Convergence acceleration of modified Fourier series in one or more dimensions

Ben Adcock
DAMTP, Centre for Mathematical Sciences
University of Cambridge
Wilberforce Rd, Cambridge CB3 0WA
United Kingdom

November 25, 2009

Abstract

Modified Fourier series have recently been introduced as an adjustment of classical Fourier series for the approximation of nonperiodic functions defined on d-variate cubes. Such approximations offer a number of advantages, including uniform convergence. However, like Fourier series, the rate of convergence is typically slow.

In this paper we extend Eckhoff's method to the convergence acceleration of multivariate modified Fourier series. By suitable augmentation of the approximation basis we demonstrate how to increase the convergence rate to an arbitrary algebraic order. Moreover, we illustrate how numerical stability of the method can be improved by utilising appropriate auxiliary functions.

In the univariate setting it is known that Eckhoff's method exhibits an auto-correction phenomenon. We extend this result to the multivariate case. Finally, we demonstrate how a significant reduction in the number of approximation coefficients can be achieved by using a hyperbolic cross index set.

Introduction

The modified Fourier basis was introduced in [17, 18] as an adjustment of the Fourier basis for the approximation of smooth, nonperiodic functions defined on $\bar{\Omega}$, where Ω is the d-variate cube $(-1,1)^d$. In the univariate case, the Fourier sine function is replaced by $\sin(n-\frac{1}{2})\pi x$, yielding the basis

$$\{\cos n\pi x, n \in \mathbb{N}\} \cup \{\sin(n-\tfrac{1}{2})\pi x, n \in \mathbb{N}_+\}.$$

The multivariate extension is obtained by Cartesian products. The advantage of this basis is that the modified Fourier expansion of a sufficiently smooth function f converges uniformly on $\bar{\Omega}$. In particular, there is no Gibbs phenomenon near the boundary [1, 17, 28].

Unfortunately, the convergence rate of the modified Fourier approximation remains relatively slow. If N is a truncation parameter, the uniform error is $\mathcal{O}\left(N^{-1}\right)$ on $\bar{\Omega}$ and $\mathcal{O}\left(N^{-2}\right)$ inside compact subsets of Ω [1, 17, 28]. Much like the Fourier case, this is due to 'jumps' in certain derivatives of the function at the endpoints $x = \pm 1$ (in the univariate case) [28]. In the multivariate setting, similar analogues hold, although the jump conditions (otherwise referred to as derivative conditions) are more complicated to express [1, 16].

For univariate Fourier expansions, provided the values of these jumps are known, there is a effective tool to accelerate convergence: namely the polynomial subtraction device [20, 22]. This idea was first considered by Krylov [21], and has been widely studied since then (see [5, 19, 23] and references therein). Polynomial subtraction is readily adapted to modified Fourier series [17] and to the multivariate case [16, 26, 24, 27].

However, such jump values are unknown in general. In typical applications only the (modified) Fourier coefficients of a given function may be known, and, even if arbitrary pointwise values of the function can be calculated, approximation via finite differences is not recommended for this purpose [23].

As noted in [10], the previous lack of robust methods for the approximation of jump values is the central reason why the polynomial subtraction technique has not been more extensively utilised (see

also [23, p.101]). In this paper, to circumvent the aforementioned problem, we use Eckhoff's method for this task [8, 9, 10]. This approach is based on the well-known observation that the (modified) Fourier coefficients themselves contain sufficient information to reconstruct the jump values (see [8] for further references). Hence, such values can be approximated by suitably defined extrapolation techniques.

Eckhoff's method was originally presented for the univariate, Fourier case. Analysis of the rate of convergence was carried out in [4]. The extension to bivariate functions was developed, without analysis, in [26, 24, 27]. The aim of this paper is to extend Eckhoff's method to the modified Fourier expansion of a function defined on the d-variate cube, and to provide analysis therein. The central result we prove demonstrates that approximating the jumps in this manner (as opposed to using their exact values) does not cause the convergence rate to deteriorate.

In fact, using approximate jump values offers at least one significant advantage. It was observed in [25] and proved in the univariate, Fourier case in [29] that Eckhoff's method exhibits an auto-correction phenomenon inside the domain. In other words, the convergence rate of the approximation based on approximate jump values is much faster in compact subsets of Ω than the approximation based on the exact values. We provide an extension of this result to the multivariate, modified Fourier setting.

Polynomial subtraction and Eckhoff's method both rely on the construction of a smooth function to interpolate the jump values. In standard implementations [4, 10, 20, 22] such a function is constructed from a certain set of polynomials (the possibility of using other functions was suggested in [10]). Though this is the most convenient choice, it leads to poor numerical stability. In Section 1 we introduce a set of trigonometric functions that improves numerical stability.

Standard multivariate approximations using Fourier series involve $\mathcal{O}\left(N^d\right)$ terms. However, it transpires that this figure can be significantly reduced to $\mathcal{O}\left(N(\log N)^{d-1}\right)$ by using a so-called hyperbolic cross index set [3, 32]. The use of such an index set does not deteriorate the convergence rate, aside from possibly a logarithmic factor [32] (for application of this index set to modified Fourier expansions see [1, 16]). In the final part of this paper we demonstrate how to incorporate such an index set into Eckhoff's method. With the aid of numerical examples, we highlight the advantage of this combined approach, namely that we are able to produce accurate approximations of multivariate functions using relatively few terms.

There are numerous devices for convergence acceleration of (univariate) Fourier expansions, including filters [31], Gegenbauer reconstruction [12, 13] and Fourier continuation methods [7], to name but a few. Without doubt, certain methods are more suitable for different applications. However, there are a number of advantages to Eckhoff's approach that warrant detailing. First, as we demonstrate in this paper, the combination of the multivariate version of this technique and hyperbolic cross index sets facilitates the construction of accurate approximations comprising only a small number of terms. Second, Eckhoff's method can be incorporated into spectral approximations of boundary value problems (see [10] and references therein for hyperbolic problems and [1, 2] for applications of modified Fourier expansions to second order boundary value problems). Finally, Eckhoff's technique is not restricted to Cartesian product domains. In theory it can be developed on any domain for which suitable orthogonal expansions are known. For example, the modified Fourier basis is known explicitly on the equilateral and right isosceles triangles [15]. The construction of accurate representation of functions on such domains is typically difficult, and Eckhoff's method may provide an attractive alternative to existing polynomial-based methods. This is an area for future investigation.

The remainder of this paper is organised as follows. In Section 1 we introduce and analyse the univariate version of Eckhoff's method for modified Fourier expansions. We then demonstrate how superior numerical results can be obtained by using a particular subtraction basis. Section 2 is devoted to the development and analysis of Eckhoff's method for functions defined on d-variate cubes. In Section 3 we extend the result of [29] concerning the existence of an auto-correction phenomenon to the multivariate version of Eckhoff's method. Finally, in Section 4 we demonstrate, without analysis, how a significant reduction in the number of approximation coefficients can be achieved. Numerical examples are provided.

The main results of this paper, namely the proof of convergence in the multivariate case, the existence of the multivariate auto-correction phenomenon and the use of a particular subtraction basis to improve numerical stability, can be readily adapted to the Fourier setting (with a little care, such results can also be applied to general *Fourier-like* expansions). However, due to the faster convergence rate, we consider modified Fourier approximations throughout.

1 The univariate version of Eckhoff's method

1.1 Definitions and basic properties

Given a function $f \in \mathrm{L}^2(-1,1)$, where $\mathrm{L}^2(-1,1)$ is the space of square integrable functions on $(-1,1)$, and truncation parameter $N \geq 2$ we define the truncated modified Fourier expansion of f by

$$\mathcal{F}_N[f](x) = \frac{1}{2}\hat{f}_0^{[0]} + \sum_{n=1}^{N-1}\left\{\hat{f}_n^{[0]}\cos n\pi x + \hat{f}_n^{[1]}\sin(n-\tfrac{1}{2})\pi x\right\} = \sum_{i=0}^{1}\sum_{n=0}^{N-1}\hat{f}_n^{[i]}\phi_n^{[i]}(x), \quad x \in [-1,1].$$

Here $\phi_0^{[0]}(x) = \frac{1}{\sqrt{2}}$, $\phi_0^{[1]}(x) = 0$ and $\phi_n^{[0]}(x) = \cos n\pi x$, $\phi_n^{[1]}(x) = \sin(n-\frac{1}{2})\pi x$ otherwise, and

$$\hat{f}_n^{[i]} = \int_{-1}^{1} f(x)\phi_n^{[i]}(x)\,\mathrm{d}x, \quad i \in \{0,1\}, \quad n \in \mathbb{N}, \tag{1.1}$$

is the n^{th} modified Fourier cosine ($i = 0$) or sine ($i = 1$) coefficient of f. As demonstrated in [17, 28] this series converges uniformly to f on $[-1,1]$ under some mild regularity assumptions (see also Section 1.2). Indeed, the coefficients $\hat{f}_n^{[i]}$ are $\mathcal{O}\left(n^{-2}\right)$ for large n (in comparison to $\mathcal{O}\left(n^{-1}\right)$ in the Fourier case).

Provided $f \in \mathrm{H}^{2k}(-1,1)$, where $\mathrm{H}^{2k}(-1,1)$ is the $2k^{\text{th}}$ classical Sobolev space ($k \in \mathbb{N}_+$), simple integration by parts of the right hand side of (1.1) yields

$$\hat{f}_n^{[i]} = \sum_{r=0}^{k-1}\frac{(-1)^{n+i}}{(\mu_n^{[i]})^{r+1}}\mathcal{A}_r^{[i]}[f] + \frac{(-1)^k}{(\mu_n^{[i]})^k}\widehat{f^{(2k)}}_n^{[i]}, \quad i \in \{0,1\}, \quad n \in \mathbb{N}_+, \tag{1.2}$$

where $\mu_n^{[0]} = n^2\pi^2$, $\mu_n^{[1]} = (n-\frac{1}{2})^2\pi^2$ and

$$(-1)^r\mathcal{A}_r^{[i]}[f] = f^{(2r+1)}(1) + (-1)^{i+1}f^{(2r+1)}(-1), \quad i \in \{0,1\}, \quad r \in \mathbb{N}. \tag{1.3}$$

The values $\mathcal{A}_r^{[i]}[f]$ are the requisite jump values for modified Fourier expansions. We say that a function f satisfies the *first k derivative conditions* if the first k such values vanish:

$$f^{(2r+1)}(1) + (-1)^{i+1}f^{(2r+1)}(-1) = 0, \quad i \in \{0,1\}, \quad r = 0,\ldots,k-1.$$

Equivalently, the first k odd derivatives of f vanish at the endpoints $x = \pm 1$. In this case the coefficients $\hat{f}_n^{[i]} = \mathcal{O}\left(n^{-2k-2}\right)$ and faster convergence of the approximation $\mathcal{F}_N[f]$ is observed (see Section 1.2).

1.2 Polynomial subtraction

If the first k such jump values are non-zero we seek to interpolate them with a function g_k. Since the function $f - g_k$ satisfies the first k derivative conditions, the new approximation $\mathcal{F}_N[f-g_k] + g_k$ converges at a faster rate to f. This is the principle of the polynomial subtraction process [20, 22].

To find a suitable function g_k we first introduce (smooth) subtraction functions $p_0^{[i]},\ldots,p_{k-1}^{[i]}$, where $p_r^{[i]}$ is even (respectively odd) if $i = 0$ ($i = 1$), that satisfy the conditions

$$\mathcal{A}_r^{[i]}\left[p_s^{[i]}\right] = \delta_{r,s}, \quad r,s = 0,\ldots,k-1, \quad i \in \{0,1\}. \tag{1.4}$$

We say that $p_0^{[i]},\ldots,p_{k-1}^{[i]}$ are *Cardinal functions* for the first k derivative conditions. With this in hand, we define g_k as follows:

$$g_k(x) = \sum_{i=0}^{1}\sum_{r=0}^{k-1}\mathcal{A}_r^{[i]}[f]p_r^{[i]}(x), \quad x \in [-1,1]. \tag{1.5}$$

Construction of appropriate Cardinal functions is commonly achieved by taking linear combinations of standard (smooth) functions $q_0^{[i]},\ldots,q_{k-1}^{[i]}$. Such functions must be chosen so that the interpolation problem

$$\text{find} \quad \{a_s^{[i]} : i \in \{0,1\}, r = 0,\ldots,k-1\} : \quad \sum_{s=0}^{k-1}a_s^{[i]}\mathcal{A}_r^{[i]}\left[q_s^{[i]}\right] = b_r^{[i]}, \quad i \in \{0,1\}, \quad r = 0,\ldots,k-1, \tag{1.6}$$

has unique solution for all choices $b_r^{[i]} \in \mathbb{R}$. We call $\{q_r^{[i]} : i \in \{0,1\}, r = 0,\ldots,k-1\}$ a *subtraction basis*.

Usually the r^{th} Cardinal function $p_r^{[i]}$ is specified to be a polynomial of degree $2(r+1)-i$ [4, 10, 22], in which case $q_r^{[i]} = x^{2(r+1)-i}$ and we refer to $\{p_r^{[i]}\}$ as *Cardinal polynomials*. This explains the name 'polynomial subtraction'. However, as we shall demonstrate, a significant advantage is gained by allowing the more general form (an idea which was suggested in [10]).

For later use we mention the following subtraction basis of trigonometric functions:

$$q_r^{[0]}(x) = \cos(r+\tfrac{1}{2})\pi x, \quad q_r^{[1]}(x) = \sin(r+1)\pi x, \quad r = 0,\ldots,k-1. \tag{1.7}$$

It is readily demonstrated that the interpolation problem (1.6) has a unique solution in this case. The functions $q_r^{[i]}$ are dual to the modified Fourier basis functions in the sense that the derivative of $\phi_n^{[i]}$ is proportional to $q_{n-1}^{[1-i]}$. This property was exploited in [1, 2] to analyse modified Fourier expansions. In the sequel, we demonstrate a practical use of this dual basis in Eckhoff's method. As we shall observe, it offers a significant numerical advantage over subtraction bases consisting of polynomials.

If g_k is given by (1.5) we define

$$\mathcal{F}_{N,k}[f](x) = \mathcal{F}_N[f-g_k](x) + g_k(x), \quad x \in [-1,1], \tag{1.8}$$

as the k^{th} *polynomial subtraction* approximation of f (for convenience we interpret $\mathcal{F}_{N,0}[f]$ as $\mathcal{F}_N[f]$). Suppose that $\|\cdot\|_\infty$ is the uniform norm on some domain Ω and that $\|\cdot\|_q$ is the $\mathrm{H}^q(\Omega)$-norm. Concerning the error of this approximation, we quote, without proof, the following two lemmas, found in [28] and [1] respectively:

Lemma 1.1. *Suppose that $k \in \mathbb{N}$, $f \in \mathrm{H}^{2k+2}(-1,1)$ and that $\mathcal{F}_{N,k}[f]$ is given by (1.8) using exact jump values. Then $\|f^{(q)} - (\mathcal{F}_{N,k}[f])^{(q)}\|_\infty$ is $\mathcal{O}\left(N^{q-2k-1}\right)$ for $q = 0,\ldots,2k$. If, additionally, $f \in \mathrm{H}^{2k+3}(-1,1)$ then convergence rate of $(\mathcal{F}_{N,k}[f])^{(q)}$ to $f^{(q)}$ is $\mathcal{O}\left(N^{q-2k-2}\right)$ uniformly in compact subsets of $(-1,1)$.*

Note that the final condition in this lemma, namely that $f \in \mathrm{H}^{2k+3}(-1,1)$, can be relaxed to the condition that $f \in \mathrm{C}^{2k+2}[-1,1]$ and $f^{(2k+2)}$ has bounded variation [28]. Moreover, when $k = 0$ Lemma 1.1 also establishes the pointwise and uniform convergence rates of $\mathcal{F}_N[f]$ to f as set out in the Introduction.

Concerning the error in the standard Sobolev norms $\|\cdot\|_q$ we have:

Lemma 1.2. *Suppose that $f \in \mathrm{H}^{2k+2}(-1,1)$ and that $\mathcal{F}_{N,k}[f]$ is as in Lemma 1.1. Then $\|f - \mathcal{F}_{N,k}[f]\|_q$ is $\mathcal{O}(N^{q-2k-\frac{3}{2}})$ for $q = 0,\ldots,2k+1$.*

1.3 Eckhoff's method for approximation of jump values

Observe that, due to the definition of the Cardinal functions $p_r^{[i]}$, we may re-write (1.2) as

$$\hat{f}_n^{[i]} = \sum_{r=0}^{k-1} \widehat{p_r}_n^{[i]} \mathcal{A}_r^{[i]}[f] + \mathcal{O}\left(n^{-2k-2}\right), \tag{1.9}$$

where, for ease of notation, we write $\widehat{p_r}_n^{[i]}$ for the modified Fourier coefficient of $p_r^{[i]}$ corresponding to $\phi_n^{[i]}$. Note that, by construction, the coefficient corresponding to $\phi_n^{[1-i]}$ is zero. Due to uniform convergence of $\mathcal{F}_N[f]$ to f, we have

$$f(x) - \mathcal{F}_N[f](x) = \sum_{i=0}^{1}\sum_{r=0}^{k-1} \mathcal{A}_r^{[i]}[f]\left(p_r^{[i]}(x) - \mathcal{F}_N[p_r^{[i]}](x)\right) + \mathcal{O}\left(N^{-2k-1}\right), \quad x \in [-1,1].$$

Now suppose that the values $\mathcal{A}_r^{[i]}[f]$ are approximated by values $\bar{\mathcal{A}}_r^{[i]}[f]$ and that g_k is constructed as in (1.5) using these approximate values. Then, it follows from (1.8) and the above expression that

$$f(x) - \mathcal{F}_{N,k}[f](x) = \sum_{i=0}^{1}\sum_{r=0}^{k-1} \left(\mathcal{A}_r^{[i]}[f] - \bar{\mathcal{A}}_r^{[i]}[f]\right)\left(p_r^{[i]}(x) - \mathcal{F}_N[p_r^{[i]}](x)\right) + \mathcal{O}\left(N^{-2k-1}\right).$$

Now consider, for example, the uniform error. Since $\|p_r^{[i]} - \mathcal{F}_N[p_r^{[i]}]\|_\infty = \mathcal{O}\left(N^{-2r-1}\right)$, to obtain an $\mathcal{O}\left(N^{-2k-1}\right)$ uniform error we require that

$$\bar{\mathcal{A}}_r^{[i]}[f] = \mathcal{A}_r^{[i]}[f] + \mathcal{O}\left(N^{2(r-k)}\right), \quad r = 0,\ldots,k-1, \quad i \in \{0,1\}. \tag{1.10}$$

In other words, rather than using exact jump values, it suffices to use (sufficiently accurate) approximations instead. To achieve this prescribed accuracy we employ Eckhoff's method [8, 9, 10], which we now describe.

Eckhoff's method is based on (1.9). In essence we seek values $\bar{\mathcal{A}}_r^{[i]}[f]$ that approximately satisfy this relation. To do so, suppose that $N \leq m(0) < \ldots < m(k-1) \leq aN$, $m(r) \in \mathbb{N}$ are given values and that $a \geq 1$ is constant. Then we define $\bar{\mathcal{A}}_r^{[i]}[f]$ as the solution of the $2k \times 2k$ linear system

$$\sum_{s=0}^{k-1} \widehat{p}_{s\,m(r)}^{\,[i]} \bar{\mathcal{A}}_s^{[i]}[f] = \hat{f}_{m(r)}^{[i]}, \quad r = 0, \ldots, k-1, \quad i \in \{0,1\}. \tag{1.11}$$

This linear system decouples into two $k \times k$ linear systems corresponding to $i = 0$ and $i = 1$, which can be solved in parallel. Henceforth we write $V^{[i]}$ for the $k \times k$ matrix with $(r,s)^{\text{th}}$ entry $\widehat{p}_{s\,m(r)}^{\,[i]}$. Note that the choice of the values $m(r)$ is essentially arbitrary. However, particular choices lead to better numerical behaviour and the auto-correction phenomenon [29], as we shall see in the sequel.

