MESHFREE METHODS USING LOCALIZED KERNEL BASES

A Dissertation

by

STEPHEN TABER ROWE

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Francis Narcowich
Co-Chair of Committee,	Joseph Ward
Committee Members,	Thomas Schlumprecht
	Fred Dahm
Head of Department,	Emil Straube

May 2015

Major Subject: Mathematics

ABSTRACT

Radial basis functions have been used to construct meshfree numerical methods for interpolation and for solving partial differential equations. Recently, a localized basis of radial basis functions has been developed on the sphere. In this dissertation, we investigate applying localized kernel bases for interpolation, approximation, and for novel discretization methods for numerically solving partial differential equations and integral equations. We investigate methods for partial differential equations on spheres using newly explored bases constructed from radial basis functions and associated quadrature methods. We explore applications of radial basis functions to anisotropic nonlocal diffusion problems and we develop theoretical frameworks for these methods.

DEDICATION

To my Mom

ACKNOWLEDGEMENTS

I would like to thank my advisers, Dr. Francis Narcowich and Dr. Joseph Ward, for their continuous support over the years. The countless hours they've spent helping me with problems, ranging from bureaucratic to technical matters, has been invaluable. Always willing to help and always supportive of any new project or interest I had, they have made the graduate school experience an enormously fruitful intellectual experience. Their continuous support, which has ranged from giving me gratuitous intellectual freedom to study numerical methods and programming languages on my own, as well as their encouragement to apply for fellowships and writing letters of recommendation is greatly appreciated.

I would like to thank Dr. Rich Lehoucq at Sandia National Laboratories for hosting me as an intern for two summers and for introducing me to interesting new projects that has driven my research at Texas A&M and for supporting my exploration of different computer methods and math methods while at Sandia. He encouraged my initial attempt at a research project and pointed out technical details that helped drive the eventual development of an entire chapter of my thesis. He spent many hours of his time helping me learn to more effectively write and communicate my ideas, which has been a surprisingly non-trivial part of the research experience for me.

I would like to thank Sandia National Laboratories for generously providing me with the Sandia National Laboratories Campus Executive Fellowship, which has had a fundamentally transformative effect on my research and career plans. I would also like to thank the numerous people in the College of Science and Department of Mathematics at Texas A&M University and at Sandia National Laboratories that

provided me help and support as part of the fellowship.

I would be remiss in my acknowledgments if I omitted my friends that have been invaluable over the years. Justin Owen has been a great friend and always there for me throughout the grad school experience. My friend and office-mate Sam Scholze has always made working at school an entertaining experience. He also introduced me to my hobby of weight lifting, which has evolved into the best stress relief hobby I've had at graduate school.

Dr. Larson in the Department of Mathematics at Texas A&M introduced me to research for the first time when I was a participant in his Matrix Analysis and Wavelets Research Experience for Undergraduates. Without him, I would certainly not be where I am today mathematically.

I would like to thank the Department of Mathematics at Texas A&M University for the generous support over the years.

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION AND BACKGROUND

A judicious choice of basis can make all the difference in the success of a numerical method. We explore and develop numerical methods for applied problems using a recently developed localized basis of kernel functions. The basis is constructed by linear combinations of radial basis functions. In many ways, this new basis maintains the flexible approximation powers of radial basis functions (RBFs) for scattered data while simultaneously avoiding the downsides of RBF methods. The localized basis enables interpolation of scattered data on manifolds such as spheres and we explore developing an analogous basis in $\mathbb{R}^d$. Due to the desirable properties of the basis for interpolation and approximation, we may consider constructing discretization spaces for solving partial differential equations and integral equations in a variety of settings.

We consider the following classical interpolation problem: let $\Omega \subset \mathbb{R}^d$ be a bounded set and let $\{x_i\}_{i=1}^N = X \subset \Omega$ be a set of scattered points (called centers). Let $\{y_j\}_{j=1}^N$ be a set of known data values $\big($e.g., samples of a function $y_j = f(x_j)\big)$. The interpolation problem seeks an *interpolant* $s : \Omega \to \mathbb{R}$ which satisfies $s(x_j) = y_j$.

Let $\pi_N(\mathbb{R}^d)$ denote the space of polynomials of degree at most N on $\mathbb{R}^d$. In the case $d = 1$ with $X \subset \mathbb{R}$, the interpolation problem always has a unique interpolant $s \in \pi_{N-1}(\mathbb{R})$. That is, there exists a subspace of functions which always admits a unique interpolant, regardless of the location of the centers. For $d > 1$, the Mairhuber-Curtis theorem demonstrates that it is not possible to fix a subspace of functions that always admits a unique interpolant regardless of the location of the data sites [28]. This result implies that the interpolants must be constructed by taking into account the configuration of the centers. This raises the difficult problem

of requiring a customized subspace of functions for each different set of centers.

1.1 Radial Basis Functions

One possible solution to the interpolation problem is to use *radial basis functions*. In this section, we define radial basis functions and we discuss their applications to scattered data interpolation, previous work, and other applications of radial basis functions. Let $\varphi(r) : \mathbb{R}^+ \to \mathbb{R}$ be a continuous function and define the function $\Phi : \mathbb{R}^d \to \mathbb{R}$ by $\Phi(x) = \varphi(\|x\|)$. We say such a function Φ is *radial*. Let $\Omega \subset \mathbb{R}^d$ be a region of interest where we have some given data. Let $\{x_i\}_{i=1}^N = X \subset \Omega$ be a finite set of centers, and define the space

$$V_X = \text{span}\{\Phi(x - x_i) : x_i \in X\}.$$

Given data values $\{y_j\}_{j=1}^N$, we seek an interpolant in V_X which is a linear combination of translates of Φ. This leads to a system of N equations in N unknowns with conditions

$$y_j = \sum_{i=1}^{N} c_i \varphi(\|x_j - x_i\|) \qquad \text{for } j = 1, \ldots, N.$$

If we let $A_{ij} = \varphi(\|x_i - x_j\|)$ and let $(\vec{c})_i = c_i$ and $(\vec{y})_j = y_j$, then we seek a solution to the problem $A\vec{c} = \vec{y}$. We refer to the matrix A as the interpolation matrix. This leads to the following question: what functions $\varphi : \mathbb{R}^+ \to \mathbb{R}$ necessarily generate an invertible interpolation matrix for any set of centers? This problem remains open, but the following restricted question is a classical question in analysis: what functions φ generate a positive definite interpolation matrix for all sets of centers? Bochner's theorem characterizes all positive semi-definite functions as the Fourier transform of a non-negative Borel measure. A useful corollary is that a continuous, integrable function $\Phi : \mathbb{R}^d \to \mathbb{R}$ is positive definite if and only if Φ is bounded and its Fourier

transform is non-negative and non-vanishing. One such example is the Gaussian $\varphi(r) = \exp(-\alpha r^2)$ for $\alpha > 0$ [28].

A more general notion is that of a *conditionally* positive definite function of order m. Recall that $\pi_m(\mathbb{R}^d)$ is the space of at most degree m polynomials on $\mathbb{R}^d$. We say that $\varphi : \mathbb{R}^+ \to \mathbb{R}$ is conditionally positive definite of order m on $\mathbb{R}^d$ if and only if for any set of scattered centers, the quadratic form

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \bar{\alpha}_j \varphi(\|x_i - x_j\|)$$

is positive for any set of scalars $\{\alpha_i\}_{i=1}^N$ not identically equal to zero that satisfy

$$\sum_{i=1}^{N} \alpha_i p(x_i) = 0$$

for all $p \in \pi_{m-1}(\mathbb{R}^d)$. That is, the interpolation matrix is positive definite on a subspace "orthogonal" to the polynomials. One such example is the thin plate spline $\varphi(r) = r^2 \log(r)$, which is conditionally positive definite of order 2 on every $\mathbb{R}^d$. Interpolating data on a set of centers with a conditionally positive definite function requires the addition of polynomial constraints to ensure the existence and uniqueness of an interpolant. Given centers $\{x_j\}_{j=1}^N$ and data values $\{y_j\}_{j=1}^N$, the interpolation problem is to find coefficients $\{c_j\}_{j=1}^N$ and $\{b_l\}_{l=0}^{m_d}$ such that

$$y_j = \sum_{i=1}^{N} c_i \varphi(\|x_j - x_i\|) + \sum_{l=0}^{m_d} b_l p_l(x_j)$$
$$\text{subject to } \sum_{i=1}^{N} c_i p_l(x_i) \text{ for } 0 \le l \le m_d$$

where $\{p_l\}_{l=0}^{m_d}$ form a basis for $\pi_{m-1}(\mathbb{R}^d)$.

Radial basis functions have been actively explored for the past several decades. Hardy's work with the multiquadric was one of the first explorations of radial basis functions for interpolation, dating back to work in 1971. Duchon's investigation of the thin plate spline used a variational approach of minimizing a semi-norm [5]. The thin plate spline was found to be the minimizer of a certain energy functional (which has a physical interpretation as the bending energy of a thin metal plate). Later, Meinguet pushed forward the use of thin plate spline interpolation for numerical methods [20, 21, 22].

Radial basis functions are actively being researched for discretization methods for partial differential equations. They are particularly intriguing because they do not require a mesh or triangulation and can be used for high dimensional problems. In 1990, Kansa introduced a radial basis function method for the solution of partial differential equations [16]. This work, based on the multiquadric RBF, established the first collocation method for elliptic, parabolic, and hyperbolic partial differential equations. Future work has explored the application of the Kansa method for shocks and shallow water wave equations. Engineers have reported success in using these methods for modeling high order differential equations which often are difficult for finite element methods. Unfortunately, Kansa's method does not guarantee a solution. Hon and Schaback reported an example of a differential operator, radial basis function, and set of centers which yielded a singular collocation matrix [15]. Furthermore, no error estimates have been proven for Kansa's method. In contrast to Kansa's method, the symmetric radial basis function method introduced by Fasshauer guarantees an invertible collocation matrix and provides error estimates, but at the cost of requiring basis functions to be twice as smooth as Kansa's method [6].

Galerkin methods are of interest because they provide a functional analytic frame-

work for constructing error estimates. Much work over the past several decades has yielded a wealth of theory which can be readily applied to produce error estimates provided the basis of functions being used has known interpolation error estimates. Since radial basis function methods have known error estimates for certain kernels, this suggests a Galerkin method using radial basis functions is viable theoretically. While these have been investigated, due to difficulty with numerically integrating radial basis functions to construct elements in the stiffness matrix, these methods have not been pursued as actively as collocation methods. In contrast to Kansa's collocation method, Wendland developed error estimates for a Galerkin radial basis function method for elliptic partial differential equations [27].

Radial basis function (RBF) interpolation offers interpolation of possibly highly scattered, high dimensional data. Two computational drawbacks associated with RBF interpolation are the construction of the interpolant and subsequent evaluation of the interpolant. Given N centers, constructing the interpolant requires inverting a matrix where the number of rows is $\mathcal{O}(N)$. For globally supported RBFs (e.g., the thin plate splines or Gaussians), the interpolation matrix is dense. The condition number of the interpolation matrix grows with respect to the minimum distance between two centers. Therefore, solving for the interpolant on a large set of centers requires inverting a large, dense, ill-conditioned matrix. The computational cost of evaluating the interpolant on M points is of order $\mathcal{O}(MN)$. Fast multipole methods have been investigated to reduce the computational complexity of RBF evaluation, although these methods reduce the accuracy of the interpolant [28].

Radial basis function interpolation has suffered from a so-called trade off principle. Consider, for example, the positive definite radial basis function $\varphi(r) = \exp(-\alpha r^2)$ for $\alpha > 0$. The Gaussian RBF is positive definite on any $\mathbb{R}^d$, but for numerical approximations, the choice of α is problematic. A large value of α results

in a well-conditioned interpolation matrix, but yields poor approximation of a function. On the other hand, decreasing α leads to better numerical approximation, but the condition number grows quickly with the number of centers [28]. There is no known analytic method for choosing α and some suggest *ad hoc* methods of guessing or trial-and-error methods for choosing α. This issue is not unique to the Gaussian; compactly supported Wendland functions also require a user chosen scale parameter. The thin plate spline does not require a scale parameter and recent work has demonstrated a "self-scaling" basis which scales automatically with the data density.

1.2 Native Spaces

In this section, we cover background regarding error estimates for radial basis function interpolation on compact domains in $\mathbb{R}^n$. Error estimates for radial basis function interpolation have been a subject of investigation for at least two decades. Most RBF interpolation error estimates take place in the *native space*, a reproducing kernel Hilbert space corresponding to the (conditionally) positive definite RBF kernel. We discuss these spaces and their importance as well as error estimates for functions not residing in the native space. The error estimates for functions not residing in the native space is crucial for error estimates on partial differential equations and for approximation error from interpolation of lower smoothness functions.

Let H denote a Hilbert space of functions on $\Omega \subset \mathbb{R}^n$. We say that a kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$ is a *reproducing kernel* for H if

- 1: $\Phi(\cdot, x) \in H$ for all $x \in \Omega$.
- 2: $f(x) = \langle f, \Phi(\cdot, x) \rangle$ for all $f \in H$ and all $x \in \Omega$.

We say that H is a *reproducing kernel Hilbert space* and Φ is the reproducing kernel. An example of a reproducing kernel Hilbert space is $W_2^1(\mathbb{R})$ with the reproducing

kernel $\Phi(\cdot, x) := \exp(-\frac{1}{2}|x - \cdot|)$. Not all Hilbert spaces are reproducing kernel Hilbert spaces. It can be shown that a Hilbert space H having a reproducing kernel is equivalent to the point evaluation functionals being continuous. A consequence of this result is that $L^2[a, b]$ does not have a reproducing kernel since the point evaluation functionals are not continuous.

Positive definite functions naturally relate to reproducing kernel Hilbert spaces. Given a positive definite function, we can directly construct a Hilbert space on which the function is a reproducing kernel. Let $\Phi : \Omega \times \Omega \to \mathbb{R}$ be a continuous, positive definite function We construct a space H_Φ by considering all finite linear combinations of $\Phi(\cdot, x)$. That is,we define

$$H_\Phi := \{\sum_{j=1}^{N} a_j \Phi(\cdot, x_j) : x_j \in X \text{ and } N < \infty\} \tag{1.1}$$

We note that this is a space of continuous functions (since Φ is a continuous positive definite kernel). This space may be equipped with a bilinear form

$$\langle f, \Phi(\cdot, y)\rangle_\Phi = f(y)$$

for all $f \in H_\Phi$ and all $y \in \Omega$. By taking the completion of this space and appropriately identifying the elements in the completion as continuous functions, we have a Hilbert space $\mathcal{H}_\Phi$ with a reproducing kernel Φ. The resulting Hilbert space is called the *native space* of the kernel Φ, under appropriate interpretation of the elements of the Hilbert space as continuous functions. See [28] for a more thorough discussion.

In the context of a conditionally positive definite function, the concept of a native space is significantly more technical. The interested reader is strongly encourage to consult [28] for a much more thorough and full discussion and exposition on the topic

of native spaces and reproducing kernel Hilbert spaces, as well as a complete and rigorous discussion of native spaces for conditionally positive definite functions.

We refer to the Sobolev space on a region Ω, denoted $W_p^k(\Omega)$ to be the collection of L^p functions with up to order k L^p weak derivatives. That is,

$$W_p^k(\Omega) := \{f \in L^p(\Omega) : D^\alpha f \in L^p(\Omega) \text{ for } |\alpha| \leq k\}.$$

The Sobolev space is equipped with the Sobolev norm given by

$$\|f\|_{W_p^k(\Omega)} := \left(\sum_{|\alpha| \leq k} \|D^\alpha f\|_{L^p}^p \right)^{\frac{1}{p}}.$$

For the case $p = 2$, which is frequently of interest for our purposes, the Sobolev space is a Hilbert space. In addition to the Sobolev norm on $W_2^k(\Omega)$, we also may use the Sobolev semi-norm given by

$$|f|_{W_2^k(\Omega)} := \left(\sum_{|\alpha| = k} \|D^\alpha f\|_{L^2}^2 \right)^{\frac{1}{2}}.$$

We define the Beppo-Levi space $BL(\mathbb{R}^n)$ to be the space of functions

$$BL(\mathbb{R}^n) := \{f \in L^p(\mathbb{R}^n) : |f|_{W_2^k(\mathbb{R}^n)} < \infty\}.$$

The Beppo-Levi space is a semi-Hilbert space when equipped with the above semi-norm. We note that polynomials of degree less than k are contained within the kernel of the Beppo-Levi semi-norm.

We refer to the surface splines or thin plate splines of order m, denoted ϕ_m :

$\mathbb{R}^d \to \mathbb{R}$, to be the functions

$$\phi_m(\|x\|) := \begin{cases} \|x\|^{2m-d} & d \text{ is odd} \\ \|x\|^{2m-d} \log(\|x\|) & d \text{ is even.} \end{cases}$$

These functions are *conditionally positive definite* of order m. If $\pi_{m-1}(\mathbb{R}^d)$ denotes the space of degree at most $m-1$ degree polynomials on $\mathbb{R}^d$, then the bilinear form

$$\sum_{i,j}^{N} a_i a_j \phi(\|x_i - x_j\|) > 0$$

provided that $\sum_{i=1}^{N} a_i p(x_i) = 0$ for each $p \in \pi_{m-1}(\mathbb{R}^d)$. Informally, we may interpret this as the thin plate splines being positive definite on a subspace "orthogonal" to the polynomials of degree at most $m-1$. The interpolation space for these functions takes the form

$$S(X) := \{ \sum_{x_i \in X} a_i \phi_m(\cdot - x_i) | \sum_{x_i \in X} a_i p(x_i) = 0 \text{ for all } p \in \pi_{m-1} \} + \pi_{m-1}.$$

Since the ϕ_m are conditionally positive definite, on a set of scattered centers X, the interpolation problem has a unique solution provided that the set of centers is *unisolvent*. We say a set of points X is unisolvent with respect to $\pi_{m-1}(\mathbb{R}^d)$ if the only polynomial which is zero on all of X is the zero polynomial. We note that this is a mild condition that should not cause issue (except, possibly in the case of very few points or points that lie exactly along a plane or line).

1.3 Error Estimates

We first discuss error estimates for radial basis function interpolation. Historically, these estimates were primarily restricted to functions residing in the native

space of the kernel. This restriction was suspected to be artificial, as numerical experiments with functions not smooth enough to reside in the native space (e.g., $W_2^k(\Omega)$) would still present predictable convergence rates. Characterizing the convergence rates requires knowledge of the geometry of the centers. Analogous to the finite element method, where the convergence rate of the scheme depends on the size of the largest elements (denoted by h), radial basis function interpolation convergence rates depend on the largest gaps in the distribution of the centers (a quantity aptly also denoted by h). We note that the radial basis function estimates are analogous to the error estimates typically derived for finite element schemes; typically, one expects the solution to a function in $W_2^k(\Omega)$ to converge at a rate of $\mathcal{O}(h^k)$ in the L^2 norm. Additionally, one expects the solution to converge at a rate of $\mathcal{O}(h^{k-\alpha})$ in the $W_2^\alpha(\Omega)$ norm. Such results have been developed for radial basis functions and we present them below.

The geometry of the data affects the quality of the interpolant in multiple ways. Informally, if the data is sufficiently "dense" in the domain, we expect the interpolant to provide a high degree of approximation to a smooth function. Furthermore, the distribution of the points may affect the quality of the interpolation matrix. If two centers x_i, x_j are very close, then the i and j columns of the interpolation matrix are very "similar", which causes the condition number of the matrix to be poor. Therefore, the challenge for scattered data interpolation is to have well-distributed data that is sufficiently dense, yet does not clump. We mathematically characterize these ideas with different quantities to represent the geometric properties of the data.

Let $X \subset \Omega$ be a collection of scattered centers. We define the *mesh norm* or *fill distance* h to be the radius of the largest ball in Ω that does not intersect X. We define the separation radius q to be half the minimum distance between centers.

Mathematically,

$$h := \sup_{x \in \Omega} \min_{x_i \in X} \|x - x_i\| \qquad q := \frac{1}{2} \inf_{x_i, x_j \in X} \|x_i - x_j\|. \tag{1.2}$$

See Figure 1.1 for a visual example of the mesh norm. We define the *mesh ratio* $\rho := \frac{h}{q}$. Informally, for ρ near one, the centers are nearly uniformly distributed and large ρ indicates clustering of points. Let $\{X_{h,q}\}$ be a collection of sets of centers indexed by mesh norm h, q. We say the collections of centers are *quasi-uniformly* distributed if there exists constants C_1, C_2 such that

$$C_1 q \leq h \leq C_2 q.$$

Consequently, $\sup_{\{X_{h,q}\}} \rho_{h,q} \leq \frac{C_2}{C_1}$, which implies the mesh ratio is bounded. Quasi-uniformly distributed collections of centers allow for theoretically shrinking h to zero while controlling the distributions of the centers so they do not clump arbitrarily as $h \to 0$. This is a fundamental assumption we assume throughout.

Now that we may quantify geometric properties of the centers, we present a classical radial basis function theorem for interpolation error estimates.

Theorem 1. *[28] Suppose that $\Omega \subseteq \mathbb{R}^d$ is bounded and satisfies an interior cone condition. Let $\Phi(x) = (-1)^{k+1}\|x\|^{2k}\log(\|x\|)$ and let $f \in N_\Phi(\Omega)$. Let X be a collection of quasi-uniformly distributed centers and let $I_X f$ denote the radial basis function interpolant constructed using Φ. Then, there exists $C > 0$ such that for sufficiently small h,*

$$|D^\alpha f(x) - D^\alpha I_X f| \leq C h_X^{k-|\alpha|} |f|_{N^\Phi(\Omega)}.$$

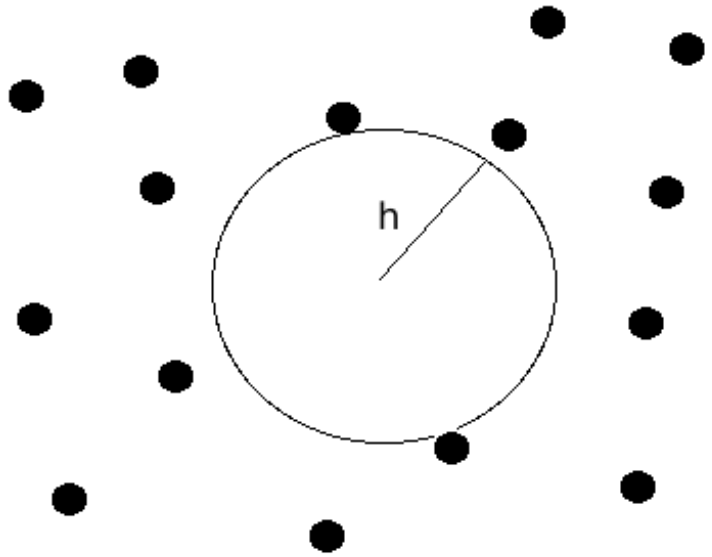


Figure 1.1: The dots represent locations of centers. The quantity h represents the radius of the largest ball that does not intersect any centers. This quantity measures how dense the data is in the region of interest.

Analogous results hold for compactly supported Wendland functions. The key observation in the above theorem is that the function f lies in the native space. This assumption is highly restrictive, so error estimates for functions not contained in the native space are of interest. The restriction of error estimates for functions only residing in the native space was a limiting factor for the theoretical development of error estimates for methods in partial differential equations. Work by Narcowich, Ward, and Wendland proved that error estimates can "escape" the native space for certain radial basis functions [26]. We present an example of such an "escape" estimate which will prove to be useful for our purposes later.

Theorem 2. *Let $\Omega \subset \mathbb{R}^d$ be a compact domain with Lipschitz boundary and satisfy*

an interior cone condition. For sufficiently small h, and $f \in W_2^k(\Omega)$ for $\frac{d}{2} < k \leq m$,

$$\|f - I_X f\|_{L^2(\Omega)} \leq Ch^k \|f\|_{W_2^k(\Omega)}$$

where $I_X f \in W_2^m(\Omega)$ is the thin plate spline interpolant [26].

We briefly remark that the restriction $k > \frac{d}{2}$ is imperative. If $k \leq \frac{d}{2}$, then f is not necessarily a continuous function; by the Sobolev embedding theorem, $W_2^k(\Omega)$ are continuous for $k > \frac{d}{2}$. Hence interpolation is a questionable operation.

The proof of this theorem relies on a very important lemma that is interesting in its own right. The so-called "Zeros lemma" relates different Sobolev norms of a function that has sufficiently many zeros within a region.

Proposition 1. *Let $\Omega \subset \mathbb{R}^n$ be a bounded and satisfy an interior cone condition. Let k be a positive integer with $0 < s \leq 1$, $1 \leq p, q \leq \infty$ and let $m \in \mathbb{N}$ satisfy $k > m + \frac{d}{p}$ for $p > 1$ or $k \geq m + d$ for $p = 1$. Let X be a set with mesh norm h that is sufficiently small. If $u \in W_p^{k+s}(\Omega)$ satisfies $u|_X = 0$, then*

$$|u|_{W_q^m(\Omega)} \leq Ch^{k+s-m-d(\frac{1}{p}-\frac{1}{q})_+} |u|_{W_p^{k+s}(\Omega)}.$$

In particular, for the choice $q = p = 2$ and $s = 0$ (i.e., a non-fractional order Sobolev space), we have $|u|_{W_2^m(\Omega)} \leq Ch^{k-m}|u|_{W_2^k(\Omega)}$. As one might suspect, by noting that a function u and its radial basis function interpolant are zero on the set of centers, an error estimate for $|u - I_X u|_{W_2^m(\Omega)}$ can be derived. Since the thin plate splines have error estimates which escape the native space and require no scale parameter, they are of particular interest for numerical methods besides interpolation. We use these approximation powers later for work involving partial differential equations and nonlocal diffusion. We briefly remark on the approximation powers of

alternative radial basis functions. The Gaussians and multiquadrics enjoy a spectral convergence result for functions in their native space [18]. That is, for $f \in N_\Phi(\Omega)$, the native space for the Gaussian kernel,

$$\|f - I_X f\|_{L^\infty(\Omega)} \leq C \exp\left(-\frac{c}{h}\right) \|f\|_{N_\Phi(\Omega)}.$$

For functions in the native space, the interpolant rapidly converges to the function at an exponential rate. However, the trade off is the small size of the native space. Elements of the native space for Gaussians and multiquadrics are necessarily analytic, which is a "small" space. In contrast, the thin plate splines guarantee convergence to less smooth functions. In particular, it has been shown that error estimates for functions outside of the native space exist.

Now that we have some understanding of the approximation powers of radial basis functions, we consider how numerically stable the construction of the interpolant is. The construction of the interpolant requires the solution of a linear system of equations, and hence the solution of a matrix equation via some iterative method or by LU decomposition. The solution of a matrix equation by iterative methods can be sensitive to noise in the data, which is a guaranteed reality since computers are involved. The condition number of a matrix measures, roughly, how much small perturbations in the data affect the solution to a problem. For a matrix A, we define the condition number $\kappa(A) := \|A\| \, \|A^{-1}\|$. In the lucky case of a symmetric, positive definite matrix, we know $\|A\| = \sup\{\lambda : \lambda \in \sigma(A)\}$ where $\sigma(A)$ is the collection of eigenvalues of A. Furthermore, $\|A^{-1}\|$ is the reciprocal of the minimal eigenvalue of A. Therefore, we have $\kappa(A) := \frac{\lambda_{\min}(A)}{\lambda_{\max}(A)}$.

Constructing an accurate solution to a linear system of equations is problematic for "very large" condition numbers. Understanding how the condition number

changes as parameters for a problem are modified is imperative. For the problem of radial basis function interpolation, the condition number can be characterized in terms of the separation radius q. (We note that, in the quasi-uniform assumption, this number can be bounded above and below by the mesh norm h). The dependence on q is perhaps not particularly surprising; as q shrinks, the centers get closer and closer. Consequently, two columns of the interpolation matrix will be very close componentwise, which implies the matrix becomes "more and more nearly linearly dependent".

We can quickly measure an upper bound on the maximal eigenvalue for kernels. To approach this problem, the Gershgorin Circle Theorem for eigenvalues provides a method to approximate upper bounds on the difference between matrix entries and eigenvalues:

$$|\lambda - a_i i| \leq \sum_{i \neq j} a_{i,j}$$

where $a_{i,j}$ are the (i,j) matrix entries. For an interpolation matrix, $a_{i,j} = \Phi(x_i, x_j)$, and consequently,

$$|\lambda_{\max}(A) - \Phi(x_i, x_i)| \leq \sum_{i \neq j} |\Phi(x_i, x_j)| \leq (N-1)\|\Phi(\cdot,\cdot)\|_{L^\infty}$$

Consequently,

$$|\lambda_{\max}| \leq N\|\Phi(\cdot,\cdot)\|_{L^\infty} \leq Cq^{-d}\|\Phi(\cdot,\cdot)\|_{L^\infty}.$$

The last inequality follows by a simple bound on the number of centers in a region $\Omega \subset \mathbb{R}^d$. A significantly more difficult problem is analyzing lower bounds for the minimal eigenvalue. We refer the interested reader to [28] for a thorough discussion, but we present some results here. For the thin plate splines $\phi(r) = (-1)^{k+1} r^{2k} \log(r)$, the minimal eigenvalue falls as q^{2k}. Consequently, as q goes to zero, the interpolation

system may become ill-conditioned. For this reason, constructing an interpolant on a large set of centers with a small separation radius may be difficult. Solving, large, dense, ill-conditioned linear systems is a non-trivial computational challenge. Furthermore, poor condition numbers degrade the performance of iterative methods. This issue is one of many which motivates our future work with an alternative basis.

We remark briefly on alternative radial basis functions. The compactly supported Wendland functions also possess an algebraic rate of change in the minimal eigenvalue. On the other hand, the Gaussians suffer from an exponential decrease in the minimal eigenvalue, which leads to unpleasantly ill-conditioned systems even for small collections of centers.

1.4 Spherical Basis Functions

The problem of interpolation on Riemannian manifolds is of interest for scientific applications. In particular, much work has been done in the case of a boundaryless, compact Riemannian manifold, such as the n-sphere $\mathbb{S}^n$. In the case of $\mathbb{S}^n$, *spherical basis function* (SBF) interpolation has been explored and allows for interpolation of scattered data. Excluding a few special cases for point distributions along platonic solids, constructing uniformly spaced points along the sphere is not possible. Methods requiring regular distributions of points are not available. Therefore, interpolation methods which allow for scattered data are imperative for spheres.

Given a set of points (centers) distributed along the n-sphere, a spherical basis function $\Phi : \mathbb{S}^n \times \mathbb{S}^n \to \mathbb{R}$ can be defined by choosing $\Phi(x, y) = \phi(x \cdot y)$ for $\phi : [-1, 1] \to \mathbb{R}$. For each center $x_j \in \{x_i\}_{i=1}^N$, the spherical basis function is the rotation to the point x_j, $\Phi_j(x) = \phi(x \cdot x_j)$. The interpolant is then formed from

linear combinations of rotations of Φ, given by

$$s(x) = \sum_{j=1}^{N} c_j \phi(x \cdot x_j).$$

Positive definite spherical basis functions, analogous to positive definite radial basis functions, yield positive definite interpolation matrices. Conditionally positive definite kernels, such as the restriction of the thin plate spline to $\mathbb{S}^n$, may also be defined in a similar fashion. While these methods offer interpolation for highly scattered data, they suffer the same drawbacks as the radial basis functions on domains in $\mathbb{R}^n$. In particular, we are interested in the restricted surface splines or polyharmonic splines

$$\phi_m(x \cdot y) := (-1)^m (1 - x \cdot y)^{m-1} \log(1 - x \cdot y)$$

for $m \geq \frac{n}{2}$.

We discuss in detail later partial differential equations on spheres, which necessitates some discussion of background. We focus on the n-sphere with an interest in differential operators, spherical harmonics, Sobolev spaces, and approximation spaces on spheres. The sphere $\mathbb{S}^n$ is a compact, boundaryless Riemannian manifold. Let $(x^1, x^2, \ldots, x^n)$ be a smooth set of local coordinates. The sphere $\mathbb{S}^n$ has a metric tensor g_{ij} and measure $d\mu = \sqrt{\det(g_{i,j})} dx^1 dx^2 \ldots dx^n$. We are particularly interested in the case $n = 2$, which is the usual sphere. In this case, we have the usual local spherical coordinates (θ, φ), where θ is the colatitude and φ is the longitude. With

these local coordinates, the metric tensor takes the form

$$g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2(\theta) \end{pmatrix}.$$

Differential operators on spheres are of particular interest since we aim to discretize and solve partial differential equations on the sphere. The covariant derivative operator, ∇, acts as the usual gradient when operating on functions expressed appropriately. The Laplace-Beltrami operator acts as a spherical analogue of the Laplacian. The Laplace-Beltrami operator is defined as $\nabla^*\nabla$, which in local coordinates may be expressed as

$$\Delta u := \frac{1}{\sqrt{\det(g_{ij})}} \sum_{i,j} \frac{\partial}{\partial x^i} \sqrt{\det(g_{ij})} g^{ij} \frac{\partial u}{\partial x^i} \tag{1.3}$$

where g^{ij} is the inverse matrix of g_{ij}. As usual, we are particularly interested in the case of $\mathbb{S}^2$, which in spherical coordinates leads to the Laplace-Beltrami operator taking the form

$$\Delta u = \frac{1}{\sin(\theta)} \frac{\partial}{\partial \theta} \left(\sin(\theta) \frac{\partial u}{\partial \theta} \right) + \frac{1}{\sin^2(\theta)} \frac{\partial^2 u}{\partial \phi^2}. \tag{1.4}$$

The eigenfunctions of the Laplace-Beltrami play an important role on the sphere and we make use of them as a basis in our numerical methods. These eigenfunctions, known as the spherical harmonics, are eigenfunctions of $-\Delta$ on $\mathbb{S}^n$ with eigenvalues $\lambda_\ell = \ell(\ell + n - 1)$. The eigenspace corresponding to λ_ℓ is spanned by a collection of orthonormal eigenfunctions denoted by $Y_{\ell,k}$ where $k = 1, \ldots, d_\ell$. We denote the eigenspace spanned by $\{Y_{\ell,k}\}_{k=1}^{d_\ell} = \mathcal{H}_\ell$. The space spanned by all eigenfunctions up to order L will be denoted by $\Pi_L := \oplus_{\ell=0}^{L} \mathcal{H}_\ell$. The dimension of the eigenspace

$\mathcal{H}_\ell \sim O(\ell^{n-1})$. Let $x, y \in \mathbb{S}^n$ and let $x \cdot y$ denote the dot product in $\mathbb{R}^{n+1}$. Then, the famous addition formula tells us

$$\sum_{k=1}^{d_\ell} Y_{\ell,k}(x) Y_{\ell,k}(y) = \frac{2\ell + n - 1}{(n-1)\omega_n} P_\ell^{\frac{n-1}{2}}(x \cdot y)$$

where ω_n is the volume of $\mathbb{S}^n$ and $P_\ell^{\frac{n-1}{2}}$ is the degree of the ℓ ultraspherical polynomial of order $\frac{n-1}{2}$. For the case $n = 2$, these polynomials are the Legendre polynomials.

The space $L^2(\mathbb{S}^n)$ is the Hilbert space of square integrable functions with respect to the measure $d\mu$. An orthonormal decomposition of L^2 functions is provided by the spherical harmonics. As a complete orthonormal set, we may expand any $f \in L^2(\mathbb{S}^n)$ via the formula

$$f = \sum_{\ell=0}^{\infty} \sum_{k=1}^{d_\ell} \hat{f}_{\ell,k} Y_{\ell,k}.$$

It then follows that the L^2 inner product of $f, g \in L^2(\mathbb{S}^n)$ is given by

$$\langle f, g \rangle_{L^2(\mathbb{S}^n)} = \sum_{\ell=0}^{\infty} \sum_{k=1}^{d_\ell} \hat{f}_{\ell,k} \overline{\hat{g}_{\ell,k}}.$$

In addition to $L^2(\mathbb{S}^n)$, we require Sobolev spaces on spheres so we may characterize the smoothness of the functions we are working with. This is invaluable for partial differential equations, as error estimates often depend on some notion of smoothness. Furthermore, regularity theorems guarantee smoothness properties of the solution provided some level of smoothness on the data. We define the Sobolev space of order m to be the collection of functions

$$H_m := \{f \in L^2(\mathbb{S}^n) : \sum_{\ell=0}^{\infty} \sum_{k=1}^{d_\ell} (1 + \lambda_\ell)^k \hat{f}_{\ell,k}^2 < \infty\} \tag{1.5}$$

which has an inner product given by

$$\langle f, g \rangle_{H_k(\mathbb{S}^n)} := \sum_{\ell=0}^{\infty} \sum_{k=1}^{d_\ell} (1+\lambda_\ell)^k \hat{f}_{k,\ell} \overline{\hat{g}_{k,\ell}}.$$

We note that we may also denote H_m by $W_2^m(\mathbb{S}^n)$ to match our notation for $\mathbb{R}^n$. Fractional order Sobolev spaces extend the definition above. For fractional τ, we may define H_τ as the space of functions such that $\|(I-\Delta)^{\frac{\tau}{2}} f\|_{L^2(\mathbb{S}^n)} < \infty$.

We note that the Laplace-Beltrami operator is a self-adjoint operator with respect to the $L^2(\mathbb{S}^n)$ inner product and $-\Delta$ is positive on the orthogonal complement of a finite dimensional subspace of spherical harmonics.

1.4.1 Conditionally Positive Definite Kernels on Spheres

A kernel κ is conditionally positive definite with respect to a finite dimensional subspace Π if, for any set of N distinct centers, the matrix $K_X := \Big(\kappa(\xi,\eta)\Big)_{\xi,\eta}$ is positive definite on the subspace of all vectors $a \in \mathbb{C}^N$ satisfying $\sum_{\xi \in X} a_\xi p(\xi) = 0$ for all $p \in \Pi$.

In the case of the sphere $\mathbb{S}^2$, we focus on the finite dimensional subspaces Π_L of degree at most L spherical harmonic polynomials. By a slight re-indexing of the $\{Y_{l,k}\}$, we may write them as a collection of orthonormal functions $\{\phi_j\}$. We now consider a class of conditionally positive definite kernels that we characterize by studying their expansion in terms of the orthonormal basis $\{\phi_j\}$. Let $\{\kappa_j\}_{j=1}^{\infty} \in \ell^2(\mathbb{N})$ with all but finitely many κ_j positive. Then, we consider the kernel κ of the form

$$\kappa(x,y) := \sum_{j \in \mathbb{N}} k(j) \phi_j(x) \bar{\phi}_j(y). \tag{1.6}$$

Such a kernel is conditionally positive definite with respect to the space $\Pi := \text{span}\{\phi_j : \kappa_j \leq 0\}$. Let J denote the set of indices so that $\kappa_j \leq 0$. To verify condi-

tionally positive definite, we consider $\sum_{\xi,\eta} \alpha_\xi \kappa(\xi,\eta)\bar{\alpha}_\eta$ for some arbitrary collection of centers X with $\xi, \eta \in X$ and some arbitrary collection of coefficients $\{\alpha_\xi\}_{\xi\in\Xi}$ such that $\sum_{j\in J} \alpha_\xi \phi_j(\xi) = 0$. Then, we compute by expanding κ in terms of the orthonormal basis

$$\sum_{\xi,\eta} \alpha_\xi \kappa(\xi,\eta)\bar{\alpha}_\eta = \sum_{j\in\mathbb{N}} \kappa_j \sum_{\xi,\eta} \alpha_\xi \phi_j(\xi)\overline{\phi_j(\eta)\alpha_\eta} = \sum_{j\notin J} \kappa_j \|\alpha_\xi \phi_j(\xi)\|_{\ell^2(X)} > 0.$$

As we know, conditionally positive definite kernels give rise to unique interpolants provided additional constraints from some finite dimensional subspace Π are included. Using a kernel of the form Equation (1.6), an interpolant to a function f is constructed by

$$I_X f(\cdot) = \sum_{\xi\in X} a_\xi \kappa(\cdot,\xi) + \sum_{j\in J} b_j \phi_j(\cdot)$$

where we know $\sum_{\xi\in X} a_\xi \phi_j(\xi) = 0$ for all $j \in J$. Constructing the interpolant to the function f by data samples $\{f(\xi)\}_{\xi\in X}$ follows by solving a matrix problem of the form

$$\begin{pmatrix} K_X & P \\ P^T & 0 \end{pmatrix} \begin{pmatrix} a \\ c \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix} \tag{1.7}$$

where $(K_X)_{\xi,\eta} = \kappa(\xi,\eta)$ and $(\Phi)_{\xi,j} := \phi_j(\xi)$. As we know, this matrix system is invertible since κ is a conditionally positive definite kernel with respect to Π. Furthermore, the interpolant can be viewed from an alternative, variational perspective. The interpolant is a minimizer of a certain variational problem involving a semi-norm induced by the coefficients.

We define the native space inner product corresponding to κ by

$$\langle f, g \rangle_\kappa = \left\langle \sum_{j \in \mathbb{N}} \hat{f}_j \phi_j, \sum_{j \in \mathbb{N}} \hat{g}_j \phi_j \right\rangle_k := \sum_{j \notin J} \frac{\hat{f}(j)\hat{g}(j)}{\kappa_j} \tag{1.8}$$

where $\hat{f}_j$ denotes the Fourier coefficient of f with respect to the orthonormal basis $\{\phi_j\}$. We note that this is certainly a *semi*-inner product and not a true inner product. Consider, for example, applying the above semi-inner product to functions in Π; the semi-inner product yields zero since the sum runs over $j \notin J$. The semi-inner product induces a semi-norm in the usual way by $|f|_k^2 = \langle f, f \rangle_k$. The interpolant constructed by solving (1.7) minimizes the semi-norm $|\cdot|_k$.

