MATHEMATICS OF COMPUTATION
Volume 94, Number 351, January 2025, Pages 381–407
https://doi.org/10.1090/mcom/3960
Article electronically published on March 20, 2024

# EXTENDING ERROR BOUNDS FOR RADIAL BASIS FUNCTION INTERPOLATION TO MEASURING THE ERROR IN HIGHER ORDER SOBOLEV NORMS

T. HANGELBROEK AND C. RIEGER

Abstract. Radial basis functions (RBFs) are prominent examples for reproducing kernels with associated reproducing kernel Hilbert spaces (RKHSs). The convergence theory for the kernel-based interpolation in that space is well understood and optimal rates for the whole RKHS are often known. Schaback added the doubling trick [Math. Comp. 68 (1999), pp. 201–216], which shows that functions having double the smoothness required by the RKHS (along with specific, albeit complicated boundary behavior) can be approximated with higher convergence rates than the optimal rates for the whole space. Other advances allowed interpolation of target functions which are less smooth, and different norms which measure interpolation error. The current state of the art of error analysis for RBF interpolation treats target functions having smoothness up to twice that of the native space, but error measured in norms which are weaker than that required for membership in the RKHS.

Motivated by the fact that the kernels and the approximants they generate are smoother than required by the native space, this article extends the doubling trick to error which measures higher smoothness. This extension holds for a family of kernels satisfying easily checked hypotheses which we describe in this article, and includes many prominent RBFs. In the course of the proof, new convergence rates are obtained for the abstract operator considered by Devore and Ron in [Trans. Amer. Math. Soc. 362 (2010), pp. 6205–6229], and new Bernstein estimates are obtained relating high order smoothness norms to the native space norm.

## 1. Introduction

A hallmark of the mathematical theory of radial basis functions (RBFs) is the well-posedness of interpolation at scattered sites. In the simplest setting, a finite set $\Xi \subset \mathbb{R}^d$ generates a basic finite dimensional space $V_\Xi = \operatorname{span}\{\phi(\cdot - x) \mid x \in \Xi\}$, using the RBF $\phi : \mathbb{R}^d \to \mathbb{R}$, which is a continuous, positive definite, radially symmetric function. Interpolation at the sites $\Xi$ is well-posed; to any function $f : \mathbb{R}^d \to \mathbb{R}$, defined at sites $\Xi$, there is a unique continuous RBF interpolant $I_\Xi f \in V_\Xi$. The current state of the art treats error measured in a variety of Sobolev norms up to a critical order determined by the RBF. The goal of this paper is to provide a new error analysis for RBF interpolation treating errors measured in Sobolev norms higher than this critical order.

Received by the editor January 4, 2023, and, in revised form, September 12, 2023, and February 8, 2024.

2020 *Mathematics Subject Classification.* Primary 41A17, 41A25, 65D12.

Research of the first author was supported by grant DMS-2010051 from the National Science Foundation.

The motivation to extend the range of error estimates to higher order Sobolev norms stems from different mathematical areas. The first motivation stems from approximation theory. To determine the exact range of parameters in which rigorous error estimates can be shown is a natural question, one which has been considered for radial basis functions in [15, 19, 25, 27].

Measuring error in higher Sobolev norms has also gained attention in the context of deep learning. See, for example, [10, 18]. There are several reasons to include derivative information in the loss function for training deep neural networks. One motivation stems from the observation that including derivative information can improve the performance of the predictive error in learning, see [9]. Another stems from the fact that machine learning techniques have become an incredibly popular tool to solve partial differential equations – this includes using deep neural networks, but also Gaussian processes and kernels, see [7]. This aspect is closely connected to the next motivation.

A particularly strong motivation comes from using RBFs as tools for mesh-free solution of PDEs. In this regard, we mention the cosmos of (pseudo-)spectral methods, although other approaches, namely Galerkin and RBF-FD methods, can also benefit. Traditionally, in spectral methods one considers (orthogonal) polynomials for which such approximation results are also available, see for instance [5, Theorem 2.2] where also error estimates in higher (weighted) Sobolev norms are discussed. We consider now radial basis functions and pseudo-spectral methods, see [16] for an introduction into the topic. See also [35] for non-standard differential operators. The overall problem is that one seeks a finite dimensional approximation to a linear differential operator $\mathcal{L}$. A main focus in pseudo-spectral methods is that the approximate operator should be applicable to a function $f$ from which only its values on a discrete set of possibly scattered points $\Xi$ are known. A common approach is to consider a kernel-based interpolation $I_\Xi f$ to the function and to consider the differential operator applied to the interpolant as discretized differential operator. To formally justify this procedure a consistency argument of the following form

$$f \approx I_\Xi(f) \Rightarrow \mathcal{L}f \approx \mathcal{L}I_\Xi f$$

is needed. Such estimates can be rigorously proven if the interpolation error measured in high (depending on $\mathcal{L}$) Sobolev norms can be controlled.

1.1. **The doubling trick.** For each positive definite RBF $\phi$, there is an associated reproducing kernel Hilbert space (RKHS) $\mathcal{N}(\phi) \subset C(\mathbb{R}^d)$, the *native space*, for which $(x, y) \mapsto \phi(x - y)$ is the reproducing kernel. The native space has an associated error analysis for interpolation which works as follows: the interpolation operator $I_\Xi$ is the $\mathcal{N}(\phi)$-orthogonal projector onto $V_\Xi$. Thus $\|I_\Xi f - f\|_{\mathcal{N}(\phi)} \le \|f\|_{\mathcal{N}(\phi)}$, which leads, thanks to the embedding $\mathcal{N}(\phi) \subset C(\mathbb{R}^d)$, to (tautological) pointwise bounds of the form

$$|f(x) - I_\Xi f(x)| \le P_\Xi(x)\|f\|_{\mathcal{N}(\phi)},$$

where $P_\Xi(x) = \max_{\|f\|_{\mathcal{N}(\phi)}=1} |I_\Xi f(x) - f(x)|$ is called the *power function*. Often, the power function can be made small when $\Xi$ is well distributed near to $x$.

This natural error estimate can be improved by the *doubling trick* for RBFs, originally described in [29]. It is the RBF version of the classical Aubin-Nitsche trick [3, 28] used in the theory of Finite Elements. Roughly, it guarantees that a function $f \in \mathcal{N}(\phi * \phi) \cap \mathcal{N}(\phi)$ which has deconvolution $v = (\widehat{f}/\widehat{\phi})^\vee$ supported in a compact set $\Omega$ satisfies

$$\|f - I_\Xi f\|_{\mathcal{N}(\phi)} \le \|f\|_{\mathcal{N}(\phi*\phi)} \|P_\Xi\|_{L_2(\Omega)}.$$

A particularly strong version of this result considers $\phi$ with $\mathcal{N}(\phi)$ norm equivalent to $W_2^m(\mathbb{R}^d)$, $m > d/2$, and $\Omega \subset \mathbb{R}^d$ a compact set satisfying an interior cone condition. In that case, $\mathcal{N}(\phi * \phi) = W_2^{2m}(\mathbb{R}^d)$, and the result of applying the doubling trick gives

$$\|f - I_\Xi f\|_{W_2^\sigma(\Omega)} \le C h^{2m-\sigma} \|f\|_{W_2^{2m}(\mathbb{R}^d)} \quad \text{for all } 0 \le \sigma \le m. \tag{1.1}$$

Here $h := \max_{x\in\Omega} \operatorname{dist}(x, \Xi)$ is the fill distance of $\Xi$ in $\Omega$.

Interestingly, in the case that the native space is $W_2^m(\mathbb{R}^d)$, the RBF $\phi$, along with the finite dimensional space $V_\Xi$, lies in $W_p^\sigma(\mathbb{R}^d)$, for $\sigma < 2m - d + d/p$. Thus, it is reasonable to ask if $I_\Xi f$ converges to $f$ in higher order Sobolev norms when $f$ satisfies the conditions required for the doubling trick. This is the problem we seek to answer.

Precisely, we will show in Theorem 5.1 that for the interpolation error for $f = \phi * \nu \in W_2^{2m}(\mathbb{R}^d)$ with $\nu \in L_2(\mathbb{R}^d)$ having support in $\Omega$ and with a suitable positive definite kernel

$$\|f - I_\Xi f\|_{W_2^\sigma(\Omega)} \le C h^m q^{m-\sigma} \|\nu\|_{L_2(\mathbb{R}^d)} \tag{1.2}$$

holds for all sufficiently dense subsets and for any $m < \sigma$ for which $\lceil \sigma \rceil < 2m - d/2$. Here we employ the separation radius $q := \frac{1}{2} \min_{\xi\in\Xi} \operatorname{dist}(\xi, \Xi \setminus \{\xi\})$. Under conditions of quasi-uniformity of $\Xi$ (i.e., when $h \le \rho q$ for some constant $\rho > 1$) and using the original result [29], this yields

$$\|f - I_\Xi f\|_{W_2^\sigma(\Omega)} \le C h^{2m-\sigma} \|\nu\|_{L_2(\mathbb{R}^d)}$$

for all $\sigma \ge 0$ such that $\lceil \sigma \rceil < 2m - d/2$.

At this point, we note that we might have obtained such approximation orders for the case $\sigma > m$ by using a smoother kernel of order greater than $\sigma$ and employing the classical result (1.1). However, following the general guideline [31, Guideline 3.11] it is favorable to use the least smooth kernel to obtain a given approximation rate.

We illustrate this by considering the family of integer order Matérn kernels. For an integer $k \ge d/2$, the kernel $\tilde{\phi}_k$ is the fundamental solution to the elliptic differential operator $(1 - \Delta)^k$ on $\mathbb{R}^d$. These kernels are discussed in greater detail in Example 3.3. We note that each $\tilde{\phi}_k$ has native space $\mathcal{N}(\tilde{\phi}_k) = W_2^k(\mathbb{R}^d)$.

For a bounded set $\Omega \subset \mathbb{R}^d$ with smooth boundary, suppose $f \in W_2^{2m}(\mathbb{R}^d)$, but with support in $\Omega$. Given a finite set $\Xi \subset \Omega$, we would like to measure the interpolation error $\|f - I_\Xi f\|_{W_2^{m+n}(\Omega)}$ for $0 \le n < m - d/2$, where $I_\Xi f$ is the kernel interpolant using a suitable Matérn kernel $\tilde{\phi}_k$. For such kernels, the resulting linear system has a theoretically derived upper bound on the largest eigenvalue of order $q^{-d}$ and theoretically derived lower bound on the smallest eigenvalue of order $q^{2k-d}$. This results in an upper bound for the condition number of order $q^{-2k}$.

Standard error estimates [33, Corollary 11.33] permit us to interpolate using the kernel $\tilde{\phi}_{2m}$; this gives $\|f - I_\Xi f\|_{W_2^{m+n}(\Omega)} = \mathcal{O}(h^{m-n})$, but the resulting linear system has a theoretically derived condition number of order $q^{-4m}$.

An improvement in the condition number could be sought by using a kernel $\tilde{\phi}_{m+n}$. In this case, one might hope to employ the doubling trick as described above;

this can be done if $\mathrm{dist}(\mathrm{supp}(f), \partial\Omega) > 0$.[1] Using $\tilde{\phi}_{m+n}$ still leads to error $\|f - I_\Xi f\|_{W_2^{m+n}(\Omega)} = \mathcal{O}(h^{m-n})$, but now with an improved condition number of order $q^{-2m-2n}$ (along with the extra requirement that $\mathrm{supp}(f)$ is compactly contained in $\Omega$).

We point out that if $f \in W_2^{2m}(\mathbb{R}^d)$ with $\mathrm{supp} f \subset \Omega$, then the function $\nu := (1-\Delta)^m f \in L_2(\mathbb{R}^d)$ has support in $\Omega$. Thus $f = \tilde{\phi}_m * \nu$ and for point sets $\Xi$ which are quasi-uniform, the result Theorem 5.1 guarantees error $\|f - I_\Xi f\|_{W_2^{m+n}(\Omega)} = \mathcal{O}(h^{m-n})$, but with a condition number of order $q^{-2m}$ which is the least one of those choices.

1.2. **Outline.** We introduce notation and present some necessary background on RBF interpolation in section 2.

In section 3 we present high order Bernstein inequalities and discuss their application to three prominent families of RBFs: surface splines (introduced in Example 3.2), Matérn kernels (described in Example 3.3), and various compactly supported kernels including Wendland's kernels of minimal degree (described in Example 3.4).

In section 4 we introduce an integral-based approximation scheme $T_\Xi : f \mapsto T_\Xi f \in V_\Xi(\phi)$ and discuss its error analysis. The application of this approximation scheme is then discussed for surface splines, Matérn kernels and compactly supported kernels.

Section 5 gives interpolation error in the case that the RBF $\phi$ is positive definite; this setting yields the result (1.2) mentioned earlier. This applies to Matérn kernels and some compactly supported RBFs. A precise discussion of the compactly supported kernels for which this works is given in section 5.2.

Section 6 gives interpolation error in the case that the RBF $\phi$ is conditionally positive definite. This requires a bit more care than the positive definite case; in particular, the results require extra hypotheses which are not present in section 5. Section 6.2 shows how these hypotheses can be met for an RBF whose Fourier transform has an algebraic singularity. Section 6.3 treats the surface splines, and derives estimates in terms of the fill distance.

In Appendix A, we prove a lemma about regularity of local polynomial reproductions which is used in section 4, but which may find use beyond the scope of this article.

## 2. Notation and background

Throughout the article, we will use $C$ as a generic positive constant whose value may change from line to line.

[1] The doubling result cannot be applied directly, since $f \in W_2^{2m}(\mathbb{R}^d)$ is not assumed to be in $W_2^{2(m+n)}(\mathbb{R}^d)$. Still, if $\mathrm{dist}(\mathrm{supp}(f), \partial\Omega) > 0$ and $h$ is sufficiently small, one may find $F_h \in W_2^{2(m+n)}(\mathbb{R}^d)$ with $\mathrm{supp}(F_h) \subset \Omega$, such that $\|f - F_h\|_{W_2^{m+n}} \le Ch^{m-n}\|f\|_{W_2^{2m}(\mathbb{R}^d)}$ holds, along with $\|F_h\|_{W_2^{2(m+n)}(\mathbb{R}^d)} \le Ch^{-2n}\|f\|_{W_2^{2m}(\mathbb{R}^d)}$. (By mollification – see [12].) Then $\|f - I_\Xi f\|_{W_2^{m+n}(\mathbb{R}^d)} \le \|f - F_h\|_{W_2^{m+n}(\mathbb{R}^d)} + \|F_h - I_\Xi F_h\|_{W_2^{m+n}(\mathbb{R}^d)} + \|I_\Xi(F_h - f)\|_{W_2^{m+n}(\mathbb{R}^d)}$, with each term controlled by $Ch^{m-n}\|f\|_{W_2^{2m}(\mathbb{R}^d)}$; for the first term, this follows from the approximation error; for the second term it follows by applying (1.2) to obtain

$$\|F_h - I_\Xi F_h\|_{W_2^{m+n}(\mathbb{R}^d)} \le Ch^{m+n}\|F_h\|_{W_2^{2(m+n)}(\mathbb{R}^d)} \le Ch^{m-n}\|f\|_{W_2^{2m}(\mathbb{R}^d)};$$

for the third term, $\|I_\Xi(F_h - f)\|_{W_2^{m+n}(\mathbb{R}^d)} \le \|F_h - f\|_{W_2^{m+n}(\mathbb{R}^d)}$ follows because $I_\Xi$ is orthogonal projection.

Denote by $x \mapsto |x|$ the Euclidean norm in $\mathbb{R}^d$ and let $(x, y) \mapsto \mathrm{dist}(x, y) = |x - y|$ be the corresponding distance, which we extend to act on subsets of $\mathbb{R}^d$ in a natural way (in particular, $\mathrm{dist}(x, A) = \inf\{\mathrm{dist}(x, a) \mid a \in A\}$). Let $\Omega \subset \mathbb{R}^d$ and $\Xi \subset \Omega$ a finite subset. Define the *separation radius* by $q_\Xi := \frac{1}{2}\min_{\xi\in\Xi} \mathrm{dist}(\xi, \Xi \setminus \{\xi\})$, the *fill distance* by $h_{\Xi,\Omega} := \sup_{x\in\Omega} \mathrm{dist}(x, \Xi)$, and the *mesh ratio* by $\rho_{\Xi,\Omega} := \frac{h_{\Xi,\Omega}}{q_\Xi}$. When the underlying sets are clear from context, we will simply write $h$, $q$ and $\rho$.

