NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2005; **12**:9–32
Published online 30 June 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/nla.390

Analysis of a circulant based preconditioner for a class of lower rank extracted systems

S. Salapaka[1,†], A. Peirce[2,*,‡] and M. Dahleh[3,¶]

[1] *University of Illinois, 362C, Mechanical Engrg. Bldg., 1206 W. Green St., Urbana, IL 61801, U.S.A.*
[2] *Department of Mathematics, The University of British Columbia, 121-1984, Mathematics Road, Vancouver, BC, Canada V6T 1Z2*
[3] *Department of Mechanical Engineering, UC Santa Barbara, U.S.A.*

SUMMARY

This paper proposes and studies the performance of a preconditioner suitable for solving a class of symmetric positive definite systems, $A_p x = b$, which we call *lower rank extracted systems* (*LRES*), by the preconditioned conjugate gradient method. These systems correspond to integral equations with convolution kernels defined on a union of many line segments in contrast to only one line segment in the case of Toeplitz systems. The $p \times p$ matrix, A_p, is shown to be a principal submatrix of a larger $N \times N$ Toeplitz matrix, A_N. The preconditioner is provided in terms of the inverse of a $2N \times 2N$ circulant matrix constructed from the elements of A_N. The preconditioner is shown to yield clustering in the spectrum of the preconditioned matrix similar to the clustering results for iterative algorithms used to solve Toeplitz systems. The analysis also demonstrates that the computational expense to solve LRE systems is reduced to $O(N \log N)$. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS: convolution integral equations; preconditioning; domain geometry

1. INTRODUCTION

In this paper, we discuss the solution of a class of symmetric positive definite linear systems, $A_p x = b$, which we call *lower rank extracted systems* (*LRES*). The coefficient matrix, A_p, has the form given by $A_p = L_p^{\mathrm{T}} A_N L_p$, where A_N is an $N \times N$ Toeplitz matrix having the following structure:

$$A_N \triangleq \begin{pmatrix} a_0 & a_1 & \cdots & a_{N-2} & a_{N-1} \\ a_1 & \ddots & \ddots & \ddots & a_{N-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & a_1 \\ a_{N-1} & \cdots & \cdots & a_1 & a_0 \end{pmatrix}$$

*Correspondence to: A. Pierce, Department of Mathematics, The University of British Columbia, 121-1984, Mathematics Road, Vancouver, BC, Canada V6T 1Z2.
†E-mail: salapaka@uiuc.edu
‡E-mail: peirce@math.ubc.ca
¶He passed away on the 29th July 2000.

Received 12 November 2002
Revised 15 June 2003

and the extraction matrix, L_p, is a $N \times p$ submatrix of an $N \times N$ permutation matrix; i.e. A_p is a principal submatrix of A_N. Similar to Toeplitz systems, LRES arise in the numerical modelling of convolution type integral equations. The difference is that typically in LRES the domain of integration is a union of disjoint line segments. Therefore, Toeplitz systems, which represent the convolution type integral equations on one contiguous line segment, can be considered a special case of LRES. In this way, the Toeplitz systems and LRES have a very close relationship: on one hand the class of Toeplitz systems form a subclass of LRES, while on the other hand, the embedding $A_p = L_p^{\mathrm{T}} A_N L_p$ implies that each LRE system can be viewed as a subsystem of a Toeplitz system. LRES appear in a wide range of scientific and engineering models, for instance in the field of image processing, in the modelling of interacting cracks, in the modelling of tabular mining excavations [1], and in the field of telecommunications in the modelling of elements in planar array antennae [2].

Their close relation to Toeplitz systems makes it possible to exploit various techniques from the vast literature for Toeplitz systems to solve them. Toeplitz systems have been studied for a long time in mathematics due to their role in trigonometric moment problems, in Szegö theory of orthogonal polynomials on the unit circle, and in various other function theoretic subjects [3]. They also arise in the solution of partial differential equations (in fluid dynamics and inverse heat equations), in the solution of convolution type integral equations [1], and in minimum realization problems in control theory, and in the areas of stochastic filtering and digital signal processing [4–6]. Even though most of these problems practically extend to LRES, since the domains of integration are not always connected, not much attention has been given to LRES. The main contribution of this paper is that it proposes a solution to a large class of LRES which guarantees low computational expense (in the order of $N \log N$ computations, where N is the size of the associated Toeplitz matrix A_N).