2. LAGRANGE FUNCTIONS

Radial basis functions present numerous advantages for numerical methods, such as impressive interpolation and approximation powers. Well-understood convergence rates for interpolation as well as characterizations of the stability of interpolation matrices support the notion that, theoretically, radial basis functions may be potent tools for developing numerical methods for problems involving scattered, irregular data. However, difficulties arise in the implementation of the methods. Solving ill-conditioned, dense linear systems can be a non-trivial computational burden. We present methods that enable one to maintain all the benefits of radial basis functions while simultaneously reducing the computational overhead. To achieve this, we change from the basis of translates of a kernel $\Phi(\|x-y\|)$ to a basis of functions which interpolate one at a single point and zero elsewhere. These functions we refer to as *Lagrange functions* (others may refer to these as *cardinal functions*).

Lagrange functions and local Lagrange functions are the primary objects of interest for constructing numerical methods. The choice of basis can impact the efficacy of a numerical method, and we demonstrate that the local Lagrange functions perform admirably. In this chapter, we discuss in detail Lagrange functions and local Lagrange functions. We provide background information necessary to understand their theoretical properties and we provide some numerical experiments demonstrating the theoretical properties.

2.1 Lagrange Functions on Spheres

We begin our discussion of Lagrange functions by focusing on Lagrange functions on the $n-$sphere, $\mathbb{S}^n$. The manifold $\mathbb{S}^n$ has numerous advantageous properties, most notably that it is a manifold without boundary. This fortunately reduces numerous

theoretical difficulties which arise during the case of bounded subsets of $\mathbb{R}^n$. In particular, we note that the boundary is a nuisance ubiquitous throughout the field of numerical methods. By working on a manifold without boundary, we avoid many of these issues.

We start by considering a set of quasi-uniformly distributed centers $X \subset \mathbb{S}^n$. We restrict our focus to the surface splines of order m defined by

$$\phi_m(x, y) := (1 - x \cdot y)^{m-1} \log(1 - x \cdot y).$$

We know this is a conditionally positive definite spherical basis function with respect to the space $\Pi_m(\mathbb{S}^n)$. The approximation space V_X is defined to be the collection of acceptable linear combinations of rotations of $\phi_m(x \cdot x_j)$ plus polynomials of up to a certain degree. That is,

$$V_X := \{ \sum_{x_j \in X} a_j \phi_m(x \cdot x_j) + \Pi_m : \sum_{j=1}^{N} a_j p(x_j) = 0 \ \forall \ p \in \Pi_m \}.$$

We consider changing from this basis of rotations of $\phi(x \cdot x_j)$ to a basis which is highly localized spatially. Let $x_j \in X$ and consider constructing the interpolant, denoted χ_j, that takes a value of one at x_j and zero elsewhere. That is, $\chi_j(x_i) = \delta_{i,j}$ where $\delta_{i,j} = 0$ if $i \neq j$ and 1 if $i = j$. We know that for any collection of unisolvent centers, there always exists a unique interpolant to any data condition on the centers, so we know χ_j must exist in V_X. To construct this function, we enforce for each

$\xi \in X$,

$$\chi_j(\xi) = \sum_{i=1}^{N} \alpha_{i,j}\Phi(x_i,\xi) + \sum_{\ell=0}^{m}\sum_{k=1}^{d_\ell} \beta_{\ell,k,j}Y_{\ell,k}(\xi),$$
$$0 = \sum_{i=1}^{N} \alpha_{i,j}p(x_i) \qquad \text{for all } p \in \Pi_m.$$

Consequently, enforcing this condition requires the solution of a linear system of size $\mathcal{O}(N)$. By quasi-uniformity, we know that $q \sim C\frac{1}{N^d}$, and as N grows, the condition number gets progressively worse with a decrease of the minimal eigenvalue algebraically in terms of q. Consequently, the construction of the Lagrange function requires solving a large, dense, possibly ill-conditioned linear system of equations. Furthermore, to construct the full basis, N such systems must be solved. The reader may be very suspicious of any possible numerical use or practical application of such a basis; this skepticism is warranted, and indeed, we do not suggest the use of this basis for application necessarily. We later present a computationally friendly basis which preserves many of the properties of the Lagrange basis. However, before we discuss this basis, we present background information required which we take advantage of for our purposes later.

We begin our discussion of results on Lagrange functions on spheres by presenting estimates on their norms. To do so, interpolation error estimates for spherical basis functions are used, which are based on the powerful "Zeros Lemma" for Riemannian manifolds. The first observation is that the Lagrange functions, for various quasi-uniform sets, are pointwise bounded above. This prevents the possibility of the Lagrange functions "spiking" up too high or low off of the centers as the mesh norm h gets small. Indeed, the bound depends on the mesh ratio, which is bounded under the quasi-uniform assumption.

Lemma 1. *The Lagrange functions are uniformly bounded by constants independent of h or q.*

We next note a remarkable property of the Lagrange functions on the sphere: they are highly spatially localized. Indeed, the Lagrange function centered at a point x_j exponentially decays with respect to the distance from the center x_j. One may erroneously view this as unsurprising; χ_j certainly is zero on all of the centers except one, so it being small may not seem surprising. However, functions that interpolate zero throughout might still posses wild oscillations between the centers. Apparently, the Lagrange functions for thin plate splines do not. Furthermore, this is surprising because of the structure of a thin plate spline; these functions grow with distance and are categorically not localized spatially in any way. However, the correct linear combinations of thin plate splines plus the appropriate polynomial apparently results in a highly localized function.

Proposition 2. *[13][Proposition 4.5] Suppose that $m > \frac{d}{2}$ and κ_m is a polyharmonic kernel on $\mathbb{S}^n$. There exist positive constants h_0, ν, and C independent of h and q so that for any set of quasiuniform centers X with mesh ratio ρ and sufficiently small mesh norm h, the Lagrange function centered at $\xi \in X$ satisfies*

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

This argument was developed in [13] and also in [10]. The argument is based off of a trick showing that the bulk of the Sobolev semi-norm of the Lagrange function was contained in a thin annulus about the center. This "bulk chasing" argument was inspired by a similar result due to Matveev, which he used for working with D^m splines on $\mathbb{R}^n$ [19]. For more general manifolds, the decay results can be extended but with a modification that the decay is not in terms of just $d(x,\xi)$, but in terms

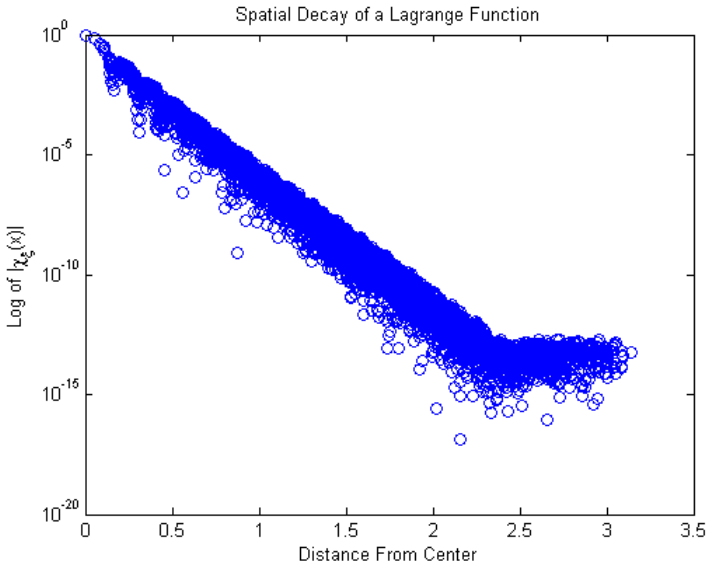


Figure 2.1: A Lagrange function constructed from 625 centers has been evaluated at 5041 points. The distance from the center of the Lagrange function to the evaluation point versus the log of the absolute value of the evaluation of the Lagrange function at the point is plotted. A clear, exponential decay is visible.

of $\min\left(d(x,\xi), r_{\mathbb{M}}\right)$ where $r_{\mathbb{M}}$ denotes the injectivity radius of the manifold. The argument for the decay largely takes place in the tangent space, which converts the problem from a problem on manifolds to a problem in $\mathbb{R}^n$, which is where the injectivity radius factors in.

In addition to the spatial decay of the Lagrange functions, they also exhibit a special Hölder continuity type of estimate.

Proposition 3. *[12] Under the assumptions of Proposition 2, for any* $0 < \epsilon \leq 1$*, there exists a constant* C *depending on the mesh ratio, the order* m *of the basis function* κ_m*, and* ϵ*, so that*

$$|\chi_j(x) - \chi_j(y)| \leq C\left(\frac{d(x,y)}{q}\right)^{\epsilon}.$$

These results suggest the basis behaves well: the basis functions are spatially localized with predictable continuity patterns. These results enable one to begin proving results about the stability of the basis. The first result that we present from [13] is one that studies the *Lebesgue constant* of the basis.

The Lebesgue constant corresponding to the collection X of centers is defined to be $L(X) := \sup_{x \in \mathbb{S}^n} \sum_{\xi \in X} |\chi_\xi(x)|$. This quantity measures the stability of the interpolation process. If this quantity grows without bound as $h \to 0$ for quasi-uniform sets, the interpolation operation becomes increasingly unstable. Consequently, the coefficients for the interpolant can grow without bound, as may happen with the case of using the basis of translates $\{\kappa(\cdot, \xi)\}$. The Lebesgue constant depends fundamentally on the choice of basis; alternative bases for the same approximation space may lead to different results. Therefore, establishing a uniformly bounded Lebesgue constant independent of mesh norm suggests the Lagrange bases provide stable methods for interpolation even as the mesh norm becomes quite small.

This should be contrasted with the case of other interpolation methods, such as interpolation via spherical harmonic polynomials. This method is indeed unstable [13]. It can be shown that spherical harmonic polynomial interpolation suffers from a Lebesgue constant that grows as $L^{\frac{d-1}{2}}$ where L is the highest degree of spherical harmonic used. On the real line, equidistant nodes yield an exponentially growing Lebesgue constant, which suggests interpolation can become problematic.

Proposition 4. *[13] Under the assumptions of Proposition 2, the Lebesgue constant is bounded by a constant depending only on the kernel, the mesh ratio ρ, and the manifold.*

A bounded Lebesgue constant provides a wonderful result which suggests that interpolation with the SBFs we employ is a near best approximation in L^∞. In

general, given a finite dimensional subspace of continuous functions, it does not follow that using these functions to interpolate data provides in any way optimal *approximation*. Polynomial interpolation can wildly oscillate yielding diverging L^∞ error. However, as a result of a bounded Lebesgue constant, an interpolation scheme yields *near optimal* approximation. Let $I_X f$ denote the interpolant and V_X the approximation space using the centers X and a kernel κ_m with bounded Lebesgue constant. Then, given any other function g in the approximation space, we know $I_X g = g$, and we can bound

$$\begin{aligned}\|f - I_X f\|_{L^\infty} &= \|f - g + g - I_X f\|_{L^\infty} = \|f - g + I_X(f - g)\|_{L^\infty} \\ &\leq \|f - g\|_{L^\infty} + \|I_X(f - g)\|_{L^\infty} \\ &= (1 + L)\text{dist}(f, V_X) = (1 + L) \inf_{g \in V_X} \|f - g\|_{L^\infty}.\end{aligned}$$

What this suggests is that interpolation is *near optimal* in the L^∞ norm. This result is analogous to, for example, Cea's lemma, which states the Galerkin solution to a bilinear problem is near optimal in the Hilbert space norm of choice. This invaluable result is often the first step in an error estimate for a discretization method for partial differential equations.

Combining techniques from the exponential decay of the Lagrange functions as well as the bounded Lebesgue constant properties led to the development of norm inequalities that enable one to bound the so-called "condition numbers" of the Lagrange functions. Let $\{v_\xi\}_{\xi \in X}$ be a basis for an approximation space. We define the *condition numbers* to be the values $C_{1,p}, C_{2,p}$ so that

$$C_{1,p}\|\{a_\xi\}\|_{\ell^p(X)} \leq \|\sum_{\xi \in X} a_\xi v_\xi\|_{L^p} \leq C_{2,p}\|\{a_\xi\}\|_{\ell^p(X)}.$$

The constants depend in some way upon the basis and upon p. Having condition number bounds of this form suggests the basis is L^p *stable*. That is, the L^p norm of the interpolant is comparable to the ℓ^p norm of the coefficients. Besides being useful for theoretical purposes, understanding the L^p condition numbers enables one to make estimates for condition numbers of matrices arising from the discretization of partial differential equations and integral equations.

Proposition 5. *[7] Let X be a quasi-uniformly distributed set of centers. There exist constants c_1 and c_2 depending only on m, the order of the SBF, and ρ, the mesh ratio, so that for sufficiently small mesh norm h, the Lagrange functions satisfy the condition number estimates*

$$c_1 q^{\frac{d}{p}} \|\{a_j\}\|_{\ell^p(X)} \leq \| \sum_{\xi \in X} a_\xi \chi_\xi \|_{L^p(\mathbb{S}^d)} \leq c_2 q^{\frac{d}{p}} \|a\|_{\ell^p(X)}.$$

To be clear, this result is not limited to merely one collection of centers X; this holds as we shrink $h, q \to 0$, provided the sets of centers satisfy the quasi-uniformity assumptions (that is, $C_1 \leq \frac{h_X}{q_X} \leq C_2$ for some fixed C_1 and C_2 independent of X). An immediate consequence of this statement is that the L^p norm of a single Lagrange function is on the order of $q^{\frac{d}{p}}$.

Proposition 6. *[7] Let X be a set of quasi-uniformly scattered centers in $\mathbb{S}^n$ and let χ_ξ, χ_η denote Lagrange functions for the points $\xi, \eta \in X$. Assume the approximation space is generated by a conditionally positive definite kernel $\kappa(\cdot, \xi) = \sum_{j \in \mathbb{N}} \kappa_j \varphi_j(\cdot) \varphi_j(\xi)$. The Lagrange function χ_ξ has expansion*

$$\chi_\xi(\cdot) := \sum_{\eta \in X} A_{\eta,\xi} \kappa(\cdot, \eta) + p_\xi$$

where $A_{\xi,\eta} := \langle \chi_\xi, \chi_\eta \rangle_k$.

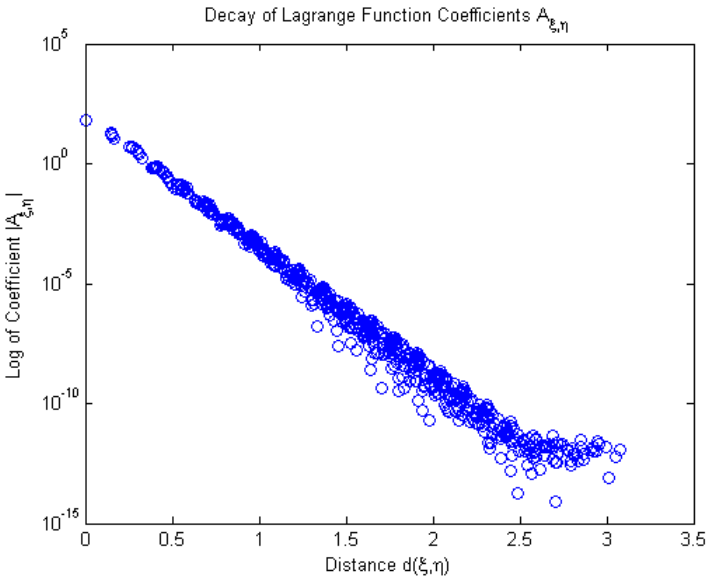


Figure 2.2: A Lagrange function centered at a point ξ constructed from 625 centers has its coefficients $A_{\xi,\eta}$ displayed. A clear, exponential decay with respect to the distance $d(\xi, \eta)$ is visible.

As a consequence of this, it has been demonstrated that the Lagrange function coefficients decay exponentially. The result follows by a type of Cauchy-Schwarz inequality that was presented in [7].

Proposition 7. *[7] Let X be a quasi-uniformly distributed set of centers on the sphere $\mathbb{S}^n$ with sufficiently small mesh norm h. Then, the coefficients of the Lagrange function χ_ξ satisfy*

$$|A_{\eta,\xi}| \leq Cq^{1-m} \exp\left(- \nu \frac{d(\xi, \eta)}{h} \right). \tag{2.1}$$

The significance of Proposition 7 is that the Lagrange functions are not solely localized spatially. The Lagrange functions, in the words of the authors of [7], have

a "small footprint" in the kernel basis. While each Lagrange function is a linear combination of all rotations of the kernel $\kappa(\cdot, \eta)$, apparently only kernels centered near ξ contribute significantly to the total value of the Lagrange function. The exponential decay forces the coefficients $A_{\xi,\eta}$ to be small for $d(\xi, \eta)$ large enough. This is highly suggestive of the idea that perhaps the Lagrange functions can be constructed in a way that takes advantage of this spatial locality while not reducing the approximation power of the functions.

The last major property we mention is the ability to switch between to relate different order Sobolev norms of linear combinations of basis functions. These are often referred to as Bernstein estimates. Let V_X be the space generated by the restricted surface splines (we sometimes refer to them as thin plate splines) $\varphi_s(t) = (-1)^{s+1}(1-t)^s \log(1-t)$.

Proposition 8. *[24] Let $g \in V_X$ which is generated by the thin plate spline for a quasi-uniform set of centers $X \subset \mathbb{S}^n$. Then, there exists a constant C independent of q, h such that*

$$\|g\|_{W_2^k(\mathbb{S}^n)} \leq Cq^{-k}\|g\|_{L^2(\mathbb{S}^n)}.$$

This result can be extend to handle L^p spaces instead, but our primary focus is on L^2.

We close this section by summarizing the key points developed by different authors over several years. The Lagrange functions of conditionally positive definite functions on manifolds are highly spatially localized as well as localized in the sense that the coefficients decay rapidly. The basis is inherently stable with a bounded Lebesgue constant as well as computable upper and lower condition number bounds. The L^p norm of a linear combination of Lagrange functions is comparable directly to the ℓ^p norm of the sequence of coefficients, with a factor of $q^{\frac{d}{p}}$ to be included.

Finally, switching between Sobolev norms induces a penalty of q^{-k}, where k is the difference between the order of the Sobolev spaces.

2.2 Local Lagrange Functions

We discuss results in this section regarding the development of a highly spatially localized, "small-footprint" basis. By small-footprint, we mean each basis element is constructed from few kernels, relative to the total number of kernels. The local basis retains many of the advantages of the full Lagrange basis discussed in Section 2.1. In particular, they decay quickly away from their center, are L^p stable, and provide near optimal approximation.

While the full Lagrange basis enjoys numerous theoretical advantages, the computational difficulty of assembling the Lagrange functions impedes their use in application. Developing a basis of functions which approximates the Lagrange functions has been a topic of some interest. Previous efforts have considered ad hoc methods of constructing Lagrange functions using few centers clustered around a center, but there was no strategy for choosing the number of centers nor for how the number of centers chosen should change as the mesh norm decreased. The results in [13] and [14] suggested that the Lagrange functions were highly localized spatially. Indeed, Proposition 2 suggests the Lagrange function is nearly zero for far away points. That the coefficients also decay exponentially, as shown in Proposition 7, suggested that a function that mimics the properties of the Lagrange function could be constructed that only requires a small number of kernels, rather than being a linear combination of every kernel. In [7], these ideas are explored and a full theory has been developed for a localized collection of Lagrange functions, referred to as local Lagrange functions; provable, theoretically supported bounds for the number of centers required for the construction of each local basis function are presented.

We begin by presenting the algorithm for the construction of the local Lagrange functions. We then discuss their theoretical properties. We encourage the interested reader to read [7] for a thorough and detailed account of the theoretical properties. Let $X \subset \mathbb{S}^2$ be a collection of quasi-uniformly scattered centers with mesh norm h and separation radius q. We focus on the thin plate spline kernel $\kappa_m(t) = (-1)^{m-1}(1-t)^m \cdot \log(1-t)$, which is a conditionally positive definite function with respect to the space Π_m of at most degree m spherical harmonics. The (full) Lagrange function at the point $\xi \in X$ is $\chi_\xi(\cdot) = \sum_\eta a_{\xi,\eta}\kappa(\cdot,\eta) + \sum_{k=1}^{L} b_{k,\xi}\psi_k(\cdot)$, which is constructed by solving an interpolation problem with the data $\chi_\xi(\eta) = \delta_{\xi,\eta}$. As we've mentioned before, this requires solving a system of size $\mathcal{O}(N)$ for each Lagrange function, where N is the cardinality of X. The *local* Lagrange function will be constructed in a similar fashion, but with fewer points.

Let $r > 0$ be a fixed number. Let $\Upsilon(\xi) := \{\eta \in X : d(\xi,\eta) < r\}$. $\Upsilon(\xi)$ is the collection of the nearest neighbors of ξ in the set X. Let K_ξ denote the matrix $K_\xi(\eta,\zeta) := \kappa(\eta,\zeta)$ for $\eta,\zeta \in \Upsilon(\xi)$ and let Ψ_ξ denote the matrix $\Psi_\xi(\eta,k) = \psi_k(\eta)$ for $\eta \in \Upsilon(\xi)$ and $\psi_k \in \Pi_m$.

We define the local Lagrange function centered at ξ, denoted $\hat{\chi}_\xi$ to be the unique function that interpolates 1 at ξ and 0 at $\eta \neq \xi \in \Upsilon(\xi)$ in the space $V_{\Upsilon(\xi)}$. That is,

$$\begin{aligned} \hat{\chi}_\xi(\cdot) &:= \sum_{\eta\in\Upsilon(\xi)} a_{\eta,\xi}\kappa(\cdot,\eta) + \sum_{k=0}^{L} b_{k,\xi}\psi_k(\cdot), \\ \text{subject to } &\sum_{\eta\in\Upsilon(\xi} a_{\eta,\xi}p(\eta) = 0 \text{ for each } p \in \Pi_m. \end{aligned} \tag{2.2}$$

To solve for $a_{\eta,\xi}$ and $b_{k,\xi}$, we solve the linear system

$$\begin{pmatrix} K_\xi & \Psi_\xi \\ \Psi_\xi^T & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \delta_{\eta,\xi} \\ 0 \end{pmatrix}. \tag{2.3}$$

Notice that the local Lagrange function is constructed only using kernels centered in $\Upsilon(\xi)$. We may estimate the number of kernels in this set using the radius r chosen for $\Upsilon(\xi)$ and the quasi-uniformity of the set of centers X. On the sphere $\mathbb{S}^d$, for some constant C_d, the ball $\mu\big(B(\xi,r)\big) = C_d r^d$. Since $\Upsilon(\eta) \subset B(\xi,r)$, we know $B(\eta,q) \subset B(\xi,r+q)$ for each η. Furthermore, $B(\eta,q) \cap B(\zeta,q) = 0$ for $\eta,\zeta \in \Upsilon(\xi)$. Consequently,

$$\mu\Big(\bigcup_{\eta\in\Upsilon(\xi)} B(\eta,q)\Big) \leq \mu\big(B(\xi,r+q)\big)$$

and consequently,

$$\#\Upsilon(\xi) C_d q^d \leq C_d (r+q)^d$$

Assume $q < r$, and consequently, $q + r < 2r$. We then have the estimate

$$\#\Upsilon(\xi) \leq 2^d \left(\frac{r}{q}\right)^d. \tag{2.4}$$

Furthermore, we may bound the total number of centers, $N = \#X$ by observing that $\mathbb{S}^n = \cup_{\xi\in X} B(\xi,h)$. For, if $x \in \mathbb{S}^n$, then $d(\xi,x) \leq h$ by the definition of h. Therefore, it follows that

$$\mu(\mathbb{S}^d) \leq \sum_{\xi\in X} \mu\big(B(\xi,h)\big) = C_d N h^d,$$

so we see that $N \sim h^{-d}$. As a consequence of these results, if we choose $r =$

$Kh|\log(h)|$ for some fixed positive $K > 0$ in the definition of $\Upsilon(\xi)$, we find

$$\#\Upsilon(\xi) \leq 2^d \left(\frac{h}{q} |\log(h)| \right)^d \leq \big(2\rho d \log(N)\big)^d. \tag{2.5}$$

Since the mesh ratio ρ is bounded for quasi-uniformly distributed sets of centers, we have $\#\Upsilon(\xi)$ scales as $M \log(N)^d$ where N is the number of centers. The significance of this result is that the size of the system to be solved for the assembly of a single local Lagrange function in Equation (2.3) is on the order of $\mathcal{O}\big(\log(N)^d\big)$. In contrast, constructing a single full Lagrange function requires solving a system of size $\mathcal{O}(N)$. We remark briefly on the computational and practical aspects of the local Lagrange functions to contrast them with the full Lagrange functions. From a computational standpoint, the local Lagrange function assembly is now practical. A log scaling of the number of centers in $\Upsilon(\xi)$ implies that the linear systems to be solved are relatively small. Furthermore, the storage required for the local Lagrange functions is far smaller. Each Lagrange function requires storing the vector $A_\xi := (a_{\xi,\eta})$ which is also of size $\mathcal{O}\big(\log(N)^d\big)$ as opposed to storing $\mathcal{O}(N)$ entries for the full Lagrange function. While this may seem insignificant at first, there are N total vectors for each ξ to be stored. Consequently, storing all the full Lagrange function coefficients requires $\mathcal{O}(N^2)$ rather than $\mathcal{O}\big(N \log(N)^d\big)$ for all of the local Lagrange functions.

In addition to significant savings, the local Lagrange functions may be constructed *in parallel*. This is a significant difference that reduces the cost of constructing all of them. While each function only requires solving a small linear system on the order of $\mathcal{O}\big(\log(N)^d\big)$, there are still N functions in total to be constructed. This remains a significant computational burden to construct. Taking advantage of the embarrassingly parallel nature of the local Lagrange assembly routine enables significant savings. Experiments in Python have demonstrated the construction of the local

Lagrange functions benefits from employing simple parallelism by distributing the assembly tasks among multiple processing cores.

We now consider the theoretical properties of the local Lagrange functions and discuss how well they compare with the full Lagrange basis. We consider their use for interpolation and for pre-conditioning problems. By their construction, the local Lagrange functions enjoy many of the properties the full Lagrange functions do. They exhibit L^p stability as well as spatial localization, although the decay is algebraic rather than exponential [7]. Most importantly, the local Lagrange functions provide near optimal L^∞ approximation [7]. That is, the difference between using a local Lagrange function or a full Lagrange function is negligible in the sup norm. However, a caveat for all of these statements is that the number of points for each local Lagrange function must be chosen correctly. Simply choosing a few nearest neighbors is inadequate; choosing all nearest neighbors within a distance $Kh|\log(h)|$ is the appropriate scalable choice, with appropriate choice of K. Choosing too few points will yield a fast assembly, but unpredictable behavior of the local Lagrange function.

Proposition 9. *[7][Proposition 6.5] Let κ_m be a thin plate spline of order m on the sphere $\mathbb{S}^2$ and let $X \subset \mathbb{S}^2$ be a collection of quasi-uniformly scattered centers. Let $K > 0$ be chosen so that $K > \frac{4m-2+2\mu}{\nu}$ and let $\Upsilon(\xi) := B(\xi, Kh|\log(h)|) \cap X$. Let $\hat{\chi}_\xi$ denote the local Lagrange function centered at ξ. Let $J = K\nu - 4m + 2 + 2\mu$. Then,*

$$\|\chi_\xi - \hat{\chi}_\xi\|_{L^\infty} \leq Ch^J \tag{2.6}$$

$$|\hat{\chi}_\xi(x)| \leq C\big(1 + d(\xi, x)\backslash h\big)^{-J}. \tag{2.7}$$

If $J > 2$, then the basis is L^p stable. There exists positive constants C_1 and C_2 so

that

$$C_1 q^{\frac{2}{p}} \|a\|_{\ell^p(X)} \leq \| \sum_{\xi} \hat{\chi}_\xi \|_{L^p(\mathbb{S}^2)} \leq C_2 q^{\frac{2}{p}} \|a\|_{\ell^p(x)}. \tag{2.8}$$

The significance of Proposition 9 is that it implies the local Lagrange basis maintains the advantages of the full Lagrange basis while being computationally tractable. Since only small systems need to be solved, the local Lagrange functions can be assembled practically (and in parallel). They are L^p stable and the difference between the full and local Lagrange function can be tuned by the parameter K, which is related to the radius of the ball used in the construction of the point set $\Upsilon(\xi)$.

The local Lagrange functions may be used directly in the form of *quasi-interpolation*. We define the quasi-interpolant of a continuous function f, denoted $Q_X f$, by sampling the function on the centers and using these values as weights for the local Lagrange functions. That is,

$$Q_X f(\cdot) = \sum_{\xi \in X} f(\xi) \hat{\chi}_\xi(\cdot).$$

The quasi-interpolant is not a true interpolant: the $\hat{\chi}_\xi(\eta) \neq \delta_{\xi,\eta}$ for all $\eta \in X$; rather, $\hat{\chi}_\xi(\eta) = \delta_{\xi,\eta}$ for $\eta \in \Upsilon(\xi)$. However, we may demonstrate that the error is quite negligible relative to the full Lagrange interpolant $I_X f$. We note

$$|f(x) - Q_X f(x)| \leq |f(x) - I_X f(x) + |Q_X(f) - I_X(f)|$$

and consequently, to guarantee that $Q_X f$ approximates f well, the error for $Q_x(f) -$

$I_X(f)$ must be controlled. We see that

$$\begin{aligned}|Q_X f(x) - I_X f(x)| &= \left| \sum_{\xi} f(\xi)\big(\chi_\xi(x) - \hat{\chi}_\xi(x)\big) \right| \\ &\leq \|f\|_{L^\infty} \sum_{\xi} \|\chi_\xi - \hat{\chi}_\xi\|_{L^\infty} \\ &\leq Cq^{-2}\|f\|_{L^\infty} h^{K\nu-4m+2+2\mu} = Ch^{K\nu-4m+2\mu}\|f\|_{L^\infty}\end{aligned}$$

where we applied the L^∞ error estimate from Proposition 9, noted that the cardinality of X is bounded above by Cq^{-2} on $\mathbb{S}^2$, and used quasi-uniformity to eliminate an h^2q^{-2} term. As a consequence of this result, we may choose K large enough to guarantee that $K\nu - 4m + 2\mu > 2m$ which guarantees optimal order approximation for functions $f \in C^{2m}(\mathbb{S}^n)$. That is, by choosing K large enough, the convergence order of the quasi-interpolant is the same as the full Lagrange interpolant. This suggests that there is no significant loss in approximation by choosing the local Lagrange functions over the full Lagrange basis. The local Lagrange functions inherit the stability of the basis as well as the near best L^∞ approximation order, provided the radius parameter K is chosen sufficiently large in the definition of $\Upsilon(\xi)$.

The local Lagrange functions also provide a *preconditioner* for spherical basis function interpolation. The full Lagrange function interpolant, $I_X f$, may be written in terms of the local Lagrange functions by solving for the coefficients $\hat{a}_\xi$ in the equation

$$I_X f := \sum_{\xi} a_\xi \kappa(\cdot, \xi) + \sum_k b_k \psi_k = \sum_{\xi} \hat{a}_\xi \hat{\chi}_\xi.$$

The objective is to construct the full Lagrange interpolant by solving for the vector of coefficients a and b. Solving for these coefficients requires solving a $\mathcal{O}(N)$ size linear system, where N is the number of centers in X.

Let A_Υ and C_Υ stand for the matrices storing the local Lagrange coefficients where $A_\Upsilon(\xi,\eta) = a_{\xi,\eta}$ and $C_\Upsilon(k,\eta) = b_{k,\eta}$, where the columns of A_Υ and C_Υ are found by solving Equation (2.3) for each ξ. Then, since $I_X f = \sum_\xi \hat{a}\hat{\chi}_\xi$ interpolates a function f on the centers of X, we have

$$\begin{pmatrix} K_X & \Psi \end{pmatrix} \begin{pmatrix} A_\Upsilon \\ C_\Upsilon \end{pmatrix} \hat{a} = f|_X \tag{2.9}$$

where $K_X(\xi,\eta) = \kappa(\xi,\eta)$ for $\xi,\eta \in X$ and $\Psi_X(\eta,k) = \psi_k(\eta)$ for $\eta \in X$ and for each spherical harmonic $\psi_k \in \Pi_m$. Therefore, we have that

$$\begin{pmatrix} a & c \end{pmatrix}^T = \begin{pmatrix} A_\Upsilon \\ C_\Upsilon \end{pmatrix} \hat{a}. \tag{2.10}$$

In this sense, the local Lagrange functions act as a right-preconditioner on the system (2.9). Consequently, the system in (2.9) can be solved for the coefficient vector $\hat{a}$ first. Then, the coefficients a and c may be computed by the matrix multiplication in (2.10). See Section 7 of [7] for details and numerical experiments. The numerical experiments considered in [7] suggest that solving (2.9) can be solved by GMRES efficiently. Using very few centers per local Lagrange function, the preconditioned system was solved via GMRES with few iterations independent of the total size of the system. This suggests the local Lagrange functions provide a highly efficient method for determining the full Lagrange function interpolant by an efficient, theoretically verifiable preconditioner.

2.3 Pointwise Convergence of Interpolants and Quasi-interpolants

We present here results which demonstrate pointwise convergence of the interpolant to arbitrary continuous functions. This differs greatly from previously men-

tioned error estimates, which always assume a level of smoothness in the functions. Here, as long as continuity is available, we are able to demonstrate convergence.

Theorem 3. *Let f be a continuous function on $\mathbb{S}^n$ and let $\{X_h\}$ be a collection of quasi-uniformly distributed centers. Then, as h converges to zero, I_{X_h} converges pointwise to f.*

Proof. Fix h and and fix $x \in \mathbb{S}^n$. Let $\{\chi_i\}$ be a set of Lagrange functions centered at x_i respectively. Let $I_X f$ be the interpolant to f constructed from the Lagrange functions. We demonstrate that $|f(x) - I_X f(x)| \sim \max\left(\omega(f, Kh\log(h), h^{\nu k}\right)$, where $\omega(f, R)$ is the modulus of continuity of f with radius R. Consequently, as $h \to 0$, $I_X f$ converges pointwise to f.

We invoke several facts regarding Lagrange functions. First, we know $\sum_{i=1}^N \chi_i(x) = 1$ for all $x \in \mathbb{S}^n$ and the Lagrange functions decay exponentially. Let $K_h := Kh|\log(h)|$ and let $B_x = \{x_i \in X : \|x_i - x\| \leq K_h\}$.

$$
\begin{aligned}
|f(x) - I_X f(x)| = |f(x) - \sum_{i=1}^N f(x_i)\chi_i(x)| &= \left| \sum_{i=1}^N \big(fx) - f(x_i)\big)\chi_i(x) \right| \\
&\leq \underbrace{\sum_{x_i \in B_x} |f(x_i) - f(x)|\,|\chi_i(x)|}_{I} + \underbrace{\sum_{x \notin B_x} |f(x_i) - f(x)|\,|\chi_i(x)|}_{II}.
\end{aligned}
$$

We first discuss term I. We note that since each $x_i \in B_x$, $|f(x_i) - f(x)| \leq \omega(f, K_h)$. Furthermore, by Proposition 4, we know $\sum_{i=1}^N |\chi_i(x)| \leq C$ is bounded independent of h or q. Consequently,

$$
\sum_{x_i \in B_x} |f(x_i) - f(x)|\,|\chi_i(x)| \leq \omega(f, K_h) \sum_{i=1}^N |\chi_i(x)| \leq C\omega(f, K_h).
$$

For term II, we may invoke the exponential decay of the Lagrange functions. By Proposition 2, we have $|\chi_i(x)| \leq C \exp\left(-\nu \frac{d(x,x_i)}{h} \right)$. Since $d(x, x_i) \geq Kh \log(h)$, we have $|\chi_i(x)| \leq h^{\nu K}$. Furthermore, we estimate the cardinality of the set $X \cap B_x^c \sim Cq^{-d}$. Therefore, we have

$$\begin{aligned}
\sum_{x_i \notin B_x} |f(x) - f(x_i)| \, |\chi_i(x)| &\leq \sum_{x_i \notin B_x} \sup_{x, x_i} |f(x) - f(x_i)| \, |\chi_i(x)| \\
&\leq 2\|f\|_{L^\infty} \sum_{x_i \notin B_x} |\chi_i(x)| \\
&\leq 2\|f\|_{L^\infty} q^{-d} h^{\nu K} \leq C h^{\nu K - d} \|f\|_{L^\infty}.
\end{aligned}$$

Putting this all together, we find

$$|f(x) - I_X f(x)| \leq C(\omega(f, K_h) + h^{\nu K - d} \|f\|_{L^\infty}).$$

Therefore, as $h \to 0$, the interpolant converges pointwise to f. □

Corollary 1. *Let f be a continuous function on $\mathbb{S}^n$ and let $\{X_h\}$ be a collection of quasi-uniformly distributed centers. Assume the local Lagrange functions are constructed using all points within a ball of radius $Kh|\log(h)|$ for appropriately large K. Then, as $h \to 0$, the quasi-interpolant $\hat{I}_X f$ converges pointwise to f.*

Proof. This follows as a direct corollary of Section 2.3 and properties of the local Lagrange functions. Fix h and a point $x \in X$. Let $I_X f$ denote the Lagrange function interpolant and let $\hat{I}_X f$ denote the local Lagrange quasi-interpolant. Then,

$$|f(x) - \hat{I}_X f(x)| = |f(x) - I_X f(x)| + |I_X f(x) - \hat{I}_X f(x)|.$$

We argue that $|I_X f(x) - \hat{I}_X f(x)|$ converges to zero. Recall that the local Lagrange

function satisfies the assumption $|\hat{\chi}_\xi(x) - \chi_\xi(x)| \leq Ch^J$ where the value of J depends on K and may be tuned by increasing K. Therefore,

$$\begin{aligned} |I_X f(x) - \hat{I}_X f(x)| &= \left| \sum_\xi f(\xi)\big(\chi_\xi(x) - \hat{\chi}_\xi(x)\big) \right| \\ &\leq \|f\|_{L^\infty} \sum_\xi \sup_{x \in \mathbb{S}^n} |\chi_\xi(x) - \hat{\chi}_\xi(x)| \\ &\leq C\|f\|_{L^\infty} q^{-n} h^J \leq C\|f\|_{L^\infty} h^{J-n}. \end{aligned}$$

Consequently, we have $|I_X f(x) - \hat{I}_X f(x)| \leq Ch^{J-n}$. By choosing K larger, if necessary, we may guarantee $J - n > 0$, and hence the pointwise difference between the quasi-interpolant and the interpolant decreases as $\mathcal{O}(h^{J-n})$. Applying this observation along with the pointwise convergence of $I_X f(x)$ to $f(x)$ yields the result. $\square$

3. LAGRANGE FUNCTION QUADRATURE

We discuss a novel quadrature method for manifolds. Given data samples of a function on a manifold, a difficult problem is approximating the integral of the function over the manifold. A robust numerical quadrature method must be capable of approximating an integral provided scattered data samples of a function. That is, given data centers, the quadrature weights must be constructed corresponding to the locations of the given data points. Relying on quadrature routines that assume specific locations of the centers is not acceptable: the method must be capable of handling possibly scattered, irregularly spaced data. Lagrange functions of certain spherical basis functions may be used to construct a quadrature routine for scattered data that yields provable error estimates for smooth data.

The absence of implementable, practical quadrature routines for radial basis functions has impeded the application of radial (spherical) basis functions for solving partial differential equations via weak formulations. Galerkin methods which rely on the weak formulation require the ability to compute integrals efficiently. Generating quadrature routines for spherical basis functions and radial basis functions enables one to practically implement a numerical Galerkin method using radial basis functions. This should be contrasted with the traditional radial basis function approach for partial differential equations, which is typically to use a collocation method. These "strong-form" methods often suffer from a dearth of theoretical justification, although they often perform quite admirably in practice. Most notably, theoretical error estimates for many radial basis function collocation methods lack error estimates.

3.1 The Quadrature Routine on Spheres

Let $X \subset \mathbb{S}^n$ be a quasi-uniformly distributed set of centers and let $\{\chi_\xi\}_{\xi \in X}$ be a collection of Lagrange functions corresponding to a kernel $\kappa(\cdot, \cdot)$, such as the restricted surface spline $\phi_s(t) := (-1)^{s+1}(1-t)^s \log(1-t)$. Our objective is to construct a collection of quadrature weights $\{w_\xi\}_{\xi \in X}$ so that, given a collection of data samples $\{f(\xi)\}_{\xi \in X}$ for a function f, we may approximate the integral of f by

$$\int_{\mathbb{S}^n} f(x)\, d\mu(x) \approx Q_X(f) := \sum_{\xi \in X} f(\xi) w_\xi.$$

The approach, first presented in [8], is to construct the quadrature weights using the Lagrange functions. Let $V_X = \text{span}\{\chi_\xi\}_{\xi \in X}$ denote the approximation space generated by the conditionally positive definite kernel κ with respect to the finite dimensional subspace of continuous functions $\Pi = \text{span}\{\psi_k\}_{k=1}^L$. For example, on $\mathbb{S}^n$ with κ generated by the thin plate splines, Π could be the span of the first degree m spherical harmonic polynomials. Let $\Psi_{\xi,k} = \psi_k(\xi)$ and let $A_{\xi,\eta} = \kappa(\xi, \eta)$ as usual.

