Available online at www.sciencedirect.com
ScienceDirect
Engineering Analysis with Boundary Elements 00 (2026) 1–16
Eng. Anal. Bound. Elements

Learning Radial Basis Functions from the German School: A Personal Perspective

Stefano De Marchi

Department of Medicine, University of Padova
Via Gallucci, 6 - 37128 Padova (Italy)

Abstract

Radial Basis Functions (RBFs) are a powerful tool for scattered data approximation, interpolation, and the numerical solution of partial differential equations. Over the past decades, a large body of work originating from the so-called German school of RBFs, led by Robert Schaback and collaborators, has made significant contributions to the theoretical understanding and practical development of kernel-based approximation methods.

This paper presents a personal perspective on some of the ideas and developments that originated within this community and have influenced subsequent research directions. In particular, we discuss greedy algorithms for center selection, stability issues related to the shape parameter, as well as the choice of the centers inside or along the boundaries of the domain, rational kernel interpolation, and the framework of variably scaled kernels. These topics represent, I consider, my most interesting contributions to the kernels community, made possible by the fruitful relationships with the German school of RBFs that has blossomed from Robert. This manuscript also tries to illustrate how theoretical insights gained in kernel approximation have led to powerful tools in numerical analysis, machine learning, and scientific computing.

Keywords: Radial basis functions, stability, greedy algorithms, variably scaled kernels
2000 MSC: 41A05, 41A20, 65D12, 65D40

1. Why this paper

Let me begin by briefly recalling the origins of my interest in RBFs and my collaboration with Robert Schaback. My first contacts with the so-called "German school" began in 1995, when I met Greg Fasshauer at the SIAM Geometric Design conference. Later in 1999, I was invited by Martin Buhmann to Giessen for a Colloquium, and, in addition to Martin, I met Oleg Davydov. In 2000, in Firenze, during the conference on "Approssimazione di Curve e Superfici" (organized together with Carla Manni, University of Roma II), I met Robert Schaback and Armin Iske, who was a postdoc at SINTEF (Oslo).

Armin was one of Robert's PhD students who made significant contributions to the study of positive-definite functions in multivariate approximation. In particular, because of my interest in finding optimal points for multivariate polynomial interpolation, I read Armin's paper "Optimal distribution of Centers for Radial Basis Function Methods" [1] in which he discussed the quality and the stability of RBFs interpolation by computing the condition number of the kernel matrix by varying the distribution of the center set. It turned out that the most promising set of centers was quasi-uniformly distributed. This was my inspiration, and I asked Robert to make a formal proof that this assertion was indeed true. After that summer, I applied for the DAAD grant, which I received, and in the summer of 2001, I made my first visit to Göttingen, where I met Holger Wendland. All other contacts came later, during visits in Germany or at conferences, such as those

in the Dolomites in Alba di Canazei, which I have had the pleasure of organizing since 2006. Germans were often attendants because they all enjoy hiking and stay in the wonderful atmosphere of the Dolomites.

This is the beginning of my scientific interest in RBFs.

About the title. I deliberately used the word *Learning* in the title. The reason is twofold. Learning is the human action of acquiring new knowledge, behaviors, skills, values, or preferences through study, experience, or instruction. But in recent decades, *machine learning* has emerged as a kind of artificial intelligence (AI) that allows computers to learn from data and improve at tasks automatically. I learned many things from the German school, and for the first time, I understood the importance of kernel-machine learning in applications other than approximation of functions and data, again thanks to one of Robert's papers [2].

1.1. Preamble

Radial basis functions (RBFs) provide a flexible and powerful tool for approximating multivariate functions from scattered data. Since the pioneering work of several researchers in the 1980s and 1990s, the theory of kernel-based approximation has developed into a mature field that connects approximation theory, numerical analysis, machine learning, and data science.

Among the communities that have significantly contributed to this development, the group often referred to as the *German school of RBFs* has played a central role. Through the work of Robert Schaback and my collaboration with him, his collaborators in Göttingen, Germany, and worldwide, several fundamental concepts were clarified, including

- greedy algorithms for selecting interpolation centers;
- stability issues associated with positive definite kernels;
- the role of the shape parameter and of the power function in the interpolation error analysis with emphasis on variably scaled kernels;
- the connection between kernel methods and modern machine learning.

The purpose of this article is to present a brief survey of these ideas, combined with a personal perspective on how they influenced not only the research developed in Padova.

2. Interpolation by polynomials

To make the paper self-contained, we present some preliminaries about the interpolation problem with polynomials of total degree n in $\mathbb{R}$ and then in $\mathbb{R}^d$, $d \geq 2$.

Let $\mathbb{P}_n(\mathbb{R})$ be the space of the univariate polynomials of total degree $\leq n$ on $\mathbb{R}$ and $C(\mathbb{R})$ the linear space of continuous functions on $\mathbb{R}$. Further, for the basis of monomials $\mathcal{M} = \{1, x, x^2, \ldots, x^n\}$ and a set $X = \{x_0, \ldots, x_n\}$ of $n+1$ distinct points, we denote by

$$Vdm(X; \mathcal{M}) = \prod_{i<j} (x_i - x_j) \tag{1}$$

the corresponding *Vandermonde determinant* which plays an important role in the unisolvency of a given set of points.

The classical univariate interpolation problem of f by polynomials of degree n can be stated as follows.

Problem 1. *Let K be a closed and bounded set of $\mathbb{R}$. Consider, X a set of $n+1$ pairwise distinct points of K, the values $\{f(x_i),\ i = 0, \ldots, n\}$ and the basis of monomials $\mathcal{M} = \{1, x, \ldots, x^n\}$. Find the polynomial $p_n = \sum_{k=0}^{n} a_k x^k$, so that*

$$p_n(x_i) = f(x_i), \quad i = 0, \ldots, n. \tag{2}$$

Being $x_i \neq x_j$, $i \neq j$, p_n is unique because $Vdm(X; \mathcal{M}) \neq 0$. Using the *Lagrange basis* $L = \{l_i,\ i = 0, ..., n\}$ with

$$l_i(x) = \prod_{i=0, i\neq j}^{n} \frac{x - x_j}{x_i - x_j} = \frac{Vdm(X_i; \mathcal{M})}{Vdm(X; \mathcal{M})}$$

where X_i is the set X in which we substitute x_i with x, we can then write

$$p_n(x) = \sum_{i=0}^{n} l_i(x) f(x_i), \quad x \in K \tag{3}$$

This process generates an interpolation error at any $x \in K$, $e_n(x) = |f(x) - p_n(x)|$ or, taking the sup-norm, $E_n = \|f - p_n\|_\infty$. Using the Lagrange form (3) of the interpolant, we can bound this error by

$$E_n \leq (1 + \Lambda_n) E_n^* \tag{4}$$

with $\Lambda_n = \sup_{x \in K} \sum_{i=0}^{n} |l_i(x)|$ the *Lebesgue constant* which only depends on n and on the node set X. As well-known, Λ_n represents the sup-norm of the linear operator (cf. e.g. [3]) $L : C(\mathbb{R}) \to \mathbb{P}_n(R)$, $Lf = \sum_{i=0}^{n} f(x_i) l_i$, where, E_n^* is the error of best-uniform approximation, that is $E_n^* := \inf_{p_n \in \mathbb{P}_n(\mathbb{R})} E_n(f)$. The function $\lambda_n = \sum_{i=0}^{n} l_i$ is the *Lebesgue function*, which indicates how much the interpolation process can "blow up" errors in the data. It is a central tool for node selection, stability analysis, and interpolation scheme design. In the *one-dimensional case*, we know

- $\Lambda_n \approx 2^n$ when the set X is made of equally spaced points of K (or even worse when X are randomly chosen);
- $\Lambda_n \approx \log(n)$ when X is made of *Chebyshev-like* points of K.

