h_s}{p^n}\right) = \prod_{j=1}^{s}\frac{1}{p^n}\sum_{h_j=0}^{p^n-1}\mathrm{wal}_{k_j}\left(\frac{h_j}{p^n}\right) = \begin{cases}1 & \text{if } p^n \mid \boldsymbol{k},\\ 0 & \text{otherwise.}\end{cases}
$$

Using this result, we obtain

$$
\begin{aligned}
\sum_{\substack{\boldsymbol{k}\in\mathbb{N}_0^s\setminus\{\boldsymbol{0}\}\\ p^n|\boldsymbol{k}}}\hat{f}(\boldsymbol{k}) &= \sum_{\boldsymbol{k}\in\mathbb{N}_0^s\setminus\{\boldsymbol{0}\}}\hat{f}(\boldsymbol{k})\frac{1}{p^{ns}}\sum_{h_1,\ldots,h_s=0}^{p^n-1}\mathrm{wal}_{\boldsymbol{k}}\left(\frac{h_1}{p^n},\ldots,\frac{h_s}{p^n}\right)\\
&= \frac{1}{p^{ns}}\sum_{h_1,\ldots,h_s=0}^{p^n-1}\sum_{\boldsymbol{k}\in\mathbb{N}_0^s\setminus\{\boldsymbol{0}\}}\hat{f}(\boldsymbol{k})\mathrm{wal}_{\boldsymbol{k}}\left(\frac{h_1}{p^n},\ldots,\frac{h_s}{p^n}\right)\\
&= \frac{1}{p^{ns}}\sum_{h_1,\ldots,h_s=0}^{p^n-1} f\left(\frac{h_1}{p^n},\ldots,\frac{h_s}{p^n}\right) - I(f).
\end{aligned}
$$

This means that the last term of (3.2) is nothing but a signed integration error of a QMC rule using a regular grid $P_n^*$ as quadrature nodes. It is shown by Dick *et al.* (2019, Theorem 3.4) that

$$
\frac{1}{p^{ns}}\sum_{h_1,\ldots,h_s=0}^{p^n-1} f\left(\frac{h_1}{p^n},\ldots,\frac{h_s}{p^n}\right) - I(f) = \sum_{\tau=1}^{\alpha-1}\frac{c_\tau(f)}{p^{\tau n}} + R_{s,\alpha,n}, \qquad (3.3)
$$

where

$$c_\tau(f) = \sum_{\substack{\boldsymbol{\tau}\in\{0,1,\dots,\alpha-1\}^s \\ |\boldsymbol{\tau}|_1=\tau}} I(f^{(\boldsymbol{\tau})}) \prod_{\substack{j=1 \\ \tau_j\neq 0}}^{s} b_{\tau_j},$$

with $|\boldsymbol{\tau}|_1 = |\tau_1| + \cdots + |\tau_s|$, and

$$\begin{aligned} |R_{s,\alpha,n}| &\leqslant \frac{\|f\|_{s,\alpha,q,r}}{p^{\alpha n}} \left[ \sum_{\emptyset\neq u\subseteq\{1,\dots,s\}} \left( \gamma_u (\alpha+1)^{|u|/q'} D_\alpha^{|u|} \right)^{r'} \right]^{1/r'} \\ &\leqslant \frac{\|f\|_{s,\alpha,q,r}}{p^{\alpha n}} \sum_{\emptyset\neq u\subseteq\{1,\dots,s\}} \gamma_u (\alpha+1)^{|u|} D_\alpha^{|u|}, \end{aligned} \tag{3.4}$$

with $q'$ and $r'$ being the Hölder conjugates of $q$ and $r$, respectively, and

$$D_\alpha = \max \left\{ |b_1|, \dots, |b_{\alpha-1}|, \sup_{x\in[0,1)} |\tilde{b}_\alpha(x)| \right\}.$$

We complete the proof by substituting (3.3) into the last term of (3.2). □

### 3.2 *An algorithm and its worst-case error bound*

Throughout this subsection let $\mathscr{S}$ be an order $\alpha$ digital $(t,s)$-sequence over $\mathbb{F}_p$ with generating matrices $C_1, \dots, C_s$. For $m, n \in \mathbb{N}$ we denote the upper-left $n \times m$ submatrices of $C_1, \dots, C_s$ by $C_1^{[n\times m]}, \dots, C_s^{[n\times m]}$ and denote a digital net with generating matrices $C_1^{[n\times m]}, \dots, C_s^{[n\times m]}$ by $P^{[n\times m]}$. In view of Remarks 2.3 and 2.12, we assume that

$$P^{[\alpha m\times m]} = P^{[(\alpha m+1)\times m]} = \cdots = P^{[\mathbb{N}\times m]},$$

where the right-most side denotes the first $p^m$ points of $\mathscr{S}$. It is easy to see that, for a finite $n$, we have

$$P^{[n\times m]} = \mathrm{tr}_n(P^{[\mathbb{N}\times m]}),$$

where the map $\mathrm{tr}_n$ is defined as in (1.1), and from Remark 2.6 we also have

$$\left\{ \boldsymbol{k} \in (P^{[n\times m]})^\perp : \boldsymbol{k} < p^n \right\} = \left\{ \boldsymbol{k} \in (P^{[\mathbb{N}\times m]})^\perp : \boldsymbol{k} < p^n \right\}. \tag{3.5}$$

Furthermore, instead of $A_N(f)$, we write

$$I(f;P) = \frac{1}{N} \sum_{\boldsymbol{x}\in P} f(\boldsymbol{x})$$

for an $N$-element point set $P \subset [0,1)^s$ to emphasize which point set is used in numerical integration.

