NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2000; **00**:1–6 *Prepared using* nlaauth.cls *[Version: 2002/09/18 v1.02]*

Spectral analysis and preconditioning techniques for radial basis function collocation matrices

This is the pre-peer reviewed version of the following article: Numer. Linear Algebra Appl. 19 (2012) pp. 31-52.

R. Cavoretto[1]*, A. De Rossi[1], M. Donatelli[2], S. Serra-Capizzano[2]

[1] *Dipartimento di Matematica, Università di Torino, Via Carlo Alberto 10, 10123 Torino, Italy*
[2] *Dipartimento di Fisica e Matematica, Università dell'Insubria, Via Valleggio 11, 22100 Como, Italy*

SUMMARY

Meshless collocation methods based on radial basis functions lead to structured linear systems which, for equispaced grid points, have almost a multilevel Toeplitz structure. In particular, if we consider partial differential equations (PDEs) in two dimensions then we find almost (up to a "low-rank" correction given by the boundary conditions) two-level Toeplitz matrices, i.e., block Toeplitz with Toeplitz blocks structures, where both the number of blocks and the block-size grow with the number of collocation points. In [D. Bini, A. De Rossi, B. Gabutti, Linear Algebra Appl., 428 (2008), 508–519] upper bounds for the condition numbers of the Toeplitz matrices approximating a one-dimensional model problem were proved. Here we refine the one-dimensional results, by explaining some numerics reported in the previous paper, and we show a preliminary analysis concerning conditioning, extremal spectral behavior, and global spectral results in the two-dimensional case for the structured part. By exploiting recent tools in the literature, a global distribution theorem in the sense of Weyl is proved also for the complete matrix-sequence, where the low-rank correction due to the boundary conditions is taken into consideration. The provided spectral analysis is then applied to design effective preconditioning techniques in order to overcome the ill–conditioning of the matrices. A wide numerical experimentation, both in the one and two-dimensional case, confirms our analysis and the robustness of the proposed preconditioners. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: Radial basis functions, collocation methods, Toeplitz matrices, preconditioning.

1. Introduction

The aim of this paper is to provide a spectral analysis and effective preconditioners for the linear systems arising from radial basis function (RBF) approximations [21, 46] of partial differential equations (PDEs). The attention for radial basis functions started in the 80's in two specific directions: the fast and accurate solution of interpolation problems and the high order approximation of PDEs by employing collocation techniques. For references on

Contract/grant sponsor: MIUR (for the last two authors); contract/grant number: 2006017542

Received 17 January 2000
Revised 18 September 2002

analytical/approximation results see e.g. [5, 18, 21, 46], while fast and effective evaluation algorithms for RBFs can be found in [3, 31]. Informally, a radial function $\phi(x) : \mathbb{R}^n \to \mathbb{R}$ is a function of the Euclidean norm $||x||$ of x, i.e., $\phi(x) = \eta(||x||)$, for $\eta(t) : \mathbb{R} \to \mathbb{R}$. Examples of special applicative interest are functions of the following form

$$\begin{array}{ll} \sqrt{t^2+c^2} & \text{multiquadric (MQ)}, \\ 1/\sqrt{t^2+c^2} & \text{inverse multiquadric (IMQ)}, \\ \mathrm{e}^{-\frac{t^2}{c^2}} & \text{Gaussian}. \end{array} \tag{1}$$

In this context c is the *shape parameter*, whose value plays a role in modeling problems with various specific features. At least numerically, it is evident that the precision of the approximation procedures based on RBFs is very high. In fact, if h denotes the maximal step size, then the approximation error behaves like $O(\lambda^{c/h})$ for the MQ based scheme and like $O(\lambda^{\sqrt{c}/h})$, when IMQ and Gaussian RBFs are considered. Here λ is a positive parameter, strictly less than one, and independent of h. The price that has to be paid concerns the increasing ill-conditioning of the related linear systems in which a growth of the order of $\mathrm{e}^{\theta c/h}$ is observed at least for large values of c/h, with θ being a positive constant independent of the finesse parameter h and also of the shape parameter c. For simplicity, we consider a Cartesian $n \times n$ grid with n^2 internal points, but the generalization to a generic $n_1 \times n_2$ grid with $n_1 \neq n_2$ is possible. Hence the real collocation matrix is the sum of a quadrantally symmetric two-level Toeplitz matrix of size n^2 and a matrix of rank $2n$, which collects the boundary conditions. We are interested in fast solution methods, in the spectral behavior of the resulting matrices, and especially in the extremal behavior (conditioning) and in the global distribution results. For the latter we need to think not to a single linear system but to a sequence of linear systems of increasing dimensions, where the dimension n^2 is related to a finesse parameter, as usually occurs in the approximation of PDEs. Such kind of spectral knowledge is then employed for suggesting appropriate $O(n^2 \log n)$ preconditioners for Krylov methods. We should mention a first important step in understanding the spectral behavior of the considered matrices was done in [9], where the remarkable link with Toeplitz sequences generated by a symbol was exploited. Related results concerning the analysis of spectral properties of Toeplitz matrices encountered in the interpolation with radial functions and the design of preconditioners can be found in [1, 2].

More in detail, concerning the one-dimensional setting, we give a more precise analysis than in [9], by showing that for some choices of RBF, e.g. IMQ e Gaussian, and for some values of the parameter c/h, the conditioning is not exponential, but grows mildly as n^2. A preliminary study concerning the spectral features of such matrices is concisely carried out in [12]. Here the spectral analysis is extended from the Toeplitz component to the whole matrix-sequence by including the boundary conditions term and some insights on the multidimensional setting are given. The spectral analysis represents a precious guide for the design and/or the choice of efficient preconditioning strategies. Therefore we propose some effective preconditioning strategies depending on the estimated condition number of the associated Toeplitz matrix. As a conclusion, we report a wide numerical experimentation for validating our analysis and our preconditioning strategy.

Finally, we remark that, although our analysis is performed only for equally spaced grid, it is our opinion that this is a first step to obtain results for quite general grids, represented by some "sufficiently regular" function of the equispaced grid. We recall that for scattered grids $\{x_i\}$ for which there exists a function $\beta(t)$ smooth enough, such that $x_i = \beta(ih)$, $i = 0, 1, \ldots, n+1$ where

$\beta(0) = 0$, $\beta(1) = 1$, the asymptotic spectral properties of the finite difference discretization of differential operators are preserved to a certain extent even though their Toeplitz structure is lost (see [38] for more details). We believe that the techniques introduced in this paper may be combined with the approach of [38] to obtain more general results (see also [40]): indeed, for large dimensions the meshfree case could behave in practice as the sufficiently regular case (see the last numerical experiments in [32]).

The paper is organized as follows. In Section 2 we recall the main tools used to prove our result, more specifically, general spectral properties of Toeplitz matrix-sequences generated by a symbol. In Section 3 we state in some detail the one-dimensional and two-dimensional problems, respectively, by emphasizing the linear algebra viewpoint and by reporting the known results. In Section 4 we analyze both the one-level and the two-level case giving a rigorous explanation of some numerics reported in [9] and we study the problem from the point of view of extremal spectral values (conditioning). In Section 5 we study the global spectral distribution of the complete matrix-sequence where also the boundary conditions are considered. Section 6 is devoted to the analysis of preconditioning techniques and to report the results of some numerical experiments, which confirm the effectiveness of the proposed choices. Finally Section 7 deals with conclusions and future work.

2. Toeplitz matrices

We start by recalling the main results concerning the spectral properties of Toeplitz matrices used to prove the asymptotic bounds. We give a compact definition of Toeplitz matrix-sequences generated by a given d-variate Lebesgue integrable symbol. Let s be a Lebesgue integrable function defined on Q^d, $Q = (0, 1)$, d positive integer, and taking values in $\mathbb{C}$. Then, for d-indices $r = (r_1, \ldots, r_d), j = (j_1, \ldots, j_d), n = (n_1, \ldots, n_d)$, $e = (1, \ldots, 1)$, $\underline{0} = (0, \ldots, 0)$, the Toeplitz matrix $T_n(s)$ of size $\hat{n} \times \hat{n}$, $\hat{n} = n_1 \cdot n_2 \cdots n_d$, is defined as $T_n(s) = [c_{r-j}]_{r,j=\underline{0}}^{n-e}$, where c_k are the Fourier coefficients of s defined by equation

$$c_j = c_{(j_1,\ldots,j_d)}(s) = \int_{Q^d} s(t_1, \ldots, t_d) \mathrm{e}^{-\mathrm{i}2\pi(j_1 t_1 + \cdots + j_d t_d)} \, dt_1 \cdots dt_d, \qquad \mathrm{i}^2 = -1, \tag{2}$$

for integers j_ℓ such that $-\infty < j_\ell < \infty$ for $1 \le \ell \le d$. The function $s(x)$ is called *symbol*. If $s(x)$ is real valued then c_{-k} is the conjugate of c_k so that T_n is Hermitian; if, in addition, $s(x)$ is even around $1/2$ with respect to every variable (i.e. $s(e/2 + y) = s(e/2 + z)$, $y \in (0, 1/2]^d$, $|z_j| = y_j$, $j = 1, \ldots, d$), then T_n is globally real symmetric and symmetric also at every level.

2.1. Extremal spectral properties

The symbol (if well defined) associated with a given Toeplitz matrix-sequence is essential for understanding in depth the spectral features, and the conditioning in the nonnegative case, of the sequence. We recall the following properties (see also [11] and references therein).

