WALSH SPACES CONTAINING SMOOTH FUNCTIONS AND QUASI-MONTE CARLO RULES OF ARBITRARY HIGH ORDER

JOSEF DICK

Abstract. We define a Walsh space which contains all functions whose partial mixed derivatives up to order $\delta \geq 1$ exist and have finite variation. In particular, for a suitable choice of parameters, this implies that certain reproducing kernel Sobolev spaces are contained in these Walsh spaces. For this Walsh space we then show that quasi-Monte Carlo rules based on digital $(t, \alpha, s)$-sequences achieve the optimal rate of convergence of the worst-case error for numerical integration. This rate of convergence is also optimal for the subspace of smooth functions. Explicit constructions of digital $(t, \alpha, s)$-sequences are given hence providing explicit quasi-Monte Carlo rules which achieve the optimal rate of convergence of the integration error for arbitrarily smooth functions.

1. Introduction

Quasi-Monte Carlo rules are quadrature rules which aim to approximate an integral $\int_{[0,1]^s} f(x) \, dx$ by the average of the $N$ function values $f(x_n)$ at the quadrature points $x_0, \ldots, x_{N-1} \in [0,1]^s$. The dimension $s$ can be arbitrarily large. The task here is to find ways of how to choose those quadrature points in order to obtain a fast convergence of the approximation to the integral. Such quadrature rules have a long history with results going back to Koksma [20] (1942/43), Korobov [21] (1959), Hlawka [17, 18] (1961/1962), Sobol [48] (1967), Faure [13] (1982), Niederreiter [29] (1987) and so on. Although extensive research has been carried out in the past decades (see for example the monographs or review articles [19, 22, 23, 26, 27, 30, 32, 46, 50, 51] or the proceedings of the biennial conference series on Monte Carlo and quasi-Monte Carlo methods [12, 31, 33, 35, 36, 37] and the references therein) explicit constructions of quadrature points are only available for two cases: in two dimensions arbitrary high convergence can be achieved by Fibonacci lattice rules for smooth periodic functions and for arbitrary high dimensions a convergence of $O(N^{-1}(\log N)^{s-1})$ can be achieved for functions of bounded variation. But for dimensions $s > 2$ no explicit constructions have been established which can fully exploit the smoothness of the integrand, though for example for periodic functions the existence of such rules has been known since the work of Korobov [21] (1959) and Hlawka [18] (1962). In this paper we give a complete solution to this problem.

Among other things we show that an explicit construction of suitable point sets and sequences can be obtained in the following way: let $d \geq 1$ be an integer and

2000 Mathematics Subject Classification. 11K38, 11K45, 65C05, 42C10.

Key words and phrases. Numerical integration, quasi-Monte Carlo, digital nets and sequences, Walsh functions.
let \( x_0, x_1, \ldots \in [0,1)^{ds} \) be the points of a digital \((t,m,ds)\)-net or digital \((t,ds)\)-sequence over a finite field \( F_q \) in dimension \( ds \) (see [30] for the definition of digital nets and sequences and see for example [13, 30, 32, 38, 48] for explicit constructions of suitable digital nets and sequences). Let \( x_n = (x_n,1, \ldots, x_n,ds) \) with \( x_{n,j} = x_{n,j,1} q^{-1} + x_{n,j,2} q^{-2} + \cdots \) and \( x_{n,j,i} \in \{0, \ldots, q-1\} \). Then for \( n \geq 0 \) we define \( y_n = (y_n,1, \ldots, y_n,s) \) with

\[
y_{n,j} = \sum_{i=1}^{\infty} \sum_{k=1}^{d} x_{n,(j-1)d+k,j} q^{-k-(i-1)d} \quad \text{for} \quad j = 1, \ldots, s.
\]

(Note that the sum over \( i \) above is often finite as \( x_{n,j,i} = 0 \) for \( i \) large enough.)

Corollary 3 (Section 5) and Corollary 4 (Section 6) now show that quasi-Monte Carlo rules using the points \( y_0, \ldots, y_{N-1} \) (with \( N = q^m \) for some \( m \geq 1 \)) achieve the optimal rate of convergence of the integration error of \( O(N^{-\vartheta} (\log N)^{d \vartheta}) \) for functions which have partial mixed derivatives up to order \( \vartheta \) which are square integrable as long as \( 1 \leq \vartheta \leq d \) (Corollary 3 is actually more general). If \( \vartheta > d \) no improvement of the convergence rate is obtained compared to functions with smoothness \( \vartheta = d \), i.e. we obtain a convergence of \( O(N^{-d} (\log N)^{d \vartheta}) \). Similar, but less general results for periodic functions compared to those in this paper have been shown in [5] by a different proof method. (The construction above is an example of a construction method which can be used. In Section 4 we outline the general algebraical properties required for the construction of suitable point sets.)

The quasi-Monte Carlo algorithm based on digital nets and sequences proposed here has also some further useful properties. For example our results also hold if one randomizes the point set by, say, a random digital shift (see for example [7, 8, 26]). (This follows easily because the worst-case error (see Section 5) is invariant with respect to digital shifts and hence we obtain the same upper bounds for randomized digital nets and sequences.) In summary the quadrature rules have the following properties (compare with [16]):

- The quadrature rules introduced in this paper are equal weight quadrature rules which achieve the optimal rate of convergence up to some \( \log N \) factors and the result holds for deterministic and randomly digitally shifted quadrature rules.
- The construction of the underlying point set is explicit and suitable point sets are available in arbitrary high dimensions and arbitrary high number of points.
- The quadrature rules automatically adjust themselves to the optimal rate of convergence \( O(N^{-\vartheta} (\log N)^{d \vartheta}) \) as long as \( 1 \leq \vartheta \leq d \).
- The underlying point set is extensible in the dimension as well as in the number of points, i.e., one can always add some coordinates or points to an existing point set such that the quality of the point set is preserved.

In the following we lay out some of the underlying principles used in this work which stem from the behaviour of the Walsh coefficients of smooth functions. Walsh functions are piecewise constant wavelets which form an orthonormal set of \( L_2([0,1]^s) \). In their simplest form, for a non-negative integer \( k \) with base 2 representation \( k = \kappa_0 + \cdots + \kappa_{m-1} 2^{m-1} \) and an \( x \in [0,1) \) with base 2 representation \( x = x_1 2^{-1} + x_2 2^{-2} + \cdots \), the \( k \)-th Walsh function in base 2 is given by

\[
\text{wal}_k(x) = (-1)^{\kappa_0 x_1 + \cdots + \kappa_{m-1} x_m}.
\]
(Later on we will use the more general definition of Walsh functions over groups.) There are some essential differences to Fourier analysis and trigonometrical polynomials. As is well known, the smoothness of a periodic function is directly related to the decay of its Fourier coefficients (see for example [53]). This is not the case for the coefficients in a Walsh series expansion. Indeed it seems almost to be the other way round. From [14, Theorem VII] for example it is known that the only absolute continuous functions whose $k$-th Walsh coefficients decay faster than $1/k$ are constant functions. Hence assuming that the Walsh coefficients decay fast forces the function to be non-smooth, i.e., to have jumps. This is also intuitively correct as observed by Fine [14]. The only values Walsh functions in base 2 take on are $+1$ and $-1$. Hence if the absolute value of the first Walsh coefficient is larger than the sum of the absolute values of the remaining Walsh coefficients then it follows that the function must have a jump at $1/2$ as the first Walsh function has a jump at $1/2$ and the remaining Walsh functions cannot smooth out this jump in time.

On the other hand the Walsh coefficients do satisfy some conditions depending on the smoothness of the function. For example if the function satisfies a Hölder condition with coefficient $0 < \lambda \leq 1$, then the $k$-th Walsh coefficient decays with order $k^{-\lambda}$. This is in analogy with trigonometric series. Differentiability assumptions on the other hand do not force the Walsh coefficients to decay faster, as we have already seen from the example of Fine [14].

Let us consider the implications of this behaviour of the Walsh coefficients for smooth functions on the problem of numerical integration. Previous results on numerical integration involving Walsh (or also Haar functions), see for example [7, 15], assumed that the coefficients decay fast, thereby allowing only non-smooth functions (except for constant functions of course) to be in the space. Digital constructions (digital nets, digital sequences) of quasi-Monte Carlo rules for numerical integration from [13, 28, 30, 38, 48] have been shown to work well for function spaces with fast decaying Walsh coefficients, but as this fast decay assumption implies that the integrands are non-smooth, it has been thought that digital nets and sequences can only achieve higher order convergence for non-smooth functions. As can be seen from [5] and this work it seems that this observation is correct for those digital nets and sequences considered until now (see [30, 32]). The digital nets and sequences for example from [13, 28, 30, 38, 48] basically work well only for non-smooth functions or functions with bounded variation (in which case they achieve a convergence of order $N^{-1}(\log N)^{s-1}$ as follows from the Koksma-Hlawka inequality, see [30]). Slight improvements for smoother functions can be achieved using randomizations (see for example [3, 39]), but no systematic exploitation of stronger smoothness assumptions has been achieved for digital nets or sequences. Arbitrary high convergence can be achieved for smooth periodic functions using lattice rules (see for example [46]), but here the problem is that good lattice rules are not explicitly known and one relies on computer search (see [16] where pros and cons of various quasi-Monte Carlo rules have been discussed). This way, the problem of finding explicit constructions for the numerical integration of smooth periodic functions introduced by Korobov in 1959, remained unsolved for dimensions $s > 2$ - explicit constructions for lattice rules have until now proven to be too difficult to find and the known explicit constructions for digital nets and sequences do not achieve higher order convergence.
A first breakthrough on this problem came in [5]. Therein the author showed that a modification of the digital construction for nets and sequences (see [30, Chapter 4]) yields explicit constructions of quasi-Monte Carlo rules which can achieve an arbitrary high convergence order by exploiting the smoothness of the function, as long as the function is periodic. The proof relies on a very involved upper bound on the Walsh coefficients of the reproducing kernel of the Korobov space (which is a reproducing kernel Hilbert space based on Fourier series).

In this work we uncover the underlying principle used implicitly in [5]. Namely we explicitly state and use the behaviour of the Walsh coefficients of smooth functions. The Walsh coefficients depend on the base $q$ digit expansion of the wavenumber $k$ and also of the point $x$ where the Walsh function is to be evaluated. Hence, maybe not surprisingly, the value of the $k$-th Walsh coefficients of smooth functions also depend on the $q$-adic expansion of $k$. Here we show that the Walsh space $E_{s, q, \vartheta, \gamma}$ introduced in Section 3 contains all functions with partial mixed derivatives up to order $\delta < \vartheta$ exist and have finite variation, where $\vartheta$ is a parameter restricting the behaviour of the Walsh coefficients of the function space $E_{s, q, \vartheta, \gamma}$. (We use a similar, though much more general, technique as Fine [14] used for showing that the Walsh coefficients of a differentiable function cannot decay faster than $1/k$.)

The concept of digital $(t, \alpha, \beta, m, s)$-nets and digital $(t, \alpha, s)$-sequences (see Section 4 and also [5] for a similar concept) is now designed to yield point sets which work well for the Walsh space $E_{s, q, \vartheta, \gamma}$, just in the same way as the digital nets and sequences from [13, 28, 30, 38, 48] are designed to work well for the spaces for example considered in [7, 15] (or as lattice rules are designed to work well for periodic Korobov spaces). Here the power of the result that the Walsh space $E_{s, q, \vartheta, \gamma}$ contains smooth functions comes into play: it follows that we can fully exploit the smoothness of an integrand using digital $(t, \alpha, m, s)$-nets or digital $(t, \alpha, s)$-sequences. As the construction of the points $y_0, y_1, \ldots$ introduced at the beginning yields explicit examples of digital $(t, \alpha, m, s)$-nets or digital $(t, \alpha, s)$-sequences as shown in Section 4 we therefore obtain explicit constructions of quasi-Monte Carlo rules which can achieve the optimal order of convergence for arbitrary smooth functions.

2. Walsh functions over groups

In this section we give the definition of Walsh functions over groups and present some essential properties. Walsh functions in base 2 were first introduced by Walsh [52], though a similar but non-complete set of functions has already been studied by Rademacher [43]. Further important results were obtained in [14]. We follow [42] in our presentation.

2.1. Definition of Walsh functions over groups. An essential tool for the investigation of digital nets are Walsh functions. A very general definition, corresponding to the most general construction of digital nets over finite rings, was given in [24]. There, Walsh functions over a finite abelian group $G$, using some bijection $\varphi$, were defined. Here we restrict ourselves to the additive groups of the finite fields $\mathbb{F}_p^r$, $p$ prime and $r \geq 1$. We restate the definitions for this special case here for the sake of convenience. In the following let $\mathbb{N}$ denote the set of positive integers and $\mathbb{N}_0$ the set of non-negative integers.
Definition 1 (Walsh functions). Let $q = p^r$, $p$ prime, $r \in \mathbb{N}$ and let $\mathbb{F}_q$ be the finite field with $q$ elements. Let $\mathbb{Z}_q = \{0, 1, \ldots, q - 1\} \subset \mathbb{Z}$ with ring operations modulo $q$ and let $\varphi : \mathbb{Z}_q \rightarrow \mathbb{F}_q$ be a bijection such that $\varphi(0) = 0$, the neutral element of addition in $\mathbb{F}_q$. Moreover denote by $\psi$ the isomorphism of additive groups $\psi : \mathbb{F}_q \rightarrow \mathbb{Z}_p$ and define $\eta := \psi \circ \varphi$. For $1 \leq i \leq r$ denote by $\pi_i$ the projection $\pi_i : \mathbb{Z}_p^r \rightarrow \mathbb{Z}_p$, $\pi_i(x_1, \ldots, x_r) = x_i$.

Let now $k \in \mathbb{N}_0$ with base $q$ representation $k = \kappa_0 + \kappa_1 q + \cdots + \kappa_{m-1} q^{m-1}$ where $\kappa_i \in \mathbb{Z}_q$ and let $x \in [0, 1)$ with base $q$ representation $x = x_1/q + x_2/q^2 + \cdots$ (unique in the sense that infinitely many $x_i$ must be different from $q - 1$). Let $\omega_p = e^{2\pi i/p}$. Then the $k$-th Walsh function over the finite field $\mathbb{F}_q$ with respect to the bijection $\varphi$ is defined by

$$w_{\varphi,k}(x) = \sum_{\pi \in \mathbb{Z}_p^r}^\omega \sum_{i=1}^r \pi_i(x_1) \pi_i(\eta(x_2)) \pi_i(\eta(x_3)) \cdots \pi_i(\eta(x_{r+1})).$$

For convenience we will in the rest of the paper omit the subscript and simply write $w_k$ if there is no ambiguity.

Multivariate Walsh functions are defined by multiplication of the univariate components, i.e., for $s > 1$, $x = (x_1, \ldots, x_s) \in [0, 1)^s$ and $k = (k_1, \ldots, k_s) \in \mathbb{N}_0^s$, we set

$$w_{\varphi,k}(x) = \prod_{j=1}^s w_{\varphi,k_j}(x_j).$$

We summarize some important properties of Walsh functions over finite fields which will be used throughout the paper. The proofs of the subsequent results can be found e.g. in [25, 41]. In the following we call $x \in [0, 1)$ a $q$-adic rational if $x$ can be represented by a finite base $q$ expansion.

Proposition 1. Let $p$, $q$, $\mathbb{F}_q$ and $\varphi$ be as in Definition 1. For $x, y$ with $q$-adic representations $x = \sum_{i=0}^\infty x_i q^{-i}$ and $y = \sum_{i=0}^\infty y_i q^{-i}$, $w \in \mathbb{Z}$ (hence the following operations are also defined for integers), define $x \oplus_{\varphi} y := \sum_{i=0}^\infty z_i q^{-i}$ where $z_i := \varphi^{-1}(\varphi(x_i) + \varphi(y_i))$ and $\otimes_{\varphi} x := \sum_{i=0}^\infty v_i q^{-i}$ where $v_i := \varphi^{-1}(\varphi(x_i))$. Further we set $x \otimes_{\varphi} y := x \oplus_{\varphi}(\otimes_{\varphi} y)$. For vectors $x, y$ we define the operations component-wise. Then we have:

1. For all $k, l \in \mathbb{N}_0$ and all $x, y \in [0, 1)$, with the restriction that if $x, y$ are not $q$-adic rationals then $x \oplus_{\varphi} y$ is not allowed to be a $q$-adic rational, we have

$$w_{\varphi,k}(x) \cdot w_{\varphi,l}(x) = w_{\varphi,k \oplus_{\varphi} l}(x), \quad w_{\varphi,k}(x) \cdot w_{\varphi,l}(y) = w_{\varphi,k \oplus_{\varphi} l}(x \oplus_{\varphi} y)$$

and

$$w_{\varphi,k}(x) \cdot w_{\varphi,l}(y) = w_{\varphi,k \oplus_{\varphi} l}(x), \quad w_{\varphi,k}(x) \cdot w_{\varphi,l}(y) = w_{\varphi,k \oplus_{\varphi} l}(x \otimes_{\varphi} y).$$

2. We have

$$\sum_{k=0}^{q-1} w_{\varphi,k}(k/q) = \begin{cases} 0 & \text{if } l \neq 0, \\ q & \text{if } l = 0. \end{cases}$$
(3) We have
\[ \int_0^1 \text{wal}_0(x) \, dx = 1 \quad \text{and} \quad \int_0^1 \text{wal}_k(x) \, dx = 0 \text{ if } k > 0. \]

(4) For all \( k, l \in \mathbb{N}_0 \) we have the following orthogonality properties:
\[ \int_{[0,1]^*} \text{wal}_k(x) \text{wal}_l(x) \, dx = \begin{cases} 1 & \text{if } k = l, \\ 0 & \text{otherwise}. \end{cases} \]

(5) For any \( f \in L_2([0,1]^s) \) and any \( \sigma \in [0,1]^s \) we have
\[ \int_{[0,1]^*} f(x) \, dx = \int_{[0,1]^*} f(x \oplus_\varphi \sigma) \, dx. \]

(6) For any integer \( s \geq 1 \) the system \( \{ \text{wal}_k : k \in \mathbb{N}_0 \} \) is a complete orthonormal system in \( L_2([0,1]^s) \).

Throughout the paper we will use the same bijection \( \varphi \) and the same finite field \( \mathbb{F}_q \) is used for Walsh functions and \( \oplus_\varphi \) and \( \oplus_\varphi \). Hence we will often write \( \oplus \) and \( \oplus \) instead of \( \oplus_\varphi \) and \( \oplus_\varphi \).

