Stable and accurate least squares radial basis function approximations on bounded domains

Ben Adcock, Daan Huybrechs and Cécile Piret

November 24, 2022

Abstract

The computation of global radial basis function (RBF) approximations requires the solution of a linear system which, depending on the choice of RBF parameters, may be ill-conditioned. We study the stability and accuracy of approximation methods using the Gaussian RBF in all scaling regimes of the associated shape parameter. The approximation is based on discrete least squares with function samples on a bounded domain, using RBF centers both inside and outside the domain. This results in a rectangular linear system. We show for one-dimensional approximations that linear scaling of the shape parameter with the degrees of freedom is optimal, resulting in constant overlap between neighbouring RBF’s regardless of their number, and we propose an explicit suitable choice of the proportionality constant. We show numerically that highly accurate approximations to smooth functions can also be obtained on bounded domains in several dimensions, using a linear scaling with the degrees of freedom per dimension. We extend the least squares approach to a collocation-based method for the solution of elliptic boundary value problems and illustrate that the combination of centers outside the domain, oversampling and optimal scaling can result in accuracy close to machine precision in spite of having to solve very ill-conditioned linear systems.

1 Introduction

1.1 Radial basis function approximations

Radial basis functions provide a convenient way to approximate functions in one or more dimensions based on translates of a single function $\phi(r)$. The argument $r = \|x - \xi\|$ is the radial (Euclidean) distance between an evaluation point $x$ and a center $\xi$. The function $\phi$ itself is called the radial basis function (RBF) (or radial function). By covering a domain with centers, functions can be approximated using linear combinations of the associated centered RBFs and this approximation is traditionally obtained by interpolation. A wide variety of radial basis functions exist with very different properties: for example, $\phi(r)$ may be smooth or non-smooth at $r = 0$, and it may be decaying or increasing for $r \to \infty$. The most commonly used smooth radial functions are listed in Table 1. Note that they all admit a shape parameter $\varepsilon$ that controls their shape. This leads to an approximation of the form

$$f(x) \approx \sum_{n=1}^{N} \lambda_n \phi(\varepsilon(x - \xi_n)).$$

(1)

Small values of the shape parameter dilate the RBF: the limit $\varepsilon \to 0$ is often referred to as the flat limit. The choice of an appropriate shape parameter has significant consequences for
both accuracy and stability of corresponding numerical methods. Most crucial is to decide how to adjust the value of $\varepsilon$ when the number of centers, $N$, grows. One can either keep $\varepsilon$ fixed, or make it grow with $N$, [24]. In this study, we will focus on sublinear and linear scalings of $\varepsilon$ with $N$. That is, we will consider respectively $\varepsilon = cN^\alpha$, for $0 < \alpha < 1$, and $\varepsilon = cN$, for some positive constant $c$. 

| Name of RBF | Abbreviation | Definition |
|-------------|--------------|------------|
| Gaussian | GA | $e^{-(\varepsilon r)^2}$ |
| Multiquadric | MQ | $\sqrt{1 + (\varepsilon r)^2}$ |
| Inverse Quadratic | IQ | $\frac{1}{1+(\varepsilon r)^2}$ |
| Inverse Multiquadric | IMQ | $\frac{1}{\sqrt{1+(\varepsilon r)^2}}$ |

Table 1: Definitions of some common radial functions $\phi(\varepsilon r)$.

1.2 Existing results and analysis

Generally speaking, smaller values of $\varepsilon$ lead to higher approximation accuracy, but worse condition numbers. Larger values of $\varepsilon$ have better conditioning, but yield less accuracy or even, in the case of bounded domains, no accuracy at all. The trade-off between small and large shape parameters was referred to in [27] as an uncertainty principle. However, this tradeoff was shown to originate not from the interpolation problem, which is well posed, but from numerically solving the problem via the standard representation [1], having to multiply large oscillatory expansion coefficients $\lambda_n$ with flat radial function translates $\phi(\varepsilon(x - \xi_n))$, in essence a sort of removable singularity [9, 12].

This has spurred a quest for algorithms to analytically remove the ill-conditioning of linear systems associated with small shape parameter values. For instance, the Contour-Padé method extends $\varepsilon$ to the complex plane and makes use of numerical contour integration to stably compute the RBF interpolant [17]. The RBF-QR method induces a change of bases via a QR factorization, exposing the problematic terms and removing them (and with them the ill-conditioning) altogether [15, 16, 14]. Other such algorithms have been introduced, since keeping a small fixed shape parameter has often been the preferred strategy (see [28, 11, 26, 12]).

Indeed, it is observed that the linear scaling regime mentioned above often yields unstable methods, [3]; however some strategies have been introduced to attain accurate approximations in this regime [22, 23]. The remaining error is known as the stagnation or saturation error. The lack of accuracy in the linear scaling regime can be avoided by using slower than linear growth such as $\varepsilon = c\sqrt{N}$, although the procedure is still unstable [25, 24]. Alternatively, the ill-conditioning of global RBF approximation systems can be avoided altogether by considering local approximations of small size only, for example as part of the computation of a local stencil for a finite difference method. Finally, the approximation accuracy can be improved by augmenting the set of centered RBFs with other basis functions, such as polynomials [10].

1.3 Contribution of the paper

We describe and analyze a numerical method which in the linear scaling regime $\varepsilon \sim cN$ is numerically stable for approximation on bounded domains. We fully quantify the setting for
the specific choice of the Gaussian RBF (see (5) below), which is smooth and rapidly decaying as \( r \to \infty \). The setting crucially relies on the combination of the following points:

- accuracy in the linear scaling regime on bounded domains is improved by adding centers outside the domain,
- numerical stability of the computations is achieved by considering discrete least squares approximation instead of interpolation, i.e., by using more sample points than centers,
- in addition the sampling points may in general be different from the centers.

These points were inspired by the analysis of function approximation in the presence of redundancy using frames in [2, 3].

These points have appeared in the literature before in the context of radial basis functions. Using least squares approximation on bounded domains with centers outside the domain was recently explored in [18, 19]. Compared to these references, we analyze the (sub)linear scalings, provide justification for the scheme and identify explicit choices of the parameters involved.

1.4 Methodology and results

It follows from the analysis in [2, 3] that least squares approximations strike a balance between the approximation error and the size of the coefficients \( \|\lambda\| \) (see (7) for the corresponding error bound). Thus, in order to analyze the accuracy of an approximation, one has to analyze these quantities simultaneously. The main theoretical results of the paper are formulated in Theorems 7 and 8. The results are stated in terms of a user-chosen desired accuracy parameter \( \tau \ll 1 \), which is used to regularize the ill-conditioned linear system.

Theorem 7 describes the limiting accuracy for \( N \to \infty \). It is shown that the error may converge to zero in the sublinear scaling regime, while a linear scaling regime \( \varepsilon \sim cN \) leads to a saturation error that depends on the choice of \( c \). However, since the result is explicit, \( c \) can be chosen to make the saturation error arbitrarily small, e.g., on the order of the small parameter \( \tau \).

Theorem 8 describes the convergence rate for smooth functions. For a sublinear scaling regime, improving the approximation error comes at a cost of a rapidly increasing coefficient norm. Hence, as a consequence of ill-conditioning, the best approximations are not numerically computable. Taking both approximation error and coefficient norm into account, Theorem 7 shows slower convergence in the sublinear scaling regime compared with the linear scaling regime. For the case of Gaussian RBF’s, an optimal value of \( c \) in the scaling \( \varepsilon = cN \) is given explicitly by (21).

Broadly speaking, the methodology of the proof relies on the two points raised above in the following way. The centers outside the domain ensure that the saturation error can be made arbitrarily small. The oversampling ensures that good approximations with bounded coefficient norm can be found in practice, using function samples only and using inexact floating point computations.