We call *Chebyshev-like* points, those points that have the so-called *arccos-distribution* which characterizes, for instance, the Chebyshev-Gauss-Lobatto points (or Chebyshev extrema) $\left\{x_k = -\cos\left(\frac{k\pi}{n}\right),\ k = 0, ..., n\right\}$ and all zeros of orthogonal polynomials on a finite interval with respect to some positive measure. All these points are *near-optimal* in the sense that their Lebesgue constant grows logarithmically with respect to the degree n.

Question 1. Fundamental question. *Are there quasi-optimal interpolation nodes explicitly known in the multivariate setting for polynomial interpolation of total degree?*

The answer is partially negative, *except for some known cases and in small dimensions* (see, for instance, the seminal paper by L. Bos [4]). The previous question was the spring that pushed us to study new families of near-optimal points, starting from the square $[-1, 1]^2$, a simple, intrinsically tensorial domain that is easy to map to other domains (see [5]). This was the main reason why we discovered the *Padua points* on the square $\Omega = [-1, 1]^2$, which are still the only set of nearly optimal points, explicitly known, for interpolation with polynomials of degree $\leq n$ on the square $[-1, 1]^2$ (cf. [5, 6]). Other sets of points that have good approximation properties and that can be defined in any spatial dimension d, are the Fekete and Leja points, but their computation is challenging and remains an open problem. Solutions have been obtained numerically via linear algebra up to dimension 3 in [7]. A generalization of these sets is known as the set of *Lissajous points* (cf. [8, 9, 10]).

- We looked for *well-distributed nodes*. We found various nodal sets for polynomial interpolation of *even* degree n in the square Ω, which turned out to be equidistributed with respect to the *Dubiner metric* [11] and which show *near-optimal Lebesgue* constant growth [5].
- We also required *efficient interpolant evaluation*: the interpolant should be constructed without solving the Vandermonde system whose complexity is $O(N^3)$, for each pointwise evaluation, with $N = \binom{n+2}{2}$ the dimension of the bivariate polynomials of total degree $\leq$. Moreover, we sought closed-form formulae.
- We required *efficient cubature formulas*: in particular, a fast computation of cubature weights for non-tensorial cubature formulae.

The last two points were inspired by the *rule of* 10 claimed by Nick L. Trefethen in [12] (also in talk given in 2009 at the Dolomites Workshops in Alba di Canazei): a good implementation should last for 10 seconds, have a 10 digits precision and does not consist of more than 10 lines of executable code.

3. Kernel-Based Interpolation

Radial Basis Functions or RBF are functions that depend only on the Euclidean distance from a center point, i.e., $\phi(\|x - c\|)$, where x is a point in a given multidimensional domain and c is the center. Common RBFs include Gaussian, multiquadric, inverse multiquadric, and polyharmonic splines. They can be globally supported, locally supported, with different smoothness properties. Among the fundamental references, we recall these books [13, 14, 15].

Suppose we wish to conduct a literature search on platforms such as Google Scholar or specialized academic databases (e.g., IEEE Xplore, SpringerLink, or Elsevier's ScienceDirect). In that case, we can identify several thousand papers on RBFs across these fields. For example, a Google Scholar search for `Radial Basis Functions` yields more than 10^6 papers, including applications in machine learning, numerical analysis, and applied mathematics.

We can approximate functions or data at scattered points in $\mathbb{R}^M, M \geq 1$ by the following setting: Let $\Omega \subseteq \mathbb{R}^M$ be a bounded set, let $\mathcal{X}_N = \{\boldsymbol{x}_i, i = 1, \ldots, N\} \subseteq \Omega$ be a set of distinct data points (also called *data sites* or nodes) and let $\mathcal{F}_N = \{f_i = f(\boldsymbol{x}_i), i = 1, \ldots, N\}$ be a set of *data values* (or measurements or function values). The approximation problem consists in finding a function $P_f : \Omega \longrightarrow \mathbb{R}$ such that $P_f(\boldsymbol{x}_i) \approx f_i$, $i = 1, \ldots, N$. When equality holds we talk about *interpolation*. To this end, we consider $P_f \in \text{span}\{\Phi(\cdot, \boldsymbol{x}_i), \boldsymbol{x}_i \in \mathcal{X}_N\}$, where $\Phi : \Omega \times \Omega \longrightarrow \mathbb{R}$ is a strictly positive definite and symmetric kernel, for which the associated *interpolation or kernel* matrix A with entries $A_{i,j} = \Phi(x_i, x_j)$ is strictly positive definite. The interpolant then assumes the form

$$P_f(\boldsymbol{x}) = \sum_{k=1}^{N} \alpha_k \Phi(\boldsymbol{x}, \boldsymbol{x}_k), \quad \boldsymbol{x} \in \Omega. \tag{5}$$

The coefficients $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)^T$ in (5) are found by solving the linear system $A\boldsymbol{\alpha} = \boldsymbol{f}$, with $\boldsymbol{f} = (f_1, \ldots, f_N)^T$, where A is strictly positive definite and symmetric, ensures the uniqueness of the solution.

Definition 1. *A function $\sigma : \mathbb{R}^d \to \mathbb{R}$ is called radial, if there exists a continuous function $\phi : [0, \infty) \to \mathbb{R}$ such that $\sigma(\boldsymbol{x}) = \phi(r)$ with $r = \|\boldsymbol{x}\|_2$.*

Using this definition, given a "basic" function ϕ and the (symmetric positive definite) kernel Φ, for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ we have a radial kernel just by considering $\Phi(\boldsymbol{x}, \boldsymbol{y}) = \phi(\|\boldsymbol{x} - \boldsymbol{y}\|_2)$. We can then indifferently use Φ or ϕ by referring to Definition 1. An example of an RBF is the well-known Gaussian function $\phi(r) = \mathrm{e}^{-\varepsilon^2 r^2}$, where ε is the shape parameter, which is indeed a *scale parameter*.

A fundamental quantity for error estimation is the *power function*

$$P_{X_N,\Phi}(\boldsymbol{x}) := \sqrt{\Phi(\boldsymbol{x}, \boldsymbol{x}) - (\boldsymbol{b}(\boldsymbol{x}))^{\mathsf{T}} \mathsf{A}^{-1} \boldsymbol{b}(\boldsymbol{x})}, \tag{6}$$

where $\boldsymbol{b} = (\Phi(\cdot, \boldsymbol{x}_1), \ldots, \Phi(\cdot, \boldsymbol{x}_N))^{\mathsf{T}}$ and A is the kernel matrix defined above. Alternatively, the power function can be expressed using the cardinal functions u_k,

$$P_{X_N,\Phi}(x) = \sqrt{\Phi(x, x) - \sum_{k=1}^{n} \Phi(x, x_k)\, u_k(x)} \tag{7}$$

This function provides a pointwise bound for the interpolation error and plays a crucial role in greedy point selection strategies. The following formula, which I learned during my first visit, appeared in [16]. Letting A^y be the kernel matrix related to the augmented dataset $X_N \cup \{\boldsymbol{y}\}$, we can express the power function as

$$P_{X_N,\Phi}(\boldsymbol{y}) = \sqrt{\frac{\det \mathsf{A}^y}{\det \mathsf{A}}}. \tag{8}$$

Notice that the power function vanishes if it is evaluated at the interpolation nodes, as emphasized by the formulation in (8). Furthermore, we can provide the following error bound for the kernel-based interpolant in (5) (see e.g., [14, Section 14.4]).

Theorem 1. *Let $\Omega \subseteq \mathbb{R}^d$ and let $\Phi \in C(\Omega \times \Omega)$ be a strictly positive definite kernel. Moreover, let $X = \{\boldsymbol{x}_k, k = 1, \dots, N\} \subset \Omega$ be a set of distinct nodes. Then*

$$|f(\boldsymbol{x}) - P_f(\boldsymbol{x})| \leq P_{X_N,\Phi}(\boldsymbol{x}) \|f\|_{\mathcal{N}_\Phi(\Omega)}, \quad \boldsymbol{x} \in \Omega,$$

where $f \in \mathcal{N}_\Phi(\Omega)$.