3. Walsh spaces containing smooth functions

In the following we investigate how the Walsh coefficients of smooth functions decay and subsequently we use this to define function classes based on Walsh functions which contain smooth functions.

3.1. The decay of the Walsh coefficients of smooth functions. A function \( f : [0,1) \to \mathbb{R} \) satisfies a Hölder condition with coefficient \( 0 < \lambda \leq 1 \) if there is a constant \( C_f > 0 \) such that
\[ |f(x) - f(y)| \leq C_f |x - y|^{\lambda} \quad \text{for all } x, y \in [0,1). \]

The right hand side of the above inequality forms a metric on \([0,1)\). When one considers the higher dimensional domain \([0,1]^s\) then \( |x - y| \) is changed to some other metric on \([0,1]^s\). Here we consider tensor product spaces and we generalize the Hölder condition to higher dimensions which is suitable for tensor product spaces in our context. Consider for example the function \( f(x) = \prod_{j=1}^s f_j(x_j) \), where \( x = (x_1, \ldots, x_s) \) and each \( f_j : [0,1) \to \mathbb{R} \) satisfies a Hölder condition with coefficient \( 0 < \lambda \leq 1 \). Then it follows that for all \( \emptyset \neq u \subseteq \{1, \ldots, s\} \) we have
\[ \prod_{j \in u} |f_j(x_j) - f_j(y_j)| \leq \prod_{j \in u} C_{f_j} \prod_{j \in u} |x_j - y_j|^\lambda \]
for all \( x_j, y_j \in [0,1) \) with \( j \in u \). But here \( \prod_{j=1}^s |x_j - y_j| \) is not a metric on \([0,1]^s\).

Note that we have
\[ \prod_{j \in u} |f_j(x_j) - f_j(y_j)| = \left| \sum_{v \subseteq u} (-1)^{|v| - |u|} \prod_{j \in v} f_j(x_j) \prod_{j \in u \setminus v} f_j(y_j) \right|, \]
which can be described in words in the following way: for given \( \emptyset \neq u \subseteq S \) let \( x_j, y_j \in [0,1) \) with \( x_j \neq y_j \) for all \( j \in u \); consider the box \( J \) with vertices \( \{a_j\}_{j \in u} : a_j = x_j \text{ or } a_j = y_j \text{ for } j \in u \}. \) Then (2) is the alternating sum of the function \( \prod_{j \in u} f_j \) at the vertices of \( J \) where adjacent vertices have opposite signs. But this sum can also be defined for functions on \([0,1]^s\) which are not of product form.
Indeed, let for a subinterval \( J = \prod_{j=1}^{s} [x_j, y_j] \) with \( 0 \leq x_j < y_j \leq 1 \) and a function \( f : [0, 1]^s \to \mathbb{R} \) the function \( \Delta(f, J) \) denote the alternating sum of \( f \) at the vertices of \( J \) where adjacent vertices have opposite signs. (Hence for \( f = \prod_{j=1}^{s} f_j \) we have \( \Delta(f, J) = \prod_{j=1}^{s} (f_j(x_j) - f_j(y_j)) \).

Let \( p \geq 1 \). Then we define the generalized variation in the sense of Vitali with coefficient \( 0 < \lambda \leq 1 \) by

\[
V_{\lambda,p}^{(s)}(f) = \sup_{\mathcal{P}} \left( \sum_{J \in \mathcal{P}} \text{Vol}(J) \left| \frac{\Delta(f, J)}{\text{Vol}(J)^{\lambda}} \right|^p \right)^{1/p},
\]

where the supremum is extended over all partitions \( \mathcal{P} \) of \([0, 1]^s\) into subintervals and \( \text{Vol}(J) \) denotes the volume of the subinterval \( J \). Note that for \( \lambda = 1 \) and \( p = 1 \) one obtains the usual definition of the variation in the sense of Vitali, see for example [30]. If we take \( p = \infty \), then we obtain a condition of the form (1) where \( u = S \) and where we can take the constant \( \prod_{j=1}^{s} C_{f_j} = V_{\lambda, \infty}^{(s)}(f) \). For \( s = 1 \) and \( p = \infty \) we obtain a Hölder condition with coefficient \( 0 < \lambda \leq 1 \).

For \( \lambda = 1 \) and if the partial derivatives of \( f \) are continuous on \([0, 1]^s\) we also have the formula

\[
V_{1,p}^{(s)}(f) = \left( \int_{[0,1]^s} \left| \frac{\partial^s f}{\partial x_1 \cdots \partial x_s} \right|^p \, dx \right)^{1/p},
\]

for all \( p \geq 1 \). Indeed we have

\[
|\Delta(f, J)| = \left| \int_{J} \frac{\partial^s f}{\partial x_1 \cdots \partial x_s}(x) \, dx \right| = \text{Vol}(J) \left| \frac{\partial^s f}{\partial x_1 \cdots \partial x_s}(\zeta_J) \right|
\]

for some \( \zeta_J \in \partial J \), which follows by applying the mean value theorem to the inequality

\[
\min_{x \in \partial J} \left| \frac{\partial^s f}{\partial x_1 \cdots \partial x_s}(x) \right| \leq \text{Vol}(J)^{-1} \left| \int_{J} \frac{\partial^s f}{\partial x_1 \cdots \partial x_s}(x) \, dx \right| \leq \max_{x \in \partial J} \left| \frac{\partial^s f}{\partial x_1 \cdots \partial x_s}(x) \right|.
\]

Therefore we have

\[
\sum_{J \in \mathcal{P}} \text{Vol}(J) \left| \frac{\Delta(f, J)}{\text{Vol}(J)^{\lambda}} \right|^p = \sum_{J \in \mathcal{P}} \text{Vol}(J) \left| \frac{\partial^s f}{\partial x_1 \cdots \partial x_s}(\zeta_J) \right|^p,
\]

which is just a Riemann sum for the integral \( \int_{[0,1]^s} \left| \frac{\partial^s f}{\partial x_1 \cdots \partial x_s} \right|^p \, dx \) and thus the equality follows.

Using Hölder’s inequality it follows that

\[
\sum_{J \in \mathcal{P}} \text{Vol}(J) \left| \frac{\Delta(f, J)}{\text{Vol}(J)^{\lambda}} \right|^p \leq \left( \sum_{J \in \mathcal{P}} \text{Vol}(J) \right)^{1/p} \left( \sum_{J \in \mathcal{P}} \left( \text{Vol}(J)^{1-1/p} \right)^{p/(p-1)} \right)^{1-1/p} \left( \sum_{J \in \mathcal{P}} \left( \text{Vol}(J)^{1-1/p} \right)^{p/(p-1)} \right)^{1-1/p}
\]

and as

\[
\left( \sum_{J \in \mathcal{P}} \left( \text{Vol}(J)^{1-1/p} \right)^{p/(p-1)} \right)^{1-1/p} = \left( \sum_{J \in \mathcal{P}} \text{Vol}(J) \right)^{1-1/p} = 1
\]

it follows that

\[
V_{\lambda,1}^{(s)}(f) \leq V_{\lambda,p}^{(s)}(f)
\]

for all \( p \geq 1 \).
Until now we did not take projections to lower dimensional faces into account (in (1) we did take projections into account as we considered all \( \emptyset \neq u \subseteq S \)).

For \( \emptyset \neq u \subseteq S \), let \( V_{\lambda, p}^{(u)}(f_u; u) \) be the generalized variation in the sense of Vitali with coefficient \( 0 < \lambda \leq 1 \) of the \( |u| \)-dimensional function \( f_u(x_u) = \int_{[0,1)^{|u|}} f(x) \, dx_{S \setminus u} \). For \( u = \emptyset \) we have \( f_\emptyset = \int_{[0,1]} f(x) \, dx_{S} \) and we define \( V_{\lambda, p}^{(u)}(f_\emptyset; \emptyset) = |f_\emptyset| \). Let \( q \geq 1 \), then

\[
(4) \quad V_{\lambda, p, q}(f) = \left( \sum_{u \subseteq S} \left( V_{\lambda, p}^{(u)}(f_u; u) \right)^q \right)^{1/q}
\]

is called the generalized variation of \( f \) on \([0,1]^s\) in the sense of Hardy and Krause. For \( \lambda = p = q = 1 \) one obtains an unanchored version of the usual definition of the variation in the sense of Hardy and Krause, see [30]. A function \( f \) for which \( V_{\lambda, p, q}(f) < \infty \) is said to be of finite variation with coefficient \( \lambda \). (We remark that in some cases it might be appropriate to leave out the term corresponding to \( u = \emptyset \) in (4), but here this term will be needed later on and hence we include it already in the definition of the variation.)

As first suggested in [47] (see also [9]) different coordinates might have different importance, hence we can also define a weighted variation. In the spirit of the weighted Sobolev spaces in [9], let \( \gamma = (\gamma_u)_{u \subseteq N} \) be a sequence of non-negative real numbers. Then we define the weighted variation \( V_{\lambda, p, q, \gamma}(f) \) of \( f \) with coefficient \( 0 < \lambda \leq 1 \) by

\[
V_{\lambda, p, q, \gamma}(f) = \left( \sum_{u \subseteq S} \gamma_u^{-1} \left( V_{\lambda, p}^{(u)}(f_u; u) \right)^q \right)^{1/q}.
\]

Note that for \( \lambda = 1 \) and \( p = q = 2 \) the weighted variation \( V_{\lambda, p, q, \gamma}(f) \) coincides with the semi-norm in a weighted Sobolev space for any function in this Sobolev space, i.e, we have the identity \( V_{\lambda, 2, 2, \gamma}(f) = \|f\|_{\text{sob}} \), where

\[
\|f\|_{\text{sob}} = \left( \sum_{u \subseteq S} \gamma_u^{-1} \int_{[0,1]^{|u|}} \left| \int_{[0,1)^{|u|}} \frac{\partial |u| f(x)}{\partial x_u} \, dx_{S \setminus u} \right|^2 \, dx_u \right)^{1/2}
\]

denotes the norm in the weighted Sobolev space (see [9] for more information on this Sobolev space).

A slightly simpler version of the following lemma was shown in [41, Lemma 4]. Note that a multivariate version of the subsequent lemma can easily be obtained using the same arguments, but for ease of notation we show the result only in dimension one.

**Lemma 1.** Let \( f \in L_1([0,1]) \) and let \( k \in \mathbb{N} \) with \( k = \kappa q^a - 1 + k' \) where \( a \in \mathbb{N}, \kappa \in \{1, \ldots, q-1\}, 0 \leq k' < q^{a-1} \) and let \( 0 \leq c < q^{a-1} \). Then

\[
\left| \int_{[c q^{a-1}, (c+1) q^{-a+1}]} f(x) \, dx \right| \leq q^{-a+1} \sup_{x,y \in [c q^{a-1}, (c+1) q^{-a+1}] \setminus \{x,y\} \in \{1, \ldots, q^{-1}\}} |f(x) - f(y)|.
\]
Proof. We have \( \overline{\text{wal}}_k = \overline{\text{wal}}_{kq^{-a-1}} \) and the function \( \overline{\text{wal}}_k \) is constant on each subinterval \([rq^{-a}, (r + 1)q^{-a}]\). Hence we have

\[
\left| \int_{[rq^{-a}, (r + 1)q^{-a}]} f(x) \overline{\text{wal}}_k(x) \, dx \right| = \left| \int_{[rq^{-a}, (r + 1)q^{-a}]} f(x) \overline{\text{wal}}_{kq^{-a-1}}(x) \, dx \right|.
\]

Note that the function \( \overline{\text{wal}}_{kq^{-a-1}} \) is constant on each of the subintervals \([rq^{-a}, (r + 1)q^{-a}]\) for \( r = 0, \ldots, q^a - 1 \). Without loss of generality we may assume that \( c = 0 \), for all other \( c \) the result follows by the same arguments. Thus we have

\[
\int_{[0,q^{-a+1}]} f(x) \overline{\text{wal}}_{kq^{-a-1}}(x) \, dx = \sum_{r=0}^{q-1} \overline{\text{wal}}_k(r/q) \int_{[rq^{-a}, (r+1)q^{-a}]} f(x) \, dx.
\]

Let now \( a_r = \int_{[rq^{-a}, (r+1)q^{-a}]} f(x) \, dx \) and further let \( A = q^{-1} \sum_{r=0}^{q-1} a_r \). Then we have

\[
\sum_{r=0}^{q-1} \overline{\text{wal}}_k(r/q)a_r \leq \sum_{r=0}^{q-1} \overline{\text{wal}}_k(r/q)A + \sum_{r=0}^{q-1} \overline{\text{wal}}_k(r/q)(a_r - A) \leq \sum_{r=0}^{q-1} |a_r - A|,
\]

since \( \sum_{r=0}^{q-1} \overline{\text{wal}}_k(r/q) = 0 \). We have

\[
|a_r - A| = q^{-1} \sum_{t=0}^{q-1} |a_r - a_t| \leq \max_{0 \leq t < q} |a_r - a_t|.
\]

Therefore we have

\[
\left| \int_{[0,q^{-a+1}]} f(x) \overline{\text{wal}}_{kq^{-a-1}}(x) \, dx \right| \leq \sum_{r=0}^{q-1} \max_{0 \leq t < q} |a_r - a_t| \leq q \max_{0 \leq t < q} |a_r - a_t|.
\]

We have

\[
|a_r - a_t| = \left| \int_{[rq^{-a}, (r+1)q^{-a}]} (f(x) - f(x + (t-r)q^{-a})) \, dx \right| \leq q^{-a} \sup_{x \in [rq^{-a}, (r+1)q^{-a}]} |f(x) - f(x + (t-r)q^{-a})|.
\]

The result follows. \( \square \)

The following lemma is a generalization of [41, Proposition 6].

Lemma 2. Let \( 0 < \lambda \leq 1 \) and let \( f \in L_2([0,1]^s) \) satisfy \( V_{s,1,\lambda}(f) < \infty \). Then for any \( k \in \mathbb{N}_0^s \setminus \{ \emptyset \} \) the \( k \)-th Walsh coefficient of \( f \) satisfies

\[
|\hat{f}(k)| \leq q^{\lfloor w \rfloor - \lambda} \sum_{s \in \mathbb{N}^s} \sum_{a_j \in \{1, \ldots, a_j \}} V_{s,1}^{\lfloor w \rfloor}(f_u; u),
\]

where \( k = (k_1, \ldots, k_s) \), \( u = \{1 \leq j \leq s : k_j \neq 0\} \) and for \( j \in u \) we have \( k_j = k_j q^{a_j - 1} + k_j' \), where \( \kappa_j \in \{1, \ldots, q - 1\} \), \( a_j \in \mathbb{N} \) and \( 0 \leq k_j' < q^{a_j - 1} \).

Proof. As \( f \in L_2([0,1]^s) \) it follows that \( f \) can be represented by a Walsh series, i.e., we have

\[
f(x) = \sum_{k \in \mathbb{N}_0^s} \hat{f}(k) \text{wal}_k(x),
\]

where the \( k \)-th Walsh coefficient \( f(k) \) is given by

\[
\hat{f}(k) = \int_{[0,1]} f(x) \overline{\text{wal}}_k(x) \, dx.
\]
First note that it suffices to show the result for $k \in \mathbb{N}^s$, as otherwise we only need to replace the function $f$ with the function $f_u(x_u) = \int_{[0,1)^s \setminus \{x\}} f(x) \, dx |_{S^u}$. Hence let now $k \in \mathbb{N}^s$ be given and let $k' = (k_1', \ldots, k_s')$. Then we have

$$\left| \int_{[0,1)^s} f(x) \, \text{wal}_k(x) \, dx \right| \leq \sum_{0 \leq c_i < q^{s+1}} \cdots \sum_{0 \leq c_s < q^{s+1}} q^{-\sum_{j=1}^{s} (a_j-1)} \sup_{J} |\Delta(f, J)|,$$

where the supremum is taken over all boxes $J = \prod_{j=1}^{s} [d_j, e_j) \subseteq \prod_{j=1}^{s} [c_j q^{-a_j+1}, (c_j + 1) q^{-a_j+1})$ with $q^{s} |e_j - d_j| \in \{1, \ldots, q - 1\}$. Now we have

$$q^{-\sum_{j=1}^{s} (a_j-1)} \sup_{J} |\Delta(f, J)| \leq \sup_{\mathcal{P}_c} \sum_{I \in \mathcal{P}_c} \text{Vol}(I)^{1-\lambda} |\Delta(f, I)| q^{-\sum_{j=1}^{s} (a_j-1)} \leq \text{Vol}(I)^{1-\lambda} |\Delta(f, I)|,$$

where the supremum on the right hand side is taken over all partitions $\mathcal{P}_c$ of the cube $\prod_{j=1}^{s} [c_j q^{-a_j+1}, (c_j + 1) q^{-a_j+1})$ and where each $I \in \mathcal{P}_c$ is of the form $I = \prod_{j=1}^{s} [x_j, y_j)$ with $q^{s} |y_j - x_j| \in \{1, \ldots, q - 1\}$. We have $q^{-\sum_{j=1}^{s} a_j} \leq \text{Vol}(I) \leq q^{-\sum_{j=1}^{s} (a_j-1)}$ and therefore

$$q^{-\sum_{j=1}^{s} (a_j-1)} \leq \text{Vol}(I)^{1-\lambda} \leq \text{Vol}(I)^1 \leq q^{s-\lambda \sum_{j=1}^{s} (a_j-1)}$$

and hence

$$q^{-\sum_{j=1}^{s} (a_j-1)} \sup_{J} |\Delta(f, J)| \leq q^{s-\lambda \sum_{j=1}^{s} (a_j-1)} \sup_{\mathcal{P}_c} \sum_{I \in \mathcal{P}_c} \text{Vol}(I)^{1-\lambda} |\Delta(f, I)|.$$

Note that

$$\sum_{0 \leq c_i < q^{s+1}} \cdots \sum_{0 \leq c_s < q^{s+1}} \text{Vol}(I)^{1-\lambda} |\Delta(f, I)| \leq \sup_{\mathcal{P}} \sum_{J \in \mathcal{P}} \text{Vol}(J)^{1-\lambda} |\Delta(f, J)|,$$

where the supremum on the left hand side is taken over all partitions $\mathcal{P}$ of the cube $\prod_{j=1}^{s} [c_j q^{-a_j+1}, (c_j + 1) q^{-a_j+1})$ into subintervals and the supremum on the right hand side is taken over all partitions of $[0,1)^s$ into subintervals. Thus the result follows. \hfill \Box

For the next lemma we will need the following two functions. For $\kappa \in \{1, \ldots, q-1\}$ let now

$$\nu_{\kappa} = \sum_{r=0}^{q-1} r \text{wal}_\kappa(r/q).$$

If $q$ is chosen to be a prime number and the bijections $\varphi$ and $\eta$ are chosen to be the identity, then $\nu_{\kappa} = q(e^{2\pi i \kappa/q} - 1)^{-1}$, see [7, Appendix A].
Further we define the function $\zeta_c(x) = \sum_{x=0}^{x=a-1} \text{wal}(r/q)$, where $a \geq 1$ and $x = x_1q^{-1}+x_2q^{-2}+\cdots$ and where for $x_a = 0$ we set $\zeta_c(x) = 0$. The function $\zeta_c$ depends on $x$ only through $x_a$, thus it is a step-function which is constant on the intervals $[cq^{-a}, (c+1)q^{-a})$ for $c = 0, \ldots, q^a-1$. By [41, Proposition 5] it follows that $\zeta_c$ can be represented by a finite Walsh series. Indeed, there are numbers $c_0, \ldots, c_{q-1}$ such that

$$\zeta_c(x) = \sum_{z=0}^{q-1} c_z \text{wal}_{2^{q-1}}(x).$$

Again, if $q$ is chosen to be a prime number and the bijections $\varphi$ and $\eta$ are chosen to be the identity, then $\zeta_c(x) = (1 - \text{wal}_{q^{a-1}}(x))(1 - \text{wal}(1/q))^{-1}$, i.e., $c_0 = (1 - \text{wal}(1/q))^{-1}$, $c_1 = (\text{wal}(1/q) - 1)^{-1}$ and $c_z = 0$ for $z \neq 0, l$.