1.5 Overview of the paper

We introduce the LS-RBF method in [2]. The relevant results of [2, 3] are reviewed in [3]. The main theoretical results for univariate problems using the Gaussian RBF are formulated in Theorems 7 and 8. The theory is illustrated with a range of experiments in [4].
2 The LS-RBF method

We now introduce the Least Squares Radial Basis Function (LS-RBF) method. We distinguish between the centers $\xi_n$ of the RBF’s, $n = 1, \ldots, N$, and the sampling points $x_m \in \Omega$, $m = 1, \ldots, M$. In the least squares approach, $M > N$. The sampling points belong to a domain $\Omega \subset \mathbb{R}^d$ on which the function $f$ to be approximated is defined. The centers may also be outside the domain.

The approximation to $f$ is

$$\tilde{f} = \sum_{n=1}^{N} \lambda_n \phi(\epsilon(\cdot - \xi_n)) = \sum_{n=1}^{N} \lambda_n \phi_n,$$

(2)

where $\lambda = (\lambda_n)_{n=1}^{N}$ is given by the discrete least squares solution

$$\lambda = \arg\min_{\zeta \in \mathbb{R}^N} \|A\zeta - b\|_\ell^2,$$

(3)

with

$$A = (\phi_n(x_m))_{m,n=1}^{M,N} \in \mathbb{R}^{M \times N}, \quad b = (f(x_m))_{m=1}^{M} \in \mathbb{R}^M.$$  

(4)

In our analysis we focus on the Gaussian RBF

$$\phi(r) = e^{-r^2},$$

(5)

but in the numerical experiments we also consider other RBF’s.

Depending on the distribution of the centers, the samples and the shape parameter, the matrix $A$ may be ill-conditioned. In that case, we solve $Ax = b$ using a direct solver with regularization threshold $\tau$. In this paper we analyze the use of a truncated singular value decomposition, in which all singular values below $\tau$ are discarded. In practice, similar results are obtained using a rank-revealing QR decomposition with pivoting.

The solution using regularization with threshold $\tau$ is denoted by $\lambda^{(\tau)}$, in contrast with the exact solution $\lambda$ of the least squares problem $\lambda$. However, in the following we will frequently omit the superscript.

Remark 1. In the analysis further on we have normalized the radial basis functions, as well as the matrix $A$. We have applied these normalizations in our implementation too, hence we describe them here. Thus, in theory and in practice, we (i) replace $\phi(\epsilon(\cdot - \xi_n))$ in (2) by $\sqrt{\epsilon} \phi(\epsilon(\cdot - \xi_n))$, and we (ii) multiply both $A$ and $b$ by $\sqrt{\frac{T}{N}}$. Normalization of the RBF’s affects the size of $\|\lambda\|$, the relevance of which is motivated in §3. The normalization of $A$ and $b$ also affects our results because we use an absolute threshold $\tau$ in the regularization. The rationale of this scaling and its possible generalizations are elaborated on in Remark 2.

3 Numerical approximations with redundancy

The ill-conditioning of the least squares matrix $A$ indicates that the RBF’s may be close to linearly dependent, or in other words, nearly redundant. The approximation of functions using redundancy was extensively analyzed in [2, 3]. We briefly review the main results, which will guide our analysis.
Define the synthesis operator $T_N : \mathbb{C}^N \to L^2(\Omega)$ by

$$T_N \lambda = \sum_{n=1}^N \lambda_n \phi_n.$$  

That is, $T_N$ associates with a set of coefficients the corresponding expansion in the RBF basis, which was defined in (2).

The solution vector $\mathbf{A}(\tau)$ of the linear system $Ax = b$, with $A$ and $b$ given by (4), and using SVD regularization with threshold $\tau$, satisfies [3, Theorem 1.3]

$$\| f - T_N \mathbf{A}(\tau) \|_{L^2(\Omega)} \leq \inf_{z \in \mathbb{C}^N} \{ \| f - T_N z \|_{L^2(\Omega)} + c_{M,N}^\tau \| f - T_N z \|_M + \tau d_{M,N}^\tau \| z \|_2 \}. \quad (6)$$

Here, $c_{M,N}^{\tau}$ and $d_{M,N}^{\tau}$ are constants that depend on the other parameters of the RBF approximation problem ($M$, $N$ and $\tau$) and in addition on the specific choice of sampling points and centers. The quantity $\| \cdot \|_M$ is a discrete norm defined over the sample points $x_m$. In particular, if the sample points $x_m$ fill the domain $\Omega$ quasi-uniformly as $M \to \infty$, one has that $c_{M,N}^{\tau}, d_{M,N}^{\tau} \leq 1$ and $\| f - T_N z \|_M \leq \| f - T_N z \|_{L^2(\Omega)}$ for sufficiently large $M$. Hence, under such conditions, the error of the least-squares approximation is determined by the quantity

$$E_{N,\tau}(f) = \inf_{z \in \mathbb{C}^N} \left\{ \| f - T_N z \|_{L^2(\Omega)} + \tau \| z \|_2 \right\}.$$  

Importantly, this term indicates that an accurate RBF approximation will be found if an accurate approximation exists with a coefficient vector $z$ with small discrete norm $\| z \|$. The numerically computed solution will be as accurate as any such vector $z$, regardless of the condition number of $A$, up to the threshold $\tau$.

In view of the generic error estimate (7), in our analysis it is sufficient to show the existence of accurate RBF approximations with corresponding bounds on the sizes of their expansion coefficients.

Remark 2. We elaborate briefly on the behaviour of the constants $c_{M,N}^{\tau}$ and $d_{M,N}^{\tau}$ in (6) for increasing $M$, and their relation to the choice of sampling points, in order to better motivate the simpler form (7). It is shown in [3, Proposition 3.10] that these constants satisfy $c_{M,N}^{\tau}, d_{M,N}^{\tau} \leq \frac{1}{\sqrt{A_{M,N}}} \, \sqrt{\lambda_{M,N}}$ where

$$A_{M,N} = \inf_{g \in H_N} \| g \|_M^2 \| g \|_{H_N} \L_1$$

with in the setting of this paper $H_N = \text{span}\{ \phi_n \}_{n=1}^N$ and

$$\| g \|_M^2 = \frac{1}{M} \sum_{m=1}^M g(x_m)g(x_m).$$

Ideally, one ensures that the discrete and continuous norms agree for large $M$:

$$\| g \|_M^2 \to \| g \|_{L^2(\Omega)}^2 \quad \text{or} \quad \frac{1}{M} \sum_{m=1}^M g(x_m)g(x_m) \to \int_\Omega g(x)g(x)dx, \quad M \to \infty. \quad (8)$$

In that case both constants $c_{M,N}^{\tau}$ and $d_{M,N}^{\tau}$ simply tend to 1. This is the case for all examples in this paper because we use equispaced points and (8) becomes a Riemann sum. For non-equispaced points, one may consider weighting each sampling point $x_m$ with a weight $\sqrt{w_m}$.

In that case $\| g \|_M^2 = \sum_{m=1}^M w_m g(x_m)g(x_m)$, and a proper choice of weights can ensure that a weighted analogue of (8) holds.
Remark 3. The previous remark pertains to the oversampling limit $M \to \infty$. For fixed $M$ and $N$, [3, Proposition 3.10] also shows that $c_{M,N}^2, d_{M,N}^2 \lesssim \frac{1}{\varepsilon^2}$. For $M$ close to $N$, this bound may be sharp and both constants may be large. Thus, without ‘sufficient’ oversampling [6] does not demonstrate accuracy of the approximation. In particular, there is no reason to assume that an RBF interpolant ($M = N$), rather than a least squares approximation ($M > N$), would yield an accurate function approximation in general. We will show with numerical experiments that $M \sim \gamma N$ for some $\gamma > 1$ appears to be sufficient in our setting and we generally choose $\gamma = 2$. It remains an open problem to prove that such linear oversampling is indeed sufficient.