Here, $\mathcal{N}_\Phi(\Omega)$ indicates the *Native space*, that is, the functional space which is the Reproducing Kernel Hilbert Space associated with Φ and $\|\cdot\|_{\mathcal{N}_\Phi(\Omega)}$ the corresponding norm inherited from inner product.

Despite the dependency on the power function, and thus on the chosen set of nodes, we observe that Theorem 1 provides a pointwise error bound that does not give any indication concerning the *mesh size*, a term used in abuse of notation in the scattered data interpolation context. To obtain further details in this direction, we need the *fill distance* (see e.g., [14, Section 14.1]), which is the radius of the maximal ball that doesn't contain centers

$$h_{X,\Omega} := \sup_{\boldsymbol{x}\in\Omega} \min_{\boldsymbol{x}_k\in X} \|\boldsymbol{x} - \boldsymbol{x}_k\|.$$

4. Greedy Algorithms and Optimal Centers

The "Fundamental question" stated in Question 1 was rephrased in the RBF framework as: *"Are there any special sets of centers for RBF's interpolation"?*. Our understanding of the problem led to the papers [16, 17]. The most influential algorithmic development was the *greedy strategy* based on the power function.

Starting from the set $X_1 = \{x_1\}$, $x_1 \in \Omega \subset \mathbb{R}^d$, the next point is selected as

$$x_{j+1} := \arg\max_{x\in\Omega} \{P_{X_j,\Phi}(x)\}, \ j \geq 1.$$

This approach produces well-distributed point sets and often achieves near-optimal approximation rates [17]. But it turned out that the greedy maximization of the power function adds a new point at the center of the largest hole in the data. This strategy is indeed *independent* of the kernel Φ. That is why we developed a new greedy method, called *geometric greedy* that, starting from $X_0 = \emptyset$ by defining the distance in Ω to X_0 as any positive value not smaller than $\mathtt{diam}(\Omega)$ (the diameter of Ω), the new point at the step j, is such that

$$x_{j+1} := \max_{x\in\Omega\setminus X_j} \{dist(x, X_j)\}, \ j \geq 1.$$

so that $X_{j+1} = X_j \cup \{x_{j+1}\}$.

This geometric greedy technique was shown to be similar to the construction of *Leja sequences* (cf. e.g. [18, 7]). Moreover, thanks to the identity (8), we obtained additional insight into the geometric interpretation of greedy point selection algorithms, since maximizing the power function corresponds to maximizing the growth of the determinant of the interpolation matrix. For a comparison of the two methods, we present in Figure 2 two pictures taken from [17]. From the pictures, it is evident that the geometric greedy produces points that are more uniformly distributed, especially given that the original discretization of $[-1, 1]^2$ was a grid.

The greedy algorithm was then named *P-greedy*, since it minimizes the power function as an error indicator and turned out to be independent of the target function. Later, in [19], a proof of the convergence error of the P-greedy approach was provided. More data-dependent greedy algorithms in kernel spaces, such as the f-greedy [20], $f \cdot P$ [21], and f/P [22], are known to provide fast-converging interpolants and are extremely easy to implement and efficient to run: their convergence rates were proved in [23]. All greedy algorithms share one common approach: minimizing some error indicator. In Table 1, we compile these greedy methods and their corresponding error indicators or *selection criteria*, which are optimized to generate the points. Concerning the convergence rates of all families of greedy methods, we refer the readers to the comprehensive paper [23], in which the authors

- firstly, defined a new scale of greedy algorithms by introducing a real parameter $\beta \in [0, \infty)$ for interpolation, that comprises all the existing ones in a unique analysis, where the degree of dependency of the selection criterion on the functional data was quantified;

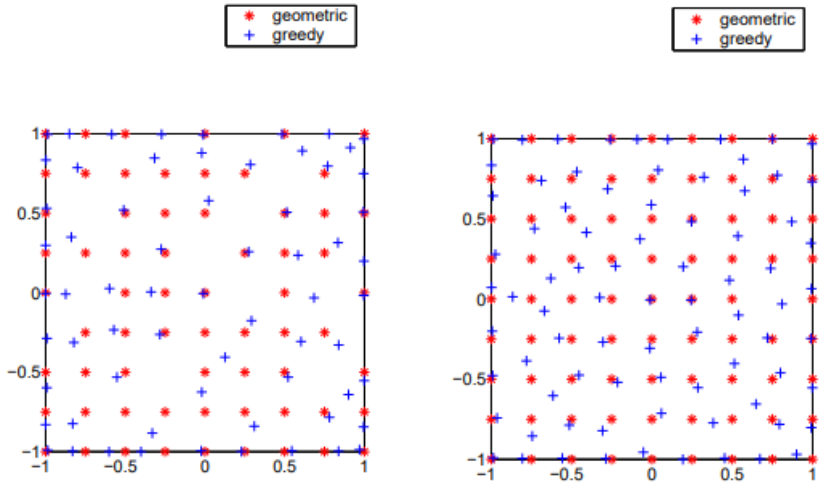

Figure 1. **Left**: 65 optimal points for the Gaussian with scale 1. **Right**: 80 points for Wendland's C^2. In both figures, the points are computed either by the geometric greedy algorithm (*) or the greedy algorithm (+).

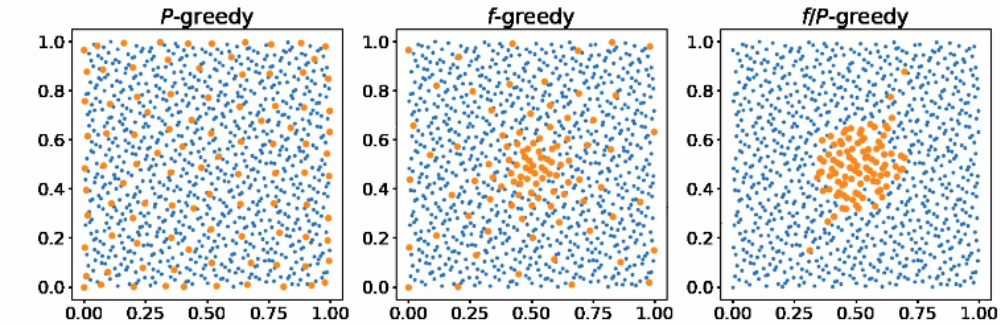

Figure 2. In orange we display 100 P, f and f/P greedy points extracted from 1000 Halton points, for the C^2 Matérn kernel and the function $f(x,y) = \|(x,y) - (0.5, 0.5)\|^{3/2}$

- then, they proved new convergence rates that account for this degree, showing that, up to a logarithmic factor, target data-dependent selection strategies yield faster convergence (see Figure 2, which illustrates the target data-dependent selection on the right)

It is worth mentioning that the convergence of generalized kernel-based interpolation methods under minimalistic assumptions on both the kernel and the target function has recently been proved by Albrect and Iske in [24], where they also prove the convergence of popular greedy data selection algorithms for totally bounded sets of sampling functionals. A typical implementation works as follows.

name	data-dependent	error indicator
$P-$greedy	no	P_n
$f-$greedy	yes	$\|E\| = \|f - P_f\|$
$f \cdot P$	yes	$\|E\| \cdot P_n$
f/P	yes	$\|E\|/P_n$

Table 1. Summarizing table of the greedy methods

```
- Choose an initial point x1
for n = 1,...,N
  - compute P_n(x) on candidate points
```

```
  - choose x_{n+1} maximizing
        P_n(x)            (P-greedy)
        |r_n(x)|          (f-greedy)
        |r_n(x)|/P_n(x)   (f/P-greedy)
   - update Cholesky factorization
end
```

We observe that, since we often deal with symmetric, positive-definite matrices, using it incremental Cholesky updates makes the computational cost of modifying them, without recomputing the entire factorization from scratch, roughly $O(N^2)$ instead of $O(N^3)$.

