Order-preserving derivative approximation with periodic radial basis functions.

Edward Fuselier · Grady B. Wright

Abstract In this exploratory paper we study the convergence rates of an iterated method for approximating derivatives of periodic functions using radial basis function (RBF) interpolation. Given a target function sampled on some node set, an approximation of the mth derivative is obtained by m successive applications of the operator "interpolate, then differentiate" - this process is known in the spline community as *successive splines* or *iterated splines*. For uniformly spaced nodes on the circle, we give a sufficient condition on the RBF kernel to guarantee that, when the error is measured only at the nodes, this iterated method approximates all derivatives with the same rate of convergence. We show that thin-plate spline, power function, and Matérn kernels restricted to the circle all satisfy this condition, and numerical evidence is provided to show that this phenomena occurs for some other popular RBF kernels. Finally, we consider possible extensions to higher-dimensional periodic domains by numerically studying the convergence of an iterated method for approximating the surface Laplace (Laplace-Beltrami) operator using RBF interpolation on the unit sphere and a torus.

1 Introduction

A situation that arises in many applications is that of approximating derivatives of a function given only samples of it at some set of nodes. This problem is, for example, central to collocation-type approaches for numerically solving differential

Edward Fuselier
Department of Mathematics and Computer Science
High Point University
High Point, NC 27262
E-mail: efuselie@highpoint.edu

Grady B. Wright
Department of Mathematics
Boise State University
Boise, ID 83725-1555
E-mail: wright@math.boisestate.edu

equations. In this particular application, the desired derivative is often approximated by first interpolating the function and then evaluating the derivative of this interpolant at the nodes. Error estimates for the resulting approximation are commonly given over a continuum, but in a collocation method the numerical solutions are typically only maintained on a discrete set, namely the data sites. This implies that when studying a derivative approximation scheme, it often makes sense to ask: *what happens at the nodes*?

In this paper we investigate this question for derivative approximations obtained from periodic radial basis function (RBF) interpolation. We consider two techniques for approximating derivatives. The first is the obvious approach of first interpolating the data and then computing the derivative of the interpolant. We refer to this as the *direct* technique. The second technique uses an *iterated* approach, which is a term used by some in the spline community to describe the process of computing the mth derivative by m successive applications of the operator "interpolate, then differentiate." In the case of one dimension and $m = 2$, these two techniques can be described formally as follows: Let X be a finite set of data points, and let I_X be the interpolation operator for some interpolation scheme. Given a target function f, we approximate the first derivative of f with $\mathcal{D}_X f := (I_X f)'$. The direct approach of approximating f'' is to simply differentiate the interpolant again, i.e. $(I_X f)''$. The iterated approach of computing f'' is to interpolate the approximation of first derivative and then differentiate it, i.e. $(I_X(I_X f)')' = \mathcal{D}_X^2 f$. Higher derivatives are obtained by naturally repeating this process. We refer to the operator $\mathcal{D}_X^m$ as an *iterated differential operator* of order m.

Periodic odd-degree spline interpolation on uniform grids yields iterated derivatives that, curiously, enjoy the same rate of convergence for all derivatives of the target function at the interpolation nodes (see [17] and the references therein). Further, these rates are faster than the direct method of taking higher order derivatives of the interpolant directly. For example, periodic cubic splines have the property that both $\mathcal{D}_X f$ and $\mathcal{D}_X^2 f$ approximate f' and f'', respectively, to $\mathcal{O}(N^{-4})$ at the equally spaced nodes X with $\#X = N$, while $f''|_X - (I_X f)''|_X$ remains only $\mathcal{O}(N^{-2})$. For lack of a better term, we will refer to these faster than expected convergence rates as *superconvergence*. We aim to show that iterated derivatives coming from some periodic RBF interpolation methods also experience superconvergence on uniformly spaced nodes.

It is well-known that many spline interpolants can be built from shifts of a single reproducing kernel (with possibly an added element coming from some low-dimensional space of functions), that is

$$I_X f(x) = \sum_{x_j \in X} c_j \phi(x, x_j) + p_f(x), \tag{1}$$

where ϕ is the kernel, and the c_j's and function p_f are chosen by enforcing interpolation on X and possibly some extra side conditions. Golomb, for example, showed that for equally spaced nodes on $[0, 1]$, periodic spline interpolants of degree $2r - 1$, $r \in \mathbb{N}$, can be written as [10, Equation (2.10)].

$$I_X f(x) = \sum_{x_j \in X} c_j b_{2r}(x - x_j) + d,$$

where d is a constant and the kernel b_{2r} is the periodic extension of the $2r^{\text{th}}$ Bernoulli polynomial [1, Ch. 23]. Considering interpolants of the form (1), we will

determine sufficient conditions on the kernel's Fourier coefficients to guarantee the preservation of order in iterated derivatives at the nodes. These results hold for the kernels b_{2r} in particular, and thus generalize the asymptotic preservation of order in iterated derivatives in the spline case to other kernel interpolation methods. Additionally, they apply to the restriction of some popular radial basis functions defined on $\mathbb{R}^d$ to the circle— namely the power functions, thin plate splines, and Matérn kernels.

Iterated derivatives may also be advantageous beyond the univariate setting. In [8], an iterated approach was presented to approximate the surface Laplacian (Laplace-Beltrami operator) $\Delta_\mathbb{M}$ using RBFs, and this approximation was then used to solve nonlinear reaction-diffusion equations on various periodic surfaces (that is, surfaces with no boundary). In numerical examples from [8], faster than expected convergence rates at the nodes were observed. Here we consider a more extensive set of examples to better understand the phenomena of accelerated convergence at the data sites. We also compare the iterated approach to the direct approach of analytically applying $\Delta_\mathbb{M}$ to an RBF interpolant, which has been applied in the past to problems on the sphere [9, 22]. Various node families are used in these experiments, and we will see that the iterated approach yields faster convergence rates in several cases. On the sphere, faster convergence is not observed at all the data sites considered, but accelerated convergence is observed for some node families at sites away from regions where the nodes are less uniform.

The paper is organized as follows. In section 2, we establish the necessary notation and definitions for kernel interpolants on the circle, associated function spaces, and other theoretical machinery. We discuss the approximation of derivatives at the nodes in section 3 - the main result on iterated derivatives is given in Theorem 3, where we establish that iterated derivatives of any order based on certain kernel interpolants approximate the derivatives of smooth periodic functions to the same order at the data sites. Examples of kernels for which this result holds are given in section 4. Section 5 contains numerical experiments on the circle, which serve to verify our results and to study some kernels not treated in the theory. In section 6, we discuss an iterated approach to approximating the Laplacian on two-dimensional surfaces, and numerically compare the convergence of the iterated method with a direct approach on the unit sphere and a torus.

2 Preliminaries

We denote the circle by $\mathbb{T}$. Here a point on the circle is associated with its angle, i.e. $\mathbb{T} := \mathbb{R} \mod 2\pi$. The distance between two points $x, y \in \mathbb{T}$ is $d(x, y) = \arccos(\cos(x - y))$. Given a function g and a finite subset Y of its domain, we will use $\|g\|_{\ell_\infty(Y)}$ to denote $\max_{y \in Y} |g(y)|$.

For the harmonic analysis required we will follow the notation in [12]. Given a vector $y = (y_0, y_1, \ldots, y_{N-1})$, its discrete Fourier transform (DFT) $\widetilde{y}$ and its inverse are related by the equations

$$\widetilde{y}_k = \sum_j y_j e^{\frac{-2\pi i j k}{N}}, \qquad y_j = \frac{1}{N} \sum_k \widetilde{y}_k e^{\frac{2\pi i k j}{N}}.$$

Note that these formulas are indexed by N-periodic integer sequences on the interval $[0, N-1]$, but any interval of length $N-1$ can be used. We will work with

the (nearly) symmetric interval centered around 0:

$$\mathcal{J}_N := \left\{ k \in \mathbb{Z} \ : \ -\left\lfloor \frac{N-1}{2} \right\rfloor \le k \le \left\lceil \frac{N-1}{2} \right\rceil \right\}.$$

Note that $|k| \le N/2$ for any index $k \in \mathcal{J}_N$. Also, for $n, k \in \mathbb{Z}$, we will use the shorthand $n \equiv k$ for $n \equiv k \mod N$.

Throughout most of the paper X will be a set of N equally spaced nodes (angles) on $\mathbb{T}$, specifically $X = \{2\pi j/N\}_{j=0}^{N-1} = \{2\pi j/N\}_{j \in \mathcal{J}_N}$. We denote a function $f : \mathbb{T} \to \mathbb{R}$ sampled on X by f_X, and we define its corresponding DFT as $\widetilde{f} := \widetilde{f_X}$. We will use the following convention for the Fourier expansion of $f \in L_2(\mathbb{T})$:

$$f(x) = \sum_{n \in \mathbb{Z}} \widehat{f}(n) e^{inx}, \quad \text{where} \quad \widehat{f}(n) = \begin{cases} \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) e^{inx} \, dx & n \neq 0 \\ \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \, dx & n = 0. \end{cases}$$

Lastly, the relationship between the $\widehat{f}$ and $\widetilde{f}$ can be seen by decomposing f into frequency bands, given by $\sum_{n \equiv k} \widehat{f}(n)$ for $k \in \mathcal{J}_N$. It can be shown that

$$\widetilde{f}_k = N \sum_{n \equiv k} \widehat{f}(n). \tag{2}$$

2.1 Interpolation with Periodic Basis Functions

We consider interpolation by shifts of a single kernel $\phi : \mathbb{T} \times \mathbb{T} \to \mathbb{R}$, with possibly an added requirement that the method reproduce trigonometric polynomials up to a certain degree, say $q - 1$. More specifically, given a finite set of distinct points $Y = \{y_j\} \subset \mathbb{T}$ and continuous function $f : \mathbb{T} \to \mathbb{R}$, we seek an interpolant of the form

$$I_Y f(x) = \sum_{y_k \in Y} c_k \phi(x, y_k) + \sum_{|k| < q} \alpha_k e^{ikx}, \tag{3}$$

where the coefficients $c_k \in \mathbb{R}$, $\alpha_k \in \mathbb{C}$ are chosen by enforcing interpolation on Y, $I_Y f|_Y = f|_Y$, with the side conditions:

$$\sum_{y_j \in Y} c_j e^{iky_j} = 0 \quad \text{for all } k \in \mathbb{Z} \text{ with } |k| < q. \tag{4}$$

To guarantee the existence and uniqueness of such an interpolant, we will use conditionally positive definite kernels (defined below). However, since polynomial reproduction is being included we need conditions on the interpolation points to guarantee uniqueness. We say that Y is *q-unisolvent* if the zero polynomial is the only trigonometric polynomial of degree at most $q - 1$ that vanishes on Y. With that, we define the following:

Definition 1 Given a nonnegative integer q, we say that $\phi : \mathbb{T} \times \mathbb{T} \to \mathbb{R}$ is *conditionally positive definite* (CPD) of order q if given any finite set of distinct q-unisolvent points $Y = \{y_1, \ldots, y_N\} \subset \mathbb{T}$ and any nonzero $\mathbf{c} \in \mathbb{R}^N$ satisfying (4) we have

$$\sum_{j,k=1}^{N} c_j c_k \phi(y_j, y_k) > 0.$$

If ϕ is CPD of order 0 (that is, there are no side conditions (4)), then we say that ϕ is *positive definite*.

Remark 1 Given a CPD ϕ, one can show that for a real target the coefficients α_k in (3) satisfy $\alpha_k = \overline{\alpha_{-k}}$, which implies a real-valued interpolant.

