HIROSHIMA MATH. J.
47 (2017), 155–179

Bounds on Walsh coefficients by dyadic difference and a new Koksma-Hlawka type inequality for Quasi-Monte Carlo integration

Takehito YOSHIKI

(Received November 8, 2016)
(Revised December 7, 2016)

ABSTRACT. In this paper we give a new Koksma-Hlawka type inequality for Quasi-Monte Carlo (QMC) integration. QMC integration of a function $f : [0,1)^s \to \mathbb{R}$ by a finite point set $\mathscr{P} \subset [0,1)^s$ is the approximation of the integral $I(f) := \int_{[0,1)^s} f(\mathbf{x}) d\mathbf{x}$ by the average $I_{\mathscr{P}}(f) := \frac{1}{|\mathscr{P}|} \sum_{\mathbf{x} \in \mathscr{P}} f(\mathbf{x})$. We treat a certain class of point sets $\mathscr{P}$ called digital nets. A Koksma-Hlawka type inequality is an inequality providing an upper bound on the integration error $\mathrm{Err}(f;\mathscr{P}) := I(f) - I_{\mathscr{P}}(f)$ of the form $|\mathrm{Err}(f;\mathscr{P})| \leq C \cdot \|f\| \cdot D(\mathscr{P})$. We can obtain a Koksma-Hlawka type inequality by estimating bounds on $|\hat{f}(\mathbf{k})|$, where $\hat{f}(\mathbf{k})$ is a generalized Fourier coefficient with respect to the Walsh system. In this paper we prove bounds on the Walsh coefficients $\hat{f}(\mathbf{k})$ by introducing an operator called 'dyadic difference' $\partial_{i,n}$. By converting dyadic differences $\partial_{i,n}$ to derivatives $\frac{\partial}{\partial x_i}$, we get a new bound on $|\hat{f}(\mathbf{k})|$ for a function f whose mixed partial derivatives up to order α in each variable are continuous. This new bound is smaller than the known bound on $|\hat{f}(\mathbf{k})|$ under some instances. The new Koksma-Hlawka type inequality is derived using this new bound on the Walsh coefficients.

1. Introduction and the main results

Quasi-Monte Carlo (QMC) integration of a function $f : [0,1)^s \to \mathbb{R}$ by a finite point set $\mathscr{P} \subset [0,1)^s$ is the approximation of the integral $I(f) := \int_{[0,1)^s} f(\mathbf{x}) d\mathbf{x}$ by the average $I_{\mathscr{P}}(f) := \frac{1}{|\mathscr{P}|} \sum_{\mathbf{x} \in \mathscr{P}} f(\mathbf{x})$ (see [10, 20, 24] for details). We want to find a quadrature point set $\mathscr{P}$ making the absolute value of the integration error $|\mathrm{Err}(f;\mathscr{P})| := |I(f) - I_{\mathscr{P}}(f)|$ small for a set of functions f. This problem is formulated as follows: We consider a function space H with norm $\|f\|_H$ and the worst case error $\sup_{\|f\|_H \leq 1} |\mathrm{Err}(f;\mathscr{P})|$ by a QMC rule using a point set $\mathscr{P}$ (for example, see [10, 17] for details). Then,

The work was supported by the Program for Leading Graduate Schools, MEXT, Japan and Australian Research Council Discovery Projects funding scheme (project number DP150101770).
2010 *Mathematics Subject Classification*. 65D30, 65D32.
Key words and phrases. Quasi-Monte Carlo integration, Walsh coefficients, Koksma-Hlaka inequality.

it holds that, for any $f \in H$,

$$|\mathrm{Err}(f;\mathscr{P})| \leq \|f\|_H \times \sup_{\|f\|_H \leq 1} |\mathrm{Err}(f;\mathscr{P})|. \tag{1}$$

Thus in order to make the integration error $|\mathrm{Err}(f;\mathscr{P})|$ small, it suffices to obtain quadrature point sets $\mathscr{P}$ making the worst case error $\sup_{\|f\|_H \leq 1}|\mathrm{Err}(f;\mathscr{P})|$ small. Since dealing with the worst case error directly is not easy, we often consider a manageable upper bound $W_H(\mathscr{P})$ on it;

$$|\mathrm{Err}(f;\mathscr{P})| \leq \|f\|_H \times W_H(\mathscr{P}). \tag{2}$$

Here we call $W_H(\mathscr{P})$ a figure of merit of $\mathscr{P}$ and these types of inequalities (2) are called Koksma-Hlawka type inequalities (for example, see [18] for details).

We often treat a point set $\mathscr{P}$ called 'digital net' (for example, see [10, 20]). A digital net $\mathscr{P}$ is defined as follows. Let $n, m \geq 1$ and $b \geq 2$ be integers with $n \geq m$. Let $0 \leq h < b^m$ be an integer and $C_1, \dots, C_s$ be $n \times m$ matrices over the finite group $\mathbb{Z}_b = \mathbb{Z}/b\mathbb{Z}$. We write the b-adic expansion $h = \sum_{j=1}^{m} h_j b^{j-1}$ and take a vector $\mathbf{h} = (h_1, \dots, h_m) \in (\mathbb{Z}_b^m)^\top$, where h_j is considered to be an element in $\mathbb{Z}_b$. For $1 \leq i \leq s$, we define a vector $(y_{h,i,1}, \dots, y_{h,i,n}) = \mathbf{h} \cdot (C_i)^\top$ and a real number $x_i(h) = \sum_{1 \leq j \leq n} y_{h,i,j} b^{-j} \in [0,1)$, where $y_{h,i,j}$ is considered to be an element of $\{0, \dots, b-1\} \subset \mathbb{Z}$. Then we define a digital net $\mathscr{P}$ by $\{\mathbf{x}_0, \dots, \mathbf{x}_{b^m-1}\}$ where $\mathbf{x}_h = (x_i(h))_{1 \leq i \leq s}$. In order to analyze QMC rules by digital nets, we use a dual net $\mathscr{P}^\perp$ [9, 21]:

$$\mathscr{P}^\perp := \{\mathbf{k} = (k_1, \dots, k_s) \in (\mathbb{N} \cup \{0\})^s \mid C_1^\top \vec{k}_1 + \cdots + C_s^\top \vec{k}_s = \mathbf{0} \in \mathbb{Z}_b^m\},$$

where $\vec{k}_i = (\kappa_{i,1}, \dots, \kappa_{i,n})^\top$ for k_i with b-adic expansion $k_i = \sum_{j \geq 1} \kappa_{i,j} b^{j-1}$. Here $\kappa_{i,j}$ is considered to be an element of $\mathbb{Z}_b$. Throughout this paper, when we take a point set $\mathscr{P}$, we assume that $\mathscr{P}$ is a digital net with $b = 2$.

In the classical theory, many researchers studied the integration error of a function f with bounded variation (or function with square integrable partial derivatives up to first order in each variable) (for example, see [10, 18]). In particular, they constructed many types of digital nets which achieve the optimal rate of convergence utilizing the theory of dual nets and Koksma-Hlawka type inequalities.

An extension to smooth periodic functions was established in [3], while a further extension to smooth (non-periodic) functions was shown in [4]. The QMC rules constructed in these papers, called higher order QMC rules, achieve the optimal rate of convergence. These rules are also constructed by obtaining a Koksma-Hlawka type inequality and analyzing properties of digital nets and their dual nets for higher order QMC rules. See also [5] for more background on higher order QMC rules.

Our main result in this paper is obtaining new Koksma-Hlawka type inequalities of digital nets for smooth function spaces, which leads to making practical optimal higher order QMC rules. One concrete application of this improvement is to substantially improve the constants in the bounds on integration error in [4], which is crucial in problems in uncertainty quantification [8, 12]. In particular, Gantner and Schwab [12] point out that the large constants from [4] cause problems in the CBC construction of interlaced polynomial lattice rules, which are construction methods to obtain a point set whose worst case error achieves the optimal order [7, 13]. To avoid this problem, it is suggested to use much smaller constants which are more realistic in [12]. This paper provides the theoretical justification for doing so.

We explain the details of our main result by comparing our result to the one in [8]. Dick et al. [8] introduced a smooth function space whose functions f satisfy that their norms (5) (see below) are finite. If f is a function whose mixed partial derivatives up to order α in each variable are continuous, then f is contained in this space. This space has some parameters called weights $\{\gamma_v\}_{v \subset S} \subset \mathbb{R}_{>0} = \{x \in \mathbb{R} : x > 0\}$, where $S := \{1, \dots, s\}$, which model the importance of different coordinate projections [25].

To state the results in [8], we introduce projected dual spaces and weight functions which correspond to the subsets $v \subset S$ as follows. For $\mathbf{k}_v \in \mathbb{N}^{|v|}$, let $(\mathbf{k}_v; \mathbf{0}) \in (\mathbb{N} \cup \{0\})^s$ denote the vector whose jth component is k_j if $j \in v$ and 0 otherwise. The dual space which corresponds to the subset $\varnothing \neq v \subset S$ is defined by $\mathscr{P}_v^\perp := \{\mathbf{k}_v \in \mathbb{N}^{|v|} \mid \mathbf{k} = (\mathbf{k}_v; \mathbf{0}) \in \mathscr{P}^\perp\}$ (note that none of the components in v is 0). The weight on a projected dual space is defined in the following way.

$$\mu_\alpha(l_1, \dots, l_{|v|}) = \sum_{i=1}^{|v|} \sum_{j \le \alpha} (a_{i,j} + 1)$$

for l_i with dyadic expansion $l_i = \sum_{j=1}^{N_i} 2^{a_{i,j}}$, with $a_{i,1} > \cdots > a_{i,N_i}$. Let $1 \le r$, $r', q \le \infty$ with $1/r + 1/r' = 1$. Dick et al. [8] showed the following bound on the worst case error (they also showed the results for a digital net with $b \ge 2$):

$$\sup_{\|f\|_{s,\alpha,\gamma,q,r} \le 1} |\mathrm{Err}(f; \mathscr{P})| \le e_{s,\alpha,\gamma,r'}(\mathscr{P}),$$

with

$$e_{s,\alpha,\gamma,r'}(\mathscr{P}) = \left(\sum_{\varnothing \neq v \subset S} \left(C_\alpha^{|v|} \gamma_v \sum_{\mathbf{k}_v \in \mathscr{P}_v^\perp} 2^{-\mu_\alpha(\mathbf{k}_v)} \right)^{r'} \right)^{1/r'}. \tag{3}$$

This implies the following inequality of the form (2):

$$|\mathrm{Err}(f; \mathscr{P})| \le \|f\|_{s,\alpha,\gamma,q,r} \times e_{s,\alpha,\gamma,r'}(\mathscr{P}), \tag{4}$$

where

$$\|f\|_{s,\alpha,\gamma,q,r} := \left(\sum_{u \subseteq S} \left(\gamma_u^{-q} \sum_{v \subseteq u} \sum_{\boldsymbol{\tau}_{u \setminus v} \in \{1,\dots,\alpha-1\}^{|u \setminus v|}} \right.\right.$$

$$\left.\left. \int_{[0,1]^{|v|}} \left| \int_{[0,1]^{s-|v|}} f^{(\boldsymbol{\alpha}_v, \boldsymbol{\tau}_{u \setminus v}, \mathbf{0})}(\mathbf{x}) \mathrm{d}\mathbf{x}_{S \setminus v} \right|^q \mathrm{d}\mathbf{x}_v \right)^{r/q} \right)^{1/r}, \qquad (5)$$

with the obvious modifications if q or r is infinite.[1] Here $(\boldsymbol{\alpha}_v, \boldsymbol{\tau}_{u \setminus v}, \mathbf{0})$ denotes a vector $(v_j)_{j=1}^s$ with $v_j = \alpha$ for $j \in v$, $v_j = \tau_j$ for $j \in u \setminus v$, and $v_j = 0$ for $j \notin u$. And we write $f^{(n_1,\dots,n_s)} = \frac{\partial^{n_1+\cdots+n_s} f}{\partial x_1^{n_1} \dots \partial x_s^{n_s}}$.