Theorem 1. [33, 10] *If $s(x) \ge 0$ almost everywhere and not identically constant over Q^d, d positive integer, then T_n is positive definite for any n and its eigenvalues belong to $(\inf s(x), \sup s(x))$, where* inf *and* sup *are intended up to zero Lebesgue measure sets. Moreover, $\lambda_{\min}^{(n)}$ is a decreasing sequence converging to $\inf s(x)$, and $\lambda_{\max}^{(n)}$ is an increasing sequence converging to $\sup s(x)$, where $\lambda_{\min}^{(n)}$ and $\lambda_{\max}^{(n)}$ are the minimal and the maximal eigenvalues of*

T_n, respectively. Furthermore, if $s(x)$ is locally twice differentiable around its infimum points with positive second derivative in at least one of them, then $\lambda_{\min}^{(n)} - \inf s(x) \sim c \sum_{i=1}^{d} n_i^{-2}$, with c positive constant independent of n. Analogously, if $s(x)$ is locally twice differentiable around its supremum points with negative second derivative in at least one of them, then $\sup s(x) - \lambda_{\max}^{(n)} \sim c \sum_{i=1}^{d} n_i^{-2}$. Therefore the spectral condition number $\mu(T_n) = \lambda_{\max}^{(n)}/\lambda_{\min}^{(n)}$ is an increasing sequence converging to $\sup s(x)/\inf s(x)$.

2.2. Distribution spectral properties

Concerning the case of matrix-sequences an important notion is that of spectral distribution in the eigenvalue or singular value sense, linking the collective behavior of the eigenvalues or singular values of all the matrices in the sequence to a given function (or to a measure). The notion goes back to Weyl and has been investigated by many authors in the Toeplitz and Locally Toeplitz context (see the book by Böttcher and Silbermann [11] where many classical results by the authors, Szegö, Avram, Parter, Widom, Tyrtyshnikov, and many others can be found, and more recent results in [23, 40, 45, 43]). A noteworthy application of the considered spectral distribution theory concerns finer estimates of the convergence rate of Krylov methods (see [4]). Here we report the definition of spectral distribution only in the eigenvalue sense, since our analysis is devoted to eigenvalues. The case of singular values represents a plain variation on the theme. For any function F defined on $\mathbb{C}$ and for any complex valued $m \times m$ matrix A, the symbol $\Sigma_\lambda(F, A)$ stands for the mean $\Sigma_\lambda(F, A) := \frac{1}{m} \sum_{j=1}^{m} F(\lambda_j(A))$, with $\lambda_j(A)$, $j = 1, \ldots, m$, denoting the eigenvalues of A.

Definition 2. *Let $\mathcal{C}_0(\mathbb{C})$ be the set of continuous functions with bounded support defined over the complex numbers $\mathbb{C}$, d a positive integer, and θ be a complex valued measurable function defined on a set $G \subset \mathbb{R}^d$ of finite and positive Lebesgue measure $\mu(G)$. Here G will be often equal to Q^d, $Q = (0, 1)$, so that $\mathrm{e}^{\mathrm{i}2\pi\overline{G}} = \mathbb{T}^d$ with $\mathbb{T}$ denoting the complex unit circle. A matrix sequence $\{A_k\}$ is said to be* distributed (in the sense of the eigenvalues) as the pair (θ, G), *or to* have the distribution function θ, *which is denoted by $\{A_k\} \sim_\lambda (\theta, G)$, if, $\forall F \in \mathcal{C}_0(\mathbb{C})$, the following limit relation holds*

$$\lim_{k\to\infty} \Sigma_\lambda(F, A_k) = \frac{1}{\mu(G)} \int_G F(\theta(t))\, dt, \qquad t = (t_1, \ldots, t_d). \tag{3}$$

Finally we say that two sequences $\{A_k\}$ and $\{B_k\}$ are equally distributed *in the sense of eigenvalues if, $\forall F \in \mathcal{C}_0(\mathbb{C})$, we have*

$$\lim_{k\to\infty} [\Sigma_\lambda(F, B_k) - \Sigma_\lambda(F, A_k)] = 0.$$

For multilevel Toeplitz sequences $\{T_n(s)\}$ generated by an integrable d variate symbol s the eigenvalues are not explicitly known, but we know the distribution in the sense of Definition 2 at least when s is real valued implying that every $T_n(s)$ is Hermitian; see [43]. More precisely we have

$$\{T_n(s)\} \sim_\lambda (s, Q^d), \qquad Q = (0, 1). \tag{4}$$

2.3. Approximating class of sequences

In Subsection 2.2 we already reported the definition of spectral distribution. Here we are interested in operative results and especially in tools for deducing the spectral distribution

of a *difficult* matrix-sequence under perturbations or when approximated by *more elementary* matrix-sequences. We will employ these results for finding the spectral distribution of the complete collocation matrix-sequence, Toeplitz plus the low-rank correction due to the boundary conditions, starting from its Toeplitz counterpart. We introduce the notion of *approximating class of sequences* (a.c.s.) and then we give a theorem for dealing with this concept (see [36, 23, 37]).

Definition 3. [36] *Suppose a sequence of matrices $\{A_k\}$ of increasing size d_k is given. We say that $\{\{B_{k,m}\} : m \in \mathbb{N}^+\}$, $B_{k,m}$ of size d_k, is an* approximating class of sequences (a.c.s.) *for $\{A_k\}$ if, for all sufficiently large $m \in \mathbb{N}$, the following splitting can be written*

$$A_k = B_{k,m} + R_{k,m} + N_{k,m} \quad \text{for all } k > k_m, \tag{5}$$

with

$$\text{rank } R_{k,m} \leq d_k\, c(m), \quad \|N_{k,m}\| \leq \omega(m), \tag{6}$$

where $\|\cdot\|$ is the spectral norm (maximal singular value), k_m, $c(m)$ and $\omega(m)$ depend only on m and, moreover,

$$\lim_{m\to\infty} \omega(m) = 0, \quad \lim_{m\to\infty} c(m) = 0. \tag{7}$$

Theorem 4. [37] *Let $\{\{B_{k,m}\}, m \in \mathbb{N}^+\}$ be an a.c.s. for $\{A_k\}$ ($A_k \in M_{d_k}(\mathbb{C})$) such that $E_{k,m} = N_{k,m} + R_{n,m}$, $B_{k,m}$ are Hermitian, d_k is increasing with k, and*

$$\begin{aligned} &\{B_{k,m}\} \sim_\lambda (\theta_m, G), \qquad 0 < \mu(G) < \infty, \\ &\lim_{m\to\infty} \theta_m = \theta \quad \text{in measure on } G. \end{aligned} \tag{8}$$

We assume

$$\sup_m \sup_k \|B_{k,m}\| = \widetilde{C}, \qquad \sup_m \sup_k \|E_{k,m}\| = \widehat{C}, \tag{9}$$

where $\widetilde{C}$, $\widehat{C}$ are positive constants. Moreover, let $\|E_{k,m}\|_1 \leq c(m)d_k$ with $c(m) \xrightarrow[m\to\infty]{} 0$ ($\|\cdot\|_1$ being the trace norm, i.e., Schatten 1 norm or equivalently the l^1 norm of the vector of singular values; see [6]*). Then θ is real valued and*

$$\{A_k\} \sim_\lambda (\theta, G). \tag{10}$$

It is worth mentioning that if the perturbation is also Hermitian then the result in (10) is valid without any norm assumption in (9) (see [42, 36]).

3. The Poisson problem with radial basis functions

This section is devoted to present the structure of the matrices coming from the one-dimensional and two-dimensional radial function approximated Poisson problem. Regarding the one-dimensional setting, we also report the estimates of the conditioning derived in [9].

3.1. The one-dimensional case

In [9] the authors provided explicit asymptotic estimates, as function of c/h, c being the shape parameter, h being the step size, to the condition number $\mu(T_n)$ of the Toeplitz matrix T_n related to the approximated one-dimensional model problem

$$\begin{cases} u''(x) = f(x), \quad x \in (0,1), \\ u(0) = u_0, \quad u(1) = u_1, \end{cases} \tag{11}$$

with the collocation technique over a grid of equally spaced points and based on the MQ, IMQ, and Gaussian radial functions, respectively.

For n positive integer, let $x_0 = 0 < x_1 < \cdots < x_n < x_{n+1} = 1$ and define $\phi(x) = \eta(|x|)$ where $\eta(t)$ is any function in the class (1). We are looking for an approximation to the solution $u(x)$ of (11) in the vector space spanned by the functions $\phi(x - x_i)$, $i = 0, 1, \ldots, n+1$. Set $v = \sum_{j=0}^{n+1} v_j \phi(x - x_j)$, replace u with v in (11) and get $\sum_{j=0}^{n+1} v_j \phi''(x - x_j) = f(x)$, $x \in (0,1)$, with $v(0) = u_0$, $v(1) = u_1$. Setting $x = x_i$ in the latter equation for $i = 1, 2, \ldots, n$ yields the linear system

$$\begin{cases} \sum_{j=0}^{n+1} \phi(x_0 - x_j) v_j = u_0, \\ \sum_{j=0}^{n+1} \phi''(x_i - x_j) v_j = f(x_i), \quad i = 1, 2, \ldots, n, \\ \sum_{j=0}^{n+1} \phi(x_{n+1} - x_j) v_j = u_1, \end{cases} \tag{12}$$

whose matrix $A_{n+2} = (a_{i,j})_{i,j=0,n+1} \in \mathbb{R}^{(n+2)\times(n+2)}$ is such that $a_{0,j} = \phi(x_0 - x_j)$, $a_{n+1,j} = \phi(x_{n+1} - x_j)$ for $j = 0, \ldots, n+1$ and $a_{i,j} = \phi''(x_i - x_j)$ for $i = 1, \ldots, n$, $j = 0, \ldots, n+1$. Let us denote by $T_n = (\phi''(x_i - x_j))_{i,j=1,n}$ the submatrix of A_{n+2} obtained by removing its first and last row and column. In the case where the set $x_i = ih$, $i = 0, \ldots, n+1$, for $h = 1/(n+1)$, forms a grid of equally spaced points in the interval Q the matrix $T_n = (\phi''((i-j)h))$ is a symmetric Toeplitz matrix, i.e., its entries are function of $i - j$. Moreover A_{n+2} is a rank-2 correction to a symmetric Toeplitz matrix.