4 Analysis in the one-dimensional case for Gaussian RBFs

In this section, we study the behaviour of (7) for the Gaussian RBF (5). For simplicity, we consider the one-dimensional case, where $\Omega \subset (-T, T)$ is the domain. Our analysis relies crucially on the radius of $\Omega$, defined as

$$B = \max_{x \in \Omega} |x|,$$  \hspace{1cm} (9)

Note that $B < T$ because $\Omega$ is a subset of $[-T, T]$. We also consider equally-spaced centers on $[-T, T]$ given by

$$\xi_n = \frac{nT}{N}, \quad n = -N, \ldots, N.$$  \hspace{1cm} (10)

For convenience, we assume the approximation uses these $2N + 1$ centers (as opposed to $N$ in (2)), indexed from $-N$ to $N$. That is,

$$f \approx T_N \lambda = \sum_{n=-N}^{N} \lambda_n \sqrt{\varepsilon} \phi(\varepsilon(\cdot - nT/N)).$$  \hspace{1cm} (11)

Note that we now also normalize the RBF functions by a factor $\sqrt{\varepsilon}$, so that the $L^2$-norms of the corresponding terms in the sum are independent of $\varepsilon$.

We define the Fourier transform of $f \in L^2(\mathbb{R})$ by

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(x) e^{-i\omega x} \, dx, \quad \omega \in \mathbb{R},$$ \hspace{1cm} (12)

with inverse

$$\check{g}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(\omega) e^{i\omega x} \, d\omega, \quad x \in \mathbb{R},$$ \hspace{1cm} (13)

for $g \in L^2(\mathbb{R})$. We also recall Plancherel’s formula

$$\|\hat{f}\|_{L^2(\mathbb{R})} = \sqrt{2\pi} \|f\|_{L^2(\mathbb{R})}.$$ \hspace{1cm} (14)

4.1 Preliminary results

Our main results are contained in §4.2 and §4.3. First, in this subsection, we state and prove a number of preliminary results.
Lemma 4. Let \( f \in L^2(\Omega) \), let \( g \in L^2(\mathbb{R}) \) be any extension of \( f \) to the real line, and let \( \lambda \in \ell^2(\mathbb{Z}) \) be arbitrary. Then

\[
E_{N,\tau}(f) \leq \|g - h\|_{L^2(\mathbb{R})} + \left( \sqrt{\frac{2\pi BN}{T}} \text{erfc}(\sqrt{2\varepsilon}(T - B)) + \tau \right) \|\lambda\|_{\ell^2},
\]

where \( E_{N,\tau}(f) \) is as in (7), \( h = \sum_{n=-\infty}^{\infty} \lambda_n \sqrt{\varepsilon} \phi(\varepsilon(-nT/N)) \in L^2(\mathbb{R}) \) and \( \text{erfc} \) is the complementary error function. In particular, if

\[
\varepsilon \geq \frac{1}{\sqrt{2}(T - B)} \left\lfloor \log \left( \frac{\sqrt{2\pi BN}}{T\tau^2} \right) \right\rfloor,
\]

then

\[
E_{N,\tau}(f) \leq \|g - h\|_{L^2(\mathbb{R})} + 2\tau \|\lambda\|_{\ell^2}.
\]

The motivation behind this lemma is the following. Analyzing \( E_{N,\tau}(f) \) directly requires constructing a finite series \( \sum_{n=-N}^{N} \lambda_n \phi_n \) which approximates \( f \) over \( \Omega \) and whose coefficients \( \lambda_n \) do not grow too large. This can be difficult, due to the requirement that the series be finite and the fact that the domain \( \Omega \) is not the whole real line. Instead, this lemma states that it is sufficient to study how well any extension \( g \in L^2(\mathbb{R}) \) of \( f \) can be approximated over the whole real line via an infinite series whose terms are not too large. As we shall see later, by moving to infinite series on the real line we are able to employ Fourier analysis techniques.

Proof of Lemma 4. We first show that the function \( h \in L^2(\mathbb{R}) \). Let \( y = Nx/T \) and \( \hat{\psi}(y) = \phi(\varepsilon Ty/N) \). Then \( h \in L^2(\mathbb{R}) \) if and only if the function \( \sum_{n=-\infty}^{\infty} \lambda_n \psi(\cdot - n) \in L^2(\mathbb{R}) \) for \( \lambda \in \ell^2(\mathbb{Z}) \). By Lemma 9.2.2 and Theorem 9.2.5 of [7], this holds provided

\[
\sum_{k \in \mathbb{Z}} |\hat{\psi}(2\pi(\omega + k))|^2 \leq C, \quad \text{a.e. } \omega \in [0, 1],
\]

for some constant \( C > 0 \). Observe that

\[
\sum_{k \in \mathbb{Z}} |\hat{\psi}(2\pi(\omega + k))|^2 = \frac{N^2}{\varepsilon^2 T^2} \sum_{k \in \mathbb{Z}} |\hat{\phi}(2\pi N(\omega + k)/\varepsilon T)|^2.
\]

Since \( \phi \) is the Gaussian, we have

\[
\hat{\phi}(\omega) = \sqrt{\pi} e^{-\omega^2/4}.
\]

and therefore

\[
\sum_{k \in \mathbb{Z}} |\hat{\psi}(2\pi(\omega + k))|^2 = \frac{N^2 \pi}{\varepsilon^2 T^2} \sum_{k \in \mathbb{Z}} \exp(-2\pi^2 N^2(\omega + k)^2/(\varepsilon^2 T^2)).
\]

This sum is clearly uniformly bounded in \( \omega \). Hence (14) holds for some suitable \( C \). We conclude that \( h \in L^2(\mathbb{R}) \).

With this in hand, we let \( z = (\lambda_n)_{n=-N}^{N} \) in (7) to obtain

\[
E_{N,\tau}(f) \leq ||f - h||_{L^2(\Omega)} + ||h - T_N z||_{L^2(\Omega)} + \tau ||\lambda||_{\ell^2}
\]

\[
\leq ||g - h||_{L^2(\mathbb{R})} + \sum_{|n|>N} |\lambda_n| \sqrt{T} \|\phi(\varepsilon(-nT/N))\|_{L^2(\mathbb{R})} + \tau ||\lambda||_{\ell^2}.
\]
Consider the middle term. By [9], we have
\[ \| \phi(\epsilon \cdot - nT/N) \|_{L^2(\Omega)} \leq \sqrt{2B} e^{-\epsilon^2(B-|nT/N|^2)} . \]
Therefore
\[
\sum_{|n|>N} |\lambda_n| \sqrt{\epsilon} \| \phi(\epsilon \cdot - nT/N) \|_{L^2(\Omega)} \leq \sqrt{4B}\|\lambda\|_{\ell^2} \sqrt{\epsilon} \sqrt{\int_{N}^{\infty} e^{-2\epsilon^2(xT/N-B)^2} \, dx} \\
= \sqrt{4B}\|\lambda\|_{\ell^2} \sqrt{\epsilon} \sqrt{\frac{N}{\sqrt{2T\epsilon}}} \int_{\sqrt{2\epsilon}(T-B)}^{\infty} e^{-y^2} \, dy \\
= \sqrt{\frac{2\pi BN}{T}} \text{erfc}(\sqrt{2\epsilon}(T-B))\|\lambda\|_{\ell^2} .
\]
This completes the proof of the first result.