Define the space of polynomials by $\mathcal{P}(\mathbb{R}^d)$ and the subspace of polynomials of degree $m$ or less by $\mathcal{P}_m(\mathbb{R}^d)$. The space of Schwartz functions is denoted $\mathcal{S}(\mathbb{R}^d)$. The Fourier transform of a Schwartz function is

$$\mathcal{F}\gamma(\omega) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \gamma(x) e^{-ix\cdot\omega} \mathrm{d}x,$$

and for tempered distributions, the Fourier transform $\mathcal{F}u$ is the distribution which satisfies $\langle \mathcal{F}u, \gamma\rangle := \langle u, \mathcal{F}\gamma\rangle$ for all $\gamma \in \mathcal{S}(\mathbb{R}^d)$. For $m \geq 0$, define

$$\mathcal{S}_m(\mathbb{R}^d) := \{\gamma \in \mathcal{S}(\mathbb{R}^d) \mid \sup_{\omega\in\mathbb{R}^d} |\omega|^{-m}\gamma(\omega) < \infty\}.$$

If the distributional Fourier transform of $\phi$ coincides on $\mathbb{R}^d \setminus \{0\}$ with a measurable function which represents the Fourier transform on $\mathcal{S}_{2m}(\mathbb{R}^d)$ then it has *generalized Fourier transform* of order $m$. Denoting the generalized Fourier transform of $\phi$ by $\widehat{\phi} : \mathbb{R}^d \setminus \{0\} \to \mathbb{C}$, the above definition is equivalent to the identity

$$(\forall \gamma \in \mathcal{S}_{2m}(\mathbb{R}^d)) \quad \int_{\mathbb{R}^d} \widehat{\phi}(\omega)\gamma(\omega)\mathrm{d}\omega = \int_{\mathbb{R}^d} \mathcal{F}\gamma(\omega)\phi(\omega)\mathrm{d}\omega.$$

See [17, 24] for more background.

**Sobolev spaces.** We recall (see for instance [1, Definition 2.39]) the Bessel potential operators which are defined for tempered distributions via the formula

$$\mathcal{F}(\mathcal{J}_s f) = (1 + |\cdot|^2)^{s/2}\mathcal{F}f.$$

For $\tau \geq 0$ we define the Sobolev space $H^\tau$ via Bessel potentials: that is, $H^\tau$ is the space of all $u \in L_2(\mathbb{R}^d)$ such that $\mathcal{J}_\tau u \in L_2(\mathbb{R}^d)$. Its norm is

$$\|u\|^2_{H^\tau} := \int_{\mathbb{R}^d} \mathcal{F}u(\omega)(1 + |\omega|^2)^{\tau/2}\mathrm{d}\omega.$$

For any $s, \tau \in \mathbb{R}$, $\mathcal{J}_s : H^\tau \to H^{\tau-s}$ is an isometry between Sobolev spaces.

The space $\dot{H}^\tau$ consists of distributions $u$ for which there exists $p \in \mathcal{P}$ such that $(-\Delta)^{\tau/2}u - p \in L_2(\mathbb{R}^d)$. If $u$ has generalized Fourier transform of some order, then

$$|u|_{\dot{H}^\tau_2} := \left\| |\cdot|^\tau \widehat{u} \right\|_{L_2(\mathbb{R}^d)}.$$

In order to work on compact sets $\Omega \subset \mathbb{R}^d$ we also need Sobolev spaces on domains. For $k \in \mathbb{N}$, we define the Sobolev space $W^k_2(\Omega)$ to be all functions $u$ with distributional derivatives $D^\alpha u \in L_2(\Omega)$ for all $|\alpha| \leq k$. Associated with these spaces are the seminorms

$$|u|^2_{W^k_2(\Omega)} := \sum_{|\alpha|=k} \frac{k!}{\alpha!}\|D^\alpha u\|^2_{L_2(\Omega)}$$

and norms $\|u\|^2_{W_2^k(\Omega)} := \sum_{j=0}^k |u|^2_{W_2^k(\Omega)}$. For fractional order Sobolev spaces, we use the seminorm

$$|u|^2_{W_2^{k+s}(\Omega)} = \sum_{|\alpha|=k} \int_\Omega \int_\Omega \frac{|D^\alpha u(x) - D^\alpha u(y)|}{|x-y|^{d+2s}} \mathrm{d}x\mathrm{d}y$$

and norm $\|u\|^2_{W_2^{k+s}(\Omega)} := \sum_{j=0}^k |u|^2_{W_2^k(\Omega)} + |u|^2_{W_2^{k+s}(\Omega)}$. It is well known that $W_2^\tau(\mathbb{R}^d)$ and $H^\tau$ have equivalent norms. To ensure equivalence between seminorms, we use Lemma 1.

**Lemma 1.** *If $u$ has generalized Fourier transform of order $\lfloor\tau\rfloor/2$, then the seminorms $|u|_{W_2^\tau(\mathbb{R}^d)}$ and $|u|_{\dot{H}^\tau}$ are equivalent.*

Before embarking on the proof, we make the following observation which allows us to relate generalized and distributional Fourier transforms: for $\kappa \in \mathcal{S}(\mathbb{R}^d)$ and multi-integer $\alpha$, $\mathcal{F}^{-1}(D^\alpha\kappa)(\omega) = (-i\omega)^\alpha \mathcal{F}^{-1}(\kappa)(\omega)$, so $\mathcal{F}^{-1}(D^\alpha\kappa) \in \mathcal{S}_{|\alpha|}(\mathbb{R}^d)$.

**Lemma 2.** *If $f$ has generalized Fourier transform of order $m$, and $|\alpha| \geq 2m$, then the distributional Fourier transform of $D^\alpha f$ is represented by the locally integrable function: namely, $\mathcal{F}(D^\alpha f)(\xi) = (i\omega)^\alpha \widehat{f}(\xi)$ almost everywhere.*

*Proof.* Let $\gamma \in \mathcal{S}(\mathbb{R}^d)$ and define $\psi$ via $\mathcal{F}\psi = D^\alpha \mathcal{F}\gamma$. By the above argument, $\psi = \mathcal{F}^{-1}(D^\alpha\mathcal{F}\gamma) \in \mathcal{S}_{|\alpha|}(\mathbb{R}^d)$ and $\psi(\omega) = (-i\omega)^\alpha\gamma(\omega)$ holds. Then

$$(D^\alpha f)(\mathcal{F}\gamma) = (-1)^{|\alpha|}\int_{\mathbb{R}^d} f(x) D^\alpha(\mathcal{F}\gamma)(x)\mathrm{d}x = (-1)^{|\alpha|}\int_{\mathbb{R}^d} f(x)\mathcal{F}\psi(x)\mathrm{d}x$$

holds by the definition of the distributional derivatives and by the definition of $\psi$. Thus $(D^\alpha f)(\mathcal{F}\gamma) = (-1)^{|\alpha|}\int_{\mathbb{R}^d}\widehat{f}(\xi)\psi(\xi)\mathrm{d}\xi$ follows by the definition of the generalized Fourier transform. Applying the formula for $\psi$ gives

$$(D^\alpha f)(\mathcal{F}\gamma) = (-1)^{|\alpha|}\int_{\mathbb{R}^d}\widehat{f}(\xi)(-i\omega)^\alpha\gamma(\omega)\mathrm{d}\xi = \int_{\mathbb{R}^d}\widehat{f}(\xi)(i\omega)^\alpha\gamma(\omega)\mathrm{d}\xi$$

and the result follows. □

*Proof of Lemma 1.* Suppose $u$ has generalized Fourier transform of order $\lfloor\tau\rfloor/2$.

In case $\tau = k \in \mathbb{N}$, it follows that $\int_{\mathbb{R}^d}|D^\alpha u(x)|^2\mathrm{d}x = \int_{\mathbb{R}^d}|\mathcal{F}(D^\alpha u)(\omega)|^2\mathrm{d}\omega = \int_{\mathbb{R}^d}|\omega^\alpha\widehat{u}(\omega)|^2\mathrm{d}\omega$ for any $|\alpha| = k$ (the first equality is Plancherel's theorem, and the second is Lemma 2). It follows that $|u|_{\dot{W}_2^k(\mathbb{R}^d)} \sim |u|_{\dot{H}^k}$.

In case $\tau \notin \mathbb{N}$ with $s = \tau - \lfloor\tau\rfloor$, we use [13, Proposition 3.4] to show that, for a multi-index $|\alpha| = \lfloor\tau\rfloor$, $|D^\alpha u|_{\dot{H}^s} \sim |D^\alpha u|_{W_2^s(\mathbb{R}^d)}$. Applying this to the definition gives $|u|^2_{W_2^\tau(\mathbb{R}^d)} \sim \sum_{|\alpha|=\lfloor\tau\rfloor}|D^\alpha u|^2_{\dot{H}^s}$. The equivalence $\sum_{|\alpha|=\lfloor\tau\rfloor}|D^\alpha u|^2_{\dot{H}^s} \sim |u|^2_{\dot{H}^\tau}$ follows because $\mathcal{F}(D^\alpha u)(\omega) = (i\omega)^\alpha\widehat{u}(\omega)$ for $\omega \neq 0$. □

**Radial basis functions and native spaces.** A function $\phi : \mathbb{R}^d \to \mathbb{R}$ is *conditionally positive definite of order $m_0$* (hereafter abbreviated by CPD) if the following holds: for any finite $\Xi \subset \mathbb{R}^d$, the *collocation matrix*

$$\Phi_\Xi = \big(\phi(\xi - \zeta)\big)_{\xi,\zeta\in\Xi}$$

is strictly positive definite on the subspace

$$\Big\{a : \Xi \to \mathbb{R} \;\;|\;\; (\forall p \in \mathcal{P}_{m_0-1})\sum_{\xi\in\Xi} a_\xi p(\xi) = 0\Big\}.$$

If $m_0 \le 0$, then $\phi$ is *positive definite.* A function which is CPD and symmetric with respect to rotations is called a *radial basis function* (RBF).

A version of Bochner's theorem [33, Theorem 8.12] asserts that if $\phi$ is continuous and increases at most algebraically (so that $|\phi(x)| \le C|x|^\alpha$ for some $\alpha \in \mathbb{R}$) and if $\phi$ has a continuous generalized Fourier transform of order $m_0$ which satisfies that $\widehat{\phi} > 0$ on some open set, then $\phi$ is CPD of order $m_0$. Although more will come later, this assumption will be in place throughout the article. It is worth noting that both the order of the generalized Fourier transform and the order of conditional positive definiteness have the nesting property: if $\phi$ has order $m_0$ then it has order $m_0 + 1$.

For a CPD function of order $m_0$ there is an associated function space, called the *native space*, $\mathcal{N}(\phi)$, which consists of continuous functions. One may find its construction in [30, 33]. The space has a semi-inner product $(f, g) \mapsto \langle f, g\rangle_{\mathcal{N}(\phi)}$ with nullspace $\mathcal{P}_{m_0-1}$. It is complete in the sense that the quotient $\mathcal{N}(\phi)/\mathcal{P}_{m_0-1}$ is a Hilbert space. We denote the induced seminorm by $f \mapsto |f|_{\mathcal{N}(\phi)}$.

If $\phi$ is positive definite (i.e., $m_0 \le 0$), the nullspace is trivial, and $\mathcal{N}(\phi)$ is a Hilbert space. In this case, $f \mapsto |f|_{\mathcal{N}(\phi)}$ is a norm.

It is worth noting that the native space depends both on the function $\phi$ and the order $m_0$; this is relevant because of the nesting property described above, so a given CPD function will generate infinitely many native spaces (one for each order).

For any functional of the form $\sum_{\xi\in\Xi} a_\xi \delta_\xi$ supported on $\Xi \subset \mathbb{R}^d$ which annihilates $\mathcal{P}_{m_0-1}$ we have for $f \in \mathcal{N}(\phi)$ that

$$\sum_{\xi\in\Xi} a_\xi f(\xi) = \left\langle f, \sum_{\xi\in\Xi} a_\xi \phi(\cdot - \xi)\right\rangle_{\mathcal{N}(\phi)}. \tag{2.1}$$

For $\Xi \subset \mathbb{R}^d$, we define the finite dimensional space

$$V_\Xi(\phi) := \Big\{\sum_{\xi\in\Xi} a_\xi \phi(\cdot - \xi) \mid \sum_{\xi\in\Xi} a_\xi \delta_\xi \perp \mathcal{P}_{m_0-1}\Big\} + \mathcal{P}_{m_0-1}.$$

For any $\Xi \subset \mathbb{R}^d$, we have $V_\Xi(\phi) \subset \mathcal{N}(\phi)$.

If $\Xi$ is *unisolvent* with respect to $\mathcal{P}_{m_0-1}$ (meaning that if $p \in \mathcal{P}_{m_0-1}$ vanishes on $\Xi$, then it is identically zero), then the interpolation operator

$$I_\Xi : \mathcal{N}(\phi) \to V_\Xi(\phi) \quad \text{where} \quad (I_\Xi f)|_\Xi = f|_\Xi \text{ for all } f \in \mathcal{N}(\phi)$$

is well-defined. It is the orthogonal projector onto $V_\Xi(\phi)$ with respect to the $\mathcal{N}(\phi)$ semi-inner product. Note that, like the native space, the interpolation operator depends on the order $m_0$ of conditional positive definiteness as well as on $\phi$ (as well as on $\Xi$).

For a CPD function $\phi$ which has a continuous generalized Fourier transform of order $m_0$, the native space can be expressed as the space of continuous functions $f$ which are tempered distributions, which have a generalized Fourier transform of order $m_0/2$, and for which $\int_{\mathbb{R}^d} |\widehat{f}(\omega)|^2/\widehat{\phi}(\omega)\mathrm{d}\omega < \infty$. In this case, the formula

$$\langle f, g\rangle_{\mathcal{N}(\phi)} = \int_{\mathbb{R}^d} \widehat{f}(\omega)\overline{\widehat{g}(\omega)}(\widehat{\phi}(\omega))^{-1}\mathrm{d}\omega \tag{2.2}$$

holds. See [30, 33] for details.

## 3. Higher order Bernstein inequalities

Bernstein estimates for RBF approximation have been demonstrated in [27], and more recently [20] for bounded regions. The existing literature treats the case that the weaker norm is $L_2(\mathbb{R}^d)$. In this section we present Bernstein inequalities where the weaker norm is the native space. These hold for RBFs which have the following property:

**Assumption 1.** We assume $\phi$ to be an RBF whose generalized Fourier transform satisfies the following two inequalities almost everywhere

$$\begin{aligned} C_1(1+|\omega|^2)^{-\tau} \le \widehat{\phi}(\omega) &\qquad \text{in } \mathbb{R}^d, \\ \widehat{\phi}(\omega) \le C_2|\omega|^{-2\tau} &\qquad \text{in } \mathbb{R}^d \setminus B(0,r_0) \end{aligned}$$

for some exponent $\tau > d/2$ and constants $0 < C_1 \le C_2$.

This guarantees the continuous embedding $H^\tau \subset \mathcal{N}(\phi)$. However, it does not quite imply $\mathcal{N}(\phi) \subset \dot{H}^\tau$, since $\widehat{\phi}$ may have a sharper singularity at $\omega = 0$ than $\mathcal{O}(|\omega|^{-2\tau})$. Of course, this assumption permits $\widehat{\phi}$ to have no singularity at all.