Let us define $g = c/h$. A direct inspection shows that the Toeplitz matrix T_n is associated with the continuous symbol

$$s(x) = c_0 + 2 \sum_{k=1}^{+\infty} c_k \cos(2\pi k x) \tag{13}$$

which, up to a multiplicative term dependent on h, is formally defined by the sequence $\{c_k\}$ with

$$c_k = \begin{cases} \dfrac{g^2}{(g^2 + k^2)^{3/2}} & \text{for the MQ,} \\ \dfrac{g^2 - 2k^2}{(g^2 + k^2)^{5/2}} & \text{for the IMQ,} \\ \mathrm{e}^{-\frac{k^2}{g^2}} (g^2 - 2k^2)/g^4 & \text{for the Gaussian.} \end{cases}$$

According to [9], there exists a function $\rho(g)$ of g such that for any n we observe $\mu(T_n) < \rho(g)$, where equality is reached only for $n \to \infty$. Furthermore, an interesting asymptotical estimate

$\gamma(g)$ of $\rho(g)$, for $g \to \infty$, was proved:

$$\rho(g) \approx \gamma(g) = \begin{cases} (\mathrm{e}^{\pi})^g/(\pi\sqrt{2g}) & \text{for the MQ,} \\ (\mathrm{e}^{2\pi})^g/(2\mathrm{e}^2\pi^{3/2}g^{3/2}) & \text{for the IMQ,} \\ (\mathrm{e}^{\pi^2})^{g^2}/(2e\pi^2g^2) & \text{for the Gaussian.} \end{cases} \tag{14}$$

Here $\phi(x) \approx \psi(x)$ if $\lim_{x\to\infty} \phi(x)/\psi(x) = 1$, while $\phi(x) \sim \psi(x)$ means asymptotical equivalence that is the existence of two positive constants r and R independent of x, such that $r\phi(x) \le \psi(x) \le R\phi(x)$ for every x large enough.

The proof of these bounds is based on classical spectral properties of Toeplitz matrices and on some tools of the theory of special functions like the modified Bessel functions of the first kind of order 0 and 1. The key step relies in the explicit expression of the symbol associated with the matrix T_n, so that the well-known asymptotic spectral properties of symmetric Toeplitz matrices are used for giving an upper-bound to the condition number of T_n.

From the numerical experiments performed in [9], the above asymptotic bounds are very strict even for small values of g when $\rho(g)$ is small, while the quantity $\rho(g)$ becomes an extremely pessimistic upper-bound when $\rho(g)$ is moderate (say e.g. $g = 3, 4$). For larger values of g the picture changes again since the conditioning becomes exponential in g, but almost independent of n. Moreover there does not seem to be a large difference between the condition number of T_n and of A_n. However in [9], an important question was not answered: in which cases the quantity $\rho(g)$ is a good approximation (and not only a mere upperbound) of the condition number $\mu(T_n)$? In which case, if not numerically, $\rho(g)$ captures at least the asymptotic order of the conditioning? A rigorous explanation of some of these phenomena is given in Section 4 which extends the preliminary analysis in [12].

3.2. The two-dimensional case

In our setting, we consider the Poisson equation

$$\begin{cases} \frac{\partial^2}{\partial x^2}u + \frac{\partial^2}{\partial y^2}u = f(x,y), & \text{for } (x,y) \in \Omega = (0,1)^2, \\ u(x,y) = g(x,y), & \text{for } (x,y) \in \partial\Omega. \end{cases}$$

We endow the square Ω with a grid of knots $z_{i,j} = (x_i, y_j) = (hi, hj)$, $i,j = 0,\ldots,n+1$, $h = 1/(n+1)$, and, by imposing the collocation conditions

$$\left(\frac{\partial^2}{\partial x^2}\phi(x,y) + \frac{\partial^2}{\partial y^2}\phi(x,y)\right)_{x=x_i,y=y_j} = f(x_i,y_j),$$

one arrives at a system of linear equations where the structured part of the matrix is a two-level Toeplitz matrix defined (in the Szegö sense [25]) by the symbol

$$\begin{aligned} s(x,y) = c_{0,0} + 2\textstyle\sum_{k=1}^{\infty}(c_{k,0}\cos(2\pi kx) + c_{0,k}\cos(2\pi ky)) + \\ 4\textstyle\sum_{k=1}^{\infty}\sum_{j=1}^{\infty} c_{k,j}\cos(2\pi kx)\cos(2\pi jy). \end{aligned} \tag{15}$$

It is important to point out that the collocation matrix is the sum of a symmetric two-level Toeplitz matrix (block Toeplitz with Toeplitz blocks) with coefficients $c_{i,j}$ and a matrix of rank $2n$ which collects the boundary conditions.

The values of the coefficients $c_{i,j}$ are determined by the radial function $\phi(x, y)$ used to approximate the solution. In particular, setting $g = c/h$ we find

$$c_{i,j} = \begin{cases} \frac{1}{(i^2+j^2+g^2)^{1/2}} + \frac{g^2}{(i^2+j^2+g^2)^{3/2}} & \text{for the MQ,} \\ \frac{i^2+j^2-2g^2}{(i^2+j^2+g^2)^{5/2}} & \text{for the IMQ,} \\ \frac{4}{g^4}(i^2+j^2-g^2)\mathrm{e}^{-\frac{i^2+j^2}{g^2}} & \text{for the Gaussian,} \end{cases}$$

up to a multiplicative term dependent of h.

Unlike in the one-dimensional case, the symbol $s(x, y)$ is not defined for the MQ: in fact, the series (15) is not convergent for $x = 0$ or $y = 0$. It can be easily verified that the series associated with the IMQ and the Gaussian functions are convergent so that the Toeplitz matrix machinery can be in principle applied for these two classes of radial functions, by taking into consideration the associated continuous symbol. The challenge concerns the case of MQ radial functions where the symbol $s(x, y)$ is discontinuous at $x = 0$, $y = 0$ (it diverges to $+\infty$), but it seems to be a smooth function in the rest of the domain. However, the asymptotical spectral behavior of Toeplitz sequences is well understood far beyond the continuous setting, since the symbol is required to be simply Lebesgue integrable.

4. Analysis of the conditioning and of the extremal spectrum

With reference to Subsection 3.1 and in particular to equation (13), we observe that $s(x)$ is a smooth positive function (see [9]). Therefore, when applying Theorem 1, instead of inf and sup we can use min and max. More in detail, in [9] the ratio $\rho(g) = \max s(x)/\min s(x)$ and its asymptotic estimate $\gamma(g)$ have been computed (see (14)), in the case of MQ, IMQ, and Gaussian function, respectively. We observe that the asymptotic estimates are very precise even for small values of g. However the main interest relies in the evaluation of the actual condition number of $T_n = T_n(s)$.

In Table I we compare the values of $\gamma(g)$ with the actual condition numbers of the Toeplitz matrices T_n, for several values of n in the case $g = 1$, $g = 2$, respectively. It is interesting to point out that for IMQ and for the Gaussian with $g = 2$ the values of $\mu(T_n)$ are far from the asymptotic estimate $\gamma(g)$, even for moderately large values of n.

The explanation relies completely in Theorem 1. Indeed the conditioning is given by

$$\frac{\max s(x) - c_1/n^2}{\min s(x) + c_2/n^2}, \tag{16}$$

with c_1 and c_2 positive constants independent of n. Hence the approximation $\max s(x)/\min s(x)$ is numerically accurate when $\min s(x)$ is far away from zero, but it is not correct when, for larger values of g, the minimum of $s(x)$ exponentially approaches zero. In that case a more reasonable approximation is given by

$$n^2 \cdot \max s(x)/c_2, \tag{17}$$

i.e., by approximating $\min s(x)$ with zero and by neglecting the term c_1/n^2 since $\max s(x)$ is positive and dominating. In reality the columns labeled by $g = 2$ in Table I, for IMQ and Gaussian, show exactly the predicted growth: when the size n doubles, the value of the

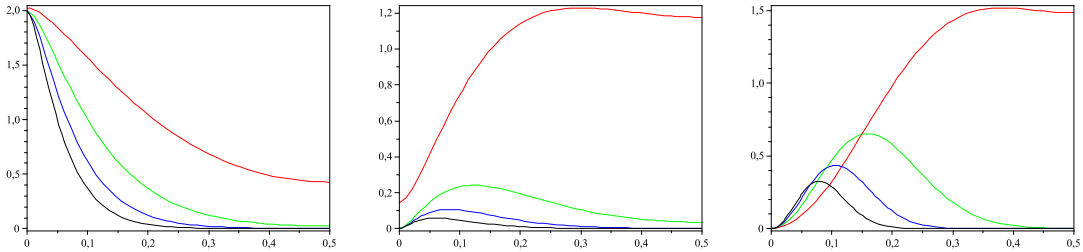
Table I. Values of the spectral condition number $\mu(T_n)$ for different values of n and g. The last line reports the asymptotic estimate $\gamma(g)$ provided in [9].

	MQ		IMQ		Gaussian	
n	g=1	g=2	g=1	g=2	g=1	g=2
50	4.7	78.8	7.5	19.0	217	337
100	4.7	80.1	8.1	52.3	338	400
200	4.7	80.4	8.3	147	395	1543
400	4.7	81.0	8.3	375	413	6069
$\gamma(g)$	5.21	85	6.5	1200	360	6.5×10^{14}

conditioning grows by a factor 4, which is coherent with the given guess of an asymptotic growth proportional to n^2.