Now let us consider the following algorithm:

ALGORITHM 3.3 Let $\mathscr{S}$ be an order $\alpha$ digital $(t,s)$-sequence over $\mathbb{F}_p$. For $m \in \mathbb{N}$ and $f\colon [0,1)^s \to \mathbb{R}$ do the following:

1. For $0 \leqslant i < \alpha$, compute

$$I^{(1)}_{m+i}(f) := I\left(f; P^{[(m+i)\times(m+i)]}\right).$$

2. For $1 \leqslant \tau < \alpha$, let

$$I^{(\tau+1)}_{m+i}(f) := \frac{p^\tau I^{(\tau)}_{m+i+1}(f) - I^{(\tau)}_{m+i}(f)}{p^\tau - 1} \quad \text{for } 0 \leqslant i < \alpha - \tau.$$

3. Return $I^{(\alpha)}_m(f)$ as an approximation of $I(f)$.

We emphasize that Algorithm 3.3 uses only digital nets with *square* generating matrices, which significantly reduces the necessary precision of points from $\alpha m$ (see Remarks 2.9 and 2.12) to $m$. Since the resulting estimate $I^{(\alpha)}_m(f)$ is given by a weighted sum of QMC rules with different sizes of nodes, $I^{(1)}_m(f), \ldots, I^{(1)}_{m+\alpha-1}(f)$, this quadrature rule is a linear algorithm with the total number of function evaluations

$$N = p^m + \cdots + p^{m+\alpha-1}.$$

As a main result of this paper, we show that our quadrature rule $I^{(\alpha)}_m(f)$ achieves the almost optimal rate of convergence of the worst-case error in $W_{s,\alpha,q,r}$.

THEOREM 3.4 Let $\alpha \in \mathbb{N}$, $\alpha \geqslant 2$ and $1 \leqslant q, r \leqslant \infty$. When $\alpha m > t$ holds the worst-case error of the algorithm $I^{(\alpha)}_m(f)$ in $W_{s,\alpha,q,r}$ is bounded above by

$$\sup_{\substack{f \in W_{s,\alpha,q,r} \\ \|f\|_{s,\alpha,q,r} \leqslant 1}} \left| I(f) - I^{(\alpha)}_m(f) \right| \leqslant \sum_{\emptyset \neq u \subseteq \{1,\ldots,s\}} \gamma_u U_{|u|,\alpha,t} \frac{(\log_p N)^{\alpha|u|}}{N^\alpha},$$

where $N = p^m + \cdots + p^{m+\alpha-1}$ and $U_{|u|,\alpha,t} > 0$ for all $\emptyset \neq u \subseteq \{1,\ldots,s\}$.

In order to prove Theorem 3.4 we need some preparations.

LEMMA 3.5 Let $P_{m,\alpha m}$ be an order $\alpha$ digital $(t,m,s)$-net over $\mathbb{F}_p$ or be the first $p^m$ points of an order $\alpha$ digital $(t,s)$-sequence over $\mathbb{F}_p$ such that $\alpha m > t$. For a nonempty subset $u \subseteq \{1,\ldots,s\}$, we write

$$(P_{m,\alpha m})^\perp_u = \{\boldsymbol{k}_u \in \mathbb{N}^{|u|} : (\boldsymbol{k}_u, \mathbf{0}) \in P^\perp_{m,\alpha m}\}.$$

Then we have

$$\sum_{\boldsymbol{k}_u \in (P_{m,\alpha m})^\perp_u} p^{-\mu_\alpha(\boldsymbol{k}_u)} \leqslant E_{|u|,\alpha} \frac{(\alpha m - t + 2)^{\alpha|u|}}{p^{\alpha m - t}},$$

where

$$E_{|u|,\alpha} = p^{\alpha|u|}\left(\frac{1}{p} + \left(\frac{p}{p-1}\right)^{\alpha|u|}\right).$$

*Proof.* See Dick (2008, Lemma 5.2) and Dick & Pillichshammer (2010, Lemma 15.20) for the proof. □

LEMMA 3.6 For $n \in \mathbb{N}$ and $\boldsymbol{k}, \boldsymbol{l} \in \mathbb{N}_0^s$ with $\boldsymbol{k} < p^n$, we have

$$\mu_\alpha(\boldsymbol{k} + p^n \boldsymbol{l}) \geqslant \mu_\alpha(\boldsymbol{k}) + \mu_\alpha(\boldsymbol{l}).$$

*Proof.* Noting that

$$\mu_\alpha(\boldsymbol{k} + p^n \boldsymbol{l}) = \sum_{j=1}^{s} \mu_\alpha(k_j + p^n l_j)$$

and

$$\mu_\alpha(\boldsymbol{k}) + \mu_\alpha(\boldsymbol{l}) = \sum_{j=1}^{s} \left(\mu_\alpha(k_j) + \mu_\alpha(l_j)\right),$$

it suffices to prove the statement for the one-dimensional case:

$$\mu_\alpha(k + p^n l) \geqslant \mu_\alpha(k) + \mu_\alpha(l),$$

for any $k, l \in \mathbb{N}_0$ with $k < p^n$. Since the result is trivial if $l = 0$ we assume $l > 0$. We denote the $p$-adic expansions of $k$ and $l$ by

$$k = \kappa_1 p^{c_1 - 1} + \cdots + \kappa_v p^{c_v - 1},$$
$$l = \iota_1 p^{d_1 - 1} + \cdots + \iota_w p^{d_w - 1},$$

with $\kappa_1, \ldots, \kappa_v, \iota_1, \ldots, \iota_w \in \{1, \ldots, p-1\}$, $c_1 > \cdots > c_v > 0$ and $d_1 > \cdots > d_w > 0$, respectively. Since $k < p^n$ we have $c_1 \leqslant n$ and $v \leqslant n$. If $w < \alpha$, we have

$$\begin{aligned}\mu_\alpha(k + p^n l) &= \mu_\alpha(\kappa_1 p^{c_1-1} + \cdots + \kappa_v p^{c_v-1} + \iota_1 p^{d_1+n-1} + \cdots + \iota_w p^{d_w+n-1}) \\ &= \sum_{i=1}^{w}(d_i + n) + \sum_{i=1}^{\min(\alpha-w,v)} c_i \geqslant \sum_{i=1}^{w} d_i + \sum_{i=1}^{\min(\alpha,v)} c_i = \mu_\alpha(l) + \mu_\alpha(k).\end{aligned}$$

On the other hand, if $w \geqslant \alpha$, we have

$$\mu_\alpha(k + p^n l) = \sum_{i=1}^{\alpha}(d_i + n) \geq \sum_{i=1}^{\alpha} d_i + \sum_{i=1}^{\min(\alpha,v)} c_i = \mu_\alpha(l) + \mu_\alpha(k).$$

Thus, we complete the proof. □

Now we are ready to prove Theorem 3.4.