Under this assumption, [33, Theorem 12.3] applies (see also [26]), with a $\tau$ dependent constant $C_0$:

$$\lambda_{\min}(\Phi_\Xi) := \min_{\sum_{\xi\in\Xi}^N a_\xi\delta_\xi \perp \mathcal{P}_{m_0-1}} \sum_{\xi,\zeta\in\Xi} a_\xi a_\zeta \phi(\xi-\zeta) \ge C_0 C_1 q^{2\tau-d}\|a\|^2_{\ell_2(\Xi)}. \tag{3.1}$$

This can be used to prove a bandlimited approximation result as in [27, Lemma 3.3]. To this end, for $\sigma > 0$, define for a tempered distribution $u$, the function $u_\sigma := (\widehat{u}\chi_{B(0,\sigma)})^\vee$. By the identity (2.2), if $u \in \mathcal{N}(\phi)$, then $u_\sigma \in \mathcal{N}(\phi)$ as well.

**Lemma 3.** *If $\phi$ satisfies Assumption 1, then there is $\kappa > 0$ so that for any finite set of points $\Xi \subset \mathbb{R}^d$, if $\sigma > \max(r_0, \kappa/q)$, then $|u-u_\sigma|_{\mathcal{N}(\phi)} \le \frac{1}{2}|u|_{\mathcal{N}(\phi)}$ for all $u \in V_\Xi(\phi)$.*

*Proof.* The proof follows that of [27, Lemma 3.3], with a simple modification to treat the requirement that $\sigma > r_0$. □

As in [27, Theorem 5.1] this gives rise to a Bernstein estimate. In contrast to the result in [27], this uses a higher order smoothness norm on the right hand side.

**Theorem 3.1.** *Suppose $\phi$ satisfies Assumption 1, and that $0 \le s < \tau - d/2$. Then there is a constant $C$ so that for any $\Xi \subset \mathbb{R}^d$ with separation radius $q < 1$,*

$$|\mathcal{J}_s u|_{\mathcal{N}(\phi)} \le C q^{-s} |u|_{\mathcal{N}(\phi)}$$

*holds for any $u \in V_\Xi(\phi)$.*

Note that every polynomial space $\mathcal{P}_m$ is an invariant subspace of $\mathcal{J}_s$. This follows from $\langle \mathcal{J}_s p, \mathcal{F}\psi\rangle = \langle \mathcal{F}p, (1+|\cdot|^2)^{s/2}\psi\rangle$, which is a consequence of the definition of $\mathcal{J}_s$. In particular, one may and then using the product rule in conjunction with the characterization

$$\mathcal{F}(\mathcal{P}_m) = \{\mathcal{F}q \mid q \in \mathcal{P}_m\} = \operatorname{span}\{\delta_0 D^\alpha \mid |\alpha| \le m\}$$

to obtain $\mathcal{J}_s\mathcal{P}_m = \mathcal{P}_m$. Since $\mathcal{P}_{m_0-1}$ is the nullspace of the native space seminorm, we have $|\mathcal{J}_s(u+p)|_{\mathcal{N}(\phi)} = |\mathcal{J}_s u|_{\mathcal{N}(\phi)}$. Thus for $\phi$ satisfying Assumption 1, $|\mathcal{J}_s u|_{\mathcal{N}(\phi)}$ can be calculated, via (2.2), as an integral on the Fourier domain.

*Proof.* The function $\psi_{\tau-s} := \mathcal{J}_{2s}\phi$ satisfies Assumption 1, with $\tau - s$ in place of $\tau$, as can be observed from its generalized Fourier transform $\widehat{\psi_{\tau-s}}(\omega) = (1+|\omega|^2)^s\widehat{\phi}(\omega)$. An application of (2.2) gives $|\mathcal{J}_s u|^2_{\mathcal{N}(\phi)} = \int_{\mathbb{R}^d} |\sum_{\xi\in\Xi} c_\xi e^{i\langle\omega,\xi\rangle}|^2(1+|\omega|^2)^s\widehat{\phi}(\omega)d\omega$ for any $u = \sum_{\xi\in\Xi} c_\xi\phi(\cdot-\xi) \in V_\Xi(\phi)$, which provides the identity

$$|\mathcal{J}_s u|^2_{\mathcal{N}(\phi)} = \left|\sum_{\xi\in\Xi} c_\xi\psi_{\tau-s}(\cdot-\xi)\right|^2_{\mathcal{N}(\psi_{\tau-s})} = |\tilde{u}|^2_{\mathcal{N}(\psi_{\tau-s})},$$

where we define $\tilde{u} := \sum_{\xi\in\Xi} c_\xi\psi_{\tau-s}(\cdot-\xi)$. Let $\sigma = 2\max(\kappa/q, r_0)$. Then Lemma 3 guarantees that $|\tilde{u}|^2_{\mathcal{N}(\psi_{\tau-s})} \le 4|(\tilde{u})_\sigma|^2_{\mathcal{N}(\psi_{\tau-s})}$. Finally, we have

$$\begin{aligned} |\,(\tilde{u})_\sigma\,|^2_{\mathcal{N}(\psi_{\tau-s})} &= \int_{|\omega|<\sigma}\left|\sum_{\xi\in\Xi} c_\xi e^{i\langle\omega,\xi\rangle}\right|^2\widehat{\phi}(\omega)(1+|\omega|^2)^s\,\mathrm{d}\omega \\ &\le \left(\frac{1+4\kappa^2}{4\kappa^2}\right)^s\sigma^{2s}\int_{|\omega|<\sigma}\left|\sum_{\xi\in\Xi} c_\xi e^{i\langle\omega,\xi\rangle}\right|^2\widehat{\phi}(\omega)\mathrm{d}\omega \le Cq^{-2s}\,|u|^2_{\mathcal{N}(\phi)}\,, \end{aligned}$$

since $q<1$. In the first inequality, we have used the fact that $\sigma$ is bounded below by $\sigma \ge 2\kappa > 0$, so $1+\sigma^2 \le \left(\frac{1+4\kappa^2}{4\kappa^2}\right)\sigma^2$. The second inequality follows automatically if $\sigma = 2\kappa/q$; if $r_0 > \kappa/q$, then the result holds with a slightly larger, $r_0$ dependent constant, because $\sigma^{2s} \le (2r_0)^{2s} \le (2r_0)^{2s}q^{-2s}$. ☐

This applies to a number of prominent RBF families.

**Example 3.2.** The surface spline $\phi_m$, the fundamental solution to $\Delta^m$ on $\mathbb{R}^d$ for $m > d/2$, is CPD of order $\lfloor m-d/2\rfloor+1$, and has generalized Fourier transform $\widehat{\phi_m}(\omega) = |\omega|^{-2m}$ of order $\lfloor m-d/2\rfloor+1$. It is, however, customary to consider $\phi_m$ as CPD of order $m_0 = m$, in which case [33, Theorem 10.43] shows that the native spaces is the Beppo-Levi space

$$\mathrm{BL}_m(\mathbb{R}^d) = \{f \in L_{2,loc}(\mathbb{R}^d) \mid (\forall|\alpha| = m)\ D^\alpha f \in L_2(\mathbb{R}^d)\}.$$

The seminorm for this space is $|f|_{\mathcal{N}(\phi_m)} = |f|_{W_2^m(\mathbb{R}^d)}$. Then Theorem 3.1 states that for $u \in V_\Xi(\phi_m)$ and $s < m-d/2$, we have $|\mathcal{J}_s u|_{\dot{H}^m} \le Cq^{-s}|u|_{W_2^m(\mathbb{R}^d)}$.

**Example 3.3.** The Matérn kernels $\tilde{\phi}_\tau$, $\tau > d/2$, known also as Bessel potential kernels, are the fundamental solutions to the (possibly) fractional operator $(1-\Delta)^\tau$ on $\mathbb{R}^d$. They are strictly positive definite, with native space $\mathcal{N}(\tilde{\phi}_\tau) = H^\tau$. For any $u \in V_\Xi(\tilde{\phi}_\tau)$ we have $\|u\|_{H^{\tau+s}} \le Cq^{-s}\|u\|_{H^\tau}$ as long as $s < \tau - d/2$. These RBFs are discussed further in Example 4.2.

**Example 3.4.** Various compactly supported RBFs, including Wendland's compactly supported RBFs of minimal degree, denoted $\phi_{k,d}$ (where $k$ is a parameter derived from its construction, but related to its smoothness) satisfy Assumption 1. Each kernel $\phi_{k,d}$ is strictly positive definite, and has native space $\mathcal{N}(\phi_{k,d}) = H^{k+(d+1)/2}(\mathbb{R}^d)$. Theorem 3.1 states that for any $u \in V_\Xi(\phi_{k,d})$ we have

$$\|u\|_{H^{k+(d+1)/2+s}(\mathbb{R}^d)} \le Cq^{-s}\|u\|_{H^{k+(d+1)/2}(\mathbb{R}^d)}$$

as long as $s < k+1/2-d/2$. These are discussed again in Example 4.3 in the next section.

## 4. RBF approximation with Sobolev norms

We now give Jackson estimates for the spaces $V_\Xi(\phi)$ using the norm $\|\cdot\|_{H^{m+n}(\mathbb{R}^d)}$, with $0 \le n < m - d/2$. Our first result involves a version of the approximation scheme developed in [11] for RBFs which are fundamental solutions to differential operators. This scheme was used to get approximation results with error measured in $L_p(\mathbb{R}^d)$; we expand this slightly to error in Sobolev norms, and for RBFs satisfying a more general set of conditions. Specifically, we show that it provides strong results for target functions $u \in H^{2m}$ having deconvolution $(\widehat{u}/\widehat{\phi})^\vee$ supported in $\Omega$.

For this, we make a basic assumption about the radially symmetric function $\phi$. Namely, that it is a smooth perturbation of a type of (essentially) homogeneous function. To make this definition we introduce the function $\mathsf{h}_s$ for $s \ge 0$ as

$$\mathsf{h}_s(x) = \begin{cases} |x|^s & s \notin 2\mathbb{N}, \\ |x|^s \log|x| & s \in 2\mathbb{N}. \end{cases}$$

By [22, (3.1)], it follows that

$$D^\alpha \mathsf{h}_s(x) = p_{s-|\alpha|}(x)\log(x) + q_{s-|\alpha|}(x) \tag{4.1}$$

with $q_{s-|\alpha|}$ a homogeneous, rational function of degree $s-|\alpha|$, and $p_{s-|\alpha|}$ a homogeneous polynomial of degree $s-|\alpha|$, which is zero when $s \notin 2\mathbb{Z}$ or when $s-|\alpha| < 0$.

**Assumption 2.** Suppose $s$ and $L$ are positive, with $s > d/2$ and $L > s+d$. We assume $\phi \in C(\mathbb{R}^d)\cap C^{s+d-1}(\mathbb{R}^d\setminus\{0\})$ is radially symmetric, and there is a constant $r_0 > 0$ so that $\phi|_{\mathbb{R}^d\setminus B(0,r_0)} \in C^L(\mathbb{R}^d \setminus B(0,r_0))$ and the following two conditions hold

(1) there is a constant $C$ so that for all multi-indices $|\beta| = L$ and $|x| > r_0$,

$$|D^\beta \phi(x)| \le C|x|^{s-|\beta|},$$

(2) there exist functions $u, v \in C^L(\overline{B(0,r_0)})$ so that for $|x| < r_0$

$$\phi(x) = u(x) + \mathsf{h}_s(x)v(x).$$

Although neither Assumption 1 nor 2 implies the other, if $\phi$ is to satisfy both simultaneously, with $v(0) \ne 0$ it must follow that $s = 2\tau - d$. To see this, decompose $\phi = \phi_1+\phi_2+\phi_3$ into three radially symmetric components such that $\phi_1$ is supported near the origin, $\phi_2$ is supported in a neighborhood of $\{x \mid |x| = r_0\}$ and $\phi_3$ is supported away from $\{x \mid |x| \le r_0\}$. Assumption 2 guarantees that $\widehat{\phi_2}(\xi)$ and $\widehat{\phi_3}(\xi)$ are both $o(|\xi|^{-(s+d)})$, while $\widehat{\phi_1}(\xi) \sim |\xi|^{-(s+d)}$ due to fact that it behaves locally like $\mathsf{h}_s$ near the origin.

By item (1) of Assumption 2, $D^\beta\phi_3 \in L_1(\mathbb{R}^d)$ for all multi-indices $|\beta| = L$, so $|\widehat{\phi_3}(\xi)| \le C|\xi|^{-L}$. Because $\phi_2 \in C^{s+d-1}(\mathbb{R}^d)$ and it has derivatives $D^\beta\phi_2(x)$ of order $|\beta| = s+d$ which extend continuously to the boundary $\{x \mid |x| = r_0\}$, its distributional derivatives of order $|\beta| = s+d$ are in $L_1(\mathbb{R}^d)$, so $|\widehat{\phi_2}(\xi)| = o(|\xi|^{-(s+d)})$ as $\xi \to \infty$.

Finally, by expanding $v$ from (2) in a Taylor series about the origin, we have $v(x) = \sum_{j=0}^L c_j|x|^j + R(x)$ (the terms corresponding to odd values of $j$ vanish, but for the present argument, there is no advantage in omitting these terms). Near to the origin, $\phi_1(x) = u(x) + \sum_{j=0}^L c_j \mathsf{h}_{s+j}(x) + \mathsf{h}_s(x)R(x)$ which implies that

$C_1|\xi|^{-s-d} \le \widehat{\phi_1}(\xi) \le C_2|\xi|^{-s-d}$ for $0 < C_1 \le C_2 < \infty$, since $\mathsf{h}_s$ has distributional Fourier transform $\widehat{\mathsf{h}}_s(\xi) \propto |\xi|^{-s-d}$ on $\mathbb{R}^d \setminus \{0\}$, and the other components of $\phi_1$ have Fourier transform which decays more rapidly.

**Example 4.1.** The family of surface splines given in Example 3.2 is defined for $m > d/2$ as $\phi_m(x) = C_{m,d}\mathsf{h}_{2m-d}(x)$. Thus, they satisfy Assumption 2 with $s = 2m - d$. Item (1) follows from (4.1) and the remark following it, while item (2) follows with $u = 0$ and constant $v$.

**Example 4.2.** The Matérn kernels $\tilde{\phi}_\tau(x) = |x|^{\tau-d/2}K_{\tau-d/2}(|x|)$ satisfy Assumption 2 with $s = 2\tau - d$. Here $K_\mu$ is a modified Bessel function [14, 10.25]. Each $\tilde{\phi}_\tau$ is in $C^\infty(\mathbb{R}^d \setminus \{0\})$ and satisfies the decay condition $|D^\alpha \tilde{\phi}_\tau(x)| \le C_M|x|^{-M}$ for all $M$ and all $\alpha$. Furthermore, item (2) holds by using the convergent power series expansion

$$\tilde{\phi}_\tau(x) = \sum_{j=0}^{\infty} a_j|x|^{2j} + \mathsf{h}_{2\tau-d}(|x|)\sum_{j=0}^{\infty} b_j|x|^{2j}$$

which is valid for all $\mu > d/2$. When $\mu - d/2 \in \mathbb{N}$, this is given in both [14, Eq. 10.31.1] and [2, Eq. 9.6.11]. When $\mu - d/2$ is fractional, it follows from either [14, Eq. 10.27.4 and Eq. 10.25.2] or [2, Eq. 9.6.2 and Eq. 9.6.10].

**Example 4.3.** The compactly supported Wendland kernels of minimal degree $\phi_{k,d}$, described in [33, Chapter 9], satisfy Assumption 2 only in dimension $d = 2$. Indeed, for $d \in \mathbb{N}$, $\phi_{k,d} \in C^{2k+\lfloor d/2\rfloor+1}(\mathbb{R}^d \setminus \{0\})$, so when $d = 2$, $\phi_{k,2} \in C^{2k+2}(\mathbb{R}^2 \setminus \{0\})$. In this case, $s = 2k + 1$ and $r_0 = 1$. Item (1) holds because $\mathrm{supp}(\phi_{k,d}) = B(0, 1)$.

The fact that item (2) holds follows from [33, Theorem 9.12]. Specifically, $\phi_{k,2}(x) = p(|x|)$ for a polynomial $p(r) = \sum_{j=0}^{3k+2} a_j r^j$ whose first $k$ odd coefficients are zero. I.e., $a_{2k+1}$ is the first non-zero coefficient of an odd power.

There do exist a number of compactly supported RBFs which satisfy Assumptions 1 and 2, however (the *generalized Wendland functions* studied in [8]). These are discussed in section 5.2.