For the second result, observe that \( \text{erfc}(z) \leq e^{-z^2} \). This follows from [8, Eqn. (7.8.2)], noting that \( 1/(x + \sqrt{x^2 + 4/\pi}) < 1 \) for \( x > 0 \). Hence, due to condition (13) on \( \epsilon \),
\[
\frac{\sqrt{2\pi BN}}{T} \text{erfc}(\sqrt{2\epsilon}(T-B)) \leq \frac{\sqrt{2\pi BN}}{T} \exp(-2\epsilon^2(T-B)^2) \leq \tau^2 .
\]
The result now follows immediately from the first part. \( \square \)

This lemma states that when \( \epsilon \) grows mildly with \( N \), we reduce the problem of estimating \( E_{N,\tau}(f) \) to that of studying the approximation of a function \( g \in L^2(\mathbb{R}) \) via the corresponding infinite expansion \( h \) in the RBF system. Note that the growth condition (13) is primarily for convenience and it is not a fundamental restriction in the analysis: it merely allows us to simplify the bulky expression in the first part of the lemma, which would otherwise propagate into all bounds later on. The condition will be satisfied in all examples considered later.

With this in hand, to estimate \( E_{N,\tau}(f) \) we now construct a suitable function \( h \) via Fourier analysis. This is the topic of the following two lemmas. Specifically, in view of Plancherel’s identity, we pick \( h \) so that \( \hat{h} \) agrees with \( \hat{g} \) for all low frequencies.

**Lemma 5.** Let \( f \in L^2(\Omega) \) and let \( g \in L^2(\mathbb{R}) \) be any extension of \( f \) to the real line. If \( \epsilon \) satisfies (13) then, for any \( K \leq N \),
\[
E_{N,\tau}(f) \lesssim \sqrt{\int_{|\omega|>K\pi/T} |\hat{g}(\omega)|^2 \, d\omega} \\
+ \left( \frac{1}{\sqrt{e^{N^2\pi^2/(2T^2\epsilon^2)} - 1}} + \sqrt{\frac{T\epsilon}{N\tau}} \right) \sqrt{\int_{|\omega|\leq K\pi/T} \frac{|\hat{g}(\omega)|^2}{|\phi(\omega/\epsilon)|^2} \, d\omega} .
\]

**Proof.** Consider the function \( F : [-N\pi/T, N\pi/T] \to \mathbb{C} \) defined by
\[
F(\omega) = \frac{\sqrt{\epsilon} \hat{g}(\omega)}{\phi(\omega/\epsilon)} \|_{-K\pi/T,K\pi/T}(\omega) .
\]
Note that $F \in L^2([-N\pi/T, N\pi/T])$ since $\hat{g} \in L^2(\mathbb{R})$ and $\hat{\phi}$ is strictly positive and bounded. Hence $F$ has the Fourier series

$$F(\cdot) = \sum_{n=-\infty}^{\infty} \lambda_n e^{-inT/\pi},$$

which converges in $L^2([-N\pi/T, N\pi/T])$ with coefficients $\lambda \in \ell^2(\mathbb{Z})$ satisfying, by Parseval’s identity,

$$\|\lambda\|_{\ell^2}^2 = \frac{T}{2N\pi} \|F\|_{L^2([-N\pi/T, N\pi/T])}^2 = \frac{T\varepsilon}{2N\pi} \int_{|\omega| \leq K\pi/T} |\hat{\phi}(\omega/\varepsilon)|^2 d\omega. \quad (15)$$

Let $h = \sum_{n=-\infty}^{\infty} \lambda_n \sqrt{\varepsilon} \phi(\varepsilon(\cdot - nT/N))$ and notice that $h \in L^2(\mathbb{R})$ by Lemma 4. Observe that

$$\hat{h}(\omega) = \frac{\hat{\phi}(\omega/\varepsilon)}{\sqrt{\varepsilon}} \sum_{n=-\infty}^{\infty} \lambda_n e^{-i\omega nT/\pi}.$$

In particular,

$$\hat{h}(\omega) = \hat{g}(\omega) I([-K\pi/T, K\pi/T])(\omega), \quad |\omega| \leq N\pi/T.$$

Hence

$$\|g - h\|_{L^2(\mathbb{R})} = \frac{1}{\sqrt{2\pi}} \|\hat{g} - \hat{h}\|_{L^2(\mathbb{R})}$$

$$= \frac{1}{\sqrt{2\pi}} \sqrt{\int_{|\omega| > K\pi/T} |\hat{g}(\omega) - \hat{h}(\omega)|^2 d\omega}$$

$$\leq \frac{1}{\sqrt{2\pi}} \left( \sqrt{\int_{|\omega| > K\pi/T} |\hat{g}(\omega)|^2 d\omega} + \sqrt{\int_{|\omega| > N\pi/T} |\hat{h}(\omega)|^2 d\omega} \right).$$

Consider the second term. We have

$$\int_{|\omega| > N\pi/T} |\hat{h}(\omega)|^2 d\omega = \sum_{l \neq 0} \int_{(2l-1)N\pi/T}^{(2l+1)N\pi/T} \varepsilon^{-1} |\hat{\phi}(\omega/\varepsilon)|^2 |F(\omega)|^2 d\omega$$

$$\leq 2\pi \|F\|_{L^2([-N\pi/T, N\pi/T])}^2 \varepsilon^{-1} e^{-(2l-1)^2N^2\pi^2/(2T^2\varepsilon^2)}.$$

Replacing $(2l-1)^2$ by $l$ and summing the resulting geometric series gives

$$\int_{|\omega| > N\pi/T} |\hat{h}(\omega)|^2 d\omega \lesssim \varepsilon^{-1} \|F\|_{L^2([-N\pi/T, N\pi/T])}^2 \frac{1}{e^{N^2\pi^2/(2T^2\varepsilon^2)} - 1}.$$

Hence, combining this with the previous estimate and (15), we obtain

$$\|g - h\|_{L^2(\mathbb{R})} + \tau \|\lambda\|_{\ell^2} \lesssim \sqrt{\int_{|\omega| > K\pi/T} |\hat{g}(\omega)|^2 d\omega}$$

$$+ \left( \sqrt{\frac{T}{N} \tau^2 + \frac{\varepsilon^{-1/2}}{\sqrt{e^{N^2\pi^2/(2T^2\varepsilon^2)} - 1}} \|F\|_{L^2([-N\pi/T, N\pi/T])}^2} \right).$$

Hence the result now follows by applying Lemma 4 once more.
The previous lemma involves an additional parameter $K$. In practice, we make the specific choice $K = \min\{N, \varepsilon\}$. This leads to the following:

**Lemma 6.** Let $f \in L^2(\Omega)$, and let $g \in L^2(\mathbb{R})$ be any extension of $f$ to the real line. If $\varepsilon$ satisfies (13) then

$$E_{N,T}(f) \lesssim \sqrt{\int_{|\omega| > \min\{N, \varepsilon\} \pi/T} |\hat{g}(\omega)|^2 d\omega} + \left( \frac{1}{\sqrt{e^{N^2 \pi^2/(2T^2 \varepsilon^2)}}} + \sqrt{\frac{T \varepsilon}{N \tau}} \right) e^{\pi^2/(4T^2)} \|g\|^2_{L^2(\mathbb{R})}.$$  

**Proof.** We use Lemma 5 with $K = \min\{N, \varepsilon\}$. Observe that

$$\int_{|\omega| \leq K \pi/T} \frac{|\hat{g}(\omega)|^2}{|\hat{\phi}(\omega/\varepsilon)|^2} d\omega \lesssim e^{K^2 \pi^2/(2T^2 \varepsilon^2)} \|g\|^2_{L^2(\mathbb{R})}.$$  

The term $K^2/\varepsilon^2 \leq 1$. Hence

$$\int_{|\omega| \leq K \pi/T} \frac{|\hat{g}(\omega)|^2}{|\hat{\phi}(\omega/\varepsilon)|^2} d\omega \lesssim e^{\pi^2/(2T^2)} \|g\|^2_{L^2(\mathbb{R})}.$$  

The result now follows from Lemma 5.  