The kernel ϕ is *zonal* if it only depends on the distance between its arguments, i.e. $\phi(x, y) = \psi(d(x, y))$ for some univariate function ψ. Since in this case $\phi(x, y) = \phi(x - y, 0)$, we will use the shorthand $\phi(x - y) := \phi(x, y)$. For a zonal real-valued kernel ϕ, $\widehat{\phi}(n)$ is real and $\widehat{\phi}(n) = \widehat{\phi}(-n)$ for all $n \in \mathbb{Z}$.

A sufficient (but not necessary) condition on a zonal kernel ϕ to guarantee positive definiteness is that ϕ can be written as

$$\phi(x) = \sum_{n=0}^{\infty} a_n \cos(nx), \qquad \sum_{n=0}^{\infty} a_n < \infty,$$

where $a_n > 0$ for all n [23]. This is equivalent to $\widehat{\phi}(n) > 0$ for all $n \in \mathbb{Z}$ with $\sum_{n \in \mathbb{Z}} \widehat{\phi}(n) < \infty$. It is also true that if a zonal kernel satisfies $\widehat{\phi}(n) > 0$ for all $|n| \geq q$ and $\sum_{|n| \geq q} \widehat{\phi}(n) < \infty$, then ϕ is CPD of order q.

Definition 2 We will say that a zonal kernel ϕ is a *periodic basis function* (PBF) *of order q* if the Fourier coefficients $\{\widehat{\phi}(n) : |n| \geq q\}$ are strictly positive and summable.

When dealing with equally spaced points $X \subset \mathbb{T}$, the interpolation coefficients take a convenient form. Starting with the interpolation conditions, expanding the kernel in a Fourier series and exchanging sums gives

$$f(x_j) = \frac{1}{N} \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \widetilde{c}_k \widetilde{\phi}_k e^{ikx_j} + \frac{1}{N} \sum_{|k| < q} N \alpha_k e^{ikx_j}$$

Also, note that in this case the side condition (4) is equivalent to $\widetilde{c}_k = 0$ for all $|k| < q$. Thus after comparing to the inverse DFT formula, we see that the coefficients satisfy the following:

$$\widetilde{c}_k = \begin{cases} 0 & |k| < q \\ \widetilde{f}_k / \widetilde{\phi}_k & |k| \geq q \end{cases}, \quad \text{and} \quad \alpha_k = N^{-1} \widetilde{f}_k, \quad |k| < q.$$

After expanding $\widetilde{f}$ in terms of frequency bands, the interpolant takes the following form, which we will make use of later:

$$I_X f(x) = \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{N}{\widetilde{\phi}_k} \sum_{l \equiv k} \widehat{f}(l) \sum_{n \equiv k} \widehat{\phi}(n) e^{inx} + \sum_{|k| < q} \sum_{l \equiv k} \widehat{f}(l) e^{ikx}. \qquad (5)$$

2.2 Wiener Spaces

The set of functions on the circle whose associated Fourier sequences are absolutely summable form the normed linear space of Wiener functions, which we denote by $\mathcal{A}$. The norm for $f \in \mathcal{A}$ is given by

$$\|f\|_{\mathcal{A}} := \sum_{n \in \mathbb{Z}} |\widehat{f}(n)|.$$

It is well known that $\mathcal{A} \subset C(\mathbb{T})$, the set of continuous functions on the circle. We will be interested in smoother spaces, linked to the kernel ϕ, based on this idea. Given a PBF ϕ of order q, we define the semi-norm

$$|f|_{\mathcal{A}_\phi} := \sum_{|n| \geq q} \frac{\left|\widehat{f}(n)\right|}{\widehat{\phi}(n)},$$

and denote the collection of all continuous functions for which this semi-norm is finite by $\mathcal{A}_\phi$.

Given a function $f \in \mathcal{A}_\phi$ and uniform nodes X, there is a nice connection between the DFT $\widetilde{f}$ and the Wiener semi-norms of f, namely that

$$\sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \left|\frac{\widetilde{f}_k}{\widetilde{\phi}_k}\right| \leq |f|_{\mathcal{A}_\phi}. \tag{6}$$

Indeed, expanding the DFT's into frequency bands gives

$$\sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \left|\frac{\widetilde{f}_k}{\widetilde{\phi}_k}\right| = \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \sum_{n \equiv k} \frac{\left|\widehat{f}(n)\right|}{\sum_{l \equiv k} \widehat{\phi}(l)} \leq \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \sum_{n \equiv k} \frac{\left|\widehat{f}(n)\right|}{\widehat{\phi}(n)} = \sum_{|n| \geq q} \frac{\left|\widehat{f}(n)\right|}{\widehat{\phi}(n)} = |f|_{\mathcal{A}_\phi}.$$

Also, weighting the discrete Fourier coefficients of a function by powers of their associated frequencies has the effect of measuring derivatives of that function. For $m \in \mathbb{N}$, note that $\widehat{f^{(m)}}(n) = (in)^m \widehat{f}(n)$, where $f^{(m)}$ is the mth derivative of f, and for $k \in \mathcal{J}_N$, $|k| \leq |n|$ for all $n \equiv k$. Using this, one can expand into frequency bands as above to get

$$\sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \left|\frac{\widetilde{f}_k k^m}{\widetilde{\phi}_k}\right| \leq |f^{(m)}|_{\mathcal{A}_\phi}. \tag{7}$$

2.3 Periodic Basis Function Approximation

The interpolation error for PBFs (and many other kernel-based approximation schemes) is governed by the rate at which the kernel's Fourier coefficients decay. It has been shown in [3, Theorem 5], for example, that for positive definite ϕ and equally spaced $X \subset \mathbb{T}$, if $f \in \mathcal{A}_\phi$ one has

$$\|I_X f - f\|_{L_\infty(\mathbb{T})} = \mathcal{O}\left(\widehat{\phi}(N/2)\right). \tag{8}$$

A similar result also holds when ϕ is CPD. We will concern ourselves with kernels providing algebraic convergence rates, thus we assume that for all nonzero n with $|n| \geq q$ (and for all n if $q = 0$) that

$$c\,|n|^{-s} \leq \widehat{\phi}(n) \leq C\,|n|^{-s}, \qquad \text{for some } s > 2, \tag{9}$$

where c, C are positive constants.[1]

---

[1] The condition $s > 2$ guarantees that the Fourier coefficients of ϕ' are absolutely summable, thus allowing one to differentiate the Fourier series of any interpolant term-by-term.

For PBFs satisfying (9), interpolation on uniform nodes gives $\mathcal{O}(N^{-s})$ approximation. In general, each derivative taken of the interpolation error multiplies the error bound by a factor of N. However, this rule of thumb applies when the error is measured over the whole domain.

3 Differential Approximation at the Nodes

In this section we explore error bounds for derivative approximation using kernel interpolants on equally spaced nodes. Section 3.1 focuses on the first and second derivatives at the nodes, and in section 3.2 we establish that for kernels satisfying a mild assumption (see Assumption 1), iterated derivatives of all orders converge at the same rate.

3.1 Derivatives of Interpolants

As (8) suggests, the decay of the kernel's Fourier coefficients paints a complete picture of the interpolant's global approximation power, so one might hope that (9) combined with the inherent symmetries of the problem (namely uniformly spaced nodes and a zonal kernel) would be enough to achieve higher approximation rates in derivatives at the nodes. Unfortunately, this is not the case - a kernel satisfying (9) that loses an order of N in the first derivative at the nodes is presented in the following example.

Example 1 Consider two PBFs, ψ_1 and ψ_2, each of whose Fourier coefficients are either of the form $1/n^4$ or $2/n^4$. The kernel ψ_1 is constructed by choosing successive pairs of its Fourier coefficients from the curves $1/n^4$ and $2/n^4$ in the following way (see Figure 1(a)):

$$\psi_1(x) = 1 + \sum_{n=1}^{\infty} a_n \cos(nx), \quad a_n = \begin{cases} 1/n^4 \ n = 1, 2, 5, 6, 9, 10, \ldots \\ 2/n^4 \ n = 3, 4, 7, 8, 11, 12, \ldots. \end{cases}$$

This kernel certainly satisfies (9) with $s = 4$, so the global interpolation error saturates at $\mathcal{O}(N^{-4})$. The Fourier coefficients of ψ_2 are determined by

$$\psi_2(x) = 1 + \sum_{n=1}^{\infty} b_n \cos(nx), \quad b_n = \begin{cases} 1/n^4 \ n \text{ odd}, \\ 2/n^4 \ n \text{ even}. \end{cases}$$

The global interpolation error for ψ_2 also saturates at $\mathcal{O}(N^{-4})$. To test approximation in the first derivative at the nodes, both of these kernels were used to interpolate the smooth periodic target function

$$f(x) = e^{-4\cos(x)} \sin(4x) \tag{10}$$

at $N = 200, 225, \ldots, 500$ uniform nodes on the circle. The relative infinity error was approximated on a fine mesh, and the relative infinity error in the first derivative at the nodes was also measured. Figure 1 shows that while both kernels give 4th order global convergence (as expected), ψ_2 also yields fourth order convergence in the first derivative at the nodes, whereas ψ_1 does not.
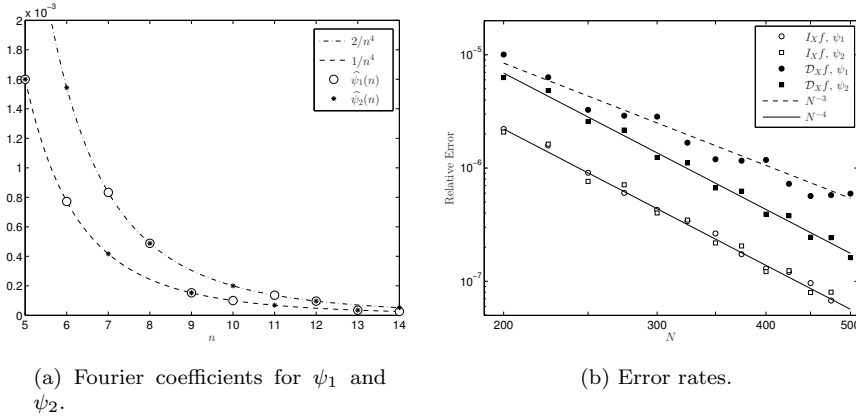
(a) Fourier coefficients for ψ_1 and ψ_2.

(b) Error rates.

Fig. 1 Relative ℓ_∞ interpolation error and relative error in the first derivative of the target function (10) using kernels ψ_1 and ψ_2. The interpolation error was computed globally on a fine mesh, while the error in the first derivative was measured only at the nodes.

Thus (9) alone will not predict higher convergence in iterated derivatives at the data sites. However, kernels satisfying the following extra condition will.

Assumption 1 *Let ϕ be an order q PBF satisfying* (9). *Suppose that there is a constant C, depending only on ϕ, such that for all $n, k \in \mathbb{Z}$ with $|k| \le |n|/2$ we have*

$$\left| \widehat{\phi}(n-k) - \widehat{\phi}(n+k) \right| \le C|k||n|^{-s-1}.$$

Note that we are assuming that certain symmetric differences of the Fourier coefficients decay slightly faster than the Fourier coefficients themselves. Everything we need proceeds from this fact, and some well-known kernels that posses this property are discussed in section 4.

Now we focus on bounding the error in the first derivative of the interpolant at the nodes. We will frequently use the following lemma.