*Proof of Theorem* 3.4 Let $f \in W_{s,\alpha,q,r}$. For each $0 \leqslant i < \alpha$ Corollary 3.2 gives

$$I^{(1)}_{m+i}(f) = I(f) + \sum_{\boldsymbol{l} \in \mathbb{N}_0^s} \sum_{\substack{\boldsymbol{k} \in (P^{[(m+i)\times(m+i)]})^{\perp} \\ \boldsymbol{0} \neq \boldsymbol{k} < p^{m+i}}} \hat{f}(\boldsymbol{k} + p^{m+i}\boldsymbol{l}) + \sum_{\tau=1}^{\alpha-1} \frac{c_\tau(f)}{p^{\tau(m+i)}} + R_{s,\alpha,m+i}.$$

Using the result shown in Dick *et al.* (2019, Lemma 2.9 and Corollary 2.11), we have

$$I^{(\alpha)}_{m}(f) = I(f) + \sum_{i=0}^{\alpha-1} w_i \left( \sum_{\boldsymbol{l} \in \mathbb{N}_0^s} \sum_{\substack{\boldsymbol{k} \in (P^{[(m+i)\times(m+i)]})^{\perp} \\ \boldsymbol{0} \neq \boldsymbol{k} < p^{m+i}}} \hat{f}(\boldsymbol{k} + p^{m+i}\boldsymbol{l}) + R_{s,\alpha,m+i} \right),$$

with

$$w_i = \prod_{j=1}^{\alpha-i-1} \left( \frac{-1}{p^j - 1} \right) \prod_{j=1}^{i} \left( \frac{p^j}{p^j - 1} \right) \quad \text{for } 0 \leqslant i \leqslant \alpha - 1,$$

where the empty product is set to 1. Here we note that $\sum_{i=0}^{\alpha-1} w_i = 1$ (see Dick *et al.* (2019, Lemma 2.10). It follows from the triangle inequality and the decomposition

$$(P^{[(m+i)\times(m+i)]})^{\perp} \setminus \{\boldsymbol{0}\} = \bigcup_{\emptyset \neq u \subseteq \{1,\dots,s\}} (P^{[(m+i)\times(m+i)]})^{\perp}_u,$$

that

$$\begin{aligned}
&\left| I^{(\alpha)}_{m}(f) - I(f) \right| \\
&\leqslant \sum_{i=0}^{\alpha-1} |w_i| \left( \sum_{\boldsymbol{l} \in \mathbb{N}_0^s} \sum_{\substack{\boldsymbol{k} \in (P^{[(m+i)\times(m+i)]})^{\perp} \\ \boldsymbol{0} \neq \boldsymbol{k} < p^{m+i}}} |\hat{f}(\boldsymbol{k} + p^{m+i}\boldsymbol{l})| + |R_{s,\alpha,m+i}| \right) \\
&= \sum_{i=0}^{\alpha-1} |w_i| \left( \sum_{\emptyset \neq u \subseteq \{1,\dots,s\}} \sum_{v \subseteq u} \sum_{\boldsymbol{l}_v \in \mathbb{N}^{|v|}} \sum_{\substack{(\boldsymbol{k}_v, \boldsymbol{k}_{u\setminus v}) \in (P^{[(m+i)\times(m+i)]})^{\perp}_u \\ \boldsymbol{0} \neq \boldsymbol{k}_v < p^{m+i} \\ 0 < \boldsymbol{k}_{u\setminus v} < p^{m+i}}} |\hat{f}(\boldsymbol{k}_v + p^{m+i}\boldsymbol{l}_v, \boldsymbol{k}_{u\setminus v}, \boldsymbol{0})| + |R_{s,\alpha,m+i}| \right) \\
&\leqslant \sum_{i=0}^{\alpha-1} |w_i| \left( \sum_{\emptyset \neq u \subseteq \{1,\dots,s\}} \sum_{\boldsymbol{l}_u \in \mathbb{N}_0^{|u|}} \sum_{\substack{\boldsymbol{k}_u \in (P^{[(m+i)\times(m+i)]})^{\perp}_u \\ \boldsymbol{0} \neq \boldsymbol{k}_u < p^{m+i}}} |\hat{f}(\boldsymbol{k}_u + p^{m+i}\boldsymbol{l}_u, \boldsymbol{0})| + |R_{s,\alpha,m+i}| \right).
\end{aligned}$$

Using Lemmas 2.16 and 3.6 the inner double sum above for a given $\emptyset \neq u \subseteq \{1,\dots,s\}$ is bounded by

$$\sum_{\boldsymbol{l}_u \in \mathbb{N}_0^{|u|}} \sum_{\substack{\boldsymbol{k}_u \in (P^{[(m+i)\times(m+i)]})_u^\perp \\ \boldsymbol{0}\neq \boldsymbol{k}_u < p^{m+i}}} |\hat{f}(\boldsymbol{k}_u + p^{m+i}\boldsymbol{l}_u, \boldsymbol{0})| \leqslant \gamma_u \|f_u\|_{s,\alpha,q,r} C_\alpha^{|u|} \sum_{\boldsymbol{l}_u \in \mathbb{N}_0^{|u|}} \sum_{\substack{\boldsymbol{k}_u \in (P^{[(m+i)\times(m+i)]})_u^\perp \\ \boldsymbol{k}_u < p^{m+i}}} p^{-\mu_\alpha(\boldsymbol{k}_u + p^{m+i}\boldsymbol{l}_u)}$$

$$\leqslant \gamma_u \|f\|_{s,\alpha,q,r} C_\alpha^{|u|} \sum_{\boldsymbol{l}_u \in \mathbb{N}_0^{|u|}} p^{-\mu_\alpha(\boldsymbol{l}_u)} \sum_{\substack{\boldsymbol{k}_u \in (P^{[(m+i)\times(m+i)]})_u^\perp \\ \boldsymbol{k}_u < p^{m+i}}} p^{-\mu_\alpha(\boldsymbol{k}_u)}.$$

Applying (3.5) and Lemma 3.5, the inner sum over $\boldsymbol{k}_u$ is bounded by

$$\sum_{\substack{\boldsymbol{k}_u \in (P^{[(m+i)\times(m+i)]})_u^\perp \\ \boldsymbol{k}_u < p^{m+i}}} p^{-\mu_\alpha(\boldsymbol{k}_u)} = \sum_{\substack{\boldsymbol{k}_u \in (P^{[(m+i)\times(m+i)]})_u^\perp \\ \boldsymbol{k}_u < p^{m+i}}} p^{-\mu_\alpha(\boldsymbol{k}_u)} \leqslant \sum_{\boldsymbol{k}_u \in (P^{[\mathbb{N}\times(m+i)]})_u^\perp} p^{-\mu_\alpha(\boldsymbol{k}_u)}$$

$$\leqslant E_{|u|,\alpha} \frac{(\alpha(m+i)-t+2)^{\alpha|u|}}{p^{\alpha(m+i)-t}}.$$