We first note that integrating kernels on $\mathbb{S}^n$ is invariant under rotation.

Lemma 2. *[8] The integral $J(y) := \int_{\mathbb{S}^n} \kappa(x, y) d\mu(x)$ is independent of y.*

As a consequence of Lemma 2, the integral of $\kappa(\cdot, \xi)$ is independent of ξ. We define the following quantities:

$$\begin{cases} J_0 = \int_{\mathbb{S}^n} \kappa(x, \xi) d\mu(x), \\ J_k = \int_{\mathbb{S}^n} \psi_k(x) d\mu(x) \qquad \text{for } k = 1, \ldots, L, \\ J = (J_1, \ldots, J_L). \end{cases}$$

Consider an arbitrary $s = \sum_{\xi \in X} s(\xi) \chi_\xi(\cdot) \in V_X$. We may decompose s in the basis

of rotations of κ plus an appropriate polynomial from Π by

$$s(\cdot) := \sum_{\xi \in X} a_\xi \kappa(\cdot, \xi) + \sum_{k=1}^{L} b_k \psi_k(\cdot).$$

Let a and b denote the vectors consisting, respectively, of the values a_ξ and b_ξ. We see that the integral of s satisfies

$$\int_{\mathbb{S}^n} s(x) d\mu(x) = \sum_{\xi \in X} s(\xi) \int_{\mathbb{S}^n} \chi_\xi(x) d\mu(x) := \sum_{\xi} s(\xi) c_\xi \tag{3.1}$$

where we have defined $c_\xi := \int_{\mathbb{S}^n} \chi_\xi(x) d\mu(x)$. Let c denote the vector of the c_ξ values; consequently, $\int_{\mathbb{S}^n} s(x) d\mu(x) := c^T s|_X$. Next, observe that the vectors a and b are the solution to the problem

$$\underbrace{\begin{pmatrix} A & \Psi \\ \Psi & 0 \end{pmatrix}}_{\mathbf{A}} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} s \\ 0 \end{pmatrix}. \tag{3.2}$$

Now, we repeat the computation from Equation (3.1), but this time in the basis of rotations of κ. We compute

$$c^T s|_X = \sum_{\xi} a_\xi \underbrace{\int_{\mathbb{S}^n} \kappa(x, \xi) d\mu(x)}_{J_0} + \sum_{k=1}^{L} b_k \underbrace{\int_{\mathbb{S}^n} \psi_k(x) d\mu(x)}_{J_k}. \tag{3.3}$$

Note that since $\mathbf{A}$ is self-adjoint, so is $\mathbf{A}^{-1}$. Furthermore, by applying Equation (3.2),

we find

$$\begin{aligned} c^T s|_X &= \begin{pmatrix} J_0 \mathbf{1}^T & J^T \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \begin{pmatrix} J_0 \mathbf{1}^T & J^T \end{pmatrix} \mathbf{A}^{-1} \begin{pmatrix} s|_X \\ 0 \end{pmatrix} \\ &= \left(\mathbf{A}^{-1} \begin{pmatrix} J_0 1 \\ J \end{pmatrix} \right)^T \begin{pmatrix} s|_X \\ 0 \end{pmatrix}. \end{aligned}$$

Consequently, we identify that the vector c is obtained via

$$\begin{pmatrix} c \\ d \end{pmatrix} = \mathbf{A}^{-1} \begin{pmatrix} J_0 1 \\ J \end{pmatrix}.$$

What this implies is that the values of c_ξ are formed by the usual interpolation matrix, albeit with a slightly modified right hand side. Now, choose $s := \chi_\xi$ and note that $\chi_\xi|_X := \delta_{\xi,\eta}$. Therefore, we get exactly that $s|_X c^T = c_\xi$, and by (3.3), we see that

$$c_\xi = \int_{\mathbb{S}^n} \chi_\xi(x) d\mu(x). \tag{3.4}$$

We may bound c_ξ above by the L^1 norm of χ_ξ by noting that

$$|c_\xi| = \left| \int_{\mathbb{S}^n} \chi_\xi(x) d\mu(x) \right| \leq \|\chi_\xi\|_{L^1(\mathbb{S}^n)}.$$

Putting these ideas together, we arrive at the quadrature construction result derived in [8]. We list here only the result for $\mathbb{S}^n$, although this may be generalized to other

manifolds.

Proposition 10. *[8] Let κ be a conditionally positive definite kernel with respect to Π on $\mathbb{S}^n$ and suppose that c and d are vectors which are the solutions to the problem*

$$\begin{pmatrix} A & \Psi \\ \Psi & 0 \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} J_0 1 \\ J \end{pmatrix}.$$

Then, for any $s(x) \in V_X$,

$$\int_{\mathbb{S}^n} s(x) d\mu(x) = \sum_{\xi \in X} c_\xi s(\xi).$$

As a result of this, we may define the *Lagrange Function Quadrature Rule.*

Definition 1. Let X be a collection of scattered centers on $\mathbb{S}^n$, let κ be an appropriate conditionally positive definite kernel (e.g., the thin plate spline), and let f be a continuous function on $\mathbb{S}^n$. We define the quadrature rule $Q_{V_X}(f)$ which approximates

$$Q_{V_X}(f) := \sum_{\xi} f(\xi) c_\xi \approx \int_{\mathbb{S}^n} f(x) d\mu(x)$$

where the c_ξ are defined in (3.4).

Now that we have a quadrature routine, we discuss the construction of the quadrature weights. While (3.4) tells us the weights may be computed by integrating a Lagrange function χ_ξ, we emphasize that this is not the method one should use for the assembly of the weights. First, this necessitates the construction of the Lagrange function as a pre-processing step. This is a computationally intense task which should be avoided if possible. Furthermore, each weight requires the integral of the Lagrange function to be computed. This means each weight requires a highly accu-

rate quadrature of the Lagrange function, which is rather circular since our objective is to obtain the weights in the first place. Instead, the weights should be constructed by solving the linear system of equations in Proposition 10.

In the event of a positive definite kernel, the quadrature weights are constructed by directly solving the interpolation problem for the constant function $f(x) := \frac{1}{J_0}$. Our focus is on the case of conditionally positive definite functions with respect to some space Π. We need to construct the values c. First, we split c into two orthogonal pieces $c = c_{||} + c_{\perp}$ where $c_{||}$ is the orthogonal projection of c onto the range of Ψ. Let $P := \Psi(\Psi^T\Psi)^{-1}\Psi^T$ be the orthogonal projection onto the range of Ψ. We know by Proposition 10 that $J = \Psi^T c$, and consequently, $c_{||} = Pc = \Psi(\Psi^T\Psi)^{-1}J$. This is simply matrix multiplication, and consequently we have $c_{||}$. We may attempt to solve for $c_{\perp}$ now. First, since $Pc_{\perp} = 0$, we have $\Psi^T c_{\perp} = 0$. We know

$$Ac + \Psi d = J_0 1$$

and consequently, by splitting c into its orthogonal components, we arrive at

$$Ac_{\perp} + \Psi d = J_0 1 - Ac_{||} = J_0 1 - A\Psi(\Psi^T\Psi)^{-1}J$$

and we further have the constraint $\Psi^T c_{\perp} = 0$. Consequently, we may solve for $c_{\perp}$ by solving the interpolation problem

$$\begin{pmatrix} A & \Psi \\ \Psi^T & 0 \end{pmatrix} \begin{pmatrix} c_{\perp} \\ d \end{pmatrix} = \begin{pmatrix} J_0 1 - A\Psi(\Psi^T\Psi)J \\ 0 \end{pmatrix}.$$

See [8] for more details as well as a discussion of how to solve for $c_{\perp}$ only without the need to solve for d. This system may be pre-conditioned using the local Lagrange

functions as discussed in Section 2.2.

We note that the quadrature formula perfectly reproduces the integrals for functions in the space V_X. We now aim to study how the quadrature error varies with the space V_X. As usual, we consider quasi-uniformly distributed sets X with mesh norm $h := h_X$. Then, we may derive the following quadrature error estimate immediately.

Lemma 3. *[8]Let X be a quasi-uniformly distributed set of centers and let κ_m be a polyharmonic kernel with respect to the space Π of continuous functions on $\mathbb{S}^n$. Then, for $f \in W_2^k(\mathbb{S}^n)$ where $\frac{n}{2} < k \leq m$,*

$$\left| \int_{\mathbb{S}^n} f(x)d\mu(x) - Q_{V_X}(f) \right| \leq Ch^k \|f\|_{W_2^k(\Omega)}.$$

Proof.

$$\begin{aligned}
\left| \int_{\mathbb{S}^n} f(x)d\mu(x) - Q_{V_X}(f) \right| &= \left| \int_{\mathbb{S}^n} f(x)d\mu(x) - \sum_{\xi} f(\xi)c_\xi \right| \\
&= \left| \int_{\mathbb{S}^n} f(x)d\mu(x) - \sum_{\xi} f(\xi) \int_{\mathbb{S}^n} \chi_\xi(x)d\mu(x) \right| \\
&= \left| \int_{\mathbb{S}^n} \big(f(x) - \sum_{\xi} f(\xi)\chi_\xi(x)\big)d\mu(x) \right|.
\end{aligned}$$

We see that this is simply the L^1 error between f and its interpolant in V_X, $I_X f = \sum_\xi f(\xi)\chi_\xi$. Applying the error estimates as usual yields the result. Alternatively, an application of Cauchy-Schwarz yields a constant multiple of the L^2 error, and the usual L^2 error estimate may be applied which yields an order h^k convergence rate. □

An optimal error estimate is presented for functions in C^{2m} in [8]. By applying the "intermediate doubling trick" from [25], we extend that result to handle values

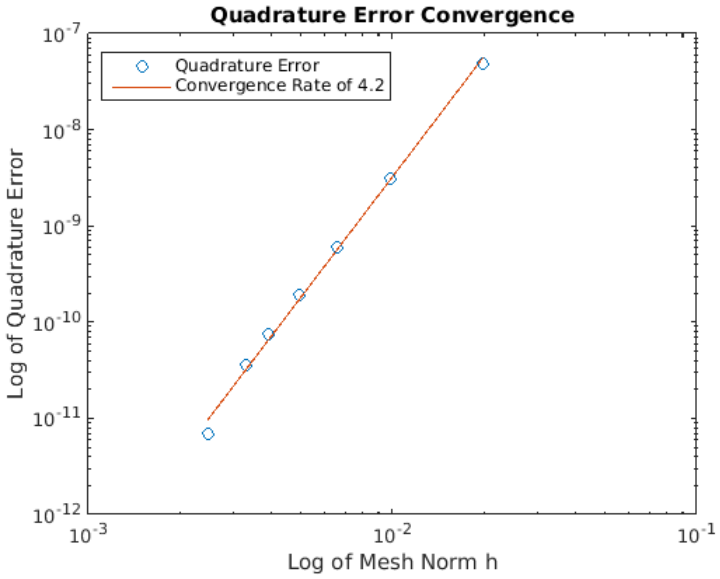


Figure 3.1: The function $f(\theta) = \cos(\theta)\exp(\cos(\theta))$ is integrated on the sphere numerically with icosahedral nodes ranging from 2562 nodes to 163842 nodes. The quadrature error decays at a rate of $\mathcal{O}(h^4)$.

between m and $2m$.

Lemma 4. *Under the same assumptions as Lemma 3, consider $f \in W_2^k(\mathbb{S}^n)$ for $\frac{n}{2} < k \leq 2m$. Then,*

$$\left| \int_{\mathbb{S}^n} f(x)d\mu(x) - Q_{V_X(}(f) \right| \leq Ch^k \|f\|_{W_2^k(\Omega)}.$$

Proof. The proof follows in an identical fashion as Lemma 3 with the application of the improved error estimate [25]. □

4. NONLOCAL DIFFUSION

The purpose of this section is to introduce a meshfree method for the solution of an anisotropic nonlocal diffusion equation. We take advantage of a recently developed approximation and interpolation scheme to construct an approximation space to discretize the weak form of a nonlocal diffusion problem. We present a new quadrature method unique to the discretization that provides a method to generate a sparse stiffness matrix. Evaluating the entries in the stiffness matrix follows by pointwise evaluations of a kernel and multiplication by quadrature weights. Computing entries in the stiffness matrix for a piecewise polynomial finite element discretization is a non-trivial computational challenge currently. For a problem in $\mathbb{R}^n$, evaluating the stiffness matrix entries requires $2n$-iterated integrals over partial element volumes. The paper [1] studied radial basis function methods for the discretization of the nonlocal diffusion equation by employing the use of a localized basis and an associated quadrature routine. The new approach we present here reduces the computational difficulty of both the construction of the quadrature weights and the evaluation of the solution on a set of points. The method we previously developed in [1] was primarily to study whether or not we could employ radial basis function techniques to nonlocal diffusion problems. Having established that, we pursued a superior computational method along with a goal of placing the method on a more sound theoretical basis. The approach we present here maintains the same benefits of the radial basis function method in [1] without the need to solve large, dense linear systems. In addition to the computational improvements, we now consider anisotropic nonlocal diffusion equations and we demonstrate that both the continuum and discrete problems are well-posed. The first approach we describe here is joint work with Lehoucq.

4.1 Nonlocal Vector Calculus

We present background material and topics from the nonlocal vector calculus required to define the nonlocal diffusion equation. The nonlocal vector calculus provides nonlocal analogues of classical operators such as the gradient, divergence, and curl operators. The nonlocal vector calculus was developed in [9].

Let $\nu(\boldsymbol{x},\boldsymbol{y}), \alpha(\boldsymbol{x},\boldsymbol{y}) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^k$ where α is an anti-symmetric mapping, i.e., $\alpha(\boldsymbol{x},\boldsymbol{y}) = -\alpha(\boldsymbol{y},\boldsymbol{x})$. The nonlocal divergence operator $\mathcal{D}$ acts on ν by

$$\big(\mathcal{D}\nu\big)(\boldsymbol{x}) := \int_{\mathbb{R}^n} \big(\nu(\boldsymbol{x},\boldsymbol{y}) + \nu(\boldsymbol{y},\boldsymbol{x})\big) \cdot \alpha(\boldsymbol{x},\boldsymbol{y})\, d\boldsymbol{y}.$$

The adjoint operator $\mathcal{D}^*$ acts on $u(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ pointwise by

$$\mathcal{D}^*(u)(\boldsymbol{x},\boldsymbol{y}) = -\big(u(\boldsymbol{y}) - u(\boldsymbol{x})\big)\alpha(\boldsymbol{x},\boldsymbol{y}) \qquad \text{for } \boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n,$$

where $\mathcal{D}^*u : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^k$. We remark that $\mathcal{L}$ may be considered a nonlocal analogue of the classical Laplacian. Section 4.2.1 discusses the relationship between nonlocal operators and differential operators in more detail.

For an open subset $\Omega \subset \mathbb{R}^n$, we define the interaction domain

$$\Omega_{\mathcal{I}} := \{\boldsymbol{y} \in \mathbb{R}^n \backslash \Omega : \alpha(\boldsymbol{x},\boldsymbol{y}) \neq 0 \text{ for some } \boldsymbol{x} \in \Omega\}. \tag{4.1}$$

Given functions $f : \Omega \to \mathbb{R}$ and $g : \Omega_{\mathcal{I}} \to \mathbb{R}$, we want to solve the weak formulation of the steady-state nonlocal diffusion problem

$$\begin{cases} \mathcal{L}u = f & \text{on } \Omega, \\ \quad u = g & \text{on } \Omega_{\mathcal{I}}, \end{cases} \tag{4.2}$$

where $\Theta(\boldsymbol{x}, \boldsymbol{y})$ is a second-order tensor satisfying $\Theta = \Theta^T$, and the nonlocal diffusion operator is defined to be

$$\mathcal{L}u(\boldsymbol{x}) = 2 \int_{\Omega \cup \Omega_{\mathcal{I}}} \big(u(\boldsymbol{y}) - u(\boldsymbol{x})\big)\alpha(\boldsymbol{x}, \boldsymbol{y}) \cdot \big(\Theta(\boldsymbol{x}, \boldsymbol{y}) \cdot \alpha(\boldsymbol{x}, \boldsymbol{y})\big)\, d\boldsymbol{y}\,, \quad \boldsymbol{x} \in \Omega\,. \tag{4.3}$$

The nonlocal diffusion model places conditions over a positive measure volume, which we refer to as a *volume constraint*. This should be contrasted with a partial differential equation, which places constraints on a measure zero boundary. The volume constraint is sufficient to guarantee that the weak formulation of (4.2) is well-posed, provided certain conditions on the kernel. For the case of integrable kernels, [3] demonstrates that (4.2) is well-posed on the constrained energy space $L^2_c(\Omega \cup \Omega_{\mathcal{I}}) = \{u \in L^2(\Omega \cup \Omega_{\mathcal{I}}) : u|_{\Omega_{\mathcal{I}}} = 0 \text{ a.e.}\}$.

Let $s \in [0, \frac{1}{2}]$ and consider nonlocal operators $\mathcal{L}$ that map $H^s(\Omega \cup \Omega_{\mathcal{I}})$ to its dual space $\big($or $L^2_c(\Omega \cup \Omega_{\mathcal{I}})$ to $L^2_c(\Omega \cup \Omega_{\mathcal{I}})\big)$. Imposing boundary conditions on $\partial\Omega$ rather than volume constraints on $\Omega \cup \Omega_{\mathcal{I}}$ is not possible. For $s \in [0, \frac{1}{2}]$ trace operators do not exist, and hence imposing $u|_{\partial\Omega} = g$ is not well defined. The nonlocal diffusion problem is well-posed with the addition of a volume constraint, even for $s \in [0, \frac{1}{2}]$, provided certain conditions on the kernel. Let $u \in L^2(\Omega \cup \Omega_{\mathcal{I}})$ and let $\gamma_\epsilon := \alpha \cdot \Theta \cdot \alpha$ be a radial kernel with support radius, or horizon, ϵ. Under general conditions, as $\epsilon \to 0$, the solution u_ϵ of (4.2) converges to the solution of

$$\begin{cases} \nabla \cdot \mathbf{C}\, \nabla u = f & \text{on } \Omega \\ \qquad\qquad u = g & \text{on } \partial\Omega\,, \end{cases} \tag{4.4}$$

where $\mathbf{C}$ is a diffusion tensor. For details, further exposition on nonlocal operators and comparisons between classical and nonlocal diffusion operators, as well as additional theory, the interested reader would be wise to consult [3, Section 3 pp.674-678].

The aforementioned paper also discusses comparisons between nonlocal calculus and the usual vector calculus in far more detail than we present here. The recent paper [4] explores the connections between nonlocal problems of the form (4.4) with a Neumann boundary condition and a smoothed particle hydrodynamic approximation.

We now begin by demonstrating that the solution u of the nonlocal diffusion equation (4.2) is the minimizer of a variational problem, the weak formulation of (4.2). Let $\Omega \subset \mathbb{R}^n$ be an open region and let $\Omega_{\mathcal{I}}$ be the corresponding interaction domain as defined in (4.1). The energy functional is defined to be

$$\begin{aligned} E(u; f) := \frac{1}{2} \int_{\Omega \cup \Omega_{\mathcal{I}}} \int_{\Omega \cup \Omega_{\mathcal{I}}} \mathcal{D}^*(u)(\boldsymbol{x}, \boldsymbol{y}) \cdot \big(\Theta(\boldsymbol{x}, \boldsymbol{y}) \cdot \mathcal{D}^*(u)(\boldsymbol{x}, \boldsymbol{y})\big)\, d\boldsymbol{x}\, d\boldsymbol{y} \\ - \int_{\Omega} f(\boldsymbol{x}) u(\boldsymbol{x})\, dx \end{aligned}$$

where f is a given function defined on Ω. Let $g(\boldsymbol{x})$ denote a function defined on $\Omega_{\mathcal{I}}$ and let $E_c(u; g)$ denote the constraint functional

$$E_c(u; g) := \int_{\Omega_{\mathcal{I}}} \big(u(\boldsymbol{x}) - g(\boldsymbol{x})\big)^2\, d\boldsymbol{x}. \tag{4.5}$$

We are interested in the constrained minimization problem of finding u such that

$$\min E(u; f) \qquad \text{subject to} \qquad E_c(u; g) = 0.$$

The constraint functional enforces a nonlocal Dirichlet volume constraint, which may be viewed as a nonlocal analogue to Dirichlet boundary conditions for differential equations. By choosing appropriate test functions v that satisfy $E_c(v; 0) = 0$, the

necessary conditions for the minimization problem are found to be

$$\int_{\Omega\cup\Omega_{\mathcal{I}}}\int_{\Omega\cup\Omega_{\mathcal{I}}} \mathcal{D}^*(u)(\boldsymbol{x},\boldsymbol{y})\cdot\big(\Theta(\boldsymbol{x},\boldsymbol{y})\cdot\mathcal{D}^*(v)(\boldsymbol{x},\boldsymbol{y})\big)\,d\boldsymbol{y}\,d\boldsymbol{x} = \int_{\Omega} f(\boldsymbol{x})v(\boldsymbol{x})\,d\boldsymbol{x}. \quad (4.6)$$

To relate (4.6) to (4.2), we make use of a nonlocal analogue of Green's first identity. Define the interaction operator $\mathcal{N}(\nu):\mathbb{R}^n\to\mathbb{R}$ by

$$\mathcal{N}(\nu)(\boldsymbol{x}) := -\int_{\Omega\cup\Omega_{\mathcal{I}}}\big(\nu(\boldsymbol{x},\boldsymbol{y})+\nu(\boldsymbol{y},\boldsymbol{x})\big)\cdot\alpha(\boldsymbol{x},\boldsymbol{y})\,d\boldsymbol{y} \qquad \text{for } \boldsymbol{x}\in\Omega_{\mathcal{I}}.$$

The nonlocal Green's first identity is

$$\int_{\Omega} v\mathcal{D}\big(\Theta\cdot\mathcal{D}^*u\big)\,d\boldsymbol{x} - \int_{\Omega\cup\Omega_{\mathcal{I}}}\int_{\Omega\cup\Omega_{\mathcal{I}}}\big(\mathcal{D}^*v\big)\cdot\big(\Theta\cdot\mathcal{D}^*u\big)\,d\boldsymbol{y}\,d\boldsymbol{x} = \int_{\Omega_{\mathcal{I}}} v\mathcal{N}\big(\Theta\cdot D^*u\big)\,d\boldsymbol{x}. \quad (4.7)$$

We apply (4.7) to (4.6) and we note that $v=0$ in $\Omega_{\mathcal{I}}$ by definition of the test functions to obtain

$$\int_{\Omega} v(\boldsymbol{x})\mathcal{D}\big(\Theta\cdot\mathcal{D}^*u\big)(\boldsymbol{x})\,d\boldsymbol{x} = \int_{\Omega} f(\boldsymbol{x})v(\boldsymbol{x})\,d\boldsymbol{x}. \quad (4.8)$$

Since this holds for arbitrary test function v, the minimizer u satisfies

$$\begin{aligned} -\mathcal{L}u = \mathcal{D}\Theta\cdot\mathcal{D}^*u &= f \qquad \text{on } \Omega,\\ u &= g \qquad \text{on } \Omega_{\mathcal{I}}. \end{aligned}$$

4.2 Discretization of the Variational Problem

Let $\Omega\subset\mathbb{R}^n$ be an open region and let $\Omega_{\mathcal{I}}$ denote the interaction domain corresponding to Ω, as defined in (4.1). Let $u,v\in L^2(\Omega\cup\Omega_{\mathcal{I}})$, $f\in L^2(\Omega)$, and $g\in L^2(\Omega_{\mathcal{I}})$.

The nonlocal bilinear form $a(\cdot,\cdot)$ is defined to be

$$a(u,v) := \frac{1}{2}\int_{\Omega\cup\Omega_{\mathcal{I}}}\int_{\Omega\cup\Omega_{\mathcal{I}}} \mathcal{D}^*(u)(\boldsymbol{x},\boldsymbol{y})\cdot\big(\Theta(\boldsymbol{x},\boldsymbol{y})\cdot\mathcal{D}^*(u)(\boldsymbol{x},\boldsymbol{y})\big)\,d\boldsymbol{y}\,d\boldsymbol{x}. \tag{4.9}$$

The nonlocal bilinear form acts as a semi-inner product on the space $L^2(\Omega\cup\Omega_{\mathcal{I}})$ and the semi-inner product induces a semi-norm $|||u||| = \sqrt{a(u,u)}$ on $L^2(\Omega\cup\Omega_{\mathcal{I}})$, which is equivalent to the $L^2(\Omega\cup\Omega_{\mathcal{I}})$ norm for functions restricted to the constrained energy space

$$L^2_c(\Omega\cup\Omega_{\mathcal{I}}) := \{u\in L^2(\Omega\cup\Omega_{\mathcal{I}}) : |||u||| < \infty \text{ and } u|_{\Omega_{\mathcal{I}}} = 0 \text{ a.e.}\}.$$

The problem we aim to solve is to find $u\in L^2_c(\Omega\cup\Omega_{\mathcal{I}})$ such that for all $v\in L^2_c(\Omega\cup\Omega_{\mathcal{I}})$,

$$a(u,v) = \int_\Omega f(\boldsymbol{x})v(\boldsymbol{x})\,d\boldsymbol{x}. \tag{4.10}$$

To show that the problem is well-posed, the bilinear form must be coercive and bounded in the Hilbert space $L^2_c(\Omega\cup\Omega_{\mathcal{I}})$, and the linear functional must be continuous with respect to the same Hilbert space; the Lax-Milgram theorem implies then that the problem (4.10) is well-posed. These theoretical properties for the forms appearing in (4.10) have been verified in [3, Lemma 4.7], which implies that the anisotropic nonlocal diffusion problem is well-posed on the space $L^2_c(\Omega\cup\Omega_{\mathcal{I}})$. The problem can be discretized by introducing a finite-dimensional subspace $V_h = \text{span}\{\phi_i\}_{i=1}^N \subset L^2_c(\Omega\cup\Omega_{\mathcal{I}})$. The corresponding discrete problem seeks $u_h = \sum_{i=1}^N c_i\phi_i \in V_h$ such that for all $v_h\in V_h$,

$$a(u_h,v_h) = \int_\Omega f(\boldsymbol{x})v_h(\boldsymbol{x})\,d\boldsymbol{x}.$$

By choosing $v_h = \phi_i$ for each $i = 1,\dots,N$, a linear system of equations for the coefficients of u_h can be created. The linear system $\mathbf{A}\boldsymbol{c} = \boldsymbol{b}$ produced by these

choices of v_h has entries given by

$$\mathbf{A}_{i,j} = a(\phi_i, \phi_j) \qquad \boldsymbol{b}_i = \int_\Omega f(\boldsymbol{x})\phi_i(\boldsymbol{x})\, d\boldsymbol{x}. \tag{4.11}$$

Our goal is to choose a basis that produces a well-conditioned, sparse stiffness matrix. In 4.5, we present a discretization using a localized basis of radial basis functions that has numerous computational benefits. We also introduce another method that makes use of Lagrange multipliers to enforce any type of volume constraint.

4.2.1 Classical Differential Operators as Limits of Nonlocal Operators

We now discuss relationships between nonlocal operators and classical differential operators. Let $u, v \in L^2(\Omega \cup \Omega_{\mathcal{I}})$ and let $\gamma_\epsilon := \alpha \cdot \Theta \cdot \alpha$ be a radial kernel with support radius ϵ. Under general conditions, the nonlocal operator (4.3) converges to a second order elliptic operator, see, e.g., [3] and the references provided. We provide an example of a nonlocal limit that justifies the interpretation of the nonlocal operator as a nonlocal analogue of the Laplacian [3]. Define $(\boldsymbol{C}_\epsilon)_{i,j} := \int_{B_\epsilon(0)} \gamma_\epsilon(\|\boldsymbol{x}\|)\boldsymbol{x}_i\boldsymbol{x}_j\, d\boldsymbol{x}$ for $i, j = 1, \ldots, n$. Recall that $\gamma_\epsilon(\|\boldsymbol{x}\|)$ is compactly supported on a ball of radius ϵ. It can be shown [3] that the compact support of γ_ϵ implies that

$$\lim_{\epsilon \to 0} \int_{\Omega \cup \Omega_{\mathcal{I}}} \int_{\Omega \cup \Omega_{\mathcal{I}}} \big(u(\boldsymbol{y}) - u(\boldsymbol{x})\big)\big(v(\boldsymbol{y}) - v(\boldsymbol{x})\big)\gamma(\|\boldsymbol{x} - \boldsymbol{y}\|)\, d\boldsymbol{y}\, d\boldsymbol{x} = \int_\Omega \nabla v(\boldsymbol{x}) \cdot \big(\boldsymbol{C} \cdot \nabla u(\boldsymbol{x})\big)\, d\boldsymbol{x}.$$

4.3 Lagrange Functions and Local Lagrange Functions

We discuss a recently investigated localized basis that may be used for interpolation and approximation using linear combinations of thin plate splines in this section. Let $X \subset \Omega$ be a set of N quasi-uniformly distributed centers and let $\varphi(r)$ denote the surface spline of order m. For each $\boldsymbol{x}_i \in X$, there exists a unique interpolant χ_i that

satisfies $\chi_i(\boldsymbol{x}_j) = \delta_{i,j}$; see Chapter 2 for theoretical properties of this basis on spheres. Identically to the case of the sphere, we refer to the basis $\{\chi_i(\boldsymbol{x})\}_{i=1}^N$ as the *Lagrange basis* and χ_i as the *Lagrange function* centered at $\boldsymbol{x}_i$. The Lagrange basis allows the interpolant to a continuous function f to be written as a linear combination of the basis elements as $I_X f = \sum_{i=1}^N f(\boldsymbol{x}_i)\chi_i(\boldsymbol{x})$. The χ_i functions are constructed by solving the $(N + n_L) \times (N + n_L)$ linear system described in Chapter 2 for each χ_i, where n_L denotes the number of basis elements in the polynomial space Π_m corresponding to the kernel φ. The assembly of the Lagrange functions is a computational issue that has limited the exploration and use of Lagrange functions. Our previous work explored the use of Lagrange functions for discretizing nonlocal diffusion problems [1].

Rather than work directly with the computationally inefficient Lagrange functions, we discuss the construction of a basis that offers nearly identical properties, but is constructed far more efficiently; we refer to these functions as *local Lagrange functions*. The discretization we later introduce for the solution of nonlocal diffusion problems in Section 4.5 uses local Lagrange functions for compact domains $\Omega \subset \mathbb{R}^n$. Let $X \subset \Omega$ be a quasi-uniformly distributed collection of scattered centers with mesh norm h and separation radius q and let $K > 0$ be a fixed constant. The method for constructing local Lagrange functions we discuss requires additional centers outside of the domain Ω to produce a a larger set of points $\Xi \supset X$. Let $\tilde{\Omega} = \{\boldsymbol{x} \in \mathbb{R}^n : d(\boldsymbol{x}, \Omega) \leq Kh|\log(h)|\}$. A set of centers Ξ can be constructed such that $\Xi \cap \Omega = X$ and Ξ has mesh norm h in $\tilde{\Omega}$. For each $\boldsymbol{x}_i \in X$, let $\Upsilon_i = \{\boldsymbol{y} \in \Xi : d(\boldsymbol{x}_i, \boldsymbol{y}) \leq Kh|\log(h)|\}$. We define the *local Lagrange function*

centered at $\boldsymbol{x}_i$ to be the function b_i, which has the form

$$b_i(\boldsymbol{x}) = \sum_{\boldsymbol{y}\in\Upsilon_i} \alpha_{\eta,i}\varphi(\|\boldsymbol{x}-\boldsymbol{y}\|) + \sum_{l=1}^{n_L} \beta_{l,i} p_l(\boldsymbol{x}) \tag{4.12}$$

and the coefficients are constructed by solving

$$\begin{pmatrix} \boldsymbol{S}_i & \boldsymbol{P} \\ \boldsymbol{P}^T & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha}^i \\ \boldsymbol{\beta}^i \end{pmatrix} = \begin{pmatrix} \boldsymbol{e}_i \\ 0 \end{pmatrix} \tag{4.13}$$

where $\boldsymbol{S}_i(\boldsymbol{y},\boldsymbol{z}) = \varphi(\|\boldsymbol{y}-\boldsymbol{z}\|)$ for $\boldsymbol{y},\boldsymbol{z}\in\Upsilon_i$, $\boldsymbol{P}(\boldsymbol{y},l) = p_l(\boldsymbol{y})$ and $\boldsymbol{e}_i(\boldsymbol{y}) = \delta(\boldsymbol{x}_i,\boldsymbol{y})$. The cardinality of Υ_i can be estimated by using the separation radius q and a volume estimate. Applying quasi-uniformity and noting that every center is separated by at least q, we estimate

$$\#|\Upsilon_i| \leq \frac{\mu\big(B(\boldsymbol{x}_i, Kh|\log(h)|)\big)}{\mu\big(B(\boldsymbol{x}_i,q)\big)} \sim \frac{K^n h^n}{Cq^n}|\log(h)|^n \leq \tilde{C}\rho^n|\log(N)|^n.$$

For quasi-uniformly distributed sets of centers, $\frac{h}{q} := \rho$ is bounded above and below by fixed constants. Therefore, constructing a local Lagrange function requires solving linear systems of size $O\big(\log(N)^n\big)$ as opposed to $O(N)$ for the full Lagrange functions. We demonstrate that the local Lagrange functions provide approximation rates analogous to known approximation rates for globally supported Lagrange functions.

Lemma 5. *Let $\frac{n}{2} < k \leq m$ and let $\Omega\cup\Omega_I \subset \mathbb{R}^n$. Let $f\in W_2^k(\Omega\cup\Omega_I)$ be a compactly supported function such that $f|_{\Omega_I} = 0$. Then, for sufficiently large K, the quasi-interpolant $\tilde{I}_X(f) = \sum_{i=1}^N f(x_i)b_i$ satisfies*

$$\|f - \tilde{I}_X f\|_{L^2(\Omega)} \leq Ch^k \|f\|_{W_2^k(\Omega)}.$$

Proof. We assume the set of centers $\Xi \subset \Omega \cup \Omega_I$ with $X := \Xi \cap \Omega$. Let χ_i be the Lagrange functions centered at $\boldsymbol{x}_i$ and b_i denote the local Lagrange function centered at $\boldsymbol{x}_i$. Then,

$$\begin{aligned}\|u - \sum_{i=1}^{N} u(\boldsymbol{x}_i) b_i\|_{L^2(\Omega\cup\Omega_I)} \leq \|u - \sum_{i=1}^{N} u(\boldsymbol{x}_i)\chi_i\|_{L^2(\Omega\cup\Omega_I)} \\ + \|\sum_{i=1}^{N} u(\boldsymbol{x}_i)(\chi_i - b_i)\|_{L^2(\Omega\cup\Omega_I)}.\end{aligned}$$

We note that $\sum_{i=1}^{N} u(\boldsymbol{x}_i)\chi_i$ is the Lagrange function interpolant to u using the set of centers in $\Xi \subset \Omega \cup \Omega_I$, and hence we may apply radial basis function error estimates on $\Omega \cup \Omega_I$ to find

$$\|u - \sum_{i=1}^{N} u(\boldsymbol{x}_i)\chi_i\|_{L^2(\Omega\cup\Omega_I)} \leq Ch^k \|u\|_{W_2^k(\Omega\cup\Omega_I)}.$$

Next, we apply Theorem 4.10 [11] to bound $\|b_\xi - \chi_\xi\|_{L^2(\Omega\cup\Omega_I)}$. Noting that $N \leq Cq^{-d}$ for quasi-uniformly distributed sets and applying the Sobolev embedding theorem to bound $\|u\|_{L^\infty(\Omega\cup\Omega_I)} \leq C\|u\|_{W_2^k(\Omega\cup\Omega_I)}$, we compute

$$\begin{aligned}\|\sum_{i=1}^{n} u(\boldsymbol{x}_i)(b_i - \chi_i)\|_{L^2(\Omega\cup\Omega_I)} &\leq Cq^{-n}\|u_i\|_{\ell^2(N)} \sup_i \|b_i - \chi_i\|_{L^2(\Omega\cup\Omega_I)} \\ &\leq q^{-2n}\|u\|_{L^\infty(\Omega\cup\Omega_I)} h^{K\nu\backslash 2-4m+2n-2\tau-1} \\ &\leq Ch^{K\nu\backslash 2-4m-2\tau-1}\|u\|_{W_2^k(\Omega\cup\Omega_I)}.\end{aligned}$$

Therefore, for sufficiently large K, the exponent on the h term is at least as large as k. Combining the two inequalities yields the result. □

We refer to the Lagrange function at $\boldsymbol{x}_i$ as the *full* or *global* Lagrange function to

contrast it with the local Lagrange function at $\boldsymbol{x}_i$, which is constructed using only points in a neighborhood of $\boldsymbol{x}_i$. The paper [7] introduces a theoretical framework for local Lagrange functions, where decay properties, quasi-interpolation convergence rates, and preconditioners were studied. The local Lagrange basis may be assembled in parallel by solving small (relative to the number of centers) linear systems. Both the parallel advantages as well as the requirement to only solve small linear systems stands in stark contrast with previous radial basis function methods that necessitate solving large, dense linear systems. We make use of recent work in [11], which has extended theoretical properties of the local Lagrange basis to compact domains in $\mathbb{R}^n$.

4.4 Local Lagrange Quadrature

We introduce a quadrature method for compactly supported functions in Ω that is essential for the implementation of the Galerkin method we introduce later for the assembly of the stiffness matrix for the nonlocal diffusion discretization in Section 4.5. Let $f \in W_2^\beta(\Omega)$ be compactly supported in Ω and let $X \subset \Omega$ be a collection of N centers. Let $\chi_i(\boldsymbol{x})$ be a globally supported Lagrange function centered at $\boldsymbol{x}_i \in X$ and let b_i be a local Lagrange function centered at $\boldsymbol{x}_i$. We define the quadrature weight at $\boldsymbol{x}_i$ to be $w_i = \int_\Omega \chi_i(\boldsymbol{x})\, d\boldsymbol{x}$ and the *Lagrange function quadrature rule* to be $Q_X(f) = \sum_{i=1}^N f(\boldsymbol{x}_i) w_i$. Similarly, we define the *local* quadrature weight at $\boldsymbol{x}_i$ to be $\hat{w}_i = \int_\Omega b_i(\boldsymbol{x})\, d\boldsymbol{x}$ and the local quadrature method $\hat{Q}_X(f) = \sum_{i=1}^N f(\boldsymbol{x}_i)\hat{w}_i$. As we expect, the quadrature error decreases as the mesh norm decreases analogous to the convergence rates observed on the sphere.

Lemma 6. *Let $f \in W_2^\beta(\Omega)$ be compactly supported for $\frac{n}{2} < \beta \leq m$. Then, for sufficiently large K,*

$$\left| \int_\Omega f(\boldsymbol{x}) - \hat{Q}_X f \right| \leq C h^\beta \|f\|_{W_2^\beta(\Omega)}.$$

Proof. The result follows by the Cauchy-Schwarz inequality along with Lemma 5.

$$\begin{aligned}\left| \int_\Omega f(\boldsymbol{x})\, d\boldsymbol{x} - \hat{Q}_X(f) \right| &= \left| \int_\Omega f(\boldsymbol{x})\, d\boldsymbol{x} - \sum_{i=1}^N f(\boldsymbol{x}_i)\hat{w}_i \right| \\ &\leq \int_\Omega |f(\boldsymbol{x}) - \sum_{i=1}^N f(\boldsymbol{x}_i) b_i(\boldsymbol{x})|\, d\boldsymbol{x} \leq \sqrt{\mu(\Omega)} \|f - \hat{I}_X f\|_{L^2(\Omega)} \\ &\leq C\sqrt{\mu(\Omega)} h^\beta \|f\|_{W_2^\beta(\Omega)}.\end{aligned}$$

□

The Lagrange function quadrature weights can be constructed without assembling all of the Lagrange functions. Let $\varphi(r)$ denote the surface spline of order m on $\mathbb{R}^n$. For a set of centers $X \subset \Omega$, the Lagrange function quadrature weights are the solution to the linear system of equations

$$\begin{pmatrix} \boldsymbol{T} & \boldsymbol{P} \\ \boldsymbol{P}^T & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{w} \\ \boldsymbol{d} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\nu} \\ \boldsymbol{\eta} \end{pmatrix} \tag{4.14}$$

where $T_{i,j} = \varphi(\|\boldsymbol{x}_i - \boldsymbol{x}_j\|)$, $P_{i,l} = p_l(\boldsymbol{x}_i)$, $\boldsymbol{\nu}_i = \int_\Omega \varphi(\|\boldsymbol{x} - \boldsymbol{x}_i\|)\, d\boldsymbol{x}$, $\boldsymbol{\eta}_l = \int_\Omega p_l(\boldsymbol{x})\, d\boldsymbol{x}$ and w_i is the quadrature weight at $\boldsymbol{x}_i$. This requires solving a dense, symmetric linear system, where the size of the system grows as $\mathcal{O}(N)$. In the case of the sphere, the system (4.14) can be preconditioned by using the local Lagrange functions as described in [8]. We present an alternative method of producing quadrature weights by using the local Lagrange functions directly. These weights may be assembled by taking advantage of GPU to compute many of the necessary integrals quickly and in parallel. This enables rapid assembly of the quadrature weights.