Based on these bounds on the integration error, Dick constructed 'interlaced digital nets' to obtain a point set with small integration error [3, 4]. He showed that the worst case error of this type of point sets achieves the order $O(N^{-\alpha}(\log N)^{s\alpha})$ in terms of the cardinality N of a point set [4]. This is known to be optimal up to log terms [23]. In [1, 2, 7, 8, 12, 13], there are component-by-component (CBC) algorithms to obtain point sets which achieve the same order. This construction is also based on the bounds of the form (4).

This paper gives a new Koksma-Hlawka type inequality of digital nets to bound the integration error of smooth functions which improves upon (4) in some instances. Here we use the following notation:

$$\mu'_\alpha(l_1, \dots, l_{|v|}) = \sum_{i=1}^{|v|} \sum_{j \le \alpha} (a_{i,j} + 2)$$

for $l_i = \sum_{j=1}^{N_i} 2^{a_{i,j}}$ instead of Dick's weight function μ_α.

Theorem 1. *Let $\alpha \in \mathbb{N} \cup \{\infty\}$ such that $\alpha \ge 2$. Assume that the function f has continuous mixed partial derivatives up to order α in each variable x_i on $[0,1]^s$, and $1 \le p, q, q' \le \infty$ such that $1/q + 1/q' = 1$. Then we have*

$$|\mathrm{Err}(f; \mathscr{P})| \le \|f\|_{\mathscr{B}_\alpha, \gamma, p, q'} \times \mathscr{W}_{\alpha,\gamma,q}(\mathscr{P}),$$

where

$$\mathscr{W}_{\alpha,\gamma,q}(\mathscr{P}) = \left(\sum_{\emptyset \ne v \subset S} \left(\gamma_v \sum_{\mathbf{k}_v \in \mathscr{P}_v^\perp} 2^{-\mu'_\alpha(\mathbf{k}_v)} \right)^q \right)^{1/q},$$

$$\|f\|_{\mathscr{B}_\alpha,\gamma,p,q'} = \left(\sum_{\emptyset \ne v \subset S} \left(\gamma_v^{-1} 2^{|v|/p} \sup_{\boldsymbol{\alpha}'_v \in \{1,\dots,\alpha\}^{|v|}} \|f^{(\boldsymbol{\alpha}'_v)}\|_p \right)^{q'} \right)^{1/q'}, \qquad (6)$$

[1] The norm in [8, Definition 3.3] has been corrected in arXiv:1309.4624v3. The correct version is restated here in Eq. (5).

and where

$$\|f^{(\boldsymbol{\alpha}'_v)}\|_p = \left(\int_{[0,1)^{|v|}} \left| \int_{[0,1)^{s-|v|}} f^{(\boldsymbol{\alpha}'_v,\mathbf{0})}(\mathbf{x}) d\mathbf{x}_{S\setminus v} \right|^p d\mathbf{x}_v \right)^{1/p},$$

with the obvious modifications if either p, q or q' is infinite. Here $(\boldsymbol{\alpha}'_v, \mathbf{0})$ denotes a sequence $(v_j)_j$ with $v_j \in \{1, \dots, \alpha\}$ for $j \in v$ and $v_j = 0$ for $j \notin v$ and $f^{(n_1,\dots,n_s)} = \frac{\partial^{n_1+\cdots+n_s} f}{\partial x_1^{n_1} \dots \partial x_s^{n_s}}$.

This result yields a significant improvement of (4). In particular, this is crucial when using the bound in a CBC algorithm, since a large constant (as it appears in [8, Theorem 3.5]) may make it impractical to perform the CBC construction. For instance, [12, Section 4.1] write that "*The resulting large values of the worst-case error bounds* [referring to the large constants in [8, Theorem 3.5]] *have been found to lead to generating vectors with bad projections.*" Additionally, we also include the case $\alpha = \infty$ which has not been studied before in the context of digital nets. In [8, Theorem 3.5], the case $\alpha = \infty$ is not included since in this case the constant C_α appearing in (4) is infinite. Furthermore, we can define a modified version of Walsh Figure of Merit (WAFOM) [19] when we consider this new bound (see [15] for details). WAFOM is a computable figure of merit to find good quadrature point sets for QMC rules for integrands with large enough smoothness α (see [15, 16, 19] for details).

Theorem 1 is based on the estimation of $\mathrm{Err}(f; \mathscr{P})$ by '(dyadic) Walsh coefficients'. In this paper, we make elaborate works for obtaining bounds on them. Dyadic Walsh coefficients are defined as follows (see [11, 22] for details).

Definition 1 (Walsh functions and Walsh coefficients). Let $f : [0,1)^s \to \mathbb{R}$ and $\mathbf{k} = (k_1, \dots, k_s) \in (\mathbb{N} \cup \{0\})^s$. We define the $\mathbf{k}$-th dyadic Walsh function $\mathrm{wal}_{\mathbf{k}}$ by

$$\mathrm{wal}_{\mathbf{k}}(\mathbf{x}) := \prod_{i=1}^{s} (-1)^{\left(\sum_{j\ge 1} a_{i,j} b_{i,j}\right)},$$

where for $1 \le i \le s$, we write the dyadic expansion of k_i by $k_i = \sum_{j\ge1} a_{i,j} 2^{j-1}$ and x_i by $x_i = \sum_{j\ge 1} b_{i,j} 2^{-j}$, where for each i, infinitely many digits $b_{i,j}$ are 0.

Using Walsh functions, we also define the $\mathbf{k}$-th dyadic Walsh coefficient $\hat{f}(\mathbf{k})$ as follows:

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

Using the Walsh coefficients $\hat{f}(\mathbf{k})$, the integration error $\mathrm{Err}(f;\mathscr{P})$ by a digital net $\mathscr{P}$ can be represented as follows [10, Chapter 15]:

$$\mathrm{Err}(f;\mathscr{P}) = \sum_{\mathbf{k}\in\mathscr{P}^{\perp}\setminus\{\mathbf{0}\}} \hat{f}(\mathbf{k}) = \sum_{\varnothing\neq v\subset S}\sum_{\mathbf{k}_v\in\mathscr{P}_v^{\perp}} \hat{f}(\mathbf{k}_v;\mathbf{0}).$$

The proof of Theorem 1 is facilitated by the following improved bound on the Walsh coefficients of smooth functions.

Theorem 2. *We assume the same assumptions as in Theorem* 1. *Let* $\varnothing\neq v\subset\{1,\dots,s\}=S$. *For* $\mathbf{k}_v\in\mathbb{N}^{|v|}$, *we have*

$$|\hat{f}(\mathbf{k}_v;\mathbf{0})| \le 2^{|v|/p}\cdot 2^{-\mu'_{\alpha}(\mathbf{k}_v)}\cdot \|f^{(\min(\alpha,\mathbf{N}_{\mathbf{k}_v}))}\|_p, \tag{7}$$

where $1\le p\le\infty$ *and* $\|\cdot\|_p$ *is the norm defined in Theorem* 1. *Here we define the symbol* $\min(\alpha,\mathbf{N}_{\mathbf{l}}) = (\min(\alpha,N_1),\dots,\min(\alpha,N_{|v|}))$ *for* $\mathbf{l}=(l_1,\dots,l_{|v|})$ *with dyadic expansion* $l_i=\sum_{j=1}^{N_i}2^{a_{i,j}}$.

This inequality follows from the formula for the Walsh coefficients by dyadic differences, which are defined in Section 3.

Now we compare Theorem 2 with [6, Theorem 14] and its higher dimensional analogue in [8] (both of two papers also showed the results for b-adic Walsh coefficients with $b\ge 2$). Our bound (7) includes the case $\alpha=\infty$ for the case $b=2$ (they treat only the case α is finite). Further (7) is better than [6, Theorem 14] in some instances. For example, assume that $s=1$. Then for $N_1\ge\alpha$, if we multiply our bound by $(5/3)^{\alpha-2}$ our bound is still smaller than the bound by Dick for any $k_1=\sum_{j=1}^{N_1}2^{a_{1,j}}$ (see [10, chapter 14]). If $N_1<\alpha$ it is in general not clear which bound is better. In higher dimensions $s>1$, we can also compare our result with the bound in [8, In the proof of Theorem 3.5]. As in the univariate case, if we multiply our result by $c^{\alpha|v|}$ for some instant $c>1$, it is still smaller than the bound in [8, In the proof of Theorem 3.5].

The remainder of this paper is organized as follows. In Section 2 we give necessary definitions and their properties. We give the proof of Theorem 2 in Section 3 using two lemmas, which allows readers to understand the outline of this proof. In Section 4 we show the proofs of lemmas to complete the proof of Theorem 2. In Section 5, we give the proof of Theorem 1.

In the following, we denote $\mathbb{N}\cup\{0\}$ by $\mathbb{N}_0$ and continue to use the symbol $f^{(n_1,\dots,n_s)}$ instead of $\frac{\partial^{n_1+\cdots+n_s}f}{\partial x_1^{n_1}\dots\partial x_s^{n_s}}$ for $(n_i)_{i=1}^s\in\mathbb{N}_0^s$. Further we assume that $k\in\mathbb{N}_0$ has dyadic expansion with $k=\sum_{j=1}^{N}2^{a_j}$ where N is some integer and $a_1,\dots,a_N\in\mathbb{N}_0$ satisfying $a_1>\cdots>a_N$. Here we set $N=0$, $\{a_j\}=\varnothing$ for $k=0$.

2. Preliminaries

In this section, we introduce some operators and functions, and give their properties.