Regarding the sum over $\boldsymbol{l}_u \in \mathbb{N}_0^{|u|}$ Goda (2016, Lemma 7) gives

$$\sum_{\boldsymbol{l}_u \in \mathbb{N}_0^{|u|}} p^{-\mu_\alpha(\boldsymbol{l}_u)} = \left(\sum_{l \in \mathbb{N}_0} p^{-\mu_\alpha(l)}\right)^{|u|} = A_\alpha^{|u|},$$

where

$$A_\alpha = 1 + \sum_{w=1}^{\alpha-1} \prod_{i=1}^{w} \left(\frac{p-1}{p^i-1}\right) + \left(\frac{p^\alpha - 1}{p^\alpha - p}\right) \prod_{i=1}^{\alpha} \left(\frac{p-1}{p^i-1}\right).$$

All together, the inner double sum for $\emptyset \neq u \subseteq \{1,\dots,s\}$ is bounded by

$$\sum_{\boldsymbol{l}_u \in \mathbb{N}_0^{|u|}} \sum_{\substack{\boldsymbol{k}_u \in (P^{[(m+i)\times(m+i)]})_u^\perp \\ 0<\boldsymbol{k}_u < p^{m+i}}} |\hat{f}(\boldsymbol{k}_u + p^{m+i}\boldsymbol{l}_u, \boldsymbol{0})| \leqslant \gamma_u \|f\|_{s,\alpha,q,r} A_\alpha^{|u|} C_\alpha^{|u|} E_{|u|,\alpha} \frac{(\alpha(m+i)-t+2)^{\alpha|u|}}{p^{\alpha(m+i)-t}}.$$

Recall that the total number of function evaluations is $N = p^m + \cdots + p^{m+\alpha-1}$. Using the above result and (3.4) the integration error for $f \in W_{s,\alpha,q,r}$ is bounded by

$$\left|I_m^{(\alpha)}(f) - I(f)\right| \leqslant \sum_{i=0}^{\alpha-1} |w_i| \sum_{\emptyset\neq u\subseteq\{1,\dots,s\}} \gamma_u \|f\|_{s,\alpha,q,r} \frac{p^t A_\alpha^{|u|} C_\alpha^{|u|} E_{|u|,\alpha} (\alpha(m+i)-t+2)^{\alpha|u|} + (\alpha+1)^{|u|} D_\alpha^{|u|}}{p^{\alpha(m+i)}}$$

$$\leqslant \|f\|_{s,\alpha,q,r} \sum_{\emptyset\neq u\subseteq\{1,\dots,s\}} \gamma_u \frac{\left(\log_p N\right)^{\alpha|u|}}{N^\alpha}$$

$$\times \left(p^t A_\alpha^{|u|} C_\alpha^{|u|} E_{|u|,\alpha} + (\alpha+1)^{|u|} D_\alpha^{|u|}\right) \sum_{i=0}^{\alpha-1} |w_i| \frac{N^\alpha}{p^{\alpha(m+i)}} \frac{(\alpha(m+i)-t+2)^{\alpha|u|}}{\left(\log_p N\right)^{\alpha|u|}}. \quad (3.6)$$

For any $0 \leqslant i < \alpha$, we have

$$\frac{N^\alpha}{p^{\alpha(m+i)}} \leqslant \frac{(\alpha p^{m+\alpha-1})^\alpha}{p^{\alpha(m+i)}} = \left(\alpha p^{\alpha+1-i}\right)^\alpha,$$

and

$$\begin{aligned}(\alpha(m+i)-t+2)^{\alpha|u|} &\leqslant (\alpha(m+i+2/\alpha))^{\alpha|u|} \leqslant (\alpha(m+i+1))^{\alpha|u|} \\ &\leqslant (2\alpha(m+i))^{\alpha|u|} \leqslant \left(2\alpha \log_p N\right)^{\alpha|u|}.\end{aligned}$$

Thus, the inner sum of (3.6) is bounded independently of $m$ as

$$\sum_{i=0}^{\alpha-1} |w_i| \frac{N^\alpha}{\left(\log_p N\right)^{\alpha|u|}} \frac{(\alpha(m+i)-t+2)^{\alpha|u|}}{p^{\alpha(m+i)}} \leqslant (2\alpha)^{\alpha|u|} \sum_{i=0}^{\alpha-1} |w_i| \left(\alpha p^{\alpha+1-i}\right)^\alpha.$$

This leads to a worst-case error bound:

$$\sup_{\substack{f\in W_{s,\alpha,q,r}\\ \|f\|_{s,\alpha,q,r}\leqslant 1}} \left| I_m^{(\alpha)}(f) - I(f)\right| \leqslant \sum_{\emptyset\neq u\subseteq\{1,\dots,s\}} \gamma_u U_{|u|,\alpha,t} \frac{\left(\log_p N\right)^{\alpha|u|}}{N^\alpha},$$

where

$$U_{|u|,\alpha,t} = (2\alpha)^{\alpha|u|}\left(p^t A_\alpha^{|u|} C_\alpha^{|u|} E_{|u|,\alpha} + (\alpha+1)^{|u|} D_\alpha^{|u|}\right) \sum_{i=0}^{\alpha-1} |w_i| \left(\alpha p^{\alpha+1-i}\right)^\alpha.$$

Hence, we complete the proof.