The following lemma appeared in a simpler form in [14].

**Lemma 3.** For $k \in \mathbb{N}_0$ and $x \in [0, 1)$ define $J_k(x) = \int_0^x \text{wal}(y) \, dy$. For $k \geq 1$ let $k = lq^{a-1} + k'$ where $l \in \{1, \ldots, q - 1\}$, $a \geq 1$ and $0 \leq k' < q^a - 1$. Then $J_k$ can be represented by a Walsh series which is given by

$$J_k(x) = q^{-a} \left( \sum_{z=0}^{q-1} c_z \text{wal}_{2^{q-1+k}}(x) + 2^{-1} \text{wal}_k(x) + \sum_{c=1}^{\infty} \sum_{c=1}^{q-1} q^{-c-1} v_c \text{wal}_{eq^{a-1}+k}(x) \right).$$

Further we have

$$J_0(x) = 1/2 + \sum_{c=1}^{\infty} \sum_{c=1}^{q-1} q^{-c-1} v_c \text{wal}_{eq^{a-1}}(x).$$

Proof. Let $k = lq^{a-1} + k'$ with $a \geq 1$, $0 \leq k' < q^{a-1}$ and $l \in \{1, \ldots, q - 1\}$. The function $\text{wal}_{eq^{a-1}}(y)$ is constant on each interval $[rq^{-a}, (r + 1)q^{-a})$ and $\text{wal}_k(y)$ is constant on each interval $[eq^{-a+1}, (c + 1)q^{-a+1})$. We have $\text{wal}_k(y) = q^{-a} \text{wal}_{2^{q^{a-1}}}(y)$. For any $0 \leq c < q^{a-1}$ we have

$$\int_{[eq^{-a+1}, (c+1)q^{-a+1})} \text{wal}_k(y) \, dy = \frac{\text{wal}_k(eq^{-a+1})}{\text{wal}_{2^{q^{a-1}}}(y)} \int_{[eq^{-a+1}, (c+1)q^{-a+1})} \text{wal}_{eq^{a-1}}(y) \, dy = \frac{\text{wal}_k(eq^{-a+1})}{\text{wal}_{2^{q^{a-1}}}(y)} \sum_{r=0}^{q^{a-1}-1} \text{wal}_l(r/q) = 0.$$

Thus we have

$$J_k(x) = \text{wal}_k(x) J_{eq^{a-1}}(x).$$

Let $x = x_1q^{-1} + x_2q^{-2} + \cdots$ and $y = x_a q^{-1} + x_a 2q^{-2} + \cdots$, then we have

$$J_{eq^{a-1}}(x) = q^{-a} \sum_{r=0}^{x_a-1} \text{wal}_l(r/q) + q^{-a} \text{wal}(x_a/q)y.$$

We now investigate the Walsh series representation of the function $J_{eq^{a-1}}(x)$. First note that $\text{wal}_l(x_a/q) = \text{wal}_{eq^{a-1}}(x)$. Further, by a slight adaption of [7, eq. (30)] we obtain

$$y = 1/2 + \sum_{c=1}^{\infty} \sum_{c=1}^{q-1} q^{-c-1} v_c \text{wal}_{eq^{a-1}}(x).$$

(5)
As \( \text{wal}_{xq^{-1}}(y) = \text{wal}_{xq^{-1}}(x) \) we obtain
\[
y = 1/2 + \sum_{c=1}^{\infty} \sum_{k=1}^{q^{-1}} q^{-c-1} u_k \text{wal}_{xq^{-1}}(x).
\]

As noted above, the Walsh series of \( \zeta_{a}(x) = \sum_{r=0}^{s-1} \text{wal}(r/q) \), where for \( x_a = 0 \) we set \( \zeta_{a}(x) = 0 \) can be written as
\[
\zeta_{a}(x) = \sum_{z=0}^{q^{-1}} c_z \text{wal}_{zq^{-1}}(x).
\]

Altogether we obtain
\[
q^a J_{kq^{-1}}(x) = \sum_{z=0}^{q^{-1}} c_z \text{wal}_{zq^{-1} + k'}(x) + 2^{-1} \text{wal}_{kq^{-1}}(x)
\]
and therefore
\[
q^a J_k(x) \equiv \sum_{z=0}^{q^{-1}} c_z \text{wal}_{zq^{-1} + k'}(x) + 2^{-1} \text{wal}_{kq^{-1}}(x) + \sum_{c=1}^{\infty} \sum_{k=1}^{q^{-1}} q^{-c-1} u_k \text{wal}_{xq^{-1}+c-1+k}(x).
\]

The result for \( k = 0 \) follows easily from (5). \( \square \)

Note that Lemma 3 can easily be generalized to arbitrary dimensions \( s \), since for \( k = (k_1, \ldots, k_s) \in \mathbb{N}_0^s \) we have for any \( x = (x_1, \ldots, x_s) \in [0,1]^s \) that
\[
J_k(x) = \int_{[0,x]} \text{wal}_k(y) \, dy = \prod_{j=1}^s J_{k_j}(x_j),
\]
where \( [0,x] = \prod_{j=1}^s [0,x_j] \).

**Lemma 4.** Let \( f \in L_2([0,1]^s) \) be given by
\[
f(x) = \sum_{k \in \mathbb{N}_0^s} \hat{f}(k) \text{wal}_k(x)
\]
and let
\[
F(x) = \int_{[0,x]} f(y) \, dy = \sum_{k \in \mathbb{N}_0^s} \hat{F}(k) \text{wal}_k(x),
\]
where \( [0,x] = \prod_{j=1}^s [0,x_j] \) with \( x = (x_1, \ldots, x_s) \). Let \( k = (k_1, \ldots, k_s) \in \mathbb{N}_0^s \) and let \( U = \{ 1 \leq j \leq s : k_j \neq 0 \} \). For \( j \in U \) let \( k_j = l_j q^a_j - 1 + k'_j \) and \( 0 \leq k'_j < q^a_j - 1 \) and further let \( k' = (k'_1, \ldots, k'_s) \) where \( k'_j = 0 \) for \( j \notin U \). Then we have
\[
\hat{F}(k) = q^{-\sum_{j \in U} a_j} \sum_{U \subseteq \nu \subseteq S} (-1)^{|\nu|} \sum_{h_{\nu} \in \mathbb{N}_0^{|\nu|}} \hat{f}(k' + (h_{\nu}, 0)) \chi_{U,v,k}(h_{\nu}),
\]
where for \( h_{\nu} = (h_j)_{j \in \nu} \in \mathbb{N}_0^{|\nu|} \) we set \( \chi_{U,v,k}(h_{\nu}) = \prod_{j \in U} \mu_j k_j(h_j) \prod_{j \in \nu \setminus U} \nu(h_j) \). Let \( \mu_j k_j(0) = c_0, \mu_j k_j(k_j - k'_j) = c_0 + 2^{-1}, \mu_j k_j(z q^a_j - 1) = c_2 \) for \( 1 \leq z < q \) and \( z \neq l_j \). Further let \( \mu_j k_j(h + k_j - k'_j) = q^{-i-1} u_z \) for \( h = z q^{i+a_j-1} \) with \( i \in \mathbb{N} \) and \( z \in \{ 1, \ldots, q - 1 \} \). For all other cases not yet covered we set \( \mu_j k_j(h) = 0 \).
Further let \( \nu(0) = 2^{-1} \) and for \( h = zq^{i-1} \) with \( i \in \mathbb{N} \) and \( z \in \{1, \ldots, q-1\} \) let 
\( \nu(h) = q^{-1} v_z. \) For all other \( h \) let \( \nu(h) = 0. \)

**Proof.** Using integration by parts in each coordinate, Fubini’s theorem and \( J_k(0) = J_k(1) = 0 \) for any \( k \in \mathbb{N} \) it follows that

\[
\hat{F}(k) = \int_{[0,1]^v} F(x) \text{Wal}_k(x) \, dx = \sum_{U \subseteq \mathbb{S}} (-1)^{|U|} \int_{[0,1]^v} J_{k_U}(x) f(x) \, dx,
\]

where for \( k = (k_1, \ldots, k_v) \) and \( x = (x_1, \ldots, x_v) \) we have \( J_{k_U}(x) = \prod_{j \in U} J_{k_j}(x_j). \)

Using the Walsh series expansion of \( J_k \) given by Lemma 3 we can now express the Walsh coefficient \( \hat{F}(k) \) as a sum of the Walsh coefficients \( f(h) \). We have

\[
\hat{F}(k) = q^{-\sum_{j \in U} a_j} \sum_{U \subseteq \mathbb{S}} (-1)^{|U|} \sum_{h_v \in \mathbb{N}_0^{|U|}} \hat{f}(k' + (h_v,0)) \chi_{U,v,k}(h_v),
\]

where \( (h_v,0) \) denotes the \( v \)-dimensional vector whose \( j \)-th component is \( h_j \) for \( j \in v \) and 0 otherwise and where \( \chi_{U,v,k}(h_v) \) is as defined above. \(\square\)

The above lemma shows how the Walsh coefficients of a function \( F = \int f \) can be obtained from the Walsh coefficients of \( f \). We need the following definition.

**Definition 2.** Let \( k = k(v; a_1, \ldots, a_v) = \kappa_1 q^{a_1-1} + \cdots + \kappa_v q^{a_v-1} \) with \( v \geq 1 \), \( \kappa_1, \ldots, \kappa_v \in \{1, \ldots, q-1\} \) and \( 1 \leq a_v < \cdots < a_1 \) be a natural number. For \( k = 0 \) we set \( v = 0 \), i.e., \( k(0) = 0 \). A function \( B : \mathbb{N}_0 \to \mathbb{R} \) is called \( q \)-adically non-increasing if \( B(k) = B(k(v; a_1, \ldots, a_v)) \) is non-increasing in \( v \) and each \( a_i \) for \( i = 1, \ldots, v \), that is, for any \( v \geq 0 \) we have

\[
B(k(v; a_1, \ldots, a_v)) \geq B(k(v+1; a_1', \ldots, a_{v+1}'))
\]

with \( 1 \leq a_{v+1}' < \cdots < a_1' \) and \( a_1, \ldots, a_v \in \{a_1', \ldots, a_{v+1}'\} \) and for an arbitrary \( 1 \leq i \leq v \) we have

\[
B(k(v; a_1, \ldots, a_v)) \geq B(k(v; a_1, \ldots, a_i-1, a_i+1, a_i+1, \ldots, a_v))
\]

provided that \( a_i + 1 < a_{i+1} \) in case \( 1 \leq i < v \).

In the following lemma we give a bound on the Walsh coefficients of \( F \) if \( f \) satisfies some smoothness condition.

**Lemma 5.** Let \( B : \mathbb{N}_0^v \to [0, \infty) \) be a \( q \)-adically non-increasing function in each variable. Let \( f \in L_2([0,1]^v) \) be given by

\[
f(x) = \sum_{k \in \mathbb{N}_0^v} \hat{f}(k) \text{Wal}_k(x)
\]

and let the Walsh coefficients of \( f \) satisfy

\[
|\hat{f}(k)| \leq B(k) \quad \text{for all } k \in \mathbb{N}_0^v.
\]

Let

\[
F(x) = \int_{[0,x]} f(y) \, dy = \sum_{k \in \mathbb{N}_0^v} \hat{F}(k) \text{Wal}_k(x),
\]

where \( [0,x] = \prod_{j=1}^v [0,x_j] \) with \( x = (x_1, \ldots, x_v) \). Let \( k = (k_1, \ldots, k_v) \in \mathbb{N}_0^v \setminus \{0\} \) and let \( U = \{1 \leq j \leq s : k_j \neq 0\} \). For \( j \in U \) let \( k_j = l_j q^{v_j-1} + k'_j \) and
0 \leq k'_j < q^a_j - 1 \) and further let \( k' = (k'_1, \ldots, k'_n) \) where \( k'_j = 0 \) for \( j \notin U \). Then there is a constant \( C_{s,U} > 0 \) independent of \( k \) such that

\[ |\hat{F}(k)| \leq C_{s,U} q^{-\sum_{j \in U} a_j} \mathcal{B}(k'). \]

Proof. Using Lemma 4 we obtain that

\[ |\hat{F}(k)| \leq q^{-\sum_{j \in U} a_j} \mathcal{B}(k') \sum_{U \subseteq \subseteq S} \sum_{h_v \in \mathbb{N}_0^n} |\chi_{U,v,k}(h_v)|, \]

as \( |\hat{f}(k' + (h_v, 0))| \leq \mathcal{B}(k' + (h_v, 0)) \leq \mathcal{B}(k') \) for all values of \( h_v \in \mathbb{N}_0^n \) for which \( \chi_{U,v,k}(h_v) \neq 0 \), since \( \mathcal{B} \) is \( q \)-adically non-increasing in each variable.

Thus it remains to bound \( \sum_{U \subseteq \subseteq S} \sum_{h_v \in \mathbb{N}_0^n} |\chi_{U,v,k}(h_v)| \) independently of \( k \). But this follows easily from Lemma 4.

For the case where \( q \) is chosen to be a prime number and the bijections \( \varphi \) and \( \eta \) are chosen to be the identity, then an explicit constant \( C > 0 \) can be obtained. Using the notation from Lemma 4 we have

\[ \sum_{h \in \mathbb{N}_0} |\mu_{j,k}(h)| = |1 - \omega_q^{-j}|^{-1} + 2^{-1} |1 + \omega_q^{-l}||\omega_q^{-l} - 1|^{-1} \]

\[ + \sum_{i=1}^{\infty} q^{-i} \sum_{z=1}^{q-1} |e^{2\pi iz/q} - 1|^{-1} \leq 3(2 - 2 \cos(2\pi/q))^{-1/2} \]

and

\[ \sum_{h=0}^{\infty} |\nu(h)| = 2^{-1} + \sum_{i=1}^{\infty} q^{-i} \sum_{z=1}^{q-1} |e^{2\pi iz/q} - 1|^{-1} \leq 2^{-1} + (2 - 2 \cos(2\pi/q))^{-1/2}. \]

Therefore we have

\[ \sum_{U \subseteq \subseteq S} \sum_{h_v \in \mathbb{N}_0^n} |\chi_{U,v,k}(h_v)| \]

\[ \leq 3^{[U]} (2 - 2 \cos(2\pi/q))^{-[U]/2} (3/2 + (2 - 2 \cos(2\pi/q))^{-1/2})^{s-[U]} \]

and hence we can choose

\[ C_{s,U} = 3^{[U]} (2 - 2 \cos(2\pi/q))^{-[U]/2} (3/2 + (2 - 2 \cos(2\pi/q))^{-1/2})^{s-[U]} \]

in Lemma 5 for this case.

We use the above results now to establish an upper bound on the Walsh coefficients of a polynomial. The proof will give a glimpse on how the argument will work for more general function classes.

Lemma 6. Let \( k = \kappa_1 q^{a_1-1} + \cdots + \kappa_v q^{a_v-1} \) with \( v \geq 1, \kappa_1, \ldots, \kappa_v \in \{1, \ldots, q-1\} \) and \( 1 \leq a_0 < \cdots < a_1 \). For \( v = 0 \) let \( k = 0 \). Let \( f : [0,1] \to \mathbb{R} \) be the polynomial \( f(x) = f_0 + f_1 x + \cdots + f_s x^s \) with \( f_s \neq 0 \) and let \( \hat{f}(k) \) denote the \( k \)-th Walsh coefficient of \( f \). Then for \( v \geq 0 \) there are constants \( 0 < C_{f,i,v} < \infty \) such that

\[ |\hat{f}(k)| \leq C_{f,i,v} q^{-a_1 - \cdots - a_v}, \]

where we can choose \( C_{f,i,v} = 0 \) for \( v > i \).
Proof. Let \( f(x) = f_0 + f_1 x + \cdots + f_i x^i \), where \( i = \deg(f) \) (that is, \( f_i \neq 0 \)). Then we have \( f^{(i)}(x) = i! f_i \neq 0 \). As \( f^{(i)} \) is a constant function, its Walsh series representation is simply given by \( f^{(i)}(x) = i! f_i \). Now we use Lemma 5. The dimension \( s \) is in our case 1 and we can choose the function \( B_1 \) by \( B_1(0) = i! f_i \) and for \( k > 0 \) we set \( B_1(k) = 0 \). Note the function \( B_1 \) defined this way is a \( q \)-adically non-increasing function. Then it follows that there is a constant \( C_1 > 0 \) such that the Walsh coefficients of the \( f_0 \) = \( \bar{f}_0 \) such that

\[
\| f \|_{L_2}^2 = \sum_{v=1}^{\infty} \frac{1}{v^s} \sum_{\tau \subset \{0, \ldots, \delta\}} \left( \int |f^{(\tau)}(x)|^2 \, dx \right)^{1/2},
\]

where for \( u(\tau) = 0 \) we set \( V_{\Lambda_{p,q,1}}^{(0)}(f^{(\tau)}, \mathcal{S}_{\Lambda_{p,q,1}}) = |f^{(\tau)}(0)| \). The weights \( \gamma \) are introduced to modify the importance of various coordinate projections and were first introduced in [47], see also [9, 10]. If for some \( u' \subseteq S \) the weight \( \gamma_{\nu} = 0 \), then we assume that the function \( f \) satisfies \( V_{\Lambda_{p,q,1}}^{(w)}(f^{(\tau)}(\cdot, \mathcal{S}_{\Lambda_{p,q,1}})) = |f^{(\tau)}(0)| \) for all \( \tau \in \{0, \ldots, \delta\}^s \) with \( v(\tau) = u' \) and in (8) we formally set \( 0/0 = 0 \).