2.1. Definitions of dyadic differences and important functions. We begin with the introduction of the dyadic difference $\partial_{i,n}$ and the weight function $\mu'_{\mathbf{u}}$.

Definition 2 (dyadic difference). Let $s, n, i \in \mathbb{N}$ with $i \leq s$. For a function $g : [0,1)^s \to \mathbb{R}$, we define the operator 'dyadic difference' $\partial_{i,n}$ by

$$\partial_{i,n}(g)(x_1, \ldots, x_s) := \frac{g(x_1, \ldots, x_i \oplus 2^{-n}, \ldots, x_s) - g(x_1, \ldots, x_i, \ldots, x_s)}{2^{-n}},$$

where $(x_1, \ldots, x_s) \in [0,1)^s$ and we write $z \oplus 2^{-n} := z + 2^{-n}(-1)^{z_n}$ for z having dyadic expansion $z = \sum_{j=1}^{\infty} z_j 2^{-j}$, where infinitely many digits z_j are 0.

Let $S = \{1, \ldots, s\}$. Let $\mathbf{k} = (k_1, \ldots, k_s) \in \mathbb{N}_0^s$, $\mathbf{u} = (u_1, \ldots, u_s) \in (\mathbb{N}_0 \cup \{\infty\})^s$ with $k_i = \sum_{j=1}^{N_i} 2^{a_{i,j}}$. For any vector $\mathbf{k}$ we define the subset $v = \{i \in S \mid k_i \neq 0\}$. In this paper we use the symbol $\mathbf{d}_{\mathbf{k},\mathbf{u}}$ to denote the composition of the operators $\{\partial_{i, a_{i,j}+1}\}_{i \in v, 1 \leq j \leq \min(N_i, u_i)}$.

Remark 1. Since any two dyadic differences commute, $\mathbf{d}_{\mathbf{k},\mathbf{u}}$ is defined independent of the order of a composition.

Definition 3 (the new weight function $\mu'_{\mathbf{u}}(\mathbf{k})$). We use the same symbols as in Definition 2. The weight function $\mu'_{\mathbf{u}}(\mathbf{k})$ of $\mathbf{k}$ is defined by

$$\mu'_{\mathbf{u}}(\mathbf{k}) := \sum_{i \in v, 1 \leq j \leq \min(N_i, u_i)} (a_{i,j} + 2),$$

and we define $\mu'_{\mathbf{u}}(\mathbf{0}) = 0$.

For any vector $\mathbf{k} = (k_i)_{i=1}^s$ we write $\mathbf{k}_v = (k_i)_{i \in v}$, where the subset v appeared in Definition 2. Further, we write $\mathbf{k} = (\mathbf{k}_v; 0)$. When $u_i = \alpha \in \mathbb{N} \cup \{\infty\}$ for every i, $\mu'_{\mathbf{u}}(\mathbf{k}_v; \mathbf{0})$ corresponds with $\mu'_{\alpha}(\mathbf{k}_v)$, which already appeared in Theorem 1.

We also need the following two functions χ_n and $W(\mathbf{k})$. These functions have important roles in order to prove the main results.

Definition 4. For $n \in \mathbb{N}$, we define the function $\chi_n : [0,1)^2 \to \mathbb{R}$ by

$$\chi_n(x, y) := \begin{cases} 2^n & \text{if } y \in [\min(x, x \oplus 2^{-n}), \max(x, x \oplus 2^{-n})], \\ 0 & \text{otherwise,} \end{cases}$$

where $x, y \in [0,1)$.

Using this, we define the 1-dimensional function $W(k) : [0,1) \to \mathbb{R}$ inductively by

$$W(0)(y) := 1,$$

$$W(2^{n_1})(y) := \int_0^1 \chi_{n_1+1}(x, y)dx,$$

$$W(2^{n_1} + \cdots + 2^{n_{N+1}})(y) := \int_0^1 \chi_{n_{N+1}+1}(x, y)W(2^{n_1} + \cdots + 2^{n_N})(x)dx,$$

where $y \in [0,1)$ and $n_1, \dots n_{N+1} \in \mathbb{N}_0$ satisfies $n_1 > \cdots > n_{N+1}$.

The s-dimensional function $W(\mathbf{k}) : [0,1)^s \to \mathbb{R}$ for a vector $\mathbf{k} = (k_1, \dots, k_s) \in \mathbb{N}_0^s$ is defined as follows:

$$W(\mathbf{k})(\mathbf{y}) := \prod_{i=1}^{s} W(k_i)(y_i), \qquad \mathbf{y} = (y_i)_{i=1}^s \in [0,1)^s.$$

REMARK 2. By this definition, $W(k)$ is continuous on $[0,1)$ for any $k \in \mathbb{N}_0$.

2.2. Evaluation of the L^p-norm of $W(k)$. In Section 3, the bounds on the Walsh coefficients in Theorem 2 are deduced by using Hölder's inequality for derivatives of f and the function $W(k)$, where we need the L^p-norm of $W(k)$. In this subsection we give a bound on this value.

2.2.1. *Important properties of χ_n and $W(k)$.* We show the important properties of χ_n and $W(k)$ necessary to calculate $\|W(k)\|_{L^p}$.

We first show the following technical lemma in terms of χ_n.

LEMMA 1. *Let $n \in \mathbb{N}$ and $c, \hat{c} \in \mathbb{N}_0$ satisfy $c, \hat{c} < 2^{n-1}$.*

(1) *Let $x, y \in [0, 2^{-n+1})$, then we have*

$$\chi_n(x + c2^{-n+1}, y + c2^{-n+1}) = \chi_n(x, y).$$

(2) *Let $x \in [c2^{-n+1}, (c+1)2^{-n+1})$ and $y \in [\hat{c}2^{-n+1}, (\hat{c}+1)2^{-n+1})$ with $c \neq \hat{c}$. Then we have $\chi_n(x, y) = 0$.*

PROOF.

(1) We have $(x + c2^{-n+1}) \oplus 2^{-n} = (x \oplus 2^{-n}) + c2^{-n+1}$. Thus the result follows from the fact that $y \in [\min(x, x \oplus 2^{-n}), \max(x, x \oplus 2^{-n})]$ is equivalent to

$$y + c2^{-n+1} \in [\min(x + c2^{-n+1}, (x + c2^{-n+1}) \oplus 2^{-n}),$$
$$\max(x + c2^{-n+1}, (x + c2^{-n+1}) \oplus 2^{-n})].$$

(2) Let $x \in [d2^{-n}, (d+1)2^{-n})$ for $d \in \mathbb{N}_0$ where $d = 2c$ or $2c+1$. When $d = 2c$, it holds that $c2^{-n+1} \leq x < x \oplus 2^{-n} = x + 2^{-n} < (c+1)2^{-n+1}$. In the case $d = 2c+1$, it holds that $c2^{-n+1} \leq x \oplus 2^{-n} = x - 2^{-n} < x < (c+1)2^{-n+1}$. Thus, in both cases, we have

$$[\min(x, x \oplus 2^{-n}), \max(x, x \oplus 2^{-n})] \subset [c2^{-n+1}, (c+1)2^{-n+1}).$$

Then if $y \in [\hat{c}2^{-n+1}, (\hat{c}+1)2^{-n+1})$ with $\hat{c} \neq c$, we have

$$y \notin [\min(x, x \oplus 2^{-n}), \max(x, x \oplus 2^{-n})],$$

which implies $\chi_n(x, y) = 0$. □

In the next Lemma, we rewrite $\chi_n(x, y)$ using a characteristic function of x for a fixed y.

Lemma 2. *Let $y \in [0,1)$, $n \in \mathbb{N}$ and $c \in \mathbb{N}_0$ satisfy $y \in [2^{-n}c, 2^{-n}(c+1))$. If $c = 2c'$ for some integer c', we have*

$$\chi_n(x, y) = \begin{cases} 2^n & \text{if } x \in [2^{-n}c, y] \cup [2^{-n}(c+1), y + 2^{-n}], \\ 0 & \text{otherwise}. \end{cases}$$

And if $c = 2c'+1$ for some integer c', we have

$$\chi_n(x, y) = \begin{cases} 2^n & \text{if } x \in [y - 2^{-n}, 2^{-n}c) \cup [y, 2^{-n}(c+1)), \\ 0 & \text{otherwise}. \end{cases}$$

Proof. We only prove the case $c = 2c'$ here since the case $c = 2c'+1$ follows from the same argument. In this case, by Item 2 of Lemma 1, we have that for all $y \in [2^{-n+1}c', 2^{-n+1}(c'+1))$,

$$\chi_n(x, y) = 0, \qquad x \notin [2^{-n+1}c', 2^{-n+1}(c'+1)). \tag{8}$$

Let $x \in [0,1)$ and $d \in \mathbb{N}_0$ satisfy $x \in [2^{-n}d, 2^{-n}(d+1))$. We calculate $\chi_n(x, y)$ in the three cases: $d \notin \{2c', 2c'+1\}$, $d = 2c'$ and $d = 2c'+1$.

We consider the case $d \notin \{2c', 2c'+1\}$. By Condition (8), we have

$$\chi_n(x, y) = 0, \qquad x \in [2^{-n}d, 2^{-n}(d+1)).$$

In the case $d = 2c'$, that is, $x \in [2^{-n+1}c', 2^{-n}(2c'+1))$, since $x \oplus 2^{-n} = x + 2^{-n}$, the fact that $\chi_n(x, y) = 2^n$ is equivalent to

$$x \leq y \leq x + 2^{-n}, \qquad 2^{-n+1}c' \leq y < 2^{-n}(2c'+1).$$

So we have

$$\chi_n(x, y) = \begin{cases} 2^n & \text{if } x \in [2^{-n+1}c', y], \\ 0 & \text{if } x \in (y, 2^{-n}(2c'+1)). \end{cases}$$

When $d = 2c' + 1$, by a similar argument to the case $d = 2c'$, we have

$$\chi_n(x, y) = \begin{cases} 2^n & \text{if } x \in [2^{-n}(2c'+1), y + 2^{-n}], \\ 0 & \text{if } x \in (y + 2^{-n}, 2^{-n}(2c'+2)). \end{cases}$$

By combining these cases, we have the result. $\square$

In the last lemma, we show that the function $W(k)$ is periodic.

Lemma 3. *We have that $W(k)$ is a periodic function with period 2^{-a_N} for a positive integer $k = \sum_{i=1}^{N} 2^{a_i}$.*

Proof. We proceed by induction on N. We prove the result for $k = 2^{a_1}$. Let $c \in \mathbb{N}_0$ satisfy $c < 2^{a_1}$. We have that for $y \in [0, 2^{-a_1})$,

$$\begin{aligned} W(2^{a_1})(y + c2^{-a_1}) &= \int_0^1 \chi_{a_1+1}(x, y + c2^{-a_1})dx \\ &= \int_{c2^{-a_1}}^{(c+1)2^{-a_1}} \chi_{a_1+1}(x, y + c2^{-a_1})dx \\ &= \int_0^{2^{-a_1}} \chi_{a_1+1}(z + c2^{-a_1}, y + c2^{-a_1})dz \\ &= \int_0^{2^{-a_1}} \chi_{a_1+1}(z, y)dz = \int_0^1 \chi_{a_1+1}(z, y)dz = W(2^{a_1})(y). \end{aligned}$$

The second and fifth equalities follow from Item 2 of Lemma 1, the forth equality follows from Item 1 of Lemma 1, and we use the change of variables $x = z + c2^{-a_1}$ in the third equality.