The local quadrature weights are constructed by computing the integrals of the

translates $\varphi(\|\boldsymbol{x} - \boldsymbol{x}_i\|)$. Recall that by equation (4.12)

$$b_i(\boldsymbol{x}) = \sum_{\boldsymbol{y} \in \Upsilon_i} \alpha_{\boldsymbol{y},\boldsymbol{x}_i} \varphi(\|\boldsymbol{x} - \boldsymbol{y}\|) + \sum_{l=1}^{n_L} \beta_{l,i} p_l(\boldsymbol{x})$$

and consequently,

$$\hat{w}_i = \sum_{\boldsymbol{y} \in \Upsilon_i} \alpha_{\boldsymbol{y},\boldsymbol{x}_i} \int_\Omega \varphi(\|\boldsymbol{x} - \boldsymbol{y}\|)\, d\boldsymbol{x} + \sum_{l=1}^{n_L} \beta_{l,i} \int_\Omega p_l(\boldsymbol{x})\, d\boldsymbol{x}. \tag{4.15}$$

The construction of the local quadrature weights does not require the solution of a large linear system, in contrast to the quadrature method introduced in [1]. However, (4.15) does require that the local Lagrange function coefficients are computed before the weights can be constructed. After constructing the local Lagrange functions, the weights can be assembled in parallel.

4.5 Galerkin Radial Basis Function Method

We propose a new method for the discretization of the nonlocal diffusion problem by using local Lagrange functions to be a local Lagrange Galerkin method. The local Lagrange quadrature method of Section 4.4 enables rapid assembly of the stiffness matrix entries. Applying the local Lagrange quadrature produces a sparse stiffness matrix, where the sparsity pattern of the stiffness matrix is determined by the horizon ϵ of the kernel. The quadrature formula for the entries only requires a pointwise evaluation of the kernel and multiplication by the quadrature weights. This is in contrast to a piecewise polynomial finite element method for a $\Omega \subset \mathbb{R}^n$ which requires the evaluation of $2n$-iterated integrals over partial element volumes. The resulting quadrature problem is a nontrivial computational challenge. Such issues relating to evaluating numerical integrals over partial element volumes do not arise in the Galerkin radial basis function method.

We previously explored a Galerkin radial basis function method using full Lagrange functions and an associated Lagrange function quadrature rule [1]. The major difference between the method of [1] and the method we have described is in the discretization space and the assembly of the quadrature weights. Assembling the quadrature weights using the full Lagrange functions necessitates the solution of a single large, dense linear system where the number of rows grows as $\mathcal{O}(N)$ where N is the number of basis functions in the discretization. Evaluating the solution requires the solution of yet another dense linear system of size $\mathcal{O}(N)$. The local Lagrange function method we have discussed here requires solving small linear systems of size $\mathcal{O}\big(\log(N)^n\big)$ for centers in $\mathbb{R}^n$ in contrast to large linear systems in the full Lagrange function method.

4.5.1 Local Lagrange Discretization

Let Ω be an open region in $\mathbb{R}^n$ and $\Omega_{\mathcal{I}}$ be the corresponding interaction domain. Let $X \subset \Omega \cup \Omega_{\mathcal{I}}$ be a set of quasi-uniformly scattered centers with mesh norm h. An extended set of centers $X' \supset X$ can be constructed such that $X' \cap \big(\Omega \cup \Omega_{\mathcal{I}}\big) = X$ and $h(X') = h$, and $\sup_{\boldsymbol{x}' \in X', \boldsymbol{x}_i \in X} \|\boldsymbol{x}' - \boldsymbol{x}_i\| \leq Kh|\log(h)|$ for a fixed user chosen integer $K > 0$. For each $\boldsymbol{x}_i \in X$, we construct b_i, the local Lagrange function centered at $\boldsymbol{x}_i$. Let $V_h = \text{span}\{b_i : \boldsymbol{x}_i \in \Omega\}$. The space $V_h \not\subset L^2_c(\Omega \cup \Omega_{\mathcal{I}})$ because the local Lagrange functions in V_h are necessarily nonzero in $\Omega_{\mathcal{I}}$; however, they are provably small in Ω_I. We now seek to augment the space V_h to construct a discretization space that is conforming. We replace b_i with $\tilde{b}_i = b_i \mathbb{1}_\Omega$, where $\mathbb{1}_\Omega$ is an indicator function for Ω. Since the space $\tilde{V}_h$ is conforming with respect to the bilinear form a from Equation (4.9), there exists $u_h \in \tilde{V}_h$ such that $a(u_h, v_h) = \int_\Omega f(\boldsymbol{x}) v_h(\boldsymbol{x})\, d\boldsymbol{x}$ for all $v_h \in \tilde{V}_h$. We establish an error estimate that demonstrates how the smoothness of the solution affects the convergence rate. The interpolation error estimates for

the local Lagrange functions for interpolation match the approximation error for the Galerkin solution.

Proposition 11. *Let $u \in W_2^k(\Omega)$ for $k > \frac{n}{2}$ be the solution to the nonlocal problem (4.10). Let u_h be the discrete solution from the restricted local Lagrange method. Then, for sufficiently small h and for sufficiently large K,*

$$\|u - u_h\|_{L^2(\Omega\cup\Omega_{\mathcal{I}})} \leq Ch^k \|u\|_{W_2^k(\Omega\cup\Omega_{\mathcal{I}})} \tag{4.16}$$

Proof. By the Lax-Milgram theorem, the discrete solution u_h satisfies $\|u-u_h\|_{L^2(\Omega\cup\Omega_{\mathcal{I}})} \leq C \inf_{v_h\in U_h} \|u - v_h\|_{L^2(\Omega\cup\Omega_{\mathcal{I}})}$. By setting $v_h = \sum_{i=1}^N u(\boldsymbol{x}_i)\tilde{b}_i$, we have

$$\begin{aligned}\|u - u_h\|_{L^2(\Omega\cup\Omega_{\mathcal{I}})} &\leq C\|u - \sum_{i=1}^N u(\boldsymbol{x}_i)\tilde{b}_i\|_{L^2(\Omega\cup\Omega_{\mathcal{I}})} = C\|u - \sum_{i=1}^N u(\boldsymbol{x}_i)b_i\|_{L^2(\Omega)} \\ &\leq C\|u - \sum_{i=1}^N u(\boldsymbol{x}_i)b_i\|_{L^2(\Omega\cup\Omega_{\mathcal{I}})}.\end{aligned}$$

Since $u \in L_c^2(\Omega \cup \Omega_{\mathcal{I}})$ it is compactly supported and hence we may apply Lemma 5 to compute

$$\|u - u_h\|_{L^2(\Omega\cup\Omega_{\mathcal{I}})} \leq Ch^k \|u\|_{W_2^k(\Omega\cup\Omega_{\mathcal{I}})}.$$

□

Let $\boldsymbol{A}_{i,j} := a(\tilde{b}_i, \tilde{b}_j)$ denote the stiffness matrix generated by applying the bilinear form to the local Lagrange functions in $\tilde{V}_h$, we demonstrate that the condition number is bounded independent of the mesh norm h or the separation radius q.

Lemma 7. *The condition number of the discrete stiffness matrix $\boldsymbol{A}$ is bounded above by a constant independent of h and q.*

Proof. Let $\mathbf{A}$ denote the $N \times N$ symmetric stiffness matrix and let $\mathbf{c} \in \mathbb{R}^N$. Then,

$$\langle \mathbf{Ac}, \mathbf{c} \rangle = \sum_{i=1}^{N} \Big(\sum_{j=1}^{N} A_{i,j} c_i) c_j \Big) = a\Big(\sum_{i=1}^{N} c_i \tilde{b}_i, \sum_{j=1}^{N} c_j \tilde{b}_j \Big).$$

By the coercivity of the bilinear form and since $\sum_{i=1}^{N} c_i \tilde{b}_i \in L_c^2(\Omega \cup \Omega_{\mathcal{I}})$, there exists λ_1, λ_2 such that

$$\lambda_1 \| \sum_{i=1}^{N} c_i \tilde{b}_i \|_{L^2(\Omega \cup \Omega_{\mathcal{I}})} \leq a\Big(\sum_{i=1}^{N} c_i \tilde{b}_i, \sum_{j=1}^{N} c_j \tilde{b}_j \Big) \leq \lambda_2 \| \sum_{i=1}^{N} c_j \tilde{b}_j \|_{L^2(\Omega \cup \Omega_{\mathcal{I}})}.$$

It follows that since $\tilde{b}_i = 0$ on $\Omega_{\mathcal{I}}$ and $\tilde{b}_i|_{\Omega} = b_i$,

$$\lambda_1 \| \sum_{i=1}^{N} c_i \tilde{b}_i \|_{L^2(\Omega \cup \Omega_{\mathcal{I}})} = \lambda_1 \| \sum_{i=1}^{N} c_i b_i \|_{L^2(\Omega)}.$$

By [11, Proposition 5.3] and [11, Theorem 4.12], there exists C_Ω and $C_{\Omega \cup \Omega_{\mathcal{I}}}$ independent of h and q such that

$$C_\Omega q^n \|\mathbf{c}\|_{\ell^2(N)} \leq \| \sum_{i=1}^{N} c_i b_i \|_{L^2(\Omega)} \qquad \| \sum_{i=1}^{N} c_i b_i \|_{L^2(\Omega \cup \Omega_{\mathcal{I}})} \leq C_{\Omega \cup \Omega_{\mathcal{I}}} q^n \|\mathbf{c}\|_{\ell^2(N)}.$$

Then, we bound

$$\operatorname{cond}(\mathbf{A}) \leq \frac{\lambda_{\max}(\boldsymbol{A})}{\lambda_{\min}(\boldsymbol{A})} \leq \frac{C_{\Omega \cup \Omega_{\mathcal{I}}} \lambda_2}{C_\Omega \lambda_1}.$$

□

4.5.1.1 Assembling the stiffness matrix by quadrature

Assembling the stiffness matrix requires evaluation of $2n$ iterated integrals for problems on $\mathbb{R}^n$. We introduce a practical method to assemble the elements of the stiffness matrix by taking advantage of the local Lagrange quadrature. We form

the discrete stiffness matrix by evaluating $a(\tilde{b}_i, \tilde{b}_j)$ for each $\boldsymbol{x}_i, \boldsymbol{x}_j \in X \cap \Omega$. The integrals are evaluated by applying the local Lagrange quadrature rule introduced in Section 4.4. The stiffness matrix $\mathbf{A}_{i,j} = a(\tilde{b}_i, \tilde{b}_j)$ is approximated by

$$\mathbf{A}_{i,j} \approx 2\delta_{i,j}\hat{w}_i \int_{\Omega \cup \Omega_{\mathcal{I}}} \gamma(\boldsymbol{x}, \boldsymbol{x}_i)\, d\boldsymbol{x} - 2\hat{w}_i\hat{w}_j\gamma(\boldsymbol{x}_i, \boldsymbol{x}_j). \tag{4.17}$$

The integral involving $\gamma(\boldsymbol{x}, \boldsymbol{x}_i)$ may be computed analytically for some kernels or by any form of quadrature. We compute the values for the source term $\boldsymbol{b}_i$ from (4.30) by applying the Lagrange function quadrature rule to approximate

$$\boldsymbol{b}_i \approx f(\boldsymbol{x}_i)\hat{w}_i \mathbb{1}_{\Omega}(\boldsymbol{x}_i). \tag{4.18}$$

A crucial detail that may be overlooked above is that applying the quadrature rule produces a *sparse* stiffness matrix. The entries $\mathbf{A}_{i,j}$ are zero for centers such that $\|\boldsymbol{x}_i - \boldsymbol{x}_j\| \geq \epsilon$ due to the compact support of γ. The support horizon ϵ of the kernel γ and the mesh norm of the centers h determines the number of nonzero entries per row. If the local Lagrange quadrature method is not used, the stiffness matrix is dense due to the nonzero values the local Lagrange functions assume throughout Ω. We can prove that the density of nonzero elements in the stiffness matrix is bounded independent of h, q; if this were not the case, then for small mesh norm values h, the sparsity could deteriorate for smaller mesh norm values. If this occurred, the problem could degenerate into a dense system; this is fortunately not the case.

Lemma 8. *Let $\{X\}_{h,q}$ be a collection of quasi-uniformly distributed centers in $\mathbb{R}^n$. Then, the ratio of the number of nonzero entries per row to the total number of columns is bounded independent of h, q.*

Proof. Fix $X := X_{h,q}$ and fix $\boldsymbol{x}_i \in X$. Recall that $\|\boldsymbol{x}_i - \boldsymbol{x}_j\| \geq \epsilon$, $\mathbf{A}_{i,j} = 0$ by

(4.17). Let $N_i = \{\boldsymbol{x}_j : \|\boldsymbol{x}_j - \boldsymbol{x}_i\| \leq \epsilon\}$. Let C_n denote the constant so that a ball of radius r has volume $C_n r^n$. The number of nonzero entries on row i is the same as the cardinality of N_i, which we compute by estimating the number of centers in N_i. We bound the cardinality of N_i, denoted $\#N_i$, above by noting that every center is separated by at least q, so

$$C_n\big(\#N_i\big)q^n = \cup_{\boldsymbol{x}_j \in N_i} \mu(B(\boldsymbol{x}_i, q)) \geq \mu\big(B_\epsilon(\boldsymbol{x}_i)\big) = C_n \epsilon^n$$

which implies $\#N_i \leq \epsilon^n q^{-n}$. The density per row is computed by $\frac{\#N_i}{N} \leq \frac{\epsilon^n q^{-n}}{N}$. We bound N by noting that we may cover Ω with balls of radius h by $\Omega \subset \cup_{\boldsymbol{x}_j \in X} B(\boldsymbol{x}_j, h)$. Consequently, $\mu(\Omega) \leq N C_n h^n$, which implies $N \geq \frac{\mu(\Omega)}{C_n h^n}$. Therefore,

$$\frac{N_i}{N} \leq \frac{\epsilon^n q^{-n}}{\mu(\Omega) C_n^{-n} h^{-n}} = \frac{C_n \epsilon^n}{\mu(\Omega)} \frac{h^n}{q^n}.$$

The result follows by recalling that the quasi-uniformity assumption bounds the mesh ratio $\frac{h}{q}$. □

Let $\tilde{u}_h$ denote the solution to the discretized linear system assembled by quadrature from (4.17) and (4.18). Let u_h denote the solution to the the problem $a(u_h, v_h) = \int_\Omega f(\boldsymbol{x}) v_h(\boldsymbol{x})\, d\boldsymbol{x})$ as described in Section 4.5.1. We desire an estimate that predicts the convergence rate of $\tilde{u}_h$ to u in terms of h, as in Proposition 11. However, this requires a thorough analysis of the affect of quadrature on the solution to the resulting linear system of equations. By applying the triangle inequality and Proposition 11, we may estimate

$$\|u - \tilde{u}_h\|_{L^2} \leq \|u - u_h\| + \|u_h - \tilde{u}_h\|_{L^2} \leq C h^k \|u\|_{W_2^k} + \|u_h - \tilde{u}_h\|_{L^2}.$$

Both u_h and $\tilde{u}_h$ are linear combinations of local Lagrange functions with coefficients $\{\alpha_i\}_{i=1}^N$ and $\{\tilde{\alpha}_i\}_{i=1}^N$ respectively. In the numerical experiments we present in Section 4.6, we only produce the coefficients $\{\tilde{\alpha}_i\}_{i=1}^N$ since we apply quadrature to assemble the linear system of equations. The error between u_h and $\tilde{u}_h$ may be quantified by

$$\|u_h - \tilde{u}_h\|_{L^2} = \|\sum_{i=1}^{N}(\alpha_i - \tilde{\alpha}_i)b_i\|_{L^2} \leq Cq^n\|\alpha_i - \tilde{\alpha}_i\|_{\ell^2(N)}.$$

We do not currently have an estimate to bound $\|\alpha_i - \tilde{\alpha}_i\|_{\ell^2(N)}$. Despite the lack of theoretical justification, we demonstrate in Section 4.6 that the discrete solution produced by solving the linear system assembled by using quadrature follows an estimate of the form in Proposition 11. These results suggest $\|u - \tilde{u}_h\|_{L^2} \sim \|u - u_h\|_{L^2} \leq Ch^k\|u\|_{W_2^k}$. One possible method to solve this issue is to introduce a larger set of centers Y to use as quadrature nodes. The disadvantage to this method is that the quadrature evaluations will require more than a single point evaluation which removes one of the advantages of the one point quadrature method we use.

4.6 Numerical Results

We present numerical results for experiments using the local Lagrange function method we discussed in Section 4.5. The topics we address are local Lagrange function construction, L^2 error computations, and condition number computations. We validate the theoretical prediction for L^2 convergence and condition numbers with observed results from numerical experiments. We study two dimensional problems of the form (4.10) with a radial kernel Φ and two different anisotropy functions κ; see Section 4.6.1 and Section 4.6.2. As required by the method, we only solve problems with zero Dirichlet volume constraints. The testing domains are $\Omega \cup \Omega_{\mathcal{I}}$ where

$\Omega = (0,1) \times (0,1)$ and $\Omega_{\mathcal{I}} = [-\frac{1}{4}, \frac{5}{4}] \times [-\frac{1}{4}, \frac{5}{4}] \backslash \Omega$. The computational results we present here are computed in MATLAB and the condition numbers of the sparse stiffness matrices are approximated by the *condest* function, since the matrices are in sparse format. If the matrices are in full format, the *cond* or *rcond* commands may be used to compute the condition number or reciprocal condition number, respectively. The sparse linear system is solved with either directly or with an iterative method. For a direct solution, MATLAB's backslash operator is used, or by conjugate gradient with a specified tolerance of 10^{-9} for an iterative method. The number of iterations required for convergence to the specified tolerance with conjugate gradient was independent of the mesh norm h. This is not particularly surprising since the matrices are well-conditioned with condition numbers independent of the mesh norm h. We briefly remark here that if we instead force $\epsilon \to 0$, the condition number decays with ϵ.

The local Lagrange functions are computed using the surface spline $\varphi(r) = r^2 \log(r)$. To ensure the local Lagrange functions provide sufficient quasi-interpolation, we choose more points than may be necessary; each local Lagrange function is constructed using approximately $11 \log(N)^2$ nearest neighbor centers, where N is the total number of centers in $\Omega \cup \Omega_{\mathcal{I}}$. The stiffness matrix for the nonlocal problem only requires Lagrange functions centered in Ω; however, the surface splines centered in $\Omega_{\mathcal{I}}$ are required for the construction of the local Lagrange functions. See Section 4.3 for a discussion on the construction method and the need for points outside of Ω.

We choose a kernel $\gamma(\boldsymbol{x}, \boldsymbol{y}) = \big(\kappa(\boldsymbol{x}) + \kappa(\boldsymbol{y})\big)\Phi(\|\boldsymbol{x} - \boldsymbol{y}\|)$ with fixed horizon ϵ; the function κ provides spatial variation which makes the problem anisotropic. We choose a solution $u \in L_c^2(\Omega \cup \Omega_{\mathcal{I}})$ is chosen for each numerical experiment and manufacture the source function f by computing $\mathcal{L}u(\boldsymbol{x}_i) = f(\boldsymbol{x}_i)$ for each center

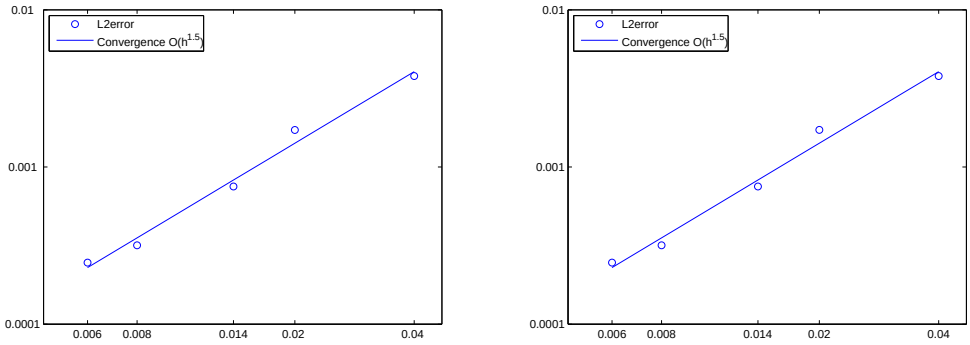
$\boldsymbol{x}_i$. The values of $f(\boldsymbol{x}_i)$ are computed by using large numbers of tensor products of Gauss-Legendre nodes to approximate the integral in Equation (4.3); the large number of quadrature nodes used is to guarantee that the value of f at each center is as accurate as possible; errors in this step are independent of the discretization method and merely errors in the manufacturing step.

The experiments are aimed at evaluating the L^2 convergence of the discrete solution by constructing sets of uniformly spaced centers and sets of scattered centers with various mesh norms. Uniformly spaced collections of centers X_h are constructed using uniformly spaced centers with spacing $h = .04, .02, .014, .008$, and $.006$. To study the effect of non-uniform, irregular sets of centers, collections of scattered centers $\tilde{X}_h$ are constructed by modifying centers in X_h by a random perturbation of magnitude at most $\frac{2h}{15}$. For all of the X_h and $\tilde{X}_h$, Local Lagrange functions are constructed to build the discretization space. The convergence of the discrete solution u_h to the solution u is evaluated by plotting the log of the L^2 norm of the error $\|u_h - u\|_{L^2(\Omega \cup \Omega_{\mathcal{I}})}$ against the log of the mesh norm h and viewing a best fit line of the data. By our theoretical predictions in Proposition 11, We expect for $u \in W_2^k(\Omega \cup \Omega_{\mathcal{I}})$ that $\|u - u_h\|_{L^2(\Omega)} \leq Ch^k \|u\|_{W_2^k(\Omega)}$.

4.6.1 Linear Anisotropy Experiment

We choose solution a sinusoidal, continuous function u and a kernel γ with a linear anisotropy function κ and radial function Φ given by

$$\begin{cases} u(\boldsymbol{x}) = \sin(2\pi x_1)\sin(2\pi x_2)\mathbb{1}_{\Omega}(\boldsymbol{x}) \\ \kappa(\boldsymbol{x}) = 1 + x_1 + x_2 \\ \Phi(\|\boldsymbol{x} - \boldsymbol{y}\|) = \exp\left(- \frac{1}{1 - \frac{1}{\epsilon^2}\|\boldsymbol{x} - \boldsymbol{y}\|^2} \right) \end{cases} \tag{4.19}$$



(a) Uniformly Spaced Centers (b) Scattered Centers

Figure 4.1: The log of h versus the log of the L^2 error for the linear anisotropic experiment with functions given by (4.19) is displayed.

and we discretize (4.10) with local Lagrange functions.

The observed L^2 convergence rates with respect to the mesh norm h for the uniformly spaced and scattered centers experiments are provided in Figure 4.1. Table 4.1 displays the condition numbers of the stiffness matrices for each experiment. As we expect from our theoretical predictions, the observed condition numbers of the stiffness matrices are invariant under decreases in the mesh norm h. This numerically validates the claims that the method produces well-conditioned matrices, even for large sets of centers.

4.6.2 Exponential Anisotropy Experiment

We choose solution a u and a kernel γ with an exponential anisotropy function κ and radial function Φ given by

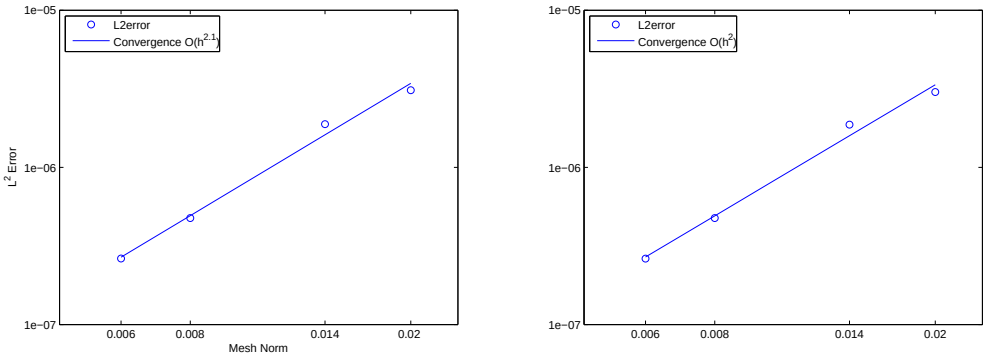
$$\begin{cases} u(\boldsymbol{x}) = \Big(x_1(x_1-1)\Big)^{\frac{3}{2}}\Big(x_2(x_2-1)\Big)^{\frac{3}{2}}\mathbb{1}_{\Omega}(\boldsymbol{x}) \\ \kappa_2(\boldsymbol{x}) = \exp(x_1+x_2) \\ \Phi(\|\boldsymbol{x}-\boldsymbol{y}\|) = \exp\left(-\frac{1}{1-\frac{1}{\epsilon^2}\|\boldsymbol{x}-\boldsymbol{y}\|^2}\right) \end{cases} \tag{4.20}$$

and we discretize (4.10) with local Lagrange functions.

Figure 4.2 provides the L^2 convergence plots for the experiments involving u_2 and κ_2. The L^2 error rate matches the expected convergence rate predicted by Proposition 11. We observe a convergence rate of $\mathcal{O}(h^2)$ in both the uniformly spaced centers and the scattered centers experiments, which matches our expectation. Table 4.1 displays the condition numbers for the discrete stiffness matrices of various values for h; once again, the condition numbers of the discrete stiffness matrices do not increase as the mesh norm decreases, which matches the prediction in Lemma 7. Consequently, as we vary the anisotropy functions, the stiffness matrices still adhere to a fixed condition number (although the number varies depending on the anisotropy function, it is independent of the mesh norm).

4.6.2.1 Vanishing Nonlocality

We present numerical results as part of an investigation of the effects of shrinking the horizon ϵ, which is the radius of support of the kernel γ. The solution of the nonlocal problem converges to the solution of (4.4) as ϵ decreases under appropriate conditions on the kernel and anisotropy function. We are focused on anisotropic



(a) Uniformly Spaced Centers (b) Scattered Centers

Figure 4.2: The log of h versus the log of the L^2 error for the exponential anisotropy experiment with functions given by (4.20) is displayed.

Table 4.1: The mesh norm h, number of rows n of the stiffness matrix, and the estimated condition number for the stiffness matrix with the linear anisotropy (4.19) and the exponential anisotropy (4.20). The condition numbers of the stiffness matrices does not increase as h decreases.

		Approximate Condition Number	
h	n	Linear	Exponential
2.83e-2	625	58	89
1.41e-2	2500	59	90
9.9e-3	5041	59	90
5.7e-3	15625	60	92
4.2e-3	27889	60	92

kernels of the form

$$\gamma_\epsilon(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{\epsilon^3}\big(\kappa(\boldsymbol{x})+\kappa(\boldsymbol{y})\big)\Phi(\frac{1}{\epsilon}\|\boldsymbol{x}-\boldsymbol{y}\|), \tag{4.21}$$

where $\Phi(\frac{1}{\epsilon}\|\boldsymbol{x}\|)$ is a compactly supported radial function with support radius ϵ; the ϵ value determines the horizon of the nonlocal operator. We study the idea of approximating the solution to an anisotropic differential equation by solving an anisotropic nonlocal problem with sufficiently small horizon ϵ. Ideally, as ϵ decreases to zero, the the discrete solution to the anisotropic nonlocal problem should be near the solution to the differential equation. The numerical experiments we present demonstrate that the discrete solution to the anisotropic nonlocal problem indeed converges to the solution of the anisotropic differential equation.

To analytically study the relationship between nonlocal operators and differential operators, we employ a Taylor series expansion argument. We relate the differential operator D to a nonlocal operator that approximates D in the small horizon limit. We assume that $\kappa, u : \mathbb{R} \to \mathbb{R}$ are smooth functions for the analysis we employ. Fix $x \in \Omega$ and apply a Taylor series expansion in a ball $B_\epsilon(x)$ to obtain for some $\zeta, \eta \in B_\epsilon(x)$

$$\begin{aligned} u(y)-u(x) &= u'(x)(y-x)+\frac{1}{2}u''(x)(y-x)^2+\frac{1}{6}u''(\zeta)(y-x)^3 \\ \kappa(x)+\kappa(y) &= 2\kappa(x)+\kappa'(x)(y-x)+\kappa''(x)(y-x)^2+\frac{1}{6}\kappa''(\eta)(y-x)^3. \end{aligned}$$

Let $\mathcal{L}_\epsilon$ denote the nonlocal operator with kernel γ_ϵ. Then, for smooth u, it follows

that

$$\mathcal{L}_\epsilon u(x) = \frac{1}{\epsilon^3}\big(2u'(x)\kappa(x)\int_{-\epsilon}^{\epsilon} z\Phi\left(\frac{1}{\epsilon}|z|\right)dz + u'(x)\kappa'(x)\int_{-\epsilon}^{\epsilon} z\Phi\left(\frac{1}{\epsilon}|z|\right)dz\big)$$
$$+ \frac{1}{\epsilon^3}\big(2u''(x)\kappa(x)\int_{-\epsilon}^{\epsilon} z^2\Phi\left(\frac{1}{\epsilon}|z|\right)dz + u''(x)\kappa'(x)\int_{-\epsilon}^{\epsilon} z^3\Phi\left(\frac{1}{\epsilon}|z|\right)dz + ...\big)$$

where we have truncated the expression to exclude any of the $(y-x)^3$ terms. These terms will be of an order that is vanishingly small relative to other terms later, so we exclude them now. The $z\Phi(\frac{1}{\epsilon}|z|)$ integrals vanish due to the fact that $z\Phi(\frac{1}{\epsilon}|z|)$ is an odd function. We also exclude any integrals involving z^3 since

$$\frac{1}{\epsilon^3}\int_{-\epsilon}^{\epsilon} z^3\Phi\left(\frac{1}{\epsilon}|z|\right) \leq \frac{1}{2}\epsilon\|\Phi\|_{L^\infty(\Omega)},$$

which is $O(\epsilon)$. After eliminating these terms, we find

$$\mathcal{L}_\epsilon u(x) \approx 2\big(u'(x)k'(x) + u''(x)k(x)\big)\int_{-\epsilon}^{\epsilon} z^2\Phi\left(\frac{|z|}{\epsilon}\right)dz$$
$$= 2\big(u'(x)k'(x) + u''(x)k(x)\big)\int_{-1}^{1}\tau^2\Phi(|\tau|)d\tau.$$

Consequently, as ϵ decreases to zero, it follows that

$$\mathcal{L}_\epsilon u(x) \to \rho\big(u'(x)k'(x) + u''(x)k(x)\big)$$
$$\rho := 2\int_{-1}^{1}\tau^2\Phi(|\tau|)\,d\tau$$

We numerically study a discretization method by using Lagrange functions to form an approximation space to solve the problem $\mathcal{L}_\epsilon u_\epsilon = f$ for anisotropic nonlocal operators. Let u denote the solution to $Du = f$ for the differential operator D and let h be the mesh norm corresponding to the set of centers X used to construct the

Lagrange functions. We solve $\mathcal{L}_\epsilon u_\epsilon = f$ by discretizing the problem with Lagrange functions to construct an approximate solution $u_{\epsilon,h}$, where the ϵ is varied to a sufficiently small quantity to find sufficient accuracy. Our experiments provide evidence that as $\epsilon \to 0$, $\|u - u_{\epsilon,h}\|_{L^2(\Omega \cup \Omega_{\mathcal{I}})} \sim O(\epsilon^2)$.

For our experiments, we consider the radial kernel

$$\Phi\left(\frac{1}{\epsilon}\|x\|\right) = \left(1 - \frac{1}{\epsilon^2}\|x\|^2\right)\mathbb{1}_{\|x\|<\epsilon}(x)$$

for the nonlocal operator and we consider two separate anisotropy functions $\kappa(x, y)$. A linear anisotropic function of the form $\kappa_1(x, y) = 1 + x + y$ is used for the first experiment and the second experiment uses an exponential anisotropy function $\kappa_2(x, y) = \exp(x + y)$. The kernel γ_ϵ is defined as in (4.21) for each of the different choices of κ. For these experiments, we are not interested in exploring mesh norm convergence; the mesh norm $h = .000075$ is fixed for the experiments and we instead consider a range of ϵ values $.075, .0625, .05, .04$, and $.035$. With these mesh norm values, we construct a grid of uniformly spaced centers on $[0, 1]$ and we discretize the problem $\mathcal{L}_\epsilon u_\epsilon = f$ with Lagrange functions constructed using the uniformly spaced centers. The discrete solution $u_{\epsilon,h}$ is computed as described in Section 4.5.

We choose

$$u(x) = \big(1 - \cos(2\pi x)\big)\mathbb{1}_{[0,1]}(x) \tag{4.22}$$

and we analytically compute $Du = f$, where D is the differential operator that $\mathcal{L}_\epsilon$

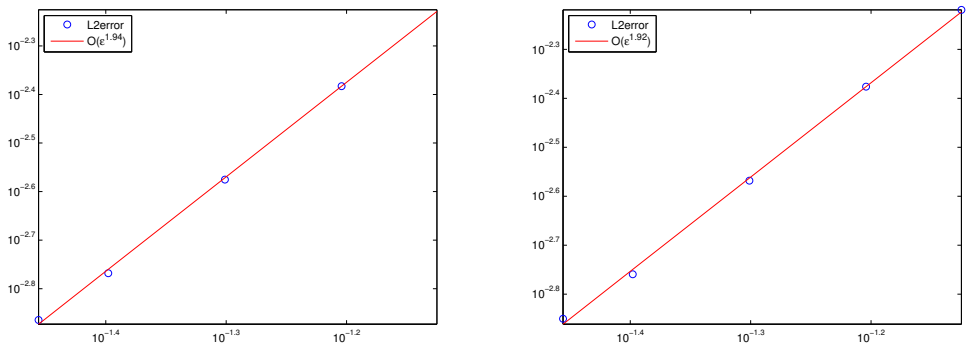
converges to. We compute,

$$f(x) = \begin{cases} -2\pi\big(\sin(2\pi x) + 2\pi(1+x)\cos(2\pi x)\big) & \text{for } \kappa_1 \\ -\exp(x)\big(2\pi\sin(2\pi x) + 4\pi^2\cos(2\pi x)\big) & \text{for } \kappa_2 \end{cases}$$

In contrast to the experiments in Section 4.6.1 and Section 4.6.2, we do not vary the source function f or the mesh norm h. Rather, we vary ϵ to analyze how the discrete solution to the nonlocal problem relates to the solution to the differential equation as ϵ shrinks. We manufacture f by applying the differential operator D to the chosen solution u to construct $Du = f$, where u is the fixed function (4.22). Figure 4.3 provides evidence that for both κ_1 and κ_2, the L^2 error $\|u - u_{\epsilon,h}\|_{L^2[0,1]}$ converges at about $O(\epsilon^2)$. This observation matches the error we expect from the Taylor series analysis we used previously. As a consequence of these experiments, the results suggest it is possible to approximate the solution to an anisotropic differential equation by discretizing and solving an anisotropic nonlocal volume constrained equation. Whether this is computationally advantageous or not is another question entirely, but it is nonetheless an interesting experiment. Indeed, if further work demonstrates that the nonlocal problem is easier to discretize and solve, then the methods discussed here provide a new avenue of investigation for solution methods for anisotropic differential equations.

4.6.3 Quadrature Experiment

We present a numerical experiment for the proposed local Lagrange quadrature method introduced in Section 4.4. The tests are computed on the set $\Omega \cup \Omega_{\mathcal{I}}$ where $\Omega = (0,1) \times (0,1)$ and $\Omega_{\mathcal{I}} = [-\frac{1}{4}, \frac{5}{4}] \times [-\frac{1}{4}, \frac{5}{4}] \backslash \Omega$. Let $X \subset \Omega \cup \Omega_{\mathcal{I}}$ be a collection



(a) $\kappa(x, y) = \exp(x + y)$ (b) $\kappa(x, y) = 1 + x + y$

Figure 4.3: The log of ϵ vs. the log of the L^2 error of the discrete solution $u_{\epsilon,h}$ is plotted. As ϵ goes to zero, we observe ϵ^2 convergence.

of scattered centers and let $\hat{w}_i$ denote the quadrature weight centered at the point $\boldsymbol{x}_i \in X$. We choose a function $u : \Omega \cup \Omega_{\mathcal{I}} \to \mathbb{R}$ and we investigate how the quadrature error, given by

$$\left| \int_{\Omega \cup \Omega_{\mathcal{I}}} u(\boldsymbol{x})\, d\boldsymbol{x} - \sum_{\boldsymbol{x}_i \in X} u(\boldsymbol{x}_i)\hat{w}_i \right|, \tag{4.23}$$

varies as we decrease the mesh norm. We expect the quadrature error to decrease as $O(h^k)$ by Lemma 6 depending on the smoothness of the function u. We choose the function

$$u(\boldsymbol{x}) = \big(x_1(1 - x_1)\big)^2 \big(x_2(1 - x_2)\big)^2 \mathbb{1}_{\Omega}(\boldsymbol{x}) \tag{4.24}$$

and we consider various sets of centers with mesh norms $h = .04, .02, .014, .008$, and $.006$. Figure 4.4 shows the results for an experiment using the polynomial function (4.24), which exhibits a convergence rate of $h^{2.5}$.

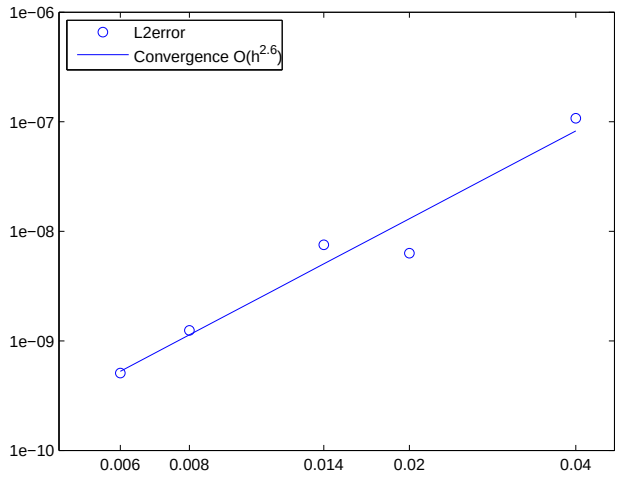


Figure 4.4: The log of the quadrature error (4.23) versus the log of the mesh norm h is displayed for the function (4.24).

4.7 Error Analysis

The purpose of this section is to provide two quadrature error estimates relevant to the quadrature method used for the nonlocal diffusion problem.

4.7.1 Two Quadrature Error Estimates

Let $f \in W_2^k(\Omega \cup \Omega_{\mathcal{I}})$ be a given function and let X be a collection of centers in $\Omega \cup \Omega_{\mathcal{I}}$. Let $\{\chi_i\}$ be a set of full Lagrange functions where χ_i is centered at x_i (i.e., $\chi_i(x_j) = \delta_{i,j}$). Let $\{w_i\}$ be the corresponding set of quadrature weights. For the conforming Local Lagrange method for nonlocal diffusion, we must compute integrals, which we approximate using the Lagrange function quadrature method:

$$\int_\Omega f(x)\chi_i(x)\,dx \approx \sum_{x_j \in X} f(x_j)\underbrace{\chi_i(x_j)}_{\delta_{i,j}} w_j = f(x_i)w_i.$$

Analysis is required in order to quantify the quadrature error

$$E_i := \left| \int_\Omega f(x)\chi_i(x)\,dx - f(x_i)w_i \right|.$$

Lemma 9. *Let $X \subset \Omega \subset \mathbb{R}^n$ have mesh norm h and separation radius q. Let $f \in W_2^k(\Omega)$ and let $\chi_i \in W_2^m(\Omega)$ with $\frac{n}{2} < k \leq m$. Then,*

$$E_i \leq Cq^{\frac{n}{2}}\|f\|_{W_2^k(\Omega)}$$

Proof. This result follows by applying by applying the interpolation error estimates and a Bernstein estimate. We know for a function $g \in W_2^k$ the quadrature error is $O(h^k\|g\|_{W_2^k(\Omega)})$ by the interpolation error estimates. Therefore, the product $f\chi_i \in W_2^k(\Omega)$ and we have

$$E_i \leq Ch^k\|f\chi_i\|_{W_2^k(\Omega)}.$$

This looks promising (order k approximation) but the right hand side has dependence on χ_i which we must remove. We estimate

$$\|f\chi_i\|_{W_2^k(\Omega)} \leq \|f\|_{W_2^k(\Omega)}\|\chi_i\|_{W_2^k(\Omega)} \tag{4.25}$$

and by estimates from, for example, [11], it follows that

$$\|\chi_i\|_{W_2^k(\Omega)} \leq Cq^{\frac{n}{2}-m}. \tag{4.26}$$

Applying (4.26) to (4.25) enables us to predict

$$\begin{aligned} E_i &\leq Ch^k \|f\chi_i\|_{W_2^k(\Omega)} \\ &\leq Ch^k q^{\frac{n}{2}-k} \|f\|_{W_2^k(\Omega)} \\ &\leq Ch^{\frac{n}{2}} \|f\|_{W_2^k(\Omega)} \end{aligned}$$

where we have used the quasi-uniformity assumption that there exists fixed constants such that $C_1 \leq \frac{h}{q} \leq C_2$. □

The result Lemma 9 is undesirable. We may slightly strengthen the result from $h^{\frac{n}{2}}$ approximation to h^n.

Lemma 10. *Under the same assumptions as in Lemma 9, we may show that*

$$E_i \leq Ch^n \omega(f, Kh\log(h)) + Ch^k \|f\|_{L^\infty}$$

We begin by rewriting the error as the integral of differences of Lagrange functions. We are using the fact here that $\sum \chi_i = 1$.