Our interest is to approximate functions $f$ having the form $f = \phi * \nu + p$, for $\nu \in L_2(\mathbb{R}^d)$, with $\mathrm{supp}(\nu)$ contained in a compact set $\Omega$, and $p \in \mathcal{P}_{m_0-1}$. We note that $\phi$ is sufficiently smooth to allow differentiation under the integral sign:

$$D^\alpha \int_\Omega \nu(z)\phi(x - z)\mathrm{d}z = \int_\Omega \nu(z)D^\alpha\phi(x - z)\mathrm{d}z$$

whenever $|\alpha| < s + d$, by compactness of $\Omega$, integrability of $\nu$, and continuity of $D^\alpha\phi$.

4.1. **Approximation scheme.** We consider an approximation scheme similar to the one presented in [11]. For this, we consider a compact set $\Omega \subset \mathbb{R}^d$, a finite subset $\Xi \subset \Omega$, and a sufficiently regular *local polynomial reproduction*. The latter is a map $a(\cdot,\cdot) : \Xi \times \Omega \to \mathbb{R}$ which satisfies the following conditions:

- for every $z \in \Omega$ if $\mathrm{dist}(\xi, z) > Kh$ then $a(\xi, z) = 0$
- for every $z \in \Omega$, $\sum_{\xi\in\Xi}|a(\xi, z)| \le \Gamma$
- for every $p \in \mathcal{P}_L$ and $z \in \Omega$, $\sum_{\xi\in\Xi} a(\xi, z)p(\xi) = p(z)$
- for every $\xi \in \Xi$, $a(\xi, \cdot)$ is measurable.

If $\Omega$ satisfies an interior cone condition and $h$ is sufficiently small, then [33, Theorem 3.14] guarantees existence of a local polynomial reproduction which has the first three of these four properties. In Appendix A, we present the modification to [33, Theorem 3.14] which is needed to get the fourth condition (actually, we show that each $a(\xi,\cdot)$ can be chosen to be infinitely smooth).

For any function $f$ which can be decomposed as $f = \phi * \nu + p$, with $\nu \in L_2(\mathbb{R}^d)$ having support in $\Omega$, and $p \in \mathcal{P}$, we define the approximation scheme $T_\Xi$ as

$$T_\Xi f(x) := \sum_{\xi\in\Xi}\left(\int_\Omega a(\xi,z)\nu(z)\mathrm{d}z\right)\phi(x-\xi) + p(x).$$

*Remark* 4.4. If $L \ge m$ and $\nu \perp \mathcal{P}_m$, then the coefficients $A_\xi = \int_\Omega a(\xi,z)\nu(z)\mathrm{d}z$ satisfy

$$\sum_{\xi\in\Xi} A_\xi p(\xi) = \int_\Omega \nu(z)\sum_{\xi\in\Xi} a(\xi,z)p(\xi)\mathrm{d}z = \int_\Omega \nu(z)p(z)\mathrm{d}z = 0$$

for any $p \in \mathcal{P}_m$. In particular, if $\phi$ is CPD of order $m_0$ and $f = \nu * \phi + p$ with $\nu \perp \mathcal{P}_{m_0-1}$ and $p \in \mathcal{P}_{m_0-1}$, then we have $T_\Xi f \in V_\Xi(\phi)$.

4.2. **Approximation error.** In order to calculate the error $\|D^\alpha f - D^\alpha T_\Xi f\|_{L_2(\mathbb{R}^d)}$, we introduce, for each multi-index $\alpha$ with $|\alpha| < s+d$, the error kernel $E^{(\alpha)} : \mathbb{R}^d \times \Omega \to \mathbb{R}$, where

$$E^{(\alpha)}(x,z) := \left|D^\alpha\phi(x-z) - \sum_{\xi\in\Xi} a(\xi,z)D^\alpha\phi(x-\xi)\right|.$$

To analyze the error kernel, we make use of polynomial reproduction in the following way:

**Lemma 4.** *Let $L$ be a non-negative integer. Suppose that $w \subset \mathbb{R}^d$, $W$ is a neighborhood of $w$, $\tilde{X} \subset W$ is a finite set, and $\tilde{a} \in \mathbb{R}^{\tilde{X}}$ satisfies $\sum_{\zeta\in\tilde{X}} \tilde{a}_\zeta p(\zeta) = p(w)$ for all $p \in \mathcal{P}_L$, along with $|\zeta - w| > r \Rightarrow \tilde{a}_\zeta = 0$. For any positive integer $M$, with $M \le L+1$, if $U$ is $M$-times continuously differentiable in a neighborhood of $\overline{B(w,r)}$, then we have*

$$\left|U(w) - \sum_{\zeta\in\tilde{X}} \tilde{a}_\zeta U(\zeta)\right| \le \frac{\|\tilde{a}\|_{\ell_1(\tilde{X})}}{M!} r^M \max_{|\beta|=M} \|D^\beta U\|_{L_\infty(B(w,r))}. \tag{4.2}$$

*Proof.* We can express $U(\zeta) = P(\zeta) + R(\zeta)$, with $P$ the Taylor polynomial of degree $M-1$ centered at $w$. Thus, $P = \sum_{|\beta|<M} \frac{1}{\beta!} D^\beta U(w)(\cdot - w)^\beta$. For $\zeta \in B(w,r)$, the remainder satisfies

$$|R(\zeta)| \le \frac{1}{M!}|\zeta - w|^M \max_{|\beta|=M} \|D^\beta U\|_{L_\infty(B(w,r))}.$$

Then $|U(w) - \sum_{\zeta\in\tilde{X}} \tilde{a}_\zeta U(\zeta)| \le \|\tilde{a}\|_{\ell_1} \max_{|\zeta-w|\le r} |R(\zeta)|$, and the result follows. □

**Lemma 5.** *Suppose $\phi$ satisfies Assumption 2. Then the error kernel satisfies, for $|x-z| > 2Kh$, the estimate*

$$E^{(\alpha)}(x,z) \le \begin{cases} Ch^{s-|\alpha|}\left(\frac{|x-z|}{h}\right)^{s-L} & |x-z| \notin [r_0 - Kh, r_0 + Kh], \\ Ch^{s+d-1-|\alpha|} & r_0 - Kh \le |x-z| \le r_0 + Kh. \end{cases}$$

*Proof.* We split this into three cases according to the size of $|x-z|$. Case 1 treats the punctured space $|x-z| > r_0+Kh$, Case 2 treats the annulus $r_0-Kh \le |x-z| \le r_0+Kh$, and Case 3 treats the inner annulus $2Kh \le |x-z| < r_0+Kh$.

In each case, we use Lemma 4 applied to $U = D^\alpha\phi(x-\cdot)$ at the point $w=z$ in $W=\Omega$ using the point set $\tilde{X} = \Xi$ and the vector $\tilde{a} \in \mathbb{R}^{\tilde{X}}$ defined by $\tilde{a}_\zeta = a(\zeta, z)$. By local polynomial reproduction, the hypotheses of Lemma 4 hold with $r = Kh$ and $\|\tilde{a}\|_{\ell_1(\tilde{X})} = \Gamma$.

The only difference between the cases lies in the smoothness $M$ enjoyed by $U$.

*Case* 1 (Assume $|x-z| > r_0+Kh$). In this case, $U = D^\alpha\phi(x-\cdot)$ is $M = L-|\alpha|$ times continuously differentiable on $\mathbb{R}^d \setminus \overline{B(x,r_0)}$. Under these conditions, we have $E^{(\alpha)}(x,z) = |U(w) - \sum_{\zeta\in\tilde{X}} \tilde{a}_\zeta U(\zeta)|$, so by (4.2), it follows that

$$E^{(\alpha)}(x,z) \le Ch^M \max_{|\beta|=M} \|D^\beta U\|_{L_\infty(B(z,Kh))} \le Ch^{L-|\alpha|} \max_{|\gamma|=L} \|D^\gamma\phi\|_{L_\infty(B(x-z,Kh))}.$$

We note that $\kappa := \min\{|\eta| \mid \eta \in B(x-z,Kh)\} \ge r_0$, so the first item of Assumption 2 applies, giving $\max_{|\gamma|=L} \|D^\gamma\phi\|_{L_\infty(B(x-z,Kh))} \le C\kappa^{s-L}$. Since $\kappa \ge \frac{1}{2}|x-z|$, we have $\max_{|\gamma|=L} \|D^\gamma\phi\|_{L_\infty(B(x-z,Kh))} \le C|x-z|^{s-L}$ which implies that

$$E^{(\alpha)}(x,z) \le Ch^{L-|\alpha|}|x-z|^{s-L} = Ch^{s-|\alpha|}\left(\frac{|x-z|}{h}\right)^{s-L}.$$

*Case* 2 (Assume that $r_0-Kh \le |x-z| \le r_0+Kh$). In this case, Assumption 2 guarantees continuity of $D^\beta\phi$ on $\mathbb{R}^d\setminus\{0\}$ for $|\beta| \le s+d-1$. Thus $U = D^\alpha\phi(x-\cdot)$ has uniformly bounded derivatives of order $M = s+d-1-|\alpha|$. In this case, Lemma 4 guarantees that

$$\begin{aligned} E^{(\alpha)}(x,z) &\le Ch^{s+d-1-|\alpha|} \max_{|\beta|=s+d-1-|\alpha|} \|D^\beta U\|_{L_\infty(B(z,Kh))} \\ &\le Ch^{s+d-1-|\alpha|} \|\phi\|_{C^{s+d-1}\left(\overline{B(0,2r_0)\setminus B(0,r_0/2)}\right)}. \end{aligned}$$

*Case* 3 (Assume that $2Kh < |x-z| < r_0-Kh$). Recall that item (2) of Assumption 2 states that $\phi(x) = u(x) + \mathsf{h}_s(x)v(x)$ in this region. To treat this case, we consider the $u$ and $\mathsf{h}_s v$ components separately.

By Assumption 2, we have $U = D^\alpha u(x-\cdot)$ has smoothness $M = L-|\alpha|$ over the set $B(x,r_0)$, which contains $B(x,r_0-Kh)\setminus\overline{B(x,2Kh)}$. Thus Lemma 4 guarantees that $|D^\alpha u(x-z) - \sum_{\xi\in\Xi} a(\xi,z)D^\alpha u(x-\xi)| \le Ch^{L-|\alpha|}$. Since $|x-z| < r_0$, we have

$$\left|D^\alpha u(x-z) - \sum_{\xi\in\Xi} a(\xi,z)D^\alpha u(x-\xi)\right| \le Cr_0^{L-s}h^{s-|\alpha|}\left(\frac{|x-z|}{h}\right)^{s-L}. \tag{4.3}$$

Similarly, letting $U = D^\alpha(\mathsf{h}_s v)(x-\cdot)$, Lemma 4 gives

$$\begin{aligned} &\left|D^\alpha(\mathsf{h}_s v)(x-z) - \sum_{\xi\in\Xi} a(\xi,z)D^\alpha(\mathsf{h}_s v)(x-\xi)\right| \\ &\le Ch^{L-|\alpha|} \max_{|\beta|=L} \|D^\beta(\mathsf{h}_s v)\|_{L_\infty(B(x-z,Kh))}, \end{aligned}$$

because $\max_{|\beta|=L-|\alpha|} \|D^\beta U\|_{L_\infty(B(z,Kh))} \le \max_{|\beta|=L} \|D^\beta(\mathsf{h}_s v)\|_{L_\infty(B(x-z,Kh))}$. We can estimate $\|D^\beta(\mathsf{h}_s v)\|_{L_\infty(B(x-z,Kh))}$ by using the inequality

$$\|D^\beta(\mathsf{h}_s v)\|_{L_\infty(B(x-z,Kh))} \le C_\beta \sum_{\gamma\le\beta} \|D^\gamma \mathsf{h}_s\|_{L_\infty(B(x-z,Kh))} \|D^{\beta-\gamma}v\|_{L_\infty(B(x-z,Kh))},$$

which follows with a $\beta$ dependent constant from the Leibniz rule. By (4.1), there is a constant so that for any $|\gamma| \le L$, $|D^\gamma \mathsf{h}_s(x)| \le C|x|^{s-L}$ on $B(0,r_0)$. Since $|x-z| > 2Kh$, it follows that $\inf\{|\zeta| \mid \zeta \in B(x-z, Kh)\} > \frac{1}{2}|x-z|$, so

$$\Big|D^\alpha(\mathsf{h}_s v)(x-z) - \sum_{\xi\in\Xi} a(\xi,z) D^\alpha(\mathsf{h}_s v)(x-\xi)\Big| \le C h^{L-|\alpha|}|x-z|^{s-L}. \tag{4.4}$$

The result in Case 3 follows by combining (4.3) and (4.4). □

**Lemma 6.** *Suppose $\phi$ satisfies Assumption 2. Then for $0<|\alpha|<s+d/2$, and $0<|x-z|<2Kh$, the error kernel satisfies*

$$E^{(\alpha)}(x,z) \le C\Big(h^{L-|\alpha|} + |x-z|^{s-|\alpha|} + \sum_{\xi\in\Xi}|a(\xi,z)||x-\xi|^{s-|\alpha|}\Big).$$

*Proof.* Assumption 2 allows us to split $E^{(\alpha)}(x,z)$ into a totally smooth part and a homogenous part $E^{(\alpha)}(x,z) \le E_S + E_H$ where

$$E_S := \Big|D^\alpha u(x-z) - \sum_{\xi\in\Xi} a(\xi,z) D^\alpha u(x-\xi)\Big|,$$
$$E_H := \Big|D^\alpha(\mathsf{h}_s v)(x-z) - \sum_{\xi\in\Xi} a(\xi,z) D^\alpha(\mathsf{h}_s v)(x-\xi)\Big|.$$

The smooth part is treated as in the proof of Lemma 5. In particular, Lemma 4 ensures that

$$\Big|D^\alpha u(x-z) - \sum_{\xi\in\Xi} a(\xi,z) D^\alpha u(x-\xi)\Big| \le C h^{L-|\alpha|}. \tag{4.5}$$

To treat $E_H$, we use the Leibniz rule and smoothness of $v$, to obtain

$$\begin{aligned} E_H &\le C\sum_{\gamma\le\alpha}\big(|D^\gamma \mathsf{h}_s(x-z)| + \sum_{\xi\in\Xi}|a(\xi,z)||D^\gamma \mathsf{h}_s(x-\xi)|\big) \\ &\le C\Big(|x-z|^{s-|\alpha|} + \sum_{\xi\in\Xi}|a(\xi,z)||x-\xi|^{s-|\alpha|}\Big), \end{aligned} \tag{4.6}$$

where the second estimate follows from (4.1). Combining (4.5) and (4.6) gives the result. □

**Theorem 4.5.** *Suppose $\phi$ satisfies Assumption 2, $f = \nu * \phi + p$, with $p \in \mathcal{P}$, and $\nu \in L_2(\mathbb{R}^d)$ having support in a bounded, open set $\Omega$ having Lipschitz boundary. Then for $\sigma$ with $\lceil\sigma\rceil < s+d/2$, the approximation error satisfies*

$$\|f - T_\Xi f\|_{W_2^\sigma(\mathbb{R}^d)} \le C h^{s+d-\sigma}\|\nu\|_{L_2(\mathbb{R}^d)}.$$

*Proof.* We begin by considering an integer $\sigma < s+d/2$. Let $\alpha$ be a multi-index with $|\alpha| = \sigma$. Then we have $\|D^\alpha f - D^\alpha T_\Xi f\|_{L_2(\mathbb{R}^d)} = \Big(\int_{\mathbb{R}^d}\big|\int_{\mathbb{R}^d} E^{(\alpha)}(x,z)\nu(z)\mathrm{d}z\big|^2 \mathrm{d}x\Big)^{1/2}$ by differentiating under the integral. Defining quantities $A$ and $B$ as

$$A := \Big\|\int_{|\cdot - z|>2Kh} E^{(\alpha)}(\cdot,z)\nu(z)\mathrm{d}z\Big\|_{L_2(\mathbb{R}^d)},$$
$$B := \Big\|\int_{|\cdot - z|<2Kh} E^{(\alpha)}(\cdot,z)\nu(z)\mathrm{d}z\Big\|_{L_2(\mathbb{R}^d)},$$

we split the error into two parts: $\|D^\alpha f - D^\alpha T_\Xi f\|_{L_2(\mathbb{R}^d)} \le A+B$. This corresponds to splitting the error kernel as $E^{(\alpha)} = E_1 + E_2$, where

$$E_1(x,z) := E^{(\alpha)}(x,z)\chi_{\{(x,z)||x-z|>2Kh\}}(x,z),$$
$$E_2(x,z) := E^{(\alpha)}(x,z)\chi_{\{(x,z)||x-z|<2Kh\}}(x,z).$$

We may control $E_1$ by Lemma 5 and $E_2$ by Lemma 6.