Nonsingularity of the linear system (1.11) can be immediately guaranteed:

Theorem 1.3. *For sufficiently large N the linear system (1.11) is nonsingular. In particular, if $p_0^{[i]}, \ldots, p_{k-1}^{[i]}$ are Cardinal polynomials or arise from the subtraction basis (1.7), then (1.11) is nonsingular for all N.*

Proof. Suppose first that $P_0^{[i]}, \ldots, P_{k-1}^{[i]}$ (for clarity we use this notation) are Cardinal polynomials. Then, since $\widehat{P}_{s\,m(r)}^{\,[i]} = (-1)^{m(r)+i}(\mu_{m(r)}^{[i]})^{-s-1}$, $V^{[i]} = D^{[i]}\tilde{V}^{[i]}$, where $D^{[i]}$ is the diagonal matrix with entries $(-1)^{m(r)}(\mu_{m(r)}^{[i]})^{-1}$ and $\tilde{V}^{[i]}$ is the Vandermonde matrix with entries $(\mu_{m(r)}^{[i]})^{-s}$. Nonsingularity (for all N) now follows immediately.

Suppose now that $p_0^{[i]}, \ldots, p_{k-1}^{[i]}$ are arbitrary Cardinal functions. Then, since $p_r^{[i]} = P_r^{[i]} + (p_r^{[i]} - P_r^{[i]})$ we may write $V^{[i]} = W^{[i]} + (V^{[i]} - W^{[i]})$, where $W^{[i]}$ is the matrix with $(r,s)^{\text{th}}$ entry $\widehat{P}_{s\,m(r)}^{\,[i]}$. To prove the result, it now suffices to show that $(W^{[i]})^{-1}(V^{[i]} - W^{[i]}) = o(1)$. Note that the s^{th} column of $V^{[i]} - W^{[i]}$ has entries $\widehat{(p_s - P_s)}_{m(r)}^{\,[i]}$. Since $p_s^{[i]} - P_s^{[i]}$ obeys the first k derivative conditions, it can be shown that $(W^{[i]})^{-1}$ applied to this vector, which is just the vector of Eckhoff's approximation of the jump values of $p_s^{[i]} - P_s^{[i]}$, is $o(1)$ (see Theorem 1.4). Using this, we deduce the result.

Suppose now that the functions $q_r^{[i]}$ are given by (1.7). Then, due to (1.6), it suffices to prove non-singularity of the matrices with $(r,s)^{\text{th}}$ entries

$$\widehat{q}_{s\,m(r)}^{\,[0]} = \frac{2(-1)^{m(r)+s+1}(s+\frac{1}{2})}{\left\{m(r)^2 - (s+\frac{1}{2})^2\right\}\pi}, \quad \widehat{q}_{s\,m(r)}^{\,[1]} = \frac{2(-1)^{m(r)+s+1}s}{\left\{(m(r)-\frac{1}{2})^2 - s^2\right\}\pi}.$$

After appropriate pre-multiplication by non-singular diagonal matrices, we obtain matrices with $(r,s)^{\text{th}}$ entries

$$\left\{m(r)^2 - (s+\tfrac{1}{2})^2\right\}^{-1}, \quad \left\{(m(r)-\tfrac{1}{2})^2 - s^2\right\}^{-1},$$

respectively. These are Cauchy matrices: hence, nonsingularity follows immediately. □

With the values $\bar{\mathcal{A}}_r^{[i]}[f]$ given as the solution of (1.11) we refer to the resulting approximation $\mathcal{F}_{N,k}[f] = \mathcal{F}_N[f - g_k] + g_k$ as the k^{th} *Eckhoff approximation* of f.

The standard construction of Eckhoff's approximation [4, 10] uses the Cardinal functions $p_r^{[i]}$ and values $\bar{\mathcal{A}}_r^{[i]}[f]$ given by (1.11). Indeed, this is the most simple form to consider for analysis. However, for computational purposes, it is often more convenient to use the subtraction basis $q_r^{[i]}$, without resorting to Cardinal functions. In this case

$$g_k(x) = \sum_{i=0}^{1} \sum_{r=0}^{k-1} \tilde{\mathcal{A}}_r^{[i]}[f] q_r^{[i]}(x), \quad x \in [-1,1],$$

and the values $\tilde{\mathcal{A}}_r^{[i]}[f]$ are specified by the linear system

$$\sum_{s=0}^{k-1} \widehat{q}_{s\,m(r)}^{\,[i]} \tilde{\mathcal{A}}_s^{[i]}[f] = \hat{f}_{m(r)}^{[i]}, \quad r = 0, \ldots, k-1, \quad i \in \{0,1\}. \tag{1.12}$$

The resulting approximation is identical to the Cardinal function formulation.

1.3.1 Convergence rate of Eckhoff's approximation

Analysis of Eckhoff's method in the univariate, Fourier setting was carried out in [4]. Using virtually identical techniques, the following result can be established for the modified Fourier case:

Theorem 1.4. *Suppose that $m(r) = c(r)N + \mathcal{O}(1)$, where $c(r) \geq 1$ and that at most $l \leq k$ of the $c(r)$ are equal. Suppose further that $2K \geq l + 1$ and that $f \in \mathrm{H}^{2(k+K)}(-1,1)$. Then the coefficients $\bar{\mathcal{A}}_r^{[i]}[f]$ obtained by Eckhoff's method satisfy (1.10).*

This result was originally proved in [4] for the Cardinal basis comprised of polynomials. However, it is easily extended to the general case. Using this result we deduce the following:

Theorem 1.5. *Suppose that l, K and f are as in Theorem 1.4, and that $\mathcal{F}_{N,k}[f]$ utilises jump values approximated by Eckhoff's method. Then $\|f - \mathcal{F}_{N,k}[f]\|_q$ is $\mathcal{O}(N^{q-2k-\frac{3}{2}})$ for $q = 0, \ldots, 2k+1$.*

Proof. Suppose that we write $\mathcal{F}^e_{N,k}[f]$ and $\mathcal{F}_{N,k}[f]$ for the approximations based on the exact jump values $\mathcal{A}_r^{[i]}[f]$ and their approximations $\bar{\mathcal{A}}_r^{[i]}[f]$ respectively. In view of Lemma 1.2 it suffices to consider the difference $\mathcal{F}^e_{N,k}[f] - \mathcal{F}_{N,k}[f]$. We have

$$\|\mathcal{F}^e_{N,k}[f] - \mathcal{F}_{N,k}[f]\|_q \leq \sum_{i=0}^{1}\sum_{r=0}^{k-1} \left|\mathcal{A}_r^{[i]}[f] - \bar{\mathcal{A}}_r^{[i]}[f]\right| \left\|p_r^{[i]} - \mathcal{F}_N[p_r^{[i]}]\right\|_q. \tag{1.13}$$

Now suppose that a smooth function h satisfies the first r derivative conditions. It can be shown that $\left\|h - \mathcal{F}_N[h]\right\|_q = \mathcal{O}(N^{q-2r-\frac{3}{2}})$ for all $q \in \mathbb{N}$. Substituting this result with $h = p_r^{[i]}$ into (1.13) and using Theorem 1.4 immediately yields the result. □

Theorem 1.6. *Suppose that f and $\mathcal{F}_{N,k}[f]$ are as in Theorem 1.5. Then $\|f^{(q)} - (\mathcal{F}_{N,k}[f])^{(q)}\|_\infty$ is $\mathcal{O}\left(N^{q-2k-1}\right)$ for $q = 0, \ldots, 2k$.*

Proof. This follows immediately from Theorem 1.5 and the Sobolev inequality

$$\|h\|_\infty \leq c\sqrt{\|h\|\|h\|_1}, \quad \forall h \in \mathrm{H}^1(-1,1),$$

where c is a positive constant independent of h. □

These results, in comparison with those of Section 1.2, demonstrate that Eckhoff's method for approximating jump values does not deteriorate the convergence rate of the approximation. However, as we describe in Section 3, for certain choices of the values $m(r)$, Eckhoff's approximation offers at least one significant advantage in this respect.

Another consequence of these results is that, for certain choices of $m(r)$, Eckhoff's method requires additional smoothness to obtain the same convergence rate as the approximation based on the exact jump values. However, whenever the $c(r)$ are distinct, the smoothness requirement is identical.

In [4] the authors also compare the size of the error constants in $\|f - \mathcal{F}^e_{N,k}[f]\|_0$ and $\|f - \mathcal{F}_{N,k}[f]\|_0$. They demonstrate that approximating the jump values in this manner not only leads to the same convergence rate, but also that the error constant is not increased unduly. For this reason, we address only the asymptotic order of convergence throughout the remainder of this paper.

1.3.2 Choice of the values $m(r)$

The values $m(r) \geq N$ can be chosen arbitrarily, provided they are distinct and satisfy $m(r) = c(r)N + \mathcal{O}(1)$. Numerous choices are possible, such as

$$m(r) = N + r, \quad r = 0, \ldots, k-1. \tag{1.14}$$

In this case $c(r) = 1$ for all r, so that the function f being approximated must have $\mathrm{H}^{3k+1}(-1,1)$ or $\mathrm{H}^{3k+2}(-1,1)$-regularity (depending on whether k is odd or even) to ensure convergence. Other choices that require only $\mathrm{H}^{2k+2}(-1,1)$-regularity are also possible, such as

$$m(r) = (r+1)N, \quad r = 0, \ldots, k-1, \tag{1.15}$$

or, given some arbitrary value $\omega = 2, 3, \ldots$,

$$m(r) = \omega^r N, \quad r = 0, \ldots, k-1. \tag{1.16}$$

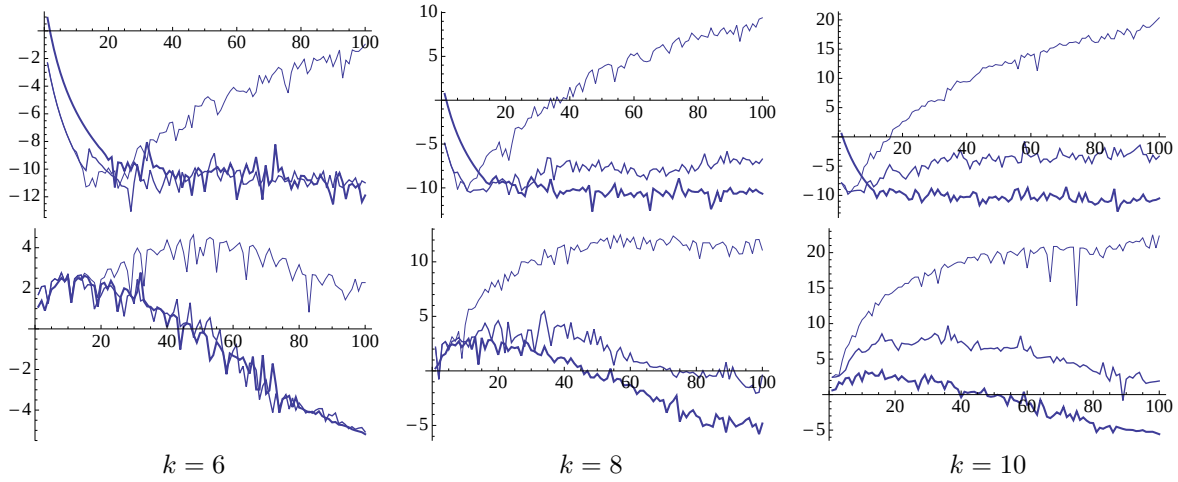


Figure 1: Log error $\log_{10} \|f(1) - \mathcal{F}_{N,k}[f]\|_\infty$ against $N = 1, \ldots, 100$ for Eckhoff's approximation using three different bases: Cardinal polynomial basis (thinnest line), Chebyshev polynomial basis (thicker line) and the dual basis (1.7) (thickest line). Here $f(x) = \cosh 6x$ (top), $f(x) = 5\mathrm{e}^{\cos 5\pi(1-x^2)}$ (bottom) and $m(r) = N + r$, $r = 0, \ldots, k-1$. Numerical results obtained in standard precision, using the *LinearSolve* routine in *Mathematica*.

One immediate disadvantage of these choices is they do not lead to a full auto-correction phenomenon (see Section 3). Further, the values $\hat{f}_n^{[i]}$, $n = 0, \ldots, N-1$, $n = m(r)$, $r = 0, \ldots, k-1$, required to form the approximation are not contiguous for (1.15) and (1.16), in contrast to (1.14). Finally, as we demonstrate in the forthcoming section, such values also lead to inferior numerical stability in comparison to (1.14).

1.3.3 Practical solution

The matrix $V^{[i]}$ is ill-conditioned. In fact, since $V^{[i]}$ is of the form $D^{[i]}\tilde{V}^{[i]}$, where $\tilde{V}^{[i]}$ is a Vandermonde matrix, the condition number is $\mathcal{O}\left(N^{2k+l-3}\right)$ for any choice of the values $m(r)$, where l is the number of equal values $c(r)$. This can be proved using well-known bounds for the norm of the inverse of a Vandermonde matrix [11]. Nonetheless, reasonably accurate numerical results are often obtained using the Björk–Pereyra algorithm [6]. In this manner, the values $\bar{\mathcal{A}}_r^{[i]}[f]$ can be found in $\mathcal{O}\left(k^2\right)$ operations.

However, increased numerical stability is obtained by replacing the Cardinal basis $\{p_r^{[i]}\}$ with an appropriately chosen subtraction basis $\{q_r^{[i]}\}$. The linear system to solve, namely (1.12), is often much more mildly conditioned (though asymptotically the same order), leading to better numerical results.

A significant improvement is offered by choosing $q_r^{[i]}$ as the $(2(r+1)-i)^{\text{th}}$ Chebyshev polynomial. This is a fairly standard approach, and the underlying matrix of the linear system is a *generalised Vandermonde* matrix [14]. However, this can be further improved upon by using the basis of dual functions (1.7). In Figure 1 we give numerical results for this basis and the Chebyshev and Cardinal polynomial bases applied to several functions. We observe that the approximation based on (1.7) offers the smallest error. Moreover, unlike the Cardinal polynomial basis, the error remains bounded. Note that the functions used here exhibit two features, large derivatives and high oscillation, which make their approximation prone to numerical errors. However, simply by replacing the subtraction basis we are able to obtain vastly superior approximations.

Regardless of the particular problem, the functions (1.7) offer a vast improvement in terms of the condition number of the linear system. As mentioned, the condition number scales like N^{2k+l-3} independently of the subtraction functions used. However, a vast reduction in the constant occurs when using (1.7). For $k = 10$ and values (1.14), the L^∞ condition number constant is roughly 3×10^{-16} for the linear system based on (1.7). In comparison, for the Chebyshev and Cardinal polynomial bases, these figures are 1×10^{-3} and 3×10^{3} respectively, the latter being roughly 10^{19} times larger. This effect is perhaps not surprising: the underlying matrix of the linear system (1.12) is a Cauchy matrix (see Theorem 1.3). Typically such matrices, though ill-conditioned themselves, are less poorly conditioned than Vandermonde matrices [14]. Note that such a linear system can also be solved in $\mathcal{O}\left(k^2\right)$ operations.

In all numerical results thus far, we have used the values (1.14). Seemingly, the condition number of the linear system (1.11) can be vastly improved from $\mathcal{O}\left(N^{3(k-1)}\right)$ to $\mathcal{O}\left(N^{2(k-1)}\right)$ by using the values

N	25	50	100	150	200
$m(r) = N + r$	1.215×10^{24}	1.808×10^{31}	7.398×10^{38}	2.784×10^{43}	5.335×10^{46}
$m(r) = (r+1)N$	8.688×10^{30}	2.147×10^{36}	5.552×10^{41}	8.185×10^{44}	1.451×10^{47}
$m(r) = 2^r N$	2.933×10^{42}	7.206×10^{47}	1.861×10^{53}	2.742×10^{56}	4.859×10^{58}

Table 1: L^∞ condition number of the linear system (1.12) using the functions (1.7) with $k = 10$ and values $m(r)$ given by (1.14)–(1.16). All values to 4 significant figures.

(1.15) or (1.16) instead. However, though true in theory, in practice the constant is so overbearingly large that it nullifies this effect. In Table 1 we give figures for the condition number of this linear system using the values (1.14)–(1.16). We observe that N exceeds 200 before the values (1.15) begin to offer an advantage (for the values (1.16) the scenario is much worse). However, since $k = 10$ in this example, any reasonable function will be well-resolved by Eckhoff's approximation for a much smaller value of N.

We mention in passing that, regardless of the subtraction functions employed, numerical results can often be further improved by solving over-determined least squares problems. This approach is fairly standard [4, 10]. For practical purposes, the least squares systems are solved by singular value decompositions, which can be found to high accuracy for Cauchy matrices [14].

This completes the study of the univariate version of Eckhoff's method. The remainder of this paper will focus on the development and analysis of the multivariate extension. It is not within the scope of this paper to address the issue of numerical stability in this context. In the Conclusion we mention a number of outstanding challenges herein, which require future investigation.

2 Eckhoff's method for multivariate expansions

In this section, we extend Eckhoff's method to functions defined on the d-variate cube $\bar{\Omega} = [-1, 1]^d$. To do so, it is first necessary to introduce multivariate modified Fourier expansions and the multivariate polynomial subtraction technique. The reader is referred to [1, 18] and [16] for further details.

2.1 Multivariate modified Fourier expansions

Suppose that $f \in \mathrm{L}^2(\Omega)$. The N^{th} truncated modified Fourier series of f can be written in the following succinct form:

$$\mathcal{F}_N[f](x) = \sum_{i \in \{0,1\}^d} \sum_{n \in I_N} \hat{f}_n^{[i]} \phi_n^{[i]}(x), \quad x = (x_1, \ldots, x_d) \in \bar{\Omega}. \tag{2.1}$$

Here $i = (i_1, \ldots, i_d)$, $n = (n_1, \ldots, n_d)$ and $\phi_n^{[i]}(x) = \phi_{n_1}^{[i_1]}(x_1) \ldots \phi_{n_d}^{[i_d]}(x_d)$. $I_N \subseteq \mathbb{N}^d$ is some finite index set. Throughout this section we assume that I_N is the full index set

$$I_N = \big\{ n \in \mathbb{N}^d : 0 \leq n_1, \ldots, n_d \leq N - 1 \big\}. \tag{2.2}$$

Note that $|I_N| = \mathcal{O}\left(N^d\right)$. In Section 4, we consider a different choice of index set, which greatly reduces this complexity without unduly affecting the convergence rate.

The bivariate case will serve as our primary example. In this setting (2.1) is

$$\begin{aligned}
\mathcal{F}_N[f](x_1, x_2) =& \frac{1}{4} \hat{f}_{0,0}^{[0,0]} + \frac{1}{2} \sum_{n_1=1}^{N-1} \left\{ \hat{f}_{n_1,0}^{[0,0]} \cos n_1 \pi x_1 + \hat{f}_{n_1,0}^{[1,0]} \sin(n_1 - \tfrac{1}{2}) \pi x_1 \right\} \\
&+ \frac{1}{2} \sum_{n_2=0}^{N-1} \left\{ \hat{f}_{0,n_2}^{[0,0]} \cos n_2 \pi x_2 + \hat{f}_{0,n_2}^{[0,1]} \sin(n_2 - \tfrac{1}{2}) \pi x_2 \right\} \\
&+ \sum_{n_1,n_2=1}^{N-1} \left\{ \hat{f}_{n_1,n_2}^{[0,0]} \cos n_1 \pi x_1 \cos n_2 \pi x_2 + \hat{f}_{n_1,n_2}^{[0,1]} \cos n_1 \pi x_1 \sin(n_2 - \tfrac{1}{2}) \pi x_2 \right. \\
&\qquad \left. + \hat{f}_{n_1,n_2}^{[1,0]} \sin(n_1 - \tfrac{1}{2}) \pi x_1 \cos n_2 \pi x_2 + \hat{f}_{n_1,n_2}^{[1,1]} \sin(n_1 - \tfrac{1}{2}) \pi x_1 \sin(n_2 - \tfrac{1}{2}) \pi x_2 \right\}.
\end{aligned}$$

2.1.1 Expansion of multivariate modified Fourier coefficients

The multivariate coefficients $\hat{f}_n^{[i]} = \int_\Omega f(x)\phi_n^{[i]}(x)\,\mathrm{d}x$, $i \in \{0,1\}^d$, $n \in \mathbb{N}^d$, are $\mathcal{O}\left(n^{-2}\right)$ for large n (provided f is sufficiently smooth), where $n^{-2} = (n_1 \dots n_d)^{-2}$. In fact, for all $n \in \mathbb{N}^d$, $|\hat{f}_n^{[i]}| \lesssim (\bar{n}_1 \dots \bar{n}_d)^{-2} = \bar{n}^{-2}$, where $\bar{m} = \max\{m, 1\}$ for $m \in \mathbb{N}$. Here, and for the remainder of this paper, we use the symbol $A \lesssim B$ to mean that there exists a constant c independent of N such that $A \leq cB$.