Now we assume that the lemma holds for any $k_N = \sum_{i=1}^{N} 2^{a_i}$. Then we prove the result for the case $k = k_N + 2^{a_{N+1}}$ satisfying $a_{N+1} < a_N$. Let $c \in \mathbb{N}_0$ satisfy $c < 2^{a_{N+1}}$. Then we have that for $y \in [0, 2^{-a_{N+1}})$,

$$\begin{aligned} &W(k_N + 2^{a_{N+1}})(y + c2^{-a_{N+1}}) \\ &\quad = \int_0^1 \chi_{a_{N+1}+1}(x, y + c2^{-a_{N+1}}) \cdot W(k_N)(x)dx \\ &\quad = \int_{c2^{-a_{N+1}}}^{(c+1)2^{-a_{N+1}}} \chi_{a_{N+1}+1}(x, y + c2^{-a_{N+1}}) \cdot W(k_N)(x)dx \\ &\quad = \int_0^{2^{-a_{N+1}}} \chi_{a_{N+1}+1}(z + c2^{-a_{N+1}}, y + c2^{-a_{N+1}}) \cdot W(k_N)(z + c2^{-a_{N+1}})dz. \end{aligned}$$

The second equality follows from Item 2 of Lemma 1, and we use the change of variables $x = z + c2^{-a_{N+1}}$ in the last equality. By the induction assumption and the fact $a_{N+1} < a_N$, we have that $W(k_N)(z + d2^{-a_{N+1}}) = W(k_N)(z)$ for $z \in [0, 2^{-a_{N+1}})$ and an integer d satisfying $0 \le d < 2^{a_{N+1}}$. Thus we obtain

$$\begin{aligned} &W(k_N + 2^{a_{N+1}})(y + c2^{-a_{N+1}}) \\ &\quad = \int_0^{2^{-a_{N+1}}} \chi_{a_{N+1}+1}(z + c2^{-a_{N+1}}, y + c2^{-a_{N+1}}) \cdot W(k_N)(z)dz. \end{aligned}$$

Then we continue the computation as follows:

$$\begin{aligned} W(k_N + 2^{a_{N+1}})(y + c2^{-a_{N+1}}) &= \int_0^{2^{-a_{N+1}}} \chi_{a_{N+1}+1}(z, y) \cdot W(k_N)(z)dz \\ &= \int_0^1 \chi_{a_{N+1}+1}(z, y) \cdot W(k_N)(z)dz \\ &= W(k_N + 2^{a_{N+1}})(y). \end{aligned}$$

The first equality follows from Item 1 of Lemma 1 and the second equality follows from Item 2 of Lemma 1. □

2.2.2. *Bound on the L^p norm of the function $W(k)$.* Now we prove the following bound on $\|W(k)\|_{L^p}$ using the properties of $W(k)$ we proved above.

Lemma 4. *Let $k \in \mathbb{N}$. We have that $W(k)(x) \ge 0$ for any $x \in [0, 1)$ and, for $1 \le p \le \infty$, we have*

$$\|W(k)\|_{L^p} \le 2^{1-1/p}.$$

Proof. By the definition of χ_n, it holds that

$$\chi_n(x, y) \ge 0, \qquad x, y \in [0, 1). \tag{9}$$

By using this, we can show that

$$W(k)(x) \ge 0, \qquad x \in [0, 1),\ k \in \mathbb{N} \tag{10}$$

by induction, which is easy. We omit the proof here. We use this result to prove the second statement.

Using Hölder's inequality, we have for $1 \le p < \infty$,

$$\|W(k)\|_{L^p}^p = \int_0^1 |W(k)(x)| \cdot |W(k)(x)^{p-1}| dx \le \|W(k)\|_{L^1} \|W(k)\|_{L^\infty}^{p-1}.$$

Thus we have

$$\|W(k)\|_{L^p} \le \|W(k)\|_{L^1}^{1/p} \|W(k)\|_{L^\infty}^{1-1/p}.$$

Then if we have $\|W(k)\|_{L^1} = 1$ and $\|W(k)\|_{L^\infty} = 2$, we have

$$\|W(k)\|_{L^p} \le \|W(k)\|_{L^1}^{1/p} \|W(k)\|_{L^\infty}^{1-1/p} = 2^{(1-1/p)}.$$

Therefore we need to prove that $\|W(k)\|_{L^1} = 1$ and $\|W(k)\|_{L^\infty} = 2$ for any $k \in \mathbb{N}$ to complete the proof.

We prove the case $k = \sum_{i=1}^N 2^{a_i}$ by induction on N. In the case $k = 2^{a_1}$, we have

$$\begin{aligned}\|W(2^{a_1})\|_{L^1} &= \int_0^1 \int_0^1 \chi_{a_1+1}(x,y)\,dydx \\ &= \int_0^1 \int_{\min(x, x\oplus 2^{-a_1-1})}^{\max(x, x\oplus 2^{-a_1-1})} 2^{a_1+1}\,dydx = 1.\end{aligned}$$

The first equality follows from Condition (9). We prove $\|W(2^{a_1})\|_{L^\infty} = 2$ to complete the case $k = 2^{a_1}$. By Lemma 3, we have

$$\|W(2^{a_1})\|_{L^\infty} = \sup_{y\in[0,1)} |W(2^{a_1})(y)| = \sup_{y\in[0,2^{-a_1})} |W(2^{a_1})(y)|.$$

By Condition (9), we have

$$\begin{aligned}\|W(2^{a_1})\|_{L^\infty} &= \sup_{y\in[0,2^{-a_1})} \int_0^1 \chi_{a_1+1}(x,y)\,dx \\ &= \max\left(\sup_{y\in[0,2^{-a_1-1})} \int_0^1 \chi_{a_1+1}(x,y)\,dx, \sup_{y\in[2^{-a_1-1},2^{-a_1})} \int_0^1 \chi_{a_1+1}(x,y)\,dx\right).\end{aligned}$$

We calculate the supremum on $[0, 2^{-a_1-1})$ and $[2^{-a_1-1}, 2^{-a_1})$ separately. We assume that $y \in [0, 2^{-a_1-1})$. By Lemma 2, we have

$$\int_0^1 \chi_{a_1+1}(x,y)\,dx = \int_0^y 2^{a_1+1}\,dx + \int_{2^{-a_1-1}}^{2^{-a_1-1}+y} 2^{a_1+1}\,dx.$$

If we choose $y = 2^{-a_1-1}$, we can maximize the right hand side. Hence we obtain

$$\sup_{y\in[0,2^{-a_1-1})} \int_0^1 \chi_{a_1+1}(x,y)\,dx = \int_0^{2^{-a_1}} 2^{a_1+1}\,dx_1 = 2.$$

By the same argument we get the same result in the case $y \in [2^{-a_1-1}, 2^{-a_1})$. We omit the details. Therefore we have $\|W(2^{a_1})\|_{L^\infty} = 2$.

Here we assume that $\|W(k_N)\|_{L^1} = 1$ and $\|W(k_N)\|_{L^\infty} = 2$ for $k_N = \sum_{i=1}^N 2^{a_i}$. Then we prove that for the case $k = k_N + 2^{a_{N+1}}$ with $a_N > a_{N+1}$:

$\|W(k_N + 2^{a_{N+1}})\|_{L^1} = 1$ and $\|W(k_N + 2^{a_{N+1}})\|_{L^\infty} = 2$. By Condition (10) and Fubini's Theorem, we have

$$\|W(k_N + 2^{a_{N+1}})\|_{L^1} = \int_0^1 \int_0^1 \chi_{a_{N+1}+1}(x, y) W(k_N)(x) dx dy = \int_0^1 W(k_N)(x) dx.$$

By Condition (10) and the assumption on k_N, we obtain

$$\|W(k_N + 2^{a_{N+1}})\|_{L^1} = \int_0^1 W(k_N)(x) dx = \|W(k_N)\|_{L^1} = 1.$$

We also prove $\|W(k_N + 2^{a_{N+1}})\|_{L^\infty} = 2$ as follows. By Condition (10) and Lemma 3, we have

$$\begin{aligned}
\|W(k_N + 2^{a_{N+1}})\|_{L^\infty} &= \sup_{y \in [0, 2^{-a_{N+1}})} W(k_N + 2^{a_{N+1}})(y) \\
&= \max\Bigg(\sup_{y \in [0, 2^{-a_{N+1}-1})} \int_0^1 \chi_{a_{N+1}+1}(x, y) W(k_N)(x) dx, \\
&\qquad \sup_{y \in [2^{-a_{N+1}-1}, 2^{-a_{N+1}})} \int_0^1 \chi_{a_{N+1}+1}(x, y) W(k_N)(x) dx \Bigg).
\end{aligned}$$

By Lemma 2 and Condition (10), we can calculate the supremum as in the case $k = 2^{a_1}$:

$$\left.\begin{array}{l} \sup_{y \in [0, 2^{-a_{N+1}-1})} \int_0^1 \chi_{a_{N+1}+1}(x, y) W(k_N)(x) dx \\ \sup_{y \in [2^{-a_{N+1}-1}, 2^{-a_{N+1}})} \int_0^1 \chi_{a_{N+1}+1}(x, y) W(k_N)(x) dx \end{array}\right\} = \int_0^{2^{-a_{N+1}}} 2^{a_{N+1}+1} W(k_N)(x) dx.$$

Then we obtain

$$\begin{aligned}
\|W(k_N + 2^{a_{N+1}})\|_{L^\infty} &= \int_0^{2^{-a_{N+1}}} 2^{a_{N+1}+1} W(k_N)(x) dx \\
&= 2^{a_{N+1}+1} \sum_{i=0}^{2^{a_N - a_{N+1}} - 1} \int_{i 2^{-a_N}}^{(i+1) 2^{-a_N}} W(k_N)(x) dx \\
&= 2^{a_{N+1}+1} \sum_{i=0}^{2^{a_N - a_{N+1}} - 1} \int_0^{2^{-a_N}} W(k_N)(x) dx \\
&= 2^{a_N+1} \cdot \int_0^{2^{-a_N}} W(k_N)(x) dx.
\end{aligned}$$

The third equality follows from Lemma 3. Thus we have

$$\begin{aligned}\|W(k_N+2^{a_{N+1}})\|_{L^\infty} &= 2^{a_N+1}\cdot\int_0^{2^{-a_N}} W(k_N)(x)dx\\ &= 2\sum_{i=0}^{2^{a_N}-1}\int_{i2^{-a_N}}^{(i+1)2^{-a_N}} W(k_N)(x)dx\\ &= 2\cdot\int_0^1 W(k_N)(x)dx,\end{aligned}$$

where the second equality follows from Lemma 3. By the assumption on k_N and Condition (10), it follows that

$$\int_0^1 W(k_N)(x)dx = \|W(k_N)\|_{L^1} = 1$$

and hence we obtain

$$\|W(k_N+2^{a_{N+1}})\|_{L^\infty} = 2\cdot\|W(k_N)\|_{L^1} = 2. \qquad \square$$

3. Bounds on the Walsh coefficients by derivatives

In this section, we show the bounds on the Walsh coefficients $\hat{f}(\mathbf{k})$ in Theorem 2 using the following Lemmas 5 and 6, which we prove in the next section. Lemma 5 gives the formula relating $\hat{f}(\mathbf{k})$ to dyadic differences $\partial_{i,n}f$. Via this lemma, we can obtain the formula relating $\hat{f}(\mathbf{k})$ to derivatives $\frac{\partial f}{\partial x_i}$ in Lemma 6. The bounds on the Walsh coefficients $\hat{f}(\mathbf{k})$ in Theorem 2 are deduced by this formula.

