Improved Estimates for Condition Numbers of Radial Basis Function Interpolation Matrices

Benedikt Diederichs and Armin Iske

University of Hamburg, Department of Mathematics, Bundesstraße 55, 20146 Hamburg, Germany

{benedikt.diederichs,armin.iske}@uni-hamburg.de

Abstract

We improve existing estimates for the condition number of matrices arising in radial basis function interpolation. To this end, we refine lower bounds on the smallest eigenvalue and upper bounds on the largest eigenvalue, where our upper bounds on the largest eigenvalue are independent of the matrix dimension (i.e., the number of interpolation points). We show that our theoretical results comply with recent numerical observations concerning the condition number of radial basis function interpolation matrices.

Keywords: Radial basis functions, multivariate scattered data interpolation, matrix condition numbers.

2010 MSC: 65D05, 41A05, 41A63

1. Introduction

Radial basis functions (RBF), or, radial kernel functions, are powerful tools for meshfree interpolation from multivariate scattered data [7, 8, 21]. To motivate the RBF interpolation method we consider the scattered data interpolation problem: on given values $f_X = (f(x_1), \dots, f(x_N))^T \in \mathbb{R}^N$ taken from a function $f : \mathbb{R}^d \to \mathbb{R}$, $d \geq 1$, at a scattered set $X = \{x_1, \dots, x_N\} \subset \mathbb{R}^d$ of pairwise distinct points, find a function $s : \mathbb{R}^d \to \mathbb{R}$ satisfying the interpolation conditions $s_X = f_X$, i.e.,

$$s(x_k) = f(x_k) \qquad \text{for } k = 1, \dots, n. \tag{1}$$

According to the RBF interpolation scheme, the interpolant s in (1) is required to be of the form

$$s(x) = \sum_{j=1}^{N} c_j \Phi(x - x_j) + p(x) \qquad \text{for } x \in \mathbb{R}^d, \tag{2}$$

where Φ is a radially symmetric kernel function, i.e., $\Phi \equiv \phi(\|\cdot\|_2)$ for a radial kernel $\phi : [0,\infty) \to \mathbb{R}$, and where $\|\cdot\|_2$ is the Euclidean norm on $\mathbb{R}^d$.

To be more precise, we require that Φ is *conditionally positive definite* of order m on $\mathbb{R}^d$, $\Phi \in \mathbf{CPD}(m)$, i.e., for any point set $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^d$ of size $N = |X|$, the *kernel matrix*

$$A_{\Phi,X} = (\Phi(x_j - x_k))_{1 \le j,k \le N} \in \mathbb{R}^{N \times N} \tag{3}$$

is positive definite on the linear subspace

$$L_X = \left\{ c = (c_1, \ldots, c_N)^T \in \mathbb{R}^N \, : \sum_{j=1}^{N} c_j p(x_j) = 0 \text{ for all } p \in \Pi_{m-1}^d \right\} \subset \mathbb{R}^N,$$

where Π_{m-1}^d is the space of all d-variate polynomials of degree at most $m-1$. Moreover, p in (2) is in this case an element of the polynomial space Π_{m-1}^d.

Note that for $m = 0$ we have $L_X = \mathbb{R}^N$ and $\Pi_{m-1}^d = \{0\}$. In this case we say that Φ is *positive definite* on $\mathbb{R}^d$, i.e., $\Phi \in \mathbf{PD} := \mathbf{CPD}(0)$.

We remark that for $\Phi \in \mathbf{CPD}(m)$ the interpolation problem $s_X = f_X$ has a unique solution s of the form (2), under constraints $c = (c_1, \ldots, c_N)^T \in L_X$ and under the assumption that the interpolation points X are Π_{m-1}^d-unisolvent. In fact, the interpolation problem $s_X = f_X$ leads us, for a fixed basis $\{p_1, \ldots, p_Q\}$ of Π_{m-1}^d, to the linear system

$$\begin{aligned} \sum_{j=1}^{N} c_j \Phi(x_k - x_j) + \sum_{\ell=1}^{Q} b_\ell p_\ell(x_k) &= f(x_k) \qquad \text{for } k = 1, \ldots, N \\ \sum_{j=1}^{N} c_j p_\ell(x_j) &= 0 \qquad\qquad \text{for } \ell = 1, \ldots, Q \end{aligned} \tag{4}$$

with unknown coefficients $c = (c_1, \ldots, c_N)^T \in \mathbb{R}^N$ for the major part of p in (2) and coefficients $b = (b_1, \ldots, b_Q)^T \in \mathbb{R}^Q$ for its polynomial part. We can rewrite the system (4) in matrix form as

$$\begin{pmatrix} A_{\Phi,X} & P_X \\ P_X^T & 0 \end{pmatrix} \begin{pmatrix} c \\ b \end{pmatrix} = \begin{pmatrix} f_X \\ 0 \end{pmatrix}, \tag{5}$$

where the *polynomial matrix*

$$P_X = (p_\ell(x_j))_{1 \le j \le N; 1 \le \ell \le Q} \in \mathbb{R}^{N \times Q}$$

is injective, due to the assumed Π_{m-1}^d-unisolvence of the interpolation points X.

Now the goal of this paper is to investigate the numerical stability of the linear system (5). Note that for $\Phi \in \mathbf{CPD}(m)$ there are positive eigenvalues $\Lambda \ge \lambda > 0$ of $A_{\Phi,X}$ satisfying

$$\lambda \|c\|_2^2 \le c^T A_{\Phi,X} c \le \Lambda \|c\|_2^2 \qquad \text{for all } c \in L_X,$$

where the *spectral condition number* $\kappa_2(A_{\Phi,X}) = \Lambda/\lambda$ accounts for the sensitivity of the system (5), see [18]. Therefore, to analyze the numerical stability of (5), we are interested in both *good* lower bounds for λ and *good* upper bounds for Λ.

We remark that the analysis of spectral properties of RBF matrices already has a long history. The first results are due to Ball [2, 3] in 1992 and these were subsequently refined by Narcowich & Ward in [13, 14, 15] and Schaback in [16, 17]. A concise account was provided by Schaback in [18], with focus on lower bounds for λ. Although upper bounds for Λ, relying on Gerschgorin's theorem, are also provided in [18], these are rather crude.