Vital to the construction and analysis of the multivariate version of Eckhoff's method is the expansion of the coefficients $\hat{f}_n^{[i]}$. Such coefficients admit an expansion similar to that of the univariate coefficients given in (1.2). For this we need some additional notation. Suppose that $[d]$ is the set of ordered tuples of length at most d with entries in $\{1, \dots, d\}$. We define $[d]^* = [d] \cup \{\emptyset\}$. For $t \in [d]$ we write $|t|$ for the length (number of elements) in t, so that $t = (t_1, \dots, t_{|t|})$ and $1 \leq t_1 < \dots < t_{|t|} \leq d$. We also write $\bar{t} \in [d]$ for the ordered tuple of length $d - |t|$ of elements not in t. For $j \in \{1, \dots, d\}$ we say that $j \in t$ if $j = t_l$ for some $l = 1, \dots, |t|$. Given $x = (x_1, \dots, x_d)$ we also define $x_t = (x_{t_1}, \dots, x_{t_{|t|}})$.

For a multi-index $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$, we define $|\alpha| = \sum_{j=1}^d \alpha_j$, $|\alpha|_\infty = \max_{j=1,\dots,d} \alpha_j$ and the differentiation operator $\mathrm{D}^\alpha = \partial_{x_1}^{\alpha_1} \dots \partial_{x_d}^{\alpha_d}$. If $\alpha = (r, r, \dots, r)$, $r \in \mathbb{N}$, we also write D^r. If $t \in [d]$ and $r \in \mathbb{N}$ we set $\mathrm{D}_t^r = \partial_{x_{t_1}}^r \dots \partial_{x_{t_{|t|}}}^r$.

Given $j = 1, \dots, d$, $r_j = 0, \dots, k-1$ and $i_j \in \{0,1\}$ we define $\mathcal{B}_{r_j}^{[i_j]}[f]$ by

$$\begin{aligned}(-1)^{r_j}\mathcal{B}_{r_j}^{[i_j]}[f](x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_d) =& \partial_{x_j}^{2r_j+1} f(x_1, \dots, x_{j-1}, 1, x_{j+1}, \dots, x_d) \\ &+ (-1)^{i_j+1}\partial_{x_j}^{2r_j+1} f(x_1, \dots, x_{j-1}, -1, x_{j+1}, \dots, x_d).\end{aligned}$$

For $t \in [d]^*$, $r_t = (r_{t_1}, \dots, r_{t_{|t|}}) \in \mathbb{N}^{|t|}$ and $i_t = (i_{t_1}, \dots, i_{t_{|t|}}) \in \{0,1\}^{|t|}$ we define $\mathcal{B}_{r_t}^{[i_t]}[f]$ as the composition

$$\mathcal{B}_{r_t}^{[i_t]}[f](x_{\bar{t}}) = \mathcal{B}_{r_{t_1}}^{[i_{t_1}]}\left[\mathcal{B}_{r_{t_2}}^{[i_{t_2}]}\left[\dots\left[\mathcal{B}_{r_{t_{|t|}}}^{[i_{t_{|t|}}]}[f]\right]\dots\right]\right],$$

with the understanding that $\mathcal{B}_{r_t}^{[i_t]}[f] = f$ when $t = \emptyset$. Note that the operators $\mathcal{B}_{r_j}^{[i_j]}[f]$, $j \in t$, commute with each other and with differentiation in the variable $x_{\bar{t}}$. Finally, given $t \in [d]^*$, $r_t \in \mathbb{N}^{|t|}$, $i \in \{0,1\}^d$ and $n_{\bar{t}} = (n_{\bar{t}_1}, \dots, n_{\bar{t}_{|\bar{t}|}}) \in \mathbb{N}^{|\bar{t}|}$, we define $\mathcal{A}_{r_t, n_{\bar{t}}}^{[i]}[f] \in \mathbb{R}$ by

$$\mathcal{A}_{r_t, n_{\bar{t}}}^{[i]}[f] = (-1)^{k|\bar{t}|}\prod_{j \notin t}\left(\mu_{n_j}^{[i_j]}\right)^{-k}\int \mathcal{B}_{r_t}^{[i_t]}[\mathrm{D}_{\bar{t}}^{2k} f](x_{\bar{t}})\phi_{n_{\bar{t}}}^{[i_{\bar{t}}]}(x_{\bar{t}})\,\mathrm{d}x_{\bar{t}}. \tag{2.3}$$

Note that the final integral is just the modified Fourier coefficient of the function $\mathcal{B}_{r_t}^{[i_t]}[\mathrm{D}_{\bar{t}}^{2k} f](x_{\bar{t}})$ corresponding to indices $i_{\bar{t}}$ and $n_{\bar{t}}$. For this reason, we have the bound

$$\left|\mathcal{A}_{r_t, n_{\bar{t}}}^{[i]}[f]\right| \lesssim \prod_{j \notin t}\bar{n}_j^{-2k-2} = \bar{n}_{\bar{t}}^{-2k-2}, \quad \forall n_{\bar{t}} \in \mathbb{N}^{|\bar{t}|}, \quad i \in \{0,1\}^d. \tag{2.4}$$

Note that $\mathcal{A}_{r_t, n_{\bar{t}}}^{[i]}[f]$ also depends on k. However, to simplify notation we will not make this dependence explicit. We may now derive an expansion for $\hat{f}_n^{[i]}$. After k integrations by parts in each variable (see also [1]), we obtain

$$\hat{f}_n^{[i]} = \sum_{t \in [d]^*}\sum_{|r_t|_\infty = 0}^{k-1}\mathcal{A}_{r_t, n_{\bar{t}}}^{[i]}[f](-1)^{|n_t| + |i_t|}\prod_{j \in t}\left(\mu_{n_j}^{[i_j]}\right)^{-(r_j+1)}. \tag{2.5}$$

As we establish in the sequel, the values $\mathcal{A}_{r_t, n_{\bar{t}}}^{[i]}[f]$, $t \in [d]$, are the appropriate generalisation of the univariate 'jumps' $\mathcal{A}_r^{[i]}[f]$ given in (1.3). The task of approximating these values to sufficient accuracy is the content of the remainder of this paper.

Suppose that $p_0^{[i]}, \dots, p_{k-1}^{[i]}$ are the Cardinal functions introduced in Section 1.2. Given $t \in [d]$, $i_t \in \{0,1\}^{|t|}$ and $r_t \in \{0, \dots, k-1\}^{|t|}$ we define $p_{r_t}^{[i_t]}(x_t) = \prod_{j \in t} p_{r_j}^{[i_j]}(x_j)$. With this in hand, we may rewrite (2.5) as

$$\hat{f}_n^{[i]} = \sum_{t \in [d]^*}\sum_{|r_t|_\infty = 0}^{k-1}\mathcal{A}_{r_t, n_{\bar{t}}}^{[i]}[f]\widehat{p_{r_t n_t}^{[i_t]}} + \mathcal{O}\left(n^{-2k-2}\right). \tag{2.6}$$

Provided the functions $p_r^{[i]}$ are Cardinal polynomials the final term of (2.6) vanishes. To simplify matters we assume this to be the case throughout the remainder of this paper, unless specified otherwise.

Continuing with the bivariate case as our primary example, we observe that (2.6) reduces to

$$\hat{f}_{n_1,n_2}^{[i_1,i_2]} = \sum_{r_1,r_2=0}^{k-1} \mathcal{A}_{r_1,r_2}^{[i_1,i_2]}[f]\widehat{p_{r_1}}_{n_1}^{[i_1]}\widehat{p_{r_2}}_{n_2}^{[i_2]} + \sum_{r_1=0}^{k-1} \mathcal{A}_{r_1,n_2}^{[i_1,i_2]}[f]\widehat{p_{r_1}}_{n_1}^{[i_1]} + \sum_{r_2=0}^{k-1} \mathcal{A}_{r_2,n_1}^{[i_1,i_2]}[f]\widehat{p_{r_2}}_{n_2}^{[i_2]} + \mathcal{A}_{n_1,n_2}^{[i_1,i_2]}[f],$$

when $d = 2$.

We have not yet established smoothness conditions for the expansion (2.5)–(2.6) to be valid. This requires introduction of the following spaces:

2.1.2 Sobolev spaces of dominating mixed smoothness

As described in greater detail in [1], modified Fourier expansions are best studied in so-called Sobolev spaces of *dominating mixed smoothness* [30, 32]. Given $q \in \mathbb{N}$, we define the q^{th} such space by

$$\mathrm{H}^q_{\text{mix}}(\Omega) = \{f : \mathrm{D}^\alpha f \in \mathrm{L}^2(\Omega),\ \forall \alpha \in \mathbb{N}^d : |\alpha|_\infty \leq q\},$$

with norm $\|f\|^2_{q,\text{mix}} = \sum_{|\alpha|_\infty \leq q} \|\mathrm{D}^\alpha f\|^2$.

The importance of such spaces in the study of modified Fourier expansions is immediately emphasised by the observation that $\mathcal{F}_N[f]$ converges uniformly to f on $\bar{\Omega}$ provided $f \in \mathrm{H}^1_{\text{mix}}(\Omega)$ [1]. Returning to the topic of the previous section, it is readily seen that (2.5)–(2.6) are valid for functions $f \in \mathrm{H}^{2k}_{\text{mix}}(\Omega)$.

Though we shall use such spaces throughout, we shall not discuss their properties in greater detail. We refer to [30, 32] for further reading, and to [1] for use of such spaces in the study of multivariate modified Fourier expansions.

2.2 Multivariate polynomial subtraction

As described in [16], to accelerate convergence, it suffices to interpolate the exact Neumann data of the function f on the boundary. In other words, given $k \in \mathbb{N}_+$, we seek a function g_k such that

$$\partial_{x_j}^{2r+1} g_k\big|_{x_j=\pm 1} = \partial_{x_j}^{2r+1} f\big|_{x_j=\pm 1}, \quad \forall j = 1, \ldots, d, \quad r = 0, \ldots, k-1,$$

or, in the notation of the previous section,

$$\mathcal{B}_{r_j}^{[i_j]}[g_k] = \mathcal{B}_{r_j}^{[i_j]}[f], \quad i_j \in \{0,1\}, \quad r_j = 0, \ldots, k-1, \quad j = 1, \ldots, d. \tag{2.7}$$

As in the univariate case, we say that the function $f - g_k$ satisfies the first k derivative conditions, and, as we shall observe, this guarantees faster convergence of the approximation $\mathcal{F}_{N,k}[f] = \mathcal{F}_N[f - g_k] + g_k$. Once more we refer to $\mathcal{F}_{N,k}[f]$ as the k^{th} polynomial subtraction approximation of f.

A suitable function g_k is given by the following lemma:

Lemma 2.1. *Suppose that $f \in \mathrm{H}^{2k}(\Omega)$ and that*

$$g_k(x) = \sum_{t \in [d]} \sum_{i_t \in \{0,1\}^{|t|}} \sum_{|r_t|_\infty = 0}^{k-1} (-1)^{|t|+1} \mathcal{B}_{r_t}^{[i_t]}[f](x_{\bar{t}}) p_{r_t}^{[i_t]}(x_t), \quad x \in \bar{\Omega}. \tag{2.8}$$

Then g_k satisfies (2.7).

Proof. It suffices to prove that g_k satisfies (2.7) with $j = 1$, $i_j = 0$ and $r_j = s$. We split the terms of (2.8) corresponding to different $t \in [d]$ into the three following cases: (i) $t = (1)$, (ii) $t = (1, u)$, where $u \in [d]$, $1 \notin u$, and (iii) $t = u$, where $1 \notin u$.

Consider case (i). The contribution of the corresponding term to $\mathcal{B}_s^{[0]}[g_k]$ is

$$\sum_{i_1=0}^{1} \sum_{r_1=0}^{k-1} \mathcal{B}_s^{[0]} \left[\mathcal{B}_{r_1}^{[i_1]}[f](x_2, \ldots, x_d) p_{r_1}^{[i_1]}(x_1)\right](x_2, \ldots, x_d) = \mathcal{B}_s^{[0]}[f](x_2, \ldots, x_d),$$

where the second equality follows directly from the properties of the Cardinal functions $p_r^{[i]}$. It now suffices to prove that the contributions corresponding to cases (ii) and (iii) cancel. For case (ii) the contribution is

$$\sum_{i_u\in\{0,1\}^{|u|}}\sum_{|r_u|_\infty=0}^{k-1}\sum_{i_1=0}^{1}\sum_{r_1=0}^{k-1}(-1)^{|u|}\mathcal{B}_s^{[0]}\left[\mathcal{B}_{r_t}^{[i_t]}[f](x_{\bar{t}})p_{r_t}^{[i_t]}(x_t)\right](x_2,\ldots,x_d)$$

$$=\sum_{i_u\in\{0,1\}^{|u|}}\sum_{|r_u|_\infty=0}^{k-1}(-1)^{|u|}\mathcal{B}_{(s,r_u)}^{[(0,i_u)]}[f](x_{\bar{u}})p_{r_u}^{[i_u]}(x_u),$$

where $(0,i_u)=(0,i_{u_1},\ldots,i_{u_{|u|}})$ and $(s,r_u)=(s,r_{u_1},\ldots,r_{u_{|u|}})$. It is readily seen that this is precisely the negative of the contribution of case (iii). □

Concerning the error of polynomial subtraction, we have the following result, proved in [1]:

Theorem 2.2. *Suppose that $f\in \mathrm{H}_{mix}^{2k+2}(\Omega)$ and that $\mathcal{F}_{N,k}[f]$ is the k^{th} polynomial subtraction approximation to f. Then $\|f-\mathcal{F}_{N,k}[f]\|_q$ is $\mathcal{O}(N^{q-2k-\frac{3}{2}})$ for $q=0,\ldots,2k+1$ and $\|\mathrm{D}^\alpha(f-\mathcal{F}_{N,k}[f])\|_\infty$ is $\mathcal{O}\left(N^{|\alpha|_\infty-2k-1}\right)$ for $|\alpha|_\infty\leq 2k$. If, additionally, $f\in \mathrm{H}_{mix}^{2k+3}(\Omega)$ then $\mathrm{D}^\alpha f(x)-\mathrm{D}^\alpha\mathcal{F}_{N,k}[f](x)$ is $\mathcal{O}\left(N^{|\alpha|_\infty-2k-2}\right)$ uniformly in compact subsets of Ω.*

As in the univariate case, we interpret $\mathcal{F}_{N,0}[f]$ as just $\mathcal{F}_N[f]$. When $k=0$, this theorem also establishes the rate of convergence of the multivariate modified Fourier expansion $\mathcal{F}_N[f]$.

For $d=2$, the function g_k is given by

$$g_k(x)=\sum_{i_1=0}^{1}\sum_{r_1=0}^{k-1}p_{r_1}^{[i_1]}(x_1)\mathcal{B}_{r_1}^{[i_1]}[f](x_2)+\sum_{i_2=0}^{1}\sum_{r_2=0}^{k-1}\mathcal{B}_{r_2}^{[i_2]}[f](x_1)p_{r_2}^{[i_2]}(x_2)$$
$$-\sum_{i_1,i_2=0}^{1}\sum_{r_1,r_2=0}^{k-1}\mathcal{B}_{r_1}^{[i_1]}\left[\mathcal{B}_{r_2}^{[i_2]}[f]\right]p_{r_1}^{[i_1]}(x_1)p_{r_2}^{[i_2]}(x_2).$$

We remark in passing that the phrase 'polynomial subtraction' is a misnomer: the function g_k is no longer a polynomial for $d\geq 2$. Herein lies the main problem with this device. Computation of the function g_k, as given by (2.8), requires knowledge of the exact derivatives of the function f over $(d-1)$-dimensional subsets of the boundary.

One approach to alleviate this problem, which we now introduce since it will be used in the sequel, is to approximate these lower dimensional functions using polynomial subtraction (an approach mentioned briefly, but not analysed, in [16]). To do so requires knowledge of functions over $(d-2)$-dimensional subsets of the boundary. However, we may repeat the same process, replacing exact functions by polynomial subtraction approximations, until we obtain an approximation that uses only derivative values over the 0-dimensional subsets of the boundary consisting of the vertices $(\pm1,\pm1,\ldots,\pm1)$ and modified Fourier coefficients of higher dimensional derivative functions.

To differentiate between the two approaches, we refer to the approximation based on (2.8) as *exact* polynomial subtraction and the approximation obtained by the above process as *approximate* polynomial subtraction. We write g_k^e, $\mathcal{F}_{N,k}^e[f]$ and g_k^a, $\mathcal{F}_{N,k}^a[f]$ respectively. Note that for $d=1$ both approximations coincide.

In the $d=2$ case we merely replace the univariate functions $\mathcal{B}_{r_1}^{[i_1]}[f]$ and $\mathcal{B}_{r_2}^{[i_2]}[f]$ by their k^{th} polynomial subtraction approximation. This yields the new function g_k^a given by

$$g_k^a(x)=\sum_{i_1=0}^{1}\sum_{r_1=0}^{k-1}p_{r_1}^{[i_1]}(x_1)\mathcal{F}_{N,k}\left[\mathcal{B}_{r_1}^{[i_1]}[f]\right](x_2)+\sum_{i_2=0}^{1}\sum_{r_2=0}^{k-1}\mathcal{F}_{N,k}\left[\mathcal{B}_{r_2}^{[i_2]}[f]\right](x_1)p_{r_2}^{[i_2]}(x_2)$$
$$-\sum_{i_1,i_2=0}^{1}\sum_{r_1,r_2=0}^{k-1}\mathcal{B}_{r_1}^{[i_1]}\left[\mathcal{B}_{r_2}^{[i_2]}[f]\right]p_{r_1}^{[i_1]}(x_1)p_{r_2}^{[i_2]}(x_2).$$

For $d\geq 3$ we define the new approximation inductively. If $\mathcal{F}_{N,k}^a[\cdot]$ has been obtained for $d-1$, we define the d-variate approximate polynomial subtraction function g_k^a by

$$g_k^a(x)=\sum_{t\in[d]}\sum_{i_t\in\{0,1\}^{|t|}}\sum_{|r_t|_\infty=0}^{k-1}(-1)^{|t|+1}\mathcal{F}_{N,k}^a\left[\mathcal{B}_{r_t}^{[i_t]}[f]\right](x_{\bar{t}})p_{r_t}^{[i_t]}(x_t),\quad x\in\bar{\Omega}. \tag{2.9}$$

For the approximate polynomial subtraction function (2.9) to be a potential alternative to its exact counterpart (2.8), it is necessary to demonstrate that the convergence rate is not deteriorated. We have:

Lemma 2.3. *Suppose that $f \in \mathrm{H}^{2k+2}_{mix}(\Omega)$ and that $\mathcal{F}^a_{N,k}[f]$ is the k^{th} approximate polynomial subtraction approximation of f. Then $\|f - \mathcal{F}^a_{N,k}[f]\|_q$ is $\mathcal{O}(N^{q-2k-\frac{3}{2}})$ for $q = 0, \dots, 2k+1$ and $\|\mathrm{D}^\alpha(f - \mathcal{F}^a_{N,k}[f])\|_\infty$ is $\mathcal{O}\left(N^{-2k-1}\right)$ for $|\alpha|_\infty \leq 2k$. If, additionally, $f \in \mathrm{H}^{2k+3}_{mix}(\Omega)$ then $\mathrm{D}^\alpha f(x) - \mathrm{D}^\alpha \mathcal{F}^a_{N,k}[f](x)$ is $\mathcal{O}\left(N^{|\alpha|_\infty - 2k-2}\right)$ uniformly in compact subsets of Ω.*

Proof. By Theorem 2.2 it suffices to consider the difference $\mathcal{F}^e_{N,k}[f] - \mathcal{F}^a_{N,k}[f]$. We use induction on d. For $d = 1$ there is nothing to prove. Now suppose that the result holds for $d-1$. We have

$$\mathcal{F}^e_{N,k}[f](x) - \mathcal{F}^a_{N,k}[f](x) = g^e_k(x) - g^a_k(x) - \mathcal{F}_N\left[g^e_k - g^a_k\right](x).$$

Since $\widehat{\mathcal{F}^a_{N,k}[h]}^{[i]}_n = \widehat{\mathcal{F}^e_{N,k}[h]}^{[i]}_n = \hat{h}^{[i]}_n$ for all $i \in \{0,1\}^d$, $n \in I_N$ and arbitrary function h, it follows that $\mathcal{F}_N[g^e_k - g^a_k] = 0$. Hence

$$\begin{aligned}\mathcal{F}^e_{N,k}[f](x) - \mathcal{F}^a_{N,k}[f](x) &= g^e_k(x) - g^a_k(x) \\ &= \sum_{t \in [d]} \sum_{i_t \in \{0,1\}^{|t|}} \sum_{|r_t|_\infty = 0}^{k-1} (-1)^{|t|+1} \left\{\mathcal{B}^{[i_t]}_{r_t}[f](x_{\bar{t}}) - \mathcal{F}^a_{N,k}\left[\mathcal{B}^{[i_t]}_{r_t}[f]\right](x_{\bar{t}})\right\} p^{[i_t]}_{r_t}(x_t).\end{aligned}$$

If $f \in \mathrm{H}^{2k+2}_{\mathrm{mix}}(\Omega)$ then it can be shown that $\mathcal{B}^{[i_t]}_{r_t}[f] \in \mathrm{H}^{2k+2}_{\mathrm{mix}}(-1,1)^{|\bar{t}|}$ [1]. Since $|t| \geq 1$, we may use the induction hypothesis on each such term to obtain the result. □

In its present form (2.9) the approximate subtraction function is not fit for practical purposes. Instead we seek a version of g^a_k that is not inductively defined. This is provided by the following lemma:

Lemma 2.4. *The approximate polynomial subtraction function g^a_k is given by*

$$g^a_k(x) = \sum_{i \in \{0,1\}^d} \sum_{t \in [d]} \sum_{|r_t|_\infty = 0}^{k-1} \sum_{|n_{\bar{t}}|_\infty = 0}^{N-1} \mathcal{A}^{[i]}_{r_t, n_{\bar{t}}}[f] p^{[i_t]}_{r_t}(x_t) \phi^{[i_{\bar{t}}]}_{n_{\bar{t}}}(x_{\bar{t}}), \tag{2.10}$$

where the terms $\mathcal{A}^{[i]}_{r_t, n_{\bar{t}}}[f]$ are given by (2.3).