Note that for \( \lambda = 1 \) and \( p = q = r = 2 \) the functional \( N_{\Lambda_{p,q,1}}^{(u)} \) is just the norm in a weighted reproducing kernel Sobolev space with continuous partial mixed derivatives up to order \( \delta + 1 \) in each variable. In one dimension the unweighted
norm in this reproducing kernel Sobolev space is given by

\[ \langle f, g \rangle_{\text{sob}, an, \delta+1} = f(0)g(0) + \cdots + f^{(\delta-1)}(0)g^{(\delta-1)}(0) \]

\[ + \int_0^1 f^{(\delta)}(x) \, dx \int_0^1 g^{(\delta)}(x) \, dx + \int_0^1 f^{(\delta+1)}(x)g^{(\delta+1)}(x) \, dx \]

and for higher dimensions one just takes the weighted tensor product of the one dimensional reproducing kernel Sobolev spaces (see [10] for examples of weighted tensor product reproducing kernel Sobolev spaces). Let the \( s \) dimensional weighted inner product be denoted by \( \langle \cdot, \cdot \rangle_{\text{sob}, an, s, \delta+1, \gamma} \) and the corresponding norm \( \| \cdot \|_{\text{sob}, an, s, \delta+1, \gamma} \) (indeed if the partial mixed derivatives up to order \( \delta + 1 \) of \( f \) are continuous on \([0,1]^{s} \) then we have \( \| f \|_{\text{sob}, an, s, \delta+1, \gamma} = N_{\delta, 1.2, 2, 2, \gamma}(f) \)).

**Theorem 1.** Let \( \delta \geq 1, 0 < \lambda \leq 1, p, q, r \geq 1 \) and a sequence \( \gamma = (\gamma_v)_{v \in \mathbb{N}} \) of nonnegative real numbers be given. Let \( f : [0,1]^{s} \to \mathbb{R} \) be such that the partial mixed derivatives up to order \( \delta \) in each variable exist and are continuous and such that

\[ N_{\delta, \lambda, p, q, r, \gamma}(f) < \infty. \]

Then for any \( k \in \mathbb{N}_0^s \setminus \{0\} \) it follows that there is a constant \( C_{f, q, s, \gamma} \) independent of \( k \) such that

\[ |\hat{f}(k)| \leq C_{f, q, s, \gamma} q^{-\mu_{\delta, \lambda}(k)}, \]

where for \( k = (k_1, \ldots, k_u) \) we set \( \mu_{\delta, \lambda}(k) = \sum_{j=1}^{s} \mu_{\delta, \lambda}(j) \) with \( \mu_{\delta, \lambda}(j) = 0 \) and for \( k = (k_1 \alpha_1 - 1 + \cdots + k_u \alpha_u - 1 \} \) \( \frac{1}{\lambda} \) with \( v \geq 1, k_1, \ldots, k_u \in \{1, \ldots, q - 1\} \) and \( 1 \leq a_u < \cdots < a_1 \) we set \( \mu_{\delta, \lambda}(k) = a_1 + \cdots + a_v \) if \( v \leq \delta \) and \( \mu_{\delta, \lambda}(k) = a_1 + \cdots + a_v + \lambda a_{v+1} \) if \( v > \delta \).

**Proof.** In order to prove the result we use the Taylor series expansion of the function \( f \). We have

\[ f(x) = \sum_{\tau \in \{0, \ldots, \delta-1\}^{n}} \frac{x^{\tau}}{\tau!} f^{(\tau)}(0) + \sum_{0 \neq u \subseteq S} \sum_{\tau(S \setminus u) \in \{0, \ldots, \delta-1\}^{\tau(S \setminus u)}} ((\delta-1)!)^{-|u|} \]

\[ \times \frac{x^{\tau(S \setminus u)}}{\tau(S \setminus u)!} \int_{[0,1]^{\tau(S \setminus u)}} f^{(\delta, \tau(S \setminus u))}(y_u, 0_{S \setminus u})(x_u - y_u)^{\delta-1} \, dy_u, \]

(10)

where \( x^{\tau(S \setminus u)} = \prod_{j \in S \setminus u} x_j^{\tau_j} \), \( \tau(S \setminus u)! = \prod_{j \in S \setminus u} \tau_j! \) and \( (x_u - y_u)^{\delta-1} = \prod_{j \in u} (x_j - y_j)^{\delta-1} \).

First note that the first sum in (10) is a polynomial in \( x \) and therefore the Walsh coefficients of this polynomial satisfy the desired bound by Lemma 6.

Now we consider the second sum. Let \( \emptyset \neq u \subseteq S \) with \( u = \{j_1, \ldots, j_{|u|}\} \) be given. Then for \( j \notin u \) the Walsh coefficients satisfy the desired bound by Lemma 6. Hence it remains to consider the Walsh coefficients of \( G_u(x_u) = \int_{[0,1]^{\tau(S \setminus u)}} f^{(\delta, \tau(S \setminus u))}(y_u, 0_{S \setminus u})(x_u - y_u)^{\delta-1} \, dy_u \)

\[ = \int_{[0,1]^{\tau(S \setminus u)}} f^{(\delta, \tau(S \setminus u))}(y_u, 0_{S \setminus u})(x_u - y_u)^{\delta-1} \, dy_u, \]

where \( (x - y)_+ = \max(0, x - y) \) and \( (x_u - y_u)_+ = \prod_{j \in u} (x_j - y_j)_+ \).
By differentiating the function $G_u$ in each variable $0 \leq k < \delta$ times we obtain (see [49, pp. 153,154])

$$\frac{\partial^{|u|}}{\partial x^u} G_u(x_u) = \left( \frac{(\delta - 1)!}{(\delta - 1 - k)!} \right)^{|u|} \int_{[0,x_u)} f^{(\delta_u,\tau_S\setminus u)}(y_u,0_{S\setminus u})(x_u - y_u)^{\delta - 1 - k} \, dy_u.$$ 

Hence $\frac{\partial^{|u|}}{\partial x^u} G_u(x_u) = 0$ if there is at least one $j \in u$ such that $x_j = 0$. Further we have

$$\frac{\partial^{|u|}}{\partial x^u} G_u(x_u) = ((\delta - 1))^{\delta_u} \int_{[0,x_u)} f^{(\delta_u,\tau_S\setminus u)}(x_u,0_{S\setminus u}).$$

From $N_{\delta,\lambda,p,q,\tau,\gamma}(f) < \infty$ it follows that the Walsh coefficients of $f(\delta_u,\tau_S\setminus u)(x_u,0_{S\setminus u})$ decay with order $\mu_{\gamma,0+\lambda}(k)$ in each variable. Further we have

$$G_u(x_u) = \int_{[0,x_u)} \int_{[0,y_1)} \cdots \int_{[0,y_{u-1})} G_{u}^{(\delta_u)}(y_\delta) \, dy_\delta \cdots dy_1$$

as the function $G_u$ and its derivatives are 0 if at least one $x_j = 0$ for $j \in u$, i.e., we have

$$\int_{[0,x_u)} G_{u}^{(\tau)}(y) \, dy = \sum_{v \subseteq u} (-1)^{|v|} G_{u}^{(\tau-1)}(x_v,0_{u \setminus v}) = G_{u}^{(\tau-1)}(x_u).$$

Hence it follows by repeated use of Lemma 5 that the desired bound holds for $G_u$ and thus the result follows from (10).

For the case where $q$ is chosen to be a prime number and the bijections $\varphi$ and $\eta$ are chosen to be the identity we can also obtain an explicit constant in Theorem 1. Indeed, using Lemma 2, Lemma 5 together with the explicit constant (6) and Lemma 6 together with the explicit constant (7) we obtain that the constant $C_{\delta,\lambda}$ can be chosen as

$$C_{f,q,s,\gamma} = \sum_{\tau \in \{0,\ldots,\delta-1\}^s} |f^{(\tau)}(0)| \hat{C}_{\tau,\tau^1+\cdots+\tau_s} + \sum_{\emptyset \neq u \subseteq S} q^{|u|} C_{s,u}^{\delta} \sum_{\tau \in \{0,\ldots,\delta-1\}^s} \hat{C} \sum_{\tau \in \{0,\ldots,\delta-1\}^s} |V_{\lambda,1}^{(|u|)}(f^{(\delta,\tau_S\setminus u)}(\cdot,0_{S\setminus u}))|,$$

where $\hat{C} = 1$ for $2 \leq q < 6$ and $\hat{C} = (2 - 2 \cos(2\pi/q))^{-1/2}$ for $q > 6$ (note that for $q > 6$ we have $\hat{C} > 1$) and

$$C_{s,u} = 3^{|u|}(2 - 2 \cos(2\pi/q))^{-|u|/2} (3^2 + (2 - 2 \cos(2\pi/q))^{-1/2})^{s-|u|}.$$ 

As noted above, under certain conditions we can write $V_{\lambda,1}^{(|u|)}$ also as an integral, see (3).

We give a further useful estimation of the constant by separating the dependence of the function from the constants. This way we obtain

$$C_{f,q,s,\gamma} \leq C_{\delta,\lambda,p,q,\tau,\gamma}(f),$$
where
\[ C_{\delta,q,s,\gamma} = \sum_{\tau \in \{0, \ldots, \delta-1\}^s} \gamma_{u}\hat{C}_{\tau^1 + \cdots + \tau_s} \]
\[ + \sum_{\emptyset \neq u \subseteq S} q^{|u|} C_{\delta,u}^s \sum_{\tau \in \{0, \ldots, \delta-1\}^s, \tau \not\in \gamma_{u}} \gamma_{u\tau}\hat{C}_{\sum_{j \in S\setminus u}\tau_j}. \]

Consider now the case where the weights are of product form (see [10]), i.e., there is a sequence of positive real numbers \((\gamma_j)_{j \in \mathbb{N}}\) such that \(\gamma_v = \prod_{j \in v} \gamma_j\) for all \(v \subset \mathbb{N}\) and for \(v = \emptyset\) we set \(\gamma_v = 1\). If now \(q\) is prime with \(2 \leq q < 6\), then
\[ C_{\delta,q,s,\gamma} = \prod_{j=1}^{s} (1 + \gamma_j(\delta - 1)) + \prod_{j=1}^{s} \left[ (1 + \gamma_j(\delta - 1))(3/2 + \hat{C}) + \gamma_j \beta \hat{C} \right] \]
\[ - \prod_{j=1}^{s} \left[ (1 + \gamma_j(\delta - 1))(3/2 + \hat{C}) \right]. \]

For example for \(q = 2\), \(\delta = 1\) and product weights we obtain
\[ C_{1,2,s,\gamma} = 1 - 2^s + 2^s \prod_{j=1}^{s} (1 + 6\gamma_j). \]

The approach used here for prime \(q\) and \(\varphi\) the identity map can also be used for arbitrary prime powers \(q\) and arbitrary mappings \(\varphi\) with \(\varphi(0) = 0\). Hence we obtain the following corollary.

**Corollary 1.** Under the assumptions of Theorem 1 there exists a constant \(C_{\delta,q,s,\gamma} > 0\) independent of \(k\) and \(f\) such that
\[ |\hat{f}(k)| \leq C_{\delta,q,s,\gamma} \mu_{\delta,\lambda,p,q,(s)\gamma}(f) q^{-2\mu_\delta + \lambda(k)} \quad \text{for all } k \in \mathbb{N}_0^s. \]

**Remark 2.** The results in this section also hold for the following generalization. In the definition of \(N_{\delta,\lambda,p,q,(s)\gamma}\) we anchored the function and its derivatives at \(0\), i.e., we used \(W_{\lambda,p,q,(s)\gamma}^f(\cdot,0_{S\setminus u})\). This can be generalized by choosing an arbitrary \(a \in [0,1]^s\) and using \(W_{\lambda,p,q,(s)\gamma}^f(\cdot, a_{S\setminus u})\) in the definition of \(N_{\delta,\lambda,p,q,(s)\gamma}\). It can be shown that in this case we also have Theorem 1 and Corollary 1.

### 3.2. Convergence of the Walsh series
For our purposes here we need strong assumptions on the convergence of the Walsh series \(S(f)(x) = \sum_{k \in \mathbb{N}_0^s} \hat{f}(k) \text{wal}_k(x)\) to the function \(f\), i.e., we require that the partial series \(S_m(f)(x) = \sum_{k \in \mathbb{N}_0^s, \|k\|_m} \hat{f}(k) \text{wal}_k(x)\) converges to \(f(x)\) at every point \(x \in [0,1]^s\) as \(m \to \infty\). (Note that the Walsh series \(S(f)\) for the functions considered in this paper is always absolutely convergent, i.e., \(\sum_{k \in \mathbb{N}_0^s} |\hat{f}(k)| < \infty\), hence the Walsh series \(S(f)(x)\) is uniformly bounded by \(\sum_{k \in \mathbb{N}_0^s} |\hat{f}(k)|\) and therefore \(S(f)(x)\) itself converges at every point \(x \in [0,1]^s\).)

This is necessary as we want to approximate the integral at function values \(x\), and for our analysis we deal with the Walsh series rather than the function itself, hence it is paramount that the function and its Walsh series coincide at every point \(x \in [0,1]^s\).

As the functions considered here are at least differentiable it follows that they are continuous and using the argument in [14, p. 373] it follows that the Walsh
series really converges at every point \( x \in [0, 1]^s \) to the function value \( f(x) \). Indeed, for a given \( x \in [0, 1]^s \) we have

\[
S_q^m(f)(x) = \sum_{k \in \{0, \ldots, q^m-1\}^s} \hat{f}(k)\text{wal}_k(x) = \text{Vol}(J_x)^{-1} \int_{J_x} f(x) \, dx,
\]

where \( J_x = \prod_{j=1}^s [q^{-m} [q^m x_j], q^{-m} [q^m x_j] + q^{-m}] \). The last equality follows from

\[
\sum_{k \in \{0, \ldots, q^m-1\}^s} \hat{f}(k)\text{wal}_k(x)
= \int_{[0,1]^s} f(y) \sum_{k \in \{0, \ldots, q^m-1\}^s} \text{wal}_k(x)\text{wal}_k(y) \, dy
= \text{Vol}(J_x)^{-1} \int_{J_x} f(x) \, dx.
\]

As the function \( f \) is continuous it immediately follows that \( S_q^m(f)(x) \) converges to \( f(x) \) as \( m \) goes to infinity and the result follows.

### 3.3. A function space based on Walsh functions containing smooth functions.

In this section we use the above results to define a function space based on Walsh functions which contains smooth functions for smoothness conditions considered in the previous section.

Let \( \vartheta > 1 \) be a real number and \( q \) a prime power. First let \( r_{\vartheta, q}(0) = 1 \). Let \( k \in \mathbb{N} \) with \( k = \kappa_1 q^{a_1-1} + \cdots + \kappa_{\vartheta} q^{a_{\vartheta}-1} \) with \( v \geq 1, \kappa_1, \ldots, \kappa_{\vartheta} \in \{1, \ldots, q-1\} \) and \( 1 \leq a_v < \cdots < a_{\vartheta} \). For \( v \leq \lfloor \vartheta \rfloor \) we set \( r_{\vartheta, q}(k) = q^{-a_1-\cdots-a_v} \) and for \( v > \vartheta \) we set \( r_{\vartheta, q}(k) = q^{-a_1-\cdots-a_v}(q^{-\vartheta}) a_{\vartheta+1} \). For vectors \( k \in \mathbb{N}_0^s \) we set \( r_{\vartheta, q}(k) = \prod_{j=1}^s r_{\vartheta, q}(k_j) \). Note that we have \( r_{\vartheta, q}(k) = q^{-\mu_{\vartheta, q}(k)} \).

Now we define a function space \( \mathcal{E}_{s, q, \vartheta, \gamma} \subseteq \mathcal{L}_2([0, 1]^s) \) with norm \( \| \cdot \|_{\mathcal{E}_{s, q, \vartheta, \gamma}} \) given by

\[
\|f\|_{\mathcal{E}_{s, q, \vartheta, \gamma}} = \max_{\gamma_u \geq 0} \gamma_u^{-1} \sup_{k_u \in \mathbb{N}^{|u|}} \frac{|\hat{f}(k_u, 0_{S\setminus u})|}{r_{\vartheta, q}(k_u)},
\]

where again for \( \gamma_u = 0 \) we assume that \( \hat{f}(k_u, 0_{S\setminus u}) = 0 \) for all \( k_u \in \mathbb{N}^{|u|} \).

The following result follows now directly from Corollary 1.

**Corollary 2.** Let \( \delta \geq 1, 0 < \lambda \leq 1, p, q, \vartheta \geq 1 \) and a sequence \( \gamma = (\gamma_u)_{u \in \mathbb{N}} \) of non-negative real numbers be given. Then there exists a constant \( C_{\delta, q, \vartheta, \gamma} > 0 \) such that for every function \( f : [0, 1]^s \to \mathbb{R} \) which has continuous partial mixed derivatives up to order \( \delta \) we have

\[
\|f\|_{\mathcal{E}_{s, q, \vartheta, \delta, \gamma}} \leq C_{\delta, q, \vartheta, \gamma} N_{\delta, \lambda, p, q, \vartheta, \gamma}(f).
\]

Again, using (11) an explicit constant in Corollary 2 can be obtained for \( q \) prime and \( \vartheta \). For all other cases (i.e., arbitrary prime powers \( q \) and mappings \( \varphi \) with \( \varphi(0) = 0 \)) explicit constants can be obtained as well, but in this case the constant may also depend on the particular choice of \( q \) and \( \varphi \). Further, as noted already above, for \( \lambda = 1 \) and \( p = q = \vartheta = 2 \) the functional \( N_{\delta, \lambda, p, q, \vartheta, \gamma} \) coincides with the norm in a certain Sobolev space (the one dimensional inner product for this Sobolev space is given by (9) and for higher dimensions one just considers tensor products of the one dimensional space) and hence it follows that \( \mathcal{E}_{s, q, \vartheta, \delta, \gamma} \) contains certain Sobolev spaces. Hence Corollary 2 shows that if we want to prove results for smooth functions it is enough to consider \( \mathcal{E}_{s, q, \vartheta, \gamma} \).
A function \( f \in \mathcal{E}_{s,q,t}\) can be written as a sum of their anova terms \( f = \sum_{u \subseteq S} f_u \) (see [11]). For a function \( f \in \mathcal{E}_{s,q,t}\) given by \( f(x) = \sum_{k \in \mathbb{N}_0^t} \hat{f}(k) \text{wal}_k(x) \) the anova term \( f_u \) corresponding to a subset \( u \subseteq S \) is simply given by
\[
 f_u(x_u) = \sum_{k_u \in \mathbb{N}^{\lceil \omega \rceil}} \hat{f}(k_u, 0_{S\setminus u}) \text{wal}_{k_u}(x_u).
\]

If for some \( u \subseteq S \) we have \( \gamma_u = 0 \), then this implies that the anova term corresponding to \( u \) satisfies \( f_u \equiv 0 \). Hence the Walsh space \( \mathcal{E}_{s,q,t} \) consists only of functions whose anova term belonging to a subset is zero for all subsets \( u \) with \( \gamma_u = 0 \) (see also [10]).

