University of Padova

Department of Mathematics "Tullio Levi-Civita"

PhD Thesis in Mathematics

computational mathematics curriculum

Kernel-based approximation of functions: three problems and their efficient solutions

Supervisor
Prof. Stefano De Marchi
University of Padova

PhD Candidate
Mohammad Karimnejad Esfahani

Student ID
1237431

Academic Year
March 2024

To my family and friends, who supported me during all these years.

Abstract

Scattered data approximation in multiple dimensions based on the positive definite kernels, especially Radial Basis Functions (RBFs), is frequently used in the modern approximation theory. RBF methods are known for their strong performance in function approximation and interpolation problems. This research addresses key challenges in multivariate approximation by utilizing the properties of RBFs, offering new insights and methodologies that improve the accuracy and conditioning of approximation processes. Specifically, it is known that the accuracy of the approximation depends on the selected set of bases with which the underlying function is reconstructed. Therefore, in this thesis, we follow the idea of finding a new set of bases that can improve the accuracy of the approximation, either through improving the conditioning of the interpolation matrix or by incorporating the features of the underlying function into the selected bases, such as discontinuities.

To be more detailed, to overcome the ill-conditioning of the interpolation matrix resulting from the conditionally positive definite kernels, we present various sets of bases constructed by different types of decomposition applied in the interpolation matrix. Another class of bases is constructed using the Mercer expansion of the reproducing kernel corresponding to any conditionally positive definite kernels.

Secondly, we design a moving least-squares approach for scattered data approximation that incorporates the discontinuities of the underlying functions into the weight functions. Thus, the newly constructed basis measures the influence of the data sites on the approximant, not only with regard to their distance from the evaluation point but also with respect to the discontinuities of the underlying function.

Eventually, we use Direct RBF Partition of Unity to set up the differentiation matrix to solve the time-dependent PDEs intrinsic to the surface. Taking this approach, the bases are modified depending on the patches that include the evaluation points. Combined with the closest point representation of the surface, the preliminary results show an improvement in terms of accuracy.

The findings of this research contribute to the broader understanding of kernel methods in approximation theory, offering valuable perspectives for future research and development. By advancing the methodologies and applications of RBFs, this thesis paves the way for more accurate and efficient approximation techniques that can be applied across various scientific and engineering disciplines.

Contents

Listing of figures

Listing of tables

1 Preliminaries

> Truth is much too complicated to allow anything but approximations.
>
> John von Neumann

1.1 Historical Background

The fundamental problem of approximation theory is to reconstruct a possibly complicated function, called the *target function*, by more straightforward, easier-to-compute functions called the *interpolants* or *approximants*. The origin of approximation theory can be traced back to ancient times, say, the formula for approximating the square root of a number, usually attributed to the Babylonians or the Greek mathematicians, such as *Euclid* and *Archimedes*, who made the foundation for geometric approximations and methods for calculating areas and volumes. Their work set the stage for later developments in the field. In the 17th century, *Lagrange* and *Newton* independently developed interpolation methods using polynomials, providing powerful tools for estimating functions between given data points. This era laid the groundwork for subsequent polynomial approximation techniques. The 18th and 19th centuries marked a shift toward harmonic analysis, with Fourier's groundbreaking work on representing functions as infinite series of sines and cosines. This period also saw the emergence of wavelet theory, offering a powerful alternative to traditional Fourier methods. The development of fast Fourier transform (FFT) algorithms in the 20th century greatly enhanced computational efficiency in approximation tasks. The mid-20th century witnessed the integration of approximation theory with functional analysis, yielding powerful insights into the convergence properties of approximation methods. Contributions from luminaries like *Bernstein, Tchebycheff*, and *Kolmogorov* paved the way for understanding the limitations and possibilities of various approximation techniques. To mention a selection of classical sources on approximation theory (from the second half of the last century), we refer to [CL09, Dav75, Mal99, Pow81, DeV98].

The latter half of the 20th century and the 21st century witnessed a surge in computational power, enabling the application of approximation theory to real-world problems in diverse fields. In particular, in the present age, computers are applied almost anywhere in computational science, engineering, and finance. Thus, it becomes increasingly important to implement mathematical functions for efficient evaluation in computer programs. Such a need motivated using and analyzing non-polynomial approximation techniques, including splines, wavelets, and radial basis functions (RBFs). Specifically, RBFs gained prominence for their ability to handle complex data structures. The readers can find vast discussion about mentioned approximation techniques in [DB78, HH13, Chu97, Dau92, I+18, Buh03, Fas07, FM15, FF15a, Wen04].

Regardless of which kind of approximation is employed, various reasons justify why *approximations* of functions are preferred over its exact mathematical form. A simple one is that, in many instances, it is not possible to implement the functions precisely because, for example, they are only represented by an infinite expansion. Furthermore, the function we want to use may not be completely known to us, maybe too expensive, or may require computer time and memory to compute in advance. This is another typical, important reason why approximations are needed. This is true even in the face of ever-increasing speed and computer memory availability, given that additional memory and speed will always increase the demand of the users and the size of the problems to be solved. Finally, the data that defines the function may have to be computed interactively or using a step-by-step approach, making it suitable to compute approximations. With those, we can then pursue further computations, for instance, or further evaluations that the user requires or displays data or functions on a screen. Such cases are standard in mathematical methods for modeling and analyzing functions. To illustrate, imagine that we are given a set of data (such as measurements along with locations at which these measurements were obtained), and we want to find a rule that allows us to deduce information about the process we are studying also at locations different from those at which we obtained our measurements. Typical examples could be a series of measurements taken over a specific period, weather data collected at weather stations, or data obtained via some physical or computer experiment involving many different input parameters (i.e., the problem is high-dimensional).

Having the applications mentioned above in mind, it immediately follows that any competing method has to be capable of dealing with a vast number of data points in an arbitrary number of space dimensions, such that the data points might bear no regularity at all, that is the data is scattered, and which might even change position with time. The first and fundamental mathematical challenge is to ensure that the scattered data fitting problem (to be formulated shortly) has a well-posed problem formulation in the sense of *Hadamard*, i.e., that it not only has a solution but that this solution is also unique and that small changes in the data lead only to small changes in the solution.

Mathematically speaking, let $f\colon \Omega \to \mathbb{R}$ with $\Omega \subset \mathbb{R}^d$. One generally takes the assumption that f lies in a (probably infinite-dimensional) linear space $\mathcal{F}$, i.e., $f \in \mathcal{F}$. For the construc-

tion of a concrete approximation method for elements $f \in \mathcal{F}$, we first fix a suitable *finite dimensional* subset $\mathcal{S} \subset \mathcal{F}$, from which we seek an approximation $s_f \in \mathcal{S}$ to f. Therefore, if $\{s_1, \ldots, s_n\}$ is a set of basis for S, the interpolation s_f is represented as $s_f = \sum_{i=1}^{n} c_i s_i$. Therefore, the first question is how these bases must be selected.
We will see that the selected basis must depend on the given data location. This enters us into the realm of meshfree methods, which have gained much attention in recent years. Thus, much of the work concerned with meshfree approximation methods is interdisciplinary - at the interface between mathematics and numerous application areas. Meshfree methods are often better suited to cope with changes in the geometry of the domain of interest (e.g., free surfaces and large deformations) than classical discretization techniques such as finite differences, finite elements, or finite volumes. Another obvious advantage of meshfree discretizations is - of course - their independence from a mesh. Mesh generation is still the most time-consuming part of any mesh-based numerical simulation. Since meshfree discretization techniques are based only on a set of independent points, these costs of mesh generation are eliminated. Therefore, meshfree approximation methods can be seen to provide a new generation of numerical tools.

1.2 Linear spaces

This chapter mainly discusses basic mathematical methods and numerical algorithms for interpolation and approximation of functions in one variable. Therefore, all materials of this chapter can be found in the textbooks mentioned in the previous chapter.
Let V be a linear space over the set of scalars $\mathbb{K}$, with $v_1, \ldots, v_n$ being vectors belonging to V.

Definition 1. *We say $\{v_i\}_{i=1}^n \subset V$ are linearly dependent if there are scalars $\alpha_i \in \mathbb{K},\ 1 \leq i \leq n$, with at least one nonzero α_i such that*

$$\sum_{i=1}^{n} \alpha_i v_i = 0. \tag{1.1}$$

We say $v_1, \ldots, v_n$ are linearly independent if they are not linearly dependent; in other words, if (1.1) *implies $\alpha_i = 0$ for $i = 1, 2, \ldots, n$.*

This suggests that $v_1, \ldots, v_n$ demonstrate linear independence if and only if no vector can be expressed as a linear combination of the others, highlighting the lack of redundancy within the set.

Definition 2. *The span of $v_1, \ldots, v_n \in V$ is defined to be the set of all the linear combinations of these vectors:*

$$span\{v_1, \ldots, v_n\} = \left\{ \sum_{i=1}^{n} \alpha_i v_i | \alpha_i \in \mathbb{K},\ 1 \leq i \leq n \right\}.$$

If $W = \text{span}\{v_1, \ldots, v_n\} \subseteq V$, then $\{v_i\}_{i=1}^n$ forms a set of *bases* for the linear space W. Obviously, the set of bases is not unique. However, the dimension of W, that is, n, is always

the same. Furthermore, the linear space V is said to be *finite dimensional* if there exists a finite maximal set of independent vectors that reproduce V. Conversely, if such a finite set of bases for V does not exist, then V is considered infinite-dimensional.

The norm function defined on the V can be viewed as a mean to measure the distance between two elements of V.

Definition 3. *For a linear space V, a mapping $\|\cdot\|\colon V \to [0,\infty)$ is said to be a norm for V, if the following properties are satisfied for any $v \in V$.*

- $\|v\| = 0$ *if and only if* $v = 0$ *(definiteness),*
- $\|\alpha v\| = |\alpha|\|v\|$ *for all* $v \in V$ *and all* $\alpha \in \mathbb{R}$.
- $\|u + v| \leq \|u\| + \|v\|$ *for all* $u, v \in V$.

The space V equipped with the norm $\|\cdot\|$ is called a normed linear or normed space.

In particular, a sequence u_n in a normed space V is called *Cauchy* sequence if and only if for every $\varepsilon > 0$ there exists an integer N such that $\|u_m - u_n\| < \varepsilon$ holds whenever $m, n > N$. Besides, a subset W of normed space V is said to be *dense* in V if each $v \in V$ is the limit of a sequence of elements of W. Moreover, the normed space V is called *separable* if it has a countable dense subset.

Definition 4. *A normed space V is said to be* **complete** *and a* **Banach** *space if every Cauchy sequence in V converges to a limit in V.*

Remark 1. *[AF03, Chap. 1] Every normed space V is either a Banach space or a dense subset of Banach space S called the completion of V whose norm satisfies*

$$\|v\|_S = \|v\|_V \quad v \in V$$

It is important to note that defining different norms on any arbitrary linear space is possible. Each norm offers a unique size measure for a given vector within the space, leading to varying forms of convergence. Consequently, there is an interest in investigating whether these diverse norms can be interconnected or related. To this end, consider two different norms $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$, defined on the linear space V.

Definition 5. *We say, two norms $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$ are equivalent if there exist positive constants c_1, c_2 such that*

$$c_1\|v\|_{(1)} \leq \|v\|_{(2)} \leq c_2\|v\|_{(1)} \quad \forall v \in V. \tag{1.2}$$

With two such equivalent norms, a sequence $\{u\}_n$ converges in one norm if and only if it converges in the other norm:

$$\lim_{n\to\infty} \|u_n - u\|_{(1)} = 0 \iff \lim_{n\to\infty} \|u_n - u\|_{(2)} = 0. \tag{1.3}$$

Theorem 1. *[AH05, Theorem 1.2.13] Over a finite-dimensional space, any two norms are equivalent*

Notice that over an **infinite dimensional** space, however, such a statement is no longer valid (see [AH05, Chap. 1.2]).

Among all normed spaces, we turn our attention towards a specific type of normed space, known as *Hilbert* spaces. To this end, we review the concept of inner product spaces, where a norm can be established through the inner product, and the concept of *orthogonality* between two elements can be introduced.

Definition 6. *Let V be a linear space over $\mathbb{K}$. An inner product $(\cdot,\cdot)$ is a function from $V \times V$ to $\mathbb{K}$ with the following properties.*

- *For any $v \in V$, $(v,v) \geq 0$ and $(v,v) = 0$ if and only if $v = 0$.*
- *For any $u, v \in V$, $(u,v) = \overline{(v,u)}$.*
- *For any $u, v, w \in V$, any $\alpha, \beta \in \mathbb{K}$, $(\alpha u + \beta v, w) = \alpha(u,w) + \beta(v,w)$.*

The space V together with the inner product $(\cdot,\cdot)$ is called an *inner product space*. In particular, moreover, an inner product $(\cdot,\cdot)$ induces a norm through the formula

$$\|v\| = \sqrt{(v,v)} \quad v \in V. \tag{1.4}$$

It is easy to verify that such a norm satisfies the required properties. As an example, the inner product for the space $\mathbb{R}^d$ is

$$(x,y) = \sum_{i=1}^{d} x_i, y_i = \mathbf{y}^T\mathbf{x}, \quad \mathbf{x} = (x_1, ..., x_d)^T, \mathbf{y} = (y_1, ..., y_d)^T \in \mathbb{R}^d,$$

which induces the norm

$$\|\mathbf{x}\| = \sqrt{(\mathbf{x},\mathbf{x})} = \Big(\sum_{i=1}^{d} |x_i|^2\Big)^{1/2}.$$

Such a norm is known as the *Euclidean* norm, and the linear space equipped with a norm defined in (1.4), is called *Euclidean* space.

Theorem 2. *[I+18, Thm 3.9] Let V be a Euclidean space with inner product $(\cdot,\cdot)$ and norm $\|\cdot\| = (\cdot,\cdot)^{1/2}$. Then the* **parallelogram identity**

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2 \quad \forall u, v \in V \tag{1.5}$$

holds. Besides, If V is a Euclidean space over the real numbers $\mathbb{R}$, then the **polarization identity** *holds:*

$$(u,v) = 1/4(\|u+v\|^2 + \|u-v\|^2) \quad \forall u, v \in V. \tag{1.6}$$

For the statement in the above theorem, the converse is also true, known as the theorem of *Jordan and von Neumann*.

Theorem 3. *[I+18, Thm 3.10] Let V be a linear space with norm $(\cdot,\cdot)$, for which the parallelogram identity (1.5) holds. Then there is an inner product $(\cdot,\cdot)$ on V, so that*

$$(v,v) = \|v\|^2 \quad \forall v \in V, \tag{1.7}$$

that is, V is an Euclidean space.

The two theorems above imply that within any normed linear space, there can be only one inner product that generates the norm. Conversely, it's important to note that not every norm can be derived from an inner product. With the concept of an inner product at our disposal, we are able to define the angle between two non-zero vectors u and v.

$$\theta = \arccos\left[\frac{(u,v)}{\|u\|\|v\|}\right]. \tag{1.8}$$

Definition 7. *Two vectors u and v are said to be orthogonal if $(u,v) = 0$. An element $v \in V$ is said to be orthogonal to a subset $U \subset V$ if $(u,v) = 0$ for any $u \in U$.*

This notion leads us to an important set of bases:

Definition 8. *Let V be an inner product space. We say $\{v_i\}_{i\geq 1} \subset V$ forms an orthonormal system if*

$$(v_i, v_j) = \delta_{i,j}, \quad i,j \geq 1. \tag{1.9}$$

If the orthonormal system is a basis of V, then it is called an orthonormal basis for V.

In case the basis is not of length 1, they are called *orthogonal* basis instead. Equipped with the inner product spaces concept, we are now ready to define the *Hilbert* space.

Definition 9. *A complete inner product space is called a* **Hilbert** *space.*

Employing either orthogonal or orthonormal bases for a Hilbert space, we can decompose a vector as the linear combination of the basis elements.

Theorem 4. *[AH05, Theorem 1.3.11] Suppose $\{v_j\}_{j=1}^{\infty}$ is an orthonormal basis in a Hilbert space V. Then, we have the following conclusions.*

(a) For any $v \in V$, the series $\sum_{j=1}^{\infty}(v, v_j)v_j$ converges in V.

(b) if $v = \sum_{j=1}^{\infty} a_j v_j \in V$ then, $a_j = (v, v_j)$.

(c) A series $\sum_{j=1}^{\infty}$ converges in V if and only if $\sum_{j=1}^{\infty}|a_j^2| < \infty$.

The expansion expressed in b is called the *Fourier series* for v. There is a detailed discussion regarding this subject in any numerical analysis textbooks, including [Dav75, AH05, Lin19]. In what follows, we introduce some particular spaces that later appear necessary in this text.

1.3 Important function spaces

As the primary objective of approximation theory is to reconstruct unknown functions, it is essential to determine the space to which the target function belongs. To achieve this, we briefly introduce some well-known function spaces.

1.3.1 Spaces of continuously differentiable functions

Among these spaces, those of continuous and continuously differentiable functions hold substantial importance in approximation theory. In numerous instances, the underlying function that requires reconstruction belongs to one of these spaces.

Let Ω be an open, bounded, connected subset of $\mathbb{R}^d$. Recall that, a generic point in $\mathbb{R}^d$ is denoted by $\mathbf{x} = (x_1, ..., x_d)^T$. A multi-index is an ordered collection of d non-negative integers, $\delta = (\delta_1, \ldots, \delta_d)$. The quantity $|\delta| = \sum_{i=1}^{d} \delta_i$ is said to be the *length* of δ. Consequently, If v is an m-times differentiable function, then for any δ with $|\delta| \leq m$,

$$D^{\delta} v(\mathbf{x}) = \frac{\partial^{|\delta|} v(\mathbf{x})}{\partial x_1^{\delta_1} \ldots \partial x_d^{\delta_d}} \tag{1.10}$$

is the $\delta-$th order partial derivative. This will appear as a handy notation for partial derivatives. For example,

$$\frac{\partial v}{\partial x_1} = D^{\delta} v \quad \text{for } \delta = (1, 0, ..., 0). \tag{1.11}$$

Correspondingly, the set of all the derivatives of order m of a function v can be written as $\{D^{\delta} v \mid |\delta| = m\}$.

The space $C(\Omega)$ consists of all real-valued continuous functions on Ω. However, since Ω is open, a function from the space $C(\Omega)$ is not necessarily bounded. Thus, in many cases, one prefers to deal with $C(\bar{\Omega})$, which is the space of continuous functions that are continuous up to the boundary, meaning that any function in $C(\bar{\Omega})$ is bounded. So, the space $C(\bar{\Omega})$ equipped with its *canonical* norm

$$\|v\|_{C(\bar{\Omega})} = \sup\{|v(\mathbf{x})| \mid \mathbf{x} \in \Omega\} = \max\{v(\mathbf{x}) \in \bar{\Omega}\},$$

forms a normed space. Accordingly, take any $m \in \mathbb{Z}_+$, $C^m(\bar{\Omega})$ denotes the space of functions which, together with their derivatives of order less than or equal to m, are continuous and bounded:

$$C^m(\bar{\Omega}) = \{v \in C(\bar{\Omega}) | D^{\delta} v \in C(\bar{\Omega}) \text{ for } |\delta| \leq m\}.$$

The space $C^m(\bar{\Omega})$ is a Banach space with the norm

$$\|v\|_{C^m(\bar{\Omega})} = \max_{|\partial| \leq m} \|D^{\partial} v\|_{C(\bar{\Omega})}. \tag{1.12}$$

Generally, for $m = 0$, $C(\bar{\Omega})$ is used rather than $C^0(\bar{\Omega})$. Besides, the space of bounded and infinitely differentiable functions is defined as $\cap_{m=0}^{\infty} C^m(\bar{\Omega})$ and is denoted by $C^{\infty}(\bar{\Omega})$. For more information regarding these spaces, see [AF03, Chap. 1]. To proceed further, we need the following notions.

Given a function v on Ω, its support is defined to be

$$\text{supp } v = \{\mathbf{x} \in \Omega | \bar{v}(\mathbf{x}) \neq 0\}.$$

Definition 10. *A subset S on a normed space V is called compact if every sequence of points in S has a subsequence converging in V to an element S. Compact sets are closed and bounded, but the converse is true only in case V is finite-dimensional*

Consequently, the subspace $C_0(\Omega)$ and $C_0^{\infty}(\Omega)$ consist of all those functions in $C(\Omega)$ and $C^{\infty}(\Omega)$, respectively, that have compact support in Ω. We end this subsection with the following lemma, which is the result of *Stone–Weierstrass* theorem (see [AF03, Thm 3.1.2]).

Lemma 1. *If Ω is bounded in $\mathbb{R}^d$, then the set $\mathbb{P}$ of all $d-$variate polynomials having rational-complex coefficients is dense in $C(\bar{\Omega})$.*

Intuitively, this means that any continuous function f on a closed interval can be approximated arbitrarily well by algebraic polynomials no matter how badly f behaves. Mathematically speaking for any $f \in C[a, b]$ and $\varepsilon > 0$ there is a polynomial $p \in \mathbb{P}$ satisfying $\|p - f\|_{\infty} < \varepsilon$.

1.3.2 L^p SPACES

The next classes of linear spaces, which are important in approximation theory, are the L^p spaces. Note that we do not introduce formally and rigorously the concepts of measurable sets and functions, as they are out of the scope of this text; instead, the readers can find a complete development of Lebesgue measure and integration in any standard textbook on real analysis; for example, see [RF10, Rud87].

Let $\Omega \subset \mathbb{R}^d$ be a non-empty open set and p be a positive real number. We denote by $L^p(\Omega)$ the class of all measurable functions u defined on Ω for which

$$\int_{\Omega} |u(\mathbf{x})|^p d\mathbf{x} < \infty. \tag{1.13}$$

In the study of $L^p(\Omega)$ spaces, we identify functions that are equal almost everywhere in Ω. It means that the elements of $L^p(\Omega)$ are thus equivalence classes of measurable functions satisfy-

ing (1.13). Now the functional $\|\cdot\|$ defined

$$\|v\|_{L^p(\Omega)} = \left[\int_\Omega |v(\mathbf{x})|^p dx\right]^{1/p} < \infty \tag{1.14}$$

is a norm on $L^p(\Omega)$ provided $1 \leq p < \infty$. It is clear that $\|u\| \geq 0$ and $\|u\|_p = 0$ if and only if $u = 0$. Besides, $\|cu\|_p = |c|\|u\|_p$, $c \in \mathbb{R}$. The triangle inequality is the result of the following lemma.

Lemma 2. *(Minkowski inequality) If* $1 \leq p < \infty$

$$\|u + v\|_{L^p(\Omega)} \leq \|u\|_{L^p(\Omega)} + \|v\|_{L^p(\Omega)} \quad \forall u, v \in L^p(\Omega), \tag{1.15}$$

This implies that for any $p \in [1, \infty)$, $\|\cdot\|_{L^p}$ is indeed a norm. Moreover, a function u that is measurable on Ω is said to be essentially bounded on Ω if there is a constant K such that $\|u(\mathbf{x})\| \leq K$ almost everywhere on Ω. The greatest lower bound of such constants K is called the essential supremum of $|u|$ on Ω and is denoted by $ess\ \sup_{\mathbf{x}\in\Omega} |u(\mathbf{x})|$. We denote by $L^\infty(\Omega)$ the vector space of all functions u that are essentially bounded on Ω, functions being once again identified if they are equal almost everywhere on Ω. It is easy to verify that the functional $\|\cdot\|_\infty$ defined by

$$\|u\|_\infty = ess \sup_{\mathbf{x}\in\Omega} |u(\mathbf{x})|$$

is a norm on $L^\infty(\Omega)$.

Theorem 5. *[AFo3, Theorem 2.16] For* $p \in [1, \infty]$, $L^P(\Omega)$ *is a Banach space.*

Specifically, $L^2(\Omega)$ is a *Hilbert* space with respect to the inner product

$$(u, v) = \int_\Omega u(\mathbf{x})\overline{v(\mathbf{x})} d\mathbf{x}.$$

Theorem 6. *[AHo5, Theorem 1.5.6] Let* $\Omega \subset \mathbb{R}^d$ *be an open set,* $1 \leq p < \infty$. *Then the space* $C_0^\infty(\Omega)$ *is dense in* $L^p(\Omega)$*; in other words, for any* $v \in L^p(\Omega)$*, there exists a sequence* $\{v_n\} \subset C_0^\infty(\Omega)$ *such that*

$$\|v_n - v\|_{L^p(\Omega)} \to 0 \quad as\ n \to \infty, \tag{1.16}$$

For any $m \in \mathbb{Z}_+$*, the inclusions* $C_0^\infty \subset C^m(\bar{\Omega}) \subset L^p(\Omega)$ *hold, meaning that the space* $C^m(\bar{\Omega})$ *is also dense in* $L^p(\Omega)$.

1.3.3 ℓ^p SPACE

We denote by ℓ^p the set of doubly infinite sequence $v = \{v_i\}_{i=1}^{\infty}$ for which

$$\|v\|_{\ell^p} = \begin{cases} \left(\sum_{i=-\infty}^{\infty} |v_i|^p \right)^{1/p} & \text{if } 0 < p < \infty \\ \sup_{-\infty \le i \le \infty} |v_i| & \text{if } p = \infty \end{cases}$$

is finite. If $1 \le p \le \infty$, then ℓ^p is Banach space with respect to its norm. Moreover, the ℓ^p spaces are embedded into one another as follows. If $0 < p \le \infty$, then $\ell^p \to \ell^q$ and $\|v\|_{\ell^q} \le \|v\|_{\ell^p}$. For more information about discrete norm ℓ^p, see [AF03, Chap. 2].

1.4 OPERATORS ON NORMED SPACES

In formulating mathematical problems, one often needs to solve or minimize a linear system. Therefore, to assess the theoretical solvability of the obtained linear system, one should know the properties of the linear operators inherent in the problem.

1.4.1 LINEAR OPERATORS

Given two sets V and W, an operator T from V to W is a rule that assigns to each element in a subset of V a unique element in W. In addition, an operator is sometimes called a mapping, a transformation, or a function. Generally, it is assumed that both V and W are linear spaces. Besides, addition and scalar multiplication of operators are defined similarly to that of ordinary functions. Here, we focus on a particular type of operators, called *linear* operators.

Definition 11. *Let V and W be two linear spaces. An operator $L : V \to W$ is said to be linear if*

$$L(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 L(v_1) + \alpha_2 L(v_2) \quad \forall v_1, v_2 \in V, \forall \alpha_1, \alpha_2 \in \mathbb{R}, \tag{1.17}$$

For instance, it can be verified that the identity operator $I : V \to V$, defined as $I(v) = v$ for all $v \in V$, is a linear operator.

Let $\mathcal{L}(V, W)$ denote the set of all continuous linear operators from a normed space V to another normed space W. When $W = V$, we use $\mathcal{L}(V)$ instead of $\mathcal{L}(V, V)$. Particularly, if $L \in \mathcal{L}(V, W)$, it is meaningful to define the norm $\|L\|_{V,W}$ as follows:

$$\|L\|_{V,W} = \sup_{0 \ne v \in V} \frac{\|Lv\|_W}{\|v\|_V} \tag{1.18}$$

This definition allows us to verify that the set $\mathcal{L}(V, W)$ forms a linear space, and the norm (1.18) establishes a norm over this space. Commonly referred to as the operator norm of L, it

possesses the following compatibility property:

$$\|Lv\|_W \leq \|L\|_{V,W}\|v\|_V \quad \forall v \in V. \tag{1.19}$$

As an example, let $V = \mathbb{R}^n$, $W = \mathbb{R}^m$, and $L(v) = A\mathbf{v}$, where $A = (a_{ij}) \in \mathbb{R}^{m\times n}$. In this case, if the norms on V and W are $\|\cdot\|_\infty$ then the operator norm is the matrix $\infty-$norm,

$$\|A\|_\infty = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}|. \tag{1.20}$$

Moreover, if $\|\cdot\|_1$ is chosen instead, then the operator norm is the matrix $1-$norm,

$$\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|. \tag{1.21}$$

In particular, if $\|\cdot\|_2$ is taken, then the operator norm is known as the *spectral* norm that is

$$\|A\|_2 = \sqrt{r_\sigma(AA^*)},$$

where A^* denotes the conjugate transpose of A. For a square matrix B, $r_\sigma(B)$ denotes the spectral radius of the matrix B, $r_\sigma(B) = \max_{\lambda\in\sigma(B)}|\lambda|$, and the $\sigma(B)$ denotes the spectrum of B, the set of all the eigenvalues of B.

A significant aspect of operator theory comes into sight in approximation theory when solving the equation $Lv = w$, where $L \in \mathcal{L}(V, W)$.

Theorem 7. *[AH05, Theorem2.4.3] Let V and W be Banach spaces. If $L \in \mathcal{L}(V, W)$ is a bijection, then $L^{-1} \in \mathcal{L}(W, V)$.*

This implies that the solution v can be derived as $L^{-1}w$. Consequently, the *existence* and *uniqueness* of solutions $v \in V$ for all $w \in W$ suggests the *stability* of the solution v. This suggests that slight changes in the input data w lead to only minor adjustments in the solution v. To be more precise, let $Lv = w$ and $L\hat{v} = \hat{w}$. We then have

$$v - \hat{v} = L^{-1}(w - \hat{w}),$$

and then

$$\|v - \hat{v}\| \leq \|L^{-1}\|\|w - \hat{w}\|. \tag{1.22}$$

As $w - \hat{w}$ becomes small, so must $v - \hat{v}$.
Specifically, the expression $\|L^{-1}\|$ in (1.22) establishes a connection between the magnitude of the error in the data w and the magnitude of the error in the solution v More significantly, an important approach involves examining the relative changes in the two errors, guiding us

towards:

$$\frac{\|v-\hat{v}\|}{\|v\|} \leq \|L^{-1}\|\|L\|\frac{\|w-\hat{w}\|}{\|w\|}.$$

The quantity $\|L^{-1}\|\|L\|$ is called the *condition number* of the equation and is denoted by $cond(L)$ and links the relative errors in the data w and in the solution v. Specifically, we always have $cond(L) \geq 1$ as

$$\|L^{-1}\|\|L\| \geq \|L^{-1}L\| = \|I\| = 1.$$

Problems with a minor condition number are called *well-conditioned*, whereas those with a significant condition number *ill-conditioned*. In a related vein, problems in which L^{-1} is bounded (along with L) are called *well-posed* or *stable*; they can still be ill-conditioned, however. Later, we will see that such ill-conditioning can lead to inaccurate solutions; therefore, an important issue is finding proper ways to overcome ill-conditioning.

1.4.2 Riesz representation theorem

An important special class of linear operators is when they take on scalar values. Let V be a normed space, and $W = \mathbb{K}$, the set of scalars associated with V. The elements in $\mathcal{L}(V, \mathbb{K})$ are called *linear functionals*. This space is usually denoted as V^*, which is called the *dual space* of V. Usually, we use lowercase letters, such as ℓ, to denote a linear functional. Specifically, on *Hilbert* spaces, linear functionals are limited in the forms they can take.

Theorem 8. *(Riesz representation theorem) Let V be a real Hilbert space, $\ell \in V^*$. Then there is a unique $u \in V$ for which*

$$\ell(v) = (u, v) \quad \forall v \in V.$$

In addition,

$$\|\ell\| = \|u\|.$$

Let $\Omega \subset \mathbb{R}^d$ be open bounded. $V = L^2(\Omega)$ is a Hilbert space. By the *Riesz representation theorem*, there is a one-to-one correspondence between V and V^*, meaning that $\ell \in V^*$ with $u \in V$. In this sense, $(L^2(\Omega))^* = L^2(\Omega)$.

1.5 Function approximation

Equipped with the notions reviewed in the previous sections, we can represent the interpolation problem in an abstract setting. An abstract approximation problem can be stated in the following setting. Suppose V_n is an $n-$dimensional subspace of V, with a basis $\{v_1, \ldots, v_n\}$. Let $\ell_i \in V',\ 1 \leq i \leq n$, be n linear continuous functionals. Given n data $\ell_i f \in \mathbb{R},\ 1 \leq i \leq n$, we look for a function s_f such that

$$s_f = \sum_{i=1}^{n} c_i v_i, \tag{1.23}$$

with the property

$$\ell_i s_f = b_i \quad 1 \leq i \leq n. \tag{1.24}$$

In this, we say s_f interpolates f, and the property in (1.24) is known as *interpolation conditions*. Note that, we used the notation s_f instead of s to underline that it is an interpolant to the unknown function f. The coefficients c_i in (1.23), are determined by solving the linear system resulted from (1.24) i.e.,

$$\begin{pmatrix} \ell_1 v_1 & \cdots & \ell_1 v_n \\ \vdots & \ddots & \vdots \\ \ell_n v_1 & \cdots & \ell_n v_n \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} \ell_1 f \\ \vdots \\ \ell_n f \end{pmatrix}. \tag{1.25}$$

The following theorem states when the final linear system is uniquely solvable.

Theorem 9. *[Dav75, Theorem 2.2.2] Let a linear space V have dimension n and let $\ell_1, \ldots, \ell_n$ be n elements of V^*. The above interpolation problem possesses a solution for arbitrary values $b_1, \ldots, b_n$ if and only if the ℓ_i are independent in V^*. The solution will be unique*

The question of an error analysis in the abstract framework is rather difficult. For a general discussion of such error analysis, see [Dav75, Chap. 3]. Instead, in order to simplify the problem let $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\} \subset \Omega \subset \mathbb{R}^d$ be pairwise distinct points known as *interpolation point* or *data site*, and $\ell_i,\ 1 \leq i \leq n$ be the point evaluation functionals w.t.r to X meaning that $\ell_i s_f = s_f(\mathbf{x}_i)$ with collected function values $\ell_i f = f(\mathbf{x}_i),\ 1 \leq i \leq n$ in one data vector, denoted by $\mathbf{f}_X = (f_1, \ldots, f_n)^T \in \mathbb{R}^n$, where $f_i = f(\mathbf{x}_i),\ 1 \leq i \leq n$. Now, the interpolation system in (1.25) can be rewritten as

$$\begin{pmatrix} v_1(\mathbf{x}_1) & \cdots & v_n(\mathbf{x}_1) \\ \vdots & \ddots & \vdots \\ v_1(\mathbf{x}_n) & \cdots & v_n(\mathbf{x}_n) \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} f(\mathbf{x}_1) \\ \vdots \\ f(\mathbf{x}_n) \end{pmatrix}$$

or in the short form

$$V\mathbf{c} = \mathbf{f}_X, \tag{1.26}$$

with V referred as the *interpolation* matrix. We shall investigate the conditions that ensure the solvability of (1.26).