To prove this lemma we need the following notation. Given $t \in [d]$ we write $[t]$ for the set of tuples $u \in [d]$ with $u \subseteq t$ (in other words, if $j \in u$ then $j \in t$ for $j = 1, \dots, d$). We write $[t]^* = [t] \cup \{\emptyset\}$ and $\bar{u} \in [t]^*$ for the tuple of elements in t but not in u. Further, given $t, u \in [d]^*$ we write $t \cup u \in [d]^*$ for the ordered tuple of elements $j = 1, \dots, d$ in t or in u, $t \cap u$ for the tuple of elements in both t and u and $t \backslash u$ for the tuple of elements in t but not in u.

Proof of Lemma 2.4. We prove this result by induction on d. For $d = 1$, since $g^a_k = g^e_k$ and $\mathcal{A}^{[i]}_r[f] = \mathcal{B}^{[i]}_r[f]$, there is nothing to prove. Now assume that the result holds for $d-1$. Then, by definition

$$g^a_k(x) = \sum_{t \in [d]} \sum_{i_t \in \{0,1\}^{|t|}} \sum_{|r_t|_\infty = 0}^{k-1} (-1)^{|t|+1} \mathcal{F}^a_{N,k}\left[\mathcal{B}^{[i_t]}_{r_t}[f]\right](x_{\bar{t}}) p^{[i_t]}_{r_t}(x_t). \tag{2.11}$$

Since $\mathcal{B}^{[i_t]}_{r_t}[f]$ is a function of at most $(d-1)$ variables, we may use the induction hypothesis to derive an expression for $\mathcal{F}^a_{N,k}\left[\mathcal{B}^{[i_t]}_{r_t}[f]\right](x_{\bar{t}})$. To do so, we require several observations. First, we note that

$$\mathcal{A}^{[i_{\bar{t}}]}_{r_u, n_{\bar{u}}}\left[\mathcal{B}^{[i_t]}_{r_t}[f]\right] = (-1)^{k|\bar{u}|} \prod_{j \in \bar{u}} (\mu^{[i_j]}_{n_j})^{-k} \int \mathcal{B}^{[i_u]}_{r_u}\left[\mathrm{D}^{2k}_{\bar{u}} \mathcal{B}^{[i_t]}_{r_t}[f]\right] \phi^{[i_{\bar{u}}]}_{n_{\bar{u}}}(x_{\bar{u}})\, \mathrm{d}x_{\bar{u}}, \quad \forall u \in [\bar{t}]^*.$$

Since $\bar{u} = \bar{t} \backslash u = \overline{t \cup u}$ and the operators $\mathcal{B}^{[i_u]}_{r_u}$ and $\mathcal{B}^{[i_t]}_{r_t}$ commute with each other and with differentiation in the independent variables, this gives

$$\mathcal{A}^{[i_{\bar{t}}]}_{r_u, n_{\bar{u}}}\left[\mathcal{B}^{[i_t]}_{r_t}[f]\right] = (-1)^{k|\bar{u}|} \prod_{j \in \bar{u}} (\mu^{[i_j]}_{n_j})^{-k} \int \mathcal{B}^{[i_{t\cup u}]}_{r_{t\cup u}}\left[\mathrm{D}^{2k}_{\bar{u}} f\right] \phi^{[i_{\bar{u}}]}_{n_{\bar{u}}}(x_{\bar{u}})\, \mathrm{d}x_{\bar{u}} = \mathcal{A}^{[i]}_{r_{t\cup u}, n_{\overline{t\cup u}}}[f].$$

Our next observation is as follows: if h is a function of at most $(d-1)$ variables, and g_k^a is the approximate polynomial subtraction function for h, then

$$\mathcal{F}_N[h - g_k^a](x) = \sum_{i\in\{0,1\}^{d-1}} \sum_{|n|_\infty=0}^{N-1} \mathcal{A}_n^{[i]}[h]\phi_n^{[i]}(x), \quad x \in [-1,1]^{d-1},$$

where $\mathcal{A}_n^{[i]}[h]$ is the value $\mathcal{A}_{r_t,n_{\bar{t}}}^{[i]}[h]$ given by (2.3) with $t = \emptyset$. This follows immediately from the induction hypothesis and equations (2.5) and (2.10).

Returning to $\mathcal{B}_{r_t}^{[i_t]}[f]$ and using these observations, we obtain

$$\begin{aligned}\mathcal{F}_{N,k}^a\left[\mathcal{B}_{r_t}^{[i_t]}[f]\right](x_{\bar{t}}) = &\sum_{i_{\bar{t}}\in\{0,1\}^{|\bar{t}|}} \sum_{|n_{\bar{t}}|_\infty=0}^{N-1} \mathcal{A}_{r_t,n_{\bar{t}}}^{[i]}[f]\phi_{n_{\bar{t}}}^{[i_{\bar{t}}]}(x_{\bar{t}}) \\ &+ \sum_{i_{\bar{t}}\in\{0,1\}^{|\bar{t}|}} \sum_{u\in[\bar{t}]} \sum_{|r_u|_\infty=0}^{k-1} \sum_{|n_{\bar{u}}|_\infty=0}^{N-1} \mathcal{A}_{r_{t\cup u},n_{\overline{t\cup u}}}^{[i]}[f]p_{r_u}^{[i_u]}(x_u)\phi_{n_{\bar{u}}}^{[i_{\bar{u}}]}(x_{\bar{u}}).\end{aligned}$$

Substituting this into (2.11) gives

$$\begin{aligned}g_k^a(x) = \sum_{i\in\{0,1\}^d} \sum_{t\in[d]} (-1)^{|t|+1} \Bigg\{ &\sum_{|r_t|_\infty=0}^{k-1} \sum_{|n_{\bar{t}}|_\infty=0}^{N-1} \mathcal{A}_{r_t,n_{\bar{t}}}^{[i]}[f]p_{r_t}^{[i_t]}(x_t)\phi_{n_{\bar{t}}}^{[i_{\bar{t}}]}(x_{\bar{t}}) \\ &+ \sum_{u\in[\bar{t}]} \sum_{|r_{t\cup u}|_\infty=0}^{k-1} \sum_{|n_{\bar{u}}|_\infty=0}^{N-1} \mathcal{A}_{r_{t\cup u},n_{\overline{t\cup u}}}^{[i]}[f]p_{r_{t\cup u}}^{[i_{t\cup u}]}(x_{t\cup u})\phi_{n_{\bar{u}}}^{[i_{\bar{u}}]}(x_{\bar{u}}) \Bigg\}. \qquad (2.12)\end{aligned}$$

To complete the proof, it suffices to show that, for any $v \in [d]$, the coefficient of $\mathcal{A}_{r_v,n_{\bar{v}}}^{[i]}[f]p_{r_v}^{[i_v]}(x_v)\phi_{n_{\bar{v}}}^{[i_{\bar{v}}]}(x_{\bar{v}})$ in (2.12) is precisely 1. The first term of (2.12) gives a contribution of $(-1)^{|v|+1}$. For the second, the terms that give contributions satisfy $t \cup u = v$. Since $t, u \neq \emptyset$ and there are $\binom{|v|}{l}$ possible choices of such u with $|u| = l$, the contribution of the second term is

$$(-1)^{|v|}\binom{|v|}{1} + \ldots + \binom{|v|}{|v|-1} = (-1)^{|v|+1}\sum_{l=1}^{|v|-1}\binom{|v|}{l}(-1)^l = 1 - (-1)^{|v|+1}.$$

Summing together this and the previous contributions now yields the result. □

The result of this lemma not only gives an explicit way to compute the k^{th} approximate polynomial subtraction function, it also demonstrates that the values $\mathcal{A}_{r_t,n_{\bar{t}}}^{[i]}[f]$ are the requisite multivariate 'jumps' that need to be approximated. This indicates the appropriate generalisation of Eckhoff's method, which we consider in the next section.

We mention in passing that, although the approximate polynomial subtraction process achieves a significant improvement over exact polynomial subtraction, it still requires explicit knowledge of these values. In general such values are unknown. Since there are

$$2^d\sum_{j=1}^{d}\binom{d}{j}k^jN^{d-j} = 2^d\left\{(k+N)^d - k^d\right\} = \mathcal{O}\left(kN^{d-1}\right), \quad k \ll N,$$

values in total, the need for a method of approximation becomes more vital as d increases.

2.3 The multivariate version of Eckhoff's method

We now extend Eckhoff's method to the multivariate setting. The bivariate version of this method was originally developed, without analysis, in [26, 24, 27]. In this section, we first establish an extension for general d, and then provide pertinent analysis.

As indicated by the form of the approximate polynomial subtraction function g_k^a it suffices to approximate the values $\mathcal{A}_{r_t,n_{\bar{t}}}^{[i]}[f]$ by values $\bar{\mathcal{A}}_{r_t,n_{\bar{t}}}^{[i]}[f]$. To this end, we define the subtraction function

$$g_k(x) = \sum_{i\in\{0,1\}^d} \sum_{t\in[d]} \sum_{|r_t|_\infty=0}^{k-1} \sum_{|n_{\bar{t}}|_\infty=0}^{N-1} \bar{\mathcal{A}}_{r_t,n_{\bar{t}}}^{[i]}[f]p_{r_t}^{[i_t]}(x_t)\phi_{n_{\bar{t}}}^{[i_{\bar{t}}]}(x_{\bar{t}}), \qquad (2.13)$$

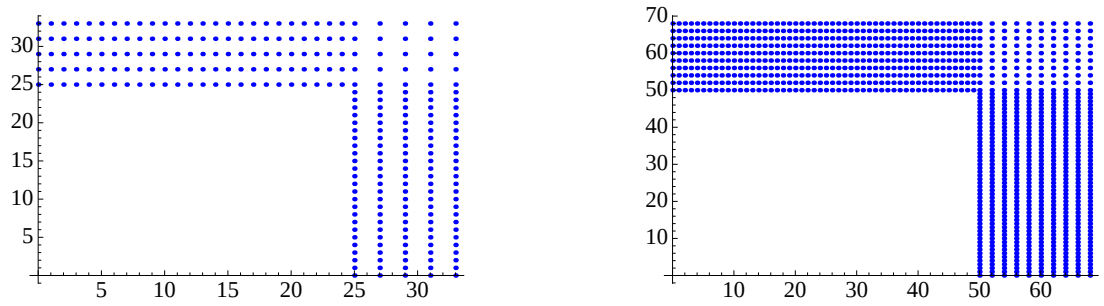


Figure 2: (left) Index set M_5 with $N = 25$ and $m(r) = N + 2r$. (right) Index set M_{10} with $N = 50$ and $m(r) = N + 2r$.

and the approximation $\mathcal{F}_{N,k}[f] = \mathcal{F}_N[f - g_k] + g_k$. In the univariate setting it follows from (1.11) that the function g_k satisfies the condition

$$\widehat{g_k}_n^{[i]} = \hat{f}_n^{[i]}, \quad n = m(0), \ldots, m(k-1), \quad i \in \{0,1\}. \tag{2.14}$$

For the d-variate extension we enforce a similar condition. Suppose that we define the finite index set $M_k \subseteq \mathbb{N}^d$ by

$$M_k = \bigcup_{t \in [d]} \left\{ n = (n_1, \ldots, n_d) \in \mathbb{N}^d : n_j = m(r_j),\ r_j = 0, \ldots, k-1, j \in t, |n_{\bar{t}}|_\infty < N \right\}. \tag{2.15}$$

We now impose the condition

$$\widehat{g_k}_n^{[i]} = \hat{f}_n^{[i]}, \quad \forall n \in M_k, \quad i \in \{0,1\}^d. \tag{2.16}$$

For $d = 1$ (2.16) reduces to (2.14). For $d = 2$ we obtain the following system of equations

$$\begin{aligned}
\widehat{g_k}_{m(r_1),m(r_2)}^{[i]} &= \widehat{f}_{m(r_1),m(r_2)}^{[i]}, \quad r_1, r_2 = 0, \ldots, k-1, \quad i \in \{0,1\}^2, \\
\widehat{g_k}_{m(r_1),n_2}^{[i]} &= \widehat{f}_{m(r_1),n_2}^{[i]}, \quad r_1 = 0, \ldots, k-1, \quad n_2 = 0, \ldots, N-1, \quad i \in \{0,1\}^2, \\
\widehat{g_k}_{n_1,m(r_2)}^{[i]} &= \widehat{f}_{n_1,m(r_2)}^{[i]}, \quad n_1 = 0, \ldots, N-1, \quad r_2 = 0, \ldots, k-1, \quad i \in \{0,1\}^2.
\end{aligned} \tag{2.17}$$

Figure 2 shows a typical form of the index set M_k for $d = 2$. Note that, as in the univariate case, the system of equations (2.16) completely decouples for different values of $i \in \{0,1\}^d$.

For both practical and analytical purposes, we need to expand the left hand side of (2.16). Given $u \in [d]$, $s_u \in \{0, \ldots, k-1\}^{|u|}$ and $n_{\bar{u}} \in \{0, \ldots, N-1\}^{|\bar{u}|}$, the corresponding term in g_k is

$$\bar{\mathcal{A}}_{s_u, n_{\bar{u}}}^{[i]}[f] p_{s_u}^{[i_u]}(x_u) \phi_{n_{\bar{u}}}^{[i_{\bar{u}}]}(x_{\bar{u}}) = \bar{\mathcal{A}}_{r_u, n_{\bar{u}}}^{[i]}[f] \prod_{j \in u} p_{s_j}^{[i_j]}(x_j) \prod_{j \notin u} \phi_{n_j}^{[i_j]}(x_j).$$

This term gives a non-zero contribution to the left hand side of (2.16) precisely when $t \subseteq u$, where $t \in [d]$ is the tuple corresponding to $n \in M_k$. Hence

$$\widehat{g_k}_n^{[i]} = \sum_{\substack{u \in [d] \\ t \subseteq u}} \sum_{|s_u|_\infty = 0}^{k-1} \bar{\mathcal{A}}_{s_u, n_{\bar{u}}}^{[i]}[f] \prod_{j \in u} \widehat{p_{s_j}}_{n_j}^{[i_j]} = \sum_{|s_t|_\infty = 0}^{k-1} \prod_{j \in t} V_{r_j, s_j}^{[i_j]} \left\{ \sum_{t \subseteq u} \sum_{|s_{u \setminus t}|_\infty = 0}^{k-1} \bar{\mathcal{A}}_{s_u, n_{\bar{u}}}^{[i]}[f] \prod_{j \in u \setminus t} \widehat{p_{s_j}}_{n_j}^{[i_j]} \right\}, \tag{2.18}$$

Here r_j is the index used in the definition (2.15) of M_k and $V^{[i]}$ is the matrix introduced in Section 1.3.

For $d = 2$, we may expand the system of equations (2.17) using (2.18) to give, for each $i \in \{0,1\}^d$,

$$\begin{aligned}
&\sum_{s_1, s_2 = 0}^{k-1} V_{r_1, s_1}^{[i_1]} V_{r_2, s_2}^{[i_2]} \bar{\mathcal{A}}_{s_1, s_2}^{[i]}[f] = \hat{f}_{m(r_1), m(r_2)}^{[i]}, \quad r_1, r_2 = 0, \ldots, k-1, \\
&\sum_{s_1 = 0}^{k-1} V_{r_1, s_1}^{[i_1]} \left\{ \bar{\mathcal{A}}_{s_1, n_2}^{[i]}[f] + \sum_{s_2 = 0}^{k-1} \bar{\mathcal{A}}_{s_1, s_2}^{[i]}[f] \widehat{p_{s_2}}_{n_2}^{[i_2]} \right\} = \hat{f}_{m(r_1), n_2}^{[i]}, \quad r_1 = 0, \ldots, k-1, \quad n_2 = 0, \ldots, N-1, \\
&\sum_{s_2 = 0}^{k-1} V_{r_2, s_2}^{[i_2]} \left\{ \bar{\mathcal{A}}_{s_2, n_1}^{[i]}[f] + \sum_{s_1 = 0}^{k-1} \bar{\mathcal{A}}_{s_1, s_2}^{[i]}[f] \widehat{p_{s_1}}_{n_1}^{[i_1]} \right\} = \hat{f}_{n_1, m(r_2)}^{[i]}, \quad n_1 = 0, \ldots, N-1, \quad r_2 = 0, \ldots, k-1.
\end{aligned}$$

In this case, it is obvious how to solve these equations. We first obtain $\bar{\mathcal{A}}^{[i]}_{r_1,r_2}[f]$ from the first equation, then use this to find $\bar{\mathcal{A}}^{[i]}_{r_1,n_2}[f]$ and $\bar{\mathcal{A}}^{[i]}_{r_2,n_1}[f]$ explicitly. The same can be done in $d \geq 3$ dimensions. Starting with the equation corresponding to $t = (1, 2, \ldots, d)$, we find $\bar{\mathcal{A}}^{[i]}_{r_t}[f]$. Using this, we solve the d equations corresponding to $|t| = d-1$, then those corresponding to $|t| = d-2$, and so on. Continuing in this manner, we obtain all the coefficients $\bar{\mathcal{A}}^{[i]}_{r_t,n_{\bar{t}}}[f]$. Though straightforward in theory, the construction of Eckhoff's approximation is likely to become increasingly cumbersome to implement for large d. However, it is certainly practical for $d = 2, 3$ as we demonstrate in the sequel by numerical example.

Observe that, to find the coefficients $\bar{\mathcal{A}}^{[i]}_{r_t,n_{\bar{t}}}[f]$, we have to solve linear systems involving the matrix $V^{[i]}$. One immediate benefit of Eckhoff's approach is that the coefficients can be found by solving essentially one-dimensional linear systems. Since we need to solve many such systems, it is easiest to find $(V^{[i]})^{-1}$ first. This can be achieved using methods outlined in Section 1.3.3. Note that existence and uniqueness of a solution to these linear systems is completely determined by the non-singularity of the matrix $V^{[i]}$ (see Theorem 1.3).

In the univariate case, the complexity of forming Eckhoff's approximation is $\mathcal{O}\left(\max\{k^2, kN\}\right)$. In the multivariate setting, it is readily seen that this figure is $\mathcal{O}\left(\max\{k^{d+1}, k^d N^d\}\right)$. Typically $k \ll N$ so this reduces to $k^d N^d$. In comparison, forming the approximation $\mathcal{F}_N[f]$ involves $\mathcal{O}\left(N^d\right)$ operations, so the increase in complexity is relatively mild for moderate values of k. Nonetheless, the value N^d grows exponentially with d. In Section 4 we demonstrate how this figure can be reduced significantly without affecting the convergence rate of $\mathcal{F}_{N,k}[f]$ unduly.

