

## AN INVERSE THEOREM FOR COMPACT LIPSCHITZ REGIONS IN $\mathbb{R}^d$ USING LOCALIZED KERNEL BASES

T. HANGELBROEK, F. J. NARCOWICH, C. RIEGER, AND J. D. WARD

ABSTRACT. While inverse estimates in the context of radial basis function approximation on boundary-free domains have been known for at least ten years, such theorems for the more important and difficult setting of bounded domains have been notably absent. This article develops inverse estimates for finite dimensional spaces arising in radial basis function approximation and meshless methods. The inverse estimates we consider control Sobolev norms of linear combinations of a localized basis by the  $L_p$  norm over a bounded domain. The localized basis is generated by forming local Lagrange functions for certain types of RBFs (namely Matérn and surface spline RBFs). In this way it extends the boundary-free construction recently presented by Fuselier, Hangelbroek and Narcowich [*Localized bases for kernel spaces on the unit sphere*, SIAM J. Numer. Anal. 51 (2013), no. 5, 2358-2562].

### 1. INTRODUCTION

This article presents a construction for localized bases generated by radial basis functions (RBFs) in the presence of a boundary and develops analytic properties of this basis, most notably inverse inequalities. Such inequalities are an essential tool in the numerical solution of PDEs by finite element and related methods (see [3, 10, 11]) notably in proving inf-sup (Babuška-Brezzi) conditions, which play a central role for mixed element and saddle point problems [1, 13, 14, 21]. They are also prevalent in approximation theory (where they are called 'Bernstein inequalities'); specifically they are used to obtain characterization of approximation spaces as interpolation spaces by way of  $K$ -functionals [4].

The type of localized basis investigated in this article was introduced very recently for the boundary-free setting (e.g., on a manifold without boundary) and has already been employed to deliver strong results in function approximation and scattered data fitting [8], numerical quadrature [9] and solution of PDEs [24] and

---

Received by the editor August 25, 2015, and, in revised form, May 19, 2016, October 14, 2016, and December 28, 2016.

2010 *Mathematics Subject Classification*. Primary 41A17, 41A27, 41A63.

*Key words and phrases*. Radial basis functions, Sobolev spaces, Bernstein and Nikolskii inequalities, trace estimate.

The first author's research was supported by grant DMS-1413726 from the National Science Foundation.

The second author's research was supported by grant DMS-1514789 from the National Science Foundation.

The third author's research was supported by Collaborative Research Centre (SFB) 1060: The Mathematics of Emergent Effects, of the Deutsche Forschungsgemeinschaft.

The fourth author's research was supported by grant DMS-1514789 from the National Science Foundation.

integral equations [19]. Indeed, in [19], Lehoucq and Rowe have applied the localized basis investigated in this article to obtain a Galerkin solution to a constrained integral equation, and they have used the  $L_p$  stability of the basis (presented in this paper in Section 4) to obtain norm bounds on the stiffness matrix associated with this problem.

The inverse estimates we consider treat finite dimensional spaces of functions, bounding strong (Sobolev) norms by weak (Lebesgue) norms:

$$(1.1) \quad \|s\|_{W_p^\sigma(\Omega)} \leq Ch^{-\sigma} \|s\|_{L_p(\Omega)} \quad (\text{or } \|s\|_{C^\sigma(\Omega)} \leq Ch^{-\sigma} \|s\|_{L_\infty(\Omega)} \text{ for } p = \infty),$$

where  $\Omega$  is a bounded subset in  $\mathbb{R}^d$ , subject to mild conditions on  $\partial\Omega$  and  $h$  is the fill distance (also known as mesh ratio) of the finite set of points used to generate our finite dimensional space (see Section 2.1 for a precise description). In one sense, these estimates can be viewed as providing an operator norm bound (from  $L_p \rightarrow L_p$ ) of differential operators restricted to this finite dimensional space. In another sense, they give precise equivalences between different norms in terms of a simple measure of the complexity (given by the parameter  $N$  above) of the finite dimensional space. Direct consequences of these inverse estimates include trace estimates and Bernstein-Nikolskii inequalities.

This topic has been considered in the boundary-free setting by a number of authors, we list [26], [27], [22], [31], [12] (although there are certainly others). The inequalities we consider here are similar, but depend only on the norm of a basic function over a bounded region.<sup>1</sup> Without a doubt this type of estimate is significantly more challenging when a boundary is present and has, to the best of our knowledge, remained elusive. Indeed, such inverse inequalities seem to have been absent for meshless methods in general (not only radial basis function approximation; cf. the discussion in [21, Section 7]).

In this article we consider two prominent families of radial basis functions: the Matérn (or Whittle-Matérn) and surface spline kernels. Generalizations to other kernels and other settings (namely, compact Riemannian manifolds) are considered in the manuscript [15].

The conventional finite dimensional space associated with a positive definite RBF  $\phi$  and a finite set  $X \subset \mathbb{R}^d$  has the form  $S(X) = \text{span}_{\eta \in X} \phi(\cdot - \eta)$ ; for a conditionally positive definite RBF,  $S(X)$  involves polynomials; see Section 2.5.2. A common set-up for a host of numerical problems invites the user to employ the basis of sampled kernels  $\phi(\cdot - \eta)$ ,  $\eta \in X$  as one would use polynomials, splines, finite elements, etc.; that is to say, as test functions for Galerkin or collocation methods, or as basis functions to solve interpolation, quadrature or other basic problems.

For a basic interpolation problem, using  $S(X)$  to interpolate data sampled at the point set  $X$ , the ensuing interpolation matrix will be positive definite, thanks to the kernel's positive definiteness, but if  $X$  is sampled densely, the interpolation matrix will become dense<sup>2</sup>.

Instead of using the basis of kernels, one may attempt to use another basis for  $S(X)$ ; one for which basic matrices (Gram, collocation, stiffness, interpolation) exhibit off-diagonal decay. Univariate splines provide a prime example of this phenomenon: for a fixed  $k$ , the shifted truncated powers  $(x - t_j) \mapsto x_+^k$  provide, in

<sup>1</sup>A previous result in the setting of a bounded region was presented in [28], but these estimates significantly undershoot the precise exponent  $-\sigma$  in (1.1).

<sup>2</sup>One may attempt to circumvent this problem by dilating the kernel; this is often done, but will generally result in degraded rates of approximation.

conjunction with polynomials of degree  $k$  or less, a basis for the spline space with breakpoints at  $t_j$ , but this basis is known to be poorly localized. However, the  $B$ -spline basis is well-localized, with elements having support which is not only compact, but *stationary* in the sense that it shrinks with the spacing of the breakpoints.

We are concerned with an analogous *localization problem* for radial basis functions:

Is there a basis for  $S(X)$  where the various elements exhibit a fast rate of *stationary* decay?

If an alternative basis is available for which the interpolation matrices are sparse, we say the basis is well-localized. For the Matérn and surface-spline kernels, the Lagrange function  $\chi_\eta$  is well-localized in a neighborhood of  $\eta$  where the points from  $X$  are distributed quasi-uniformly. If this is not the case, for instance if  $\eta$  occurs near to the boundary of the convex hull of  $X$ , localization is lost.

This issue can be circumvented by using only the Lagrange basis elements  $\chi_\xi$  that have centers  $\xi$  in a subset  $\Xi \subset X$ , where  $\Xi$  is chosen from  $X$  so that the Lagrange functions  $\chi_\xi$ ,  $\xi \in \Xi$ , are localized. Using these elements we may define  $V_\Xi := \text{span}_{\xi \in \Xi} \chi_\xi$ , which is of course a subspace of  $S(X)$ . To avoid a possible point of confusion, we emphasize that  $V_\Xi \neq S(\Xi)$ . The former space requires all basis functions centered in  $X$  for its construction; the latter, only those in  $\Xi$ .

After this initial streamlining, it is important to note that even though  $\chi_\xi$ ,  $\xi \in \Xi$ , is spatially localized, its construction still requires *all* of the points in  $X$ . Thus finding the  $\chi_\xi$ 's is computationally expensive. In [8], *local* Lagrange functions  $\{b_\xi\}_{\xi \in \Xi}$  were introduced. Constructing them is done by first choosing points  $\Upsilon(\xi) \subset X$  in a small neighborhood of  $\xi \in \Xi$ , and then finding the Lagrange function  $b_\xi \in S(\Upsilon(\xi)) \subset S(X)$ . Since  $\Upsilon(\xi)$  will contain many fewer points than  $X$ , it will be much less expensive to find  $b_\xi$ . Finally, we define  $\tilde{V}_\Xi = \text{span}_{\xi \in \Xi} b_\xi$ , which is a subspace of  $S(X)$ . We remark that  $\chi_\xi \neq b_\xi$  and  $V_\Xi \neq \tilde{V}_\Xi$ . However, they are close—a fact that will prove important in what follows.

We now turn to the connection between the set  $\Omega$  and the spaces described above. At the start, we are given a quasi-uniform set  $\Xi \subset \Omega$ . The enlarged set  $X$  is *not* given. Rather, an extension is constructed from  $\Xi$ , using a method—described in Section 2.3—that preserves the key geometric properties of  $\Xi$ . The extension, which will be denoted by  $\tilde{\Xi}$  later (instead of  $X$ ), is contained in a bounded region  $\tilde{\Omega}$  that contains  $\Omega$  and is roughly speaking about twice the size of  $\Omega$ . It is for this setup that we get estimates of the form (1.1) for  $s \in V_\Xi$  or  $\tilde{V}_\Xi$ . (See Theorem 5.1.)

**1.1. Overview and outline.** We begin by giving a basic explanation and background on RBFs used in this article. This is done in Section 2.

In Section 3, we introduce the Lagrange basis (the functions generating the space  $V_\Xi$ ) and provide estimates that control the Sobolev norm (i.e.,  $W_p^\sigma(\Omega)$ ) of a function in  $V_\Xi$  by the  $\ell_p$  norm on the Lagrange coefficients and in addition by the  $L_p$  norm of  $s$ . That is, for  $s = \sum_{\xi \in \Xi} a_\xi \chi_\xi$  we show

$$\|s\|_{W_p^\sigma(\Omega)} \leq C(\#\Xi)^{1/p-\sigma/d} \|(a_\xi)_{\xi \in \Xi}\|_{\ell_p(\Xi)} \quad \text{and} \quad \|s\|_{W_p^\sigma(\Omega)} \leq C(\#\Xi)^{-\sigma/d} \|s\|_{L_p(\Omega)}.$$

Such a result has not appeared previously.

Section 4 introduces the other stable basis considered in this paper: the local Lagrange basis, which generates the space  $\tilde{V}_\Xi$ . We give sufficient conditions to

prove existence and stability of such a basis. We give estimates that control the Sobolev norm (i.e.,  $W_p^\sigma(\Omega)$ ) of a function in  $\tilde{V}_\Xi$  by the  $\ell_p$  norm on the local Lagrange coefficients and by the  $L_p$  norm of the function. This result is presented in Theorem 4.10. Next we compare the sequence norm with the  $L_p$  norm of an expansion  $s = \sum_{\xi \in \Xi} a_\xi \chi_\xi \in V_\Xi$  or  $s = \sum_{\xi \in \Xi} a_\xi \chi_\xi \in \tilde{V}_\Xi$  over the domain  $\Omega$ . We thus obtain

$$\|(a_\xi)_{\xi \in \Xi}\|_{\ell_p(\Xi)} \sim C(\#\Xi)^{-1/p} \|s\|_{L_p(\Omega)}.$$

In the final section we give our main inverse estimates. For  $s \in \tilde{V}_\Xi$  we have

$$\|s\|_{W_p^\sigma(\Omega)} \leq C(\#\Xi)^{-\sigma/d} \|s\|_{L_p(\Omega)},$$

and we use this to demonstrate trace estimates for that space.

## 2. BACKGROUND: RBF APPROXIMATION ON BOUNDED DOMAINS

We begin by describing the basic elements used in this article, starting with geometric properties of point sets, a discussion of the the underlying domain, smoothness spaces on the domain, and finishing with some background about the radial basis functions which we use.

**2.1. Point sets.** Given a set  $D \subset \mathbb{R}^d$  and a discrete, possibly infinite, set  $X \subset D$ , we define its *fill distance* (or *mesh norm*)  $h$ , the *separation radius*  $q$  and the *mesh ratio*  $\rho$  to be:

$$(2.1) \quad h(X, D) := \sup_{x \in D} \text{dist}(x, X), \quad q(X) := \frac{1}{2} \inf_{\xi \in X} \text{dist}(\xi, X \setminus \{\xi\}), \quad \rho(X, D) := \frac{h(X, D)}{q(X)},$$

where in defining  $\rho(X, D)$  we assume that  $q(X) > 0$ .

When there is no chance of confusion, we drop dependence in these parameters on  $X$  and  $D$  (referring simply to  $h, q$  and  $\rho$ ).

*Remark 2.1.* A finite fill distance  $h$  guarantees that the set  $D$  is covered by the family of balls  $B(\xi, h) := \{x \in D \mid \text{dist}(x, \xi) < h\}$ ,  $\xi \in X$ . A positive separation radius  $q$  guarantees that  $B(\xi, q) \cap B(\zeta, q) = \{\}$  for distinct  $\zeta, \xi \in \Xi$ . The mesh ratio, which automatically satisfies  $\rho \geq 1$ , measures the uniformity of the distribution of  $X$  in  $D$ . The larger  $\rho(X, D)$  is, the less uniform the distribution is. If  $\rho$  is finite, then we say that the point set  $X$  is quasi-uniformly distributed (in  $D$ ), or simply that  $X$  is quasi-uniform.

Note that, for a compact subset  $D$  and a nonempty, finite subset  $X \subset D$ , the fill distance and separation radius are both positive and finite  $0 < q < h < \infty$ . Consequently,  $\rho$  is finite, too.

Many of the results in this article depend in some way on the geometry of the point set  $X$ ; often this emerges in an estimate, where a constant depends on  $\rho$ . In most cases, (as one may expect) the strength of the estimate degrades as  $\rho$  increases. Throughout the paper, we have attempted make this control explicit, by factoring, whenever possible, the constant into a part which is totally independent of the point set, and another, which is a function of  $\rho$ .

It is often useful to estimate certain sums over  $X$ . Assume that  $q(X) > 0$ . If  $f : [0, \infty) \rightarrow [0, \infty)$  is a positive, decreasing, continuous function, then

$$(2.2) \quad \sum_{\zeta \in X} f(\text{dist}(\zeta, \xi)) \leq f(0) + C \sum_{n=1}^{\infty} n^{d-1} f(nq),$$

where  $C$  depends only on the spatial dimension  $d$ . This is easily established by introducing annuli centered at  $\xi$ , with inside radius  $nq$  and outside radius  $(n+1)q$ ,  $n \geq 1$ . The number of points contained in each annulus is proportional to  $n^{d-1}$ , and the contribution to the sum from each  $n$ ,  $n \geq 1$ , is less than  $n^{d-1}f(nq)$ . Hence, (2.2) holds.

**2.2. The domain  $\Omega$ .** We now consider a bounded region  $\Omega \subset \mathbb{R}^d$  containing a finite point set  $\Xi$  with  $h = h(\Xi, \Omega)$  and  $q = q(\Xi)$  as defined above. This presents two challenges.

The first concerns  $\Xi$ —although we may expect it to be finely sampled (often referred to as *sufficiently dense*, meaning that  $h(\Xi, \Omega)$  is small) in  $\Omega$ , it will not be so in a neighborhood of  $\Omega$ . To construct the localized bases to be used in what follows, we need a larger set  $X \subset \mathbb{R}^d$  so that  $X \cap \Omega = \Xi$ . In other words, we require some extra points to lie outside of  $\Omega$  (in fact, when working with local Lagrange functions  $b_\xi$ , it suffices to consider only a very small extension  $\Upsilon \subset \{x \in \mathbb{R}^d \mid \text{dist}(x, \Omega) < Kh \log h\}$ ). See Figure 1 for an illustration of this. This assumption is in place to guarantee decay of the basis functions; in other words, it is only a tool for guaranteeing the decay of  $\chi_\xi$  or  $b_\xi$ , and is not otherwise important for the stability estimate. It would be quite reasonable to be ‘given’ initially only the set  $\Xi \subset \Omega$  and to use this to construct  $X$ . In Lemma 2.2 below we demonstrate how to extend a given set of centers  $\Xi \subset \Omega$  in a controlled way to obtain a satisfactory set  $X$ .

The second challenge concerns the domain  $\Omega$ . For estimates relating  $\|\mathbf{a}\|_{\ell_p}$  and the  $L_p$  norm of expansions  $\|\sum_\xi a_\xi b_\xi\|$  or  $\|\sum_\xi a_\xi \chi_\xi\|$  the boundary becomes more important. The extra assumption we make on  $\Omega$ , in force throughout the article, is that  $\Omega$  satisfies an interior cone condition (see Appendix A for a discussion).

**2.3. Extending points.** Given  $\Omega$  and  $\Xi \subset \Omega$ , we wish to find an extension  $\Xi_{\text{ext}} \supset \Xi$  dense in  $\mathbb{R}^d$  so that the separation radius does not decrease and the fill distance is controlled (and, consequently,  $\rho$  does not increase). A simple constructive example is the following.

**Lemma 2.2.** *Suppose  $\Xi \subset \Omega$  has fill distance  $h(\Xi, \Omega) = h$  and separation radius  $q(\Xi) = q$ . Then there is a discrete set  $\Xi_{\text{ext}}$  so that  $\Xi_{\text{ext}} \cap \Omega = \Xi$ ,  $q(\Xi_{\text{ext}}) = q$ , and  $h(\Xi_{\text{ext}}, \mathbb{R}^d) = h(\sqrt{d}/2 + 2)$ .*

*Proof.* We proceed as follows: let  $\Xi_{\text{ext}} = \Xi \cup \{\zeta \in h\mathbb{Z}^d \mid \text{dist}(\zeta, \Omega) \geq h\}$ . We note that  $h(h\mathbb{Z}^d, \mathbb{R}^d) = \frac{\sqrt{d}}{2}h$  and  $q(h\mathbb{Z}^d) = h$ . It follows immediately that  $q(\Xi_{\text{ext}}) = q$ . If  $x \in \mathbb{R}^d$  is within  $(\frac{\sqrt{d}}{2} + 1)h$  of  $\Omega$ , then  $\text{dist}(x, \Xi) \leq (\frac{\sqrt{d}}{2} + 2)h$ . On the other hand, if  $x \in \mathbb{R}^d$  satisfies  $\text{dist}(x, \Omega) > (\frac{\sqrt{d}}{2} + 1)h$ , then there is  $\zeta \in h\mathbb{Z}^d$  with  $\text{dist}(x, \zeta) < \frac{\sqrt{d}}{2}h$  so that  $\text{dist}(\zeta, \Omega) > h$  (and  $\zeta$  is therefore in  $\Xi_{\text{ext}}$ ).  $\square$

*Remark 2.3.* We note that other extensions exist which do not increase  $h$ . For example, [15, Lemma 5.1] extends points so that  $h(\Xi_{\text{ext}}, \mathbb{R}^d) = h$  and  $q(\Xi_{\text{ext}}) = \min(q, h/2)$ . As an expository convenience, we use an extension  $\Xi_{\text{ext}}$  of  $\Xi$  to  $\mathbb{R}^d$

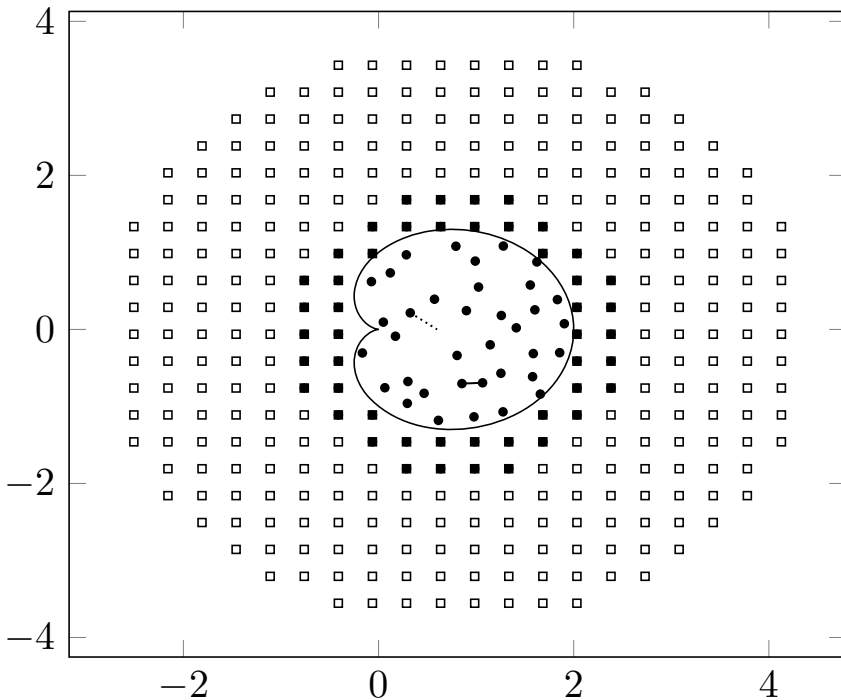


FIGURE 1. This image shows the domain  $\Omega$  (the region inside the cardioid) with a set of points  $\Xi \subset \Omega$  indicated with  $\bullet$ . The dotted line segment indicates  $h$ —the greatest distance between a point of  $\Omega$  and one of  $\Xi$ . The solid line segment indicates  $2q$ , the nearest neighbor distance in  $\Xi$ . The elements in the extended point set  $\tilde{\Xi} \setminus \Xi$  are denoted with a square—these are the centers used to construct  $\chi_\xi$  (discussed in Section 3). The points  $\blacksquare$  denote the points of  $\Upsilon$ , which are used to construct  $b_\xi$  (this is done in Section 4).

which does not increase  $h$ . In practice, an extension could be used which might not precisely preserve the geometry of the point set (such as the elementary one in Lemma 2.2). This will not change the results in this paper, other than by modifying slightly the constants. We leave it to the reader to make the (very simple) modifications necessary to treat other extensions (which would increase  $h$  and  $\rho$ ).

We construct the extended point set in an extended neighborhood

$$(2.3) \quad \tilde{\Xi} := \Xi_{\text{ext}} \cap \tilde{\Omega}, \quad \text{where} \quad \tilde{\Omega} := \{x \in \mathbb{R}^d \mid \text{dist}(x, \Omega) \leq \text{diam}(\Omega)\}$$

and where  $\Xi_{\text{ext}}$  is constructed according to the method of Remark 2.3.

**2.4. Smoothness spaces on  $\Omega$ .** In order to present a suitably robust family of inverse estimates, we employ a scale of spaces depending on a positive, occasionally fractional, smoothness parameter; as in [1], for integer values of this parameter, we use the conventional Sobolev spaces, while for fractional values we use fractional spaces, which involve a Hölder-like seminorm.

For a domain  $\Omega \subset \mathbb{R}^d$ , the integer order Sobolev space is defined in the conventional way. For  $1 \leq p < \infty$  and  $m \in \mathbb{N}$ , we have the seminorm and norm

$$|u|_{W_p^m(\Omega)}^p := \sum_{|\alpha|=m} \binom{m}{\alpha} \int_{\Omega} |D^\alpha u(x)|^p dx, \quad \|u\|_{W_p^m(\Omega)}^p := \sum_{k=0}^m \binom{m}{k} |u|_{W_p^k(\Omega)}^p.$$

Note that for the first expression (the Sobolev seminorm), we use the binomial coefficient with multi-integers  $\binom{m}{\alpha} = \frac{m!}{\alpha_1! \dots \alpha_d!}$  while for the second we use a standard binomial coefficient  $\binom{m}{k} = \frac{m!}{k!(m-k)!}$ . Although other weights would give equivalent norms, resulting in the same Sobolev spaces equipped with the same topology, these choices of coefficients will be necessary to obtain the specific reproducing kernels we desire (see Sections 2.5.1 and 2.5.2).