We define the symbols used in the statements.

Definition 5. We use the same symbols as in Definition 2. For $\mathbf{k}$ and $\mathbf{u}$, we define

$$k_{i,>}^{u_i} := \begin{cases}\sum_{j>u_i} 2^{a_{i,j}} & \text{if } i\in v,\\ 0 & \text{if } i\in S\backslash v,\end{cases} \qquad k_{i,\le}^{u_i} := \begin{cases}\sum_{j\le u_i} 2^{a_{i,j}} & \text{if } i\in v,\\ 0 & \text{if } i\in S\backslash v,\end{cases}$$

and

$$\begin{aligned}&\mathbf{k}_{>}^{\mathbf{u}} := (k_{i,>}^{u_i})_{i\in S}, && \mathbf{k}_{\le}^{\mathbf{u}} := (k_{i,\le}^{u_i})_{i\in S},\\ &\mathbf{k}_{v>}^{\mathbf{u}} := (k_{i,>}^{u_i})_{i\in v}, && \mathbf{k}_{v\le}^{\mathbf{u}} := (k_{i,\le}^{u_i})_{i\in v},\\ &\min(\mathbf{u},\mathbf{N}_{\mathbf{k}}) := (\min(u_i,N_i))_{i\in S}, && \min(\mathbf{u},\mathbf{N}_{\mathbf{k}_v}) := (\min(u_i,N_i))_{i\in v},\\ &|\min(\mathbf{u},\mathbf{N}_{\mathbf{k}})|_{l^1} := \sum_{i\in S}\min(u_i,N_i).\end{aligned}$$

When we analyze Walsh coefficients, it is suitable to use dyadic differences. In fact, the $\mathbf{k}$-th Walsh coefficient $\hat{f}(\mathbf{k})$ can be represented by $\partial_{i,n} f$ as follows:

Lemma 5. *Let* $f \in L^1([0,1)^s)$, $\mathbf{k} \in \mathbb{N}_0^s$ *and* $\mathbf{u} \in (\mathbb{N}_0 \cup \{\infty\})^s$. *Then we have* $\mathbf{d}_{\mathbf{k},\mathbf{u}} f \in L^1([0,1)^s)$ *and the formula:*

$$\hat{f}(\mathbf{k}) = (-1)^{|\min(\mathbf{u}, \mathbf{N_k})|_{l^1}} 2^{-\mu'_{\mathbf{u}}(\mathbf{k})} \widehat{\mathbf{d}_{\mathbf{k},\mathbf{u}} f}(\mathbf{k}). \tag{11}$$

Notice that $\mathbf{d}_{\mathbf{k},\mathbf{u}}$ *is the symbol introduced in Definition* 2.

Formula (11) means that dyadic differences connect the $\mathbf{k}$-th Walsh coefficient $\hat{f}(\mathbf{k})$ to the weight function $\mu'_{\mathbf{u}}(\mathbf{k})$ for $f \in L^1([0,1)^s)$.

Dyadic differences $\partial_{i,n} f$ are similar to derivatives $\frac{\partial f}{\partial x_i}$. Using the assumption that f has continuous partial mixed derivatives, we can continue to compute the right hand side in this formula to change $\partial_{i,n} f$ into $\frac{\partial f}{\partial x_i}$.

Lemma 6. *We write* $\mathbf{u} = (u_1, \ldots, u_s) \in (\mathbb{N}_0 \cup \{\infty\})^s$. *We assume that a function* f *satisfies that its mixed partial derivatives up to order* u_i *in each variable* x_i *are continuous on* $[0,1]^s$. *Then for any vector* $\mathbf{k} \in \mathbb{N}_0^s$, *we have*

$$\hat{f}(\mathbf{k}) = (-1)^{|\min(\mathbf{u}, \mathbf{N_k})|_{l^1}} 2^{-\mu'_{\mathbf{u}}(\mathbf{k})} \int_{[0,1)^s} f^{(\min(\mathbf{u}, \mathbf{N_k}))}(\mathbf{x}) \cdot W(\mathbf{k}^{\mathbf{u}}_{\leq})(\mathbf{x}) \cdot \mathrm{wal}_{\mathbf{k}^{\mathbf{u}}_{>}}(\mathbf{x}) d\mathbf{x}.$$

Then we can get the following bound on $|\hat{f}(\mathbf{k})|$ which is a little more general than Theorem 2:

Lemma 7. *We assume the same assumptions as in Lemma* 6. *Then we have*

$$|\hat{f}(\mathbf{k})| \leq 2^{|v|/p} \cdot 2^{-\mu'_{\mathbf{u}}(\mathbf{k})} \cdot \|f^{(\min(\mathbf{u}, \mathbf{N}_{\mathbf{k}_v}))}\|_p < \infty,$$

where $1 \leq p \leq \infty$ *and* $\|\cdot\|_p$ *is the norm defined in Theorem* 1.

Proof. For $\mathbf{k} \in \mathbb{N}_0^s$, the i-th component of $\mathbf{k}^{\mathbf{u}}_{\leq}$ equals 0 for $i \notin v$. Since $W(0)(x) = 1$ and $\mathrm{wal}_0(x) = 1$ for any $x \in [0,1)$, we have for $\mathbf{x} = (x_i)_{i=1}^s \in [0,1)^s$,

$$W(\mathbf{k}^{\mathbf{u}}_{\leq})(\mathbf{x}) = \prod_{i=1}^{s} W(k^{u_i}_{i,\leq})(x_i) = \prod_{i \in v} W(k^{u_i}_{i,\leq})(x_i) = W(\mathbf{k}^{\mathbf{u}}_{v\leq})(\mathbf{x}_v)$$

and

$$\mathrm{wal}_{\mathbf{k}^{\mathbf{u}}_{\leq}}(\mathbf{x}) = \prod_{i=1}^{s} \mathrm{wal}_{k^{u_i}_{i,\leq}}(x_i) = \prod_{i \in v} \mathrm{wal}_{k^{u_i}_{i,\leq}}(x_i) = \mathrm{wal}_{\mathbf{k}^{\mathbf{u}}_{v\leq}}(\mathbf{x}_v),$$

where we write $\mathbf{x}_v = (x_i)_{i \in v}$. Combining those facts with Lemma 6, we have

$$|\hat{f}(\mathbf{k})| \leq 2^{-\mu'_{\mathbf{u}}(\mathbf{k})} \left| \int_{[0,1)^s} f^{(\min(\mathbf{u},\mathbf{N}_{\mathbf{k}}))}(\mathbf{x}) \cdot W(\mathbf{k}^{\mathbf{u}}_{\leq})(\mathbf{x}) \cdot \mathrm{wal}_{\mathbf{k}^{\mathbf{u}}_{>}}(\mathbf{x}) d\mathbf{x} \right|$$
$$= 2^{-\mu'_{\mathbf{u}}(\mathbf{k})} \left| \int_{[0,1)^{|v|}} \left(\int_{[0,1)^{|S\setminus v|}} f^{(\min(\mathbf{u},\mathbf{N}_{\mathbf{k}}))}(\mathbf{x}) d\mathbf{x}_{S\setminus v} \right) \right.$$
$$\left. \cdot W(\mathbf{k}^{\mathbf{u}}_{v\leq})(\mathbf{x}_v) \cdot \mathrm{wal}_{\mathbf{k}^{\mathbf{u}}_{v>}}(\mathbf{x}_v) d\mathbf{x}_v \right|.$$

By the definition of the Walsh functions, it holds that $|\mathrm{wal}_{\mathbf{k}^{\mathbf{u}}_{v>}}(\mathbf{x}_v)| = 1$ for any $\mathbf{x}_v \in [0,1)^{|v|}$. Then we have

$$|\hat{f}(\mathbf{k})| \leq 2^{-\mu'_{\mathbf{u}}(\mathbf{k})} \int_{[0,1)^{|v|}} \left| \int_{[0,1)^{|S\setminus v|}} f^{(\min(\mathbf{u},\mathbf{N}_{\mathbf{k}}))}(\mathbf{x}) d\mathbf{x}_{S\setminus v} \right| \cdot |W(\mathbf{k}^{\mathbf{u}}_{v\leq})(\mathbf{x}_v)| d\mathbf{x}_v$$
$$\leq 2^{-\mu'_{\mathbf{u}}(\mathbf{k})} \|f^{(\min(\mathbf{u},\mathbf{N}_{\mathbf{k}_v}))}\|_p \cdot \|W(\mathbf{k}^{\mathbf{u}}_{v\leq})\|_{L^{p/(p-1)}},$$

where we used Hölder's inequality in the second inequality. By Lemma 4, we have

$$\|W(\mathbf{k}^{\mathbf{u}}_{v\leq})\|_{L^{p/(p-1)}} = \prod_{i\in v} \|W(k^{u_i}_{i\leq})\|_{L^{p/(p-1)}} \leq 2^{|v|/p}.$$

Thus we obtain the result. □

As we mention in Definition 3, when $u_i = \alpha \in \mathbb{N} \cup \{\infty\}$ for every i, $\mu'_{\mathbf{u}}(\mathbf{k}) = \mu'_{\mathbf{u}}(\mathbf{k}_v; \mathbf{0}) = \mu'_{\alpha}(\mathbf{k}_v)$. Then Theorem 2 follows from this lemma as a corollary.

In the following section, we will prove Lemma 5 and 6. From now, we denote by $\prod_{i=1}^n \varphi_i$ the composition of operators $\varphi_1 \circ \cdots \circ \varphi_n$.