4. Digital \((t, \alpha, \beta, n \times m, s)\)-nets and digital \((t, \alpha, \beta, \sigma, s)\)-sequences

In this section we give the definition of digital \((t, \alpha, \beta, n \times m, s)\)-nets and digital \((t, \alpha, \beta, \sigma, s)\)-sequences. Similar point sets were introduced in [3].

4.1. The digital construction scheme. The construction of the point set used here is a slight generalization of the digital construction scheme introduced by Niederreiter, see [30], by breaking with the tradition of having square generating matrices.

**Definition 3.** Let \( q \) be a prime-power and let \( n, m, s \geq 1 \) be integers. Let \( C_1, \ldots, C_s \) be \( n \times m \) matrices over the finite field \( \mathbb{F}_q \) of order \( q \). Now we construct \( q^m \) points in \([0,1)^s\) for \( 0 \leq h \leq q^m - 1 \) let \( h = h_0 + h_1 q + \cdots + h_{m-1} q^{m-1} \) be the \( q \)-adic expansion of \( h \). Consider an arbitrary but fixed bijection \( \varphi : \{0,1,\ldots,q-1\} \rightarrow \mathbb{F}_q \). Identify with the vector \( \vec{h} = (\varphi(h_0), \ldots, \varphi(h_{m-1}))^T \in \mathbb{F}_q^m \), where \( \top \) means the transpose of the vector. For \( 1 \leq j \leq s \) multiply the matrix \( C_j \) by \( \vec{h} \), i.e.,
\[
 C_j \vec{h} := (y_{j,1}(h), \ldots, y_{j,n}(h))^T \in \mathbb{F}_q^n,
\]
and set
\[
 x_{h,j} := \frac{\varphi^{-1}(y_{j,1}(h))}{q^1} + \cdots + \frac{\varphi^{-1}(y_{j,n}(h))}{q^n}.
\]
The point set \( \{x_0, \ldots, x_{q^m-1}\} \) is called a digital net (over \( \mathbb{F}_q \)) (with generating matrices \( C_1, \ldots, C_s \)).

For \( n, m = \infty \) we obtain a sequence \( \{x_0, x_1, \ldots\} \), which is called a digital sequence (over \( \mathbb{F}_q \)) (with generating matrices \( C_1, \ldots, C_s \)).

Niederreiter’s concept of a digital \((t, m, s)\)-net and a digital \((t, s)\)-sequence will appear as a special case in the subsequent section. Further, the digital nets considered below all satisfy \( n \geq m \).

For a digital net with generating matrices \( C_1, \ldots, C_s \) let \( \mathcal{D} = \mathcal{D}(C_1, \ldots, C_s) \) be the dual net given by
\[
 \mathcal{D} = \{k \in \mathbb{N}_0^s \setminus \{0\} : C_1^T \vec{k}_1 + \cdots + C_s^T \vec{k}_s = 0\},
\]
where for \( k = (k_1, \ldots, k_s) \) with \( k_j = k_{j,0} + k_{j,1} q + \cdots \) and \( k_{j,i} \in \{0, \ldots, q-1\} \) let \( \vec{k}_j = (\varphi(k_{j,0}), \ldots, \varphi(k_{j,n-1}))^T \). Further, for \( \emptyset \neq u \subseteq S \) let \( \mathcal{D}_u = \mathcal{D}(C_j)_{j \in u} \) and \( \mathcal{D}_u^* = \mathcal{D}_u \cap \mathbb{N}^{\lceil u \rceil} \).

Note that throughout the paper Walsh functions and digital nets are defined using the same finite field \( \mathbb{F}_q \) and the same bijection \( \varphi \).

The following lemma is a slight generalization of [42, Lemma 2.5].
Lemma 7. Let \( \{x_0, \ldots, x_{q^n-1}\} \) be a digital net over \( \mathbb{F}_q \) with bijection \( \varphi \), where \( \varphi(0) = 0 \), generated by the \( n \times m \) matrices \( C_1, \ldots, C_s \) over \( \mathbb{F}_q \), \( n, m \geq 1 \). Then for any vector \( k = (k_1, \ldots, k_s) \) of nonnegative integers \( 0 \leq k_1, \ldots, k_s < q^n \) we have

\[
\sum_{h=0}^{q^n-1} r_{a, \varphi}(x_h) = \begin{cases} q^n & \text{if } k \in \mathcal{D} \cup \{0\}, \\ 0 & \text{else}, \end{cases}
\]

where 0 is the zero vector in \( \mathbb{N}_0^n \).

4.2. \((t, \alpha, \beta, n \times m, s)\)-nets and \((t, \alpha, \beta, s)\)-sequences. Digital \((t, \alpha, \beta, m, s)\)-nets and digital \((t, \alpha, \beta, n \times m, s)\)-nets were first introduced in [5]. Those point sets were used for quasi-Monte Carlo rules which achieve the optimal rate of convergence of the worst-case error in Korobov spaces (which are reproducing kernel Hilbert spaces of smooth periodic functions). By a slight generalization of digital \((t, \alpha, \beta, m, s)\)-nets we will show that those digital nets also achieve the optimal convergence of the worst-case error in the space \( E_t, \alpha, \beta, m, s \) the periodicity assumption necessary in [5]. Further, if the derivatives up to order \( \delta \) also have bounded variation with coefficient \( 0 < \lambda \leq 1 \), then we have shown that such functions are in \( E_{s,q,\vartheta,\delta} \). We will show that digital \((t, \alpha, n \times m, s)\)-nets and digital \((t, \alpha, s)\)-sequences work well for numerical integration in the space \( E_{s,q,\vartheta,\delta} \) as long as \( \alpha \geq \lceil \vartheta \rceil \) and hence also for smooth functions. In the following we repeat some definitions and results from [5] and give the definition of digital \((t, \alpha, \beta, n \times m, s)\)-nets and digital \((t, \alpha, \beta, \vartheta, s)\)-sequences.

For a real number \( \vartheta > 1 \) the definition of the Walsh space \( E_{s,q,\vartheta,\delta} \) suggests to define the following metric \( \mu_{q,\vartheta}(k, l) = \mu_{q,\vartheta}(k \odot l) \) on \( \mathbb{N}_0^n \), which is an extension of the metric introduced in [28], see also [44] (the metric for \( \vartheta = 1 \) can be used for Walsh spaces for example considered in [7]; for this case one basically obtains the metric in [28, 44]). Let \( \mu_{q,\vartheta}(0) = 0 \). Let \( k \in \mathbb{N} \) be given by \( k = k_1 q^{a_1} + \cdots + k_s q^{a_s} \) where \( 0 \leq a_0 < \cdots < a_1 \) and \( \alpha \in \{1, \ldots, q-1\} \). Then for \( v < \vartheta \) let \( \mu_{q,\vartheta}(k) = a_1 + \cdots + a_\vartheta \) and for \( v \geq \vartheta \) let \( \mu_{q,\vartheta}(k) = a_1 + \cdots + a_\vartheta + (v - \lfloor v \rfloor) a_\vartheta \). For a \( k \in \mathbb{N}_0^n \) with \( k = (k_1, \ldots, k_s) \) let \( \mu_{q,\vartheta}(k) = \mu_{q,\vartheta}(k_1) + \cdots + \mu_{q,\vartheta}(k_s) \). Then we have \( r_{q,\vartheta}(k) = q^{-\mu_{q,\vartheta}(k)} \). (Note that this definition of \( \mu_{q,\vartheta} \) is the same as used in Theorem 1.) As we will see later, in order to obtain a small worst-case error in the Walsh space \( E_{s,q,\vartheta,\delta} \) we need digital nets for which \( \min\{\mu_{q,\vartheta}(k) : k \in \mathcal{D} \} \) is large. By translating this property into a linear independence property of the row vectors of the generating matrices \( C_1, \ldots, C_s \) we arrive at the following definition.

Definition 4. Let \( n, m, \alpha \geq 1 \) be natural numbers, let \( 0 < \beta \leq \alpha n/m \) be a real number and let \( 0 \leq t \leq \beta n \) be a natural number. Let \( \mathbb{F}_q \) be the finite field of prime power order \( q \) and let \( C_1, \ldots, C_s \in \mathbb{F}_q^{n \times m} \) with \( C_j = (c_{j,1}, \ldots, c_{j,n})^\top \). If for
all } 1 \leq i_{j,\nu} < \cdots < i_{j,1} \leq m, \text{ where } 0 \leq \nu_j \leq n \text{ for all } j = 1, \ldots, s, \text{ with }
\sum_{j=1}^{s} i_{j,1} + \cdots + i_{j,\min(\nu_j,\alpha)} + \cdots + i_{s,1} + \cdots + i_{s,\min(\nu_s,\alpha)} \leq \beta n - t
\text{the vectors }
\begin{align*}
c_{1,i_1,\nu_1}, \ldots, c_{1,i_1,1}, \ldots, c_{s,i_s,\nu_s}, \ldots, c_{s,i_s,1}
\end{align*}
\text{are linearly independent over } \mathbb{F}_q \text{ then the digital net with generating matrices }
C_1, \ldots, C_s \text{ is called a digital } (t, \alpha, \beta, n \times m, s)-\text{net over } \mathbb{F}_q. \text{ Further we call a digital }
(t, \alpha, \beta, n \times m, s)-\text{net over } \mathbb{F}_q \text{ a digital } (t, \alpha, n \times m, s)-\text{net over } \mathbb{F}_q.
\text{If } t \text{ is the smallest non-negative integer such that the digital net generated by }
C_1, \ldots, C_s \text{ is a digital } (t, \alpha, \beta, n \times m, s)-\text{net, then we call the digital net a strict digital }
(t, \alpha, \beta, n \times m, s)-\text{net or a strict digital } (t, \alpha, n \times m, s)-\text{net if } \alpha = \beta.

\textbf{Remark 3.} Using duality theory (see [34]) it follows that for a digital } (t, \alpha, \beta, n \times m, s)-\text{net we have } \min_{k \in D} \mu_{q, \alpha}(k) > \beta n - t \text{ and for a strict digital } (t, \alpha, \beta, n \times m, s)-\text{net we have } \min_{k \in D} \mu_{q, \alpha}(k) = \beta n - t + 1. \text{ Hence digital } (t, \alpha, \beta, n \times m, s)-\text{nets with high quality have a large value of } \beta n - t.

Digital } (t, \alpha, \beta, n \times m, s)-\text{nets do not exist for arbitrary choices of the parameters } t, \alpha, \beta, n, m, s, \text{ see [5]. The digital nets considered in [5] had the restriction that } n = m \text{ and special attention was payed to those digital nets with high quality, i.e., where } \alpha = \beta. \text{ In this paper, a special role will be played by those digital nets for which } n = \alpha m \text{ and } \beta = 1. \text{ The restriction on the linear independence of the digital nets comprises now } n - t = \alpha m - t \text{ row vectors, which is the same as in [5], with the only difference that the size of the generating matrices is now bigger as now each generating matrix has } n = \alpha m \text{ rows. As those digital nets play a special role in this work we have the following definition.}

\textbf{Definition 5.} A digital } (t, \alpha, \beta, n \times m, s)-\text{net over } \mathbb{F}_q \text{ is called a digital } (t, \alpha, \beta, n \times m, s)-\text{net over } \mathbb{F}_q. \text{ A strict digital } (t, \alpha, \beta, \alpha m \times m, s)-\text{net over } \mathbb{F}_q \text{ is called a strict digital } (t, \alpha, \beta, \alpha m \times m, s)-\text{net over } \mathbb{F}_q.

We can also define sequences of points for which the first } q^n \text{ points form a digital } (t, \alpha, \beta, n \times m, s)-\text{nets.

\textbf{Definition 6.} Let } \alpha, \sigma \geq 1 \text{ and } t \geq 0 \text{ be integers and let } 0 < \beta \leq \alpha / \sigma \text{ be a real number. Let } \mathbb{F}_q \text{ be the finite field of prime power order } q \text{ and let } C_1, \ldots, C_s \in \mathbb{F}_q^{\infty \times \infty} \text{ with } C_j = (c_{j,1}, c_{j,2}, \ldots)^\top. \text{ Further let } C_j, \sigma m \times m \text{ denote the left upper } \sigma m \times m \text{ submatrix of } C_j. \text{ If for all } m > t / (\beta \sigma) \text{ the matrices } C_1, \sigma m \times m, \ldots, C_s, \sigma m \times m \text{ generate a digital } (t, \alpha, \beta, \sigma m \times m, s)-\text{net} \text{ then the digital sequence with generating matrices } C_1, \ldots, C_s \text{ is called a digital } (t, \alpha, \beta, \sigma m \times m, s)-\text{sequence over } \mathbb{F}_q. \text{ Further we call a digital } (t, \alpha, \beta, \sigma, s)-\text{sequence over } \mathbb{F}_q \text{ a digital } (t, \alpha, \beta, \sigma, s)-\text{sequence over } \mathbb{F}_q.
\text{If } t \text{ is the smallest non-negative integer such that the digital sequence generated by } C_1, \ldots, C_s \text{ is a digital } (t, \alpha, \beta, \sigma, s)-\text{sequence, then we call the digital sequence a strict digital } (t, \alpha, \beta, \sigma, s)-\text{sequence or a strict digital } (t, \alpha, s)-\text{sequence if } \alpha = \sigma \text{ and } \beta = 1.

For short we will often write } (t, \alpha, \beta, \sigma, n \times m, s)-\text{net instead of digital } (t, \alpha, \beta, \sigma, n \times m, s)-\text{net over } \mathbb{F}_q. \text{ The same applies to the other notions defined above.

\textbf{Remark 4.} Note that the definition of a digital } (t, 1, m \times m, s)-\text{net coincides with the definition of a digital } (t, m, s)-\text{net and the definition of a digital } (t, 1, s)-\text{sequence.
coincides with the definition of a digital \((t, s)\)-sequence as defined by Niederreiter [30]. Further note that the \(t\)-value depends on \(\alpha, \beta\) and \(\sigma\), i.e., \(t = t(\alpha, \beta, \sigma)\) or \(t = t(\alpha)\) if \(\alpha = \beta = \sigma\).

The definition of \((t, \alpha, s)\)-sequences here differs slightly from the definition in [5]. Indeed the definition of a \((t, \alpha, s)\)-sequence in [5] corresponds to a \((t, \alpha, \alpha, 1, s)\)-sequence in the terminology of this paper, whereas here we call a \((t, \alpha, 1, \alpha, s)\)-sequence a \((t, \alpha, s)\)-sequence. On the other hand note that the condition of linear independence in Definition 4 is the same in both cases, i.e., the sum \(i_{1,1} + \cdots + i_{1,\min(\nu_1, \alpha)} + \cdots + i_{s,1} + \cdots + i_{s,\min(\nu_s, \alpha)}\) needs to be bounded by \(om - t\) for all \(m\) for \((t, \alpha, 1, \alpha, s)\)-sequences and also for \((t, \alpha, \alpha, 1, s)\)-sequences.

4.3. Some properties of \((t, \alpha, \beta, n \times m, s)\)-nets and \((t, \alpha, \beta, \sigma, s)\)-sequences.

The properties of such digital nets and sequences shown in [5] also hold here. For example it was shown there that a digital \((t, \alpha, m, s)\)-net is also a digital \((\lceil \alpha \rceil, \alpha', m, s)\)-net for all \(1 \leq \alpha' \leq \alpha\) and every digital \((t, \alpha, s)\)-sequence is also a digital \((\lceil \alpha \rceil, \alpha', s)\)-sequence for all \(1 \leq \alpha' \leq \alpha\). In the same way we have the following theorem.

**Theorem 5.** Let \(P \) be a digital \((t, \alpha, \beta, n \times m, s)\)-net over \(F_q\) and let \(S\) be a digital \((t, \alpha, \beta, \sigma, s)\)-sequence over \(F_q\). Then we have:

(i) \(P\) is a digital \((t', \alpha', \beta', n \times m, s)\)-net for all \(1 \leq \beta' \leq \beta\) and all \(t \leq t' \leq \beta m\) and \(S\) is a digital \((t', \alpha', \beta', \sigma, s)\)-sequence for all \(1 \leq \beta' \leq \beta\) and all \(t \leq t'\).

(ii) \(P\) is a digital \((t', \alpha', \beta', n \times m, s)\)-net for all \(1 \leq \alpha' \leq m\) where \(\beta' = \beta \min(\alpha, \alpha')/\alpha\) and \(t' = \lceil \min(\alpha, \alpha')/\alpha \rceil\) and \(S\) is a digital \((t', \alpha', \beta', \sigma, s)\)-sequence for all \(1 \leq \alpha' \leq m\) where \(\beta' = \beta \min(\alpha, \alpha')/\alpha\) and where \(t' = \lceil \min(\alpha, \alpha')/\alpha \rceil\).

(iii) Any digital \((t, \alpha, n \times m, s)\)-net is a digital \((\lceil \alpha \rceil, \alpha', n \times m, s)\)-net for all \(1 \leq \alpha' \leq \alpha\) and every digital \((t, \alpha, s)\)-sequence is a digital \((\lceil \alpha \rceil, \alpha', s)\)-sequence for all \(1 \leq \alpha' \leq \alpha\).

(iv) If \(C_1, \ldots, C_n \in \mathbb{Z}_p^{n \times m}\) are the generating matrices of a digital \((t, \alpha, \beta, n \times m, s)\)-net then the matrices \(C_1^{(n)}, \ldots, C_s^{(n)}\), where \(C_j^{(n)}\) consists of the first \(n'\) rows of \(C_j\), generate a digital \((t, \alpha, \beta, n' \times m, s)\)-net for all \(1 \leq n' \leq n\).