Lemma 1 *Let ϕ satisfy Assumption 1. Then there is a constant C such that for all $N \in \mathbb{N}$ and all $k \in \mathcal{P}_N$ we have*

$$\left| \sum_{n \equiv k} (n-k)\widehat{\phi}(n) \right| \le C|k|N^{-s}.$$

Proof Since $k \in \mathcal{P}_N$, we have that $|k| \le N/2$, so the only nonzero terms in the sum $\sum_{n \equiv k}(n-k)\widehat{\phi}(n)$ occur when $|n| \ge \lceil N/2 \rceil \ge N/2$. Making the substitution $n = Nj + k$ or $n = -Nj + k$ for $n \equiv k$ gives us

$$\sum_{n \equiv k}(n-k)\widehat{\phi}(n) = \sum_{\substack{n \equiv k \\ n \ge N/2}}(n-k)\widehat{\phi}(n) + \sum_{\substack{n \equiv k \\ n \le -N/2}}(n-k)\widehat{\phi}(n)$$

$$= \sum_{j=1}^{\infty} Nj\left(\widehat{\phi}(Nj+k) - \widehat{\phi}(Nj-k) \right),$$

where in the last equality we have used the identity $\widehat{\phi}(-n) = \widehat{\phi}(n)$. Now we use Assumption 1 to get

$$
\left| \sum_{n \equiv k} (n-k)\widehat{\phi}(n) \right| \leq \sum_{j=1}^{\infty} Nj \left| \widehat{\phi}(Nj+k) - \widehat{\phi}(Nj-k) \right|
$$

$$
\leq \sum_{j=1}^{\infty} Nj \left(C|k|(Nj)^{-s-1} \right) = C|k|N^{-s} \sum_{j=1}^{\infty} j^{-s}.
$$

Since $s > 2 > 1$, the series above converges and the result follows.

Before moving on, we will need a few formulas for the interpolant and its derivative at the nodes. Let $x_j \in X$. Using (5) and the fact that $e^{inx_j} = e^{ikx_j}$ for $n \equiv k$, we get

$$
I_X f|_{x_j} = \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{Ne^{ikx_j}}{\widetilde{\phi}_k} \sum_{l \equiv k} \widehat{f}(l) \sum_{n \equiv k} \widehat{\phi}(n) + \sum_{|k| < q} \sum_{l \equiv k} \widehat{f}(l)e^{ikx_j}. \tag{11}
$$

Also, we may differentiate the Fourier expansion (5) term by term to get

$$
(I_X f(x))' = \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{N}{\widetilde{\phi}_k} \sum_{l \equiv k} \widehat{f}(l) \sum_{n \equiv k} in\,\widehat{\phi}(n)e^{inx} + \sum_{|k| < q} \sum_{l \equiv k} \widehat{f}(l)ik\,e^{ikx},
$$

which for $x_j \in X$ implies

$$
\mathcal{D}_X f|_{x_j} = \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{iNe^{ikx_j}}{\widetilde{\phi}_k} \sum_{l \equiv k} \widehat{f}(l) \sum_{n \equiv k} n\,\widehat{\phi}(n) + \sum_{|k| < q} \sum_{l \equiv k} \widehat{f}(l)ik\,e^{ikx_j}. \tag{12}
$$

Recall that if the global pointwise interpolation error is $\mathcal{O}(N^{-s})$, one should expect the error in the first derivative to decay like $\mathcal{O}(N^{-s+1})$. Now we are ready to prove that the approximation rate in the first derivative is actually $\mathcal{O}(N^{-s})$ at the nodes.

Theorem 2 *Let ϕ be an order q PBF satisfying Assumption 1. Then for all f such that $f' \in \mathcal{A}_\phi$, we have*

$$
\|(I_X f)' - f'\|_{\ell_\infty(X)} = \|\mathcal{D}_X f - f'\|_{\ell_\infty(X)} \leq CN^{-s}|f'|_{\mathcal{A}_\phi},
$$

where C depends only on ϕ.

Proof Let $x_j \in X$, and note that $f'|_{x_j} = I_X f'|_{x_j}$. Thus the result will follow once we bound $(\mathcal{D}_X f - I_X f')|_{x_j}$. Using (11) and the fact that $\widehat{f'}(l) = il\,\widehat{f}(l)$, we have

$$
I_X f'|_{x_j} = \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{Nie^{ikx_j}}{\widetilde{\phi}_k} \sum_{l \equiv k} \widehat{f}(l) \sum_{n \equiv k} l\,\widehat{\phi}(n) + \sum_{|k| < q} \sum_{l \equiv k} \widehat{f}(l)il\,e^{ikx_j}.
$$

Subtracting this from (12) gives us

$$(\mathcal{D}_X f - I_X f')|_{x_j} = \underbrace{\sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{N i e^{ikx_j}}{\widetilde{\phi}_k} \sum_{l \equiv k} \widehat{f}(l) \sum_{n \equiv k} (n - l) \, \widehat{\phi}(n)}_{I} + \sum_{|k| < q} \sum_{l \equiv k} \widehat{f}(l) i (k - l) \, e^{ikx_j}.$$

(13)

Now we add and subtract $k \in \mathcal{J}_N$ from the inner-most sum of I to write the above expression as $A + B$, where

$$A = \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{i N e^{ikx_j}}{\widetilde{\phi}_k} \sum_{l \equiv k} \widehat{f}(l) \sum_{n \equiv k} (n - k) \widehat{\phi}(n),$$

$$B = \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{i N e^{ikx_j}}{\widetilde{\phi}_k} \sum_{l \equiv k} \widehat{f}(l) \sum_{n \equiv k} (k - l) \widehat{\phi}(n) + \sum_{|k| < q} \sum_{l \equiv k} \widehat{f}(l) i (k - l) \, e^{ikx_j}.$$

First we bound $|A|$. Consolidating the sum over l into $\widetilde{f}_k/N$ and applying Lemma 1, we get

$$|A| \leq \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \left| \frac{\widetilde{f}_k}{\widetilde{\phi}_k} \right| \left| \sum_{n \equiv k} (n - k) \widehat{\phi}(n) \right| \leq C N^{-s} \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \left| \frac{\widetilde{f}_k \, k}{\widetilde{\phi}_k} \right| \leq C N^{-s} |f'|_{\mathcal{A}_\phi},$$

where in the last step we used (7).

To bound $|B|$, we collapse the sum over n into $\widetilde{\phi}_k/N$, cancel, and combine the two sums to get

$$B = \sum_{k \in \mathcal{J}_N} i e^{ikx_j} \sum_{l \equiv k} \widehat{f}(l)(k - l).$$

Since $k \in \mathcal{J}_N$, the inner sum is only nonzero when $|l| \geq N/2$, giving us

$$|B| \leq \sum_{k \in \mathcal{J}_N} \sum_{\substack{l \equiv k \\ |l| \geq N/2}} |\widehat{f}(l)(k - l)| \leq 2 \sum_{k \in \mathcal{J}_N} \sum_{\substack{l \equiv k \\ |l| \geq N/2}} |\widehat{f}(l) \, l| = 2 \sum_{|n| \geq N/2} |\widehat{f}(n) \, n|,$$

where we have used the fact that $|k| \leq N/2 \leq |l|$ for all indices l in the inner sum. Now, since $\widehat{\phi}$ decays as in (9), we get that $1 \leq C N^{-s}/\widehat{\phi}(n)$ for all $|n| \geq N/2$, which gives us

$$|B| \leq 2 \sum_{|n| \geq N/2} |\widehat{f}(n) \, n| \leq C N^{-s} \sum_{|n| \geq N/2} \frac{|\widehat{f}(n) \, n|}{\widehat{\phi}(n)} \leq C N^{-s} |f'|_{\mathcal{A}_\phi}.$$

This finishes the proof.

Given that the first derivative is approximated better on the nodes than over all of $\mathbb{T}$, it is reasonable then to expect improvements at the nodes for second order derivatives too. However, this is not the case: the second derivative converges in general like $\mathcal{O}(N^{-s+2})$ globally, and below we show that there is no improvement in this rate at the nodes.

Proposition 1 *Let ϕ be an order q PBF satisfying (9) with $s > 3$. Consider the target function $f = e^{imx}$, where $m \in \mathbb{Z}$ satisfies $q < |m| \leq N/4$ and $N \in \mathbb{N}$ is large enough so that such an m exists.[2] Then we have*

$$\|(I_X f)'' - f''\|_{\ell_\infty(X)} \geq C N^{-s+2},$$

where C depends only on ϕ.

Proof Let $x_j \in X$ and note that $f''(x_j) = I_X f''(x_j)$. Also, note that since $f = e^{imx}$ with $m \in \mathcal{J}_N$, we have $\widehat{f''}_k = -k^2 \widetilde{f}_k$ and $\widetilde{f}_k = N\delta_{k,m}$ (this implies in particular that the polynomial terms in the interpolants of f and f'' are zero). This gives us

$$(I_X f)''|_{x_j} = -\sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{\widetilde{f}_k}{\widetilde{\phi}_k} e^{ikx_j} \sum_{n \equiv k} n^2 \widehat{\phi}(n) = -\frac{N e^{imx_j}}{\widetilde{\phi}_m} \sum_{n \equiv m} n^2 \widehat{\phi}(n)$$

and

$$(I_X f'')|_{x_j} = \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{\widetilde{f''}_k}{\widetilde{\phi}_k} e^{ikx_j} \sum_{n \equiv k} \widehat{\phi}(n) = -\frac{N e^{imx_j}}{\widetilde{\phi}_m} \sum_{n \equiv m} m^2 \widehat{\phi}(n).$$

Thus the error at a node x_j is bounded from below by

$$\left| ((I_X f)'' - I_X f'')|_{x_j} \right| = \left| \frac{N}{\widetilde{\phi}_m} \sum_{n \equiv m} (n^2 - m^2)\widehat{\phi}(n) \right| = \frac{N}{\widetilde{\phi}_m} \sum_{\substack{n \equiv m \\ |n| \geq N/2}} (n^2 - m^2)\widehat{\phi}(n)$$

$$\geq \frac{1}{\|\phi\|_{\mathcal{A}}} \sum_{\substack{n \equiv m \\ |n| \geq N/2}} (n^2 - m^2)\widehat{\phi}(n),$$

where we have used the fact that $\widetilde{\phi}_m = N \sum_{n \equiv m} \widehat{\phi}(n) \leq N\|\phi\|_{\mathcal{A}}$. Since $|m| \leq N/4$ and ϕ satisfies (9), we get

$$\left| ((I_X f)'' - f'')|_{x_j} \right| \geq \frac{1}{\|\phi\|_{\mathcal{A}}} ((N+m)^2 - m^2)\widehat{\phi}(N+m) = \frac{1}{\|\phi\|_{\mathcal{A}}} (N^2 + 2Nm)\widehat{\phi}(N+m)$$

$$\geq \frac{N^2}{2\|\phi\|_{\mathcal{A}}} \widehat{\phi}(N+m) \geq \frac{cN^2}{2\|\phi\|_{\mathcal{A}}} |N+m|^{-s} \geq \frac{c(4/5)^s}{2\|\phi\|_{\mathcal{A}}} N^{-s+2},$$

where c is the constant from (9). This finishes the proof.

3.2 Iterated Differential Approximation

In this section we investigate how well the iterated differential operator $\mathcal{D}_X^m f$ approximates $f^{(m)}$ at the nodes, where $m \geq 2$. If ϕ satisfies Assumption 1, we will show that one gets the same rate of convergence in every iterated derivative (even if $\phi^{(m)}$ does not exist!). First we derive a Fourier expansion of $\mathcal{D}_X^m f$.

---

[2] This paper focuses on real interpolation, but PBFs can certainly handle complex-valued targets. Here the target being a complex exponential greatly simplifies the calculations in the proof.