By integrating $E_1(x,z)$ with respect to either $x$ or $z$, we obtain an estimate for the $L_p$ norm of the integral operator $\mathcal{E}_1 : g \mapsto \int_{\mathbb{R}^d} g(z)E_1(x,z)\mathrm{d}z$. In particular, for $1 \le p \le \infty$,

$$\begin{aligned}\|\mathcal{E}_1\|_{L_p(\mathbb{R}^d)\to L_p(\mathbb{R}^d)} \le Ch^{s-|\alpha|} &\int_{2Kh<|y|<r_0-Kh} (|y|/h)^{s-L}\,\mathrm{d}y \\ &+ Ch^{s+d-1-|\alpha|}\mathrm{vol}(\{y \mid r_0 - Kh < |y| < r_0 + Kh\}) \\ &+ Ch^{s-|\alpha|}\int_{r_0+Kh<|y|<\infty} (|y|/h)^{s-L}\,\mathrm{d}y.\end{aligned}$$

To treat the first term on the right hand side, we use by homogeneity of the integrand and a change to polar coordinates to get the estimate

$$\begin{aligned}h^{s-|\alpha|}\int_{2Kh<|y|<r_0-Kh} (|y|/h)^{s-L}\,\mathrm{d}y &\le Ch^{L-|\alpha|}\int_{2Kh}^{r_0} r^{s+d-1+L}\mathrm{d}r \\ &\le Ch^{L-|\alpha|}(2Kh)^{s+d-L} = Ch^{s+d-|\alpha|}.\end{aligned}$$

The second term involves the volume of an annulus having thickness proportional to $h$; so $\mathrm{vol}(\{y \mid r_0 - Kh < |y| < r_0 + Kh\}) \le Ch$. The third term can again be estimated by using homogeneity of the integrand followed by a change to polar coordinates;

$$\begin{aligned}h^{s-|\alpha|}\int_{r_0+Kh<|y|<\infty} (|y|/h)^{s-L}\,\mathrm{d}y &\le h^{L-|\alpha|}\int_{r_0}^{\infty} r^{d+1+s-L}\mathrm{d}r \\ &\le Ch^{L-|\alpha|} \le Ch^{s+d-|\alpha|}\end{aligned}$$

(because $L > s+d$ by assumption). So $\|\mathcal{E}_1\|_{L_p(\mathbb{R}^d)\to L_p(\mathbb{R}^d)} \le Ch^{s+d-|\alpha|}$. In particular, this holds for $p=2$, which gives

$$A \le Ch^{s+d-|\alpha|}\|\nu\|_{L_2(\mathbb{R}^d)}. \tag{4.7}$$

By Lemma 6, $E_2(x,z) \le C(h^{L-|\alpha|} + |x-z|^{s-|\alpha|} + \sum_{\xi\in\Xi}|a(\xi,z)||x-\xi|^{s-|\alpha|})$ for $x,z$ satisfying $|x-z| < 2Kh$. This allows us to estimate $B$ with three integrals, each generated by one of the above terms. Defining $B_1$, $B_2$ and $B_3$ as

$$B_1 := \left\|\int_{|\cdot-z|<2Kh} h^{L-|\alpha|}|\nu(z)|\mathrm{d}z\right\|_{L_2(\mathbb{R}^d)},$$
$$B_2 := \left\|\int_{|\cdot-z|<2Kh} |\cdot-z|^{s-|\alpha|}|\nu(z)|\mathrm{d}z\right\|_{L_2(\mathbb{R}^d)},$$
$$B_3 := \left\|\int_{|\cdot-z|<2Kh} \sum_{\xi\in\Xi}|a(\xi,z)|\,|\cdot-\xi|^{s-|\alpha|}|\nu(z)|\mathrm{d}z\right\|_{L_2(\mathbb{R}^d)}.$$

By Hölder's inequality, we then have

$$B \le C(B_1+B_2+B_3). \tag{4.8}$$

The first two parts can be controlled by the method used for $E_1$, giving

$$B_1 \le Ch^{L+d-|\alpha|}\|\nu\|_{L_2(\mathbb{R}^d)}, \tag{4.9}$$

$$B_2 \le Ch^{s+d-|\alpha|}\|\nu\|_{L_2(\mathbb{R}^d)}, \tag{4.10}$$

since $s-|\alpha| > -d/2 > -d$.

To handle $B_3$, we apply Hölder's inequality to the sum $\sum_{\xi\in\Xi}|a(\xi,z)|\,|\cdot-\xi|^{s-|\alpha|}$, writing $|a(\xi,z)| = \sqrt{|a(\xi,z)|}\sqrt{|a(\xi,z)|}$, and $\|a(\cdot,z)\|_{\ell_1(\Xi)} = \sum_{\xi\in\Xi}|a(\xi,z)|$ to obtain

$$(B_3)^2 \le \int_{\mathbb{R}^d}\Big|\int_{B(x,2Kh)}\|a(\cdot,z)\|_{\ell_1(\Xi)}^{1/2}\Big(\sum_{\xi\in\Xi}|a(\xi,z)|\,|x-\xi|^{2(s-|\alpha|)}\,|\nu(z)|^2\Big)^{1/2}\mathrm{d}z\Big|^2\mathrm{d}x.$$

Applying Hölder's inequality to the inner integral gives

$$(B_3)^2 \le \int_{\mathbb{R}^d}\Big(\int_{B(x,2Kh)}\|a(\cdot,\zeta)\|_{\ell_1(\Xi)}\mathrm{d}\zeta\Big)\times \int_{B(x,2Kh)}\sum_{\xi\in\Xi}|a(\xi,z)|\,|x-\xi|^{2(s-|\alpha|)}\,\big|\nu(z)\big|^2\mathrm{d}z\mathrm{d}x.$$

By the estimate $\int_{B(x,2Kh)}\|a(\cdot,\zeta)\|_{\ell_1(\Xi)}\mathrm{d}\zeta \le \int_{B(x,2Kh)}\Gamma\mathrm{d}\zeta \le Ch^d$, we have

$$(B_3)^2 \le Ch^d\int_{\mathbb{R}^d}\int_{B(x,2Kh)}\sum_{\xi\in\Xi}|a(\xi,z)|\,|x-\xi|^{2(s-|\alpha|)}\,|\nu(z)|^2\mathrm{d}z\mathrm{d}x.$$

Because $a(\xi,z)=0$ when $|z-\xi|>Kh$ and $z\in B(x,2Kh)$, the inner sum is taken only over $\xi\in\Xi$ which are within $3Kh$ from $x$. We use this to switch the order of sums and integrals:

$$\begin{aligned}(B_3)^2 &\le Ch^d\int_{\mathbb{R}^d}\sum_{|\xi-x|<3Kh}|x-\xi|^{2(s-|\alpha|)}\left(\int_{B(x,2Kh)}|a(\xi,z)|\big|\nu(z)\big|^2\mathrm{d}z\right)\mathrm{d}x\\ &\le Ch^d\sum_{\xi\in\Xi}\left(\int_{B(\xi,3Kh)}|x-\xi|^{2(s-|\alpha|)}\,\mathrm{d}x\right)\left(\int_{\mathbb{R}^d}|a(\xi,z)|\big|\nu(z)\big|^2\mathrm{d}z\right).\end{aligned}$$

The last integral can be made larger by increasing the domain of integration to $\mathbb{R}^d$. At this point, we observe that $\int_{B(\xi,2Kh)}|x-\xi|^{2(s-|\alpha|)}\,\mathrm{d}x \le Ch^{2s-2|\alpha|+d}$. This leaves

$$\begin{aligned}(B_3)^2 &\le Ch^{2s-2|\alpha|+2d}\int_{\mathbb{R}^d}\|a(\cdot,z)\|_{\ell_1(\Xi)}|\nu(z)|^2\mathrm{d}z\\ &\le C\Gamma h^{2s+2d-2|\alpha|}\|\nu\|^2_{L_2(\mathbb{R}^d)}.\end{aligned}\tag{4.11}$$

The bound $B \le Ch^{s+d-|\alpha|}\|\nu\|_{L_2(\mathbb{R}^d)}$ follows from the decomposition (4.8) and estimates (4.9), (4.10) and (4.11). Combining this fact with (4.7) completes the proof in case $\sigma\in\mathbb{N}$.

For fractional $\sigma$ with $\lceil\sigma\rceil < s+d$, we simply interpolate between integer order Sobolev spaces, using $\sigma_1 = \lfloor\sigma\rfloor$ and $\sigma_2 = \lceil\sigma\rceil$, so that $\sigma = \theta\sigma_2 + (1-\theta)\sigma_1$. This can be done by using Hölder's inequality to estimate the Fourier characterization of the $H^\sigma$ norm, or to by way of the Gagliardo-Nirenberg inequality.

In either case, we have the estimate $\|F\|_{W_2^\sigma(\mathbb{R}^d)} \le C\|F\|^{1-\theta}_{W_2^{\sigma_1}(\mathbb{R}^d)}\|F\|^{\theta}_{W_2^{\sigma_2}(\mathbb{R}^d)}$, which ensures

$$\|f-T_\Xi f\|_{W_2^\sigma} \le \big(Ch^{s+d-\sigma_1}\|\nu\|_{L_2(\mathbb{R}^d)}\big)^{1-\theta}\big(Ch^{s+d-\sigma_2}\|\nu\|_{L_2(\mathbb{R}^d)}\big)^{\theta}.$$

The result follows because $h^{(s+d-\sigma_1)(1-\theta)}h^{(s+d-\sigma_2)\theta} = h^{s+d-\sigma}$. □

## 5. Interpolation with positive definite RBFs

With the aid of the Bernstein estimates from section 3, we show that the approximation rate of Theorem 4.5 is inherited by RBF interpolation: for this, we consider an RBF $\phi$ having a native space which is norm equivalent to $H^\tau$, and a target function for which the doubling result of [29] applies.

We measure the interpolation error $\|f - I_\Xi f\|_{H^\sigma(\mathbb{R}^d)}$ for suitable values of $\sigma > 0$.

### 5.1. Main result for positive definite RBFs.

**Theorem 5.1.** *Suppose $\tau > d/2$ and $\phi$ is a positive definite RBF with native space equivalent to the Sobolev space $H^\tau(\mathbb{R}^d)$. Suppose, further, that $\phi$ satisfies Assumption 2 with $s = 2\tau - d$. If $\Omega \subset \mathbb{R}^d$ is compact and satisfies an interior cone condition, then there is a constant $C$ so that the following holds. For any $f \in H^{2\tau}(\mathbb{R}^d)$ which satisfies $f = \phi * \nu$ with $\nu \in L_2(\mathbb{R}^d)$ supported in $\Omega$, for any sufficiently dense subset $\Xi \subset \Omega$ and for $\sigma > 0$ satisfying $\lceil\sigma\rceil < 2\tau - d/2$, the inequality*

$$\|f - I_\Xi f\|_{H^\sigma(\mathbb{R}^d)} \le Ch^\tau q^{\tau-\sigma}\|\nu\|_{L_2(\mathbb{R}^d)}$$

*holds.*

Before proving Theorem 5.1, some remarks are in order.

Using the mesh-ratio $\rho = h/q$ gives $\|f - I_\Xi f\|_{W_2^\sigma(\Omega)} \le C\rho^{\tau-\sigma}h^{2\tau-\sigma}\|\nu\|_{L_2(\mathbb{R}^d)}$. If $\Xi$ is such that $q$ and $h$ are kept roughly on par, i.e., if $\Xi$ is quasi-uniform with controlled mesh ratio, then this extends previous doubling results in, which held for $\sigma \le \tau$.

In other words, the novelty of Theorem 5.1 is that it holds in case $\tau < \sigma$ and $\lceil\sigma\rceil < 2\tau - d/2$.

Because $\mathcal{N}(\phi)$ is norm equivalent to $H^\tau(\mathbb{R}^d)$, it follows that there exist constants $0 < C_1 \le C_2$ so that for all $f \in H^\tau(\mathbb{R}^d)$,

$$C_1 \int_{\mathbb{R}^d} |\widehat{f}(\omega)|^2/\widehat{\phi}(\omega)\mathrm{d}\omega \le \int_{\mathbb{R}^d} |\widehat{f}(\omega)|^2(1+|\omega|^2)^\tau \mathrm{d}\omega \le C_1 \int_{\mathbb{R}^d} |\widehat{f}(\omega)|^2/\widehat{\phi}(\omega)\mathrm{d}\omega.$$

By employing an approximate identity, this shows that $\widehat{\phi}(\omega) \sim (1+|\omega|^2)^{-\tau}$, so Assumption 1 holds automatically.

*Proof.* By the above comment, we consider $\sigma$ which satisfies $\sigma > \tau$ and $\lceil\sigma\rceil < 2\tau - d/2$. By hypothesis, $\phi$ satisfies Assumption 1. Thus Theorem 3.1 applies to $I_\Xi f - T_\Xi f \in V_\Xi(\phi)$, and $\|\mathcal{J}_{\sigma-\tau}(I_\Xi f - T_\Xi f)\|_{\mathcal{N}(\phi)} \le Cq^{\tau-\sigma}\|I_\Xi f - T_\Xi f\|_{\mathcal{N}(\phi)}$ holds, which implies

$$\|I_\Xi f - T_\Xi f\|_{H^\sigma} \le Cq^{\tau-\sigma}\|I_\Xi f - T_\Xi f\|_{\mathcal{N}(\phi)}.$$

Theorem 4.5 gives $\|f - T_\Xi f\|_{\mathcal{N}(\phi)} \le C\|f - T_\Xi f\|_{H^\tau} \le Ch^\tau\|\nu\|_{L_2(\mathbb{R}^d)}$, while the standard doubling argument given in the proof of [29, Theorem 5.1] shows that

$$\|f - I_\Xi f\|^2_{\mathcal{N}(\phi)} \le \|f - I_\Xi f\|_{L_2(\Omega)}\|\nu\|_{L_2(\mathbb{R}^d)}.$$

Since the inequality $\|f - I_\Xi f\|_{L_2(\Omega)} \le Ch^\tau\|f - I_\Xi f\|_{W_2^\tau(\mathbb{R}^d)}$ holds by standard arguments (see [33, Theorem 11.32], or the original version [25, Theorem 2.12]), we have, by norm equivalence of the spaces $\mathcal{N}(\phi) \sim H^\tau \sim W_2^\tau(\mathbb{R}^d)$, that

$$\|f - I_\Xi f\|_{\mathcal{N}(\phi)} \le Ch^\tau\|\nu\|_{L_2(\mathbb{R}^d)}.$$

Thus, the triangle inequality gives $\|I_\Xi f - T_\Xi f\|_{\mathcal{N}(\phi)} \le Ch^\tau \|\nu\|_{L_2(\mathbb{R}^d)}$, and

$$\|I_\Xi f - T_\Xi f\|_{H^\sigma} \le Cq^{\tau-\sigma} h^\tau \|\nu\|_{L_2(\mathbb{R}^d)} \tag{5.1}$$

follows. On the other hand, a direct application of Theorem 4.5 gives

$$\|f - T_\Xi f\|_{H^\sigma} \le Ch^{2\tau-\sigma} \|\nu\|_{L_2(\mathbb{R}^d)}. \tag{5.2}$$

Together, (5.2) and (5.1) give

$$\|f - I_\Xi f\|_{H^\sigma} \le \|f - T_\Xi f\|_{H^\sigma} + \|T_\Xi f - I_\Xi f\|_{H^\sigma} \le Ch^\tau q^{\tau-\sigma}\|\nu\|_{L_2(\mathbb{R}^d)}$$

and the result follows. □

5.2. **A note on compactly supported RBFs.** As pointed out in Example 4.3, the compactly supported RBFs of minimal degree constructed in [33, Chapter 9] do not satisfy Assumption 2, unless $d = 2$. This is precisely because of the behavior at the boundary of the support of $\phi_{k,d}$. This can be addressed by following the same construction, but using a radial polynomial of slightly higher degree. The requirements of Assumption 2 may also be satisfied by other compactly supported RBFs, of which there are many, one may find other constructions in [4, 6, 34]. Furthermore, it may be possible to prove Theorem 5.1 for the classical Wendland functions with $d > 2$ by improving the error analysis of the $T_\Xi$ scheme (in such a way that Assumption 2 is weakened), or by using a different approach altogether.