The previous analysis has shown that for $g = 2$ and IMQ or Gaussian symbols the growth of the conditioning do depend on n by the formula $n^2 \cdot \max s(x)/c_2$ (see (16) and its approximation (17)). However, when g becomes larger than 2, the surprise is that we observe another change in the picture. The conditioning becomes extreme: we really appreciate the exponential growth of $\frac{\max s(x)}{\min s(x)}$, but there is no longer dependency on n. How to explain this phenomenon? The reason relies again in formula (16). In fact for larger g since the point of minimum is unique we recall that Kac, Murdoch, and Szegö [27] gave the expression of c_2 as the second derivative of s in the minimum point which has to be positive by local convexity. Therefore the explanation of the latter phenomenon could be given in terms of $c_2 = c_2(g)$: if $c_2(g)$ is positive but rapidly converging to zero as a function of g, then the quantity $\frac{\max s(x)}{\min s(x)}$ really captures the conditioning of T_n. In other words, the observed behavior can be explained again by formula (16), but we need to show that at the minimum point, which is 0.5 for MQ and the origin for IMQ and Gaussian, not only the first derivative is zero but the second derivative is a very small positive number. Unfortunately, the latter statement is completely false in the IMQ and Gaussian setting, while the desired behavior is observed for the MQ radial basis functions (see also Figure 1 (a)). However, concerning IMQ and Gaussian RBFs, for $g \geq 3$ it becomes clear from Figures 1 (b) and 1 (c) that $x = 0$ is not the only minimal point, at least numerically. A further minimal point shows up at $x = 0.5$ and the function has locally the expected behavior. In fact, around $x = 0.5$, the graph of the function becomes flatter and flatter as g increases. This visual evidence is supported by the following data: $s'(0.5)$ is equal to 0 and the numerical values of $s(0.5) - s(0)$ and $s''(0.5)$ reported in Tables II tend to zero very fast with g (look at the values for $g \geq 2$). Hence it becomes evident that $x = 0.5$ is a further point of local minimum (numerically global) with the desired features.

In reality, the value of the second derivative at $x = 0.5$ in our setting shows that $c_2(g)$, as a function of g, rapidly collapses to zero exponentially and therefore for larger g, the expression in the denominator $\min s(x) + c_2/n^2$ can be approximated by $\min s(x)$ since c_2 collapses to zero as $\min s(x)$ but the division by n^2 makes it negligible. Furthermore, the quantity $\max s(x) - c_1/n^2$ can be approximated as usual by $\max s(x)$, since this maximum is always well separated from zero. In conclusion in the present setting, for $g \geq 3$, the true approximation of the conditioning is given by $\max s(x)/\min s(x)$ which is extremely high with



(a) MQ (b) IMQ (c) Gaussian

Figure 1. Case 1D – Graph of $s(x)$ for $g = 1, 2, 3, 4$.

Table II. Case 1D – Values of $s(x)$, and $s''(x)$ for IMQ and Gaussian functions.

	$g = 2$	$g = 3$	$g = 4$
IMQ: $s(0.5) - s(0)$	3.6406×10^{-2}	1.4040×10^{-3}	7.2263×10^{-5}
$s''(0.5)$	$3.0658 \times 10^{+0}$	3.1383×10^{-1}	2.3285×10^{-2}
Gaussian: $s(0.5) - s(0)$	1.8096×10^{-3}	1.1907×10^{-8}	5.0081×10^{-16}
$s''(0.5)$	$2.1205 \times 10^{+0}$	8.3464×10^{-5}	1.1702×10^{-11}

respect to g, but essentially constant with respect to n.

The analysis given so far is confirmed also numerically (see Tables III–V for the 1D case; similar results are obtained in the 2D case as well). More in detail, the tables report spectral conditioning (μ), the minimal eigenvalue ($\lambda_{\min}$), and the maximal eigenvalue ($\lambda_{\max}$) of the Toeplitz matrix. We note that for $g = 2$ and IMQ or Gaussian functions, $\mu = \mu(T_n)$ grows about quadratically with respect to the size n of the linear system. In the other cases, the same quantity grows exponentially with respect to g but is practically constant as a function of n. Clearly, for $g = 3$ and IMQ or Gaussian functions, we have again a small growing of μ with respect to n since the eigenvalues of T_n vary continuously in g.

Remark 5. *As n tends to infinity, the values of minimal and maximal eigenvalues tend to given values. These quantities can be considered an accurate approximation of the extremal values (infimum and supremum) of the symbol. The only exception is observed in the 2D case for MQ radial basis functions where, as n doubles also the value of the maximal eigenvalue approximately doubles. This exceptional behavior is not a surprise since the symbol defined in the present case is unbounded at $x = 0$ or $y = 0$ (see the end of Subsection 3.2). Indeed for $n = m$ and setting the real size of the matrix T_n is $N = n^2$, $\lambda_{max}(T_n)$ grows as $N^{0.52}$. The latter fully agrees with the Riemann-Lebegue lemma (see [7]) for which we expect $\lambda_{max}(T_n) = o(N)$. Therefore we have an indication that the related symbol s is Lebesgue integrable in Q^2 and that*

Table III. Case 1D – Condition numbers, smallest and largest eigenvalues for MQ by varying $g = 1, 2, 3, 4$.

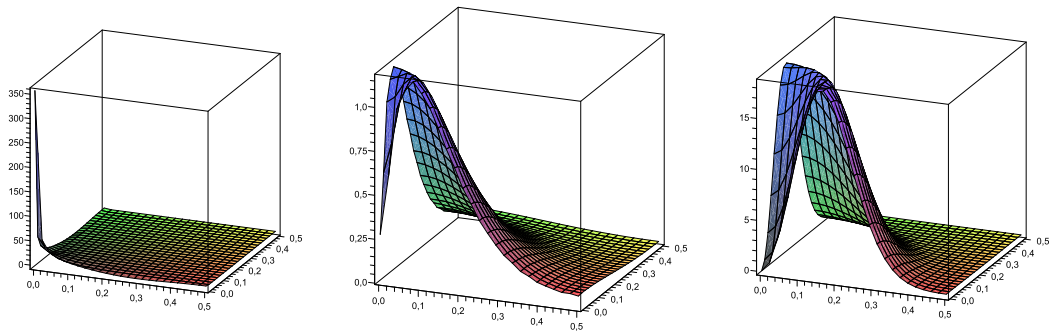
n		$g = 1$	$g = 2$	$g = 3$	$g = 4$
100	$\lambda_{\min}$	4.2761×10^{-1}	2.4848×10^{-2}	1.2952×10^{-3}	6.4264×10^{-5}
	$\lambda_{\max}$	$2.0215 \times 10^{+0}$	$1.9892 \times 10^{+0}$	$1.9790 \times 10^{+0}$	$1.9671 \times 10^{+0}$
	μ	$4.7273 \times 10^{+0}$	$8.0055 \times 10^{+1}$	$1.5280 \times 10^{+3}$	$3.0610 \times 10^{+4}$
200	$\lambda_{\min}$	4.2750×10^{-1}	2.4816×10^{-2}	1.2912×10^{-3}	6.3900×10^{-5}
	$\lambda_{\max}$	$2.0238 \times 10^{+0}$	$1.9967 \times 10^{+0}$	$1.9932 \times 10^{+0}$	$1.9891 \times 10^{+0}$
	μ	$4.7341 \times 10^{+0}$	$8.0457 \times 10^{+1}$	$1.5437 \times 10^{+3}$	$3.1128 \times 10^{+4}$
400	$\lambda_{\min}$	4.2747×10^{-1}	2.4809×10^{-2}	1.2902×10^{-3}	6.3811×10^{-5}
	$\lambda_{\max}$	$2.0246 \times 10^{+0}$	$1.9990 \times 10^{+0}$	$1.9979 \times 10^{+0}$	$1.9966 \times 10^{+0}$
	μ	$4.7361 \times 10^{+0}$	$8.0578 \times 10^{+1}$	$1.5485 \times 10^{+3}$	$3.1289 \times 10^{+4}$

Table IV. Case 1D – Condition numbers, smallest and largest eigenvalues for IMQ by varying $g = 1, 2, 3, 4$.

n		$g = 1$	$g = 2$	$g = 3$	$g = 4$
100	$\lambda_{\min}$	1.5136×10^{-1}	4.6028×10^{-3}	1.2882×10^{-3}	4.8509×10^{-5}
	$\lambda_{\max}$	$1.2283 \times 10^{+0}$	2.4086×10^{-1}	1.0672×10^{-1}	5.9915×10^{-2}
	μ	$8.1149 \times 10^{+0}$	$5.2330 \times 10^{+1}$	$8.2843 \times 10^{+1}$	$1.2351 \times 10^{+3}$
200	$\lambda_{\min}$	1.4733×10^{-1}	1.6392×10^{-3}	1.2370×10^{-3}	4.8285×10^{-5}
	$\lambda_{\max}$	$1.2284 \times 10^{+0}$	2.4108×10^{-1}	1.0693×10^{-1}	6.0116×10^{-2}
	μ	$8.3378 \times 10^{+0}$	$1.4707 \times 10^{+2}$	$8.6439 \times 10^{+1}$	$1.2450 \times 10^{+3}$
400	$\lambda_{\min}$	1.4606×10^{-1}	6.4213×10^{-4}	3.9723×10^{-4}	4.8230×10^{-5}
	$\lambda_{\max}$	$1.2284 \times 10^{+0}$	2.4113×10^{-1}	1.0698×10^{-1}	6.0170×10^{-2}
	μ	$8.4104 \times 10^{+0}$	$3.7552 \times 10^{+2}$	$2.6932 \times 10^{+2}$	$1.2476 \times 10^{+3}$

Table V. Case 1D – Condition numbers, smallest and largest eigenvalues for Gaussian by varying $g = 1, 2, 3, 4$.

n		$g = 1$	$g = 2$	$g = 3$	$g = 4$
100	$\lambda_{\min}$	4.4831×10^{-3}	1.6285×10^{-3}	1.3234×10^{-8}	8.0507×10^{-16}
	$\lambda_{\max}$	$1.5174 \times 10^{+0}$	6.5085×10^{-1}	4.3294×10^{-1}	3.2375×10^{-1}
	μ	$3.3846 \times 10^{+2}$	$3.9966 \times 10^{+2}$	$3.2713 \times 10^{+7}$	$4.0213 \times 10^{+14}$
200	$\lambda_{\min}$	3.8382×10^{-3}	4.2218×10^{-4}	1.2199×10^{-8}	5.4847×10^{-16}
	$\lambda_{\max}$	$1.5175 \times 10^{+0}$	6.5174×10^{-1}	4.3424×10^{-1}	3.2542×10^{-1}
	μ	$3.9537 \times 10^{+2}$	$1.5437 \times 10^{+3}$	$3.5597 \times 10^{+7}$	$5.9333 \times 10^{+14}$
400	$\lambda_{\min}$	3.6744×10^{-3}	1.0743×10^{-4}	1.1975×10^{-8}	5.0898×10^{-16}
	$\lambda_{\max}$	$1.5176 \times 10^{+0}$	6.5197×10^{-1}	4.3458×10^{-1}	3.2587×10^{-1}
	μ	$4.1301 \times 10^{+2}$	$6.0690 \times 10^{+3}$	$3.6289 \times 10^{+7}$	$6.4023 \times 10^{+14}$