REMARK 3.7 Algorithm 3.3 is naturally extensible with respect to $m$ in the following way: for $m_{\min}, m_{\max} \in \mathbb{N}$, $m_{\max} - m_{\min} \geqslant \alpha$ and $f\colon [0,1)^s \to \mathbb{R}$ do the following:

1. For $m_{\min} \leqslant i \leqslant m_{\max}$ compute
$$I_i^{(1)}(f) := I\left(f; P^{[i\times i]}\right).$$
2. For $1 \leqslant \tau < \alpha$ let
$$I_i^{(\tau+1)}(f) := \frac{p^\tau I_{i+1}^{(\tau)}(f) - I_i^{(\tau)}(f)}{p^\tau - 1} \quad \text{for } m_{\min} \leqslant i \leqslant m_{\max} - \tau.$$

Then we obtain a sequence of the approximate values $I_{m_{\min}}^{(\alpha)}, I_{m_{\min}+1}^{(\alpha)}, \dots, I_{m_{\max}-\alpha+1}^{(\alpha)}$. If one wants to increase $m_{\max}$ by 1 it suffices to compute $I_{m_{\max}+1}^{(1)}(f)$ instead of whole $I_{m_{\max}-\alpha+1}^{(1)}(f), \dots, I_{m_{\max}+1}^{(1)}(f)$. This is a key advantage as compared to another possible algorithm, which we introduce below.

### 3.3 *Another possible algorithm*

It is clear from the proof of Theorem 3.4 that, in order to vanish the main terms of (3.1), i.e.,

$$\sum_{\tau=1}^{\alpha-1} \frac{c_\tau(f)}{p^{\tau n}}$$

by applying Richardson extrapolation recursively, there is no need to set $n = m$ and to change them at the same time. In fact, we can fix $m$ and change $n$ only, although the resulting algorithm is no longer extensible in $m$.

ALGORITHM 3.8 Let $P_{m,\alpha m}$ be an order $\alpha$ digital $(t, m, s)$-net over $\mathbb{F}_p$ with generating matrices $C_1, \ldots, C_s \in \mathbb{F}_p^{\alpha m\times m}$. For $f\colon [0,1)^s \to \mathbb{R}$ do the following:

1. For $0 \leqslant i < \alpha$ compute

$$J_{m+i}^{(1)} := I\left(f; P_{m,\alpha m}^{[(m+i)\times m]}\right),$$

where $P_{m,\alpha m}^{[(m+i)\times m]}$ denotes a digital net with generating matrices $C_1^{[(m+i)\times m]}, \ldots, C_s^{[(m+i)\times m]}$.

2. For $1 \leqslant \tau < \alpha$ let

$$J_{m+i}^{(\tau+1)} := \frac{p^\tau J_{m+i+1}^{(\tau)} - J_{m+i}^{(\tau)}}{p^\tau - 1} \quad \text{for } 0 \leqslant i < \alpha - \tau.$$

3. Return $J_m^{(\alpha)}$ as an approximation of $I(f)$.

We see that the total number of function evaluations used in $J_m^{(\alpha)}$ is $N = \alpha p^m$. Similarly to Algorithm 3.3, this alternative algorithm achieves the almost optimal rate of convergence as shown below. Since we can prove the result exactly in the same way as Theorem 3.4 we omit the proof.

THEOREM 3.9 Let $\alpha \in \mathbb{N}$, $\alpha \geqslant 2$ and $1 \leqslant q, r \leqslant \infty$. The worst-case error of the algorithm $J_m^{(\alpha)}(f)$ in $W_{s,\alpha,q,r}$ is bounded above by

$$\sup_{\substack{f\in W_{s,\alpha,q,r}\\ \|f\|_{s,\alpha,q,r}\leqslant 1}} \left|I(f) - J_m^{(\alpha)}(f)\right| \leqslant \sum_{\emptyset\neq u\subseteq\{1,\ldots,s\}} \gamma_u V_{|u|,\alpha,t} \frac{(\log_p N)^{\alpha|u|}}{N^\alpha},$$

where $N = \alpha p^m$ and

$$V_{|u|,\alpha,t} = \left(\frac{\alpha}{\log_p 2}\right)^{\alpha|u|} \left(p^t A_\alpha^{|u|} C_\alpha^{|u|} E_{|u|,\alpha} + (\alpha+1)^{|u|} D_\alpha^{|u|}\right) \sum_{i=0}^{\alpha-1} |w_i|,$$

for all $\emptyset \neq u \subseteq \{1, \ldots, s\}$.

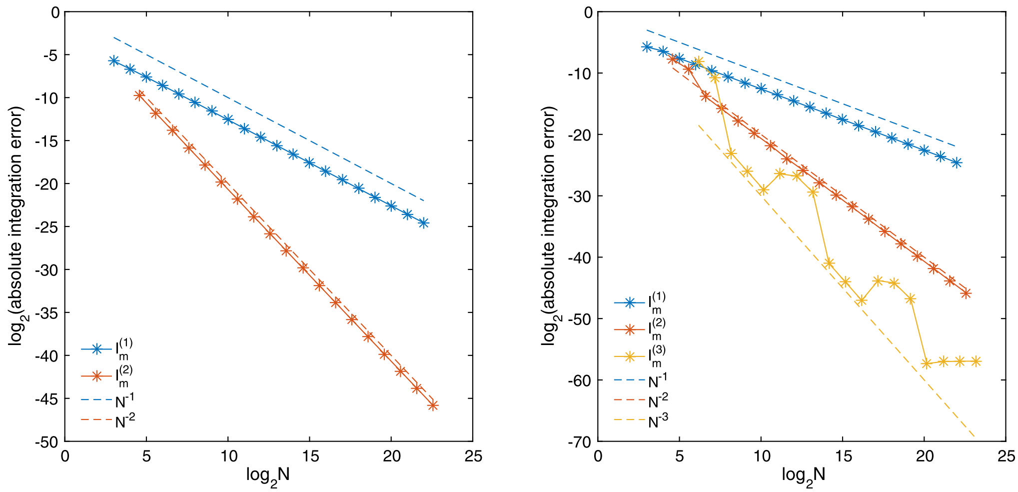


FIG. 1. Integration error for $f_1$: $\alpha = 2$ (left) and $\alpha = 3$ (right).

## 4. Numerical experiments

Finally, we conduct some numerical experiments to confirm the effectiveness of our extrapolation-based quadrature rules. For all the experiments we set the base $p = 2$ and use a MATLAB implementation of higher-order Sobol' nets and sequences from Dick (2007, 2008).