Remark 2. *In case $\ell_i\ 1 \leq i \leq n$ differ from the point evaluation functionals, and the given interpolation points are not necessarily distinct, we enter the realm of Hermite interpolation. The exact formulation of Hermite interpolation is out of the scope of this thesis, and the readers are referred to [Lin19, Chap 2] for more information.*

1.5.1 Linear least squares

Let us begin with a more general setting, where the number of bases is less than the given data, meaning that we have $V_m = span\{v_1, \dots, v_m\}$ with $m \leq n$. In the linear least squares approximation problem, we look for the best $s_f \in V_m$ which minimizes among all $v \in V_m$ the pointwise sum of square errors on X, so that

$$\|s_f - \mathbf{f}_X\|_2^2 \leq \|v - \mathbf{f}_X\|_2^2 \quad \forall v \in V_m. \tag{1.27}$$

To solve the minimization problem, we represent s_f as a unique linear combination of the basis functions in V_m, similar to (1.23) with the difference that now $1 \leq i \leq m$. Thereby, the linear least squares approximation problem can be reformulated as an equivalent minimization problem of the form

$$\min_{\mathbf{c}\in\mathbb{R}^m} \|V\mathbf{c} - \mathbf{f}_X\| \tag{1.28}$$

where $\mathbf{c} = (c_1, \dots, c_m)^T \in \mathbb{R}^m$, and $V \in \mathbb{R}^{m\times n}$ is known as the generalized *Vandermonde* matrix. The solution of the linear equation system is obtained via

$$V^T V\mathbf{c} = V^T \mathbf{f}_X \tag{1.29}$$

referred as *Gaussian normal equation*. If V has full rank, i.e., $rank(V) = m$, then the symmetric matrix $V^T V$ is *positive definite*, (see definition (16)), meaning that the existence of the unique solution is guaranteed. This gives us more freedom compared with (1.26) when it comes to selecting the set of bases since the solvability was ensured regardless of the $\{v_1, \dots, v_m\}$.

We mention that finding the coefficients $\mathbf{c}$ either through solving the linear system in (1.26) or (1.29) poses numerical challenges due to the ill-conditioning of the linear system. For a more in-depth error analysis on linear least squares approximation, see the textbook [Bjö96].
In what follows, we briefly review the required criteria of the approximation space and its norm, which guarantees the well-posedness of the problem (1.25), regardless of the selected set of basis $\{v_i\}$.

1.5.2 Best approximations

Let $\mathcal{F}$ be a linear space, equipped with a norm $\|\cdot\|$, and S be a non-empty subset of $\mathcal{F}$. To approximate one $f \in \mathcal{F}\backslash S$ by elements from S, we are interested in finding an approximant $s_f \in S$, whose distance to f is minimal among all elements from S. The **existence** and **uniqueness** of s_f is known as the *best approximation problem*. We will see that it is the properties of the linear space $\mathcal{F}$, its defined norm, and the approximation space S that are determinant. In particular, we restrict ourselves to the best approximation problem only in the **Hilbert** spaces.

Definition 12. *Let $\mathcal{F}$ and S be as above. The set S is said to be convex if*

$$u, v \in S \to \lambda u + (1-\lambda)v \in S \quad \forall \lambda \in (0,1).$$

Informally, the convexity of the set S is characterized by the property that the line segment joining any two elements of S is also contained in S.

Definition 13. *Let S be a convex set in a linear space $\mathcal{F}$. A function $f : S \to \mathbb{R}$ is said to be convex if*

$$f(\lambda u + (1-\lambda)v) \leq \lambda f(u) + (1-\lambda) f(v) \;\; \forall u, v \in S, \forall \lambda \in [0,1]. \tag{1.30}$$

The function f is **strictly convex** *if the above inequality is strict for $u \neq v$ and $\lambda \in (0,1)$.*

Now, we are ready for the following theorem.

Theorem 10. *[I$^+$18, Theorem 3.14] Let $\mathcal{F}$ be a Hilbert space with inner product $(\cdot,\cdot)$ and norm $\|\cdot\| = (\cdot,\cdot)^{1/2}$. Moreover, let $S \subset \mathcal{F}$ be a closed and convex subset of $\mathcal{F}$. Then there exists for any $f \in \mathcal{F}$ a best approximation $s_f \in S$ to f.*

Proof. Let $(s_n)_{n\in\mathbb{N}} \subset S$ be a minimal sequence in S, i.e.,

$$\|s_n - f\| \to \inf_{s\in S}(f, S) \quad \text{for } n \to \infty$$

with the minimal distance $\eta = \inf_{s\in S}(f, S)$. From the parallelogram identity (1.5), we obtain the estimate

$$\begin{aligned} \|s_n - s_m\|^2 &= 2\|s_n - f\|^2 + 2\|s_m - f\|^2 - 4\|\frac{s_n + s_m}{2} - f\|^2 \\ &\leq 2\|s_n - f\|^2 + 2\|s_m - f\|^2 - 4\eta^2. \end{aligned}$$

Therefore, for any $\varepsilon > 0$ there is one $N = N(\varepsilon) \in \mathbb{N}$ satisfying

$$\|s_n - s_m\| < \varepsilon \quad \text{for all } n, m \geq N,$$

i.e., $(s_n)_{n\in\mathbb{N}}$ is a **Cauchy sequence** in the *Hilbert* space $\mathcal{F}$, and therefore convergent in $\mathcal{F}$. Since S is a closed set, the limit element s^* lies in s, and we have

$$\eta = \lim_{n\to\infty} \|s_n - f\| = \|s^* - f\|,$$

i.e., $s^* = s_f \in S$ is a best approximation to f. □

We see that the parallelogram identity must hold; otherwise, the proof's scheme is under question. This highlights why approximation in Euclidean spaces fundamentally differs from other inner product spaces.

In the following discussion, we briefly review the uniqueness of the best approximant.

Definition 14. *A norm $\|\cdot\|$ is called* **strictly convex** *on $\mathcal{F}$, if the unit ball $\mathcal{B} = \{u \in \mathcal{F} \mid \|u\| \leq 1\} \subset \mathcal{F}$ is strictly convex.*

One should notice that not every norm is strictly convex. However, we can formulate those which are induced an inner product as follows

Theorem 11. *[I$^+$18, Theorem 3.28] Every Euclidean norm is strictly convex.*

The following theorem states *sufficient* conditions for the uniqueness of the best approximation.

Theorem 12. *[I$^+$18, Theorem 3.37] Let $\mathcal{F}$ be a linear space, equipped with a strictly convex norm $\|\cdot\|$. Moreover, assume $S \subset \mathcal{F}$ is convex and $f \in \mathcal{F}$. If there exists a best approximation $s_f \in S$ to f, then s_f is unique.*

Given the preceding discussion, the following theorem is an immediate conclusion.

Theorem 13. *[AH05, Theorem 3.4.3] Assume $S \subset \mathcal{F}$ is a non-empty, closed, convex subset of a Hilbert space $\mathcal{F}$. Then for any $f \in \mathcal{F}$, there is a unique element s_f such that*

$$\|f - s_f\| = \inf_{s \in \mathcal{F}} \|f - s\|$$

i.e., s_f is the best approximation to f.

Before ending this subsection, we present the following theorems regarding the L^p and ℓ^p norms.

Theorem 14. *[I$^+$18, Theorem 3.32 & Corollary 3.35] For $1 < p < \infty$, the L^p- and ℓ^p-norms are strictly convex on L^p and $\mathbb{R}^d$.*

Subsequently, we explore how to construct the best approximant.

1.5.3 Reconstruction of the Best Approximant

In this subsection, we intend to explain further advantages of inner product space to characterize and construct the best approximation. So the main purpose of this subsection is to yield for finite-dimensional approximation spaces $S \subset \mathcal{F}$ constructive methods to compute best approximations by orthogonal projection $\Pi : \mathcal{F} \to \mathcal{S}$ of $f \in \mathcal{F}$ on S.
Intuitively speaking, assume s_1 and s_2 be two non zero and independent elements of S. We look for the scalar $\lambda \in \mathbb{R}$ such that λs_2 is the projection of s_1 on s_2. Then

$$\lambda s_2 \perp s_1 - \lambda s_2.$$

Considering (1.8), this means

$$(\lambda s_2, s_1 - \lambda s_2) = 0.$$

Therefore, we have $\lambda = \frac{(s_2,s_1)}{(s_2,s_2)}$. This means that if $s_2 \neq 0$,

$$\text{projection of } s_1 \text{ on } s_2 = \frac{(s_2, s_1)}{(s_2, s_2)} s_2.$$

Consequently, it seems reasonable to expect the approximation error, i.e., $|f - s_f|$, to be an orthogonal projection onto the approximation space. Without going into details, we immediately represent the following theorem, which characterizes the best approximation,

Theorem 15. *[AH05, Theorem 3.4.6] Assume S is a complete subspace of an inner product space $\mathcal{F}$. Then for any $f \in \mathcal{F}$, there is a unique element $s_f \in S$ such that*

$$\|f - s_f\| = \inf_{s \in S} \|f - s\|$$

Moreover, the best approximation $s_f \in S$ is characterized by the property

$$(f - s_f, s) = 0 \quad \forall s \in S, \tag{1.31}$$

In other terms, the best approximation $s_f \in S$ to $f \in \mathcal{F}$ is the unique orthogonal projection of f onto S. Fixing $\{s_1, \ldots, s_n\}$ as a set of bases for S, and employing the representation in (1.23) for s_f, we have

$$\begin{aligned} (f - \sum_{i=1}^{n} c_i s_i, s_j) = 0 \quad 1 \leq j \leq n \\ \sum_{i=1}^{n} c_i (s_i, s_j) = (f, s_j) \quad 1 \leq j \leq n \end{aligned}$$

Therefore, the vector of the coefficients can be obtained by solving the following linear system.

$$G\mathbf{c} = \mathbf{b}, \tag{1.32}$$

with the *Gram matrix* $G = ((s_j, s_i))_{1 \leq i,j \leq n} \in \mathbb{R}^{n \times n}$, the unknown coefficient vector $\mathbf{c} = (c_1, \ldots, c_n)^T \in \mathbb{R}^n$ and the right hand side $\mathbf{b} = ((f, s_1), \ldots, (f, s_n))^T \in \mathbb{R}^n$. In particular, due to the existence and uniqueness of the best approximation, the linear system (1.32) must be uniquely solvable. We specialize this statement on G as follows.

Theorem 16. *[I+18, Theorem 4.4] The Gram matrix G is symmetric positive definite (see definition (16)).*

Though we do not mention the exact proof, we highlight that the symmetry of G is concluded from the symmetry of the inner product. At the same time, the positive definiteness

of G is an immediate result of the properties of the inner product. Given our investigations in this, the problem of Euclidean approximation by finite-dimensional approximation spaces S seems to be solved. But note that we have not posed any essential or preferred conditions on the set of basis $\{s_1, \ldots, s_n\}$, yet.

1.5.4 Orthogonal approximation

Now we discuss a particular case when the set of basis poses the property of **orthogonality**. For an orthogonal basis $\{s_1, \ldots, s_n\}$ of S, i.e.,

$$(s_i, s_j) = \begin{cases} 0 & \text{for } i \neq j \\ \|s_i\|^2 & \text{for } i = j \end{cases} \tag{1.33}$$

the Gram matrix G is a diagonal matrix,

$$G = diag(\|s_1\|^2, \ldots, \|s_n\|^2), \tag{1.34}$$

in which case, the solution

$$\mathbf{c} = \left(\frac{(f, s_1)}{\|s_1\|^2}, \cdots, \frac{(f, s_n)}{\|s_n\|^2} \right) \in \mathbb{R}^n.$$

Notice that if $\{s_1, \ldots, s_n\}$ are **orthonormal** instead, $\|s_i\|^2 = 1,\ 1 \leq i \leq n$; therefore the inner product of the function values and the bases functions would be enough to compute the approximant. We formulate our discussion in the following theorem, similar to Theorem (4).

Theorem 17. *[I+18, Theorem 4.5] Let $\mathcal{F}$ be an Euclidean space with inner product $(\cdot, \cdot)$. Moreover, let $S \subset \mathcal{F}$ be a finite-dimensional linear subspace with orthogonal basis $\{s_1, \ldots, s_n\}$. Then, for any $f \in \mathcal{F}$,*

$$s_f = \sum_{i=1}^{n} \frac{(f, s_j)}{\|s_j\|} s_i \ \in S \tag{1.35}$$

is the unique best approximation to f.

Generally speaking, obtaining the interpolant representation as in (1.35), i.e., choosing a set of basis functions with orthogonality properties, is not always straightforward. For instance, applying the *Gram–Schmidt* procedure yields the familiar *Legendre* polynomials when utilizing monomials as the bases. However, constructing orthonormal polynomials through this method can be cumbersome in numerous scenarios. Thus, in many cases, the linear system (1.26) is preferred over dealing with (1.35). Consequently, the existence, uniqueness, and optimality of the approximation hinge on the invertibility and conditioning of the interpolation matrix in (1.26). Hence, the following chapter identifies bases that ensure the appropriate solution of equation (1.26).

1.6 Thesis outline

In this direction, in the second next 2, we review a specific type of function approximation, known as kernel-based approximation, particularly with Radial Basis Functions (RBF). We mainly focus on the properties and characteristics of such an approximation rather than the construction of these radial kernels. The last three chapters describe my research and new contributions to the field during my Ph.D.

In Chapter 3, we explored two potential methods for finding data-dependent orthonormal bases that are orthonormal for both discrete and native space norms. The first method involves kernel matrix factorization, while the second is based on approximating the eigenpairs of the linear operator linked to the reproducing kernel, as Mercer's theorem states. To identify an optimal low-rank basis, we use the truncated Singular Value Decomposition technique with a coefficient matrix whose rank is lower than the original matrix. Our study also delves into error analysis, duality, and stability and includes several numerical experiments.

Chapter 4 is dedicated to developing a new numerical scheme based on the Moving Least Squares approach to approximate the discontinuous underlying function, such that the discontinuities are incorporated into the approximant. We also provide an error estimate on a suitable piecewise Sobolev Space. The numerical experiments comply with the theoretical convergence rate.

Finally, in Chapter 5, we suggest a new numerical scheme based on the closest point representation of a surface to solve time-dependent PDE problems intrinsic to the surfaces.

2

Kernel Based Approximation Methods

> Mathematics may be defined as the subject in which we never know what we are talking about, nor whether what we say is true.
>
> Bertrand Russell

2.1 Kernels as Basis function

This chapter discusses a class of basis functions known as *Radial Basis Functions* to construct the approximation. Afterward, we investigate the relation of these bases with the reproducing kernel of Hilbert spaces. The materials of this chapter are mainly derived from [Fas07, Wen04, Buh03, FM15].

We continue to use our assumptions i.e., $S \subset \Omega \subset \mathbb{R}^d$ where $S = span\{s_1, \ldots, s_n\}$ for $n \in \mathbb{N}$, and $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ and $\mathbf{f}_X = (f_1, \ldots, f_n)$ are the given data. As mentioned before, the unique solution of linear system (1.26) exists if and only if the interpolation matrix $V_{ij} = s_i(\mathbf{x}_j)$ is *non-singular*. Consequently, we look for the bases that ensure this for any data sites X. By introducing *Haar* spaces, we specialize our assumptions on S and $\{s_1, \ldots, s_n\}$ as follows.

Definition 15. *Let the finite-dimensional linear function space $S \subset C(\Omega)$ have a basis $\{s_1, \ldots, s_n\}$. Then S is a* **Haar** *space on Ω if*

$$det V \neq 0$$

for any set of distinct points X. (V is defined as (1.26)).

Note that the existence of a *Haar* space guarantees the invertibility of the interpolation matrix. Now, we express the well-known theorem referred to as **Mairhuber-Curtis** theorem.

Theorem 18. *If $\Omega \subset \mathbb{R}^d$, $d \geq 2$, contains an interior point, then there exist no Haar spaces of continuous functions except for one-dimensional ones.*

Proof. Let $d \geq 2$ and assume that S is a *Haar* space with basis $\{s_1, \ldots, s_n\}$ with $n \geq 2$. We need to show that this leads to a contradiction. Then, by the definition of a Haar space, we have

$$detV \neq 0.$$

Consider a closed path P in Ω that connects only $\mathbf{x}_1$ and $\mathbf{x}_2$. This is feasible since, by assumption, Ω contains an interior point. By moving $\mathbf{x}_1$ and $\mathbf{x}_2$ continuously along the path P without disturbing any of the other $\mathbf{x}_j$, we effectively swap the positions of $\mathbf{x}_1$ and $\mathbf{x}_2$. This swap causes the first and second rows of the determinant of V to be exchanged, thus changing the sign of the determinant. Given that the determinant is a continuous function of $\mathbf{x}_1$ and $\mathbf{x}_2$, it follows that $detV$ must be zero at some point along the path P. This contradicts the assumption of $detV \neq 0$. □

The Mairhuber-Curtis theorem states that to guarantee the well-posedness of a multivariate scattered data interpolation problem, one cannot pre-select a set of basis functions to solve the interpolation problem for arbitrary scattered data. For instance, it is not feasible to achieve unique interpolation using (multivariate) polynomials of degree m for data provided at arbitrary locations in $\mathbb{R}^2$. Instead, the basis must be chosen based on the locations of the data. To represent such a class of basis, we begin with the following definition,

Definition 16. *A real symmetric matrix $A \in \mathbb{R}^{n\times n}$ is called* **positive definite** *if its associated quadratic form is positive for any nonzero coefficient vector $\mathbf{c} = (c_1, \ldots, c_n)^T \in \mathbb{R}^n$, i.e.,*

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_j c_i A_{i,j} > 0.$$

The matrix is called **positive semi-definite** *if the quadratic form is allowed to be nonnegative.*

A significant characteristic of positive definite matrices is that all of their eigenvalues are positive; hence, a positive definite matrix is always non-singular. Consequently, our focus will primarily be kernels that produce positive definite matrices.

Definition 17. *A symmetric kernel Φ is called positive definite on $\Omega \times \Omega$ if its assocciated interpolation matrix V with entries $V_{ij} = \Phi(\mathbf{x}_i, \mathbf{x}_j),\ i, j = 1, \ldots, n$ is positive definite for any $n \in \mathbb{N}$ and X of distinct points.*

From now on, we use the notation $V_{\Phi,X}$ instead of V for the interpolation matrix to insist that this matrix has resulted from kernel Φ and not any arbitrary set of basis. In this particular case (1.26) could be rewritten as

$$V_{\Phi,X}\mathbf{c} = \mathbf{f}_X. \tag{2.1}$$

In the literature $V_{\Phi,X}$ is also called the **kernel** matrix.
Besides, knowing that the interpolant must depend on the data sites, we denote the interpolant with $s_{f,X}$ instead of s_f. Now, we summarize the properties of kernel interpolation.

Theorem 19. *[I+18, Theorem 8.3] For Φ being a PD (positive definite) kernel, and $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \subset \mathbb{R}^d$, the following statements are true.*

1. *The interpolation matrix $V_{\Phi,X}$ is positive definite i.e., linear system in (1.26) is uniquely solvable.*
2. *If $s_{f,X} \in S$ vanishes on X, i.e., if $s_{f,X}(X) = 0$, then $s_{f,X} = 0$.*
3. *The interpolation problem has a unique solution $s_{f,X} \in S$ represented as $s_{f,X} = \sum_{i=1}^{n} c_i \Phi(\cdot, \mathbf{x}_i)$, whose coefficient vector $\mathbf{c} = (c_1, \ldots, c_n)^T \in \mathbb{R}^n$ is determined by the unique solution of the linear system $V_{\Phi,X}\mathbf{c} = \mathbf{f}_X$.*

Before continuing, we investigate another class of kernels known as *conditionally positive definite* (CPD) kernels. To pave the way, we need the following notion. Besides, we denote the space of $d-$dimensional polynomial of degree at most $m-1$ by $\mathbb{P}^d_{m-1}$.

Definition 18. *We call a set of points $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \subset \mathbb{R}^d$, $\mathbb{P}^d_{m-1}-$unisolvent if the only polynomial of total degree at most $m-1$ interpolating zero data on X is the zero polynomial.*

Definition 19. *Let $\Phi : \Omega \times \Omega \to \mathbb{R}$ be a continuous symmetric kernel. It is said that Φ is a conditionally positive semi-definite kernel of order m on $\Omega \subset \mathbb{R}^d$ if, for all $n \in \mathbb{N}$, all pairwise distinct centers $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \subset \Omega$, and all $\alpha \in \mathbb{R}^n$ satisfying*

$$\sum_{j=1}^{n} \alpha_j p(\mathbf{x}_j) = 0, \quad p \in \mathbb{P}^d_{m-1},$$

the quadratic form

$$\sum_{i,j=1}^{n} \alpha_i \alpha_j \Phi(\mathbf{x}_i, \mathbf{x}_j) \geq 0.$$

Moreover, Φ is said to be conditionally positive definite (CPD) of order m if equality holds only for $\alpha = 0$. Notice that letting $m = 0$ then Φ is positive (semi)-definite

In this case letting $p_1, \ldots, p_q$ be a basis for the polynomial space $\mathbb{P}^d_{m-1}$, the interpolant can be represented as

$$s_{f,X}(\mathbf{x}) = \sum_{j=1}^{n} c_j \Phi(\mathbf{x}, \mathbf{x}_j) + \sum_{j=1}^{q} d_j p_j(\mathbf{x}), \ \forall\, \mathbf{x} \in \mathbb{R}^d. \tag{2.2}$$

In order to compute the coefficients c_j and d_j in (2.2), the interpolation condition leads us to the linear system

$$\begin{bmatrix} V_{\Phi,X} & P \\ P^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_X \\ 0 \end{bmatrix}, \tag{2.3}$$

with $V_{\Phi,X} = [\Phi(\mathbf{x}_i, \mathbf{x}_j)]_{1 \leq i,j \leq n}$, $P_{ij} = p_j(\mathbf{x}_i), 1 \leq i \leq n,\ 1 \leq j \leq q$, and $\mathbf{f}_X = [f(\mathbf{x}_i)]_{1 \leq i \leq n}$.

Theorem 20. *[Wen04, Theorem 8.21] Suppose that Φ is CPD kernel of order m and X is a $\mathbb{P}^d_{m-1}-$unisolvent set of centers. Then the system* (2.3) *is uniquely solvable.*

It is obvious that the addition of polynomial terms to the expansion guarantees polynomial reproduction, i.e., if the data come from a polynomial of total degree less than m, then they are fitted by that polynomial.
Now, we classify the general notion of kernel functions with some additional structure.

Definition 20. *A kernel is called*

- *translation-invariant if it depends only on the difference of the two arguments, i.e., if there is a function $\varphi(\mathbf{x}, \mathbf{y}) : \mathbb{R}^d \to \mathbb{R}$ such that*

$$\Phi(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x} - \mathbf{y}) \quad \textit{for all } \mathbf{x}, \mathbf{y} \in \Omega. \tag{2.4}$$

- *zonal if it depends only on the inner product of the two arguments, i.e., if there is a function $\varphi(\mathbf{x}, \mathbf{y}) : \mathbb{R} \to \mathbb{R}$ such that*

$$\Phi(\mathbf{x}, \mathbf{y}) = \phi\left(\langle \mathbf{x}, \mathbf{y} \rangle\right) \quad \textit{for all } \mathbf{x}, \mathbf{y} \in \Omega. \tag{2.5}$$

- *radial if it depends only on the Euclidian norm of the difference of the two arguments, i.e., if there is a function $\varphi(\mathbf{x}, \mathbf{y}) : [0, \infty) \to \mathbb{R}$ such that*

$$\Phi(\mathbf{x}, \mathbf{y}) = \phi(\|\mathbf{x} - \mathbf{y}\|_2) \quad \textit{for all } \mathbf{x}, \mathbf{y} \in \Omega. \tag{2.6}$$

Here, we delve into the special case of radial kernel referred to as *radial basis functions* (RBFs), which may be one of the most popular types of kernels discussed in the literature. In particular, for a radial function Φ and three distinct points $\mathbf{z}_1, \mathbf{z}_2, \mathbf{x}$ we have

$$\|\mathbf{z}_1 - \mathbf{x}\| = \|\mathbf{z}_2 - \mathbf{x}\| \to \Phi(\mathbf{z}_1, \mathbf{x}) = \Phi(\mathbf{z}_2, \mathbf{x}), \quad \mathbf{x}, \mathbf{z}_1, \mathbf{z}_2 \in \mathbb{R}^d.$$

In other words, the value of Φ at any point at a certain fixed distance from the origin (or any other fixed center point) is constant.

2.2 Radial Basis Functions

Employing the Radial Basis Function, the approximation space can be spanned by $\{\phi(\|\cdot - \mathbf{x}_i\|);\ i = 1, \ldots, n\}$, possibly added by polynomial space, that is generating the basis function using one basic function ϕ. Radial function interpolants are invariant under all Euclidean transformations (i.e., translations, rotations, and reflections). Moreover, the application of RBFs to the interpolation problem benefits from the fact that the interpolation problem becomes insensitive to the dimension d of the space in which the data sites lie because we work with the same

Table 2.1: Matérn functions for various choice of β, [Fas07, Table 4.3]

$\beta = \frac{d+1}{2}$	$\beta = \frac{d+3}{2}$	$\beta = \frac{d+5}{2}$
$e^{-\varepsilon r}$	$(1+\varepsilon r)e^{-\varepsilon r}$	$(3+3\varepsilon r+\varepsilon^2 r^2)e^{-r}$

univariate function ϕ for all choices of d. Here, we avoid the fundamental discussion regarding the construction of the radial basis function, and we refer the readers to [Fas07, Wen04, Buh03]. Instead, we recall some of the most common and well-known PD and CPD RBFs. Before proceeding, we underline the additional parameter $\varepsilon > 0$ known as *shape parameter* appearing in some of the introduced RBFs. We shall discuss its effects on both the accuracy and stability of RBF approximations. Letting $r = \|\mathbf{x} - \mathbf{y}\|$ for any $\mathbf{x}, \mathbf{y} \in \Omega$, we list some globally supported RBFs :

- The **Gaussian** function
$$\phi_\varepsilon(r) = \exp(-\varepsilon^2 r^2),$$
which is positive definite (and radial) on $\mathbb{R}^d$ for any d.

- The **generalized inverse multiquadrics**
$$\phi_\varepsilon(r) = (1+(\varepsilon r)^2)^{-\beta} \quad \beta > 0,$$
which are positive definite on $\mathbb{R}^d$ for all d and infinitely differentiable.

- Another example of PD kernels is given by the class of **Matérn** function,
$$\phi_\varepsilon(r) = \frac{K_{\beta-\frac{d}{2}}(\varepsilon r)(\varepsilon r)^{\beta-\frac{d}{2}}}{2^{\beta-1}\Gamma(B)} \quad \beta > \frac{d}{2}.$$
Here K_ν is the *modified Bessel* function of the second kind of order ν. Some simple representatives of the family of Matérn functions are listed in Table(2.1).

To proceed, we investigate another critical class of positive definite RBFs: *compactly supported* Radial Basis Functions. Similarly, we avoid the characterization and construction of these functions.

Definition 21. *Let ϕ be such that $t \to t\varphi(t) \in L_1[0,\infty)$. Then we define the integral operator $\mathcal{I}$ via*
$$(\mathcal{I}\phi)(r) = \int_r^\infty t\phi(t)dt, \quad r \geq 0.$$

Employing the operator $\mathcal{I}$ represented firstly in [Zm95], the well-known and popular family of *Wendland's compactly supported functions* [Wen95, Wen98] is obtained accordingly,

Definition 22. *Let $\phi_\ell(r) = (1-r)_+^\ell$ with $\ell \geq \lfloor \frac{d}{2} \rfloor + 1$, one can define*
$$\phi_{d,k} = \mathcal{I}^k \phi_{\lfloor d/2 \rfloor + k + 1}.$$

It turns out that the functions $\phi_{d,k}$ are all supported on $[0,1]$ and have a polynomial representation there.

Theorem 21. *[Fas07, Theorem 11.4] The functions $\phi_{d,k}$, $k = 0, 1$ have the form*

$$\phi_{d,0,\varepsilon}(r) = (1-\varepsilon r)_+^\ell,$$
$$\phi_{d,1,\varepsilon}(r) = (1-\varepsilon r)^{\ell+1}[(\ell+1)(\varepsilon r)+1]$$

For more information regarding other classes of CS functions (e.g., Wu's CS functions), see ([Wen04, Chap. 9]) or [Fas07, Chap. 11].

In what follows, we list some of the most commonly used CPD RBFs.

- generalized Multiquadrics (MQ)

  $$\phi_\varepsilon(r) = (-1)^{\lceil\beta\rceil}\left(1+(\varepsilon r)^2\right)^\beta$$

  which are conditionally positive definite of order $\lceil\beta\rceil$ for $0 < \beta \notin \mathbb{N}$.

- Radial powers

  $$\phi(r) = (-1)^{\lceil\frac{\beta}{2}\rceil} r^\beta$$

  which are conditionally positive definite of order $\lceil\beta/2\rceil$ for $0 < \beta \notin 2\mathbb{N}$.

- Thin-plate splines (TPS)

  $$\phi(r) = (-1)^{\beta+1} r^{2\beta}\log(r)$$

  which are conditionally positive definite of order $\beta+1$ with $\beta \in \mathbb{N}$.

Back to the shape parameter ε, generally speaking, it can be thought of as an inverse length scale, meaning that a smaller value of ε (i.e., larger variance) causes the function to become *flatter* while increasing ε leads to a more peaked RBF and therefore localizes its influence. In a nutshell, in the framework of RBF approximation, optimizing the shape parameter value is always required as it affects both the accuracy and stability of the approximant. However, in many cases, finding the best value for the shape parameter can become a challenging task. See, e.g., [FZ07] and the reference therein. To illustrate, it is known (see, e.g., [Fas07, Chap. 16]) that employing Gaussian RBFs, the best accuracy is obtained when $\varepsilon \to 0$. On the other hand, for small values of ε, the basic functions increasingly resemble constant functions. Consequently, both the rows and columns of the matrix $V_{\Phi,X}$ in equation (2.1) tend to become more similar so that the matrix becomes almost singular - even for well-separated data sites, which leads to an ill-conditioned linear system. This phenomenon, firstly mentioned in [S+95], can be generalized to the other types of RBFs and is known as the *uncertainty* or *trade-off principle*.

2.3 Native spaces of kernels

In this subsection, we collect some basic facts regarding a specific space of functions associated with each PD or CPD radial basis function, called the *native space*, and establish a connection to reproducing kernel Hilbert spaces. We will not give any proof but refer to the literature instead. Fundamental is the monograph [Weno4], which contains most of the facts collected here and offers some broader background.

Definition 23. *Let $\mathcal{H}$ be a real Hilbert space of functions $f : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ with inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$. A function $K : \Omega \times \Omega \to \mathbb{R}$ is called reproducing kernel for $\mathcal{H}$ if*

1. *$K(\cdot, \mathbf{x}) \in \mathcal{H}$ for all $\mathbf{x} \in \Omega$,*
2. *$f(\mathbf{x}) = \langle f, K(\cdot, \mathbf{x}) \rangle_{\mathcal{H}}$ for all $f \in \mathcal{H}$ and all $\mathbf{x} \in \Omega$.*

The second property in the above definition inspires the name reproducing kernel, yielding its uniqueness.

Theorem 22. *The reproducing kernel in a reproducing kernel Hilbert space is uniquely defined.*

Proof. Assume there are two reproducing kernels K_1, K_2. By reproducing property, we have

$$\langle f, K_1(\cdot, \mathbf{y}) - K_2(\cdot, \mathbf{y}) \rangle_{\mathcal{H}} = 0$$

for all $f \in \mathcal{H}(\Omega)$ and all $\mathbf{y} \in \Omega$. Inserting $f = K_1(\cdot, \mathbf{y}) - K_2(\cdot, \mathbf{y})$ concludes the proof. □

Theorem 23. *[Weno4, Theorem. 10.2] Suppose that $\mathcal{H}(\Omega)$ is a Hilbert space of functions $f : \Omega \to \mathbb{R}$. The following statements are equivalent:*

- *The point evaluation functionals are continuous, i.e., $\delta_{\mathbf{x}} \in \mathcal{H}(\Omega)$ for all $\mathbf{x} \in \Omega$.*
- *$\mathcal{H}(\Omega)$ has a reproducing kernel.*

It is possible to show that there is a one-to-one correspondence between reproducing kernel *Hilbert* spaces and positive semi-definite kernels. We shall explain this connection in more detail. Recall that $\mathcal{H}^*$ is the dual of $\mathcal{H}$; the space of bounded linear functionals on $\mathcal{H}$.

Theorem 24. *[Weno4, Theorem. 10.4] Suppose $\mathcal{H}$ is a reproducing kernel Hilbert function space with reproducing kernel $K : \Omega \times \Omega$. Then K is semi-positive definite. Moreover, K is positive definite if and only if the point evaluation functionals $\delta_{\mathbf{x}}$ are linearly independent in $\mathcal{H}^*$.*

This theorem provides one direction of the connection between positive definite kernels and reproducing kernels. Since in this thesis, we mainly work with the RBFs, either PD or CPD radial functions; we want to know how to construct a reproducing kernel Hilbert space associated with those basic functions. This opens up our way to the notion of *native space*. Thus, we replace the kernel K with PD radial function Φ. Consequently, we start by:

$$H_\Phi(\Omega) := span\{\Phi(\cdot, \mathbf{y}) \ : \ \mathbf{y} \in \Omega\}, \tag{2.7}$$

equipped with the bilinear form

$$\langle \sum_{j=1}^{N} \alpha_j \Phi(\cdot, \mathbf{x}_j), \sum_{k=1}^{M} \beta_k \Phi(\cdot, \mathbf{y}_k) \rangle_\Phi := \sum_{j=1}^{N} \sum_{k=1}^{M} \alpha_j \beta_k \Phi(\mathbf{x}_j, \mathbf{y}_k),$$

where $N, M = \infty$ are allowed. By this construction $H_\Phi(\Omega)$ forms a pre-Hilbert space concerning the norm $\|\cdot\|_\Phi$. As a result, one should look for the completion of $H_\Phi(\Omega)$ denoted by $\mathcal{H}_\Phi(\Omega)$, i.e., the Hilbert function space with reproducing kernel Φ. To this end, it can be shown [Weno4, Lemma 10.8] that the map

$$\mathcal{R} : \mathcal{H}_\Phi(\Omega) \to C(\Omega), \ \mathcal{R}(f)(\mathbf{x}) := (f, K(\cdot - \mathbf{x}))_\Phi \tag{2.8}$$

is well-defined and injective.

Definition 24. *The native space $\mathcal{N}_\Phi(\Omega)$ of positive definite kernel Φ is given by*

$$\mathcal{N}_\Phi(\Omega) = \mathcal{R}(\mathcal{H}_\Phi(\Omega)) \tag{2.9}$$

equipped with the inner product

$$(f, g)_{\mathcal{N}_\Phi(\Omega)} := (\mathcal{R}^{-1} f, \mathcal{R}^{-1} g)_\Phi. \tag{2.10}$$

Accordingly, Theorem [Weno4, Theorem 10.10] states that $\mathcal{N}_\Phi$ is indeed a Hilbert function space with reproducing kernel Φ. As a result, we have

$$\|f\|_\Phi = \|f\|_{\mathcal{N}_\Phi}, \quad \text{for all } f \in H_\Phi(\Omega),$$

where $H_\Phi(\Omega)$ is defined as in (2.7). The following theorems are related to the uniqueness of native space associated with Φ.