Lemma 2 *Let ϕ be an order q PBF and let $m \in \mathbb{N}$. Then we have*

$$\mathcal{D}_X^m f(x) = \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{\widetilde{f}_k}{\widetilde{\phi}_k} (iC_k)^{m-1} \sum_{n \equiv k} in \, \widehat{\phi}(n) e^{inx} + \sum_{|k| < q} \sum_{l \equiv k} \widehat{f}(l)(ik)^m \, e^{ikx},$$

where the constants C_k are given by $C_k := \left(\sum_{n \equiv k} n \widehat{\phi}(n) \right) \left(\sum_{n \equiv k} \widehat{\phi}(n) \right)^{-1}.$

Proof Starting with (12), collect terms into $\widetilde{f}_k$ and expand $\widetilde{\phi}$ in terms of $\widehat{\phi}$ to get

$$\mathcal{D}_X f|_{x_j} = N^{-1} \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \left(\widetilde{f}_k i \frac{\sum_{n \equiv k} n \, \widehat{\phi}(n)}{\sum_{n \equiv k} \widehat{\phi}(n)} \right) e^{ikx_j} + N^{-1} \sum_{|k| < q} \widetilde{f}_k (ik) \, e^{ikx_j}$$

$$= N^{-1} \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \widetilde{f}_k (iC_k) e^{ikx_j} + N^{-1} \sum_{|k| < q} \widetilde{f}_k (ik) \, e^{ikx_j}.$$

This shows that for $k \in \mathcal{J}_N$ one has

$$\widetilde{\mathcal{D}_X f}_k = \begin{cases} \widetilde{f}_k iC_k & |k| \geq q, \\ \widetilde{f}_k ik & |k| < q. \end{cases}$$

Thus for $\mathcal{D}_X^2 f = \mathcal{D}_X(\mathcal{D}_X f)$ we get

$$\mathcal{D}_X^2 f(x) = \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{\widetilde{\mathcal{D}_X f}_k}{\widetilde{\phi}_k} \sum_{n \equiv k} in \, \widehat{\phi}(n) e^{inx} + N^{-1} \sum_{|k| < q} \widetilde{\mathcal{D}_X f}_k (ik) \, e^{ikx}$$

$$= \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{\widetilde{f}_k}{\widetilde{\phi}_k} (iC_k) \sum_{n \equiv k} in \, \widehat{\phi}(n) e^{inx} + N^{-1} \sum_{|k| < q} \widetilde{f}_k (ik)^2 \, e^{ikx}$$

This process can be repeated and the result follows by a simple induction argument.

In the analysis that follows we will need to bound the constants C_k.

Lemma 3 *Let ϕ satisfy Assumption 1 and let $k \in \mathcal{J}_N$ with $|k| \geq q$. Then $|C_k| \leq C|k|$, where C depends only on ϕ.*

Proof We can express C_k as

$$C_k = k + \frac{1}{\sum_{n \equiv k} \widehat{\phi}(n)} \sum_{n \equiv k} (n - k) \widehat{\phi}(n).$$

Lemma 1 and (9) gives

$$|C_k| \leq |k| + \frac{C|k|N^{-s}}{\sum_{n \equiv k} \widehat{\phi}(n)} < |k| + \frac{C|k|N^{-s}}{\widehat{\phi}(k)} \leq |k| + C|k|N^{-s}k^s.$$

Since $k \in \mathcal{J}_N$, we have $|k| \leq N/2$, so $N^{-s}k^s$ is bounded. The result follows.

We also require a generalization of Lemma 1.

Lemma 4 *Let ϕ satisfy Assumption 1. Let $k \in \mathcal{J}_N$ with $|k| \geq q$, and let $m \in \mathbb{N}$. Then we have*

$$\sum_{n \equiv k} (nC_k^{m-1} - k^m)\widehat{\phi}(n) \leq C|k|^m N^{-s},$$

where C is independent of k and N.

Proof We will proceed by induction on m. The base case holds because of Lemma 1. Now assume that the result holds for the mth case and consider

$$\sum_{n \equiv k} (nC_k^m - k^{m+1})\widehat{\phi}(n) = \sum_{n \equiv k} (nC_k^m - nkC_k^{m-1})\widehat{\phi}(n) + \sum_{n \equiv k} (nkC_k^{m-1} - k^{m+1})\widehat{\phi}(n).$$

The desired bound of the right-most term follows by induction. For the left, we use the definition of C_k and apply Lemma 1 to get

$$\left| \sum_{n \equiv k} (nC_k^m - nkC_k^{m-1})\widehat{\phi}(n) \right| = |C_k^{m-1}| \, |C_k - k| \left| \sum_{n \equiv k} n\,\widehat{\phi}(n) \right|$$

$$= \frac{|C_k^{m-1}|}{\sum_{l \equiv k} \widehat{\phi}(l)} \left| \sum_{l \equiv k} l\widehat{\phi}(l) - k\sum_{l \equiv k} \widehat{\phi}(l) \right| \left| \sum_{n \equiv k} n\widehat{\phi}(n) \right|$$

$$= |C_k|^m \left| \sum_{l \equiv k} (l - k)\widehat{\phi}(l) \right| \leq C|k|N^{-s}|C_k|^m.$$

Now apply Lemma 3 to finish the proof.

Now we are ready to prove the following.

Theorem 3 *Let ϕ be a PBF satisfying Assumption 1 and let $m \in \mathbb{N}$. Then for all f such that $f^{(m)} \in \mathcal{A}_\phi$ we have*

$$\|\mathcal{D}_X^m f - f^{(m)}\|_{\ell_\infty(X)} \leq CN^{-s}|f^{(m)}|_{\mathcal{A}_\phi},$$

where C depends only on ϕ and m.

Proof We proceed as in the proof of Theorem 2. Let $x_j \in X$, and note that the result will follow by bounding $(\mathcal{D}_X^m f - I_X f^{(m)})|_{x_j}$. Using (11) and $\widehat{f^{(m)}}(l) = (il)^m \widehat{f}(l)$, we have

$$I_X f^{(m)}|_{x_j} = \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{Ni^m e^{ikx_j}}{\widetilde{\phi}_k} \sum_{l \equiv k} \widehat{f}(l) \sum_{n \equiv k} l^m \widehat{\phi}(n) + \sum_{|k| < q} \sum_{l \equiv k} \widehat{f}(l)(il)^m e^{ikx_j}.$$

Also, Lemma 2 gives

$$\mathcal{D}_X^m f|_{x_j} = \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{\widetilde{f}_k}{\widetilde{\phi}_k} (iC_k)^{m-1} \sum_{n \equiv k} in\,\widehat{\phi}(n)e^{inx_j} + \sum_{|k| < q} \sum_{l \equiv k} \widehat{f}(l)(ik)^m e^{ikx_j}$$

$$= \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{Ni^m e^{ikx_j}}{\widetilde{\phi}_k} \sum_{l \equiv k} \widehat{f}(l) \sum_{n \equiv k} nC_k^{m-1}\,\widehat{\phi}(n) + \sum_{|k| < q} \sum_{l \equiv k} \widehat{f}(l)(ik)^m e^{ikx_j}.$$

Thus we can write the difference $(D_\phi^m f - I_X f^{(m)})|_{x_j} = A + B$, with

$$A = \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{N i^m e^{ikx_j}}{\widetilde{\phi}_k} \sum_{l \equiv k} \widehat{f}(l) \sum_{n \equiv k} \left(n C_k^{m-1} - k^m \right) \widehat{\phi}(n),$$

$$B = \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \frac{N i^m e^{ikx_j}}{\widetilde{\phi}_k} \sum_{l \equiv k} \widehat{f}(l) \sum_{n \equiv k} (k^m - l^m) \widehat{\phi}(n) + \sum_{|k| < q} \sum_{l \equiv k} \widehat{f}(l) i^m e^{ikx_j} (k^m - l^m).$$

First we bound $|A|$. Consolidating the sum over l back into $\widetilde{f}_k$ and applying Lemma 4, we get

$$|A| \leq \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \left| \frac{\widetilde{f}_k}{\widetilde{\phi}_k} \right| \left| \sum_{n \equiv k} (n C^{m-1} - k^m) \widehat{\phi}(n) \right| \leq C N^{-s} \sum_{\substack{k \in \mathcal{J}_N \\ |k| \geq q}} \left| \frac{\widetilde{f}_k k^m}{\widetilde{\phi}_k} \right| \leq C N^{-s} |f^{(m)}|_{\mathcal{A}_\phi},$$

where in the last step we used (7).

To bound $|B|$, we collapse the sum over n into $\widetilde{\phi}_k$, cancel, then combine the terms to get

$$B = i^m e^{ikx_j} \sum_{k \in \mathcal{J}_N} \sum_{l \equiv k} \widehat{f}(l)(k^m - l^m).$$

Since $k \in \mathcal{J}_N$, the inner sum is only nonzero when $|l| \geq N/2$, giving us

$$|B| \leq \sum_{k \in \mathcal{J}_N} \sum_{\substack{l \equiv k \\ |l| \geq N/2}} |\widehat{f}(l)(k^m - l^m)| \leq 2^m \sum_{k \in \mathcal{J}_N} \sum_{\substack{l \equiv k \\ |l| \geq N/2}} |\widehat{f}(l) \, l^m| = 2^m \sum_{|n| \geq N/2} |\widehat{f}(n) \, n^m|,$$

where we have used the fact that $|k| \leq N/2 \leq |l|$ for all indices l in the inner sum. Now, since $\widehat{\phi}$ decays as in (9), we get that $1 \leq C N^{-s}/\widehat{\phi}(n)$ for all $|n| \geq N/2$, which gives us

$$|B| \leq 2^m \sum_{|n| \geq N/2} |\widehat{f}(n) \, n^m| \leq C N^{-s} \sum_{|n| \geq N/2} \frac{|\widehat{f}(n) \, n^m|}{\widehat{\phi}(n)} \leq C N^{-s} |f^{(m)}|_{\mathcal{A}_\phi}.$$

This finishes the proof.

4 Examples of Superconvergent Kernels

We now discuss classes of kernels satisfying Assumption 1, thus guaranteeing super-convergence in iterated derivatives. To motivate what comes next, recall example 1, where the kernel ψ_1 did not experience superconvergence in the first derivative, while ψ_2 did. Going back to Assumption 1, note that the terms in the symmetric differences of $\widehat{\psi}_2$ always lie on the same curve: either $1/n^4$ or $2/n^4$. However, the terms in the same symmetric differences of $\widehat{\psi}_1$ always come from different curves: one term will lie on $1/n^4$ and the other on $2/n^4$. The resulting "gaps" turn out to be too big to give any added convergence. If however the coefficients do not have such freedom for increasing n, the PBF will get added convergence in its iterated derivatives. We make this precise in the proposition below.

Proposition 2 *Let ϕ be a PBF and suppose that there is a positive constant C such that*

$$\widehat{\phi}(n) = C|n|^{-s} + \mathcal{O}(|n|^{-s-1}). \tag{14}$$

Then ϕ satisfies (9) and Assumption 1, and hence experiences superconvergence on uniform node sets in its iterated derivatives.

Proof Such a kernel obviously satisfies (9). Let $n \in \mathbb{Z}$ and let k be such that $|k| \leq |n|/2$. The estimate in Assumption 1 holds when $k = 0$, so we assume that $|k| \geq 1$. Also, note that $n - k$ and $n + k$ share the same sign, so we assume without loss of generality that $n - k$ and $n + k$ are both nonnegative. The mean value theorem gives us

$$|f(x) - f(y)| \leq \|f'\|_{L_\infty([x,y])}|x - y|,$$

which we will apply to the function $f(x) = x^{-s}$ on the interval with endpoints $n - k$ and $n + k$. Note that since $f'(x) = -s/x^{-s-1}$ is monotone, it will achieve its maximum at one of the endpoints. Therefore we get

$$|(n - k)^{-s} - (n + k)^{-s}| \leq s \left|\frac{n}{2}\right|^{-s-1} 2|k| = s2^{s+1}|n|^{-s-1}|k|, \tag{15}$$

where we have used the fact that $\max\{|n - k|, |n + k|\} \geq |n|/2$. Using (14) in conjunction with (15) and the fact that $1 \leq |k|$ gives us

$$\left|\widehat{\phi}(n - k) - \widehat{\phi}(n + k)\right| \leq C \left|(n - k)^{-s} - (n + k)^{-s}\right| + \mathcal{O}(|n/2|^{-s-1})$$
$$\leq s2^{s+1}|n|^{-s-1}|k| + C|k|n^{-s-1} = C|k||n|^{-s-1}.$$

Thus ϕ satisfies Assumption 1.