Very recently, numerical estimates for the condition number of $A_{\Phi,X}$ were obtained by Boyd & Gildersleeve [6]. But there is still a large gap between the bounds from their numerical experiments and the theoretical bounds in the RBF literature. For the special case of uniform grids, *optimal* bounds on λ and Λ are due to Baxter [4], where there is also a large gap between his bounds in [4] for gridded data and those bounds from the RBF literature for scattered data.

This paper reduces existing gaps between those bounds quite significantly. To explain this further, we remark that existing lower bounds on λ are of the form

$$\lambda \ge G(q), \tag{6}$$

where

$$q \equiv q(X, \|\cdot\|_2) = \frac{1}{2} \min_{j \neq k} \|x_j - x_k\|_2$$

is the (Euclidean) *separation radius* of X, and where $G : [0, \infty) \to [0, \infty)$ is a monotonically increasing function depending on Φ, but not on X and not on $N = |X|$. Due to the monotonicity of G and the norm equivalence in finite dimensional normed linear spaces, it is clearly possible to express the bound in (6) for separation radii in other norms on $\mathbb{R}^d$. In fact, it will be convenient to express our results w.r.t the separation radius in the ∞-norm,

$$q_\infty \equiv q(X, \|\cdot\|_\infty) := \frac{1}{2} \min_{j \neq k} \|x_j - x_k\|_\infty.$$

To discuss one of our results only very briefly, we consider the *Gaussian* kernel

$$\Phi(x) = e^{-\beta\|x\|_2^2} \qquad \text{for } \beta > 0.$$

In this case, the *best* bound on λ of the form (6) known so far is given by

$$G(q) = C_d q^{-d} e^{-M_d^2/(\beta q^2)},$$

where C_d and $M_d^2 = 40.71d^2$ are constants depending on d (cf. [21, Table 12.1]). In this paper, we will reduce this best bound (up to an arbitrarily small $\varepsilon > 0$) to $M_d = d\pi/4$ for the Euclidean separation radius q and to $M_d = \sqrt{d}\pi/4$ for the separation radius q_∞. We detail our results by Example 2.6.

The outline of this paper is as follows. We improve the current best-known bounds on the spectral condition number $\kappa_2(A_{\Phi,X}) = \Lambda/\lambda$ by refining existing lower bounds of the form (6) for the smallest eigenvalue λ (in Section 2), on the one hand, and by improving existing upper bounds for the largest eigenvalue Λ (in Section 3), on the other hand. The latter leads us to refined upper bounds on Λ that are independent of $N = |X|$. In Section 4 we finally discuss the theoretical bounds of this paper in comparison with bounds obtained from the numerical experiments in [6].

2. Improved lower bounds for the smallest eigenvalue

In our subsequent analysis, we assume that $\Phi : \mathbb{R}^d \to \mathbb{R}$ is continuous and of at most polynomial growth around infinity, so that Φ has a generalized Fourier transform $\hat{\Phi}$ in the sense of tempered distributions [9, 10]. Moreover, we assume that $\hat{\Phi}$ is non-vanishing and non-negative on $\mathbb{R}^d \setminus \{0\}$, allowing for an algebraic singularity at zero. In this case, it is well-known that the identity

$$\sum_{j,k=1}^{N} c_j c_k \Phi(x_j - x_k) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{\Phi}(\omega) \left| \sum_{j=1}^{N} c_j e^{i x_j \cdot \omega} \right|^2 d\omega \tag{7}$$

holds for all $(c_1, \ldots, c_N)^T \in L_X$ (see e.g. [21, Corollary 8.13]).

We remark that the above assumptions on Φ and $\hat{\Phi}$ are highly relevant as they hold for large classes of radial kernel functions (cf. [21, Section 8.2]). Now the identity (7) plays a central role in the ground-breaking work of Narcowich and Ward [13, 14, 15], where they construct a suitable *minorant* $\Psi : \mathbb{R}^d \to \mathbb{R}$, whose Fourier transform $\hat{\Psi} : \mathbb{R}^d \to \mathbb{R}$ satisfies

$$\hat{\Phi}(\omega) \geq \hat{\Psi}(\omega) \geq 0 \qquad \text{for all } \omega \in \mathbb{R}^d \setminus \{0\}, \tag{8}$$

so that they can conclude the estimate

$$\begin{aligned} \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{\Phi}(\omega) \left| \sum_{j=1}^{N} c_j e^{i x_j \cdot \omega} \right|^2 d\omega \quad &\geq \quad \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{\Psi}(\omega) \left| \sum_{j=1}^{N} c_j e^{i x_j \cdot \omega} \right|^2 d\omega \\ &= \quad \sum_{j,k=1}^{N} c_j c_k \Psi(x_j - x_k). \end{aligned}$$

Using Gerschgorin's theorem, the smallest eigenvalue λ of the kernel matrix $A_{\Phi,X}$ in (3) could be bounded from below by

$$\lambda \geq \min_{1 \leq k \leq N} \left(\Phi(0) - \sum_{\substack{j=1 \\ j \neq k}}^{N} |\Phi(x_j - x_k)| \right). \tag{9}$$

Note that the estimate for λ in (9) can only be rather crude, unless $A_{\Phi,X}$ is diagonally dominant. But Narcowich & Ward were able to construct a minorant

Ψ satisfying (8), such that the matrix

$$B_{\Psi,X} = \left(\Psi(x_j - x_k)\right)_{1\le j,k\le N} \in \mathbb{R}^{N\times N}$$

is diagonally dominant. We remark that this construction can also be found in the work [2] of Ball for the case $\Phi(x) = \|x\|_2$. This has later inspired Baxter [5] to extend Ball's techniques to polyharmonic splines.