(v) Any digital \((t, \alpha, \beta, \sigma, s)\)-sequence is a digital \((t, \alpha, \beta, \sigma', s)\)-sequence for all \(1 \leq \sigma' \leq \sigma\).

We further consider a special case which is not contained in [5], namely where \(n = om\) with \(\alpha \geq 1\) (in [5] we assumed that \(n = m\)). Let \(C \in \mathbb{F}_q^{om \times m}\) with \(C = (c_{1,1}, \ldots, c_{om})^T\), where \(c_l\) denotes the \(l\)-th row of \(C\). We obtain \(\alpha\) matrices \(C_1, \ldots, C_\alpha \in \mathbb{F}_q^{m \times m}\) by defining \(C_j = (c_{j,1}, \ldots, c_{j,om})^T\). We have the following theorem.

**Theorem 6.** Let \(C_1, \ldots, C_\alpha \in \mathbb{F}_q^{m \times m}\) be the generating matrices of a digital \((t, \alpha, \beta, om \times m, s)\)-net. Let \(C_j = (c_{j,1}, \ldots, c_{j,om})^T\). Then the matrices

\[
C_{1,1}, \ldots, C_{1,\alpha}, \ldots, C_{s,1}, \ldots, C_{s,\alpha} \in \mathbb{F}_q^{m \times m}
\]

given by \(C_{j,i} = (c_{j,i+1}, \ldots, c_{j,i+\alpha(m-1)})^T\) are generating matrices of a digital \((t', 1, \beta, m \times m, \alpha s)\)-net with

\[
t' = \left\lfloor \frac{t}{\alpha} \right\rfloor.
\]
Proof. Let $d_{j,i} ≥ 0$ be such that $\sum_{j=1}^{s} \sum_{i=1}^{\alpha} d_{j,i} = βm - t'$. Then we need to show that the set of vectors
\[ \{ c_{j,i+α} : j = 1, \ldots, s; i = 1, \ldots, α; 0 ≤ r < d_{j,i} \} \]
is a linearly independent set.

Let $u_j = \{ i + αr : 1 ≤ i ≤ α, 0 ≤ r < d_{j,i} \}$ and let $a_{j,1}, \ldots, a_{j,α}$ denote the largest $α$ values in $u_j$. Note that the choice of $d'_{j,1}, \ldots, d'_{j,α}$ which maximizes the sum $\sum_{j=1}^{s} \sum_{i=1}^{α} a_{j,i}$ is obtained by $d'_{j,1} = \cdots = d'_{j,α-1} = 0$ for $j = 1, \ldots, s$ and by choosing $d'_{j,α}$ such that $\sum_{j=1}^{s} d'_{j,α} = βm - t'$.

Thus it follows that
\[ \sum_{j=1}^{s} \sum_{i=1}^{α} a_{j,i} ≤ \sum_{j=1}^{s} (αα(d'_{j,α} - 1)) = α \sum_{j=1}^{s} d'_{j,α} = α(βm - t'). \]
But as $t ≤ αt'$ it follows that $α(βm - t') ≤ αβm - t$. Thus
\[ \sum_{j=1}^{s} \sum_{i=1}^{α} a_{j,i} ≤ αβm - t' \]
from which follows that the vectors $c_{j,i+α} : j = 1, \ldots, s; i = 1, \ldots, α; 0 ≤ r < d_{j,i}$ are linearly independent. \( \square \)

The following result now follows directly from Theorem 6.

**Theorem 7.** Let $C_1, \ldots, C_s ∈ \mathbb{F}_q^{∞×∞}$ be the generating matrices of a digital $(t, α, β, α, s)$-sequence. Let $C_j = (c_{j,1}, c_{j,2}, \ldots)^T$. Then the matrices
\[ C_{1,1}, \ldots, C_{1,α}, \ldots, C_{s,1}, \ldots, C_{s,α} ∈ \mathbb{F}_q^{∞×∞} \]
given by $C_{j,i} = (c_{j,1}, c_{j,1+α}, c_{j,1+2α}, \ldots)^T$ generate a digital $(t', 1/β, 1, αs)$-sequence with
\[ t' = \left\lfloor \frac{t}{α}\right\rfloor. \]

**4.4. Constructions of $(t, α, β, n × m, s)$-nets and $(t, α, σ, s)$-sequences.** In this section we show how explicit examples of $(t, α, β, n × m, s)$-nets and $(t, α, β, σ, s)$-sequences can be constructed. The idea for the construction is based on the construction method presented in [5].

Let $d ≥ 1$ and let $C_{1}, \ldots, C_{sd}$ be the generating matrices of a digital $(t, m, sd)$-net. Note that many explicit examples of such generating matrices are known, see for example [13, 30, 38, 48] and the references therein. For the construction of a $(t, α, m, s)$-net any of the above mentioned explicit constructions can be used, but as will be shown below the quality of the $(t, α, m, s)$-net obtained depends on the quality of the underlying digital $(t, m, sd)$-net on which our construction is based on.

Let $C_j = (c_{j,1}, \ldots, c_{j,m})^T$ for $j = 1, \ldots, sd$, i.e., $c_{j,l}$ are the row vectors of $C_j$. Now let the matrix $C_j^{(d)}$ be made of the first rows of the matrices $C_{(j-1)d+1}, \ldots, C_{jd}$, then the second rows of $C_{(j-1)d+1}, \ldots, C_{jd}$ and so on. The matrix $C_j^{(d)}$ is then an $dm × m$ matrix, i.e., $C_j^{(d)} = (c_{j,v}, \ldots, c_{j,v})^T$ where $c_{j,l} = c_{u,v}$ with $l = (v-j)d+u$, $1 ≤ v ≤ m$ and $(j-1)d < u ≤ jd$ for $l = 1, \ldots, dm$ and $j = 1, \ldots, s$. The following result is a slight generalization of [5, Theorem 3] and can be obtained using the same proof technique.
Theorem 8. Let \( d \geq 1 \) be a natural number and let \( C_1, \ldots, C_{sd} \) be the generating matrices of a digital \((t', m, sd)\)-net over the finite field \( \mathbb{F}_q \) of prime power order \( q \). Let \( C_1^{(d)}, \ldots, C_{s}^{(d)} \) be defined as above. Then for any \( \alpha \geq 1 \) the matrices \( C_1^{(d)}, \ldots, C_{s}^{(d)} \) are generating matrices of a digital \((t, \alpha, \min(1, \alpha/d), dm \times m, s)\)-net over \( \mathbb{F}_q \) with

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

The above construction and Theorem 8 can easily be extended to \((t, \alpha, \beta, \sigma, s)\)-sequences. Indeed, let \( d \geq 1 \) and let \( C_1, \ldots, C_{sd} \) be the generating matrices of a digital \((t, sd)\)-sequence. Again many explicit generating matrices are known, see for example [13, 30, 38, 48]. Let \( C_j = (c_{j,1}, c_{j,2}, \ldots)^T \) for \( j = 1, \ldots, sd \), i.e., \( c_{j,t} \) are the row vectors of \( C_j \). Now let the matrix \( C_j^{(d)} \) be made of the first rows of the matrices \( C_{(j-1)d+1}, \ldots, C_{jd} \), then the second rows of \( C_{(j-1)d+1}, \ldots, C_{jd} \) and so on, i.e.,

\[
C_j^{(d)} = (c_{(j-1)d+1,1}, \ldots, c_{jd,1}, c_{(j-1)d+1,2}, \ldots, c_{jd,2}, \ldots)^T.
\]

The following theorem states that the matrices \( C_1^{(d)}, \ldots, C_{s}^{(d)} \) are the generating matrices of a digital \((t, \alpha, \min(1, \alpha/d), d, s)\)-sequence, compare with [5, Theorem 4].

Theorem 9. Let \( d \geq 1 \) be a natural number and let \( C_1, \ldots, C_{sd} \) be the generating matrices of a digital \((t', sd)\)-sequence over the finite field \( \mathbb{F}_q \) of prime power order \( q \). Let \( C_1^{(d)}, \ldots, C_{s}^{(d)} \) be defined as above. Then for any \( \alpha \geq 1 \) the matrices \( C_1^{(d)}, \ldots, C_{s}^{(d)} \) are generating matrices of a digital \((t, \alpha, \min(1, \alpha/d), d, s)\)-sequence over \( \mathbb{F}_q \) with

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

The last result shows that \((t, \alpha, \beta, \sigma, m, s)\)-nets indeed exist for \( \beta = 1 \) and any \( 0 < \sigma \leq \alpha \) and for \( m \) arbitrarily large. We have even shown that digital \((t, \alpha, \beta, \sigma, \alpha m \times m, s)\)-nets exist which are extensible in \( m \) and \( s \). This can be achieved by using an underlying \((t', sd)\)-sequence which is itself extensible in \( m \) and \( s \). If the \( t' \)-value of the original \((t', m, s)\)-net or \((t', s)\)-sequence is known explicitly then we also know the \( t \)-value of the digital \((t, \alpha, \beta, \sigma, \alpha m \times m, s)\)-net or \((t, \alpha, \beta, \sigma, s)\)-sequence. Furthermore it has also been shown how such digital nets can be constructed in practise. Further results on such sequences are established in [5].

Remark 10. Theorem 6 and Theorem 7 show that the \( t \)-value of a digital \((t, \alpha, \alpha m \times m, s)\)-net or digital \((t, \alpha, s)\)-sequence grows at least of order \( \alpha \). On the other hand, Theorem 8 and Theorem 9 provide explicit constructions for which the \( t \)-value grows at most with order \( \alpha^2 \) (by choosing \( d = \alpha \) in the constructions above). Hence the optimal dependence of \( t \) on \( \alpha \) is at this stage not known. The optimal dependence of \( t \) on the dimension \( s \) is \( t = O(s) \), as this is also true for classical digital \((t, s)\)-sequences, see [38].

By comparing Theorem 6 with Theorem 8 for digital nets and Theorem 7 with Theorem 9 for digital sequences it follows that the difference of the \( t \)-value between the best possible construction of a digital \((t, \alpha, \alpha m \times m, s)\)-net or digital \((t, \alpha, s)\)-sequence with the construction proposed here is at most \( s(\alpha - 1)\alpha/2 \) and hence our constructions are close to the best possible constructions.
Let $d_q(s, \alpha)$ be the smallest value of $t$ such that there exists a digital $(t, \alpha, s)$-sequence over $\mathbb{F}_q$. Then it follows that there are constants $c_1, c_2 > 0$ such that

$$c_1 s \alpha \leq d_q(\alpha, s) \leq c_2 s \alpha^2.$$ 

5. **Numerical integration in the Walsh space $E_{s,q,\vartheta,\gamma}$**

In this section we investigate numerical integration in the Walsh space $E_{s,q,\vartheta,\gamma}$ using quasi-Monte Carlo rules

$$Q_{q^m,s}(f) = \frac{1}{q^m} \sum_{n=0}^{q^m-1} f(x_n),$$

where $x_0, \ldots, x_{q^m-1}$ are the points of a digital $(t, \alpha, \beta, m, s)$-net over $\mathbb{F}_q$. More precisely, we want to approximate the integral

$$I_s(f) = \int_{[0,1]^s} f(x) \, dx$$

by the quasi-Monte Carlo rule $Q_{q^m,s}(f)$. As a quality measure for our rule we introduce the worst-case error in the next section.

5.1. **The worst-case error in the Walsh space $E_{s,q,\vartheta,\gamma}$**

The worst-case error for the Walsh space $E_{s,q,\vartheta,\gamma}$ using the quasi-Monte Carlo rule $Q_{q^m,s}$ is given by

$$e(Q_{q^m,s}, E_{s,q,\vartheta,\gamma}) = \sup_{f \in E_{s,q,\vartheta,\gamma}, \|f\|_{E_{s,q,\vartheta,\gamma}} \leq 1} |I_s(f) - Q_{q^m,s}(f)|.$$ 

The initial error is given by

$$e(Q_{0,s}, E_{s,q,\vartheta,\gamma}) = \sup_{f \in E_{s,q,\vartheta,\gamma}, \|f\|_{E_{s,q,\vartheta,\gamma}} \leq 1} |I_s(f)|.$$ 

In the following we use digital nets generated by the matrices $C_1, \ldots, C_s$ as quadrature points for the quadrature rule $Q_{q^m,s}$. Let $f \in E_{s,q,\vartheta,\gamma}$. Using Lemma 7 it follows that

$$|I_s(f) - Q_{q^m,s}(f)| = \left| \sum_{k \in D} \hat{f}(k) \right| \leq \sum_{k \in D} |\hat{f}(k)| = \sum_{\emptyset \neq u \subseteq S} \sum_{k_u \in D_u^*} |\hat{f}(k_u, 0_{S\setminus u})|.$$ 

Now we have $|\hat{f}(k_u, 0_{S\setminus u})| \leq \gamma_u r_q,\vartheta(\kappa_u) \|f\|_{E_{s,q,\vartheta,\gamma}}$ and thus we obtain

$$|I_s(f) - Q_{q^m,s}(f)| \leq \|f\|_{E_{s,q,\vartheta,\gamma}} \sum_{\emptyset \neq u \subseteq S} \gamma_u \sum_{k_u \in D_u^*} r_q,\vartheta(\kappa_u).$$ 

By choosing $\hat{f}(k_u, 0_{S\setminus u}) = \gamma_u r_q,\vartheta(\kappa_u)$ for all $u$ and $k_u$ we can also obtain equality in (12). Thus we have

$$e(Q_{q^m,s}, E_{s,q,\vartheta,\gamma}) = \sum_{\emptyset \neq u \subseteq S} \gamma_u \sum_{k_u \in D_u^*} r_q,\vartheta(\kappa_u).$$ 

From the last formula we can now see that essentially a large value of $\min \{\mu_q,\vartheta(k) : k \in D\}$ guarantees a small worst-case error. Further it can be shown that

$$e(Q_{0,s}, E_{s,q,\vartheta,\gamma}) = \gamma_q.$$
We have shown the following theorem.

**Theorem 11.** The initial error for multivariate integration in the Walsh space $E_{s,q,\vartheta,\gamma}$ is given by

$$e(Q_{0,s}, E_{s,q,\vartheta,\gamma}) = \gamma_0$$

and the worst-case error for multivariate integration in the Walsh space $E_{s,q,\vartheta,\gamma}$ using a digital net as quadrature points is given by

$$e(Q_{q^n,s}, E_{s,q,\vartheta,\gamma}) = \sum_{\emptyset \neq u \subseteq S} \sum_{\kappa_u \in D_u^\vartheta} r_{q,\theta}(k_u).$$

In the following lemma we establish an upper bound on the sum $\sum_{\kappa_u \in D_u^\vartheta} r_{q,\theta}(k_u)$ for digital $(t, \alpha, \beta, n \times m, s)$-nets over $F_q$. The proof is similar to [5, Lemma 6].

**Lemma 8.** Let $\vartheta > 1$ be a real number, $q \geq 2$ be a prime power, $C_1, \ldots, C_s \in F_q^{n \times m}$ be the generating matrices of a digital $(t, [\vartheta], \beta, n \times m, s)$-net over $F_q$ with $0 < \beta < 1$ and let $D_u^\vartheta = D_u^\vartheta((C_j)_{j \in u})$. For all $\emptyset \neq u \subseteq S$ we have: if $\vartheta$ is not an integer it follows that

$$\sum_{\kappa_u \in D_u^\vartheta} r_{q,\theta}(k_u) \leq C'_{|u|,q,\theta}(\beta n - t + [\vartheta])|u|^{|\vartheta|-1}q^{1-|\vartheta|}|u|(|\vartheta| - 1)$$

where

$$C'_{|u|,q,\theta} = q^{|\vartheta| - 1}(q - q^{1-|\vartheta|})^{1-(|\vartheta| - 1)}q^{-|\vartheta|}$$

and if $\vartheta$ is an integer it follows that

$$\sum_{\kappa_u \in D_u^\vartheta} r_{q,\theta}(k_u) \leq C'_{\beta n - t + \vartheta}|q|^{-\vartheta}q^{-\vartheta}$$

where

$$C'_{|u|,q,\theta} = q^{|\vartheta| - 1}(q - q^{1-|\vartheta|})^{1-(|\vartheta| - 1)}q^{-|\vartheta|}.$$

**Proof.** To simplify the notation we prove the result only for $u = S$. For all other subsets the result follows by the same arguments.

We first consider the case where $\vartheta > 1$ is not an integer. We partition the set $D_S^\vartheta$ into parts where the highest digits of $k_j$ are prescribed and we count the number of solutions of $C_1^T \kappa_1 + \cdots + C_s^T \kappa_s = 0$. For $j = 1, \ldots, s$ let now $i_{j,1} < \cdots < i_{j,1}$ with $i_{j,1} \geq 1$. Note that we now allow $i_{j,1} < 1$, in which case the contributions of those $i_{j,1}$ are to be ignored. This notation is adopted in order to avoid considering many special cases. Further we write $i_s([\vartheta]) = (i_{1,[\vartheta]}, \ldots, i_{s,[\vartheta]})$ and define

$$D_S^\vartheta(i_{s,[\vartheta]}) = \{ k \in D_S^\vartheta : k_j = [\kappa_{j,1}b_{i_{j,1}}^{i_{j,1}-1} + \cdots + \kappa_{j,1}b_{i_{j,1}}^{i_{j,1}-1} + l_j] \text{ with } 0 \leq l_j < b_{i_{j,1}}^{i_{j,1}-1} \text{ and } 1 \leq \kappa_{j,1} < b \text{ for } j = 1, \ldots, s \}$$

where $[\cdot]$ just means that the contributions of $i_{j,1} < 1$ are to be ignored. Let $\mu(i_s([\vartheta])) = i_{1,1} + \cdots + i_{1,[\vartheta]} - 1 + (\vartheta - [\vartheta]) i_{1,[\vartheta]} + \cdots + i_{s,1} + \cdots + i_{s,1} - 1 + (\vartheta - [\vartheta]) i_{s,1}.$

Then we have

$$\sum_{k \in D_S^\vartheta} r_{q,\theta}(k_S) = \sum_{i_{1,1}=1}^{\infty} \cdots \sum_{i_{s,1}=1}^{\infty} \frac{D_S^\vartheta(i_s([\vartheta]))}{q^{|\mu(i_s([\vartheta]))|}}.$$

Some of the sums above can be empty in which case we just set the corresponding summation index $i_{j,1} = 0$. 