2.4 Analysis of Eckhoff's method

To commence our analysis of the multivariate version of Eckhoff's method, we require the following two lemmas, the first of which is a generalisation of Theorem 1.4:

Lemma 2.5. *Suppose that $h \in \mathrm{H}^{2(k+K)}_{mix}(\Omega)$, where $2K \geq l+1$ and l is the number of equal values $c(r)$, and that $t \in [d]$. Suppose further that*

$$\mathcal{B}^{[i_j]}_{r_j}[h] = 0, \quad r_j = 0, \ldots, k-1, \quad i_j \in \{0,1\}, \quad j \in t,$$
$$\mathcal{B}^{[i_j]}_{r_j}[h] = 0, \quad r_j = 0, \ldots, \alpha_j - 1, \quad i_j \in \{0,1\}, \quad j \notin t,$$

where $\alpha_{\bar{t}} \in \mathbb{N}^{|\bar{t}|}$, and that the values $\mathcal{E}^{[i_t]}_{r_t,n_{\bar{t}}}$, $r_t \in \{0, \ldots, k-1\}^{|t|}$, $n_{\bar{t}} \in \mathbb{N}^{|\bar{t}|}$ are defined by

$$\sum_{|s_t|_\infty = 0}^{k-1} \prod_{j \in t} V^{[i_j]}_{r_j,s_j} \mathcal{E}^{[i_t]}_{s_t,n_{\bar{t}}} = \hat{h}^{[i]}_n,$$

where $n_j = m(r_j)$, $r_j = 0, \ldots, k-1$ when $j \in t$ and $n_j \in \mathbb{N}$ otherwise. Then we have

$$\left|\mathcal{E}^{[i_t]}_{r_t,n_{\bar{t}}}\right| \lesssim N^{2(|r_t|-k|t|)} \bar{n}_{\bar{t}}^{-2\alpha_{\bar{t}}-2},$$

where $\bar{n}_{\bar{t}}^{-2\alpha_{\bar{t}}-2} = \prod_{j \notin t} \bar{n}_j^{-2\alpha_j-2}$.

Proof. For each $j \in t$ we may expand $\hat{h}^{[i]}_n$ a total of $(k+K)$ times with respect to n_j using the univariate expansion (1.2). We now apply $(V^{[i_j]}_{r_j,s_j})^{-1}$ to the result and use Theorem 1.4. □

Somewhat counter-intuitively, we first perform our analysis of Eckhoff's method for a function f that satisfies the first k derivative conditions, i.e

$$\mathcal{B}^{[i_j]}_{r_j}[f] = 0, \quad i_j \in \{0,1\}, \quad r_j = 0, \ldots, k-1, \quad j = 1, \ldots, d. \tag{2.19}$$

To do so, we require the following two lemmas:

Lemma 2.6. *Suppose that $t \in [d]$, $r_t \in \{0, \ldots, k-1\}^{|t|}$, $n_{\bar{t}} \in \{0, \ldots, N-1\}^{|\bar{t}|}$ and*

$$\mathcal{E}^{[i]}_{r_t,n_{\bar{t}}}[f] = \sum_{\substack{u \in [d] \\ t \subseteq u}} \sum_{|r_{u \setminus t}|_\infty = 0}^{k-1} \left(\mathcal{A}^{[i]}_{r_u,n_{\bar{u}}}[f] - \bar{\mathcal{A}}^{[i]}_{r_u,n_{\bar{u}}}[f]\right) \prod_{j \in u \setminus t} \widehat{p_{r_j}}^{[i_j]}_{n_j}. \tag{2.20}$$

Then

$$\sum_{|s_t|_\infty=0}^{k-1} \prod_{j\in t} V_{r_j,s_j}^{[i_j]} \mathcal{E}_{s_t,n_{\bar{t}}}^{[i]}[f] = -\sum_{\substack{u\in[d]^* \\ t\not\subseteq u}} \sum_{|s_u|_\infty=0}^{k-1} \mathcal{A}_{s_u,n_{\bar{u}}}^{[i]}[f]\widehat{p_{s_u n_u}}^{[i_u]}, \quad n \in M_k, \quad i \in \{0,1\}^d. \tag{2.21}$$

Proof. Consider the right hand side of (2.16). Using the expansion (2.6) gives

$$\hat{f}_n^{[i]} = \sum_{\substack{u\in[d] \\ t\subseteq u}} \sum_{|s_u|_\infty=0}^{k-1} \mathcal{A}_{s_u,n_{\bar{u}}}^{[i]}[f]\widehat{p_{s_u n_u}}^{[i_u]} + \sum_{\substack{u\in[d]^* \\ t\not\subseteq u}} \sum_{|s_u|_\infty=0}^{k-1} \mathcal{A}_{s_u,n_{\bar{u}}}^{[i]}[f]\widehat{p_{s_u n_u}}^{[i_u]}, \quad n \in M_k, \quad i \in \{0,1\}^d.$$

Equating this with (2.18) and rearranging gives the result. □

Lemma 2.7. *Suppose that $f \in \mathrm{H}_{mix}^{2(k+K)}(\Omega)$ satisfies the first k derivative conditions (2.19) and that $\mathcal{E}_{r_t,n_{\bar{t}}}^{[i]}[f]$ is given by (2.20). Then $\left|\mathcal{E}_{r_t,n_{\bar{t}}}^{[i]}[f]\right| \lesssim N^{2(|r_t|_\infty-k)}\bar{n}_{\bar{t}}^{-2k-2}$.*

Proof. Since f obeys the first k derivative conditions, $\mathcal{A}_{r_t,n_{\bar{t}}}^{[i]}[f] = 0$ when $t \neq \emptyset$. Hence

$$\sum_{|s_t|_\infty=0}^{k-1} \prod_{j\in t} V_{r_j,s_j}^{[i_j]} \mathcal{E}_{s_t,n_{\bar{t}}}^{[i]}[f] = -\mathcal{A}_n^{[i]}[f], \quad n \in M_k, \quad i \in \{0,1\}^d.$$

Since $\hat{f}_n^{[i]} = \mathcal{A}_n^{[i]}[f]$ in this case, an application of Lemma 2.5 now yields the result. □

Due to Theorem 2.3, to estimate the convergence rate of the multivariate Eckhoff approximation $\mathcal{F}_{N,k}[f]$, it suffices to consider the difference $\mathcal{F}_{N,k}^a[f] - \mathcal{F}_{N,k}[f]$, where $\mathcal{F}_{N,k}^a[f]$ is the approximate polynomial subtraction approximation introduced in Section 2.2. For this we need the following lemma, which demonstrates the importance of the quantity (2.20):

Lemma 2.8. *We have*

$$\begin{aligned}&\mathcal{F}_{N,k}^a[f](x) - \mathcal{F}_{N,k}[f](x) \\ &\qquad = \sum_{i\in\{0,1\}^d} \sum_{t\in[d]} \sum_{|r_t|_\infty=0}^{k-1} \sum_{|n_{\bar{t}}|_\infty=0}^{N-1} \mathcal{E}_{r_t,n_{\bar{t}}}^{[i]}[f]\phi_{n_{\bar{t}}}^{[i_{\bar{t}}]}(x_{\bar{t}}) \prod_{j\in t} \left\{ p_{r_j}^{[i_j]}(x_j) - \mathcal{F}_N[p_{r_j}^{[i_j]}](x_j) \right\}. \end{aligned} \tag{2.22}$$

Proof. We may write

$$\mathcal{F}_{N,k}^a[f](x) - \mathcal{F}_{N,k}[f](x) = h_k(x) - \mathcal{F}_N[h_k](x), \tag{2.23}$$

where h_k is the smooth function

$$h_k(x) = \sum_{i\in\{0,1\}^d} \sum_{t\in[d]} \sum_{|r_t|_\infty=0}^{k-1} \sum_{|n_{\bar{t}}|_\infty=0}^{N-1} \left(\mathcal{A}_{r_t,n_{\bar{t}}}^{[i]}[f] - \bar{\mathcal{A}}_{r_t,n_{\bar{t}}}^{[i]}[f] \right) p_{r_t}^{[i_t]}(x_t)\phi_{n_{\bar{t}}}^{[i_{\bar{t}}]}(x_{\bar{t}}).$$

To prove the result, it suffices to demonstrate that the right hand sides of (2.22) and (2.23) have equal modified Fourier coefficients for all indices $i \in \{0,1\}^d$ and $n \in \mathbb{N}^d$. It is readily shown that both have vanishing coefficients whenever $n \in I_N$, so we consider the case $n \notin I_N$. In this setting, there is some $u \in [d]$ such that $n_j \geq N$ whenever $j \in u$ and $n_j = 0, \ldots, N-1$ otherwise. By identical arguments to those used to obtain (2.18), it can be shown that the coefficient the right hand side of (2.23), namely $\widehat{h_{k_n}}^{[i]}$, is

$$\widehat{h_{k_n}}^{[i]} = \sum_{|r_u|_\infty=0}^{k-1} \widehat{p_{r_u n_u}}^{[i_u]} \mathcal{E}_{r_u,n_{\bar{u}}}^{[i]}[f]. \tag{2.24}$$

We now consider the corresponding coefficient of (2.22). For each $t \in [d]$, due to the function $\phi_{n_{\bar{t}}}^{[i_{\bar{t}}]}$, we must have that $u \subseteq t$, otherwise, the corresponding term vanishes. However, due to the product, we must also have that $t \subseteq u$ for a non-zero contribution. Hence, $t = u$ and the modified Fourier coefficient of (2.22) reduces to (2.24), completing the proof. □

We are now able to provide an error estimate for a function f that obeys the first k derivative conditions:

Lemma 2.9. *Suppose that $f \in \mathrm{H}^{2(k+K)}_{mix}(\Omega)$ obeys the first k derivative conditions (2.19), where $2K \geq l+1$ and l is the number of equal $c(r)$, and that $\mathcal{F}_{N,k}[f]$ is the multivariate Eckhoff approximation of f. Then $\|D^\alpha(f - \mathcal{F}_{N,k}[f])\|_\infty$ is $\mathcal{O}\left(N^{|\alpha|_\infty - 2k-1}\right)$ for $|\alpha|_\infty \leq 2k$ and $\|(f - \mathcal{F}_{N,k}[f])\|_q$ is $\mathcal{O}(N^{q-2k-\frac{3}{2}})$ for $q = 0, \ldots, 2k+1$.*

Proof. It suffices to consider the difference $\mathcal{F}^a_{N,k}[f] - \mathcal{F}_{N,k}[f]$. Using Lemma 2.8, the bound derived in Lemma 2.6 and the fact that $\|(p_r^{[i]} - \mathcal{F}_N[p_r^{[i]}])^{(q)}\|_\infty = \mathcal{O}\left(N^{q-2r-1}\right)$, $r \in \mathbb{N}_0$, $q \in \mathbb{N}_0$, we deduce that

$$\begin{aligned}
&\|D^\alpha(\mathcal{F}^a_{N,k}[f] - \mathcal{F}_{N,k}[f])\|_\infty \\
&\qquad \lesssim \sum_{i \in \{0,1\}^d} \sum_{t \in [d]} \sum_{|r_t|_\infty = 0}^{k-1} \sum_{|n_{\bar{t}}|_\infty = 0}^{N-1} \left|\mathcal{E}^{[i]}_{r_t, n_{\bar{t}}}[f]\right| \left\|D^{\alpha_{\bar{t}}} \phi^{[i_{\bar{t}}]}_{n_{\bar{t}}}\right\|_\infty \prod_{j \in t} \left\| \left(p^{[i_j]}_{r_j} - \mathcal{F}_N[p^{[i_j]}_{r_j}]\right)^{(\alpha_j)} \right\|_\infty \\
&\qquad \lesssim \sum_{t \in [d]} \sum_{|r_t|_\infty = 0}^{k-1} \sum_{|n_{\bar{t}}|_\infty = 0}^{N-1} \bar{n}_{\bar{t}}^{\alpha_{\bar{t}} - 2k-2} N^{2(|r_t|_\infty - k)} \prod_{j \in t} N^{\alpha_j - 2r_j - 1}.
\end{aligned}$$

Since $|\alpha|_\infty \leq 2k$, we have $\bar{n}_{\bar{t}}^{\alpha_{\bar{t}} - 2k-2} \leq \bar{n}_{\bar{t}}^{-2}$. Hence

$$\|D^\alpha(\mathcal{F}^a_{N,k}[f] - \mathcal{F}_{N,k}[f])\|_\infty \lesssim \sum_{t \in [d]} N^{2(|r_t|_\infty - k)} N^{|\alpha_t|_\infty - 2|r_t| - 2|t|} \lesssim N^{|\alpha|_\infty - 2k-1},$$

which gives the result for the uniform error. The result for the $\mathrm{H}^q(\Omega)$ norm is proved in an identical manner. □

With this in hand we are able to deduce the main result of this section:

Theorem 2.10. *Suppose that $f \in \mathrm{H}^{2(k+K)}_{mix}(\Omega)$, where $2K \geq l+1$ and l is the number of equal $c(r)$, and that $\mathcal{F}_{N,k}[f]$ is the multivariate Eckhoff approximation of f. Then $\|\mathrm{D}^\alpha(f - \mathcal{F}_{N,k}[f])\|_\infty$ is $\mathcal{O}\left(N^{|\alpha|_\infty - 2k-1}\right)$ for $|\alpha|_\infty \leq 2k$ and $\|f - \mathcal{F}_{N,k}[f]\|_q$ is $\mathcal{O}(N^{q-2k-\frac{3}{2}})$ for $q = 0, \ldots, 2k+1$.*

Proof. We proceed by induction on d. Since the $d = 1$ result has been proved, we assume that the result holds for $d-1$. Suppose that g^e_k is the exact polynomial subtraction function (2.8) so that $f - g^e_k$ satisfies the first k derivative conditions. Writing $f = (f - g^e_k) + g^e_k$ and using linearity of $\mathcal{F}_{N,k}[\cdot]$ we deduce from Lemma 2.9 that it suffices to consider the error $g^e_k - \mathcal{F}_{N,k}[g^e_k]$.

The function g^e_k is a finite sum of functions $h(x)$ of the form $h_1(x_t)h_2(x_{\bar{t}})$, $t \in [d]$, $|t| < d$, where $h_1 \in \mathrm{H}^{2(k+K)}_{\mathrm{mix}}(-1,1)^{|t|}$ and $h_2 \in \mathrm{H}^{2(k+K)}_{\mathrm{mix}}(-1,1)^{|\bar{t}|}$. Using linearity once more, we deduce that it suffices to prove the result for h. In the usual manner, we consider the difference $\mathcal{F}^a_{N,k}[h] - \mathcal{F}_{N,k}[h]$, where $\mathcal{F}^a_{N,k}[h]$ is the approximate polynomial subtraction approximation of h. A simple argument verifies that

$$\mathcal{F}^a_{N,k}[h] = \mathcal{F}^a_{N,k}[h_1]\mathcal{F}^a_{N,k}[h_2], \quad \mathcal{F}_{N,k}[h] = \mathcal{F}_{N,k}[h_1]\mathcal{F}_{N,k}[h_2].$$

Noting that $a_1b_1 - a_2b_2 = (a_1 - a_2)b_1 + a_2(b_1 - b_2)$ we write

$$\mathcal{F}^a_{N,k}[h] - \mathcal{F}_{N,k}[h] = \left(\mathcal{F}^a_{N,k}[h_1] - \mathcal{F}_{N,k}[h_1]\right)\mathcal{F}^a_{N,k}[h_2] + \mathcal{F}_{N,k}[h_1]\left(\mathcal{F}^a_{N,k}[h_2] - \mathcal{F}_{N,k}[h_2]\right).$$

By induction

$$\begin{aligned}
\left\|D^{\alpha_t}\left(\mathcal{F}^a_{N,k}[h_1] - \mathcal{F}_{N,k}[h_1]\right)\right\|_\infty &\lesssim N^{|\alpha_t|_\infty - 2k-1}, \quad \|D^{\alpha_t}\mathcal{F}_{N,k}[h_1]\|_\infty \lesssim 1, \\
\left\|D^{\alpha_{\bar{t}}}\left(\mathcal{F}^a_{N,k}[h_2] - \mathcal{F}_{N,k}[h_2]\right)\right\|_\infty &\lesssim N^{|\alpha_{\bar{t}}|_\infty - 2k-1}, \quad \left\|D^{\alpha_{\bar{t}}}\mathcal{F}^a_{N,k}[h_2]\right\|_\infty \lesssim 1.
\end{aligned}$$

Hence

$$\left\|D^\alpha\left(\mathcal{F}^a_{N,k}[h] - \mathcal{F}_{N,k}[h]\right)\right\|_\infty \lesssim N^{|\alpha_t|_\infty - 2k-1} + N^{|\alpha_{\bar{t}}|_\infty - 2k-1} \lesssim N^{|\alpha|_\infty - 2k-1},$$

as required. The result for the $\mathrm{H}^q(\Omega)$ norm can be proved in an identical manner. □

As in the univariate setting, we arrive at the same conclusion: approximating jump values with Eckhoff's method does not deteriorate the convergence rate. However, additional smoothness is once more required for the multivariate version of Eckhoff's method over approximation by polynomial subtraction, unless the values $c(r)$, $r = 0, \ldots, k-1$, are distinct. Nonetheless, as we now consider, there is an advantage to choosing equal values $c(r)$: namely, a much faster convergence rate inside the domain Ω.

3 The auto-correction phenomenon

As demonstrated in Lemmas 1.1 and 2.3, the polynomial subtraction approximation has a convergence rate one power of N faster inside the domain than on the boundary. It turns out that, for the particular choice of the values $m(r) = N+r$, Eckhoff's approximation possesses the much faster convergence rate of $\mathcal{O}\left(N^{-3k-2}\right)$ away from the boundary—a full $\mathcal{O}\left(N^k\right)$ faster than the corresponding approximation based on exact jump values. This auto-correction phenomenon was observed numerically in [25] and proved in the univariate, Fourier case in [29]. The aim of this section is to extend this result to the multivariate modified Fourier setting.

In previous sections, we observed that Eckhoff's approximation decouples into terms corresponding to each particular value of i. The analysis of each such term can be handled separately, and, since the analysis is virtually identical, it suffices to consider only one particular value. For the remainder of this section, we assume that f only has non-zero modified Fourier coefficients when $i = (0, 0, \ldots, 0)$. Accordingly, we drop the $[i]$ superscript.

Since uniform convergence of Eckhoff's approximation on $\bar{\Omega}$ is guaranteed by Theorem 2.10, we write

$$f(x) - \mathcal{F}_{N,k}[f](x) = \sum_{n \notin I_N} \hat{v}_n \phi_n(x) = \sum_{t \in [d]} \sum_{\substack{n_j \geq N \\ j \in t}} \sum_{|n_{\bar{t}}|_\infty = 0}^{N-1} \hat{v}_n \phi_n(x), \quad x \in \bar{\Omega}, \tag{3.1}$$

where $v(x) = f(x) - g_k(x)$ and g_k is given by (2.13). Following the same method of proof as in [29], we seek to expand the right hand side of (3.1) using the so-called *Abel transformation*. Given a sequence $a_m \in \mathbb{R}$, $m \in \mathbb{N}$, we define the operator $\triangle_{r,n}$, $r, n \in \mathbb{N}$, by

$$\triangle_{0,n}[a_m] = a_n, \quad \triangle_{r+1,n}[a_m] = \triangle_{r,n}[a_m] + \triangle_{r,n+1}[a_m], \quad r, n \in \mathbb{N}.$$

It is easily seen that

$$\triangle_{r,n}[a_m] = \sum_{s=0}^{r} \begin{pmatrix} r \\ s \end{pmatrix} a_{n+s}, \quad r, n \in \mathbb{N}. \tag{3.2}$$

Now suppose that $a_m \in \mathbb{R}$, $m \in \mathbb{N}^d$. We write $\triangle^j_{r,n}$, $j = 1, .., d$, for the above operator acting on the j^{th} entry of n. Further, given $t \in [d]$, $r \in \mathbb{N}^{|t|}$ and $n \in \mathbb{N}^{|t|}$ we define $\triangle^t_{r,n}$ by the composition

$$\triangle^t_{r,n}[a_m] = \triangle^{t_1}_{r_{t_1},n_{t_1}} \left[\triangle^{t_2}_{r_{t_2},n_{t_2}} \left[\ldots \triangle^{t_{|t|}}_{r_{t_{|t|}},n_{t_{|t|}}} [a_m] \right] \right].$$

It follows from (3.2) that

$$\triangle^t_{r,n}[a_m] = \sum_{s_{t_1}=0}^{r_{t_1}} \cdots \sum_{s_{t_{|t|}}=0}^{r_{t_{|t|}}} \begin{pmatrix} r_{t_1} \\ s_{t_1} \end{pmatrix} \ldots \begin{pmatrix} r_{t_{|t|}} \\ s_{t_{|t|}} \end{pmatrix} a_{(n+s;m)}, \tag{3.3}$$

where $(n+s;m)$ has j^{th} entry $n_j + s_j$ if $j \in t$ and m_j otherwise.