For fractional orders  $\sigma = m + \delta \notin \mathbb{N}$  with  $0 < \delta < 1$  the norm for  $W_p^\sigma(\Omega)$  is obtained from  $\|u\|_{W_p^\sigma(\Omega)}^p := \|u\|_{W_p^m(\Omega)}^p + |u|_{W_p^\sigma(\Omega)}^p$ , where we add the Slobodeckij seminorm

$$|u|_{W_p^\sigma(\Omega)}^p := \sum_{|\alpha|=m} \int_{\Omega} \int_{\Omega} \frac{|D^\alpha u(x) - D^\alpha u(y)|^p}{|x - y|^{d+p\delta}} dx dy.$$

We note that when  $\sigma = m + \delta$  is fractional<sup>3</sup>,  $W_p^\sigma(\Omega)$  is the Besov space  $B_{p,p}^\sigma(\Omega)$  (this is [5, Theorem 6.7]). In particular,  $W_p^\sigma(\Omega) = B_{p,p}^\sigma(\Omega) = [W_p^m(\Omega), W_p^{m+1}(\Omega)]_{\delta,p}$  serves as the  $[\delta, p]$  (real) interpolation space between  $W_p^m(\Omega)$  and  $W_p^{m+1}(\Omega)$  (see [30, 1.6.2] for a definition and basic results).

Of particular importance is the fact that, for  $2 \leq p < \infty$  and  $m \in \mathbb{N}$ , we have the continuous embedding  $W_2^m(\Omega) \subset W_p^s(\Omega)$  for all  $s \leq m - (d/2 - d/p)$ .

Throughout the paper, we make the (not unusual) modification  $W_\infty^m(\Omega) = C^m(\bar{\Omega})$  when  $p = \infty$  and  $m \in \mathbb{N}$ . For fractional order spaces when  $p = \infty$  (discussed in Section 5), we use the Hölder space  $C^s(\bar{\Omega})$ , for which the seminorm  $\max_{|\alpha|=\lfloor s \rfloor} \sup_{x,y \in \bar{\Omega}} \frac{|D^\alpha u(x) - D^\alpha u(y)|^p}{|x-y|^\delta}$  is finite for  $\delta = s - \lfloor s \rfloor$ . In this case,  $W_2^m(\Omega) \subset C^s(\bar{\Omega})$  for all  $s < m - d/2$ .

**2.4.1. Scaling and fractional Sobolev spaces.** For an open set  $O \subset \mathbb{R}^d$ , let us introduce the notation  $O_R := \{x \mid x/R \in O\}$ . The following lemma shows how the fractional Sobolev seminorm scales with  $R$ .

**Lemma 2.4.** *Suppose  $1 \leq p \leq \infty$ ,  $s \in [0, \infty)$  and  $u \in W_p^s(O_R)$ . Let  $U : O \rightarrow \mathbb{C} : x \mapsto u(Rx)$ . Then*

$$|u|_{W_p^s(O_R)} = CR^{d/p-s} |U|_{W_p^s(O)}.$$

*Proof.* We consider the case  $1 \leq p < \infty$  and  $s = k + \delta$ ,  $0 < \delta < 1$ , since the cases where  $s$  is an integer and  $p = \infty$  follow similarly, but are much easier. For  $RX = x$ , the chain rule gives us  $D^\alpha u(x) = R^{-|\alpha|} U(X)$  and

$$\begin{aligned} |u|_{W_p^s(O_R)}^p &= R^{d-p\delta-pk} \sum_{|\alpha|=k} \int_O \int_O \frac{|D^\alpha U(X) - D^\alpha U(Y)|^p}{|X - Y|^{d+p\delta}} dY dX \\ &= R^{d-ps} |U|_{W_p^s(O)}^p. \end{aligned} \quad \square$$

<sup>3</sup>When  $\sigma = m$  is an integer, we have  $W_2^m(\Omega) = B_{2,2}^m(\Omega)$ , although  $W_p^m(\Omega) \subsetneq B_{p,p}^m(\Omega)$  for  $p > 2$  and  $W_p^m(\Omega) \supsetneq B_{p,p}^m(\Omega)$  for  $p < 2$ .

2.4.2. *Subadditivity and fractional Sobolev spaces.* Carstensen and Faermann [2] have pointed out that the  $p$ th power  $|u|_{W_p^\sigma(\Omega)}^p$  of the fractional Sobolev seminorm fails to be subadditive. This is in contrast to the ( $p$ th power) integral order seminorms, which are obtained from integrals of nonnegative functions, and are easily seen to be subadditive.

The following lemma is a modification of a result of Faermann ([7, Lemma 3.1]) which we use as a tool to treat the issue of nonsubadditivity. This will be used in what follows.

**Lemma 2.5.** *Suppose  $\mathcal{V} = \{\tilde{v}_j \mid j \in \mathcal{N}\}$  is a countable family of subsets  $\tilde{v}_j \subset \Omega$  covering  $\Omega$  with finite overlap: i.e.,  $\Omega \subset \bigcup_{j \in \mathcal{N}} \tilde{v}_j$  and there is  $M > 0$  so that  $\max_{x \in \Omega} \sum_{j \in \mathcal{N}} \chi_{\tilde{v}_j}(x) \leq M$ . Suppose further that there exist sets  $v_j \supset \tilde{v}_j$  so that the complements  $w_j := \Omega \setminus v_j$  each are a fixed positive distance from the corresponding sets  $\tilde{v}_j$ : i.e., there is  $H > 0$  so that for every  $j \in \mathcal{N}$ ,  $\inf_{x \in \tilde{v}_j, y \in w_j} |x - y| \geq H$ .*

*Let  $1 \leq p < \infty$  and  $s \in (0, \infty) \setminus \mathbb{N}$  with  $k = \lfloor s \rfloor$  and  $\delta = s - \lfloor s \rfloor$ . Then for any  $u \in W_p^s(\Omega)$  we have*

$$(2.4) \quad \|u\|_{W_p^s(\Omega)}^p \leq \left( \sum_{j \in \mathcal{N}} |u|_{W_p^s(v_j)}^p \right) + CMH^{-p\delta} \|u\|_{W_p^k(\Omega)}^p.$$

*Proof.* By subadditivity of the outer integral, we have that

$$|u|_{W_p^s(\Omega)}^p \leq \sum_{|\alpha|=k} \sum_{j \in \mathcal{N}} \int_{\tilde{v}_j} \int_{\Omega} \frac{|D^\alpha u(x) - D^\alpha u(y)|^p}{|x - y|^{d+p\delta}} dy dx.$$

For fixed  $\alpha$  and  $j$ , the fact that  $v_j \cup w_j = \Omega$  permits us to bound the double integral  $\int_{\tilde{v}_j} \int_{\Omega} \frac{|D^\alpha u(x) - D^\alpha u(y)|^p}{|x - y|^{d+p\delta}} dy dx$  by

$$\int_{\tilde{v}_j} \int_{v_j} \frac{|D^\alpha u(x) - D^\alpha u(y)|^p}{|x - y|^{d+p\delta}} dy dx + \int_{\tilde{v}_j} \int_{w_j} \frac{|D^\alpha u(x) - D^\alpha u(y)|^p}{|x - y|^{d+p\delta}} dy dx.$$

Summing over the first terms gives the bound  $\sum_{j \in \mathcal{N}} |u|_{W_p^s(v_j)}^p$ , since  $\tilde{v}_j \subset v_j$ , and so this gives the first part of the right-hand side of (2.4).

Consider the sum of the second terms. Applying the quasi-triangle inequality  $(a + b)^p \leq C_p(a^p + b^p)$  to the numerator  $|D^\alpha u(x) - D^\alpha u(y)|^p$ , we can control the double integral  $\int_{\tilde{v}_j} \int_{w_j} \frac{|D^\alpha u(x) - D^\alpha u(y)|^p}{|x - y|^{d+p\delta}} dy dx$  by

$$C_p \left( \int_{\tilde{v}_j} \int_{w_j} \frac{|D^\alpha u(x)|^p}{|x - y|^{d+p\delta}} dy dx + \int_{\tilde{v}_j} \int_{w_j} \frac{|D^\alpha u(y)|^p}{|x - y|^{d+p\delta}} dy dx \right) =: J_{j,1} + J_{j,2}.$$

We have that  $\bigcup_{j \in \mathcal{N}} (\tilde{v}_j \times w_j) \subset \{(x, y) \in \Omega^2 \mid |x - y| > H\}$ . By symmetry, we have also that  $\bigcup_{j \in \mathcal{N}} (w_j \times \tilde{v}_j) \subset \{(x, y) \in \Omega^2 \mid |x - y| > H\}$ . Using the finite overlap, we have that

$$\left. \begin{aligned} & \sum_{j \in \mathcal{N}} \chi_{[\tilde{v}_j \times w_j]} \\ & \sum_{j \in \mathcal{N}} \chi_{[w_j \times \tilde{v}_j]} \end{aligned} \right\} \leq M \chi_{\{(x,y) \in \Omega^2 \mid |x-y| > H\}}.$$

Consequently, for any nonnegative, integrable  $g : \Omega \times \Omega \rightarrow \mathbb{R}$ , we have

$$\left. \begin{aligned} & \sum_{j \in \mathcal{N}} \int_{\tilde{v}_j} \int_{w_j} g(x, y) dx dy \\ & \sum_{j \in \mathcal{N}} \int_{w_j} \int_{\tilde{v}_j} g(x, y) dx dy \end{aligned} \right\} \leq M \int_{\{(x,y) \in \Omega^2 \mid |x-y| > H\}} g(x, y) dx dy$$



as well. Setting  $g(x, y) = \frac{|D^\alpha u(x)|^p}{|x-y|^{d+p\delta}}$  and applying Fubini–Tonelli allows us to control  $\sum_{j \in \mathcal{N}} (J_{j,1} + J_{j,2})$  by

$$\begin{aligned} \sum_{j \in \mathcal{N}} (J_{j,1} + J_{j,2}) &\leq 2M \int_{\{(x,y) \in \Omega^2 \mid |x-y| > H\}} \frac{|D^\alpha u(x)|^p}{|x-y|^{d+p\delta}} dx dy \\ &\leq CMH^{-p\delta} \int_{\Omega} |D^\alpha u(x)|^p dx. \end{aligned}$$

(For the last estimate, we have used the fact that  $\int_{|x-y| > H} \frac{1}{|x-y|^{d+p\delta}} dy \leq C_{p\delta} H^{-pd}$  for all  $x$ .) □

**2.5. Radial basis functions.** There are two families of radial basis functions considered in this article: the Matérn functions and the surface splines. Both families (under the right conditions) admit exponentially decaying basis functions; this is mentioned in Section 3.1. They also admit rapidly constructed localized basis functions (having polynomial decay); this is demonstrated in Sections 4.3.1 and 4.3.2. The results we present in Sections 5 hold for these families.

Two features common to both families are:

- (1) For any finite set of points  $\Xi \subset \mathbb{R}^d$  the interpolation problem is well posed. This means that for any data  $(\xi, y_\xi)_{\xi \in \Xi}$ , there exists a unique interpolant  $s$  generated by the RBF.
- (2) The RBF is a reproducing kernel for a (semi-)Hilbert space, called the *native space*, and the unique interpolant to  $(\xi, y_\xi)_{\xi \in \Xi}$  is the *best* interpolant in this space: it has the least (semi-)norm among all interpolants to the data.

We include both families (which are in some ways quite similar) because both are often in use, practically. The first is prized for the RBF’s rapid decay and strict positive definiteness; the second is included for its dilation invariance and its historical significance. Of course, there are many other prominent families of RBFs, each with its own distinguishing features (some are infinitely smooth, some are compactly supported, etc.). Rather than give a broad overview, we introduce the specific families employed in this paper and direct the interested reader to [32] for a comprehensive introduction to RBF theory. At this point it is unclear if the algorithm for constructing localized bases works for other families; the arguments we employ rely heavily on the RBF’s role as the fundamental solution to an elliptic partial differential operator.

**2.5.1. Matérn kernels.** The Matérn function of order  $m > d/2$  is defined as

$$(2.5) \quad \kappa_m : \mathbb{R}^d \rightarrow \mathbb{R} : x \mapsto CK_{m-d/2}(|x|) |x|^{m-d/2}.$$

Here  $C$  is a constant depending on  $m$  and  $d$ , and  $K_\nu$  is a Bessel function of the second kind.

The Matérn function is positive definite, which means that for every finite set  $X \subset \mathbb{R}^d$ , the *collocation* matrix

$$K_X := (\kappa_m(\xi - \zeta))_{\xi, \zeta \in X}$$

is strictly positive definite.

The guaranteed invertibility of  $K_X$  is of use in solving interpolation problems; given  $\mathbf{y} \in \mathbb{R}^X$ , one finds  $\mathbf{a} \in \mathbb{R}^X$  so that  $K_X \mathbf{a} = \mathbf{y}$ . It follows that  $\sum_{\xi \in X} a_\xi \kappa_m(\cdot - \xi)$  is the unique interpolant to  $(\xi, y_\xi)_{\xi \in X}$  in  $S(X) := \text{span}_{\xi \in X} \kappa_m(\cdot - \xi)$ .

It is the reproducing kernel for the Hilbert space  $\mathcal{N}(\kappa_m) = W_2^m(\mathbb{R}^d)$  equipped with the (standard) inner product

$$\begin{aligned} \langle f, g \rangle_{W_2^m(\mathbb{R}^d)} &= \int_{\mathbb{R}^d} \sum_{j=0}^m \binom{m}{j} \langle D^j f(x), D^j g(x) \rangle dx \\ &= \sum_{|\beta| \leq m} \binom{m}{|\beta|} \binom{|\beta|}{\beta} \int_{\mathbb{R}^d} D^\beta f(x) D^\beta g(x) dx, \end{aligned}$$

where  $D^j f$  is the tensor (i.e., the  $j$ -dimensional array) of partial derivatives of order  $j$ . Being the reproducing kernel means simply that  $f(x) = \langle f, \kappa_m(x - \cdot) \rangle_{W_2^m(\mathbb{R}^d)}$  for all  $x \in \mathbb{R}^d$  and all  $f \in W_2^m(\mathbb{R}^d)$ . It can be shown that among all functions interpolating the data  $(\xi, y_\xi)_{\xi \in X}$ , the interpolant  $\sum_{\xi \in X} a_\xi \kappa_m(\cdot - \xi)$  (i.e., where  $\mathbf{a}$  is the solution of  $K_X \mathbf{a} = \mathbf{y}$ ) has the smallest  $W_2^m(\mathbb{R}^d)$  norm.

2.5.2. *Surface splines.* For  $m > d/2$ , the *surface spline* is

$$(2.6) \quad \phi_m : \mathbb{R}^d \rightarrow \mathbb{R} : x \mapsto C \begin{cases} |x|^{2m-d}, & d \text{ is odd,} \\ |x|^{2m-d} \log |x|, & d \text{ is even.} \end{cases}$$

The surface spline of order  $m$  is conditionally positive definite (CPD) with respect to  $\Pi_{m-1}$ , the space of polynomials of degree  $m - 1$ . This means that the quadratic form  $\mathbb{R}^X \rightarrow \mathbb{R} : \mathbf{a} \mapsto \langle \mathbf{a}, K_X \mathbf{a} \rangle = \sum_{\xi \in X} \sum_{\zeta \in X} \phi_m(\xi - \zeta) a_\xi a_\zeta$  is positive for every finite set  $X \subset \mathbb{R}^d$  and for all nonzero  $\mathbf{a} \in \mathbb{R}^X$  which satisfy  $\sum_{\xi \in X} a_\xi p(\xi) = 0$  for all  $p \in \Pi_{m-1}$ . (In other words, it is positive definite on a subspace of  $\mathbb{R}^X$  of finite codimension; namely, the annihilator of  $\Pi_{m-1} |_{X \cdot}$ .)

One may solve interpolation problems using the finite dimensional space

$$S(X) := \left\{ \sum_{\xi \in X} a_\xi \phi_m(\cdot - \xi) \mid \sum_{\xi} a_\xi p(\xi) = 0 \text{ for all } p \in \Pi_{m-1} \right\} + \Pi_{m-1}$$

provided that data sites  $X$  are *unisolvent*: i.e., so that if  $p \in \Pi_{m-1}$  satisfies  $p(\xi) = 0$  for all  $\xi \in X$ , then  $p = 0$ . Let  $\{p_1, \dots, p_N\}$  be a basis for  $\Pi_{m-1}$  and construct the  $\#X \times N$  Vandermonde matrix  $\Phi = (p_j(\xi))_{\xi \in X, j=1, \dots, N}$ . For data  $\mathbf{y} \in \mathbb{R}^X$  one finds  $\mathbf{a} \in \mathbb{R}^X$  and  $\mathbf{c} \in \mathbb{R}^N$  so that

$$\begin{pmatrix} K_X & \Phi \\ \Phi^T & 0 \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \mathbf{c} \end{pmatrix} = \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \end{pmatrix}.$$

It follows that  $s_X := \sum_{\xi \in X} a_\xi \phi_m(\cdot - \xi) + \sum_{j=1}^N c_j p_j$  is the unique interpolant to  $(\xi, y_\xi)_{\xi \in X}$  in  $S(X)$ .

The surface spline  $\phi_m$  is the reproducing kernel for the semi-Hilbert space

$$D^{-m}L_2 = \{f \in C(\mathbb{R}^d) \mid \forall |\alpha| = m, D^\alpha f \in L_2(\mathbb{R}^d)\}$$

(sometimes called the Beppo–Levi space), which is a semi-Hilbert space (a vector space having a semidefinite inner product with nullspace  $\Pi_{m-1}$ , so that the quotient  $D^{-m}L_2/\Pi_{m-1}$  is a Hilbert space). The space  $D^{-m}L_2$  is endowed with the semidefinite product

$$\langle f, g \rangle_{D^{-m}L_2} = \int_{\mathbb{R}^d} \langle D^m f(x), D^m g(x) \rangle dx = \sum_{\beta=m} \binom{m}{\beta} \int_{\mathbb{R}^d} D^\beta f(x) D^\beta g(x) dx.$$

Although  $\phi_m \notin D^{-m}L_2$  (its  $m$ th derivatives behave, roughly, like  $\mathcal{O}(|x|^{m-d})$ , which is not square integrable, since  $2m > d$ ), with a little effort, one may show that the spaces  $S(X)$  are contained in  $D^{-m}L_2$ . The RBF  $\phi_m$  is its reproducing kernel in the sense that for  $X \subset \mathbb{R}^d$  and two functions  $f_1, f_2 \in D^{-m}L_2$  where  $f_2$  has the form  $f_2 = \sum_{\xi \in X} a_\xi \phi(\cdot - \xi) + p \in S(X)$  we have  $\langle f_1, f_2 \rangle_{D^{-m}L_2} = \sum_{\xi \in X} a_\xi f_1(\xi)$ . The interested reader will find a material on surface splines and conditionally positive definite RBFs in [32, Chapter 8].

As in the case of Matérn kernels, the unique interpolant residing in  $S(X)$  has the smallest  $D^{-m}L_2$  seminorm among all interpolants to the data  $(\xi, y_\xi)_{\xi \in X}$ .

**2.5.3. Labeling kernels.** In most cases in this article, the Matérn and surface spline RBFs exhibit similar behaviors. Because our results often depend only on a single parameter  $m$  indexing the RBF, we use the abbreviated notation  $k_m$  to stand for either  $\kappa_m$  or  $\phi_m$ .

In both cases, the function  $k_m$  has  $L_p$  smoothness less than  $2m - d + d/p$  (i.e., for any bounded set  $\Omega$ ,  $k_m \in W_p^\sigma(\Omega)$  for all  $\sigma < 2m - d + d/p$ ). It follows that any finite linear combination of shifts of  $k_m$  has the same regularity. Denote the space of such linear combinations as

$$S(X) := \left\{ \begin{array}{l} \text{span}_{\xi \in X} \kappa_m(\cdot - \xi), \\ \left\{ \sum_{\xi \in X} a_\xi \phi_m(\cdot - \xi) \mid \sum_{\xi} a_\xi p(\xi) = 0 \ \forall p \in \Pi_{m-1} \right\} + \Pi_{m-1}, \end{array} \right. \quad \begin{array}{l} k_m = \kappa_m, \\ k_m = \phi_m. \end{array}$$

Then we have

$$S(X) \subset W_p^\sigma(\Omega) \quad \text{for all} \quad \sigma < 2m - d(1 - 1/p).$$

Likewise, we let  $\mathcal{N}(k_m)$  represent either of the two native spaces:  $W_2^m(\mathbb{R}^d)$  or  $D^{-m}L_2(\mathbb{R}^d)$ . We note that both satisfy the continuous embedding  $W_2^m(\mathbb{R}^d) \subset \mathcal{N}(k_m) \subset W_{2,loc}^m(\mathbb{R}^d)$ . In this case, the functions in the native space have a lower  $L_p$  regularity, with

$$\mathcal{N}(k_m) \subset \begin{cases} W_p^s(\Omega) & \text{for } s \leq m - (d/2 - d/p)_+, \quad 1 \leq p < \infty, \\ C^s(\bar{\Omega}) & \text{for } s < m - d/2. \end{cases}$$

### 3. LAGRANGE FUNCTIONS AND FIRST BERNSTEIN INEQUALITIES

In this section we investigate some further results about the RBF  $k_m$ ; namely, we consider analytic properties of the Lagrange functions. These have been presented in [17], but we explain them below for the sake of completeness.

After this we give a first class of Bernstein estimates, valid for linear combinations of Lagrange functions.

**3.1. Lagrange functions.** For a finite  $X \subset \mathbb{R}^d$ , there exists a family of (uniquely defined) functions  $(\chi_\xi)_{\xi \in X}$  satisfying  $\chi_\xi \in S(X)$  and  $\chi_\xi(\zeta) = \delta(\xi, \zeta)$  for all  $\zeta \in X$ . We may take the  $\mathcal{N}(k_m)$  inner product of two Lagrange functions  $\chi_\xi, \chi_\zeta \in S(X)$ , noting that they have the form  $\chi_\xi = \sum_{\eta \in X} A_{\eta,\xi} k_m(\cdot - \eta) + p$  and  $\chi_\zeta = \sum_{\eta \in X} A_{\eta,\zeta} k_m(\cdot - \eta) + \tilde{p}$  (in the case of Matérn functions  $k_m = \kappa_m$ , we have  $p = \tilde{p} = 0$ ), to obtain

$$(3.1) \quad \langle \chi_\xi, \chi_\zeta \rangle_{\mathcal{N}(k_m)} = \langle \chi_\xi, \sum_{\eta \in X} A_{\eta,\zeta} k_m(\cdot - \eta) + \tilde{p} \rangle_{\mathcal{N}(k)} = \sum_{\eta \in X} A_{\eta,\zeta} \chi_\xi(\eta) = A_{\xi,\zeta}.$$

*Lagrange function coefficients.* We can make the following ‘bump estimate’ which uses a bump function  $\psi_{\xi,q} = \psi(\frac{\cdot-\xi}{q}) : \mathbb{R}^d \rightarrow [0,1]$  that is compactly supported in  $B(\xi,q)$  and satisfies  $\psi_{\xi,q}(\xi) = 1$  on a neighborhood of  $q$ . We have

$$(3.2) \quad \|\chi_\xi\|_{\mathcal{N}(k_m)} \leq \|\psi_{\xi,q}\|_{\mathcal{N}(k_m)} \leq C\|\psi_{\xi,q}\|_{W_2^m(\mathbb{R}^d)} \leq Cq^{\frac{d}{2}-m}.$$

This follows because  $\chi_\xi$  is the best interpolant to  $\zeta \rightarrow \delta(\xi,\zeta)$ . As a consequence, Lagrange coefficients are uniformly bounded:

$$(3.3) \quad |A_{\xi,\zeta}| = |\langle \chi_\xi, \chi_\zeta \rangle_{\mathcal{N}(k)}| \leq Cq^{d-2m}.$$

*Better decay.* For the kernels considered in this article, and more generally for the framework given in [17] and [18], to get desired estimates for Lagrange functions over a compact region  $\Omega \subset \mathbb{R}^d$  the interpolatory conditions must be satisfied on a point set that is suitably dense in a fairly large neighborhood of  $\Omega$ . To handle this, we use the quasi-uniform extension  $\tilde{\Xi}$  developed in Section 2.2. This brings us to the definition of  $V_{\Xi}$ .