Proof.

$$\begin{aligned} E_i &= \left| \int_\Omega f(x)\chi_i(x)\,dx - f(x_i)w_i \right| = \left| \int_\Omega f(x)\chi_i(x)\,dx - f(x_i)\int_\Omega \chi_i(x)\,dx \right| \\ &= \left| \int_\Omega \big(f(x) - f(x_i)\big)\chi_i(x)\,dx \right| \end{aligned}$$

Let $K > 0$ be a fixed integer and let h denote the mesh norm of X. Let $B := \{x \in \Omega : \|x - x_i\| \leq Kh|\log(h)|\}$ and let $B^c := \Omega \backslash B$. We split the region of integration into two pieces, B and B^c, and use different approximations on each piece of the

integral. That is,

$$E_i = \left| \int_B \big(f(x) - f(x_i)\big)\chi_i(x)\,dx + \int_{B^c} \big(f(x) - f(x_i)\big)\chi_i(x)\,dx \right|.$$

We first consider the integral over B. Recall that the modulus of continuity $\omega(f, R) := \sup_{\|x-y\|\le R} |f(x) - f(y)|$ and $\omega(f, R) \to 0$ as $R \to 0$. Furthermore, it has been shown that the L^1 norm of χ_i is of order $q^n \sim h^n$. With these pieces of information, we compute

$$\begin{aligned}
\left| \int_B \big(f(x) - f(x_i)\big)\chi_i(x)\,dx \right| &\le \int_B |f(x) - f(x_i)||\chi_i(x)|\,dx \\
&\le \omega(f, Kh|\log(h)|) \int_B |\chi_i(x)|\,dx \\
&\le C\omega(f, Kh|\log(h)|)h^n
\end{aligned}$$

Now we consider the integral over B^c. Here we invoke the exponential decay of the Lagrange functions. It has been shown in [11] that

$$|\chi_i(x)| \le C \exp(-\frac{\nu}{h} d(x, x_i))$$

Consequently, $\sup_{x\in B^c} |\chi_i(x)| \le Ch^{K\nu}$. Therefore, we have

$$\begin{aligned}
\int_{B^c} |\big(f(x) - f(x_i)\big)||\chi_i(x)|\,dx &\le 2\|f\|_{L^\infty(\Omega)} h^{K\nu} \int_{B^c} dx \\
&\le 2\|f\|_{L^\infty(\Omega)} \mu(\Omega) h^{K\nu}
\end{aligned}$$

where $\mu(\Omega)$ is the volume of the region of integration. Combining the two integral estimates on B and B^c yields the result. If desired, we could replace the $\|f\|_{L^\infty}$ by bounding it above by $\|f\|_{W_2^k(\Omega)}$. K can be chosen to be larger than n (we had

freedom to choose it) and $\nu > 0$ is a constant. $\square$

Lemma 10 provides at least h^n convergence, which is an improvement over Lemma 9 which yielded just an $h^{\frac{n}{2}}$. However, we should note the h^n is multiplied by an $\omega(f, Kh|\log(h)|)$, a term which decreases to zero. A further refined estimate would require some control on how $\omega(f, R)$ decays. We note that we suspect this error estimate is still pessimistic.

This quadrature error problem is an open problem in radial basis functions which needs to be addressed. The results appear exceedingly pessimistic, especially for products $\chi_i\chi_j$. However, in practice, the quadrature performs quite well. We study one additional example of a quadrature error estimate by making use of Taylor series.

Lemma 11. *Let $f \in C^m(\Omega)$ (f has m continuous derivatives) and $\Omega \subset \mathbb{R}$. Then, the quadrature error is on the order of $h^{\frac{3}{2}}$.*

Proof. Let $x_i \in X$ be a center, let $K > 0$ be a fixed (chosen) integer and let $K_h := Kh|\log(h)|$. We need to estimate

$$\left| \int_\Omega f(x)\chi_i(x)\,dx - f(x_i)\int_\Omega \chi_i(x)\,dx \right| = \left| \int_\Omega \big(f(x) - f(x_i)\big)\chi_i(x)\,dx \right|.$$

We expand f into a Taylor series about x_i to obtain

$$f(x) = f(x_i) + f'(x_i)(x - x_i) + \frac{1}{2}f''(x_i)(x - x_i)^2 + \cdots + \frac{1}{m!}f^m(y)(x - y).$$

Therefore, it follows that

$$f(x) - f(x_i) = f'(x)(x - x_i) + \frac{1}{2}f''(x_i)(x - x_i)^2 + \cdots + \frac{1}{m!}f^m(y)(x - y).$$

Let $B := \{x \in \Omega : \|x - x_i\| \leq K_h\}$ and $B^c := \Omega \backslash B$. Then, we split the error into

two integrals over B and B^c respectively

$$\left| \int_\Omega \big(f(x) - f(x_i)\big)\chi_i(x) \right| \leq \left| \int_B \big(f(x) - f(x_i)\big)\chi_i(x)\, dx \right| + \left| \int_{B^c} \big(f(x) - f(x_i)\big)\chi_i(x)\, dx \right|.$$

As in the proof of Lemma 10, we may bound the integral over B^c using the exponential decay of the Lagrange function to find

$$\left| \int_{B^c} \big(f(x) - f(x_i)\big)\chi_i(x)\, dx \right| \leq Ch^{K\nu}\|f\|_{L^\infty}.$$

The new approach focuses on the integral over the ball B. We assume h is sufficiently small so that the Taylor expansion on B is valid, and we compute on B that

$$|f(x) - f(x_i)| = |f'(x_i)(x - x_i) + \cdots + \frac{f^{(m)}(y)}{m!}(x - y)^m|.$$

We apply the triangle inequality to get a term involving a polynomial $|(x-x_i)^k\chi_i(x)|$ for $k = 1, \ldots, m$. We note that since the integral is over B, $|x - y| \leq 2K_h$ for all $x, y \in B$. Therefore, we compute, for example,

$$\begin{aligned} \int_B |(x - x_i)^k||\chi_i(x)|\, dx &\leq \sup_x |(x - x_i)^k| \int_B \chi_i(x)\, dx \\ &\leq C(K_h)^k h^n \sim O(h^{k+n}|\log(h)|) \end{aligned}$$

where we invoked the fact that the L^1 norm of χ_i is on the order of h^n. Therefore, the error is $O(K_h h^n) = O(h^{n+1}|\log(h)|)$. Since we are in $\mathbb{R}$, $n = 1$, and therefore we have $O(h^2|\log(h)|)$. □

4.7.2 Local Lagrange Quadrature Error Estimates

The estimates in Lemma 9 and Lemma 10 both use the full Lagrange functions rather than the local Lagrange functions. We may extend the lemmas to get analogues of both Lemma 9 and Lemma 10 for the local Lagrange functions.

Lemma 12. *Under the same assumptions as in Lemma 9 with $\hat{\chi}_i$ denoting the local Lagrange functions, we have*

$$\left| \int_\Omega f(x)\hat{\chi}_i(x)\,dx - f(x_i)\hat{w}_i \right| \leq Ch^n \|f\|_{W_2^k(\Omega)}.$$

Proof. We begin by adding and subtracting terms within $|\int_\Omega f(x)\hat{\chi}_i(x)\,dx - f(x_i)\hat{w}_i|$. We add and subtract, $\int_\Omega f(x)\chi_i(x)\,dx$ and $f(x_i)w_i$. Therefore, we have

$$\begin{aligned}\left| \int_\Omega f(x)\hat{\chi}_i(x)\,dx - f(x_i)\hat{w}_i \right| \leq & \left| \int_\Omega f(x)\big(\hat{\chi}_i(x) - \chi_i(x)\big)\,dx \right| \\ & + \left| \int_\Omega \big(f(x) - f(x_i)\big)\chi_i(x)\,dx \right| \\ & + \left| f(x_i) \int_\Omega (\chi_i(x) - \hat{\chi}_i(x)\,dx \right|.\end{aligned}$$

We have broken up the local Lagrange quadrature error into three pieces: an error between the integrals of f times a Lagrange function vs. a local Lagrange function, the quadrature error using the full Lagrange function, and the difference between the local Lagrange quadrature weight and the full Lagrange quadrature.

We apply Corollary 4.11 [11] to bound $\|\chi_i - \hat{\chi}_i\|_{L^2} \leq Ch^J$ for a $J > 0$. The value of J can be increased by increasing the number of centers used in the local Lagrange function construction.

$$\left| \int_\Omega f(x)\big(\hat{\chi}_i(x) - \chi_i(x)\big)\,dx \right| \leq \|f\|_{L^2(\Omega)} \|\hat{\chi}_i - \chi_i\|_{L^2(\Omega)} \leq Ch^J \|f\|_{L^2(\Omega)}.$$

The next term we estimate is the quadrature error for the full Lagrange function, which we may apply Lemma 9 to find

$$\left| \int_\Omega \big(f(x) - f(x_i)\big)\chi_i(x)\,dx \right| \leq Ch^{\frac{n}{2}} \|f\|_{W_2^k(\Omega)}.$$

The last term we estimate controls the error between the quadrature weight from the full Lagrange function and the quadrature weight from the local Lagrange function. We have

$$\begin{aligned}
\left| f(x_i)\left(\int_\Omega \chi_i(x) - \hat{\chi}_i(x)\,dx \right) \right| &\leq \sup |f(x_i)| \int_\Omega \|\chi_i(x) - \hat{\chi}_i(x)\,dx \\
&\leq \|f\|_{L^\infty(\Omega)} \|\chi_i - \hat{\chi}_i\|_{L^1(\Omega)} \\
&\leq Ch^J \|f\|_{L^\infty(\Omega)}.
\end{aligned}$$

Therefore, we have bounded the local Lagrange quadrature error above by

$$\left| \int_\Omega f(x)\chi_i(x)\,dx - f(x)\hat{w}_i \right| \leq C\left(h^J\big(\|f\|_{L^2(\Omega)} + \|f\|_{L^\infty(\Omega)}\big) + h^{\frac{n}{2}} \|f\|_{W_2^k(\Omega)} \right)$$

We may choose J by tuning the value K used in the construction of the local Lagrange functions so that $J > \frac{n}{2}$. Therefore, we have the quadrature error is $O(h^{\frac{n}{2}})$. □

Corollary 2. *The estimate in Lemma 10 is valid for local Lagrange functions as well, provided sufficiently many centers are used in the construction of the local Lagrange function*

Proof. The proof is identical to Lemma 12 but we invoke Lemma 10 in the proof rather than Lemma 9. For $J > n$, the result holds. □

4.7.3 The Quadrature Error for Solutions to Nonlocal Diffusion Problems

We address the difficulty associated with getting an error estimate for the discrete solution we construct to the nonlocal diffusion problem. Let $u \in W_2^k(\Omega \cup \Omega_{\mathcal{I}})$ denote the solution which we aim to approximate. Let u_h, which is a linear combination of local Lagrange functions, denote the approximate solution that would be constructed *if there were no quadrature error*. This is not an interpolant; this would be the result of solving a linear system of equations where the matrices and vectors are constructed *without quadrature error*. The error estimate we derived currently shows that $\|u - u_h\| \sim O(h^k)$. As we know, this is inadequate because we do not construct u_h. Rather, we construct $\tilde{u}_h$, which is also a linear combination of local Lagrange functions like u_h, by solving a linear system of equations. The matrices and vectors for the solution $\tilde{u}_h$ are constructed by solving $\tilde{A}\tilde{u} = \tilde{b}$ where $\tilde{A}$ and $\tilde{b}$ are the arrays assembled by employing the one point quadrature methods. What we want is the following:

Conjecture 1. $\|u - \tilde{u_h}\|_{L^2} \leq Ch^k \|u\|_{W_2^k}$

This is significantly different from the current estimate, as this takes into account quadrature error. Numerically, we have strong evidence for this conjecture. One obvious approach is to simply apply the triangle equality, which yields

$$\begin{aligned} \|u - \tilde{u_h}\|_{L^2} &= \|u - u_h + u_h - \tilde{u}_h\|_{L^2} \\ &\leq \|u - u_h\|_{L^2} + \|u_h - \tilde{u}_h\|_{L^2} \\ &\leq Ch^k \|u\|_{W_2^k} + \underbrace{\|u_h - \tilde{u}_h\|_{L^2}}_{?} \end{aligned}$$

It does not follow from either of Lemma 9 or Lemma 10 that $\|u_h - \tilde{u}_h\|_{L^2} \leq Ch^k \|u\|_{W_2^k}$ unfortunately. It doesn't even follow that the error is $O(h^n)$ (where recall

that n is the spatial dimension). The reason is because the error $\|u_h - \tilde{u}_h\|_{L^2}$ is not a direct quadrature error. Let $u_h = \sum a_i \chi_i$ and $\tilde{u}_h = \sum \tilde{a}_i \chi_i$. Then,

$$\begin{aligned} \|u_h - \tilde{u}_h\|_{L^2} &= \| \sum (a_i - \tilde{a}_i \chi_i \|_{L^2} \\ &\leq C q^n \| \{a_i - \tilde{a}_i\} \|_{\ell^2} \end{aligned}$$

The coefficients $\{\tilde{a}_i\}$ are computed by building the matrix equations using equations the quadrature methods and solving the resulting linear systems. The coefficients $\{a_i\}$ are the coefficients that we would get if we constructed the linear system of equations without any quadrature error and then solved the system of equations. Therefore, problem of comparing u_h and $\tilde{u}_h$ is the following: first we build a linear system of equations, which has error built into it because we use quadrature, then we solve that linear system of equations to get $\tilde{u}_h$. A simple quadrature error estimate of the form given in either Lemma 9 or Lemma 10 is unfortunately insufficient; we either need a more sophisticated analysis on the quadrature errors in the matrices and we must understand how the solution of the linear system of equations is affected by this perturbation in the matrix, or we need a Strang's lemma. There is an "obvious" alternative. We simply could mimic the technique used in the elliptic partial differential equations solution method and introduce a dense set of quadrature nodes Y with mesh norm h_Y. Then, the framework used for the theory of quadrature error in the elliptic PDE method could be applied and we could derive comparable numerical results. We do not do this for one specific reason: introducing a larger point set Y destroys the advantage of one point quadrature for off diagonal entries that the method currently has.

4.7.4 An Alternative Nonlocal Diffusion Method

In the previous sections, we investigated a discretization of an anisotropic nonlocal diffusion problem with Dirichlet volume constraints by constructing a basis of local Lagrange functions. Since the discretization space was non-conforming with respect to the Hilbert space used for the variational formulation of the problem, a modified discretization space was constructed by enforcing the Dirichlet volume constraint on the local Lagrange functions. The method has proven to be effective in numerous ways. The method is conforming, it produces a sparse, symmetric, positive definite matrix. It can be assembled with ease on a CUDA GPU and inverted quickly with conjugate gradient, also on a GPU. Furthermore, the number of iterations required for inversion is independent of the mesh norm.

However, the previous method suffers from a severe drawback. Despite being simply inelegant, it is restricted to only work for Dirichlet volume constraints that are zero in the interaction domain. Neumann or Robin volume constraints were not explored and the basis would not reproduce functions with nonzero Dirichlet volume constraints. We present here a method that is capable of handling any volume constraint. The method employs the same concepts of local Lagrange functions and local Lagrange quadrature, although it does not require arbitrarily forcing the basis functions to be zero on the interaction domain. The drawback to the new method we present here is the method is *non-conforming*. This does not seem to impact the quality of the solution. The solution is enforced by means of Lagrange multipliers. The work we present here is joint work with Lehoucq, Narcowich, and Ward.

We study an alternative variational form that imposes volume constraints by Lagrange multipliers. We introduce an alternative variational form because the discretization spaces we propose are not contained in the constrained energy space.

This results in a non-conforming method; Lagrange multipliers provide a method to approximate the Dirichlet volume constraint. For functions $u \in L^2(\Omega \cup \Omega_{\mathcal{I}})$ and $\lambda \in L^2(\Omega_I)$, define the form $B : L^2(\Omega \cup \Omega_{\mathcal{I}}) \times L^2(\Omega_I) \to \mathbb{R}$

$$B(u, \lambda) = \int_{\Omega_I} \lambda(\boldsymbol{y}) u(\boldsymbol{y}) \, d\boldsymbol{y}.$$

Consider the following problem: find $u \in L^2(\Omega \cup \Omega_{\mathcal{I}})$ and $\lambda \in L^2(\Omega_I)$ such that

$$\begin{aligned} a(u, v) + B(v, \lambda) &= \int_{\Omega} v(\boldsymbol{x}) b(\boldsymbol{x}) \, d\boldsymbol{x} \qquad \text{for all } u \in L^2(\Omega \cup \Omega_{\mathcal{I}}) \\ B(u, w) &= \int_{\Omega_I} w(\boldsymbol{y}) g(\boldsymbol{y}) \, d\boldsymbol{y} \qquad \text{for all } w \in L^2(\Omega_I). \end{aligned} \tag{4.27}$$

Lemma 13. *There exists positive constants α_0, β_0 such that*

$$|a(u, v)| \leq \alpha_0 \|u\|_U \|v\|_U \text{ for all } u, v \in U$$

$$B(u, q) \leq \beta_0 \|u\|_U \|q\|_{\Pi} \text{ for all } u \in U \, , q \in \Pi$$

Proof. By the assumptions on the kernel,

$$\begin{aligned} |a(u, v)| &= \frac{1}{2} \int_{\Omega \cup \Omega_{\mathcal{I}}} \int_{\Omega \cup \Omega_{\mathcal{I}}} \big(u(\boldsymbol{x}) - u(\boldsymbol{y})\big)\big(v(\boldsymbol{x}) - v(\boldsymbol{y})\big)\big(\kappa(\boldsymbol{x}) + \kappa(\boldsymbol{y})\big)\Phi(\|\boldsymbol{x} - \boldsymbol{y}\|) \, d\boldsymbol{y} \, d\boldsymbol{x} \\ &\leq 2\|\kappa\|_{L^\infty(\Omega \cup \Omega_{\mathcal{I}})} \left(2 \left| \int_{\Omega \cup \Omega_{\mathcal{I}}} u(\boldsymbol{x}) v(\boldsymbol{x}) \int_{\Omega \cup \Omega_{\mathcal{I}}} \Phi(\|x - y\|) \, d\boldsymbol{y} \, d\boldsymbol{x} \right| \right. \\ &\quad \left. + 2 \left| \int_{\Omega \cup \Omega_{\mathcal{I}}} u(\boldsymbol{x}) \int_{\Omega \cup \Omega_{\mathcal{I}}} v(\boldsymbol{y}) \Phi(\|\boldsymbol{x} - \boldsymbol{y}\|) \, d\boldsymbol{y} \, d\boldsymbol{x} \right| \right) \\ &\leq 4\|\kappa\|_{L^\infty(\Omega \cup \Omega_{\mathcal{I}})} (\gamma_2 + \gamma_3) \|u\|_{L^2(\Omega \cup \Omega_{\mathcal{I}})} \|v\|_{L^2(\Omega \cup \Omega_{\mathcal{I}})} \end{aligned}$$

The form B is bounded by applying the Cauchy-Schwarz inequality

$$B(u, q) = \int_{\Omega_I} u(\boldsymbol{x}) q(\boldsymbol{x}) \, d\boldsymbol{x} \leq \|u\|_{L^2(\Omega_I)} \|q\|_{L^2(\Omega_I)} \leq \|u\|_{L^2(\Omega \cup \Omega_{\mathcal{I}})} \|q\|_{L^2(\Omega_I)}.$$

□

We now verify that the anisotropic problem is well-posed. Let $U := L^2(\Omega \cup \Omega_{\mathcal{I}})$, $\Pi := L^2(\Omega_I)$ and $Z := \{u \in U : B(u, q) = 0 \text{ for all } q \in \Pi\}$.

Lemma 14. *There exists a constant $\beta > 0$ such that*

$$\beta \|q\|_\Pi \leq \sup_{u \in U} \frac{B(u, q)}{\|u\|_U}.$$

Proof. Fix $q \in L^2(\Omega_I)$ and choose u such that $u|_{\Omega_I} = q$ almost everywhere and $u|_\Omega = 0$. Then,

$$B(u, q) = \int_{\Omega_I} u(\boldsymbol{x}) q(\boldsymbol{x}) \, d\boldsymbol{x} = \int_{\Omega_I} q(\boldsymbol{x})^2 \, d\boldsymbol{x}$$

and $\|u\|_U = \|u\|_{L^2(\Omega \cup \Omega_{\mathcal{I}})} = \|q\|_{L^2(\Omega_I)}$. Therefore, we have (with $\beta = 1$),

$$\|q\|_\Pi = \|q\|_\Pi \frac{\|q\|_\Pi}{\|q\|_\Pi} = \frac{B(u, q)}{\|u\|_U} \leq \sup_{u \in U} \frac{B(u, q)}{\|u\|_U}.$$

□

The bilinear form a is coercive when restricted to the space Z. The coercivity of the bilinear form on Z, the boundendess of a and b, and the inf-sup condition together implies that the problem (4.28) is well-posed [2].

We introduce finite dimensional discretization spaces $U_h = \text{span}\{\phi_i\}_{i=1}^N \subset L^2(\Omega \cup \Omega_{\mathcal{I}})$ and $\Lambda_h = \text{span}\{\psi_k\}_{k=1}^{N_I} \subset L^2(\Omega_I)$. The discrete problem is to find $u_h = \sum_{i=1}^N \alpha_i \phi_i \in U_h$ and $\lambda_h = \sum_{k=1}^{N_I} \beta_k \psi_k \in \Lambda_h$ such that

$$\begin{aligned} a(u_h, v_h) + B(v_h, \lambda_h) &= \int_\Omega v_h(\boldsymbol{x}) b(\boldsymbol{x}) \, d\boldsymbol{x} && \text{for all } v_h \in U_h \\ B(u_h, w_h) &= 0 && \text{for all } w_h \in \Lambda_h. \end{aligned} \tag{4.28}$$

This results in a linear system of equations

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}^T & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix} = \begin{pmatrix} \boldsymbol{b} \\ 0 \end{pmatrix} \tag{4.29}$$

where $\boldsymbol{\alpha}_i = \alpha_i, \boldsymbol{\beta}_k = \beta_k$, and the entries in the matrices above are

$$\mathbf{A}_{i,j} = a(\phi_i, \phi_j), \qquad \mathbf{B}_{i,k} = B(\phi_i, \psi_k), \qquad \boldsymbol{b}_i = \int_\Omega \phi_i b \, dx. \tag{4.30}$$

We introduce the space $Z_h = \{v_h \in V_h : \int_{\Omega_I} v_h b_i = 0 \text{ for all } b_i \in \Lambda_h\}$. Our goal is to verify that the bilinear form a is coercive on Z_h and that the spaces V_h and Λ_h satisfy a discrete inf-sup condition. That is, we need to argue there exists α_0 independent of h, q such that $a(z_h, z_h) \geq \alpha_0 \|z_h\|^2_{L^2(\Omega \cup \Omega_\mathcal{I})}$ and that there exists a $\beta > 0$ independent of h, q such that

$$\beta \|w_h\|_{L^2(\Omega_I)} \leq \sup_{v_h \in V_h} \frac{B(v_h, w_h)}{\|v_h\|_{L^2(\Omega \cup \Omega_\mathcal{I})}}$$

holds for all $w_h \in \Lambda_h$.

Proposition 12. *Let X be a quasi-uniformly distributed set of centers and $V_h = span\{b_\xi : \xi \in X\}$ and $\Lambda_h = span\{b_\eta : \eta \in X \cap \Omega\}$. Then, there exists a constant β independent of h, q such that*

$$\beta \leq \inf_{w_h \in \Lambda_h} \sup_{v_h \in V_h} \frac{B(v_h, w_h)}{\|w_h\|_{L^2(\Omega_I)} \|v_h\|_{L^2(\Omega \cup \Omega_\mathcal{I})}}.$$

Proof. We begin by fixing $w_h = \sum_{\eta \in X \cap \Omega} w_\eta b_\eta \in \Lambda_h \subset V_h$. Since $w_h \in V_h$, we may bound

$$\frac{B(w_h, w_h)}{\|w_h\|_{L^2(\Omega \cup \Omega_\mathcal{I})}} \leq \sup_{v_h \in V_h} \frac{B(v_h, w_h)}{\|v_h\|_{L^2(\Omega \cup \Omega_\mathcal{I})}}.$$

Recall that $B(w_h, w_h) = \int_{\Omega_I} w_h(y) w_h(y)\, dy = \|w_h\|^2_{L^2(\Omega_I)}$. If we show there exists a $C > 0$ independent of h, q such that $C\|w_h\|_{L^2(\Omega \cup \Omega_{\mathcal{I}})} \leq \|w_h\|_{L^2(\Omega_I)}$, the result follows. We apply [11, Theorem 6.2] to bound $\|w_h\|_{L^2(\Omega_I)} \geq C_{\Omega_I} q^{\frac{d}{2}} \|\vec{w}\|_{\ell^2}$. Furthermore, there exists a constant $C_{\Omega \cup \Omega_{\mathcal{I}}}$ such that $\|w_h\|_{L^2(\Omega \cup \Omega_I} \leq C_{\Omega \cup \Omega_{\mathcal{I}}} q^{\frac{d}{2}} \|\vec{w}\|_{\ell^2}$ by [11, Theorem 4.12]. Therefore, we may set $C := \frac{C_{\Omega_I}}{C_{\Omega \cup \Omega_{\mathcal{I}}}}$. □

A discrete coercivity result may be attained assuming that h is sufficiently small.

Theorem 4. *Assume that for $0 < h < h_0$ that $dist_{L^2(\Omega_I)}(1, \Pi_h) \leq \epsilon |\Omega_I|^{\frac{1}{2}}$ for some constant $0 < t \leq 1$ and $\epsilon \leq \frac{(1-t)\sqrt{\rho}}{1+\sqrt{1-\rho}}$. If $u_h \in Z_h$, then,*

$$a(u_h, u_h) \geq \frac{t\rho\lambda\delta^{d+2}}{1 + \sqrt{1-\rho}} \|u_h\|^2_{L^2(\Omega)}.$$

The constants λ and δ may be found in [23].

As a consequence of Proposition 12 and Theorem 4, it follows that the discretized problem is well-posed.

Proposition 13. *Let V_h, Π_h, Z_h, h_0 all be as defined above. For $h < h_0$, there exists unique functions $u_h \in V_h$ and $p_h \in \Pi_h$ that solve*

$$\begin{aligned} a(u_h, v_h) + B(v_h, \lambda_h) &= \int_\Omega v_h(\boldsymbol{x}) b(\boldsymbol{x})\, d\boldsymbol{x} && \text{for all } v_h \in U_h \\ B(u_h, w_h) &= 0 && \text{for all } w_h \in \Lambda_h. \end{aligned} \tag{4.31}$$

5. ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS ON SPHERES

The objective of this chapter is to introduce a Galerkin method for the solution of elliptic partial differential equations on manifolds. In particular, we study heat conduction on the sphere $\mathbb{S}^2$ in an inhomogeneous, anisotropic medium. That is, the equation we aim to discretize and solve is

$$Lu = -\text{div}(\mathbf{a} \cdot \nabla u) + bu = f \tag{5.1}$$

where div is the divergence operator and ∇ is the covariant derivative operator on the sphere, a is a rank two symmetric, positive definite tensor, and $f \in W_2^k(\mathbb{S}^2)$ for some $k \geq 0$. Most of the content of this section, as well as additional theory on the use of local Lagrange functions, may be found in [25]. The work presented here is joint work with Narcowich and Ward.

The primary tool for the discretization will be the surface splines $\kappa_m(x, y) := (-1)^m(1 - x \cdot y)^{m-1} \log(1 - x \cdot y)$. The conditionally positive definite surface spline has a reproducing kernel Hilbert space equivalent to the Sobolev space $W_2^m(\mathbb{S}^2)$.

Solving partial differential equations on spheres with spherical basis functions it not a new idea. However, introducing a meshfree Galerkin method with spherical basis functions along with a practical quadrature routine is novel. Previous methods relying on SBFs have relied on collocation schemes. For example, see [17].

We may expand (5.1) as

$$Lu = -\frac{1}{\sqrt{\det(g_{ij})}} \sum_{i,j} \frac{\partial}{\partial x_i} \sqrt{\det(g_{ij})} a^{ij}(x) \frac{\partial u}{\partial x_j} + b(x)u = f$$

We place appropriate assumptions on the data in (5.1). We assume that $b(x) \in$

$C^\infty(\mathbb{S}^2)$ and that a is a positive definite rank 2 tensor such that there exists positive constants c_1, c_2 such that

$$c_1 \sum_{i,j} g^{ij}(x) v_i v_j \leq \sum_{i,j} a^{ij}(x) \leq c_2 \sum_{i,j} g^{ij}(x) v_i v_j \tag{5.2}$$

for any vectors $v \in T_x(\mathbb{S}^2)$, the tangent space of x. We further assume that there exists a constant $b_1 \geq b_0 > 0$ such that $b_0 \leq b(x) \leq b_1$. In the event we choose $a = g$, then the problem reduces to

$$Lu = -\Delta u + bu = f.$$

In the event that $a \neq g$ has spatial variation, we refer to this as the anisotropic problem.

To solve (5.1), we place the problem into weak form as the first step. We begin by multiplying (5.1) by a test function $v \in W_2^1(\mathbb{S}^2)$ and integrating by parts to place the equation into weak form. We define the bilinear form $a(u, v)$ and the linear function $\ell(v)$ operating on $u, v \in W_2^1(\mathbb{S}^2)$ by

$$\begin{aligned} a(u, v) &:= \langle Lu, v \rangle_{L^2(\mathbb{S}^2)} \int_{\mathbb{S}^2} a^{ij}(x) \sum_{i,j} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} + b(x) u(x) \, d\mu(x), \\ \ell(v) &:= \int_{\mathbb{S}^2} f(x) v(x) \, d\mu(x). \end{aligned} \tag{5.3}$$

We trust the reader will not confuse $a(u, v)$, the bilinear form, with the similar nomenclature a^{ij} for the rank two tensor. We may now formally state the problem we aim to discretize and solve. Let $a(\cdot, \cdot)$ and $\ell(\cdot)$ be as defined in (5.3). We seek

$u \in W_2^1(\mathbb{S}^2)$ so that for every $v \in W_2^1(\mathbb{S}^2)$,

$$a(u, v) = \ell(v). \tag{5.4}$$

We need to guarantee that a unique solution exists for this problem. This follows by an application of the famous Lax-Milgram Lemma. To apply the Lax-Milgram lemma, the bilinear form $a(u, v)$ must be coercive and bounded and the linear function $\ell(v)$ must be bounded. We recall that we say a bilinear form $a : H \times H \to \mathbb{R}$ is coercive on a Hilbert space H if $a(u, u) \geq \alpha \|u\|_H^2$ for some positive α.

Proposition 14. *Let a and ℓ denote the bilinear form and linear functional, respectively, defined in* (5.3)*. Then, a is coercive and bounded on $W_2^1(\mathbb{S}^2)$ and ℓ is a bounded linear functional on $W_2^1(\mathbb{S}^2)$.*

Proof. Recall the assumption that were placed upon a^{ij} and $b(x)$. We know there exists c_1, c_2 relating g^{ij} and a^{ij} and that there exists $0 < b_0 \leq b(x) \leq b_1$. Therefore, we may bound the bilinear form above and below

$$\begin{aligned}
\min(c_1, b_0)\|u\|^2_{W_2^1(\mathbb{S}^2)} &\leq \int_{\mathbb{S}^2} c_1 \sum_{i,j} g^{ij} \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j} + b_0 u^2(x)\, d\mu(x) \leq a(u, v), \\
\max(c_2, b_1)\|u\|_{W_2^1(\mathbb{S}^2)} &\geq \int_{\mathbb{S}^2} c_2 \sum_{i,j} g^{ij} \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j} + b_1 u^2(x)\, d\mu(x) \geq a(u, v).
\end{aligned}$$

□

Consequently, the bilinear form is coercive, and actually defines an inner product equivalent to the $W_2^1(\mathbb{S}^2)$ inner product.

Lemma 15. *Let $\ell(v)$ denote the linear functional $\ell(v)$ is continuous in the $W_2^1(\mathbb{S}^2)$ norm.*

Proof. This result follows by a quick application of the Cauchy-Schwarz inequality. We see

$$\ell(v) = \int_{\mathbb{S}^2} f(x)v(x)\,d\mu(x) \leq \|f\|_{L^2(\mathbb{S}^2)}\|v\|_{L^2(\mathbb{S}^2)} \leq \|f\|_{L^2(\mathbb{S}^2)}\|v\|_{W_2^1(\mathbb{S}^2)}.$$

□

As a consequence of (14) and (15), we may apply the Lax-Milgram lemma to conclude that there exists a unique $u \in W_2^1(\mathbb{S}^2)$ such that $a(u,v) = \ell(v)$ for all $v \in W_2^1(\mathbb{S}^2)$. Consequently, a solution exists and is unique and we may pursue a discretization of the problem to approximate the solution u. We note a regularity condition that is due to the ellipticity of the differential operator.

Proposition 15. *[25] Let L be the differential operator as described in* (5.1) *along with our assumptions on the tensor a^{ij} and the function b. Then, if u is any distributional solution to the problem $Lu = f$, where $f \in W_2^s(\mathbb{S}^2)$ for $s \geq 0$, then for any $t < s-1$, there exists a constant C_t so that $u \in W_2^{s+2}(\mathbb{S}^2)$ and $\|u\|_{W_2^{s+2}} \leq C_t(\|u\|_{W_2^s} + \|u\|_{W_2^t})$. Furthermore, $\|u\|_{W_2^{s+2}(\mathbb{S}^2)} \leq C\|Lu\|_{W_2^s(\mathbb{S}^2)}$.*

This regularity result guarantees the amount of smoothness u will posses. Intuitively, u should have at least two more derivatives than f, which is exactly what the regularity result confirms. Such results are not surprising and known due to the C^∞ coefficients in (5.1) and since the operator L is elliptic.

As a result of the properties of the bilinear form, we may note a corollary of the regularity proposition that enables us to enhance our error estimates later.

Corollary 3. *Let $V_h \subset W_2^1(\mathbb{S}^2)$ be a closed subspace. Let u_h denote the Galerkin solution to the problem $a(u_h, v_h) = \ell(v_h)$ for all $v_h \in V_h$. Let w denote the solution*

to the problem $Lw = u - u_h$ and let w_h denote the corresponding Galerkin solution in v_h. Then,

$$\|u - u_h\|_{L^2}^2 \leq M\|u - u_h\|_{W_2^1(\mathbb{S}^2)}\|v - v_h\|_{W_2^1(\mathbb{S}^2)}.$$

Proof. We first note that the bilinear form a was constructed by integration by parts of $\langle Lu, v\rangle_{L^2(\mathbb{S}^2)} := a(u, v)$. Then, we use this to evaluate

$$a(w, u - u_h) = \langle Lw, u - u_h\rangle_{L^2(\mathbb{S}^2)} = \|u - u_h\|_{L^2(\mathbb{S}^2)}^2.$$

We note that since u_h is the Galerkin solution corresponding to u, Galerkin orthogonality implies that $a(w_h, u - u_h) = 0$. Applying the Galerkin orthogonality, additivity of the bilinear form and boundedness of the bilinear form, we evaluate

$$\begin{aligned}\|u - u_h\|_{L^2(\mathbb{S}^2)}^2 &= a(w, u - u_h) = a(w - w_h, u - u_h)\\ &\leq M\|w - w_h\|_{W_2^1(\mathbb{S}^2)}\|u - u_h\|_{W_2^1(\mathbb{S}^2)}.\end{aligned}$$

□

5.1 Error Estimates with Spherical Basis Functions

We employ spherical basis functions for the discretization method we develop to approximate the solution to the problem $Lu = f$. Consider a spherical basis function ϕ on $\mathbb{S}^2$ that is either positive definite or conditionally positive definite. Let X be a collection of quasi-uniformly scattered centers on $\mathbb{S}^2$ and let

$$V_X := V_{\phi,X} = \left\{ \sum_{\xi\in X} a_\xi \phi(\cdot, \xi) : \sum_{\xi} a_\xi p(\xi) = 0 \text{ for all } p \in \Pi \right\} + \Pi$$

where Π is the space of continuous functions corresponding to the conditionally positive definite kernel ϕ. For this section, we assume Π is the span of spherical harmonics up to a fixed degree L (not to be confused with the differential operator L). We place some restrictions on the behavior of the kernel to guarantee that we may employ known interpolation error estimates. Let $\tau > 1$ and assume that there exists positive constants c, C so that the Fourier coefficients of ϕ satisfy

$$c(1+\lambda_\ell)^{-\tau} \leq \hat{\phi}_\ell \leq C(1+\lambda_\ell)^{-\tau} \text{ for all } \ell \geq L+1 \tag{5.5}$$

where L is the degree of the highest order spherical harmonic in Π.

Let $H^k := W_2^k(\mathbb{S}^2)$ for shorthand. In the weak formulation, the Hilbert space we search in is H^1. For sufficient $\tau > 1$, the space $V_X \subset H^1$, and consequently we may considered the discretized following discretized problem: Find $u_X \in V_X$ such that for all $v_X \in V_X$

$$a(u_X, v_X) = \ell(v_X).$$

Since $V_X \subset H^1$, the discretization is conforming, and consequently an application of Lax-Milgram once again guarantees such a solution exists. We refer to the solution u_X as the *discrete solution* in contrast to the solution (or *exact solution*) u. Furthermore, we know by Cea's Lemma that the discrete solution is a near-optimal approximation to u in the discretization space V_X. That is,

$$\|u - u_X\|_{H^1} \leq C \inf_{v_X \in V_X} \|u - v_X\|_{H^1}.$$

Consequently, by choosing the interpolant $I_X u \in V_X$, it follows that

$$\|u - u_X\|_{H^1} \leq C\|u - I_X u\|_{H^1}.$$

By identical reasoning, for the problem $Lw = u - u_X$, we know the Galerkin solution $w_X \in V_X$ satisfies

$$\|w - w_X\|_{H^1} \leq C\|w - I_X w\|_{H^1}.$$

Regarding the regularity of each solution, if the data $f \in H^s$, then by the regularity result Proposition 15, it follows that $u \in H^{s+2}$. For $\tau > 1$, the interpolant $I_X u$ is in the space $H^{\tau+\alpha}$ for $\alpha < \tau - 1$. As a result, the difference $u - I_X u \in H^\sigma$ where $\sigma = \min(s+2, \tau+\alpha)$. We may once again apply elliptic regularity Proposition 15 to the problem $Lw = u - I_X u$ to find that since $u - I_X u \in H^\sigma$, $w \in H^{\sigma+2}$.

We consider the kernel $\phi = \kappa_m$, the restricted surface spline that is in the Sobolev space $W_2^m(\mathbb{S}^2)$. Let χ_ξ denote the Lagrange function centered at ξ constructed by linear combinations of $\kappa_m(\cdot, \eta)$ plus an appropriate polynomial from Π_m, the space of degree m spherical polynomials. The space V_X is the space we employ to discretize the problem (5.1). With these regularity results combined with the approximation powers of the discretization spaces, we may derive error estimates for the problem.

Lemma 16. *Let $\tau > \frac{n}{2}$ (for $\mathbb{S}^n$ or $\tau > 1$ for $\mathbb{S}^2$). Let X be a collection of scattered centers with mesh norm h. Let w be the solution to the problem $Lw = u - I_X u$. Then,*

$$\|w - w_X\|_{H^1} \leq Ch^2\|u - I_X u\|_{H^3}.$$

Proof. We know $w \in H^{\sigma+2}$ where $\sigma = \min(s+2, \tau+\alpha)$ for $\tau > \frac{n}{2}$ and $\alpha < \tau - \frac{n}{2}$. Then, as a result, $\sigma > 1$, so $\sigma + 2 > 3$. Therefore, $w \in H^3$. As a result, we may apply the interpolation error estimates for the kernel ϕ, which allows us to estimate

$$\|w - I_x w\|_{H^1} \leq Ch^2\|w\|_{H^3}.$$

Furthermore, by the regularity result Proposition 15, it follows

$$\|w\|_{H^3} \leq C\|Lw\|_{H^1} = C\|u - I_X u\|_{H^1}.$$

□

We focus on the specific case of $n = 2$ since the sphere $\mathbb{S}^2$ is our primary manifold of interest. For general $\mathbb{S}^n$, the key replacements are that $\tau > \frac{n}{2}$ and $\alpha < \tau - \frac{n}{2}$.

Theorem 5. *Let $n = 2$ and let $Lu = f$ for $f \in H^s$ and $s \geq 0$. Let u_X denote the Galerkin solution to the problem $Lu = f$ in the space V_X constructed from scattered centers with mesh norm h and a kernel ϕ that satisfies* (5.5)*. Then,*

$$\|u - u_X\|_{L^2} \leq \begin{cases} Ch^{s+2}\|u\|_{H^{s+2}} & \text{for } s \leq 2\tau - 2 \\ Ch^{2\tau}\|u\|_{H^{2\tau}} & \text{for } 2\tau - 2 < s \end{cases} \tag{5.6}$$

Proof. The interpolation error estimates derived in [25] (Proposition A.3) show that on $\mathbb{S}^n$, if $\tau > \frac{n}{2}$ and $\frac{n}{2} < \mu \leq 2\tau$, and $\beta \leq \min(\mu, \tau)$, then

$$\|u - I_X u\|_{H^\beta} \leq Ch^{\mu-\beta}\|u\|_{H^\mu}. \tag{5.7}$$

In the context of our problem, $u \in H^{s+2}$ where $s \geq 0$. The highest order approximation power possible is 2τ, which occurs in the event $2\tau < s + 2$, and consequently, the two cases to consider are $s \leq 2\tau - 2$ and $s > 2\tau - 2$.