Theorem 25. *[Weno4, Theorem 10.11] Suppose that Φ is a symmetric positive definite kernel. Suppose further that $\mathcal{G}$ is a reproducing kernel Hilbert space of functions on Ω with reproducing kernel Φ. Then $\mathcal{G}$ is the native space $\mathcal{N}_\Phi(\Omega)$ for Φ, and the inner products are the same.*

As in the case of a positive definite kernel, we proceed to construct the native space of a conditionally positive kernel. However, since this discussion is very technical, we only mentioned

a few results. Similar to (2.7), we start with the linear space

$$H_\Phi(\Omega) := span\{\sum_{j=1}^{N} \alpha_j \Phi(\cdot, \mathbf{x}_j) : N \in \mathbb{N}, \alpha \in \mathbb{R}^N,\ \mathbf{x}_1, \dots, \mathbf{x}_N \in \Omega,$$
$$\text{with} \sum_{j=1}^{N} \alpha_j p(\mathbf{x}_j) = 0 \text{ for all } p \in \mathbb{P}^d_{m-1}\}.$$

Notice that such a linear space is different from (2.7). However, we keep using the same notation, H_Φ, for the sake of simplicity. Analogous to (2.8), let

$$R : \mathcal{H}_\Phi(\Omega) \to C(\Omega),$$
$$R(f(\mathbf{x})) = f(\mathbf{x}) - \Pi f(\mathbf{x}) = f(\mathbf{x}) - \sum_{k=1}^{q} f(\xi_k) l_k(\mathbf{x}),$$

where $l_k,\ 1 \le k \le q$, are the Lagrange basis of $\mathbb{P}^d_{m-1}$ for the points $\Xi = \{\xi_1, \dots, \xi_q\}$ which is assumed to be a $\mathbb{P}^d_{m-1}$-unisolvent subset of X with $\dim(\mathbb{P}^d_{m-1}) = q$

Definition 25. *The native space corresponding to a symmetric kernel Φ that is CPD of order m on Ω is defined by*

$$\mathcal{N}_\Phi(\Omega) = R(\mathcal{H}_\Phi(\Omega)) \oplus \mathbb{P}^d_{m-1},$$

equipped with the inner product

$$\langle f, g \rangle_{\mathcal{N}_\Phi} = \langle f, g \rangle + \sum_{k=1}^{q} f(\xi_k) g(\xi_k),$$

where

$$\langle f, g \rangle = \left\langle R^{-1}(f - \Pi f), R^{-1}(g - \Pi g) \right\rangle_\Phi .$$

With this inner product, $\mathcal{N}_\Phi(\Omega)$ becomes a reproducing-kernel Hilbert space with the reproducing kernel

$$\begin{aligned} K(\mathbf{x}, \mathbf{y}) \ &= \Phi(\mathbf{x}, \mathbf{y}) - \sum_{k=1}^{q} l_k(\mathbf{x}) \Phi(\xi_k, \mathbf{y}) - \sum_{r=1}^{q} l_r(\mathbf{y}) \Phi(\mathbf{x}, \xi_r) \\ &\quad + \sum_{k=1}^{q} \sum_{r=1}^{q} l_k(\mathbf{x}) l_r(\mathbf{y}) \Phi(\xi_k, \xi_r) + \sum_{k=1}^{q} l_k(\mathbf{x}) l_k(\mathbf{y}). \end{aligned} \tag{2.11}$$

We underline that the kernel K used here is a PD kernel (since it is a reproducing kernel) with

built-in polynomial precision. Thus we have

$$\mathcal{N}_\Phi(X) = \operatorname{span}\{K(\cdot, \mathbf{x}_j) : \ \mathbf{x}_j \in X\} \subset \mathcal{N}_\Phi(\Omega).$$

It means that starting from a CPD radial kernel, we find another PD kernel such that the native space of Φ can be spanned by its translation over the given data sites X.

2.3.1 Properties of the Native Spaces

Here, we briefly review some properties of the native spaces. To keep our notation consistent, we let Φ be a PD radial function that is a reproducing kernel for its native space according to the above discussion; however, notice that, in general, the reproducing kernel is not necessarily a radial function.

We begin by noticing that point evaluation functionals in a reproducing kernel Hilbert space, including the native space, is bounded, i.e., for every $f \in \mathcal{N}_\Phi(\Omega)$ and $\mathbf{x} \in \Omega$ we have

$$|f(\mathbf{x})| \leq \sqrt{\Phi(\mathbf{x}, \mathbf{x})} \|f\|_{\mathcal{N}_\Phi(\Omega)}.$$

Moreover, using the previous notation for X and $\mathbf{f}_X$, we can characterize the interpolant by variational principles (recall the property in equation (1.31)).

Theorem 26. *[Wen04, Lemma 10.24] With the notation from above, we get*

$$(f - s_{f,X}, s)_{\mathcal{N}_\Phi(\Omega)} = 0 \quad \textit{for all } s \in \mathcal{N}_\Phi(\Omega). \tag{2.12}$$

This yields an important lemma in the native space.

Lemma 3. *[Wen04, Corollary 10.25] Under the assumption in Theorem* (26)*, we get*

$$\|f\|^2_{\mathcal{N}_\Phi(\Omega)} = \|f - s_{f,X}\|^2_{\mathcal{N}_\Phi(\Omega)} + \|s_{f,X}\|^2_{\mathcal{N}_\Phi(\Omega)} \tag{2.13}$$

for all $f \in \mathcal{N}_\Phi(\Omega)$, This yields immediately

$$\|f - s_{f,X}\|_{\mathcal{N}_\Phi(\Omega)} \leq \|f\|_{\mathcal{N}_\Phi(\Omega)}, \ \textit{and} \ \|s_{f,X}\|_{\mathcal{N}_\Phi(\Omega)} \leq \|f\|_{\mathcal{N}_\Phi(\Omega)}. \tag{2.14}$$

The following theorems present the optimality property.

Theorem 27. *[Fas07, Theorem 18.1] Suppose Φ is a CPD kernel of an arbitrary order m and X is $\mathbb{P}^d_{m-1}-$unisolvent. Then the interpolant $s_{f,X}$ to function values $\mathbf{f}_X$ is the minimum-(semi)norm interpolant i.e., we have:*

$$|s_{f,X}|_{\mathcal{N}_\Phi(\Omega)} = \min_{s \in \mathcal{N}_\Phi(\Omega)} \|s\|_{\mathcal{N}_\Phi(\Omega)}, \tag{2.15}$$

for all s such that $s(X) = \mathbf{f}_X$ i.e., the interpolation condition is satisfied.

Notice that, in case $m = 1$ that is Φ is a PD kernel, the semi-norms in (2.15) become the norm.

Theorem 28. *[Fas07, Theorem 18.2] Let*

$$H_\Phi(X) = \{h = \sum_{j=1}^{n} c_j \Phi(\cdot, \mathbf{x}_j) + p \ : \ p \in \mathbb{P}^d_{m-1} \\ \text{and } \sum_{j=1}^{n} c_j q(\mathbf{x}_j) = 0 \ \ \forall q \in \mathbb{P}^d_{m-1}\}$$

where Φ is a CPD kernel of an arbitrary order m with X being $\mathbb{P}^d_{m-1}-$unisolvent set. Then the interpolant $s_{f,X}$ to $\mathbf{f}_X$ is the best approximation to f from $H_\Phi(X)$ in $\mathcal{N}_\Phi(\Omega)$, i.e.,

$$|f - s_{f,X}|_{\mathcal{N}_\Phi(\Omega)} \leq |f - h|_{\mathcal{N}_\Phi(\Omega)}$$

In what follows, we briefly investigate the error bounds of RBF interpolation.

2.4 Error Bounds in Terms of Power Function

Various error estimate types are suggested for RBF interpolation, each possessing different proprieties (see, e.g., [S+95, WS93, NWW05, NWW06]). Here, we mainly focus on error estimates given in the native space corresponding to the RBF employed for the interpolation process.

Let Φ be an arbitrary PD radial kernel. The matrix form of the interpolant can be represented as

$$s_{f,X}(\mathbf{x}) = \Phi_X(\mathbf{x})^T \mathbf{c}, \quad \mathbf{x} \in \Omega \subset \mathbb{R}^d, \tag{2.16}$$

with $\Phi_X(\mathbf{x})^T = [\Phi(\mathbf{x}, \mathbf{x}_1), \ldots, \Phi(\mathbf{x}, \mathbf{x}_n)]$, $\mathbf{c} = [c_1, \ldots, c_n]^T$ be the vector of coefficients as before, and $\mathbf{x}$ be an arbitrary point in the domain referred as the *evaluation point*. Recalling (2.1), we know that the $\mathbf{c} = V^{-1}_{\Phi,X}\mathbf{f}_X$ with $V_{\Phi,X}$ being the kernel matrix. Substituting this in (2.16), we have

$$\begin{aligned} s_{f,X}(\mathbf{x}) &= \Phi_X(\mathbf{x})^T (V^{-1}_{\Phi,X}\mathbf{f}_X) \\ &= \mathbf{f}_X^T (V^{-1}_{\Phi,X}\Phi_X(\mathbf{x})) \end{aligned} \tag{2.17}$$

We underline that (2.17) can be viewed as the *cardinal* form of the interpolant with the term $V^{-1}_{\Phi,X}\Phi_X(\mathbf{x})$ being the *cardinal basis function*. Moreover, notice that

$$\begin{aligned} \mathbf{f}_X^T = f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)) &= (\langle f, \Phi(\cdot, \mathbf{x}_1)\rangle_{\mathcal{N}_\Phi(\Omega)}, \ldots, \langle f, \Phi(\cdot, \mathbf{x}_n\rangle_{\mathcal{N}_\Phi(\Omega)}) \\ &= \langle f, (\Phi(\cdot, \mathbf{x}_1), \ldots, \Phi(\cdot, \mathbf{x}_n))\rangle_{\mathcal{N}_\Phi(\Omega)} \\ &= \langle f, \Phi_X^T(\cdot)\rangle_{\mathcal{N}_\Phi(\Omega)}. \end{aligned} \tag{2.18}$$

Substituting (2.18) in (2.17) and using the reproduction property of the reproducing kernel Hilbert space, we have

$$\begin{aligned}|f(\mathbf{x}) - s(\mathbf{x})| &= |\langle f, \Phi(\cdot,\mathbf{x})\rangle_{\mathcal{N}_\Phi(\Omega)} - \langle f, \Phi_X(\cdot)^T V_{\Phi,X}^{-1}\Phi_X(\mathbf{x})\rangle_{\mathcal{N}_\Phi(\Omega)}| \\ &= \left|\langle f, \Phi(\cdot,\mathbf{x}) - \Phi_X(\cdot)^T V_{\Phi,X}^{-1}\Phi_X(\mathbf{x})\rangle_{\mathcal{N}_\Phi(\Omega)}\right| \\ &\leq \|f\|_{\mathcal{N}_\Phi(\Omega)}\|\Phi(\cdot,\mathbf{x}) - \Phi_X(\cdot)^T V_{\Phi,X}^{-1}\Phi_X(\mathbf{x})\|_{\mathcal{N}_\Phi(\Omega)} \\ &= \|f\|_{\mathcal{N}_\Phi(\Omega)} P_{\Phi,X}(\mathbf{x}), \end{aligned} \tag{2.19}$$

with the *power function* $P_{\Phi,X}(\mathbf{x}) = \sqrt{\Phi(\mathbf{x},\mathbf{x}) - \Phi_X(\mathbf{x})^T V_{\Phi,X}^{-1}\Phi_X(\mathbf{x})}$.
Due to the error bound provided in (2.19), the impact of data smoothness, measured by the native space norm, is evaluated independently of the chosen basis function and the data site distribution. However, notice that the shape parameter ε can still affect both terms since the native space norm is affected by the shape parameter.

The error bound in (2.19) can be improved to a more straightforward formula. To express this, we begin with the following definition.

Definition 26.

1. *The* ***fill distance*** *is defined as*

$$h_{X,\Omega} = \sup_{\mathbf{x}\in\Omega} \min_{1\leq j\leq n} \|\mathbf{x} - \mathbf{x}_j\|_2.$$

2. *The* ***separation distance***

$$q_X = \frac{1}{2}\min_{i\neq j}\|\mathbf{x}_i - \mathbf{x}_j\|.$$

3. *The set of data sites X is said to be* ***quasi-uniform*** *with respect to a constant $c_{qu} > 0$ if*

$$q_X \leq h_{X,\Omega} \leq c_{qu} q_X. \tag{2.20}$$

According to [Weno4], the power function can be bounded in terms of the fill distance, so we have

$$|f(\mathbf{x}) - s(\mathbf{x})| \leq C h_{X,\Omega}^{\beta} \|f\|_{\mathcal{N}_\Phi(\Omega)}, \tag{2.21}$$

for some appropriate exponent β which depends on the smoothness of the kernel Φ; see [Weno4, Chap. 11].

One drawback of this power function approach is that error bounds apply only for the functions f that belong to the native space of Φ but not to the functions that belong to larger classes of functions. Besides, the error bounds suggested in either (2.19) or (2.21) can serve mostly for theoretical use. Note that we do not know the function f as it is supposed to be approximated,

and so, right-hand side norm $\|f\|_{\mathcal{N}_\Phi(\Omega)}$ can not be computed directly. Therefore, it might be more appropriate to address the error analysis of the RBF interpolation via the *sampling inequalities* in the *Sobolev spaces* since these estimates use the whole range of L_p norms and do not need the power function approach. Here, we avoid further discussion regarding the sampling inequalities as they are out of the context of this thesis. However, we shall return to the Sobolev spaces in chapter (4), where we discuss the error bounds of our suggested *approximation* scheme. We end this section by referring the interested readers to [NWW06] or [Wen04, chap. 11] for a detailed discussion regarding Sobolev bounds for functions with scattered zeros.

3

Full-rank orthonormal bases for conditionally positive definite kernel-based spaces

> There is no sense in being precise when you do not even know what you are talking about
>
> John von Neumann

3.1 Introduction

According to our discussion in the previous chapter, we saw that the interpolation linear systems either in (2.1) or (2.3) are built to be well-posed for every data distribution, but yet we did not discuss the condition number of this system. It is well-known (see e.g [DMS10]) that the interpolation problem in the subspace $\mathcal{N}_\Phi(X) = \text{span}\{\Phi(\cdot, \mathbf{x}_j) : \mathbf{x}_j \in X\} \subset \mathcal{N}_\Phi(\Omega)$ spanned by the basis of *translates* $\Phi(\cdot, \mathbf{x}_j)$, $1 \leq j \leq n$, is numerically unstable due to the ill-conditioning of the *kernel matrix* $V_{\Phi,X}$. Therefore, it is natural to devise strategies to prevent such instabilities by either preconditioning the system (see, e.g., [BLM11]) or by finding a better basis for the approximation space we are using. The latter case gives rise to *stable algorithms* introduced in [FW04] for the particular case of multiquadric kernels, the RBF-QR algorithm of [FP08], and its extension in [FLF11], and RBF-GA algorithm introduced in [FLP13]. Different approach was taken in [FM12] to to find $L_2(\mathbb{R}^d, \rho)$ following *Mercer's* theorem. However, such an idea could be restricted only to Gaussian RBFs or any other PD radial kernel such that their eigenfunction expansion is known. Another approach for the construction of better alternate bases for PD kernels has been introduced in [MS09] and was extended later in [PS11].

Their main idea is to produce $\mathcal{N}_\Phi$-orthonormal and discretely orthonormal data-dependent bases in the subspace $\mathcal{N}_\Phi(X)$ based on numerical using different matrix factorizations of the kernel matrix, such as SVD or Cholesky factorization. This has led to different bases with different properties, but all of them are data-dependent since one needs to form the kernel matrix, which is dependent on the given data site X. Stability issues, recursive compatibility, duality, and orthogonality properties were investigated for these new bases. Following such an idea, in [DMS13], a particular orthonormal basis built on a *weighted singular value decomposition* of the kernel matrix has been introduced. These bases are also related to a discretization of the compact integral operator T_Φ given by Mercer's theorem and provide a connection with the continuous bases that arise from an eigendecomposition of T_Φ. Although effective, this basis is computationally expensive to compute, so in [DMS15], the authors discussed methods related to *Krylov* subspaces to compute this basis in a fast way. Coming back to the CPD kernels, the linear system (2) may also suffer from ill-conditioning for some constellations of the interpolation points (see [11]). However, in contrast with the PD case, the literature contains very few contributions that address finding more stable bases for CPD kernels. An exception is [PS13] in which the authors tried to extend the previous work in [PS11] to the CPD case. But in their idea, having a full orthonormal basis of n functions is impossible. Explicitly, it is shown that one cannot simply use factorization techniques due to the augmented polynomial space, and therefore, some care needs to be taken. Another approach was taken in [BCM99] to find bases that are, in a certain sense, homogeneous, meaning that they are not sensitive to poorly scaled problems. Some numerical results regarding these homogeneous bases have also been reported in [Fas07, Chap. 34].

Contribution : In this chapter, we work with the reproducing kernel K given in equation (2.11) for the associated native Hilbert space $\mathcal{N}_\Phi(\Omega)$ with Φ being a CPD radial kernel, and Ω being non-empty subset of $\mathbb{R}^d$. We give a well-organized matrix formulation of the kernel matrix denoted by $\mathbf{K}$ by constructing the matrices corresponding to cardinal basis from monomials. Then, we present two possible ways to find full-rank data-dependent orthonormal bases that are discrete ℓ_2 and $\mathcal{N}_\Phi$-orthonormal. The kernel matrix $\mathbf{K}$ factorization gives the first approach, and the next one is based on the eigenpairs approximation of the linear operator associated with the reproducing kernel K given by Mercer's theorem. In the sequel, we employ the truncated singular value decomposition technique to find an optimal low-rank basis with the coefficient matrix whose rank is less than that of the original matrix. Special attention is also given to error analysis, duality, and stability. Some numerical experiments are also provided.

3.2 Matrix formulation

We start by providing an explicit representation of the matrix $\mathbf{K}$ by considering a cardinal basis of the polynomial-based space. Let X be the set of data sites as before. Any set of bases

$u_0, \ldots, u_n$ of $d-$variate polynomials up to degree $m-1$ can be arranged into a row vector

$$U(\mathbf{x}) = (u_0(\mathbf{x}), \ldots, u_q(\mathbf{x}))$$

for any evaluation point $\mathbf{x} \in \Omega$. Recall that q was the dimension of $\mathbb{P}^d_{m-1}$. In this case, it makes sense to expect that new set of bases U can be expressed by the monomials $\widetilde{m}(\mathbf{x}) = \left[\widetilde{m}_1(\mathbf{x}), \ldots, \widetilde{m}_q(\mathbf{x})\right]$; which are the common bases for $\mathbb{P}^d_{m-1}$. Therefore,

$$u_k = \sum_{k=0}^{q} \widetilde{m}_j(\mathbf{x})\mathbf{c}_{jk}, \quad 0 \leq k \leq q.$$

One can let U to be the Lagrange basis for $\mathbb{P}^d_{m-1}$ denoted by $l(\cdot) = \left[l_1(\cdot), \ldots, l_q(\cdot)\right]$, constructed via (cf. e.g. [MB19])

$$l(\mathbf{x}) = \widetilde{m}(\mathbf{x}) \cdot C_l, \quad \forall \mathbf{x} \in \Omega, \tag{3.1}$$

where $C_l \in \mathbb{R}^{q\times q}$ is known as the *construction* matrix. In this setting, let $[l_j(\xi_i)]_{1\leq i,j\leq q} = I$, and $V = [\widetilde{m}_j(\xi_i)]_{1\leq i,j\leq q}$ be the Vandermonde matrix. Recalling (3.1), the construction matrix is obtained by $C_l = V^{-1}$. Besides, if the Lagrange basis is needed at another set of evaluation points, say $Y = \{\mathbf{y}_1, \ldots, \mathbf{y}_s\}$, by equation (3.1) we get

$$V^T L_Y^T = V_Y^T,$$

where $L_Y = [l_j(\mathbf{y}_i)]$ and $V_Y = \left[\widetilde{m}_j(\mathbf{y}_i)\right]$ with $1 \leq i \leq s$ and $1 \leq j \leq q$. Hence, for the kernel matrix $\mathbf{K} = K(\mathbf{x}_i, \mathbf{x}_j)$ with

$$\begin{aligned} K(\mathbf{x}_i, \mathbf{x}_j) &= \Phi(\mathbf{x}_i, \mathbf{x}_j) - \sum_{k=1}^{q} l_k(\mathbf{x}_i)\Phi(\xi_k, \mathbf{x}_j) - \sum_{r=1}^{q} l_r(\mathbf{x}_j)\Phi(\mathbf{x}_i, \xi_r) \\ &\quad + \sum_{k=1}^{q}\sum_{r=1}^{q} l_k(\mathbf{x}_i)l_r(\mathbf{x}_j)\Phi(\xi_k, \xi_r) + \sum_{k=1}^{q} l_k(\mathbf{x}_i)l_k(\mathbf{x}_j), \quad ij = 1, \ldots, \end{aligned} \tag{3.2}$$

we get

$$\mathbf{K} = V_{\Phi,X} - L_1 \cdot A_1 - A_2 \cdot L_1^T + L_1 \cdot A_3 \cdot L_1^T + L_1 \cdot L_1^T,$$

with $V_{\Phi,X}$ as in equation (2.3) and

$$\begin{aligned} L_1 &= [l_k(\mathbf{x}_i)]_{\substack{1 \leq i \leq n \\ 1 \leq k \leq q}}, \quad A_1 = \left[\Phi(\xi_k, \mathbf{x}_j)\right]_{\substack{1 \leq k \leq q \\ 1 \leq j \leq n}}, \\ A_2 &= [\Phi(\mathbf{x}_i, \xi_r)]_{\substack{1 \leq i \leq n \\ 1 \leq r \leq q}}, A_3 = [\Phi(\xi_k, \xi_r)]_{\substack{1 \leq k \leq q \\ 1 \leq r \leq q}}. \end{aligned}$$

We recall that the points $\Xi = \{\xi_1, \ldots, \xi_q\}$ must be $\mathbb{P}^d_{m-1}$-unisolvent subset of X.

3.3 Full-rank orthonormal bases

In what follows, we investigate suitable bases for subspaces of $\mathcal{N}_\Phi(\Omega)$ when Φ is CPD. The bases will be given in different forms, and we will explore their variety and prove the results of their connection. Hence, let Φ be a fixed CPD kernel with corresponding reproducing kernel K in (3.2), $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$ a fixed set of centers, and $U = [u_1, \ldots, u_n]$ a general data-dependent basis such that

$$\mathcal{N}_\Phi(X) = \text{span}\{K(\cdot, \mathbf{x}_j) : \mathbf{x}_j \in X\} = \text{span}\{u_1, ..., u_n\} \subset \mathcal{N}_\Phi(\Omega).$$

Following [PS11], any element of the basis U can be written as a linear combination of the translates $K(\cdot, \mathbf{x}_j),\ j = 1, ..., n$ via the construction matrix C

$$u_i = \sum_{j=1}^{n} K(\cdot, \mathbf{x}_j) c_{ji}, \quad 1 \leq i \leq n, \tag{3.3}$$

or in matrix form

$$E = \mathbf{K}C, \tag{3.4}$$

where $E = [u_j(\mathbf{x}_i))]_{1 \leq i,j \leq n}$ and $C = [c_{ji}]_{1 \leq i,j \leq n}$. Therefore, the construction matrix enables us to shift from one set of bases to another while each possesses various properties. For the sake of simplicity, instead of $\mathcal{N}_\Phi(\Omega)$ and discrete $\ell_2(X)$, we use the shorter notation $\mathcal{N}_\Phi$ and ℓ_2, respectively.

Theorem 29. *The $\mathcal{N}_\Phi$ and ℓ_2 Gramian matrices associated with the general basis U are symmetric and positive definite with full-rank n.*

Proof. The Gramian matrices associated with the basis U corresponding to $\mathcal{N}_\Phi$ and ℓ_2 are

$$\begin{aligned} G_{\mathcal{N}_\Phi} &= [\langle u_i, u_j \rangle_{\mathcal{N}_\Phi}]_{1 \leq i,j \leq n} = C^T \mathbf{K} C, \\ G_{\ell_2} &= [\langle u_i, u_j \rangle_{\ell_2}]_{1 \leq i,j \leq n} = \left(\sum_{k=1}^{n} u_i(\mathbf{x}_k) u_j(\mathbf{x}_k) \right)_{1 \leq i,j \leq n} = E^T E = C^T \mathbf{K}^2 C. \end{aligned}$$

The evaluation matrix E is necessarily full-rank because the basis must allow unique interpolation on X. Since $C = \mathbf{K}^{-1}E$ the construction matrix C is also full-rank, resulting the same for $G_{\mathcal{N}_\Phi}$ and G_{ℓ_2}.The matrices $G_{\mathcal{N}_\Phi}$ and G_{ℓ_2} are clearly symmetric. Now, since $\mathbf{K}$ is a positive definite matrix then, for all nonzero vectors $\mathbf{z} \in \mathbb{R}^n$, we have

$$\mathbf{z}^T \cdot G_{\mathcal{N}_\Phi} \cdot \mathbf{z} = (C\mathbf{z})^T \mathbf{K} (C\mathbf{z}) > 0,$$

and since E is a full-rank matrix similarly

$$\mathbf{z}^T \cdot G_{\ell_2} \cdot \mathbf{z} = \langle E\mathbf{z}, E\mathbf{z} \rangle = \|E\mathbf{z}\|^2{}_2 > 0.$$

We have then show that $G_{\mathcal{N}_\Phi}$ and G_{ℓ_2} are positive definite. □

Remark 3. *Suppose that we construct the evaluation matrix E through the augmented system*

$$\begin{bmatrix} E_{n\times n} \\ 0_{q\times n} \end{bmatrix} = \begin{bmatrix} A_{n\times n} & P_{n\times q} \\ P^T_{q\times n} & 0_{q\times q} \end{bmatrix} \begin{bmatrix} \widetilde{C}_{n\times n} \\ \widetilde{D}_{q\times n} \end{bmatrix}.$$

The moment conditions $P^T\widetilde{C} = 0$, reveals that the $n \times n$ matrix $\widetilde{C}$ has rank $n - q$ and the evaluation matrix E is necessarily rank n, since the basis must allow unique interpolation on X. Then the Gramian matrix $G_{\mathcal{N}_\Phi} = \widetilde{C}^T E$ is symmetric and positive semi-definite with rank $n - q$. So, it is impossible to have a full orthonormal basis if $q > 0$.

The above remark highlights why we prefer reproducing kernels over the standard bases with augmented polynomials. Besides, one should note that the space $\mathcal{N}_\Phi(X)$ is fixed, meaning that the new bases are given with respect to the data sites X. This clarifies why the new bases U are data-dependent.

In the following, we address two possible ways to find data-dependent orthonormal bases corresponding to the CPD kernel Φ on a domain $\Omega \subset \mathbb{R}^d$.

3.3.1 Matrix decomposition approach

According to [PS11], equation (3.4) reveals that one can find data-dependent basis U from the decomposition of the symmetric and positive definite matrix $\mathbf{K}$ corresponding to the CPD kernel Φ as

$$\mathbf{K} = EC^{-1}.$$

The kind of matrix decomposition depends on the Hilbert space in which we are looking for orthonormal bases. We can characterize $\mathcal{N}_\Phi$ and discretely ℓ_2 orthonormal bases based on the Gramian matrices as follows:

1) For $\mathcal{N}_\Phi-$orthonormal bases, we have

$$G_{\mathcal{N}_\Phi} = I \leftrightarrow C^T\mathbf{K}C = I \leftrightarrow \mathbf{K} = (C^{-1})^T C^{-1} \leftrightarrow E = (C^{-1})^T.$$

Then, there are two important cases.

i) The Choleskey decomposition $\mathbf{K} = LL^T$ with a nonsingular lower triangular matrix L which leads to the Newton basis with a different normalization [MS09]. In this case $E = L$ and $C = (L^T)^{-1}$.

ii) The singular value decomposition (SVD) decomposition of the form $\mathbf{K} = QDQ^T$ with an orthogonal matrix Q and a diagonal matrix D having the eigenvalues of $\mathbf{K}$ on its diagonal. In this case $E = Q\sqrt{D}$ and $C = Q\left(\sqrt{D}\right)^{-1}$.

2) For ℓ_2-orthonormal bases, we have

$$G_{\ell_2} = I \leftrightarrow C^T\mathbf{K}^T\mathbf{K}C = I \leftrightarrow \mathbf{K}C = Q \leftrightarrow \mathbf{K} = QC^{-1} \leftrightarrow E = Q.$$

Also, here, there are two important special cases.

i) The standard QR decomposition $\mathbf{K} = QR$ into an orthogonal matrix Q and an upper triangular matrix R will lead to a basis with $E = Q$ and $C = R^{-1}$.
ii) The SVD of $\mathbf{K} = QDQ^T$ which leads to $E = Q$ and $C = QD^{-1}$.

3.3.2 Eigenpairs approximation approach

We discuss another family of orthonormal bases based on the eigenvalues and eigenfunctions of *Hilbert-Schmidt* operator [FM15, Chap 2] associated with the reproducing kernel K given in (2.11). Mercer's theorem expresses the connection of such a linear operator with the infinite series representation of a positive definite kernel.

Theorem 30. ***(Mercer's theorem)*** *Let K be a continuous positive definite kernel that satisfies*

$$\int_\Omega K(\mathbf{x},\mathbf{y})v(\mathbf{x})v(\mathbf{y})d\mathbf{x}d\mathbf{y} \geq 0, \quad \forall v \in L_2(\Omega),\ \mathbf{x},\mathbf{y} \in \Omega.$$

Then K can be represented by

$$K(\mathbf{x},\mathbf{y}) = \sum_{j=1}^{\infty} \lambda_j \widetilde{u}_j(\mathbf{x})\widetilde{u}_j(\mathbf{y}), \tag{3.5}$$

where λ_j are the eigenvalues such that $\lambda_j \to 0$ as $j \to \infty$, and $\widetilde{u}_j$ are the L_2-orthonormal eigenfunctions of the operator $T_K : L_2(\Omega) \to L_2(\Omega)$ given by

$$T_K(v)(\mathbf{x}) = \int_\Omega K(\mathbf{x},\mathbf{y})v(\mathbf{y})d\mathbf{y}, \quad v \in L_2(\Omega),\ \mathbf{x} \in \Omega.$$

Moreover, this representation is absolutely and uniformly convergent.

Theorem (30) can lead to another characterization of the Native space $\mathcal{N}_\Phi(\Omega)$ as

$$\mathcal{N}_\Phi(\Omega) = \left\{ f : f = \sum_{j=1}^{\infty} c_j \widetilde{u}_j \right\},$$

where the kernel K itself is in $\mathcal{N}_\Phi(\Omega)$ because of the eigenfunction expansion (3.5). The reproducing property of the kernel K should be checked by the following equation

$$\langle f(\cdot), K(\cdot,\mathbf{x})\rangle_{\mathcal{N}_\Phi} = \left\langle \sum_{j=1}^{\infty} c_j \widetilde{u}_j(\cdot), \sum_{i=1}^{\infty} \lambda_i \widetilde{u}_i(\cdot)\widetilde{u}_i(\mathbf{x}) \right\rangle_{\mathcal{N}_\Phi} = \sum_{j=1}^{\infty} c_i \widetilde{u}_i(\mathbf{x}) = f(\mathbf{x}),$$

which leads to the $\mathcal{N}_\Phi$-orthogonality of the eigenfunctions

$$\langle \widetilde{u}_i, \widetilde{u}_j\rangle_{\mathcal{N}_\Phi} = \frac{\delta_{ij}}{\sqrt{\lambda_i}\sqrt{\lambda_j}}. \tag{3.6}$$

The inner product for $\mathcal{N}_\Phi(\Omega)$ is then given by

$$\langle f, g\rangle_{\mathcal{N}_\Phi} = \left\langle \sum_{j=1}^{\infty} c_j \widetilde{u}_j, \sum_{i=1}^{\infty} d_i \widetilde{u}_i \right\rangle_{\mathcal{N}_\Phi} = \sum_{j=1}^{\infty} \frac{c_j d_j}{\lambda_j}.$$

Equation (3.6) reveals two important cases for the basis functions.

i) Basis functions

$$\{u_j\}_{j=1}^{\infty} = \left\{\sqrt{\lambda_j}\widetilde{u}_j\right\}_{j=1}^{\infty}, \quad \|u_j\|^2_{\mathcal{N}_\Phi} = 1, \quad \|u_j\|^2_{L_2} = \lambda_j, \tag{3.7}$$

which is orthonormal in $\mathcal{N}_\Phi(\Omega)$ and orthogonal in $L_2(\Omega)$.

ii) Basis functions

$$\{v_j\}_{j=1}^{\infty} = \{\widetilde{u}_j\}_{j=1}^{\infty}, \quad \|v_j\|^2_{\mathcal{N}_\Phi} = \frac{1}{\lambda_j}, \quad \|v_j\|^2_{L_2} = 1, \tag{3.8}$$

which is orthogonal in $\mathcal{N}_\Phi(\Omega)$ and orthonormal in $L_2(\Omega)$.