A comprehensive survey of methods to solve Toeplitz systems (especially iterative methods) has been presented in Reference [7]. There are also a number of non-iterative algorithms, such as the Levinson algorithm [8–11], that have been proposed, which reduce the computational effort to the order of N^2 operations. Over the last decade, significant attention has been given to using the *preconditioned conjugate gradient method* (*PCGM*) [12, 13]. Many algorithms based on this method reduce the computational effort to the order of $N \log N$ operations [14–20]. In this method, $P_N A_N \bar{x} = P_N \bar{b}$ is solved instead of $A_N \bar{x} = \bar{b}$. The matrix P_N is chosen so that the matrix $P_N A_N$ has its spectrum clustered, which ensures better convergence rates.

In this paper, we use the PCGM to solve the LRES and propose a preconditioner, P_p, to solve them more efficiently. This preconditioner has been motivated by one used in Reference [1] for solving interacting crack problems that arise in modelling mining excavations. For interacting crack problems, there is a requirement to model a sequence of such sub-problems in which the interaction between sub-blocks at one step determines the extent of sub-vectors at a subsequent step. One option is to set up a new system matrix for each new set of interacting sub-blocks. However, by treating each such subsystem as embedded in the larger system with system matrix A_N, we avoid this set-up process at each stage of the calculation and also derive a considerable computational advantage from the preconditioner. It is remarkable that the preconditioner constructed by using the encompassing Toeplitz matrix yields such an efficient clustering of the eigenvalues associated with the multiple interacting sub-problems. Indeed, the extraction operators that we introduce make it possible to capture the required information about the higher frequency modes associated with each of the

subcracks/subexcavations. These extraction operators define the geometry of the interacting crack problem.

In the case of Toeplitz systems, this preconditioner reduces to one of the preconditioners studied in Reference [14]. When compared to other iterative schemes, it has significantly better clustering characteristics and therefore, better convergence rates. In Reference [14], an elegant analysis of the performance of this preconditioner for Toeplitz systems is presented. Furthermore, the elements of the preconditioner are shown to be approximations of the Fourier coefficients of the reciprocal of the generating function, a result which is not only theoretically interesting, but also provides scope for extensions to larger classes of systems. Similar preconditioners to solve band Toeplitz matrices [21] and block-Toeplitz Toeplitz-block (BTTB) systems [22] have been proposed. It should be noted that even though LRES have Toeplitz subblocks, they are fundamentally different from BTTB systems, both in structure and concept. A clear difference in their structure is that LRE matrices are not block-Toeplitz while BTTB systems are. Also, LRES have rectangular subblocks (which are subblocks of Toeplitz matrices) while 'Toeplitz-blocks' of BTTB are square. Furthermore, BTTB systems are a generalization of Toeplitz systems to two dimensions and in this context they represent double-integral equations defined over 2-dimensional rectangular domains. On the other hand, the LRES that we present here denote integral equations defined over 1-dimensional domains. As a consequence, the coefficient matrices in BTTB systems are generated by functions of two variables, $f(\theta_1, \theta_2)$ while the LRES are specified by univariable functions, $f(\theta)$. It is possible to generalize LRE systems to two dimensions (2-dimensional LRES) so that they represent double-integral equations defined over 2-dimensional domains which are not necessarily rectangular. In this case, the relation between the 2-dimensional LRES and the BTTB systems is analogous to the relation between the 1-dimensional LRES and the Toeplitz systems. In this paper, we present the analysis of only 1-dimensional LRES. These structural and conceptual properties also differentiate LRES from other systems considered in the literature to solve rectangular Toeplitz matrices [7, 23, 24] or rectangular block-Toeplitz matrices.