The following result of Komornik & Loreti [11, Theorem 8.1] leads us to improved lower bounds on λ by dropping the restriction $\hat{\Psi} \ge 0$ in (8), in which case it is possible to construct Ψ such that $B_{\Psi,X}$ is diagonal. To this end, we let

$$B_r^p := \left\{x \in \mathbb{R}^d : \|x\|_p \le r\right\} \subset \mathbb{R}^d \qquad \text{for } r > 0 \text{ and } 1 \le p \le \infty$$

denote the closed ball around zero with radius r with respect to the p-norm. It will be convenient to let $B_r := B_r^2 = \{x \in \mathbb{R}^d : \|x\|_2 \le r\}$ for $r > 0$. Moreover, $H^1(\Omega) \equiv W^{1,2}(\Omega)$ denotes the usual Sobolev space on $\Omega \subset \mathbb{R}^d$ consisting of all functions $u \in L^2(\Omega)$, all of whose first order weak derivatives are in $L^2(\Omega)$, and $H_0^1(\Omega) \equiv W_0^{1,2}(\Omega)$ is the topological closure of $C_0^\infty(\Omega)$ in $H^1(\Omega)$. For more details concerning the Sobolev spaces $W^{m,p}(\Omega)$ we refer to [1, Chapter III].

Theorem 2.1 (Komornik & Loreti). *For $q > 0$ and $1 \le p \le \infty$ let $\mu_p > 0$ be the smallest eigenvalue of $-\Delta$ in the Sobolev space $H_0^1(B_q^p)$. Moreover, let $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^d$ be a finite point set with separation radius $q_p \ge q$. Then, the inequality*

$$\int_{B_R} \left| \sum_{j=1}^N c_j e^{i x_j \cdot \omega} \right|^2 d\omega \ge k_p(q,R)\|c\|_2^2 \qquad \forall c \in \mathbb{R}^N$$

holds for every $R > \sqrt{\mu_p}$, where $k_p(q,R)$ is a positive constant depending only on p, q, d and R, but not on X or $N = |X|$. □

To obtain improved lower bounds on λ, the key idea is the construction of a minorant Ψ whose support $\mathrm{supp}(\Psi)$ is contained in the (Euclidean) ball B_{2q}, i.e., $\mathrm{supp}(\Psi) \subset B_{2q}$, and whose Fourier transform $\hat{\Psi}$ satisfies

$$\chi_{B_R}(\omega) \ge \hat{\Psi}(\omega) \qquad \text{for all } \omega \in \mathbb{R}^d, \tag{10}$$

where χ_{B_R} is the indicator function of B_R. In this case, we have, for any $c \in \mathbb{R}^N$,

$$\begin{aligned}\int_{B_R} \left| \sum_{j=1}^N c_j e^{ix_j \cdot \omega} \right|^2 d\omega &\geq \int_{B_R} \left| \sum_{j=1}^N c_j e^{ix_j \cdot \omega} \right|^2 \hat{\Psi}(\omega)\, d\omega \\ &= (2\pi)^d \sum_{j,k=1}^N c_j c_k \Psi(x_j - x_k) \\ &= (2\pi)^d \Psi(0) \|c\|_2^2. \qquad (11)\end{aligned}$$

Now we use the result of Theorem 2.1 to improve existing lower bounds on λ.

Theorem 2.2. *Let $X \subset \mathbb{R}^d$ be a finite point set with separation radius $q_p \geq q$. Then, the bound*

$$\lambda \geq (2\pi)^{-d} k_p(q,R) \varphi_0(R/2) \qquad (12)$$

holds for every $R > \sqrt{\mu_p}$, where φ_0 is the monotonically decreasing function

$$\varphi_0(r) := \inf_{\|\omega\|_2 \leq 2r} \hat{\Phi}(\omega).$$

Proof: For the kernel matrix $A_{\Phi,X}$ in (3), where $X = \{x_1, \ldots, x_N\}$, by using the representation (7) and the assumption $\hat{\Phi} \geq 0$, we obtain the inequality

$$(2\pi)^d c^T A_{\Phi,X} c = \int_{\mathbb{R}^d} \hat{\Phi}(\omega) \left| \sum_{j=1}^N c_j e^{ix_j \cdot \omega} \right|^2 d\omega \geq \varphi_0(R/2) \int_{B_R} \left| \sum_{j=1}^N c_j e^{ix_j \cdot \omega} \right|^2 d\omega$$

for all $c \in L_X$, so that the stated bound in (12) follows from Theorem 2.1. □

We remark that for $p = 2$ the values for μ_2 and $k_2(q,R)$ are rather difficult to compute. For μ_2, we find the identity $\mu_2^2 = \rho_d/q$, where ρ_d is the first positive root of the Bessel function $J_{d/2-1}$ (cf. the second remark after [11, Theorem 8.1]), whereas $k_p(q,R)$ has (to the best of our knowledge) not been calculated so far.

But the case $p = \infty$ is much easier, where we obtain the following results.

Corollary 2.3. *Let $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^d$ be a finite point set with separation radius q_∞. Then, for any $R > \sqrt{d}\pi/(2q_\infty)$ we have the estimate*

$$\int_{B_R} \left| \sum_{j=1}^N c_j e^{ix_j \cdot \omega} \right|^2 d\omega \geq \left(1 - \frac{d\pi^2}{4q_\infty^2 R^2}\right) q_\infty^{-d} (2\pi)^d \frac{\pi^{2d}}{4^{2d}} \|c\|_2^2 \qquad \forall c \in \mathbb{R}^N.$$

Proof: The eigenfunction of $-\Delta$ in $H_0^1(B_{q_\infty}^\infty)$ corresponding to the smallest eigenvalue $\mu_\infty = d\pi^2/(4q_\infty^2) > 0$ is given by

$$H(x) = \prod_{j=1}^{d} \cos\left(\frac{\pi x_j}{2q_\infty}\right) \qquad \text{for } x \in B_{q_\infty}^\infty.$$

We extend H to a function on $\mathbb{R}^d$ by letting $H(x) = 0$ for $x \in \mathbb{R}^d \setminus B_{q_\infty}^\infty$. Then, the function $G = (R^2 + \Delta)(H * H)$ has compact support with $\mathrm{supp}(G) \subset B_{2q_\infty}^\infty$. We denote the Fourier transform of G by g, so that

$$g(\omega) = (R^2 - \|\omega\|_2^2)\hat{H}^2(\omega) \qquad \text{for all } \omega \in \mathbb{R}^d.$$