Unfortunately, in most cases, eigenpairs of the operator T_K are not known analytically. The exception is the Gaussian kernel, which is a PD kernel by definition. On the other hand, to our knowledge, no research has been conducted on investigating the analytical form of the eigenpairs related to CPD kernels in (2.11). Thus it will be required to approximate them using numerical schemes. This leads to the following eigenvalue problem on X

$$\int_\Omega K(\mathbf{x}_i, \mathbf{y})\widetilde{u}_j(\mathbf{y})d\mathbf{y} = \lambda_j \widetilde{u}_j(\mathbf{x}_i), \quad i = 1, \ldots, n, \forall j > 0,$$

which can be discretized by using the *symmetric Nyström method* [AH05] which is a cubature rule $(X, W)_n$, $n \in \mathbb{N}$, for the set of distinct points X and a set of positive weights $W = \{w_r\}_{r=1}^n$ such that

$$\int_\Omega f(\mathbf{y})d\mathbf{y} \approx \sum_{r=1}^{n} f(\mathbf{x}_r)w_r, \quad \forall f \in \mathcal{N}_\Phi(\Omega).$$

This leads to

$$\sum_{r=1}^{n} K(\mathbf{x}_i, \mathbf{x}_r)\widetilde{u}_j(\mathbf{x}_r)w_r \approx \lambda_j \widetilde{u}_j(\mathbf{x}_i), \quad i,j = 1,\ldots,n, \tag{3.9}$$

with a set of positive weights $\{w_r\}_{r=1}^{n}$. Now, it suffices to solve the following discrete eigenvalue problem in order to find the approximation of the eigenvalues and eigenfunctions (evaluated on X) of T_K. In other terms, we have

$$(\mathbf{K}W)\widetilde{\mathbf{e}}^{(j)} = \lambda_j \widetilde{\mathbf{e}}^{(j)}, \quad j = 1,\ldots,n,$$

with

$$\begin{aligned} W &= \operatorname{diag}(w_r), \\ \widetilde{\mathbf{e}}^{(j)} &= \left[\widetilde{u}_j(\mathbf{x}_i))\right]_{1 \le i \le n} . \end{aligned} \tag{3.10}$$

Then, the continuum eigenvalue problem reduces to the solution of an unsymmetric eigenvalue problem

$$(\mathbf{K}W)\widetilde{\mathbf{e}} = \lambda\widetilde{\mathbf{e}}, \tag{3.11}$$

for the unsymmetric matrix $\mathbf{K}W$. One possible way to deal with (3.11) would be to make some manipulation to convert the unsymmetric problem of (3.11) to the following symmetric one

$$(\sqrt{W}\mathbf{K}\sqrt{W})(\sqrt{W}\cdot\widetilde{\mathbf{e}}) = \lambda(\sqrt{W}\cdot\widetilde{\mathbf{e}}).$$

Now, the SVD decomposition for the symmetric matrices, which is nothing but a unitary diagonalization, leads to

$$\sqrt{W}\mathbf{K}\sqrt{W} = QDQ^T, \tag{3.12}$$

where $D = \operatorname{diag}(\lambda_j)$,

$$Q = \left[\sqrt{W}\widetilde{\mathbf{e}}^{(1)}, \ldots, \sqrt{W}\widetilde{\mathbf{e}}^{(n)}\right],$$

is an orthogonal matrix w.r.t the Euclidean norm. Equations (3.7), (3.8), and (3.10) lead to the evaluation matrices

i) $E_1 = [u_j(\mathbf{x}_i))]_{1\le i,j\le n} = \left[\sqrt{\lambda_j}\widetilde{u}_j(\mathbf{x}_i))\right]_{1\le i,j\le n} = \left[\sqrt{\lambda_1}\widetilde{\mathbf{e}}^{(1)}, \ldots, \sqrt{\lambda_n}\widetilde{\mathbf{e}}^{(n)}\right] = \left(\sqrt{W}\right)^{-1} Q\sqrt{D},$

ii) $E_2 = [v_j(\mathbf{x}_i))]_{1\le i,j\le n} = [\widetilde{u}_j(\mathbf{x}_i))]_{1\le i,j\le n} = \left[\widetilde{\mathbf{e}}^{(1)}, \ldots, \widetilde{\mathbf{e}}^{(n)}\right] = \left(\sqrt{W}\right)^{-1} Q,$

According to (3.4) and (3.12), the corresponding construction matrices can be derived

i) $C_U = \mathbf{K}^{-1}E_1 = \sqrt{W}QD^{-1}Q^T\sqrt{W}\left(\sqrt{W}\right)^{-1}Q\sqrt{D} = \sqrt{W}Q\left(\sqrt{D}\right)^{-1}$.

ii) $C_V = \mathbf{K}^{-1}E_2 = \sqrt{W}QD^{-1}Q^T\sqrt{W}\left(\sqrt{W}\right)^{-1}Q = \sqrt{W}QD^{-1}$.

By considering the discretized scaled inner product

$$\langle f,g\rangle^2_{\ell_{2,w}} = \sum_{j=1}^{n} w_j f(\mathbf{x}_j)g(\mathbf{x}_j) \approx \langle f,g\rangle^2_{L_2(\Omega)} = \int_\Omega f(\mathbf{x})g(\mathbf{x})d\mathbf{x},$$

we have $\mathcal{N}_\Phi-$orthonormal and $\ell_{2,w}-$orthogonal basis functions

$$\begin{aligned} u_j(\mathbf{x}) &= \sum_{i=1}^{n} K(\mathbf{x},\mathbf{x}_i)c_{ij} = \sum_{i=1}^{n} K(\mathbf{x},\mathbf{x}_i)\frac{\sqrt{w_i}}{\sqrt{\lambda_j}}Q(i,j) \\ &= \sum_{i=1}^{n} K(\mathbf{x},\mathbf{x}_i)\frac{\sqrt{w_i}}{\sqrt{\lambda_j}}\frac{\sqrt{w_i}}{\sqrt{\lambda_j}}E_1(i,j) = \frac{1}{\lambda_j}\sum_{i=1}^{n} w_i K(\mathbf{x},\mathbf{x}_i)u_j(\mathbf{x}_i), \end{aligned}$$

with $\|u_j\|^2_{\ell_{2,w}} = \lambda_j$, and $\mathcal{N}_\Phi-$orthogonal and $\ell_{2,w}-$orthonormal basis functions

$$\begin{aligned} v_j(\mathbf{x}) &= \sum_{i=1}^{n} K(\mathbf{x},\mathbf{x}_i)c_{ij} = \sum_{i=1}^{n} K(\mathbf{x},\mathbf{x}_i)\frac{\sqrt{w_i}}{\lambda_j}Q(i,j) \\ &= \sum_{i=1}^{n} K(\mathbf{x},\mathbf{x}_i)\frac{\sqrt{w_i}}{\lambda_j}\sqrt{w_i}E_2(i,j) = \frac{1}{\lambda_j}\sum_{i=1}^{n} w_i K(\mathbf{x},\mathbf{x}_i)v_j(\mathbf{x}_i), \end{aligned}$$

with $\|v_j\|^2_{\mathcal{N}_\Phi} = \frac{1}{\lambda_j}$, for $1 \le j \le n$.

Remark 4. *According to this theory, we can deduce that all data-dependent bases, which are both discretely and $\mathcal{N}_\Phi$-orthogonal, are scaled SVD bases derived from the eigenpairs approximation approach. It is also clear that the SVD bases given in section 3.3.1 are special cases of the general bases given in this section.*

Remark 5. *The reader should note that the expansion series in* (3.5) *is valid only for the PD kernels. So, working with the associated reproducing kernel of a CPD kernel rather than the standard bases enables us to use Mercer's theorem and find two additional classes of bases.*

Remark 6. *We point out that to construct the new bases; we require that at least the weights $\{w_r\}_{r=1}^n$ are positive and able to reproduce constants that is $\sum_{j=1}^{n} w_j = |\Omega|$. Obviously, a higher order cubature formula will lead to a better approximation of the eigenbasis $\{\widetilde{u}_j\}_{j>0}$, so in the limit; we could expect that our basis will be able to reproduce $\{\widetilde{u}_j\}_{j>0}$ and each function in $\mathcal{N}_\Phi(\Omega)$. Nevertheless, at a finite stage with fixed n, also assuming that we know an almost exact cubature formula, we are still approximating $f \in \mathcal{N}_\Phi(\Omega)$ with a projection into $\mathcal{N}_\Phi(X)$. In principle, it is also possible to use weights not related to a cubature rule. Still, in this way, no connection can be expected between $\mathcal{N}_\Phi(\Omega)$ and the eigenbasis $\{\widetilde{u}_j\}_{j>0}$.*

3.3.3 Low-rank approximation

As previously mentioned, a vital characteristic of a reproducing kernel, which is the inevitability of the kernel matrix, has been compromised in the transition from theory to practical implementation. This is primarily because the *numerical rank* of K is frequently significantly lower than n, making the kernel matrix K ill-conditioned. In other words, for many kernels, the eigenvalues in (3.5) decrease rapidly toward zero, implying a very good low-rank approximation to the kernel. Notice that vectors involved in kernel representation in (3.5) are of infinite size and so need to be truncated at some finite length M, possibly much smaller than n. Thus, in many cases, linear systems arising from kernels are ill-conditioned, but they also have a low-rank subsystem that performs like the full system. Considering (3.5), we have the following theorem from [NW08].

Theorem 31. *Let $K : \Omega \times \Omega \to \mathbb{R}$ be a PD kernel with Mercer series (3.5). Then, M-term truncation*

$$K_M(\mathbf{x}, \mathbf{y}) = \sum_{n=1}^{M} \lambda_n \widetilde{u}_j(\mathbf{x}) \widetilde{u}_j(\mathbf{y}), \tag{3.13}$$

for a fixed $\mathbf{x}$ provides the best M-term least squares approximation of $K(\mathbf{x}, \mathbf{y})$ from $L_2(\Omega)$.

The summation (3.13) yields the best $M-$term approximation of each kernel matrix in $L_2(\Omega)$ norm, but this is not necessarily the best low-rank approximation in the 2-norm sense. Therefore, we consider SVD low-rank representation (truncated SVD) of the kernel matrix $\mathbf{K}$, which is obtained by discarding all but the k largest eigenvalues and the corresponding eigenvectors and is represented as

$$\mathbf{K}_k = Q_k D_k Q_k^T, \tag{3.14}$$

where $Q_k \in \mathbb{R}^{n \times k}$ and $D_k \in \mathbb{R}^{k \times k}$. It means that $\mathbf{K}_k$ is the projection of $\mathbf{K}$ onto the space spanned by the top k eigenvectors of $\mathbf{K}$. The following states that the above approximation is the best rank-k approximation in both *Frobenius* and *spectral* norm.

Theorem 32. ***(Eckart-Young** [EY36]) Let A_k be the rank-k approximation of $A \in \mathbb{R}^{n \times n}$ achieved by the truncated SVD. Then A_k is the closest rank-k matrix to A, i.e.*

$$\min_{\text{rank}(G)=k} \|A - G\|_F = \|A - A_k\|_F = \sqrt{\sigma_{k+1}^2 + \cdots + \sigma_n^2},$$

where σ_i's denote singular values of A and G is an arbitrary rank-k matrix.

Remark 7. *We remark that the SVD also gives the best low-rank approximation in the spectral norm, i.e.*

$$\min_{\text{rank}(G)=k} \|A - G\|_2 = \|A - A_k\|_2 = \sigma_{k+1}.$$

So, the rank-reduced system will be very close to the exact system but with a more well-behaved linear system with a better-conditioned value matrix [GVL13]. Accordingly, the evaluation matrix of the new bases can be represented as

$$\begin{aligned} E_{1k} &= Q_k\sqrt{D_k}, \\ E_{2k} &= Q_k, \end{aligned} \tag{3.15}$$

such that $E_{1k}, E_{2k} \in \mathbb{R}^{n\times k}$.

Remark 8. *One should notice that in order to consider the truncated SVD, it requires that there exists a well-defined gap in the singular values, i.e., $\frac{\sigma_{k+1}}{\sigma_k}$ must be large enough. Otherwise, the determination of the optimal rank k would be complicated, and the low-rank approximation of matrix* $\mathbf{K}$ *is meaningless. A detailed discussion is available in [GVL13, Chap 12.2].*

Although we know from Theorem 30 that the continuous eigenvalues accumulate to zero, the speed with which they decay is clearly not the same for the different kernels. Although depending on the choice of the shape parameter ε, they present a fast decay to zero (Gaussian), a medium decay (inverse multiquadric), and a slow decay (cubic Matérn kernel).

3.4 Application to interpolation

3.4.1 General interpolant

Having derived different types of data-dependent bases U, the interpolant $s_f \in \mathcal{N}_\Phi(X)$ to vector values $\mathbf{f}$ of some function f, can be represented as

$$s_f(\mathbf{x}) = \sum_{j=1}^{n} \alpha_j u_j(\mathbf{x}), \tag{3.16}$$

where the coefficients α_j are determined by solving the linear system

$$E\alpha = \mathbf{f}, \tag{3.17}$$

where $\alpha = [\alpha_j]_{1\le j\le n}$ and $E = \left[u_j(\mathbf{x}_i)\right]_{1\le,i,j\le n}$ can be one of the evaluation matrices obtained in Subsections 3.3.1 or 3.3.2. Once the coefficient vector α is calculated through (3.17), one can obtain the approximate function values $\mathcal{F}_Y \approx \mathbf{f}_Y = [f(\mathbf{y}_i)]_{1\le i\le s}$ at the set of test points $Y = \{\mathbf{y}_1, \ldots, \mathbf{y}_s\}$ by

$$\mathcal{F}_Y = E_Y \cdot \alpha, \tag{3.18}$$

where $E_Y = [u_j(\mathbf{y}_i)]$ with $1 \leq i \leq s$ and $1 \leq j \leq n$ is obtained by (3.4) as

$$E_Y = \mathbf{K}_Y C,$$

where C is the corresponding construction matrix and $\mathbf{K}_Y = [K(\mathbf{y}_i, \mathbf{x}_j)]$ can be computed via the same procedure explained at the end of Section (3.2).

Theorem 33. *The evaluation matrices of the $\mathcal{N}_\Phi$ and $\ell_{2,w}-$orthonormal basis functions are better conditioned than the kernel matrix* $\mathbf{K}$.

Proof. If E is the evaluation matrix corresponding to a $\ell_{2,w}-$orthonormal basis then it is an orthogonal matrix and so $\text{cond}_{2,w}(E) = 1$. Now let E be the evaluation matrix corresponding to a $\mathcal{N}_\Phi-$orthonormal basis derived from the general scaled SVD bases, then

$$E = \left(\sqrt{W}\right)^{-1} Q\sqrt{D}.$$

Moreover according to (3.12), we have

$$\mathbf{K} = (\sqrt{W})^{-1} Q\sqrt{D}\sqrt{D}Q^T(\sqrt{W})^{-1} = EE^T,$$

and

$$(\sqrt{D})^{-1} Q^T\sqrt{W}EE^T\sqrt{W}Q(\sqrt{D})^{-1} = I.$$

Therefore

$$\widehat{Q} = (\sqrt{D})^{-1} Q^T\sqrt{W}E,$$

is an orthogonal matrix w.r.t the norm $\|\cdot\|_{\ell_{2,w}}$, which in turn gives

$$E = \left(\sqrt{W}\right)^{-1} Q\sqrt{D}\widehat{Q},$$

that is nothing but the SVD of the matrix E. Therefore the spectral condition number of E is the square root of the spectral condition number of $\mathbf{K}$. The same theory can be used for the Newton basis functions given in 3.3.1. □

Remark 9. *If linear maps $\mathcal{L}$ like derivatives have to be evaluated, we use the system*

$$\mathcal{L}E_Y = \mathcal{L}\mathbf{K}_Y C,$$

where $\mathcal{L}E_Y = [\mathcal{L}u_j(\mathbf{y}_i)]$ and $\mathcal{L}\mathbf{K}_Y = [\mathcal{L}\mathbf{K}(\mathbf{y}_i, \mathbf{x}_j)]$ with $1 \leq i \leq s,\ 1 \leq j \leq n$ is given by applying the operator $\mathcal{L}$ to the reproducing kernel (3.2) and doing the the same procedure explained at the end of Section 3.3.

3.4.2 Error bound

In this section, we provide the error estimate for the approximation given in (3.18). First, the following stability issue is proved.

Theorem 34. *For a fixed CPD kernel Φ, fixed set of center points $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, general data-dependent basis U, and $f \in \mathcal{N}_\Phi$, the following stability estimate holds for the approximate function values $\mathcal{F}_Y$ at the set of test points $Y = \{\mathbf{y}_1, \ldots, \mathbf{y}_s\}$,*

$$\|\mathcal{F}_Y\|_2^2 \leq s \cdot \rho(\widetilde{\mathbf{K}}) \cdot \mathrm{cond}_2(\mathrm{G}_{\mathcal{N}_\Phi}) \cdot \|f\|_{\mathcal{N}_\Phi}^2,$$

where $\mathrm{cond}_2(\mathrm{G}_{\mathcal{N}_\Phi})$ *is the spectral condition number of the $\mathcal{N}_\Phi-$Gramian, ρ is the spectral radius, and $\widetilde{\mathbf{K}} = [K(\mathbf{y}_i, \mathbf{y}_j)]_{1\leq i,j\leq s}$ for the corresponding reproducing kernel K.*

Proof. Since Frobenius norm is compatible with the Euclidean norm, according to (3.18), we have

$$\|\mathcal{F}_Y\|_2^2 = \|E_Y \cdot \alpha\|_2^2 \leq \|E_Y\|_F^2 \|\alpha\|_2^2 . \tag{3.19}$$

Now according to (3.16), we get

$$\alpha^T G_{\mathcal{N}_\Phi} \alpha = \|s_f\|_{\mathcal{N}_\Phi}^2 \leq \|f\|_{\mathcal{N}_\Phi}^2.$$

Since

$$\alpha^T G_{\mathcal{N}_\Phi} \alpha = \langle \alpha, G_{\mathcal{N}_\Phi} \alpha \rangle \leq \|\alpha\|_2 \|G_{\mathcal{N}_\Phi} \alpha\|_2 \leq \|\alpha\|_2^2 \|G_{\mathcal{N}_\Phi}\|_2 = \rho(G_{\mathcal{N}_\Phi}) \|\alpha\|_2^2,$$

then

$$\rho(G_{\mathcal{N}_\Phi}) \|\alpha\|_2^2 \leq \|f\|_{\mathcal{N}_\Phi}^2.$$

Therefore

$$\|\alpha\|_2^2 \leq \|f\|_{\mathcal{N}_\Phi}^2 \rho(G_{\mathcal{N}_\Phi}^{-1}). \tag{3.20}$$

Moreover, we have

$$\mathbf{K}_Y^{(i)} \mathbf{K}^{-1} \left(\mathbf{K}_Y^{(i)}\right)^T = K(\mathbf{y}_i, \mathbf{y}_i) - P_{\Phi,X}^2(\mathbf{y}_i), \quad i = 1, \ldots, s,$$

where $\mathbf{K}_Y^{(i)}$ is the $i-$th row of the matrix $\mathbf{K}_Y$ and $P_{\Phi,X}$ is the so-called power function. Now according to (3.4), we get

$$E_Y^{(i)} C^{-1} \mathbf{K}^{-1} \left(C^{-1}\right)^T \left(E_Y^{(i)}\right)^T = K(\mathbf{y}_i, \mathbf{y}_i) - P_{\Phi,X}^2(\mathbf{y}_i), \quad i = 1, \ldots, s,$$

where $E_Y^{(i)}$ is the i-th row of the matrix E_Y. Therefore

$$E_Y^{(i)} \left(G_{\mathcal{N}_\Phi}\right)^{-1} \left(E_Y^{(i)}\right)^T = K(\mathbf{y}_i, \mathbf{y}_i) - P^2_{\Phi,X}\mathbf{y}_i) \leq K(\mathbf{y}_i, \mathbf{y}_i), \quad i = 1, \ldots, s,$$

which leads to

$$\|E_Y^{(i)}\|_2^2 \leq K(\mathbf{y}_i, \mathbf{y}_i)\rho(G_{\mathcal{N}_\Phi}), \quad i = 1, \ldots, s,$$

Now since

$$\|E_Y\|_F^2 = \sum_{i=1}^{s} \|E_Y^{(i)}\|_2^2,$$

we have

$$\|E_Y\|_F^2 \leq \mathrm{tr}(\widetilde{\mathbf{K}})\rho(G_{\mathcal{N}_\Phi}) \leq s \cdot \rho(\widetilde{\mathbf{K}}) \cdot \rho(G_{\mathcal{N}_\Phi}). \tag{3.21}$$

So by substituting (3.20) and (3.21) in (3.19), the proof is completed. □

Theorem 35. *For a fixed CPD kernel Φ, general data-dependent basis U, and $f \in \mathcal{N}_\Phi$, the following error bound holds*

$$\|\mathbf{f}_Y - \mathcal{F}_Y\|_2^2 \leq \left(s \cdot \rho(\widetilde{\mathbf{K}}) - \rho(G_{\mathcal{N}_\Phi}^{-1}) \|E_Y\|_F^2\right) \|f\|_{\mathcal{N}_\Phi}^2.$$

Proof.

$$\begin{aligned}
\|\mathbf{f}_Y - \mathcal{F}_Y\|_2^2 &= \sum_{i=1}^{s} |f(\mathbf{y}_i) - s_f(\mathbf{y}_i)|^2 \\
&\leq \sum_{i=1}^{s} P^2_{\Phi,X}(\mathbf{y}_i) \|f\|_{\mathcal{N}_\Phi}^2 \\
&= \sum_{i=1}^{s} \left(K(\mathbf{y}_i, \mathbf{y}_i) - E_Y^{(i)} \left(G_{\mathcal{N}_\Phi}\right)^{-1} \left(E_Y^{(i)}\right)^T\right) \|f\|_{\mathcal{N}_\Phi}^2 \\
&\leq \left(\mathrm{tr}(\widetilde{\mathbf{K}}) - \sum_{i=1}^{s} \left(\rho(G_{\mathcal{N}_\Phi}^{-1}) \|E_Y^{(i)}\|_2^2\right)\right) \|f\|_{\mathcal{N}_\Phi}^2 \\
&\leq \left(s \cdot \rho(\widetilde{\mathbf{K}}) - \rho(G_{\mathcal{N}_\Phi}^{-1}) \|E_Y\|_F^2\right) \|f\|_{\mathcal{N}_\Phi}^2.
\end{aligned}$$

□

Remark 10. *For the pointwise behaviour of the $\mathcal{N}_\Phi-$orthonormal basis U, the bounds obtained*

above become

$$\begin{aligned} |s_f(\mathbf{y})| &\leq \sqrt{K(\mathbf{y},\mathbf{y})} \cdot \|f\|_{\mathcal{N}_\Phi}, \\ |f(\mathbf{y}) - s_f(\mathbf{y})| &\leq \sqrt{K(\mathbf{y},\mathbf{y}) - \|U(\mathbf{y})\|_2^2} \cdot \|f\|_{\mathcal{N}_\Phi}, \end{aligned}$$

for fixed $\mathbf{y} \in \Omega$.

3.5 Duality

The goal of this section is to construct new class of bases that are dual to the general data-dependent bases $U = [u_1, \ldots, u_n]$, proposed for the finite-dimensional inner product subspace $\mathcal{N}_\Phi(X)$ of the native space $\mathcal{N}_\Phi(\Omega)$ associated to the CPD kernel Φ. The dual space $\mathcal{N}_\Phi{}^*$ consists of all linear functionals on $\mathcal{N}_\Phi$. Consider the dual functionals η_i such that

$$\eta_i(\alpha_1 u_1 + \cdots + \alpha_n u_n) = \alpha_i, \quad \alpha_i \in \mathbb{R},\ i = 1, \ldots, n,$$

which in turn leads to

$$\eta_i(u_j) = \delta_{ij}.$$

Then any linear functional $\eta \in \mathcal{N}_\Phi{}^*$ can be written as

$$\eta = \eta(u_1)\eta_1 + \eta(u_2)\eta_2 + \cdots + \eta(u_n)\eta_n.$$

Now by the *Riesz Representation* Theorem (8), every linear functional on $\mathcal{N}_\Phi{}^*$ has a representer in $\mathcal{N}_\Phi$. That is, for each η_i, there exists $d_i \in \mathcal{N}_\Phi$ such that

$$\eta_i(u_j) = \langle u_j, d_i \rangle = \delta_{ij}. \tag{3.22}$$

Therefore, we associate $\Lambda = \left[\eta_1, \ldots, \eta_n\right]$ with the representers $\mathcal{D} = [d_1, \ldots, d_n]$. Since Λ is linearly independent in $\mathcal{N}_\Phi{}^*$ and dual to U, then the so-called dual basis $\mathcal{D}$ is linearly independent in $\mathcal{N}_\Phi$ and also dual to U. Now let $(\mathcal{N}_\Phi, U, \mathcal{D})$ with basis $U = [u_1, \ldots, u_n]$ and dual basis $\mathcal{D} = [d_1, \ldots, d_n]$, then we can view the basis U as the map

$$\begin{aligned} U\colon\ \mathbb{R}^n &\to \mathcal{N}_\Phi \\ \alpha &\mapsto U(\alpha) = \sum_{j=1}^{n} \alpha_j u_j, \end{aligned}$$

and likewise, the dual basis $\mathcal{D}$ as

$$\begin{aligned} \mathcal{D}: \quad & \mathbb{R}^n \to \mathcal{N}_\Phi \\ & \alpha \mapsto \mathcal{D}(\alpha) = \sum_{j=1}^{n} \alpha_j d_j. \end{aligned}$$

Also the following dual map for identifying the dual space $\mathcal{N}_\Phi{}^*$ with $\mathcal{N}_\Phi$

$$\begin{aligned} \mathcal{D}^*: \quad & \mathcal{N}_\Phi \to \mathbb{R}^n \\ & f \mapsto \mathcal{D}^*(f) = \left[\langle f, d_1\rangle, \ldots, \langle f, d_n\rangle\right]^T. \end{aligned}$$

Then according to (3.22), $\mathcal{D}$ is dual to U exactly when

$$\mathcal{D}^*(U) = \left[\langle u_j, d_i\rangle_{\mathcal{N}_\Phi}\right]_{1\le i,j\le n} = I.$$

Theorem 36. *Let U be a general data-dependent basis, then for $(\mathcal{N}_\Phi, U, \mathcal{D})$, the dual basis $\mathcal{D}$ can be expressed in terms of the basis U as*

$$\mathcal{D} = U\mathcal{C},$$

where

$$\mathcal{C} = (U^*(U))^{-1},$$

is a symmetric, positive definite, and full-rank $n \times n$ matrix.

Proof. Let $\mathcal{D} = U\mathcal{C}$, then by applying U^* to both sides, we get

$$U^*(\mathcal{D}) = U^*(U)\mathcal{C},$$

which leads to

$$\mathcal{C} = (U^*(U))^{-1}U^*(\mathcal{D}).$$

Since $\mathcal{D}$ is dual to U, this reduces to $\mathcal{C} = (U^*(U))^{-1}$, which is nothing but the inverse of the $\mathcal{N}_\Phi$-Gramian matrix as

$$\mathcal{C} = \left(\left[\langle u_i, u_j\rangle_{\mathcal{N}_\Phi}\right]_{1\le i,j\le n}\right)^{-1} = \left(G_{\mathcal{N}_\Phi}\right)^{-1},$$

that is symmetric and positive definite with rank n. □

Remark 11. *For $(\mathcal{N}_\Phi, T, \mathcal{D})$, with the basis of translates*

$$T = \left[K(\cdot, \mathbf{x}_1), \ldots, K(\cdot, \mathbf{x}_n))\right],$$

we have

$$\mathcal{C} = (T^*(T))^{-1} = \mathbf{K}^{-1},$$

then

$$\mathcal{D} = T\mathbf{K}^{-1}.$$

So, the Lagrange bases and the bases T of translates are a dual pair.

Remark 12. *Among all data-dependent bases, the $\mathcal{N}_\Phi$-orthonormal bases are exactly those which are self-dual, since $\mathcal{C} = I$.*

Theorem 37. *Let U be a general data-dependent basis, then for $(\mathcal{N}_\Phi, U, \mathcal{D})$, the dual basis $\mathcal{D}$ can be expressed in terms of the basis T of translates as*

$$\mathcal{D} = T(E^T)^{-1},$$

where E is the evaluation matrix.

Proof. According to Theorem 36 and equation (3.3), we have

$$\mathcal{D} = U\mathcal{C} = TC(C^T\mathbf{K}C)^{-1} = T\mathbf{K}^{-1}(C^T)^{-1} = T(E^T)^{-1}.$$

□

Theorem 38. *Let V be the $\ell_{2,w}-$orthonormal basis functions proposed in Section 3.3.2, then for $(\mathcal{N}_\Phi, V, \mathcal{D})$, the dual basis $\mathcal{D}$ can be expressed in terms of the basis T of translates as*

$$\mathcal{D} = T\sqrt{W}Q.$$

Proof. According to the above theorem, we get

$$\mathcal{D} = T(E_2^T)^{-1} = T(Q^T(\sqrt{W})^{-1})^{-1} = T\sqrt{W}Q.$$

□

3.6 Numerical Experiments

For the numerical experiments, we consider three different underlying functions and three different types of CPD kernels, all of order 2, namely

- Generalized MQ RBF with $\beta = \frac{3}{2}$,
- Cubic RBF, $\phi_c(r) = r^3$, which is shape parameter free,
- Thin plate spline RBF, $\phi_{tps}(r) = r^2\log(r)$, which is shape parameter free too,

n	standard gMQ	Reproducing Kernel	SVD basis
20	1.9687e +17	1.5468e+17	3.9330e+08
50	9.3105e+17	2.3023e+18	1.0946e+08
80	1.2990e+19	5.5690e+18	1.0174e+08
110	1.2637e+19	3.0754e+18	1.2553e+08
150	5.0118e+19	1.911e+19	1.3472e+08

Table 3.1: 2-norm condition number of the interpolation matrix for different bases; Test problem 1.

where $r = \|\mathbf{x} - \mathbf{y}\|_2$ with $\mathbf{x}, \mathbf{y} \in \Omega \subset \mathbb{R}^d$. In the following subsection, **standard basis** refers to any of the above RBFs appended by polynomial space of the required degree (2.2), and **Reproducing kernel** refers to the corresponding PD kernel in (2.11). Besides, by truncated SVD basis, we mean the basis explained in subsection 3.3.3 such that the evaluation is selected to be E_{1k} in (3.15).

Moreover, working with generalized MQ RBF, one always needs to find the optimal value of shape parameter ε, which depends on the number and constellation of the data sites. In particular, ε values significantly affect the accuracy and stability of the interpolation process. However, we skip this task and always let $\varepsilon = 1$ since our numerical experiments show that with the suggested alternate bases, we obtain good accuracy even without optimizing the shape parameter. Moreover, in order to compute the accuracy of the interpolation, the *root mean square error* (RMSE) is computed as

$$RMSE = \sqrt{\frac{1}{s}\sum_{i=1}^{s}(f(\mathbf{z}_i) - s_f(\mathbf{z}_i))^2}, \tag{3.23}$$

where $\{\mathbf{z}_i\}_{i=1}^{s}$ is the set of evaluation points.

3.6.1 Test problem 1

Let us consider the *Runge* function

$$f(x) = \frac{1}{1 + 25x^2}, \quad x \in [-1, 1].$$

We reconstruct f using of *uniformly* distributed center points with different sizes $n = \{20, 50, 80, 110, 150\}$. Regarding the size of data sites, our interpolant is evaluated over an equispaced point set on Ω with size $s = 5n$. To evaluate the reproducing kernel (2.11), we let $\Xi = \{0, 1\}$ form the Lagrange linear basis for the polynomial space. Here we use $\mathcal{N}_\Phi-$orthonormal SVD bases, i.e., evaluation matrix $E = Q\sqrt{D}$. Table 3.1 shows the ℓ_2 condition number of the interpolation matrix using different bases. It is observable that the SVD bases lead to better conditioning. Figure 3.1(left) shows how more stable bases lead to better accuracy, particularly for an underlying function that is prone to inaccurate interpolation due

n	standard gMQ	Reproducing Kernel	Truncated SVD
9	2.5275e+05	7.5530e+04	274.8278
25	9.9028e+08	4.5283e+08	2.1280e+04
81	2.7716e+15	1.4380e+15	2.4357e+05
289	2.8478e+19	8.9075e+18	5.1423e+05
1089	5.6001e+20	6.1388+19	9.5946e+05
4225	3.9644e+21	1.1233e+22	1.9277e+06
10000	5.5799e+22	3.2801e+21	2.8522e+06

Table 3.2: 2-norm condition number of interpolation matrix for different bases; Test problem 2.

to its intrinsic oscillatory behavior.

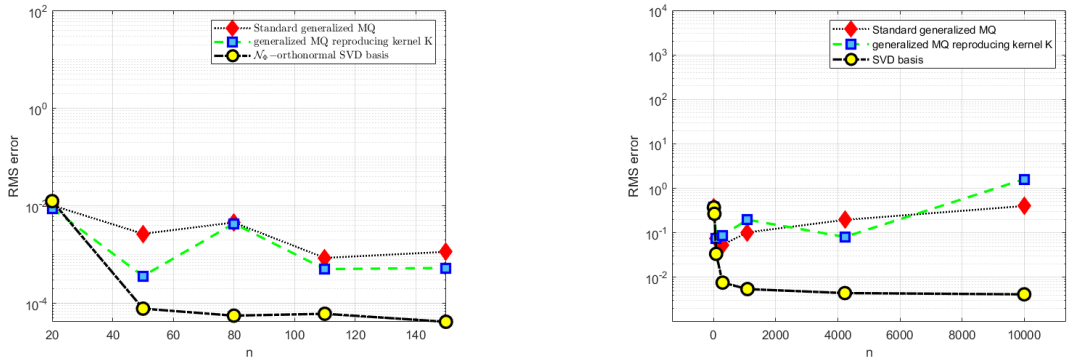

Figure 3.1: RMSE of Runge's (left) and Franke's functions (right) approximants using different bases; Test problems 1 and 2.

3.6.2 Test problem 2

For the second test problem, we take the *Franke* function, [Fas07, Chap 2] defined on $\Omega = [0,1]^2 \subset \mathbb{R}^2$ as the target function. The interpolation this time is done at the sequence of *Halton* center points with different sizes $n = \{9, 25, 81, 289, 1089, 4225, 10000\}$. Moreover, let $\Xi = \{(0,0), (0,1), (1,0)\}$ representing the Lagrange linear polynomials. Similarly, for each n, the interpolant is evaluated over a uniform grid with size $s = 2n$ on the domain of interest. We consider truncated $\mathcal{N}_\Phi-$orthonormal SVD basis, with threshold $\delta = 10^{-9}$, obtained by trial and error. It means zeroing all the eigenvalues of the interpolation matrix $\mathbf{K}$, which are smaller than δ. Table 3.2 shows the ℓ_2 condition number of the interpolation matrix for different bases. In Figure (3.1)(left), we show the RMSE of the interpolation using these three different bases. Once more, we recall that we avoided any shape parameter value optimization algorithm, and we just let $\varepsilon = 1$.

3.6.3 Test problem 3

Here, we reconstruct the oscillatory function $f(\mathbf{x}, \mathbf{y}) = \cos(20(\mathbf{x}+\mathbf{y}))$ defined on the unit disk with center $(0,0)$. To do so, consider the data set X consisting of 3000 *Halton* points on the

unit disk (see Figure 3.2(left)), Ξ as in the previous example, and the truncation sequence $k = \{20, 100, 500, 1100, 1800, 2400\}$ meaning that, in the first experiment, we take 20 singular values resulted from the SVD decomposition of the kernel matrix $\mathbf{K}$ in (3.2). We use ϕ_c and ϕ_{tps} RBFs to approximate $f(\mathbf{x}, \mathbf{y})$. To measure the accuracy of the reproduction process, we computed the RMSE on an equally spaced grid of evaluation points with size $s = 6000$ on the domain. The right plot in Figure 3.2) shows that, indeed, there is an excellent low-rank approximation to the problem.

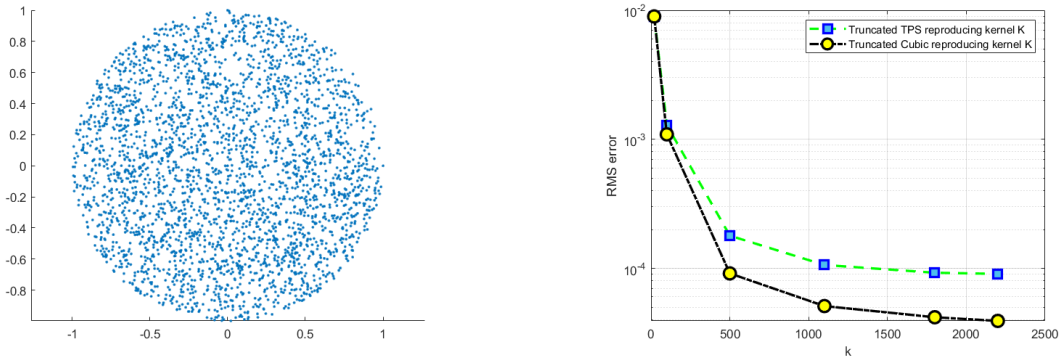

Figure 3.2: Data sites X (left), and the RMSE resulted from Truncated SVD approximation for two different bases (right); Test problem 3.