We recall here some aspects of Wendland's construction which can be used to construct compactly supported RBFs that satisfy Assumption 2.

For a measurable function, $f : (0,\infty) \to \mathbb{R}$ which is integrable with respect to $\mathrm{d}\mu = s\mathrm{d}s$, we define $\mathcal{I}f(r) := \int_r^\infty sf(s)\mathrm{d}s$. The operator $\mathcal{I}$ has an intertwining property with the Fourier transform: the $d+2$-dimensional Fourier transform of a suitably integrable radial function $f$ equals, as a radial function, the $d$-dimensional Fourier transform of $\mathcal{I}f$, see e.g. [32, Lemma 2.1]: i.e.,

$$\begin{aligned} &r^{-d/2}\int_0^\infty f(t)t^{(d+2)/2}J_{d/2}(rt)\mathrm{d}t \\ &= r^{-(d-2)/2}\int_0^\infty \mathcal{I}f(t)t^{d/2}J_{(d-2)/2}(rt)\mathrm{d}t \quad \text{for all } r>0. \end{aligned}$$

Define $\psi_\ell(0,\infty) \to \mathbb{R}$ by $\psi_\ell := (1-\cdot)_+^\ell$. Then for spatial dimension $d$, and integer $\ell \ge \lfloor d/2 \rfloor + 1$, the function $x \mapsto \psi_\ell(|x|) = (1-|x|)_+^\ell$ is radial, positive definite and supported in $B(0,1)$. Via Bochner's theorem and the above intertwining formula, $x \mapsto (\mathcal{I}^k\psi_\ell)(|x|)$ is positive definite as well, see also [32, Eq. (5)].

The RBFs of minimal degree described in Example 4.3 are defined as $\phi_{k,d} := \mathcal{I}^k\psi_\ell$, with $\ell = k + \lfloor d/2 \rfloor + 1$. For general $k, \ell$, and $f : [0,1] \to \mathbb{R}$, a simple induction gives the identity $\mathcal{I}^k f(r) = \frac{2^{1-k}}{\Gamma(\alpha)}\int_r^1 tf(t)(t^2-r^2)^{k-1}\mathrm{d}t$ for $r \le 1$. In particular, the family of functions

$$\mathcal{I}^k\psi_\ell(r) := \frac{2^{1-k}}{\Gamma(\alpha)}\int_r^1 t(1-t)^\ell(t^2-r^2)^{k-1}\mathrm{d}t$$

can be extended to non-integer values of $k$ and $\ell$; such "generalized Wendland functions" have been introduced and studied in [8].

By collecting known results from [33] and [8], Proposition 1 shows that, for $\ell \ge k + d$, each kernel $\mathcal{I}^k\psi_\ell$ satisfies the hypotheses of Theorem 5.1. In particular, we have the compatibility between Sobolev order $m$ and homogeneity parameter $s$ from Assumption 2: namely $s = 2m - d$ since both quantities equal $2k+1$.

**Proposition 1.** *For integers $k, \ell$ satisfying $\ell \geq k+d$, the function*

$$\psi_{\ell,k} : \mathbb{R}^d \to \mathbb{R} : x \mapsto \mathcal{I}^k \psi_\ell(|x|)$$

*is a compactly supported RBF which satisfies Assumption 2 with $s = 2k+1$. Its native space, $\mathcal{N}(\psi_{\ell,k})$, is norm equivalent to $W_2^m(\mathbb{R}^d)$ with $m = k + \frac{d+1}{2}$.*

*Proof.* Smoothing properties of the operator $\mathcal{I}$ given in [33, Lemma 9.8] guarantee that $\mathcal{I}^k\psi_\ell \in C^{\ell+k}\big((0,\infty)\big)$. Since $s = 2k+1$ and $\ell \geq k+d$, it follows that $\ell + k \geq s + d - 1$, so $\mathcal{I}^k\psi_\ell \in C^{s+d-1}\big((0,\infty)\big)$, as required.

Because $\phi$ has support in $B(0,1)$, item (1) holds with $r_0 = 1$.

Since each application of $\mathcal{I}$ increases the polynomial degree by 2, $\mathcal{I}^k\psi_\ell$ is polynomial of degree $2k+\ell$, and, as observed in [33, Theorem 9.12], the first $k$ odd-degree coefficients in the monomial expansion of $\mathcal{I}^k\psi_\ell$ vanish. This also follows directly from the formula in [8, Theorem 3.2]. By splitting into even and odd degree powers, we obtain

$$\mathcal{I}^k\psi_\ell(|x|) = \sum_{j=0}^{2k+\ell} d_j|x|^j = \left(\sum_{j=0}^{k+\lfloor \ell/2 \rfloor} d_{2j}|x|^{2j}\right) + |x|^{2k+1}\left(\sum_{j=0}^{\lfloor (\ell-1)/2 \rfloor} d_{2k+1+2j}|x|^{2j}\right)$$

so item (2) holds with $s = 2k+1$.

The fact that $\mathcal{N}(\psi_{\ell,k}) = W_2^{k+\frac{d+1}{2}}(\mathbb{R}^d)$ has been observed in [8, Corollary 2.4]. Specifically, the $d$-dimensional Fourier transform of $\psi_{\ell,k}$ is shown to satisfy $\widehat{\psi_{\ell,k}}(\xi) \sim (1+|\xi|)^{-(d+2k+1)}$ in [8, Eqn. (2.3)]. □

## 6. Interpolation using conditionally positive definite RBFs

The CPD case requires an extra assumption and has a slightly different error estimate. For various reasons, the target function $f = \nu * \phi + p \in \mathcal{N}(\phi)$ must satisfy the polynomial annihilation condition $\nu \perp \mathcal{P}_{m_0-1}$, which is equivalent to the vanishing moment condition $\widehat{\nu}(\xi) = \mathcal{O}(|\xi|^{m_0})$. Furthermore, the error estimate is initially in terms of the quantity $\mathcal{E}_{\Omega,\Xi}(f)$, which can be refined in a few ways (this is discussed after the proof).

Section 6.1 provides the analogous result to Theorem 5.1 for CPD kernels. Sections 6.2 and 6.3 give instances where the annihilation condition is guaranteed to hold and provide bounds for the quantity $\mathcal{E}_{\Omega,\Xi}(f)$ in terms of the fill distance.

### 6.1. Main result for conditionally positive definite RBFs.

**Theorem 6.1.** *Suppose $\phi$ is an RBF which is CPD of order $m_0$ and which satisfies Assumptions 1 and 2, with $s = 2\tau - d$. If $\Omega \subset \mathbb{R}^d$ is compact and satisfies an interior cone condition, then there is a constant $C$ so that the following holds. If $\Xi \subset \Omega$ is a sufficiently dense set, and if $f \in \mathcal{N}(\phi)$ satisfies the decomposition $f = \phi * \nu + p$ where $p \in \mathcal{P}_{m_0-1}$, and $\nu \in L_2(\mathbb{R}^d)$ is a function having two properties,* supp$(\nu) \subset \Omega$ *and $\nu \perp \mathcal{P}_{m_0-1}$, then*

$$|\mathcal{J}_{\sigma-\tau}(f - I_\Xi f)|_{\mathcal{N}(\phi)} \leq C q^{\tau-\sigma}(h^\tau + \mathcal{E}_{\Omega,\Xi}(f))\|\nu\|_{L_2(\mathbb{R}^d)}$$

*with $\lceil \sigma \rceil < 2\tau - d/2$. Here $\mathcal{E}_{\Omega,\Xi}(f) := \frac{\|f - I_\Xi f\|_{L_2(\Omega)}}{|f - I_\Xi f|_{\mathcal{N}(\phi)}}$.*

*Proof.* The estimates

$$|\mathcal{J}_{\sigma-\tau}(f - T_\Xi f)|_{\mathcal{N}(\phi)} \leq C h^{2\tau-\sigma}\|\nu\|_{L_2(\mathbb{R}^d)} \quad \text{and} \quad |f - T_\Xi f|_{\mathcal{N}(\phi)} \leq C h^\tau \|\nu\|_{L_2(\mathbb{R}^d)}$$

follow from Theorem 4.5 and the embedding $H^\tau = W_2^\tau(\mathbb{R}^d) \subset \mathcal{N}(\phi)$ which implies the estimate $|\mathcal{J}_{\sigma-\tau}(f - T_\Xi f)|_{\mathcal{N}(\phi)} \lesssim \|\mathcal{J}_{\sigma-\tau}(f - T_\Xi f)\|_{H^\tau} \lesssim \|f - T_\Xi f\|_{W_2^\sigma(\mathbb{R}^d)}$. We can treat the interpolation error in the native space by using a similar 'doubling' argument to that of Theorem 5.1. Orthogonality gives

$$|f - I_\Xi f|^2_{\mathcal{N}(\phi)} = \langle f, f - I_\Xi f\rangle_{\mathcal{N}(\phi)} = \int_{\mathbb{R}^d} \widehat{\nu}(\omega)\big(\widehat{f}(\omega) - \widehat{I_\Xi f}(\omega)\big)\mathrm{d}\omega.$$

Some care is necessary to apply a Plancherel-like result, since $\widehat{f}(\omega) - \widehat{I_\Xi f}(\omega)$ is only a generalized Fourier transform (and also not necessarily in $L_2(\mathbb{R}^d)$). The identity

$$\int_{\mathbb{R}^d} \widehat{\nu}(\omega)\big(\widehat{f}(\omega) - \widehat{I_\Xi f}(\omega)\big)\mathrm{d}\omega = \int_\Omega \nu(x)\big(f(x) - I_\Xi f(x)\big)\mathrm{d}x$$

is handled in Lemma 7. Applying Cauchy-Schwarz gives

$$|f - I_\Xi f|^2_{\mathcal{N}(\phi)} \le \|\nu\|_{L_2(\Omega)}\|f - I_\Xi f\|_{L_2(\Omega)} \le \|\nu\|_{L_2(\Omega)}|f - I_\Xi f|_{\mathcal{N}(\phi)}\mathcal{E}_{\Omega,\Xi}(f).$$

Dividing gives $|f - I_\Xi f|_{\mathcal{N}(\phi)} \le \mathcal{E}_{\Omega,\Xi}(f)\|\nu\|_{L_2(\mathbb{R}^d)}$ and applying the triangle inequality gives $|(T_\Xi f - I_\Xi f)|_{\mathcal{N}(\phi)} \le (Ch^\tau + \mathcal{E}_{\Omega,\Xi}(f))\|\nu\|_{L_2(\mathbb{R}^d)}$. Because $\nu \perp \mathcal{P}_{m_0-1}$, it follows from Remark 4.4 that $T_\Xi f \in V_\Xi(\phi)$. Since $T_\Xi f - I_\Xi f \in V_\Xi(\phi)$, we may apply Theorem 3.1 to obtain

$$|\mathcal{J}_{\sigma-\tau}(T_\Xi f - I_\Xi f)|_{\mathcal{N}(\phi)} \le Cq^{\tau-\sigma}|(T_\Xi f - I_\Xi f)|_{\mathcal{N}(\phi)} \le Cq^{\tau-\sigma}(h^\tau + \mathcal{E}_{\Omega,\Xi})\|\nu\|_{L_2(\mathbb{R}^d)}$$

and the result follows. □

Under some extra conditions on the RBF, $\mathcal{E}_{\Omega,\Xi}(f)$ can be controlled by the fill distance, yielding a result similar to the positive definite case. This is discussed below. However, even without extra hypotheses, the term $\mathcal{E}_{\Omega,\Xi}(f)$ can be estimated by the power function $P_\Xi(x) = \sup_{|f|_{\mathcal{N}(\phi)}=1}|f(x) - I_\Xi f(x)|$, which can be estimated by [33, Theorem 11.9].

**Corollary 1.** *Suppose $\phi$ satisfies the requirements of Theorem 6.1. Then*

$$|\mathcal{J}_{\sigma-\tau}(f - I_\Xi f)|_{\mathcal{N}(\phi)} \le Cq^{\tau-\sigma}h^{\tau-d/2}\|\nu\|_{L_2(\mathbb{R}^d)}.$$

*Proof.* Because $I_\Xi$ is idempotent, we have

$$\begin{aligned}\mathcal{E}_{\Omega,\Xi}(f) &= \frac{\|f - I_\Xi f - I_\Xi(f - I_\Xi f)\|_{L_2(\Omega)}}{|f - I_\Xi f|_{\mathcal{N}(\phi)}}\\ &\le \frac{\mathrm{vol}(\Omega))^{1/2}\max_{x\in\Omega}|f(x) - I_\Xi f(x) - I_\Xi(f(x) - I_\Xi f(x))|}{|f - I_\Xi f|_{\mathcal{N}(\phi)}}\\ &\le (\mathrm{vol}(\Omega))^{1/2}\max_{x\in\Omega}P_\Xi(x).\end{aligned}$$

By [33, Theorem 11.9], for $\ell > m_0 - 1$, there exist positive constants $c_1, c_2$ so that

$$P_\Xi(x) \le C\min_{p\in\mathcal{P}_\ell}\|\phi - p\|^{1/2}_{L_\infty(B(0,c_2h))}.$$

Since $\phi$ satisfies Assumption 2 (in particular item (2)) it follows that, if we take $\ell > 2\tau - d$, we have $\min_{p\in\mathcal{P}_\ell}\|\phi - p\|_{L_\infty(B(0,c_2h))} \le Ch^{s/2} = Ch^{2\tau-d}$. The result follows by plugging the estimate $\mathcal{E}_{\Omega,\Xi}(f) \le Ch^{\tau-d/2}$ into Theorem 6.1. □

We now make the additional assumption necessary to refine $\mathcal{E}_{\Omega,\Xi}(f)$ by using the zeros lemma. We assume that $\widehat{\phi}(\xi) \le C|\xi|^{-2\tau}$ in a neighborhood of the origin, which without loss is $B(0, r_0)\setminus\{0\}$, where $r_0$ is the constant from Assumption 1 (by continuity of $\widehat{\phi}$). Together with Assumption 1, this is equivalent to assuming the continuous embedding $\mathcal{N}(\phi) \subset \dot{H}^\tau$. We also assume $m_0 \le \tau$, which permits us to compare the $W_2^\tau(\mathbb{R}^d)$ seminorm appearing in the zeros estimate and the homogeneous seminorm of $\dot{H}^\tau$.