### 4.1 *Low-dimensional cases*

First, let us consider the simplest case $s = 1$. The test function we use is

$$f_1(x) = x^3 \left( \log x + \frac{1}{4} \right).$$

While the third derivative of $f_1$ is in $L_q([0, 1))$ for any finite $q \geqslant 1$ the fourth derivative is not in $L_1([0, 1))$, implying that $f_1 \notin W_{1,4,q,r}$ but $f_1 \in W_{1,3,q,r}$. Note that $I(f_1) = 0$. Figure 1 shows the absolute integration error obtained by using Algorithm 3.3 (and Remark 3.7) with $\alpha = 2$ and $\alpha = 3$. In both cases the integration error of $I_m^{(1)}$ achieves the convergence of nearly order $N^{-1}$. We can see that the order of convergence of the integration error is improved from $N^{-1}$ to $N^{-2}$ by applying Richardson extrapolation. In case of $\alpha = 3$ the recursive application of Richardson extrapolation further improves the order of convergence to approximately $N^{-3}$. This convergence behaviour is in good agreement with our theoretical result.

Next let us consider a bi-variate test function

$$f_2(x, y) = \left( \frac{1}{2} - xy \right)^6 \mathbf{1}_{xy \leqslant 1/2},$$

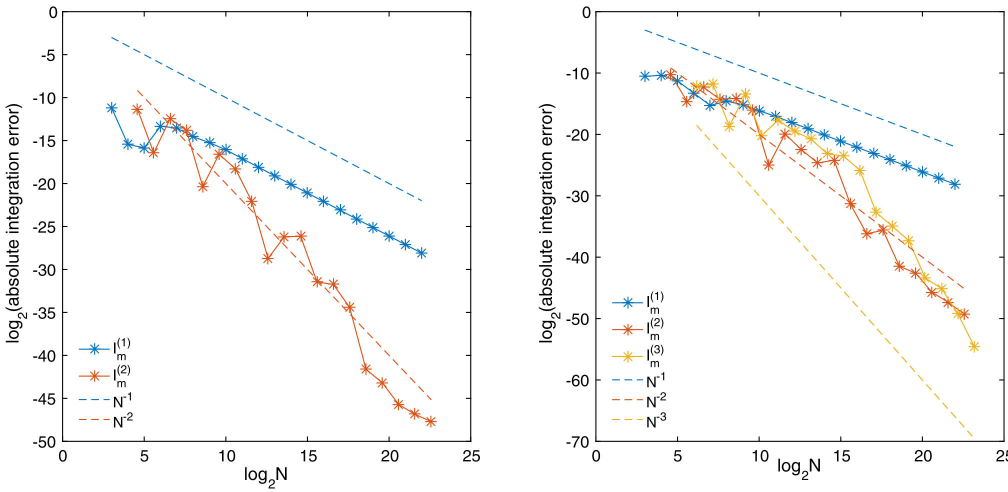


FIG. 2. Integration error for $f_2$: $\alpha = 2$ (left) and $\alpha = 3$ (right).

where $\mathbf{1}_A$ denotes the indicator function of an event $A$. The derivative $f_2^{(3,3)}$ has a discontinuity along the curve $xy = 1/2$, but is in $L_q([0,1)^2)$ for any $q$, which ensures $f_2 \in W_{2,3,q,r}$. Note that we have

$$I(f_2) = \frac{1}{896}\left(\frac{363}{140} + \log 2\right).$$

Figure 2 shows the absolute integration error by Algorithm 3.3 (and Remark 3.7) with $\alpha = 2$ and $\alpha = 3$. Similarly to the result for $f_1$, the integration errors of $I_m^{(1)}$ and $I_m^{(2)}$ achieve the convergence of nearly order $N^{-1}$ and $N^{-2}$, respectively. For the case $\alpha = 3$, after the recursive application of Richardson extrapolation, the error decays asymptotically with the order $N^{-3}$. However, the magnitude of the error itself is almost comparable to that for $I_m^{(2)}$ in this range of $N$. In fact, as can be seen from the right plot of Fig. 3, QMC rules using order 3 Sobol' sequences achieve the convergence of order $N^{-3}$ only asymptotically, and the performances of order 2 and 3 Sobol' sequences are comparable. This implies that, on the right-hand side of (3.1), $c_1(f)/p^n$ is the most dominant term, but $c_2(f)/p^{2n}$ is not the only secondary dominant term and is comparable to

$$\sum_{\boldsymbol{l}\in\mathbb{N}_0^s}\sum_{\substack{\boldsymbol{k}\in P_{m,n}^{\perp}\\ \boldsymbol{0}\neq\boldsymbol{k}<p^n}} \hat{f}(\boldsymbol{k}+p^n\boldsymbol{l}).$$

This is why our Algorithm 3.3 cannot achieve the desired rate of convergence for $\alpha = 3$ when $m$ is not large enough. Thus, improving the performance of original higher-order digital nets and sequences is important for our extrapolation-based rules to work well.

When IEEE 754 double-precision floating-point format is employed, the first $2^m$ points of order 2 Sobol' sequences can be represented with full precision for the considered range of $m \in [3, 22]$, whereas those of order 3 Sobol' sequences cannot for $m \geqslant 18$. As we see from Fig. 3, the integration error for $f_1$ using order 3 Sobol' sequences remains almost the same for $m \geqslant 18$, which is considered to be the

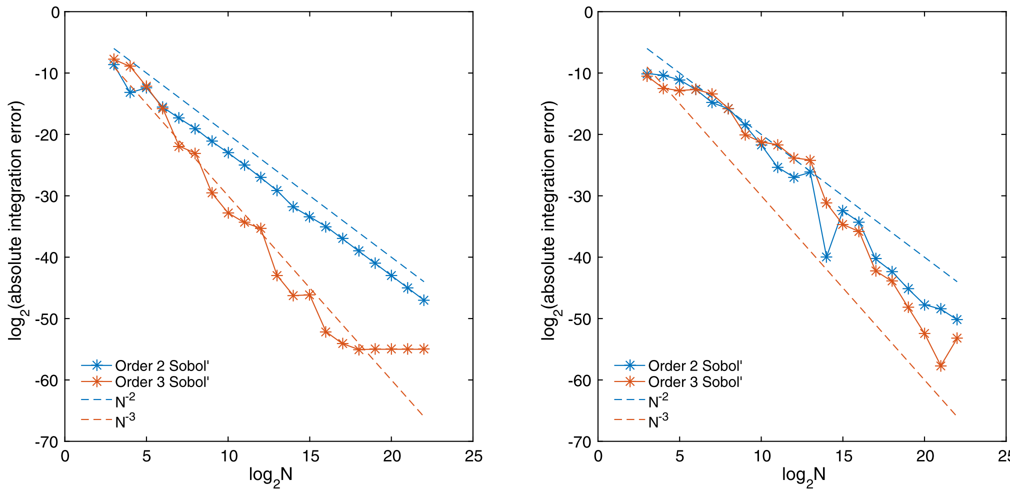


FIG. 3. Integration error by QMC rules using order 2 and 3 Sobol' sequences for $f_1$ (left) and $f_2$ (right). If $\alpha m > 52$ the truncation map $\mathrm{tr}_{52}$ is applied to all of the quadrature nodes.