4.2 \ Limiting accuracy

In this and the next section we analyze $E_{N,T}(f)$. First, we consider the limiting behaviour as $N \to \infty$ for different choices of the parameter $\varepsilon$:

**Theorem 7.** Let $f \in L^2(\Omega)$. Then:

1. If $\varepsilon \sim cN^\alpha$ for some $0 < \alpha < 1$ and $c > 0$ as $N \to \infty$ then

$$E_{N,T}(f) \lesssim \tau \sqrt{T \varepsilon/N} e^{\pi^2/(4T^2)} \|f\|_{L^2(\Omega)} + o(1), \quad N \to \infty.$$  

In particular, $E_{N,T}(f) \to 0$ as $N \to \infty$.

2. If $\varepsilon \sim cN$ as $N \to \infty$ for some $c > 0$ then

$$E_{N,T}(f) \lesssim \left( \frac{1}{\sqrt{e^{\pi^2/(2c^2 T^2)}}} + \sqrt{cT \tau} \right) e^{\pi^2/(4T^2)} \|f\|_{L^2(\Omega)} + o(1). \quad (16)$$

In particular, if $\tau < 1/2$ and $c$ satisfies

$$cT \leq \frac{\pi}{\sqrt{2 \log(1 + \tau^{-2})}}, \quad (17)$$

then, as $N \to \infty$,

$$E_{N,T}(f) \lesssim \tau e^{\pi^2/(4T^2)} \|f\|_{L^2(\mathbb{R})} + o(1).$$

10
This result shows the following. First, whenever $\varepsilon$ grows asymptotically more slowly than $N$ the error of the RBF approximation is guaranteed to approach zero as $N \to \infty$. On the other hand, if $\varepsilon \sim cN$ then the error only decreases down to some finite value, with that value being dependent of the SVD regularization $\tau$, the extension domain size $T$ and the constant $c$. In particular, if the product $cT$ is too large, the best achievable accuracy may be significantly larger than $\tau$. On the other hand, by choosing $cT$ according to $\tau$ as in (17) we ensure an asymptotic accuracy of order $\tau$.

Proof. We use Lemma 6 with $g \in L^2(\mathbb{R})$ the extension of $f$ by zero. Notice that \[ \|g\|_{L^2(\mathbb{R})} = \|f\|_{L^2(\Omega)}. \] Since $\varepsilon$ satisfies (13) we have, in all cases, \[ \int_{|\omega|>\min\{N,\varepsilon\} \pi/T} |\hat{g}(\omega)|^2 \, d\omega \to 0, \quad N \to \infty. \]

Moreover, if $\varepsilon \sim cN$, then \[ \frac{1}{\sqrt{e^{N^2\pi^2/(2c^2T^2)}} - 1} + \frac{\varepsilon T}{N} \sim \frac{\varepsilon T}{N} \tau, \quad N \to \infty, \]

and if $\varepsilon \sim cN$ then \[ \frac{1}{\sqrt{e^{N^2\pi^2/(2c^2T^2)}} - 1} + \frac{\varepsilon T}{N} \sim \frac{1}{\sqrt{e^{N^2\pi^2/(2c^2T^2)}} - 1} + \sqrt{cT\tau} \sim 1 \sqrt{\frac{\pi}{4c^2T^2}} - 1 + \sqrt{cT\tau}, \quad N \to \infty. \]

This gives the result. \qed

4.3 Error decrease for smooth functions

The previous result determines the limiting behaviour, but we are also interested in how fast the term $E_{N,\tau}(f)$ reaches that limit. This, naturally, depends on the regularity of the function being approximated. In this section, we study the behaviour for functions in the Sobolev spaces $H^k(\Omega)$. In order to obtain precise bounds, we now use exact regimes such as $\varepsilon = cN^\alpha$, rather than the asymptotic scaling $\varepsilon \sim cN^\alpha$ of the previous theorem.

Theorem 8. Let $f \in H^k(\Omega)$. Then

1. Suppose that $\varepsilon = cN^\alpha$ for some $0 < \alpha < 1$ and $c > 0$. Then \[ E_{N,\tau}(f) \lesssim \left[ (cN^\alpha \pi/T)^{-k} + \left( e^{-\pi^2N^{2(1-\alpha)/(4c^2T^2)}} + \sqrt{cT\tau/N^{1-\alpha}/2} \right) e^{\pi^2/(4T^2)} \right] \|f\|_{H^k(\Omega)}, \]

for all $N$ satisfying \[ N^{1-\alpha} \geq \frac{\sqrt{2\log(2)}T}{\pi}, \quad cN^\alpha \geq \frac{1}{\sqrt{2(T-B)}} \log \left( \frac{\sqrt{2\pi BN}}{T\gamma^2} \right). \]
2. Suppose that \( \varepsilon = cN \) for some \( c > 0 \). Then
\[
E_{N, \tau}(f) \lesssim \left( \min\{c, 1\} N \pi / T \right)^{-k} \left( \frac{1}{\sqrt{\varepsilon^2 / (2c^2 T^2)} - 1} + \sqrt{c T \tau} \right) \|f\|_{H^k(\Omega)},
\]
for all \( N \) satisfying
\[
cN \geq \frac{1}{\sqrt{2(T - B)}} \sqrt{\log \left( \frac{2\pi BN}{T \tau^2} \right)}. \tag{19}
\]
In particular, if \( \tau < 1/2 \) and
\[
c \leq \min \left\{ 1, \frac{\pi}{2T \sqrt{2 \log(1 + \tau^{-2})}} \right\}, \tag{20}
\]
then
\[
E_{N, \tau}(f) \lesssim \left( cN \pi / T \right)^{-k} + \tau \pi^2 / (4T^2) \|f\|_{H^k(\Omega)}.
\]

The main conclusion is that whenever \( \varepsilon \) grows more slowly than \( N \) the algebraic convergence order is suboptimal. Specifically, \( O(N^{-\alpha k}) \), as opposed to the optimal \( O(N^{-k}) \) rate. Thus, this theorem suggests choosing \( \varepsilon = cN \). Seeking to make the algebraically-decreasing term in this case as small as possible, one is tempted to take \( c \) as large as possible. However, as was already noted in Theorem 7, this limits the maximal achievable accuracy. To attain the desired \( O(\tau) \) accuracy, we need to limit \( c \) according to \( \text{Eq. } (20) \). Hence, the overall conclusion is to choose \( \varepsilon = cN \) with
\[
c = c^* = \frac{\pi}{T \sqrt{2 \log(1 + \tau^{-2})}}. \tag{21}
\]
This also suggests choosing \( T > B \) as small as possible. However, because of the condition \( \text{Eq. } (19) \), choosing \( T \) too small may result in a longer pre-asymptotic regime of the error, before the error rates of Theorem 8 kick in.