Remark 13. *We have to highlight that according to our discussion in 2.3, one always needs to make sure that the set $\Xi \subset X$, meaning that the subset used to build the Lagrange polynomials must belong to the set of data sites.*

Remark 14. *All three experiments show that the RMSE resulting from suggested better-conditioned bases, SVD bases, is improved compared with standard bases; however, increasing the number of data sites does not increase accuracy, meaning that the RMSE is stuck after some steps. This might be due to the shape parameter tuning being removed. Moreover, after some steps, the singular values of the kernel matrix are too small, so they only have subtle effects on the interpolation.*

3.7 Conclusion

Two different approaches have been presented to construct new stable bases for CPD kernel-based spaces. Both of these approaches are based on working with reproducing kernels of the corresponding Native Hilbert Space of CPD kernels. Inspired by [PS11], we used different factorizations of the kernel matrix to obtain other bases with various features. We also investigated the "natural" class of bases by the eigenpairs approximation of the linear operator associated with Mercer's theorem. The dual bases of the general data-dependent bases are also introduced.

Regarding stability, the experiments confirm the good behavior of the new bases expected from the analysis conducted in the previous sections. More precisely, employing a low-rank ap-

proximation of the kernel matrix enables the handling of approximations involving a relatively large number of points, also for not optimized shape parameters, and on quite general sets. From a numerical point of view, this procedure can be accomplished without thinning the data sites $X \subset \Omega$, but simply checking if the singular values of the kernel matrix decay under a certain tolerance. In this case, as a future work, one can consider a cost comparison and highlight the trade-off between cost and stability in a detailed manner. Another point worth considering is employing the Nystrom method in the eigenpairs approximation approach which poses new computational costs. The cubature rules to discretize the integrals in eigenpairs approximation can relate to this aspect of the problem.

Another facet of the newly established foundation through SVD factorization is its inability to employ an adaptive algorithm for singular value computation, necessitating a complete matrix factorization for every fixed point distribution. In this case, we can refer to optimized eigenvalue algorithms for finding only a subset of the full spectrum of the kernel matrix such as those presented, for example, in [DMS15, MB19]. Last but not least, as future work, one may consider Remark 9 in order to employ all these new stable bases to solve PDE problems.

4

Moving Least Squares Approximation using Variably Scaled Discontinuous Weight Functions

> The purpose of computing is insight, not numbers.
>
> Richard Hamming

4.1 Introduction

Discontinuous functions are widespread in various fields, including image reconstruction, signal processing, and engineering applications. Therefore, it is crucial to approximate and interpolate these functions accurately. In this chapter, we aim to develop a Moving Least Squares method for scattered data approximation that accounts for these discontinuities in the weight functions. The concept involves managing the impact of data points on the approximant, considering not only their distance to the evaluation point but also the discontinuities present in the underlying function.

Recall that $f : \Omega \to \mathbb{R}$ is a function sampled at some finite set of data sites $X = \{\mathbf{x}_i\}_{1=1}^{N} \subset \Omega \subset \mathbb{R}^d$ with corresponding data values $f_i = f(\mathbf{x}_i)$. Notice that in this chapter, we slightly changed our notation to use N for the number of data sites instead of n used in the previous chapters. We look to reconstruct f using the given function values, that is

$$s_{f,X}(\mathbf{x}) = \sum_{i=1}^{N} \alpha_i(\mathbf{x}) f_i, \tag{4.1}$$

with $\alpha_i, 1 \leq i \leq N$ known as generating or shape functions. In this case, one asks $s_{f,X}$ to exactly

reproduce a set of functions, let say radial basis functions. This leads to $\alpha = [\alpha_1, \ldots, \alpha_N]^T = (V_{\Phi,X}^{-1}\Phi_X(\mathbf{x}))$ which is the cardinal basis discussed in (2.17). However, now $s_{f,X}$ interpolates the data i.e. $s_{f,X}(\mathbf{x}_i) = f_i,\ 1 \leq i \leq N$. Instead, here we consider a more general framework known as *quasi-interpolation* in which $s_{f,X}$ only approximates the data, i.e., $s_{f,X}(\mathbf{x}_i) \approx f_i$. The latter case means that we prefer to let the approximant $s_{f,X}$ only nearly fits the function values, recall our discussion in subsection (1.5.1). This is useful, for instance, when the given data contains some noise or the number of data is too large. The standard approach to deal with such a problem is to compute the Least-Squares (LS) solution, recall (1.28), i.e., one minimizes the error (or cost) function

$$\sum_{i=1}^{N} [s_{f,X}(\mathbf{x}_i) - f_i]^2. \tag{4.2}$$

An alternative setting of LS is known as the weighted LS, in which (4.2) turns to

$$\sum_{i=1}^{N} [s_{f,X}(\mathbf{x}_i) - f_i]^2 w(\mathbf{x}_i), \tag{4.3}$$

which is ruled by the *weighted* discrete ℓ_2 inner product. In practice, the function $w(\mathbf{x}_i)$ is incorporated to enhance the least squares (LS) formulation for data points f_i, especially those influenced by factors such as noise. However, these methods operate globally, meaning all the data sites influence the solution at any evaluation point $\mathbf{x} \in \Omega$. Alternatively, one can opt for a localized approach by considering only the n closest data sites $\mathbf{x}_i$, $i = 1, \ldots, n$ to $\mathbf{x}$, where $n \ll N$, for a fixed evaluation point $\mathbf{x}$. Following this idea, the Moving Least Squares (MLS) method, a local adaptation of the classical weighted least-squares technique, has been developed. To be more precise, in the MLS scheme, for each evaluation point $\mathbf{x}$, one needs to solve a *weighted least-squares* problem, minimizing

$$\sum_{i=1}^{N} [s_{f,X}(\mathbf{x}_i) - f_i]^2 w(\mathbf{x}, \mathbf{x}_i) \tag{4.4}$$

by choosing the weight functions $w(\mathbf{x}, \mathbf{x}_i) : \mathbb{R}^d \times \mathbb{R}^d \longrightarrow \mathbb{R}$ to be localized around $\mathbf{x}$, so that only few data sites are taken into account. The key difference with (4.3) lies in the fact that now the weight function is indeed *moving* with respect to the evaluation point, meaning that it depends on both the $\mathbf{x}_i$ and $\mathbf{x}$. Consequently, for each evaluation point $\mathbf{x}$, a small linear system needs to be solved. Additionally, one can let $w(\cdot, \mathbf{x}_i)$ be a radial function i.e., $w(\mathbf{x}, \mathbf{x}_i) = \phi(\|\mathbf{x} - \mathbf{x}_i\|_2)$ for some non-negative univariate function $\phi : [0, \infty) \to \mathbb{R}$. Using this approach, $w(\cdot, \mathbf{x}_i)$ inherits the translation invariance property of radial basis functions. We mention that (4.4) could be generalized as well by letting $w_i(\cdot) = w(\cdot, \mathbf{x}_i)$ moves with respect to a *reference* point $\mathbf{y}$ such that $\mathbf{y} \neq \mathbf{x}$.(see e.g [Fas07, Chap 22]).

The inception of the MLS approximation technique can be attributed to Shepard's pio-

neering work [She68], where he explored approximation using constants. Subsequently, Lancaster and Salkauskas introduced the general framework of MLS in [LS81], offering an analysis of its application in smoothing and interpolating scattered data. Building upon this foundation, Bos investigated the relationship between MLS and the Backus-Gilbert approach in [BS89], demonstrating its efficacy in approximating derivatives. Since then, the MLS method has demonstrated its utility across various domains [MSD12, MS13]. Error analyses of MLS approximation have been conducted by several researchers, with Levin's work [Lev98] serving as a pivotal reference point. Wendland provided error bounds, considering the fill distance, as outlined in [Wen04, Chap. 3 & 4] and [Wen01]. Armento et al. explored theoretical aspects in [AD01], presenting L_∞ error estimates for one-dimensional cases, later extending their findings to multi-dimensional scenarios in [Arm01], focusing on the support of weight functions rather than the fill distance. More recently, Mirzaei derived error estimates for MLS approximation of functions belonging to integer or fractional-order Sobolev spaces in [Mir15], echoing the findings previously studied in [NWW05] for kernel-based interpolation.

The MLS method has rarely been used to approximate piecewise-continuous functions, i.e., functions with some discontinuities or jumps. In this case, it would be essential that the approximant takes into account the location of the discontinuities. To address these challenges in this chapter, we let the weight function be a *Variably Scaled Discontinuous Kernel* (VSDK) [DMMP20]. VSDK interpolants have been employed to mitigate the Gibbs phenomenon, outperforming classical kernel-based interpolation in [DMEM$^+$20]. Similarly, in the MLS approximation framework, the usage of VSDK weights allows the construction of data-dependent approximants (as discussed in [Lev98]) that are able to overcome the performances of classical MLS approximants, as indicated by careful theoretical analysis and then assessed by various numerical experiments.

This chapter is organized as follows. Section 4.2 recalls necessary notions of the MLS, VSDKs, and Sobolev spaces. Section 4.3 presents this work's original contribution, consisting of variably scaled discontinuous weights for reconstructing discontinuous functions in the framework of MLS approximation. The error analysis shows that the MLS-VSDKs approximation can outperform classical MLS schemes as the discontinuities of the underlying function are assimilated into the weight function. In Section 4.4, we discuss some numerical experiments that support our theoretical findings, and in Section 4.5, we draw some conclusions.

4.2 Preliminaries on MLS and VSKs

This section reviews the basic concepts and notions required in this chapter.

4.2.1 Moving Least Squares (MLS) approximation

In this subsection, we introduce the MLS method more precisely and investigate its properties. To this end, let $\mathbb{P}_m^d$ indicates the space of d-variate polynomials of degree at most $m \in \mathbb{N}$, with

basis $\{p_1, \ldots, p_q\}$ and dimension $q = \binom{m+d}{d}$. Notice that we used $m-1$ instead of m in the previous chapters. This change will appear useful to have simpler notation later.

As the standard formulation, the MLS approximant looks for the best-weighted approximation to f at the evaluation point $\mathbf{x}$ in $\mathbb{P}_m^d$, concerning the discrete ℓ_2 norm induced by the weighted inner product

$$\langle f, g \rangle_{w_\mathbf{x}} = \sum_{i=1}^{N} w(\mathbf{x}_i, \mathbf{x}) f(\mathbf{x}_i) g(\mathbf{x}_i). \tag{4.5}$$

Mathematically speaking, the MLS approximant will be the linear combination of the polynomial basis, i.e.,

$$s_{f,X}(\mathbf{x}) = \sum_{j=1}^{q} c_j(\mathbf{x}) p_j(\mathbf{x}), \tag{4.6}$$

where the coefficients are obtained by solving by constructing the Gram matrix concerning the inner product in (4.5) and solving the linear system in (1.32). The dependency of $c_j(\mathbf{x})$ to the evaluation point $\mathbf{x}$ highlights the local nature of the approximant $s_{f,X}$ given in (4.6). Moreover, from (4.5) we observe that the weight function $w_i(\mathbf{x}) = w(\mathbf{x}, \mathbf{x}_i)$ controls the influence of the center $\mathbf{x}_i$ over the approximant, so it should be *small* when evaluated at a point that is far from $\mathbf{x}$, that is it should decay to zero fast enough. So, compactly supported RBFs, or decaying globally supported RBFs like Gaussian RBFs, introduced in the subsection (2.2), can be a good candidate for the weight function.

To relate the representation in (4.6) to *quasi interpolant* of the form (4.1), let $I(\mathbf{x}) = \{i \in \{1, \ldots, N\}, \|\mathbf{x} - \mathbf{x}_i\|_2 \leq r\}$ be the family of indices of the centers X, for which $w_i(\mathbf{x}) > 0$, with $|I| = n \ll N$. Therefore, only the centers $\mathbf{x}_i \in I$ influence the approximant $s_{f,X}(\mathbf{x})$. Consequently, we can reformulate (4.4) as

$$\min_{\alpha_i} \left\{ \sum_{i \in I(\mathbf{x})} [s_{f,X}(\mathbf{x}_i) - f_i]^2 w(\mathbf{x}, \mathbf{x}_i) \right\} \tag{4.7}$$

Theorem 39. *[Wen04, Theorem 4.3] Suppose that for every $\mathbf{x} \in \Omega$ the set $\{\mathbf{x}_j : j \in I(\mathbf{x})\}$ is $\mathbb{P}_m^d$-unisolvent. In this situation, problem (4.7) is uniquely solvable and the solution $s_{f,X} = p^*(\mathbf{x})$ can be represented as*

$$s_{f,X} = \sum_{i \in I(\mathbf{x})} \alpha_i(\mathbf{x}) f(\mathbf{x}_i) \tag{4.8}$$

where the coefficients $\alpha_i(\mathbf{x})$ are determined by minimizing the quadratic form

$$\sum_{i \in I(\mathbf{x})} \alpha_i^2(\mathbf{x}) \frac{1}{w(\mathbf{x}_i, \mathbf{x})} \tag{4.9}$$

subject to the polynomial reproduction constraints

$$\sum_{i\in I(\mathbf{x})} p(\mathbf{x}_i)\alpha_i(\mathbf{x}) = p(\mathbf{x}), \quad \textit{for all } p \in \mathbb{P}^d_m.$$

According to [Weno4, Corollary 4.4]) the basis α_i can be obtained as follows:

Lemma 4. *The basis functions α_i are given by*

$$\alpha_i(\mathbf{x}) = w(\mathbf{x}, \mathbf{x}_i) \sum_{k=1}^{q} \lambda_k(\mathbf{x}) p_k(\mathbf{x}_i), \quad i \in I(\mathbf{x}), \tag{4.10}$$

where $\lambda_k(\mathbf{x})$ are the unique solution of

$$\sum_{k=1}^{q} \lambda_k(\mathbf{x}) \sum_{i\in I(\mathbf{x})} w(\mathbf{x}, \mathbf{x}_i) p_k(\mathbf{x}_i) p_s(\mathbf{x}_i) = p_\ell(\mathbf{x}), \quad 1 \le \ell \le q. \tag{4.11}$$

Furthermore, thanks to equation (4.10), it is observable that the behavior of $\alpha_i(\mathbf{x})$ is heavily influenced by the behavior of the weight functions $w_i(\mathbf{x})$, in particular it includes continuity and the support of the basis functions $\alpha_i(\mathbf{x})$. In other terms, $\{\alpha_i(\mathbf{x})\}$ depends on the set X, but they might not even be continuous. Another significant feature is that the weight functions $w_i(\mathbf{x})$, which are singular at the data sites, lead to cardinal bases functions, meaning that the MLS scheme interpolates the data (for more details, see [Lev98, Theorem 3]). Consequently, the matrix representation of (4.10) and (4.11) is

$$\begin{aligned} \alpha(\mathbf{x}) &= W(\mathbf{x}) P^T \lambda(\mathbf{x}), \\ \lambda(\mathbf{x}) &= (PW(\mathbf{x})P^T)^{-1}\mathbf{p}(\mathbf{x}), \end{aligned}$$

where $\alpha(\mathbf{x}) = [\alpha_1(\mathbf{x}), ..., \alpha_n(\mathbf{x})]^T$, $W(\mathbf{x}) \in \mathbb{R}^{n\times n}$ is the diagonal matrix carrying the weights $w_i(\mathbf{x})$ on its diagonal, $P \in \mathbb{R}^{q\times n}$ such that its k-th row contains p_k evaluated at data sites in $I(\mathbf{x})$, and $\mathbf{p}(\mathbf{x}) = [p_1(\mathbf{x}), ..., p_q(\mathbf{x})]^T$. More explicitly, the basis functions are given by

$$\alpha(\mathbf{x}) = W(\mathbf{x}) P^T (PW(\mathbf{x})P^T)^{-1}\mathbf{p}(\mathbf{x}). \tag{4.12}$$

In the MLS literature, it is known that a local polynomial basis shifted to the evaluation point $\mathbf{x} \in \Omega$ leads to a more stable method (see, e.g., [Weno4, Chap. 4]). Accordingly, we let the polynomial bases be $\{1, (\cdot - \mathbf{x}), (\cdot - \mathbf{x})^2, \ldots, (\cdot - \mathbf{x})^m\}$, meaning that different bases for each evaluation point are employed. In this case, since with standard monomials bases, we have $p_1 \equiv 1$ and $p_k(0) = 0$ for $2 \le k \le q$, then $\mathbf{p}(\mathbf{x}) = [1, 0, \ldots, 0]^T$.

We shall review some technical notions concerning the geometry of Ω that will appear useful in the error analysis of the MLS scheme.

1. A set $\Omega \subset \mathbb{R}^d$ is said to satisfy an **interior cone condition** if there exists an angle $\Theta \in$

$(0, \pi/2)$ and a radius $r > 0$ so that for every $\mathbf{x} \in \Omega$ a unit vector $\xi(\mathbf{x})$ exists such that the cone

$$C(\mathbf{x}, \xi, \Theta, r) = \{\mathbf{x} + t\mathbf{y} : \mathbf{y} \in \mathbb{R}^d, \|\mathbf{y}\|_2 = 1, \cos(\Theta) \leq \mathbf{y}^T\xi, t \in [0, r]\}$$

is contained in Ω.

2. A domain $\mathcal{D} \subset \mathbb{R}^d$ is said to be **star-shaped** with respect to a ball $B = B(\mathbf{y}, \rho) = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x} - \mathbf{y}\| \leq \rho\}$ if for every $\mathbf{x} \in \mathcal{D}$ the closed convex hull of $\mathbf{x} \cup B$ is contained in $\mathcal{D}$.If Ω is bounded, i.e., $\Omega \subset B(\mathbf{y}, R)$ for some $R > 0$, then the chunkiness parameter γ is defined to be the ratio of the diameter $diam_\Omega$ to the radius ρ_Ω of the largest ball relative to which is star-shaped, i.e., $\gamma = \frac{diam_\Omega}{\rho_max}$.

Equipped with the above notions, we express the following theorem

Theorem 40. *[Wen04, Theorem 4.7] Suppose that $\Omega \subset \mathbb{R}^d$ is compact and satisfies an interior cone condition with angle $\theta \in (0, \pi/2)$ and radius $r > 0$. Fix $m \in \mathbb{N}$. Let h_0, C_1 and C_2 denote the constants. Suppose that $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset \Omega$ satisfies (2.20) and $h_{X,\Omega} < h_0$. Let $\delta = 2C_2 h_{X,\Omega}$. Then the bases functions $a_j(\mathbf{x})$ from the Lemma (4) provide local polynomial reproduction, i.e.*

1. $\sum_{j=1}^N \alpha_j(\mathbf{x}) p(\mathbf{x}_j) = p(\mathbf{x})$ *, for all* $p \in \mathbb{P}_m^d$, $\mathbf{x} \in \Omega$,
2. $\sum_{j=1}^N |\alpha_j| \leq \tilde{C}_1$,
3. $\alpha(\mathbf{x}) = 0$ *if* $\|\mathbf{x} - \mathbf{x}_j\|_2 > \tilde{C}_2 h_{X,\Omega}$,

with certain constants $\tilde{C}_1$, $\tilde{C}_2$ that can be derived explicitly.

The crucial point in the definition is that the constants involved are independent of the data sites. The first and the third conditions justify the name local polynomial reproduction. In particular, the third condition is a consequence of the compact support of w. The second condition is important for the approximation property of the associated quasi-interpolant.

Since the MLS scheme reproduces polynomials at least locally, one expects that the MLS scheme's approximation order inherits the order of polynomial approximation.

Theorem 41. *[Wen04, Corollary 4.8] In the situation of Theorem (40) define Ω^* to be the closure of $\bigcup_{\mathbf{x} \in \Omega} B(\mathbf{x}, 2C_2 h_0)$. Then, there exists a constant $c > 0$ that can be computed explicitly, such that for all $f \in C^{m+1}(\Omega^*)$ and all quasi-uniform $X \subset \Omega$ with $h_{X,\Omega} < h_0$ the approximation error is bounded as follows:*

$$\|f - s_{f,X}\|_{L_\infty}(\Omega) \leq c h_{X,\Omega}^{m+1} \max_{|\delta|=m} \|D^\delta f\|_{C^{m+1}}(\Omega^*). \tag{4.13}$$

The norm on the right-hand side of inequality (4.13) implies that the target function f is continuous of order m, i.e., $f \in C^{m+1}(\Omega^*)$, but here we are looking to approximate the functions that possess some discontinuity. Thus, it makes sense that f lies in another function space.

With this motivation, we review basic notions of *Sobolev* spaces and investigate the error bounds in such spaces.

4.2.2 Sobolev spaces and error estimates for MLS

For any positive integer k and $1 \leq p \leq \infty$ the *Sobolev* space is defined as

$$W_p^k(\Omega) = \{u \in L^p \ : \ D^\delta u \in L^p \text{ for } 0 \leq |\delta| \leq k\}$$

where $D^\delta u$ is the weak *distributional* (weak) derivative of u. Equipped with norms

$$\|u\|_{w_p^k(\Omega)} := \Big(\sum_{|\delta| \leq k} \|D^\delta u\|^p_{L^p(\Omega)} \Big)^{1/p}. \quad 1 \leq p < \infty \tag{4.14}$$

$$\|u\|_{w_\infty^k(\Omega)} := \max_{0 \leq \delta \leq k} \|D^\delta u\|_\infty \tag{4.15}$$

W_p^k is called Sobolev space over Ω Notice that, $W_p^0(\Omega) = L^p(\Omega)$. Moreover, letting $0 < s < 1$, the *fractional-order* Sobolev space $W_p^{k+s}(\Omega)$ is the space of the functions u for which semi-norm and norm are defined as

$$\|u\|_{W_p^{k+s}(\Omega)} := \Big(\|u\|_{W_p^k(\Omega)} + |u|_{W_p^{k+s}(\Omega)} \Big)^{1/p},$$

with $|u|_{w_p^{k+s}(\Omega)} := \Big(\sum_{|\delta|=k} \int_\Omega \int_\Omega \frac{|D^\delta u(\mathbf{x}) - D^\delta u(\mathbf{y})|^p}{|\mathbf{x}-\mathbf{y}|^{d+ps}} d\mathbf{x} d\mathbf{y} \Big)^{1/p}$.

The spaces $W_p^k(\Omega)$ were introduced by Sobolev [Sob38]. Afterward, many related spaces under different symbols and names were studied. The readers can find a vast discussion in [Sob08, AF03]. We will not investigate these generalizations as they are unrelated to this work, but we will only review some basic properties and definitions regarding these spaces.

Theorem 42. *[Bre08, Theorem 3.3] The Sobolev space $W_p^k(\Omega)$ is a Banach space.*

Theorem 43. *[AF03, Theorem 3.6] The Sobolev space $W_2^k(\Omega)$ is a separable Hilbert space with inner product*

$$(u, v)_k = \sum_{0 \leq |\delta| \leq k} (D^\delta u, D^\delta v),$$

where $(u, v) = \int_\Omega u(\mathbf{x}) \bar{v}(\mathbf{x}) d\mathbf{x}$ is the inner product on $L^2(\Omega)$.

Moreover, given the number of indices defining Sobolev spaces, it is natural to expect that there are inclusion relations to provide some ordering among them.

Theorem 44. *[AF03, Prop 1.4.1] Suppose that Ω is any domain, k and z are non-negative integers satisfying $k \leq z$, and p is any real number satisfying $1 \leq p \leq \infty$. Then $W_p^z(\Omega) \subset W_p^k(\Omega)$.*

Theorem 45. *[AF03, Prop 1.4.2] Suppose that Ω is a bounded domain, k is a non-negative integer, and p and q are real numbers satisfying $1 \leq p \leq q \leq \infty$. Then $W_q^k(\Omega) \subset W_p^k(\Omega)$.*

If u lies in a Sobolev space, D^δ may not exist in the usual (pointwise) sense. Now, we turn our attention to introducing a polynomial approximation of degree m for a function in a Sobolev space. Let B be a ball with respect to which D is *star-shaped*[Bre08, Def 4.2.2] having radius $\rho \geq \frac{1}{2}\rho_{\max}$. Then

$$Q_m u(\mathbf{x}) := \sum_{|\delta|\leq m} \frac{1}{\delta!} \int D^\delta u(\mathbf{y})(\mathbf{x}-\mathbf{y})^\delta \varphi(\mathbf{y}) d\mathbf{y}, \tag{4.16}$$

where $\varphi(\mathbf{y}) \geq 0$ is a C^∞ *bump* function supported in B satisfying both $\int_B \varphi(\mathbf{y})d\mathbf{y} = 1$ and $\max \varphi \leq C\rho^{-d}$. Sobolev error bounds for a function approximated by the *averaged Taylor polynomial* defined in (4.16), on a star-shaped domain are discussed in [Bre08, NWW05, NWW06]. These bounds are usually used for analyzing the finite element method (FEM); however, in [Mir15], the author suggested Sobolev error bound for the MLS scheme.

Consequently, let D^δ be a derivative operator such that $|\delta| \leq m$ (we recall that m is the maximum degree of the polynomials). Under some mild conditions regarding the weight functions, [Mir15, Theorem 3.11] shows that $\{D^\delta \alpha_i(\mathbf{x})\}_{1\leq i\leq n}$ forms a *local polynomial reproduction* in a sense that there exist constants $h_0,\ C_{1,\delta},\ C_2$ such that for every evaluation point $\mathbf{x}$

- $\sum_{i=1}^N D^\delta \alpha_i(\mathbf{x}) p(\mathbf{x}_i) = D^\delta p(\mathbf{x}) \quad$ for all $p \in \mathbb{P}_m^d$
- $\sum_{i=1}^N |D^\delta \alpha_i(\mathbf{x})| \leq C_{1,\delta} h_{X,\Omega}^{-|\delta|}$
- $D^\delta \alpha_i(\mathbf{x}) = 0$ provided that $\|\mathbf{x}-\mathbf{x}_i\|_2 \geqslant 2C_2 h_{X,\Omega}$

for all X with $h_{X,\Omega} \leq h_0$.

Theorem 46. *[Mir15, Theorem 3.12] Suppose that $\Omega \subset \mathbb{R}^d$ is a bounded set with a Lipschitz boundary. Let m be a positive integer, $0 \leq s < 1,\ p \in [1,\infty)\ , q \in [1,\infty]$ and let δ be a multi-index satisfying $m > |\delta| + d/p$ for $p > 1$ and $m \geqslant |\delta| + d$ for $p = 1$. If $f \in W_p^{m+s}(\Omega)$, there exist constants $C > 0$ and $h_0 > 0$ such that for all $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset \Omega$ which are quasi-uniform with $h_{X,\Omega} \leq \min\{h_0, 1\}$, the error estimate holds*

$$\|f - s_{f,X}\|_{W_q^{|\delta|}(\Omega)} \leq C h_{X,\Omega}^{m+s-|\delta|-d(1/p-1/q)_+} \|f\|_{W_p^{m+s}(\Omega)}. \tag{4.17}$$

when the polynomial bases are shifted to the evaluation point $\mathbf{x}$ and scaled with respect to the fill distance $h_{X,\Omega}$, and $w_i(\cdot)$ is positive on $[0,1/2]$, supported in $[0,1]$ such that its even extension is non-negative and continuous on $\mathbb{R}$.

Remark 15. *The above error bound holds also when $s = 1$. However, recalling the definition of (semi-)norms in fractional-order Sobolev space, we see that in this case, we reach an integer-order Sobolev space of $m+1$. Therefore, it requires that $m+1 > |\delta| + d/p$ for $p > 1$ or $m+1 \geqslant |\delta|$ for $p = 1$ in order that (4.17) holds true. The key point is that in this case, the polynomial space is still $\mathbb{P}_m^d$ and not $\mathbb{P}_{m+1}^d$.*

4.2.3 Variably Scaled Discontinuous Kernels (VSDKs)

Recalling the shape parameter ε in subsection (2.2), one might be interested in applying a different shape parameter to each basis function in the RBF interpolation expansion. Taking Gaussian RBF as an example, the approximant (interpolant) can be represented as

$$s_{f,X} = \sum_{i=1}^{N} c_i \exp(-\varepsilon_i^2 \|\cdot - \mathbf{x}_i\|^2).$$

However, in such instances, the interpolant is not created using a single kernel, requiring a reevaluation of the theoretical framework. As far as we are aware, the most promising approach to explore this approach, particularly from a theoretical standpoint, is [BLRS15], where Variably Scaled Kernels (VSKs) were first introduced. The basic idea behind them is to map the data sites from $\mathbb{R}^d$ to $\mathbb{R}^{d+1}$ via a scaling function $\psi : \Omega \longrightarrow \mathbb{R}$ and to construct an augmented approximation space in which the data sites are $\{(\mathbf{x}_i, \psi(\mathbf{x}_i))\ i = 1, ..., N\}$ (see [MS13, Def. 2.1]). Though the first goal of doing so was getting a *better* nodes distribution in the augmented dimension, later on in [DMMP20], the authors came up with the idea of also encoding the behavior of the underlying function f inside the scale function ψ. To be more precise, the key idea for the target function f that possesses some jumps is as follows.

Definition 27. *Let $\mathcal{P} = \{\Omega_1, ..., \Omega_n\}$ be a partition of Ω and let $\beta = (\beta_1, ..., \beta_n)$ be a vector of real distinct values. Moreover, assume that all the jump discontinuities of the underlying function f lie on $\bigcup_{j=1}^{n} \partial\Omega_j$. The piecewise constant scaling function $\psi_{\mathcal{P},\beta}$ with respect to the partition $\mathcal{P}$ and the vector β is defined as*

$$\psi_{\mathcal{P},\beta}(\mathbf{x})|_{\Omega_j} = \beta_j, \ \mathbf{x} \in \Omega.$$

Successively, let Φ^ε be a positive definite radial kernel on $\Omega \times \Omega$ that depends on the shape parameter $\varepsilon > 0$. A variably scaled discontinuous kernel on $(\Omega \times \mathbb{R}) \times (\Omega \times \mathbb{R})$ is defined as

$$\Phi^\varepsilon_\psi(\mathbf{x}, \mathbf{y}) = \Phi^\varepsilon\big(\Psi(\mathbf{x}), \Psi(\mathbf{y})\big), \quad \mathbf{x}, \mathbf{y} \in \Omega. \tag{4.18}$$

such that $\Psi(\mathbf{x}) = (\mathbf{x}, \psi(\mathbf{x}))$.

Moreover, we point out that if Φ^ε is (semi-)positive definite then so is Φ^ε_ψ, and if Φ^ε and ψ are continuous then so is Φ^ε_ψ [BLRS15, Theorem 2.2]. Figure 4.1 shows two different choices for the discontinuous scale function for the univariate case. In any case, it matters that the discontinuities of the target function f are assimilated into the kernel Φ^ε_Ψ.

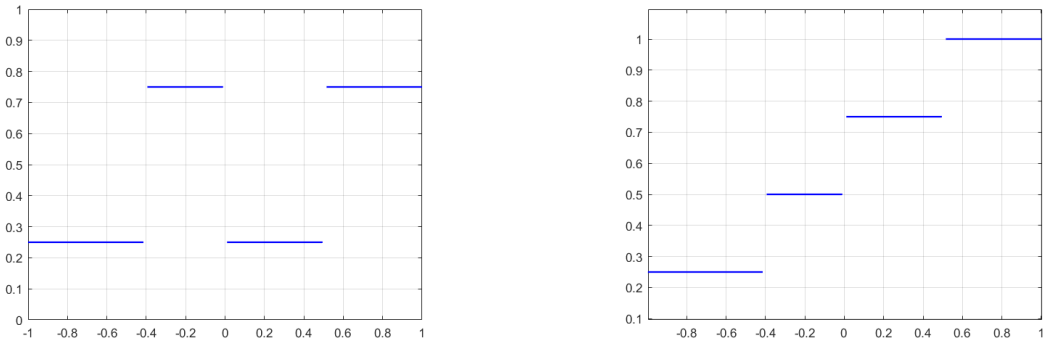

Figure 4.1: Discontinuous scale functions.

4.3 MLS-VSDKs

Let f be a function with some discontinuities defined on Ω, $\mathcal{P}$ and $\psi_{\mathcal{P},\beta}$ as in Definition 27. We look for the MLS approximant with *variably scaled discontinuous weight function* such that

$$w_\psi(\mathbf{x}, \mathbf{x}_i) = w(\Psi(\mathbf{x}), \Psi(\mathbf{x}_i)), \tag{4.19}$$

where w can be any decaying globally supported or compactly supported RBF. Taking the latter, let $B(\mathbf{x}, r)$ and $I(\mathbf{x})$ as before. Now, the MLS-VSDK approximant is represented as

$$s^\Psi_{f,X}(\mathbf{x}) = \sum_{i \in I(\mathbf{x})} \alpha^\Psi_i(\mathbf{x}) f_i, \tag{4.20}$$

where the shape functions are decided by solving

$$\min_{\alpha_i} \left\{ \sum_{i \in I(\mathbf{x})} [s^\Psi_{f,X}(\mathbf{x}_i) - f_i]^2 w_\psi(\mathbf{x}, \mathbf{x}_i) \right\}. \tag{4.21}$$

Notice that $I(\mathbf{x})$ is defined with respect to the non-scaled weight function w, meaning that the scale function applies after selecting the $\mathbf{x}_i$ that influences the approximante. In this case, the weight function is not necessarily radial now. Thus we define $I^\Psi(\mathbf{x})$ to be the set of indices for which $w_\psi(\mathbf{x}, \mathbf{x}_i) > 0$.