In Section 2, we motivate the need to study LRE problems by giving two examples of physical models which are represented by LRES. In Section 3, we formulate the basic problem and introduce the circulant and preconditioner matrices. We define clustering and establish some fundamental properties of the circulant matrices and their relation to the preconditioner. The main idea that we exploit in this paper is the same as the one used in Reference [14] to propose and analyse preconditioners for Toeplitz systems. More precisely, we show that the eigenvalues of the circulant matrix associated with the LRE system approximate its generating function, f, at certain points; and that the elements of the preconditioner are approximations of the Fourier coefficients of $1/f$. These properties are then used to establish the clustering and convergence properties of the preconditioner for the LRES. Section 4 provides the results of some simulations. First, we give simulation results of the application of the proposed preconditioner to an example of an LRE system that represents a multi crack problem in mining. We study and quantify the performance of the proposed preconditioner for this problem. Then, we show the persistence of the performance of the algorithm for different generating functions, different sizes of the matrices and different shapes of the domains. We also provide results of their performance for an LRE matrix associated with a divergent sequence to study the robustness of this algorithm. Finally, in Section 5, we present some concluding remarks.

Notation

- $[T]_{p,q}$ is the element in the pth row and the qth column of the matrix T.
- δ_k is the kronecker-delta function, $\delta_k = 1$ if $k = 1$, and $\delta_k = 0$ if $k \neq 0$.
- I is the identity matrix (its dimension is fixed by the context it appears in).
- C^q is the class of q times continuously differentiable functions on the unit circle.
- x^N is a vector of length $2N$ given by $(x_{-(N-1)} \ \cdots \ x_0 \ \cdots \ x_N)^{\mathrm{T}}$.
- $\|v\| = (\sum v_i^2)^{1/2}$ is the Euclidean norm of the vector v. The dimension of v is determined from the context it appears in.
- $\|T\|$ and $\|T\|_{\mathrm{F}}$ are the induced and the Frobenius (Hilbert–Schmidt) norms of the operator T.
- T_N^x, N in $\mathbb{N}$, $x = \{x_n\}_{n=-\infty}^{\infty}$ is an $N \times N$ symmetric Toeplitz matrix given by

$$\begin{pmatrix} x_0 & x_1 & \cdots & x_{N-2} & x_{N-1} \\ x_1 & \ddots & \ddots & \ddots & x_{N-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & x_1 \\ x_{N-1} & \cdots & \cdots & x_1 & x_0 \end{pmatrix}$$

- H_N^x, $N \in \mathbb{N}$, $x = \{x_n\}_{n=-\infty}^{\infty}$ is an $N \times N$ symmetric Hankel matrix given by

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_{N-1} & x_N \\ x_2 & \iddots & \iddots & \iddots & x_{N-1} \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ x_N & x_{N-1} & \cdots & x_2 & x_1 \end{pmatrix}$$

- J_N is an $N \times N$ counter identity matrix given by

$$\begin{pmatrix} 0 & \cdots & \cdots & 0 & 1 \\ 0 & \cdots & \iddots & 1 & 0 \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ 0 & \iddots & \iddots & \cdots & 0 \\ 1 & 0 & \cdots & \cdots & 0 \end{pmatrix}$$

- C_N^x is a $2N \times 2N$ circulant matrix associated with the Toeplitz matrix T_N^x given by

$$\begin{pmatrix} x_0 & x_1 & \cdots & x_2 & x_1 \\ x_1 & \ddots & \ddots & \ddots & x_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & x_1 \\ x_1 & \cdots & \cdots & x_1 & x_0 \end{pmatrix} = \begin{pmatrix} T_N^x & J_N H_N^x \\ H_N^x J_N & T_N^x \end{pmatrix}$$

- $D_L^x(N,m,n)$ is an $m \times n$ matrix given by

$$(I_m \ 0 \ \cdots \ \cdots \ \cdots) \begin{pmatrix} \cdots & 0 & x_N & \cdots & x_2 & x_1 \\ \cdots & 0 & \ddots & \ddots & \ddots & x_2 \\ & \ddots & \ddots & \ddots & \vdots & \\ & & \ddots & \ddots & x_N & \\ & & & 0 & 0 & \\ & & & \vdots & \vdots & \end{pmatrix} \begin{pmatrix} \vdots \\ \vdots \\ \vdots \\ \vdots \\ 0 \\ I_n \end{pmatrix}$$