4. Formulas on the Walsh coefficients

Here we prove the formulas on the Walsh coefficients $\hat{f}(\mathbf{k})$ in Lemmas 5 and 6. In order to proceed with the proof, we introduce the following symbols:

DEFINITION 6. We use the same notation as in Definition 2. Let $\mathbf{p} = (p_i)_{i\in v}$, $\mathbf{q} = (q_i)_{i\in v} \in \mathbb{N}^{|v|}$ with $1 \leq p_i \leq q_i \leq N_i$. For given $\mathbf{k} \in \mathbb{N}_0^s$, we use the following symbols in the proofs below

$$\mathbf{d}^{\mathbf{q}}_{\mathbf{p}} := \prod_{i\in v, p_i \leq j \leq q_i} \partial_{i, a_{i,j}+1},$$

where the product means the composition of operators. Notice that, for a function f, $\mathbf{d}_{\mathbf{k},\mathbf{u}}f$ is written as $\mathbf{d}_{\mathbf{p}}^{\mathbf{q}}f$ where $\mathbf{p} = (1, \dots, 1)$ and $\mathbf{q} = (\min(N_i, u_i))_{i \in v}$.

4.1. Formula relating the Walsh coefficients to the dyadic differences. We prove the formula in Lemma 5, which represents the relationship between the Walsh coefficients $\hat{f}(\mathbf{k})$ and the dyadic differences $\partial_{i,n} f$.

Proof. We prove only the case $s = 1$ here. In the case $s > 1$, we obtain the result by applying the same method in a component-wise fashion.

We easily obtain the first statement as follows. Let $k_1 = \sum_{j=1}^{N_1} 2^{a_j}$. Since $\partial_{1,a_j+1} f$ is the sum of $f(\cdot \oplus 2^{-a_j-1}) \in L^1([0,1))$ and $f \in L^1([0,1))$, we have $\partial_{1,a_j+1} f \in L^1([0,1))$. By repeating this argument, we have $\mathbf{d}_{k_1,u_1} f \in L^1([0,1))$.

We show the second statement inductively. We omit the case $k_1 = 0$ or $u_1 = 0$ since the proof is easy. We show the case $u_1 = 1$:

$$\hat{f}(k_1) = (-1) \cdot 2^{-a_j-2} \cdot \widehat{\partial_{1,a_j+1} f}(k_1), \tag{12}$$

where $k_1 = \sum_{j=1}^{N_1} 2^{a_j}$. By changing variables $x_1 \mapsto x_1 \oplus 2^{-a_j-1}$, we have

$$\begin{aligned}
&\int_0^1 f(x_1 \oplus 2^{-a_j-1}) \cdot \mathrm{wal}_{k_1}(x_1) dx_1 \\
&\quad = \sum_{c=0}^{2^{a_j}-1} \left(\int_{2^{-a_j-1}\cdot 2c}^{2^{-a_j-1}\cdot(2c+1)} f(x_1 + 2^{-a_j-1}) \cdot \mathrm{wal}_{k_1}(x_1) dx_1 \right. \\
&\qquad \left. + \int_{2^{-a_j-1}\cdot(2c+1)}^{2^{-a_j-1}\cdot(2c+2)} f(x_1 - 2^{-a_j-1}) \cdot \mathrm{wal}_{k_1}(x_1) dx_1 \right) \\
&\quad = \sum_{c=0}^{2^{a_j}-1} \left(\int_{2^{-a_j-1}\cdot(2c+1)}^{2^{-a_j-1}\cdot(2c+2)} f(x_1) \cdot \mathrm{wal}_{k_1}(x_1 - 2^{-a_j-1}) dx_1 \right. \\
&\qquad \left. + \int_{2^{-a_j-1}\cdot 2c}^{2^{-a_j-1}\cdot(2c+1)} f(x_1) \cdot \mathrm{wal}_{k_1}(x_1 + 2^{-a_j-1}) dx_1 \right) \\
&\quad = \int_0^1 f(x_1) \cdot \mathrm{wal}_{k_1}(x_1 \oplus 2^{-a_j-1}) dx_1 \\
&\quad = \int_0^1 f(x_1) \cdot \mathrm{wal}_{k_1}(x_1) \cdot \mathrm{wal}_{k_1}(2^{-a_j-1}) dx_1 \\
&\quad = -\int_0^1 f(x_1) \cdot \mathrm{wal}_{k_1}(x_1) dx_1,
\end{aligned}$$

where the last three identities follow from the definition of the Walsh functions. Using this calculation, we obtain

$$\widehat{\partial_{1,a_j+1} f}(k_1) = -2 \cdot 2^{a_j+1} \cdot \int_0^1 f(x_1) \cdot \mathrm{wal}_{k_1}(x_1) dx_1 = (-1) \cdot 2^{a_j+2} \cdot \hat{f}(k_1).$$

We write $U = \min(u_1, N_1)$. Using (12) inductively, we obtain

$$\begin{aligned}\widehat{\mathbf{d}_{k_1,u_1} f}(k_1) &= \widehat{\mathbf{d}_1^U f}(k_1) = (-1) \cdot 2^{a_1+2} \cdot \widehat{\mathbf{d}_2^U f}(k_1) \\ &= (-1)^2 \cdot 2^{\sum_{j=1}^{2}(a_j+2)} \cdot \widehat{\mathbf{d}_3^U f}(k_1) = \cdots = (-1)^U \cdot 2^{\mu'_{u_1}(k_1)} \cdot \hat{f}(k_1),\end{aligned}$$

which is the result. Notice that $\mathbf{d}_p^q$ for $1 \le p \le q$ is one dimensional version introduced in Definition 6. □

4.2. Formula relating the Walsh coefficients to derivatives. In this subsection we prove the formula in Lemma 6, which represents the relationship between the Walsh coefficients $\hat{f}(\mathbf{k})$ and the derivatives $\frac{\partial f}{\partial x_i}$. This is done by unveiling the relationship between derivatives and dyadic differences.

To proceed with the following proofs, we define some symbols.

DEFINITION 7. We use the same symbols as in Definition 2 and 6. We define the following functions

$$w_{i,n}(x_1, \dots, x_s) := \mathrm{wal}_{2^{n-1}}(x_i), \qquad (x_1, \dots, x_s) \in [0,1)^s,$$

$$\mathbf{w}_{\mathbf{p}}^{\mathbf{q}}(x_1, \dots, x_s) := \prod_{i \in v, p_i \le j \le q_i} w_{i,a_{i,j}+1}(x_1, \dots, x_s), \qquad (x_1, \dots, x_s) \in [0,1)^s,$$

where the product means the multiplication of function values. Notice that, using this notation, we can rewrite the $\mathbf{k}$-th Walsh function as follows: $\mathrm{wal}_{\mathbf{k}} = \mathbf{w}_{(1,\dots,1)}^{(N_i)_{i \in v}}$.

Further we define the following operators

$$w\partial_{i,n}(g) := w_{i,n} \cdot \partial_{i,n}(g) \qquad \text{for } g : [0,1)^s \to \mathbb{R},$$

$$\mathbf{wd}_{\mathbf{p}}^{\mathbf{q}} := \prod_{i \in v, p_i \le j \le q_i} w\partial_{i,a_{i,j}+1},$$

where $w_{i,n} \cdot \partial_{i,n}(g)$ means the product as a function and the product symbol in the second line means the composition of operators.

4.2.1. *Important properties of dyadic differences.* In order to prove Lemma 6, we show some properties of dyadic differences $\partial_{i,n}$. We first show that the function $w_{j,m}$ and the operator $\partial_{i,n}$ commute.

Lemma 8. *When* $(i,n) \neq (j,m) \in \mathbb{N}^2$, *for a function* $g : [0,1)^s \to \mathbb{R}$, *we have the following identity as a function on* $[0,1)^s$*:*

$$w_{j,m} \cdot \partial_{i,n}(g) = \partial_{i,n}(g \cdot w_{j,m}).$$

Proof. We omit the proof here since it is easy. □

Next we prove that $\mathbf{wd}_{\mathbf{p}}^{\mathbf{q}}$ is an operator on $L^1([0,1)^s)$.

Lemma 9. *We use the same symbols as in the above definition. For a function* $g \in L^1([0,1)^s)$, *we have* $\mathbf{wd}_{\mathbf{p}}^{\mathbf{q}} g \in L^1([0,1)^s)$.

Proof. By the above lemma, we have that $\mathbf{wd}_{\mathbf{p}}^{\mathbf{q}} g$ equals $\mathbf{w}_{\mathbf{p}}^{\mathbf{q}} \cdot \mathbf{d}_{\mathbf{p}}^{\mathbf{q}} g$ as a function on $[0,1)^s$. By the definition of the Walsh functions, we see $|\mathbf{w}_{\mathbf{p}}^{\mathbf{q}}(\mathbf{x})| = 1$ for $\mathbf{x} \in [0,1)^s$. Since $\mathbf{d}_{\mathbf{p}}^{\mathbf{q}} g$ is the sum of the functions in $L^1([0,1)^s)$ as in the proof of Lemma 5, we have $\mathbf{d}_{\mathbf{p}}^{\mathbf{q}} g \in L^1([0,1)^s)$. Thus the product $\mathbf{wd}_{\mathbf{p}}^{\mathbf{q}} g (= \mathbf{w}_{\mathbf{p}}^{\mathbf{q}} \cdot \mathbf{d}_{\mathbf{p}}^{\mathbf{q}} g)$ is in $L^1([0,1)^s)$. □

The following Lemma is the key to proving Lemma 6, which replaces $w\partial_{i,n}(g)$ with the derivative $\frac{\partial g}{\partial x_i}$.

Lemma 10. *Let* $n, s, i \in \mathbb{N}$ *satisfy* $s \geq i$. *Let* $g : [0,1]^s \to \mathbb{R}$, *as a function of the ith component* x_i, *satisfy*

$$g \in C^1([2^{-n+1}c, 2^{-n+1}(c+1))), \qquad c = 0, \ldots, 2^{n-1} - 1. \tag{13}$$

Then for any $\mathbf{z} = (z_1, \ldots, z_s) \in [0,1)^s$, *we have*

$$(w\partial_{i,n}(g))(\mathbf{z}) = \int_0^1 \frac{\partial g}{\partial x_i}(z_1, \ldots, z_{i-1}, y, z_{i+1}, \ldots, z_s) \cdot \chi_n(z_i, y) dy.$$

Notice that, since $\frac{\partial g}{\partial x_i}$ *is defined on the support of* $\chi_n(z_i, y)$, *this integrand is defined for* $y \in [0,1)$.