**Definition 3.1.** For a compact set  $\Omega \subset \mathbb{R}^d$  and a finite subset  $\Xi \subset \Omega$ , let  $\tilde{\Xi}$  be the extension to  $\{x \in \mathbb{R}^d \mid \text{dist}(x,\Omega) \leq \text{diam}(\Omega)\}$  given in (2.3) in Section 2.3. Then for the system of Lagrange functions  $(\chi_\xi)_{\xi \in \tilde{\Xi}}$  generated by  $k_m$  over  $\tilde{\Xi}$ , let

$$V_{\Xi} := \text{span}\{\chi_\xi \mid \xi \in \Xi\}.$$

In particular  $V_{\Xi} \subset S(\tilde{\Xi})$ .

For  $\xi$  in the original set  $\Xi$ , we have the improved estimates:

$$(3.4) \quad \|\chi_\xi\|_{W_2^m(\mathbb{R}^d \setminus B(\xi,R))} \leq Cq^{d/2-m} \exp\left(-\mu \frac{R}{h}\right) \text{ for all } 0 < R < \text{dist}(\xi, \partial\Omega).$$

This is demonstrated in Appendix A, specifically in Lemma A.3. This leads to improved estimates. For  $\xi \in \Xi$  and all  $x \in \Omega$ ,

$$(3.5) \quad |\chi_\xi(x)| \leq C\rho^{m-d/2} \exp\left(-\mu \frac{|x-\xi|}{h}\right).$$

Likewise, for  $\xi, \zeta \in \Xi$ ,

$$(3.6) \quad |A_{\xi,\zeta}| \leq Cq^{d-2m} \exp\left(-\nu \frac{|\xi-\zeta|}{h}\right).$$

**3.2. Stability of the Lagrange-function basis for  $V_{\Xi}$  on  $\Omega$ .** Recall that  $V_{\Xi} = \text{span}\{\chi_\xi\}_{\xi \in \Xi}$ , where  $\Xi$  is a subset of all of the centers in  $\tilde{\Xi}$ . We begin by defining the operator  $T : \mathbb{C}^{\Xi} \rightarrow V_{\Xi}$  by  $T\mathbf{a} = \sum_{\xi \in \Xi} a_\xi \chi_\xi =: s$ . In other words,  $T$  is the *synthesis* operator, which takes a set of coefficients  $\{a_\xi\}_{\xi \in \Xi}$  and outputs a function  $s \in V_{\Xi}$ . Because the basis in consideration is the Lagrange basis, the coefficients satisfy  $T\mathbf{a}(\xi) = s(\xi) = a_\xi$  for  $\xi \in \Xi$  and therefore, for the basis  $\{\chi_\xi\}_{\xi \in \Xi}$ , the operator  $T$  is an interpolation operator.

If we use the  $\ell_p(\Xi)$  norm for  $\mathbb{C}^{\Xi}$  and  $L_p(\Omega)$  for  $V_{\Xi}$ , then the stability of the basis, relative to these norms, is measured by comparing  $\|\mathbf{a}\|_{\ell_p(\Xi)}$  and  $\|s\|_{L_p(\Omega)}$ . We show this with the following proposition, which indicates that if  $s \in V_{\Xi}$  is small (relative to  $L_p(\Omega)$ ), then its coefficients are small in  $\ell_p(\Xi)$  (and likewise, if the coefficients of  $s$  are small in  $\ell_p(\Xi)$ , so too is the norm of  $s \in L_p(\Omega)$ ).

**Proposition 3.2** (Lagrange Basis Stability). *There exists a constant  $h_0 = h_0(m, d)$ , so that for  $\Xi \subset \Omega$  satisfying  $h(\Xi, \Omega) < h_0$  and  $1 \leq p \leq \infty$ , we have constants  $c_1 = c_1(m, \rho, \Omega) > 0$  and  $c_2 = c_2(m, \Omega)$  so that*

$$(3.7) \quad c_1 \|\mathbf{a}\|_{\ell_p(\Xi)} \leq q^{-d/p} \|\sum_{\xi \in \Xi} a_\xi \chi_\xi\|_{L_p(\Omega)} \leq c_2 \rho^{m+d/p} \|\mathbf{a}\|_{\ell_p(\Xi)}.$$

We remark that the dependence in the lower constant  $c_1$  on  $\rho$  can be made explicit. This is worked out in Lemma B.6.

*Proof.* The proof is given in Appendix B. Lemma B.1 provides the upper bound and Lemma B.6 gives the lower bound. □

Another way to think of this inequality is as an  $L_p(\Omega)$  Marcinkiewicz-Zygmund (MZ) inequality. Such inequalities are used to relate the  $L_1$  norm of a trigonometric polynomial to the  $\ell_1$  norm of the polynomial evaluated on some fixed, finite set. MZ inequalities have also been developed for spherical polynomials on  $\mathbb{S}^d$  [23]. For spherical polynomials in  $\mathbb{S}^d$ , there is another type of inequality, a Nikolskii inequality. On  $\mathbb{S}^d$ , these have the form  $\|S\|_{L_p} \leq CL^{d(\frac{1}{r}-\frac{1}{p})_+} \|S\|_{L_r(\mathbb{S}^d)}$ , for any degree  $L$  spherical polynomial. Our result below establishes such an inequality for  $V_\Xi$ .

**Corollary 3.3** (Nikolskii Inequality). *With the assumptions and notation of Proposition 3.2, and with  $1 \leq p, r \leq \infty$ , we have that*

$$(3.8) \quad \|s\|_{L_p(\Omega)} \leq Cq^{-d(\frac{1}{r}-\frac{1}{p})_+} \|s\|_{L_r(\Omega)}, \quad s \in V_\Xi,$$

with  $C = C(m, \rho, \Omega, p, r)$ .

*Proof.* Recall that, for  $\mathbf{a} \in \mathbb{C}^\Xi$ ,  $\|\mathbf{a}\|_{\ell_p(\Xi)} \leq N^{(\frac{1}{p}-\frac{1}{r})_+} \|\mathbf{a}\|_{\ell_r(\Xi)}$ , where  $N = \#\Xi$ . Since  $N \sim q^{-d}$ , this inequality implies that  $\|\mathbf{a}\|_{\ell_p(\Xi)} \leq C_{\Omega, r, p} q^{-d(\frac{1}{p}-\frac{1}{r})_+} \|\mathbf{a}\|_{\ell_r(\Xi)}$ . From this and (3.7), we thus have

$$\|s\|_{L_p(\Omega)} \leq C_{\rho, m, \Omega, p, r} q^{d(\frac{1}{p}-\frac{1}{r}-\frac{1}{p})_+} \|\mathbf{a}\|_{\ell_r(\Xi)} \leq C_{\rho, m, \Omega, p, r} q^{d(\frac{1}{p}-\frac{1}{r}-\frac{1}{p}-\frac{1}{r})_+} \|s\|_{L_r(\Omega)}.$$

The result follows from the identity  $x - (x)_+ = -(-x)_+$ . □

**3.3. Bernstein type estimates for (full) Lagrange functions.** In this section we will provide a Bernstein (or inverse) theorem relating Sobolev norms of functions in  $V_\Xi$  to the corresponding  $\ell_p$  norms on the coefficients. This is the key to controlling the Sobolev norm of the function in  $V_\Xi$  by its  $L_p(\Omega)$  norm.

Before proceeding, we first prove two lemmas.

**Lemma 3.4.** *Let  $1 \leq p \leq \infty$ ,  $s > d/2$  and  $0 \leq \sigma \leq s$ . Suppose  $O$  is a fixed open set with Lipschitz boundary and as before  $O_R = \{x \mid x/R \in O\}$ . Suppose further that  $W_2^s(O)$  is embedded continuously in  $W_p^\sigma(O)$  (where we take  $C^\sigma$  in case  $p = \infty$ ). Then there is a constant  $C$  depending on  $O$ ,  $p$ ,  $s$  and  $\sigma$  so that if  $U \in W_p^\sigma(O)$  and if the set  $X$  of zeros of  $U$  in  $O$  are sufficiently dense that  $h(X, O) < 1$  and Lemma A.1 applies, then for  $u(x) = U(x/R)$ , we have*

$$|u|_{W_p^\sigma(O_R)} \leq CR^{s-\sigma+d/p-d/2} |u|_{W_2^s(O_R)}.$$

*Proof.* Lemma 2.4 shows that  $|u|_{W_p^\sigma(O_R)} = R^{d/p-\sigma}|U|_{W_p^\sigma(O)}$ . Because  $\|U\|_{W_p^\sigma(O)} \leq C\|U\|_{W_2^s(O)}$ , we have that

$$|u|_{W_p^\sigma(O_R)} \leq CR^{d/p-\sigma}\|U\|_{W_2^s(O)} = CR^{d/p-\sigma} \left( \sum_{j=0}^k \binom{k}{j} |U|_{W_2^j(O)}^2 + |U|_{W_2^s(O)}^2 \right)^{1/2}$$

(in case  $s$  is fractional; we leave the necessary modification to the reader in case  $s \in \mathbb{N}$ ). We now apply the zeros estimate [26, Theorem 4.2] to each term on the right-hand side, obtaining

$$|u|_{W_p^\sigma(O_R)} = CR^{d/p-\sigma}|U|_{W_2^s(O)},$$

since  $|U|_{W_2^j(O)}^2 \leq Ch^{2(s-j)}|U|_{W_2^s(O)}^2$  for all  $j$ . Applying Lemma 2.4 again, (this time with  $p = 2$ ), yields the desired estimate.  $\square$

**Lemma 3.5.** *Suppose  $\Xi$  is sufficiently dense (meaning  $h(\Xi, \Omega) \leq h_0$  for a constant  $h_0 = h_0(d, m) > 0$ ) and  $\eta \in \Xi$ . Decompose  $\Xi$  into disjoint annuli  $\Xi = \bigcup_{j=0}^\infty \Xi_j(\eta)$ : where  $\Xi_j(\eta) := \{\zeta \in \Xi \mid 2^{j-1}h \leq \text{dist}(\zeta, \eta) \leq 2^j h\}$  for  $j > 0$  and  $\Xi_0(\eta) := \{\zeta \in \Xi \mid 0 \leq \text{dist}(\zeta, \eta) \leq h\}$ .*

*We have, for  $2 \leq p < \infty$  and  $0 \leq \sigma \leq m - d/2 + d/p$  that*

$$(3.9) \quad \left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{W_p^\sigma(B(\eta, 3h))}^p \leq Cp^{p(m+d/2)-d} h^{d-p\sigma} \sum_{j=0}^\infty 2^{j(d+1)(p-1)} e^{-\nu p 2^{j-1}} \sum_{\xi \in \Xi_j(\eta)} |a_\xi|^p$$

with  $C = C(p, \sigma, m, d)$  and  $\nu = \nu(m, d)$ .

*Proof.* Repeatedly applying the quasi-triangle inequality

$$\|a + b\|^p \leq 2^{p-1}(\|a\|^p + \|b\|^p)$$

to this sum gives

$$(3.10) \quad \left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{W_p^\sigma(B(\eta, 3h))}^p \leq \sum_{j=0}^\infty 2^{(j+1)(p-1)} \left\| \sum_{\xi \in \Xi_j(\eta)} a_\xi \chi_\xi(x) \right\|_{W_p^\sigma(B(\eta, 3h))}^p.$$

Observe that  $\#\Xi_j(\eta) \leq \omega_d \rho^{d 2^{jd}}$  (where the constant  $\omega_d$  depends on  $d$ ), so the generalization of the above quasi-triangle inequality  $\left\| \sum_{j=1}^n a_j \right\|^p \leq n^{p-1} \sum_{j=1}^n \|a_j\|^p$  gives

$$(3.10) \quad \left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{W_p^\sigma(B(\eta, 3h))}^p \leq \sum_{j=0}^\infty 2^{(j+1)(p-1)} (\omega_d \rho^{d 2^{jd}})^{p-1} \sum_{\xi \in \Xi_j(\eta)} |a_\xi|^p \|\chi_\xi\|_{W_p^\sigma(B(\eta, 3h))}^p.$$

For  $\text{dist}(\xi, \eta) = R$  sufficiently large, we have the inclusion

$$B(\eta, 3h) \subset \tilde{\Omega} \setminus B(\xi, R - 3h).$$

Applying the zeros lemma [25, Theorem 1.1] gives

$$\|\chi_\xi\|_{W_p^\sigma(B(\eta, 3h))} \leq \|\chi_\xi\|_{W_p^\sigma(\tilde{\Omega} \setminus B(\xi, R-3h))} \leq Ch^{m-\sigma-(\frac{d}{2}-\frac{d}{p})_+} \|\chi_\xi\|_{W_2^m(\tilde{\Omega} \setminus B(\xi, R-3h))}.$$

Applying the energy estimate (3.4) and noting that  $d/2 - d/p \geq 0$  gives us

$$\|\chi_\xi\|_{W_p^\sigma(B(\eta, 3h))} \leq Ch^{m-\sigma-(d/2-d/p)_+} q^{d/2-m} e^{-\nu R/h} = Cp^{m-d/2} h^{d/p-\sigma} e^{-\nu R/h}.$$

We note that for  $\xi$  in the annular set  $\Xi_j(\eta)$ ,  $\text{dist}(\xi, \eta) = R \geq h2^{j-1}$ , so  $e^{-\nu R/h} \leq e^{-\frac{\nu}{2}2^j}$ . Applying this to (3.10) gives

$$\begin{aligned}
 (3.11) \quad & \left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{W_\rho^\sigma(B(\eta, 3h))}^p \\
 & \leq C \omega_d^{p-1} \rho^{p(m-d/2)+d(p-1)} \sum_{j=0}^\infty 2^{j(d+1)(p-1)} \sum_{\xi \in \Xi_j(\eta)} |a_\xi|^p h^{d-p\sigma} e^{-\frac{\nu}{2}p2^j} \\
 & \leq C \rho^{p(m+d/2)-d} h^{d-p\sigma} \sum_{j=0}^\infty 2^{j(d+1)(p-1)} e^{-\frac{\nu}{2}p2^j} \sum_{\xi \in \Xi_j(\eta)} |a_\xi|^p. \quad \square
 \end{aligned}$$

Note that when  $p = \infty$ , we use only integer smoothness  $\sigma = k \in \mathbb{Z}$  and the standard space  $C^k(\bar{\Omega})$  of  $k$  times integral functions over  $\bar{\Omega}$ .

**Theorem 3.6.** *For a sufficiently dense set  $\Xi$  (meaning  $h(\Xi, \Omega) \leq h_0$  for a constant  $h_0 = h_0(d, m) > 0$ ) we have, for  $0 \leq \sigma \leq m - (d/2 - d/p)_+$  when  $1 \leq p < \infty$  (or  $\sigma \in \mathbb{N}$  with  $0 \leq \sigma < m - d/2$  if  $p = \infty$ ),*

$$(3.12) \quad \left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{W_\rho^\sigma(\Omega)} \leq C \rho^{m+d/2+d/p} h^{d/p-\sigma} \|\mathbf{a}\|_{\ell_p(\Xi)}$$

with  $C = C(p, \sigma, m, d)$ .

*Proof.* This is handled in four cases:  $p = \infty$ ,  $2 \leq p < \infty$ ,  $p = 1$  and  $1 < p < 2$ .

*Case 1:*  $p = \infty$ . If  $\sigma \in \mathbb{Z}$ , we simply need to bound  $\max_{|\alpha|=\sigma} \sum_{\xi \in \Xi} \|D^\alpha \chi_\xi\|_\infty$ . To do this, consider a point  $x \in \Omega$  and a ball  $B(x, r) \subset \bar{\Omega}$  with  $r = h \max(16m^2, 1/h_1)$  and  $h_1$  is the constant from the zeros Lemma A.1. In this case, we use a Bramble-Hilbert argument involving the averaged Taylor polynomial  $Q^m \chi_\xi$  of degree  $m - 1$  described in Brenner-Scott [1]. The estimate

$$\|D^\alpha Q^m \chi_\xi\|_{L_\infty(B(x,r))} \leq C r^{m-|\alpha|-d/2} |\chi_\xi|_{W_2^m(B(x,r))}$$

holds by applying [17, (3.9)]. In order to bound  $\|D^\alpha(\chi_\xi - Q^m \chi_\xi)\|_{L_\infty(B(x,r))}$ , we use Lemma 2.4 (with  $U((y - \xi)/r) = \chi_\xi(y) - Q^m \chi_\xi(y)$  and the embedding  $W_2^{m-|\alpha|}(B(0, 1)) \subset C(B(0, 1))$ ) to obtain

$$|(\chi_\xi - Q^m \chi_\xi)|_{C^{|\alpha|}(B(x,r))} \leq C r^{-|\alpha|} |U|_{C^{|\alpha|}(B(0,1))} \leq C r^{-|\alpha|} |U|_{W_2^m(B(0,1))}.$$

Rescaling gives the estimate

$$\|D^\alpha(\chi_\xi - Q^m \chi_\xi)\|_{L_\infty(B(x,r))} \leq C r^{-|\alpha|} \left( \sum_{j=0}^m r^{2j-d} \binom{m}{j} |\chi_\xi - Q^m \chi_\xi|_{W_2^j(B(x,r))}^2 \right)^{1/2}.$$

Each seminorm in this last expression can be estimated by the Bramble-Hilbert Lemma, allowing us to bound the above by  $C r^{m-|\alpha|-d/2} |\chi_\xi|_{W_2^m(B(x,r))}$ . Together with the estimate on  $D^{|\alpha|} Q^m \chi_\xi$ , and recalling that  $r = Kh$ , we have

$$\begin{aligned}
 |D^\alpha \chi_\xi(x)| & \leq \|D^\alpha Q^m \chi_\xi\|_{L_\infty(B(x,r))} + \|D^\alpha(\chi_\xi - Q^m \chi_\xi)\|_{L_\infty(B(x,r))} \\
 & \leq C h^{m-|\alpha|-d/2} |\chi_\xi|_{W_2^m(B(x,r))}.
 \end{aligned}$$

From here, we apply the energy estimate (3.4) to obtain

$$(3.13) \quad |D^\alpha \chi_\xi(x)| \leq C h^{m-|\alpha|-\frac{d}{2}} q^{\frac{d}{2}-m} e^{-\nu \frac{|x-\xi|}{h}} = C \rho^{m-\frac{d}{2}} h^{-|\alpha|} e^{-\nu \frac{|x-\xi|}{h}}.$$

The sum over  $\Xi$  can be carried out over annular regions  $\Xi_j(x) = \{\xi \in \Xi \mid jh \leq \text{dist}(\xi, x) < (j+1)h\}$ . This leaves

$$\begin{aligned} \sum_{j=0}^{\infty} \sum_{\xi \in \Xi_j(x)} |D^\alpha \chi_\xi(x)| &\leq C\rho^{m-\frac{d}{2}}h^{-|\alpha|} \sum_{j=0}^{\infty} \sum_{\xi \in \Xi_j(x)} e^{-\nu\frac{|x-\xi|}{h}} \\ &\leq C\rho^{m-\frac{d}{2}}h^{-|\alpha|} \sum_{j=0}^{\infty} \rho^d(j+1)^d e^{-\nu j} \\ &\leq C\rho^{m+\frac{d}{2}}h^{-|\alpha|}. \end{aligned}$$

In the last inequality, we use the fact that the sum  $\sum_{j=0}^{\infty} (j+1)^d e^{-\nu j} = C$  depends on  $d$  and  $m$  (but not  $\rho$ ).

*Case 2:*  $2 \leq p < \infty$ . We treat this case in two stages. At first, we consider  $\sigma = k \in \mathbb{N}$ , treating fractional Sobolev exponents for later.

*Case 2i:*  $\sigma = k \in \mathbb{N}$ . By subadditivity, the Sobolev norm may be taken over overlapping balls

$$\left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{W_p^k(\Omega)}^p \leq \sum_{\eta \in \Xi} \left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi(x) \right\|_{W_p^k(B(\eta, h))}^p.$$

Applying Lemma 3.5 to the norm over each ball  $B(\eta, h) \subset B(\eta, 3h)$ . gives

$$\left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{W_p^k(\Omega)}^p \leq C\rho^{pm+pd/2-d}h^{d-p\sigma} \sum_{j=0}^{\infty} 2^{j(d+1)(p-1)} e^{-\nu p 2^{j-1}} \sum_{\eta \in \Xi} \sum_{\xi \in \Xi_j(\eta)} |a_\xi|^p.$$

We may exchange summation between  $\xi$  and  $\eta$ , noting that  $\eta \in \Xi_j(\xi)$  iff  $\xi \in \Xi_j(\eta)$ . This implies the estimate  $\sum_{\eta \in \Xi} \sum_{\xi \in \Xi_j(\eta)} |a_\xi|^p = \sum_{\xi \in \Xi} \sum_{\eta \in \Xi_j(\xi)} |a_\xi|^p \leq \omega_d \rho^d 2^{jd} \sum_{\xi \in \Xi} |a_\xi|^p$ . Consequently,

$$(3.14) \quad \left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{W_p^k(\Omega)}^p \leq C\omega_d^{p-1} \rho^{p(m+d/2)} h^{-pk+d} \left( \sum_{j=0}^{\infty} 2^{j(d+1)p} e^{-\frac{\nu}{2} p 2^j} \right) \sum_{\xi \in \Xi} |a_\xi|^p.$$

The result follows by summing the series and taking the  $p$ th root.

*Case 2ii:*  $\sigma \notin \mathbb{N}$ . Let  $\sigma = k + \delta$  with  $0 < \delta < 1$ , and employ Lemma 2.5, using the neighborhoods  $\{B(\eta, h) \cap \Omega \mid \eta \in \Xi\}$  as  $\{\tilde{v}_j \mid j \in \mathcal{N}\}$  and  $\{B(\eta, 3h) \cap \Omega \mid \eta \in \Xi\}$  for  $\{v_j \mid j \in \mathcal{N}\}$ . Note that for this choice of cover,  $M \leq C\rho^d$ . Indeed, for any  $x \in \Omega$ , enumerate the centers  $\{\xi \in \Xi \mid \text{dist}(\xi, x) < h\}$  as  $\xi_1, \dots, \xi_n$ . Then  $x \in B(\xi_j, h)$  for each  $j = 1, \dots, n$ . Because the balls  $B(\xi_j, q)$  are disjoint, it follows that  $n(c_{d,1}q^d) = \text{vol}(\bigcup_{j=1}^n B(x_j, q) \cap B(x, h)) \leq c_{d,2}h^d$ , so  $n \leq C_d(h/q)^d$ . (Here  $c_{d,2}h^d$  is the volume of the ball of radius  $h$ , and  $c_{d,1}q^d$  is the volume of the portion in  $B(x, h)$  of any ball which is centered in  $B(x, h)$ .) This guarantees that

$$\left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{W_p^\sigma(\Omega)}^p \leq \sum_{\eta \in \Xi} \left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{W_p^\sigma(B(\eta, 3h))}^p + C\rho^d h^{-p\delta} \left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{W_p^k(\Omega)}^p.$$

We apply (3.14) to bound the second term, which gives

$$C\rho^d h^{-p\delta} \left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{W_p^k(\Omega)}^p \leq C\rho^{p(m+d/2)+d} h^{-pk+d-p\delta} |a_\xi|^p.$$



The first term is handled precisely as the integer case  $\sigma = k$ , which has been discussed above, leaving

$$\sum_{\eta \in \Xi} \left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{W_\rho^\sigma(B(\eta, 3h))}^p \leq C \rho^{p(m+d/2)} h^{d/p-\sigma} \|\mathbf{a}\|_{\ell_p(\Xi)}.$$

*Case 3.* We again consider the proof in two steps, first for the case of integer smoothness, where the Sobolev norm is subadditive on sets, and then in the fractional case, where we can apply Lemma 2.5.<sup>4</sup>

As an initial simplification, note that the triangle inequality gives  $\|s\|_{W_1^\sigma(\Omega)} \leq \|\mathbf{a}\|_{\ell_1(\Xi)} (\max_{\xi \in \Xi} \|\chi_\xi\|_{W_1^\sigma(\Omega)})$ , so we need only to consider the size of  $\|\chi_\xi\|_{W_1^\sigma(\Omega)}$ .

*Case 3i:*  $\sigma = k \in \mathbb{N}$ . We proceed, as in Case 2, by first considering  $\sigma = k \in \mathbb{N}$  and using subadditivity of the norm. For any integer  $K$ , we have  $\|\chi_\xi\|_{W_1^k(\Omega)} \leq \|\chi_\xi\|_{W_1^k(B(\xi, Kh))} + \|\chi_\xi\|_{W_1^k(\Omega \setminus B(\xi, Kh))}$ .