Theorem 47. *Let w be as (4.19), such that for every evaluation point $\mathbf{x} \in \Omega$, the set $\{\mathbf{x}_i : i \in I(\mathbf{x})\}$ is $\mathbb{P}^d_m$−unisolvent. In this setting, the MLS-VSDK scheme in (4.21) is uniquely solvable if for the augmented data sites $\{(\mathbf{x}_i, \psi(\mathbf{x}_i)\}_{i=1}^N$, we have $w_\psi(\cdot, \mathbf{x}_i) > 0$ for all $i \in I(\mathbf{x})$.*

Proof. Considering that we are looking for the best approximation in polynomial space, one must be able to find a representation for $s_{f,X}$ such that $s_{f,X} = \sum_{j=1}^q c_j p_j$. Let us fix the following

notation,

$$\begin{aligned}
\mathbf{c} &= (c_1, \ldots, c_q)^T \\
\mathbf{f} &= (f(\mathbf{x}_i);\ i \in I(\mathbf{x})) \\
P &= (p_j(\mathbf{x}_i))_{i \in I(\mathbf{x}), 1 \leq j \leq q}. \\
W(\mathbf{x}) &= diag(w_\psi(\mathbf{x}, \mathbf{x}_i) : i \in I(\mathbf{x})) \\
\mathbf{p}(\mathbf{x}) &= (p_1(\mathbf{x}), \ldots, p_Q(\mathbf{x}))^T.
\end{aligned}$$

Now, we have to minimize the equation (4.21), which leads us to

$$C(\mathbf{c}) = \mathbf{f}^T W(\mathbf{x})\mathbf{f} - 2\mathbf{f}^T W(\mathbf{x})P\mathbf{c} + \mathbf{c}^T P^T W(\mathbf{x})P\mathbf{c}. \tag{4.22}$$

Since $C(\mathbf{c})$ is a quadratic function in $\mathbf{c}$, we get the unique solution if $P^T W(\mathbf{x})P$ is positive definite. We notice that,

$$\mathbf{c}^T P^T W(\mathbf{x})P\mathbf{c} = \| W^{1/2}(\mathbf{x})P\mathbf{c}\|_2.$$

Consequently, $P^T W(\mathbf{x})P$ is positive semi-definite. Since $W(\mathbf{x})$ carries only positive weights, $\mathbf{c}^T P^T W(\mathbf{x})P\mathbf{c} = 0$ implies that $P\mathbf{c} = 0$. The unisolvency of $I(\mathbf{x})$ forces $\mathbf{c}$ to be zero. So, $P^T W(\mathbf{x})P$ is positive definite, and the solution uniquely exists. Besides, to treat the Backus-Gilbert representation of MLS in (4.20), it requires that $(W(\mathbf{x}))^{-1}$ exists and be positive definite. This is guaranteed since w_ψ are positive. □

Remark 16. *Some remarks are in order:*

- *The shape functions in* (4.21) *are indeed data-dependent thanks to* (4.19) *since they are modified with respect to the underlying function f.*
- *In case that the weight function w is radial, then it is straightforward to verify* $I^\Psi(\mathbf{x}) \subseteq I(\mathbf{x})$. *By definition, we have*

  $$\|\Psi(\mathbf{x}) - \Psi(\mathbf{x}_i)\|^2 = \|\mathbf{x} - \mathbf{x}_i\|^2 + \|\psi(\mathbf{x}) - \psi(\mathbf{x}_i)\|^2,$$

  which is obviously greater than

  $$\|\mathbf{x} - \mathbf{x}_i\|^2.$$

  On the other hand, the requirement in Theorem (47), $w_\psi(\cdot, \mathbf{x}_i) > 0$ *for all* $i \in I(\mathbf{x})$, *requires that* $|I(\mathbf{x})| = |I^\Psi(\mathbf{x})|$. *These imply* $I(\mathbf{x}) = I^\Psi(\mathbf{x})$.

The last remark poses a significant interpretation that is, the value of the scaling function ψ must be selected in a way that $\|\Psi(\mathbf{x}) - \Psi(\mathbf{x}_i)\|$ remains in $B(\mathbf{x}, r)$ so that $W(\mathbf{x})$ in (4.22) carries only positive values.

Lemma 5. *Under the assumption of Theorem* (40), *the MLS-VSDK scheme* (4.21) *satisfies polynomial reproduction property.*

Proof. Since polynomials are continuous functions, then $w_\psi(\cdot, \mathbf{x}_i) = w(\cdot, \mathbf{x}_i)$ for any $\mathbf{x} \in \Omega$ and $I(\mathbf{x})$. Then, according to Theorem (40), the polynomial reproduction property is concluded. □

Since the bases functions are data-dependent, one might expect that the space in which we express the error bound should also be data-dependent. Towards this idea, for $k \in \mathbb{Z},\ 0 \leq k$, and $1 \leq p \leq \infty$, we define the *piecewise* Sobolev Spaces

$$\mathcal{W}_p^k(\Omega) = \{f : \Omega \to \mathbb{R} \text{ s.t. } f_{|\Omega_j} \in W_p^k(\Omega_j), \quad j \in \{1, ..., n\}\},$$

where $f_{|\Omega_j}$ denotes the restriction of f to Ω_j, and $W_p^k(\Omega_j)$ denote the Sobolev space on Ω_i. We endow $\mathcal{W}_p^k(\Omega)$ with the norm

$$\|f\|_{\mathcal{W}_p^k(\Omega)} = \sum_{j=1}^{n} \|f\|_{W_p^k(\Omega_j)}. \tag{4.23}$$

Theorem 48. $\|\cdot\|_{\mathcal{W}_p^k(\Omega)}$ *defined in* (4.23) *defines a norm.*

Proof. Let $u \in \mathcal{W}_p^k(\Omega)$. Since $\|u\|_{W_p^k}$ is a nonnegative value, then $\|u\|_{\mathcal{W}_p^k} = 0$ implies $u = 0$. The second property of Definition (3) holds obviously. For the third property let $v \in \mathcal{W}_p^k(\Omega)$,

$$\begin{aligned}
\|u + v\|_{\mathcal{W}_p^k(\Omega)} &= \sum_{j=1}^{n} \|u + v\|_{W_p^k(\Omega_j)} \\
&\leq \sum_{j=1}^{n} \left(\|u\|_{W_p^k(\Omega_j)} + \|v\|_{W_p^k(\Omega_j)} \right) \\
&\leq \sum_{j=1}^{n} \left(\|u\|_{W_p^k(\Omega_j)} \right) + \sum_{j=1}^{n} \left(\|v\|_{W_p^k(\Omega_j)} \right) \\
&= \|u\|_{\mathcal{W}_p^k(\Omega)} + \|v\|_{\mathcal{W}_p^k(\Omega)}
\end{aligned}$$

which concludes the proof. □

In particular, $\mathcal{L}^p(\Omega)$ denotes the special case of $\mathcal{W}_p^0(\Omega)$ by $\mathcal{L}^p(\Omega)$. Moreover, it could be shown that for any partition of Ω, the standard Sobolev space $W_p^k(\Omega)$ is contained in $\mathcal{W}_p^k(\Omega)$ (see [DMEM+20] and reference therein).

For the error analysis of MLS-VSDK, we require the concept of *Lipschitz boundary* conditions. See [AF03] for the exact definition.

Proposition 1. *Assume, that every set $\Omega_j \in \mathcal{P}$ satisfies Lipschitz boundary conditions. Let $\mathcal{P}$ be as in Definition 27 and set the derivative order $\delta = 0$. Then, by using Theorem* (46)*, the error satisfies the inequality*

$$\|f - s_{f,X}^{\psi}\|_{L^2(\Omega_j)} \leq C_j h_{\Omega_j}^{m+1-d(1/p-1/2)_+} \|f\|_{W_p^{m+1}(\Omega_j)}, \qquad \textit{for all} \quad \Omega_j \in \mathcal{P} \tag{4.24}$$

with h_{Ω_j} the fill distance with respect to Ω_j.

Proof. Recalling Definition 27 we know that the discontinuities of f and subsequently $w_i(\cdot)$ are located only at the boundary and not on the domain Ω_j, meaning that $w_i(\cdot)$ is continuous inside Ω_j. Furthermore, the basis $\{\alpha_i(\mathbf{x})\}_{1\leq i\leq n}$ forms a local polynomial reproduction, i.e., there exists a constant C such that $\sum_{i=1}^{N}|\alpha_i|\leq C$. Letting $s=1$ and $q=2$, by noticing that $W_q^0(\Omega_j)=L^q(\Omega_j)$, then the error bound (4.24) is an immediate consequence of Theorem (46). □

From the above proposition, it could be understood that $s_{f,X}^{\psi}$ behaves similarly to $s_{f,X}$ in the domain Ω_j, where there is no discontinuity. This agrees with Definition (27). Consequently, it is required to extend the error bound (4.24) to the whole domain Ω.

Theorem 49. *Let f, $\mathcal{P}$, $\psi_{\mathcal{P},\beta}$ be as before, and the weight functions as in* (4.19)*. Then, for $m>|\partial|+d/p$ (equality also holds for $p=1$), and $f\in\mathcal{W}_p^{m+1}(\Omega)$, for the MLS-VSDK approximant $s_{f,X}^{\psi}$ the error can be bounded as follows:*

$$\|f-s_{f,X}^{\psi}\|_{\mathcal{L}^2(\Omega)}\leq Ch^{m+1-d(1/p-1/2)_+}\|f\|_{\mathcal{W}_p^{m+1}(\Omega)} \tag{4.25}$$

Proof. By Proposition (1), we know that (4.24) holds for each Ω_j. Let h_{X,Ω_i} and C_i be the fill distance and a constant associated with each Ω_i, respectively. Then, we have

$$\sum_{j=1}^{n}\|f-s_{f,X}^{\psi}\|_{L^2(\Omega_j)}\leq\sum_{j=1}^{n}C_jh_{X,\Omega_i}^{m+1-d(1/p-1/2)_+}\|f\|_{W_p^{m+1}(\Omega_j)}.$$

By definition we get $\sum_{j=1}^{n}\|f-s_{f,X}^{\psi}\|_{L^2(\Omega_j)}=\|f-s_{f,X}^{\psi}\|_{\mathcal{L}^2(\Omega)}$. Moreover, letting $C=\max\{C_1,\ldots,C_n\}$ and $h=max\{h_{X,\Omega_1},\ldots,h_{X,\Omega_n}\}$ then the right hand side can be bounded by

$$Ch^{m+1-d(1/p-1/2)_+}\|f\|_{\mathcal{W}_p^{m+1}(\Omega)}.$$

Putting these together, we conclude. □

Some remarks are in order.

1. One might notice that the error bound in (4.17) is indeed local (the basis functions are local by assumption), meaning that if f is less smooth in a subregion of Ω, say it possesses only $m'\leq m$ continuous derivatives there, then the approximant (interpolant) has order $m'+1$ in that region and this is the best we can get. On the other hand, according to (4.25), thanks to the definition of piecewise Sobolev space, the regularity of the underlying function in the interior of the subdomain Ω_j matters. In other words, as long as f possesses regularity of order m in subregions, say Ω_j and Ω_{j+1}, the approximant order of $m+1$ is achievable, regardless of the discontinuities on the boundary of Ω_j and Ω_{j+1}.

2. Another interesting property of the MLS-VSDK scheme is that it is indeed data-dependent. To clarify, for the evaluation point $\mathbf{x}\in\Omega_j$ take two data sites $\mathbf{x}_i,\ \mathbf{x}_{i+1}\in B(\mathbf{x},r)$ with

the same distance from $\mathbf{x}$ such that $\mathbf{x}_i \in \Omega_j$ and $\mathbf{x}_{i+1} \in \Omega_{j+1}$. Due to the definition (4.18), $w_\psi(\mathbf{x}, \mathbf{x}_{i+1})$ decays to zero faster than $w_\psi(\mathbf{x}, \mathbf{x}_i)$ i.e., the data sites from the same subregion Ω_j pay more contribution to the approximant (interpolant) $s^{\psi}_{f,X}$, rather than the one from another subregion Ω_{j+1} beyond a discontinuity line. On the other hand, in the classical MLS scheme, this does not happen as the weight function gives the same value to both $\mathbf{x}_i$ and $\mathbf{x}_{i+1}$.

3. We highlight that in the MLS-VSDK scheme, we do not scale polynomials, so the polynomial space $\mathbb{P}^d_m$ is not changed. We scale only the weight functions, and thus, in case the given function values bear discontinuities, the basis functions $\{\alpha_i(\cdot)\}_{i=1}^n$ are modified.

We end this section by recalling that, to achieve the MLS approximation convergence order, it is necessary to operate in a stationary setting where the shape parameter ε is scaled relative to the fill distance. This leads to peaked basis functions for densely spaced data and flat basis functions for more widely spaced data. In other words, the weight functions' local support and the basis functions must be adjusted based on the $h_{X,\Omega}$ using the shape parameter ε. This requirement also applies to the MLS-VSDK scheme, which means that even after scaling w_i, we still need to consider the shape parameter's effect. This is different from the VS(D)Ks interpolation, where $\varepsilon = 1$ was kept fixed [DMMP20, BLRS15].

4.4 Numerical experiments

In this section, we compare the performance of the MLS-VSDK and the classical MLS methods. We consider the polynomials up to degree 1 in all numerical tests. Considering the evaluation points as $Z = \{z_1, \ldots, z_s\}$ we compute root mean square error as in (3.23), and maximum error by

$$MAE = \max_{z_i \in Z} |f(z_i) - s_{f,X}(z_i)|.$$

We consider four different weight functions to verify the convergence order of $s^{\psi}_{f,x}$ to a given f, as presented in Theorem 49.

1. $w^1(\mathbf{x}, \mathbf{x}_i) = (1 - \varepsilon\|\mathbf{x} - \mathbf{x}_i\|)^4_+ \cdot (4\varepsilon\|\mathbf{x} - \mathbf{x}_i\| + 1)$, which is the well-known C^2 *Wendland* function. Since each w^1_i is locally supported on the open ball $B(0, 1)$, then it verifies the conditions required by Theorem (49).

2. $w^2(\mathbf{x}, \mathbf{x}_i) = \exp(-\varepsilon\|\mathbf{x} - \mathbf{x}_i\|^2)$, i.e. the Gaussian RBF. We underline that when Gaussian weight functions are employed, with decreasing separation distance of the approximation centers, the calculation of the basis functions in (4.12) can be badly conditioned. Therefore, to make the computations stable, in this case, we regularize the system by adding a small multiple, say $\lambda = 10^{-8}$, of the identity to the diagonal matrix W.

3. $w^3(\mathbf{x}, \mathbf{x}_i) = \exp(-\varepsilon\|\mathbf{x} - \mathbf{x}_i\|)(15 + 15\|\mathbf{x} - \mathbf{x}_i\| + 6\|\mathbf{x} - \mathbf{x}_i\|^2 + \|\mathbf{x} - \mathbf{x}_i\|^3)$, that is a C^6 *Matérn* function.

4. $w^4(\mathbf{x}, \mathbf{x}_i) = (\exp(\varepsilon\|\mathbf{x} - \mathbf{x}_i\|)^2 - 1)^{-1}$, suggested in [Lev98], which enjoys an additional feature which leads to interpolatory MLS, since it possesses singularities at the centers.

One might notice that w^2, w^3, and w^4 are not locally supported. However, the key point is that they are all decreasing with the distance from the centers, and so, in practice, one can overlook the data sites that are so far from the center $\mathbf{x}$. As a result, one generally considers a *local stencil* containing n nearest data sites of the set Z of evaluation points. While there is no clear theoretical background concerning the stencil size, in MLS literature, one generally lets $n = 2 \times q$ (see, e.g., [Bay19]). However, it might be possible that in some special cases, one could achieve better accuracy using different stencil sizes. This aspect is covered by our numerical tests, which are outlined below.

1. In Section 4.4.1, we present an example in the one-dimensional framework, where the stencil size is fixed to be $n = 2 \times q$. Moreover, we consider w^1, w^2, and w^3.
2. In Section 4.4.2, we move to the two-dimensional framework, and we keep the same stencil size. Here, we restrict the test to the weight function w^1 and verify Theorem (49).
3. In Section 4.4.3, we remain in the two-dimensional setting, but the best accuracy is achieved with $n = 20$. Moreover, in addition to w^2 and w^3, we test the interpolatory case by considering w^4 as a weight function.
4. In Section 4.4.4, we present two-dimensional experiments where the data sites have been perturbed via some white noise. We fix $n = 25$ and w^2, w^3 are involved.

4.4.1 Example 1

On $\Omega = (-1, 1)$, we assess MLS approximant for

$$f_1(x) = \begin{cases} e^{-x}, & -1 < x < -0.5 \\ x^3, & -0.5 \leq x < 0.5, \\ 1, & 0.5 \leq x < 1 \end{cases}$$

with discontinuous scale function

$$\psi(x) = \begin{cases} 1, & x \in (-1, 0.5) \text{ and } [0.5, 1) \\ 2, & x \in [-0.5, 0.5)\,. \end{cases}$$

We note that the function ψ is defined only by two cases. The important fact is that it has a jump as f_1. To evaluate the approximant, consider the evaluation grid of equispaced points with a step size of $5.0e-4$. Tables 4.1 and 4.2 include RMSE of f_1 approximation using w^1 as the weight function.

Again, in order to investigate the convergence rate, consider two sets of uniform and Halton nodes with the size from Table 4.1. To generalize our results to globally supported weight

number of centers	ε value	RMSE MLS-VSDK	RMSE classic MLS
9	0.25	3.58e-1	3.95e-1
17	0.5	1.99e-1	3.02e-1
33	1	3.10e-3	2.17e-1
65	2	8.42e-4	1.54e-1
257	4	5.67e-5	7.68e-2
513	8	1.43e-5	5.35e-2

Table 4.1: Comparison of the RMSE for f_1 approximation at *uniform* data sites.

number of centers	ε value	RMSE MLS-VSDK	RMSE classic MLS
9	0.25	3.53e-1	3.77e-1
17	0.5	1.99e-1	3.01e-1
33	1	3.08e-3	2.17e-1
65	2	8.39e-4	1.54e-1
257	4	5.67e-5	7.73e-2
513	8	1.43e-5	5.41e-2

Table 4.2: Comparison of the RMSE for f_1 approximation at *Halton* data sites.

functions, we consider w^2 and w^3, Gaussian and Matérn C^6 radial functions, respectively. For the uniform data sites, let the shape parameter values be $\varepsilon_{GA}^{U} = [5, 20, 40, 80, 160, 320]$ and $\varepsilon_{Mat}^{U} = [5, 10, 20, 40, 80, 160]$ for w^2 and w^3. Our computation shows convergence rates of 2.54 and 2.26 for the MLS-VSDK scheme, shown in Figure 4.2. Accordingly, for *Halton* points, let $\varepsilon_{Mat}^{H} = [5, 10, 20, 50, 200, 400]$, $\varepsilon_{GA}^{H} = [10, 20, 30, 50, 100, 200]$. The corresponding convergence rates are 2.38 and 2.33. On the other hand, in both cases, the standard

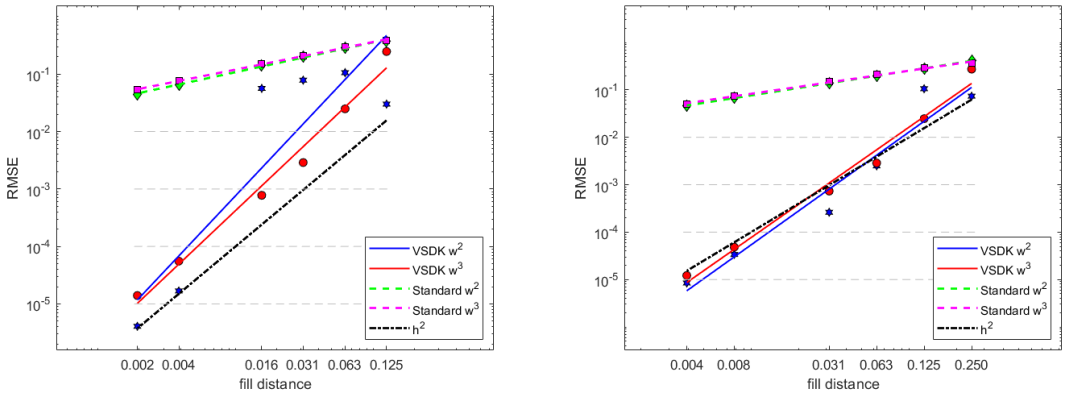

Figure 4.2: Convergence rates for approximating f_1 with MLS-VSDK and MLS-Standard schemes using *uniform* data sites (left) and *Halton* data sites (right).

MLS scheme can hardly reach an approximation order of 1 using non-scaled weight functions.

4.4.2 Example 2

Consider on $\Omega = (-1,1)^2$ the discontinuous function

$$f_2(x,y) = \begin{cases} \exp(-(x^2+y^2)), & x^2+y^2 \leq 0.6 \\ x+y, & x^2+y^2 > 0.6 \end{cases}$$

and the discontinuous scale function

$$\psi(x,y) = \begin{cases} 1, & x^2+y^2 \leq 0.6 \\ 2, & x^2+y^2 > 0.6 \end{cases}$$

We take the grid of equispaced points with mesh size $1.00e-2$ for the evaluation points. Figure 4.3 shows both the *RMSE* and *absolute error* for the classical *MLS* and *MLS-VSDK* approximation of f_2 sampled from $1089 = 33^2$ uniform data sites taking w^1 as the weight function. Figure 4.3 shows that using classical MLS, the approximation error significantly increases near the discontinuities, while using the MLS-VSDK scheme, the approximant can overcome this issue.

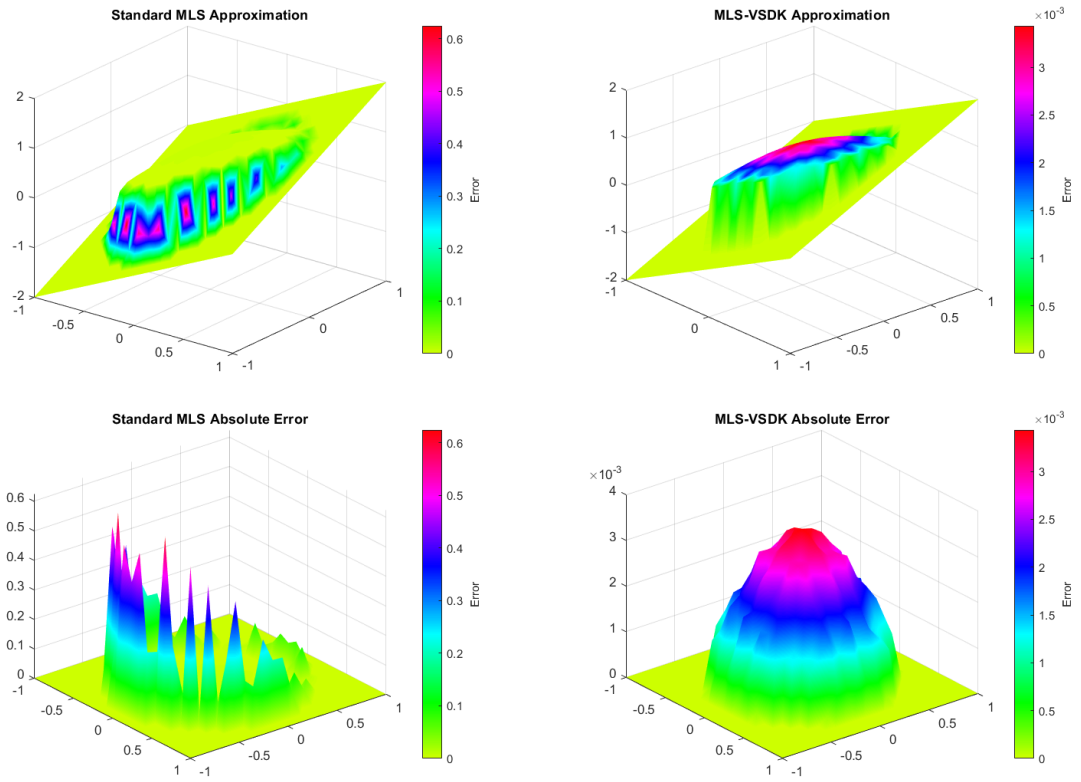

Figure 4.3: RMSE and abs-error of f_2 the MLS (left) and the MLS-VSDK (right) approximation schemes using w^1 weight function

We consider increasing sets of $\{25, 81, 289, 1089, 4225, 16641\}$ Halton and uniform points

as the data sites to investigate the convergence rate. To find an appropriate value for the shape parameter, we fix an initial value and multiply it by a factor of 2 at each step. Thus, let $\varepsilon = [0.25, 0.5, 1, 2, 4, 8]$ be the vector of the shape parameter, which is modified concerning the number of the centers in both cases of uniform and Halton data sites. The left plot of Figure 4.4 shows a convergence rate of 2.58 for the MLS-VSDK and only 0.66 for classical MLS methods, while these values are 2.04 and 0.70 in the right plot.

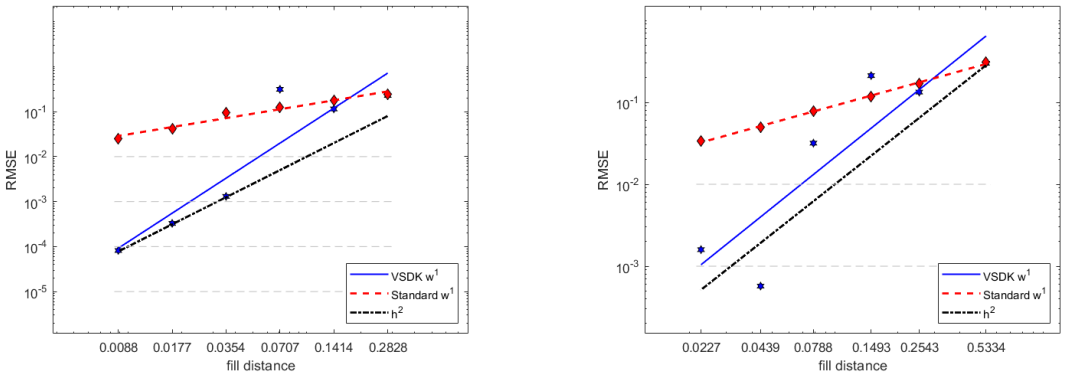

Figure 4.4: Convergence rates for approximation of function f_2 with MLS-VSDK and MLS standard schemes using *Uniform* data sites (left) and *Halton* data sites (right).

4.4.3 Example 3

Consider the following function

$$f_3(x,y) = \begin{cases} 2\big(1-\exp(-(y+0.5)^2)\big), & |x| \leq 0.5, \ , |y| \leq 0.5. \\ 4(x+0.8), & -0.8 \leq x \leq -0.65, |y| \leq 0.8. \\ 0.5, & 0.65 \leq x \leq 0.8, |y| \leq 0.2 \\ 0, & \text{otherwise.} \end{cases}$$

defined on $\Omega = (-1,1)^2$. Regarding the discontinuities of f_3, the scale function is considered to be

$$\psi(x,y) = \begin{cases} 1, & |x| \leq 0.5, \ , |y| \leq 0.5. \\ 2, & -0.8 \leq x \leq -0.65, |y| \leq 0.8. \\ 3, & 0.65 \leq x \leq 0.8, |y| \leq 0.2 \\ 0, & \text{otherwise.} \end{cases}$$

Moreover, let the centers and evaluation points be the same as the Example 4.4.1. Table 4.3 and 4.4 shows RMSE of MLS-VSDK and conventional MLS approximation of f_3 using w^4 which interpolates the data. We underline that our experiments show that the stencil of size $n = 20$ leads to the best accuracy. Figure 4.5 shows **RMSE** and **Absolute Error** for *standard MLS* and *MLS-VSDK approximation* of f_3 sampled from 1089 uniform points using w_4 as weight

number of centers	ε value	RMSE MLS-VSDK	RMSE classic MLS
25	1	3.67e-1	1.47e+0
81	2	3.68e-1	8.86e-1
289	4	1.49e-2	7.44e-1
1089	8	4.23e-3	7.72e-1
4225	16	1.06e-3	6.64e-1
16641	32	2.65e-4	5.25e-1

Table 4.3: RMSE of f_3 interpolation with *uniform* data sites.

number of centers	ε value	RMSE MLS-VSDK	RMSE classic MLS
25	1	8.84e-1	1.53e+0
81	2	8.95e-2	1.05e+0
289	4	1.42e-2	8.74e-1
1089	8	4.18e-3	6.48e-1
4225	16	1.09e-3	6.68e-1
16641	32	3.02e-4	7.07e-1

Table 4.4: RMSE of f_3 interpolation with *Halton* data sites.

function. Once again, Figure 4.5 shows how the MLS-VSDK scheme can improve the accuracy by reducing the error near the jumps.

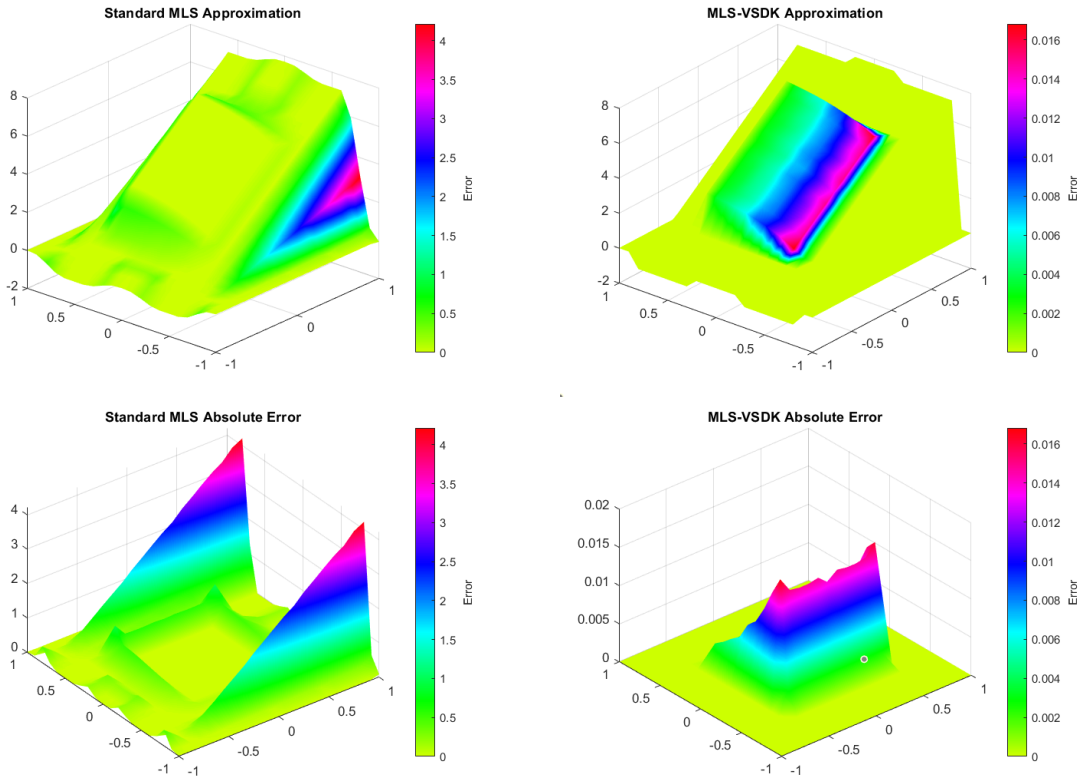

Figure 4.5: RMSE and abs-error of f_3 MLS(left) and MLS-VSDK(right) approximation (interpolation) schemes using w^4 weight function

Eventually, letting $\varepsilon^U_{GA} = [2, 4, 8, 16, 32, 64]$ and $\varepsilon^U_{Mat} = [10, 20, 40, 80, 160, 320]$, Figure 4.6 shows that h^2 convergence is achievable. To be more precise, the rate of convergence in the left plot is 2.54 and 2.69 for w_2 and w_3, respectively. On the other hand, letting $\varepsilon^H_{GA} = [1, 2, 4, 8, 16, 32]$ and ε^H_{Mat} as the Uniform case, convergence rates of 2.50 and 2.73 is achievable when *Halton* data sites are employed.

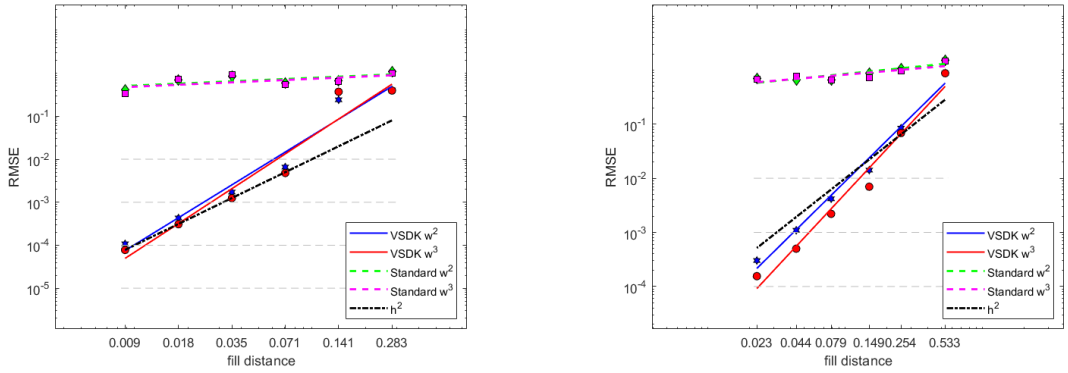

Figure 4.6: Convergence rates for approximation of function f_3 with MLS-VSDK and MLS standard schemes using *Uniform* data sites (left) and *Halton* data sites (right).

4.4.4 Example 4

In applications, the discontinuities are likely to be unknown. To overcome this problem, one can consider the edge detector method to extract the discontinuities. However, in this way, the approximation depends also on the performance of the edge detector method as well [DMEM+20]. Toward this direction, in our last experiment, we assume the location of the discontinuities is not precisely known. This is modeled by adding some noise drawn from the standard normal distribution multiplied by 0.01 to the edges of $\Omega_i \in \mathcal{P}$. We take the test function f_2 and the data sites in Section 4.4.2. We fix $n = 25$, and $\varepsilon_{GA} = [0.25, 0.5, 1, 2, 4, 8]$, $\varepsilon_{Mat} = [1, 2, 4, 816, 32]$ for both *Halton* and *uniform* centers. Figure 4.7 shows that the suggested MLS-VSDK can still obtain a reasonable convergence rate compared to the classical MLS, even when the discontinuities are not known precisely.

4.5 Conclusions

To approximate a discontinuous function using scattered data values, we studied a new technique based on discontinuously scaled weight functions, which we called the MLS-VSDK scheme, which is the application of discontinuous scaled weight functions to the MLS. It enabled us to move toward a data-dependent scheme, meaning that MLS-VSDK can encode the behavior of the underlying function. We obtained a theoretical Sobolev-type error estimate, which justifies why MLS-VSDK can outperform conventional MLS. The numerical experiments con-

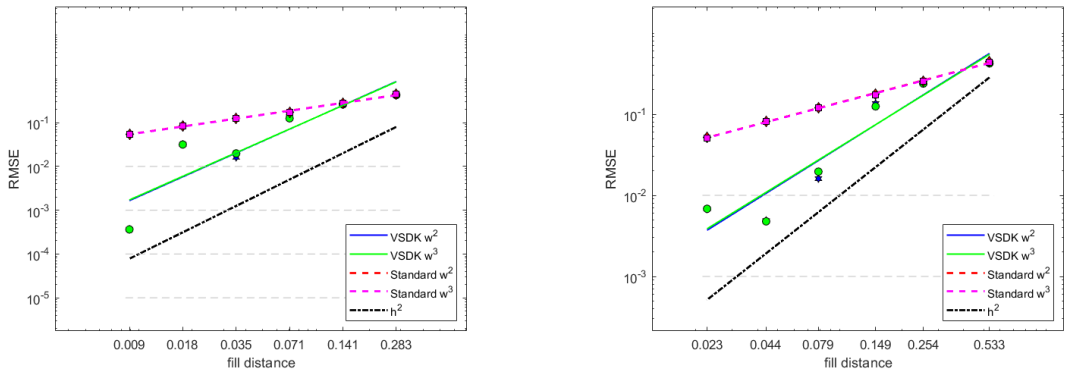

Figure 4.7: Convergence rates for approximation of function f_2, based on noisy given data values, with MLS-VSDK and MLS standard schemes using *Uniform* data sites (left) and *Halton* data sites (right).

firmed the theoretical convergence rates. Besides, our numerical tests showed that the suggested scheme can reach high accuracy even if the position of the data values is slightly perturbed.