Next are some examples of kernels satisfying (14).

4.1 Bernoulli Splines

The univariate *Bernoulli splines* serve as kernels for periodic spline interpolation operators and can be defined on $[0, 2\pi]$ by the Fourier series:

$$B_s(x) = \sum_{|n|>0} \frac{1}{(in)^s} e^{inx}, \tag{16}$$

where $s \in \mathbb{N}$ [18]. For even s we have

$$B_s(x) = \sum_{|n|>0} \frac{1}{(in)^s} e^{inx} = \frac{2}{i^s} \sum_{n=1}^{\infty} \frac{1}{n^s} \cos(nx), \tag{17}$$

and thus up to a constant factor, B_s is CPD of order 1 and clearly satisfies (14), so Theorem 3 applies. Superconvergence in iterated derivatives for odd-degree splines is well-known and has been shown using different methods (see, for example [17]). Theorem 3 gives an alternate proof.

4.2 Radial Basis Functions

We now consider classes of PBFs obtained by restricting Euclidean radial basis functions (RBFs) to the circle (for the reader unfamiliar with RBFs, we recommend the books [4] or [20]). Let $\Phi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be an RBF defined on $\mathbb{R}^d$, $d \geq 2$. These functions are conditionally positive definite (with an amended subspace of multivariate polynomials, say of order q), radial, and univariate in the sense that $\Phi(\mathbf{x}, \mathbf{y}) = \varphi(\|\mathbf{x} - \mathbf{y}\|)$ for some univariate φ. Thus we will frequently express RBFs as $\Phi(\mathbf{x}, \mathbf{y}) = \Phi(r)$, where $r = \|\mathbf{x} - \mathbf{y}\|$. It is well known that if an RBF is conditionally positive definite on $\mathbb{R}^d$ of order q, then its restriction to the sphere $\mathbb{S}^{d-1}$ gives a CPD zonal kernel, also of order q [2]. In particular, the restriction of any RBF Φ to the circle is a zonal PBF: if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ are both length 1, then

$$\Phi(\|\mathbf{x} - \mathbf{y}\|) = \Phi\left(\sqrt{2(1 - \cos(\theta))}\right) =: \phi(\theta)$$

where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$.[3] Below we consider classes of these functions, and show that the associated zonal PBFs satisfy (14). Important to us will be the interplay between the Fourier transform of the kernel Φ in $\mathbb{R}^d$ and Fourier coefficients of the associated PBF ϕ on the circle (see [15,24]).

4.2.1 Thin Plate Splines and Power Kernels

Consider the family of radial kernels on $\mathbb{R}^d$ given by

$$\Phi_\beta(r) = C_\beta \begin{cases} r^\beta & \beta > 0, \beta \notin 2\mathbb{N}, \\ r^\beta \log(r) & \beta \in 2\mathbb{N}, \end{cases}$$

where C_β a constant. When $\beta \notin 2\mathbb{N}$, the *power kernel* $\Phi_\beta = C_\beta r^\beta$ is conditionally positive definite of order $q = \lceil \frac{\beta}{2} \rceil$ on $\mathbb{R}^d$, and in the other case we get the *thin plate spline* $\Phi_\beta(r) = C_\beta r^\beta \log(r)$, which is CPD of order $q = \beta/2 + 1$ [20, Corollary 8.18]. Asymptotic formulas for the spherical Fourier coefficients of these kernels were worked out in [15]. In particular, when one of these RBFs is restricted from $\mathbb{R}^2$ to the circle, using [15, Equation (4.12)] (with the substitutions $n = 2$, $2s - 2 = \beta$) we immediately get

$$\widehat{\phi_\beta}(n) = \frac{\Gamma(\beta + 1)}{2^{\beta+2} \pi^2 \Gamma^2 \left(\frac{\beta}{2} + 1\right)} n^{-\beta-1} \left(1 + \mathcal{O}(n^{-1})\right),$$

so Proposition 2 applies.

4.2.2 Matérn Kernels

Next we consider the Matérn kernels, also known as Sobolev splines. The Matérn kernel of order $\beta > d/2$ on $\mathbb{R}^d$, which we denote by K_β, satisfies

$$\mathcal{K}_\beta(\xi) = \frac{1}{(1 + |\xi|^2)^\beta},$$

---

[3] A capital letter will be used to denote an RBF, with the corresponding PBF denoted by the same letter in lower case.

where here $\mathcal{K}_\beta$ is the the Euclidean Fourier transform of K_β over $\mathbb{R}^d$ [4, Section 4.4]. We will denote the restriction of K_β to $\mathbb{T}$ by k_β. The following relates $\widehat{k_\beta}$ with $\mathcal{K}_\beta$ (with $d = 2$):

$$\widehat{k_\beta}(n) = \int_0^\infty t\mathcal{K}_\beta(t)J_n^2(t)dt = \int_0^\infty \frac{t}{(1+t^2)^\beta}J_n^2(t)dt, \tag{18}$$

where J_n is the order n Bessel function of the first kind [15, Proposition 3.1]. Estimates on $\widehat{k_\beta}(n)$ were carried out in [15] to show that $\widehat{k_\beta}(n) \sim n^{-2\beta+1}$ for large n. We repeat some of those arguments here to show that this kernel satisfies Proposition 2 with $s = -2\beta + 1$.

First, for $t > 1$ the following expansion is valid:

$$\frac{t}{(1+t^2)^\beta} = \frac{1}{t^{2\beta-1}} - \frac{\beta}{t^{2\beta+1}} + \mathcal{O}(t^{-2\beta-3}).$$

Thus we may rewrite the integral in (18) as

$$\widehat{k_\beta}(n) = \int_0^\infty \frac{t}{(1+t^2)^\beta}J_n^2(t)dt = A + B + C,$$

where

$$A = \int_0^\infty t^{-2\beta+1}J_n^2(t)dt, \qquad B = \int_0^1 \left(\frac{t}{(1+t^2)^\beta} - \frac{1}{t^{2\beta-1}}\right)J_n^2(t)dt,$$

$$C = \int_1^\infty \mathcal{O}(t^{-2\beta-1})J_n^2(t)dt.$$

The term A was computed in [15, Equation (4.10)]. The result is:

$$A = \frac{\Gamma(2\beta-1)}{2^{2\beta-3}\Gamma^2(\beta)}n^{-2\beta+1}(1 + \mathcal{O}(n^{-1})),$$

for n sufficiently large. The term B can be handled with the bound $|J_n(t)| \leq \frac{2^{-n}|t|^n e^{|\mathrm{Im}(t)|}}{\Gamma(n+1)}$ [19, Page 49, Equation (1)]—since $\mathrm{Im}(t) = 0$ this implies that B decays faster than any power of n^{-1}. The term C can be bounded by an integral similar to A:

$$|C| = \int_1^\infty \mathcal{O}(t^{-2\beta-1})J_n^2(t) \leq c\int_0^\infty t^{-2\beta-1}J_n^2(t)dt = \mathcal{O}(n^{-2\beta-1}).$$

It follows that

$$\widehat{k_\beta}(n) = \frac{\Gamma(2\beta-1)}{2^{2\beta-3}\Gamma^2(\beta)}n^{-2\beta+1}(1 + \mathcal{O}(n^{-1})),$$

Thus k_β also satisfies Proposition 2 with $s = -2\beta + 1$.

5 Numerical Experiments on the Circle

In this section we present several numerical experiments. The first verifies Theorem 3 and Proposition 1 for a thin plate spline, while the next two give evidence of superconvergence for the restriction of a Wendland kernel and the inverse multiquadric (IMQ) kernel to the circle. Note that the theory of section 4 does not cover the latter two kernels.

In all the experiments that follow, we use the same smooth target function

$$g(x) = e^{-4\cos(x)}\sin(4(x-1)), \tag{19}$$

and for each kernel we calculate the relative pointwise error at the interpolation nodes in the first, second and sixth iterated derivatives, as well as the relative error in using $(I_X g)''$ to approximate g''. For the first two kernels, we present results for both nonuniform node sets and uniform node sets of increasing size. The former node sets are used to illustrate the differences in convergence rates when the nodes are not uniform. These nonuniform node sets are generated so that roughly a third of the nodes are equally spaced on the upper half of the circle, and the rest are equally spaced on the lower half.[4] This results in a "mesh ratio" $\rho = h/q \sim 2$, where h is half of the largest gap between two successive nodes and q is half of the shortest gap between successive nodes.

5.1 A Thin Plate Spline

The main purpose of this experiment is to verify the results of Theorem 3. We use the thin plate spline

$$\Phi_4(r) = r^4 \log(r),$$

which is C^3 and conditionally positive definite of order 3. The discussion in section 4.2.1 shows that the restriction of this kernel to the circle has Fourier coefficients satisfying (9) with $s = 5$. Therefore if this kernel experiences superconvergence in its iterated derivatives the errors should all converge as $\mathcal{O}(N^{-5})$ on uniform nodes.

Figures 2(a) and 2(b) show the convergence results for non-uniform and uniform nodes, respectively. In the non-uniform case (Figure 2(a)), we see that the convergence rate is not preserved for the iterated derivatives with the error in the first derivative converging like $\mathcal{O}(N^{-4})$, the second like $\mathcal{O}(N^{-3})$, and the sixth not converging at all. We also see that $\mathcal{D}_X^2 g$ and $(I_X g)''$ approximate g'' at the same rate, however the errors are smaller for the iterated derivatives.

The results for the uniform nodes (Figure 2(b)) exactly match the theory with all iterated derivatives providing $\mathcal{O}(N^{-5})$ approximations. We also see that $(I_X g)''$ only provides a $\mathcal{O}(N^{-3})$ approximation of g'', as suggested by Proposition 1. It is interesting to note that, despite the fact that the kernel is only C^3, the sixth iterated derivative still approximates $g^{(6)}$ like $\mathcal{O}(N^{-5})$. However, for the largest node sets the sixth iterated derivative approximation begins to degrade. This should perhaps be expected since the construction of $\mathcal{D}_X^6 g$ involves six applications of the kernel interpolation operator, which is not perfectly conditioned. This prompted

---

[4] More precisely, given N, we first choose $n_1 = \lceil N/3 \rceil$ equally spaced points on the interval $[0, \pi - \pi/n_1]$. Next choose $n_2 = N - n_1 = \lfloor 2N/3 \rfloor$ equally spaced points on $[\pi, 2\pi - \pi/n_2]$. This gives node sets of size N with $\rho \sim 2$ (and $\rho = 2$ when N is divisible by 3).
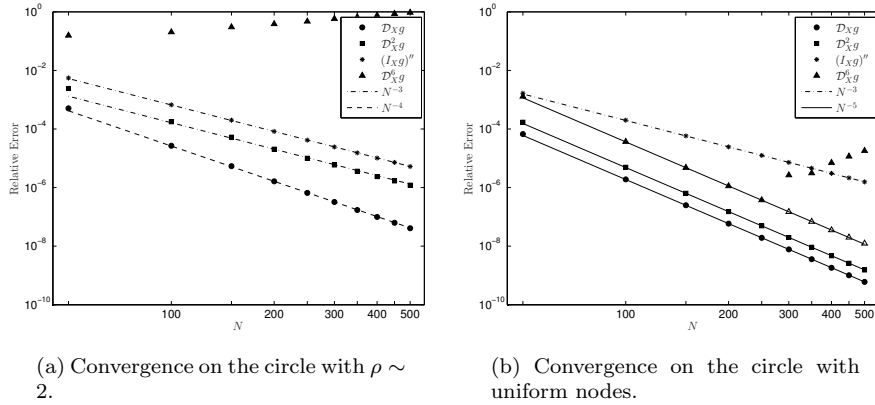
(a) Convergence on the circle with $\rho \sim 2$.