The first term satisfies  $\|\chi_\xi\|_{W_1^k(B(\xi, Kh))} \leq \omega_d(Kh)^{d/2} \|\chi_\xi\|_{W_2^k(B(\xi, Kh))}$ . For  $K$  sufficiently large (a constant depending only on  $d$ ), the zeros estimate [26, Theorem 4.2] gives

$$(3.15) \quad \|\chi_\xi\|_{W_1^k(B(\xi, Kh))} \leq \omega_d(Kh)^{d/2} h^{m-k} \|\chi_\xi\|_{W_2^m(\tilde{\Omega})} \leq CK^{d/2} \rho^{m-d/2} h^{d-k}.$$

The second term may be controlled by decomposing  $\Omega \setminus B(\xi, Kh) = \bigcup_{\ell=K}^\infty A_\ell$  in annuli (i.e., by taking  $A_\ell := \{x \in \Omega \mid \ell h \leq \text{dist}(x, \xi) \leq (\ell + 1)h\}$ ). Subadditivity gives

$$\begin{aligned} \|\chi_\xi\|_{W_1^k(\Omega \setminus B(\xi, Kh))} &\leq \sum_{\ell=K}^\infty \|\chi_\xi\|_{W_1^k(A_\ell)} \\ &\leq \sum_{\ell=K}^\infty (\text{vol}(A_\ell))^{1/2} \|\chi_\xi\|_{W_2^k(A_\ell)} \\ &\leq \sum_{\ell=K}^\infty C((\ell + 1)h)^{d/2} h^{m-\sigma} \|\chi_\xi\|_{W_2^m(A_\ell)}. \end{aligned}$$

In the final line we have applied the zeros estimate (and simultaneously estimated the volume of the annulus  $A_\ell$ ). At this point, we can apply the energy estimate (3.4) to obtain

$$(3.16) \quad \|\chi_\xi\|_{W_1^k(\Omega \setminus B(\xi, Kh))} \leq \sum_{\ell=K}^\infty C((\ell + 1)h)^{d/2} h^{m-\sigma} q^{d/2-m} e^{-\nu \ell} \leq C \rho^{m-d/2} h^{d-\sigma}.$$

Combining (3.16) with (3.15), gives the desired result for  $\sigma = k \in \mathbb{N}$ .

*Case 3ii:*  $\sigma \notin \mathbb{N}$ . To handle the fractional case  $\sigma = k + \delta$ , we apply Lemma 2.5, with an initial decomposition  $\tilde{v}_1 = B(\xi, Kh)$ ,  $\tilde{v}_2 = \Omega \setminus B(\xi, Kh)$ ,  $v_1 = B(\xi, (K + 1)h)$  and  $v_2 = \Omega \setminus B(\xi, (K - 1)h)$ . Observe that these are disjoint, so the overlap constant is  $M = 1$ . Thus we have

$$|\chi_\xi|_{W_1^\sigma(\Omega)} \leq |\chi_\xi|_{W_1^\sigma(B(\xi, (K+1)h))} + |\chi_\xi|_{W_1^\sigma(\Omega \setminus B(\xi, (K-1)h))} + Ch^{-\delta} \|\chi_\xi\|_{W_1^k(\Omega)}.$$

---

<sup>4</sup>In this case, because  $0 \leq \sigma \leq m$ , one could just as easily adopt the strategy of proving the result for the extrema  $\sigma = 0$  and  $\sigma = m$ , and then using interpolation of operators to bound the synthesis operator  $T : \ell_1(\Xi) \rightarrow W_1^\sigma(\Omega)$ , noting that  $W_1^\sigma(\Omega)$  is the  $(\sigma/m, 1)$  interpolation space between  $L_1(\Omega)$  and  $W_1^m(\Omega)$ .

We can further decompose the middle term  $|\chi_\xi|_{W_1^\sigma(\Omega \setminus B(\xi, (K-1)h))}$  in annuli by applying Lemma 2.5 a second time. This time, we let  $\tilde{v}_\ell := \{x \in \Omega \mid 2^\ell(K-1)h \leq \text{dist}(x, \xi) < 2^{\ell+1}(K-1)h\}$  for  $\ell = 0, 1, \dots$ . The annuli  $\{\tilde{v}_\ell \mid \ell \in \mathbb{N}\}$  partition  $\Omega \setminus B(\xi, (K-1)h)$ , so the overlap constant  $M$  is  $M = 1$ ; in fact, we need only the first  $\ell_0 = 1 + \log_2(\text{diam}(\Omega)/(K-1)h)$  annuli.

Define the neighborhoods of  $\tilde{v}_\ell$  as  $v_\ell := \{x \in \Omega \mid 2^{\ell-1}(K-1)h \leq \text{dist}(x, \xi) < 2^{\ell+2}(K-1)h\}$ . The sets  $w_\ell = \Omega \setminus v_\ell$  satisfy  $\text{dist}(\tilde{v}_\ell, w_\ell) \geq \frac{1}{2}(K-1)h \geq h$ . Lemma 2.5 shows that

$$|\chi_\xi|_{W_1^\sigma(\Omega \setminus B(\xi, (K-1)h))} \leq \left( \sum_{\ell=0}^{\ell_0} |\chi_\xi|_{W_1^\sigma(v_\ell)} \right) + Ch^{-\delta} \|\chi_\xi\|_{W_1^k(\Omega)}.$$

Since  $\|\chi_\xi\|_{W_1^\sigma(\Omega)} = \|\chi_\xi\|_{W_1^k(\Omega)} + |\chi_\xi|_{W_1^\sigma(\Omega)}$  and  $h < 1$ , this leaves

$$(3.17) \quad \|\chi_\xi\|_{W_1^\sigma(\Omega)} \leq |\chi_\xi|_{W_1^\sigma(B(\xi, (K+1)h))} + \left( \sum_{\ell=0}^{\ell_0} |\chi_\xi|_{W_1^\sigma(v_\ell)} \right) + Ch^{-\delta} \|\chi_\xi\|_{W_1^k(\Omega)},$$

which we must estimate.

*Estimating the third term in (3.17):* The final term is easiest to control: Case 3i gives the estimate

$$(3.18) \quad h^{-\delta} \|\chi_\xi\|_{W_1^k(\Omega)} \leq C\rho^{m-d/2} h^{d-k-\delta}.$$

*Estimating the first term in (3.17):* The first term in (3.17) is controlled in a similar way to (3.15). Begin by setting  $R := (K+1)h$ ,  $u = \chi_\xi(\cdot - \xi)$  and  $U = u(R(\cdot))$ . Applying Lemma 3.4 with  $O = B(0, 1)$  gives

$$\begin{aligned} |\chi_\xi|_{W_1^\sigma(B(\xi, (K+1)h))} &= |u|_{W_1^\sigma(B(0, R))} \\ &\leq CR^{d/2+m-\sigma} |u|_{W_2^m(B(0, R))} \\ &= Ch^{d/2+m-\sigma} |\chi_\xi|_{W_2^m(B(\xi, (K+1)h))}. \end{aligned}$$

The bump estimate (3.2) then gives

$$(3.19) \quad \begin{aligned} |\chi_\xi|_{W_1^\sigma(B(\xi, (K+1)h))} &\leq Ch^{d/2+m-\sigma} |\chi_\xi|_{W_2^m(\tilde{\Omega})} \\ &\leq Ch^{d/2+m-\sigma} q^{d/2-m} \\ &\leq C\rho^{m-d/2} h^{d-\sigma}. \end{aligned}$$

*Estimating the middle term in (3.17):* To handle the series appearing in (3.17), we proceed as in the last paragraph, applying, for each  $\ell$ , Lemma 3.4, now with  $u = \chi_\xi(\cdot - \xi)$ ,  $R = 2^{\ell+2}(K-1)h$  and  $U = u(R\cdot)$ . In this case  $O = B(0, 1) \setminus B(0, 1/8)$ . The scaling lemma uses the embedding  $W_2^\sigma(O) \subset W_1^\sigma(O)$  which incurs an embedding constant  $C$  which is independent of  $\ell$ . Namely,

$$\begin{aligned} |\chi_\xi|_{W_1^\sigma(v_\ell)} &= |u|_{W_1^\sigma(O_R)} \\ &\leq CR^{d/2+m-\sigma} |u|_{W_2^m(O_R)} \\ &= C(2^{\ell+2}(K-1))^{d/2+m-\sigma} h^{d/2+m-\sigma} \|\chi_\xi\|_{W_2^m(v_\ell)}. \end{aligned}$$

Since  $v_\ell$  is contained in  $\tilde{\Omega} \setminus B(\xi, 2^{\ell-1}(K-1)h)$ , we have

$$|\chi_\xi|_{W_1^\sigma(v_\ell)} \leq C2^{\ell(d/2+m-\sigma)} h^{d/2+m-\sigma} \|\chi_\xi\|_{W_2^m(\tilde{\Omega} \setminus B(v_\ell))}.$$

Now we apply the energy estimate (3.4) which gives

$$|\chi_\xi|_{W_1^\sigma(v_\ell)} \leq C 2^{\ell(d/2+m-\sigma)} h^{d/2+m-\sigma} q^{d/2-m} e^{-\mu(K-1)2^{\ell-1}}.$$

Observing that the infinite series  $\sum_{\ell=0}^\infty (2^{\ell(d/2+m-\sigma)} e^{-\mu(K-1)2^{\ell-1}})$  converges to a constant depending only on  $d$  and  $m$ , we can bound the middle term:

$$(3.20) \quad \left( \sum_{\ell=0}^{\ell_0} |\chi_\xi|_{W_1^\sigma(v_\ell)} \right) \leq C h^{m+d/2-\sigma} q^{d/2-m} = C \rho^{m-d/2} h^{d-\sigma}.$$

The case  $p = 1$  follows from the estimates (3.19), (3.20), (3.18) and (3.17).

*Case 4:*  $1 < p < 2$ . In this case, we use Riesz-Thorin to estimate the norm of the operator  $T : \ell_p(\Xi) \rightarrow W_p^\sigma(\Omega)$ , where  $T$  is the synthesis operator  $T\mathbf{a} = \sum_{\xi \in \Xi} a_\xi \chi_\xi$  (i.e., the linear map which takes coordinate space  $\mathbb{C}^\Xi$  into the vector space  $V_\Xi$ ). Letting  $\theta = 2(\frac{1}{p} - \frac{1}{2})$  (so that  $\frac{1}{p} = \theta \cdot 1 + (1 - \theta) \cdot \frac{1}{2}$ ) gives

$$\begin{aligned} \left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{W_p^\sigma(\Omega)} &\leq \left( C \rho^{m-d/2} h^{d-\sigma} \right)^\theta \left( C \rho^{m-d/2} h^{d/2-\sigma} \right)^{1-\theta} \|\mathbf{a}\|_{\ell_p(\Xi)} \\ &\leq C \rho^{m-d/2} h^{d/p-\sigma} \|\mathbf{a}\|_{\ell_p(\Xi)}. \end{aligned} \quad \square$$

Using Proposition 3.2, we may replace the discrete norm  $\|\mathbf{a}\|_{\ell_p(\Xi)}$  by its equivalent  $h^{-d/p} \|s\|_{L_p}$ , and so obtain an  $L_p$  version of Theorem 3.6.

**Corollary 3.7.** *With the assumptions of Theorem 3.6, we have*

$$(3.21) \quad \left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{W_p^\sigma(\Omega)} \leq C h^{-\sigma} \left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{L_p(\Omega)}$$

with  $C = C(p, \sigma, m, \rho, \Omega)$ .

Explicit dependence of  $C$  on  $\rho$  can be obtained from (3.12) and Lemma B.6.

#### 4. LOCAL LAGRANGE FUNCTIONS

We now consider a new class of functions  $b_\xi \in S(\tilde{\Xi})$ ,  $\xi \in \Xi$ , constructed in a local and cost-effective way, employing only a small set of centers in  $\tilde{\Xi}$  that are near  $\xi$ . For each  $\xi \in \Xi$ , this small set is called the local *footprint* of  $\xi$  and denoted by  $\Upsilon(\xi) \subset \tilde{\Xi}$  (see Definition 4.1). Each  $b_\xi$  is a Lagrange interpolant, centered at  $\xi$ , for points in  $\Upsilon(\xi)$ . The set  $\Upsilon(\xi)$  is chosen to give  $b_\xi$  fast decay away from  $\xi$ , although not exponential decay. The size of the footprint is controlled by a parameter  $K > 0$ .

Unlike the *full* Lagrange functions, the local versions do not satisfy interpolatory conditions throughout  $\tilde{\Xi}$ . There is no guarantee that they will have zeros outside of the set  $\Upsilon(\xi)$ —as a result the operator  $T\mathbf{a} = \sum_{\xi \in \Xi} a_\xi b_\xi$  does not satisfy  $T\mathbf{a}(\xi) = a_\xi$ . It is only a *quasi-interpolant* (approximation on the sphere with this operator was considered in [8]).

As in [8] the analysis of this new basis is considered in two steps. First, an intermediate basis function  $\tilde{\chi}_\xi$  is constructed and studied: the *truncated Lagrange function*. These functions employ the same footprint as  $b_\xi$  (i.e., they are members of  $S(\Upsilon(\xi))$ ) but their construction is global rather than local. This topic is considered in Section 4.2. Then, a comparison is made between the truncated Lagrange function and the local Lagrange function. The error between local and truncated Lagrange functions is controlled by the size of the coefficients in the representation of  $b_\xi - \tilde{\chi}_\xi$  using the standard (kernel) basis for  $S(\Upsilon(\xi))$ . This is considered in Section 4.3.

**4.1. Footprint and local Lagrange function.**

**Definition 4.1.** For a compact set  $\Omega \subset \mathbb{R}^d$  and a finite subset  $\Xi \subset \Omega$ , let  $\tilde{\Xi}$  be the extension to  $\{x \in \mathbb{R}^d \mid \text{dist}(x, \Omega) \leq \text{diam}(\Omega)\}$  given in Section 2.3. For a positive parameter  $K$ , define  $\Upsilon(\xi) := \{\zeta \in \tilde{\Xi} \mid |\xi - \zeta| \leq Kh|\log h|\}$  for each  $\xi \in \Xi$ . Then for the system of local Lagrange functions  $(b_\xi)_{\xi \in \tilde{\Xi}}$ , where each  $b_\xi$  is the Lagrange function centered at  $\xi$ , generated by  $k_m$  over  $\Upsilon(\xi)$ , let

$$\tilde{V}_\Xi := \text{span}\{b_\xi \mid \xi \in \Xi\}.$$

Note in particular that  $\tilde{V}_\Xi \subset S(\tilde{\Xi})$ . Indeed, it is contained in a slight expansion of  $S(\Xi)$ . Namely,  $\tilde{V}_\Xi \subset S(\mathcal{Y})$ , where  $\mathcal{Y} := \bigcup_{\xi \in \Xi} \Upsilon(\xi) \subset \{\xi \in \tilde{\Xi} \mid \text{dist}(\xi, \Omega) \leq Kh|\log h|\}$ .

The construction of each  $b_\xi$  depends only on its nearby neighbors in  $\Upsilon(\xi)$ , so the majority of points in  $\tilde{\Xi}$  are unnecessary from a computational point of view. However, the (analytic properties of) full Lagrange functions  $\chi_\xi$  generated by  $k_m$  over  $\tilde{\Xi}$  will still be of use in proving theorems, so we will continue to refer to the extended set  $\tilde{\Xi}$ , even though much of it plays no role in the construction of the functions  $b_\xi$ .

In our main result, we make use of the following:

Let  $(\chi_\xi)_{\xi \in \Xi}$  be the family of ‘full’ Lagrange functions constructed by  $k_m$  using the extended point set  $\tilde{\Xi}$ . For any  $J > 0$ , the family  $(b_\xi)_{\xi \in \Xi}$  satisfies

$$(4.1) \quad \|\chi_\xi - b_\xi\|_{W_{\mathcal{F}}(\Omega)} \lesssim h^J \text{ for all } \xi \in \Xi.$$

To obtain this result, we show that for a given  $J$  there is a  $K > 0$ , which governs the size of the footprint, ensuring that  $\|\chi_\xi - b_\xi\|_\infty = \mathcal{O}(h^J)$  holds. The value of  $K$  depends linearly on  $J$ , as well as some fixed constants involving  $m$  and  $d$ .

In the following two sections, we show that this result holds for Matérn (in Lemma 4.7) and surface spline radial basis functions (in Lemma 4.9). Specifically, this holds for any prescribed value of  $J$ , where  $J$  depends linearly on  $K$ , as given in Definition 4.1.

**4.2. Intermediate construction: Truncated Lagrange functions.** For a (full) Lagrange function  $\chi_\xi = \sum_{\zeta \in \tilde{\Xi}} A_{\xi, \zeta} k(\cdot, \zeta) + p \in S(\tilde{\Xi})$  on the point set  $\tilde{\Xi}$ , the truncated Lagrange function  $\tilde{\chi}_\xi := \sum_{\zeta \in \Upsilon(\xi)} \tilde{A}_{\xi, \zeta} k(\cdot, \zeta) + p$  is a function in  $S(\Upsilon(\xi))$  obtained by omitting the coefficients outside of  $\Upsilon(\xi)$  and slightly modifying the remaining coefficients  $\mathbf{A}_\xi = (A_{\xi, \zeta})$ . (For positive definite kernels, no modification is necessary, and the construction is quite simple.)

The cost of truncating can be measured using the norm of the omitted coefficients (the tail).

**Lemma 4.2.** *Let  $K > (4m - 2d)/\nu$  and for each  $\xi \in \Xi$ , let  $\Upsilon(\xi) = \{\zeta \in \tilde{\Xi} \mid |\xi - \zeta| \leq Kh|\log h|\}$ . Then*

$$\sum_{\zeta \in \tilde{\Xi} \setminus \Upsilon(\xi)} |A_{\xi, \zeta}| \leq C\rho^{2m} h^{K\nu/2 + d - 2m}$$

with  $C = C(m, d)$ .

*Proof.* The inequality (3.6) guarantees that the tail of the coefficients obey

$$\sum_{\zeta \in \tilde{\Xi} \setminus \Upsilon(\xi)} |A_{\xi, \zeta}| \leq Cq^{d-2m} \sum_{|\xi - \zeta| \geq Kh|\log h|} \exp\left(-\nu \frac{\text{dist}(\xi, \zeta)}{h}\right).$$

By observing that for  $\zeta \in \tilde{\Xi} \setminus \Upsilon(\xi)$ , we have  $q^d \leq C \text{vol}(B(\zeta, q) \setminus B(\xi, Kh|\log h|))$  with a constant  $C$  that depends only on the spatial dimension  $d$ . (Note that for most  $\zeta$ , the above set is simply  $B(\zeta, q)$ , while for those  $\zeta$  which are near the boundary of  $B(\xi, Kh|\log h|)$  the set contains a half-ball), we can control the above sum by an integral, namely

$$\begin{aligned} \sum_{\zeta \in \tilde{\Xi} \setminus \Upsilon(\xi)} |A_{\xi, \zeta}| &\leq Cq^{d-2m} \sum_{|\xi - \zeta| \geq Kh|\log h|} \exp\left(-\nu \frac{|\xi - \zeta|}{h}\right) \\ &\leq Cq^{-2m} \sum_{|\xi - \zeta| \geq Kh|\log h|} \int_{y \in \tilde{B}(\xi, q)} \exp\left(-\nu \frac{|\xi - \zeta|}{h}\right) dy \\ (4.2) \quad &\leq Cq^{-2m} \int_{y \in \mathbb{R}^d \setminus B(\xi, Kh|\log h|)} \exp\left(-\nu \frac{|\xi - \zeta|}{h}\right) dy. \end{aligned}$$

(Here we employ the sets  $\tilde{B}(\xi, q) := B(\xi, q) \setminus B(\xi, Kh|\log h|)$ , which contain at least half volume of  $B(\xi, q)$ .) In the final inequality, we have used the fact that the sets  $B(\zeta, q) \setminus B(\xi, Kh|\log h|)$  are disjoint and that for  $y \in B(\zeta, q)$ ,  $\text{dist}(\xi, y) \leq \text{dist}(\xi, \zeta) + q \leq \text{dist}(\xi, \zeta) + h$ , which implies  $-\text{dist}(\xi, \zeta) \leq -\text{dist}(\xi, y) + h$  (leading to a small increase in the constant; a factor of  $e^\nu$ ).

Applying a polar change of variables in the final integral gives the inequality

$$\sum_{\zeta \in \tilde{\Xi} \setminus \Upsilon(\xi)} |A_{\xi, \zeta}| \leq Cq^{-2m} \int_{Kh|\log h|}^\infty \exp\left(-\nu \frac{r}{h}\right) r^{d-1} dr.$$

We simplify this estimate by splitting  $\nu = \nu/2 + \nu/2$  and writing

$$\begin{aligned} \sum_{\zeta \in \tilde{\Xi} \setminus \Upsilon(\xi)} |A_{\xi, \zeta}| &\leq Ch^d q^{-2m} \left( \int_{K|\log h|}^\infty r^{d-1} \exp\left(-K|\log h| \frac{\nu}{2}\right) \exp\left(-r \frac{\nu}{2}\right) dr \right) \\ &\leq Ch^d q^{-2m} h^{K\nu/2} = C\rho^{2m} h^{K\nu/2 + d - 2m}. \end{aligned}$$

The lemma follows. □

4.2.1. *Bounds for truncated functions: Matérn functions.* Let  $\|\cdot\|_Z$  be a norm on  $S(\tilde{\Xi})$  for which a universal constant  $\Gamma$  exists so that  $\sup_{z \in \Omega} \|k_m(\cdot - z)\|_Z \leq \Gamma$ . Since  $\|k_m(\cdot - z)\|_Z$  is finite and bounded independent of  $z$ , we have

$$(4.3) \quad \|\chi_\xi - \tilde{\chi}_\xi\|_Z \leq \Gamma \sum_{\zeta \in \tilde{\Xi} \setminus \Upsilon(\xi)} |A_{\xi, \zeta}| \leq C\Gamma \rho^{2m} h^{K\nu/2 - 2m + d}.$$

In particular, we have the following.

**Lemma 4.3.** *Let  $m > d/2$  and consider the Matérn radial basis function  $k_m = \kappa_m$  described in (2.5). For  $1 \leq p < \infty$  and  $\sigma < 2m - d + \frac{d}{p}$  we have*

$$\|\chi_\xi - \tilde{\chi}_\xi\|_{W_p^\sigma(\mathbb{R}^d)} \leq \sum_{\zeta \in X \setminus \Upsilon(\xi)} |A_{\xi, \zeta}| \|\kappa_m(\cdot, \zeta)\|_{W_p^\sigma(\mathbb{R}^d)} \leq C\rho^{2m} h^{K\nu/2 + d - 2m}$$

with  $C = C(m, d)$ .

For  $p = \infty$ , the above result holds for the Hölder space  $W_\infty^\sigma(\mathbb{R}^d)$  replaced with  $C^\sigma(\mathbb{R}^d)$ .

*Proof.* By examination of (2.5), we see that  $\kappa_m \in W_p^\tau(\mathbb{R}^d)$  for  $1 \leq p < \infty$  and  $\tau < 2m - d + d/p$ , while for  $p = \infty$ ,  $\kappa_m \in C^\tau(\mathbb{R}^d)$  with  $\tau < 2m - d$ . In either case, the smoothness norm is translation invariant, so it follows that

$$\|\kappa_m(\cdot - z)\|_{W_p^\tau(\mathbb{R}^d)} \leq C_{\tau,p} \quad \text{and} \quad \|\kappa_m(\cdot - z)\|_{C^\tau(\mathbb{R}^d)} \leq C_{\tau,\infty}$$

hold. The result follows from (4.3).  $\square$

**4.2.2. Bounds for truncated functions: Surface splines.** When  $k_m = \phi_m$  (i.e., a surface spline, and therefore conditionally positive definite), the argument of the previous section is a little more complicated. Given a Lagrange function  $\chi_\xi = \sum_{\zeta \in X} A_{\zeta,\xi} k_m(\cdot, \zeta) + p$ , simply truncating coefficients does not yield a function in  $S(\Upsilon(\xi))$ . That is,  $(A_{\zeta,\xi})_{\zeta \in \Upsilon(\xi)}$  does not necessarily satisfy the side condition  $\sum_{\zeta \in \Upsilon(\xi)} A_{\zeta,\xi} p(\zeta) = 0$  for all  $p \in \Pi_{m-1}$ .