5

Direct RBF Partition of Unity scheme for solving time-dependent PDEs on surfaces using closest point surface representation

> The more I think about language, the more it amazes me that people ever understand each other at all.
>
> Kurt Gödel

5.1 Introduction

Many applications in the natural and applied sciences require the solutions of partial differential equations (PDEs) on surfaces or more general manifolds. Because analytical solutions are rarely possible, one must approximate the solution using numerical schemes. A great effort has been made to develop numerical methods for many important classes of PDEs. Most existing surface PDE solvers can be classified into two types. The first, referred to *intrinsic methods*, solves the PDE directly on the manifold using either a mesh [Hol01, DGJ03], a parametrization of the manifold [FH05], or an explicit reconstruction of the manifold [LZ13]. Another category of methods, referred to as *embedding methods*, involves the extension of the surface PDE to a PDE defined on a small region surrounding the surface. This new PDE is formulated to match the original problem on the surface; see, e.g., [BCOS01, CT18].

The *closest point representation* of the surface is a method that practices the latter idea (see, e.g., [MM12] about the theoretical foundation of closest point methods). To be more precise, the Closest Point (CP) methods decouple surface geometry and PDE evolution via a closest point extension step, yielding a method that involves standard Cartesian grid methods in the embedding space. In the classical formulation of the CP method [RM08, MR08], the discretization

is carried out in a neighborhood of the surface using standard finite difference (FD) schemes and barycentric Lagrangian interpolation. Followingly, in [MR10], the authors proposed the *implicit* closest point method to provide a stable approximation of the Laplace–Beltrami and other higher-order surface differential operators. Besides, it is well-known the issue arising from the multivariate polynomial *Finite Difference* (FD) scheme can be bypassed by using the Radial Basis Function to generate stencil weights in the FD formulas [FLB+12, FF15b, SK19]. As a result, further research includes an explicit CP formulation using finite difference schemes derived from RBFs [PLR18] and a least-squares implicit formulation of the CP method in [PLPR19].

Another approach for localization and evolving RBF into a fast technique applicable to large-scale problems is the RBF partition of unity (RBF-PU) method. The first combination of a PU method with RBF interpolation goes back to [Wen04, Wen02]. For more about RBF-PU, we refer the readers to [DMMP19, DMMPR19] and the reference inside. In a recent paper [Mir21], a new method, *Direct RBF Partition Unity* (D-RBF-PU), was introduced to simplify the implementation and increase the efficiency of the standard RBF-PU method. This method directly approximates PDE operators, bypassing the need to differentiate against PU weight functions and enabling some useful discontinuous weight functions. Although related to the RBF-FD method, this method is much faster at setting up stiffness and mass matrices. This approach is particularly efficient for studying PDEs, where computing and implementing surface derivatives of PU weight functions is a challenging problem.

This chapter suggests a fresh numerical approach to tackling partial differential equations intrinsic to smooth orientable surfaces. The method employed is the direct radial basis function partition of unity (D-RBF-PU), which incorporates the Closest Point function. Our preliminary results show that it can overcome the existing RBF-FD CP method. In section (5.2), we review basic notions regarding the surface intrinsic differential geometry and cosest point functions. Section (5.3) is dedicated to two explicit CP methods existing in the literature. Afterward, in section (5.4), we suggest a new numerical scheme based on solving the time-dependent PDE problems intrinsic to an arbitrary surface, and in section (5.5) includes some preliminary numerical results.

5.2 Preliminaries

This section reviews the concepts of surface intrinsic differential operators, closest point surface representation and functions extension, and several closest point principles.

5.2.1 Surface intrinsic differential operators

Suppose there is a smooth surface $\mathcal{S}$ embedded in $\mathbb{R}^3$ and a scalar function $u(\mathbf{y})$ defined on it. Let $\tilde{u}(\mathbf{x})$ be an extension of $u(\mathbf{y})$ defined in $\mathbb{R}^3$, such that $\tilde{u}(\mathbf{y}) = u(\mathbf{y})$ for any $\mathbf{y} \in \mathcal{S}$. In this

case, the surface gradient of u at $\mathbf{y} \in \mathcal{S}$, is defined as follows:

$$\nabla_{\mathcal{S}} u(\mathbf{y}) = \nabla_{\mathcal{S}} \tilde{u}(\mathbf{y}) = \nabla \tilde{u}(\mathbf{y}) - \big(\mathbf{n}(\mathbf{y}).\nabla \tilde{u}(\mathbf{y})\big)\mathbf{n}(\mathbf{y}), \tag{5.1}$$

with $\mathbf{n}(\mathbf{y})$ being the unit normal vector to $\mathcal{S}$ at $\mathbf{y}$. Accordingly, let $\mathbf{P}(\mathbf{y}) = \mathbf{I} - \mathbf{n}(\mathbf{y})\mathbf{n}(\mathbf{y})^T$ be the projection matrix at $\mathbf{y} \in \mathcal{S}$ where $\mathbf{I}$ is the identity matrix. Consequently, the surface gradient $\nabla_{\mathcal{S}} u(\mathbf{y})$ in (5.1) can be rewritten as

$$\nabla_{\mathcal{S}} u(\mathbf{y}) = \nabla_{\mathcal{S}} \tilde{u}(\mathbf{y}) = P(\mathbf{y}) \nabla \tilde{u}(\mathbf{y}).$$

We can also denote the surface gradient in components by

$$\nabla_{\mathcal{S}} u(\mathbf{y}) = \big(\mathcal{D}_1 u(\mathbf{y}), \mathcal{D}_2 u(\mathbf{y}), \mathcal{D}_3 u(\mathbf{y})\big), \quad \mathbf{y} \in \mathcal{S},$$

where $\mathcal{D}_i, \quad i = 1, 2, 3$ are the components of the surface gradient. Following such an idea, the surface Laplacian, or the *Laplace-Beltrami* function of a scalar function $u(\mathbf{y})$ with extension $\tilde{u}(\mathbf{y})$ is defined as

$$\Delta_{\mathcal{S}} u(\mathbf{y}) = \Delta_{\mathcal{S}} \tilde{u}(\mathbf{y}) = \nabla_{\mathcal{S}} . \nabla_{\mathcal{S}} \tilde{u}(\mathbf{y}) = \sum_{i=1}^{3} \mathcal{D}_i \mathcal{D}_i \tilde{u}(\mathbf{y}) \quad \mathbf{y} \in \mathcal{S}.$$

After some manipulations (see [Che15]) we get

$$\Delta_{\mathcal{S}} \tilde{u}(\mathbf{y}) = \Delta \tilde{u}(\mathbf{y}) - \kappa(\mathbf{y}) \frac{\partial \tilde{u}}{\partial \mathbf{n}}(\mathbf{y}) - \frac{\partial^2 \tilde{u}}{\partial \mathbf{n}^2}(\mathbf{y}) \quad \mathbf{y} \in \mathcal{S},$$

where $\kappa(\mathbf{y}) = \nabla . \mathbf{n}(\mathbf{y})$ is the mean curvature at $\mathbf{y}$, $\frac{\partial \tilde{u}}{\partial \mathbf{n}}(\mathbf{y}) = \nabla \tilde{u}(\mathbf{y}).\mathbf{n}(\mathbf{y})$, and $\frac{\partial^2 \tilde{u}}{\partial \mathbf{n}^2}(\mathbf{y}) = \nabla\big(\nabla \tilde{u}(\mathbf{y}).\mathbf{n}(\mathbf{y})\big).\mathbf{n}(\mathbf{y})$.
Such an idea can be generalized to intrinsic differential operators on manifolds of any dimension and co-dimension; see, e.g., [MM12].

5.2.2 Closest Point Surface Representation

When dealing with numerical methods for solving partial differential equations on surfaces, it is crucial to have an accurate representation of the surface. One common approach for representing a surface involves using two local parameters (or one in the case of a curve). For instance, a circle with radius R can be represented using the arc length s as a parameter, resulting in the following parametrization: $\mathcal{S}(s) = R\left(\cos(\frac{s}{R}), \sin(\frac{s}{R})\right)$. However, obtaining an appropriate parametrization is significantly more challenging for more complex surfaces.

An alternative approach is to embed $\mathcal{S}$ in a higher dimensional space $\mathbb{R}^d$. For example, a curve can be embedded in $\mathbb{R}^2$ or $\mathbb{R}^3$, while a surface can be embedded in $\mathbb{R}^3$ or higher. This technique allows us to solve numerical problems, such as PDEs, on the surface by computing

the points belonging to the embedding space $\mathbb{R}^d$, rather than the surface itself. To accomplish this, we need a representation of the surface in $\mathbb{R}^d$, which leads us to the closest point representation of the surface, defined as follows:

Definition 28. *For a given surface $\mathcal{S}$, the closest point function $cp : \mathbb{R}^d \to \mathbb{R}^d$ takes point $\mathbf{x} \in \mathbb{R}^d$ and returns a point $cp(\mathbf{x}) \in \mathcal{S} \subset \mathbb{R}^d$ which is closest in Euclidean distance to $\mathbf{x}$. In other words, $cp(\mathbf{x}) = \min_{\mathbf{q}\in\mathcal{S}} \|\mathbf{q} - \mathbf{x}\|_2$.*

If the surface $\mathcal{S}$ is smooth, there exists a neighborhood $B(\mathcal{S})$ of $\mathcal{S}$ such that $\forall \mathbf{x} \in B(\mathcal{S})$, $cp(\mathbf{x})$ is unique; see e.g., [Hir12, MM12]. Later, in our numerical scheme, we refer to this neighborhood as the *computational tube*. For points that are not in $B(\mathcal{S})$, there might be multiple points on the surface $\mathcal{S}$ that is equidistant from $\mathbf{x}$; in this way, the function $cp(\mathbf{x})$ is defined to return an arbitrarily chosen closest point. For further details, please refer to [MR10].

The closest point representation of a surface $\mathcal{S}$ refers to the knowledge of $cp(\mathbf{x})$ value for all $\mathbf{x} \in \mathbb{R}^d$ or at least on those points utilized in a computation. So, this representation is implicit, as the surface $\mathcal{S}$ is known only through the closest point function. In particular, unlike the level set representation, the CP representation does not require any knowledge of the interior or exterior of the surface, making it easier to represent surfaces with boundaries or non-orientable surfaces. As a result, we see that this representation is not limited by the geometry or dimension of the surfaces, making it a highly advantageous feature of the closest point representation.

Depending on the grid points, different techniques can be applied to obtain a CP representation for computation. Analytical formulae are an option for simple geometries such as spheres or tours. Another approach is to minimize the square distance function by *Newton's method* for parametrized surfaces [CM15, MBR11], or compute the closest points from other representations of the surface such as a triangulation [MR10], a point cloud [LLZ11], or a level set function [CM15]. For more regarding the literature, refer to [MR08, MR10] and the reference therein.

The closest point representation of a surface leads to a natural extension of surface data, namely *the closest point extension*. As we shall see later, the method extends the surface function to a narrow band $B(\mathcal{S})$ so that the extended function is constant along the normals to the surface. Therefore, considering the discussion in the subsection (5.2), we can replace the surface intrinsic differential operators with the corresponding Cartesian differential operators to formulate an embedding equation in the narrow band in $\mathbb{R}^d$.

Definition 29. *Let $\mathcal{S}$ be a smooth surface embedded in $\mathbb{R}^d$, and let $B(\mathcal{S}) \subset \mathbb{R}^d$ be a tubular neighborhood of $\mathcal{S}$ in which the closest point function is uniquely defined. Then the closest point extension of a scalar function $u : \mathcal{S} \to \mathbb{R}$ is function $\tilde{u} : B(\mathcal{S}) \to \mathbb{R}$ so that $\tilde{u}(\mathbf{x}) = u(cp(\mathbf{x}))$ for $\mathbf{x} \in B(\mathcal{S})$. The closest point extension of a vector field $v : \mathcal{S} \to \mathbb{R}^d$ is $\tilde{v} : B(\mathcal{S}) \to \mathbb{R}^d$ so that $\tilde{v}_i(\mathbf{x}) = v_i(cp(\mathbf{x}))$ for $\mathbf{x} \in B(\mathcal{S})$ and $i = 1, \ldots, d$.*

The closest point extensions $\tilde{u}$ and $\tilde{v}$ defined in Definition (29) are constant in the normal

directions to the surface, at least within a neighborhood of a smooth surface. This leads to simplified derivative calculations in the embedding space, [MM12, RM08].

Gradient principle : For a scalar function u defined on $\mathcal{S} \subset \mathbb{R}^d$, let $\tilde{u}(\mathbf{x}) = u(cp(\mathbf{x}))$ defined in $B(\mathcal{S}) \subset \mathbb{R}^d$ be the closest point extension of u, then $\nabla_{\mathcal{S}} u(\mathbf{y}) = \nabla \tilde{u}(\mathbf{y})$ for $\mathbf{y} \in \mathcal{S}$.

Divergence principle : For a vector field v defined on $\mathcal{S} \subset \mathbb{R}^d$, let $\tilde{v}(\mathbf{x}) = v(cp(\mathbf{x}))$ defined in $B(\mathcal{S}) \subset \mathbb{R}^d$ be the closest point extension of v, then $\nabla_{\mathcal{S}}.v(\mathbf{y}) = \nabla.\tilde{v}(\mathbf{y})$ for $\mathbf{y} \in \mathcal{S}$.

Laplacian principle : For a scalar function u defined on $\mathcal{S} \subset \mathbb{R}^d$, let $\tilde{u}(\mathbf{x}) = u(cp(\mathbf{x}))$ defined in $B(\mathcal{S}) \subset \mathbb{R}^d$ be the closest point extension of u, $\Delta_{\mathcal{S}} u(\mathbf{y}) = \Delta \tilde{u}(\mathbf{y})$ for $\mathbf{y} \in \mathcal{S}$.

As mentioned before, for a scalar function u, since its closest point extension $\tilde{u}$ is constant along the normals to $\mathcal{S}$, we have $(\mathbf{n}.\nabla \tilde{u}) = 0$. The extended proof for the general manifolds with arbitrary dimension and co-dimension could be found in [MM12].

Remark 17. *We highlight that $\tilde{u}$ can be defined in the whole embedding space $\mathbb{R}^d$. Still, for theoretical and practical reasons, it is preferred to do so in a narrow band $B(\mathcal{S})$ surrounding the surface.*

5.3 Closest Point Numerical Scheme

Let $\mathcal{S}$ be a surface and a PDE defining a flow on the surface over time in terms of *intrinsic* in-surface differential operators. For example, the diffusion equation on a surface $\mathcal{S}$,

$$\begin{aligned} u_t(t, \mathbf{y}) &= \Delta_{\mathcal{S}} u(t, \mathbf{y}) \\ u(0, \mathbf{y}) &= u_0(\mathbf{y}) \end{aligned} \tag{5.2}$$

where $\mathbf{y} \in \mathcal{S}$ and $t \geq 0$.

This PDE is time-dependent, and we wish to propagate its solution on the surface over time. Using the Laplace principle, we extend the surface PDE to the following embedding equation,

$$\begin{aligned} \tilde{u}_t(t, \mathbf{x}) &= \Delta \tilde{u}(t, cp(\mathbf{x})) \\ \tilde{u}(0, \mathbf{x}) &= u_0(cp(\mathbf{x})), \end{aligned} \tag{5.3}$$

where $\mathbf{x} \in B(\mathcal{S})$ and $t \geq 0$ and $\tilde{u}$ is the CP extension of u.

Here, we review two explicit CP methods to deal with (5.2) investigated in [RM08] and [PLR18] respectively. We refer the readers to [MR10, vGMM13] for other CP implicit methods.

5.3.1 Explicit *Ruuth-Merriman* CP Approach

Considering our discussion in section (5.2), we know that the solution of the (5.2) and (5.3) agree at the surface. In more detail, if $u(t, \mathbf{y})$ is the solution (5.2) for $\mathbf{y} \in \mathcal{S}$ and $\tilde{u}(t, \mathbf{x})$ is a solution of (5.3) for $\mathbf{x} \in B(\mathcal{S})$, then $u(t, \mathbf{y}) = \tilde{u}(t, \mathbf{y})$ for $t \geq 0$ and points on the surface $\mathbf{y} \in \mathcal{S}$. Therefore, it appears that we only need to solve the embedding equation (5.3).

Starting from initial conditions $\tilde{u}(0, \mathbf{x}) = u_0(cp(\mathbf{x}))$, which corresponds to the closest point extension of the surface data u_0, we have

$$\Delta\tilde{u}(0, cp(\mathbf{x})) = \Delta\tilde{u}(0, \mathbf{x}), \tag{5.4}$$

which means that at $t = 0$, we can evolve equation (5.3) simply by

$$\tilde{u}_t(t, \mathbf{x}) = \Delta\tilde{u}(t, \mathbf{x}). \tag{5.5}$$

However, after $t > 0$, as $\tilde{u}(t, \mathbf{x})$ is evolved by (5.5), it no longer equals $\tilde{u}(t, cp(\mathbf{x}))$; so we need to compute the closest point extension $\tilde{u}(t, \mathbf{x}) := \tilde{u}(t, cp(\mathbf{x}))$ again if we still want to evolve equation (5.3) simply by (5.5). We remark that here, the closest point extension is an interpolation step, and the order of the interpolation should be sufficiently high so that interpolation errors do not dominate the solution.

We must still discretize in the spatial direction for a complete numerical scheme. Discretizing the above idea along the time direction by the *forward Euler* scheme leads to the *explicit Ruuth-Merriman* approach. Thus, starting from the closest point extension of the initial data, we perform the following two steps at each time step.

1. **(Evolving step)** Perform a forward Euler time step with step size Δt in the embedding space:
$$w^{n+1} = \tilde{u}^n(\mathbf{x}) + \Delta t.\Delta\tilde{u}^n(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d; \tag{5.6}$$
2. **(Extension step)** Perform a closest point extension for each point in the embedding space:
$$\tilde{u}^{n+1}(\mathbf{x}) = w^{n+1}(cp(\mathbf{x})), \quad \mathbf{x} \in \mathbb{R}^d. \tag{5.7}$$

In [RM08], both of these steps are done using barycentric Lagrange interpolation with polynomial up to degree $p = q + r - 1$, where q is the order of finite differences schemes, and r is the order of the derivatives. Another approach is taken in [PLR18] where the *RBF-generated Finite Difference* (RBF-FD) scheme was applied to perform these two steps. To make this work self-contained, we briefly review this method as well.

5.3.2 Explicit RBF-CPM Method

The general idea of the RBF-FD method was first developed in [TS03, WF06]. Suppose we are provided with data of the form $\mathcal{Z} = \{\mathbf{z}_i\}_{i=1}^N \subset \Omega$ and corresponding function values

$\{f(\mathbf{z}_i)\}_{i=1}^N$, with f being the target function defined on $\Omega \subset \mathbb{R}^d$. We wish to find the value of $(\mathcal{L}f)(\xi)$, where $\mathcal{L}$ is the linear differential operator defined on some normed linear space and $\xi \in \Omega$ being the evaluation point. For local approximation of $(\mathcal{L}f)(\xi)$ with the RBF-FD method, let $\mathcal{Z}_\xi = \{\mathbf{z}_j\}_{j\in J_\xi} \subset \mathcal{Z}$ denote the set of n nearest data sites of ξ where J_ξ is the set of indices such that

$$J_\xi = \{j \in \{1, ..., N\};\ \mathbf{z}_j \in \mathcal{Z}_\xi\}.$$

Analogous to (4.1), we intend to find $(\mathcal{L}f)(\xi)$ by calculating a linear combination of the function values belonging to the stencil associated with ξ, i.e.,

$$\mathcal{L}f(\xi) = \sum_{j\in J_\xi} w_j f(\mathbf{z}_j), \tag{5.8}$$

where the weights w_j are known as the *differentiation weights*. For the sake of simplicity, following (4.1) instead of $\mathcal{L}s_{f,X}(\xi)$ we use the notation $\mathcal{L}f(\xi)$. Obtaining the differentiation weights $\mathbf{w}^T = [w_1, \ldots, w_{|\mathcal{Z}_\xi|}]$ is done similarly to the generating functions in (4.1) so,

$$\mathbf{w}^T = \mathcal{L}\Phi^T_{\mathcal{Z}_\xi}(\xi) V^{-1}_{\Phi,\mathcal{Z}_\xi}, \tag{5.9}$$

with $\Phi_{\mathcal{Z}_\xi}(\xi)$ defined analogously to (2.16) i.e., the basis Φ is translated over $\mathcal{Z}_\xi$ and is evaluated at ξ, but now the operator $\mathcal{L}$ is applied to each basis function Φ. Given a set of evaluation points $\Xi = \{\xi_1, ..., \xi_M\}$ the final linear system would be

$$W\mathbf{f}_\mathcal{Z} \approx \mathbf{f}^\mathcal{L}_\Xi \tag{5.10}$$

where $\mathbf{f}_\mathcal{Z}$ and $\mathbf{f}^\mathcal{L}_\Xi$ are the column vectors of the function values and the approximated solution respectively, $W \in \mathbb{R}^{M\times N}$ is the *differentiation matrix* such that its rows contain the differentiation weights from (5.9) associated to each $\{\xi_k\}_{k=1}^M$.
Notice that the RBF-FD is one of the ways to generate the differentiation weights. Another approach to do so is the MLS scheme, which is investigated in [MSD12, MS13].

Remark 18. *Regarding the linear system* (5.10) *some remarks are in order:*

- *Recalling* (4.1)*, the differential operator is not applied directly to the shape functions. It means that one does not first form the interpolant $s_{f,X}$ and then apply the linear operator $\mathcal{L}$. Such an idea is related to the direct discretization explained in [Sch13].*

- *Since the stencil weights are computed locally, they need to be mapped into the appropriate columns of the differentiation matrix W.*

- *Depending on the value of shape parameter ε, the interpolation matrix $V_{\Phi,\mathcal{Z}_\xi}$ could be highly ill-conditioned. So, the RBF-FD scheme could be very inaccurate. See, e.g., [FFBB16].*

- *Given a PDE problem, the right-hand-side of* (5.10) *is known, and* $\mathbf{f}_{\mathcal{Z}}$ *is desired, so one needs to compute the inverse of the differentiation matrix* W*, requiring* W *to be sparse and well-conditioned.*
- *Last but not least, one can let* $\mathcal{Z}$ *be the same as* Ξ*, since on one side, we have the function values and on the other, the operator* $\mathcal{L}$ *applied to the function values.*

Having introduced the RBFD-FD method, we now review the RBF-FD-CP method investigated in[PLR18]. To keep our notation in harmony with the explained RBF-FD methodology, we use $\mathbf{z}$ and ξ instead of $\mathbf{x}$ and $\mathbf{y}$ in (5.2) and (5.3).
Let $\mathcal{Z} = \{\mathbf{z}_j\}_{j=1}^N \subset \Omega$ be a collection of Cartesian grid points on a small tubular domain $B(\mathcal{S})$ containing the surface $\mathcal{S}$. Then, we define surface data points via $\xi_j = cp(\mathbf{z}_j)$ to form a set $\Xi = \{\xi_j\}_{j=1}^N \subset \mathcal{S}$; see Figure (5.1) for a schematic demonstration.

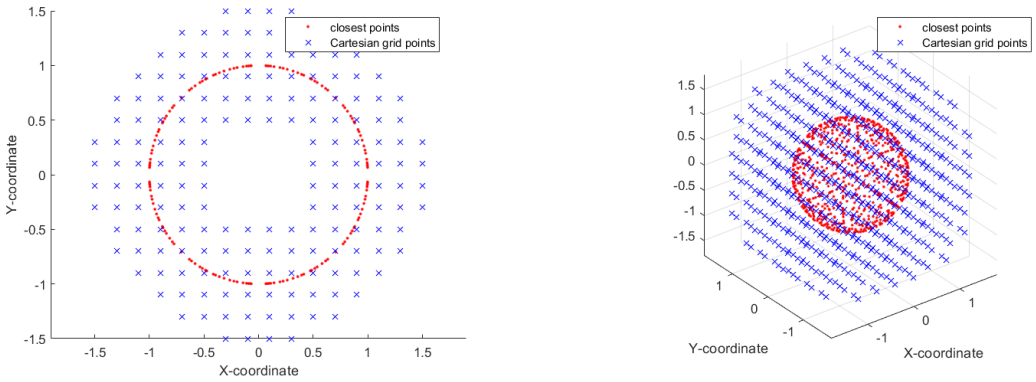

Figure 5.1: Cartesian grid and its closest point on the surface; unit circle (left) and unit sphere (right).

Applying the equivalence of the Laplacian property yields the relation

$$\Delta_{\mathcal{S}} u(\Xi) = \Delta \tilde{u}(\Xi) \tag{5.11}$$

where $\tilde{u}$ is the constant-along-normal extension of u. Now, the RBF-FD method gives the approximation as follows:

$$\Delta \tilde{u}(\xi_j) \approx \mathbf{w}^T \tilde{u}(\mathcal{Z}_{\xi_j}).$$

The differentiation matrix can be assembled analogous (5.10) i.e.,

$$\Delta_{\mathcal{S}} u(\Xi) = \Delta \tilde{u}(\Xi) \approx W \tilde{u}(\mathcal{Z}). \tag{5.12}$$

We are now ready to discretize the initial equation (5.2), intrinsic to the surface. Using the *forward Euler scheme* with spatial discretization at $\Xi \subset \mathcal{S}$ as before, we have

$$u(\Xi, t^{n+1}) = u(\Xi, t^n) + \Delta t \Delta_{\mathcal{S}} u(\Xi, t^n) + \mathcal{O}(\Delta^2), \tag{5.13}$$

for $t^n = n\Delta t$. Recall that, in the proposed setup of [PLR18], the initial grid points $\mathcal{Z} \subset B(\mathcal{S})$ is regular, whereas $\Xi = cp(\mathcal{Z})$ could be highly nonuniform. Considering (5.12), we know that (5.13) is equivalent to the discrete equation of the constant-along-normal extended function, and therefore (5.13) becomes

$$u(\Xi, t^{n+1}) = \tilde{u}(\Xi, t^n) + \Delta t \Delta \tilde{u}(\Xi, t^n) + \mathcal{O}(\Delta^2). \tag{5.14}$$

In [PLR18] the authors do not use *pointwise* projection, i.e., $\tilde{u}(\Xi, t^n) = \tilde{u}(\mathcal{Z}, t^n)$ due to instability issues. Instead, they let $u(\Xi) = \tilde{u}(\Xi) \approx P\tilde{u}(\mathcal{Z})$ where P is a projection (extension) matrix. Putting all these together, the approximate solution can be updated from time t^n to t^{n+1} by

$$\tilde{u}(\Xi, t^{n+1}) = (P + \Delta t\ W)\tilde{u}(\mathcal{Z}, t^n). \tag{5.15}$$

Notice that the projection matrix P is nothing but the RBF interpolation matrix.

Remark 19. *We highlight that the explicit closest point method explained in the previous subsection evaluates the derivatives of the function u on the grid points $\{x_j\}$ in the embedding space. On the other hand, the RBF-FD method calculates the derivatives directly on the closest points $\{y_j\}$ on the surface. This eliminates the interpolation step in evaluating derivatives, thereby eliminating a potential source of error and computational cost.*

5.4 Direct RBF-PU CP method

An alternative approach for generating the differentiation weights in (5.8) is the *partition of unity* (PU) method, first discussed in [BM97, MB97]. The basic idea is to start with a partition of the open and bounded domain $\Omega \subset \mathbb{R}^d$ into M subdomains Ω_j such that $\Omega \subset \cup_{j=1}^{M} \Omega_j$ with some mild overlap among the subdomains. Associated with these subdomains, we choose a partition of unity, i.e., a family of compactly supported, non-negative, continuous functions ω supported on the closure of Ω_j such that at every point ξ in Ω we have $\sum_{j=1}^{M} \omega_j(\xi) = 1$. Now, the global approximant of f is formed by joining the local approximant, belonging to each subdomain, via the weight associated with each subdomain. Letting $\mathcal{Z}_j = \mathcal{Z} \bigcap \Omega_j$, we have

$$s_{f,\mathcal{Z}} = \sum_{j=1}^{M} \omega_j s_{f,\mathcal{Z}_j}, \tag{5.16}$$

such that by $s_{f,\mathcal{Z}_j}$ we mean an approximant restricted to Ω_j. In the classical PU method, applying any linear differential operator $\mathcal{L}$ to $s_{f,\mathcal{Z}}$ leads to

$$\mathcal{L}s_{f,\mathcal{Z}} = \sum_{j=1}^{M} \mathcal{L}\left(\omega_j s_{f,\mathcal{Z}_j}\right). \tag{5.17}$$

For more information on PDE's solution using the RBF-PU method, see, e.g., [DMMPR19] and the reference therein. To make (5.17) well-defined, it is essential that both ω_j and $s_{f,\mathcal{Z}_j}$ are smooth enough. In particular, in contrast to the RBF-FD method, here in the classical PU scheme, the approximant (interpolant) was constructed first, and then the linear operator was applied. It means, recalling (4.1), the linear operator is acting on the shape functions directly. Consequently, computing (5.17) will be challenging as it requires applying a kind of Leibniz's rule for the operators $\mathcal{L}$ on the PU weights and the local approximate. Such a challenge may be why the PU approach has rarely been employed for surface PDEs.

To overcome such an issue, the *Direct PU* method for solving boundary value problems was first investigated in [Mir21] and then was employed to deal with surface PDEs in [MM23]. The first notable difference is that the PU weights are not required to be differentiated using the direct method. In other words, instead of working with a global interpolant and applying the linear operator, one applies the operator directly to the local interpolant and then joins all these local interpolants together using the PU weights.

In what follows, we describe our suggested numerical scheme to solve time-dependent PDEs intrinsic to a surface, employing CP representation of the surface and Direct RBF-PU to discretize the problem.

5.4.1 Description of the method

In this section, we introduce an explicit Direct RBF-PU closest point method for solving PDEs on the surfaces.

Again, consider the diffusion problem in (5.2) and its correspondence embedded equation in (5.3). Like the previous section, let $\mathcal{Z}$ and Ξ be a collection of Cartesian grid points and their associated closest points on the surface. The main difference compare to the RBF-FD-CP methods is that now, we need another set of points denoted by $C = \{\zeta_1, ..., \zeta_M\} \subset \Omega$ which are the centers of the patches $\{\Omega_\ell\}_{\ell=1}^{M}$ that makes a covering for the Ω. Figure (5.2) shows two different types of domain decomposition; while the number of patches is fixed, the radius of patches is different. Even though in Figure (5.2) the patch centers are distributed in the computational tube, this is not necessary, so instead, one can let the patch centers on the surface so we have $C \subset \mathcal{S}$. In both cases, one must always check that the patches contain at least one surface data point. Otherwise, considering such a patch would not be very sensible; see the right plot in Figure (5.2). We shall return to the explanation of grid and surface point generation later.

Now, we deploy the methodology to obtain a D-RBF-PU approximation. Analogous to (5.16), let $\mathcal{Z}_j = \mathcal{Z} \cap \Omega_j$, and $\Xi_j = \Xi \cap \Omega_j$ i.e., the Cartesian grid points and the surface closest points inside the $j-$patch $\Omega_j,\ 1 \leq j \leq M$. In D-RBF-PU, the value $\Delta\tilde{u}$ is directly

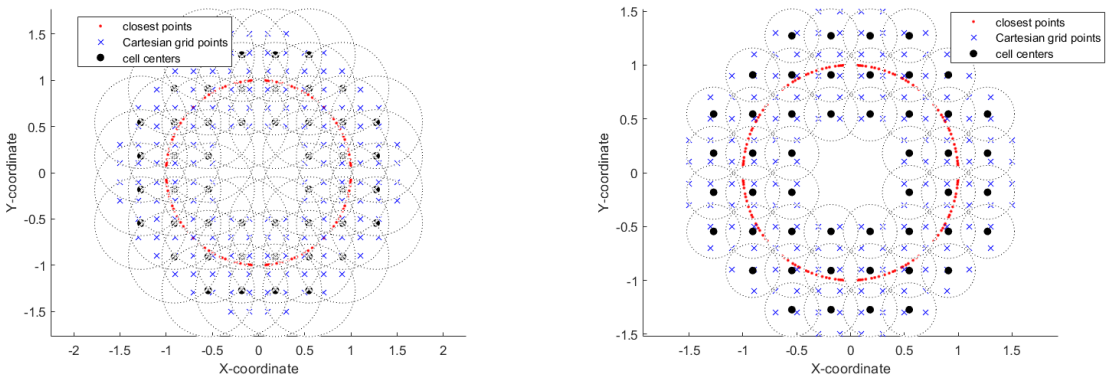

Figure 5.2: Domain decomposition with cell radius $r = 1/2$ (left), and cell radius $r = 1/4$ (right).

approximated by the PU method without any detour via local approximants $s^{\Delta}_{f,\mathcal{Z}_j}$ of $\tilde{u}$, say

$$\Delta\tilde{u}(\Xi) \approx \sum_{j=1}^{M} \omega_j(\Xi_j) s^{\Delta}_{f,\mathcal{Z}_j}(\Xi_j) := s^{\Delta}_{f,\mathcal{Z}}, \tag{5.18}$$

where $s^{\Delta}_{f,\mathcal{Z}_j}$ are the local approximants of $\Delta\tilde{u}$ on patches Ω_j. The resulting global approximation is denoted by $s^{\Delta}_{f,\mathcal{Z}}$, which is different from what is obtained in the classical RBF-PU.
To compute the local approximant $s_{f,\mathcal{Z}_j}$, we fix r to be the radius of each patch. Followingly, let J_j and I_j be the set of indices of the data sites and the surface points in the $j-$th patch, respectively, i.e.

$$J_j = \{i \in \{1, 2, \ldots, N\} : \mathbf{z}_i \in \mathcal{Z}_j\}.$$

and

$$I_j = \{i \in \{1, 2, \ldots, N\} : \xi_i \in \Xi_j\}.$$

Now, the local interpolant on $j-$th patch is represented as

$$s^{\Delta}_{f,\mathcal{Z}_j}(\Xi_j) = \left(\Delta\Phi^T_{\mathcal{Z}_j}(\Xi_j) V^{-1}_{\Phi,\mathcal{Z}_j}\right) \tilde{u}(\mathcal{Z}_j) \tag{5.19}$$

$$=: \mathbf{w}^T \tilde{u}(\mathcal{Z}_j). \tag{5.20}$$

One might notice that for a single surface point $\xi \in \Xi_\ell$ on the surface, Direct RBF-PU can be viewed as the RBF-FD method with the stencil $\mathcal{Z}_j$. However, the main difference here is that each patch can contain more than one surface point; $\Delta\Phi^T_{\mathcal{Z}_j}(\Xi_j) \in \mathbb{R}^{|I_j|\times|J_j|}$ such that each row is defined as in (5.9). Also, the patches can have some overlap, meaning that one surface point can belong to more than one patch. In the latter case, when assembling the final differentiation matrix, one must add the differentiation weights derived from different patches for each surface point ξ. Taking all these into consideration, the global interpolant (5.18) can be written

$$\Delta\tilde{u}(\Xi) = \sum_{\ell=1}^{M} \mathcal{W}_\ell \left(\Delta\Phi^T_{\mathcal{Z}_\ell}(\Xi_\ell) V^{-1}_{\Phi,\mathcal{Z}_\ell}\right) \tilde{u}(\mathcal{Z}_\ell), \tag{5.21}$$

where $\mathcal{W}_\ell \in \mathbb{R}^{|I_\ell|\times|I_\ell|}$ is a diagonal matrix carrying the PU weights $\omega(\xi_i, \zeta_\ell)$, $\xi_i \in \Xi_\ell$ on its diagonal. Eventually, considering the global indices of differentiation weights obtained in (5.21), one can assemble the final differentiation matrix:

$$\Delta_{\mathcal{S}} u(\Xi) = \Delta\tilde{u}(\Xi) = W\tilde{u}(\mathcal{Z}). \tag{5.22}$$

Time discretization can be done using a forward Euler scheme as in (5.13). Similar to the RBF-FD-CPM, the extension matrix in equation (5.15) is required to keep the scheme stable.