This is exactly the construction of Komornik and Loreti. Note that $g(\omega) \leq g(0)$ for all $\omega \in \mathbb{R}^d$, since $\hat{H}^2 > 0$. We define $\hat{\Psi}$ as required in (10) by

$$\hat{\Psi}(\omega) := g(\omega)/g(0) \leq \chi_{B_R}(\omega) \qquad \text{for } \omega \in \mathbb{R}^d.$$

By rather elementary calculations, we obtain

$$\begin{aligned} g(0) &= R^2 \left(\int_{\mathbb{R}^d} H(x)\,dx\right)^2 = R^2 q_\infty^{2d} \frac{4^{2d}}{\pi^{2d}} \\ G(0) &= (R^2 - \mu_\infty) \int_{\mathbb{R}^d} H^2(x)\,dx = \left(R^2 - \frac{d\pi^2}{4q_\infty^2}\right) q_\infty^d \end{aligned}$$

and, finally, by using (11),

$$k_\infty(q,R) = (2\pi)^d \Psi(0) = \left(1 - \frac{d\pi^2}{4q_\infty^2 R^2}\right) q_\infty^{-d} (2\pi)^d \frac{\pi^{2d}}{4^{2d}}.$$

□

We can now give explicit constants in the lower bound (12) for λ.

Corollary 2.4. *Under the assumptions of Theorem 2.2 we have the estimate*

$$\lambda \geq \frac{2d^{d/2}}{d+2} \frac{\pi^{2d}}{4^{2d}} \varphi_0\left(\frac{(d+1)\pi}{4q}\right) q^{-d} = C(d)\varphi_0\left(\frac{(d+1)\pi}{4q}\right) q^{-d}. \tag{13}$$

Proof: Note that $q_\infty \geq q/\sqrt{d}$, i.e., X is separated by $q/\sqrt{d}$ in the ∞-norm. Applying Corollary 2.3 with $R = \sqrt{d(d+2)}\pi/(2q)$ and by using the same idea as in the proof of Theorem 2.2 we obtain the stated estimate (13), where we used the inequality $d+1 > \sqrt{d(d+2)}$ and the monotonicity of φ_0. □

Remark 2.5. *If we choose* $R = (d\pi+\varepsilon)/(2q)$ *in the application of Corollary 2.3 (rather than* $R = \sqrt{d(d+2)}\pi/(2q)$*), then the lower bound for* λ *in (13) becomes*

$$\lambda \geq C(\varepsilon, d)\varphi_0\left(\frac{d\pi+\varepsilon}{4q}\right)q^{-d}$$

for some constant $C(\varepsilon, d) > 0$ *satisfying* $C(\varepsilon, d) = \mathcal{O}(\varepsilon)$*, for* $\varepsilon \to 0$.

If the separation radius q_∞ *in the* ∞*-norm is of interest, then this simplifies our calculations. In this case, a direct application of Corollary 2.4 gives the bound*

$$\lambda \geq \frac{1}{d+1}\frac{\pi^{2d}}{4^{2d}}\varphi_0\left(\frac{\sqrt{d+1}\pi}{4q_\infty}\right)q_\infty^{-d}.$$

With letting $R = (\sqrt{d}\pi+\varepsilon)/(2q_\infty)$ *(in the application of Corollary 2.3), we find*

$$\lambda \geq C(\varepsilon, d)\varphi_0\left(\frac{\sqrt{d}\pi+\varepsilon}{4q_\infty}\right)q_\infty^{-d}$$

for some constant $C(\varepsilon, d) > 0$ *satisfying* $C(\varepsilon, d) = \mathcal{O}(\varepsilon)$*, for* $\varepsilon \to 0$. □

Let us finally consider relevant examples. For the situation of radial kernels Φ whose (generalized) Fourier transforms $\hat{\Phi}$ have algebraic decay around infinity, e.g. for polyharmonic splines and for compactly supported RBFs, the *optimal* asymptotic order for the estimate (6) is already known (see the discussion around [21, Corollary 12.8]), where the optimal orders, as stated in [21, Table 12.1] comply with our results. But for the Gaussians and for the multiquadrics, our result in Corollary 2.4 improves the currently known best lower bounds on λ. Details are given in the following two examples.

Example 2.6. *The Fourier transform of the Gaussian* $\Phi(x) = e^{-\beta\|x\|_2^2}$, $\beta > 0$, *is given as* $\hat{\Phi}(w) = (\pi/\beta)^{d/2}e^{-\|w\|_2^2/(4\beta)}$*. Therefore, we have*

$$\varphi_0(R) = (\pi/\beta)^{d/2}e^{-R^2/\beta}.$$

Using Corollary 2.4, we obtain the estimate

$$\lambda \geq C(d)q^{-d}\left(\frac{\pi}{\beta}\right)^{d/2}e^{-\frac{(d+1)^2\pi^2}{16q^2\beta}}.$$

Following along the lines of Remark 2.5 concerning the choice of R, we get

$$\lambda \geq C(\varepsilon, d) q^{-d} \left(\frac{\pi}{\beta}\right)^{d/2} e^{-\frac{(d\pi+\varepsilon)^2}{16q^2\beta}} \qquad \text{for } \varepsilon > 0 \tag{14}$$

for some constant $C(\varepsilon, d) > 0$ satisfying $C(\varepsilon, d) = \mathcal{O}(\varepsilon)$, for $\varepsilon \to 0$. □

Example 2.7. *For the (inverse) multiquadrics $\Phi(x) = (\gamma^2 + \|x\|_2^2)^{\beta/2}$, where $\beta \in \mathbb{R} \setminus 2\mathbb{N}$ and $\gamma \neq 0$, we have*

$$\varphi_0(R) \geq \tilde{C}(d, \beta) \gamma^{(\beta+d-1)/2} e^{-2\gamma R} R^{-(\beta+d+1)/2},$$

with an explicitly known constant $\tilde{C}(d, \gamma, \beta)$ (see [21, Corollary 12.5]).