(b) Convergence on the circle with uniform nodes.

Fig. 2 Relative ℓ_∞ error (measured at the nodes) for derivative approximations of (19) using the restriction of the thin plate spline Φ_4 to the circle. The plot on the left shows the results using node sets with $\rho \sim 2$, and on the right the same experiments are repeated on uniform nodes. The unfilled triangles in 2(b) represent the results using quad-precision.

a repeat of the experiment using quad-precision (32-digits) instead of double-precision, and in this case the error continues to decrease like $\mathcal{O}(N^{-5})$.

5.2 A Wendland kernel

Here we consider the restriction of the following Wendland kernel to the circle:

$$\Phi_{3,2}(r) := (1-r)_+^6 (35r^2 + 18r + 1),$$

where $(x)_+ = x$ if $x \geq 0$ and is zero otherwise. This kernel is C^4 and positive definite on $\mathbb{R}^d$ for $d \leq 3$. Additionally, It can be shown that $\Phi_{3,2}$ restricted to the circle has Fourier coefficients satisfying (9) with $s = 6$.

The results for the non-uniform and uniform nodes are presented in Figures 3(a) and 3(b), respectively. We see trends similar to those in the previous example. The rate of approximation for the iterated derivatives is not preserved in the nonuniform case, and $\mathcal{D}_X^2 g$ and $(I_X g)''$ approximate g'' at the same rate. In the uniform node case, Figure 3(b) indicates that the iterated derivatives are experiencing superconvergence with the first, second and sixth iterated derivatives providing $\mathcal{O}(N^{-6})$ approximations. This exactly matches the decay rate of the Fourier coefficients of $\Phi_{3,2}$. Finally, we see $(I_X g)''$ is still only providing an $\mathcal{O}(N^{-4})$ approximation of g'', as in the non-uniform node case.

5.3 Inverse Multiquadric

We now consider the inverse multiquadric (IMQ) kernel, given by:
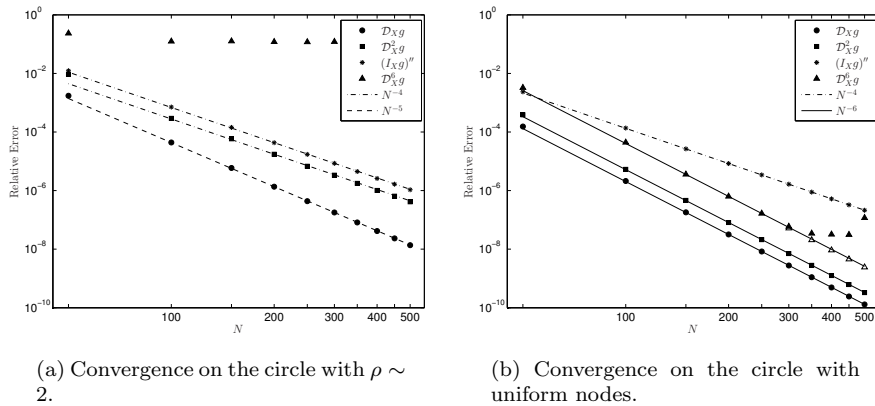
$$\Phi(r) = \frac{1}{\sqrt{1 + (\varepsilon\, r)^2}},$$

(a) Convergence on the circle with $\rho \sim$ 2.

(b) Convergence on the circle with uniform nodes.

Fig. 3 Relative ℓ_∞ error (measured at the nodes) for derivative approximations of (19) using the restriction of $\Phi_{3,2}$ to the circle. The plot on the left shows the results using node sets with $\rho \sim 2$, and on the right the same experiments are repeated on uniform nodes. The unfilled triangles in 3(b) represent the results using quad precision (32 digit accuracy).

where ε is the *shape parameter*. This positive definite, C^∞ kernel is known to enjoy convergence rates higher than any polynomial order for smooth target functions [20, Section 11.4]. It has generally been reported in the literature that there is typically an optimal value of ε that produces the best accuracy in the IMQ interpolants and that this value tends to decrease with increasing smoothness of the target function [4]. Furthermore, in the limit of $\varepsilon \to 0$, it can be shown using the arguments from [5] that the IMQ interpolant (and many other smooth RBF interpolants) restricted to the circle converges to a trigonometric interpolant. However, for small values of ε, one must use a "stable algorithm" to compute the interpolant such as the RBF-QR procedure of Fornberg and Piret [5].

Figures 4(a) and 4(b) show the convergence results for the IMQ kernel on uniform nodes for $\varepsilon = 0.9$ and $\varepsilon = 3.6$, respectively, which are considered relatively small and relatively large values for the shape parameter. Unlike the previous results, these are plotted on a semilog scale to highlight the exponential convergence rates. Also, since the IMQ kernel is C^∞, we have included results for $(I_X g)^{(6)}$ to compare to $\mathcal{D}_X^6 g$.

For the $\varepsilon = 0.9$ case in Figure 4(a), which were obtained using the RBF-QR algorithm, we see that both the iterated and direct derivatives converge rapidly for all orders. While the error in the iterated derivatives is slightly larger than the direct derivatives (for most values of N), the convergence is just slightly faster and seems to match the convergence of the first derivative. The similarity in the results for the iterated and direct derivatives for this relatively small value of ε is most likely related to the IMQ interpolant being very close to a trigonometric interpolant. In the latter case, the iterated and direct derivatives give the same result when N is odd (as is the case here).

The results in Figure 4(b) for $\varepsilon = 3.6$ show a much more pronounced difference in the convergence rates for the iterated and direct derivatives. All iterated derivatives appear to be experiencing superconvergence, while convergence of the direct
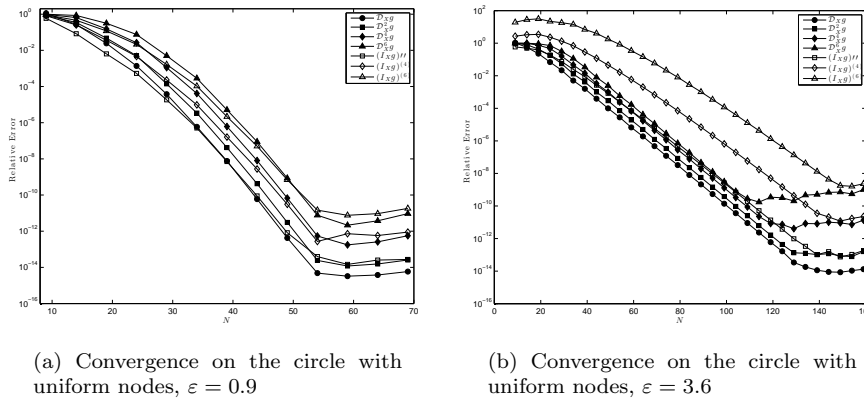
(a) Convergence on the circle with uniform nodes, $\varepsilon = 0.9$

(b) Convergence on the circle with uniform nodes, $\varepsilon = 3.6$

Fig. 4 Semilog plot of relative ℓ_∞ error (measured at the nodes) for iterated and direct derivative approximations of (19) using the restricted IMQ kernel with (a) $\varepsilon = 0.9$ and (b) $\varepsilon = 3.6$. All results are for uniform nodes on the circle.

derivatives slows with increasing order of the derivative. Although not presented, we did several more experiments that indicated this superconvergence behavior for the iterated derivatives with the IMQ kernel at other values of ε.

We note, as commented above, that we should expect that the results for $\varepsilon = 0.9$ to be better than $\varepsilon = 3.6$ for this example, since the target function is very smooth. This may not always be the case in practice, and these results indicate that when using the IMQ kernel iterated derivatives may provide much better approximations to higher order derivatives than the direct approach.

6 Iterated Derivatives on Periodic Surfaces

In this section we consider an iterated approximation of the surface Laplacian (Laplace-Beltrami operator) $\Delta_{\mathbb{M}}$, where $\mathbb{M} \subset \mathbb{R}^3$ denotes the surface in question. This is the prominent differential operator in time-dependent diffusion and reaction-diffusion equations on surfaces, which have a number of applications. Working out the surface Laplacian analytically can be a cumbersome task for a general surface. Below, we describe the iterated approach taken in [8] to approximate this operator, which only requires having nodes on the surface, the corresponding (numerical) normal vectors to the surface at the nodes, and the Euclidean first order derivatives of the kernel— this significantly simplifies the task of approximating $\Delta_{\mathbb{M}}$. Next we consider numerical examples on a torus and a unit sphere, where we study the convergence rates of the iterated and a direct approach to approximating $\Delta_{\mathbb{M}}$.

6.1 Iterated Surface Laplacian

We let $\nabla_{\mathbb{M}}$ and $\nabla_{\mathbb{M}}\cdot$ denote the surface gradient and surface divergence, respectfully. These operators can be expressed entirely in Cartesian coordinates as follows.

Let $\mathbf{x} = (x, y, z)$ be a point on $\mathbb{M}$ and $\mathbf{n} = (n^x, n^y, n^z)$ the normal vector to $\mathbb{M}$ at $\mathbf{x}$. The surface gradient at $\mathbf{x}$ is then given as

$$\nabla_{\mathbb{M}} := \mathbf{P}\nabla = \left(\mathbf{I} - \mathbf{n}\mathbf{n}^T\right)\nabla, \tag{20}$$

where $\mathbf{P}$ is a 3-by-3 matrix that projects any vector in $\mathbb{R}^3$ at $\mathbf{x}$ to a vector tangent to $\mathbb{M}$ at $\mathbf{x}$. Letting $\mathbf{e}^x$, $\mathbf{e}^y$, $\mathbf{e}^z$, be the standard unit vectors in x, y, and z directions in $\mathbb{R}^3$, we can re-write (20) in component form as

$$\nabla_{\mathbb{M}} := \begin{bmatrix} (\mathbf{e}^x \cdot \mathbf{P})\nabla \\ (\mathbf{e}^y \cdot \mathbf{P})\nabla \\ (\mathbf{e}^z \cdot \mathbf{P})\nabla \end{bmatrix} = \begin{bmatrix} (\mathbf{e}^x - n^x\mathbf{n})\cdot\nabla \\ (\mathbf{e}^y - n^y\mathbf{n})\cdot\nabla \\ (\mathbf{e}^z - n^z\mathbf{n})\cdot\nabla \end{bmatrix} = \begin{bmatrix} \mathbf{p}^x \cdot \nabla \\ \mathbf{p}^y \cdot \nabla \\ \mathbf{p}^z \cdot \nabla \end{bmatrix} = \begin{bmatrix} \mathcal{G}^x \\ \mathcal{G}^y \\ \mathcal{G}^z \end{bmatrix}, \tag{21}$$

where $\mathbf{p}^x$ denotes the first row of $\mathbf{P}$, etc. The surface divergence of a smooth vector field $\mathbf{f} = (f^x, f^y, f^z) : \mathbb{M} \to \mathbb{R}^3$ at a point $\mathbf{x} \in \mathbb{M}$ can then be expressed as

$$\nabla_{\mathbb{M}} \cdot \mathbf{f} := (\mathbf{P}\nabla) \cdot \mathbf{f} = \mathcal{G}^x f^x + \mathcal{G}^y f^y + \mathcal{G}^z f^z. \tag{22}$$

Let $X = \{\mathbf{x}_j\}_{j=1}^N \subset \mathbb{M}$ and $\Phi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ be an RBF on $\mathbb{R}^3$. A function f can be interpolated at X using

$$I_X f = \sum_{x_j \in X} c_j \Phi(\|\mathbf{x} - \mathbf{x}_j\|). \tag{23}$$

The approximation properties of these types of interpolants on surfaces $\mathbb{M} \subset \mathbb{R}^d$ has previously been studied in [7]. One approach to compute an approximation to $\Delta_{\mathbb{M}} f$ is by analytically applying it to the interpolant (23), i.e.

$$\Delta_{\mathbb{M}} I_X f \sim \Delta_{\mathbb{M}} f. \tag{24}$$

An alternative is the following iterated approach. Motivated by the fact that $\Delta_{\mathbb{M}} = \nabla_{\mathbb{M}} \cdot \nabla_{\mathbb{M}}$, we replace the first order differentials with the process "interpolate, then differentiate." That is, first f is interpolated using (23), and then the Cartesian components of the surface gradient (21) are computed of $I_X f$. Next, each component of the resulting vector field is interpolated using (23), and the surface divergence (22) is then applied to these interpolants. This "iterated Laplacian" of f can be expressed as

$$L_X f := \nabla_{\mathbb{M}} \cdot \mathbf{I}_X \left(\nabla_{\mathbb{M}} I_X f\right) \sim \Delta_{\mathbb{M}} f, \tag{25}$$

where $\mathbf{I}_X \mathbf{g}$ is obtained by applying I_X to each component of a 3-dimensional vector field $\mathbf{g}$. For general surfaces, this iterated approach is simpler to construct than the direct approach (24) in that it does not require computing derivatives of the normal vectors and higher order derivatives of the radial kernel ϕ. More details on constructing L_X, including `MATLAB` code, can be found in [8].