Proof. Let $c' \in \mathbb{N}_0$ satisfy $z_i \in [2^{-n}c', 2^{-n}(c'+1))$. We consider the two cases: $c' = 2c$ and $c' = 2c+1$ for some integer c. We only calculate the case $c' = 2c$ since the other case can be calculated in the same way. In this case, by the fact $w_{i,n}(\mathbf{z}) = 1$ for $z_i \in [2^{-n}c', 2^{-n}(c'+1))$ and Assumption (13), we have

$$\begin{aligned}
(w\partial_{i,n}(g))(\mathbf{z}) &= w_{i,n}(\mathbf{z}) \cdot (\partial_{i,n}(g))(\mathbf{z}) \\
&= 1 \cdot 2^n \cdot (g(z_1, \ldots, z_i + 2^{-n}, \ldots, z_s) - g(z_1, \ldots, z_s)) \\
&= \int_{z_i}^{z_i + 2^{-n}} 2^n \cdot \frac{\partial g}{\partial x_i}(z_1, \ldots, z_{i-1}, y, z_{i+1}, \ldots, z_s) dy \\
&= \int_0^1 \frac{\partial g}{\partial x_i}(z_1, \ldots, z_{i-1}, y, z_{i+1}, \ldots, z_s) \cdot \chi_n(z_i, y) dy.
\end{aligned}$$

The last equality follows from $[z_i, z_i + 2^{-n}] = [\min(z_i, z_i \oplus 2^{-n}), \max(z_i, z_i \oplus 2^{-n})]$ and the definition of χ_n. $\square$

4.2.2. *Proof of Lemma* 6. Now we prove Lemma 6 using the lemmas above and Lemma 5.

PROOF. We prove the case $s = 1$. We omit the case $k_1 = 0$ or $u_1 = 0$ since the proof is easy. We write $k_1 = \sum_{j=1}^{N} 2^{a_j}$ and $U = \min(u_1, N)$ here. We assume that $u_1 \geq 1$, then we calculate $\widehat{\mathbf{d}_{k_1,u_1} f}(k_1)$ as follows:

$$\widehat{\mathbf{d}_{k_1,u_1} f}(k_1) = \widehat{\mathbf{d}_1^U f}(k_1) = \int_0^1 (\mathbf{d}_1^U f)(x_1) \cdot \mathbf{w}_1^N(x_1) dx_1$$
$$= \int_0^1 \mathbf{wd}_1^U (f \cdot \mathbf{w}_{U+1}^N)(x_1) dx_1,$$

where we used Lemma 8 in the third equality. Notice that $\mathbf{d}_p^q$ is one dimensional version defined in Definition 6 and $\mathbf{w}_p^q$, $\mathbf{wd}_p^q$ are one dimensional versions introduced in Definition 7, where $1 \leq p \leq q$. Using the assumption on f and the definition of the Walsh functions, we have that

$$f \cdot \mathbf{w}_{U+1}^N \in C^{u_1}([2^{-a_{U+1}-1} c, 2^{-a_{U+1}-1}(c+1))),$$

and we have

$$\frac{d}{dx_1}(f \cdot \mathbf{w}_{U+1}^N) = \left(\frac{df}{dx_1} \cdot \mathbf{w}_{U+1}^N\right) \qquad \text{on } [2^{-a_{U+1}-1} c, 2^{-a_{U+1}-1}(c+1)),$$

with $0 \leq c \leq 2^{a_{U+1}+1} - 1$.

Let $n \geq 1$ and $0 \leq c' < 2^n$ be integers. By the definition of $w\partial_{1,n}$, we have that, if $g \in C^1([c'2^{-n}, (c'+1)2^{-n}))$, it holds that $w\partial_{1,n} g \in C^1([c'2^{-n}, (c'+1)2^{-n}))$ and $\frac{d}{dx_1}(w\partial_{1,n} g) = w\partial_{1,n}\left(\frac{dg}{dx_1}\right)$ on $[c'2^{-n}, (c'+1)2^{-n})$.

Since $2^{-a_{U+1}} > 2^{-a_U}$, if we take $g = f \cdot \mathbf{w}_{U+1}^N$ and $n = a_U + 1$ here, we have

$$\frac{d}{dx_1}(w\partial_{1,a_U+1}(f \cdot \mathbf{w}_{U+1}^N)) = w\partial_{1,a_U+1}\left(\frac{df}{dx_1} \cdot \mathbf{w}_{U+1}^N\right)$$
$$\text{on } [c2^{-a_U-1}, (c+1)2^{-a_U-1}),$$

where $0 \leq c \leq 2^{a_U+1} - 1$. Applying this argument inductively, we have

$$\frac{d}{dx_1}(\mathbf{wd}_2^U(f \cdot \mathbf{w}_{U+1}^N)) = \mathbf{wd}_2^U\left(\frac{df}{dx_1} \cdot \mathbf{w}_{U+1}^N\right) \qquad \text{on } [c2^{-a_2-1}, (c+1)2^{-a_2-1}),$$

where $0 \leq c \leq 2^{a_2+1} - 1$. Since $2^{-a_2-1} \geq 2^{-a_1}$, we can take $n = a_1 + 1$ and $g = \mathbf{wd}_2^U(f \cdot \mathbf{w}_{U+1}^N)$ in Lemma 10. Then we continue the computation of

$\widehat{\mathbf{d}_{k_1,u_1}f}(k_1)$ as follows:

$$\begin{aligned}\widehat{\mathbf{d}_{k_1,u_1}f}(k_1) &= \int_0^1 w\partial_{1,a_1+1}(\mathbf{wd}_2^U(f\cdot \mathbf{w}_{U+1}^N))(x_1)dx_1\\ &= \int_0^1\left(\int_0^1 \chi_{a_1+1}(x_1,y)\cdot \mathbf{wd}_2^U\left(\frac{df}{dy}\cdot \mathbf{w}_{U+1}^N\right)(y)dy\right)dx_1.\end{aligned}$$

By the definition of the Walsh functions, we have that $\mathbf{w}_{U+1}^N \in L^1([0,1))$. Using this fact and the assumption on f, we have $\frac{df}{dx_1}\cdot \mathbf{w}_{U+1}^N \in L^1([0,1))$. Therefore if we take $g = \frac{df}{dx_1}\cdot \mathbf{w}_{U+1}^N$ in Lemma 9 and consider the fact that $|\chi_{a_1+1}(x_1,y)| \le 2^{a_1+1}$ for any $x_1, y \in [0,1)$, we see that the integrand $\chi_{a_1+1}\cdot \mathbf{wd}_2^U\left(\frac{df}{dx_1}\cdot \mathbf{w}_{U+1}^N\right)$ in the last line is in $L^1([0,1)^2)$. Thus we can use Fubini's Theorem as follows:

$$\begin{aligned}\widehat{\mathbf{d}_{k_1,u_1}f}(k_1) &= \int_0^1\left(\int_0^1 \chi_{a_1+1}(x_1,y)dx_1\right)\cdot \mathbf{wd}_2^U\left(\frac{df}{dy}\cdot \mathbf{w}_{U+1}^N\right)(y)dy\\ &= \int_0^1 \mathbf{wd}_2^U\left(\frac{df}{dy}\cdot \mathbf{w}_{U+1}^N\right)(y)\cdot W(2^{a_1})(y)dy\\ &= \int_0^1 \mathbf{wd}_2^U\left(\frac{df}{dx_1}\cdot \mathbf{w}_{U+1}^N\right)(x_1)\cdot W(2^{a_1})(x_1)dx_1.\end{aligned}$$

By repeating the argument we have

$$\begin{aligned}\widehat{\mathbf{d}_{k_1,u_1}f}(k_1) &= \int_0^1 w\partial_{1,a_2+1}\left(\mathbf{wd}_3^U\left(\frac{df}{dx_1}\cdot \mathbf{w}_{U+1}^N\right)\right)(x_1)\cdot W(2^{a_1})(x_1)dx_1\\ &= \int_0^1\left(\int_0^1 \chi_{a_2+1}(x_1,y)\cdot \mathbf{wd}_3^U\left(\frac{d^2f}{d^2y}\cdot \mathbf{w}_{U+1}^N\right)(y)dy\right)\cdot W(2^{a_1})(x_1)dx_1\\ &= \int_0^1\left(\int_0^1 \chi_{a_2+1}(x_1,y)\cdot W(2^{a_1})(x_1)dx_1\right)\cdot \mathbf{wd}_3^U\left(\frac{d^2f}{dy^2}\cdot \mathbf{w}_{U+1}^N\right)(y)dy\\ &= \int_0^1 \mathbf{wd}_3^U\left(\frac{d^2f}{dy^2}\cdot \mathbf{w}_{U+1}^N\right)(y)\cdot W(2^{a_1}+2^{a_2})(y)dy\\ &= \int_0^1 \mathbf{wd}_3^U\left(\frac{d^2f}{dx_1^2}\cdot \mathbf{w}_{U+1}^N\right)(x_1)\cdot W(2^{a_1}+2^{a_2})(x_1)dx_1\\ &= \cdots\\ &= \int_0^1\left(\frac{d^Uf}{dx_1^U}\cdot \mathbf{w}_{U+1}^N\right)(x_1)\cdot W(k_{1,\le}^{u_1})(x_1)dx_1.\end{aligned}$$

Combining Lemma 5 we have the result for $s = 1$. By calculating in a component-wise manner, we have the result for the case $s \geq 1$. We omit that case here. $\square$

In fact, we can determine the sign of $\hat{f}(\mathbf{k})$ for a special case.

COROLLARY 1. *Let $f \in C^\infty([0,1]^s)$ and $\mathbf{k} \in \mathbb{N}_0^s$. We use the symbol N_i appearing in Definition 2. Then if $f^{(N_1,\dots,N_s)}(\mathbf{x}) \geq 0$ for every $\mathbf{x} \in [0,1)^s$, we have $\hat{f}(\mathbf{k}) \cdot (-1)^{\sum_{i=1}^s N_i} \geq 0$.*

PROOF. By Lemma 4 and the fact $W(0)(x) = 1$ for any $x \in [0,1)$, we have that $W(\mathbf{k})(\mathbf{x}) = \prod_{i=1}^s W(k_i)(x_i) \geq 0$ for any $\mathbf{x} = (x_i)_{i=1}^s \in [0,1)^s$. By combining this fact and the assumption on $f^{(N_1,\dots,N_s)}$, we have that the product $f^{(N_1,\dots,N_s)} \cdot W(\mathbf{k})$ is positive over $[0,1)^s$. Thus, by Lemma 6 with $u_i = \infty$ we have

$$\hat{f}(\mathbf{k}) \cdot (-1)^{\sum_{i=1}^s N_i} = 2^{-\mu'_\infty(\mathbf{k})} \cdot \int_{[0,1)^s} f^{(N_1,\dots,N_s)}(\mathbf{x}) \cdot W(\mathbf{k})(\mathbf{x}) d\mathbf{x} \geq 0,$$

which is the result. $\square$

5. Koksma-Hlawka inequality for smooth functions

In this section, we prove the Koksma-Hlawka type inequalities for smooth functions in Theorem 1, which show the bounds on the QMC integration error. We first show that under the assumptions in Theorem 1 we have $f(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{N}_0^s} \hat{f}(\mathbf{k}) \operatorname{wal}_{\mathbf{k}}(\mathbf{x})$ for every $\mathbf{x} \in [0,1)^s$. To do so, we use [14, Lemma 18]. This lemma states that if f is continuous on $[0,1]^s$ and $\sum_{\mathbf{k} \in \mathbb{N}_0^s} |\hat{f}(\mathbf{k})| < \infty$ then

$$f(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{N}_0^s} \hat{f}(\mathbf{k}) \operatorname{wal}_{\mathbf{k}}(\mathbf{x}) \tag{14}$$

for every $\mathbf{x} \in [0,1)^s$. We now show that f is continuous on $[0,1]^s$ and $\sum_{\mathbf{k} \in \mathbb{N}_0^s} |\hat{f}(\mathbf{k})| < \infty$.