(a) MQ (b) IMQ (c) Gaussian

Figure 2. Case 2D – Graph of $s(x, y)$ for $g = 2$

the singularity at $(0, 0)$ *is of the type*

$$\frac{1}{[x^2 + y^2]^\alpha}$$

with α *close to* 0.52 *(see also Figure 2 (a)).*

5. Spectral distribution of the complete matrix-sequence

In Section 4 we have studied in some detail the matrices $T_n = T_n(s)$ However the real collocation matrix is a rank 2 correction of T_n in the case of $d = 1$ and is a rank $2m$ correction when $d = 2$ and $m = n_1 = n_2$. Moreover by the Szegö distribution result in (4) we know that $\{T_n(s)\} \sim_\lambda (s, Q^d)$ with $Q = (0, 1)$ and $d = 1, 2$. Now by exploiting Theorem 4 we would like to deduce the same distribution result for the real sequences $\{A_n\}$ arising in dD RBF collocation when $d = 1, 2$. The latter amounts in proving that the trace norm of the correction divided by the size of the matrix tends to zero as n tends to infinity (in the 2D case we assume just for simplicity that $n_1 = n_2$).

We start by considering the one-dimensional setting. Denoting by B_n the rank 2 correction matrix associated with the boundary conditioning and c the shape parameter, we fix $s_1 = \sqrt{\lambda_1}$ and $s_2 = \sqrt{\lambda_2}$ the only nonzero singular values of B_n, where λ_1 and λ_2 are the only nonzero eigenvalues of $B_n B_n^T$. We recall that in our setting $\|B_n\|_1 = s_1 + s_2$ with $h = 1/(n + 1)$ and $c = g/(n + 1)$, where $\|\cdot\|_1$ is the Schatten 1 norm given by the sum of all singular values (see Theorem 4). Indeed the only nontrivial block of $B_n B_n^T$ is a 2-by-2 circulant matrix

$$C = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$$

with $a, b > 0$. Therefore $\lambda_1 = a + b$, $\lambda_2 = a - b$, and $\|B_n\|_1 = \sqrt{a + b} + \sqrt{a - b}$. We consider now in detail the three basis separately.

MQ: We find

$$a = \frac{(n+2)g^2}{(n+1)^4} + \frac{(n+2)(2n+3)}{6(n+1)^3} = O\left(n^{-1}\right)$$

$$b = \frac{1}{(n+1)^4} \sum_{j=0}^{n+1} \sqrt{(j^2+g^2)((n+1)-j)^2+g^2} = O\left(n^{-1}\right),$$

thus $\|B_n\|_1 = O\left(n^{-\frac{1}{2}}\right)$.

IMQ: Using the second order accurate trapezoidal rule with mesh width 1 for the integral

$$\int_0^n \frac{1}{x^2+g^2}\,\mathrm{d}x = O\left(n^{-1}\right),$$

we have that

$$a = \frac{1}{(n+1)^4} \sum_{j=0}^{n+1} \frac{1}{j^2+g^2} = O\left(n^{-4}\right) \int_0^n \frac{1}{x^2+g^2}\,\mathrm{d}x + O\left(n^{-6}\right) = O\left(n^{-5}\right).$$

Similarly, we observe

$$b = \frac{1}{(n+1)^4} \sum_{j=0}^{n+1} \frac{1}{\sqrt{(j^2+g^2)((n+1)-j)^2+g^2}} = O\left(n^{-5}\right)$$

and thus $\|B_n\|_1 = O\left(n^{-\frac{5}{2}}\right)$.

Gaussian: By the error function, we have that

$$\int_0^n \mathrm{e}^{-\frac{2x^2}{g^2}}\,\mathrm{d}x = O(n)$$

and using again the trapezoidal rule with mesh width 1, we obtain

$$a = \frac{1}{4(n+1)^4} \sum_{j=0}^{n+1} \mathrm{e}^{-\frac{2j^2}{g^2}} = O\left(n^{-3}\right).$$

Similarly, it holds

$$b = \frac{1}{4(n+1)^4} \sum_{j=0}^{n+1} \mathrm{e}^{-\frac{j^2}{g^2}} \mathrm{e}^{-\frac{((n+1)-j)^2}{g^2}} = O\left(n^{-3}\right)$$

and thus $\|B_n\|_1 = O\left(n^{-\frac{3}{2}}\right)$.

Without repeating the details, we just observe that the very same calculations give the desired bounds also for $d = 2$.

Finally, a plain application of Theorem 4 leads to $\{A_n\} \sim_\lambda (s, Q^d)$, for $d = 1, 2$, that is the same spectral distribution of the Toeplitz counterpart, i.e. $\{T_n(s)\}$. Furthermore, by invoking Lemma 3.2 and Theorem 3.5 in [23], the complex set given by the open ball centered in the range of s with radius $\epsilon_n = \|B_n\| \le \|B_n\|_1 \le 2\|B_n\|$ contains all the eigenvalues of $T_n(s)$ for every n.

6. Preconditioning and numerical experiments

We propose several preconditioning strategies and we will show how to use in a practical sense the spectral estimates proposed so far for the choice of the best techniques. Indeed, we will show that the choice of the preconditioner depends on the condition number of T_n that can be estimated by employing the theoretical tools developed in the previous sections.

Firstly, we apply the preconditioned conjugate gradient (PCG) to solve the linear system where the coefficient matrix is the Toeplitz matrix T_n generated by the symbols associated to the considered choices MQ, IMQ, and Gaussian. We consider only the Toeplitz part of the coefficient matrix because the low-rank correction induced by the boundary conditions can be ignored since the spectral feature are decided by the Toeplitz component, as proved in Section 5. Finally, we provide some results for the complete case where also the boundary conditions are imposed. In such case the preconditioned GMRES or the CG for the preconditioned least square problem is used instead of the PCG, since the coefficient matrix is not longer symmetric and, in some cases, numerically singular.

The reported tables furnish the iteration count for reaching the preassigned accuracy. The symbol # indicates that the requested accuracy is not obtained. The desired tolerance is 10^{-7} in all the numerics. The data vector is the vector of all ones.

6.1. Numerical experiments in 1D

We fix 3000 iterations as the maximum number of PCG steps and test several classes of preconditioners.

6.1.1. Algebra preconditioners For each of the considered functions (MQ, IMQ, and Gaussian), in Tables VI-VIII we provide the number of iterations requested by the PCG method with preconditioner chosen in the given algebra: we emphasize the dependency on the matrix size n when the shape parameter c is fixed. We also report the spectral condition number $\mu(T_n)$ and the iteration count for the pure CG without any preconditioning. For the Gaussian case the values of c are chosen smaller due to the higher ill-conditioning of the involved matrices.

With "opt" and "nat" we denote the optimal and the natural preconditioner, respectively [20]. Let $\mathcal{M} = \{Q_n D_n Q_n^H \mid D_n \text{ is a diagonal matrix}\}$ be the matrix algebra diagonalized by the unitary transform Q_n. A noteworthy example is that of the circulant algebra where Q_n is the Fourier transform, in such case we denote $\mathcal{M}$ by $\mathcal{C}$. The optimal preconditioner of a matrix A is defined as the unique minimizer of the functional $B \to F_A(B) \equiv \|B - A\|_F$, where $\|\cdot\|_F$ is the Frobenius norm (Schatten 2 norm or equivalently the l^2 norm of the vector of singular values; see [6]) and where B ranges in the vector space $\mathcal{M}$, having the structure of algebra. When $A = T_n(s)$, the diagonal matrix D_n can be seen as the sampling of a linear positive operator approximating s. Furthermore, the natural preconditioner of $T_n(s)$ is defined case by case, depending on the given algebra, and usually the matrix D_n can be regarded as a sampling of (a slight variation of) the Fourier approximation of s. In the circulant context, the optimal preconditioner was explicitly defined by Tony Chan in [17], while the natural one was introduced by Gilbert Strang in [41] and we refer to these papers for the explicit construction. In the case of a real symbol s, a further algebra of interest is the τ algebra [8], also known as DST-I. It is the algebra of matrices diagonalized by the orthogonal discrete sine transform of

Table VI. Algebra preconditioners for MQ, number of PCG iterations for $c = 10^{-2}$.

n	$\mu(T_n)$	No prec.	τ_{opt}	τ_{nat}	$\mathcal{C}_{\text{opt}}$	$\mathcal{C}_{\text{nat}}$	R. Chan	Super
200	8.045×10^{1}	50	5	5	6	6	6	7
400	3.129×10^{4}	584	7	6	17	17	6	199
600	1.381×10^{7}	#	29	50	173	196	71	#
800	6.434×10^{9}	#	250	600	1680	#	1257	#
1000	3.090×10^{12}	#	2473	#	#	#	#	#

Table VII. Algebra preconditioners for IMQ, number of PCG iterations for $c = 10^{-2}$.

n	$\mu(T_n)$	No prec.	τ_{opt}	τ_{nat}	$\mathcal{C}_{\text{opt}}$	$\mathcal{C}_{\text{nat}}$	R. Chan	Super
200	1.470×10^{2}	29	7	6	9	7	7	38
400	1.247×10^{3}	211	9	7	11	9	10	879
600	3.625×10^{5}	2399	14	11	64	47	11	#
800	1.258×10^{8}	#	89	166	645	#	63	#
1000	4.818×10^{10}	#	811	#	#	#	847	#

Table VIII. Algebra preconditioners for Gaussian, number of PCG iterations for $c = 4 \times 10^{-3}$.

n	$\mu(T_n)$	No prec.	τ_{opt}	τ_{nat}	$\mathcal{C}_{\text{opt}}$	$\mathcal{C}_{\text{nat}}$	R. Chan	Super
200	2.847×10^{1}	42	4	3	6	4	4	15
400	9.435×10^{3}	102	6	5	15	8	8	85
600	1.922×10^{4}	796	8	6	30	13	13	652
800	6.827×10^{8}	#	167	10	1682	30	30	#
1000	6.903×10^{14}	#	#	33	#	83	82	#

the first kind $Q = [q_{ij}] \in \mathbb{R}^{n\times n}$ such that $q_{i,j} = \sqrt{\frac{2}{n+1}} \sin(\frac{ij\pi}{n+1})$, $i, j = 1, \ldots, n$.