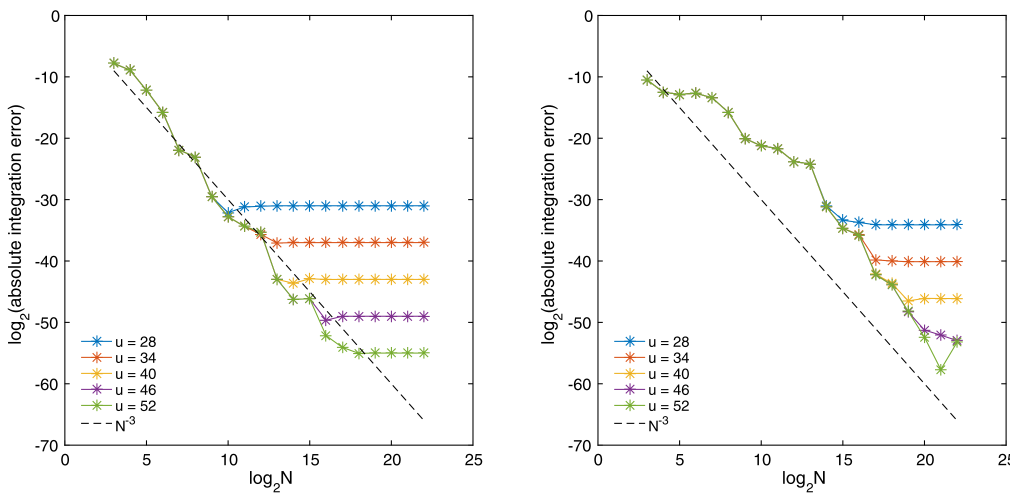


FIG. 4. Effect of precision $u$ for QMC rules using order 3 Sobol' sequences for $f_1$ (left) and $f_2$ (right). If $\alpha m > u$ the truncation map $\mathrm{tr}_u$ is applied to all of the quadrature nodes.

consequence of rounding the quadrature nodes. In fact, by changing the maximum precision from 52 to lower values $u$, the switch of the convergence behaviour from the $O(N^{-3})$ decay to the plateau happens for smaller $m$ as shown in the left plot of Fig. 4. Since our extrapolation-based quadrature rules do not suffer from the rounding problem in this range of $m$, the error continues to decay even for $m \geqslant 18$ as is clear from the right plot of Fig. 1. However, such a switch of the convergence behaviour for order 3 Sobol' sequences cannot be clearly observed for $f_2$, see the right plot of Fig. 3. Changing the maximum precision from 52 to lower values $u$ does yield such a switch, but for larger $m$ as compared to the case for $f_1$. Hence, whether or not the round-off error is comparable to the integration error when $\alpha m$ goes beyond

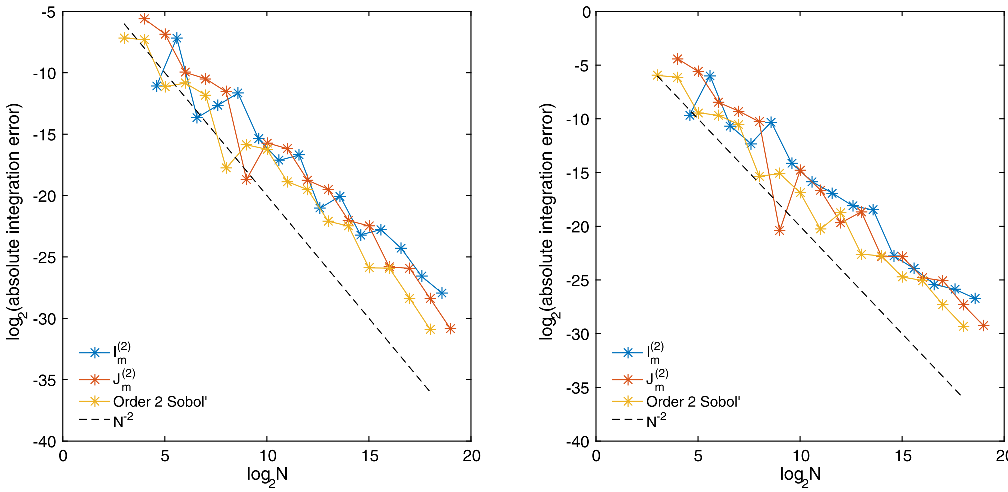


FIG. 5. Comparison of integration errors by three algorithms for $f_3$ with $c_1 = 1.3$ (left) and $f_4$ with $c_2 = 1$ (bottom).

an available precision, depends on the considered integrand, and in general it seems quite difficult to distinguish the round-off error from the integration error. Again, we would like to emphasize that our extrapolation-based quadrature rules are free from such difficulty unless $m$ is large enough, say $m = 52$, which is an important advantage compared to the original higher-order digital nets and sequences.

### 4.2 *High-dimensional cases*

Let us move on to the high-dimensional setting. Following Dick *et al.* (2019) and Gantner & Schwab (2016), we consider the following two test functions:

$$f_3(\boldsymbol{x}) = \prod_{j=1}^{s}\left[1 + \gamma_j\left(x_j^{c_1} - \frac{1}{1+c_1}\right)\right],$$

$$f_4(\boldsymbol{x}) = \exp\left(c_2\sum_{j=1}^{s}\gamma_j x_j\right),$$

with parameters $c_1 > 0, c_2 \neq 0$, for which we have $I(f_3) = 1$ and

$$I(f_4) = \prod_{j=1}^{s}\left[\frac{1}{c_2\gamma_j}\left(\exp(c_2\gamma_j) - 1\right)\right].$$

When $1 < c_1 < 2$ the second derivative of the function $x \mapsto x^{c_1}$ is not absolutely continuous, but is in $L_q([0,1))$ for any $q < 1/(2-c_1)$, which means that $f_3 \in W_{s,2,q,r}$ for $q < 1/(2-c_1)$. On the other hand, $f_4$ is analytic and belongs to $W_{s,\alpha,q,r}$ for any $\alpha \geqslant 2$ and $q \geqslant 1$. As stated by Gantner & Schwab (2016), $f_4$ is designed to mimic the behaviour of parametric solution families of partial differential equations. We employ this test function to see potential applicability to such problems.