5. On the stability of kernel approximation

In [25], we resolved a fundamental "paradox" in kernel (RBF) interpolation: "Kernel interpolation may be numerically ill-conditioned, but it is stable at the functional level".

That is, even if the kernel matrix is badly conditioned, the interpolation operator is stable when analyzed in appropriate function spaces. Recalling that the power function controls the approximation error. In contrast, the Lebesgue function controls stability, and both depend on the geometry of X; we gave new insights into the stability of kernel-based interpolation.

Specifically, given data $f(x_j)$ at scattered points $X = \{x_1, \ldots, x_N\}$, letting the interpolant

$$s_{f,X}(x) = \sum_{j=1}^{N} f(x_j)\, u_j(x),$$

where u_j are the *cardinal functions* (see above (7)). The goal was to study the stability of the mapping from data values to the interpolant.

Theorem 2. *The interpolation at sufficiently many quasi-uniform points satisfies the following stability bounds:*

$$\|s_{f,X}\|_{L_\infty(\Omega)} \le C\,(\|f\|_{\ell_\infty(X)} + \|f\|_{\ell_2(X)}).$$

and

$$\|s_{f,X}\|_{L_2(\Omega)} \le h_{X,\Omega}^{d/2}\|f\|_{\ell_2(X)}$$

That was the key result showing that the interpolation operator is continuous and stable, with a stability constant C independent of X. Moreover, the Lagrange functions $u_j(x)$ are uniformly bounded. This prevents uncontrolled amplification of errors in the data. We showed that for quasi-uniform points, the Lebesgue constant $\Lambda_N = \max_{x\in\Omega} \sum_{j=1}^{N} |u_j(x)|$ has a growth like

$$\Lambda_N = O(\sqrt{N}). \tag{9}$$

This implies moderate error amplification, much better than the exponential growth seen in polynomial interpolation on equispaced points.

The choice of "good" centers is intrinsically related to the choice of the bases. We should remember that the most popular methods for the interpolant can be grouped into the following categories.

1. *RBF-QR methods*: it is rooted in a particular decomposition of the kernel, and it has been developed so far to treat the Gaussian kernel [26, 27, 28, 29].
2. *Hilbert-Schmidt Singular Value Decomposition (HS-SVD)*: it has been developed to stably compute the RBF interpolants [30, 15]. In principle, this technique can be applied to any kernel, provided that the HS eigenvalues and eigenvectors are known. However, these quantities are far from being easy to compute, and in practice, they only work for the Gaussian function.
3. *WSVD bases*: it is a more general approach that applies to any RBF, consisting of computing a weighted SVD decomposition which produces stable bases [31]. This paper stems from two important papers by Schaback on basis change. The first [32] introduced the so-called Newton basis that turns out to be orthogonal in the Hilbert space in which the kernel is reproducing. The second [33] among various results on bases change, presented an efficient algorithm for computing Newton's bases.

5.1. Numerical experiments

For kernels of *finite smoothness*, such as Matérn or Sobolev kernels, the experiments confirm the theoretical bound (9). However, numerical results indicate an even more favorable behavior: the Lebesgue constants appear to remain nearly uniformly bounded as the number of data points increases. Moreover, the maximum of the Lebesgue function is typically attained in the interior of the domain. This shows that in practice, kernel interpolation with finite-smoothness kernels is highly stable (see Figure 3, Left).

The situation changes significantly for *infinitely smooth kernels*, such as the Gaussian. The experiments reveal several important differences. In particular, the Lebesgue function develops large peaks near the boundary of the domain. The Lebesgue constant grows more rapidly and does not appear to be uniformly bounded. The interior of the domain remains relatively stable, while instability is concentrated near corners or edges. These results indicate that instability for Gaussian kernels is primarily a boundary phenomenon (see Figure 3, Right).

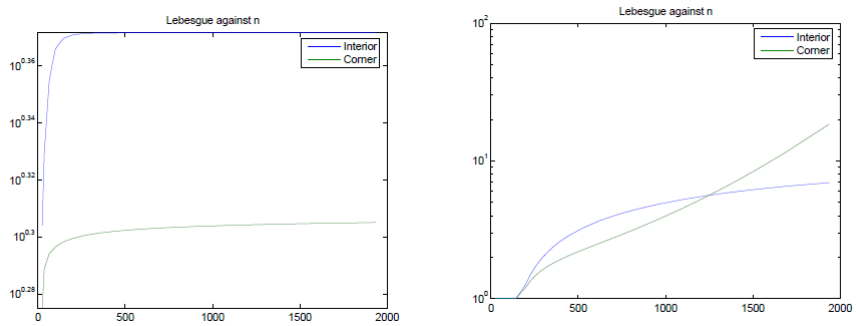

Figure 3. The growth of the Lebesgue constant at interior and boundary points. Left: for Sobolev/Matérn kernel. Right: for the Gaussian kernel

We can also investigate the effect of the *shape parameter*. The kernel scaling, typically introduced via a parameter c in $\Phi_c(x) = \Phi(x/c)$, plays a crucial role in stability.

For Gaussian kernels, we realized that increasing the scale leads to the so-called *flat limit*. In this regime, the kernel interpolant approaches polynomial interpolation, and the associated Lagrange basis functions become increasingly oscillatory and large. This explains why large-scale Gaussian kernels lead to severe numerical instability: polynomial interpolation on scattered points does not generally yield uniformly bounded basis functions.

The behavior of *cardinal functions* is also interesting. The shape of the Lagrange basis functions provides direct insight into stability. For Gaussian kernels, the basis functions become highly oscillatory and exhibit large amplitudes, especially near the boundary. For Matérn/Sobolev kernels, the basis functions remain well localized and uniformly bounded.

The experiments also compare interpolation on regular grids and scattered data sets. For finite smoothness kernels, the change from regular to scattered points does not significantly affect stability. The overall behavior remains consistent, confirming robustness with respect to point distribution. However, geometric considerations still play a role, particularly regarding boundary coverage. A major finding is the strong influence of *boundary effects on stability*. For Gaussian kernels, large errors and peaks of the Lebesgue function occur near domain boundaries and corners. The situation worsens when no data points are placed on the boundary (e.g., interpolation on a circle with interior points only). This highlights the importance of adequately sampling the boundary to ensure stability. To improve stability, adaptive point selection strategies have proven effective. In particular, the use of greedy algorithms based on maximizing the power function shows the following: a dramatic reduction in the peaks of the Lebesgue function, and a more uniform and stable distribution of interpolation points.

Interestingly, selecting points based on the Lebesgue function itself is less effective than using the power function.

Another effective stabilization strategy is oversampling near the boundary. Adding more points near the boundary significantly reduces instability in Gaussian kernels. This effect is particularly strong for analytic

kernels; for kernels with limited smoothness, it is essential when applied to nontrivial domains such as circles and cusped cardioids, whereas it is not when using highly smooth kernels.

Additional experiments have been conducted on nontrivial domains, such as circles and cusped cardioids, see Figures 4 and 5. The results confirm that the qualitative behavior observed in simpler domains persists. Kernel smoothness remains the dominant factor, and boundary effects continue to play a crucial role.

In conclusion, the numerical experiments strongly support the theoretical findings on the stability of kernel-based interpolation, which can be summarized as follows: (i) kernels of finite smoothness provide stable and reliable interpolation; (ii) infinitely smooth kernels, while powerful, require careful handling; (iii) instability is closely linked to boundary effects and kernel scaling; (iv) adaptive point selection and boundary oversampling are effective stabilization techniques. Overall, we demonstrated that stability is governed not only by the ill-conditioning of the kernel matrix but also by the behavior of the Lebesgue function, the choice of kernel, and the geometry of the data points.