Such preconditioners are quite effective, especially when the conditioning is not too bad. As expected from the analysis in [34], the best results are observed when using the τ algebra with special reference to the natural preconditioner. Moreover, it is worth mentioning that each step of the PCG with τ preconditioning is less expensive than with circulant preconditioning since we can use only real arithmetic. Concerning the circulant preconditioning, the Raymond Chan preconditioner [13] is very effective while the superoptimal [44] could be even worse than no preconditioning (this is not a real surprise due to recent results in the literature, see [19] and references therein).

6.1.2. Band Toeplitz preconditioners We now consider the band Toeplitz preconditioning obtained using the Fourier series of degree r or the Cesaro sum again of degree r [7]. The preconditioner shows the form $T_n(p_r)$ with $p_r(x) = \alpha_0 + 2\sum_{k=1}^{r} \alpha_k \cos(kx)$, and $\alpha_k = c_k$ in the Fourier case and $\alpha_k = \frac{r-k}{r} c_k$ in the case of Cesaro averaging.

The solution of a band linear system with bandwidth $2r + 1$ has a computational cost of $O(r^2 n)$ with standard Gaussian elimination; under mild assumptions, such a cost can be reduced when using multigrid to $O(rn)$ or even more by using sophisticated cyclic reduction schemes. In any case, choosing $r = O(\log(n))$ the cost of a single PCG step amounts to

Table IX. Band Toeplitz preconditioners for MQ, number of PCG iterations for $c = 10^{-2}$.

	Fourier		Cesaro	
n	$r = 6$	$r = \lceil 2\log_2(n) \rceil$	$r = 6$	$r = \lceil 2\log_2(n) \rceil$
200	9	4	13	9
400	1483	474	115	61
600	#	#	1329	651
800	#	#	#	#
1000	#	#	#	#

Table X. Band Toeplitz preconditioners for IMQ, number of PCG iterations for $c = 10^{-2}$.

	Fourier		Cesaro	
n	$r = 6$	$r = \lceil 2\log_2(n) \rceil$	$r = 6$	$r = \lceil 2\log_2(n) \rceil$
200	12	7	17	13
400	444	142	48	29
600	#	#	442	257
800	#	#	#	2666
1000	#	#	#	#

Table XI. Band Toeplitz preconditioners for Gaussian: number of PCG iterations for $c = 4 \times 10^{-3}$.

	Fourier		Cesaro	
n	$r = 6$	$r = \lceil 2\log_2(n) \rceil$	$r = 6$	$r = \lceil 2\log_2(n) \rceil$
200	2	2	14	10
400	3	2	48	32
600	92	2	220	119
800	#	2	#	#
1000	#	348	#	#

$O(n\log^2(n))$. In the numerics in Tables IX-XI, we report the numerical results for two choices of the band-parameter r. More specifically we set $r = 6$ in the first case and $r = \lceil 2\log_2(n) \rceil$ in the second case. We observe that, thanks to the positivity of f, $T_n(p_r)$ is always positive definite when the Cesaro averaging is employed, while the positivity (and even the nonsingularity) is not guaranteed when using the Fourier partial sum.

The Cesaro approximation is quite effective for any choice of the RBFs, even if better results are obtained by the algebra preconditioning. The Fourier approximation is sometimes not a good choice due to potential singularity of non definiteness. For the Gaussian case, if the band is large enough, then the results are excellent since the Fourier coefficients have a fast decay (see the analysis in [34, 30]). Indeed, while for MQ and IMQ the logarithmic growth does not imply substantial advantages, for the Gaussian basis we observe that the number of iterations is equal to 2 independently of n (see Table XI).

6.1.3. Mixed preconditioners In general, when the condition number is larger than 10^5, the previous techniques start to suffer. However, by combining band and algebra matrices

Table XII. Mixed preconditioners for MQ, number of PCG iterations for $c = 10^{-2}$.

	Fourier		Cesaro	
n	$r = 6$	$r = \lceil 2\log_2(n) \rceil$	$r = 6$	$r = \lceil 2\log_2(n) \rceil$
200	#	9	#	#
400	59	88	469	923
600	28	28	24	25
800	32	23	33	31
1000	30	21	33	32

Table XIII. Mixed preconditioners for IMQ, number of PCG iterations for $c = 10^{-2}$.

	Fourier		Cesaro	
n	$r = 6$	$r = \lceil 2\log_2(n) \rceil$	$r = 6$	$r = \lceil 2\log_2(n) \rceil$
200	48	51	48	48
400	38	38	38	38
600	33	34	34	34
800	25	22	29	29
1000	26	21	27	26

(see [14, 35, 28]) we are able to propose effective preconditioners. Moreover such mixed techniques have good performances especially in the case of high ill-conditioning. We define a preconditioner of the form $C_n(h)T_n(p_r)C_n(h)$ where $C_n(h)$ is the natural (Strang) circulant preconditioner associated to $h = \sqrt{f/p_r}$ and where p_r is chosen as in Subsection 6.1.2, see [28]. The related results are reported in Tables XII-XIV. In that setting, quite surprisingly, the circulant approach seems more effective than that associated with the τ algebra. With regard to the implementation, since the Fourier coefficients of h cannot be given in general in close form, we employed a trapezoidal rule and FFT computations. For avoiding breakdown due to division by zero, we have set any value to $v = 2.2204 \times 10^{-16}$ if the value at hand is less in modulus than v.

Mixed preconditioner shows very good performances especially when the conditioning is high. On the other hand, the choice of the band part (Fourier or Cesaro, r fixed or growing as $\log(n)$) does not seem important for the convergence speed. Again the only exception, due to the fast decay of the Fourier coefficients, is the case of the Gaussian basis. In Table XIV, we notice that the use of the Fourier series with logarithmic growing degree is especially effective from a numerical viewpoint.

As a conclusion, we can propose an algorithmic strategy. The first step is to estimate theoretically the conditioning of the involved matrices using the tools introduced in the previous sections. If its value is less than 10^5, then it is convenient to choose a pure natural τ preconditioner. Otherwise the mixed preconditioner with the Fourier series for the band part and the natural circulant choice for the algebra part could be advocated.

6.2. Numerical experiments in 2D

In the two-dimensional setting, we consider optimal and natural preconditioners in the circulant algebra. In such case the natural preconditioner is known in literature as Strang

Table XIV. Mixed preconditioners for Gaussian: number of PCG iterations for $c = 5 \times 10^{-3}$.

	Fourier		Cesaro	
n	$r = 6$	$r = \lceil 2\log_2(n) \rceil$	$r = 6$	$r = \lceil 2\log_2(n) \rceil$
200	2	2	316	442
400	51	2	39	39
600	37	2	35	35
800	26	2	27	27
1000	18	14	19	18

preconditioner. We compare block circulant (BC) and block circulant with circulant blocks (BCCB) preconditioners (see [16, 15]). The computational cost for solving a linear system with BCCB amounts to $O(N\log(N))$, with $N = n_1 n_2$, that is the cost of the two-level FFT for the first and the third preconditioner. For BC the situation is different since the blocks are considered without structure so that the cost is $O(Nn_1\log(n_2))$. However, more sophisticated techniques could be used in this case to invert Toeplitz blocks.

In principle two-level band preconditioners are not as easy to handle as in the one-level setting. When using block Gaussian Elimination the cost grows as Nn_2^2 and the same holds while using the standard Gaussian Elimination for band structures. On the other hand, the use of specialized multigrid techniques amounting in a linear cost of $O(n_1 + n_2)$ can make these preconditioners especially appealing: the drawback is that multigrid procedures are never straightforward to implement. Alternatively, one may consider a rank 1 approximation obtained from the real stencil via TSVD: this idea very popular in the image restoration setting leads to the best Frobenius norm approximation of given preassigned rank [26]. More precisely, we have to deal with a preconditioner in tensor form i.e. of the form $C \otimes R$: the related linear systems can be plainly solved by decoupling. In other words we have to solve n_2 one-level Toeplitz linear systems of size n_1 and n_1 one-level Toeplitz linear systems of size n_2. In conclusion the cost will be $O(N(n_1 + n_2))$, which is competitive with that needed for treating BC systems. In addition, if we solve the one-level Toeplitz systems by exploiting the techniques emphasized in the previous section, then we obtain PCG methods sharing the same asymptotic cost. Furthermore, a mixed technique can be obtained by combining the rank 1 approximation with one of matrix algebra preconditioners.

In Tables XV-XVII we show the iteration count when increasing the matrix dimensions and by considering all the basic RBFs discussed so far. The maximal allowed number of iterations is fixed as 1000. The results for BCCB Strang preconditioner are not reported since they are not competitive with those related to the BCCB optimal preconditioner.

The results are not as good as in the one-level case. A reason relies in a structural difficulty emphasized in a series of negative results (e.g. spectral equivalence and strong clustering are impossible to achieve, see [39, 29] and references therein). Moreover, the MQ case is the most difficult to treat because we observe simultaneously very small eigenvalues and large eigenvalues since the function goes from zero to infinity (refer to Remark 5 where it is noted that the symbol is unbounded at $x = 0$ or $y = 0$). However, as in the 1D case, if the estimated condition number is small, then the BCCB preconditioning should be used, otherwise the BC preconditioning or the rank 1 approximation have to be preferred. Indeed, when solving the preconditioned linear system, even if the latter strategies require a slightly higher computational cost with

Table XV. 2D preconditioning for MQ, number of PCG iterations for $c = 0.1$.