The result for *restricted surface splines* on even dimensional spheres ( $\mathbb{S}^{2n}$ ) has been developed in [8, Proposition 6.1]. We now present a similar estimate for surface splines on  $\mathbb{R}^d$  where the truncated Lagrange function is corrected by perturbing its coefficients slightly. This is done by using the orthogonal projector having range  $\perp (\Pi_{m-1} |_{\Upsilon(\xi)})$ . Keeping this perturbation small is essential to our later results, so we must estimate it. We use the following result about Gram matrices for polynomials sampled on finite point sets.

*Gram matrices for polynomials restricted to point sets.* Let  $N = \dim \Pi_{m-1}$  and consider  $X \subset \mathbb{R}^d$  a finite point set. For a basis  $\{p_1, \dots, p_N\}$  of  $\Pi_{m-1}$ , denote by  $\Phi_X$  the (Vandermonde-type) matrix with  $N$  columns and  $\#X$  rows whose  $j$ th column is  $p_j$  restricted to  $X$ . In other words,

$$(4.4) \quad \Phi_X \in M_{(\#X) \times N}(\mathbb{R}) \quad \text{with} \quad (\Phi_X)_{\xi,j} = p_j(\xi).$$

**Lemma 4.4.** *For every  $m \in \mathbb{N}$ , and any radius  $r > 0$ , point  $x \in \mathbb{R}^d$  and point set  $X \subset B(x, r)$  with fill distance  $h \leq h_0 r$ , where  $h_0 = h_0(m)$ , the inverse of the Gram matrix  $G_X = \Phi_X^T \Phi_X \in M_{N \times N}(\mathbb{R})$  has norm bounded by*

$$\|G_X^{-1}\|_{2 \rightarrow 2} \leq Cr^{-2(m-1)}$$

with  $C = C(m, d)$ .

*Proof.* From [32, Theorem 3.8 and Corollary 3.11], we have that if  $X \subset B(x, r)$  has fill distance  $h \leq h_0 r$ , then  $X$  is a *norming set* for  $B(x, r)$  with norming constant 2. (Here  $h_0 = c_{m-1}$ , from [32, Corollary 3.11].) This means that for every  $p \in \Pi_{m-1}$ ,  $\|p\|_{L_\infty(B(x,r))} \leq 2\|p|_X\|_{\ell_\infty(X)}$ .

The norm of the Gram matrix can be controlled by

$$\|G_X^{-1}\|_{2 \rightarrow 2} = \left( \min_{\|\mathbf{a}\|=1} \langle G_X \mathbf{a}, \mathbf{a} \rangle \right)^{-1} \quad \text{and} \quad \langle G_X \mathbf{a}, \mathbf{a} \rangle = \|\Phi_X \mathbf{a}\|_{\ell_2(X)}^2 = \|R_X V \mathbf{a}\|_{\ell_2(X)}^2,$$

where  $V \mathbf{a} := \sum_{j=1}^N a_j p_j$  and  $R_X$  is the restriction operator  $R_X V \mathbf{a} = \sum_{j=1}^N a_j p_j|_X$ . For  $h$  sufficiently small, the norming set property ensures that

$$\|p\|_{L_\infty(B(x,r))} \leq 2\|R_X p\|_{\ell_\infty(X)} \leq 2\|R_X p\|_{\ell_2(X)}.$$

On the other hand, we have the following growth properties of polynomials  $\Pi_{m-1}$ : there exists a constant  $C_m > 0$  so that for every  $0 < r < 1$ ,  $\|p\|_{L_\infty(B(x,1))} \leq C_m r^{-(m-1)} \|p\|_{L_\infty(B(x,r))}$ . Returning to the basis  $(p_1, \dots, p_N)$ , we have

$$\|\mathbf{a}\|_{\ell_2(N)} \leq C_{m,d} \left\| \sum_{j=1}^N a_j p_j \right\|_{L_\infty(B(x,1))} \leq C_{m,d} \left(\frac{1}{r}\right)^{m-1} \left\| \sum_{j=1}^N a_j p_j \right\|_{L_\infty(B(x,r))}.$$

This gives  $\|\mathbf{a}\|_{\ell_2(N)} \leq C r^{-(m-1)} \left\| \sum_{j=1}^N a_j p_j \right\|_{L_\infty(B(x,r))}$ , and the result follows.  $\square$

A bound similar to this for  $\mathbb{S}^{d-1}$  using spaces of spherical harmonics in place of  $\Pi_{m-1}$  has been demonstrated in [8, Lemma 6.4], while [15] gives general conditions for the auxiliary space of a CPD kernel.

*Modifying coefficients.* We use the matrix  $\Phi_{\Upsilon(\xi)}$  to construct the orthogonal projector  $P = \Phi_{\Upsilon(\xi)} (\Phi_{\Upsilon(\xi)}^T \Phi_{\Upsilon(\xi)})^{-1} \Phi_{\Upsilon(\xi)}^T$ , which has range  $\Pi|_{\Upsilon(\xi)}$  and kernel  $\perp(\Pi|_{\Upsilon(\xi)})$ . For a fixed  $\xi$ , denote the truncated coefficients  $(A_{\zeta,\xi})_{\zeta \in \Upsilon(\xi)} \in \mathbb{R}^{\Upsilon(\xi)}$  by  $\mathbf{A}_\xi$ . In order to satisfy the side conditions, we generate the modified coefficients  $\tilde{\mathbf{A}}_\xi = (\tilde{A}_{\zeta,\xi}) \in \mathbb{R}^{\Upsilon(\xi)}$  via

$$\tilde{\mathbf{A}}_\xi = \mathbf{A}_\xi - P \mathbf{A}_\xi.$$

In other words,  $\tilde{\mathbf{A}}_\xi$  is the orthogonal projection of  $\mathbf{A}_\xi$  onto  $\perp(\Pi|_{\Upsilon(\xi)})$ . Define the ‘truncated’ Lagrange function as

$$\tilde{\chi}_\xi := \sum_{\zeta \in \Upsilon(\xi)} \tilde{A}_{\zeta,\xi} \phi_m(\cdot - \zeta) + p.$$

**Lemma 4.5.** *Let  $m > d/2$  and consider the surface spline radial basis function  $k_m = \phi_m$  described in (2.6). For sufficiently small  $h$  we have*

$$(4.5) \quad \|\mathbf{A} - \tilde{\mathbf{A}}\|_{\ell_2(\Upsilon(\xi))} \leq C \rho^{2m} h^{K\nu/2+1-3m+d} |\log h|^{1-m}$$

with  $C = C(m, d)$ .

*Proof.* Using  $\mathbf{A}_\xi - \tilde{\mathbf{A}}_\xi = P \mathbf{A}_\xi$  and  $P = \Phi_{\Upsilon(\xi)} G_{\Upsilon(\xi)}^{-1} \Phi_{\Upsilon(\xi)}^T$ , we can estimate the  $\ell_2$  norm of the difference of the coefficients as

$$\|\mathbf{A}_\xi - \tilde{\mathbf{A}}_\xi\|_{\ell_2(\Upsilon(\xi))} = \langle \Phi_{\Upsilon(\xi)}^T \mathbf{A}_\xi, G_{\Upsilon(\xi)}^{-1} \Phi_{\Upsilon(\xi)}^T \mathbf{A}_\xi \rangle^{1/2} \leq \|G_{\Upsilon(\xi)}^{-1}\|_{2 \rightarrow 2}^{1/2} \|\Phi_{\Upsilon(\xi)}^T \mathbf{A}_\xi\|_{\ell_2(N)}.$$

Since  $\sum_{\zeta \in \Xi} A_{\zeta,\xi} p(\zeta) = 0$  for all polynomials  $p \in \Pi_{m-1}$ , we have  $\Phi_{\Upsilon(\xi)}^T \mathbf{A}_\xi = -(\sum_{\zeta \in \Xi \setminus \Upsilon(\xi)} A_{\zeta,\xi} p_j(\zeta))_{j=1}^N$ .

Applying the estimate (2.2), the  $\ell_2(N)$  norm of  $\Phi^T \mathbf{A}$  is controlled by

$$\begin{aligned} \|\Phi_{\Upsilon(\xi)}^T \mathbf{A}_\xi\|_{\ell_2(N)} &\leq \|\Phi_{\Upsilon(\xi)}^T \mathbf{A}_\xi\|_{\ell_1(N)} \leq \sum_{j=1}^N \left| \sum_{\zeta \in \Xi \setminus \Upsilon(\xi)} A_{\zeta,\xi} p_j(\zeta) \right| \\ &\leq \sum_{\zeta \in \Xi \setminus \Upsilon(\xi)} |A_{\zeta,\xi}| \sum_{j=1}^N |p_j(\zeta)|. \end{aligned}$$

In the first estimate we use the inequality  $\sum |c_j|^2 \leq (\sum |c_j|)^2$ . Applying Hölder’s inequality and (3.6) to the right-hand side gives

$$\begin{aligned} \|\Phi_{\Upsilon(\xi)}^T \mathbf{A}_\xi\|_{\ell_2(N)} &\leq N \sum_{\zeta \in \tilde{\Xi} \setminus \Upsilon(\xi)} \left( \max_{j=1, \dots, N} |p_j(\zeta)| \right) |A_{\zeta, \xi}| \\ &\leq Cq^{d-2m} \sum_{\zeta \in \tilde{\Xi} \setminus \Upsilon(\xi)} \left( \max_{j=1, \dots, N} |p_j(\zeta)| \right) \exp\left(-\nu \frac{\text{dist}(\xi, \zeta)}{h}\right), \end{aligned}$$

where we have absorbed  $N$  (recall that  $N = \dim \Pi_{m-1}$  depends on  $m$  and  $d$ ) into the constant  $C$ .

We now recall the argument in (4.2) which allows us to estimate the above sum by an integral:

$$\begin{aligned} \|\Phi_{\Upsilon(\xi)}^T \mathbf{A}_\xi\|_{\ell_2(N)} &\leq Cq^{-2m} \int_{Kh \log h}^\infty \max_{j=1, \dots, N} \left( \|p_j\|_{L_\infty(B(\xi, z))} \right) e^{-\nu z/h} dz \\ &\leq Cq^{-2m} \int_{Kh \log h}^\infty \max(1, z^{m-1}) e^{-\nu z/h} dz \\ (4.6) \qquad &\leq C\rho^{2m} h^{K\nu/2+d-2m}. \end{aligned}$$

In (4.6) we have used a change to polar coordinates, as in Lemma 4.2.

Estimate (4.5) follows by combining Lemma 4.4 (using  $r = Kh|\log h|$ ) with (4.6). □

As in the positive definite case, we are able to control the truncation error measured in suitable smoothness norms—the only requirement is that the kernel is bounded. In the conditionally positive definite case, the kernel may be unbounded, so we measure the norm over the bounded region  $\Omega$ . Specifically, the surface spline  $\phi_m \in W_{p, \text{loc}}^\sigma(\mathbb{R}^d)$  for all  $\sigma < 2m - d + \frac{d}{p}$  (as well as  $C_{\text{loc}}^\sigma(\mathbb{R}^d)$  for  $\sigma < 2m - d$ ). There is  $\Gamma < \infty$  (depending on  $\sigma, p, m$  and  $\Omega$ ) so that for  $\zeta \in \tilde{\Omega}$ ,  $\|\phi_m(\cdot - \zeta)\|_{W_p^\sigma(\Omega)} \leq \Gamma$ .

**Lemma 4.6.** *For  $1 \leq p < \infty$  and  $\sigma < 2m - d + \frac{d}{p}$ ,*

$$\|\tilde{\chi}_\xi - \chi_\xi\|_{W_p^\sigma(\Omega)} \leq C\rho^{2m+d/2} h^{K\nu/2+1-3m+d} |\log h|^{d/2+1-m}$$

with  $C = C(\sigma, m, p, \Omega)$ .

A similar result holds for  $p = \infty$ , replacing  $W_p^\sigma(\Omega)$  by  $C^\sigma(\bar{\Omega})$  for  $\sigma < 2m - d$ .

*Proof.* The Sobolev estimate holds by considering

$$\begin{aligned} \|\tilde{\chi}_\xi - \chi_\xi\|_{W_p^\sigma(\Omega)} &\leq \sum_{\zeta \in \Upsilon(\xi)} |A_{\xi, \zeta} - \tilde{A}_{\xi, \zeta}| \|\phi_m(\cdot - \zeta)\|_{W_p^\sigma(\Omega)} \\ &\quad + \sum_{\zeta \notin \Upsilon(\xi)} |A_{\xi, \zeta}| \|\phi_m(\cdot - \zeta)\|_{W_p^\sigma(\Omega)}. \end{aligned}$$

The first term can be bounded by introducing the constant

$$\Gamma := \max_{\zeta \in \tilde{\Omega}} \|\phi_m(\cdot - \zeta)\|_{W_p^\sigma(\Omega)} < \infty,$$



which gives  $\sum_{\zeta \in \Upsilon(\xi)} |A_{\xi, \zeta} - \tilde{A}_{\xi, \zeta}| \|\phi_m(\cdot - \zeta)\|_{W_p^\sigma(\Omega)} \leq \Gamma \|\mathbf{A}_\xi - \tilde{\mathbf{A}}_\xi\|_{\ell_1(\Upsilon(\xi))}$ . Employing (4.5) yields

$$\begin{aligned} \sum_{\zeta \in \Upsilon(\xi)} |A_{\xi, \zeta} - \tilde{A}_{\xi, \zeta}| \|\phi_m(\cdot - \zeta)\|_{W_p^\sigma(\Omega)} &\leq C\Gamma \rho^{2m} h^{\frac{K\nu}{2} + 1 - 3m + d} |\log h|^{1-m} (\#\Upsilon(\xi))^{\frac{d}{2}} \\ &\leq C\Gamma \rho^{2m+d/2} h^{\frac{K\nu}{2} + 1 - 3m + d} |\log h|^{1-m+\frac{d}{2}}. \end{aligned}$$

For the second inequality we have used the estimate  $\#\Upsilon(\xi) \leq C\rho^d |\log h|^d$ .

The second term  $\sum_{\zeta \notin \Upsilon(\xi)} |A_{\xi, \zeta}| \|\phi_m(\cdot - \zeta)\|_{W_p^\sigma(\Omega)} \leq \Gamma \sum_{\zeta \notin \Upsilon(\xi)} |A_{\xi, \zeta}|$  is treated with Lemma 4.2 to obtain  $\sum_{\zeta \notin \Upsilon(\xi)} |A_{\xi, \zeta}| \|\phi_m(\cdot - \zeta)\|_{W_p^\sigma(\Omega)} \leq C\Gamma \rho^{2m} h^{\frac{K\nu}{2} - 2m + d}$ .  $\square$

**4.3. Local Lagrange functions.** In this section we consider a locally constructed function  $b_\xi$ . Our main goal is to show that for  $\Xi \subset \Omega$ , there exist functions  $b_\xi$  defined on  $\mathbb{R}^d$ , so that  $\|\sum_{\xi \in \Xi} a_\xi b_\xi\|_{W_p^\sigma(\Omega)} \leq Ch^{\frac{d}{p} - \sigma} \|\mathbf{a}\|_{\ell_p(\Xi)}$ .

At this point, a standard argument bounds the error between  $b_\xi$  and  $\tilde{\chi}_\xi$  (this argument is essentially the same one used on the sphere in [8]). This works by measuring the size of  $b_\xi - \tilde{\chi}_\xi \in S(\Upsilon(\xi))$ .

**4.3.1. Bounds for local Lagrange functions: Matérn functions.** For the positive definite case, the argument is fairly elementary. For  $\zeta \in \Upsilon(\xi)$ , let  $y_\zeta := b_\xi(\zeta) - \tilde{\chi}_\xi(\zeta)$ . Observe that  $b_\xi - \tilde{\chi}_\xi = \sum_{\zeta \in \Upsilon(\xi)} a_\zeta k_m(\cdot - \zeta) \in S(\Upsilon(\xi))$ , where  $\mathbf{a} = (a_\zeta)$  and  $\mathbf{y} = (y_\zeta)$  are related by  $\mathbf{K}_{\Upsilon(\xi)} \mathbf{a} = \mathbf{y}$ . The matrix  $(\mathbf{K}_{\Upsilon(\xi)})^{-1}$  has entries  $(A_{\zeta, \eta})_{\zeta, \eta \in \Upsilon(\xi)}$ .

For a kernel of order  $m$ , the entries of the matrix  $\mathcal{A} = (A_{\zeta, \eta})_{\zeta, \eta \in \Upsilon(\xi)}$  can be estimated by (3.3):  $|A_{\zeta, \eta}| \leq Cq^{d-2m}$ . It follows that  $(\mathbf{K}_{\Upsilon(\xi)})^{-1}$  has  $\ell_1$  matrix norm

$$\left\| (\mathbf{K}_{\Upsilon(\xi)})^{-1} \right\|_{1 \rightarrow 1} \leq C(\#\Upsilon(\xi))q^{d-2m} \leq C\rho^{2m} |\log h|^d h^{d-2m}.$$

(Here we have used the estimate  $\#\Upsilon(\xi) \leq C\rho^d |\log h|^d$ .) Consequently  $\|\mathbf{y}\|_1 \leq (\#\Upsilon(\xi)) \|\mathbf{y}\|_\infty$ .

Because  $y_\zeta = \chi_\xi(\zeta) - \tilde{\chi}_\xi(\zeta)$  for  $\zeta \in \Upsilon(\xi)$  and  $\|\chi_\xi - \tilde{\chi}_\xi\|_\infty \leq C\|\chi_\xi - \tilde{\chi}_\xi\|_{W_2^m(\mathbb{R}^d)}$  we have

$$\sum_{\zeta \in \Upsilon(\xi)} |a_\zeta| \leq \left\| (\mathbf{K}_{\Upsilon(\xi)})^{-1} \right\|_{1 \rightarrow 1} \|\mathbf{y}\|_1 \leq C\rho^{2m+d} h^{d-2m} |\log h|^{2d} \|\chi_\xi - \tilde{\chi}_\xi\|_{W_2^m(\mathbb{R}^d)}.$$

For a generic norm  $\|\cdot\|_Z$  for which  $\max_{z \in \bar{\Omega}} \|k_m(\cdot - z)\|_Z \leq \Gamma$  we have  $\|b_\xi - \tilde{\chi}_\xi\|_Z \leq \Gamma \sum_{\zeta} |a_\zeta|$ . We now have the counterpart to Lemma 4.3, which shows that (4.1) holds for Matérn kernels.

**Lemma 4.7.** *For  $k_m = \kappa_m$  and for  $1 \leq p \leq \infty$  and  $\sigma < 2m - d + d/p$  we have*

$$(4.7) \quad \|b_\xi - \chi_\xi\|_{W_p^\sigma(\mathbb{R}^d)} \leq C\rho^{4m+d} h^{K\nu/2+2d-4m} |\log h|^{2d}$$

with  $C = C(m, d)$ .

Setting  $|\log h|^{2d} \leq Ch^{-1}$  (either by finding a sufficiently small  $h^*$  so that this holds for  $h < h^*$ , or by increasing the constant, or both), and by employing a simple interpolation inequality, we have

$$(4.8) \quad \|b_\xi - \chi_\xi\|_{W_p^\sigma(\mathbb{R}^d)} \leq C\rho^{4m+d} h^J, \quad J = K\nu/2 + 2d - 4m - 1.$$

4.3.2. *Bounds for local Lagrange functions: Surface splines.* As in the previous section, we are guided by the estimates for local Lagrange functions on the sphere [8, Proposition 5.2].

In this case we have  $\tilde{\chi}_\xi - b_\xi = \sum_{\zeta \in \Upsilon(\xi)} a_\zeta \phi_m(\cdot - \zeta) + \sum_{j=1}^N c_j p_j \in S(\Upsilon(\xi))$ . The vectors  $\mathbf{a} = (a_\zeta)_{\zeta \in \Upsilon(\xi)}$  and  $\mathbf{c} = (c_j)_{j=1, \dots, N}$  are related to  $\mathbf{y} = (y_\zeta)_{\zeta \in \Upsilon(\xi)} = (\tilde{\chi}_\zeta - b_\zeta)_{\zeta \in \Upsilon(\xi)}$  by

$$\begin{pmatrix} \mathbf{K}_{\Upsilon(\xi)} & \Phi \\ \Phi^T & 0_{N \times N} \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \mathbf{c} \end{pmatrix} = \begin{pmatrix} \mathbf{y} \\ 0_{N \times 1} \end{pmatrix},$$

where  $\mathbf{K}_{\Upsilon(\xi)}$  is the collocation matrix and  $\Phi$  is the Vandermonde matrix introduced in (4.4). The norms of  $\mathbf{a}$  and  $\mathbf{c}$  can be controlled by  $\|\mathbf{y}\|_{\ell_2(\Upsilon(\xi))}$ . This is demonstrated in [8, Proposition 5.2], which shows that  $\|\mathbf{a}\|_{\ell_2(\Upsilon(\xi))} \leq \vartheta^{-1} \|\mathbf{y}\|_{\ell_2(\Upsilon(\xi))}$  where  $\vartheta$  is the minimal positive eigenvalue of  $P^\perp \mathbf{K}_{\Upsilon(\xi)} P^\perp$ . Recall that  $P^\perp = \text{Id} - P$ ,  $P = \Phi(\Phi^T \Phi)^{-1} \Phi^T$  is the projector introduced in Section 4.2.2.

We make the following observation, which is [8, Proposition 5.2]:

$$(4.9) \quad \|\mathbf{a}\|_{\ell_2(\Upsilon(\xi))} \leq \vartheta^{-1} \|\mathbf{y}\|_{\ell_2(\Upsilon(\xi))},$$

$$(4.10) \quad \|\mathbf{c}\|_{\ell_2(N)} \leq 2 \max_{\eta, \zeta \in \Upsilon(\xi)} |\phi_m(\eta - \zeta)| \|\mathbf{G}_{\Upsilon(\xi)}^{-1}\|^{1/2} \vartheta^{-1} (\#\Upsilon(\xi)) \|\mathbf{y}\|_{\ell_2(\Upsilon(\xi))}.$$

It is possible to estimate the size of  $\vartheta$  by considering the matrix of kernel coefficients for the Lagrange functions  $b_{\eta, \Upsilon(\xi)} = \sum_{\zeta \in \Upsilon(\xi)} A_{\zeta, \eta} k_m(\cdot, \zeta) + \sum_{j=1}^N B_{j, \eta} \phi_j$ .

**Lemma 4.8.** *For  $\vartheta$ , the minimal positive eigenvalue of  $P^\perp \mathbf{K}_{\Upsilon(\xi)} P^\perp$ , we have  $\vartheta^{-1} = \|\mathcal{A}\|_{2 \rightarrow 2}$ , where  $\mathcal{A} = (A_{\zeta, \eta})_{\zeta, \eta \in \Upsilon(\xi)}$  is the matrix of kernel coefficients for the Lagrange functions in  $S(\Upsilon(\xi))$ .*

*Proof.* Writing  $\mathcal{B} = (B_{j, \eta})_{j=1, \dots, N}$  it follows that  $\mathbf{K}_{\Upsilon(\xi)} \mathcal{A} + \Phi \mathcal{B} = \text{Id}$ . From this we have  $P^\perp = P^\perp \mathbf{K}_{\Upsilon(\xi)} \mathcal{A}$  and  $\ker \mathcal{A} \subset \ker P^\perp$ . On the other hand, each column of  $\mathcal{A}$  satisfies the side condition  $\sum_{\eta \in \Upsilon(\xi)} A_{\zeta, \eta} p(\eta) = 0$  for all  $p \in \Pi$ , so  $\text{ran} \mathcal{A} \subset \text{ran} P^\perp$ . From this it follows that  $\ker \mathcal{A} = \ker P^\perp$  and  $\text{ran} \mathcal{A} = \text{ran} P^\perp$ .