Interested people can play with the Matlab code available on GitHub at the link `https//github.com/demarchi17/StabilityKernels`.

Example 1. *We considered a cardioid domain and selected* 104 *nearly uniform and P-greedy points for the Wendland* C^2 *kernel with a scale parameter of* 30*, shown in Figure 4. The corresponding Lebesgue functions are depicted in Figure 5. Some results done later in the Ph.D. thesis by Christian Rieger (University of Göttingen 2008 [34]), suggested that an* $O(h^2)$ *oversampling in a strip close to the boundary has a positive effect on the Lebesgue constant growth.*

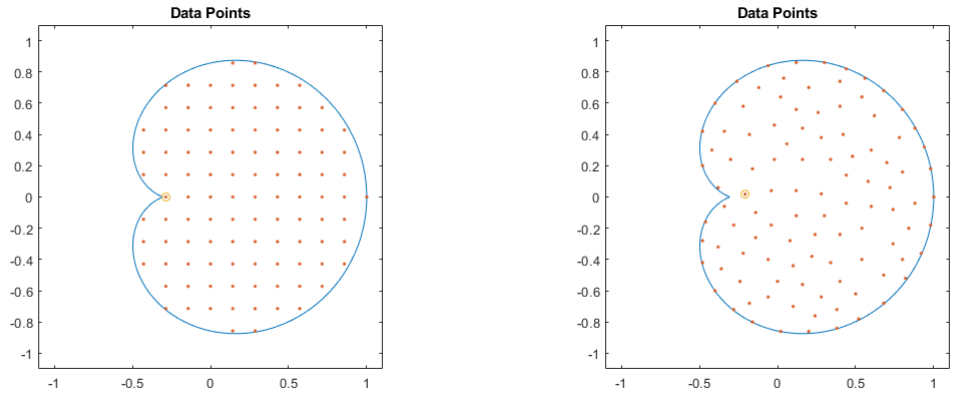

Figure 4. The cardiod with 104 points uniformly and P-greedy distributed

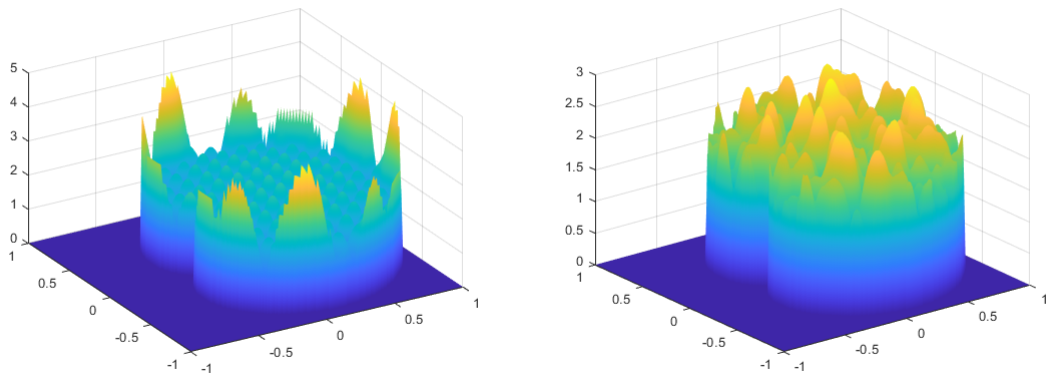

Figure 5. The Lebesgue function on the cardiod with 104 points uniformly and P-greedy distributed

6. Rational RBF Interpolants

To improve the computational stability of kernel-based approximations, we introduced the framework of *rational radial basis functions* (RRBFs). The idea was suggested by the *Rescaled Localized RBF method* presented by [35] for solving PDEs on non-cartesian and non-conforming grids.

These methods extend classical kernel interpolation by considering nonlinear approximants expressed as ratios of RBF expansions. In particular, we focus on the *eigen-rational kernel-based scheme*, where the denominator is constructed via a suitable eigenvector of the kernel matrix (cf. [36, 37]).

Definition

Let $X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subset \Omega \subset \mathbb{R}^d$ and let Φ be a strictly positive definite kernel. A rational RBF approximant is defined as

$$\mathscr{R}(\boldsymbol{x}) = \frac{R^{(1)}(\boldsymbol{x})}{R^{(2)}(\boldsymbol{x})} = \frac{\displaystyle\sum_{i=1}^{N} \alpha_i \Phi(\boldsymbol{x}, \boldsymbol{x}_i)}{\displaystyle\sum_{j=1}^{N} \beta_j \Phi(\boldsymbol{x}, \boldsymbol{x}_j)}, \tag{10}$$

provided that $R^{(2)}(\boldsymbol{x}) \neq 0$ for all $\boldsymbol{x} \in \Omega$. The coefficients $\boldsymbol{\beta}$ defining the denominator are obtained by solving the generalized eigenvalue problem (cf. e.g. [36])

$$\Lambda \boldsymbol{q} = \lambda \Theta \boldsymbol{q}, \tag{11}$$

where

$$\Lambda = \frac{1}{\|\boldsymbol{f}\|_2^2} D^T A^{-1} D + A^{-1}, \qquad \Theta = \frac{1}{\|\boldsymbol{f}\|_2^2} D^T D + I_N.$$

Here, A is the kernel matrix and $D = \operatorname{diag}(f_1, \ldots, f_N)$. The denominator $R^{(2)}$ is then defined as the RBF interpolant of the eigenvector $\boldsymbol{q}$.

We provided some error estimates. Let $\mathcal{H}_\Phi(\Omega)$ denote the RKHS associated with Φ, and as before $P_{X,\Phi}$ be the power function. Then, under suitable regularity assumptions, the RRBF approximation satisfies the estimate

$$|f(\boldsymbol{x}) - \mathscr{R}(\boldsymbol{x})| \leq \frac{C}{|R^{(2)}(\boldsymbol{x})|} P_{X,\Phi}(\boldsymbol{x}) \, \|f\|_{\mathcal{H}_\Phi}, \tag{12}$$

for all $\boldsymbol{x} \in \Omega$, where $C > 0$ is a constant independent of f.

Theorem 3 (Stability of RRBF Interpolation)**.** *Assume that, Φ is strictly positive definite, the point set X is quasi-uniform, and the denominator satisfies* $\inf_{\boldsymbol{x}\in\Omega} |R^{(2)}(\boldsymbol{x})| \geq c_0 > 0$*. Then, the Rational RBF interpolation operator*

$$T_X^{\mathrm{rat}} : \ell^\infty(X) \to L^\infty(\Omega)$$

is bounded, and

$$\|T_X^{\mathrm{rat}}\| \leq \frac{\Lambda_X}{c_0}, \tag{13}$$

where Λ_X is the Lebesgue constant of the corresponding RBF space.

Sketch of proof. The proof follows from writing the rational interpolant in Lagrange form:

$$\mathscr{R}(\boldsymbol{x}) = \frac{\sum_{j=1}^{N} f(\boldsymbol{x}_j) u_j(\boldsymbol{x})}{R^{(2)}(\boldsymbol{x})}.$$

Taking absolute values yields

$$|\mathscr{R}(\boldsymbol{x})| \leq \frac{\sum_{j=1}^{N} |f(\boldsymbol{x}_j)| \, |u_j(\boldsymbol{x})|}{|R^{(2)}(\boldsymbol{x})|} \leq \frac{\Lambda_X}{c_0} \|f\|_{\ell^\infty(X)}.$$

□

The effective Lebesgue function of the RRBF scheme can be written as

$$\widetilde{\Lambda}_X(\boldsymbol{x}) = \frac{\Lambda_X(\boldsymbol{x})}{|R^{(2)}(\boldsymbol{x})|}.$$

This shows that the denominator acts as a *stabilizing factor*, reducing error amplification in regions where Λ_X is large (e.g. near boundaries or in the flat kernel regime).