From Corollary 2.4, we obtain

$$\lambda \geq C(d, \beta) \gamma^{(\beta+d-1)/2} q^{-(d-\beta-1)/2} e^{-2\gamma(d+1)\pi/4q}$$

for some constant $C(d, \beta) > 0$. □

3. Improved upper bounds for the largest eigenvalue

Now let us turn to upper bounds for the largest eigenvalue Λ of the kernel matrix $A_{\Phi,X}$ in (3). To this end, we restrict ourselves to $\Phi \in \mathbf{PD}$. It has been common practice to estimate Λ by using the Gerschgorin theorem,

$$\Lambda \leq \max_{1 \leq k \leq N} \sum_{j=1}^{N} |\Phi(x_k - x_j)|, \tag{15}$$

whose straightforward application immediately yields the estimate $\Lambda \leq N\Phi(0)$. But this estimate can only be very crude, as the Gerschgorin theorem gives good estimates (15) only for diagonally dominant matrices. On the other hand, the other estimate $\Lambda \leq N\Phi(0)$ can only be good for $\Phi(x_k - x_j) \approx \Phi(0)$, in which case $A_{\Phi,X}$ is not at all diagonally dominant.

Moreover, note that the estimate $\Lambda \leq N\Phi(0)$ depends on the size $N = |X|$ of the point set X, but not on the separation radius q. To combine upper bounds on Λ with our lower bounds on λ (from the previous section), we want to trade the dependence on N for a dependence on q. To this end, one can evaluate the sum of the upper bound in (15) more carefully, and this has been done in [15].

Lemma 3.1. *Let $\varphi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ be a monotonically decreasing function, and let $X \subset \mathbb{R}^d$ be a finite point set with separation radius q. Then, we have*

$$\sum_{x \in X} \varphi(\|y - x\|_2) \leq \varphi(0) + \sum_{n=1}^{\infty} 3d(n+2)^{d-1}\varphi(nq) \qquad \text{for every } y \in X. \quad (16)$$

Proof: For $\varphi(r) = e^{-\beta r^2}$ this is covered by [15, Lemma 2.1]. The general case works in exactly the same way. □

Remark 3.2. *The bound given in* (16) *is finite if and only if*

$$\Phi(x) = \varphi(\|x\|) \in L^1(\mathbb{R}^d).$$

In this case, $\hat{\Phi}$ is continuous on $\mathbb{R}^d$. As shown in [15] for $\Phi \in$ **PD** *on $\mathbb{R}^d$ for all $d \in \mathbb{N}$, an upper bound for Λ can only be obtained independently of X and $N = |X|$, iff $\Phi \in L^1(\mathbb{R}^d)$.* □

For the Gaussian kernel, where $\varphi(r) = e^{-\beta r^2}$, for $\beta > 0$, the infinite sum on the right hand side in (16) is convergent. More explicitly, in this case we obtain the bound

$$\Lambda \leq 1 + 3d \sum_{n=1}^{\infty} (n+2)^{d-1} e^{-\beta n^2 q^2}. \quad (17)$$

For the multiquadrics kernel, where $\varphi(r) = (\gamma^2 + r^2)^{\beta/2}$ for $\beta \in \mathbb{R} \setminus 2\mathbb{N}$ and $\gamma \neq 0$, the infinite sum on the right hand side in (16) is convergent for $\beta < -d$, and the resulting upper bound on Λ in (16) is independent of N. But for $\beta \geq -d$ there is no upper bound on Λ independent of N (see Remark 3.2).

Although we can obtain bounds on Λ via (16) that are independent of N, they are rather crude. Moreover, the infinite sum in (16) is usually difficult to evaluate. A more comprehensive analysis concerning the asymptotic behaviour of the sum in (16) is given in [15].

We now derive upper bounds on Λ by using Fourier techniques that are similar to those for our lower bounds on λ from the previous section. We start with the following upper bound for exponential sums.

Lemma 3.3. *Let $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^d$ be a finite point set with separation radius q_∞. Then, the estimate*

$$\int_{[-R,R]^d} \left| \sum_{j=1}^{N} c_j e^{ix_j \cdot \omega} \right|^2 d\omega \leq \left(2R + \frac{\pi}{q_\infty} \right)^d \|c\|_2^2 \qquad \forall c \in \mathbb{R}^N \tag{18}$$

holds for every $R > 0$.

Proof: The special case $d = 1$ goes back to Selberg [19], see also [20]. Selberg constructs a function $\psi : \mathbb{R} \to \mathbb{R}$ satisfying

$$\psi(x) \geq \chi_{[-R,R]}, \qquad \operatorname{supp}(\hat{\psi}) \subset [-q_\infty, q_\infty], \qquad \psi(0) = 2R + \pi/q_\infty.$$

The cases $d > 1$, related to [12], are treated as follows. By basic calculations, we find that the tensor

$$\psi_\otimes(x_1, \ldots, x_d) := \psi(x_1) \cdots \psi(x_d) \geq \chi_{[-R,R]^d}(x_1, \ldots, x_d)$$

is localized in the frequency domain in $[-q_\infty, q_\infty]^d$ and $\psi_\otimes(0) = (2R + \pi/q_\infty)^d$, so that we can proceed as in the proof of Theorem 2.1. □

Note that the bound in (18) continues to hold, if we translate $[-R, R]^d$ by an arbitrary vector. This observation leads us to the following key result.

Theorem 3.4. *Let $\hat{\Phi} : \mathbb{R}^d \to \mathbb{R}$ be a continuous function with monotonically decreasing $\hat{\Phi}(x) = \varphi(\|x\|_2)$. Then, for $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^d$ with separation radius q_∞ the estimate*

$$\int_{\mathbb{R}^d} \left| \sum_{j=1}^{N} c_j e^{ix_j \cdot \omega} \right|^2 \hat{\Phi}(\omega)\, d\omega \leq 2^d \left(R + \frac{\pi}{q_\infty} \right)^d \sum_{k \in \mathbb{N}_0^d} \hat{\Phi}(Rk) \|c\|_2^2 \qquad \forall c \in \mathbb{R}^N \tag{19}$$

holds for every $R > 0$.