**Corollary 2.** *Suppose $\phi$ satisfies the requirements of Theorem 6.1 with $m_0 \le \tau$ and that $\widehat{\phi}(\xi) \le C|\xi|^{-2\tau}$ for $0 < |\xi| < r_0$. Then*

$$|\mathcal{J}_{\sigma-\tau}(f - I_\Xi f)|_{\mathcal{N}(\phi)} \le C q^{\tau-\sigma} h^\tau \|\nu\|_{L_2(\mathbb{R}^d)}.$$

*Proof.* Because $m_0 \in \mathbb{N}$, we have $m_0 \le \lfloor \tau \rfloor$. By the zeros estimate [21, Theorem A.4], we have $\|f - I_\Xi f\|_{L_2(\Omega)} \le C h^\tau |f - I_\Xi f|_{W_2^\tau(\mathbb{R}^d)}$. Because $f - I_\Xi f \in \mathcal{N}(\phi)$, [33, Theorem 10.21] ensures it has generalized Fourier transform of order $m_0/2 \le \lfloor \tau \rfloor/2$ (note that $m_0$ is an integer). Lemma 1 applies and guarantees that the $W_2^\tau(\mathbb{R}^d)$ and $\dot{H}^\tau$ seminorms are identical. Consequently, we have $\|f - I_\Xi f\|_{L_2(\Omega)} \le C h^\tau |f - I_\Xi f|_{\dot{H}^\tau}$. Because $\widehat{\phi}(\xi) \le C|\xi|^{-2\tau}$ for all $\xi \ne 0$, we have

$$\|f - I_\Xi f\|_{L_2(\Omega)} \le C h^\tau |f - I_\Xi f|_{\mathcal{N}(\phi)}.$$

It follows that $\mathcal{E}_{\Omega,\Xi}(f) \le C h^\tau$, and the result follows. □

**Lemma 7.** *Suppose $\phi$, $\Xi$ and $f = \nu * \phi + p$ satisfy the hypotheses of Theorem 6.1. Then*

$$\int_{\mathbb{R}^d} \widehat{\nu}(\omega)\big(\widehat{f}(\omega) - \widehat{I_\Xi f}(\omega)\big)\mathrm{d}\omega = \int_{\mathbb{R}^d} \nu(x)\big(f(x) - I_\Xi f(x)\big)\mathrm{d}x.$$

*Proof.* We achieve this by mollification. Let $\kappa \in C_c^\infty(\mathbb{R}^d)$ be a smooth function which equals 1 in $B(0, 1/2)$ and vanishes outside of $B(0, 1)$. Then $\widehat{\nu_R} := \kappa(\cdot/R)\widehat{\nu}$ is a smooth test function supported in $B(0, R)$ (since $\widehat{\nu}$ is entire), hence a Schwartz function satisfying $\widehat{\nu}_R(\xi) = \mathcal{O}(|\xi|^{m_0})$; here we have used the polynomial annihilation assumption placed on $\nu$. Because $f - I_\Xi f \in \mathcal{N}(\phi)$, it has a generalized Fourier transform of order $m_0/2$, so

$$\int_{\mathbb{R}^d} \widehat{\nu_R}(\omega)\big(\widehat{f}(\omega) - \widehat{I_\Xi f}(\omega)\big)\mathrm{d}\omega = \int_{\mathbb{R}^d} \nu_R(x)\big(f(x) - I_\Xi f(x)\big)\mathrm{d}x.$$

Since $\int_{\mathbb{R}^d} |\widehat{\nu}(\omega)||\widehat{f}(\omega) - \widehat{I_\Xi f}(\omega)|\mathrm{d}\omega \le |f|_{\mathcal{N}(\phi)}|f - I_\Xi f|_{\mathcal{N}(\phi)} < \infty$, dominated convergence guarantees that $\lim_{R\to\infty} \int_{\mathbb{R}^d} |\widehat{\nu}_R(\omega) - \widehat{\nu}(\omega)|\big|\widehat{f}(\omega) - \widehat{I_\Xi f}(\omega)\big|\mathrm{d}\omega = 0$. The fact that $f - I_\Xi f \in \mathcal{N}(\phi)$ also guarantees that it is continuous and has slow growth. Thus, for any compact set $K$, we have

$$\lim_{R\to\infty} \int_K |\nu_R(x) - \nu(x)|\big|f(x) - I_\Xi f(x)\big|\mathrm{d}x = 0.$$

If $K \supset \Omega$ then $\nu_R(x) - \nu(x) = \nu_R(x)$ when $x \in \mathbb{R}^d \setminus K$. Writing $\nu_R$ as a convolution, namely $\nu_R(x) = R^d \int \nu(y)\kappa^\vee(R(x - y))\mathrm{d}y$, it follows that $|\nu_R(x)| \le CR^d(1 + R\,\mathrm{dist}(x, \Omega))^{-L}$, where we have used that $\kappa$ is a Schwartz function and $\nu \in L_1$ is supported in $\Omega$. Because $f$ and $I_\Xi f$ have algebraic growth, the estimate

$$\begin{aligned}\int_{\mathbb{R}^d\setminus K} |\nu_R(x) - \nu(x)|\big|f(x) - I_\Xi f(x)\big|\mathrm{d}x &\le CR^d \int_{\mathbb{R}^d\setminus K} \frac{|x|^{m_1}}{(1 + R\mathrm{dist}(x, \Omega))^L}\mathrm{d}x \\ &\le CR^{d-L} \int_{\mathbb{R}^d\setminus K} |x|^{m_1 - L}\mathrm{d}x \to 0\end{aligned}$$

holds and the lemma follows. □

In the next subsections, we consider two applications of Theorem 6.1. The first considers a RBF where $\widehat{\phi}$ has an algebraic singularity at the origin which determines the order of conditional positive definiteness. The second treats surface splines, considered in Examples 3.2 and 4.1, which do not satisfy the hypotheses of Theorem 5.1.

6.2. **RBFs with algebraic singularities.** In this subsection we assume $\widehat{\phi}$ has a singularity similar to $|\omega|^{-\beta_0-d}$ near the origin. If the other conditions of Theorem 6.1 hold, then Lemma 8 shows that $\nu \perp \mathcal{P}_{\lfloor \beta_0/2 \rfloor}$.

Because $|\cdot|^{-\beta_0-d+\alpha}$ is locally integrable if and only if $\alpha > \beta_0$, it follows that if $\phi$ has a generalized Fourier transform of order $m_0$, then $2m_0 > \beta_0$, since the function $\omega \mapsto |\omega|^{2m_0}\widehat{\phi}(\omega)$ must be locally integrable. Consequently, if $m_0$ is minimal in the sense that $m_0 = \lfloor \beta_0/2 \rfloor + 1$ then $\nu \perp \mathcal{P}_{\lfloor \beta_0/2 \rfloor}$ implies $\nu \perp \mathcal{P}_{m_0-1}$.

We note that this is sufficient to treat surface splines of order $m$ having the unconventional order $m_0 = \lfloor m - d/2 \rfloor + 1$; i.e., with auxiliary polynomial space $\mathcal{P}_{\lfloor m-d/2 \rfloor}$. As mentioned in Example 3.2, $\widehat{\phi_m} = |\xi|^{-2m}$, so $\beta_0 = 2m - d$ in this case. The conventional situation of surface splines with CPD order $m_0 = m$ is treated in the next section.

Another example which this section treats, which is relevant to the pseudo-spectral methods mentioned in section 1, is the case of a differential operator like $\mathcal{L} = 1 - \Delta$ applied to $\phi_m$. In that case, one can see that Assumption 1 holds from the Fourier transform: $\widehat{\mathcal{L}\phi_m}(\xi) = (1 + |\xi|^2)|\xi|^{-2m}$, although the singularity at 0 does not match the decay at infinity. Assumption 2 holds in this case, too, as can be easily checked. Finally, Corollary 2 does not apply in case, because the singularity $|\xi|^{-2m}$ is sharper than the decay at infinity $|\xi|^{2-2m}$. In this case, one could use Corollary 1.

**Lemma 8.** *Suppose $\phi$ is CPD of order $m_0$ for which there is a neighborhood $B(0,r)$ of the origin where the following two conditions hold:*

- *there is $C$ so that $\widehat{\phi}(\omega) \le C|\omega|^{-\beta_0-d}$ a.e. in $B(0,r)$*
- $\int_{B(0,r)} \widehat{\phi}(\xi)|\omega|^{\beta_0}|\log \omega|^{-1}\mathrm{d}\omega = \infty$.

*If $f \in \mathcal{N}(\phi)$ has the form $f = \nu * \phi + p$, with $\nu \in L_2(\mathbb{R}^d)$ having compact support, then $\nu \perp \mathcal{P}_{\lfloor \beta_0/2 \rfloor}$.*

Note that the above hypotheses are met if there are constants $0 < c \le C < \infty$ such that $c|\omega|^{-\beta_0-d} \le \widehat{\phi}(\omega) \le C|\omega|^{-\beta_0-d}$ a.e. in $B(0,r)$.

*Proof.* Assume without loss that $r < 1$. By [33, Theorem 10.21], since $f \in \mathcal{N}(\phi)$ it has a generalized Fourier transform which satisfies $\widehat{f}/(\widehat{\phi})^{-1/2} \in L_2(\mathbb{R}^d)$. By Hölder's inequality,

$$\int_{B(0,r)} |\widehat{f}(\omega)||\omega|^{\beta_0/2}|\log \omega|^{-1}\mathrm{d}\omega$$
$$\le \left(\int_{\mathbb{R}^d} |\widehat{f}(\omega)|^2/\widehat{\phi}(\omega)\mathrm{d}\omega\right)^{1/2}\left(\int_{B(0,r)} |\omega|^{\beta_0}|\log \omega|^{-2}\widehat{\phi}(\omega)\mathrm{d}\omega\right)^{1/2}$$

holds, so $\omega \mapsto |\widehat{f}(\omega)|\omega|^{\beta_0/2}|\log \omega|^{-1} \in L_1(B(0,r))$.

Since the support of $\nu$ is compact, $\widehat{\nu}$ is entire. Let $k \in \mathbb{N}$ be the smallest integer for which there is a multi-index $\alpha$ such that $D^\alpha\widehat{\nu}(0) \neq 0$. By Taylor's theorem, we can write $\widehat{\nu}(z) = H_k(z) + R(z)$, where $R(z) = o(|z|^k)$ as $z \to 0$. Here $H_k$ is the Taylor polynomial of degree $k$ at 0; it happens to be homogeneous because of the minimality of $k$.

For $\Theta_k \in C^\infty(\mathbb{S}^{d-1})$ defined by $H_k(z) = |z|^k \Theta_k(z/|z|)$, the set

$$\mathcal{C} := \left\{ \zeta \in \mathbb{S}^{d-1} \;\mid\; |\Theta_k(\zeta)| > \frac{1}{2}\|\Theta_k\|_\infty \right\}$$

is open and non-empty. Thus, in the cone $\{z \in \mathbb{R}^d \mid z/|z| \in \mathcal{C}\}$, we have that $|H_k(z)| \geq \frac{1}{2}\|\Theta_k\|_\infty |z|^k$. Since $R(z) = o(|z|^k)$, there is $r_0 > 0$, and a corresponding neighborhood $\mathcal{R} := \{z \mid |z| \leq r_0,\ z/|z| \in \mathcal{C}\}$ such that

$$\Big(\forall z \in \mathcal{R}\Big) \quad |\widehat{\nu}(z)| \geq \frac{1}{4}\|\Theta_k\|_\infty |z|^k.$$

Since $\widehat{\nu}(\omega) = \widehat{f}(\omega)/\widehat{\phi}(\omega)$, we have, for $\omega \in \mathcal{R}$, that

$$\begin{aligned} |\widehat{f}(\omega)||\omega|^{\beta_0/2}|\log\omega|^{-1} &= |\widehat{\nu}(\omega)|\widehat{\phi}(\omega)|\omega|^{\beta_0/2}|\log\omega|^{-1} \\ &\geq \frac{\|\Theta_k\|_\infty}{4}\widehat{\phi}(\omega)|\omega|^{\beta_0/2+k}|\log\omega|^{-1}. \end{aligned}$$

Since $\mathcal{C}$ has positive measure, the integrability of the right hand side guarantees that $k > \beta_0/2$. Because $k$ is an integer, $k \geq 1 + \lfloor \beta_0/2 \rfloor$, and the result follows. □

6.3. **Surface splines.** Suppose now that $\phi_m$ is the fundamental solution to $\Delta^m$ on $\mathbb{R}^d$. Then $\phi_m$ is CPD order $m_0 = m$, with $\mathcal{N}(\phi_m) = \mathrm{BL}_m(\mathbb{R}^d)$. We also assume the boundary of $\Omega$ is $C^\infty$ (rather than merely Lipschitz), and express its outer normal by $\vec{n} : \partial\Omega \to \mathbb{S}^{d-1}$. In this case, we replace the condition

$$f = \nu * \phi_m + p \in \mathrm{BL}_m(\mathbb{R}^d) \text{ with } \nu \in L_2(\mathbb{R}^d),\ \mathrm{supp}(\nu) \subset \Omega \text{ and } \nu \perp \mathcal{P}_{m-1} \tag{6.1}$$

by a stronger version:

$$\text{the unique Beppo-Levi extension } f_e \text{ of } f|_\Omega \text{ is in } W_{2,loc}^{2m}(\mathbb{R}^d) \tag{6.2}$$

which will ensure that the conclusion of Theorem 6.1 holds.

If $f \in W_2^{2m}(\Omega)$, then [23, Theorem 8.2] shows that the Beppo-Levi extension $\mathrm{BL}_m(\mathbb{R}^d)$ (i.e., the native space extension) can be written as $f_e = \phi_m * \nu_f + p$. Indeed, as described in [23, Section 8.2], we have that

$$\nu_f * \phi_m = \int_\Omega \Delta^m f(\alpha)\phi_m(\cdot - \alpha)\mathrm{d}\alpha + \sum_{j=0}^{m-1} \int_{\partial\Omega} N_j f(\alpha)\lambda_{j,\alpha}\phi_m(x-\alpha)\mathrm{d}\sigma(\alpha), \tag{6.3}$$

where

$$\lambda_j f = \begin{cases} \mathrm{Tr}\Delta^{j/2} f & j \text{ is even,} \\ D_{\vec{n}}\Delta^{(j-1)/2} & j \text{ is odd} \end{cases}$$

and the operators $N_j : W_2^{2m}(\Omega) \to L_2(\partial\Omega)$ are from [23, Theorem 2.4].

**Lemma 9.** *If $\Omega$ has $C^\infty$ boundary, then the condition (6.2) implies (6.1).*

*Proof.* Suppose (6.2) holds. Let $\overline{\Omega} \subset B(0,R)$, for some $R > 0$. Then because $f_e \in W_2^{2m}(B(0,R))$, the trace theorem guarantees that $\lambda_k f_e \in W_2^{2m-j-1}(\partial\Omega)$ for $0 \leq k \leq 2m-1$; in particular, the trace $\lambda_k^+$ from $\Omega$ coincides with the trace $\lambda_k^-$ from $\mathbb{R}^d \setminus \overline{\Omega}$.

By the jump conditions [23, Corollary 3.4] for layer potentials

$$V_j g = \int_{\partial\Omega} g(\alpha)\lambda_{j,\alpha}\phi_m(\cdot - \alpha)\mathrm{d}\sigma(\alpha),$$

which state that $\lambda_j^+ V_j g - \lambda_j^- V_j g = (-1)^j g$, we have that $N_j f = 0$. Thus (6.3) consists only of one term, and $\nu_f = \Delta^m f \in L_2(\mathbb{R}^d)$, which is supported in $\Omega$.

Finally, [23, Lemma 8.1] guarantees that $\nu_f \perp \mathcal{P}_{m-1}$. □

In case (6.2) holds, Corollary 2 applies and

$$|\mathcal{J}_{\sigma-m}(f - I_\Xi f)|_{\dot{H}^m} \le Cq^{m-\sigma}h^m\|\nu\|_{L_2(\Omega)}.$$

Furthermore, because $f$ and $I_\Xi f$ have generalized Fourier transforms of order $m/2$, we can use Lemma 1 to ensure that $|f - I_\Xi f|_{W_2^\sigma(\mathbb{R}^d)} \sim |f - I_\Xi f|_{H^\sigma}$ whenever $\sigma \ge m$, so for $m \le \sigma$ with $\lceil\sigma\rceil < 2m - d/2$, we have

$$|f - I_\Xi f|_{W_2^\sigma(\Omega)} \le C|f - I_\Xi f|_{\dot{H}^\sigma} \le Cq^{m-\sigma}h^m\|\nu\|_{L_2(\Omega)}. \tag{6.4}$$

Here we have used that $|u|_{\dot{H}^\sigma} \le |\mathcal{J}_{\sigma-m}(u)|_{\dot{H}^m}$ when $\sigma \ge m$. In particular, if the point set $\Xi$ is quasi-uniform with mesh ratio $\rho$, we have, with $\rho$ dependent constant,

$$|f - I_\Xi f|_{W_2^\sigma(\mathbb{R}^d)} \le Ch^{2m-\sigma}\|\nu\|_{L_2(\Omega)}.$$

*Remark* 6.2. A necessary and sufficient condition for $f \in W_2^{2m}(\Omega)$ to satisfy (6.2) is that $f \in \cap_{j=0}^{m-1}\ker(N_j)$; this is [23, Corollary 8.3].