When combined with RRBFs, greedy sampling further improves the denominator behavior, leading to a better conditioning of the rational scheme, reduced peaks in $\widetilde{\Lambda}_X$, and improved boundary resolution.

With Martin Buhmann and Emma Perracchione [36], we studied an *eigen-rational kernel-based scheme* that was introduced for multivariate interpolation within a meshfree framework. Let

$$\Phi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$$

be a *conditionally positive definite* kernel and denote by $\bar{\Phi}$ its associated strictly positive definite kernel. For instance, if Φ is a generalized multiquadric of order 2, then $\bar{\Phi}$ can be chosen as the inverse multiquadric. In the strictly positive definite case, we simply take $\bar{\Phi} = \Phi$, as for example with Matérn kernels.

Given a dataset $\mathcal{X}_N = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subset \Omega$ and corresponding function values $\mathcal{F}_N = \{f_1, \ldots, f_N\}$, we define the rational interpolant

$$\hat{P}_f(\boldsymbol{x}) = \frac{\displaystyle\sum_{i=1}^{N} \alpha_i \Phi(\boldsymbol{x}, \boldsymbol{x}_i) + \sum_{m=1}^{L} \gamma_m p_m(\boldsymbol{x})}{\displaystyle\sum_{k=1}^{N} \beta_k \bar{\Phi}(\boldsymbol{x}, \boldsymbol{x}_k)} = \frac{P_g(\boldsymbol{x})}{P_h(\boldsymbol{x})}, \tag{14}$$

where P_g and P_h are kernel-based interpolants constructed from auxiliary data g_i and h_i, respectively.

To get a well-defined interpolant, we proceed as follows. Once values h_i are prescribed, we define $g_i = f_i h_i$ and construct P_g as the kernel interpolant of $\boldsymbol{g} = (g_1, \ldots, g_N)^T$. By construction, the rational function $\hat{P}_f$ interpolates the data, i.e.

$$\hat{P}_f(\boldsymbol{x}_i) = f_i, \quad i = 1, \ldots, N.$$

The use of a strictly positive definite kernel in the denominator ensures that $P_h(\boldsymbol{x}) \neq 0$ for all $\boldsymbol{x} \in \Omega$, so making the interpolant well-defined.

Proposition 1. *The eigen-rational interpolation method can be interpreted as a rational RBF expansion, where the denominator depends on the largest eigenvalue of the kernel matrix.*

Proof. We express the interpolant in terms of cardinal functions. Let $\{u_j\}_{j=1}^N$ and $\{\bar{u}_j\}_{j=1}^N$ be the cardinal bases associated with Φ and $\bar{\Phi}$, respectively. Then

$$P_g(\boldsymbol{x}) = \sum_{j=1}^{N} g_j u_j(\boldsymbol{x}), \quad P_h(\boldsymbol{x}) = \sum_{j=1}^{N} h_j \bar{u}_j(\boldsymbol{x}).$$

Substituting into (14), we obtain

$$\hat{P}_f(\boldsymbol{x}) = \frac{\sum_{j=1}^{N} h_j f_j u_j(\boldsymbol{x})}{\sum_{k=1}^{N} h_k \bar{u}_k(\boldsymbol{x})}.$$

This can be rewritten as

$$\hat{P}_f(\boldsymbol{x}) = \sum_{j=1}^{N} f_j \hat{u}_j(\boldsymbol{x}), \quad \hat{u}_j(\boldsymbol{x}) := \frac{h_j u_j(\boldsymbol{x})}{\sum_{k=1}^{N} h_k \bar{u}_k(\boldsymbol{x})}.$$

It follows immediately that $\hat{u}_j(\boldsymbol{x}_i) = \delta_{ij}$, hence the interpolatory property holds.

If $\Phi = \bar{\Phi}$ is strictly positive definite, the functions $\{\hat{u}_j\}$ form a partition of unity. Indeed,

$$\sum_{j=1}^{N} \hat{u}_j(\boldsymbol{x}) = \frac{\sum_{j=1}^{N} h_j \bar{u}_j(\boldsymbol{x})}{\sum_{k=1}^{N} h_k \bar{u}_k(\boldsymbol{x})} = 1.$$

To derive stability estimates, we consider

$$\sum_{j=1}^{N} \hat{u}_j(\boldsymbol{x})^2 = \sum_{j=1}^{N} \left(\frac{h_j u_j(\boldsymbol{x})}{P_h(\boldsymbol{x})} \right)^2 \leq \frac{\|\boldsymbol{h}\|_\infty^2}{P_h(\boldsymbol{x})^2} \sum_{j=1}^{N} u_j(\boldsymbol{x})^2.$$

From [13, Theorem 12.1], we have

$$1 + \sum_{j=1}^{N} u_j(\boldsymbol{x})^2 \leq \frac{\mathcal{P}^2_{K,X_N}(\boldsymbol{x})}{\omega}, \tag{15}$$

where $\omega > 0$ is the smallest eigenvalue of the kernel matrix associated with $X_N \cup \{\boldsymbol{x}\}$. This provides a bound for the rational basis functions and completes the proof. □

Classical estimates in terms of Lebesgue constants and convergence rates with respect to the fill distance show that eigen-rational interpolation behaves comparably to standard kernel-based methods: see [36] for a detailed analysis and examples.

7. Variably Scaled Kernels

Variably Scaled Kernels (VSK), first introduced in the seminal paper by Bozzini et al. [38], incorporate spatially dependent scaling functions into the kernel definition.

Definition 2. *Let $\Phi : \mathbb{R}^{(d+1)\times(d+1)} \longrightarrow \mathbb{R}$ be a continuous strictly positive definite radial kernel and let $\psi : \mathbb{R}^d \longrightarrow \mathbb{R}$ be a scaling function. A variably scaled kernel Φ^Ψ on $\mathbb{R}^{d\times d}$ is defined as*

$$\Phi^\Psi(\boldsymbol{x}, \boldsymbol{y}) = \Phi((\boldsymbol{x}, \psi(\boldsymbol{x})), (\boldsymbol{y}, \psi(\boldsymbol{y}))), \tag{16}$$

for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$.

The shape parameter of the underlying kernel Φ in Definition 2 is chosen for simplicity equal to 1. For simplicity, we may omit this information from the notation to avoid complicated expressions.

Remark 7.1. *The VSK for radial kernels ϕ becomes*

$$\Phi^\Psi(\boldsymbol{x}, \boldsymbol{y}) = \phi\left(\sqrt{\|\boldsymbol{x} - \boldsymbol{y}\|^2 + |\psi(\boldsymbol{x}) - \psi(\boldsymbol{y})|^2} \right),$$

where ϕ is the usual univariate function related to Φ. Hence, if Φ is radial on $\mathbb{R}^{(d+1)\times(d+1)}$ so is Φ^Ψ on $\mathbb{R}^{d\times d}$. Moreover, it comes easily to see that if Φ is (strictly) positive definite, so is Φ^Ψ.