6.2 Setup for the Numerical Experiments

In all experiments that follow, we use the C^4 Matérn radial kernel

$$\Phi(r) = e^{-\varepsilon r}\left((\varepsilon r)^2 + 3(\varepsilon r) + 3\right). \tag{26}$$

Although not presented, similar results were observed for other kernels of finite smoothness. For most experiments, we set the shape parameter to $\varepsilon = 5$. Given a surface $\mathbb{M}$, node set X, and target function f, the relative errors for the direct Laplacian (24) and iterated Laplacian (25) will be calculated as follows:

$$\frac{\|\Delta_{\mathbb{M}} f - \Delta_{\mathbb{M}} I_X f\|_{\ell_\infty(X)}}{\|\Delta_{\mathbb{M}} f\|_{\ell_\infty(X)}} \text{ and } \frac{\|\Delta_{\mathbb{M}} f - L_X f\|_{\ell_\infty(X)}}{\|\Delta_{\mathbb{M}} f\|_{\ell_\infty(X)}}.$$

For kernels of finite smoothness, like (26), typically the error decays algebraically with the largest "gap" in the node set X, which is given by the mesh norm: $h := \sup_{x \in \mathbb{M}} \text{dist}(x, X)$. Note that the surfaces considered are two-dimensional, so that if the node sets are quasi-uniform then $1/h \sim \sqrt{N}$, where N is the total number of nodes.

6.3 Experiments on a Torus

The torus we consider is given by the implicit equation:

$$\mathbb{M} = \left\{ \mathbf{x} = (x, y, z) \in \mathbb{R}^3 \;\middle|\; \left(1 - \sqrt{x^2 + y^2}\right)^2 + z^2 - \frac{1}{9} = 0 \right\}, \tag{27}$$

which leads to the following natural parameterization:

$$x = \left(1 + \frac{1}{3}\cos(p)\right)\cos(t), \; y = \left(1 + \frac{1}{3}\cos(p)\right)\sin(t), \; z = \frac{1}{3}\sin(p), \quad 0 \le p, t \le 2\pi.$$

The surface Laplacian in this intrinsic coordinate system is

$$\Delta_{\mathbb{M}} f(p, t) = \frac{1}{\left(1 + \frac{1}{3}\cos(p)\right)^2}\frac{\partial^2 f}{\partial t^2} + \frac{9}{\left(1 + \frac{1}{3}\cos(p)\right)}\frac{\partial}{\partial p}\left(\left(1 + \frac{1}{3}\cos(p)\right)\frac{\partial f}{\partial p}\right).$$

The target function we use in the experiments is given, in Cartesian coordinates, as

$$f(x, y, z) = \cos(10z + 3)e^{-x^2} - \sin(4x + 5y).$$

This function was chosen because of it smoothness and the fact that it does not have any obvious symmetries on $\mathbb{M}$. We omit the analytical surface Laplacian of f for brevity, but note that a MATLAB m-file representing $\Delta_{\mathbb{M}} f$ can be downloaded from [6].

We use two families of nodes in the experiments. The first is gridded in the parameter space (see Figure 5(a)), and determined as follows. Given m, choose m equally spaced angles on $[0, 2\pi)$ (in t) and $3m$ equally spaced angles on $[0, 2\pi)$ (in p) - then take a direct product to obtain $N = 3m^2$ points. The experiments were carried out with $N = 300, 1200, \ldots, 19200$ gridded nodes (i.e. $m = 10, 20, ..., 80$). These points remain more or less uniformly spaced on the torus as N grows. The
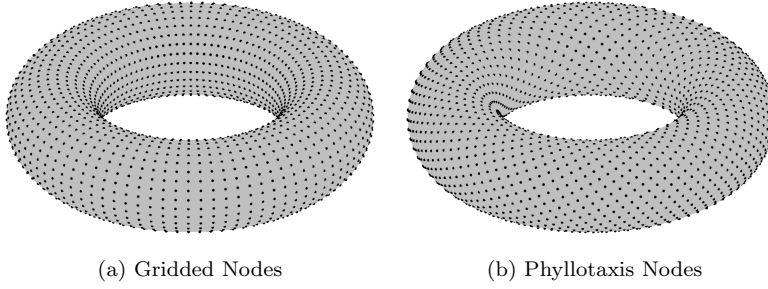
(a) Gridded Nodes   (b) Phyllotaxis Nodes

Fig. 5 Node sets on the torus.

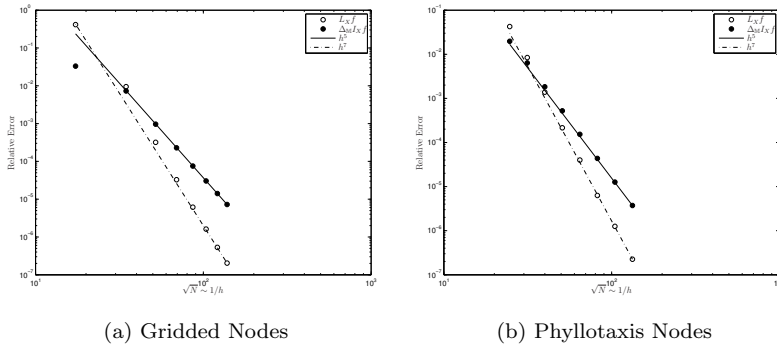(a) Gridded Nodes   (b) Phyllotaxis Nodes

Fig. 6 Relative error at the nodes between $\Delta_{\mathbb{M}} f$ and the approximations (24) and (25) using the C^4 Matérn kernel on the torus.

second family of nodes considered are *phyllotaxis spiral* nodes, which are sometimes used to mimic certain botanical features on surfaces. These can be generated on any surface of revolution based on the ideas in [16]. When N is a Fibonacci number these nodes happen to be highly uniform, so we used phyllotaxis nodes with cardinallity $610, 987, \ldots, 17711$. These nodes are also available to download from [6].

Figures 6(a) and 6(b) display the convergence rates for the iterated and direct surface Laplacian of the RBF interpolant for the gridded and phyllotaxis node families, respectively. We see that for both families of nodes, the direct surface Laplacian of the interpolant is converging like $\mathcal{O}(h^5)$, while the iterated Laplacian is converging a full two-orders faster, i.e. like $\mathcal{O}(h^7)$. The observed convergence rate at the nodes for the iterated version is the same as the expected global convergence rate of the RBF interpolant of this target function [7], indicating the iterated version is experiencing superconvergence.

6.4 Experiments on the Unit Sphere

We next compare the iterated and direct Laplacian on the unit sphere, $\mathbb{S}^2$. As target functions we use the spherical harmonics (see [14]) $Y_{6,0}$ and $Y_{6,4}$, which are plotted in Figure 7.
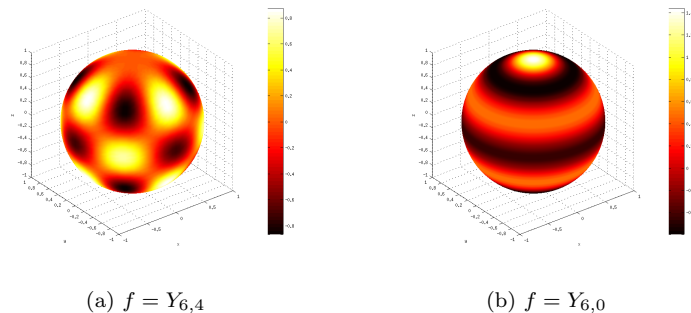

(a) $f = Y_{6,4}$ (b) $f = Y_{6,0}$

Fig. 7 Spherical harmonic target functions for the sphere.

We use three families of nodes in the experiments. The first family is the *minimum energy* (ME) nodes of Womersley and Sloan [21]. These form highly uniform node sets, and do not lie on a regular grid; see Figure 8(a) for a plot of these nodes for $N = 2500$. For the experiments with these node sets, we cardinalities of $N = 49, 100, 225, 400, 625, 1024, 2500, 3600, 5776, 7225, 10000, 16384$. The second family is the gridded nodes obtained from a direct product of m equally spaced nodes in latitude and and m equally spaced nodes in longitude. We use $m = 10, 20, \ldots, 130$, which results in node sets of size $N = 74, 344, 814, \ldots, 14044, 16514$, once duplicates on the prime meridian and poles are removed. The $N = 1484$ set of these nodes is displayed in Figure 8(b), and we clearly see that these are not quasi-uniform due to the clustering near the poles. Lastly, we use the so called *Fibonacci* family of nodes [11], which are based on the idea of phyllotaxis spiral nodes. As in the torus case, the node sets used correspond to Fibonacci numbers: $N = 377, 610, \ldots, 17711$. These node sets and `MATLAB` code to generate them, are available at [6].

We first focus on the results for the ME nodes. Figures 9(a) and 9(b) show the convergence rates for the $Y_{6,4}$ and $Y_{6,0}$ target functions, respectively, using the iterated and direct surface Laplacian. We see that for the smaller node sets it appears that the iterated surface Laplacian for both targets is converging like $\mathcal{O}(h^7)$, while the direct Laplacian is converging like $\mathcal{O}(h^5)$. However, as the node sets grow we see that both the iterated and direct Laplacians clearly converge like $\mathcal{O}(h^5)$ for both targets. Although, the errors in the iterated method are still much smaller.

A different trend is observed for the other two node families; namely, that convergence in the iterated Laplacian $L_X f$ was considerably faster for one of the targets, but not the other. Figures 11(a) and 12(a) display the convergence rates for the $Y_{6,4}$ target function using the gridded and Fibonacci nodes, respectively. We
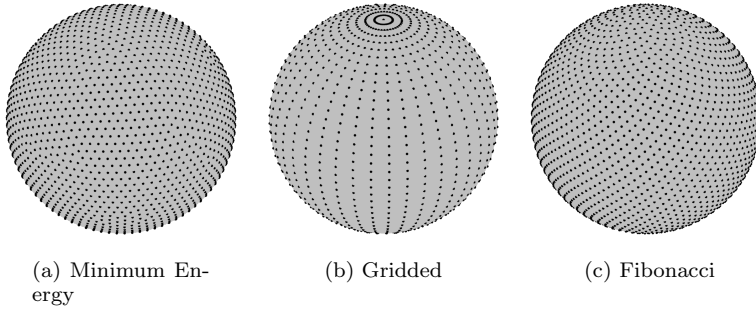
(a) Minimum En-   (b) Gridded   (c) Fibonacci
ergy

Fig. 8 Node sets on the sphere used in the numerical experiments.