Because  $P^\perp \mathcal{A} = \mathcal{A}$  we have  $P^\perp = P^\perp \mathbf{K}_{\Upsilon(\xi)} \mathcal{A} = P^\perp \mathbf{K}_{\Upsilon(\xi)} P^\perp \mathcal{A}$ , and the nonzero spectrum of  $\mathcal{A}$  is the reciprocal of the nonzero spectrum of  $P^\perp \mathbf{K}_{\Upsilon(\xi)} P^\perp$ . In other words,  $\vartheta^{-1} = \max_{\lambda \in \sigma(\mathcal{A})} |\lambda|$ .  $\square$

Applying Gershgorin's theorem to  $\mathcal{A}$ , whose entries are  $A_{\zeta, \eta} = \langle b_{\zeta, \Upsilon(\xi)}, b_{\eta, \Upsilon(\xi)} \rangle$  and therefore satisfy  $|A_{\zeta, \eta}| \leq C q^{d-2m}$ , we have  $\vartheta^{-1} \leq C(1 + \#\Upsilon(\xi)) q^{d-2m}$ . By (4.9) we have

$$(4.11) \quad \|\mathbf{a}\|_{\ell_2(\Upsilon(\xi))} \leq C \rho^{2m} h^{d-2m} |\log h|^d \|\mathbf{y}\|_{\ell_2(\Upsilon(\xi))}.$$

We will apply (4.10) to estimate  $\|\mathbf{c}\|_{\ell_2(N)}$ , noting that Lemma 4.4 ensures that  $\|\mathbf{G}_{\Upsilon(\xi)}^{-1}\|^{1/2} \leq C(Kh |\log h|)^{-(m-1)}$ . We have shown in the previous paragraph that  $\vartheta^{-1} \leq C \rho^{2m} h^{d-2m} |\log h|^d$  and a counting argument gives  $(\#\Upsilon(\xi)) \leq C \rho^{2m} |\log h|^d$ . Thus the right-hand side of (4.10) can be estimated as

$$\begin{aligned} & 2 \max_{\eta, \zeta \in \Upsilon(\xi)} |\phi_m(\eta - \zeta)| \|\mathbf{G}_{\Upsilon(\xi)}^{-1}\|^{1/2} \vartheta^{-1} (\#\Upsilon(\xi)) \|\mathbf{y}\|_{\ell_2(\Upsilon(\xi))} \\ & \leq C(2Kh |\log h|)^{2m-d} \left( (Kh |\log h|)^{-(m-1)} \right) (\rho^{2m} h^{d-2m} |\log h|^d) (\rho^d |\log h|^d). \end{aligned}$$

Applying (4.10) gives

$$(4.12) \quad \|\mathbf{c}\|_{\ell_2(N)} \leq C\rho^{2m+d}h^{-(m-1)}|\log h|^{m+1+d}\|\mathbf{y}\|_{\ell_2(\Upsilon(\xi))}.$$

We are now in a position to prove that (4.1) holds for surface splines.

**Lemma 4.9.** *Let  $k_m = \phi_m$ , the surface spline RBF on  $\mathbb{R}^d$ , and let  $J > 0$ . For  $\Xi \subset \Omega$ , form the local Lagrange functions  $b_\xi \in \Upsilon(\xi)$ , with  $\Upsilon(\xi) = \tilde{\Xi} \cap B(\xi, Kh|\log h|)$ , where  $J = K\frac{\nu}{2} - 5m + d + 1$ . Then for  $1 \leq p \leq \infty$  and  $\sigma < 2m - d + \frac{d}{p}$ , and for sufficiently small  $h$ ,*

$$\|b_\xi - \chi_\xi\|_{W_p^\sigma(\Omega)} \leq C\rho^{4m+2d}h^J$$

with  $C = C(\sigma, m, p, \Omega)$ .

*Proof.* We use the triangle inequality  $\|b_\xi - \chi_\xi\|_{W_p^\sigma(\Omega)} \leq \|b_\xi - \tilde{\chi}_\xi\|_{W_p^\sigma(\Omega)} + \|\tilde{\chi}_\xi - \chi_\xi\|_{W_p^\sigma(\Omega)}$ , noting that the second term has been estimated in Lemma 4.5, and that the first can be controlled as

$$\|b_\xi - \tilde{\chi}_\xi\|_{W_p^\sigma(\Omega)} \leq \|\mathbf{a}\|_{\ell_1(\Upsilon(\xi))} \max_{z \in \Upsilon(\xi)} \|\phi_m(\cdot - z)\|_{W_p^\sigma(\Omega)} + \|\mathbf{c}\|_{\ell_1(N)} \max_{1 \leq j \leq N} \|p_j(\cdot)\|_{W_p^\sigma(\Omega)}.$$

From (4.11) we have  $\|\mathbf{a}\|_{\ell_1(\Upsilon(\xi))} \leq \sqrt{\#\Upsilon(\xi)}\|\mathbf{a}\|_{\ell_2(\Upsilon(\xi))}$  and counting argument gives  $\#\Upsilon(\xi) \leq C_{m,d}\rho^{d|\log h|^d}$ , so

$$\begin{aligned} \|\mathbf{a}\|_{\ell_1(\Upsilon(\xi))} &\leq C\rho^{2m+d/2}h^{d-2m}|\log h|^{3d/2}\|\mathbf{y}\|_{\ell_\infty(\Upsilon(\xi))} \\ &\leq C\rho^{4m+d}h^{K\nu/2-5m+2d+1}|\log h|^{2d-(m-1)}, \end{aligned}$$

where we have employed the result of Lemma 4.6 and the embedding  $W_2^m \subset L_\infty$  to estimate

$$\|\mathbf{y}\|_{\ell_\infty(\Upsilon(\xi))} \leq \|b_\xi - \tilde{\chi}_\xi\|_{L_\infty(B(\xi, Kh|\log h|))} \leq C_{d,m}\rho^{2m+\frac{d}{2}}h^{\frac{K\nu}{2}+1-3m+d}|\log h|^{1-m+\frac{d}{2}}.$$

Similarly, from (4.12), we have

$$\begin{aligned} \|\mathbf{c}\|_{\ell_1(N)} &\leq C\rho^{2m+d+d/2}h^{-(m-1)}|\log h|^{m+1+d+d/2}\|\mathbf{y}\|_{\ell_\infty(\Upsilon(\xi))} \\ &\leq C\rho^{4m+2d}h^{K\nu/2-4m+2+d}|\log h|^{2+2d}. \end{aligned}$$

Because  $\max_{z \in \Upsilon(\xi)} \|\phi_m(\cdot - z)\|_{W_p^\sigma(\Omega)}$  and  $\max_{1 \leq j \leq N} \|\phi_j(\cdot)\|_{W_p^\sigma(\Omega)}$  are bounded by a constant  $\Gamma$  which depends only on  $\Omega, m, p$  and  $\sigma$ , we have

$$\|b_\xi - \chi_\xi\|_{W_p^\sigma(\Omega)} \leq \Gamma C\rho^{4m+2d}h^{K\nu/2-5m+1+d}|\log h|^{2+2d}.$$

The lemma follows for  $h$  sufficiently small that  $|\log h|^{2d+2} < h^{-1}$ . □

**4.4. Bernstein type estimate for local Lagrange functions.** In this section we discuss the local Lagrange ( $b_\xi$ ) functions generated by  $k_m$  and the centers  $\tilde{\Xi}$ . We develop partial Bernstein inequalities similar to (3.12), where for functions  $\sum_{\xi \in \Xi} a_\xi b_\xi \in \tilde{\Xi}$ , smoothness norms  $\|s\|_{W_p^\sigma}$  are controlled by an  $\ell_p$  norm on the coefficients:  $\|\mathbf{a}\|_{\ell_p(\Xi)}$ .

**Theorem 4.10.** *Consider the family of local Lagrange functions generated with  $K > \frac{10m-2}{\nu}$ . For  $0 \leq \sigma \leq m - (d/2 - d/p)_+$  when  $1 \leq p < \infty$  (or  $\sigma \in \mathbb{N}$  and  $0 \leq \sigma < m - d/2$  when  $p = \infty$ ), we have*

$$(4.13) \quad \left\| \sum_{\xi \in \Xi} a_\xi b_\xi \right\|_{W_p^\sigma(\Omega)} \leq C\rho^{4m+3d}h^{d/p-\sigma}\|\mathbf{a}\|_{\ell_p(\Xi)},$$

where  $C = C(\sigma, p, m, \Omega)$ .

*Proof.* We start with the basic splitting

$$s := \sum_{\xi \in \Xi} a_\xi b_\xi = \left( \sum_{\xi \in \Xi} a_\xi \chi_\xi \right) + \left( \sum_{\xi \in \Xi} a_\xi (b_\xi - \chi_\xi) \right) =: G + B.$$

Applying the Sobolev norm gives  $\|s\|_{W_p^\sigma(\Omega)} \leq \|G\|_{W_p^\sigma(\Omega)} + \|B\|_{W_p^\sigma(\Omega)}$ . From (3.12), we have

$$\|G\|_{W_p^\sigma(\Omega)} \leq C\rho^{m+d/2+d/p}h^{d/p-\sigma}\|\mathbf{a}\|_{\ell_p(\Xi)} \leq C\rho^{4m+3d}h^{d/p-\sigma}\|\mathbf{a}\|_{\ell_p(\Xi)}.$$

Taking the  $L_p$  norm of  $B$ , we have

$$\left\| \sum_{\xi \in \Xi} a_\xi (b_\xi - \chi_\xi) \right\|_{W_p^\sigma(\Omega)} \leq \max_{\xi \in \Xi} \|b_\xi - \chi_\xi\|_{W_p^\sigma(\Omega)} \sum_{\xi \in \Xi} |a_\xi|.$$

We control the  $\ell_1$  norm by using Hölder’s inequality  $\|\mathbf{a}\|_1 \leq (\#\Xi)^{\frac{p-1}{p}}\|\mathbf{a}\|_p$  and  $\#\Xi \leq C_\Omega\rho^d h^{-d}$ . Using Lemma 4.9 (or Lemma 4.7 in case  $k_m = \kappa_m$ ), we arrive at the desired inequality

(4.14)

$$\left\| \sum_{\xi \in \Xi} a_\xi (b_\xi - \chi_\xi) \right\|_{W_p^\sigma(\Omega)} \leq C\rho^{4m+3d}h^{J-d(\frac{p-1}{p})}\|\mathbf{a}\|_{\ell_p(\Xi)} \leq C\rho^{4m+3d}h^{d/p-\sigma}\|\mathbf{a}\|_{\ell_p(\Xi)}$$

because the choice of  $K$  ensures  $J \geq d - \sigma$ . The theorem follows.  $\square$

For  $s \in \tilde{V}_\Xi = \text{span}_{\xi \in \Xi} b_\xi$  we may replace the discrete norm  $\|\mathbf{a}\|_{\ell_p(\Xi)}$  by its equivalent  $h^{-d/p}\|s\|_{L_p}$ , as we now show.

**Proposition 4.11** (Local Basis Stability and Nikolskii Inequality). *For every  $\rho_0 \geq 1$  there exists a constant  $h_0 > 0$  so that if  $\Xi \subset \Omega$  has fill distance  $h(\Xi, \Omega) \leq h_0$  and mesh ratio  $\rho \leq \rho_0$ , then the family of local Lagrange functions generated with  $K > \frac{10m-2}{\nu}$  satisfies the bounds*

$$(4.15) \quad c\|\mathbf{a}\|_{\ell_p(\Xi)} \leq q^{-d/p}\|s\|_{L_p(\Omega)} \leq C\rho^{m+d/p}\|\mathbf{a}\|_{\ell_p(\Xi)}$$

for all  $s = \sum_{\xi \in \Xi} a_\xi b_\xi \in \tilde{V}_\Xi$ , with  $c = c(\rho, m, \Omega)$  and  $C = C(\Omega, m)$ . In addition, for  $1 \leq p, r \leq \infty$ , we have

$$(4.16) \quad \|s\|_{L_p(\Omega)} \leq Cq^{-d(\frac{1}{r}-\frac{1}{p})_+}\|s\|_{L_r(\Omega)}$$

with  $C = C(p, r, \rho, m, \Omega, \cdot)$ .

*Proof.* The upper bound follows from the previous theorem, with  $\sigma = 0$ . To obtain the lower bound, note that

$$q^{-d/p}\|s\|_{L_p(\Omega)} = q^{-d/p}\left\| \sum_{\xi \in \Xi} a_\xi (b_\xi - \chi_\xi) + \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{L_p(\Omega)}.$$

Consequently, by (4.1), (3.7) and (4.14), we have

$$\begin{aligned} q^{-d/p}\|s\|_{L_p(\Omega)} &\geq q^{-d/p}\left( \left\| \sum_{\xi \in \Xi} a_\xi \chi_\xi \right\|_{L_p(\Omega)} - C\rho^{4m+3d}h^{J-d(\frac{p-1}{p})}\|\mathbf{a}\|_{\ell_p(\Xi)} \right) \\ &\geq (c_1 - C\rho^{4m+3d}h^{J-d})\|\mathbf{a}\|_{\ell_p(\Xi)}, \end{aligned}$$

where  $c_1 = c_1(\rho, \Omega, m)$  is the constant from Proposition 3.2. Let  $h_0 > 0$  be such that  $c_1 - C\rho^{4m+3d}h_0^{J-d} \geq \frac{1}{2}c_1$ . This guarantees the same holds for all  $0 < h \leq h_0$ . The proof of the Nikolskii inequality is, *mutatis mutandis*, that of Corollary 3.3.  $\square$

5. MAIN RESULTS AND COROLLARIES

At this point we can prove the inverse inequality for local Lagrange functions in  $\tilde{V}_\Xi$ .

**Theorem 5.1.** *Suppose  $\Omega \subset \mathbb{R}^d$  is a bounded Lipschitz region. For  $m > d/2$  and for every  $\rho_0 > 0$  there exists a constant  $h_0 > 0$ , so that if  $\Xi \subset \Omega$  has mesh ratio  $\rho \leq \rho_0$ , fill distance  $h \leq h_0$ , and if  $\tilde{\Xi} \subset \tilde{\Omega}$  is a suitable extension of  $\Xi$  (as mentioned in Remark 2.3), then for all  $s = \sum_{\xi \in \Xi} a_\xi b_\xi \in \tilde{V}_\Xi$  the following holds. For  $1 \leq p < \infty$  and all  $0 \leq \sigma \leq m - (d/2 - d/p)_+$ , or for  $p = \infty$  and an integer  $\sigma < m - d/2$ , we have*

$$\|s\|_{W_p^\sigma(\Omega)} \leq Ch^{-\sigma} \|s\|_{L_p(\Omega)}$$

with  $C = C(m, \rho, \Omega)$ .

*Proof.* This is an immediate combination of Theorem 4.10 and Proposition 4.11.  $\square$

**5.1. Restriction to the boundary.** Immediate applications of Theorem 5.1 are the following ‘trace’ estimates. (Since the elements of  $\tilde{V}_\Xi$  are continuous, it is appropriate to consider these results about restriction to the boundary.) To make sense of these, we first need to describe Sobolev spaces on the boundary  $\partial\Omega$ .

**5.1.1. Smoothness spaces on  $\partial\Omega$ .** We use the common tactic of employing a partition of unity with corresponding changes of variable to flatten the boundary. (As in [30, 1.11] and [20], for instance.) The details of the partition of unity and change of variable depends on the smoothness of the boundary, and this influences the types of Sobolev spaces we can define (namely, the maximum order of smoothness is governed by the smoothness of the boundary).

For a domain whose boundary is Lipschitz we consider a partition of unity  $(\psi_j)_{j=1}^N$  of  $\partial\Omega$ , where each  $\psi_j : \partial\Omega \rightarrow [0, 1]$  is Lipschitz, and let  $(U_j, h_j)_{j=1}^N$  be a corresponding collection of bi-Lipschitz charts so that each  $U_j$  is an open set in  $\partial\Omega$  containing the closure of  $\text{supp}(\psi_j)$  and each  $h_j : U_j \rightarrow \mathcal{O}_j \subset \mathbb{R}^{d-1}$  is a bijective Lipschitz function. Then for  $1 \leq p < \infty$  and  $0 < \sigma \leq 1$ , the Sobolev space  $W_p^\sigma(\partial\Omega)$  consists of functions  $f \in L_p(\partial\Omega)$  such that

$$(5.1) \quad \|f\|_{W_p^\sigma(\partial\Omega)}^p := \sum_{j=1}^N \|(\psi_j \circ (h_j^{-1}))(f \circ (h_j^{-1}))\|_{W_p^\sigma(\mathcal{O}_j)}^p$$

is finite.

For higher orders of smoothness, we simply increase the smoothness of the boundary, and the partition of unity and chart. For  $\sigma < M$ , let  $(\psi_j)_{j=1}^N$  be a  $C^M$  partition of unity of  $\partial\Omega$ , and let  $(U_j, h_j)_{j=1}^N$  be a collection  $C^M$  charts. Then  $W_p^\sigma(\partial\Omega)$  consists of functions for which the norm (5.1) is finite.

We note that this transporting of norms from Euclidean space to manifold by way of partition of unity and pull-back can be carried out for other smoothness spaces. In particular, it holds as well for the Besov class (see again [30] and [20]). For this reason, it follows that for fractional  $\sigma$ ,  $W_p^\sigma(\partial\Omega) = B_{p,p}^\sigma(\partial\Omega)$  with equivalence of norms (as in the Euclidean case).

**5.2. Trace estimates.** We may use Theorem 5.1 to obtain the following trace estimate for functions in  $\tilde{V}_\Xi$ . This is nonstandard because the norms of the trace are bounded by  $L_p$  norms rather than Sobolev norms.

**Corollary 5.2.** *Under hypotheses of Theorem 5.1, for  $s \in \tilde{V}_\Xi$  we have, for  $1 \leq p \leq \infty$  and  $0 < \sigma \leq m - 1/p - (d/2 - d/p)_+$ ,*

$$\|s\|_{W_p^\sigma(\partial\Omega)} \leq Ch^{-\sigma-1/p} \|s\|_{L_p(\Omega)}$$

with  $C = C(m, \rho, \Omega)$ .

*Proof.* For  $\sigma > 0$  we have that  $W_p^{\sigma+1/p}(\Omega) = B_{p,p}^{\sigma+1/p}(\Omega)$  and  $W_p^\sigma(\partial\Omega) = B_{p,p}^\sigma(\partial\Omega)$ .

It follows that  $\text{Tr} : W_p^{\sigma+1/p}(\Omega) \rightarrow W_p^\sigma(\partial\Omega)$  is bounded by the trace theorem (one will find a suitable one for smooth boundaries in [29, 3.3.3], and for Lipschitz boundaries in [20, Theorem 2.1]) so

$$\|s\|_{W_p^\sigma(\partial\Omega)} \leq C_\Omega \|s\|_{W_p^{\sigma+1/p}(\Omega)} \leq C_{\Omega,\rho} h^{-\sigma-1/p} \|s\|_{L_p(\Omega)}.$$

The first inequality is from the trace theorem, while the second follows from Lemma B.6. □

We can get a similar estimate for  $\sigma = 0$ , although this requires a modified trace result.

**Lemma 5.3.** *Suppose  $\Omega$  is compact with  $C^1$  boundary. For  $1 < p < \infty$  there is a constant  $C_p$  so that for all  $u \in C^1(\bar{\Omega})$  and  $\epsilon > 0$  we have*

$$\|u\|_{L_p(\partial\Omega)} \leq C_p(\epsilon^{-q/p} \|u\|_{L_p(\Omega)}^p + \epsilon \|u\|_{W_p^1(\Omega)}^p)$$

with  $q = \frac{p}{p-1}$ .

*Proof.* Note that in this case, we consider Sobolev norms over  $\Omega$ , so for the norms on the right-hand side, we make use of the definition given in Section 2.4. The  $L_p(\partial\Omega)$  norm on the left is with respect to surface measure, but this can be estimated in a standard way (by partition of unity and change of variables).

We begin by proving a trace result for  $\Omega = \mathbb{R}_+^d = \mathbb{R}^{d-1} \times [0, \infty)$ . For  $u \in C^1(\mathbb{R}_+^d)$  having compact support and  $x' \in \mathbb{R}^{d-1}$ , let  $r_{x'}$  be the first positive zero of  $t \mapsto |u(x', t)|$ . Then

$$\begin{aligned} |u(x', 0)|^p &\leq \int_0^{r_{x'}} \partial_d |u(x', x_d)|^p dx_d \\ &\leq p \int_0^\infty |u(x', x_d)|^{p-1} |\partial_d(u(x', x_d))| dx_d \\ &\leq \int_0^\infty C(\epsilon) |u(x', x_d)|^{(p-1)\frac{p}{p-1}} + \epsilon |\partial_d(u(x', x_d))|^p dx_d. \end{aligned}$$

The last line uses Young's inequality  $ab \leq C(\epsilon)a^q + \epsilon b^p$  with  $C(\epsilon) = q^{-1}(\epsilon p)^{-q/p}$ . Integrating this over  $\mathbb{R}^{d-1}$ , we have

$$(5.2) \quad \|u\|_{L_p(\mathbb{R}^{d-1})}^p \leq C_p \epsilon^{-q/p} \|u\|_{L_p(\mathbb{R}_+^d)}^p + \epsilon \|u\|_{W_p^1(\mathbb{R}_+^d)}^p.$$

Now let  $(\Psi_j)_{j=1}^N$  be a finite collection of nonnegative, compactly supported,  $C^1$  functions so that  $\sum \Psi_j = 1$  in a small neighborhood of  $\partial\Omega$ . Let  $(U_j)_{j=1}^N$  denote a corresponding collection of open sets so that  $\text{supp}(\Psi)_j \subset U_j$  and so that there is  $h_j : U_j \rightarrow B(0, \epsilon_j)$ , an open ball in  $\mathbb{R}^d$ .

For  $f \in C^1(\bar{\Omega})$  and  $1 \leq j \leq N$ , the product  $\Psi_j f$  is compactly supported and (extending by 0) we have  $u_j := (\Psi_j f) \circ (h_j^{-1}) \in C_1(\mathbb{R}_+^d)$ . Applying (5.2) to  $(\Psi_j f) \circ h_j^{-1}$  gives  $\|u_j\|_{L_p(\mathbb{R}^{d-1})}^p \leq C_p \epsilon^{-q/p} \|u_j\|_{L_p(\mathbb{R}^d)}^p + \epsilon \|u_j\|_{W_p^1(\mathbb{R}^d)}^p$ . Because  $\Psi_j$  and  $h_j^{-1}$  are  $C^1$  over compact sets, their norms can be bounded independent of  $j$ . By applying chain and product rules, it follows that

$$(5.3) \quad \sum_{j=1}^N \|u_j\|_{L_p(\mathbb{R}^{d-1})}^p \leq C_{p,q} (\epsilon^{-q/p} \|u\|_{L_p(\Omega)}^p + \epsilon \|u\|_{W_p^1(\Omega)}^p)$$

with an increased constant which depends on  $\max_{1 \leq j \leq N} \|(h_j)^{-1}\|_{C_1(h_j^{-1}(\text{supp}(\Psi_j))^-)}$  and  $\max_{1 \leq j \leq N} \|\Psi_j\|_{C_1(\mathbb{R}^d)}$  as well as that of (5.2). Because  $(\Psi_j|_{\partial\Omega})$  is a partition of unity for  $\partial\Omega$ , the left-hand side of (5.3) controls the  $L_p$  norm of  $u|_{\partial\Omega}$ , which gives the  $\epsilon$ -modified trace inequality

$$\|u\|_{L_p(\partial\Omega)}^p \leq C_{p,q} (\epsilon^{-q/p} \|u\|_{L_p(\Omega)}^p + \epsilon \|u\|_{W_p^1(\Omega)}^p). \quad \square$$

**Corollary 5.4.** *Let  $\Omega$  be a bounded domain with  $C^1$  boundary and assume the hypotheses of Theorem 5.1. For  $s \in \tilde{V}_\Xi$ , for  $1 \leq p \leq \infty$  and for  $1 + (d/2 - d/p)_+ \leq m$ , we have that*

$$\|s\|_{L_p(\partial\Omega)} \leq Ch^{-1/p} \|s\|_{L_p(\Omega)}$$

with  $C = C(p, \rho, m, \Omega)$ .