Proof: We regard the point sets

$$Q_k = \{Rk + R(\varepsilon_1, \ldots, \varepsilon_d)^T : \varepsilon_j \in [0, 1)\} \qquad \text{for } k = (k_1, \ldots, k_d) \in \mathbb{N}_0^d.$$

Note that $\mathbb{R}^d_{\geq 0} = \bigcup_{k \in \mathbb{N}_0^d} Q_k$. Therefore, we can cover $\mathbb{R}^d$ by the union of

$$Q_{k,j} = \operatorname{diag}(j) Q_k \qquad \text{for } k \in \mathbb{N}_0^d \text{ and } j \in \{\pm 1\}^d.$$

Note that $\min_{Q_{k,j}} \|x\|_2 = R\|k\|_2$.

Now by using Lemma 3.3, we obtain for any $c \in \mathbb{R}^N$ the estimate

$$\begin{aligned}
\int_{\mathbb{R}^d} \left| \sum_{j=1}^N c_j e^{ix_j \cdot \omega} \right|^2 \hat{\Phi}(\omega)\, d\omega &= \sum_{\substack{k \in \mathbb{N}_0^d \\ j \in \{\pm 1\}^d}} \int_{Q_{k,j}} \left| \sum_{j=1}^N c_j e^{ix_j \cdot \omega} \right|^2 \hat{\Phi}(\omega)\, d\omega \\
&\leq 2^d \sum_{k \in \mathbb{N}_0^d} \max_{\omega \in Q_k} \hat{\Phi}(\omega) \left(R + \frac{\pi}{q_\infty} \right)^d \\
&= 2^d \sum_{k \in \mathbb{N}_0^d} \hat{\Phi}(Rk) \left(R + \frac{\pi}{q_\infty} \right)^d .
\end{aligned}$$

□

We remark that it is possible to obtain the result of Theorem 3.4 also for *non-decreasing* $\hat{\Phi}$. To this end, the right-hand side in (19) should be replaced by a sum over the maxima of $\hat{\Phi}$ on Q_k. Another possibility is to find a decreasing upper bound for $\hat{\Phi}$. Finally, we recall that for smooth Φ we have fast decay for $\hat{\Phi}$ around infinity, so that in this case the convergence of the sum in the right hand side of (19) is fast.

Remark 3.5. *The result of Theorem 3.4 cannot apply to a conditionally positive definite radial kernel* $\Phi \in \mathbf{CPD}(m)$*, since its generalized Fourier transform* $\hat{\Phi} = \varphi(\|\cdot\|_2)$ *has an algebraic singularity at zero, and so in this case the sum in the right hand side of* (19) *is divergent. Yet it is possible to obtain an estimate of the form* (19) *by a more refined evaluation of the integral over* Q_0*. Indeed, for* $c \in L_X$ *that integral is finite, i.e.,*

$$\int_{Q_0} \left| \sum_{x \in X} c_x e^{ix \cdot \omega} \right|^2 \hat{\Phi}(\omega)\, d\omega < \infty.$$

□

It is interesting to compare Theorem 3.4 with the techniques developed in [4]. If $X \subset \mathbb{Z}^d$, the exponential sums in (19) are actually trigonometric polynomials and thus periodic in $[0,2\pi)^d$. In particular, we obtain

$$\begin{aligned}\int_{\mathbb{R}^d} \left|\sum_{j=1}^N c_j e^{ix_j\cdot\omega}\right|^2 \hat{\Phi}(\omega)\,d\omega &= \sum_{n\in\mathbb{Z}^d}\int_{[0,2\pi)^d} \left|\sum_{j=1}^N c_j e^{ix_j\cdot\omega}\right|^2 \hat{\Phi}(\omega+2\pi n)\,d\omega \\ &= \int_{[0,2\pi)^d} \left|\sum_{j=1}^N c_j e^{ix_j\cdot\omega}\right|^2 \sum_{n\in\mathbb{Z}^d}\hat{\Phi}(\omega+2\pi n)\,d\omega,\end{aligned}$$

assuming $\hat{\Phi}$ is decaying sufficiently fast. Now we let $\sigma(\omega)=\sum_{n\in\mathbb{Z}^d}\hat{\Phi}(\omega+2\pi n)$ and obtain, using the orthogonality of $e^{ix_j\cdot\omega}$ for $x_j\in\mathbb{Z}^d$

$$\inf_{\omega\in[0,2\pi)^d}\sigma(\omega)\|c\|_2^2 \le (2\pi)^{-d}\int_{\mathbb{R}^d}\left|\sum_{j=1}^N c_j e^{ix_j\cdot\omega}\right|^2 \hat{\Phi}(\omega)\,d\omega \le \sup_{\omega\in[0,2\pi)^d}\sigma(\omega)\|c\|_2^2.$$

These estimates are actually sharp, whenever σ is continuous, as choosing $\left|\sum_{j=1}^N c_j e^{ix_j\cdot\omega}\right|^2$ to be the Fejér kernel shifted to the point where σ attains its minimum or maximum respectively, realizes exactly these bounds.

Now in [4] it is shown that whenever Φ is a completely monotone function, the minimum of σ is at $\pi(1,\dots,1)$ while the maximum is at zero. This gives

$$\int_{\mathbb{R}^d}\left|\sum_{j=1}^N c_j e^{ix_j\cdot\omega}\right|^2 \hat{\Phi}(\omega)\,d\omega \le (2\pi)^d\sum_{n\in\mathbb{Z}^d}\hat{\Phi}(2\pi n)\|c\|_2^2.$$

Their Fourier transform is monotonically decreasing and letting $R=2\pi$ and $q_\infty=1/2$ we see, comparing with (19), that our estimate is optimal in these instances up to a constant depending on d alone.

3.1. Optimal upper bound for the Gaussian kernel

We apply our result of Theorem 3.4 to the Gaussian kernel $\Phi(x)=e^{-\beta\|x\|_2^2}$, for $\beta>0$. By using (19), we obtain the estimate

$$\Lambda \le \left(\frac{R}{\pi}+\frac{1}{q_\infty}\right)^d \sum_{k\in\mathbb{N}_0^d}\hat{\Phi}(Rk) = \left(\frac{R}{\pi}+\frac{1}{q_\infty}\right)^d\left(\frac{\pi}{\beta}\right)^{d/2}\sum_{k\in\mathbb{N}_0^d} e^{-R^2\|k\|_2^2/(4\beta)}, \tag{20}$$

where we have used the representation (7).