Before using this transform, we need some additional notation. Given $x, y \in \mathbb{R}^d$ we write $x.y = x_1y_1 + \ldots x_dy_d$, and, if $y = (c, c, \ldots, c)$ has equal entries, just $x.c$. Moreover, given $u \in [t]^*$, $r_u \in \mathbb{N}^{|u|}$ and $k \in \mathbb{N}$ we define $(r_u; k) \in \mathbb{N}^{|t|}$ by the condition that the j^{th} entry of $(r_u; k)$, which we write $(r_u; k)_j$, takes value r_j if $j \in u$ and k otherwise.

Lemma 3.1. *Suppose that $g \in \mathrm{H}^1_{mix}(\Omega)$, $t \in [d]$ and that $x \in \Omega$. Then, for $k \in \mathbb{N}$ and $n_{\bar{t}} \in \mathbb{N}^{|\bar{t}|}$, we have*

$$\sum_{\substack{n_j \geq N \\ j \in t}} \hat{g}_n \phi_{n_t}(x_t) =$$

$$\mathrm{Re}\left\{ \sum_{u \in [t]^*} \sum_{|r_u|_\infty = 0}^{k} \mathrm{e}^{i\pi x_u . (N-1)} \prod_{j \in t} (1 + \mathrm{e}^{-i\pi x_j})^{-(r_u;k)_j - 1} \sum_{\substack{n_j \geq N \\ j \in \bar{u}}} \triangle^t_{(r_u;k+1),(n_{\bar{u}};N)} [\hat{g}_m] \mathrm{e}^{i\pi n_{\bar{u}} . x_{\bar{u}}} \right\},$$

where $(n_{\bar{u}}; N) \in \mathbb{N}^{|t|}$ has j^{th} entry n_j if $j \in \bar{u}$ and N otherwise and $m_{\bar{t}} = n_{\bar{t}}$.

Proof. We proceed by induction on $|t|$. Suppose first that $|t| = 1$ and, without loss of generality, that $d = 1$. The verification of the lemma in this case is very standard (see also [29]). We have

$$\sum_{n\geq N} \hat{g}_n \mathrm{e}^{in\pi x} = \sum_{n\geq N} \left(\triangle_{1,n}[\hat{g}_m] - \hat{g}_{n+1}\right) \mathrm{e}^{in\pi x} = \sum_{n\geq N} \triangle_{1,n}[\hat{g}_m]\mathrm{e}^{in\pi x} - \mathrm{e}^{-i\pi x} \sum_{n\geq N} \hat{g}_n \mathrm{e}^{in\pi x} + \hat{g}_N \mathrm{e}^{i(N-1)\pi x}.$$

Rearranging gives

$$\sum_{n\geq N} \hat{g}_n \mathrm{e}^{in\pi x} = \frac{\mathrm{e}^{i(N-1)\pi x}}{1+\mathrm{e}^{-i\pi x}} \hat{g}_N + \frac{1}{1+\mathrm{e}^{-i\pi x}} \sum_{n\geq N} \triangle_{1,n}[\hat{g}_m]\mathrm{e}^{in\pi x},$$

which provides the result for $k = 0$. Iterating this process yields the result for general k.

Now let $t \in [d]$ be of length $|t| \geq 2$. Write $t = (t_1, \ldots_{t_{|t|}})$ and $\tau = (t_2, \ldots_{t_{|t|}})$. We have

$$\sum_{\substack{n_j\geq N \\ j\in t}} \hat{g}_n \phi_{n_t}(x_t) = \sum_{n_{t_1}\geq N} \phi_{n_{t_1}}(x_{t_1}) \sum_{\substack{n_j\geq N \\ j\in \tau}} \hat{g}_n \phi_{n_\tau}(x_\tau).$$

By the induction hypothesis we obtain

$$\begin{aligned}\sum_{\substack{n_j\geq N \\ j\in t}} \hat{g}_n \phi_{n_t}(x_t) = \mathrm{Re} \sum_{n_{t_1}\geq N} \mathrm{e}^{in_{t_1}\pi x_{t_1}} \Bigg\{ \sum_{u\in[\tau]^*} \mathrm{e}^{i\pi x_u.(N-1)} \sum_{|r_u|_\infty=0}^{k} \prod_{j\in t} (1+\mathrm{e}^{-i\pi x_j})^{-(r_u;k)_j-1} \\ \times \sum_{\substack{n_j\geq N \\ j\in \bar{u}}} \triangle^{\tau}_{(r_u;k+1),(n_{\bar{u}};N)}[\hat{g}_m] \mathrm{e}^{i\pi n_{\bar{u}}.x_{\bar{u}}} \Bigg\} \\ = \mathrm{Re} \sum_{u\in[\tau]^*} \mathrm{e}^{i\pi x_u.(N-1)} \sum_{|r_u|_\infty=0}^{k} \prod_{j\in\tau} (1+\mathrm{e}^{-i\pi x_j})^{-(r_u;k)_j-1} \\ \times \sum_{\substack{n_j\geq N \\ j\in \bar{u}}} \mathrm{e}^{i\pi n_{\bar{u}}.x_{\bar{u}}} \sum_{n_{t_1}\geq N} \mathrm{e}^{in_{t_1}\pi x_{t_1}} \triangle^{\tau}_{(r_u;k+1),(n_{\bar{u}};N)}[\hat{g}_m]. \end{aligned} \tag{3.4}$$

Using the result for $|t| = 1$ yields

$$\begin{aligned}\sum_{n_{t_1}\geq N} \mathrm{e}^{in_{t_1}\pi x_{t_1}} \triangle^{\tau}_{(r_u;k+1),(n_{\bar{u}};N)}[\hat{g}_n] = \sum_{r_{t_1}=0}^{k} \mathrm{e}^{i\pi x_{t_1}(N-1)} (1+\mathrm{e}^{i\pi x_{t_1}})^{-r_{t_1}-1} \triangle^{t_1}_{r_{t_1},N} \left[\triangle^{\tau}_{(r_u;k+1),(n_{\bar{u}};N)}[\hat{g}_m]\right] \\ + \sum_{n_{t_1}\geq N} (1+\mathrm{e}^{-i\pi x_{t_1}})^{-k-1} \triangle^{t_1}_{k+1,n_{t_1}} \left[\triangle^{\tau}_{(r_u;k+1),(n_{\bar{u}};N)}[\hat{g}_m]\right]. \end{aligned} \tag{3.5}$$

If we substitute (3.5) into (3.4) we obtain the result. Note that if $v \in [t]^*$ then either $v \in [\tau]^*$ or $v = (t_1, u)$ for some $u \in [\tau]^*$. The two terms of (3.5) correspond respectively to these scenarios. □

The crux of the auto-correction phenomenon is the following trivial observation:

Lemma 3.2. *Suppose that $v = f - g_k$, where g_k is given by (2.13), and that the values $m(r) = N + r$, $r = 0, \ldots, k-1$. Then $\triangle^t_{r_t,n_t}[\hat{v}_m] = 0$ for all $|r_t|_\infty \leq k-1$, $|n_t|_\infty \leq N$, $|m_{\bar{t}}|_\infty \leq N$ and $t \in [d]$.*

Proof. By construction $\hat{v}_n = 0$ for $|n|_\infty \leq N + k - 1$. We now use (3.3) to obtain the result. □

We may now re-write (3.1) as

$$f(x) - \mathcal{F}_{N,k}[f](x) = \sum_{t\in[d]} \sum_{|n_{\bar{t}}|_\infty=0}^{N-1} h_{n_{\bar{t}}}(x_t)\phi_{n_{\bar{t}}}(x_{\bar{t}}), \tag{3.6}$$

where $h_{n_{\bar{t}}}(x_t)$ is obtained from the expansion derived in Lemma 3.1:

$$h_{n_{\bar{t}}}(x_t) = \mathrm{Re}\left\{ \sum_{u\in[t]^*} \sum_{|r_u|_\infty=0}^{k} \mathrm{e}^{i\pi x_u.(N-1)} \prod_{j\in t} (1+\mathrm{e}^{-i\pi x_j})^{-(r_u;k)_j-1} \sum_{\substack{n_j\geq N \\ j\in\bar{u}}} \triangle^{t}_{(r_u;k+1),(n_{\bar{u}};N)}[\hat{v}_m] \mathrm{e}^{i\pi n_{\bar{u}}.x_{\bar{u}}} \right\}.$$

Consider the term of $h_{n_{\bar{t}}}$ corresponding to $u = t$ separately. This is

$$\mathrm{e}^{i\pi x_t.(N-1)} \sum_{|r_t|_\infty=0}^{k} \prod_{j\in t} (1+\mathrm{e}^{-i\pi x_j})^{-r_j-1} \triangle^t_{r_t,N}[\hat{v}_m],$$

where we write $\triangle^t_{r_t,N}$ instead of the full expression $\triangle^t_{r_t,(N,\ldots,N)}$. Using Lemma 3.2, all terms of this expression where $|r_t|_\infty < k$ are zero. Hence, we define

$$H_{n_{\bar{t}}}(x_t) = \mathrm{e}^{i\pi x_t.(N-1)} \sum_{|r_t|_\infty=k} \prod_{j\in t} (1+\mathrm{e}^{-i\pi x_j})^{-r_j-1} \triangle^t_{r_t,N}[\hat{v}_m], \tag{3.7}$$

where $m_{\bar{t}} = n_{\bar{t}}$, and

$$G_{n_{\bar{t}}}(x_t) = \sum_{\substack{u\in[t]^*\\ u\neq t}} \sum_{|r_u|_\infty=0}^{k} \mathrm{e}^{i\pi x_u.(N-1)} \prod_{j\in t} (1+\mathrm{e}^{-i\pi x_j})^{-(r_u;k)-1} \sum_{\substack{n_j\geq N\\ j\in\bar{u}}} \triangle^t_{(r_u;k+1),(n_{\bar{u}};N)}[\hat{v}_m] \mathrm{e}^{i\pi n_{\bar{u}}.x_{\bar{u}}}, \tag{3.8}$$

so that the function h_{n_t} may be expressed as $h_{n_{\bar{t}}}(x_t) = \mathrm{Re}\left\{G_{n_{\bar{t}}}(x_t) + H_{n_{\bar{t}}}(x_t)\right\}$. To derive an estimate for the error $f(x) - \mathcal{F}_{N,k}[f](x)$ we first need bounds for the functions $G_{n_{\bar{t}}}$ and $H_{n_{\bar{t}}}$. We derive such bounds in the sequel. First, however, it is useful to consider the case $d = 1$ to demonstrate elements of the multivariate proof. This is given in a similar form in [29].

3.1 The case $d = 1$

For $d = 1$, using (3.1) and the characterisation given in Lemma 3.1 with $t = (1)$, we may write

$$\begin{aligned} f(x) - \mathcal{F}_{N,k}[f](x) &= \sum_{n\geq N} \hat{v}_n \phi_n(x) \\ &= \mathrm{Re}\left\{ \sum_{r=0}^{k} \frac{\mathrm{e}^{i(N-1)\pi x}}{(1+\mathrm{e}^{-i\pi x})^{r+1}} \triangle_{r,N}[\hat{v}_m] + \frac{1}{(1+\mathrm{e}^{-i\pi x})^{k+1}} \sum_{n\geq N} \triangle_{k+1,n}[\hat{v}_m] \mathrm{e}^{in\pi x} \right\}. \end{aligned}$$

In light of Lemma 3.2, $\triangle_{r,N}[\hat{v}_m] = 0$ for $r = 0, \ldots, k-1$, so this reduces to

$$\begin{aligned} f(x) - \mathcal{F}_{N,k}[f](x) &= \mathrm{Re}\left\{ \frac{\mathrm{e}^{i(N-1)\pi x}}{(1+\mathrm{e}^{-i\pi x})^{-k-1}} \triangle_{k,N}[\hat{v}_m] + \frac{1}{(1+\mathrm{e}^{-i\pi x})^{k+1}} \sum_{n\geq N} \triangle_{k+1,n}[\hat{v}_m] \mathrm{e}^{in\pi x} \right\} \\ &= \mathrm{Re}\left\{H(x) + G(x)\right\}, \end{aligned} \tag{3.9}$$

where $G(x)$ and $H(x)$ are the univariate forms of $G_{n_{\bar{t}}}$ and $H_{n_{\bar{t}}}$. Note that for $d = 1$ there is only one $t \in [d]$, namely $t = (1)$, and trivially $\bar{t} = \emptyset$.

We now seek bounds for G and H. To do so, we require the following lemma, which is given in a similar form in [29]:

Lemma 3.3. *Suppose that p_s, $s = 0, \ldots, k-1$ are the univariate Cardinal polynomials for the first k derivative conditions. Then*

$$\triangle_{r,n}[\widehat{p}_{sm}] = \widehat{p}_{sn} \frac{(2s+r+1)!(-1)^r}{(2s+1)!n^r} + \mathcal{O}\left(n^{-2s-r-3}\right), \quad \forall r \in \mathbb{N}, \quad s = 0, \ldots, k-1, \quad n \to \infty.$$

Proof. By construction, $\widehat{p}_{sm} = (-1)^m (m\pi)^{-2(s+1)}$. Using (3.2) we obtain

$$\begin{aligned} \triangle_{r,n}[\widehat{p}_{sm}] &= \sum_{l=0}^{r} \binom{r}{l} \frac{(-1)^{n+l}}{((n+l)\pi)^{2(s+1)}} = \frac{(-1)^n}{(n\pi)^{2(s+1)}} \sum_{l=0}^{r} \binom{r}{l} \frac{(-1)^l}{(1+\frac{l}{n})^{2(s+1)}} \\ &= \frac{(-1)^n}{(n\pi)^{2(s+1)}} \sum_{l=0}^{r} (-1)^l \binom{r}{l} \left\{ \sum_{p=0}^{r} \left(\frac{l}{n}\right)^p \binom{2s+p+1}{p} + \mathcal{O}\left(n^{-r-1}\right) \right\} \\ &= \widehat{p}_{sn} \sum_{p=0}^{r} n^{-p} \binom{2s+p+1}{p} \sum_{l=0}^{r} (-1)^l \binom{r}{l} l^p + \mathcal{O}\left(n^{-2s-r-3}\right). \end{aligned}$$

It is readily seen that

$$\sum_{l=0}^{r}(-1)^l \begin{pmatrix} r \\ l \end{pmatrix} l^p = \begin{cases} 0 & p = 0, \ldots, r-1, \\ (-1)^r r! & p = r. \end{cases}$$

Substituting this into the previous expression now gives the result. □

We now consider the coefficients $\hat{v}_m$. Using the asymptotic expansion (1.2) and the form of the univariate function g_k, we obtain

$$\hat{v}_n = \sum_{r=0}^{k-1} \left(\mathcal{A}_r[f] - \bar{\mathcal{A}}_r[f]\right) \widehat{p_r}_n + \sum_{r=k}^{k+K-1} \mathcal{A}_r[f]\widehat{p_r}_n + \mathcal{O}\left(n^{-2(k+K+1)}\right),$$

provided $f \in \mathrm{H}^{2(k+K+1)}(-1,1)$. In particular

$$\triangle_{s,n}[\hat{v}_m] = \sum_{r=0}^{k-1} \left(\mathcal{A}_r[f] - \bar{\mathcal{A}}_r[f]\right) \triangle_{s,n}[\widehat{p_r}_m] + \sum_{r=k}^{k+K-1} \mathcal{A}_r[f]\triangle_{s,n}[\widehat{p_r}_m] + \mathcal{O}\left(n^{-2(k+K+1)}\right).$$

By Theorem 1.4 and Lemma 3.3 we have

$$|\triangle_{s,n}[\hat{v}_m]| \lesssim \sum_{r=0}^{k-1} N^{2(r-k)}\bar{n}^{-2r-s-2} + \bar{n}^{-2k-s-2} + \bar{n}^{-2(k+K+1)}.$$

In particular, provided $2K \geq k+1$, we obtain $|\triangle_{k,N}[\hat{v}_m]| \lesssim N^{-3k-2}$ and $|\triangle_{k+1,n}[\hat{v}_m]| \lesssim N^{-3k-1}n^{-2}$. Recalling the definitions of G and H given in (3.9), this yields

$$|G(x)| \lesssim N^{-3k-2}, \quad |H(x)| \lesssim N^{-3k-2}, \quad x \in (-1,1),$$

provided $f \in \mathrm{H}^{3(k+1)}(-1,1)$. From this we immediately obtain the univariate result:

Theorem 3.4. *Suppose that $\mathcal{F}_{N,k}[f]$ is the univariate Eckhoff approximation of $f \in \mathrm{H}^{3(k+1)}(\Omega)$ using the values $m(r) = N + r$, $r = 0, \ldots, k-1$. Then $f(x) - \mathcal{F}_{N,k}[f](x)$ is $\mathcal{O}\left(N^{-3k-2}\right)$ uniformly for x in compact subsets of Ω.*

3.2 Bounds for $G_{n_{\bar{t}}}$ and $H_{n_{\bar{t}}}$

We commence with the following preliminary result:

Lemma 3.5. *Suppose that $t \in [d]$, $r_t \in \mathbb{N}^{|t|}$, $2K \geq k+1$ and that the function $h \in \mathrm{H}^{2(k+K)+1}_{mix}(\Omega)$, satisfies the first k derivative conditions. Then*

$$\left|\triangle_{r_t,n_t}[\hat{h}_n]\right| \lesssim \bar{n}^{-2k-2}\prod_{j\in t}\bar{n}_j^{-2r_j} = \bar{n}^{-2k-2}\bar{n}_t^{-2r_t}.$$

Proof. It suffices to consider $t = (1, \ldots, d)$ and use induction on d. Consider $d = 1$ and a univariate function h. Since h obeys the first k derivative conditions, we have

$$\hat{h}_n = \sum_{s=k}^{k+K-1} \mathcal{A}_s[h]\widehat{p_s}_n + \mathcal{O}\left(n^{-2(k+K)-1}\right),$$

Hence, using Lemma 3.3, we obtain

$$\left|\triangle_{r,n}[\hat{h}_n]\right| \lesssim \sum_{s=k}^{k+K-1} |\mathcal{A}_s[h]\triangle_{r,n}[\widehat{p_s}_n]| + \bar{n}^{-2(k+K)-1} \lesssim \bar{n}^{-r-2k-2} + \bar{n}^{-2(k+K)-1}.$$

This gives the result for $d = 1$. Now assume that the result holds for all functions of at most $(d-1)$ variables. Then, if h is function of d variables and $t = (1, \ldots, d)$, we have

$$\begin{aligned}\triangle^t_{r_t,n_t}[\hat{h}_n] &= \sum_{u\in[d]}\sum_{|s_u|_\infty=k}^{k+K-1} \triangle^t_{r_t,n_t}\left[\mathcal{A}_{s_u,n_{\bar{u}}}[h]\widehat{p_{s_u}}_{n_u}\right] + \mathcal{O}\left(n^{-2(k+K)-1}\right) \\ &= \sum_{u\in[d]}\sum_{|s_u|_\infty=k}^{k+K-1} \triangle^{\bar{u}}_{r_{\bar{u}},n_{\bar{u}}}\left[\mathcal{A}_{s_u,n_{\bar{u}}}[h]\right]\triangle^u_{r_u,n_u}\left[\widehat{p_{s_u}}_{n_u}\right] + \mathcal{O}\left(n^{-2(k+K)-1}\right).\end{aligned}$$

Using Lemma 3.3, we deduce that

$$\left|\triangle^u_{r_u,n_u}\left[\widehat{p_{s_u n_u}}\right]\right| \lesssim \prod_{j\in u}\left|\triangle^j_{r_j,n_j}[\widehat{p_{s_j n_j}}]\right| \lesssim \bar{n}_u^{-r_u-2s_u-2}. \tag{3.10}$$

Furthermore $\mathcal{A}_{s_u,n_{\bar{u}}}[h]$ is the modified Fourier coefficient of a function of the variables $x_{\bar{u}}$ that satisfies the first k derivative conditions. Since $|\bar{u}| < d$, we may use the induction hypothesis and (3.10) to give

$$\left|\triangle^t_{r_t,n_t}[\hat{h}_n]\right| \lesssim \sum_{u\in[d]}\sum_{|s_u|_\infty=k}^{k+K-1} \bar{n}_{\bar{u}}^{-r_{\bar{u}}-2k-2}\bar{n}_u^{-r_u-2s_u-2} + \bar{n}^{-2(k+K)-1} \lesssim \bar{n}_t^{-r_t}\bar{n}^{-2k-2},$$

as required. □

With this in hand, we may estimate the functions $G_{n_{\bar{t}}}$ and $H_{n_{\bar{t}}}$. We have:

Lemma 3.6. *Suppose that $f \in \mathrm{H}^{3(k+1)}_{mix}(\Omega)$. Then the function $H_{n_{\bar{t}}}$ defined by (3.7) satisfies $|H_{n_{\bar{t}}}(x_t)| \lesssim N^{-3k-2}\bar{n}_{\bar{t}}^{-2}$ uniformly for x_t in compact subsets of $(-1,1)^{|t|}$.*

Proof. We first observe that, for $n \in \mathbb{N}^d$ such that $n_j \geq N$ whenever $j \in t$ and $n_j = 0,\dots,N-1$ otherwise, $\hat{v}_n$ satisfies

$$\hat{v}_n = \sum_{|s_t|_\infty=0}^{k-1} \mathcal{E}_{s_t,n_{\bar{t}}}[f]\widehat{p_{s_t n_t}} + \sum_{\substack{v\in[d]^*\\ t\not\subseteq v}}\sum_{|s_v|_\infty=0}^{k-1} \mathcal{A}_{s_v,n_{\bar{v}}}[f]\widehat{p_{s_v n_v}}. \tag{3.11}$$

We now substitute the two terms of (3.11) into the definition of $H_{n_{\bar{t}}}$ given in (3.7) and consider them separately. For the first term we have

$$\triangle^t_{r_t,N}\left[\mathcal{E}_{s_t,n_{\bar{t}}}[f]\widehat{p_{s_t n_t}}\right] = \mathcal{E}_{s_t,n_{\bar{t}}}[f]\triangle^t_{r_t,N}\left[\widehat{p_{s_t n_t}}\right] = \mathcal{E}_{s_t,n_{\bar{t}}}[f]\prod_{j\in t}\triangle^j_{r_j,N}[\widehat{p_{s_j n_j}}].$$

Using Lemma 2.6 and (3.10) we obtain the bound

$$\left|\triangle^t_{r_t,N}\left[\mathcal{E}_{s_t,n_{\bar{t}}}[f]\widehat{p_{s_t n_t}}\right]\right| \lesssim N^{2(|s_t|_\infty-k)}\prod_{j\in t}N^{-2s_j-r_j-2}\bar{n}_{\bar{t}}^{-2} \lesssim N^{-2k-|r_t|-2|t|}\bar{n}_{\bar{t}}^{-2}.$$

Since $|r_t| \geq |r_t|_\infty = k$ and $|t| \geq 1$, we obtain the required bound for the first term.