*Proof.* For  $p = 1$  the theorem follows directly from the boundedness of trace from  $W_1^1(\Omega)$  to  $L_1(\partial\Omega)$  (see [6, Theorem 1, Chapter 5.5]) and by repeating the argument of Theorem B.6.

For  $1 < p < \infty$ , we apply Lemma 5.3 with  $\epsilon = h^{p-1}$  (so that  $\epsilon^{-q/p} = h^{-\frac{p-1}{p}}$ ) followed by Theorem B.6. Thus,

$$\begin{aligned} \|s\|_{L_p(\partial\Omega)}^p &\leq C(h^{-1} \|s\|_{L_p(\Omega)}^p + h^{p-1} \|s\|_{W_p^1(\Omega)}^p) \\ &\leq C(h^{-1} \|s\|_{L_p(\Omega)}^p + h^{-1} \|s\|_{L_p(\Omega)}^p) \end{aligned}$$

and the result follows by taking the  $p$ th root. □

#### APPENDIX A. ENERGY AND POINTWISE BOUNDS ON THE LAGRANGE FUNCTION

In this section, we show that Lagrange functions for surface splines and Matérn kernels satisfy decay estimates as in Section 3.1.

We say that  $\Omega$  satisfies an interior cone condition if there are constants  $\varphi \in (0, \pi/2)$  and  $0 < R < \infty$  so that for every  $x \in \Omega$  there is a cone  $\mathcal{C}_{\vec{n}} = \{y \mid |y - x| \leq R, n \cdot (\frac{y-x}{|y-x|}) \geq \cos \varphi\}$  opening in the direction determined by the unit vector  $\vec{n}$  so that  $\mathcal{C}_{\vec{n}} \subset \Omega$ .

We recall the zeros estimate [18, Theorem A.11] for a bounded region  $\Omega$  with Lipschitz boundary (the version we cite is a streamlined modification of an earlier estimate given in [25, Theorem 2.12]).

**Lemma A.1** (Zeros estimate). *Let  $1 \leq p \leq \infty$  and  $m > d/p$  (when  $p = 1$  we may have  $m \geq d/p$ ). Suppose  $\Omega$  satisfies a cone condition with aperture  $\varphi$  and radius  $R$ . Then there are constants  $h_1$  (depending on  $m$  and  $\varphi$ ) and  $\Lambda$  (depending on  $m, d, p, \varphi$ ) so that if  $X \subset \Omega$  has fill distance  $h \leq h_1 R$  and if  $u \in W_p^m(\Omega)$  satisfies  $u|_X = 0$ , then*

$$\|u\|_{W_p^k(\Omega)} \leq \Lambda h^{m-k} \|u\|_{W_p^m(\Omega)}$$

and

$$\|u\|_{L^\infty(\Omega)} \leq \Lambda h^{m-d/p} \|u\|_{W_p^m(\Omega)}.$$

An important feature of this lemma is that the density  $h$  is controlled by the cone radius  $R$ , but the constant  $\Lambda$  does not depend on  $R$ . This allows a comparison of results across sets which are geometrically related. For example, annuli  $B(x, r_2) \setminus B(x, r_1)$  satisfy cone conditions with aperture  $\varphi$  independent of  $r_2$  and  $r_1$ , and with cone radius equal to half the thickness  $\frac{r_2-r_1}{2}$ . Thus, the above result holds for any point set with  $h \leq h_1(r_2 - r_1)/2$ . With almost no modification, this result extends to balls  $B(x, r)$  (where  $h \leq h_1 r$ ) and complements of balls (where there is no restriction on  $h$ ).

Consider now the annulus  $\mathbf{a}(\xi, r, t) := \{x \in \mathbb{R}^d \mid r - t < |x - \xi| \leq r\}$ . Applying Lemma A.1 with  $p = 2$  and  $k = m - 1$ , we estimate the Sobolev norm<sup>5</sup> as  $\|u\|_{W_2^m(\mathbf{a})}^2 \leq |u|_{W_2^m(\mathbf{a})}^2 + m\Lambda^2 h^2 \|u\|_{W_2^{m-1}(\mathbf{a})}^2$  which, after rearranging terms, implies that  $\|u\|_{W_2^m(\mathbf{a})}^2 \leq \frac{1}{1-m\Lambda^2 h^2} |u|_{W_2^m(\mathbf{a})}^2$  for  $u$  vanishing on  $X \subset \mathbf{a}$  with  $h \leq h_1 t/2$ . In short, if  $h \leq \min(\frac{h_1 t}{2}, h_2)$  with

$$(A.1) \quad h_2 := (\sqrt{2m}\Lambda)^{-1},$$

then

$$(A.2) \quad |u|_{W_2^k(\mathbf{a})} \leq \|u\|_{W_2^k(\mathbf{a})} \leq \Lambda h^{m-k} \|u\|_{W_2^m(\mathbf{a})} \leq 2\Lambda h^{m-k} |u|_{W_2^m(\mathbf{a})}$$

for  $u$  vanishing on  $X$ .

**Lemma A.2.** *Suppose  $m > d/2$ . There is a constant  $\nu = \nu(m, d)$  with  $\nu < 1$  such that if  $X \subset \mathbb{R}^d$  is a finite point set,  $\mathbf{a} = \mathbf{a}(\xi, r, t)$  is the annulus of outer radius  $r$ , width  $t$  and center  $\xi \in X$ , and  $X_0 = X \cap \mathbf{a}$  has fill distance  $h = h(X_0, \mathbf{a}) \leq \min(\frac{h_1 t}{2}, h_2)$ , then:*

- the Matérn Lagrange function  $\chi_\xi \in \text{span}\{\kappa_m(\cdot - \zeta) \mid \zeta \in X\}$  satisfies

$$\|\chi_\xi\|_{W_2^m(\mathbb{R}^d \setminus B(\xi, r))} \leq \nu \|\chi_\xi\|_{W_2^m(\mathbb{R}^d \setminus B(\xi, r-t))};$$

- the Lagrange function  $\chi_\xi \in S(\phi_m, X)$  for the order  $m$  surface spline satisfies

$$|\chi_\xi|_{W_2^m(\mathbb{R}^d \setminus B(\xi, r))} \leq \nu |\chi_\xi|_{W_2^m(\mathbb{R}^d \setminus B(\xi, r-t))}.$$

*Proof.* In either case, the function  $k_m$  is the reproducing kernel for a (semi-)Hilbert space (described in Sections 2.5.1 and 2.5.2), and we use the notation  $[u]_{m, Y}$  to denote  $\|u\|_{W_2^m(Y)}$  or  $|u|_{W_2^m(Y)}$ , respectively.

Let  $\tau : \mathbb{R} \rightarrow [0, 1]$  be a smooth cut-off function supported on the interval  $(-\infty, 1)$  equaling 1 on  $(-\infty, 0]$ . For  $r > t$ , we define  $\tau_{\xi, r, t} : \mathbb{R}^d \rightarrow \mathbb{R}$  as  $\tau_{\xi, r, t}(x) = \tau(\frac{1}{t}(|x - \xi| - (r - t)))$ , and note that it is a smooth function supported in the ball  $B(\xi, r)$ , and equals 1 in  $B(\xi, r - t)$ . By the chain rule, there is a bound  $\|D^\beta \tau_{\xi, r, t}\|_\infty \leq C t^{-|\beta|}$  which is independent of  $r$ .

Both  $\chi_\xi$  and  $\tau_{\xi, r, t} \chi_\xi$  are Lagrange functions on  $X$ . Thus  $[\chi_\xi]_m \leq [\tau_{\xi, r, t} \chi_\xi]_m$ . Using the additivity of  $[\cdot]_m$ , and noting that the Lagrange functions are identical on  $B(\xi, r - t)$  while  $\tau_{\xi, r, t}$  vanishes outside  $B(\xi, r)$ , we have

$$[\chi_\xi]_m^2 \leq [\tau_{\xi, r, t} \chi_\xi]_m^2 \quad \longrightarrow \quad [\chi_\xi]_{m, \mathbb{R}^d \setminus B(\xi, r-t)}^2 \leq [\tau_{\xi, r, t} \chi_\xi]_{m, \mathbf{a}(\xi, r, t)}^2.$$

---

<sup>5</sup>Recall that we use the Sobolev norm as defined in Section 2.4—in particular, the  $k$ th order partial derivatives are weighted by  $\binom{m}{k}$ .



By using Hölder’s inequality in conjunction with the product rule, we have

$$\begin{aligned}
 \int_{\mathbf{a}} |D^\alpha(\tau_{\xi,r,t}(x)\chi_\xi(x))|^2 dx &= \int_{\mathbf{a}} \left| \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^{\alpha-\beta} \tau_{\xi,r,t}(x) D^\beta \chi_\xi(x) \right|^2 dx \\
 &\leq C \sum_{\beta \leq \alpha} t^{-2|\alpha-\beta|} \int_{\mathbf{a}} |D^\beta \chi_\xi(x)|^2 dx \\
 \text{(A.3)} \qquad \qquad \qquad &\leq C \sum_{\beta \leq \alpha} h^{-2(|\alpha|-|\beta|)} \int_{\mathbf{a}} |D^\beta \chi_\xi(x)|^2 dx.
 \end{aligned}$$

In the last line we have  $t^{-|\alpha-\beta|} = t^{-|\alpha|+|\beta|} \leq h_0^{|\alpha|-|\beta|} h^{-|\alpha|+|\beta|}$  which follows from the fact that  $\beta \leq \alpha$ . Applying (A.2) to (A.3) gives, for each  $\beta \leq \alpha$ ,  $\int_{\mathbf{a}} |D^\beta \chi_\xi(x)|^2 dx \leq C^2 h^{2(m-|\beta|)} \|\chi_\xi\|_{W_2^m(\mathbf{a})}$ . This yields the inequality

$$[\tau_{\xi,r,t}\chi_\xi]_{m,\mathbb{R}^d \setminus B(\xi,r-t)}^2 \leq C \sum_{|\alpha|=m} \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} (h^{-2|\alpha|+2|\beta|}) C^2 h^{2(m-|\beta|)} |\chi_\xi|_{W_2^m(\mathbf{a})}^2.$$

Canceling powers of  $h$  and collecting constants which depend only on  $m$  and  $d$ , we have

$$[\chi_\xi]_{m,\mathbb{R}^d \setminus B(\xi,r-t)}^2 \leq [\tau_{\xi,r,t}\chi_\xi]_{m,\mathbb{R}^d \setminus B(\xi,r-t)}^2 \leq C [\chi_\xi]_{m,\mathbf{a}}^2.$$

Finally, we note that  $[\chi_\xi]_{m,\mathbf{a}}^2 = [\chi_\xi]_{m,\mathbb{R}^d \setminus B(\xi,r-t)}^2 - [\chi_\xi]_{m,\mathbb{R}^d \setminus B(\xi,r)}^2$  which yields

$$[\chi_\xi]_{m,\mathbb{R}^d \setminus B(\xi,r)}^2 \leq \frac{C-1}{C} [\chi_\xi]_{m,\mathbb{R}^d \setminus B(\xi,r-t)}^2$$

and the lemma follows with  $\nu = \sqrt{\frac{C-1}{C}} < 1$ . □

We may now iterate Lemma A.2 to get control of the ‘energy’ of the tail of the Lagrange functions.

**Lemma A.3.** *Suppose  $D \subset \mathbb{R}^d$  is bounded, and  $X \subset D$  is a finite point set with fill distance satisfying  $h(X, D) \leq h_2$ . There is  $\mu = \mu(m, d) > 0$  so that for  $R < \text{dist}(\xi, \partial D)$ :*

- the Matérn Lagrange function  $\chi_\xi \in \text{span} \{ \kappa_m(\cdot - \zeta) \mid \zeta \in X \}$  satisfies

$$\|\chi_\xi\|_{W_2^m(\mathbb{R}^d \setminus B(\xi,R))} \leq C q^{d/2-m} \exp\left(-\mu \frac{R}{h}\right);$$

- the Lagrange function  $\chi_\xi \in S(\phi_m, X)$  for the order  $m$  surface spline satisfies

$$|\chi_\xi|_{W_2^m(\mathbb{R}^d \setminus B(\xi,R))} \leq C q^{d/2-m} \exp\left(-\mu \frac{R}{h}\right);$$

*Proof.* Setting  $t = 4h/h_1$  (where  $h_1$  is the constant appearing in Lemma A.1), consider, for  $r \leq \text{dist}(\xi, \partial D)$ , an annulus  $\mathbf{a}(\xi, r, t)$  and the restricted point set  $X_0 = X \cap \mathbf{a}(\xi, r, t)$ . The slightly smaller, inner annulus  $\mathbf{a}(\xi, r-h, t-2h)$  has the property that for every  $x \in \mathbf{a}(\xi, r-h, t-2h)$ , there is  $\zeta \in X_0$  so that  $\text{dist}(x, \zeta) \leq h$  (since in that case  $\text{dist}(x, X_0) = \text{dist}(x, X)$ ). It follows that  $h(X_0, \mathbf{a}(\xi, r, t)) \leq 2h$  and therefore  $h(X_0, \mathbf{a}(\xi, r, t)) \leq \frac{h_1 t}{2}$ .

Now letting  $n = \lfloor R/t \rfloor$ , by Lemma A.2 we have that

$$|\chi_\xi|_{W_2^m(\mathbb{R}^d \setminus B(\xi,R))} \leq \nu |\chi_\xi|_{W_2^m(\mathbb{R}^d \setminus B(\xi,R-t))} \leq \dots \leq \nu^n |\chi_\xi|_{W_2^m(\mathbb{R}^d)}$$

and the last expression can be bounded by  $\nu^{-1}\nu^{\frac{h_0 R}{4h}}|\chi_\xi|_{W_2^m(\mathbb{R}^d)}$ . By the ‘bump estimate’ (3.2), we have that  $|\chi_\xi|_{W_2^m(\mathbb{R}^d)} \leq Cq^{d/2-m}$ , so the lemma follows with  $\mu = -\frac{h_0}{4}\log(\nu)$ , which is positive since  $\nu < 1$ .  $\square$

Note that if  $\Omega \subset \mathbb{R}^d$  is compact, then  $\tilde{\Omega} = \{x \in \mathbb{R}^d \mid \text{dist}(x, \Omega) \leq \text{diam}(\Omega)\}$  automatically satisfies a cone condition (with radius  $R = \text{diam}(\Omega)$  and aperture independent of  $\Omega$ ). Thus, the result (3.4) follows with  $D = \tilde{\Omega}$ ,  $X = \tilde{\Xi}$ ,  $h_0 = h_0(d, m)$  and  $\xi \in \tilde{\Xi}$ .

Because  $x \in \Omega$ ,  $\xi \in \Xi$  implies  $R = |\xi - x| \leq \text{dist}(\xi, \partial\Omega)$ , we can apply the second part of the zeros estimate Lemma A.1 to Lemma A.3 to obtain the pointwise estimate (3.5)

(A.4)

$$|\chi_\xi(x)| \leq Ch^{m-d/2}\|\chi_\xi\|_{W_2^m(\mathbb{R}^d \setminus B(\xi, \text{dist}(\xi, x)))} \leq C\rho^{m-d/2}\exp\left(-\mu\frac{\text{dist}(x, \xi)}{h}\right).$$

Note that in the second inequality we have written  $h^{m-d/2}q^{d/2-m} = \rho^{m-d/2}$ .

### APPENDIX B. STABILITY BOUNDS FOR THE LAGRANGE FUNCTION

We now demonstrate that the family of Lagrange functions for suitable kernels over a domain  $\Omega$  satisfy stability bounds of the form (3.7). This was demonstrated in [16, Proposition 3.6 and Theorem 3.7]; we follow the argument presented there, with modifications for dealing with a suitably bounded Euclidean domain, and to obtain a necessary refinement: that the threshold fill distance  $h_0$  depends only on  $m$  and  $d$  (and not on  $\rho$  or  $\Omega$ ).

**Lemma B.1.** *Suppose  $\Omega$  is a bounded domain and  $\Xi \subset \Omega$  is a finite subset with fill distance  $h \leq h_2$ , where  $h_2 = h_2(m, d)$  is the constant given in (A.1). There exists a constant  $c_2 = c_2(m, d)$  so that the family of functions  $(\chi_\xi)_{\xi \in \Xi}$  have the property that for  $s = \sum_{\xi \in \Xi} a_\xi \chi_\xi$ ,*

$$\|s\|_p \leq c_2\rho^{m+d/2}q^{d/p}\|\mathbf{a}\|_{\ell_p(\Xi)}$$

holds.

*Proof.* For  $p = \infty$ , inequality (A.4) leads to a bound on the Lebesgue constant  $\mathcal{L}$  for the  $\chi_\xi$ ’s over  $\Omega$ :

$$(B.1) \quad \mathcal{L} := \sup_{x \in \Omega} \left(\sum_{\xi \in \Xi} |\chi_\xi(x)|\right) < C\rho^{m+d/2}, \quad C = C(m, d).$$

Indeed, for fixed  $x \in \Omega$  we note that  $\sum_{\xi \in \Xi} |\chi_\xi(x)| \leq C\rho^{m-d/2} \sum_{\xi \in \Xi} e^{-\mu\frac{\text{dist}(x, \xi)}{h}}$ . By estimating in annuli, using sets  $A_n := \{\xi \in \Xi \mid h(n-1) < |x - \xi| \leq hn\}$  having  $\#A_n \leq C\left(\frac{hn}{q}\right)^d$ , we have that

$$\sum_{\xi \in \Xi} |\chi_\xi(x)| \leq C\rho^{m-d/2}\left(1 + \sum_{n=1}^{\infty} \rho^d n^d e^{-\mu n}\right) < C\rho^{m+d/2},$$

where the constant  $C = C(m, d)$  is independent of  $\Xi$ . Taking the supremum then yields (B.1). It follows that

$$\|s\|_{L_\infty(\Omega)} \leq \mathcal{L}\|s|_\Xi\|_{\ell_\infty(\Xi)} = \mathcal{L}\|\mathbf{a}\|_{\ell_\infty(\Xi)}.$$

For  $p = 1$ , we have

$$\int_{\Omega} |s(x)| dx \leq \sum_{\xi \in \Xi} |a_{\xi}| \int_{\Omega} |\chi_{\xi}(x)| dx \leq Ch^d \| \mathbf{a} \|_{\ell_1(\Xi)}.$$

Here we have used the fact that  $\| \chi_{\xi} \|_1 \leq C\rho^{m-d/2}h^d \leq C\rho^{m+d/2}q^d$ , which follows by integrating (A.4). A standard application of operator interpolation proves the other cases.  $\square$

*Preliminary estimates.* Because  $\Omega$  satisfies a cone condition with aperture  $\varphi$  and radius  $R$ , there is a constant  $\alpha$  (depending only on  $d$  and  $\varphi$ ) so that for all  $x \in \Omega$ ,

$$\alpha r^d \leq \text{vol}(B(x, r) \cap \Omega)$$

for  $r \leq R$  (the radius of the cone condition). Similarly, we have that there is a constant  $K$  (depending only on  $d$  and  $\varphi$ ) so that

$$(B.2) \quad \#(\Xi \cap B(x, r)) \leq K(r/q)^d.$$

We can use a simple modification of the zeros lemma [16, Lemma 7.1] valid for balls, which states that there exists a constant  $h_3 > 0$  depending only on  $m$  and  $d$  so that for  $h \leq h_3$ , the Hölder-like condition

$$|\chi_{\xi}(x) - \chi_{\xi}(y)| \leq C\rho^{m-d/2} \left( \frac{|x - y|}{q} \right)^{\epsilon}$$

holds for  $0 < \epsilon < m - d/2$  and with a constant  $C = C(d, m)$ .

*Remark B.2.* For the remainder of this appendix, we assume  $\Xi$  is sufficiently dense such that  $h(\Xi, \Omega) \leq h_0 := \min(h_2, h_3)$ . We note that  $h_0$  depends only on  $d$  and  $m$  (because this is true for  $h_2$  and  $h_3$ ).

This permits us to understand the structure of  $\chi_{\xi}$  around the centers  $\zeta \in \Xi$ . Namely, because  $\chi_{\xi}(\xi) = 1$ ,

$$\chi_{\xi}(x) \geq \frac{2}{3} \quad \text{for } x \text{ in } B(\xi, \gamma q)$$

whenever  $\gamma^{\epsilon} \leq 1/(3C\rho^{m-d/2})$ . For the off-center case (i.e., when  $\zeta \neq \xi$ ),

$$(B.3) \quad |\chi_{\zeta}(x)| \leq C\rho^{m-d/2}\gamma^{\epsilon} \quad \text{for } x \text{ in } B(\xi, \gamma q)$$

with a constant  $C = C(d, m)$ .

Now fix  $0 < \gamma \leq 1/(3C\rho^{m-d/2})^{1/\epsilon}$  and define  $B_{\xi} := \Omega \cap B(\xi, \gamma q)$ . The above estimate guarantees that  $\alpha(\gamma q)^d \leq \text{vol}(B_{\xi})$ , from which  $\alpha(\gamma q)^d \left(\frac{2}{3}\right)^p \leq \int_{B_{\xi}} |\chi_{\xi}(x)|^p dx$ , implies

$$\alpha(\gamma q)^d \left(\frac{2}{3}\right)^p \sum_{\xi \in \Xi} |a_{\xi}|^p \leq \sum_{\xi \in \Xi} \int_{B_{\xi}} |a_{\xi} \chi_{\xi}(x)|^p dx.$$

This is the starting point for the corresponding lower bound to Lemma B.1, since the quasi-triangle inequality  $(A + B)^p \leq 2^{p-1}(A^p + B^p)$  implies that  $|a_{\xi} \chi_{\xi}(x)|^p \leq 2^{p-1} \left( \left| \sum_{\zeta \in \Xi} a_{\zeta} \chi_{\zeta}(x) \right|^p + \left| \sum_{\zeta \neq \xi} a_{\zeta} \chi_{\zeta}(x) \right|^p \right)$  and so

$$(B.4) \quad \alpha(\gamma q)^d \left(\frac{2}{3}\right)^p \sum_{\xi \in \Xi} |a_{\xi}|^p \leq 2^{p-1} \sum_{\xi \in \Xi} \int_{B_{\xi}} \left( \left| \sum_{\zeta \in \Xi} a_{\zeta} \chi_{\zeta}(x) \right|^p + \left| \sum_{\zeta \neq \xi} a_{\zeta} \chi_{\zeta}(x) \right|^p \right) dx.$$

The desired lower bound is  $c_1 q^{d/p} \|\mathbf{a}\|_{\ell_p(\Xi)} \leq \|s\|_p$ , so we must estimate the size of the overestimated ‘off-diagonal’ terms:  $2^{p-1} \sum_{\xi \in \Xi} \int_{B_\xi} |\sum_{\zeta \neq \xi} a_\zeta \chi_\zeta(x)|^p dx$ .

*Controlling the off-diagonal terms.* This is done in two stages, by splitting

$$\sum_{\xi \in \Xi} \int_{B_\xi} \left| \sum_{\{\zeta \in \Xi: \zeta \neq \xi\}} a_\zeta \chi_\zeta(x) \right|^p dx = \sum_{\xi \in \Xi} (I_\xi + II_\xi),$$

where the first term  $I_\xi := \int_{B_\xi} \left| \sum_{\{\zeta \in \Xi: |\zeta - \xi| \geq \Gamma q\}} a_\zeta \chi_\zeta(x) \right|^p dx$  is the ‘far away’ contribution and the second  $II_\xi := \int_{B_\xi} \left| \sum_{\{\zeta \in \Xi: |\zeta - \xi| \leq \Gamma q\}} a_\zeta \chi_\zeta(x) \right|^p dx$  is the nearby contribution. These depend on an (as yet) undetermined parameter  $\Gamma > 0$ . First we use the exponential decay of (A.4) to control the far away portion of the off-diagonal part, Then we use the Hölder estimates (B.3) to bound the nearby portion.