5.4.2 Node Generation

To obtain efficient algorithms, any embedding method should treat the embedding PDE on a narrow band

$$B(\mathcal{S}) = \{\mathbf{x} : \|\mathbf{x} - cp(\mathbf{x})\| \leq \gamma\} \tag{5.23}$$

surrounding the surface, where γ is the bandwidth. There are several reasons why limiting the computation to a narrow band considerably complicates the solutions. For instance, to solve the embedding PDE on the band, artificial boundary conditions must be applied at the boundaries of the computational band. However, the choice of the bandwidth, denoted as γ, remains unclear and lacks justification through analytical arguments. On the other hand, using the closest point method within the evolution strategy can simplify narrow banding by avoiding the introduction of artificial boundaries. This approach allows for a clear separation of the evolution at the surface and its extension throughout space. For further information, please refer to [MM12]. Specifically, in [RM08] and later in [PLR18], the bandwidth γ is obtained via dealing with the Gauss circle problem [Wei04], which is finding the number of integer lattice points m inside a circle with radius R centered at the origin. Table 5.1, derived from [PLR18], shows the bandwidth and the stencil size in the RBF-FD-CP method.

As stated, in practice, to generate the computational domain, one needs to maintain a uniform mesh of grid nodes in the embedding space, i.e., a domain Ω that encompasses the surface $\mathcal{S}$. Then, it only matters to find those grid points at a distance at most γ with the surface boundary $\mathcal{S}$. The final step is to project the selected points on the surface of interest using the closest point function. It can be done via optimization techniques or a closed form of the CP function. To illustrate, assume the surface of the interest is a circle with radius R centered at the origin, meaning that it can be expressed with the parametric form

$$\mathcal{S} = \{\mathbf{x} : \mathbf{x}(\theta) = R\big(\cos(\theta), \sin(\theta)\big),\ 0 \leq \theta \leq 2\pi\} \tag{5.24}$$

Thus, we can find a closed formula for the closest point function that projects an arbitrary point $\mathbf{x} \in \Omega$ on $\mathcal{S}$, defined as

$$cp(\mathbf{x}) = \frac{R}{R'}(R' \cos(\theta), R' \sin(\theta)) \tag{5.25}$$

where $R^{'}$ is the distance of ξ from the origin in the polar coordinate system. Such a procedure can be extended to the sphere of radius R or an ellipse. In the first case, the surface is expressed as

$$\mathcal{S} = \{\mathbf{x} : \mathbf{x}(\theta, \varphi) = R\big(\cos(\theta)\cos(\varphi), \sin(\theta)\cos(\theta), \sin(\varphi)\big),\ -\pi \leq \theta \leq \pi, -\frac{\pi}{2} \leq \varphi \leq \frac{\pi}{2}\} \tag{5.26}$$

while in the latter, we have

$$\mathcal{S} = \{\mathbf{x} : \mathbf{x}(\theta) = \big(a\cos(\theta), b\sin(\theta)\big),\ 0 \leq \theta \leq 2\pi\} \tag{5.27}$$

where a, b are the radius on the X and Y axes respectively.

$\gamma\,(2D)$	$m\,(2D)$	$\gamma\,(3D)$	$m\,(3D)$
$(\sqrt{2}+\sqrt{2}/2)\,\Delta\mathbf{x}$	9	$(\sqrt{3}+\sqrt{3}/2)\,\Delta\mathbf{x}$	27
$(\sqrt{4}+\sqrt{2}/2)\,\Delta\mathbf{x}$	13	$(\sqrt{4}+\sqrt{3}/2)\,\Delta\mathbf{x}$	33
$(\sqrt{5}+\sqrt{2}/2)\,\Delta\mathbf{x}$	21	$(\sqrt{5}+\sqrt{3}/2)\,\Delta\mathbf{x}$	57
$(\sqrt{8}+\sqrt{2}/2)\,\Delta\mathbf{x}$	25	$(\sqrt{6}+\sqrt{3}/2)\,\Delta\mathbf{x}$	81

Table 5.1: Computational tube radius γ for an m-point RBF-FD stencil in two and three dimensions

5.5 Numerical results

In what follows, we present two numerical experiments, its D-RBF-PU-CP solution compared with the RBF-FD-CP method.

5.5.1 Test problem 1

In this experiment, we consider the heat equation

$$u_t = \Delta_{\mathcal{S}} u$$

intrinsic to the unit circle S centered at the origin. Following [RM08], for an initial profile $u(\theta, 0) = \sin(\theta)$, the exact solution is

$$u(\theta, t) = e^{-t}\sin(\theta).$$

Using an analytic closest point representation of the unit circle (5.25), the surface heat equation is discretized and solved using the proposed Direct-RBF-PU closest points method and RBF-FD-CPM scheme of [PLR18]. We work with unscaled (shape parameter $\varepsilon = 1$) C^4 *Matérn* RBF in both cases.

Following [PLR18], we let computational tube radius γ to be $(\sqrt{4}+\sqrt{2}/2)$ (see Table 5.1) with $m = 13$. In order to generate the Cartesian grid around S, we begin by generating a grid

on the square $\Omega = (-2,2) \times (-2,2)$. Then, we select the grid points inside the tubular domain and the closest points on $\mathcal{S}$, according to the methodology explained in 5.4.2. We consider different grid sizes $\Delta \mathbf{x} = \{0.2,\ 0.1,\ 0.05,\ 0.025,\ 0.0125,\ 0.00625\}$. Figure 1 (left) shows the computational domain with grid size $\Delta \mathbf{x} = 0.2$.
For employing the Direct-RBF-PU CP method, we must create another Cartesian grid around $\mathcal{S}$ to decide our patch centers. We use another mesh with a different grid size on Ω to ensure that patch centers differ from the RBF centers. In our numerical test the second grid contains $\{6^2,\ 12^2,\ 24^2,\ 48^2,\ 96^2,\ 192^2\}$ patches with corresponding cell radius $\{1,\ 1/2,\ 1/4,\ 1/8,\ 1/16,\ 1/32\}$. Eventually, the C^2 Wendland compactly supported function is used as the weight function. The left plot in Figure 5.3 shows the $\infty-$ norm error for different grid sizes. Moreover, Fig-

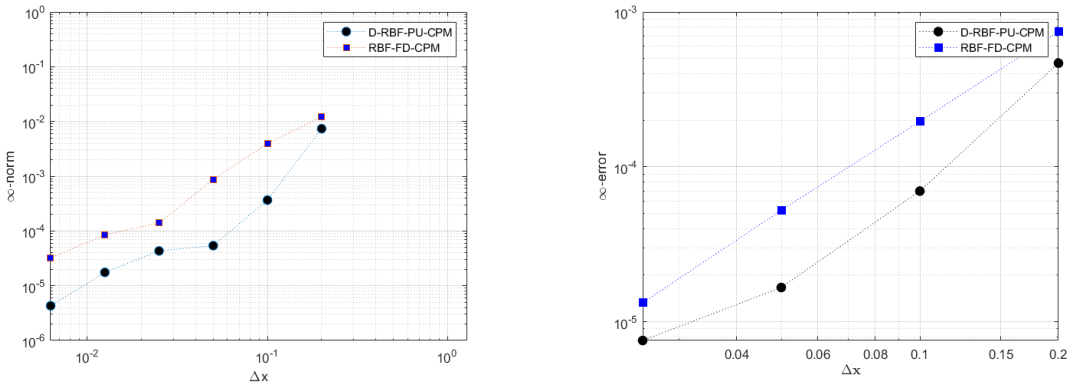

Figure 5.3: Relative error against the grid spacing $\Delta \mathbf{x}$ for the approximation of the Laplace–Beltrami operator on the unit circle (left) and sphere (right).

ure (5.4) shows the sparsity of the final differentiation matrix. Although the nonzero elements in RBF-FD are much smaller, our experiment shows that even increasing the stencil size to $m = 50$ (nz would be almost equal), the Direct-RBF-PU method still outperforms the RBF-FD method. In both cases, one should notice that optimizing different parameterize, such as ε,

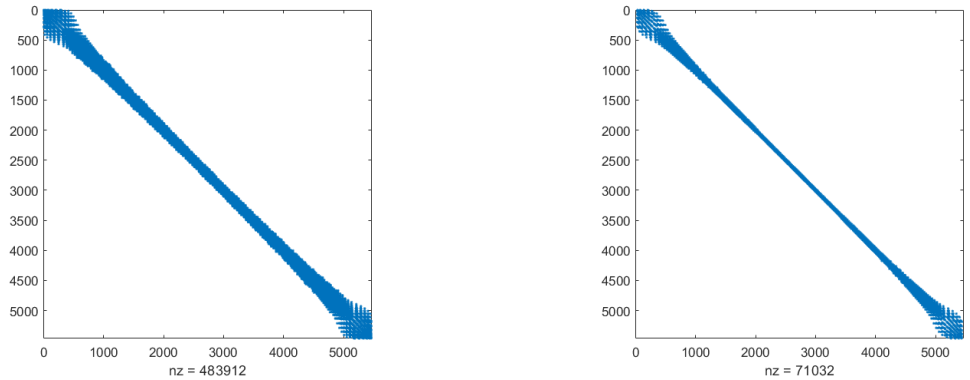

Figure 5.4: Sparsity of the final differentiation matrix for D-RBF-PU-CP method (left) and RBF-FD-CP method (right); Test problem 1

cell radius, number of patches, or stencil size, could lead to different (possibly better) accuracy.

5.5.2 Test problem 2

As the second test, we apply the D-RBF-PU method to a three-dimensional problem: the heat equation on the sphere

$$u_t = \Delta_{\mathcal{S}} u$$

intrinsic to the unit sphere S centered at the origin. For an initial profile $u(\theta, \varphi, 0) = \sin(\varphi)$, the exact solution for all times t is given by,

$$u(\theta, \varphi, 0) = e^{-2t} \sin(\varphi).$$

In this problem, we employ discontinuous weight functions such that

$$\omega_j(\xi) = \begin{cases} 1, & \text{if } j\text{-th patch is the nearest patch to } \xi \\ 0, & \text{otherwise} \end{cases} \tag{5.28}$$

which gives the total weight 1 to the nearest patch and null weights to the others. With such a weight function, the D-RBF-PU method gets very similar to the RBF-FD scheme, with the difference that the center of the patch (stencil) is not ξ necessarily. In this test, the radius of the patches is selected to be $4 \max(D)$ where D is the distance of surface data points with their nearest patch centers. Another difference with our previous example is that now we let the patch centers on the surface $\mathcal{S}$ rather than in the computational tube around it. This is possible simply by generating a grid around the surface and then projecting them on the surface.
In this example, we employ Gaussian RBF with shape parameter $\varepsilon = 0.05$; however, to avoid any potential problem of ill-conditioning, we use the stable RBF-GA method, which provides an accurate and stable algorithm and is a cheaper stabilization method over RBF-QR [FLP13]. It appears that it is of essential importance to use such stabilization methods to obtain good results in terms of accuracy. Otherwise, good accuracy is not obtainable. We let the computational tube radius γ to be $(\sqrt{5} + \sqrt{2}/2)$ (see Table 5.1) with $m = 57$ for the RBF-FD method. An analytic closest point representation of the unit sphere is used, and the surface heat equation is discretized and solved; the right plot of Figure5.3 shows the comparison of the accuracy between the proposed method and the RBF-FD method. Moreover, Table (5.2) shows the number of nonzero elements of the differentiation matrix associated with each grid size $\Delta\mathbf{x}$ for D-PU and RBF-FD schemes.

$\Delta\mathbf{x}$	N	M	$nnz.D-PU$	$nnz.RBF-FD$
0.2	2240	1304	42,448	127,680
0.1	8072	5072	159,016	460,104
0.05	31416	18824	829,608	1,790,712
0.025	125216	72680	5,403,184	7,137,312

Table 5.2: Number of nonzero elements of differentiation matrix associated to each grid size $\Delta\mathbf{x}$, total number of points N, and number patches M (only for D-RBF-PU-CP scheme).

Although the number of patches is relatively high, compared with the RBF-FD scheme, the final differentiation matrix is more sparse, and the accuracy is higher.

5.6 Conclusion

This chapter reviewed the closest point method for embedding the PDE problems intrinsic to the surface into a higher dimension. The Direct RBF-PU scheme was employed to solve the new embedded PDE. Although more test problems are required, the preliminary results show an improvement in accuracy compared with the RBF-FD method.

To continue, one can consider using PHS with appended polynomials instead of other kernels since this strategy might help overcome the instability and remove the RBF-GA strategy. Another approach is considering the eigenvalues of the operators and then, considering SVD truncated bases, removing those that are too small.

Another direction to develop this work is to consider the MLS-VSDK approach suggested in chapter (4) to generate differentiation weights instead of the suggested D-RBF-PU and deal with more complex PDE problems that could also be discontinuous. Generating the differentiation weights with the better-conditioned kernels, as outlined in the remark (9), can be another interesting direction toward the extension of this chapter.

6
Conclusion

In the first chapter, we presented the idea of function approximation in the most general setting. We confine ourselves to RBF approximation and interpolation to answer the question of which set of bases is appropriate for function approximation, especially in the multidimensional framework. We provided a reasonable connection between RBF interpolation and general interpolation, for instance, Fourier interpolation in Chapter 2. Our first contribution starts in Chapter 3 where we show how to find a new set of bases that are better conditioned compared to the standard translated RBFs. In Chapter 4 we present our answer to the natural question of whether it makes sense to approximate discontinuous functions with continuous bases. Toward this idea, we incorporated the discontinuities into the bases to achieve a better approximation scheme in terms of accuracy. Thus, we would rather consider the third and fourth chapters general works, meaning they can be applied to various problems in different frameworks. Indeed, the PDE problems intrinsic to the surface investigated in Chapter 5 are one of those.

To continue this thesis work, apart from the new suggested method based on the Direct RBF-PU method, we would like to consider the MLS-VSDK approach suggested in Chapter 4 to generate differentiation weights and deal with more complex PDE problems, those with discontinuity and in particular, the PDEs with Neumann boundary conditions. Besides, our numerical tests in Chapter 5 show the necessity of employing a stabilization method, recalling that RBF-GA was used to avoid the ill-conditioning of the final linear system. Thus, another possible approach to generating the differentiation weights can be using the better-conditioned bases outlined in Chapter 3 joined with the remark 9. Moreover, in the framework of Chapter 3, it is also worth investigating the new bases that span the native space and are L_2 orthonormal. These bases can be seen as alternate bases for the eigenfunctions associated with Mercer's series, which are not straightforward to obtain. Also, using L_2 orthonormal bases, one should be able to connect the RBF interpolation with the generalized Fourier function approximation.

A

Appendix:

All the MATLAB codes required for reproducing the numerical results reported in this thesis are available at the repository provided by the author on GitHub: `https://github.com/Mohes7395`.

The readers can use all the codes for personal use, but please mention them.

References

[AD01] Maria G Armentano and Ricardo G Durán. Error estimates for moving least square approximations. *Applied Numerical Mathematics*, 37(3):397–416, 2001.

[AF03] Robert A Adams and John JF Fournier. *Sobolev Spaces*. Elsevier, 2003.

[AH05] Kendall Atkinson and Weimin Han. *Theoretical numerical analysis*, volume 39. Springer, 2005.

[Arm01] María G Armentano. Error estimates in sobolev spaces for moving least square approximations. *SIAM Journal on Numerical Analysis*, 39(1):38–51, 2001.

[Bay19] Víctor Bayona. Comparison of moving least squares and rbf+ poly for interpolation and derivative approximation. *Journal of Scientific Computing*, 81:486–512, 2019.

[BCM99] Richard K Beatson, Jon B Cherrie, and Cameron T Mouat. Fast fitting of radial basis functions: Methods based on preconditioned gmres iteration. *Advances in Computational Mathematics*, 11:253–270, 1999.

[BCOS01] Marcelo Bertalmio, Li-Tien Cheng, Stanley Osher, and Guillermo Sapiro. Variational problems and partial differential equations on implicit surfaces. *Journal of Computational Physics*, 174(2):759–780, 2001.

[Bjö96] Åke Björck. *Numerical methods for least squares problems*. SIAM, 1996.

[BLM11] Richard K Beatson, Jeremy Levesley, and CT Mouat. Better bases for radial basis function interpolation problems. *Journal of Computational and Applied Mathematics*, 236(4):434–446, 2011.

[BLRS15] Mira Bozzini, Licia Lenarduzzi, Milvia Rossini, and Robert Schaback. Interpolation with variably scaled kernels. *IMA Journal of Numerical Analysis*, 35(1):199–219, 2015.

[BM97] Ivo Babuška and Jens M Melenk. The partition of unity method. *International journal for numerical methods in engineering*, 40(4):727–758, 1997.

[Bre08] Susanne C Brenner. *The mathematical theory of finite element methods*. Springer, 2008.

[BS89] LEONARD PETER Bos and K Salkauskas. Moving least-squares are backus-gilbert optimal. *Journal of Approximation Theory*, 59(3):267–275, 1989.

[Buh03] Martin Dietrich Buhmann. *Radial basis functions*. Cambridge University Press, Cambridge, UK, 2003.

[Che15] Yujia Chen. *Geometric multigrid and closest point methods for surfaces and general domains*. PhD thesis, Oxford University, UK, 2015.

[Chu97] Charles K Chui. *Wavelets: a mathematical tool for signal analysis*. SIAM, 1997.

[CL09] Elliott Ward Cheney and William Allan Light. *A course in approximation theory*, volume 101. American Mathematical Soc, 2009.

[CM15] Yujia Chen and Colin B Macdonald. The closest point method and multigrid solvers for elliptic equations on surfaces. *SIAM Journal on Scientific Computing*, 37(1):A134–A155, 2015.

[CT18] Jay Chu and Richard Tsai. Volumetric variational principles for a class of partial differential equations defined on surfaces and curves: In memory of heinz-otto kreiss. *Research in the Mathematical Sciences*, 5:1–38, 2018.

[Dau92] Ingrid Daubechies. *Ten lectures on wavelets*. SIAM, 1992.

[Dav75] Philip J Davis. *Interpolation and approximation*. Courier Corporation, 1975.

[DB78] Carl De Boor. *A practical guide to splines*, volume 27. springer-verlag New York, 1978.

[DeV98] Ronald A DeVore. Nonlinear approximation. *Acta numerica*, 7:51–150, 1998.

[DGJ03] Qiang Du, Max D Gunzburger, and Lili Ju. Voronoi-based finite volume methods, optimal voronoi meshes, and pdes on the sphere. *Computer methods in applied mechanics and engineering*, 192(35-36):3933–3957, 2003.

[DMEM+20] Stefano De Marchi, Wolfgang Erb, Francesco Marchetti, Emma Perracchione, and Milvia Rossini. Shape-driven interpolation with discontinuous kernels: Error analysis, edge extraction, and applications in magnetic particle imaging. *SIAM Journal on Scientific Computing*, 42(2):B472–B491, 2020.

[DMMP19] Stefano De Marchi, A Martínez, and Emma Perracchione. Fast and stable rational rbf-based partition of unity interpolation. *Journal of Computational and Applied Mathematics*, 349:331–343, 2019.

[DMMP20] Stefano De Marchi, Francesco Marchetti, and Emma Perracchione. Jumping with variably scaled discontinuous kernels (vsdks). *BIT Numerical Mathematics*, 60:441–463, 2020.

[DMMPR19] Stefano De Marchi, A Martinez, Emma Perracchione, and M Rossini. Rbf-based partition of unity methods for elliptic pdes: Adaptivity and stability issues via variably scaled kernels. *Journal of Scientific Computing*, 79:321–344, 2019.

[DMS10] Stefano De Marchi and Robert Schaback. Stability of kernel-based interpolation. *Advances in Computational Mathematics*, 32:155–161, 2010.

[DMS13] Stefano De Marchi and Gabriele Santin. A new stable basis for radial basis function interpolation. *Journal of Computational and Applied Mathematics*, 253:1–13, 2013.

[DMS15] Stefano De Marchi and Gabriele Santin. Fast computation of orthonormal basis for RBF spaces through Krylov space methods. *BIT Numerical Mathematics*, 55:949–966, 2015.

[EY36] Carl Eckart and Gale Young. The approximation of one matrix by another of lower rank. *Psychometrika*, 1(3):211–218, 1936.

[Fas07] Gregory E Fasshauer. *Meshfree approximation methods with MATLAB*, volume 6. World Scientific, 2007.

[FF15a] Bengt Fornberg and Natasha Flyer. *A primer on radial basis functions with applications to the geosciences*. SIAM, 2015.

[FF15b] Bengt Fornberg and Natasha Flyer. Solving pdes with radial basis functions. *Acta Numerica*, 24:215–258, 2015.

[FFBB16] Natasha Flyer, Bengt Fornberg, Victor Bayona, and Gregory A Barnett. On the role of polynomials in rbf-fd approximations: I. interpolation and accuracy. *Journal of Computational Physics*, 321:21–38, 2016.

[FH05] Michael S Floater and Kai Hormann. Surface parameterization: a tutorial and survey. *Advances in multiresolution for geometric modelling*, pages 157–186, 2005.

[FLB+12] Natasha Flyer, Erik Lehto, Sébastien Blaise, Grady B Wright, and Amik St-Cyr. A guide to rbf-generated finite differences for nonlinear transport: Shallow water simulations on a sphere. *Journal of Computational Physics*, 231(11):4078–4095, 2012.

[FLF11] Bengt Fornberg, Elisabeth Larsson, and Natasha Flyer. Stable computations with gaussian radial basis functions. *SIAM Journal on Scientific Computing*, 33(2):869–892, 2011.

[FLP13] Bengt Fornberg, Erik Lehto, and Collin Powell. Stable calculation of gaussian-based rbf-fd stencils. *Computers & Mathematics with Applications*, 65(4):627–637, 2013.

[FM12] Gregory E Fasshauer and Michael J McCourt. Stable evaluation of gaussian radial basis function interpolants. *SIAM Journal on Scientific Computing*, 34(2):A737–A762, 2012.

[FM15] Gregory E. Fasshauer and Michael J. McCourt. *Kernel-based approximation methods using Matlab*, volume 19. World Scientific Publishing Company, 2015.

[FP08] Bengt Fornberg and Cécile Piret. A stable algorithm for flat radial basis functions on a sphere. *SIAM Journal on Scientific Computing*, 30(1):60–80, 2008.

[FW04] Bengt Fornberg and Grady Wright. Stable computation of multiquadric interpolants for all values of the shape parameter. *Computers & Mathematics with Applications*, 48(5-6):853–867, 2004.

[FZ07] Gregory E Fasshauer and Jack G Zhang. On choosing "optimal" shape parameters for rbf approximation. *Numerical Algorithms*, 45:345–368, 2007.

[GVL13] Gene H Golub and Charles F Van Loan. *Matrix computations*. JHU press, 2013.

[HH13] Klaus Höllig and Jörg Hörner. *Approximation and modeling with B-splines*. SIAM, 2013.

[Hir12] Morris W Hirsch. *Differential topology*, volume 33. Springer Science & Business Media, 2012.

[Hol01] Michael Holst. Adaptive numerical treatment of elliptic systems on manifolds. *Advances in Computational Mathematics*, 15(1):139–191, 2001.

[I+18] Armin Iske et al. *Approximation theory and algorithms for data analysis*. Springer, 2018.

[Lev98] David Levin. The approximation power of moving least-squares. *Mathematics of computation*, 67(224):1517–1531, 1998.

[Lin19] Peter Linz. *Theoretical numerical analysis*. Courier Dover Publications, 2019.

[LLZ11] Shingyu Leung, John Lowengrub, and Hongkai Zhao. A grid based particle method for solving partial differential equations on evolving surfaces and modeling high order geometrical motion. *Journal of Computational Physics*, 230(7):2540–2561, 2011.

[LS81] Peter Lancaster and Kes Salkauskas. Surfaces generated by moving least squares methods. *Mathematics of computation*, 37(155):141–158, 1981.

[LZ13] Jian Liang and Hongkai Zhao. Solving partial differential equations on point clouds. *SIAM Journal on Scientific Computing*, 35(3):A1461–A1486, 2013.

[Mal99] Stéphane Mallat. *A wavelet tour of signal processing*. Elsevier, 1999.

[MB97] JM Melenk and I Babuška. Approximation with harmonic and generalized harmonic polynomials in the partition of unity method. *Computer Assisted Mechanics and Engineering Sciences*, 4:607–632, 1997.

[MB19] Maryam Mohammadi and Maryam Bahrkazemi. Bases for polynomial-based spaces. *Journal of Mathematical Modeling*, 7(1):21–34, 2019.

[MBR11] Colin B Macdonald, Jeremy Brandman, and Steven J Ruuth. Solving eigenvalue problems on curved surfaces using the closest point method. *Journal of Computational Physics*, 230(22):7944–7956, 2011.

[Mir15] Davoud Mirzaei. Analysis of moving least squares approximation revisited. *Journal of Computational and Applied Mathematics*, 282:237–250, 2015.

[Mir21] Davoud Mirzaei. The direct radial basis function partition of unity (d-rbf-pu) method for solving pdes. *SIAM Journal on Scientific Computing*, 43(1):A54–A83, 2021.

[MM12] Thomas Marz and Colin B Macdonald. Calculus on surfaces with general closest point functions. *SIAM Journal on Numerical Analysis*, 50(6):3303–3328, 2012.

[MM23] Reyhaneh Mir and Davoud Mirzaei. The d-rbf-pu method for solving surface pdes. *Journal of Computational Physics*, 479:112001, 2023.

[MR08] Colin B Macdonald and Steven J Ruuth. Level set equations on surfaces via the closest point method. *Journal of Scientific Computing*, 35:219–240, 2008.

[MR10] Colin B Macdonald and Steven J Ruuth. The implicit closest point method for the numerical solution of partial differential equations on surfaces. *SIAM Journal on Scientific Computing*, 31(6):4330–4350, 2010.

[MS09] Stefan Müller and Robert Schaback. A newton basis for kernel spaces. *Journal of Approximation Theory*, 161(2):645–655, 2009.

[MS13] Davoud Mirzaei and Robert Schaback. Direct meshless local petrov–galerkin (dmlpg) method: a generalized mls approximation. *Applied Numerical Mathematics*, 68:73–82, 2013.

[MSD12] Davoud Mirzaei, Robert Schaback, and Mehdi Dehghan. On generalized moving least squares and diffuse derivatives. *IMA Journal of Numerical Analysis*, 32(3):983–1000, 2012.

[NW08] Erich Novak and Henryk Woźniakowski. *Tractability of Multivariate Problems: Standard information for functionals*, volume 2. European Mathematical Society, 2008.

[NWW05] Francis Narcowich, Joseph Ward, and Holger Wendland. Sobolev bounds on functions with scattered zeros, with applications to radial basis function surface fitting. *Mathematics of Computation*, 74(250):743–763, 2005.

[NWW06] Francis J Narcowich, Joseph D Ward, and Holger Wendland. Sobolev error estimates and a bernstein inequality for scattered data interpolation via radial basis functions. *Constructive Approximation*, 24:175–186, 2006.

[PLPR19] Argyrios Petras, Leevan Ling, Cécile Piret, and Steven J Ruuth. A least-squares implicit rbf-fd closest point method and applications to pdes on moving surfaces. *Journal of Computational Physics*, pages 146–161, 2019.

[PLR18] Argyrios Petras, Leevan Ling, and Steven J Ruuth. An rbf-fd closest point method for solving pdes on surfaces. *Journal of Computational Physics*, 370:43–57, 2018.

[Pow81] Michael James David Powell. *Approximation theory and methods*. Cambridge university press, 1981.

[PS11] Maryam Pazouki and Robert Schaback. Bases for kernel-based spaces. *Journal of Computational and Applied Mathematics*, 236(4):575–588, 2011.

[PS13] Maryam Pazouki and Robert Schaback. Bases for conditionally positive definite kernels. *Journal of Computational and Applied Mathematics*, 243:152–163, 2013.

[RF10] Halsey Royden and Patrick Michael Fitzpatrick. *Real analysis*. China Machine Press, 2010.

[RM08] Steven J Ruuth and Barry Merriman. A simple embedding method for solving partial differential equations on surfaces. *Journal of Computational Physics*, 227(3):1943–1961, 2008.

[Rud87] Walter Rudin. *Real and Complex Analysis, 3rd ed*. McGraw-Hill, New York,, 1987.

[S+95] Robert Schaback et al. Multivariate interpolation and approximation by translates of a basis function. *Series In Approximations and Decompositions*, 6:491–514, 1995.

[Sch13] Robert Schaback. Direct discretizations with applications to meshless methods for pdes. *Dolomites Research Notes on Approximation*, 6(Special_Issue), 2013.

[She68] Donald Shepard. A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 23rd ACM national conference*, 1968.

[SK19] Pratik Suchde and Joerg Kuhnert. A meshfree generalized finite difference method for surface pdes. *Computers & Mathematics with Applications*, 78(8):2789–2805, 2019.

[Sob38] Sergej Lvovich Sobolev. On a theorem of functional analysis. *Mat. Sbornik*, 4:471–497, 1938.

[Sob08] Sergej Lvovich Sobolev. *Some applications of functional analysis in mathematical physics*, volume 90. American Mathematical Soc, 2008.

[TS03] Andrei I Tolstykh and DA Shirobokov. On using radial basis functions in a "finite difference mode" with applications to elasticity problems. *Computational Mechanics*, 33(1):68–79, 2003.

[vGMM13] Ingrid von Glehn, Thomas März, and Colin B Macdonald. An embedded method-of-lines approach to solving partial differential equations on surfaces. *arXiv preprint arXiv:1307.5657*, 2013.

[Wei04] Eric W Weisstein. Gauss's circle problem. *https://mathworld. wolfram. com/*, 2004.

[Wen95] Holger Wendland. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in computational Mathematics*, 4:389–396, 1995.

[Wen98] Holger Wendland. Error estimates for interpolation by compactly supported radial basis functions of minimal degree. *Journal of approximation theory*, 93(2):258–272, 1998.

[Wen01] Holger Wendland. Local polynomial reproduction and moving least squares approximation. *IMA Journal of Numerical Analysis*, 21(1):285–300, 2001.

[Wen02] Holger Wendland. Fast evaluation of radial basis functions: Methods based on partition of unity. 2002.

[Wen04] Holger Wendland. *Scattered data approximation*, volume 17. Cambridge university press, 2004.

[WF06] Grady B Wright and Bengt Fornberg. Scattered node compact finite difference-type formulas generated from radial basis functions. *Journal of Computational Physics*, 212(1):99–123, 2006.

[WS93] Zong-min Wu and Robert Schaback. Local error estimates for radial basis function interpolation of scattered data. *IMA journal of Numerical Analysis*, 13(1):13–27, 1993.

[Zm95] Wu Zong-min. Multivariate compactly supported positive definite radial functions. *Advances in Comp. Math*, 1995.

Acknowledgments

With this thesis coming to an end, it is time to share some non-mathematical words regarding my experiences and the people I met during these three years. These acknowledgments are, as usual, placed at the end of the manuscript, although it might be a bit unfair, given that nothing would have been possible without the names you will read in the following lines.

First of all, I would like to express my gratitude to my supervisor, Prof. Stefano De Marchi. It is said that when you meet a new person, you have no idea what journey you will have alongside him/her in the future, and this was exactly true about him. I can still remember the very first time when I was a master student, and I met him at a Conference in Tehran, asking for some suggestions about my research. It started with a one-month-long visit to Padova University in September 2019 to finish my master's thesis and also attend the Canazei workshop. That experience was one of the most important ones that helped me decide to continue my research and start a PhD. In fact, my master's degree was in financial mathematics, but working with him made me interested in approximation theory using kernel methods, so I started my PhD at the University of Padova in December 2020 under his supervision. During this journey, I learned much from him, not only about mathematics but also about the applications and insights required in the research. I will always be grateful to you, Professor De Marchi, for supporting me all these years and helping me grow and become a better researcher.

A huge thanks goes to Dr. Francesco Marchetti, who helped immensely with my PhD, which led to my first PhD article (chapter 4). Francesco, this PhD would have been impossible without your presence. Thanks for being an instructor, colleague, and friend in these years.

I should also thank Prof. Maryam Mohammadi, who was greatly helpful in deepening and expanding my mathematical knowledge and, more importantly, learning how to take an appropriate approach to initializing small ideas and then extending them to worthy research. Our joint work led to the project, which is explained in Chapter 3.

I would like to thank Prof. Davoud Mirzaei. I was honored to spend a visiting period at the Uppsala University of Sweden for three months, hosted by him. Not only is he a sharp and strong mathematician, but he is also a welcoming and friendly person. Although three months seemed too short, it was a great experience and an opportunity to learn a new way of thinking and evaluating the research activities. I have learned more than I imagined. All materials explained in Chapter 5 are a joint work resulting from my visit there. In particular, he helped me develop and deepen the theoretical material of Chapter 4. I must also thank him for his fruitful comments on improving this thesis.

I want to thank all the numerical analysis research groups here in Padova, particularly Dr. Giacomo Elephante, with whom I attended different conferences and workshops.

Apart from the people I had the chance to work with, there are also some people who deserve to be mentioned here since, without their emotional support, it would have been impossible for me to complete this PhD.

Above all, I have to appreciate my family: my father, my mother, my sister, and my brother, who have given me positive energy during hard times. I also have to appreciate my best friends, Farid and Hessam, Nima, and Soroush, who have accompanied me in every step of my life since we were just kids in high school. Though I was far away from you physically, part of my heart always stayed there in Tehran beside you.

Besides, I was also lucky to have some good colleagues who became good friends during these years. First, I should name Giacomo Lanaro for the morning coffee, our chess games, our deep discussion, and for letting me know more about the history of Europe and Italy. The memories and moments that I shared with you are much more to be written here.

There are some people in life who we meet late and then leave us early. Ofelia, the first time we discussed anything except daily greetings, I was already in the second year of my PhD, and you were already thinking of leaving for England to continue your career as a Postdoc. During the short time you stayed at Padova, you became such a good and close friend. Thanks for everything, L'Ofelia.

I should also thank Elisa for staying supportive in difficult situations, particularly those days when I had no home in Padova, and for the walks and breakfast in the early morning on weekends.

In the end, there are other friends with whom I had good and joyful moments; Saeed, Rahman, Mary, and Anousheh are only a few to mention. Thank you guys for being there. Besides, special thanks go to my compatriot Arghavan. I will always remember when we met downstairs to have Iranian tea for a couple of minutes, and then we ended up talking for a couple of hours. As you surely know, there has never been any friendship among us, but we have been only two compatriots. Thank you for your support in these three years and all those philosophical and historical discussions, joined with laughter and joy.

It is too hard to summarize three and a half years of living in a new country with a new culture and new language, but I better finish my words here by remembering all those who accompanied me through this journey.