Now consider the second term of (3.11) substituted into (3.7). For $v \in [d]^*$ with $t \not\subseteq v$ either (i) $v \cap t \neq \emptyset$ or (ii) $v \cap t = \emptyset$. Consider case (i) first. We have

$$\triangle^t_{r_t,N}\left[\mathcal{A}_{s_v,n_{\bar{v}}}[f]\widehat{p_{s_v n_v}}\right] = \triangle^{t\setminus v}_{r_{t\setminus v},N}\left[\mathcal{A}_{s_v,n_{\bar{v}}}[f]\right]\triangle^{t\cap v}_{r_{t\cap v},N}[\widehat{p_{s_v n_v}}].$$

Since $\mathcal{A}_{s_v,n_{\bar{v}}}[f] = \hat{h}_{n_{\bar{v}}}$, where h is a function of $x_{\bar{v}}$ that obeys the first k derivative conditions, we may apply Lemma 3.5 to give

$$\begin{aligned}\left|\triangle^t_{r_t,N}\left[\mathcal{A}_{s_v,n_{\bar{v}}}[f]\widehat{p_{s_v n_v}}\right]\right| &\lesssim \prod_{j\in t\cap v}N^{-2s_j-r_j-2}\prod_{j\in t\setminus v}N^{-2k-r_j-2}\bar{n}_{v\setminus t}^{-2s_{v\setminus t}-2}\bar{n}_{\overline{t\cup v}}^{-2k-2}\\ &\lesssim N^{-|r_t|-2|t\cap v|-2(k+1)(|t\setminus v|)}\bar{n}_{\bar{t}}^{-2} \lesssim N^{-3k-2}\bar{n}_{\bar{t}}^{-2}.\end{aligned}$$

Here the final inequality follows since, by assumption, $|t\cap v|, |t\setminus v| \geq 1$. Now consider case (ii). Since $t \cap v = \emptyset$, we have

$$\triangle^t_{r_t,N}\left[\mathcal{A}_{s_v,n_{\bar{v}}}[f]\widehat{p_{s_v n_v}}\right] = \triangle^t_{r_t,N}\left[\mathcal{A}_{s_v,n_{\bar{v}}}[f]\right]\widehat{p_{s_v n_v}}.$$

Using Lemma 3.5 and (3.10) we obtain

$$\left|\triangle^t_{r_t,N}\left[\mathcal{A}_{s_v,n_{\bar{v}}}[f]\widehat{p_{s_v n_v}}\right]\right| \lesssim \prod_{j\in t}N^{-r_j-2k-2}\prod_{j\notin v\cup t}\bar{n}_j^{-2k-2}\prod_{j\in v}\bar{n}_j^{-2s_j-2} \lesssim N^{-|r_t|_\infty-2k-2}\bar{n}_{\bar{t}}^{-2} \lesssim N^{-3k-2}\bar{n}_{\bar{t}}^{-2}.$$

This completes the proof. □

We now derive a bound for $G_{n_{\bar{t}}}$:

Lemma 3.7. *Suppose that $f \in \mathrm{H}^{3(k+1)}_{mix}(\Omega)$. Then the function $G_{n_{\bar{t}}}$ defined by (3.8) satisfies $|G_{n_{\bar{t}}}(x_t)| \lesssim N^{-3k-2}\bar{n}_{\bar{t}}^{-2}$ uniformly for x_t in compact subsets of $(-1,1)^{|t|}$.*

Proof. Since $x_t \in (-1,1)^{|t|}$ it suffices to bound

$$\sum_{\substack{u\in[t]^*\\ u\neq t}} \sum_{|r_u|_\infty=0}^{k} \sum_{\substack{n_j\geq N\\ j\in\bar{u}}} \left| \triangle^t_{(r_u;k+1),(n_{\bar{u}};N)}[\hat{v}_m] \right|, \tag{3.12}$$

by the $N^{-3k-2}\bar{n}_{\bar{t}}^{-2}$. To do so, we substitute the two terms of (3.11) into (3.12) and consider them separately. For the first term we have

$$\sum_{\substack{u\in[t]^*\\ u\neq t}} \sum_{|r_u|_\infty=0}^{k} \sum_{\substack{n_j\geq N\\ j\in\bar{u}}} \sum_{|s_t|_\infty=0}^{k-1} \left| \triangle^t_{(r_u;k+1),(n_{\bar{u}};N)} \left[\mathcal{E}_{s_t,n_{\bar{t}}}[f]\widehat{p_{s_t n_t}}\right] \right|. \tag{3.13}$$

Since $u \subseteq t$, we observe that

$$\triangle^t_{(r_u;k+1),(n_{\bar{u}};N)} \left[\mathcal{E}_{s_t,n_{\bar{t}}}[f]\widehat{p_{s_t n_t}}\right] = \mathcal{E}_{s_t,n_{\bar{t}}}[f] \prod_{j\in u} \triangle^j_{r_j,N}[\widehat{p_{s_j n_j}}] \prod_{j\in t\setminus u} \triangle^j_{k+1,n_j}[\widehat{p_{s_j n_j}}].$$

Using Lemmas 2.6 and (3.10) we deduce that

$$\left| \triangle^t_{(r_u;k+1),(n_{\bar{u}};N)} \left[\widehat{p_{s_t n_t}}\mathcal{E}_{s_t,n_{\bar{t}}}[f]\right] \right| \lesssim N^{2(|s_t|_\infty-k)}\bar{n}_{\bar{t}}^{-2} \prod_{j\in u} N^{-2s_j-r_j-2}\bar{n}_{\bar{u}}^{-2s_{\bar{u}}-k-3}.$$

Substituting this into (3.13) we obtain

$$\begin{aligned}
&\sum_{\substack{u\in[t]^*\\ u\neq t}} \sum_{|r_u|_\infty=0}^{k} \sum_{\substack{n_j\geq N\\ j\in\bar{u}}} \sum_{|s_t|_\infty=0}^{k-1} \left| \triangle^t_{(r_u;k+1),(n_{\bar{u}};N)} \left[\widehat{p_{s_t n_t}}\mathcal{E}_{s_t,n_{\bar{t}}}[f]\right] \right| \\
&\quad\lesssim \sum_{\substack{u\in[t]^*\\ u\neq t}} \sum_{|r_u|_\infty=0}^{k} \sum_{\substack{n_j\geq N\\ j\in\bar{u}}} \sum_{|s_t|_\infty=0}^{k-1} N^{2(|s_t|_\infty-k)} \prod_{j\in u} N^{-2s_j-r_j-2}\bar{n}_{\bar{t}}^{-2}\bar{n}_{\bar{u}}^{-2s_{\bar{u}}-k-3} \\
&\quad\lesssim \bar{n}_{\bar{t}}^{-2} \sum_{\substack{u\in[t]^*\\ u\neq t}} \sum_{|r_u|_\infty=0}^{k} \sum_{|s_t|_\infty=0}^{k-1} N^{2(|s_t|_\infty-k)} \prod_{j\in u} N^{-2s_j-r_j-2} \prod_{j\in t\setminus u} N^{-2s_j-k-2} \\
&\quad\lesssim \bar{n}_{\bar{t}}^{-2} \sum_{\substack{u\in[t]^*\\ u\neq t}} \sum_{|r_u|_\infty=0}^{k} N^{-2k-|r_u|-2|u|-(k+2)(|t|-|u|)} \lesssim N^{-3k-2}\bar{n}_{\bar{t}}^{-2},
\end{aligned}$$

as required. Here the last inequality follows by noting that $|t|-|u| \geq 1$.

We now consider the second term of (3.11) substituted into (3.12):

$$\sum_{\substack{v\in[d]^*\\ t\not\subseteq v}} \sum_{|s_v|_\infty=0}^{k-1} \sum_{\substack{u\in[t]^*\\ u\neq t}} \sum_{|r_u|_\infty=0}^{k} \sum_{\substack{n_j\geq N\\ j\in\bar{u}}} \left| \triangle^t_{(r_u;k+1),(n_{\bar{u}};N)} \left[\mathcal{A}_{s_v,n_{\bar{v}}}[f]\widehat{p_{s_v n_v}}\right] \right|. \tag{3.14}$$

As in the proof of Lemma 3.6, we split this into two cases: either (i) $v\cap t \neq \emptyset$ or (ii) $v\cap t = \emptyset$. Suppose that we consider case (i). Since $v\cap t \neq \emptyset$ we obtain

$$\triangle^t_{(r_u;k+1),(n_{\bar{u}};N)} \left[\mathcal{A}_{s_v,n_{\bar{v}}}[f]\widehat{p_{s_v n_v}}\right] = \triangle^{t\cap v}_{(r_{u\cap v};k+1),(n_{\bar{u}\cap v};N)} \left[\widehat{p_{s_v n_v}}\right] \triangle^{t\cap\bar{v}}_{(r_{u\cap\bar{v}};k+1),(n_{\bar{u}\cap\bar{v}};N)} \left[\mathcal{A}_{s_v,n_{\bar{v}}}[f]\right].$$

We have

$$\left| \triangle^{t\cap v}_{(r_{u\cap v};k+1),(n_{\bar{u}\cap v};N)} \left[\widehat{p_{s_v n_v}}\right] \right| \lesssim \prod_{j\in u\cap v} N^{-2s_j-r_j-2}\bar{n}_{\bar{u}\cap v}^{-2s_{\bar{u}\cap v}-k-3}\bar{n}_{v\setminus t}^{-2s_{v\setminus t}-2}.$$

Furthermore

$$\left| \triangle^{t\cap\bar{v}}_{(r_{u\cap\bar{v}};k+1),(n_{\bar{u}\cap\bar{v}};N)} \left[\mathcal{A}_{s_v,n_{\bar{v}}}[f]\right] \right| \lesssim \prod_{j\in u\cap\bar{v}} N^{-2k-r_j-2}\bar{n}_{\bar{u}\cap\bar{v}}^{-3k-3}\bar{n}_{\overline{t\cup v}}^{-2k-2}.$$

Combining these two estimates yields

$$\left|\triangle^t_{(r_u;k+1),(n_{\bar{u}};N)}\left[\mathcal{A}_{s_v,n_{\bar{v}}}[f]\widehat{p_{s_v}}_{n_v}\right]\right| \lesssim \prod_{j\in u\cap v} N^{-2s_j-r_j-2}\prod_{j\in u\cap\bar{v}} N^{-2k-r_j-2}\bar{n}_{\bar{u}\cap v}^{-2s_{\bar{u}\cap v}-k-3}\bar{n}_{\bar{u}\cap\bar{v}}^{-3k-3}\bar{n}_{\bar{t}}^{-2}.$$

Hence

$$\begin{aligned}
&\left|\sum_{|s_v|_\infty=0}^{k-1}\sum_{\substack{u\in[t]^*\\u\neq t}}\sum_{|r_u|_\infty=0}^{k}\sum_{\substack{n_j\geq N\\ j\in\bar{u}}}\triangle^t_{(r_u;k+1),(n_{\bar{u}};N)}\left[\mathcal{A}_{s_v,n_{\bar{v}}}[f]\widehat{p_{s_v}}_{n_v}\right]\right| \\
&\qquad\lesssim \sum_{|s_v|_\infty=0}^{k-1}\sum_{\substack{u\in[t]^*\\u\neq t}}\sum_{|r_u|_\infty=0}^{k}\prod_{j\in u\cap v} N^{-2s_j-r_j-2}\prod_{j\in u\cap\bar{v}} N^{-2k-r_j-2}\prod_{j\in\bar{u}\cap v} N^{-2s_j-k-2}\prod_{j\in\bar{u}\cap\bar{v}} N^{-3k-2}\bar{n}_{\bar{t}}^{-2} \\
&\qquad\lesssim \sum_{\substack{u\in[t]^*\\u\neq t}} N^{-2(k+1)|u\cap\bar{v}|}N^{-(k+2)|\bar{u}\cap v|}N^{-(3k+2)|\bar{u}\cap\bar{v}|}\bar{n}_{\bar{t}}^{-2}.
\end{aligned}$$

We claim that this term is $\lesssim N^{-3k-2}\bar{n}_{\bar{t}}^{-2}$. We have two possibilities: either $\bar{u}\cap\bar{v}\neq\emptyset$ or $\bar{u}\cap\bar{v}=\emptyset$. If $\bar{u}\cap\bar{v}\neq\emptyset$ then the result follows immediately. Now suppose that $\bar{u}\cap\bar{v}=\emptyset$. In this case, it follows that $u\cap\bar{v}\neq\emptyset$ and $\bar{u}\cap v\neq\emptyset$. Hence we also obtain the result. This completes case (i).

Next consider case (ii). Since $v\cap t=\emptyset$ we have

$$\triangle^t_{(r_u;k+1),(n_{\bar{u}};N)}\left[\mathcal{A}_{s_v,n_{\bar{v}}}[f]\widehat{p_{s_v}}_{n_v}\right] = \triangle^t_{(r_u;k+1),(n_{\bar{u}};N)}\left[\mathcal{A}_{s_v,n_{\bar{v}}}[f]\right]\widehat{p_{s_v}}_{n_v}.$$

In the standard manner we obtain

$$\begin{aligned}
\left|\triangle^t_{(r_u;k+1),(n_{\bar{u}};N)}\left[\mathcal{A}_{s_v,n_{\bar{v}}}[f]\widehat{p_{s_v}}_{n_v}\right]\right| &\lesssim \prod_{j\in u} N^{-2k-r_j-2}\bar{n}_{\bar{u}}^{-3k-3}\bar{n}_{\bar{v}\setminus t}^{-2k-2}\bar{n}_v^{-2s_v-2} \\
&\lesssim N^{-2(k+1)|u|-|r_u|_\infty}\bar{n}_{\bar{u}}^{-3k-3}\bar{n}_{\bar{t}}^{-2}.
\end{aligned}$$

Hence, in this case

$$\sum_{|s_v|_\infty=0}^{k-1}\sum_{\substack{u\in[t]^*\\u\neq t}}\sum_{|r_u|_\infty=0}^{k}\sum_{\substack{n_j\geq N\\ j\in\bar{u}}}\left|\triangle^t_{(r_u;k+1),(n_{\bar{u}};N)}\left[\mathcal{A}_{s_v,n_{\bar{v}}}[f]\widehat{p_{s_v}}_{n_v}\right]\right| \lesssim \prod_{j\in\bar{u}} N^{-3k-2}\bar{n}_{\bar{t}}^{-2} \lesssim N^{-3k-2}\bar{n}_{\bar{t}}^{-2},$$

where the final inequality follows since $|\bar{u}|\geq 1$. This completes the proof. □

3.3 Analysis of the auto-correction phenomenon and numerical results

We may now prove the key result of this section:

Theorem 3.8. *Suppose that $\mathcal{F}_{N,k}[f]$ is the multivariate Eckhoff approximation of $f\in \mathrm{H}^{3(k+1)}_{mix}(\Omega)$ using the values $m(r)=N+r$, $r=0,\ldots,k-1$. Then $f(x)-\mathcal{F}_{N,k}[f](x)$ is $\mathcal{O}\left(N^{-3k-2}\right)$ uniformly for x in compact subsets of Ω.*

Proof. Substituting the bounds derived in Lemmas 3.6 and 3.7 into the expansion (3.6) immediately yields the result. □

Though the analysis in this section was carried out for the approximation based on Cardinal polynomials, it is a simple exercise to extend it to the general subtraction bases described in Section 1. Hence, we have established the existence of an auto-correction phenomenon for arbitrary dimension d and arbitrary subtraction basis. Note that for the auto-correction phenomenon we require $f\in \mathrm{H}^{3(k+1)}_{\mathrm{mix}}(\Omega)$, rather than just $f\in \mathrm{H}^{3k+1}_{\mathrm{mix}}(\Omega)$ or $f\in \mathrm{H}^{3k+2}_{\mathrm{mix}}(\Omega)$ for uniform convergence (see Theorem 2.10). This extra smoothness condition is also present for polynomial subtraction: here $\mathrm{H}^{2k+3}_{\mathrm{mix}}(\Omega)$-regularity is required to obtain an $\mathcal{O}\left(N^{-2k-2}\right)$ error away from the boundary, rather than just $\mathrm{H}^{2k+2}_{\mathrm{mix}}(\Omega)$-regularity for uniform convergence [1]. In [29] the author demonstrates that slightly different smoothness assumptions can be imposed, depending on whether k is even or odd. For simplicity, we do not make this distinction.

For general values $m(r)$ it can be shown using identical methods that an auto-correction phenomenon is present provided the first $l\leq k$ values are chosen so that $m(r)=N+r$, $r=0,\ldots,l-1$. In this case,

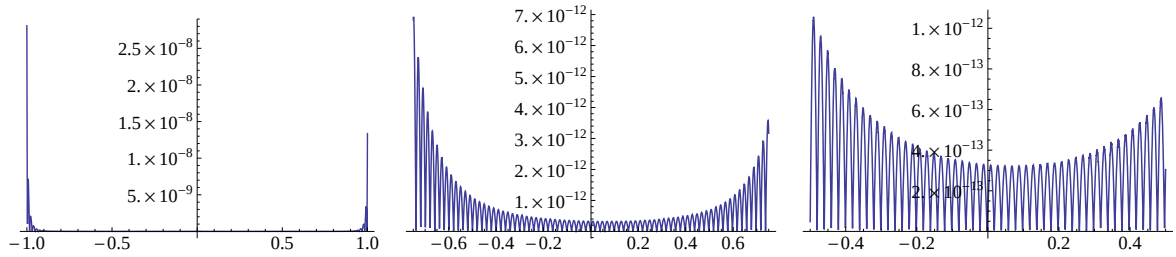


Figure 3: Graphs of $|f(x) - \mathcal{F}_{N,k}[f](x)|$ for $-1 \leq x \leq 1$ (left), $-0.75 \leq x \leq 0.75$ (middle) and $-0.5 \leq x \leq 0.5$ (right), where $N = 50$, $k = 2$ and $f(x) = x^2 \sin 5x + \cos 6x$.