The condition that f is continuous on $[0,1]^s$ follows from the assumption on f. We confirm that $\sum_{\mathbf{k} \in \mathbb{N}_0^s} |\hat{f}(\mathbf{k})| < \infty$ in the following way. If we apply Theorem 2 for $\alpha = 2$, we have

$$\begin{aligned}\sum_{\mathbf{k} \in \mathbb{N}_0^s \setminus \{\mathbf{0}\}} |\hat{f}(\mathbf{k})| &= \sum_{\emptyset \neq v \subset S} \sum_{\mathbf{k}_v \in \mathbb{N}^{|v|}} |\hat{f}(\mathbf{k}_v; \mathbf{0})| \\ &\leq \sum_{\emptyset \neq v \subset S} \sum_{\mathbf{k}_v \in \mathbb{N}^{|v|}} 2^{|v|/p} 2^{-\mu'_2(\mathbf{k}_v)} \|f^{(\min(2, \mathbf{N}_{\mathbf{k}_v}))}\|_p\end{aligned}$$

$$\le \sum_{\emptyset \neq v \subset S} \sum_{\mathbf{k}_v \in \mathbb{N}^{|v|}} 2^{|v|/p} 2^{-\mu_2'(\mathbf{k}_v)} \max_{\mathbf{n} \in \{1,2\}^{|v|}} \|f^{(\mathbf{n})}\|_p$$

$$\le 2^{s/p} \max_{\mathbf{n}' \in \{0,1,2\}^s} \|f^{(\mathbf{n}')}\|_{L^p} \sum_{\emptyset \neq v \subset S} \sum_{\mathbf{k}_v \in \mathbb{N}^{|v|}} 2^{-\mu_2'(\mathbf{k}_v)}.$$

Note that the L^p-norm $\|f\|_{L^p} := \left(\int_{[0,1)^s} |f(x)|^p dx\right)^{1/p}$ is different from the norm $\|f\|_p$ defined in Section 1. Thus we have

$$\sum_{\mathbf{k} \in \mathbb{N}_0^s} |\hat{f}(\mathbf{k})| = |\hat{f}(\mathbf{0})| + \sum_{\mathbf{k} \in \mathbb{N}_0^s \setminus \{\mathbf{0}\}} |\hat{f}(\mathbf{k})|$$

$$\le 2^{s/p} \max_{\mathbf{n}' \in \{0,1,2\}^s} \|f^{(\mathbf{n}')}\|_{L^p} \left(1 + \sum_{\emptyset \neq v \subset S} \sum_{\mathbf{k}_v \in \mathbb{N}^{|v|}} 2^{-\mu_2'(\mathbf{k}_v)}\right).$$

Since $\max_{\mathbf{n}' \in \{0,1,2\}^s} \|f^{(\mathbf{n}')}\|_{L^p} < \infty$ holds by the assumption on f, we have only to show the last summation is finite. We prove this in the following way:

$$1 + \sum_{\emptyset \neq v \subset S} \sum_{\mathbf{k}_v \in \mathbb{N}^{|v|}} 2^{-\mu_2'(\mathbf{k}_v)} = \left(1 + \sum_{k \in \mathbb{N}} 2^{-\mu_2'(k)}\right)^s$$

$$= \left(1 + \sum_{l \in \mathbb{N}_0} 2^{-l-2} + \sum_{l_1, l_2 \in \mathbb{N}_0, l_1 < l_2} \sum_{k \in \mathbb{N}_0, k < 2^{l_1}} 2^{-l_1 - l_2 - 4}\right)^s$$

$$= \left(\frac{3}{2} + \sum_{l_2 \in \mathbb{N}_0} l_2 2^{-l_2 - 4}\right)^s \le \left(\frac{3}{2} + 2^{-4} \cdot \sum_{l_2 \in \mathbb{N}_0} \left(\frac{3}{4}\right)^{l_2}\right)^s < \infty.$$

Thus we apply Formula (14) to f to get

$$|\mathrm{Err}(f;\mathcal{P})| = \left| \int_{[0,1)^s} f(\mathbf{x}) d\mathbf{x} - \frac{1}{|\mathcal{P}|} \sum_{\mathbf{x} \in \mathcal{P}} f(\mathbf{x}) \right| = \left| \hat{f}(\mathbf{0}) - \frac{1}{|\mathcal{P}|} \sum_{\mathbf{x} \in \mathcal{P}} \sum_{\mathbf{k} \in \mathbb{N}_0^s} \hat{f}(\mathbf{k}) \,\mathrm{wal}_{\mathbf{k}}(\mathbf{x}) \right|.$$

Now we introduce the character property of the Walsh functions. Let $\mathcal{P}$ be a digital net in $[0,1)^s$ where $|\mathcal{P}| = 2^m$. Then we have (see [10, Lemma 4.75])

$$\sum_{\mathbf{x} \in \mathcal{P}} \mathrm{wal}_{\mathbf{k}}(\mathbf{x}) = \begin{cases} 2^m & \text{if } \mathbf{k} \in \mathcal{P}^\perp, \\ 0 & \text{otherwise.} \end{cases}$$

Using this fact, we have

$$|\mathrm{Err}(f;\mathcal{P})| = \left| \hat{f}(\mathbf{0}) - \sum_{\mathbf{x} \in \mathcal{P}^\perp} \hat{f}(\mathbf{k}) \right| \le \sum_{\mathbf{x} \in \mathcal{P}^\perp \setminus \{\mathbf{0}\}} |\hat{f}(\mathbf{k})| = \sum_{\emptyset \neq v \subset S} \sum_{\mathbf{k}_v \in \mathcal{P}_v^\perp} |\hat{f}(\mathbf{k}_v; \mathbf{0})|.$$

Let $\alpha \in \mathbb{N} \cup \{\infty\}$ with $\alpha \geq 2$ and $1 \leq p, q, q' \leq \infty$ such that $1/q + 1/q' = 1$. Applying Theorem 2 to f, we have

$$\begin{aligned}|\mathrm{Err}(f;\mathscr{P})| &\leq \sum_{v} \sum_{\mathbf{k}_v \in \mathscr{P}_v^{\perp}} 2^{|v|/p} 2^{-\mu'_{\alpha}(\mathbf{k}_v)} \|f^{(\min(\alpha, \mathbf{N}_{\mathbf{k}_v}))}\|_p \\ &\leq \sum_{v} \left(\gamma_v^{-1} 2^{|v|/p} \sup_{\alpha_v \in \{1,\dots,\alpha\}^{|v|}} \|f^{(\alpha_v)}\|_p \right) \cdot \left(\gamma_v \sum_{\mathbf{k}_v \in \mathscr{P}_v^{\perp}} 2^{-\mu'_{\alpha}(\mathbf{k}_v)} \right) \\ &\leq \|f\|_{\mathscr{B}_{\alpha,\gamma,p,q'}} \times \mathscr{W}_{\alpha,\gamma,q}(\mathscr{P}),\end{aligned}$$

where we used Hölder's inequality in the last inequality.

Acknowledgement

The author would like to thank the referee and Prof. Josef Dick for helpful comments.

References

[1] J. Baldeaux, J. Dick, J. Greslehner, and F. Pillichshammer, Construction algorithms for higher order polynomial lattice rules, J. Complexity, **27** (2011), 281–299.

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

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

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

[5] J. Dick, On quasi-Monte Carlo rules achieving higher order convergence, Monte Carlo and Quasi-Monte Carlo Methods 2008, Springer, (2009), 73–96.

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

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

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

[9] J. Dick and F. Pillichshammer, Multivariate integration in weighted Hilbert spaces based on Walsh functions and weighted Sobolev spaces, J. Complexity, **21** (2005), 149–195.

[10] J. Dick and F. Pillichshammer, Digital Nets and Sequences, Discrepancy Theory and Quasi-Monte Carlo Integration, Cambridge University Press, Cambridge, 2010.

[11] N. J. Fine, On the Walsh functions, Trans. Amer. Math. Soc., **65** (1949), 372–414.

[12] R. N. Gantner and Ch. Schwab, Computational Higher Order Quasi-Monte Carlo Integration, To appear in Proc. MCQMC 2014, Available at http://www.sam.math.ethz.ch/sam_reports/reports_final/reports2014/2014-25.pdf.

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

[14] T. Goda, K. Suzuki, and T. Yoshiki, The b-adic tent transformation for quasi-Monte Carlo integration using digital nets, J. Approx. Theory, **194** (2015), 62–86.

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

[16] S. Harase and R. Ohori, A search for extensible low-WAFOM point sets, arXiv:1309.7828v2.

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

[18] L. Kuipers and H. Niederreiter, Uniform Distribution of Sequences, John Wiley, New York, 1974. Reprint, Dover Publications, Mineola, NY, 2006.

[19] M. Matsumoto, M. Saito, and K. Matoba, A computable figure of merit for Quasi-Monte Carlo point sets, Math. Comp., **83** (2014), 1233–1250.

[20] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, CBMS-NSF, Philadelphia, Pennsylvania, 1992.

[21] H. Niederreiter and G. Pirsic, Duality for digital nets and its applications, Acta Arith., **97** (2001), 173–182.

[22] F. Schipp, W. R. Wade, and P. Simon, Walsh Series, An Introduction to Dyadic Harmonic Analysis, Adam Hilger Ltd., Bristol, 1990.

[23] I. F. Sharygin, A lower estimate for the error of quadrature formulas for certain classes of functions, Zh. Vychisl. Mat. i Mat. Fiz., **3** (1963), 370–376.

[24] I. H. Sloan and S. Joe, Lattice Methods for Multiple Integration, Clarendon Press, Oxford, 1994.

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

Takehito Yoshiki
School of Mathematics and Statistics
The University of New South Wales, Sydney 2052*, Australia*
E-mail: takehito.yoshiki1@unsw.edu.au