By Cea's Lemma, it follows that

$$\|u - u_X\|_{H^1} \leq C\|u - I_X u\|_{H^1}.$$

We apply the interpolation error estimates to approximate the convergence order of

$u - I_X u$ in the H^1 norm. In the language of the interpolation error estimate (5.7), $\beta = 1$, and we see

$$\|u - I_X u\|_{H^1} \leq \begin{cases} Ch^{s+1}\|u\|_{H^{s+1}} \text{ for } s \leq 2\tau - 2, \\ Ch^{2\tau-1}\|u\|_{H^{2\tau}} \text{ for } s > 2\tau - 2. \end{cases} \tag{5.8}$$

To recover L^2 error estimates, we need to apply the Nitsche trick to recover an additional order of h. By Corollary 3,

$$\|u - u_X\|_{L^2}^2 \leq \|w - I_X w\|_{H^1}\|u - I_X u\|_{H^1}.$$

By applying Lemma 16, we may replace the $w - I_X w$ norm by

$$\|u - u_X\|_{L^2}^2 \leq Ch^2\|u - I_X u\|_{H^1}^2$$

which implies that

$$\|u - u_X\|_{L^2} \leq Ch\|u - I_X u\|_{H^1}.$$

Finally, we may apply (5.8) to acquire the approximation error rates

$$\|u - I_X u\|_{L^2} \leq \begin{cases} Ch^{s+2}\|u\|_{H^{s+1}} \text{ for } s \leq 2\tau - 2 \\ Ch^{2\tau}\|u\|_{H^{2\tau}} \text{ for } s > 2\tau - 2. \end{cases} \tag{5.9}$$

□

The error estimate we derived is under the assumption that the stiffness matrix is assembled precisely with no quadrature error. The issue of quadrature is technical and difficult and we study this problem in a later section. The assumptions on the

kernel are given that the kernel is (conditionally) positive definite and the conditions in (5.5). We remark that our particular interest is in the surface spline $\phi_m(x, y) := (-1)^m(1 - x \cdot y)^{m-1} \log(1 - x \cdot y)$.

5.2 Stiffness Matrix in the Lagrange Basis

The error estimates derived in Section 5.1 are valid for many kernels. While the theoretical error estimates in Theorem 5 are valid for a diverse set of kernels, the choice of kernel has impact in the practical implementation of the method. A numerically suitable, robust basis that is easy to assemble, is well-conditioned, and preferably sparse is highly desired. Optimizing these conditions requires a choice of a basis in the discretization space, V_X. We explore the use of the *Lagrange basis* for a basis for the discretization space.

The kernels we consider are the surface splines ϕ_m for $m \geq 2$ on the sphere $\mathbb{S}^2$. These spaces satisfy the error estimates in Theorem 5 and also are known to have a robust, exponentially decaying Lagrange basis. Chapter 2 provides the theoretical properties of this basis on the sphere. In particular, the basis is highly localized spatially and has a "small-footprint" in the sense that the number of kernels with non-negligible coefficients is on the order of $\mathcal{O}\big(\log(N)^2\big)$. The basis can be approximated very well by the alternative basis of *local Lagrange functions*, as described in Section 2.2.

Let $V_{m,X} := V_{\phi_m,X}$ be the discretization space generated by ϕ_m and the appropriate spherical polynomials. We choose for the basis the collection of Lagrange functions, $V_{m,X} = \text{span}\{\chi_\xi : \xi \in X\}$ that satisfy $\chi_\xi(\eta) = \delta_{\xi,\eta}$. The Galerkin solution to the problem $a(u_X, v_X) = \ell(v_X)$ for all $v_X \in V_X$ can be expressed in terms of the basis as $u_X := \sum_{\xi \in X} u_\xi \chi_\xi$. Substituting this into (5.4) along with the choice $v_X = \chi_\eta$

for each $\eta \in X$ yields the linear system of equations

$$\sum_{\xi} u_\xi a(\chi_\xi, \chi_\eta) = \ell(\chi_\eta).$$

We identify

$$\begin{aligned} \mathbf{A}_{\xi,\eta} &:= a(\chi_\xi, \chi_\eta) = \int_{\mathbb{S}^2} a^{ij}(x) \nabla\chi_\xi \cdot \nabla\chi_\eta + b(x)\chi_\xi\chi_\eta \, d\mu(x), \\ \mathbf{f}_\eta &:= \ell(\chi_\eta) = \int_{\mathbb{S}^2} f(x)\chi_\eta(x) \, d\mu(x). \end{aligned} \tag{5.10}$$

Letting $\mathbf{u}_\xi := u_\xi$, the linear system of equations we aim to solve are

$$\mathbf{A}\mathbf{u} = \mathbf{f}. \tag{5.11}$$

We refer to $\mathbf{A}$ as the *stiffness matrix*. The choice of basis affects the properties of the stiffness matrix. We show that our choice of basis offers many benefits, but the assembly of the matrix is non-trivial. Evaluating the inner products in (5.10) requires a quadrature routine, which we discuss later. Without an effective, implementable quadrature routine, the use of spherical basis functions may not be a practical choice.

We begin our exploration of the properties of the stiffness matrix by studying the condition number of the stiffness matrix. Recall that the condition number, $\kappa(\mathbf{A}) := \kappa_2(\mathbf{A}) = \|\mathbf{A}\|_2 \cdot \|\mathbf{A}^{-1}\|_2$. We note that $\mathbf{A}$ is real and self-adjoint, since $\mathbf{A}_{\xi,\eta} = a(\chi_\xi, \chi_\eta) = a(\chi_\eta, \chi_\xi) = \mathbf{A}_{\eta,\xi}$. Since the bilinear form defines an inner product, $\mathbf{A}$ is a Gram-matrix for the basis $\{\chi_\xi\}_{\xi\in X}$ in the $a(\cdot,\cdot)$ inner product and consequently $\mathbf{A}$ is positive definite. Therefore, $\kappa(\mathbf{A}) = \frac{\lambda_{\max}(\mathbf{A})}{\lambda_{\min}(\mathbf{A})}$.

We first establish a bound on $\lambda_{\min}(A) := \lambda_1$. We estimate the eigenvalue by considering the quadratic form $\langle \mathbf{A}\mathbf{c}, \mathbf{c}\rangle$ with $\|\mathbf{c}\|_{\ell^2} = 1$. The minimum and maximum values of the quadratic form correspond to the minimal and maximal eigenvalues of

$\mathbf{A}$, respectively. Let N denote the cardinality of the set of centers X. We compute

$$\begin{aligned}\langle \mathbf{Ac}, \mathbf{c}\rangle &= \sum_{i=1}^{N} (\mathbf{Ac})_i \mathbf{c}_i \\ &= \sum_{i=1}^{N} \big(\sum_{j=1}^{N} \mathbf{A}_{i,j}\mathbf{c}_j \big)\mathbf{c_i} \\ &= \sum_{i=1}^{N} \big(a(\chi_i, \chi_\eta)\mathbf{c}_j \big)\mathbf{c}_i \\ &= a\bigg(\sum_{i=1}^{N} \mathbf{c}_i\chi_i, \sum_{i=1}^{N} \mathbf{c}_j\chi_j \bigg).\end{aligned}$$

Let $v := \sum_{i=1}^{N} \mathbf{c}_i\chi_i$. Then, $\lambda_{\min}(\mathbf{A}) = \min a(v,v)$. We note that $a(v,v) := \langle Lv, v\rangle_{L^2(\mathbb{S}^2)}$, and consequently,

$$\lambda_{\min}(L)\|v\|^2_{L^2(\mathbb{S}^2)} \leq \langle \mathbf{Ac}, \mathbf{c}\rangle$$

Since $\lambda_{\min}(L)$ is independent of the choice of discretization (and positive), this is a fixed constant. We need only estimate $\|v\|_{L^2(\mathbb{S}^2)}$ in terms of the coefficients. We know by Proposition 5 that we may bound

$$\|v\|^2_{L^2(\mathbb{S}^2)} \geq Cq^2\|\mathbf{c}\|_{\ell^2(N)}.$$

Consequently, we may estimate the lower eigenvalue of $\lambda_{\min}(\mathbf{A})$ by

$$\lambda_{\min}(\mathbf{A}) = \min_{\|\mathbf{c}\|=1} \langle \mathbf{Ac}, \mathbf{c}\rangle \geq \lambda_{\min}(L)q^2\|c\|^2_{\ell^2(N)} = C\lambda_{\min}(L)q^2. \tag{5.12}$$

We investigate the upper bound in a similar fashion, although we require a Bernstein inequality to relate $\|Lv\|_{L^2}$ to $\|v\|_{L^2}$. We may not proceed as before and claim $\langle \mathbf{Ac}, \mathbf{c}\rangle \leq \lambda_{\max}(L)\|v\|^2_{L^2}$ since there exists no maximal eigenvalue of the differential

operator L. Instead, we apply Cauchy-Schwarz and use the Bernstein inequality (8) to compute

$$\begin{aligned}\lambda_{\max}(\mathbf{A}) &= \max_{\|\mathbf{c}\|=1}\langle \mathbf{A}\mathbf{c},\mathbf{c}\rangle \leq \max_{\|\mathbf{c}\|}\langle Lv, v\rangle_{L^2} \\ &\leq \max_{\|\mathbf{c}\|=1}\|Lv\|_{L^2}\|v\|_{L^2} \leq \max_{\|\mathbf{c}\|=1} Cq^{-2}\|v\|_{L^2}^2 \\ &\leq \max_{\|\mathbf{c}\|=1} Cq^{-2}q^2\|c\|_{\ell^2(N)}\end{aligned} \tag{5.13}$$

and we therefore may conclude that the maximal eigenvalue $\lambda_{\max}(\mathbf{A})$ is bounded above by a constant independent of h or q.

Theorem 6. *Let* $\mathbf{A}$ *denote the stiffness matrix assembled using the basis* $\{\chi_\xi : \xi \in X\}$ *for a quasi-uniformly distributed set of centers* X *with separation radius* q *that is sufficiently small. Then, the condition number of the stiffness matrix is bounded above by*

$$\kappa_2(\mathbf{A}) \leq Cq^{-2}$$

for some positive constant C *independent of the mesh norm* h *or* q*.*

Proof. Since $\mathbf{A}$ is a symmetric, positive definite matrix, we know the condition number is bounded above by the ratio of the maximal and minimal eigenvalues. By (5.12) and (5.13), we compute

$$\kappa_2(\mathbf{A}) = \frac{\lambda_{\max}(\mathbf{A})}{\lambda_{\min}(\mathbf{A})} \leq C\frac{1}{q^2\lambda_{\min}(L)}.$$

□

In addition to being stable, the stiffness matrix has banded properties and is essentially sparse. While each entry in the stiffness matrix is nonzero, and hence the matrix is dense, most off-diagonal entries are negligible in value. The stiffness

matrix entry $\mathbf{A}_{\xi,\eta} = a(\chi_\xi, \chi_\eta)$ satisfies a decay property so that as $d(\xi, \eta)$ increases, the value of $\mathbf{A}_{\chi_\xi,\chi_\eta}$ decays exponentially. This mimics the spatial decay properties of the Lagrange functions.

Proposition 16. *Let* $\mathbf{A}$ *denote the stiffness matrix. For sufficiently small* h*, the entries in* $\mathbf{A}$ *decay as*

$$\mathbf{A}_{\xi,\eta} \leq Ch^{-2} \exp\big(-\frac{\nu}{h} dist(\xi, \eta)\big).$$

Proof. We take advantage of the decay of the Lagrange functions, as well as the decay of the covariant derivatives of Lagrange functions. We know

$$\mathbf{A}_{\xi,\eta} = \int_{\mathbb{S}^2} \sum_{i,j} a^{ij} \nabla\chi_\xi \cdot \nabla\chi_\eta + b(x)\chi_\xi\chi_\eta \, d\mu(x).$$

We may bound the integrand above by noting

$$\left| \sum_{i,j} a^{ij} \nabla\chi_\xi \nabla\chi_\eta + b(x)\chi_\xi\chi_\eta \right| \leq C|\nabla\chi_\xi| \, |\nabla\chi_\eta| + \|b\|_{L^\infty} |\chi_\xi\| |\chi_\eta|$$

We then note that, by the exponential decay, we know

$$\begin{aligned} |\chi_\xi(x)\chi_\eta(x)| &\leq C \exp(-\frac{\nu}{h}\text{dist}(\xi, x)) \exp(-\frac{\nu}{h}\text{dist}(\eta, x)) \\ &\leq C \exp(-\frac{\nu}{h}\text{dist}(\chi, \eta)) \end{aligned}$$

where we took advantage of the fact that by the triangle inequality, $\text{dist}(x, \eta) + \text{dist}(\xi, x) \geq \text{dist}(\xi, \eta)$. Similarly, the decay of the covariant derivatives of Lagrange

functions follows with an additional q^{-1} power, and we consequently have

$$\begin{aligned} C|\nabla\chi_\xi|\,|\nabla\chi_\eta| + \|b\|_{L^\infty}|\chi_\xi\|\chi_\eta| &\leq (Cq^{-2}+1)\exp(-\frac{\nu}{h}\mathrm{dist}(\xi,\eta)) \\ &\leq \tilde{C}h^{-2}\exp(-\frac{\nu}{h}\mathrm{dist}(\xi,\eta)) \end{aligned}$$

where we have used the quasi-uniformity to flip from q to h and adjusted the constant appropriately. □

A consequence of Proposition 16 is that the stiffness matrix entries off of the diagonal decay exponentially fast. This is very advantageous, since the stiffness matrix is dense. However, exponential decay off of the diagonal suggests very few terms will have non-negligible value. Indeed, by choosing a ball of radius $r = Kh|\log(h)|$, we see that for $\eta \notin B(\xi, r)$,

$$\mathbf{A}_{\xi,\eta} \leq Ch^{-2}\exp(-\frac{\nu}{h}\mathrm{dist}(\xi,\eta)) \leq Ch^{-2}h^{K\nu}$$

and consequently, for $K\nu - 2$ sufficiently small, the stiffness matrix entries will be very small. Consequently, we may choose a cutoff radius for appropriate K and zero out entries in the stiffness matrix beyond this point. This produces a sparse matrix with comparable condition number and comparable accuracy for the construction of the Galerkin solution. Furthermore, from a practical standpoint, the stiffness matrix assembly may be greatly accelerated: only a few entries must be computed for each row. This greatly reduces the cost of assembly and makes the method more advantageous from a computational view.

5.3 Quadrature

The theoretical properties of the discretization method of Section 5.1 suggest that the use of the spherical basis functions, and in particular, Lagrange functions, produces a stable stiffness matrix with optimal approximation power. While the theoretical properties of the stiffness matrix $\mathbf{A}$ are excellent, a practical issue of concern is the assembly of the stiffness matrix numerically. Evaluating a single entry in the matrix requires computing non-trivial integrals of covariant derivatives of functions. Providing a practical numerical quadrature routine is imperative for an efficient implementation. An efficient quadrature routine is not sufficient; the qualities of the stiffness matrix must be preserved even when assembled with numerical quadrature.

We propose a quadrature routine using an additional, separate set of centers, distinct from the collection of centers used to construct the discretization space. We let X refer to the collection of centers used to build the approximation space, V_X, as in Section 5.1. We let Y denote the collection of centers, which we refer to as *quadrature nodes*. The space V_X is constructed by a kernel ϕ_m, while the space V_Y is constructed by a possibly different order kernel, ϕ_M. There is no need to enforce $M = m$, and no advantage exists in choosing $M > m$. Since lower M values result in better conditioned linear systems for interpolation (and hence, quadrature weight construction), choosing a lower value of M is recommended.

For the analysis and discussion in this section, we make an additional assumption on the tensor a^{ij}. While we previously enforced Equation (5.2), we now add the additional constraint that there exists a function $a \in C^\infty(\mathbb{S}^2)$ such that $a^{ij}(x) = a(x)g^{ij}(x)$.

Let h_X and h_Y denote the mesh norms of X and Y, respectively. We assume $h_Y \leq h_X$ and $q_Y \leq q_X$. Let $\{\chi_\xi\}_{\xi \in X}$ be the set of Lagrange functions. We refer

to the stiffness matrix $\mathbf{A}$ assembled with no numerical quadrature error (i.e., as in Section 5.1) as the *continuous* stiffness matrix. The continuous stiffness matrix is assembled entrywise by

$$A_{\xi,\eta} = \int_{\mathbb{S}^2} a(x)\nabla\chi_\xi \cdot \nabla\chi_\eta \, dx + \int_{\mathbb{S}^2} b(x)\chi_\xi(x)\chi_\eta(x) \, d\mu(x). \tag{5.14}$$

Let $\{w_z\}_{z\in Y}$ denote the quadrature weights associated with the set of centers Y. The quadrature weights enable perfect reproduction formulas for functions in V_Y as well as polynomials in the space Π_M corresponding to the kernel ϕ_m. However, it is possible that some of the computed weights may be negative in the case of highly scattered data sets Y. For quasi-uniformly distributed nodes, this is not likely the case and a reasonable assumption is that

$$w_z \geq Ch_Y^2 \tag{5.15}$$

for each $z \in Y$. All numerical experiments use quadrature nodes satisfying (5.15). On the other hand, regardless of whether (5.15) holds or not, we do know there exists a positive constant so that

$$|w_z| \leq Ch_Y^2.$$

This holds as a result of (5). Since $w_z = \int_{\mathbb{S}^2} \chi_z(x) \, d\mu(x)$ for each $z \in Y$, we apply (5) to compute

$$|w_z| \leq \|\chi_z\|_{L^1(\mathbb{S}^2)} \leq Cq^2 = C\rho^{-2}h^2.$$

We apply quadrature to (5.14) to construct a discrete stiffness matrix

$$\tilde{A}_{\xi,\eta} = \sum_{z\in Y} a(z)\nabla\chi_\xi \cdot \nabla\chi_\eta(z)w_z + \sum_{z\in Y} b(z)\chi_\xi(z)\chi_\eta(z)w_z. \tag{5.16}$$

The right hand side vector for the continuous case and for the discrete case are given by, respectively,

$$b_\xi = \int_{\mathbb{S}^2} f(x)\chi_\xi(x)\,dx \qquad \tilde{b}_\xi = \sum_{z\in Y} f(z)\chi_\xi(z)w_z. \tag{5.17}$$

Our objective is to present an approach to estimating the effects of using quadrature to evaluate the vectors and matrices for the problem. The results presented here were an initial attempt; for a detailed and superior analysis that takes into account banding the matrices and local Lagrange functions, see [25].

In the continuous case, we solve the problem $Aa = b$, and represent the solution $u_h = \sum_{\xi\in X} a_\xi\chi_\xi(x)$. The discrete solution vector is computed by solving $\tilde{A}\tilde{a} = \tilde{b}$, and the discrete solution $\tilde{u}_h = \sum_{\xi\in X} \tilde{a}_\xi\chi_\xi$. We may compute the L^2 norm error between these functions by noting that

$$\begin{aligned}\|u_h - \tilde{u}_h\|_{L^2} &= \|\sum_{\xi\in X}(a_\xi - \tilde{a}_\xi)\chi_\xi\|_{L^2} \\ &\leq Cq^{\frac{d}{2}}\|a_\xi - \tilde{a}_\xi\|_{\ell^2(X)}\end{aligned}$$

Let u denote the exact solution to the PDE. Combining the previous error estimate on u_h and $\tilde{u}_h$, we find

$$\|u - \tilde{u}_h\|_{L^2} \leq \|u - u_h\|_{L^2} + \|u_h - \tilde{u}_h\|_{L^2} \leq Ch_X^k\|u\|_{W_2^k} + Cq^{\frac{d}{2}}\|a_\xi - \tilde{a}_\xi\|_{\ell^2(X)}.$$

If we can estimate the ℓ^2 error above, then we can argue the discrete solution is no worse than the continuous solution. Furthermore, by properly tuning h_Y, q_Y, we can ensure optimal order convergence.

Lemma 17. *Let* $\tilde{A} = A + \Delta A$, $\tilde{a} = a + \Delta a$, *and* $\tilde{b} = b + \Delta b$. *Then,*

$$\|\Delta a\| \leq \|(A + \Delta A)^{-1}\|(\|\Delta A a\| + \|\Delta b\|).$$

We note that $Aa - \tilde{A}\tilde{a} = b - \tilde{b} = -\Delta b$. We then algebraically manipulate

$$\begin{aligned} Aa - \tilde{A}\tilde{a} &= Aa - (A + \Delta A)(a + \Delta a) \\ &= Aa - Aa - A\Delta a - (\Delta A)a - (\Delta A)(\Delta a) \\ &= -(A + \Delta A)(\Delta a) - (\Delta A)a \end{aligned}$$

Solving for Δa yields

$$\begin{aligned} -(A + \Delta A)(\Delta a) &= (\Delta A)a - \Delta b \\ \|\Delta a\| &= \|(A + \Delta A)^{-1}\big((\Delta A)a - \Delta b\big)\| \\ \|\Delta a\| &\leq \|(A + \Delta A)^{-1}\|\,(\|(\Delta A)a\| + \|\Delta b\|) \end{aligned}$$

We have three quantities to estimate: $\|\Delta b\|$, which we estimate using L^2 norms, $\|(A + \Delta A)^{-1}\|$, which we estimate using the reciprocal of the lowest eigenvalue of $\tilde{A}$, and $\|(\Delta A)a\|$, which we estimate using computed error estimates, row sums, and flipping from $\|a\|$ to $\|u\|$ to $\|f\|$.

Lemma 18. *Let* $f \in W_2^k(\mathbb{S}^2)$ *for* $k > \frac{d}{2}$. *Then,*

$$\|\Delta b\| \leq C h_Y^k q_X^{-k} \|f\|_{W_2^k}.$$

Proof. For fixed ξ, we have

$$\begin{aligned}\left|b_\xi - \tilde{b}_\xi\right| &= \left| \int_{\mathbb{S}^2} f(x)\chi_\xi(x)\,dx - \sum_{z\in Y} f(z)\chi_\xi(z) \right| \\ &\leq Ch_Y^k \|f\chi_\xi\|_{W_2^k}\end{aligned}$$

For $k > \frac{d}{2}$, W_2^k is an algebra and $\|f\chi_\xi\|_{W_2^k} \leq \|f\|_{W_2^k}\|\chi_\xi\|_{W_2^k}$. We estimate $\|\chi_\xi\|_{W_2^k} \leq Cq^{\frac{d}{2}-k}$ by comparing it to a worse interpolant, where the C is independent of X or Y. Consequently, we have

$$\left|b_\xi - \tilde{b}_\xi\right| \leq Ch_Y^k q_X^{\frac{d}{2}-k}\|f\|_{W_2^k}. \tag{5.18}$$

We finish by computing $\|\Delta b\|$ as follows

$$\|\Delta b\| = \sqrt{\sum_{\xi\in\Xi} \left|b_\xi - \tilde{b}_\xi\right|^2} \leq Ch_Y^k q^{\frac{d}{2}-k} q^{-\frac{d}{2}} \|f\|_{W_2^k}$$

which follows by noting $\sum_{\xi\in\Xi} 1 \leq Cq^{-d}$. □

Lemma 19. *Let $a \in C^\infty$ and $b \in C^\infty$. Then, we have*

$$\left| \int_{\mathbb{S}^2} a(x)\chi_\xi(x)\chi_\eta(x)\,dx - \sum_{z\in Y} a(z)\chi_\xi(z)\chi_\eta(z)w_z \right| \leq Ch_Y^m q_X^{\frac{d}{2}-m}.$$

Proof. Since $a \in C^\infty$ and $\chi_\xi \in W_2^m$, we have $a\chi_\xi\chi_\eta \in W_2^m$. Consequently, the quadrature error is

$$\left| \int_{\mathbb{S}^2} a(x)\chi_\xi(x)\chi_\eta(x)\,dx - Q_Y(a\chi_\xi\chi_\eta) \right| \leq Ch_Y^m \|a\chi_\xi\chi_\eta\|_{W_2^m}.$$

We estimate $\|a\chi_\xi\chi_\eta\|$ by Theorem B.1 of [25] which asserts

$$\|a\chi_\xi\chi_\eta\|_{W_2^m} \leq C(\|a\|_{L^\infty}\|\chi_\xi\chi_\eta\|_{W_2^m} + \|a\|_{W_2^m}\|\chi_\xi\chi_\eta\|_{L^\infty}).$$

We may absorb the $\|a\|_{L^\infty}$ and $\|a\|_{W_2^m}$ into the constant C. By the boundedness of the Lebesgue constant, we know that $\|\chi_\xi\|_{L^\infty}$ is bounded independent of q, h, ξ, so this term may also be absorbed into the constant. We apply Theorem B.1 again to estimate $\|\chi_\xi\chi_\eta\|_{W_2^m}$ and estimate $\|\chi_\xi\|_{W_2^m} \leq Cq_X^{\frac{d}{2}-m}$ to get

$$\begin{aligned}\|\chi_\xi\chi_\eta\|_{W_2^m} &\leq C(\|\chi_\xi\|_{L^\infty}\|\chi_\eta\|_{W_2^m} + \|\chi_\xi\|_{W_2^m}\|\chi_\eta\|_{L^\infty}) \\ &\leq Cq_X^{\frac{d}{2}-m}.\end{aligned}$$

Combining these estimates with the quadrature error estimate yields the desired result. □

Lemma 20. *Let $b \in C^\infty$. Then,*

$$\left|\int_{\mathbb{S}^2} b(x)\nabla\chi_\xi \cdot \nabla\chi_\eta(x)\,dx - Q_Y(b\nabla\chi_\xi \cdot \nabla\chi_\eta)\right| \leq Ch_Y^{m-1}q_X^{\frac{d}{2}-m-1}.$$

Proof. We proceed similarly by applying the quadrature error estimates and noting that $b\nabla\chi_\xi \cdot \nabla\chi_\eta \in W_2^{m-1}$. Then, we have

$$\left|\int_{\mathbb{S}^d} b(x)\nabla\chi_\xi \cdot \nabla\chi_\eta\,dx - \sum_{z\in Y} b(z)\nabla\chi_\xi \cdot \nabla\chi_\eta(z)\right| \leq Ch_Y^{m-1}\|b\nabla\chi_\xi \cdot \nabla\chi_\eta\|_{W_2^{m-1}}.$$

We estimate this norm by applying Theorem B.1 to get

$$\|b\nabla\chi_\xi \cdot \nabla\chi_\eta\|_{W_2^{m-1}} \leq C\big(\|b\|_{L^\infty}\|\nabla\chi_\xi \cdot \nabla\chi_\eta\|_{W_2^{m-1}} + \|b\|_{W_2^{m-1}}\|\nabla\chi_\xi \cdot \nabla\chi_\eta\|_{L^\infty}\big).$$

We may absorb the $\|b\|_{L^\infty}$ and $\|b\|_{W_2^{m-1}}$ terms into the constant since they are independent of q, h. We bound

$$\begin{aligned}\|\nabla\chi_\xi \cdot \nabla\chi_\eta\|_{L^\infty} &\leq \sup_{x\in\mathbb{S}^d} |\nabla\chi_\xi(x)|\,|\nabla\chi_\eta(x)| \\ &\leq Cq_X^{-2}\exp(-\frac{\nu}{h_X}(d(x,\xi)+d(x,\eta))) \\ &\leq Cq_X^{-2}\exp(-\frac{\nu}{h_X}d(\xi,\eta)).\end{aligned}$$

We apply Corollary B.3 from [25] next to estimate

$$\begin{aligned}\|\nabla\chi_\xi \cdot \nabla\chi_\eta\|_{W_2^{m-1}} &\leq C\big((\|\chi_\xi\|_{W_2^m}+\|\chi_\eta\|_{W_2^m})(\|\nabla\chi_\xi\|_{L^\infty}+\|\nabla\chi_\eta\|_{L^\infty})\big) \\ &\leq Cq_X^{\frac{d}{2}-m}q_X^{-1} = Cq_X^{\frac{d}{2}-m-1}.\end{aligned}$$

Combining this with previous estimates yields the result. □

Lemma 21. $|\tilde{A}_{\xi,\eta}| \leq Cq_X^{-2}\exp(-\nu\frac{d(\xi,\eta)}{h_X})$.

Proof. This follows by noting that (and dropping the non-gradient terms)

$$\begin{aligned}\tilde{A}_{\xi,\eta} &:= \sum_{z\in Y} a(z)\nabla\chi_\xi\cdot\nabla\chi_\eta(z) + b(z)\chi_\xi(z)\chi_\eta(z) \\ &\leq \|a\|_{L^\infty}\|\nabla\chi_\xi\cdot\nabla\chi_\eta(z)\|_{L^\infty}\underbrace{\sum_{z\in Y} w_z}_{C(d)} \\ &\leq Cq_X^{-2}\|a\|_{L^\infty}C(d)\exp(-\nu\frac{d(\xi,\eta)}{h_X})\end{aligned}$$

□

Corollary 4. $|\Delta A_{\xi,\eta}| \leq Cq_X^{-2}\exp(-\nu\frac{d(\xi,\eta)}{h_X})$.

Proof. This follows because the entries in both A and $\tilde{A}$ decay exponentially, and

hence the difference is bounded above by the larger of the two exponential decays. $\square$

Lemma 22. $\|\Delta A\| \leq C h_Y^{m-1} q_X^{\frac{d}{2}-m-1}$.

Proof. We first note that $\|\Delta A\| \leq \sqrt{\|\Delta A\|_1 \|\Delta A\|_\infty}$, where $\|\Delta A\|_1 = \sup_\xi \sum_\eta |A_{\xi,\eta}|$.

Fix ξ. Let $\Omega_k = \{\eta : \Gamma 2^k q_x \leq d(\xi,\eta) \leq \Gamma 2^{k+1} q_x\}$.

$$\sum_\eta |\Delta A_{\xi,\eta}| = \underbrace{\sum_{\eta \in B(\xi, \Gamma q_x)} |\Delta A_{\xi,\eta}|}_{I} + \underbrace{\sum_{k=1}^{N_q} \sum_{\eta \in \Omega_k} |\Delta A_{\xi,\eta}|}_{II}.$$

We estimate I by noting that there are $C q_X^{-d} \Gamma g_x^d = C\Gamma^d$ centers in $B(\xi, \gamma q_x)$, and the quadrature error yields $|\Delta A_{\xi,\eta}| \leq C h_Y^{m-1} q_x^{\frac{d}{2}-m-1}$. Consequently,

$$I \leq C\Gamma^d h_Y^{m-1} q_X^{\frac{d}{2}-m-1}.$$

We note the dependence on Γ. Our strategy will be to show that II can be made arbitrarily small, given appropriately chosen Γ, and hence the error will be dominated by the $h_Y^{m-1} q_X^{\frac{d}{2}-m-1}$.

We first note that $\#\Omega_k \leq C(2^{k+1}\Gamma)^d$ and for each $\eta \in \Omega_k$, $|\Delta A_{\xi,\eta}| \leq C q_X^{-2} \exp(-\nu \Gamma 2^k)$. Consequently,

$$\begin{aligned} II &:= \sum_{k=0}^{N_q} \sum_{\eta \in \Omega_k} |\Delta A_{\xi,\eta}| \\ &\leq C q_X^{-2} \sum_{k=0}^{N_q} (2^{k+1}\Gamma)^d \exp(-\nu \Gamma 2^k). \end{aligned}$$

We note that $2^{k+1} d \leq 2^{d+1} 2^{k(d+1)}$. Therefore, we have

$$II \leq C q_X^{-2} \frac{1}{\Gamma} \sum_{k=0}^{N_q} (2^k \Gamma)^{d+1} \exp(-\nu \Gamma 2^k).$$

By Lemma 3.9 of [12], we may choose Γ such that this term is smaller than $Ch_Y^{m-1}q_X^{\frac{d}{2}-m+1}\Gamma^d$. However, Γ does depend on h_Y, q_X. In particular, we choose Γ such that $\exp(-\nu\frac{\Gamma}{2}) \leq h_Y^{m-1}q_X^{\frac{d}{2}-m+1}$. Informally, Γ does not need to increase very much relative to a decrease in h_Y, as a modest increase in Γ results in a large drop in $\exp(-\nu\frac{\Gamma}{2})$. □

We now aim to prove a "discrete" Bernstein theorem. This result was of interest during the study of the error regarding quadrature error. We present the statement of the result, then prove it after a few preliminary results.

Proposition 17. *Let X be a set of centers with mesh norm and separation radius h_X, q_X respectively and Y a set of centers with mesh norm and separation radius h_Y, q_Y respectively. Let $u = \sum_{\xi\in X} u_\xi\chi_\xi$ and let $\{w_z\}_{z\in Y}$ denote the set of quadrature nodes corresponding to Y. Then, or appropriate conditions on h_Y,*

$$\sum_{z\in Y} w_z|\nabla u(z)|^2 \leq Cq_X^{-2}\|u\|_{L^2}^2.$$

Lemma 23. *Under the same assumptions of Proposition 17,*

$$|\sum_{z\in Y} w_z|\nabla u(z)|^2 - \int |\nabla u(x)|^2\,dx| \leq Ch_Y^k\||\nabla u|^2\|_{W_2^k}.$$

Proof. This follows by applying quadrature error estimates with the appropriate choice of k. □

Lemma 24.

$$\||\nabla u|^2\|_{W_2^k} \leq C\|u\|_{W_2^{k+1}}\|\|\nabla u\|_{L^\infty}$$

Proof. This follows by applying Corollary B.3 from [25]. □

Proof. We now may attempt to prove the discrete Bernstein estimate. We estimate $\sum_{z\in Y} w_z|\nabla u(z)|^2$ by

$$\sum_{z\in Y} w_z|\nabla u(z)|^2 \leq \int_{\mathbb{S}^2} |\nabla u(z)|^2\,dz + \left| \int_{\mathbb{S}^2} |\nabla u(z)|^2\,dz - \sum_{z\in Y} w_z|\nabla u(z)|^2 \right|$$

Applying the Bernstein estimate yields

$$\int_{\mathbb{S}^2} |\nabla u(z)|^2\,dz \leq Cq_X^{-2}\|u\|_{L^2}^2$$

and applying Lemma 23 yields

$$\sum_{z\in Y} w_z|\nabla u(z)|^2 \leq C_1 q_X^{-2}\|u\|_{L^2}^2 + C_2 h_Y^k \||\nabla u|^2\|_{W^k}.$$

We now bound the latter term by arguing that for all sufficiently small h_Y,

$$C_2 h_Y^k \||\nabla u|^2\|_{W_2^k} \leq \frac{1}{2} C_1 q_X^{-2}\|u\|_{L^2}^2.$$

We apply Lemma 24 and then estimate each norm separately. First, $\|u\|_{W_2^{k+1}} \leq Cq_X^{-k-1}\|u\|_{L^2}$ and $\|\nabla u\|_{L^\infty} \leq Cq_X^{-1}\|u\|_{L^\infty}$. By applying the Nikolskii inequality, we have $\|u\|_{L^\infty} \leq Cq^{-\frac{d}{2}}\|u\|_{L^2}$. Consequently, we find

$$\||\nabla u|^2\|_{L^2} \leq 2q_X^{-k-1} q_X^{-\frac{d}{2}-1}\|u\|_{L^2}^2. \tag{5.19}$$

Therefore, we have

$$\sum_{z\in Y} w_z|\nabla u(z)|^2 \leq Ch_Y^{k-1} q_X^{-k-\frac{d}{2}} q_X^{-2}\|u\|_{L^2}^2.$$

□

By choosing h_Y small enough to ensure that $Ch_Y^{k-1}q_X^{-k-\frac{d}{2}} \leq \frac{1}{2}C_1$, we are done.

5.3.1 Numerical Experiments

In this section, we discuss numerical results of various experiments that explore the computational properties of the Galerkin method. We consider different differential operators, explore the effects of the quadrature node density on the L^2 error of the discrete solution, and compute condition numbers for the discrete stiffness matrix. We also demonstrate that local Lagrange functions, as discussed in [7], provide a computationally less expensive approximation space and yield comparable error and condition numbers as the approximation space generated by the Lagrange functions. We choose the spherical basis function $\phi_3(t) = (1-t)^2\log(1-t)$ to construct the approximation space and $\phi_2(t) = (1-t)\log(1-t)$ for the quadrature weights. We use the minimum energy points for the centers X used in the approximation space $V_{\phi_3,X}$. For the quadrature nodes, we use the icosahedral nodes and quasi-minimum energy points. For each experiment, the L^2 error is computed by evaluating the discrete solution on a set of evaluation points E and applying the Lagrange function quadrature rule. The set E is 62500 quasi-minimum energy points, which is used for each experiment independent of X and Y. Let N_X and N_Y denote the number of points in X and Y respectively. We approximate h_Y by $\frac{1}{\sqrt{N_Y}}$.

We first consider the problem $-\Delta u + u = f$ with $u = \exp(\cos(\theta))$ and $f = \exp(\cos(\theta))(\cos^2(\theta) + 2z\cos(\theta))$. In the second and third columns of Table 5.1 we display the relative L^2 errors of the discrete solution for two separate experiments. To obtain the discrete stiffness matrix, we first fixed 961 centers for X and varied the number of quadrature points used in Y The quadrature points are icosahedral nodes with between $2,562$ points to $92,162$ points. We theoretically expect the L^2 error

to be $\mathcal{O}(|\log(h_Y)|^2 h_Y^4)$. In fact, the numerically observed error is $\mathcal{O}(|\log(h_Y)|^2 h_Y^{5.2})$. The experiment was repeated with $N_X = 3721$ minimum energy nodes and using the same Y. This time, ignoring the $N_Y = 2562$ outlier, $|\log(h_Y)|h_Y^{5.5}$ is observed, indicating that improvement in the theoretical errors rates is possible. The Lagrange basis was used for these two sets.

Next, we treated the problem $-\text{div}(\mathbf{a}\cdot\nabla u) + u = f$ for the case in which $\mathbf{a} = a(\theta,\phi)\mathbf{g}$, where $\mathbf{g}$ is the metric tensor for $\mathbb{S}^2$ and $a(\theta,\phi) = 1 - \frac{1}{2}\cos(\theta)$. We again chose $u = \exp(\cos(\theta))$, which results in the right hand side being $f = \big(-\frac{1}{2}(\cos^3(\theta) + \cos^2(\theta) - 5\cos(\theta) + 1) + 1\big)\exp(\cos(\theta))$.

We also consider the possibility of using a local Lagrange basis to discretize the PDE. In this case, the approximation space is $V_X = \text{span}\{\chi_\xi^{loc} \colon \xi \in X\}$, where the χ_ξ^{loc} functions are constructed using only kernels $\phi(\cdot,\eta)$ such that $\text{dist}(\xi,\eta) \leq 7h_X|\log(h_X)|$. See [7] for a detailed description of the theoretical properties of this basis. The χ_ξ^{loc}'s may be constructed in parallel by solving a small linear system. This reduces computational complexity associated with assembling the $\alpha_{\xi,\eta}$ coefficients. By appropriately tuning the number of kernels used per Lagrange function, the local Lagrange function can be made to satisfy $\|\chi_\xi - \chi_\xi^{loc}\|_{L^\infty} \sim h_X^{2m}$, where m is the smoothness of the kernel ϕ. For the anisotropic problem, the fifth and sixth columns in Table 5.1 display the results of the experiment using the local Lagrange bias. For $N_X = 961$, each local Lagrange function is constructed using about 423 centers and for $N_X = 3721$, each local Lagrange function is constructed using around 776 centers, where the number of centers used per kernel is chosen to be all centers with distance at most $7h_X|\log(h_X)|$ from the center. The computed L^2 errors from using the local basis versus the full basis are negligible. Since the local bases offer comparable L^2 error while being computationally simpler, they offer no drawbacks when compared to the full basis and certainly are a good choice for the doing the discretization step.

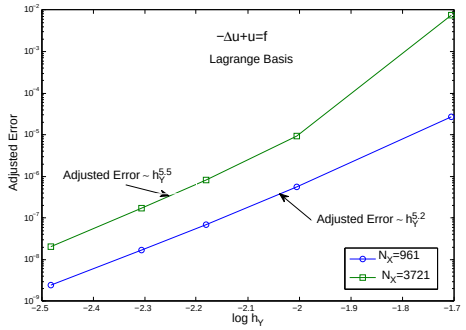
The results of the two experiments are plotted in Figure 5.1(a) and Figure 5.1(b).

A third experiment was conducted keeping Y with fixed and varying X. The result is displayed in Figure 5.1(c). In this experiment, the error *increases* with decreasing h_X. This is counterintuitive, but in complete agreement with the theory. What this illustrates is that the dominant term in the L^2 error comes from quadrature. This is no surprise and is a well-known phenomenon in Galerkin methods.

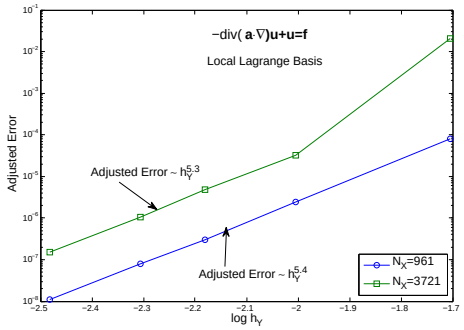
The condition number of the discrete stiffness matrix is dependent primarily on the separation radius of the centers, q_X. We theoretically predicted the condition number to be $\mathcal{O}(q_X^{-2})$, which we validated numerically. See Figure 5.1(d). In addition, the theory predicts that changing the *quadrature nodes* should not significantly alter the condition number of the stiffness matrix. Again, this result was validated.