Without loss of generality we may assume $q_\infty = 1/2$ (which then agrees with the separation radius for the integer grid). Otherwise we may rescale β by q_∞^2. Letting $R = 2\pi$ in (20), this gives the estimate

$$\Lambda \le 4^d \left(\frac{\pi}{\beta}\right)^{d/2} \sum_{k\in\mathbb{N}_0^d} e^{-\|\pi k\|_2^2/\beta}. \tag{21}$$

On the other hand, the aforementioned results by Baxter [4] give for the Gaussian case and for $X \subset \mathbb{Z}^d$ the *sharp* upper bound

$$\Lambda \le \left(\frac{\pi}{\beta}\right)^{d/2} \sum_{k\in\mathbb{Z}^d} e^{-\|\pi k\|_2^2/\beta}$$

which agrees with our bound on Λ, up to a positive constant that only depends on the dimension d.

One possibility to further estimate the sum on the right hand side of (21) is

$$\sum_{k\in\mathbb{N}_0^d} e^{-\|\pi k\|_2^2/\beta} = \left(\sum_{k=0}^{\infty} e^{-\pi^2 k^2/\beta}\right)^d \le \left(1 + \frac{\sqrt{\beta}}{2\sqrt{\pi}}\right)^d,$$

which gives the estimate

$$\Lambda \le \left(\frac{\pi}{\beta}\right)^{d/2} 2^d \left(2 + \frac{\sqrt{\beta}}{\sqrt{\pi}}\right)^d. \tag{22}$$

We finally remark that for large q_∞ (i.e., for large β) the estimate in (22) is worse than that in (17), as the upper bound in (22) does, unlike that in (17), not tend to one for $\beta \to \infty$. However, we are usually interested in small separation radii q_∞ (i.e., in small β).

3.2. Upper bounds for inverse multiquadrics

Now we turn to the inverse multiquadric $\Phi(x) = (\gamma^2 + \|x\|_2^2)^{\beta/2}$, for $\beta < -d$ and $\gamma \neq 0$, whose Fourier transform is given by

$$\hat{\Phi}(w) = C(\beta) \left(\frac{\|w\|_2}{\gamma}\right)^{-(\beta+d)/2} K_{(d+\beta)/2}(\gamma\|w\|_2).$$

Here $C(\beta)$ is a constant, depending only on β, and K_α is the modified Bessel function.

As mentioned before, the bound given in (19) is optimal up to a positive constant that only depends on the dimension d. This bound, however, is rather difficult to evaluate. Therefore, we prefer to elaborate on the asymptotic behaviour of Λ with respect to width parameter γ.

From the estimates

$$
\begin{aligned}
K_{(d+\beta)/2}(\gamma\|w\|_2) &\leq C(\beta,d)(\gamma\|w\|_2)^{(d+\beta)/2} \\
K_{(d+\beta)/2}(\gamma\|w\|_2) &\leq \sqrt{\frac{2\pi}{\gamma\|w\|_2}} e^{-\gamma\|w\|_2} e^{(\beta+d)^2/(8\gamma\|w\|_2)},
\end{aligned}
$$

to be found in [21, Lemma 5.13-5.14]. we obtain

$$
\begin{aligned}
\max_{w\in Q_0} \hat{\Phi}(w) &\leq C(\beta,d)\gamma^{\beta+d} \\
\max_{w\in Q_k} \hat{\Phi}(w) &\leq C(\beta)(\|Rk\|_2 + R\sqrt{d})^{-(\beta+d)/2} \|Rk\|_2^{-1/2} \gamma^{(\beta+d-1)/2} \\
&\quad \times e^{-\gamma\|Rk\|_2} e^{(\beta+d)^2/(8\gamma\|Rk\|_2)}
\end{aligned}
$$

If we let $R = \gamma^{-1}$ in the second inequality, this yields

$$
\begin{aligned}
\max_{w\in Q_k} \hat{\Phi}(w) &\leq C(\beta)\gamma^{\beta+d}(\|k\|_2 + \sqrt{d})^{-(\beta+d)/2} \|k\|_2^{-1/2} e^{-\|k\|_2} e^{(\beta+d)^2/(8\|k\|_2)} \\
&= C(\beta)\gamma^{\beta+d}\eta(\|k\|_2,\beta,d).
\end{aligned}
$$

Note that $\eta(\|k\|_2,\beta,d)$ is summable for all β, d. Altogether, we obtain

$$
\Lambda \leq \left(\pi^{-1} + \frac{\gamma}{q_\infty}\right)^d \gamma^\beta \left(C(\beta,d) + C(\beta) \sum_{\substack{k\in\mathbb{N}_0^d \\ k\neq 0}} \eta(\|k\|_2,\beta,d) \right),
$$

where $\gamma^\beta = \Phi(0)$.

4. Comparison with numerical estimates

We finally compare our theoretical estimates on spectral condition numbers

$$
\kappa_2(A_{\Phi,X}) = \frac{\Lambda}{\lambda}
$$

with those from the numerical experiments in [6]. We remark that the numerical results in [6] are restricted gridded data in dimensions $d = 1, 2$. Moreover, the condition numbers in [6] are measured in the ∞-norm,

$$\kappa_\infty(A_{\Phi,X}) = \|A_{\Phi,X}\|_\infty \cdot \|A^{-1}_{\Phi,X}\|_\infty.$$

On the basis of their numerical observations, several conjectures concerning the analytic expression of condition numbers $\kappa_\infty(A_{\Phi,X})$ are given in [6]. We compare with their conjectures for the Gaussian and for the multiquadric kernel.

4.1. Estimates on condition number for the Gaussian kernel

According to the numerical observation in [6] concerning the univariate case, $d = 1$, the condition number $\kappa_\infty(A_{\Phi,X})$ is independent of N, where they conjecture the representation

$$\kappa_\infty(A_{\Phi,X}) = \frac{1}{2} e^{\pi^2/(4\beta)}$$

for equispaced interpolation points X with separation radius $q_\infty = 1/2$.

This compares with our theoretical estimates for $q_\infty = 1/2$ as follows.