WALSH SPACES AND QUASI-MONTE CARLO RULES 27
Note that by the \((t, \lceil \vartheta \rceil, \beta, n \times m, s)\)-net property we have that \(|D_{\beta}^{2}(i_{s,[\vartheta]}))| = 0\) as long as \(i_{1,1} + \cdots + i_{1,[\vartheta]} + \cdots + i_{s,1} \cdots + i_{s,[\vartheta]} \leq \beta n - t\). Hence let now \(0 \leq i_{1,1}, \ldots, i_{s,[\vartheta]} \leq 1\), \(i_{j,[\vartheta]} < \cdots < i_{j,1}\) for \(j = 1, \ldots, s\) and where if \(i_{j,1} < 1\) we set \(i_{j,1} = 0\) (in which case we also have \(i_{j,1} + i_{j,2} + \cdots = 0\) and the inequalities \(i_{j,1} > \cdots > i_{j,[\vartheta]}\) are ignored) and \(i_{1,1} + \cdots + i_{1,[\vartheta]} + \cdots + i_{s,1} + \cdots + i_{s,[\vartheta]} > \beta n - t\). We now need to estimate \(|D_{\beta}^{2}(i_{s,[\vartheta]}))|\), that is we need to count the number of \(k \in D_{\beta}^{2}\) with \(k_{j} = [\kappa_{j,1} b^{j,1-1} + \cdots + \kappa_{j,[\vartheta]} b^{j,[\vartheta]-1} + l_{j}]\) such that \(C_{j}^{T} k_{1} + \cdots + C_{j}^{T} k_{s} = 0\).

There are at most \((q - 1)^{s[\vartheta]}\) choices for \(\kappa_{1,1}, \ldots, \kappa_{s,[\vartheta]}\) (we write at most because if \(i_{j,1} < 1\) then the corresponding \(\kappa_{j,1}\) does not have any effect and therefore need not to be included).

Let now \(1 \leq \kappa_{1,1}, \ldots, \kappa_{s,[\vartheta]} < q\) be given and define

\[
\bar{g} = \kappa_{1,1} c_{1,1,1}^{T} + \cdots + \kappa_{1,[\vartheta]} c_{1,1,[\vartheta]}^{T} + \cdots + \kappa_{s,1} c_{s,1,1}^{T} + \cdots + \kappa_{s,[\vartheta]} c_{s,1,[\vartheta]}^{T},
\]

where we set \(c_{j,1}^{T} = 0\) if \(l < 1\) or \(l > n\). Further let

\[
B = (c_{1,1}^{T}, \ldots, c_{1,s,[\vartheta]-1}^{T}, \ldots, c_{s,1,1}^{T}, \ldots, c_{s,s,[\vartheta]-1}^{T}).
\]

Now the task is to count the number of solutions \(\bar{t}\) of \(B\bar{t} = \bar{g}\).

As long as the columns of \(B\) are linearly independent the number of solutions can at most be \(1\). By the \((t, \lceil \vartheta \rceil, \beta, n \times m, s)\)-net property this is certainly the case if we write \((x)_{+} = \max(x, 0)\)

\[
(i_{1,[\vartheta]} - 1)_{+} + \cdots + (i_{1,[\vartheta]} - \lceil \vartheta \rceil)_{+} + \cdots + (i_{s,[\vartheta]} - 1)_{+} + \cdots + (i_{s,[\vartheta]} - \lceil \vartheta \rceil)_{+} \leq \lceil \vartheta \rceil (i_{1,[\vartheta]} + \cdots + i_{s,[\vartheta]}) \leq \beta n - t,
\]

that is, as long as

\[
i_{1,[\vartheta]} + \cdots + i_{s,[\vartheta]} \leq \beta n - t - \frac{\beta n - t}{\lceil \vartheta \rceil}.
\]

Let now \(i_{1,[\vartheta]} + \cdots + i_{s,[\vartheta]} > \frac{\beta n - t}{\lceil \vartheta \rceil}\). Then by considering the rank of the matrix \(B\) and the dimension of the space of solutions of \(B\bar{t} = \bar{g}\) it follows the number of solutions of \(B\bar{t} = \bar{g}\) is smaller or equal to \(q^{i_{1,[\vartheta]} + \cdots + i_{s,[\vartheta]} - (\beta n - t)/\lceil \vartheta \rceil}\). Thus we have

\[
|D_{\beta}^{2}(i_{s,[\vartheta]}))| = 0
\]

if \(\sum_{j=1}^{s} \sum_{l=1}^{[\vartheta]} i_{j,l} \leq \beta n - t\), we have

\[
|D_{\beta}^{2}(i_{s,[\vartheta]}))| = (q - 1)^{s[\vartheta]}
\]

if \(\sum_{j=1}^{s} \sum_{l=1}^{[\vartheta]} i_{j,l} > \beta n - t\) and \(\sum_{j=1}^{s} i_{j,[\vartheta]} \leq \frac{\beta n - t}{\lceil \vartheta \rceil}\) and finally we have

\[
|D_{\beta}^{2}(i_{s,[\vartheta]}))| = (q - 1)^{s[\vartheta]} q^{i_{1,[\vartheta]} + \cdots + i_{s,[\vartheta]} - (\beta n - t)/\lceil \vartheta \rceil}.
\]

We estimate the sum (13) now. Let \(S_{1}\) be the sum in (13) where \(i_{1,1} + \cdots + i_{1,[\vartheta]} > \beta n - t\) and \(i_{1,1} + \cdots + i_{s,[\vartheta]} \leq \frac{\beta n - t}{\lceil \vartheta \rceil}\). Let \(l_{1} = i_{1,1} + \cdots + i_{1,[\vartheta]} - 1 + \cdots + i_{s,1} + \cdots + i_{s,[\vartheta]} - 1\) and let \(l_{2} = i_{1,1} + \cdots + i_{s,[\vartheta]}\). Let \(A(l_{1} + l_{2})\) denote the number of
admissible choices of \(i_{1,1}, \ldots, i_{s,[\vartheta]}\) such that \(l_1 + l_2 = i_{1,1} + \cdots + i_{s,[\vartheta]}\). Then we have

\[
S_1 = (q - 1)^{s[\vartheta]} \sum_{l_2=0}^{\lfloor \beta_n - 1 \rfloor} \frac{1}{q^{(s[\vartheta] - 1)l_2}} \sum_{l_1=\beta n + t + 1 - l_2}^{\infty} A(l_1 + l_2) q^{-l_2}.
\]

We have \(A(l_1 + l_2) \leq \left( \frac{l_1 + l_2 + s[\vartheta] - 1}{s[\vartheta] - 1} \right)\) and hence we obtain

\[
S_1 \leq (q - 1)^{s[\vartheta]} \sum_{l_2=0}^{\lfloor \beta_n - 1 \rfloor} \frac{1}{q^{(s[\vartheta] - 1)l_2}} \sum_{l_1=\beta n + t + 1 - l_2}^{\infty} \frac{1}{q^{l_1}} \left( \frac{l_1 + l_2 + s[\vartheta] - 1}{s[\vartheta] - 1} \right).
\]

From a result by Matoušek [26, Lemma 2.18], see also [7, Lemma 6], we have

\[
(q - 1)^{s[\vartheta]} \sum_{l_1=\beta n + t + 1 - l_2}^{\infty} \frac{1}{q^{l_1}} \left( \frac{l_1 + l_2 + s[\vartheta] - 1}{s[\vartheta] - 1} \right) \leq q^{l_2 - \beta n + t + s[\vartheta]} \left( \frac{\beta n - t + s[\vartheta]}{s[\vartheta] - 1} \right)
\]

and further we have

\[
\sum_{l_2=0}^{\lfloor \beta n - 1 \rfloor} q^{l_2} \frac{1}{q^{(s[\vartheta] - 1)l_2}} = \sum_{l_2=0}^{\lfloor \beta n - 1 \rfloor} q^{l_2(s[\vartheta] - 1)} = \frac{q^{s[\vartheta] - 1} - 1}{q^{s[\vartheta] - 1} - 1} \left( \frac{\beta n - t + s[\vartheta]}{s[\vartheta] - 1} \right).
\]

Thus we obtain

\[
S_1 \leq \frac{q^{(s[\vartheta] - 1)} - 1}{q^{s[\vartheta] - 1} - 1} \left( \frac{\beta n - t + s[\vartheta]}{s[\vartheta] - 1} \right).n
\]

Let \(S_2\) be the part of (13) for which \(i_{1,1} + \cdots + i_{s,[\vartheta]} > \beta n - t\) and \(i_{1,[\vartheta]} + \cdots + i_{s,[\vartheta]} > \frac{\beta n - t}{s[\vartheta]}\), i.e., we have

\[
S_2 \leq (q - 1)^{s[\vartheta]} \sum_{i_{1,1}=1}^{\infty} \cdots \sum_{i_{s,[\vartheta]}=1}^{\infty} \cdots \sum_{i_{s,[\vartheta]}=1}^{\infty} q^{-(\beta n - t)/\vartheta} q^{(i_{1,[\vartheta]} + \cdots + i_{s,[\vartheta]})(\vartheta - 1)}.
\]

where we have the additional conditions \(i_{1,1} + \cdots + i_{s,[\vartheta]} > \beta n - t\) and \(i_{1,[\vartheta]} + \cdots + i_{s,[\vartheta]} > \frac{\beta n - t}{s[\vartheta]}\). As above let \(l_1 = i_{1,1} + \cdots + i_{1,[\vartheta]} - 1 + \cdots + i_{s,1} + \cdots + i_{s,[\vartheta]} - 1\) and let \(l_2 = i_{1,[\vartheta]} + \cdots + i_{s,[\vartheta]}\). Let \(A(l_1 + l_2)\) denote the number of admissible choices of \(i_{1,1}, i_{s,[\vartheta]}\) such that \(l_1 + l_2 = i_{1,1} + \cdots + i_{s,[\vartheta]}\). Note that \(l_1 > \lfloor \vartheta \rfloor l_2\). Then
we have $A(l_1 + l_2) \leq \left(\frac{l_1 + l_2 + s[\vartheta]}{s[\vartheta]} - 1\right)$ and hence we obtain
\[
S_2 \leq (q - 1)^{s[\vartheta]} q^{-(\beta n - \ell)/[\vartheta]} \sum_{l_2 = \left\lfloor \frac{\beta n - \ell}{[\vartheta]} \right\rfloor + 1}^{\infty} \sum_{l_1 = \lfloor \vartheta \rfloor l_2 + 1}^{\infty} \frac{1}{q^{l_2}} \left(\frac{l_1 + l_2 + s[\vartheta] - 1}{s[\vartheta]} - 1\right)
\]
\[
= (q - 1)^{s[\vartheta]} q^{-(\beta n - \ell)/[\vartheta]} \sum_{l_2 = \left\lfloor \frac{\beta n - \ell}{[\vartheta]} \right\rfloor + 1}^{\infty} \sum_{l_1 = 0}^{\infty} q^{-t_1 + t_2 - 1 - t_2 \vartheta} \left(\frac{l_1 + l_2 + \lfloor \vartheta \rfloor l_2 + 1 + s[\vartheta] - 1}{s[\vartheta]} - 1\right).
\]

By using again Matoušek [26, Lemma 2.18], see also [7, Lemma 6], we have
\[
(q - 1)^{s[\vartheta]} \sum_{l_1 = 0}^{\infty} q^{-t_1 + t_2 - 1 - t_2 \vartheta} \left(\frac{l_1 + l_2 + \lfloor \vartheta \rfloor l_2 + 1 + s[\vartheta] - 1}{s[\vartheta]} - 1\right)
\]
\[
\leq q^{s[\vartheta]} q^{t_2(1-\vartheta)} \left(\frac{l_2}{s[\vartheta]} - 1 + s[\vartheta] - 1\right)
\]
and also
\[
q^{s[\vartheta]} q^{t_2(1-\vartheta)} \left(\sum_{l_2 = \left\lceil \frac{\beta n - \ell}{[\vartheta]} \right\rceil + 1}^{\infty} q^{l_2(1-\vartheta)} \left(\frac{l_2}{s[\vartheta]} - 1 + s[\vartheta] - 1\right)\right)
\]
\[
\leq q^{s[\vartheta]} (1 - q^{(1-\vartheta)/[\vartheta]} - s[\vartheta] \left(\beta n - t + \lfloor \vartheta \rfloor - 2 + s[\vartheta]\right)) q^{-\vartheta(\beta n - \ell)/[\vartheta]}.
\]

Hence we have
\[
S_2 \leq q^{s[\vartheta]} (1 - q^{(1-\vartheta)/[\vartheta]} - s[\vartheta] \left(\beta n - t + \lfloor \vartheta \rfloor - 2 + s[\vartheta]\right)) q^{-\vartheta(\beta n - \ell)/[\vartheta]}.
\]

Note that we have $\sum_{k \in \mathcal{D}_S} r_{q, \vartheta}(k_S) = S_1 + S_2$. Let $a \geq 1$ and $b \geq 0$ be integers, then we have
\[
\left(\frac{a + b}{b}\right) = \prod_{i=1}^{b} \left(1 + \frac{a}{b}\right) \leq (1 + a)^b.
\]

Therefore we obtain
\[
S_1 \leq \frac{q^{s[\vartheta]} q^{s[\vartheta] - 1}}{1 - q^{\vartheta - [\vartheta]} (\beta n - t + 2)^{s[\vartheta] - 1}} q^{-\vartheta(\beta n - \ell)/[\vartheta]}.
\]

and
\[
S_2 \leq q^{s[\vartheta]} \left(1 - q^{(1-\vartheta)/[\vartheta]} - s[\vartheta] \left(\beta n - t + \lfloor \vartheta \rfloor\right)^{s[\vartheta] - 1} q^{-\vartheta(\beta n - \ell)/[\vartheta]}\right).
\]

Thus we have
\[
\sum_{k \in \mathcal{D}_S} r_{q, \vartheta}(k_S) \leq C_{q, s, \vartheta} (\beta n - t + \lfloor \vartheta \rfloor)^{s[\vartheta] - 1} q^{-\vartheta(\beta n - \ell)/[\vartheta]},
\]

where
\[
C_{q, s, \vartheta} = q^{s[\vartheta]} \left((q - q^{\vartheta - [\vartheta]} - 1) + (1 - q^{(1-\vartheta)/[\vartheta]} - s[\vartheta])\right).
\]

The result follows for the case $0 < \vartheta - \lfloor \vartheta \rfloor < 1$. 
Let now \( \vartheta > 1 \) be an integer. Then using the same arguments as above it can be shown that

\[
S_1 \leq (\beta n - t + 2)^{s\vartheta} q^{-(\beta n - t) - 1+s\vartheta}
\]

and

\[
S_2 \leq q^{\vartheta}(1 - q^{1/\vartheta-1}) - s\vartheta (\beta n - t + \vartheta)^{\vartheta-1} q^{-(\beta n - t)}.
\]

Thus we have

\[
\sum_{\mathbf{k}_S \in D^*_S} r_{q,\vartheta}(\mathbf{k}_S) \leq C_{s,q,\vartheta}'(\beta n - t + \vartheta)^{\vartheta} q^{-(\beta n - t)},
\]

where

\[
C_{s,q,\vartheta}' = q^{s\vartheta}(q^{-1} + (1 - q^{1/\vartheta-1})^{-s\vartheta}).
\]

The result now follows. \( \square \)

Remark 12. We note that the above lemma does not hold for \( \beta > 1 \) in general. Indeed, take for example \( u = \{1\} \), then \( \mathbf{k}_u = (k_1) \) and choose \( k_1 = q^n \). Then the digit vector of the first \( n \) digits of \( q^n \) is \( (0, \ldots, 0) \) and hence \( C_{1}^{\top} k_1 = 0 \) and hence \( \mathbf{k}_1(1) \in D^*_1 \). Thus

\[
\sum_{\mathbf{k}_1(1) \in D^*_1} r_{q,\vartheta}(\mathbf{k}_1(1)) \geq q^{-n-1}
\]

and hence a counterexample can be obtained for some choices of \( n, \beta, \vartheta \).

In [5] we did allow \( \beta > 1 \), but therein we had the additional assumption that the functions are periodic. In this case we were able to show that the Walsh coefficients \( r_{q,\vartheta}(\mathbf{k}, \mathbf{l}) = \prod_{j=1}^{n} r_{q,\vartheta}(k_j, l_j) \) of the reproducing kernel also satisfy the additional property that \( r_{q,\vartheta}(q^m k_j, q^m k_j) = r_{q,\vartheta}(q^m k_j) = q^{-2\alpha m} r_{q,\vartheta}(k_j, k_j) \) for all \( k_j, m \in \mathbb{N} \), see [5, Lemma 15]. Similarly, if we would also assume here that \( r_{q,\vartheta}(q^n k) = q^{-\vartheta n} r_{q,\vartheta}(k) \) and \( r_{q,\vartheta}(k) \) given as above if \( q \nmid k \), then the above counterexample would fail as then \( r_{q,\vartheta}(q^n) = r_{q,\vartheta}(1q^n) = q^{-\vartheta(n+1)} r_{q,\vartheta}(1) \).

Using the above lemma we can now obtain an upper bound on the worst-case error.