**Lemma B.3.** *For every  $p \in [1, \infty)$  there is a function  $F : (0, \infty) \rightarrow \mathbb{R}$  satisfying  $\lim_{t \rightarrow \infty} F(t) = 0$  so that for every  $\Gamma > 0$ , we have the inequality*

$$\sum_{\xi \in \Xi} \int_{B_\xi} \left| \sum_{\{\zeta \in \Xi: |\zeta - \xi| \geq \Gamma q\}} a_\zeta \chi_\zeta(x) \right|^p dx \leq F(\Gamma) (\gamma q)^d \rho^{p(m - \frac{d}{2})} \|\mathbf{a}\|_{\ell_p(\Xi)}^p$$

holds with  $F(\Gamma) \leq \tilde{C} e^{-\frac{\mu}{2} p \Gamma}$  with  $\tilde{C} = \tilde{C}(m, d, p)$  and  $\mu = \mu(d, m)$  the constant from (3.4).

*Proof.* We sum over the nonoverlapping dyadic regions

$$\Omega_k := \Omega_k(\xi) := \{\zeta \in \Xi \mid \Gamma 2^k q \leq \text{dist}(\xi, \zeta) \leq \Gamma 2^{k+1} q\}, \quad k = 0, 1, \dots, N_q,$$

where  $2^{N_q} \sim \frac{\text{diam}(\Omega)}{\Gamma q}$ . This means that, for  $M_k := \int_{B_\xi} \left| \sum_{\zeta \in \Omega_k} a_\zeta \chi_\zeta(x) \right|^p dx$ ,

$$I_\xi \leq \sum_{k=0}^{N_q} 2^{(p-1)(k+1)} \int_{B_\xi} \left| \sum_{\zeta \in \Omega_k} a_\zeta \chi_\zeta(x) \right|^p dx = \sum_{k=0}^{N_q} 2^{(p-1)(k+1)} M_k,$$

where the above inequality follows by iterating the quasi-triangle inequality  $|A + B|^p \leq 2^{p-1}(|A|^p + |B|^p)$  to get  $\left| \sum_{j=1}^n A_j \right|^p \leq \sum_{j=1}^n 2^{j(p-1)} |A_j|^p$ .

We now estimate the contribution from each  $M_k$ , the portion of  $II_\xi$  coming from the dyadic interval  $\Omega_k$ . By using the (generalized quasi-triangle) inequality  $|\sum_{j=1}^n A_j|^p \leq n^{p-1} \sum |A_j|^p$ , we have

$$\begin{aligned} M_k &\leq (\#\Omega_k)^{p-1} \times \sum_{\zeta \in \Omega_k} \int_{B_\xi} |a_\zeta \chi_\zeta(x)|^p dx \\ &\leq (\#\Omega_k)^{p-1} \times \max_{\zeta \in \Omega_k} \|\chi_\zeta\|_{L_1(B_\xi)}^p \times \sum_{\zeta \in \Omega_k} |a_\zeta|^p \\ (B.5) \quad &\leq (K(2^{k+1}\Gamma)^d)^{p-1} C(\gamma q)^d \rho^{p(m - \frac{d}{2})} (\exp(-\mu p \Gamma 2^k)) \sum_{\zeta \in \Omega_k} |a_\zeta|^p. \end{aligned}$$

In the final line, we have used the estimates (B.2) and (A.4).

Multiplying by  $2^{(p-1)(k+1)}$  and summing from 0 to  $N_q$ , we obtain (after rearranging some terms and combining constants which depend only on  $d$  and  $p$ )

$$I_\xi \leq C(\gamma q)^d \rho^{p(m-\frac{d}{2})} \left( \sum_{k=0}^{N_q} (2^{k(d+1)} \Gamma^d)^{p-1} \exp(-\mu p \Gamma 2^k) \left[ \sum_{\zeta \in \Omega_k} |a_\zeta|^p \right] \right).$$

We can now sum over  $\xi$ , bounding  $\sum_{\xi \in \Xi} I_\xi$  by

$$\begin{aligned} & C(\gamma q)^d \rho^{p(m-\frac{d}{2})} \left( \sum_{k=0}^{N_q} (2^{k(d+1)} \Gamma^d)^{p-1} \exp(-\mu p \Gamma 2^k) \sum_{\xi \in \Xi} \left[ \sum_{\zeta \in \Omega_k} |a_\zeta|^p \right] \right) \\ & \leq C(\gamma q)^d \rho^{p(m-\frac{d}{2})} \left( \sum_{k=0}^{N_q} (2^{k(d+1)} \Gamma^d)^{p-1} \exp(-\mu p \Gamma 2^k) (K(2^{k+1} \Gamma)^d) \left[ \sum_{\zeta \in \Xi} |a_\zeta|^p \right] \right) \\ & \leq C \frac{(\gamma q)^d \rho^{p(m-\frac{d}{2})}}{\Gamma^p} \left( \sum_{k=0}^{N_q} (2^k \Gamma)^{(d+1)p} \exp(-\mu p \Gamma 2^k) \right) \left[ \sum_{\zeta \in \Xi} |a_\zeta|^p \right]. \end{aligned}$$

In the second inequality, we have exchanged summation over  $\xi$  and  $\zeta$ . In short, we have used

$$\sum_{\xi \in \Xi} \sum_{\zeta \in \Omega_k} [|a_\zeta|^p] = \sum_{\zeta \in \Xi} \sum_{\xi \in \Xi} \left[ \chi_{\Omega_k(\xi)}(\zeta) \sum_{\zeta \in \Omega_k} |a_\zeta|^p \right]$$

in conjunction with the estimate  $\#\{\xi: \zeta \in \Omega_k(\xi)\} \leq K(2^{k+1} \Gamma)^d$  obtained from (B.2), since for  $\zeta \in \Xi$ ,  $\#\{\xi: \zeta \in \Omega_k(\xi)\} = \#\Omega_k(\zeta)$ . In the final inequality, we have used the fact that  $2^{(k+1)d} \leq 2^{k(d+1)} \times 2^{d+1}$  and that  $\Gamma^{dp} = \frac{\Gamma^{(d+1)p}}{\Gamma^p}$ . We estimate this with an integral as

$$\begin{aligned} \sum_{\xi \in \Xi} I_\xi & \leq C \left( \Gamma^{dp} \exp(-\mu p \Gamma) \right. \\ & \left. + \frac{2}{\Gamma^p} \int_\Gamma^\infty \exp(-\mu p r) r^{(d+1)p-1} dr \right) (\gamma q)^d \rho^{p(m-\frac{d}{2})} \|\mathbf{a}\|_{\ell_p(\Xi)}^p. \end{aligned}$$

Which shows that  $F(\Gamma) := C \left( \Gamma^{dp} \exp(-\mu p \Gamma) + \frac{2}{\Gamma^p} \int_\Gamma^\infty \exp(-\mu p r) r^{(d+1)p-1} dr \right)$ .

The integral term can be bounded by making a change of variable  $R = r\Gamma$  as

$$\begin{aligned} \frac{2}{\Gamma^p} \int_\Gamma^\infty \exp(-\mu p r) r^{(d+1)p-1} dr & = 2\Gamma^{dp} \int_1^\infty \exp(-\mu p \Gamma R) R^{(d+1)p-1} dR \\ & \leq C_{d,p,m} \Gamma^{dp} \exp(-\mu p \Gamma). \end{aligned}$$

Because  $\max_{\Gamma > 1} \Gamma^{dp} \exp(-\frac{\mu}{2} p \Gamma) \leq C_{d,p,m}$ , the estimate  $F(\Gamma) \leq \tilde{C} e^{-\frac{\mu}{2} p \Gamma}$  follows.  $\square$

**Lemma B.4.** *For every  $p \in [1, \infty)$  and every  $\Gamma > 0$ , we have the inequality*

$$\sum_{\xi \in \Xi} \int_{B_\xi} \left| \sum_{\{\zeta \in \Xi: \zeta \neq \xi, |\zeta - \xi| \leq \Gamma q\}} a_\zeta \chi_\zeta(x) \right|^p dx \leq C(\Gamma^d \gamma^\epsilon)^p (\gamma q)^d \rho^{p(m-\frac{d}{2})} \|\mathbf{a}\|_{\ell_p(\Xi)}^p.$$

*Proof.* Note that  $\#\{\zeta \in \Xi : \zeta \neq \xi, |\zeta - \xi| \leq \Gamma q\} \leq K\Gamma^d$ , so by the quasi-triangle inequality, we have, for each  $\xi \in \Xi$ ,

$$\begin{aligned} II_\xi &\leq \int_{B_\xi} (K\Gamma^d)^{p-1} \sum_{\text{dist}(\zeta, \xi) \leq \Gamma q} |a_\zeta \chi_\zeta(x)|^p dx \\ &\leq \int_{B_\xi} (K\Gamma^d)^{p-1} (C\rho^{m-d/2}\gamma^\epsilon)^p \sum_{\text{dist}(\zeta, \xi) \leq \Gamma q} |a_\zeta|^p dx \\ &\leq C\Gamma^{d(p-1)}\gamma^{\epsilon p}\rho^{p(m-d/2)}(\gamma q)^d \sum_{\text{dist}(\zeta, \xi) \leq \Gamma q} |a_\zeta|^p. \end{aligned}$$

In the first inequality we use the estimate on the number of centers (B.2). In the second inequality, we use the bound (B.3). The third inequality follows from the simple estimate  $\text{vol}(B_\xi) \leq C(\gamma q)^d$ .

Summing over  $\xi \in \Xi$ , we obtain:

$$\begin{aligned} \sum_{\xi \in \Xi} II_\xi &\leq C\Gamma^{d(p-1)}\gamma^{\epsilon p}(\gamma q)^d \sum_{\xi \in \Xi} \sum_{\text{dist}(\zeta, \xi) \leq \Gamma q} |s(\zeta)|^p \\ &\leq C(\Gamma^d\gamma^\epsilon)^p (\gamma q)^d \rho^{p(m-\frac{d}{2})} \|\mathbf{a}\|_{\ell_p(\Xi)}^p. \end{aligned}$$

The final estimate results by exchanging the two summations, and employing the fact that  $\#\{\xi \in \Xi : \text{dist}(\zeta, \xi) \leq \Gamma q\} \leq K\Gamma^d$ . This completes the proof of the lemma.  $\square$

**Lemma B.5.** *There exists a constant  $\gamma$  satisfying  $\gamma \geq C\rho^{\frac{d-2m}{2\epsilon}}(\log(\rho))^{-d/\epsilon}$  with  $C(d, m, p, \epsilon)$ , so that*

$$2^{p-1} \sum_{\xi \in \Xi} \int_{B_\xi} \left| \sum_{\zeta \neq \xi} a_\zeta \chi_\zeta(x) \right|^p dx \leq \frac{1}{2} \alpha (\gamma q)^d \left(\frac{2}{3}\right)^p \sum_{\xi \in \Xi} |a_\xi|^p$$

holds for all  $\mathbf{a} \in \ell_p(\Xi)$  and all  $p \in [1, \infty)$ .

*Proof.* By the quasi-triangle inequality, we have

$$\sum_{\xi \in \Xi} \int_{B_\xi} \left| \sum_{\zeta \neq \xi} a_\zeta \chi_\zeta(x) \right|^p dx \leq 2^{p-1} \left( \sum_{\xi \in \Xi} I_\xi + \sum_{\xi \in \Xi} II_\xi \right).$$

Apply Lemma B.3, and choose  $\Gamma$  so that  $\tilde{C}e^{-\frac{\mu}{2}p\Gamma} = \frac{1}{4}\alpha\left(\frac{1}{3}\right)^p\rho^{-p(m-\frac{d}{2})}$ , where  $\tilde{C}$  is the constant appearing in Lemma B.3. We note that our choice of  $\Gamma$  guarantees  $\Gamma \leq C_{d,p,m} \log(\rho)$ . By Lemma B.3,  $F(\Gamma) \leq \frac{1}{4}\alpha\left(\frac{1}{3}\right)^p\rho^{-p(m-\frac{d}{2})}$ , it then follows that

$$\sum_{\xi \in \Xi} \int_{B_\xi} \left| \sum_{\{\zeta \in \Xi : |\zeta - \xi| \geq \Gamma q\}} a_\zeta \chi_\zeta(x) \right|^p dx \leq 4^{-p}\alpha\left(\frac{2}{3}\right)^p (\gamma q)^d \sum_{\xi \in \Xi} |a_\xi|^p.$$

Now select  $\gamma$  so that the inequalities  $0 < \gamma \leq 1/(3C\rho^{m-d/2})^{1/\epsilon}$  (with  $C$  from (B.3)) and  $C(\Gamma^d\gamma^\epsilon)^p \leq \frac{1}{4}\alpha\left(\frac{2}{3}\right)^p\rho^{-p(m-\frac{d}{2})}$  (with  $C$  from Lemma B.4) both hold. The problem of choosing  $\gamma$  can be rewritten as  $\gamma^\epsilon \leq \rho^{d/2-m} \min(\frac{1}{3C}, C_{d,m,p}\Gamma^{-d})$ .

Since  $\Gamma^{-d} > C_{d,p,m}(\log(\rho))^{-d}$ , it suffices to take  $\gamma^\epsilon = C_{d,m,p}\rho^{d/2-m}(\log(\rho))^{-d}$  for some constant  $C_{d,m,p}$ . For this choice of  $\gamma$ , Lemma B.4 guarantees that

$$\sum_{\xi \in \Xi} \int_{B_\xi} \left| \sum_{\{\zeta \in \Xi: \zeta \neq \xi, |\zeta - \xi| \leq \Gamma q\}} a_\zeta \chi_\zeta(x) \right|^p dx \leq 4^{-p} \alpha \left( \frac{2}{3} \right)^p (\gamma q)^d \sum_{\xi \in \Xi} |a_\xi|^p$$

as well.

Thus,

$$\begin{aligned} 2^{p-1} \sum_{\xi \in \Xi} \int_{B(\xi, \gamma q)} \left| \sum_{\zeta \neq \xi} a_\zeta \chi_\zeta(x) \right|^p dx &\leq 4^{p-1} \sum_{\xi \in \Xi} (I_\xi + II_\xi) \\ &\leq \left( \frac{1}{4} + \frac{1}{4} \right) \alpha (\gamma q)^d \left( \frac{2}{3} \right)^p \sum_{\xi \in \Xi} |a_\xi|^p \end{aligned}$$

and the result follows with  $\gamma \geq C_{d,m,p,\epsilon} \rho^{\frac{d-2m}{2\epsilon}} (\log(\rho))^{-d/\epsilon}$ . □

**Lemma B.6.** *Suppose  $\Omega$  is a bounded domain and  $\Xi \subset \Omega$  is a finite subset with fill distance  $h \leq h_0 := \min(h_2, h_3)$ , where  $h_0 = h_0(m, d)$ . There exists a constant then the family of functions  $(\chi_\xi)_{\xi \in \Xi}$  have the property that for any  $s = \sum_{\xi \in \Xi} a_\xi \chi_\xi$ ,*

$$c_1 q^{d/p} \| \mathbf{a} \|_{\ell_p(\Xi)} \leq \| s \|_p$$

holds with  $c_1 \geq C \rho^{\frac{d(d-2m)}{2\epsilon p}} (\log(\rho))^{-\frac{d^2}{p\epsilon}}$  with  $0 < \epsilon < m - d/2$  and  $C = C(d, p, m, \epsilon)$ .

*Proof.* Since  $s(\xi) = a_\xi$ , the  $L_\infty$  case follows immediately with constant 1. For  $1 \leq p < \infty$  we use (B.4) and Lemma B.5 to make the estimate

$$\begin{aligned} \alpha (\gamma q)^d \left( \frac{2}{3} \right)^p \sum_{\xi \in \Xi} |a_\xi|^p &\leq 2^{p-1} \sum_{\xi \in \Xi} \int_{B_\xi} \left( \left| \sum_{\zeta \in \Xi} a_\zeta \chi_\zeta(x) \right|^p + \left| \sum_{\zeta \neq \xi} a_\zeta \chi_\zeta(x) \right|^p \right) dx \\ &\leq \left( 2^{p-1} \sum_{\xi \in \Xi} \int_{B_\xi} \left| \sum_{\zeta \in \Xi} a_\zeta \chi_\zeta(x) \right|^p \right) dx \\ &\quad + \frac{1}{2} \alpha (\gamma q)^d \left( \frac{2}{3} \right)^p \sum_{\xi \in \Xi} |a_\xi|^p. \end{aligned}$$

Applying  $\sum_{\xi \in \Xi} \int_{B_\xi} \left| \sum_{\zeta \in \Xi} a_\zeta \chi_\zeta(x) \right|^p \leq \int_\Omega \left| \sum_{\zeta \in \Xi} a_\zeta \chi_\zeta(x) \right|^p dx$ , the result follows with

$$c_1 = \frac{1}{3} (\alpha \gamma^d)^{1/p} \geq C_{d,m,p,\epsilon} \rho^{\frac{d(d-2m)}{2\epsilon p}} (\log(\rho))^{-\frac{d^2}{p\epsilon}}. \quad \square$$

ACKNOWLEDGMENTS

The authors are grateful to the anonymous referees for their helpful feedback which has improved the paper. Additionally, the authors would like to thank Manolis Georgoulis for pointing out the trace estimates in Section 5.2.

Some of the research was carried out during visits to the Institute for Numerical Simulation at the University of Bonn. Dr. Hangelbroek thanks the SFB 1060 for helping to support his visits.

## REFERENCES

- [1] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Texts in Applied Mathematics, vol. 15, Springer, New York, 2008. MR2373954
- [2] Carsten Carstensen and Birgit Faermann, *Mathematical foundation of a posteriori error estimates and adaptive mesh-refining algorithms for boundary integral equations of the first kind*, Engineering analysis with boundary elements **25** (2001), no. 7, 497–509.
- [3] W. Dahmen, B. Faermann, I. G. Graham, W. Hackbusch, and S. A. Sauter, *Inverse inequalities on non-quasi-uniform meshes and application to the mortar element method*, Math. Comp. **73** (2004), no. 247, 1107–1138. MR2047080
- [4] R. A. DeVore, *Nonlinear approximation*, Acta numerica, 1998, Acta Numer., vol. 7, Cambridge Univ. Press, Cambridge, 1998, pp. 51–150. MR1689432
- [5] R. A. DeVore and R. C. Sharpley, *Besov spaces on domains in  $\mathbf{R}^d$* , Trans. Amer. Math. Soc. **335** (1993), no. 2, 843–864. MR1152321
- [6] L. C. Evans, *Partial Differential Equations*, 2nd ed., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 2010. MR2597943
- [7] B. Faermann, *Localization of the Aronszajn-Slobodeckij norm and application to adaptive boundary element methods. II. The three-dimensional case*, Numer. Math. **92** (2002), no. 3, 467–499. MR1930387
- [8] E. Fuselier, T. Hangelbroek, F. J. Narcowich, J. D. Ward, and G. B. Wright, *Localized bases for kernel spaces on the unit sphere*, SIAM J. Numer. Anal. **51** (2013), no. 5, 2538–2562. MR3097032
- [9] E. Fuselier, T. Hangelbroek, F. J. Narcowich, J. D. Ward, and G. B. Wright, *Kernel based quadrature on spheres and other homogeneous spaces*, Numer. Math. **127** (2014), no. 1, 57–92. MR3191217
- [10] E. H. Georgoulis, *Inverse-type estimates on hp-finite element spaces and applications*, Math. Comp. **77** (2008), no. 261, 201–219. MR2353949
- [11] I. G. Graham, W. Hackbusch, and S. A. Sauter, *Finite elements on degenerate meshes: inverse-type inequalities and applications*, IMA J. Numer. Anal. **25** (2005), no. 2, 379–407. MR2126208
- [12] M. Griebel, C. Rieger, and B. Zwicknagl, *Multiscale approximation and reproducing kernel Hilbert space methods*, SIAM J. Numer. Anal. **53** (2015), no. 2, 852–873. MR3327356
- [13] J.-L. Guermond, *The LBB condition in fractional Sobolev spaces and applications*, IMA J. Numer. Anal. **29** (2009), no. 3, 790–805. MR2520170
- [14] J. Guzmán, A. J. Salgado, and F.-J. Sayas, *A note on the Ladyženskaja-Babuška-Brezzi condition*, J. Sci. Comput. **56** (2013), no. 2, 219–229. MR3071173
- [15] T. Hangelbroek, F. Narcowich, C. Rieger, and J. Ward, *Direct and inverse results on bounded domains for meshless methods via localized bases on manifolds*, arXiv:1406.1435. (to appear in Festschrift for the 80th birthday of Ian Sloan)
- [16] T. Hangelbroek, F. J. Narcowich, X. Sun, and J. D. Ward, *Kernel approximation on manifolds II: the  $L_\infty$  norm of the  $L_2$  projector*, SIAM J. Math. Anal. **43** (2011), no. 2, 662–684. MR2784871
- [17] T. Hangelbroek, F. J. Narcowich, and J. D. Ward, *Kernel approximation on manifolds I: bounding the Lebesgue constant*, SIAM J. Math. Anal. **42** (2010), no. 4, 1732–1760. MR2679593
- [18] 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. MR2970852
- [19] R. B. Lehoucq and S. T. Rowe, *A radial basis function Galerkin method for inhomogeneous nonlocal diffusion*, Comput. Methods Appl. Mech. Engrg. **299** (2016), 366–380. MR3434919
- [20] S. Mayboroda and M. Mitrea, *Sharp estimates for Green potentials on non-smooth domains*, Math. Res. Lett. **11** (2004), no. 4, 481–492. MR2092902
- [21] J. M. Melenk, *On Approximation in Meshless Methods*, Frontiers of numerical analysis, Universitext, Springer, Berlin, 2005, pp. 65–141. MR2178270
- [22] H. N. Mhaskar, F. J. Narcowich, J. Prestin, and J. D. Ward,  *$L^p$  Bernstein estimates and approximation by spherical basis functions*, Math. Comp. **79** (2010), no. 271, 1647–1679. MR2630006



- [23] H. N. Mhaskar, F. J. Narcowich, and J. D. Ward, *Spherical Marcinkiewicz-Zygmund inequalities and positive quadrature*, Math. Comp. **70** (2001), no. 235, 1113–1130. MR1710640
- [24] F. J. Narcowich, S. T. Rowe, and J. D. Ward, *A novel Galerkin method for solving PDES on the sphere using highly localized kernel bases*, Math. Comp. **86** (2017), no. 303, 197–231. MR3557798
- [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. MR2114646
- [26] 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. MR2239119
- [27] Christian Rieger, *Sampling inequalities and applications*, Ph.D. thesis, Universität Göttingen, 2008. <http://hdl.handle.net/11858/00-1735-0000-0006-B3B9-0>
- [28] R. Schaback and H. Wendland, *Inverse and saturation theorems for radial basis function interpolation*, Math. Comp. **71** (2002), no. 238, 669–681. MR1885620
- [29] H. Triebel, *Theory of Function Spaces*, Monographs in Mathematics, vol. 78, Birkhäuser Verlag, Basel, 1983. MR781540
- [30] Hans Triebel, *Theory of Function Spaces. II*, Monographs in Mathematics, vol. 84, Birkhäuser Verlag, Basel, 1992.
- [31] J. P. Ward,  *$L^p$  Bernstein inequalities and inverse theorems for RBF approximation on  $\mathbb{R}^d$* , J. Approx. Theory **164** (2012), no. 12, 1577–1593. MR3020939
- [32] H. Wendland, *Scattered Data Approximation*, Cambridge Monographs on Applied and Computational Mathematics, vol. 17, Cambridge University Press, Cambridge, 2005. MR2131724

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF HAWAII – MĀNOA, 2565 MCCARTHY MALL,  
HONOLULU, HAWAII

*E-mail address:* [hangelbr@math.hawaii.edu](mailto:hangelbr@math.hawaii.edu)

DEPARTMENT OF MATHEMATICS, TEXAS A & M UNIVERSITY, COLLEGE STATION, TEXAS 77843

*E-mail address:* [fnarc@math.tamu.edu](mailto:fnarc@math.tamu.edu)

INSTITUT FÜR NUMERISCHE SIMULATION, UNIVERSITÄT BONN, WEGELERSTRASSE 6, 53115 BONN,  
GERMANY

*E-mail address:* [rieger@ins.uni-bonn.de](mailto:rieger@ins.uni-bonn.de)

DEPARTMENT OF MATHEMATICS, TEXAS A&M UNIVERSITY, COLLEGE STATION, TEXAS 77843

*E-mail address:* [jward@math.tamu.edu](mailto:jward@math.tamu.edu)