• By using (14), we obtain for any $\varepsilon > 0$ the estimate

$$\kappa_2(A_{\Phi,X}) \leq C(\varepsilon) e^{(\pi+\varepsilon)^2/(4\beta)},$$

where the constant $C(\varepsilon)$ is independent of β for β small enough.

• By using (22) we obtain the estimate

$$\kappa_2(A_{\Phi,X}) \leq e^{\pi^2/(2\beta)} \left(10 + \frac{5\sqrt{\beta}}{\sqrt{\pi}}\right).$$

For the bivariate case, $d = 2$, the numerical observations in [6] on regular grids lead them to conjecture

$$\kappa_\infty(A_{\Phi,X}) \sim (1/4) e^{\pi^2/(2\beta)},$$

whereas, by using (14) and (22), we obtain for $q_\infty = 1/2$ the estimates

$$\begin{aligned} \kappa_2(A_{\Phi,X}) &\leq (d+1)2^d \left(2 + \frac{\sqrt{\beta}}{\sqrt{\pi}}\right)^d e^{(d+1)\pi^2/(4\beta)} \\ \kappa_2(A_{\Phi,X}) &= \mathcal{O}\left(e^{(\sqrt{d}\pi+\varepsilon)^2/(4\beta)}\right) \quad \text{for } \beta \to 0. \end{aligned}$$

4.2. Estimates on condition numbers for inverse multiquadrics

For the case of inverse multiquadrics $\Phi(x) = (\gamma^2 + \|x\|_2^2)^{\beta/2}$, for $\beta < -d$ and $\gamma \neq 0$, we obtain for any $\varepsilon > 0$ the estimate

$$\kappa_2(A_{\Phi,X}) \leq C(\beta, d, \varepsilon) \left(\frac{\gamma}{q_\infty}\right)^{(\beta-d+1)/2} \left(\pi^{-1} + \frac{\gamma}{q_\infty}\right)^d e^{2\gamma(\sqrt{d}\pi+\varepsilon)/(4q_\infty)}.$$

For $d = 1$ and uniformly distributed points X with $q_\infty = 1/2$, the numerical results in [6] lead them to conjecture

$$\kappa_\infty(A_{\Phi,X}) \sim (1/2)e^{\gamma\pi}$$

for $\beta = -2$, whereas we obtain in that case the estimate

$$\kappa_2(A_{\Phi,X}) = \mathcal{O}(e^{\gamma(\pi+\varepsilon)}) \quad \text{ for } \gamma \to \infty.$$

This matches (up to constant ε) the order of the asymptotic growth for $\gamma \to \infty$.

Altogether, we can conclude that for all cases discussed in this section, our theoretical estimates (concerning spectral condition numbers κ_2) are remarkably close to those in the conjectures of [6] (for κ_∞).

References

[1] R.A. Adams and J.J.F. Fournier: *Sobolev Spaces*. Pure and Applied Mathematics **140** (2nd ed.), Academic Press., Boston, MA, 2003.

[2] K. Ball: Eigenvalues of Euclidean distance matrices. *Journal of Approximation Theory* **68**, 1992, 74-82.

[3] K. Ball, N. Sivakumar, and J.D. Ward: On the sensitivity of radial basis interpolation to minimal data separation distance. *Constructive Approximation* **8**, 1992, 401–426.

[4] B.J.C. Baxter: Norm estimates for inverses of Toeplitz distance matrices. *Journal of Approximation Theory* **79**, 1994, 222–242.

[5] B.J.C. Baxter: On kernel engineering via Paley-Wiener. *Calcolo* **48**, 2011, 21–31.

[6] J.P. Boyd and K.W. Gildersleeve: Numerical experiments on the condition number of the interpolation matrices for radial basis functions. *Applied Numerical Mathematics* **61**, 2011, 443–459.

[7] M.D. Buhmann: *Radial Basis Function*. Cambridge University Press, Cambridge, UK, 2003.

[8] G.E. Fasshauer: *Meshfree Approximation Methods with Matlab*. World Scientific, Singapore, 2007.

[9] I.M. Gel'fand and N.Y. Vilenkin: *Generalized Functions*. Volume 4: *Applications of Harmonic Analysis*. Academic Press, New York, 1964.

[10] D.S. Jones: *The Theory of Generalized Functions*. Cambridge University Press, Cambridge, UK, 1982.

[11] V. Komornik and P. Loreti: *Fourier Series in Control Theory*. Springer Science and Business Media, 2005.

[12] W. Liao: MUSIC for multidimensional spectral estimation: stability and super-resolution. *IEEE Transactions on Signal Processing* **63**, 2015, 23.

[13] F.J. Narcowich and J.D. Ward: Norms of inverses and condition numbers for matrices associated with scattered data. *J. Approx. Th.* **64**, 1991, 69–94.

[14] F.J. Narcowich and J.D. Ward: Norm estimates for the inverses of a general class of scattered data radial function Interpolation matrices. *Journal of Approximation Theory* **69**, 1992, 84–109.

[15] F.J. Narcowich, N. Sivakumar, and J.D. Ward: On condition numbers associated with radial function interpolation. *J. Math. Anal. Appl.* **186**, 1994, 457–485.

[16] R. Schaback: Lower bounds for norms of inverses of interpolation matrices for radial basis functions. *J. Approximation Theory* **79**, 1994, 287–306.

[17] R. Schaback: Error estimates and condition numbers for radial basis function interpolation. *Advances in Comput. Math.* **3**, 1995, 251–264.

[18] R. Schaback: Stability of radial basis function interpolants. In: *Approximation Theory X: Wavelets, Splines, and Applications*, C.K. Chui, L.L. Schumaker, J. Stöckler (eds.). Vanderbilt Univ. Press, Nashville 2002, 433–440.

[19] A. Selberg: *Lectures on Sieves.* Collected Papers, Volume II, World Scientific, 1991.

[20] J.D. Vaaler: Some extremal functions in Fourier analysis. *Bulletin of the American Mathematical Society* **12**, 1985, 183–216.

[21] H. Wendland: *Scattered Data Approximation.* Cambridge University Press, Cambridge, UK, 2005.