(x_1, x_2)	$N = 10$	$N = 20$	$N = 30$	$N = 40$	$N = 50$
$(1, 1)$	4.958×10^{-8}	1.307×10^{-10}	3.799×10^{-12}	3.022×10^{-13}	4.202×10^{-14}
$(-1, -1)$	6.341×10^{-8}	1.372×10^{-10}	3.723×10^{-12}	2.861×10^{-13}	3.898×10^{-14}
$(\frac{1}{2}, \frac{2}{3})$	1.189×10^{-12}	4.293×10^{-15}	2.039×10^{-19}	4.673×10^{-19}	1.485×10^{-20}
$(0, 0)$	9.542×10^{-13}	1.885×10^{-16}	9.473×10^{-19}	2.037×10^{-20}	1.002×10^{-21}

Table 2: Pointwise error $|f(x_1, x_2) - \mathcal{F}_{N,k}[f](x_1, x_2)|$ for various values of (x_1, x_2) and N, where $k = 4$ and $f(x_1, x_2) = \left(\mathrm{e}^{3x_1} + \mathrm{e}^{-4x_1}\right)\left(\sin 5x_2 + \frac{1}{2}\right)$. Results to 4 significant figures.

the convergence rate away from the boundary is $\mathcal{O}\left(N^{-2k-l-2}\right)$. In particular, if $m(0) = N$, as is the case with the choices (1.15) and (1.16), then the convergence rate is $\mathcal{O}\left(N^{-2k-3}\right)$.

The auto-correction phenomenon is also exhibited by the error $f - \mathcal{F}_{N,k}[f]$ measured in the $\mathrm{L}^2(\Omega')$ norm, where Ω' is some set compactly contained in Ω. This has been studied in the univariate, Fourier case in [29]. The extension to the multivariate, modified Fourier setting is straightforward.

In Figure 3, we demonstrate the univariate auto-correction phenomenon. For the particular choice of function and parameters, the error at the endpoints is roughly 10^{-8}, whereas in the interval $[-0.5, 0.5]$ this figure is 10^{-12}. In Table 2, we present numerical results for the auto-correction phenomenon in the bivariate setting (here and henceforth, calculations are performed with additional precision, where necessary). Once more, we observe that the error inside the domain is much smaller than on the boundary.

4 Hyperbolic cross index sets and Eckhoff's method

Thus far the approximation $\mathcal{F}_{N,k}[f]$ has been based on the full index set (2.2). Though arguably the most natural index set to consider, it turns out that the truncated expansion $\mathcal{F}_N[f]$ includes a large number of terms that have an insignificant contribution to the overall sum. In view of this, an alternative approach to define I_N is to include only those terms in $\mathcal{F}_N[f]$ that are greater in absolute value than some tolerance ϵ. This is the idea of hyperbolic cross index sets [3, 32]. In many applications, modified Fourier series included, such an approach leads to a greatly reduced index set of size $|I_N| = \mathcal{O}\left(N(\log N)^{d-1}\right)$, and thereby effects a significant saving in computational effort. Moreover, the approximation $\mathcal{F}_N[f]$ converges to f at a rate comparable to that of the corresponding approximation based on the full index set (2.2). In this section, we consider the use of such a set in Eckhoff's method.

4.1 A hyperbolic cross for modified Fourier coefficients

To develop a hyperbolic cross index set for modified Fourier coefficients, we need an estimate for $|\hat{f}_n^{[i]}|$. This is provided by the bound $|\hat{f}_n^{[i]}| \lesssim (\bar{n}_1 \ldots \bar{n}_d)^{-2}$. If we set $\epsilon = N^{-2}$, then the term $\hat{f}_n^{[i]}$ is included in $\mathcal{F}_N[f]$ only if $\bar{n}_1 \ldots \bar{n}_d < N$. This leads to a hyperbolic cross index set:

$$I_N = \left\{n \in \mathbb{N}^d : \bar{n}_1 \ldots \bar{n}_d < N\right\}. \tag{4.1}$$

This set, in conjunction with modified Fourier series, has been investigated in [1, 16]. It is elementary to show that $|I_N| = \mathcal{O}\left(N(\log N)^{d-1}\right)$; a vast reduction over the full index set (2.2), for which this value is $\mathcal{O}\left(N^d\right)$. Furthermore, we have the following result, proved in [1]:

k	10^{-2}	10^{-4}	10^{-6}	10^{-8}	10^{-10}	10^{-12}	10^{-14}
1	121	1521	31329	—	—	—	—
	89	513	3053	17461	97241	—	—
2	49	121	561	1849	10201	60025	—
	49	105	297	841	2269	6269	17501
3	81	121	169	441	1225	3969	13689
	81	117	193	353	697	1333	2773
4	81	121	169	289	529	1089	2401
	81	121	165	257	397	593	1005
5	121	121	169	289	361	625	1089
	121	121	169	273	329	493	789

Table 3: Number of terms in the full (top value) and hyperbolic cross (bottom value) index set versions of Eckhoff's approximation applied to the function $f(x,y) = \mathrm{e}^{2x}(\cos 3y + \sin 2y)$ required to obtain an accuracy of $\|f - \mathcal{F}_{N,k}[f]\|_\infty < 10^{-2j}$ for $j = 1, 2, \ldots, 7$ (the dash indicates where more than $100,000$ terms are required to obtain the prescribed tolerance).

Theorem 4.1. *Suppose that $f \in \mathrm{H}^{2k+2}_{mix}(-1,1)^2$ and that $\mathcal{F}^e_{N,k}[f]$ is the exact polynomial subtraction approximation of f based on the hyperbolic cross index set (4.1). Then*

$$\|f - \mathcal{F}^e_{N,k}[f]\|_0 = \mathcal{O}\left(N^{-2k-\frac{3}{2}}(\log N)^{\frac{d-1}{2}}\right), \quad \|f - \mathcal{F}^e_{N,k}[f]\|_q = \mathcal{O}\left(N^{q-2k-\frac{3}{2}}\right), \quad q = 1, \ldots, 2k+1,$$

and $\|\mathrm{D}^\alpha(f - \mathcal{F}^e_{N,k}[f])\|_\infty = \mathcal{O}\left(N^{|\alpha|_\infty - 2k-1}(\log N)^{d-1}\right)$ for $|\alpha|_\infty \leq 2k$. If, additionally, $f \in \mathrm{H}^{2k+3}_{mix}(\Omega)$ then $\mathrm{D}^\alpha f(x) - \mathrm{D}^\alpha \mathcal{F}^e_{N,k}[f](x)$ is $\mathcal{O}\left(N^{|\alpha|_\infty - 2k-2}(\log N)^{d-1}\right)$ uniformly in compact subsets of Ω.

In view of Theorem 2.2, we conclude that replacing the full index set (2.2) by (4.1) does not affect the convergence rate of the approximation, aside from possibly a logarithmic factor (note that setting $k = 0$ in the above theorem establishes the convergence rate of $\mathcal{F}_N[f]$ to f). This, combined with the significant reduction in number of expansion terms, makes hyperbolic cross index sets greatly beneficial.

4.2 The hyperbolic cross version of Eckhoff's method

Given $n \in \mathbb{N}^d$ we define $|n|_0 = \bar{n}_1 \ldots \bar{n}_d$ so that the hyperbolic cross index set (4.1) includes only those n with $|n|_0 < N$. To adapt the multivariate version of Eckhoff's method to utilise the hyperbolic cross, we first replace the function g_k defined in (2.13) by

$$g_k(x) = \sum_{i \in \{0,1\}^d} \sum_{t \in [d]} \sum_{|r_t|_\infty = 0}^{k-1} \sum_{|n_{\bar{t}}|_0 = 0}^{N-1} \bar{\mathcal{A}}^{[i]}_{r_t, n_{\bar{t}}}[f] p^{[i_t]}_{r_t}(x_t) \phi^{[i_{\bar{t}}]}_{n_{\bar{t}}}(x_{\bar{t}}), \tag{4.2}$$

with unknowns $\bar{\mathcal{A}}^{[i]}_{r_t, n_{\bar{t}}}[f]$ that enforce the conditions $\widehat{g_k}^{[i]}_n = \hat{f}^{[i]}_n$, $n \in M_k$, where M_k is the index set

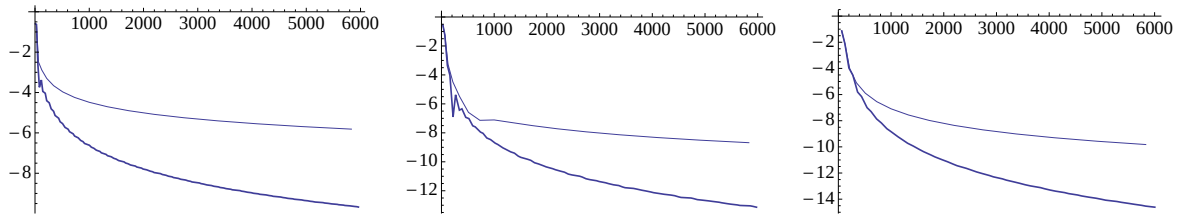
$$M_k = \bigcup_{t \in [d]} \left\{ n = (n_1, \ldots, n_d) \in \mathbb{N}^d : n_j = m(r_j),\ r_j = 0, \ldots, k-1, j \in t,\ |n_{\bar{t}}|_0 < N \right\}.$$

Note that the only difference in the definitions of g_k and M_k is the replacement of $|n_{\bar{t}}|_\infty$ by $|n_{\bar{t}}|_0$. We now define the approximation $\mathcal{F}_{N,k}[f]$ in the standard manner: namely, $\mathcal{F}_{N,k}[f] = \mathcal{F}_N[f - g_k] + g_k$, where $\mathcal{F}_N[f - g_k]$ is based on the index set (4.1).

For $d = 2$, there is no difference between the functions (2.13) and (4.2). The only difference between the two resulting approximations arises from the index set used in $\mathcal{F}_N[\cdot]$. However, for $d \geq 3$, the functions (2.13) and (4.2) are distinct, leading to further savings in the number of approximation terms.

It is readily seen that the operational cost of forming the hyperbolic cross version of Eckhoff's approximation is $\mathcal{O}\left(\max\{k^{d+1}, k^d N (\log N)^{d-1}\}\right)$. For $k \ll N$ this represents a significant reduction over the full index set version, where the corresponding figure is $\mathcal{O}\left(\max\{k^{d+1}, k^d N^d\}\right)$ (see Section 2.3). Furthermore, no specific techniques are required: as in the previous setting, we repeatedly solve one-dimensional linear systems involving the matrix $V^{[i]}$.

In Table 3, we demonstrate the improvement offered by this approximation. By means of example, we observe that, when $k = 3$, to obtain an error of less than 10^{-14} requires around $14,000$ terms for the full index set version of Eckhoff approximation, but only around $2,800$ for its hyperbolic cross counterpart.



$k = 2$ $k = 3$ $k = 4$

Figure 4: Log error $\log_{10} \|f - \mathcal{F}_{N,k}[f]\|_\infty$ against number of approximation terms for the full (thin line) and hyperbolic cross (thick line) versions of Eckhoff's method applied to (4.3).

For $d = 3$ the improvement offered is more substantial. In Figure 4 we compare the error of the full and hyperbolic cross versions of Eckhoff's method applied to the function

$$f(x_1, x_2, x_3) = \left(x_1^2 \cos 5x_1 + \tfrac{46}{125} \sin 5 - \tfrac{4}{25} \cos 5\right) (\cosh 2x_2 - \cosh 1 \sinh 1) \\ \times \left(x_3 \sin 2x_3 + \tfrac{1}{2} \cos 2 - \tfrac{1}{4} \sin 2\right). \tag{4.3}$$

For $k = 3$, using roughly $5,000$ terms, the hyperbolic cross version offers an error roughly 10^4 times smaller than the full version. For $k = 4$, the hyperbolic cross approximation obtains an error of 10^{-10} using only $1,500$ terms. The full index set approximation will not reach this value until the number of terms exceeds $6,000$.

Figure 4 also demonstrates the advantage offered by the method developed in this paper, namely the hyperbolic cross version of Eckhoff's method, over the original modified Fourier approximation. For example, to obtain an accuracy of 10^{-10} with $k = 4$ requires roughly $1,500$ terms, whereas to do the same with the modified Fourier approximation $\mathcal{F}_N[f]$ requires in excess of 10^{12} terms.

The analysis of the hyperbolic cross version of Eckhoff's approximation is beyond the scope of this paper. Numerical results indicate that the uniform convergence rate is $\mathcal{O}\left(N^{-2k-1}(\log N)^{d-1}\right)$, but this remains a conjecture. Unfortunately, numerical results also demonstrate that there is no auto-correction phenomenon for this approximation. Away from the boundary, the approximation converges at the same rate as exact polynomial subtraction. In other words, the error is $\mathcal{O}\left(N^{-2k-2}(\log N)^{d-1}\right)$.

Conclusions and future work

The aim of this paper was to examine the convergence acceleration of modified Fourier expansions. To achieve this goal we have generalised Eckhoff's method to the multivariate case, and proved that this approach yields not only faster uniform convergence but also an auto-correction phenomenon inside the domain. We have then demonstrated how a significant reduction in the number of approximation coefficients can be achieved by using a hyperbolic cross index set. The so-called hyperbolic cross version of Eckhoff's method gives accurate approximations comprising a relatively small number of terms. Finally, in the univariate setting, we have established how numerical stability can be increased by using a particular subtraction basis.

There are a number of areas for future investigation. First, as mentioned in the Introduction, Eckhoff's method can be extended to non-Cartesian product domains, provided suitable orthogonal expansions are known. Due to their applications in spectral elements, the equilateral and right isosceles triangles are two important examples that warrant future consideration.

In [1, 2] the author considers the application of modified Fourier series to the spectral approximation of second order boundary value problems. The method possesses a number of advantages, including mild conditioning of the discretization matrix and the availability of an optimal, diagonal preconditioner. However, the convergence rate is only cubic in the truncation parameter. Accelerating convergence is a subject of current investigation, including the incorporation of the methods developed in this paper into such approximations.

Finally, there are several open problems relating to both the theory and implementation of the multivariate form of Eckhoff approximation. First, as mentioned, the analysis of the hyperbolic cross version of Eckhoff's method has not yet been carried out. We intend to address this in a future paper. Second, though we have demonstrated numerically the advantage offered by the subtraction basis (1.7), we are yet to explain this effect theoretically. On a related topic, since more solves of linear systems

involving the (ill-conditioned) matrix $V^{[i]}$ are required in higher dimensions, the method is increasingly susceptible to round-off error. While we have used additional precision in the multivariate numerical examples presented herein as compensation, a complete resolution of this issue is outside the scope of this paper. Future work, most likely along the lines of incorporating (appropriately optimised) least squares procedures, is necessary to address this problem.

Acknowledgements

The original idea for this paper was presented to the author by Euan Spence (University of Bath), to whom he extends his gratitude. He would also like to thank his supervisor Arieh Iserles (DAMTP, University of Cambridge), Alfredo Deaño (DAMTP, University of Cambridge), Daan Huybrechs (Katholieke Universiteit Leuven), David Levin (Tel Aviv University) and Arnak Poghosyan (Yerevan State University). Finally, the author would like to thank the two anonymous referees for their useful comments and suggestions, which greatly improved the paper.

References

[1] B. Adcock. Multivariate modified Fourier series and application to boundary value problems. *Technical report NA2008/08, DAMTP, University of Cambridge*, 2008.

[2] B. Adcock. Univariate modied Fourier methods for second order boundary value problems. *BIT*, 49(2):249–280, 2009.

[3] K. I. Babenko. Approximation of periodic functions of many variables by trigonometric polynomials. *Soviet Math. Dokl.*, 1:513–516, 1960.

[4] A. Barkhudaryan, R. Barkhudaryan, and A. Poghosyan. Asymptotic behavior of Eckhoff's method for Fourier series convergence acceleration. *Anal. Theory Appl.*, 23(3):228–242, 2007.

[5] G. Baszenski, F.-J. Delvos, and M. Tasche. A united approach to accelerating trigonometric expansions. *Comput. Math. Appl.*, 30(3–6):33–49, 1995.

[6] A. Björck and V. Pereyra. Solution of Vandermonde systems of equations. *Math. Comp.*, 24:893–903, 1970.

[7] J. P. Boyd. A comparison of numerical algorithms for Fourier Extension of the first, second, and third kinds. *J. Comput. Phys.*, 178:118–160, 2002.

[8] K. S. Eckhoff. Accurate and efficient reconstruction of discontinuous functions from truncated series expansions. *Math. Comp.*, 61(204):745–763, 1993.

[9] K. S. Eckhoff. Accurate reconstructions of functions of finite regularity from truncated Fourier series expansions. *Math. Comp.*, 64(210):671–690, 1995.

[10] K. S. Eckhoff. On a high order numerical method for functions with singularities. *Math. Comp.*, 67(223):1063–1087, 1998.

[11] W. Gautschi. Norm estimates for inverses of Vandermonde matrices. *Numer. Math.*, 23:337–347, 1975.

[12] D. Gottlieb and C-W. Shu. On the Gibbs' phenomenon and its resolution. *SIAM Rev*, 39(4):644–668, 1997.

[13] D. Gottlieb, C-W. Shu, A. Solomonoff, and H. Vandeven. On the Gibbs phenomenon I: Recovering exponential accuracy from the Fourier partial sum of a nonperiodic analytic function. *J. Comput. Appl. Math.*, 43(1–2):91–98, 1992.

[14] N. J. Higham. *Accuracy and stability of numerical algorithms*. SIAM, 2nd edition, 2002.

[15] D. Huybrechs, A. Iserles, and S. P. Nørsett. From high oscillation to rapid approximation V: The equilateral triangle. *Technical report NA2009/04, DAMTP, University of Cambridge*, 2009.

[16] D. Huybrechs, A. Iserles, and S. P. Nørsett. From high oscillation to rapid approximation IV: Accelerating convergence. *IMA J. Num. Anal. (to appear)*, 2010.

[17] A. Iserles and S. P. Nørsett. From high oscillation to rapid approximation I: Modified Fourier expansions. *IMA J. Num. Anal.*, 28:862–887, 2008.

[18] A. Iserles and S. P. Norsett. From high oscillation to rapid approximation III: Multivariate expansions. *IMA J. Num. Anal. (to appear)*, 2009.

[19] W. B. Jones and G. Hardy. Accelerating convergence of trigonometric approximations. *Math. Comp.*, 2(111):547–560, 1970.

[20] L. V. Kantorovich and V. I. Krylov. *Approximate Methods of Higher Analysis*. Interscience, New York, 3rd edition, 1958.

[21] A. Krylov. On approximate calculations. *Lectures delivered in 1906 (in Russian). St Petersburg*, 1907.

[22] C. Lanczos. *Discourse on Fourier series*. Hafner, New York, 1966.

[23] J. N. Lyness. Computational techniques based on the Lanczos representation. *Math. Comp.*, 28(125):81–123, 1974.

[24] A. Nersessian and A. Poghosyan. Bernoulli method in multidimensional case. *Preprint in ArmNIINTI 09.03.2000 N20 Ar-00 (in Russian)*, 2000.

[25] A. Nersessian and A. Poghosyan. Fast convergence of a polynomial-trigonometric interpolation. *Preprint in ArmNIINTI 07.07.2000 N45 Ar-00 (in Russian)*, 2000.

[26] A. Nersessian and A. Poghosyan. Asymptotic errors of accelerated two-dimensional trigonometric approximations. *Proceedings of the ISAAC fourth Conference on Analysis. Yerevan, Armenia (G. A. Barsegian, H. G. W. Begehr, H. G. Ghazaryan, A. Nersessian eds), Yerevan, pp 70–78*, 2004.

[27] A. Nersessian and A. Poghosyan. The convergence acceleration of two-dimensional Fourier interpolation. *Armenian Journal of Mathematics*, 1:50–63, 2008.

[28] S. Olver. On the convergence rate of a modified Fourier series. *Math. Comp.*, 78:1629–1645, 2009.

[29] A. Poghosyan. On an autocorrection phenomenon of the Krylov–Gottlieb–Eckhoff method. *Submitted*, 2006.

[30] H.-J. Schmeißer and H. Triebel. *Topics in Fourier analysis and function spaces*. Wiley, 1987.

[31] E Tadmor. Filters, mollifiers and the computation of the Gibbs' phenomenon. *Acta Numerica*, 16:305–378, 2007.

[32] V. Temlyakov. *Approximation of Periodic Functions*. Nova Sci., New York, 1993.