*Remark* 6.3. A condition which implies (6.2) has been considered by Gutzmer and Melenk in [19]. Namely, that $f$ satisfies *natural boundary conditions*:

$$f \in W_2^{2m}(\Omega) \text{ and } D^\alpha f(x) = 0 \text{ for } x \in \partial\Omega \text{ and } m \le |\alpha| \le 2m-1. \tag{6.5}$$

The result [19, Lemma 2] shows that if $f$ satisfies (6.5), then $f$ satisfies (6.2). Thus (6.4) provides a higher order counterpart to their result, then [19, Theorem 2] shows that for sufficiently dense $\Xi \subset \Omega$,

$$|f - I_\Xi f|_{W_2^k(\Omega)} \le h^{2m-k}|f|_{W_2^{2m}(\Omega)}$$

holds for $k \le m$. We note that the results of [19] hold under more general conditions, namely for $\Omega$ having Lipschitz boundary without the assumption of quasi-uniformity on $\Xi$.

## Appendix A. Regular local polynomial reproductions

**Lemma 10.** *If $\Omega \subset \mathbb{R}^d$ is compact and satisfies an interior cone condition, then for every $L > 0$, there exists a constant $K$ depending on $L$ and the cone aperture, and $h_0 > 0$ depending on $L$ and both cone parameters, so that for any finite subset $\Xi \subset \Omega$ with $h(\Xi, \Omega) < h_0$ there is a stable, local polynomial reproduction of order $L$. I.e., there is a map $a(\cdot,\cdot) : \Xi \times \Omega \to \mathbb{R}$ which satisfies the following four conditions:*

(1) *for every $z \in \Omega$ if $\mathrm{dist}(\xi, z) > Kh$ then $a(\xi, z) = 0$*
(2) *for every $z \in \Omega$, $\sum_{\xi\in\Xi}|a(\xi, z)| \le 3$*
(3) *for every $p \in \mathcal{P}_L$ and $z \in \Omega$, $\sum_{\xi\in\Xi} a(\xi, z)p(\xi) = p(z)$*
(4) *for every $\xi \in \Xi$, $a(\xi, \cdot)$ is smooth.*

*Proof.* Let $N := \dim \mathcal{P}_L$. Select a basis $\{p_j \mid 1 \le j \le N\}$ for $\mathcal{P}_L$. Then the result [33, Lemma 3.14] guarantees the existence of a map $\tilde{a}$ which satisfies items (1)–(3). Indeed, for every $z \in \Omega$, $\sum_{\xi\in\Xi}\big|\tilde{a}(\xi,z)\big| \le 2$ and if $\operatorname{dist}(\xi,z) > \tilde{K}h$ then $\tilde{a}(\xi,z) = 0$.

Let $K = \tilde{K}+1$. Pick $y \in \Omega$. Let $\Xi_0 := \Xi \cap B(y,Kh)$. Because $\Xi_0$ is unisolvent, it contains a unisolvent subset $\Xi^\flat \subset \Xi_0$ with $\#\Xi^\flat = N$ (i.e., it contains a subset which is poised for interpolation by $\mathcal{P}_L$). Enumerate $\Xi^\flat := \{\xi_1,\dots,\xi_N\}$, and let $\Xi^\sharp := \Xi_0 \setminus \Xi^\flat$.

Consider now the $C^\infty$ function $F : \mathbb{R}^d \times \mathbb{R}^N \to \mathbb{R}^N$ defined by

$$\big(F(x,b)\big)_j = p_j(x) - \sum_{k=1}^N b_k p_j(\xi_k) - \sum_{\zeta\in\Xi^\sharp} \tilde{a}(\zeta,y)p_j(\zeta).$$

For $b_0 := \tilde{a}(\cdot,y)\,|_{\Xi^\flat}$, $F(y,b_0) = 0$, and $D_bF(x,b) = \big(\frac{\partial F_j}{\partial b_k}\big) = \big(p_j(\xi_k)\big)$ is the Vandermonde matrix for $\Xi^\flat$ and is therefore non-singular for all $x$. By the implicit function theorem, there is $B(y,r_1)$ and a smooth function $g : B(y,r_1) \to \mathbb{R}^N$ so that $g(y) = b_0 = \tilde{a}(\cdot,y)\,|_{\Xi^\flat}$ and $F(x,g(x)) = 0$ for all $x \in B(y,r_1)$.

Note that $\|g(y)\|_{\ell_1(\mathbb{R}^N)} \le 2 - \|\tilde{a}(\cdot,y)\,|_{\Xi^\sharp}\,\|_{\ell_1(\Xi^\sharp)}$. It follows from continuity of $g$ that there is $r_2 \in (0,r_1)$ so that we for all $x \in B(y,r_2)$

$$\|g(x)\|_{\ell_1(\mathbb{R}^N)} \le 3 - \|\tilde{a}(\cdot,y)\,|_{\Xi^\sharp}\,\|_{\ell_1(\Xi^\sharp)}.$$

By decreasing the radius even further, i.e., taking $r(y) := \min(r_2,h)$, we have that for every $x \in B\big(y,r(y)\big)$ and for every $\xi \in \Xi_0$, we have $\operatorname{dist}(x,\xi) \le Kh$, since $\operatorname{dist}(y,\xi) < \tilde{K}h$.

For $x \in B(y,r)$, set

$$a_y(\xi,x) := \begin{cases} \big(g(x)\big)_j & \xi = \xi_j \in \Xi^\flat,\\ \tilde{a}(\xi,y) & \xi \in \Xi^\sharp,\\ 0 & \xi \in \Xi\setminus\Xi_0,\end{cases}$$

and note that $a_y$ is a local polynomial reproduction of order $L$, locality $\tilde{K}$ and stability 3 in $B(y,r)$.

By compactness, there is a finite cover of the form $\Omega = \bigcup_{j=1}^M B(y_j,r(y_j))$. Denote by $a_j : \Xi\times\Omega \to \mathbb{R}$ the extension by zero of $a_{y_j} : \Xi \times B(y_j,r(y_j)) \to \mathbb{R}$. Let $\big(\psi_j\big)_{j=1\dots M}$ be a smooth partition of unity subordinate to this cover: i.e., consisting of functions $\psi_j : \Omega \to [0,1]$ with $\operatorname{supp}(\psi_j) \subset B(y_j,r(y_j))$ and $\sum_{j=1}^M \psi_j = 1$.

Then $a : \Xi\times\Omega \to \mathbb{R}$ defined by $a(\xi,z) := \sum_{j=1}^M \psi_j(z)a_j(\xi,z)$ is a smooth local polynomial reproduction, since

$$\sum_{\xi\in\Xi} p(\xi)a(\xi,z) = \sum_{j=1}^M \psi_j(z)\sum_{\xi\in\Xi} p(\xi)a_j(\xi,z) = \sum_{\substack{j=1\\ z\in B(y_j,r(y_j))}}^M \psi_j(z)p(z) = p(z).$$

We have also that $\sum_{\xi\in\Xi}|\sum_{j=1}^M \psi_j(z)a_j(\xi,z)| \le 3$, so $a$ has stability constant $\Gamma \le 3$. Finally, for $z \in \Omega$, if $a(z,\xi) \ne 0$, then for some $j$, $z \in B(y_j,r(y_j))$ and $a_j(\xi,z) \ne 0$. But this implies that $|z-\xi| \le Kh$. ∎

## Acknowledgments

The authors wish to thank the anonymous reviewers whose comments helped improve this manuscript.

## References

[1] H. Abels, *Pseudodifferential and Singular Integral Operators*, De Gruyter Graduate Lectures, De Gruyter, Berlin, 2012. An introduction with applications. MR2884718

[2] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards Applied Mathematics Series, No. 55, U. S. Government Printing Office, Washington, DC, 1964. For sale by the Superintendent of Documents. MR167642

[3] J.-P. Aubin, *Behavior of the error of the approximate solutions of boundary value problems for linear elliptic operators by Gelerkin's and finite difference methods*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3) **21** (1967), 599–637. MR233068

[4] M. D. Buhmann, *Radial functions on compact support*, Proc. Edinburgh Math. Soc. (2) **41** (1998), no. 1, 33–46, DOI 10.1017/S0013091500019416. MR1604361

[5] C. Canuto and A. Quarteroni, *Approximation results for orthogonal polynomials in Sobolev spaces*, Math. Comp. **38** (1982), no. 157, 67–86, DOI 10.2307/2007465. MR637287

[6] M. Charina, C. Conti, and N. Dyn, *Multivariate compactly supported $C^\infty$ functions by subdivision*, Appl. Comput. Harmon. Anal. **70** (2024), Paper no. 101630, DOI 10.1016/j.acha.2024.101630. MR4695556

[7] Y. Chen, B. Hosseini, H. Owhadi, and A. M. Stuart, *Solving and learning nonlinear PDEs with Gaussian processes*, J. Comput. Phys. **447** (2021), Paper No. 110668, 29, DOI 10.1016/j.jcp.2021.110668. MR4311012

[8] A. Chernih and S. Hubbert, *Closed form representations and properties of the generalised Wendland functions*, J. Approx. Theory **177** (2014), 17–33, DOI 10.1016/j.jat.2013.09.005. MR3128401

[9] W. M. Czarnecki, S. Osindero, M. Jaderberg, G. Swirszcz, and R. Pascanu, *Sobolev training for neural networks*, Adv. Neural Inf. Process Syst. **30** (2017).

[10] T. De Ryck, S. Lanthaler, and S. Mishra, *On the approximation of functions by tanh neural networks*, Neural Netw. **143** (2021), 732–750.

[11] R. Devore and A. Ron, *Approximation using scattered shifts of a multivariate function*, Trans. Amer. Math. Soc. **362** (2010), no. 12, 6205–6229, DOI 10.1090/S0002-9947-2010-05070-6. MR2678971

[12] S. Hilbert, *A mollifier useful for approximations in Sobolev spaces and some applications to approximating solutions of differential equations*, Math. Comp. **27** (1973), 81–89, DOI 10.2307/2005248. MR331715

[13] E. Di Nezza, G. Palatucci, and E. Valdinoci, *Hitchhiker's guide to the fractional Sobolev spaces*, Bull. Sci. Math. **136** (2012), no. 5, 521–573, DOI 10.1016/j.bulsci.2011.12.004. MR2944369

[14] F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, and B. V. Saunders (eds.), *NIST digital library of mathematical functions*, `http://dlmf.nist.gov/`, Release 1.0.18 of 2018-03-27.

[15] J. Duchon, *Sur l'erreur d'interpolation des fonctions de plusieurs variables par les $D^m$-splines* (French, with English summary), RAIRO Anal. Numér. **12** (1978), no. 4, 325–334, vi, DOI 10.1051/m2an/1978120403251. MR519016

[16] G. E. Fasshauer, *RBF collocation methods as pseudospectral methods*, Boundary Elements XXVII, WIT Trans. Model. Simul., vol. 39, WIT Press, Southampton, 2005, pp. 47–56. MR2234627

[17] I. M. Gel'fand and N. Ya. Vilenkin, *Generalized Functions. Vol. 4: Applications of Harmonic Analysis*, Academic Press, New York-London, 1964. Translated by Amiel Feinstein. MR173945

[18] I. Gühring, G. Kutyniok, and P. Petersen, *Error bounds for approximations with deep ReLU neural networks in $W^{s,p}$ norms*, Anal. Appl. (Singap.) **18** (2020), no. 5, 803–859, DOI 10.1142/S0219530519410021. MR4131039

[19] T. Gutzmer and J. M. Melenk, *Approximation orders for natural splines in arbitrary dimensions*, Math. Comp. **70** (2001), no. 234, 699–703, DOI 10.1090/S0025-5718-00-01299-0. MR1813144

[20] T. Hangelbroek, F. J. Narcowich, C. Rieger, and J. D. Ward, *An inverse theorem for compact Lipschitz regions in $\mathbb{R}^d$ using localized kernel bases*, Math. Comp. **87** (2018), no. 312, 1949–1989, DOI 10.1090/mcom/3256. MR3787398

[21] T. Hangelbroek, F. J. Narcowich, and J. D. Ward, *Polyharmonic and related kernels on manifolds: interpolation and approximation*, Found. Comput. Math. **12** (2012), no. 5, 625–670, DOI 10.1007/s10208-011-9113-5. MR2970852

[22] T. Hangelbroek and A. Lauve, *The polyharmonic Dirichlet problem and path counting* (English, with English and French summaries), J. Math. Pures Appl. (9) **102** (2014), no. 3, 449–481, DOI 10.1016/j.matpur.2013.12.001. MR3249719

[23] T. C. Hangelbroek, *On a polyharmonic Dirichlet problem and boundary effects in surface spline approximation*, SIAM J. Math. Anal. **50** (2018), no. 4, 4616–4654, DOI 10.1137/18M1167188. MR3849117

[24] D. S. Jones, *The Theory of Generalised Functions*, 2nd ed., Cambridge University Press, Cambridge-New York, 1982. MR665103

[25] F. J. Narcowich, J. D. Ward, and H. Wendland, *Sobolev bounds on functions with scattered zeros, with applications to radial basis function surface fitting*, Math. Comp. **74** (2005), no. 250, 743–763, DOI 10.1090/S0025-5718-04-01708-9. MR2114646

[26] F. J. Narcowich and J. D. Ward, *Norms of inverses and condition numbers for matrices associated with scattered data*, J. Approx. Theory **64** (1991), no. 1, 69–94, DOI 10.1016/0021-9045(91)90087-Q. MR1086096

[27] F. J. Narcowich, J. D. Ward, and H. Wendland, *Sobolev error estimates and a Bernstein inequality for scattered data interpolation via radial basis functions*, Constr. Approx. **24** (2006), no. 2, 175–186, DOI 10.1007/s00365-005-0624-7. MR2239119

[28] J. Nitsche, *Ein Kriterium für die Quasi-Optimalität des Ritzschen Verfahrens* (German), Numer. Math. **11** (1968), 346–348, DOI 10.1007/BF02166687. MR233502

[29] R. Schaback, *Improved error bounds for scattered data interpolation by radial basis functions*, Math. Comp. **68** (1999), no. 225, 201–216, DOI 10.1090/S0025-5718-99-01009-1. MR1604379

[30] R. Schaback, *Native Hilbert spaces for radial basis functions. I*, New Developments in Approximation Theory (Dortmund, 1998), Internat. Ser. Numer. Math., vol. 132, Birkhäuser, Basel, 1999, pp. 255–282. MR1724923

[31] R. Schaback and H. Wendland, *Kernel techniques: from machine learning to meshless methods*, Acta Numer. **15** (2006), 543–639, DOI 10.1017/S0962492906270016. MR2269747

[32] H. Wendland, *Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree*, Adv. Comput. Math. **4** (1995), no. 4, 389–396, DOI 10.1007/BF02123482. MR1366510

[33] H. Wendland, *Scattered Data Approximation*, Cambridge Monographs on Applied and Computational Mathematics, vol. 17, Cambridge University Press, Cambridge, 2005. MR2131724

[34] Z. M. Wu, *Compactly supported positive definite radial functions*, Adv. Comput. Math. **4** (1995), no. 3, 283–292, DOI 10.1007/BF03177517. MR1357720

[35] W. Zhao, Y.-c. Hon, and M. Stoll, *Localized radial basis functions-based pseudo-spectral method (LRBF-PSM) for nonlocal diffusion problems*, Comput. Math. Appl. **75** (2018), no. 5, 1685–1704, DOI 10.1016/j.camwa.2017.11.030. MR3766544

Department of Mathematics, University of Hawaiʻi–Mānoa, 2565 McCarthy Mall, Honolulu, Hawaii 96822

*Email address*: hangelbr@math.hawaii.edu

Department of Mathematics and Computer Science, Philipps-Universität Marburg, Hans-Meerwein-Strasse 6, 35032 Marburg, Germany

*Email address*: riegerc@mathematik.uni-marburg.de