(a) $f = Y_{6,4}$ (b) $f = Y_{6,0}$
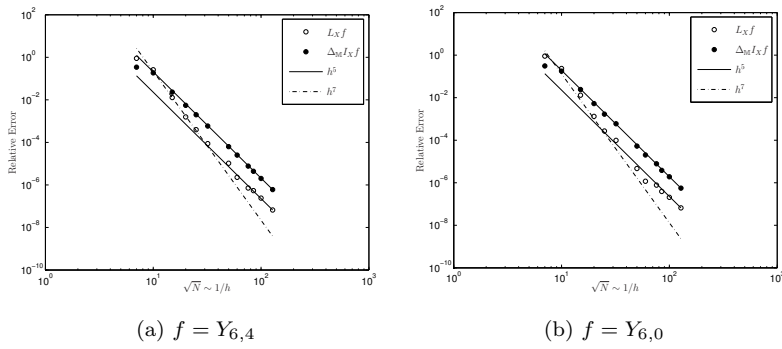
Fig. 9 Relative error at the ME nodes between $\Delta_{\mathbb{M}} f$ and the direct and iterated approxima-
tions, (24) and (25), respectively, using the C^4 Matérn kernel on the sphere with $\varepsilon = 5$. While
the iterated Laplacian appears to be converging faster for smaller node sets, eventually the
convergence rates of the two methods are comparable.

can clearly see that while the direct Laplacian is converging like $\mathcal{O}(h^5)$, the iterated
Laplacian is converging like $\mathcal{O}(h^7)$. However, for the $Y_{6,0}$ target, Figures 11(b) and
12(b) clearly show both the direct and iterated approximations converging at the
same slower rate of $\mathcal{O}(h^5)$.

It's worth noting two things regarding these results: (1) the gridded and Fi-
bonacci node sets are less uniform at the poles, and (2) from Figure 7 it is clear
that the support of $Y_{6,4}$ is away from the poles, while the support of $Y_{6,0}$ is not.
It is reasonable to then consider measuring the error away from the poles, where
the nodes are more uniform. As a follow up experiment, we repeat the previous
tests on the $Y_{6,0}$ target, but instead of measuring the relative error over the entire
node set X, we only measure it over the sites $X' \subset X$ consisting of nodes that lie
outside the polar caps of radius $\theta = \pi/16$ (see Figure 10). The outlined squares in
Figures 11(b) and 12(b) display the resulting convergence rates. Interestingly, we
see that on X' the iterated Laplacian $L_X f$ enjoys the faster convergence rate of
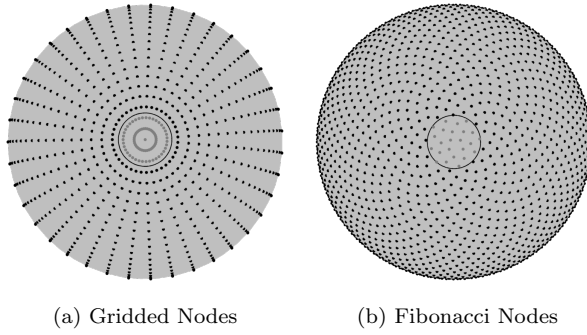$\mathcal{O}(h^7)$.

(a) Gridded Nodes   (b) Fibonacci Nodes

Fig. 10 View from the north pole of the node sets X and X' used in the experiments involving the gridded and Fibonacci nodes on the sphere. Here X', depicted in the figure by the darker spots, denotes all nodes from X outside the northern and southern polar caps of radius $\theta = \pi/16$.
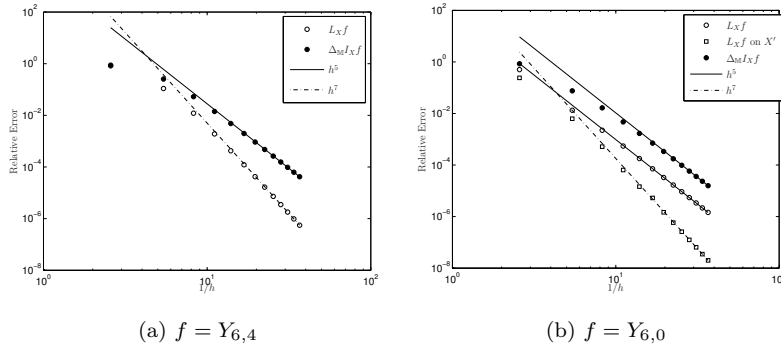

(a) $f = Y_{6,4}$ (b) $f = Y_{6,0}$

Fig. 11 Relative error at the gridded nodes between $\Delta_{\mathbb{M}} f$ and the direct and iterated approximations, (24) and (25), respectively, using the C^4 Matérn kernel. Here the shape parameter ε is fixed at $\varepsilon = 10$ to keep the conditioning of the interpolation matrices under control. The nonuniformity of the nodes necessitated a calculation of the mesh norm h, which we estimated by using the largest distance between neighboring nodes on the equator. Note that while superconvergence is not observed for the target $Y_{6,0}$ over the entire sphere, there is accelerated convergence away from the polar caps as indicated by the outlined square markers in part (b).

7 Concluding Remarks

This is the first study to compare the direct and iterated methods for approximating derivatives of periodic functions using radial basis functions. It is also the first study to investigate convergence rates that can expected with this method when errors are only measured *at the data sites*. In the univariate case and with uniform nodes, we have given a sufficient condition on the RBF kernel's Fourier coefficients to guarantee that every iterated derivative approximates the corresponding derivative of smooth periodic functions to the same high order at uniform data sites. We have established that thin plate splines, power functions, and Matérn kernels restricted to the circle all yield superconvergent approximation of derivatives at
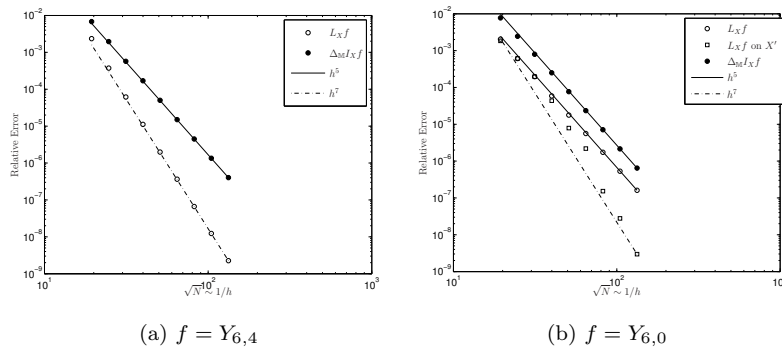
(a) $f = Y_{6,4}$ (b) $f = Y_{6,0}$

Fig. 12 Relative error at the Fibonacci nodes between $\Delta_{\mathbb{M}} f$ and the direct and iterated approximations, (24) and (25), respectively, using the C^4 Matérn kernel with $\varepsilon = 5$. Again there is accelerated convergence away from the polar caps as indicated by the outlined square markers in part (b).

uniformly spaced data sites. Our numerical experiments suggest that the same is true for Wendland kernels, and the IMQ example in section 5.3 shows the iterated approach may be beneficial for some infinitely smooth kernels too — at least for larger values of the shape parameter.

The numerical experiments presented in section 6 show that superconvergence seems to extend to iterated derivatives on higher dimensional periodic domains, although the generality is not so obvious. While this behavior was clearly observed in the torus experiments, faster convergence on the sphere was not observed at the data sites in all the examples. However, accelerated convergence was observed for the gridded and Fibonacci families where the nodes were uniform, i.e. away from the poles. This leads one to wonder how "local" this phenomenon is: if the nodes are arranged in a way amenable to superconvergence in a subregion of the domain, at which locations (if any) will faster convergence rates occur within that region? Addressing this issue will be pursued in future work.

References

1. Milton Abramowitz and Irene A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.
2. Brad J. C. Baxter and Simon Hubbert. Radial basis functions for the sphere. In *Recent progress in multivariate approximation (Witten-Bommerholz, 2000)*, volume 137 of *Internat. Ser. Numer. Math.*, pages 33–47. Birkhäuser, Basel, 2001.
3. F.-J. Delvos. Approximation properties of periodic interpolation by translates of one function. *RAIRO Modél. Math. Anal. Numér.*, 28(2):177–188, 1994.
4. Gregory E. Fasshauer. *Meshfree approximation methods with MATLAB*, volume 6 of *Interdisciplinary Mathematical Sciences*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2007. With 1 CD-ROM (Windows, Macintosh and UNIX).

5. B. Fornberg and C. Piret. A stable algorithm for flat radial basis functions on a sphere. *SIAM J. Sci. Comp.*, 30:60–80, 2007.
6. E. J. Fuselier. Nodes used in order-preserving approximation of derivatives with periodic radial basis functions, accessed 2012. `http://math.highpoint.edu/~efuselier/OrderPreservingData/`.
7. E. J. Fuselier and G. B. Wright. Scattered Data Interpolation on Embedded Submanifolds with Restricted Positive Definite Kernels: Sobolev Error Estimates. *SIAM Journal on Numerical Analysis*, 50(3):1753–1776, 2012.
8. Edward J. Fuselier and Grady B. Wright. A high-order kernel method for diffusion and reaction-diffusion equations on surfaces. *Journal of Scientific Computing*, pages 1–31, 2013.
9. Q. T. Lê Gia. Approximation of parabolic pdes on spheres using spherical basis functions. *Adv. Comput. Math.*, 22:377–397, 2005.
10. Michael Golomb. Approximation by periodic spline interpolants on uniform meshes. *J. Approximation Theory*, 1:26–65, 1968.
11. Álvaro González. Measurement of areas on a sphere using Fibonacci and latitude-longitude lattices. *Math. Geosci.*, 42(1):49–64, 2010.
12. R. A. Lorentz, F. J. Narcowich, and J. D. Ward. Collocation discretizations of the transport equation with radial basis functions. *Appl. Math. Comput.*, 145(1):97–116, 2003.
13. Thomas R. Lucas. Error bounds for interpolating cubic splines under various end conditions. *SIAM J. Numer. Anal.*, 11:569–584, 1974.
14. Claus Müller. *Spherical harmonics*, volume 17 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1966.
15. Francis J. Narcowich, Xinping Sun, and Joseph D. Ward. Approximation power of RBFs and their associated SBFs: a connection. *Adv. Comput. Math.*, 27(1):107–124, 2007.
16. J. N. Ridley. Ideal phyllotaxis on general surfaces of revolution. *Math. Biosci.*, 79(1):1–24, 1986.
17. M. J. Shelley and G. R. Baker. Order-preserving approximations to successive derivatives of periodic functions by iterated splines. *SIAM J. Numer. Anal.*, 25(6):1442–1452, 1988.
18. Joachim Stöckler. Multivariate Bernoulli splines and the periodic interpolation problem. *Constr. Approx.*, 7(1):105–122, 1991.
19. G. N. Watson. *A treatise on the theory of Bessel functions*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1995. Reprint of the second (1944) edition.
20. Holger Wendland. *Scattered data approximation*, volume 17 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2005.
21. R. S. Womersley and I. H. Sloan. Interpolation and cubature on the sphere, accessed 2012. `http://web.maths.unsw.edu.au/~rsw/Sphere/`.
22. G. B. Wright, N. Flyer, and D. Yuen. A hybrid radial basis function - pseudospectral method for thermal convection in a 3D spherical shell. *Geochem. Geophys. Geosyst.*, 11:Q07003, 2010.
23. Yuan Xu and E. W. Cheney. Strictly positive definite functions on spheres. *Proc. Amer. Math. Soc.*, 116(4):977–981, 1992.
24. W. zu Castell and F. Filbir. Radial basis functions and corresponding zonal series expansions on the sphere. *J. Approx. Theory*, 134(1):65–79, 2005.