$-\Delta u + u = f$ Lagrange Basis			$-\mathrm{div}(\mathbf{a} \cdot \nabla u) + u = f$ Local Lagrange Basis		
N_Y	$N_X = 961$	$N_X = 3721$	N_Y	$N_X = 961$	$N_X = 3721$
2562	7.86e-5	2.19e-2	2500	8.00e-5	2.10e-2
10242	2.22e-6	3.76e-5	10000	2.46e-6	3.23e-5
23042	3.34e-7	3.83e-6	22500	3.02e-7	4.78e-6
40962	8.96e-8	9.32e-7	40000	7.80e-8	1.04e-6
92162	1.50e-8	1.27e-7	90000	1.10e-8	1.49e-7

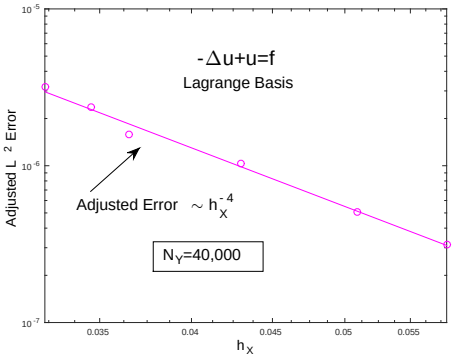
Table 5.1: Both $-\Delta u + u = f$ and $-\mathrm{div}(\mathbf{a} \cdot \nabla u) + u = f$ were numerically solved using minimum energy point sets for X and icosahedral point sets for Y. The L^2 error for all cases was $\mathcal{O}(|\log(h_Y)|^2 h_Y^{5+})$. Here, $h_Y = N_Y^{-1/2}$. For the first equation, a Lagrange basis was used, and, for the second, a *local* Lagrange basis [25].



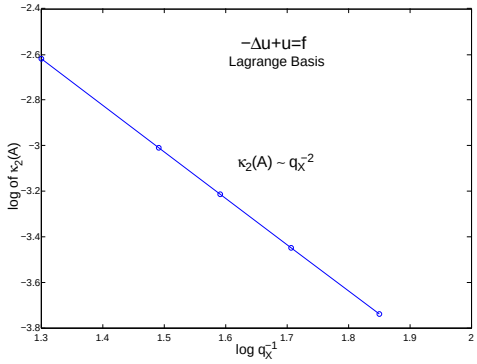
(a) Adjusted L^2 error for $-\Delta u + u = f$.



(b) Adjusted L^2 error for $-\mathrm{div}(\mathbf{a} \cdot \nabla u) + u = f$.



(c) Adjusted L^2 error for Y fixed.



(d) Condition number vs. q_X^{-1}.

Figure 5.1: In (a) and (b), semi-log plots of the errors (adjusted by removing log factors) for $-\Delta u + u = f$ and $-\mathrm{div}(\mathbf{a} \cdot \nabla u) + u = f$ are shown. The minimum energy points were used for X and icosahedral points were used for Y. In (c), a loglog plot of the L^2 error vs. h_X is plotted. For this experiment, the number of quadrature points is fixed and the number of centers used for the approximation space varies. In (d), the log of the condition number for the stiffness matrix for $-\Delta u + u = f$ is plotted [25].

6. RESTRICTED LAGRANGE FUNCTIONS

We present results regarding subsets of Lagrange functions on compact domains in $\mathbb{R}^n$. Numerical experiments suggested that, away from the boundary, Lagrange functions in $\mathbb{R}^n$ for thin plate splines behaved analogously as Lagrange functions on the sphere. Numerical evidence suggests localized spatial behavior and the Lagrange functions have a "small-footprint", i.e., only few kernels contribute significantly to each Lagrange function. This suggested that by making appropriate or simple assumptions, results from the sphere may translate over to Euclidean domains.

However, due to the existence of a boundary, where the Lagrange functions do not behave as they do on the sphere, some choice must be made to handle the boundary. Two obvious options exist: add additional points outside of the domain or restrict and study only Lagrange functions centered sufficiently in the interior of the domain. Both ideas roughly work on the same concept: we are pushing the boundary "away" so the Lagrange functions behave like they do on spheres. If you add additional points, effectively, the boundary has now been pushed out. If you restrict and only use Lagrange functions on the interior, the boundary is now "far away" from your Lagrange functions.

We focus on the choice of studying Lagrange functions centered sufficiently far from the boundary of the domain. This choice was not arbitrary and was spurred by application. The nonlocal diffusion problem of Chapter 4 works in a region where there are effectively two boundaries: an interior boundary on Ω and an exterior boundary outside of $\Omega \cup \Omega_I$. For zero Dirichlet volume constraints, we assume everything in Ω_I is uniformly zero. It is therefore reasonable to work with Lagrange functions centered only in Ω. Consequently, we may place centers all throughout

$\Omega \cup \Omega_I$, but only the Lagrange functions in Ω contribute. Therefore, we do not require additional centers outside of the domain of interest, we just need Lagrange functions centered in Ω. However, by viewing the problem slightly differently, it may occur to the reader that there is no difference between considering Ω the problem domain and Ω_I is where additional centers are added.

Before we begin, we note that these results are not a full theory and were only the beginning. A full and complete theory has been developed by others; see [11] for an excellent complete theory of Lagrange and local Lagrange functions on Euclidean domains. The viewpoint taken in [11] is to add centers outside of the domain, rather than restrict to a subset of Lagrange functions away from the boundary. We note that we have taken advantage of the results of [11] for our nonlocal diffusion work.

Serendipitously, much of the technical work for the results we present here have been developed during previous investigations of Lagrange function properties on spheres. With very simple and small modifications, and additional assumptions, we may replicate results from [13] and [12] by taking advantage of their proof techniques. The results for Lagrange functions on manifolds worked by shifting from the manifold to the tangent plane, which effectively made the problem Euclidean. However, on manifolds, a technical restriction was placed that controlled decay: the radius of injectivity. For our purposes, the radius of injectivity is mirrored as the distance from the center of the Lagrange function to the boundary. After that distance is reached, the decay of the function stops. This matches the result for manifolds.

Let $\Omega \subset \mathbb{R}^d$ and let Ξ be a collection of centers with mesh norm h and separation radius q. Recall that for an open set $U \subset \mathbb{R}^d$, we define the m^{th} Sobolev seminorm:

$$|f|_{W_2^m(U)} := \left(\sum_{|\beta|=m} \int_U |D^\beta f(x)|^2 \, dx \right)^{\frac{1}{2}}.$$

Also recall that the thin plate spline interpolant to a function f satisfies $|I_X f|_{W_2^m(\mathbb{R}^d)} \leq |u|_{W_2^m(\mathbb{R}^d)}$ for all u that satisfy $u|_X = f|_X$.

We require the following lemma that enables us to relate sums of seminorms of order less than m to the order m Sobolev seminorm.

Lemma 25. *[Zeros Lemma [13]] Let $X = \{x_1, ..., x_n\} \subset B(x, r)$ have mesh norm $h = h(X, B(x, r))$ satisfy $h \leq h_0 r$ where $h_0 = \frac{1}{32m^2}$. Assume that $m > \frac{d}{2}$. Then, there is a constant $C_{m,d,2} > 0$ such that*

$$\left(\sum_{k \leq m} r^{2(k-m)} |u|^2_{W_p^k(B(x,r))} \right)^{\frac{1}{2}} \leq C_{m,d,2} |u|_{W_2^m(B(x,r))} \tag{6.1}$$

holds for all $u \in W_2^m(B(x, r))$ vanishing on X. In addition, we have that

$$|u|_{L^\infty(B(x,r))} \leq C_{m,d,2} r^{m-\frac{d}{2}} |u|_{W_2^m(B(x,r))}. \tag{6.2}$$

Our interest is in studying the properties of Lagrange functions centered sufficiently far from the boundary. Let ξ be a center in the interior of Ω with $r_0 := r_0(\xi) = d(\partial\Omega, \xi) > 0$. We denote χ_ξ by χ. We begin by studying a very useful lemma that uses a technique developed by Matveev. Lemma 26 below follows by slight modification to Lemma 4.3 of [13] and [10] Lemma 4.1.

Lemma 26. *If Ξ has mesh norm $h < h_0 \min(1, \frac{1}{3} r_0)$. Then, there exists $\epsilon \in (0, 1)$ such that for $1 < t < \frac{r_0 h_0}{3h}$,*

$$|\chi|_{m, B^c(\xi, 3t\frac{h}{h_0})} \leq \epsilon |\chi|_{m, B^c(\xi, 3(t-1)\frac{h}{h_0})}.$$

Proof. We mimic the proof from [13] Lemma 4.3 and [10] Lemma 4.1. We present the proof in multiple steps.

Step 1: We build a C^∞ cutoff function which is zero on $B(\xi,(3t-1)\frac{h}{h_0})$. and is 1 on $B(\xi,(3t-2)\frac{h}{h_0})$. We define

$$\phi(\alpha)=\sigma(\frac{h_0}{h}d(\alpha,\xi)-3(t-1))$$

where $\sigma(T)$ is 1 for $T<1$ and zero for $T>2$. We verify now that ϕ has the desired property: choose $\alpha\in B(\xi,(3t-2)\frac{h}{h_0})$. Then, $d(\alpha,\xi)\leq(3t-2)\frac{h}{h_0}$, and so $d(\alpha,\xi)-3(t-1)\leq 3t-2-(3(t-1))=1$, so $\sigma(\alpha)=1$. For the case $d(\alpha,\xi)>(3t-1)\frac{h}{h_0}$, we have $d(\alpha,\xi)-3(t-1)\geq(3t-1)-3(t-1)=2$, and hence $\sigma(\alpha)=0$. The cutoff ϕ is nonconstant on an annulus $B(\xi,(3t-1)\frac{h}{h_0})\backslash B(\xi,(3t-2)\frac{h}{h_0})$.

Let $A(\alpha,t,r)$ denote the annulus $B(\alpha,tr)\backslash B(\alpha,(t-1)r)$. Let B_1 denote the region where ϕ equals one, i.e.,

$$B_1=B\left(\xi,(3t-2)\frac{h}{h_0}\right)$$

and let B_1^c denote its complement. We let A_1 denote the annulus on which ϕ is nonconstant

$$A_1=A\left(\xi,3t-1,\frac{h}{h_0}\right)$$

and we note that the support of ϕ is $A_1\cup B_1$.

By the variational property of χ, the Sobolev seminorm of χ is less than that of $\phi\chi$, since $\phi\chi|_\Xi=\chi|_\Xi$. Furthermore, since ϕ has compact support, its Sobolev seminorm is determined by its values on $A_1\cup B_1$. That is, we have

$$|\chi|_m^2\leq|\phi\chi|_m^2=|\phi\chi|_{m,B_1}^2+|\phi\chi|_{m,A_1}^2=|\chi|_{m,B_1}^2+|\phi\chi|_{m,A_1}^2,$$

where we used the fact that $\phi\chi=\chi$ on B_1. Therefore, by subtracting $|\chi|_{m,B_1}^2$, we can

bound the Sobolev seminorm of χ outside of B_1 by the norm of $\phi\chi$ on the annulus A_1. That is,

$$|\chi|^2_{m,B_1^c} \leq |\phi\chi|^2_{m,A_1}.$$

Our goal is to estimate the Sobolev norm of $\phi\chi$ on this annulus. We proceed in two stages. First, we apply the Leibniz rule to bound the Sobolev norm on the annulus by seminorms of χ, and then we use a lemma by breaking up the annulus into balls of a certain radius.

Estimate 1: We first try to estimate $|\phi\chi|^2_{m,A_1}$ by the Sobolev norm of χ on A_1. We first need an estimate on the value of the derivative of ϕ. We note that

$$|D^\alpha\phi(x)| = |D^\alpha\sigma\left(\frac{h_0}{h}d(\xi,\alpha) - 3(t-1)\right)| \leq \left(\frac{h_0}{h}\right)^\alpha |D^\alpha\sigma|_\infty = C(\alpha)\left(\frac{h}{h_0}\right)^{-\alpha}.$$

By applying the product rule, we compute,

$$\begin{aligned}
|\phi\chi|^2_{m,A_1} &= \sum_{|\alpha|=m}\int_{A_1}\frac{m!}{\alpha!}|D^\alpha(\phi\chi)|^2\,dx \\
&= \sum_{|\alpha|=m}\int_{A_1}\frac{m!}{\alpha!}\left|\sum_{\beta\leq\alpha}\binom{\alpha}{\beta}D^\beta\chi D^{\alpha-\beta}\phi\right|^2 dx \\
&\leq C\sum_{\alpha\leq m}\sum_{\beta\leq\alpha}\int_{A_1}\left|D^\beta\chi(x)D^{\alpha-\beta}\phi(x)\right|^2 dx.
\end{aligned}$$

We now use our estimate on $D^{\alpha-\beta}\phi(x)$ to leave only derivatives of χ inside the above expression:

$$|\phi\chi|^2_{m,A_1} \leq C\sum_{|\alpha|\leq m}\sum_{\beta\leq\alpha}\left(\frac{h}{h_0}\right)^{2(|\beta|-|\alpha|)}\int_{A_1}|D^\beta\chi(x)|^2\,dx.$$

Now, for each $k \leq m$, we can combine the $|\beta| = k$ terms to form $W_2^k(A_1)$ seminorms,

which yields

$$|\phi\chi|^2_{m,A_1} \leq C \sum_{k\leq m} \left(\frac{h}{h_0}\right)^{2(k-m)} |\chi|_{W_2^k(A_1)}.$$

Estimate 2: We want to apply Lemma 25, so we decompose the annulus into a sequence of balls $(B_j)_{j\in J}$ with the conditions that

1. Each ball is of radius $\frac{h}{h_0}$.

2. Each ball has its center in A_1, so $B_j \subset A_2$ with

$$A_2 := A(\xi, t, 3\frac{h}{h_0}).$$

3. Every $x \in A_2$ is in at most N_d balls of B_j, with N_d depending only on the spatial dimension and not on t or h.

Notice that $\frac{h}{h_0} = r$ satisfies the condition of Lemma 25 on each ball B_j, since the mesh norm on the ball is at most h. Therefore, we can relate the W_2^k seminorm on the ball to the W_2^m seminorms on the ball. Then, by summing over the seminorms of each ball, we can relate this sum of seminorms to a seminorm on the large annulus A_2.

$$\begin{aligned}
\sum_{k\leq m} \left(\frac{h}{h_0}\right)^{2(k-m)} |\chi|^2_{W_2^k(A_1)} &\leq \sum_{k\leq m}\sum_{j\in J} \left(\frac{h}{h_0}\right)^{2(k-m)} |\chi|^2_{W_2^k(B_j)} \\
&\leq C \sum_{j\in J} |\chi|^2_{W_2^m(B_j)} \\
&\leq C|\chi|^2_{W_2^m(A_2)}.
\end{aligned}$$

Therefore, we have shown $|\phi\chi|^2_{W_2^m(A_1)} \leq |\chi|^2_{W_2^m(A_2)}$. So far, we have shown that

$$|\chi|^2_{W_2^m(B_1^c)} \leq |\phi\chi|^2_{W_2^m(A_1)} \leq C|\chi|^2_{W_2^m(A_2)}.$$

That is, the seminorm of χ outside of a ball is bounded above by the seminorm on just an annulus. We now relate this to the complements of concentric balls. Recall that $B_1 = B(\xi, (3t-2)\frac{h}{h_0})$, and $B_1 \subset B(\xi, 3t\frac{h}{h_0}) := B^+$, and hence $(B^+)^c \subset B_1^c$. Therefore, we bound the seminorms on the complements of these balls. By setting $B^- = B(\xi, 3(t-1)\frac{h}{h_0})$, we note that $A(\xi, t, 3\frac{h}{h_0}) = (B^-)^c \backslash (B^+)^c$. With this information, we know

$$\begin{aligned}
|\chi|^2_{W_2^m((B^+)^C)} &\leq |\chi|^2_{W_2^m(B_1^c)} \\
&\leq K|\chi|^2_{W_2^m(A_2)} \\
&\leq K(|\chi|^2_{W_2^m((B^-)^c)} - |\chi|^2_{W_2^m((B^+)^c)}
\end{aligned}$$

Algebraic manipulation of this expression yields

$$(K+1)|\chi|^2_{W_2^m(B(\xi,3t\frac{h}{h_0}))} \leq K|\chi|^2_{W_2^m(B(\xi,3(t-1)\frac{h}{h_0}))}.$$

Dividing by $K+1$ and taking square roots yields (with $\epsilon = \sqrt{\frac{K}{K+1}}$)

$$|\chi|_{W_2^m(B\left(\xi,3t\frac{h}{h_0}\right))} \leq \epsilon|\chi|_{W_2^m(B\left(\xi,3(t-1)\frac{h}{h_0}\right))}.$$

□

One of the key tricks to the above lemma was being able to apply the Zeros lemma Lemma 25. We note that the ball must be contained within Ω, which forces

us to relate t to r_0 in the above lemma. If the ball in consideration goes outside of Ω, then there are no centers in that region; consequently, the mesh norm of the ball, h_B, grows larger than h. Therefore, we would not be able to apply the Zeros lemma. Therefore, the restriction we place that relates the radius of the ball to the distance to the boundary, r_0, is required.

Our next step is to establish an exponential decay result for Sobolev seminorms on complements of balls of a chosen radius T. We note that, as is the theme present in this section, the decay is limited until the distance to the boundary r_0 is reached.

Corollary 5. *There is a constant $\nu > 0$ such that if Ξ has mesh norm $h < h_0 \min(1, \frac{r_0}{3})$, then for $T_0 \leq T < r_0$,*

$$|\chi|_{W_2^m(B^c(\xi,T))} \leq C q^{\frac{d}{2}-m} \exp\bigg(- \nu \frac{T}{h}\bigg).$$

Proof. Set $T = 3t\frac{h}{h_0} = 3n\frac{h}{h_0} + r$ with $0 \leq r < 3\frac{h}{h_0}$, where n represents the number of times we may iterate the Lemma 26. Then,

$$|\chi|_{m,B^c(\xi,T)} \leq \epsilon |\chi|_{W_2^m(B^c(\xi,3(t-1)\frac{h}{h_0}))}. \tag{6.3}$$

By redefining $t' = t-1$, we may iterate the lemma exactly n times, and we note that $n > t-1$. Furthermore, $t = \frac{h_0}{3}\frac{T}{h}$. Therefore,

$$\begin{aligned} |\chi|_{m,B^c(\xi,T)} &\leq \epsilon^n |\chi|_{W_2^m(B^c(\xi,3(t-n)))} \\ &\leq \frac{1}{\epsilon}\epsilon^t |\chi|_{W_2^m(\mathbb{R}^d)} \\ &\leq \frac{1}{\epsilon}\big(\epsilon^{\frac{h_0}{3}}\big)^{\frac{T}{h}} |\chi|_{W_2^m(\mathbb{R}^d)}. \end{aligned}$$

By setting $\nu = |\frac{h_0}{3}\log(\epsilon)|$, we get

$$|\chi|_{W_2^m(B(\xi,T))} \leq C\exp\left(-\nu\frac{T}{h}\right)|\chi|_{W_2^m(\mathbb{R}^d)}.$$

Our last step is to approximate the seminorm of χ. Recall that in the Sobolev seminorm, χ has minimal norm of all functions vanishing on Ξ. Therefore, if we compare it with the cutoff function $\sigma(\frac{x-\xi}{q})$, which vanishes on Ξ, but takes value 1 at ξ, we know

$$|\chi|_{W_2^m(\mathbb{R}^d)} \leq \left|\sigma\left(\frac{\cdot-\xi}{q}\right)\right|_{W_2^m(\mathbb{R}^d)}.$$

By the chain rule, we may compare this with the Sobolev seminorm of $q^{\frac{d}{2}-m}|\sigma|_{W_2^m(\mathbb{R}^d)}$. The Sobolev seminorm of σ is independent of q, h, X, and is some constant. Therefore, we have

$$|\chi|_{W_2^m(B^c(\xi,T))} \leq Cq^{\frac{d}{2}-m}\exp\left(-\nu\frac{T}{h}\right).$$

□

We next aim to establish a pointwise spatial decay result for a Lagrange function centered at ξ, a center in the interior of Ω. We maintain the assumption that $r_0(\xi) := r_0 = d(\xi, \partial\Omega) > 0$. Furthermore, r_0 depends on ξ, and hence each ξ will have a different region of decay. This distinction is imperative: each Lagrange function will be different, and no statement is being made about Lagrange functions such that $d(\xi, \partial\Omega) = 0$. On the other hand, for points ξ centered such that $d(\xi, \partial\Omega) > 0$, given h sufficiently small, a decay can still be demonstrated. The spatial decay result is inspired heavily by [13] Proposition 4.5 and [10] Proposition 4.5.

Theorem 7. *Let χ_ξ (henceforth denoted as χ) be centered at ξ with $d(\xi, \partial\Omega) = r_0 >$*

0. *If Ξ has mesh norm $h < h_0 \min(1, \frac{r_0}{2})$, then for each $x \in \Omega$:*

$$|\chi(x)| \leq \begin{cases} C \exp(-\nu \frac{d(x,\xi)}{h}) & d(x,\xi) < r_0 \\ C \exp(-\nu \frac{r_0}{h}) & d(x,\xi) \geq r_0. \end{cases} \tag{6.4}$$

Proof. We remind the reader that this result is only being shown for $x \in \Omega$. We split the estimate into three pieces. We first estimate a bound independent of h or q for nearby points $d(x,\xi) < \frac{h}{h_0}$. Then, we estimate for points with distance $\frac{h}{h_0} < d(x,\xi) < r_0$. Lastly, we handle $x \in \Omega$ such that $d(x,\xi) \geq r_0$.

Nearby Points: Let $d(x,\xi) < \frac{h}{h_0}$. Note that the ball $B(x, 2\frac{h}{h_0}) \subset \Omega$, and hence its mesh norm is bounded above by $2h$. Therefore, one can apply Lemma 25. By (6.2), we have

$$|\chi(x)| \leq C \left(\frac{h}{h_0}\right)^{m-\frac{d}{2}} |\chi|_{W_2^m(B(x,2\frac{h}{h_0}))}.$$

We may bound the Sobolev seminorm contained above by $q^{\frac{d}{2}-m}$ as remarked earlier, which yields a bound $|\chi(x)| \leq C\left(\frac{h}{q}\right)^{m-\frac{d}{2}}$. Let $C_1 = C \exp(\frac{\nu}{h_0})$. We note that since $d(x,\xi) < \frac{h}{h_0}$, we may bound

$$|\chi(x)| \leq C_1 \left(\frac{h}{q}\right)^{m-\frac{d}{2}} \exp\left(-\frac{\nu}{h_0}\right) \leq C_1 \left(\frac{h}{q}\right)^{m-\frac{d}{2}} \exp\left(-\nu \frac{d(x,\xi)}{h}\right).$$

We note that this is only being done on the nearby points such that $d(\xi, x) < \frac{h}{h_0}$, and the ν term should be independent of h, q, and X.

Intermediate Points: Let $\frac{h}{h_0} < d(x,\xi) < r_0$. Then, by applying (6.2) to the ball $B(x, \frac{h}{h_0})$, we have $|\chi(x)| \leq Ch^{m-\frac{d}{2}} |\chi|_{W_2^m(B(x,\frac{h}{h_0}))}$. Let $T = d(x,\xi)$. Then, $B(x, \frac{h}{h_0}) \subset B^c(\xi, T - \frac{h}{h_0})$. Then, with $T' = T - \frac{h}{h_0}$, we may apply Corollary 5 which

yields

$$\begin{aligned}|\chi(x)| &\leq C\left(\frac{h}{q}\right)^{m-\frac{d}{2}} \exp(-\nu\frac{T'}{h}) \\ &\leq C\left(\frac{h}{q}\right)^{m-\frac{d}{2}} \exp\left(\frac{\nu}{h_0}\right) \exp\left(-\nu\frac{T}{h}\right) \\ &= C_1\left(\frac{h}{q}\right)^{m-\frac{d}{2}} \exp\left(-\nu\frac{d(x,\xi)}{h}\right).\end{aligned}$$

Distant Points: Assume $d(x,\xi) > r_0$, (but recall that $x \in \Omega$). Then, the ball $B(x, \frac{h}{h_0}) \subset B^c(\xi, r_0 - \frac{h}{h_0})$. Hence, $|\chi|_{W_2^m(B(x,\frac{h}{h_0}))} \leq |\chi|_{W_2^m(B^c(\xi,r_0-\frac{h}{h_0}))}$. Furthermore, $B(x, \frac{h}{h_0})$ satisfies r mesh norm at most $2h$. Therefore, we can apply (6.2) to get

$$|\chi(x)| \leq C\left(\frac{h}{h_0}\right)^{m-\frac{d}{2}} |\chi|_{W_2^m(B(x,\frac{h}{h_0}))}.$$

As in the previous case, we begin estimating again:

$$\begin{aligned}|\chi(x)| &\leq C\left(\frac{h}{h_0}\right)^{m-\frac{d}{2}} |\chi|_{W_2^m(B(x,\frac{h}{h_0}))} \\ &\leq C\left(\frac{h}{h_0}\right)^{m-\frac{d}{2}} |\chi|_{W_2^m(B^c(\xi,r_0-\frac{h}{h_0}))} \\ &\leq C\left(\frac{h}{q}\right)^{m-\frac{d}{2}} \exp\left(-\nu\frac{r_0-\frac{h}{h_0}}{h}\right) \\ &\leq C_1\left(\frac{h}{q}\right)^{m-\frac{d}{2}} \exp\left(-\nu\frac{r_0}{h}\right).\end{aligned}$$

□

We have so far demonstrated that a Lagrange function centered in the interior of Ω exponentially decays, at least in a ball of radius r_0, which is analogous to a decay result on the sphere. Now that we have some form of decay results, we aim to

translate results from the sphere on the Lebesgue constant, stability of the Lagrange basis, and decay of the coefficients to $\mathbb{R}^n$. As we have seen, for Lagrange functions on the boundary, there is no hope of translating decay ideas, and hence unlikely the other results follow. However, an approach to consider is a 'restricted' version of the above results. Let $\Omega \subset \tilde{\Omega}$. By choosing all centers ξ such that $d(\xi, \partial\tilde{\Omega}) > R_0$ for some fixed R_0, perhaps we can translate some results. We note in particular that the above decay result is afflicted by its dependency on $r_0(\xi)$, the distance from ξ to the boundary of Ω. However, by using the assumption of a minimal such R_0, we know Lagrange functions decay to at lowest $\exp(-\nu\frac{R_0}{h})$.

We note the similarity to the case of manifolds with a radius of injectivity $r_{\mathbb{M}}$. In the case we consider here, $r_0(\xi)$ acts as a replacement for $r_{\mathbb{M}}$ for each ξ. By restricting to a set of Lagrange functions within Ω with a guaranteed distance $d(\xi, \partial\tilde{\Omega}) = R_0$, we effectively replace $r_{\mathbb{M}}$ with R_0. We proceed by doing exactly that in our following computations. The next result is analogous to [13] Theorem 4.6, and we make parallel use of their techniques to verify the result for our situation.

Proposition 18. *Let $\Omega \subset \tilde{\Omega}$ satisfy $B(x, R_0) \subset \tilde{\Omega}$ for all $x \in \Omega$. Let Ξ be a set of centers with mesh norm h defined in $\tilde{\Omega}$. Let $\Xi' = \Xi \cap \Omega$. Let $\mathcal{L}_{\Omega,\tilde{\Omega},X} = \sup_{x\in\tilde{\Omega}} \sum_{\xi\in\Xi'} |\chi_\xi(x)|$. Then, $\mathcal{L}$ is bounded independent of h and q.*

Proof. We emphasize that we are only using Lagrange functions centered in Ω, which satisfy a minimal decay of the form

$$|\chi_\xi(x)| \leq \begin{cases} C\exp(-\nu\frac{d(x,\xi)}{h}) & d(x,\xi) < R_0 \\ C\exp(-\nu\frac{R_0}{h}) & d(x,\xi) \geq R_0 \end{cases}$$

where R_0 is the radius such that $B(\xi, R_0) \subset \tilde{\Omega}$ for all $\xi \in \Omega$.

Fix $x \in \Omega$. We need to estimate $\sum_{\xi \in \Xi'} |\chi_\xi(x)|$. We estimate this quantity by decomposing the summation into a set of nearby Lagrange functions and distant Lagrange functions. Let $\Upsilon_x = B(x, R_0) \cap X$. Then

$$\underbrace{\sum_{d(\xi,x)<R_0} |\chi_\xi(x)|}_{I} + \underbrace{\sum_{d(\xi,x)\geq R_0} |\chi_\xi(x)|}_{II}.$$

Estimate on II: We make the unrefined estimate of

$$|\chi_\xi(x)| \leq C \exp\left(- \nu \frac{R_0}{h} \right)$$

for each ξ. We note that this is a worst possible estimate, as many of these functions likely decay far better (we are not taking into account $r_0(\xi)$). Regardless, we may then estimate how many such centers are contained within this region. The number of centers is bounded above by $C\mu(\Omega \cap B(x, R_0)^c) \cdot q^{-d}$. By invoking quasi-uniformity, $q^{-d} = \rho^d h^{-d}$. Then, we note

$$\sum_{d(\xi,x)\geq R_0} |\chi_\xi(x)| \leq C\rho^d h^{-d} \exp(-\nu \frac{R_0}{h}).$$

We note that the function $x^{-d} \exp(-\alpha \frac{1}{x})$ is bounded above independent of x. This follows because the function $f(x) = x^{-d} \exp(-\frac{a}{x})$ has a critical point at $x = \frac{a}{d}$ with max value $f(x') = (\frac{a}{d})^{-d} \exp(-d)$ for $x > 0$. On the interval $(0, x')$ the function is increasing, but bounded above by $f(x')$.

Estimating I: We estimate entries in I by breaking up the sum into annuli and using the exponential decay.

First, consider the ball $B(x, h)$. On this ball the area is bounded above by h^d, and the number of centers in this region can be bounded by the product of the area

of the region times q^{-d}. This yields ρ^d centers on $B(x,h)$. For any $\xi \in B(x,h)$, we can at worst say

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

Therefore, on on the set $B(x,h)$, we estimate:

$$\sum_{\xi \in B(x,h)} |\chi_\xi(x)| \leq C\rho^{m-\frac{d}{2}} \rho^d$$

which is bounded independent of h or q (by invoking quasi-uniformity of the centers). Next, we decompose $B(x,R_0) = B(x,h) \cup (\cup_{n=1}^N A_n)$ where A_n is the annulus

$$A_n = B(x,(n+1)h) \backslash B(x,nh).$$

We estimate the number of centers on A_n to be $q^{-d}(n+1)h^d = \rho^d(n+1)$. On this annulus, we know a minimal distance of $d(x,\xi) \geq nh$, and hence

$$|\chi_\xi(x)| \leq C\exp(-\nu\frac{nh}{h}) = C\exp(-\nu n).$$

Therefore, we compute

$$\sum_{\xi \in A_n} |\chi_\xi(x)| \leq C\rho^d(n+1)\exp(-\nu n).$$

Then, summing over each annulus, we compute

$$\sum_{\xi \in B(x,R_0} |\chi_\xi(x)| \leq \sum_{n=1}^N C(n+1)\exp(-\nu n) \leq K.$$

for some fixed constant K independent of h, q. Therefore, our estimate on I and II are both bounded, and we achieve the desired result. $\square$

We reiterate the similarity to the sphere: by adding enough points (the set $\tilde{\Omega}$), and choosing sufficiently small h, we are guaranteed decay for all interior Lagrange functions. In particular, we now see that the ideas from the sphere translate almost directly. The next direction will be to seek L^2 stability of the Lagrange basis. That is, we seek to show that for a sequence of scalars α_i and Lagrange functions contained in Ω (with some minimal guaranteed distance $d(\xi, \partial\tilde{\Omega}) = R_0$) that there exists C_1, C_2 such that

$$C_1 q^{\frac{d}{2}} \|\{\alpha\}_{i=1}^n\|_{\ell^2} \leq \|\sum_{\xi\in\Omega} \alpha_\xi \chi_\xi\|_{L^2(\Omega)} \leq C_2 q^{\frac{d}{2}} \|\{\alpha\}_{i=1}^n\|_{\ell^2}.$$

We first study continuity properties of the Lagrange functions. We make use of a continuity result from [12] that was invaluable for many proofs on manifolds. The next lemma is Lemma 7.1 of [12].

Lemma 27. *Let $\epsilon > 0$, $m > \frac{d}{2} + \epsilon$, $r > 0$, and suppose the finite set of points $\Xi \subset B(0,r)$ is sufficiently dense so that $h := \max_{x\in B(0,r)} dist(x, \Xi) \leq h_0 r$. Then, there is a constant so that every $f \in W_2^m(B(0,r))$ which vanishes on Ξ and $y \in B(0,r)$ satisfy*

$$|f(y) - f(0)| \leq C\|y\|^\epsilon |r^{m-\frac{d}{2}-\epsilon} \left(\sum_{|\alpha|=m} \int_{B(0,r)} |D^\alpha f(x)|^2 \, dx \right)^{\frac{1}{2}}.$$

Corollary 6. *Let $\epsilon > 0$, $m > \frac{d}{2} + \epsilon$, $r > 0$, and suppose the finite set of points $\Xi \subset B(x,r)$ is sufficiently dense so that $h := \max_{y\in B(x,r)} dist(y, \Xi) \leq h_0 r$. Then, there is a constant so that every $f \in W_2^m(B(x,r))$ and $y \in B(x,r)$ satisfy*

$$|f(y) - f(x)| \leq C\|y\|^\epsilon |r^{m-\frac{d}{2}-\epsilon} \left(\sum_{|\alpha|=m} \int_{B(x,r)} |D^\alpha f(z)|^2 \, dz \right)^{\frac{1}{2}}.$$

Proof. For fixed x, we define $u : B(0,r) \to \mathbb{R}$ by $u(z-x) = f(z)$. $\square$

Proposition 19. *Let $0 < \epsilon \leq 1$, $R_0 > 0$, and let $\Omega \subset \tilde{\Omega} \subset \mathbb{R}^d$ satisfy $B(x, R_0) \subset \tilde{\Omega}$ for all $x \in \Omega$. For $m > \frac{d}{2} + \epsilon$, there is a constant C depending only on $m, \mathbb{R}_0, \Omega, \tilde{\Omega}, \rho$ such that*

$$|\chi_\xi(x) - \chi_\xi(y)| \leq C\left[\frac{\|x-y\|}{q}\right]^\epsilon.$$

Proof. Consider $x, y \in \tilde{\Omega}$. We first start by considering pairs x, y such that $d(x,y) \leq h$. By the exponential decay result, we know for $0 \leq T < R_0$, $\|\chi\|_{W_2^m(B^c(\xi,T))} \leq Cq^{\frac{d}{2}-m}\exp(-\nu\frac{T}{h})$. For the case $T \geq R_0$, we replace T in the previous equation with R_0.

Nearby Points: Consider points x such that $d(x,\xi) \leq \frac{h}{h_0}$. We apply Corollary 6 with $r = \frac{2h}{h_0}$ to see that

$$|\chi_\xi(x) - \chi_\xi(y)| \leq C\|x-y\|^\epsilon \left(\frac{h}{h_0}\right)^{m-\epsilon-\frac{d}{2}} |\chi_\xi|_{W_2^m(B(x,\frac{2h}{h_0}))}.$$

We may bound the seminorm by $|\chi_\xi|_{W_2^m(B(x,\frac{2h}{h_0}))} \leq Cq^{\frac{d}{2}-m}$, which yields the bound

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

For quasi-uniformly distributed sets of centers, the ratio $\frac{h}{q} = \rho$ is bounded independent of h and q, and hence the result follows for the nearby points.

Intermediate Points: Consider points x such that $\frac{h}{h_0} \leq \|x-\xi\| \leq R_0$. Note that $B(x, \frac{h}{h_0}) \subset B^c(\xi, T - \frac{h}{h_0})$. Then, by applying Corollary 6 on $B(x, \frac{h}{h_0})$, we

compute

$$\begin{aligned}
|\chi_\xi(x) - \chi_\xi(y)| &\leq Ch^{m-\epsilon-\frac{d}{2}}\|x-y\|^\epsilon |\chi_\xi|_{W_2^m(B(x,\frac{h}{h_0}))} \\
&\leq Ch^{m-\epsilon-\frac{d}{2}}\|x-y\|^\epsilon |\chi_\xi|_{W_2^m(B^c(\xi,T-\frac{h}{h_0}))} \\
&\leq Ch^{m-\epsilon-\frac{d}{2}}\|x-y\|^\epsilon q^{m-\frac{d}{2}}\exp(-\nu\frac{T}{h}) \\
&\leq C\|x-y\|^\epsilon q^{-\epsilon}\exp(-\nu\frac{T}{h}).
\end{aligned}$$

where we applied quasi-uniformity to relate $h^{-\epsilon} \sim q^{-\epsilon}$ and used Corollary 5.

Distant Points: Consider x such that $\|x-\xi\| \geq R_0$. We repeat the previous method by noting that $B(x,\frac{h}{h_0}) \subset B^c(\xi, R_0 - \frac{h}{h_0})$. This leads to

$$\begin{aligned}
|\chi_\xi(x) - \chi_\xi(y)| &\leq Ch^{m-\epsilon-\frac{d}{2}}\|x-y\|^\epsilon |\chi_\xi|_{W_2^m(B^c(\xi,R_0-\frac{h}{h_0}))} \\
&\leq Ch^{m-\epsilon-\frac{d}{2}}\|x-y\|^\epsilon q^{\frac{d}{2}-m}\exp(-\nu\frac{R_0}{h}) \\
&\leq Cq^{-\epsilon}\|x-y\|^\epsilon \exp(-\nu\frac{R_0}{h})
\end{aligned}$$

We now consider the case $d(x,y) > h$. Let $M = \sup_{d(x,y)>h} |\chi_\xi(x) - \chi_\xi(y)|$. We note that $M \leq 2\sup_x |\chi_\xi(x)| \leq 2\mathcal{L}_{\Omega,\tilde{\Omega}}$. which is bounded independent of h, q. Then, we let $\tilde{C} = \max(C, \frac{M}{\rho^\epsilon})$. Then, for points $d(x,y) > h$, we have

$$\begin{aligned}
|\chi_\xi(x) - \chi_\xi(y)| \leq M &\leq M\frac{|x-y|^\epsilon}{h^\epsilon} \\
&\leq M\frac{q^{-\epsilon}}{h^\epsilon q^{-\epsilon}}|x-y|^\epsilon \\
&\leq \frac{M}{\rho^\epsilon}q^{-\epsilon}|x-y|^\epsilon \leq \tilde{C}q^{-\epsilon}|x-y|^\epsilon.
\end{aligned}$$

□

So far, our work is focused on studying subsets of Lagrange functions centered

sufficiently far from the boundary. For shorthand, we say that a subset $\Omega \subset \tilde{\Omega}$ is r embedded if and only if for each $x \in \Omega$, $B(x, r) \subset \tilde{\Omega}$.

Corollary 7. *Let Ω be R_0 embedded in Ω and let $\Xi = X \cap \Omega$. Then, for each $\xi \in \Xi$, there exists C such that*

$$|\nabla \chi_\xi(x)| \leq \begin{cases} Cq^{-1} \exp(-\nu \frac{d(x,\xi)}{h} & d(x,\xi) \leq R_0) \\ Cq^{-1} \exp(-\nu \frac{R_0}{h}) & d(x,\xi) \geq R_0 \end{cases}$$

Proof. The directional derivative in the direction $\vec{t}$ with $\|\vec{t}\| = 1$ is computed by

$$D_{\vec{t}} \chi_\xi(x) = \lim_{\tau \to 0} \frac{\chi_\xi(x + \tau\vec{t}) - \chi_\xi(x)}{\tau}.$$

For $d(x, y) \leq h$, we know

$$|\chi_\xi(x) - \chi_\xi(y)| \leq Cq^{-1} \|x - y\| \exp\bigg(- \nu \frac{d(x,\xi)}{h} \bigg).$$

By setting $y = x + \tau\vec{t}$, we have

$$|\chi_\xi(x + \tau\vec{t}) - \chi_\xi(x)| \leq Cq^{-1} \tau \exp\bigg(- \nu \frac{\tau}{h} \bigg),$$

and hence by dividing by τ, we have

$$\left| \frac{\chi_\xi(x) - \chi_\xi(x + \tau\vec{t})}{\tau} \right| \leq Cq^{-1} \exp(-\nu \frac{d(x,\xi)}{h}).$$

Since this holds for all $\tau \leq h$, we have

$$|D_{\vec{t}} \chi_\xi(x)| \leq Cq^{-1} \exp\bigg(- \nu \frac{d(x,\xi)}{h} \bigg).$$

By noting that $D_{\vec{t}}(x) = \nabla\chi(x) \cdot \vec{t}$, we may bound $\|\nabla\chi(x)\|$ by choosing $\vec{t} = \frac{\nabla\chi(x)}{\|\nabla\chi(x)\|}$, we see

$$D_{\vec{t}}(x) = \frac{\nabla\chi(x) \cdot \nabla\chi(x)}{\|\nabla\chi(x)\|} \leq Cq^{-1} \exp\left(- \nu \frac{d(x,\xi)}{h} \right).$$

In particular, we can get partial derivative decay estimates. By choosing $\vec{t} = e_i$, the unit vector, we get $|\frac{\partial\chi_\xi}{\partial x_i}(x)| \leq Cq^{-1} \exp(-\nu\frac{d(x,\xi)}{h})$. □

Our next objective is to establish the equivalence of the L^p norm of a linear combination of Lagrange functions with the ℓ^p norm of the sequence of coefficients on the linear combination. We require an upper and lower estimate. The upper estimate we may prove by using operator interpolation techniques and the boundedness of the restricted Lebesgue constant.