We put $s = 100$ and $\gamma_j = j^{-2}$. We consider three quadrature rules: Algorithm 3.3 with $\alpha = 2$, denoted by $I_m^{(2)}$; Algorithm 3.8 with $\alpha = 2$, denoted by $J_m^{(2)}$; and QMC rules using order 2 Sobol' sequences. Figure 5 shows the comparison of the absolute integration errors obtained by these three algorithms. In fact, there is no decisive difference in performance between these algorithms, and all of them achieve the nearly desired rate of convergence, which is $O(N^{-2+\varepsilon})$ for arbitrarily small $\varepsilon > 0$. This result not only supports our theoretical result, but also indicates that Richardson extrapolation allows truncation of higher-order digital nets and sequences without sacrificing the practical performance of them even for high-dimensional cases.

**Acknowledgements**

The author would like to thank Tomohiko Hironaka and Takehito Yoshiki for useful discussions and comments. The comments and suggestions made by the anonymous referees, improving the exposition of this paper, are greatly appreciated.

**Funding**

Japan Society for the Promotion of Science Grant-in-Aid for Young Scientists (15K20964).

## References

Baldeaux, J., Dick, J., Leobacher, G., Nuyens, D. & Pillichshammer, F. (2012) Efficient calculation of the worst-case error and (fast) component-by-component construction of higher order polynomial lattice rules. *Numer. Algorithms*, **59**, 403–431.

Baldeaux, J., Dick, J. & Pillichshammer, F. (2011) Duality theory and propagation rules for higher order nets. *Discrete Math.*, **311**, 362–386.

Dick, J. (2007) Explicit constructions of quasi-Monte Carlo rules for the numerical integration of high-dimensional periodic functions. *SIAM J. Numer. Anal.*, **45**, 2141–2176.

Dick, J. (2008) Walsh spaces containing smooth functions and quasi-Monte Carlo rules of arbitrary high order. *SIAM J. Numer. Anal.*, **46**, 1519–1553.

Dick, J. (2009) The decay of the Walsh coefficients of smooth functions. *Bull. Aust. Math. Soc.*, **80**, 430–453.

Dick, J., Goda, T. & Yoshiki, T. (2019) Richardson extrapolation of polynomial lattice rules. *SIAM J. Numer. Anal.*, **57**, 44–69.

Dick, J., Kuo, F. Y., Le Gia, Q. T., Nuyens, D. & Schwab, C. (2014) Higher order QMC Petrov–Galerkin discretization for affine parametric operator equations with random field inputs. *SIAM J. Numer. Anal.*, **52**, 2676–2702.

Dick, J., Kuo, F. Y., Le Gia, Q. T. & Schwab, C. (2015) Fast QMC matrix-vector multiplication. *SIAM J. Sci. Comput.*, **37**, A1436–A1450.

Dick, J. & Pillichshammer, F. (2010) *Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration*. Cambridge: Cambridge University Press.

Faure, H. (1982) Discrépances de suites associées à un système de numération (en dimension $s$). *Acta Arith.*, **41**, 337–351.

Gantner, R. N. & Schwab, C. (2016) Computational higher order quasi-Monte Carlo integration. *Monte Carlo and Quasi-Monte Carlo Methods*. Springer Proceedings in Mathematics & Statistics. Switzerland: Springer, pp. 271–288.

Goda, T. (2015) Good interlaced polynomial lattice rules for numerical integration in weighted Walsh spaces. *J. Comput. Appl. Math.*, **285**, 279–294.

Goda, T. (2016) Quasi-Monte Carlo integration using digital nets with antithetics. *J. Comput. Appl. Math.*, **304**, 26–42.

Goda, T. & Dick, J. (2015) Construction of interlaced scrambled polynomial lattice rules of arbitrary high order. *Found. Comput. Math.*, **15**, 1245–1278.

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

Goda, T., Suzuki, K. & Yoshiki, T. (2017) Optimal order quasi-Monte Carlo integration in weighted Sobolev spaces of arbitrary smoothness. *IMA J. Numer. Anal.*, **37**, 505–518.

Goda, T., Suzuki, K. & Yoshiki, T. (2018) Optimal order quadrature error bounds for infinite-dimensional higher-order digital sequences. *Found. Comput. Math.*, **18**, 433–458.

Hinrichs, A., Markhasin, L., Oettershagen, J. & Ullrich, T. (2016) Optimal quasi-Monte Carlo rules on order 2 digital nets for the numerical integration of multivariate periodic functions. *Numer. Math.*, **134**, 163–196.

Kuo, F. Y. & Nuyens, D. (2016) Application of quasi-Monte Carlo methods to elliptic PDEs with random diffusion coefficients: a survey of analysis and implementation. *Found. Comput. Math.*, **16**, 1631–1696.

Niederreiter, H. (1986) Low-discrepancy point sets. *Monatsh. Math.*, **102**, 155–167.

Niederreiter, H. (1988) Low-discrepancy and low-dispersion sequences. *J. Number Theory*, **30**, 51–70.

Niederreiter, H. (1992) *Random Number Generation and Quasi-Monte Carlo Methods*. CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 63. Philadelphia: Society for Industrial and Applied Mathematics (SIAM).

Niederreiter, H. & Xing, C. P. (2001) *Rational Points on Curves over Finite Fields: Theory and Applications*. London Mathematical Society Lecture Note Series, vol. 285. Cambridge: Cambridge University Press.

Rosenbloom, M. Yu. & Tsfasman, M. A. (1997) Codes for the *m*-metric. *Probl. Inf. Transm.*, **33**, 55–63.

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

Sobol', I. M. (1967) The distribution of points in a cube and approximate evaluation of integrals. *Zh. Vychisl. Mat. Mat. Fiz.*, **7**, 784–802.

Tezuka, S. (1993) Polynomial arithmetic analogue of Halton sequences. *ACM Trans. Model. Comput. Simul.*, **3**, 99–107.