By considering the map $\Psi(\boldsymbol{x}) = (\boldsymbol{x}, \psi(\boldsymbol{x}))$ on Ω and in analogy with (5), we can express the VSK interpolant at the nodes $\tilde{X} = \{(\boldsymbol{x}_k, \psi(\boldsymbol{x}_k)),\ \boldsymbol{x}_k \in X\}$ as

$$R(\Psi(\boldsymbol{x})) = \sum_{k=1}^{N} c_k \Phi(\Psi(\boldsymbol{x}), \Psi(\boldsymbol{x}_k)), \tag{17}$$

with $\boldsymbol{x} \in \Omega,\ \boldsymbol{x}_k \in X$. Therefore, in order to obtain a VSK interpolant R^Ψ at the set of nodes X on Ω, it is sufficient to project back the interpolant in (17), that is,

$$R^\Psi(\boldsymbol{x}) = \sum_{k=1}^{N} c_k \Phi^\Psi(\boldsymbol{x}, \boldsymbol{x}_k) = \sum_{k=1}^{N} c_k \Phi(\Psi(\boldsymbol{x}), \Psi(\boldsymbol{x}_k)) = R(\Psi(\boldsymbol{x})).$$

An important consequence of this construction is that the analysis of the variably scaled setting is fully understood in terms of the analysis of the underlying standard kernel (which is composed with Ψ).

The following proposition in [38] states some additional properties of VSKs, which are fundamental to understanding why VSKs improve the stability of the approximation process.

Proposition 2. *Let Φ, Φ^Ψ, and ψ be as in Definition 2. We have*

(i) If Φ and ψ are continuous, so is Φ^ψ. Moreover, if $\psi : \Omega \to \psi(\Omega)$ is a bijection, then Φ^ψ inherits the positiveness properties of Φ.

*(ii) Let q_X be the separation distance between the centers (see Definition **??** above), we have*

$$q_X \leq q_{\Psi(X)}$$

for any choice of the scaling map ψ. Indeed, by the definition of the Euclidean norm, we have

$$\|\Psi(\boldsymbol{x}_i) - \Psi(\boldsymbol{x}_j)\|^2 = \|\boldsymbol{x}_i^2 - \boldsymbol{x}_j^2\|^2 + (\psi(\boldsymbol{x}_i) - \psi(\boldsymbol{x}_j))^2 \leq \|\boldsymbol{x}_i^2 - \boldsymbol{x}_j^2\|^2(1+L)^2,$$

where L is the Lipschitz constant of ψ.

(iii) Let $G_\Psi(\Omega) = \{(\boldsymbol{x}, \psi(\boldsymbol{x})) \,|\, \boldsymbol{x} \in \Omega\} \subset \Omega \times \mathbb{R}$ be the graph of the scaling function ψ. Then, the native spaces $\mathcal{N}_\Phi(G_\Psi(\Omega))$ and $\mathcal{N}_{\Phi^\Psi}(\Omega)$ are isometrically isomorphic (cf. [Th. 2, Bozzini15]).

As a consequence of Proposition 2 (ii), the variably scaled setting might improve the stability of the interpolation process by increasing the separation distance. The interpolation via Variably Scaled Kernels (VSKs) depends on the definition of an appropriate scaling function, but no fixed theoretical or numerical recipes for its construction have been provided. Recently, in [39], a user-independent tool for learning the scaling function using Discontinuous Neural Networks (called δ-NN) was presented, which partially addresses this gap.

This idea has been extended to *Variably Scaled Discontinuous Kernels (VSDK)*, simply by taking a discontinuous scaling function, which has been successfully applied to problems in image reconstruction and edge detection [40, 41]. Moreover, one can plug in the appropriate scaling function to the eigen-rational kernel-based interpolant, getting a *Variably Scaled Rational Kernel*, but this is the topic of a forthcoming paper. Here we illustrate how a VSDK works with an example.

Example 2. *Let $f : \Omega \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ be defined as*

$$f(x,y) := \begin{cases} x + y^2 + 0.5, & 2x^2 + 3y^2 \leq 0.5, \\ 0.5, & y \geq x^3 + 1, \\ -(x+y)/8, & \textit{otherwise.} \end{cases}$$

As the interpolation set, we consider the equispaced grid in Ω with a step size of $1/13$ and the C^2 Wendland kernel. We interpolated f using classical RBFs on the grid, tuning the shape parameter via LOOCV. Then, we considered the VSDK setting by choosing the discontinuous scaling map

$$\psi(x,y) := \begin{cases} 2, & 2x^2 + 3y^2 \leq 0.5, \\ 1, & y \geq x^3 + 1, \\ 0, & \textit{otherwise.} \end{cases} \tag{18}$$

Denoting by s the interpolant obtained by standard RBF interpolation and by s^Ψ the VSDK interpolant, in Table 2 we report the Maximal Approximation Error (MAE) and the RMSE as well as the SSIM (Structural SIMilarity) index (cf. [42]), which is widely used to measure the similarity between two images: the SSIM index takes values in $(-1, 1]$, with larger values indicating better quality. It is evident that the VSDK interpolant with scaling function (18) is much more accurate than the standard RBF interpolant. The reconstructions are shown in Figure 6.

Table 2. Error results of standard and VSDK interpolation of f_2.

	RBF	VSDK
MAE	1.04e + 0	1.31e − 2
RMSE	7.50e − 2	7.27e − 4
SSIM	0.861	0.999

Figure 6. Standard and VSDK interpolation of the function f. Left: the function f. Middle: reconstruction with the classical RBF with the best shape parameter. Right: VSDK reconstruction

8. Concluding Remarks

The ideas developed within the RBF community, which has in Robert Schaback one of the main references, have profoundly influenced modern approximation theory and kernel-based numerical methods. Many of these developments can be traced back to the theoretical insights and collaborative spirit fostered by Robert Schaback and his research group, illustrating the remarkable versatility of RBF methods. Kernel-based approximation methods now play a central role in many fields, including numerical solution of PDEs, machine learning and support vector machines, radial basis neural networks, topological data analysis, and medical and scientific imaging. For all of these fields, one can find contributions in Robert's list of publications.

I cannot conclude without quoting some anecdotes and special suggestions he gave me, and I believe he also gave them to other close collaborators. Research is very important to him, as are relationships. During my visits, we took many bike tours around Göttingen, which gave me a view of the surrounding fields, forests, and natural landscape. Staying in nature is fundamental because mathematics finds inspiration in looking at the geometry and patterns behind landscapes, shapes, flowers, and roads. I think that the spirit to stay in nature that Robert transferred to my already big passion for mountains allowed me to dig into the world of RBFs deeper and deeper. I realized that the shapes behind radial basis functions remind many of natural objects. Think of the bell-shaped PD functions, or the "sombrero" shape of some CPD, etc. But this is not the only message that Robert left to me. More important is his view of the University environment. He repeatedly told me that a good researcher should do 3 fundamental things, with "equally distributed" attention: research, didactics, and bureaucracy. And also in Hong Kong, when he did his presentation, he insisted that attending conferences is important for learning and sharing new ideas. And I completely agree with him. Perhaps because we are both Sagittariuses, we have a positive view of life and often dream of changing the world. Last but not least, Robert keeps relationships alive. For many years, I received his Season's greetings with sentiments like *"for you and your families, I hope for a merry Christmas and a new year with plenty of success, minimal stress, and perfect health. Greetings from Göttingen"*, and he used to add a photo of the `Göttingen Altes Rathaus Weihnachtsmarkt` or, once in a while, a collage from his holiday with Gudrun. There are not many photos of Robert online, which is why during the meeting on the occasion of his 80th birthday in Hong Kong, I took the photo in Figure 7, where he is with some of his co-authors: left to him Elisabeth Larsson (Uppsala), me and Maryam Mohammadi (Karazmi-Tehran); right to Robert there is Leevan Ling (Hong Kong) in between two his students.

I have no words to express my gratitude for having met a colleague who has been not only a mentor but also a friend. The sentence "who met a friend has found a treasure" is completely meaningful when talking about Robert.