$n_1 \times n_2$	No prec.	BC Opt.	BC Strang	BCCB Opt.	Rank 1	Rank1 + BCCB Opt.	Rank1 + BC Strang
10×10	16	9	11	9	14	7	8
20×20	107	12	41	15	42	36	63
30×30	763	50	248	54	112	194	260
40×40	#	193	#	239	294	306	379
50×50	#	745	#	969	767	579	752

Table XVI. 2D preconditioning for IMQ, number of PCG iterations for $c = 0.1$.

$n_1 \times n_2$	No prec.	BC Opt.	BC Strang	BCCB Opt.	Rank 1	Rank1 + BCCB Opt.	Rank1 + BC Strang
10×10	9	6	6	7	6	6	6
20×20	37	9	9	10	10	6	7
30×30	216	11	18	16	18	7	10
40×40	#	49	151	65	36	10	22
50×50	#	180	#	277	80	323	#

Table XVII. 2D preconditioning for Gaussian, number of PCG iterations for $c = 0.05$.

$n_1 \times n_2$	No prec.	BC Opt.	BC Strang	BCCB Opt.	Rank 1	Rank1 + BCCB Opt.	Rank1 + BC Strang
10×10	8	5	5	6	5	4	4
20×20	17	11	9	11	19	11	10
30×30	85	12	11	16	20	11	11
40×40	#	41	25	103	22	13	13
50×50	#	381	162	#	24	30	30

respect to those based on BCCB approximations, they are more robust when varying the condition number. In particular, the rank 1 preconditioning seems to be very effective, at least in the case of IMQ and Gaussian RBF also for large condition numbers. Moreover, the two independent one-level Toeplitz linear systems have symbols with a spectral behavior similar to a 1D collocation problem in the same class (MQ, IMQ or Gaussian). Therefore, they could be solved by PCG using the preconditioners proposed in Subsection 6.1 for the 1D case.

6.3. The complete case with boundary conditions

When imposing boundary conditions the coefficient matrix A_n is not longer symmetric, thus we can solve the linear system with the preconditioned GMRES or by the standard CG for normal equations (CGLS) applied to the preconditioned system: the latter has to be definitely preferred whenever the global matrix is (numerically) rank deficient. We also observe that $\{T_n\}$ indicates the global spectral behavior of $\{A_n\}$ except for few outliers whose number depends linearly on the rank of the correction term. We recall that the global distribution is not sensitive to the behavior of few extreme eigenvalues, so the spectral distribution of two

Table XVIII. 1D case and IMQ basis: condition number, extremal eigenvalues and extremal singular values of T_N and A_N for $c = 10^{-2}$.

n	200	400	600	800	1000
$\mu(T_n)$	1.4707e+002	1.2476e+003	3.6250e+005	1.2588e+008	4.8186e+010
$\mu(A_n)$	1.6200e+016	2.0548e+015	4.1116e+015	8.7834e+015	9.8260e+015
$\lvert\lambda_1(T_n)\rvert$	2.4108e-001	6.0170e-002	2.6742e-002	1.5043e-002	9.6272e-003
$\lvert\lambda_2(T_n)\rvert$	2.4108e-001	6.0170e-002	2.6742e-002	1.5042e-002	9.6272e-003
$\lvert\lambda_{n-1}(T_n)\rvert$	5.2231e-003	4.8284e-005	7.3864e-008	1.1966e-010	2.0006e-013
$\lvert\lambda_n(T_n)\rvert$	1.6392e-003	4.8230e-005	7.3771e-008	1.1950e-010	1.9979e-013
$\lvert\lambda_1(A_n)\rvert$	2.4108e-001	1.6379e-001	1.3886e-001	1.2141e-001	1.0900e-001
$\lvert\lambda_2(A_n)\rvert$	2.4108e-001	6.0170e-002	2.6742e-002	1.5043e-002	9.6272e-003
$\lvert\lambda_{n-1}(A_n)\rvert$	5.2781e-003	4.8230e-005	7.3771e-008	1.1950e-010	1.9981e-013
$\lvert\lambda_n(A_n)\rvert$	3.1672e-016	3.7535e-017	-8.5606e-017	-5.2246e-016	1.0601e-014
$\sigma_1(T_n)$	2.4108e-001	6.0170e-002	2.6742e-002	1.5043e-002	9.6272e-003
$\sigma_2(T_n)$	2.4108e-001	6.0170e-002	2.6742e-002	1.5042e-002	9.6272e-003
$\sigma_{n-1}(T_n)$	5.2231e-003	4.8284e-005	7.3864e-008	1.1966e-010	2.0006e-013
$\sigma_n(T_n)$	1.6392e-003	4.8230e-005	7.3771e-008	1.1950e-010	1.9979e-013
$\sigma_1(A_n)$	2.4465e+000	2.4480e+000	2.4485e+000	2.4487e+000	2.4489e+000
$\sigma_2(A_n)$	2.4108e-001	6.0170e-002	2.6742e-002	1.5043e-002	9.6272e-003
$\sigma_{n-1}(A_n)$	1.6398e-003	3.7294e-005	7.3772e-008	1.1950e-010	1.9979e-013
$\sigma_n(A_n)$	1.5102e-016	1.1913e-015	5.9550e-016	2.7879e-016	2.4922e-016

sequences could be the same while the conditioning could be very different. In our context, we can claim that $\{T_n\}$ establishes the essential conditioning of $\{A_n\}$, whose extreme eigenvalues could behave in a wild, irregular way. In that sense the analysis of the Toeplitz part is very instructive for designing good methods, working also for the complete linear systems. Indeed, in Table XVIII, we see that for the 1D case and the IMQ basis $\mu(A_n) \approx 10^{15}$ independently of n. Similar results hold also for the MQ and Gaussian basis. In particular, for the Gaussian basis A_n is singular and the eigenvalue zero has multiplicity two. More in detail, in the 1D case A_n is a rank two modification of T_n and the rank two correction affects only the behavior of the 2 extremal eigenvalues/singular values for all the considered basis. Moreover, this implies that using T_n as preconditioner for A_n the preconditioned GMRES or the preconditioned CGLS converges in about two iterations. It should be noted that the use of the Sherman-Morrison-Woodbury formula [24] could be also be considered: given $A \in \mathbb{C}^{m\times m}$ and $U, V \in \mathbb{C}^{m\times k}$, the Sherman-Morrison-Woodbury formula reads as

$$(A + UV^H)^{-1} = A^{-1} - A^{-1}U(I + V^H A^{-1}U)^{-1}V^H A^{-1}. \tag{18}$$

In our case $A = T_n$ is the Toeplitz part and $k = 2$, $m = n$ for $d = 1$, and $4n$, $m = n^2$ for $d = 2$.

However, for the sake of completeness, in Table XIX we apply the preconditioners that give the best results in Tables VI and VII for MQ and IMQ directly to A_n (in both cases we have numerical singularity of the coefficient matrix due to the boundary conditions). From Table XIX, we note that if the estimated $\mu(T_n)$ is not large then A_n can be preconditioned directly with the optimal or natural τ algebra preconditioner. On the other hand, if $\mu(T_n)$ is large, it is better to use T_n as preconditioner and to solve the linear system with T_n as in Subsection 6.1. For the Gaussian basis we have similar results and also the mixed preconditioners of

Table XIX. Preconditioned CGLS for A_n (1D case with boundary conditions) for MQ and IMQ: number of iterations for $c = 10^{-2}$.

	MQ					IMQ				
n	No prec.	T_n	τ_{opt}	τ_{nat}	R. Chan	No prec.	T_n	τ_{opt}	τ_{nat}	R. Chan
200	341	3	7	6	7	91	4	13	10	13
400	#	4	17	14	13	#	4	20	16	22
600	#	5	381	463	1901	#	4	80	46	65
800	#	6	#	#	#	#	5	#	#	1891

Table XX. Preconditioned CGLS for A_n (2D case with boundary conditions) for Gaussian, number of iterations for $c = 5 \times 10^{-2}$.

$n_1 \times n_2$	No prec.	T_n	BC Opt.	BCCB Opt.	rank1
10×10	81	39	54	58	56
20×20	267	91	129	137	287
30×30	1167	66	79	98	101
40×40	#	204	2307	#	490

Subsection 6.1.3 should be considered since they are especially effective, like to the Toeplitz case in Table XIV, if the band is large enough.

In the 2D case the coefficient matrix A_n is a rank $2n_1$ modification of its Toeplitz part T_n, for $n_1 = n_2$. This implies that using T_n as preconditioner for A_n the GMRES of the CGLS converge in about $2n_1$ iterations. Therefore, in Table XX we apply the preconditioners discussed in Subsection 6.2 directly to the complete matrix A_n for the Gaussian basis. According to the previous observation, the number of iterations grows like $O(n_1)$. For small values of if $\mu(T_n)$ the BCCB preconditioner is to prefer, while if $\mu(T_n)$ is large is to prefer the rank 1 preconditioner.

7. Conclusions

In the first part of this paper we have recalled the main tools used to prove our result, more specifically, general spectral properties of Toeplitz matrix-sequences generated by a symbol. By using these known facts, we have given a theoretical explanation of some numerics reported in [9] and we have proposed and discussed a more complete picture as g increases and as n increases. More precisely, we have analyzed both the one-level and the two-level case and we have studied the problem from the point of view of extremal spectral values (conditioning) and from the viewpoint of global distribution results. Such analysis gives useful information for the choice of the preconditioner. In this way, we have considered several classical and new preconditioning techniques, giving also practical indications for the more effective choice of the preconditioner according to the current setting (choice of RBF and of the shape parameter, conditioning and matrix size, 1D or 2D). The numerical results have confirmed the effectiveness of some of the proposed choices.

There are several theoretical points to be completed, like the smoothness of the symbols in a multidimensional setting. Moreover, the numerics give some important indications to be

proved theoretically, while the analysis presented in this paper could be extended to other applications of the theory of radial functions, like interpolation or other kind of PDEs.

Moreover, our spectral analysis and preconditioning strategies are considered only for equally spaced grid, but we believe that this is a first step to reach more general results for almost any kind of discretization grids.