**Proof of Theorem 8.** By the Sobolev extension theorem, there exists an extension \( g \in H^k(\mathbb{R}) \) of \( f \) satisfying \( \|g\|_{H^k(\mathbb{R})} \lesssim \|f\|_{H^k(\Omega)} \). Recall that, by definition,
\[
\|g\|_{H^k(\mathbb{R})}^2 = \int_{-\infty}^{\infty} (1 + |\omega|)^{2k} |\hat{g}(\omega)|^2 \, d\omega.
\]
Hence
\[
\sqrt{\int_{|\omega| > \min\{N, \varepsilon\} \pi / T} |\hat{g}(\omega)|^2 \, d\omega} \lesssim \left( \min\{N, \varepsilon\} \pi / T \right)^{-k} \|f\|_{H^k(\Omega)}.
\]
We use Lemma 6. The case for \( \varepsilon = cN \) follows immediately. For the case \( \varepsilon = cN^\alpha \) we merely notice that
\[
e^{N^2 \pi^2 / (2c^2 T^2)} - 1 \geq \frac{1}{2} e^{N^2 \pi^2 / (2c^2 T^2)} = \frac{1}{2} e^{N(2\pi^2/4c^2 T^2)},
\]
due to the assumption \( N \). This completes the proof. \( \square \)
5 Numerical experiments

We illustrate the methodology of this paper with a sequence of examples. For the case of univariate approximation with a Gaussian radial function, these results confirm the results of the theory.

We vary the scaling of the shape parameter \( \epsilon \) with \( N \), the shape of the domain \( \Omega \) as well as its dimension and the radial function \( \phi \). In all examples, the centers \( \{ \xi_n \} \) are located in the bounding box \( [-T, T]^d \), whereas the sampling points \( \{ x_m \} \) are restricted to lie inside \( \Omega \).

We solve the rectangular least squares problem using a truncated SVD decomposition with a regularization threshold \( \tau \). All singular values smaller than \( \tau \sigma_1 \), with \( \sigma_1 \) the largest one, are discarded.

5.1 Centers outside the computational domain and the choice of \( T \)

First, we illustrate the importance of adding centers outside the computational domain in the linear scaling regime \( \epsilon = cN \). We consider the interval \( \Omega = [-1, 1] \). Intuitively, due to the truncation at the endpoints, the radial basis functions lack sufficient density near \( \pm 1 \) when centers are confined to \([-1, 1]\). This is clearly visible in Fig. 1 in which the translations of the Gaussian radial basis function are plotted with and without centers outside \([-1, 1]\). Loosely speaking, the approximation space is enlarged by adding centers outside the domain, until the leftmost and rightmost RBF are numerically small on the computational domain itself.

More precisely, our analysis has shown that convergence sets in once \( N \) is larger than a
Figure 2: Error and magnitude of the coefficients in the linear scaling regime $\varepsilon = cN$ for the functions $f(x) = \frac{1}{1+10x^2}$ (left column) and $f(x) = 1/(x-1.2)$ (right column). Results are shown for varying values of the regularization threshold $\tau$. Different values of $T$ are used in the top row ($T = 1.5$) and bottom row ($T = 4$). The proportionality constant $c$ is related to $T$ and $\tau$ via (21). Shown with horizontal dashed lines is the estimate of the limiting accuracy (16).

minimal value reflected in the bound (19). This bound merely ensures the above-mentioned condition. Indeed, note that the rightmost RBF in (10) is $\phi(\varepsilon(x-T))$. Letting $\varepsilon = cN$ equal the right-hand side of (19) and evaluating this RBF at the point $x = 1$ yields, noting that $B = 1$ in this case and recalling the scaling by $\sqrt{\varepsilon}$,

$$\sqrt{\varepsilon} \phi(\varepsilon(1-T)) = \sqrt{cN} \phi \left( \frac{1}{\sqrt{2}} \log \left( \frac{\sqrt{2\pi}N}{T \tau^2} \right) \right) = \sqrt{cN} e^{-\frac{1}{2} \log \left( \frac{\sqrt{2\pi}N}{T \tau^2} \right)} = \sqrt{\frac{cT}{\sqrt{2\pi}}} \tau.$$

That is the maximal value of the $N$-th radial basis function on $[-1,1]$: it is on the order of $\tau$.

5.2 Scaling regimes of the shape parameter

We consider three regimes for the shape parameter $\varepsilon(N)$: (i) $\varepsilon$ grows like $cN$, (ii) it grows like $c\sqrt{N}$ or (iii) it is constant.

5.2.1 Linear scaling: $\varepsilon = cN$

Results for the linear scaling regime $\varepsilon = cN$ are shown in Fig.2 for two functions approximated on $[-1,1]$: $f(x) = 1/(1+10x^2)$, the Runge function with poles near the real line in the middle of the interval (left column), and $f(x) = 1/(x-1.2)$ with a pole near the right endpoint (right column). We have used an oversampling factor $\gamma = 2$. Results are shown for different values
of the threshold $\tau$, indicated on the figure, and with $c$ chosen depending on $T$ and $\tau$ according to our estimate of the optimal value $[21]$. In all cases, there is an initial regime in which the coefficient norm grows. Smaller values of the threshold $\tau$ result in larger coefficient norms. Note that decreasing $\tau$ leads to decreasing $c$ and decreasing $\varepsilon$ and, hence, to flatter radial basis functions. However, as the approximation converges to its best accuracy, the coefficient norm also settles down to a moderate value. This initial regime of coefficient growth can potentially be avoided by using the method introduced in [4].

Seemingly geometric convergence is observed in all cases, because the functions involved are analytic on $[-1, 1]$. In the linear scaling regime the maximal achievable accuracy is limited and depends on $c$, $T$ and $\tau$. The limiting accuracy predicted by Theorem 7, equation (16), is

$$\left(\frac{1}{\sqrt{e^{\pi^2/(2\cdot c^2 T^2)}}} + \sqrt{c T \tau}\right) e^{\pi^2/(4 T^2)}.$$

This value is included in the figures with horizontal dashed lines. There is very good agreement with this limit in all cases.

Comparing the top and bottom row illustrates that convergence rate in this example is better for smaller $T$, in agreement with Theorem 8. As long as (17) is satisfied, theory also predicts that smaller $T$ is better. We omit further experiments involving varying $T$ but note that a thorough investigation of the influence of the extension parameter $T$ was given for Fourier extension approximations in [1].

We repeat the experiment in Fig. 3 for a fixed value of $\tau = 1e-10$ and varying $c$. There is still a good agreement between the maximal achievable accuracy and the predicted accuracy limit. However, in this regime we clearly observe saturation errors: the accuracy limit can be much higher than the threshold value $\tau$. For values of $c$ where this is the case (e.g. $c = 0.5$), the coefficient norm also exhibits erratic behaviour as a function of $N$.

### 5.2.2 Sublinear scaling: $\varepsilon = c\sqrt{N}$

Results are shown in Fig. 4 for the scaling regime $\varepsilon = c\sqrt{N}$. It follows from Theorem 7 that there is no saturation error in this regime: the expected limiting accuracy for large $N$ is $O(\tau)$, independently of $c$. However, it follows from Theorem 8 that larger values of $c$ asymptotically lead to more rapid convergence. Both observations are confirmed in Fig. 4. The theory makes
Figure 4: Error and magnitude of the coefficients in the square-root scaling regime \( \varepsilon = c\sqrt{N} \) for function \( f(x) = \frac{1}{1+10x^2} \). Different values of \( T \) are used in the left and right panel. Each plot includes varying values of \( c \), independently of the fixed threshold \( \tau = 1e - 10 \).

no claim about the pre-asymptotic regime, and we note that in the right panel of the figure a larger value of \( c \) can initially be worse than a smaller one. Once the asymptotic regime of convergence is reached, larger \( c \) leads to faster decay of the error as well as to a smaller coefficient norm.