Acknowledgments. This work was carried out within the thematic group on "Approximation Theory and Applications" of the Unione Matematica Italiana (UMI) and with the support of the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 104 published on 2.2.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union – NextGenerationEU– Project Title "Computational mEthods for Medical Imaging (CEMI)" – CUP D53D23005830006 - Grant Assignment Decree No. 973 adopted on 30.6.2023 by the Italian Ministry of University and Research (MUR). I gratefully acknowledge Robert Schaback, the German school, for many years of scientific inspiration, collaboration, and friendship. Also, I am grateful to the "Padova RBF team," which has grown over the decades and made significant contributions to the RBF community.

Figure 7. Photo taken at Hong Kong conference, 5th January 2026

References

[1] A. Iske, Optimal distribution of centers for radial basis function methods, Tech. Rep. M0004, Tech. University Munchen (2000).
[2] R. Schaback, H. Wendland, Kernel techniques: from machine learning to meshless methods 15 (2006) 543–639.
[3] W. Cheney, W. Light, A course on Approximation Theory, Graduate Studies in Mathematics, vol. 101, American Mathematical Society, 2000.
[4] L. Bos, On certain configurations of points in rn which are unisolvent for polynomial interpolation, Journal of Approximation Theory 64 (3) (1991) 271–280.
[5] M. Caliari, S. De Marchi, M. Vianello, Bivariate polynomial interpolation on the square at new nodal sets, Applied Mathematics and Computation 165 (2) (2005) 261–274.
[6] L. Bos, M. Caliari, S. De Marchi, M. Vianello, Y. Xu, Bivariate lagrange interpolation at the padua points: the generating curve approach, Journal of Approximation Theory 143 (2006) 15–25.
[7] L. Bos, S. De Marchi, A. Sommariva, M. Vianello, Computing multivariate fekete and leja points by numerical linear algebra, SIAM Journal on Numerical Analysis 48 (5) (2010) 1984–1999.
[8] W. Erb, C. Kaethner, P. Dencker, M. Ahlborg, A survey on bivariate lagrange interpolation on lissajous nodes, Dolomites Research Notes on Approximation 8 (2015) 23–36.
[9] W. Erb, C. Kaethner, M. Ahlborg, T. M. Buzug, Bivariate lagrange interpolation at the node points of non-degenerate lissajous nodes, Numerische Mathematik 133 (1) (2016) 685–705.
[10] W. Erb, Bivariate lagrange interpolation at the node points of lissajous curves - the degenerate case, Applied Mathematics and Computation 289 (2016) 409–425.
[11] M. Dubiner, The theory of multi-dimensional polynomial approximation, J. Anal. Math. 67 (1995) 39–116.
[12] L. N. Trefethen, Approximation Theory and Approximation Practise, SIAM, 2013.
[13] H. Wendland, Scattered Data Approximation, Vol. 17 of Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, 2005.
[14] G. E. Fasshauer, Meshfree Approximation Methods with MATLAB, World Scientific, Singapore, 2007.
[15] G. E. Fasshauer, M. J. McCourt, Kernel-based Approximation Methods Using MATLAB, World Scientific, Singapore, 2015.
[16] S. De Marchi, On optimal center locations for radial basis function interpolation: computational aspects, Rendiconti del Seminario Matematico Universitá Politecnica di Torino 61 (2003) 343–358.
[17] S. De Marchi, R. Schaback, H. Wendland, Near-optimal data-independent point locations for radial basis function interpolation, Advances in Computational Mathematics 23 (2005) 317–330.
[18] S. De Marchi, On leja sequences: some results and applications, Applied Mathematics and Computation 152 (3) (2004) 621–647.
[19] G. Santin, B. Haasdonk, Convergence rate of the data-independent p-greedy algorithm in kernel-based approximation, Dolomites Research Notes on Approximation 10 (2017) 68–78.
[20] H. W. R. Schaback, Adaptive greedy techniques for approximate solution of large rbf systems, Numer. Algorithms 24(3) (2000) 239–254.
[21] S. Dutta, M. W. Farthing, E. Perracchione, G. Savant, M. Putti, A greedy non-intrusive reduced order model for shallow water equations, Journal of Computational Physics 439 (2021) 110378.
[22] S. Müller, Complexity and stability of kernel-based reconstructions, Ph.D. thesis, Fakultät für Mathematik und Informatik, Georg-August-Universität Göttingen (2009).

[23] T. Wenzel, G. Santin, B. Haasdonk, Analysis of target data-dependent greedy kernel algorithms: Convergence rates for f-, $f \cdot p$- and f/p-greedy, Constructive Approximation 57 (2023) 45–74.
[24] K. Albrecht, A. Iske, On the convergence of generalized kernel-based interpolation by greedy data selection algorithms, BIT Numerical Mathematics 65 (1). doi:10.1007/s10543-024-01048-3.
[25] S. De Marchi, R. Schaback, Stability of kernel-based interpolation, Adv. Comput. Math. 32.
[26] F. B., P. C, A stable algorithm for flat radial basis functions on a sphere, SIAM J. Sci. Comput. 30 (2007) 60–80.
[27] B. Fornberg, E. Larsson, N. Flyer, Stable computations with gaussian radial basis functions, SIAM Journal on Scientific Computing 33 (2) (2011) 869–892.
[28] G. E. Fasshauer, M. J. McCourt, Stable evaluation of gaussian radial basis function interpolants, SIAM Journal on Scientific Computing 34 (2) (2012) A737–A762.
[29] K. Kormann, C. Lasser, A. Yurova, Stable interpolation with isotropic and anisotropic gaussians using hermite generating function, SIAM Journal on Scientific Computing 41 (6) (2019) A3839–A3859.
[30] R. Cavoretto, G. E. Fasshauer, M. McCourt, An introduction to the hilbert-schmidt svd using iterated brownian bridge kernels, in: Numerical Algorithms, Vol. 68, 2015, pp. 393–422.
[31] S. De Marchi, G. Santin, A new stable basis for radial basis function interpolation, Journal of Computational and Applied Mathematics 253 (2013) 1–13.
[32] S. Müller, R. Schaback, A newton basis for kernel spaces, Journal of Approximation Theory 161 (2) (2009) 645–655.
[33] Bases for kernel-based spaces, J. Comput. Appl. Math. 236 (4) (2011) 575–588, international Workshop on Multivariate Approximation and Interpolation with Applications (MAIA 2010).
[34] C. Rieger, Sampling inequalities and applications, Ph.D. thesis, Georg-August-Universität zu Göttingen (2008).
[35] S. Deparis, D. Forti, A. Quarteroni, A rescaled localized radial basis function interpolation on non-cartesian and nonconforming grids, SIAM Journal on Scientific Computing 36 (6) (2014) A2745–A2762.
[36] M. D. Buhmann, S. De Marchi, E. Perracchione, Analysis of a new class of rational rbf expansions, IMA Journal of Numerical Analysis 40 (3) (2020) 1972–1993.
[37] S. De Marchi, A. Martínez, E. Perracchione, Fast and stable rational rbf-based partition of unity interpolation, Journal of Scientific Computing 79 (2019) 321–344.
[38] M. Bozzini, L. Lenarduzzi, M. Rossini, R. Schaback, Interpolation with variably scaled kernels, IMA Journal of Numerical Analysis 35 (2015) 199–219.
[39] G. Audone, F. Della Santa, E. Perracchione, S. Pieraccini, A recipe for learning variably scaled kernels via discontinuous neural networks, J. Comput. Appl. Math. 469 (2025) 1–18.
[40] S. De Marchi, F. Marchetti, E. Perracchione, Jumping with variably scaled discontinuous kernels (vsdks), BIT Numerical Mathematics 60 (2) (2020) 441–463.
[41] S. De Marchi, W. Erb, F. Marchetti, E. Perracchione, M. Rossini, Shape-driven interpolation with discontinuous kernels: Error analysis, edge extraction and applications in mpi, SIAM Journal on Scientific Computing 42 (2) (2020) B472–B491.
[42] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Transactions on Image Processing 13 (4) (2004) 600–612.