A further observation concerns some similarities observed between the present type of matrices and that arising in the solution of fist kind Fredholm equations. Indeed, the collective behavior of the symbols $s(x) = s_g(x)$ as a function of the parameter g is very interesting and gives indications on the ill-posed nature of the linear systems arising from the RBF approximation. In reality, while the RBF interpolation is well-posed, finding the expansion coefficients $\{v_j\}_{j=0}^{n+1}$ in (12) in the RBF interpolation is ill-posed, see p. 133 in [21]. Indeed as g grows the function flattens to zero in most of the domain with the exception of a smaller and smaller interval close to $x = 0$. This means that the high and middle frequency eigenvectors are related to extremely small eigenvalues and that the size of the degenerating eigenspace is very large. This is exactly what happens in the case of deconvolution problems: in that case the presence of high frequency noise makes the problem very difficult so that it is necessary resorting to a regularizing process. Here we are in the PDE world and therefore the problem given by the noise may result less important. Moreover we can give at this point a quantitative statement concerning the size of the degenerating eigenspace. The first observation is that $s_g(x) > 0$ and for any positive δ

$$\lim_{g\to\infty} \mu\{x \in Q^d : \ s_g(x) > \delta\} = 0, \qquad Q = (0,1).$$

Hence, as a consequence of the Szegö distribution results, not only an extreme ill-conditioning arises (exponential as a function of g independent of n), but this ill-conditioning is associated with a subspace of increasing dimension. In fact, by looking at Figures 1 and 2 we observe

$$\lim_{n\to\infty} \frac{\#\{\lambda > \delta\}}{n^d} = \mu\{x \in Q^d : \ s_g(x) > \delta\} \to_{g\to\infty} 0,$$

at least for $d = 1, 2$. We think that the behavior described so far is confirmed as for a generic value of d, but this will be the subject of future investigations.

Acknowledgements

Warm thanks to the referees for their suggestions and comments that helped us to improve the quality of the paper. Finally, we acknowledge that the work of the third and fourth authors was partially supported by MIUR 2008, grant number 20083KLJEZ.

REFERENCES

1. B.J.C. Baxter, *Norm estimates for inverses of Toeplitz distance matrices*, J. Approx. Theory, 79 (1994), 222–242.
2. B.J.C. Baxter, *Preconditioned conjugate gradients, radial basis functions, and Toeplitz matrices*, Comput. Math. Appl., 43 (2001), 305–318.
3. R.K. Beatson, W.A. Light, *Fast evaluation of radial basis functions: methods for 2–dimensional polyharmonic splines*, IMA J. Numer. Anal., 17 (1997), 343–372.

4. B. Beckermann, A. Kuijlaars, *Superlinear convergence of Conjugate Gradients*, SIAM J. Numer. Anal., 39 (2001), 300–329.
5. T. Belytschko, Y. Krongauz, D. Organ, M. Fleming, P. Krysl, *Meshless methods: An overview and recent developments*, Comput. Methods Appl. Mech. Eng., 139 (1996), 3–47.
6. R. Bhatia, *Matrix Analysis*, Springer, New York, 1997.
7. R. Bhatia, *Fourier Series*, Cambridge University Press, 2005.
8. D. Bini, M. Capovani, *Spectral and Computational Properties of Band Simmetric Toeplitz Matrices*, Linear Algebra Appl., 52/53 (1983), pp. 99–125.
9. D.A. Bini, A. De Rossi, B. Gabutti, *On certain (block) Toeplitz matrices related to radial functions*, Linear Algebra Appl., 428 (2008), 508–519.
10. A. Böttcher, S. Grudsky, *On the condition numbers of large semi-definite Toeplitz matrices*, Linear Algebra Appl., 279 (1998), 285–301.
11. A. Böttcher, B. Silbermann, *Introduction to Large Truncated Toeplitz Matrices*, Springer, New York, 1999.
12. R. Cavoretto, A. De Rossi, M. Donatelli, S. Serra-Capizzano, *Spectral analysis for radial basis function collocation matrices*, Numerical Mathematics and Advanced Applications Proceedings of ENUMATH 2009, the 8th European Conference on Numerical Mathematics and Advanced Applications, Uppsala, Sweden, in press.
13. R.H. Chan, *Circulant Preconditioners for Hermitian Toeplitz Systems*, SIAM J. Matrix Anal. Appl., 10 (1989), 542–550.
14. R.H. Chan, W.K. Ching, *Toeplitz-circulant preconditioners for Toeplitz systems and their applications to queueing networks with batch arrivals*, SIAM J. Sci. Comput., 17 (1996), 762–772.
15. R.H. Chan, M.K. Ng, *Conjugate gradient method for Toeplitz systems*, SIAM Rev., 38 (1996), 427–482.
16. R.H. Chan, X.Q. Jin, *An introduction to iterative Toeplitz solvers*, SIAM, Philadelphia, 2007.
17. T. Chan, *An Optimal Circulant Preconditioner for Toeplitz Systems*, SIAM J. Sci. Stat. Comput., 9 (1988), 766–771.
18. A.H.D. Cheng, M.A. Golberg, E.J. Kansa, G. Zammito, *Exponential Convergence and H-c Multiquadric Collocation Methods for Partial Differential Equations*, Numer. Methods Partial Differential Equations, 19 (2003), 571–594.
19. F. Di Benedetto, Serra-Capizzano, *A Note on the Superoptimal Matrix Algebra Operators*, Linear Multilinear Algebra, 50–4 (2002), 343–372.
20. F. Di Benedetto, S. Serra Capizzano, *A unifying approach to abstract matrix algebra preconditioning*, Numer. Math. 82 (1999), 57-90.
21. G.E. Fasshauer, *Meshfree approximation methods with Matlab*, World Scientific Publishers, Singapore, 2007.
22. A.I. Fedosoyev, M.J. Friedman, E.J. Kansa, *Improved multiquadric methods for elliptic partial differential equations via PDE collocation on the boundary*, Comput. Math. Appl., 43 (2002), 439–455.
23. L. Golinskii, S. Serra-Capizzano, *The asymptotic properties of the spectrum of non symmetrically perturbed Jacobi matrix sequences*, J. Approx. Theory, 144-1 (2007), 84–102.
24. G. Golub, C. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1983.
25. U. Grenander, G. Szegö, *Toeplitz Forms and Their Applications*, Second Edition, Chelsea, New York, 1984.
26. J. Kamm, J.G. Nagy *Kronecker Product and SVD Approximations in Image Restoration*, Linear Algebra Appl., 284 (1998), 177–192.
27. M. Kac, W.L. Murdoch, G. Szegö, *On the eigenvalues of certain Hermitian forms*, J. Rational Mech. Anal., 2(1953), 767–800.
28. D. Noutsos, P. Vassalos, *Superlinear convergence for PCG using band plus algebra preconditioners for Toeplitz systems*, Comput. Math. Appl., 56 (2008), 1255–1270.
29. D. Noutsos, S. Serra-Capizzano, P. Vassalos, *Matrix algebra preconditioners for multilevel Toeplitz systems do not insure optimal convergence rate*, Theoret. Computer Science, 315 (2004), 557–579.
30. D. Noutsos, S. Serra-Capizzano, P. Vassalos, *Block band Toeplitz preconditioners derived from generating function approximations: analysis and applications*, Numer. Math., 104 -3 (2006), 339–376.
31. D. Potts, G. Steidl, A. Nielsony, *Fast convolution with radial kernels at nonequispaced knots*, Numer. Math., 98 (2004), 329–351.
32. A. Russo, C. Tablino-Possio, *Preconditioned HSS method for finite element approximations of convection-diffusion equations*, SIAM J. Matrix Anal. Appl., 58 (2010), –.
33. S. Serra-Capizzano, *On the extreme eigenvalues of Hermitian (block) Toeplitz matrices*, Linear Algebra Appl., 270 (1998), 109–129.
34. S. Serra-Capizzano, *Toeplitz preconditioners constructed from linear approximation processes*, SIAM J. Matrix Anal. Appl., 20-2 (1998), 446–465.
35. S. Serra-Capizzano, *Superlinear PCG methods for symmetric Toeplitz systems*, Math. Comp. 68-226 (1999), 793–803.

36. S. Serra-Capizzano, *Distribution results on the algebra generated by Toeplitz sequences: a finite dimensional approach*, Linear Algebra Appl., 28 (2001), 121–130.
37. S. Serra-Capizzano, D. Sesana, *Tools for the eigenvalue distribution in a non-Hermitian setting*, Linear Algebra Appl., 430 (2009), 423–437.
38. S. Serra-Capizzano, C. Tablino-Possio, *Analysis of preconditioning strategies for collocation linear systems*, Linear Algebra Appl., 369 (2003), 41–75.
39. S. Serra-Capizzano, E. Tyrtyshnikov, *How to prove that a preconditioner can not be superlinear*, Math. Comput., 72 (2003), 1305–1316.
40. B. Silbermann, O. Zabroda, *Asymptotic behavior of generalized convolutions: an algebraic approach*, J. Integral Equ. Appl., 18-2 (2006), 169–196.
41. G. Strang, *A proposal for Toeplitz matrix calculations*, Stud. Appl. Math., 74 (1986) 171–176.
42. P. Tilli, *Locally Toeplitz matrices: spectral theory and applications*, Linear Algebra Appl., 278 (1998), 91–120.
43. P. Tilli, *A note on the spectral distribution of Toeplitz matrices*, Linear Multilin. Algebra, 45 (1998), 147–159.
44. E. Tyrtyshnikov, *Optimal and Superoptinal Circulant Precondtioners*, SIAM J. Matrix Anal. Appl., 13 (1992), 459–473.
45. E. Tyrtyshnikov, N. Zamarashkin, *Spectra of multilevel Toeplitz matrices: advanced theory via simple matrix relationships*, Linear Algebra Appl., 270 (1998), 15–27.
46. H. Wendland, *Scattered data approximation*, Cambridge Monogr. Appl. Comput. Math., vol. 17, Cambridge Univ. Press, Cambridge, 2005.