For comparison, we include a result on the approximation of a less smooth function in Fig. 5. To that end we consider the function \( f(x) = |x|^5 \). The convergence rates are algebraic in this case, but the results are otherwise similar: the linear regime offers faster convergence, up to a saturation limit. The sublinear regime leads to convergence, but does so more slowly.

5.2.3 The flat limit: \( \varepsilon = c \)

In the flat limit in \( 1-D \), one can easily show that the radial function approximant converges to a polynomial over \([-T, T]\), once a stable method has been applied. However, none of these strategies, mentioned in §1.1, are part of our setting and we will therefore not consider this case.

5.3 Families of radial basis functions

Our analysis and the previous experiments are based on the Gaussian RBF. From the point of view of the analysis developed in this paper, and the specific methodologies of proof that we employed, there is a large qualitative difference between compactly supported and non-compactly supported RBF’s. However, the main observation that inspired our analysis, namely the interplay of numerical stability with coefficient norm and approximation error, remains the same: the size of the expansion coefficients limits the potentially achievable accuracy in numerical computations.

Here, we repeat the approximation of the Runge function \( f(x) = \frac{1}{1+10x^2} \) using four different families of radial basis functions: the multiquadric (MQ), inverse quadratic (IQ), inverse multiquadric (IMQ) and the Gaussian (GA) (recall their definitions in Table 1). We always use equispaced centers on \([-T, T]\) and oversampling with an oversampling ratio of \( \gamma \).

Results are shown in Fig. 6 for \( T = 4 \) and \( \gamma = 2 \). For reference, in each figure we have included the limiting accuracy of Gaussian radial basis function as predicted by Theorem 7 in dashed lines. We note that in all cases the coefficient norm initially increases, before decreasing down to a limit. In this experiment, it does seem to be the case that the Gaussian
Figure 5: Log-log plot of the error decay using linear and sublinear $\varepsilon$ values when approximating function $f(x) = |x|^5$, for $\tau = 10^{-5}$ and for $\tau = 10^{-10}$. The decay of the error is algebraic in this case, since the target function is of finite regularity. We display the $\log$(error) versus $\log(N)$ along with some powers of $\frac{1}{N}$ to help gauge the decay. As expected, the linear scaling of $\varepsilon$ leads to a fast convergence, but does not decay past a best error value. On the other hand, a sublinear scaling of the shape parameter leads to a slower decay of the error, and never reaches a minimum error threshold.

RBF offers highest convergence rates, the highest accuracy and the smallest coefficient norm. The loss of accuracy compared to the threshold $\tau$ in the other RBF families is due to a larger coefficient norm. These results have motivated our selection of the Gaussian RBF for further analysis.

Qualitatively similar results are obtained for other choices of $T$ and $\gamma$, except that (as with the Gaussian RBF) accuracy may be lost in the absence of sufficient oversampling. We study the effect of oversampling separately further on.

5.4 The sampling rate and a rule of thumb

In our theoretical analysis, we have investigated the minimization problem (7). We have argued that, provided there is sufficient oversampling, the numerical result corresponds to this optimum. A more general minimization problem which does take sampling into account, via the constants $c_{\tau M,N}$ and $d_{\tau M,N}$, was (6).

In practice, any difference between (6) and (7) can be easily detected. As a rule of thumb, if (7) is minimized, then one expects the two terms in this minimization problem to be balanced and, thus, one expects the ratio of the approximation error over the coefficient norm to approximately equal $\tau$:

$$\frac{\|f - T_N z\|_{L^2(\Omega)}}{\|z\|_{L^2}} \approx \tau.$$ 

Conversely, if this ratio does not approximate $\tau$, then (7) was not optimally solved and one reason may be that there was not enough oversampling. This means that the error is governed by the original expression (6) in which $c_{\tau M,N}$ and $d_{\tau M,N}$ are larger than expected. The solution in this case is to increase $M$ with respect to $N$. Recall that we assume that the centers
ultimately (i.e., for large $M$) fill the domain $\Omega$ in a quasi-uniform manner, to ensure that $c^{\tau}_{M,N}, d^{\tau}_{M,N} \lesssim 1$.

The rule of thumb is illustrated in Fig. 7 where we repeat a part of Fig. 2 and we have simply added the ratio (22) to the figure. The dotted lines do approximate $\tau$ in all cases, including in the pre-asymptotic regime. This is an indication that an oversampling factor of 2 was sufficient for these two examples. Note that the constants $c^{\tau}_{M,N}, d^{\tau}_{M,N}$ can in principle also be computed numerically, see [3, §3.6].

5.5 Two-dimensional examples

The explicit analytical results of this paper are specific to the univariate Gaussian function. These results can be straightforwardly extended to a tensor product setting. However, that would lead to basis functions aligned with the coordinate axes of the form $\phi(x)\phi(y)$. Such functions are rarely used in practice because radial basis functions in 2D are, as the name implies, defined in terms of the radial distance:

$$\phi(x, y) = \phi(|x - y|).$$

In the next experiments we use the latter, standard definition of radial basis functions. Strictly speaking, the theory of this paper no longer applies. However, we illustrate that the guiding principles remain the same. In particular, we show that the linear scaling regime does lead to accurate approximations using radial basis functions in 2D, and that the choices of the parameters involved correspond to striking a balance between coefficient norm and approximation accuracy.
Figure 7: Same as the bottom row of Fig. 2 but with an extra indication of the value of $\tau$ and of the ratio of the accuracy over the coefficient norm (dotted lines). As a rule of thumb, this ratio is approximately equal to the threshold $\tau$.

Figure 8: Approximation of function $f(x, y) = \frac{1}{1+10(x^2+y^2)}$ with 100 centers and 200 sample points inside the computational domain on the left and 1,000 centers and 2,000 sample points on the right. The right panel shows the approximation error in logarithmic scale.

In an analogous manner to our one-dimensional setup, a bounding box of adjustable size $[-T_1, T_1] \times [-T_2, T_2]$ is chosen to contain the computational domain $\Omega$. The $N$ centers are placed in an hexagonal pattern inside the bounding box. The $M$ sample nodes within $\Omega$ also have a hexagonal pattern, although an investigation of optimal node distribution with possible local node refinement is out of the scope of this paper. The Gaussian radial function is uniquely fit for this setup since it is the only one that is separable in $x$ and $y$; i.e. $e^{-\epsilon^2(x-x_n)^2+(y-y_n)^2} = e^{-\epsilon^2(x-x_n)^2}e^{-\epsilon^2(y-y_n)^2}$. So, since the number of centers in $x$ and in $y$ grow as $O(\sqrt{N})$, we choose $\epsilon = c\sqrt{N}$. An optimal $c$ value for the two-dimensional case is also out of the scope of this study.

An example of this approximation scheme is shown in Fig. 8.
Figure 9: The figure shows approximation error and coefficient norm as a function of the number of degrees of freedom $N$ for the solution of the one-dimensional ODE $-u''(x) = 20 \frac{1-30x^2}{(1+10x^2)^3}$ with true solution $f(x) = \frac{1}{1+10x^2}$ (left panels), and the two-dimensional PDE $-\Delta u = 40 \frac{1-10(x^2+y^2)}{(1+10(x^2+y^2))^3}$ with true solution $f(x,y) = \frac{1}{1+10(x^2+y^2)}$ (right panels). We chose the threshold $\tau = 1e^{-15}$. In both cases high accuracy is achieved and maintained for larger values of $N$.

6 A least squares collocation scheme

We conclude the paper with some experiments to indicate the usefulness of radial basis functions in combination with least squares for the solution of elliptic boundary value problems. To that end we formulate an RBF based collocation scheme, inspired by Kansa’s method [20, 21], using more collocation points than degrees of freedom.