Lemma 28. *Let X be a collection of quasi-uniformly scattered centers in $\tilde{\Omega}$ and let $\Omega \subset \tilde{\Omega}$ such that $B(x, R_0) \subset \tilde{\Omega}$ for all $x \in \Omega$. Let $s = \sum_{\xi\in\Omega} \alpha_\xi \chi_\xi$. Then*

$$\|s\|_{L^\infty(\tilde{\Omega})} \leq C\|\alpha\|_{\ell^\infty}.$$

Proof. We note that for any $x \in \Omega$

$$\begin{aligned} |s(x)| = |\sum_{\xi\in\Omega} \alpha_\xi \chi_\xi(x)| &\leq \sum_{\xi\in\Omega} |\alpha_\xi| \, |\chi_\xi(x)| \\ &\leq \|\alpha\|_{\ell^\infty} \sum_{\xi\in\Omega} |\chi_\xi(x)| \leq \mathcal{L}_{\Omega,\tilde{\Omega}} \|\alpha\|_{\ell^\infty}. \end{aligned}$$

We note that the restricted Lebesgue constant $\mathcal{L}_{\Omega,\tilde{\Omega}}$ is bounded for fixed R_0 independent of h and q. The result only holds for linear combinations of Lagrange functions centered within Ω. □

Lemma 29. *Under the same assumptions as Lemma 28, for any $\xi \in \Omega$, there exists*

constant C independent of h, q such that

$$\int_{\tilde{\Omega}} |\chi_\xi(x)| dx \leq Cq^d.$$

Proof. We use the same argument as Proposition 18. We decompose $\tilde{\Omega}$ into $B(\xi, R_0)$ and $B^c(\xi, R_0)$. On $B^c(\xi, R_0)$, we note that $|\chi_\xi(x)| \leq C \exp(-\nu \frac{R_0}{h})$, and hence

$$\int_{B^c(\xi,R_0)} |\chi_\xi(x)| dx \leq \mu(\tilde{\Omega} \cap B^c(\xi, R_0)) \exp(-\nu \frac{R_0}{h}).$$

We note that for sufficiently small h, this will be smaller than h^d always (due to the $\exp(-\nu \frac{R_0}{h})$. This depends on R_0 and must be factored into the assumptions.

We next decompose $B(\xi, R_0)$ into $B(\xi, h)$ and a union of annuli of outer radius $(n+1)h$ and inner radius nh. On $B(\xi, h)$, we apply the simple bound (independent of ξ, h) $|\chi_\xi(x)| \leq C$, where C is independent of ξ, h, q from the decay estimates in the decay theorem. This yields an estimate

$$\int_{B(\xi,h)} |\chi_\xi(x)| \, dx \leq Ch^d \sim Cq^d.$$

Let $A_n = B(\xi, (n+1)h) \backslash B(\xi, nh)$. Then, we estimate

$$\begin{aligned} \int_{A_n \cap \Omega} |\chi_\xi(x)| \, dx &\leq C \exp(-\nu \frac{nh}{h}) \mu(A_n) \\ &\leq Cq^d \big((n+1)^d - n^d \big) \exp(-\nu n). \end{aligned}$$

We next argue that this sum is bounded independent of h, q. The sum of the terms $\big((n+1)^d - n^d\big) \exp(-\nu n)$ is bounded above by the infinite sum of those terms, which is finite. Consequently, summing the integrals over the annuli A_n yields a value Cq^d for some constant C. Combining the sums over the annuli and the sums on $B^c(\xi, R_0)$

yields a value on the order of q^d. The result does require h to be sufficiently small relative to R_0 to guarantee that the $\exp\big(-\nu\frac{R_0}{h}\big)$ is sufficiently small. □

Proposition 20. *For sufficiently small h (which depends on R_0), there exists a constant C so that any $s = \sum_{\xi\in\Omega} a_\xi\chi_\xi$ satisfies the upper bound*

$$\|s\|_{L^p(\tilde{\Omega})} \leq Cq^{\frac{d}{p}}\|\{\alpha_\xi\}\|_{\ell^p}.$$

Proof. Consider the operator $T(\{\alpha_\xi\}) = \sum_{\xi\in\Omega}\alpha_\xi\chi_\xi$. From (28), we have

$$\|\sum_{\xi\in\Omega}\alpha_\xi\chi_\xi\|_{L^\infty(\tilde{\Omega})} = \|T\{\alpha_\xi\}\|_{L^\infty(\tilde{\Omega})} \leq \mathcal{L}_{\Omega,\tilde{\Omega}}\|\{\alpha_\xi\}\|_{\ell^\infty}$$

and from (29)

$$\|\sum_{\xi\in\Omega}\alpha_\xi\chi_\xi\|_{L^1(\tilde{\Omega})} = \|T\{\alpha_\xi\}\|_{L^1(\tilde{\Omega})} \leq Cq^d\|\{\alpha_\xi\}\|_{\ell^1}.$$

By applying operator interpolation, it follows that

$$\|\sum_{\xi\in\Omega}\alpha_\xi\chi_\xi\|_{L^p(\tilde{\Omega})} = \|T\{\alpha_\xi\}\|_{L^p(\tilde{\Omega})} \leq (Cq^d)^{\frac{1}{p}}\mathcal{L}_{\Omega,\tilde{\Omega}}^{1-\frac{1}{p}}\|\{\alpha_\xi\}\|_{\ell^p} = Cq^{\frac{d}{p}}\|\{\alpha_\xi\}\|_{\ell^p}.$$

□

We now begin the lower bound estimate. This proof is modeled after the one given in Proposition 3.7 and Lemma 3.8 of [12]. Given a ball $B(\xi, r) \subset \tilde{\Omega} \subset \mathbb{R}^d$, let C_d denote the constant such that $\mathrm{vol}(B(\xi, r)) = C_d r^d$.

Consider the set $\Omega_k(\xi) = \{\zeta \in \Xi : \Gamma 2^k q \leq d(\xi, \zeta) \leq \Gamma 2^{k+1} q\}$. We want to estimate the number of centers in this region. By quasi-uniformity, the number of centers in Ω is bounded above by $\frac{\mathrm{vol}(\Omega)}{C_d q^d}$. In particular, given a subset $A \subset \Omega$, the

maximum number of centers in A is $\frac{\text{vol}(A)}{C_d q^d}$. Therefore, given an annulus centered at ξ with inner radius $\Gamma q 2^k$ and outer radius $\Gamma 2^{k+1} q$, the maximum number of centers in this region is

$$\frac{\text{vol}(A)}{C_d q^d} = C_d^{-1} q^{-d} [C_d \Gamma^d 2^{dk+d} q^d - C_d \Gamma^d 2^{dk} q^d] = 2^{kd} \Gamma^d,$$

which is independent of q.

Proposition 21. *Let $R_0 > 0$ and let $\Omega \subset \tilde{\Omega}$ be such that $B(x, R_0) \subset \tilde{\Omega}$ for each $x \in \Omega$. Let $X \subset \tilde{\Omega}$ with a set of centers with sufficiently small mesh norm h and let $\Xi = X \cap \Omega$. Then, there exists C_1 depending not on h or q (but on R_0) such that*

$$C_1 q^{\frac{d}{p}} \|\{\alpha_\xi\}\|_{\ell^p} \leq \|\sum_{\xi \in \Xi} \alpha_\xi \chi_\xi\|_{L^p(\tilde{\Omega})}.$$

Proof. We follow the strategy in [12] with minor modifications. We estimate the L^p norm of the function by integrating over a union of disjoint balls $\{B(\xi, \gamma q) : \xi \in X\}$. For $\gamma \leq 1$, these balls are necessarily disjoint by the definition of the separation radius. We compute

$$\|\sum_{\xi \in \Xi} \alpha_\xi \chi_\xi\|^p_{L^p(\tilde{\Omega})} = \int_{\tilde{\Omega}} \left| \sum_{\xi \in \Xi} \alpha_\xi \chi_\xi \right|^p dx \geq \sum_{\xi \in \Xi} \int_{B(\xi, \gamma q)} \left| \sum_{\zeta \in \Xi} \alpha_\zeta \chi_\zeta \right|^p dx$$

Let $\gamma \leq 1$, $\Gamma \geq 1$ and let $q \leq \frac{R_0}{\Gamma}$ and let $\xi \in \Xi$. We decompose

$$\int_{B(\xi, \gamma q)} \left| \sum_{\xi \neq \zeta} \alpha_\zeta \chi_\zeta(x) \right|^p dx \leq I_\xi + II_\xi + III_\xi$$

where we define

$$I_\xi := 3^{p-1} \int_{B(\xi,\gamma q)} |\sum_{0<d(\zeta,\xi)\leq \Gamma q} \alpha_\zeta \chi_\zeta(x)|^p \, dx,$$
$$II_\xi := 3^{p-1} \int_{B(\xi,\gamma q)} |\sum_{\Gamma q<d(\zeta,\xi)\leq R_0} \alpha_\zeta \chi_\zeta(x)|^p \, dx,$$
$$III_\xi := 3^{p-1} \int_{B(\xi,\gamma q)} |\sum_{d(\zeta,\xi)>R_0} \alpha_\zeta \chi_\zeta(x)|^p \, dx.$$

We require $\gamma q \leq R_0$ to ensure $B(\xi, \gamma q) \subset \tilde{\Omega}$. This will be imposed later by choosing appropriately small γ and an upper bound q_0 on q.

By choosing sufficiently small γ, we may apply Proposition 19 which implies there exists a γ such that $|\chi_\xi(x)| \geq \frac{2}{3}$ on $B(\xi, \gamma q)$ for all $\xi \in \Xi \subset \Omega$ (where we note this C depends on R_0 and ρ, but not just on q or h). For fixed $\xi \in \Xi$, it follows that

$$\int_{B(\xi,\gamma q)} |\chi_\xi(x)|^p \, dx \geq C \frac{2^p}{3^p} (\gamma q)^d. \tag{6.5}$$

We use the diagonal dominance argument in [12]. We have

$$\sum_{\xi\in\Xi} \int_{B(\xi,\gamma q)} |\sum_{\zeta\in\Xi} \alpha_\zeta \chi_\zeta(x)|^p \, dx \geq \sum_{\xi\in\Xi} \left(2^{1-p} \int_{B(\xi,\gamma q)} |\alpha_\xi \chi_\xi(x)|^p \, dx - \int_{B(\xi,\gamma q)} |\sum_{\zeta\neq\xi} \alpha_\zeta \chi_\zeta(x)|^p \, dx \right).$$

By applying (6.5), we have

$$\|s\|_p^p \geq \sum_{\xi\in\Xi} \left(C_d \frac{2}{3^p} |\alpha_\xi|^p - \int_{B(\xi,\gamma q)} |\sum_{\zeta\neq\xi} \alpha_\zeta \chi_\zeta(x)|^p \, dx \right).$$

Our objective is to bound the second term by the first term. We argue that for sufficiently small γ, the second term can necessarily be smaller than half the term

$\frac{2}{3^p} C_d (\gamma q)^d$. If this is possible, then we have the bound

$$\|s\|_p^p \geq \frac{C_d}{3^p} (\gamma q)^d \|\{\alpha_\xi\}\|_{\ell^p}^p.$$

□

Lemma 30. *For $p \geq 1$, $|\sum_{j=1}^n a_j|^p \leq n^{p-1} \sum_{j=1}^n |a_j|^p$.*

Proof. We introduce the auxiliary sequence $b_j = 1$ for each $j = 1, \ldots, n$ and apply Holder's inequality with $\frac{1}{q} = 1 - \frac{1}{p}$

$$\begin{aligned} \sum_{j=1}^n a_j b_j &\leq (\sum_{j=1}^n |a_j|^p)^{\frac{1}{p}} \big(\sum_{j=1}^n b_j^q \big)^{\frac{1}{q}} \\ &= n^{1-\frac{1}{p}} \big(\sum_{j=1}^n |a_j|^p \big)^{\frac{1}{p}} \end{aligned}$$

Taking both sides to the power of p yields the result. □

Lemma 31. *For $p \geq 1$, $|\sum_{j=1}^n a_j|^p \leq \sum_{j=1}^n 2^{j(p-1)} |a_j|^p$.*

Proof. Let $b_1 = a_1$ and $b_2 = \sum_{j=2}^n a_j$. We apply Lemma 30 to $|\sum_{k=1}^2 b_k|^p$. Then, we have

$$|a_1 + \sum_{j=2}^n a_j|^p \leq 2^{p-1} (|a_1|^p + |\sum_{j=2}^n a_j|^p).$$

Now, define $b_1 = a_2$ and $b_2 = \sum_{j=3}^n a_j$ and apply the result again. each iteration yields an additional 2^{p-1} term. Iterating this $n-1$ times yields the result. □

Lemma 32. *Let $\Xi \subset \Omega$ be a set of centers with separation radius q. Let $x \in \Omega$ and $B(x, R) \subset \Omega$ with $R > q$. Then,*

$$|\#\Xi \cap B(x, R)| \leq 2^d \frac{R^d}{q^d}.$$

Proof. Let $N = |\#\Xi \cap B(x,R)|$. By the definition of the separation radius, if $\xi, \eta \in \Xi \cap B(x,R)$, then $B(\xi,q) \cap B(\eta,q) = \emptyset$. Consequently, if we consider the inflated ball $B(x,2R)$, then $B(\xi,q) \subset B(x,2R)$ since $R \geq q$. Then, we have

$$\cup_{\xi \in \Xi \cap B(x,R)} B(\xi,q) \subset B(x,2R)$$

which implies

$$\mathrm{vol}(\cup_{\xi \in \Xi \cap B(x,R)} B(\xi,q)) = C_d N q^d \leq \mathrm{vol}(B(x,2R)) = C_d 2^d R^d.$$

Therefore, $N \leq 2^d \frac{R^d}{q^d}$. □

Lemma 33. *There exists q_0 and R such that*

$$\sum_{\xi \in \Xi} \int_{B(\xi,\gamma q)} \left| \sum_{\zeta \neq \xi} \alpha_\zeta \chi_\zeta(x)|^p \, dx \right| \leq \frac{C_d}{3^p} (\gamma q)^d \sum_{\xi \in \Xi} |\alpha_\xi|^p$$

holds for all $1 \leq p < \infty$ whenever $q < q_0$ and $\gamma < R$.

Proof. We use the strategy developed in [12]. Let $\gamma \leq 1$ and $q \leq \frac{R_0}{\Gamma}$ and $\xi \in \Xi \subset \Omega$. Let

$$\int_{B(\xi,\gamma q)} |\sum_{\zeta \neq \xi} \alpha_\zeta \chi_\zeta(x)|^p \, dx \leq (I_\xi + II_\xi + III_\xi),$$

where (applying Lemma 30 with $n = 3$) we denote

$$I_\xi := 3^{p-1} \int_{B(\xi,\gamma q)} |\sum_{0 < d(\zeta,\xi) \leq \Gamma q} \alpha_\zeta \chi_\zeta(x)|^p \, dx,$$
$$II_\xi := 3^{p-1} \int_{B(\xi,\gamma q)} |\sum_{\Gamma q < d(\zeta,\xi) \leq R_0} \alpha_\zeta \chi_\zeta(x)|^p \, dx,$$
$$III_\xi := 3^{p-1} \int_{B(\xi,\gamma q)} |\sum_{d(\zeta,\xi) > R_0} \alpha_\zeta \chi_\zeta(x)|^p \, dx.$$

Step 1: Estimating II_ξ.

We define

$$\Omega_k := \{\zeta \in \Xi : \Gamma 2^k q \leq d(\xi, \zeta) \leq \Gamma 2^{k+1} q\}, \qquad k = 0, 1, \ldots, N_q.$$

By definition of we are considering points such that $\Gamma q \leq d(\zeta, \xi) \leq R_0$. So, $2^{N_q}\Gamma q \sim R_0$. So, $N_q \sim \log(\frac{R_0}{\Gamma q})$. We apply Lemma 31 to compute

$$\begin{aligned} |\sum_{II_\xi} \alpha_\zeta \chi_\zeta(x)|^p &= \sum_{k=1}^{N_q} \sum_{\zeta \in \Omega_k} \alpha_\zeta \chi_\zeta(x)|^p \\ &\leq \sum_{k=1}^{N_q} 2^{k(p-1)} |\sum_{\zeta \in \Omega_k} \alpha_\zeta \chi_\zeta(x)|^p. \end{aligned}$$

We define $M_k = \int_{B(\xi,\gamma q)} |\sum_{\zeta \in \Omega_k} \alpha_\zeta \chi_\zeta(x)|^p \, dx$ and we have

$$I_\xi :\leq 3^{p-1} \sum_{k=0}^{N_q} 2^{(p-1)(k+1)} M_k. \tag{6.6}$$

We first need to estimate M_k, then we need to bound the sum over the M_k terms. Let N_k denote the number of centers in Ω_k. We apply Lemma 30 with $n = N_k$ and compute

$$\begin{aligned} M_k &\leq N_k^{p-1} \sum_{\zeta \in \Omega_k} \int_{B(\xi,\gamma q)} |\alpha_\zeta \chi_\zeta(x)|^p \, dx \\ &\leq N_k^{p-1} \max_{\zeta \in \Omega_k} \|\chi_\zeta\|^p_{L^\infty(B(\xi,\gamma q))} C_d(\gamma q)^d \sum_{\zeta \in \Omega_k} |\alpha_\zeta|^p. \end{aligned}$$

We estimate N_k by applying Lemma 32

$$N_k = 2^d \frac{\Gamma^d 2^{(k+1)d} q^d}{q^d} = 2^d (\Gamma 2^{k+1})^d$$

points in it. We estimate the L^∞ norm by applying Theorem 7. Since $d(\xi,\zeta) \geq \Gamma 2^k q$, for any $x \in B(\xi, \gamma q)$, we have $d(\zeta, x) \geq \Gamma 2^k q - \gamma q$. Consequently, we estimate

$$\begin{aligned}
|\chi_\zeta(x)| &\leq C \exp\left(- \nu \frac{d(\zeta,x)}{h} \right) \\
&\leq C \exp\left(- \nu \frac{\Gamma 2^k q - \gamma q}{h} \right) \\
&\leq C \exp\left(- \frac{\nu}{\rho} \Gamma 2^k \right) \exp\left(\frac{\nu}{\rho} \gamma \right) \\
&\leq C \exp\left(- \frac{\nu}{\rho} \Gamma 2^k \right) \exp\left(\frac{\nu}{\rho} \right).
\end{aligned}$$

In the above computation, we applied the assumption $\gamma \leq 1$. We absorb the $\exp(\frac{\nu}{\rho})$ into the constant C; this is allowed by quasi-uniformity (ν should not depend on h or q). With these estimates, we may bound M_k by

$$M_k \leq (2^d(\Gamma 2^{k+1})^d)^{p-1} C^p \exp(p\nu') \exp(-\nu' p \Gamma 2^k) C_d (\gamma q)^d \sum_{\zeta \in \Omega_k} |\alpha_\zeta|^p$$

where we define $\nu' := \frac{\nu}{\rho}$. Then, by applying (6.6), we have

$$II_\xi \leq 3^{p-1} \sum_{k=0}^{N_q} 2^{(p-1)(k+1)} M_k.$$

Note that

$$2^{(p-1)(k+1)} (2^{d(k+1)})^{p-1} = 2^{(d+1)(p-1)} 2^{k(d+1)(p-1)}$$

and denote $K := 3C2^d 2^{d+1}$. By summing over ξ and pulling out all of the powers of

two not depending on k, we bound

$$\sum_{\xi\in\Xi} II_\xi \le 3^{p-1}\sum_{k=0}^{N_q} 2^{(p-1)(k+1)}M_k$$
$$\le CK^{p-1}C_d(\gamma q)^d\sum_{k=0}^{N_q}(2^{k(d+1)}\Gamma^d)^{p-1}\exp(-\nu'\Gamma 2^k)\sum_{\xi\in\Xi}\sum_{\zeta\in\Omega_k(\xi)}|\alpha_\zeta|^p.$$

We need an upper bound on $\sum_{\xi\in\Xi}\sum_{\zeta\in\Omega_k(\xi)}|\alpha_\zeta|^p$. We note that we can count the number of times α_ζ appears. That is, we need to total the number of times ζ appears in each $\Omega_K(\xi)$ for each $\xi\in\Xi$. That is, we seek $\#\{\xi:\zeta\in\Omega_k(\xi)\}=\#\{\xi:\Gamma 2^k q\le d(\zeta,\xi)\le\Gamma 2^{k+1}d\}$, which can be bounded above by $\#\Omega_k(\zeta)=2^d2^{k+1}\Gamma^d$. Furthermore, we note that $2^{(k+1)d}\le 2^{k(d+1)}2^{d+1}$, which allows us to bound

$$\sum_{\xi\in\Xi}\sum_{\zeta\in\Omega_k(\xi)}|\alpha_\zeta|^p\le 2^{d+1}2^{k(d+1)}\Gamma^d2^d\sum_{\xi\in\Xi}|\alpha_\xi|^p=\frac{2^{d(k+1)}\Gamma^d}{3C}K\sum_{\xi\in\Xi}|\alpha_\xi|^p.$$

Applying this to the upper bound for $\sum II_\xi$, we compute:

$$\sum_{\xi\in\Xi} II_\xi \le CK^{p-1}C_d(\gamma q)^d\sum_{k=0}^{N_q}2^{k(d+1)}\Gamma^d)^{p-1}\exp(-\nu'\Gamma 2^k)(2^d(2^{k+1}\Gamma)^d\sum_{\xi\in\Xi}|\alpha_\xi|^p$$
$$\le\frac{1}{3}K^pC_d(\gamma q)^d\sum_{k=0}^{N_q}(2^k)^{(d+1)p}\Gamma^{d(p-1)}\Gamma^d\exp(-\nu' p\Gamma 2^k)\sum_{\xi\in\Xi}|\alpha_\xi|^p$$
$$\le\frac{C_d}{3}K^p\frac{(\gamma q)^d}{\Gamma^p}\sum_{k=0}^{N_q}(2^k\Gamma)^{(d+1)p}\exp(-\nu' p\Gamma 2^k)\sum_{\xi\in\Xi}|\alpha_\xi|^p.$$

Let $f(x)=(x\Gamma)^{(d+1)p-1}\exp(-\nu' p\Gamma x)$. After factoring, we see

$$\sum_{k=0}^{N_q}(2^k\Gamma)^{(d+1)p}\exp(-\nu' p\Gamma 2^k)=\sum_{k=0}^{N_q}2^k\Gamma(2^k\Gamma)^{(d+1)p-1}\exp(-\nu' p\Gamma 2^k)=\Gamma\sum_{k=0}^{N_q}2^kf(2^k).$$

By the Cauchy condensation test, we may bound $\sum_{k=0}^{\infty}2^kf(2^k)\le 2\sum_{k=1}^{\infty}f(k)$, pro-

vided that $f(n)$ is a positive, non-increasing sequence. For sufficiently large Γ, this is true. Now, we focus on estimating the summation $2\sum_{k=1}^{\infty} f(k)$. By the integral comparison test,

$$\begin{aligned} 2\Gamma \sum_{k=1}^{\infty} f(k) &\leq 2\Gamma f(1) + 2\Gamma \int_1^{\infty} (r\Gamma)^{(d+1)p-1} \exp(-\nu' pr\Gamma)\, dr \\ &\leq 2\Gamma^{(d+1)p} \exp(-\nu' p\Gamma) + 2\int_{\Gamma}^{\infty} \tau^{(d+1)p-1} \exp(-\nu' p\tau)\, d\tau. \end{aligned}$$

Applying this to our estimate for $\sum_{\xi\in\Xi} II_\xi$, we find

$$\sum_{\xi\in\Xi} II_\xi \leq \frac{C_d}{3} K^p \underbrace{(\Gamma^{dp} \exp(-\nu' p\Gamma) + \frac{2}{\Gamma^p}\int_{\Gamma}^{\infty} \tau^{(d+1)p-1} \exp(-\nu' p\tau)\, d\tau}_{G(\Gamma)} (\gamma q)^d \|\alpha_\xi\|_{\ell^p}^p.$$

If we can find a Γ_0 such that for all $\Gamma > \Gamma_0$, $G(\Gamma)$ is sufficiently small, then we can make $\sum_{\xi\in\Xi} II_\xi$ as small as possible. By Lemma 34, such a choice can be made. We now have chosen Γ, and for the remainder of the proof, it is fixed.

Step 2: Estimating III_ξ.

Let $\Omega_{R_0}(\xi) = \{\zeta : d(\xi,\zeta) \geq R_0\}$. We first estimate

$$\begin{aligned} III_\xi &\leq 3^{p-1} \#\Omega_{R_0}(\xi) \sum_{\zeta} \int_{B(\xi,\gamma q)} |\alpha_\zeta \chi_\zeta(x)|^p\, dx \\ &\leq 3^{p-1} (\#\Omega_{R_0}(\xi))^{p-1} \max_{\zeta\in\Omega_{R_0}(\xi)} \|\chi_\zeta\|_{L^\infty(B(\xi,\gamma q))}^p (\gamma q)^d \sum_{\zeta\in\Omega_{R_0}(\xi)} |\alpha_\zeta|^p \end{aligned}$$

We note that $\cup_{\xi\in\Xi} B(\xi,q) \subset \tilde{\Omega}$ is a disjoint union of balls, and hence we may estimate $\#\Omega_{R_0}(\xi) \leq \frac{\mathrm{vol}(\Omega)}{C_d q^d}$. On $B(\xi,\gamma q)$, we apply the exponential decay, and note

that $d(\zeta, x) \geq d(\zeta, \xi) - \gamma q \geq R_0 - \gamma q$.

$$\begin{aligned} |\chi_\zeta(x)| &\leq C \exp\left(- \nu \frac{d(\zeta, x)}{h} \right) \leq C \exp\left(- \nu \frac{R_0}{h} + \frac{\nu}{\rho} \gamma \right) \\ &= C \exp\left(\frac{\nu}{\rho} \right) \exp\left(- \nu \frac{R_0}{\rho} \right). \end{aligned}$$

We absorb the $\exp(\frac{\nu}{\rho})$ into the C term (which may slightly increase C) to obtain the estimate on $|\chi_\zeta(x)|$. Let $\nu' = \frac{\nu}{\rho}$ once again, and we now compute

$$\begin{aligned} III_\xi &\leq 3^{p-1} \left(\frac{\mathrm{vol}(\tilde{\Omega})}{C_d q^d} \right)^{p-1} C^p \exp(-\nu p \frac{R_0}{q}) (\gamma q)^d \sum_{\zeta \in \Omega_{R_0}(\xi)} |\alpha_\zeta|^p \\ &\leq C_1 C_d K^{p-1} q^{-d(p-1)} \exp(-\nu p \frac{R_0}{q}) (\gamma q)^d \sum_{\zeta \in \Omega_{R_0}(\xi)} |\alpha_\zeta|^p \end{aligned}$$

where $K := \frac{3 C_1 \mathrm{vol}(\tilde{\Omega})}{C_d}$. We now sum over $\xi \in \Xi$ and once again estimate how many times $|\alpha_\zeta|^p$ is over counted. We note that this is the same as $\#\Omega_{R_0}(\xi) \leq \frac{\mathrm{vol}(\tilde{\Omega})}{C_d q^d}$. Therefore, we have

$$\begin{aligned} \sum_{\xi \in \Xi} III_\xi &\leq C_1 C_d K^{p-1} q^{-d(p-1)} \exp(-\nu p \frac{R_0}{q}) (\gamma q)^d \sum_{\xi \in \Xi} \sum_{\zeta \in \Omega_{R_0}(\xi)} |\alpha_\zeta|^p \\ &\leq C_1 C_d K^{p-1} q^{-d(p-1)} \exp(-\nu p \frac{R_0}{q}) (\gamma q)^d \frac{\mathrm{vol}(\tilde{\Omega})}{C_d q^d} \sum_{\xi \in \Xi} |\alpha_\xi|^p \\ &\leq \frac{C_d}{3} K^p q^{-dp} \exp(-\nu p \frac{R_0}{q}) (\gamma q)^d \sum_{\xi \in \Xi} |\alpha_\xi|^p. \end{aligned}$$

We aim to make this term arbitarily small, which we may control with the $q^{-dp} \exp(-\nu p \frac{R_0}{q})$ term. We impose a condition that $q \leq q_0$, where $q_0 \leq \frac{R_0}{\Gamma}$ is small enough to ensure that $q^{-d} \exp(-\nu \frac{R_0}{q})$ is sufficiently small.

Step 3: Estimating I_ξ: At this point, we consider Γ to be known and fixed, and $q \leq q_0$, as established in Step 2. We consider, for fixed ξ, the set $\Omega_\xi = \{\zeta : d(\xi, \zeta) \leq$

$\Gamma q\}$. For $\zeta \neq \chi$, $\chi_\zeta(\xi) = 0$, and for any $x \in B(\xi, \gamma q)$, we have by Proposition 19

$$|\chi_\zeta(x)| = |\chi_\zeta(x) - \chi_\zeta(\xi)| \leq C\left(\frac{d(x,\xi)}{q}\right)^\epsilon \leq C\gamma^\epsilon.$$

We then estimate the cardinality of $\{\zeta : d(\xi,\zeta) \leq \Gamma q\}$ to be bounded above by $\frac{2^d C_d (\Gamma q)^d}{C_d q^d} = (2\Gamma)^d$. We apply Lemma 30 to I_ξ to obtain

$$\begin{aligned} I_\xi &\leq 3^{p-1}(2^d\Gamma^d)^{p-1} \sum_{0<d(\zeta,\xi)\leq\Gamma q} \int_{B(\xi,\gamma q)} |\alpha_\zeta|^p |\chi_\zeta|^p \, dx \\ &\leq 3^{p-1}(2^d\Gamma^d)^{p-1}(C\gamma^\epsilon)^p C_d(\gamma q)^d \sum_{0<d(\zeta,\xi)\leq\Gamma q} |\alpha_\zeta|^p \, dx. \end{aligned}$$

We now sum over $\xi \in \Xi$ and note that we may over-estimate the number of times α_ξ appears by $2^d\Gamma^d$, which yields

$$\sum_{\xi\in\Xi} I_\xi \leq 3^{p-1} C_d 2^{dp} \Gamma^{dp} C^p \gamma^{\epsilon p} (\gamma q)^d \|\alpha_\xi\|_{\ell^p}^p.$$

Since $q \leq q_0$ and Γ is fixed and chosen, we may choose γ small enough to shrink this term as small as required. □

Lemma 34 (Lemma 3.9 of [12])**.** *Let $\nu > 0$, $d \geq 1$. For every $\epsilon > 0$, there exists a $\Gamma_0 \geq \frac{d+1}{\nu}$ (depending on ν, d, ϵ) so that for $\Gamma \geq \Gamma_0$ and for all $p \in [1,\infty)$,*

$$\max\left(\Gamma^{dp}\exp(-\nu p\Gamma), \frac{2}{\Gamma^p}\int_\Gamma^\infty r^{(d+1)p-1}\exp(-\nu p r)\, dr\right) \leq \epsilon^p.$$

Proof. See [12] for the proof. □

Corollary 8. *Let Ω be R_0 embedded subset of Ω, and let $X \subset \tilde{\Omega}$ be a finite set of centers with mesh norm h and separation radius q and let $\Xi = X \cap \Omega$. For*

$1 \leq p \leq q \leq \infty$,

$$\|\sum_{\xi\in\Xi} \alpha_\xi \chi_\xi\|_{L^q(\tilde{\Omega})} \leq C q^{-d(\frac{1}{p}-\frac{1}{q})} \|\sum_{\xi\in\Xi} \alpha_\xi \chi_\xi\|_{L^p(\tilde{\Omega})}$$

Proof. We know by Proposition 21 that there exists c_1, c_2 independent of q such that for all $1 \leq p \leq \infty$,

$$c_1 q^{\frac{d}{p}} \|\{\alpha_\xi\}\|_{\ell^p} \leq \|\sum_{\xi\in\Xi} \alpha_\xi \chi_\xi\|_{L^p(\tilde{\Omega})} \leq c_2 q^{\frac{d}{p}} \|\{\alpha_\xi\}\|_{\ell^p}.$$

Since $p \leq q$, $\|\{\alpha_\xi\}\|_{\ell^q} \leq \|\{\alpha_\xi\}\|_{\ell^p}$, and hence we have

$$\begin{aligned} \|\sum \alpha_\xi \chi_\xi\|_{L^q} &\leq c_2 q^{\frac{d}{q}} \|\{\alpha_\xi\}\|_{\ell^q} \\ &\leq c_2 q^{\frac{d}{q}} \|\{\alpha_\xi\}\|_{\ell^p} \\ &\leq c_2 q^{\frac{d}{q}} c_1^{-1} q^{-\frac{d}{p}} \|\sum_{\xi\in\Xi} \alpha_\xi \chi_\xi\|_{L^p} \end{aligned}$$

□

The approach we are taking here is to assume we have some $\Omega \subset \tilde{\Omega}$ with scattered centers $X \subset \tilde{\Omega}$, and we focus only on Lagrange functions centered at points in $\Xi = X \cap \Omega$. The distance R_0 is pivotal to all of the arguments above and the restricted Lebesgue constant and condition numbers of the L^p stability depend on R_0, as well as the exponential decay. On the other hand, we could start with a set Ω, and add additional points outside to define a region $\tilde{\Omega}$. By tailoring the number of points we add, we can then *choose* the R_0, which determines how well the functions decay. We need to then ask the question: given a mesh norm, what should R_0 be?

Given a fixed h_0, for any $h < h_0$, by choosing $R_0 = Kh_0|\log(h_0)|$, we have

$$|\chi_\xi(x)| \leq C\exp(-\nu\frac{R_0}{h}) = C\exp(-\nu\frac{h_0|\log(h_0)|}{h}) \leq Ch_0^{\nu K}.$$

The problem is that to create $\tilde{\Omega}$ given a mesh norm h in Ω, we must add sufficiently many points into $\tilde{\Omega}$ to guarantee a mesh norm of h (which is required by all the above).

7. SUMMARY AND CONCLUSIONS

In this dissertation, we focused on the theory and application of a localized basis constructed from radial basis functions. We developed numerical methods for solving partial differential equations on manifolds and for integral equations arising from nonlocal diffusion on Euclidean domains. We developed implementable methods and studied their computational properties. The experimental results matched theoretical estimates for the different methods and suggested the computational methods we developed provide novel approaches to solving numerous applied problems.

The method we developed for nonlocal diffusion on compact Euclidean domains provides an innovative approach to solving a challenging problem. This work, joint with Rich Lehoucq, is the result of summer research supported by Sandia National Laboratories. We used recently developed results to produce a local Lagrange function discretization for a variational formulation of the nonlocal diffusion problem. A unique quadrature method was designed that works for the local Lagrange basis and provides a fast assembly of a sparse stiffness matrix. Theoretical estimates verified that the condition number of the matrix is bounded independent of the mesh norm and that the resulting matrix is sparse, provided that the cutoff local Lagrange function and the corresponding quadrature method are used. A separate method using local Lagrange functions was developed jointly with Lehoucq, Narcowich and Ward. The well-posedness of the Lagrange multiplier variational formulation is proven by a combination of a discrete inf-sup condition and a coercivity result. Possible future work includes addressing the error resulting from quadrature.

The new method for solving partial differential equations on manifolds extends the growing collection of literature exploring new methods of discretizing problems

by radial basis functions. The work, joint with Narcowich and Ward, uses the highly localized Lagrange and local Lagrange bases. A previously developed quadrature method yields an implementable assembly routine for stiffness matrices. The effects of quadrature on the L^2 error of the solution are studied. There is potential for future work to improve this method. The assembly of the stiffness matrix can likely be accelerated by using the massive parallelism offered by graphics processing units. Furthermore, only few centers contribute significantly to the matrix due to the exponential decay of the elements in the stiffness matrix. Consequently, the matrix can be made sparse by setting many entries in the stiffness matrix to be zero dependent on the positions of the respective centers.

There is much potential for radial basis function in many areas of numerical analysis and for applications. Taking advantage of parallelism in these methods as well as exploring the properties of new, efficient bases of radial basis functions is potentially a very fruitful area. We hope the results discussed in this dissertation are just the beginning of many future radial basis function techniques using local Lagrange functions.

REFERENCES

[1] S. D. Bond, R. B. Lehoucq, and S. T. Rowe. A Galerkin radial basis function method for nonlocal diffusion. Technical Report SAND 2013-10673P, 2014. To Appear in the Proceedings of the Seventh International Workshop on Meshfree Methods for Partial Differential Equations; see http://www.sandia.gov/ rblehou/papers.html.

[2] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[3] Q. Du, M. Gunzburger, R. B. Lehoucq, and Kun Zhou. Analysis and approximation of nonlocal diffusion problems with volume constraints. *SIAM Review*, 54(4):667–696, 2012.

[4] Q. Du, R.B. Lehoucq, and A.M. Tartakovsky. Integral approximations to classical diffusion and smoothed particle hydrodynamics. *Computer Methods in Applied Mechanics and Engineering*, 286:216–229, April 2015.

[5] J. Duchon. Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces. *Rev. Française Automat. Informat. Rech. Opér. Anal. Numer.*, 10:5–12, 1976.

[6] G. E. Fasshauer. Solving partial differential equations by collocation with radial basis functions. In A. Le Méhauté, C. Rabut, and L. L. Schumaker, editors, *Surface Fitting and Multiresolution Methods*, pages 131–138, Nashville, 1997. Vanderbilt University Press.

[7] E. J. 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(5), 2013.

[8] E. J. Fuselier, T. Hangelbroek, F. J. Narcowich, J. D. Ward, and G. B. Wright. Kernel based quadrature on spheres and other homogeneous spaces. *Numerische Mathematik*, 127(1):57–92, 2014.

[9] M. Gunzburger and R. B. Lehoucq. A nonlocal vector calculus with application to nonlocal boundary value problems. *SIAM Multiscale Modeling & Simulation*, 8(5):1581–1598, 2010.

[10] T. Hangelbroek. The penalized Lebesgue constant for surface spline interpolation. *Proc. Americ. Math Soc.*, 140:173–187, 2012.

[11] T. Hangelbroek, F. J. Narcowich, C. Rieger, and J. D. Ward. An inverse theorem on bounded domains for meshless methods using localized bases. *ArXiv e-prints*, June 2014.

[12] T. Hangelbroek, F. J. Narcowich, X. Sun, and J. D. Ward. Kernel approximation on manifolds II: the L_∞ norm of the L_2 projector. *SIAM J. Math. Anal.*, 43(2):662–684, 2011.

[13] T. Hangelbroek, F. J. Narcowich, and J. D. Ward. Kernel approximation on manifolds I: Bounding the lebesgue constant. *SIAM J. Math. Anal.*, 42(4):175–208, 2010.

[14] T. Hangelbroek, F. J. Narcowich, and J. D. Ward. Polyharmonic and Related Kernels on Manifolds: Interpolation and Approximation. *Found. Comput. Math.*, 12:625–670, 2012.

[15] Y. C. Hon and R. Schaback. On unsymmetric collocation by radial basis functions. *J. Appl. Math. Comp.*, 119:177–186, 2001.

[16] E. J. Kansa. Multiquadrics - a scattered data approximation scheme with applications to computational fluid-dynamics - ii: Solutions to parabolic, hyperbolic and elliptic partial differential equations. *Comput. Math. Appl.*, 19:147–161, 1990.

[17] Q.T. Le Gia, I.H. Sloan, and H. Wendland. Multiscale rbf collocation for solving pdes on spheres. *Numerische Mathematik*, 121(1):99–125, 2012.

[18] W. R. Madych and S. A. Nelson. Bounds on multivariate polynomials and exponential error estimates for multiquadric interpolation. *J. Approximation Theory*, 70:94–114, 1992.

[19] O. V. Matveev. Approximation properties of interpolating D^m-splines. *Soviet Math. Dokl.*, 44:650–655, 1992.

[20] J. Meinguet. Basic mathematical aspects of surface spline interpolation. In G. Hämmerlin, editor, *Numerische Integration (Oberwolfach 1978)*, volume 45 of *Int. Ser. Numer. Math.*, pages 211–220, Basel, 1979. Birkhäuser.

[21] J. Meinguet. A convolution approach to multivariate representation formulas. In W. Schempp and K. Zeller, editors, *Multivariate Approximation Theory (Oberwohlfach 1979)*, volume 51 of *Int. Ser. Numer. Math.*, pages 198–210, Basel, 1979. Birkhäuser.

[22] J. Meinguet. Multivariate interpolation at arbitrary points made simple. *J. Appl. Math. Phys.*, 30:292–304, 1979.

[23] B. Mengesha and T. Mengesha. Results on Nonlocal Boundary Value Problems. *Numerical Functional Analysis and Optimization*, 31(12):1301–1317, 2010.

[24] 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(271):1647–1679, 2010.

[25] 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. *ArXiv e-prints*, April 2014.

[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(2):175–186, 2006.

[27] H. Wendland. Meshless Galerkin methods using radial basis functions. *Math. Comp.*, 68:1521–1531, 1999.

[28] H. Wendland. *Scattered Data Approximation*. Cambridge University Press, Cambridge, UK, 2005.