Theorem 13. Let \( \vartheta > 1 \) be a real number and \( q \geq 2 \) be a prime power. The worst-case error for multivariate integration in the Walsh space \( \mathcal{E}_{s,q,\vartheta,\gamma} \) using a digital \((t, |\vartheta|, \beta, n \times m, s)\)-net over \( \mathbb{F}_q \) with \( 0 < \beta \leq 1 \) as quadrature points is for non-integers \( \vartheta \) bounded by

\[
e(Q_{q^{m,s},\mathcal{E}_{s,q,\vartheta,\gamma}}) \leq q^{-\vartheta \lfloor (\beta n - t)/|\vartheta| \rfloor} \sum_{\emptyset \neq \mathbf{u} \subseteq \mathcal{S}} \gamma_{\mathbf{u}} C_{q,|\mathbf{u}|,\vartheta}(\beta n - t + |\vartheta|)^{|\mathbf{u}| - |\vartheta|},
\]

where

\[
C_{q,|\mathbf{u}|,\vartheta} = q^{|\mathbf{u}| |\vartheta|} ((q - q^{-|\vartheta|})^{-1} + (1 - q^{1-|\vartheta|/|\vartheta|})^{-|\mathbf{u}| |\vartheta|}),
\]

and if \( \vartheta \) is an integer, the square worst-case error is bounded by

\[
e(Q_{q^{m,s},\mathcal{E}_{s,q,\vartheta,\gamma}}) \leq q^{-\beta n - t} \sum_{\emptyset \neq \mathbf{u} \subseteq \mathcal{S}} \gamma_{\mathbf{u}} C'_{q,|\mathbf{u}|,\vartheta}(\beta n - t + |\vartheta|)^{|\mathbf{u}| |\vartheta|},
\]

where

\[
C'_{q,|\mathbf{u}|,\vartheta} = q^{|\mathbf{u}| |\vartheta|} (q^{-1} + (1 - q^{1/\vartheta-1})^{-|\mathbf{u}| |\vartheta|}).
\]

As a direct consequence of Corollary 2 we obtain the following result.
Corollary 3. Let $\delta \geq 1$ be an integer, $0 < \lambda < 1$ and $q \geq 2$ be a prime power. Then for any function $f : [0, 1)^s \to \mathbb{R}$ whose partial mixed derivatives up to order $\delta$ exist and are continuous it follows that the integration error using a digital $(t, \delta + 1, \beta, n \times m, s)$-net over $\mathbb{F}_q$ with $0 < \beta \leq 1$ as quadrature points is for $0 < \lambda < 1$ bounded by

$$|I_s(f) - Q_{\delta, q, s}(f)| \leq q^{-(\delta + \lambda)((\beta n - t)/((\delta + 1))}C_{\delta, q, s, \lambda, p, q, r, \gamma}(f) \sum_{\emptyset \neq u \subseteq S} \gamma u C_{u, \delta, \gamma}(\beta n - t + \delta + 1)^{|u|((\delta + 1) - 1)}$$

and for $\lambda = 1$ the integration error is bounded by

$$|I_s(f) - Q_{\delta, q, s}(f)| \leq q^{-(\beta n - t)}C_{\delta, q, s, \lambda, p, q, r, \gamma}(f) \sum_{\emptyset \neq u \subseteq S} \gamma u C_{u, \delta, q, s + 1}(\beta n - t + \delta + 1)^{|u|((\delta + 1)}$$

where the constant $C_{\delta, q, s, \lambda}$ is given in Corollary 2 and the constants $C_{u, \delta, \lambda}$ and $C_{s, \delta, q, s + 1}$ are given in Theorem 13.

Explicit constructions of digital $(t, \alpha, \min(1, \alpha/d), dm \times m, s)$-nets over $\mathbb{F}_q$ for all prime powers $q$, integers $\alpha, d, m, s > 1$ are given in Section 4.4. By choosing $d = \alpha = \lfloor \delta \rfloor = \delta + 1$, by Theorem 13 and Corollary 3 we obtain a convergence of $O(q^{-\beta n m^s|\delta| + 1})$, which is optimal even for the smooth functions contained in the Walsh space $E_{s, q, \delta, \gamma}$, see [45] where a lower bound for smooth periodic functions was shown.

Remark 14. In [5, Remark 4] it was noted that if $m = n$ and $\beta > \alpha$ the $t$-value must grow with $m$ and hence the restriction $\beta \leq \alpha$ was added. A similar argument yields in our case that the $t$-value must grow with $n$ if $\beta n > \alpha m$ as Theorem 13 shows a convergence of $O(q^{-\beta n + 1})$ but the best possible convergence rate is $q^{-\alpha m}$, hence the restriction $\beta \leq \alpha m/n$ was added.

In case the smoothness of the function is not known our constructions adjust themselves automatically up to a certain degree in the following way: for the construction of the digital net we choose some value of $d \geq 1$ and construct a digital $(t, \alpha, \min(1, \alpha/d), dm \times m, s)$-net or a digital $(t, \alpha, \min(1, \alpha/d), d, s)$-sequence for all $\alpha \geq 1$. The values $\delta \geq 1$ and $0 < \lambda < 1$ determine the real smoothness of the function, which we now assume is not known. The value of $\alpha$ is the smoothness analog for the digital net, i.e., we need to choose $\alpha = \delta + 1$. First assume that $\delta + \lambda \leq d$, then $\min(1, \alpha/d) = (\delta + 1)/d$ and therefore we have $\beta = (\delta + 1)/d$. As $n = dm$ it follows that $\beta n = (\delta + 1)m$ and therefore Corollary 3 shows that we achieve a convergence of $O(q^{-(\delta + \lambda)m s^{(\delta + 1) + 1}})$, which is optimal. Now assume on the other hand that $\delta + \lambda > d$, then $\min(1, \alpha/d) = 1$ and therefore $\beta = 1$. Again we have $n = dm$ and hence $\beta n = dm$. In this case Corollary 3 shows that our construction achieves a convergence of $O(q^{-(\delta + \lambda)m s^{(\delta + 1) + 1}})$.

Note that numerical integration of functions with less smoothness, i.e., for example functions with partial mixed derivatives up to degree 1 in $L_2([0, 1)^t)$ or functions with bounded variation, has been considered in many papers and monographs, see for example [6, 7, 9, 10, 23, 30, 47, 48]. Using the notation from above, basically those results are concerned with the case where $\delta = 0$ and $\lambda = 1$, hence the results here are a direct continuation of what was previously known. The construction of
digital nets proposed here for \( d = 1 \) yields obviously digital \((t, m, s)\)-nets and \((t, s)\)-sequences as for example defined in [30]. In view of Corollary 3 and the explanation which followed it is hence not surprising that the classical examples and theory (see for example [6, 7, 17, 20, 23, 30, 42, 48]) only yielded a convergence of \( O(q^{m(-1+\varepsilon)}) \) for any \( \varepsilon > 0 \) (the \( \varepsilon \) here is used to hide the powers of \( m \)).

Note that the worst-case error in the Walsh space \( E_{s,q,\vartheta,\gamma} \) is invariant with respect to a digital shift (see [7]), hence Corollary 3 also holds for digitally shifted digital nets. The root mean square worst-case error for this case would of course be bounded by the bound in Corollary 3, as this bound holds for any digital shift, i.e., our result here is even stronger in that we have shown that even for the worst digital shift we still have the bound of Corollary 3. From this it follows that for our situation here there is, in some sense, no bad digital shift. Other more sophisticated scrambling methods which do not destroy the essential properties of the point set can be used as well (for example a digital shift of depth \( m \), see [8, 26]), see [40] for some ideas in this direction.


In this section we basically give an alternative proof of Corollary 3 for a slightly different Sobolev space. The main part of the analysis is completely different and is based on the result in [5].

We consider the Sobolev space \( \mathcal{H}_{\text{sob},s,\alpha,\gamma} \) for which \( s \geq 1 \) and \( \alpha \geq 2 \). For the 1 dimensional unweighted case (i.e. the weights are chosen to be 1) the inner product is given by

\[
(f, g)_{\mathcal{H}_{\text{sob},1,\alpha,1}} = \sum_{\tau=0}^{\alpha-1} \int_0^1 f^{(\tau)}(x) \, dx \int_0^1 g^{(\tau)}(x) \, dx + \int_0^1 f^{(\alpha)}(x) g^{(\alpha)}(x) \, dx,
\]

where \( f^{(\tau)} \) denotes the \( \tau \)-th derivative of \( f \) and where \( f^{(0)} = f \). Note the similarities with the inner product defined in (9). Indeed, the results in this section also hold for the Sobolev space with inner product given by (9). The reproducing kernel (see [1] for more information about reproducing kernels) for this space is given by

\[
K_{\mathcal{H}_{\text{sob},1,\alpha,1}}(x,y) = \sum_{\tau=0}^{\alpha} \frac{B_{\tau}(x)B_{\tau}(y)}{(\tau!)^2} + \frac{B_{2\alpha}(|x-y|)}{(2\alpha)!},
\]

where \( B_{\tau} \) denotes the Bernoulli polynomial of degree \( \tau \). For example we have \( B_0(x) = 1, B_1(x) = x - 1/2, B_2(x) = x^2 - x + 1/6 \) and so on.

The reproducing kernel for the \( s \) dimensional weighted Sobolev space \( \mathcal{H}_{\text{sob},s,\alpha,\gamma} \) is now given by

\[
K_{\mathcal{H}_{\text{sob},s,\alpha,\gamma}}(x,y) = \sum_{u \in S} \gamma_u \prod_{j \in u} \left( \sum_{\tau=1}^{\alpha} \frac{B_{\tau}(x_j)B_{\tau}(y_j)}{(\tau!)^2} + (-1)^{\alpha+1} \frac{B_{2\alpha}(|x_j - y_j|)}{(2\alpha)!} \right).
\]

For example if \( \gamma_u = \prod_{j \in C_u} \gamma_j \) the space \( \mathcal{H}_{\text{sob},s,\alpha,\gamma} \) is a tensor product space of weighted one dimensional spaces.

Note that numerical integration in the Sobolev space with \( \alpha = 1 \) has been considered in many papers, see for example [6, 7, 8] and the references therein.

As opposed to the previous section, here we restrict ourselves to \( q \) prime and \( \varphi \) the identity map. This is necessary as the results here are based on [5] where
the same assumptions where made, though it should be possible to generalize the results in [5].

As $\mathcal{K}_{sob,s,\alpha,\gamma} \in \mathcal{L}_2([0,1)^{2s})$ it follows that $\mathcal{K}_{sob,s,\alpha,\gamma}$ can be represented by a Walsh series, i.e., we have

$$\mathcal{K}_{sob,s,\alpha,\gamma}(x,y) = \sum_{k,l \in \mathbb{N}_0^s} \hat{\mathcal{K}}_{sob,s,\alpha,\gamma}(k,l) \text{wal}_k(x) \text{wal}_l(y),$$

where

$$(15) \quad \hat{\mathcal{K}}_{sob,s,\alpha,\gamma}(k,l) = \int_{[0,1)^{2s}} \mathcal{K}_{sob,s,\alpha,\gamma}(x,y) \text{wal}_k(x) \text{wal}_l(y) \, dx \, dy.$$ 

Note that if $\gamma_u = 0$ for some $u$, then $\hat{\mathcal{K}}_{sob,s,\alpha,\gamma}(k,l) = 0$ if $k_j = l_j = 0$ for $j \notin u$ and $k_j, l_j \neq 0$ for $j \in u$. Let

$$\theta_{q,\alpha}(k,l) = \int_0^1 B_{2\alpha}(|x-y|) \text{wal}_k(x) \text{wal}_l(y) \, dx \, dy.$$ 

Then it follows from [5, Lemma 2 and Lemma 4] that there exists a constant $C'_{q,\alpha} > 0$ independent of $k, l$ such that

$$(16) \quad |\hat{\mathcal{K}}_{sob,1,\alpha,1}(k,l)| \leq C'_{q,\alpha} r_{q,\alpha}(k) r_{q,\alpha}(l),$$

where $r_{q,\alpha}$ is as in Section 3.3. It follows from Lemma 6 that a similar bound holds for the polynomials $B_\tau$. Thus there is a constant $C_{q,\alpha} > 0$, depending only on $q$ and $\alpha$, such that

$$(16) \quad |\hat{\mathcal{K}}_{sob,1,\alpha,1}(k,l)| \leq C_{q,\alpha} r_{q,\alpha}(k) r_{q,\alpha}(l) \quad \text{for all } k, l \in \mathbb{N}.$$ 

In the following we consider the worst-case error for multivariate integration in the Sobolev space $\mathcal{H}_{sob,s,\alpha,\gamma}$ for $s \geq 1$ and $\alpha \geq 2$, i.e.,

$$e(Q_{q,m,s},\mathcal{H}_{sob,s,\alpha,\gamma}) = \sup_{f \in \mathcal{H}_{sob,s,\alpha,\gamma}, \|f\|_{\mathcal{H}_{sob,s,\alpha,\gamma}} \leq 1} |I_s(f) - Q_{q,m,s}(f)|.$$ 

The initial error is given by

$$e(Q_{0,s},\mathcal{H}_{sob,s,\alpha,\gamma}) = \sup_{f \in \mathcal{H}_{sob,s,\alpha,\gamma}, \|f\|_{\mathcal{H}_{sob,s,\alpha,\gamma}} \leq 1} |I_s(f)|.$$ 

Using the approach from [5, Section 2.3] it follows that

$$e^2(Q_{0,s},\mathcal{H}_{sob,s,\alpha,\gamma}) = \gamma_0$$

and

$$e^2(Q_{q,m,s},\mathcal{H}_{sob,s,\alpha,\gamma}) = -\gamma_0 + \frac{1}{q^{2m}} \sum_{n,n' = 0}^{q^{m-1}} \mathcal{K}_{sob,s,\alpha,\gamma}(x_n, x_{n'})$$

$$= \frac{1}{q^{2m}} \sum_{n,n' = 0}^{q^{m-1}} (\mathcal{K}_{sob,s,\alpha,\gamma}(x_n, x_{n'}) - \gamma_0).$$

The following result can be obtained by the same arguments as [5, Theorem 1].
Theorem 15. The square initial error for multivariate integration in the Sobolev space \( H_{\text{sob}, s, \alpha, \gamma} \) is given by
\[
e^2(Q_0, s, H_{\text{sob}, s, \alpha, \gamma}) = \gamma_0
\]
and the square worst-case error for multivariate integration in the Sobolev space \( H_{\text{sob}, s, \alpha, \gamma} \) using a digital net as quadrature points is given by
\[
e^2(Q_{m, s}, H_{\text{sob}, s, \alpha, \gamma}) = \sum_{\emptyset \neq u \subseteq \{1, \ldots, s\}} \gamma_u \sum_{k_u, l_u \in D_u^*} \hat{K}_{\text{sob}, s, \alpha, \gamma}((k_u, 0), (l_u, 0)).
\]

Using (14) and (15) and the fact that \( j_0 \int_0^1 \sum_{\tau=1}^\alpha B_\tau(x)B_\tau(y) + B_{2\alpha}(|x-y|) \, dx \, dy = 0 \) we have
\[
\hat{K}_{\text{sob}, s, \alpha, \gamma}((k_u, 0), (l_u, 0)) = \gamma_u \prod_{j \in u} \tilde{K}_\alpha(k_j, l_j),
\]
where for \( k, l \in \mathbb{N} \) we set
\[
\tilde{K}_\alpha(k, l) = \int_0^1 \int_0^1 \left( \sum_{\tau=1}^\alpha B_\tau(x)B_\tau(y) \right) \frac{B_{2\alpha}(|x-y|)}{(2\alpha)!} \, dx \, dy.
\]

The following result is comparable to [5, Lemma 3] and follows from Theorem 15 together with (16) and (17).

Lemma 9. The square worst-case error for multivariate integration in the Sobolev space \( H_{\text{sob}, s, \alpha, \gamma} \) using a digital net as quadrature points can be bounded by
\[
e^2(Q_{m, s}, H_{\text{sob}, s, \alpha, \gamma}) \leq \sum_{\emptyset \neq u \subseteq \{1, \ldots, s\}} \gamma_u C_{u, q, \alpha}^{|u|} \left( \sum_{k_u, l_u \in D_u^*} r_q(k_u) \right)^2.
\]

Using Lemma 8 we can now obtain the following theorem.

Theorem 16. Let \( \alpha > 1 \) be an integer, \( q \geq 2 \) be a prime number, \( C_1, \ldots, C_s \in \mathbb{Z}_q^{n \times m} \) be the generating matrices of a digital \((t, \alpha, \beta, n \times m, s)\)-net over \( \mathbb{Z}_q \) with \( 0 < \beta \leq 1 \). The worst-case error for multivariate integration in the Sobolev space \( H_{\text{sob}, s, \alpha, \gamma} \) using a digital \((t, \alpha, \beta, n \times m, s)\)-net over \( \mathbb{Z}_q \) with \( 0 < \beta \leq 1 \) as quadrature points is bounded by
\[
e(Q_{m, s}, H_{\text{sob}, s, \alpha, \gamma}) \leq q^{-(\beta n - t)} \left( \sum_{\emptyset \neq u \subseteq \{1, \ldots, s\}} \gamma_u C_{|u|, q, \alpha}^{|u|} (\beta n - t + \alpha)^2 |u|^{\alpha} \right)^{1/2},
\]
where
\[
C_{|u|, q, \alpha}^{|u|} = C_{q, \alpha}^{\frac{|u|}{2}} q^{|u|\alpha} (q^{-1} + (1 - q^{1/\alpha - 1})^{-|u|\alpha}).
\]

Explicit constructions of digital \((t, \alpha, \min(1, \alpha/d), m \times m, s)\)-nets over \( \mathbb{Z}_q \) for all prime numbers \( q \) and integers \( \alpha, d, m, s > 1 \) are given in Section 4.4. By choosing \( d = \alpha \), by Theorem 16 we obtain a convergence of \( O(q^{-\alpha m^{\alpha/2}}) \), which is optimal for the Sobolev space \( H_{\text{sob}, s, \alpha, \gamma} \) as it is also optimal for the subspace of periodic functions, see [5, 45].

Theorem 16 is comparable to [5, Theorem 5], but with the important difference that here we do not assume that the integrand is periodic.
Corollary 4. Let $\alpha > 1$ be an integer, $q \geq 2$ be a prime number and $S$ be a digital $(t(a), a, \min(1, a/d), d, s)$-sequence over $\mathbb{Z}_q$ for any integer $a \geq 1$. The worst-case error for multivariate integration in the Sobolev space $H_{\text{sob}, s, \alpha, \gamma}$ using the first $N = q^m$ points of $S$ as quadrature points is bounded by

$$e(Q_q^m, s, H_{\text{sob}, s, \alpha, \gamma}) \leq N^{-\min(\alpha, d)} \left( \sum_{\emptyset \neq u \subseteq \{1, \ldots, s\}} \gamma_u C_{|u|, q, \alpha}^{(1)} (\log N)^2 |u|^{\alpha} \right)^{1/2},$$

where $C_{|u|, q, \alpha}^{(1)} > 0$ is a constant depending only on $|u|, q, \alpha$.

Note that by the constructions in Section 4.4 explicit examples of such digital $(t(a), a, \min(1, a/d), d, s)$-sequences over $\mathbb{Z}_q$ for any prime $q$, and $a, d, s \geq 1$ are given.

Remark 17. As in [5, Section 6] we can also obtain results for digital nets which are randomized by a random digital shift. The bounds on the mean square worst-case error are of the same form as [5, Theorem 8] and [5, Corollary], but now without the assumption that the function have to be periodic.

REFERENCES


[32] H. Niederreiter, Constructions of \((t, m, s)\)-nets and \((t, s)\)-sequences, Finite Fields Appl., 11 (2005), 578–600.


School of Mathematics, University of New South Wales, Sydney 2052, Australia

E-mail address: josifmaths.unsw.edu.au