We wish to solve

$$-\Delta u = f, \text{ in } \Omega$$

subject to the Dirichlet boundary condition

$$u = g, \text{ on } \partial \Omega.$$  

We choose $M_\Omega$ points $\{x_m\}$ inside $\Omega$, and an additional $M_{\partial \Omega}$ points $\{y_m\}$ on the boundary. This results in a least squares system once $M = M_\Omega + M_{\partial \Omega} > N$.

We set out to solve the linear system $B\lambda = b$, where

$$B = \begin{pmatrix} -\Delta \phi_n(x_m)_{m,n=1}^{M_\Omega,N} \\ \phi_n(y_m)_{m,n=1}^{M_{\partial \Omega},N} \end{pmatrix} \in \mathbb{R}^{M \times N}, \quad b = \begin{pmatrix} f(x_m)_{m=1}^{M_\Omega} \\ g(y_m)_{m=1}^{M_{\partial \Omega}} \end{pmatrix} \in \mathbb{R}^M.$$

The upper block of the linear system is precisely the least squares matrix for the approximation of $f$ on $\Omega$, using the basis functions $-\Delta \phi_n$. The lower block is the least squares discretization of the boundary condition on $\partial \Omega$.

The approximate solution is therefore

$$f(x) \approx \sum_{n=1}^{N} \lambda_n \phi(\varepsilon(x - \xi_n)).$$

Results are shown in Fig. [9] for a one-dimensional and two-dimensional Poisson equation. We have chosen linear scaling $\varepsilon = cN$ using the proportionality constant [21] in 1D, and scaling of the form $c\sqrt{N}$ for the 2D problem, exactly as in §5.5 and similarly using a hexagonal grid.
The domain resembles the number $\pi$ and is based on a smooth parameterization of its boundary using a Fourier series of length 32. The boundary values are derived from the given analytical solution. Other parameters, such as the number of points $M_\Omega$ and $M_{\partial\Omega}$, have been chosen heuristically to ensure sufficient oversampling.

A noteworthy observation is that the results in Fig. 9 demonstrate high accuracy, which is maintained for larger values of $N$. This contrasts with other methods for PDEs based on global RBF approximations since Kansa, e.g. [6, §3.7] and [13, §4.1].

7 Concluding remarks

The analysis and numerical results of this paper have confirmed the possibility of accurate and stable approximations using the Gaussian RBF in 1D. This leaves several interesting options for further research, such as the extension of the analysis to 2D and higher-dimensional problems and to other kinds of radial basis functions. An interesting computational challenge is the rapid solution of the rectangular linear systems. Finally, it is clear that in the application of RBF’s to partial differential equations much remains to be explored.

Acknowledgements

We would like to thank Nick Trefethen for discussions on the topic of this paper, as well as for useful comments on this manuscript. This research was initiated during an inspiring colloquium meeting at Michigan Tech, for which we like to thank the institution.

References

[1] B. Adcock and D. Huybrechs. On the resolution power of Fourier extensions for oscillatory functions. J. Comput. Appl. Math., 260:312–336, 2014.

[2] B. Adcock and D. Huybrechs. Frames and numerical approximation. SIAM Review, 61:443–473, 2019.

[3] B. Adcock and D. Huybrechs. Frames and numerical approximation II: generalized sampling. J. Fourier Anal. Appl., 26(87):1–34, 2020.

[4] B. Adcock and M. Seifi. Frame approximation with bounded coefficients. Adv. Comput. Math., 47:25, 2021.

[5] J. P. Boyd. Error saturation in Gaussian radial basis functions on a finite interval. J. Comput. Appl. Math., 234(5):1435–1441, 2010.

[6] W. Chen, Z.-J. Fu, and C. S. Chen. Recent advances in radial basis function collocation methods. Springer, Heidelberg, 2014.

[7] O. Christensen. An Introduction to Frames and Riesz Bases. Birkhäuser, 2nd edition, 2016.

[8] NIST Digital Library of Mathematical Functions. http://dlmf.nist.gov/, Release 1.0.24 of 2019-09-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F.
Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.

[9] T. A. Driscoll and B. Fornberg. Interpolation in the limit of increasingly flat radial basis functions. *Comput. Math. Appl.*, 43:413–422, 2002.

[10] G. Fasshauer. *Meshfree Approximation Methods with MATLAB*. Interdisciplinary mathematical sciences. World Scientific, 2007.

[11] G. E. Fasshauer and M. J. McCourt. Stable evaluation of Gaussian radial basis function interpolants. *SIAM J. Sci. Comput.*, 34(2):737–762, 2012.

[12] B. Fornberg and N. Flyer. *A primer on Radial Basis Functions with Applications to the Geosciences*. SIAM, Philadelphia, 2015.

[13] B. Fornberg and N. Flyer. Solving PDEs with radial basis functions. *Acta Numerica*, 24:215–258, 2015.

[14] B. Fornberg, E. Larsson, and N. Flyer. Stable computations with Gaussian radial basis functions. *SIAM J. Sci. Comput.*, 33(2):869–892, 2011.

[15] B. Fornberg and C. Piret. A stable algorithm for flat radial basis functions on a sphere. *SIAM J. Sci. Comput.*, 30(1):60–80, 2007.

[16] B. Fornberg and C. Piret. On choosing a radial basis function and a shape parameter when solving a convective PDE on a sphere. *J. Comput. Phys.*, 227(5):2758–2780, 2008.

[17] B. Fornberg and G. Wright. Stable computation of multiquadric interpolants for all values of the shape parameter. *Comput. Math. Appl.*, 48(5):853–867, 2004.

[18] F. Frykland, E. Lehto, and A.-K. Tornberg. Partition of unity extension of functions on complex domains. *J. Comput. Phys.*, 375:57–79, 2018.

[19] B. Hashemi, Y. Nakatsukasa, and L. N. Trefethen. Rectangular eigenvalue problems. *Adv. Comp. Math.*, 48:1–16, 2022.

[20] E. Kansa. Multiquadrics – A scattered data approximation scheme with applications to computational fluid-dynamics – I surface approximations and partial derivative estimates. *Comput. Math. Appl.*, 19(8):127–145, 1990.

[21] E. Kansa. Multiquadrics – A scattered data approximation scheme with applications to computational fluid-dynamics – II solutions to parabolic, hyperbolic and elliptic partial differential equations. *Comput. Math. Appl.*, 19(8):147–161, 1990.

[22] V. Maz’ya and G. Schmidt. On approximate approximations using Gaussian kernels. *IMA J. Numer. Anal.*, 16(1):13–29, 1996.

[23] C. Piret. A radial basis function based frames strategy for bypassing the Runge phenomenon. *SIAM J. Sci. Comput.*, 38(4):A2262–A2282, 2016.

[24] R. B. Platte. How fast do radial basis function interpolants of analytic functions converge? *IMA J. Numer. Anal.*, 31(4):1578–1597, 2011.
[25] R. B. Platte and T. A. Driscoll. Polynomials and potential theory for Gaussian radial basis function interpolation. *SIAM J. Numer. Anal.*, 43(2):750–766, 2005.

[26] J. Rashidinia, G. Fasshauer, and M. Khasi. A stable method for the evaluation of Gaussian radial basis function solutions of interpolation and collocation problems. *Comput. Math. Appl.*, 72(1):178–193, 2016.

[27] R. Schaback. Error estimates and condition numbers for radial basis function interpolation. *Adv. Comput. Math.*, 3(3):251–264, 1995.

[28] G. B. Wright and B. Fornberg. Stable computations with flat radial basis functions using vector-valued rational approximations. *J. Comput. Phys.*, 331:137–156, 2017.