- $D_R^x(N,m,n)$ is an $m \times n$ matrix given by

$$(\cdots \ \cdots \ \cdots \ 0 \ I_m) \begin{pmatrix} \vdots & \vdots & & & \\ 0 & 0 & & & \\ x_N & \ddots & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \\ x_2 & \ddots & \ddots & \ddots & 0 \ \cdots \\ x_1 & x_2 & \cdots & x_N & 0 \ \cdots \end{pmatrix} \begin{pmatrix} I_n \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{pmatrix}$$

2. MOTIVATION

In this section, we give motivations for the need to study LRE systems and emphasize their relations with Toeplitz systems. For any LRE system, $A_p x = b$, with the domain given by

$D = \bigcup_{k=1}^{q} V_k$, the coefficient matrix, A_p is completely determined by the kernel of the associated integral equation and by the geometry of the domain. Accordingly, to every LRE system, we associate two matrices

1. An $N \times N$ Toeplitz matrix, A_N, corresponding to the Toeplitz system representing an integral equation with the same kernel as the LRE system, but its domain being an interval, V, which contains the domain D of the LRE system (i.e. V is the convex hull of the domain D). This matrix A_N contains all the information about the kernel.
2. An $N \times p$ extraction matrix, L_p (a submatrix of an $N \times N$ permutation matrix), that has q block-columns with the width of each column equal to the number of representative points in the corresponding segment of the domain D. In this matrix, the ith block-column has only one identity matrix (with all other entries in this block-column being 0); and every alternate block-row is a zero block. In this way, the matrix L_p has the complete information of the geometry of the domain D. Indeed, the extraction matrices, L_p are used to define the geometry of the LRE problems.

The coefficient matrix, A_p then satisfies the relation $A_p = L_p^{\mathrm{T}} A_N L_p$. This constitutes the main difference between the LRE and Toeplitz systems. The Toeplitz systems are completely determined by the kernel while for the LRE systems, one needs to know also the structure of the domain besides knowing the kernel.

The 1-dimensional integral equations and the corresponding LRE systems are often represent simplified models of higher-dimensional phenomena. We present two such examples to emphasize the importance of the LRE systems and to understand the concepts presented above.

Example of Microstrip Reflector Arrays [2]
Mobile and satellite communication systems require high gain antennae which are compact and light weight. This has attracted a lot of research in development of smaller antennae. The microstrip reflector array (MRA) is a result of this research effort. The MRAs are very advantageous since they are flat and therefore easier to fabricate, mount on flat surfaces, and cost effective.

In a conventional reflector antenna, the surface is appropriately designed (say parabolic) to ensure a coherent phase-front for the reflected field. To achieve the same goal, the MRA uses microstrip patches (see Figure 1(a)) connected to open or short transmission lines (tails) to form a directional beam. The patches serve as reradiators and the tails as phase shifters. In an ideal parabolic antenna, the reflected beam has a coherent phase-front since $\mathrm{FM} + \mathrm{MM}' = \mathrm{FO} + \mathrm{OO}'$ (see Figure 1(b)). This can be achieved in a MRA by properly designing the patches such that $\mathrm{FP} + \mathrm{PP}' = \mathrm{FO}'' + \mathrm{O}''\mathrm{O}' + \mathrm{d}P$, where $\mathrm{d}P$ denotes the equivalent electrical length of the patch phase shift.

One of the first steps in analysing this MRA structure is studying a z-directed linear array of microstrips (see Figure 1 (Lower)). The integral equation describing the induced current (J_z) and the incident field ($E_z^i(z)$) is,

$$\int_{\mathscr{D}} \Gamma(z - z') J_z(z')\, \mathrm{d}z' = -E_z^i(z)$$

where $\Gamma(z - z') = c(k_0^2 + \mathrm{d}^2/\mathrm{d}z^2)(\mathrm{e}^{-jk_0 r}/r)$ with $r = \sqrt{(z - z')^2 + a^2}$, and constants c, a and k_0, is a dyadic Green's function [2, 25]. On expanding J_z in terms of basis functions, $\{\phi_n(z)\}$ as

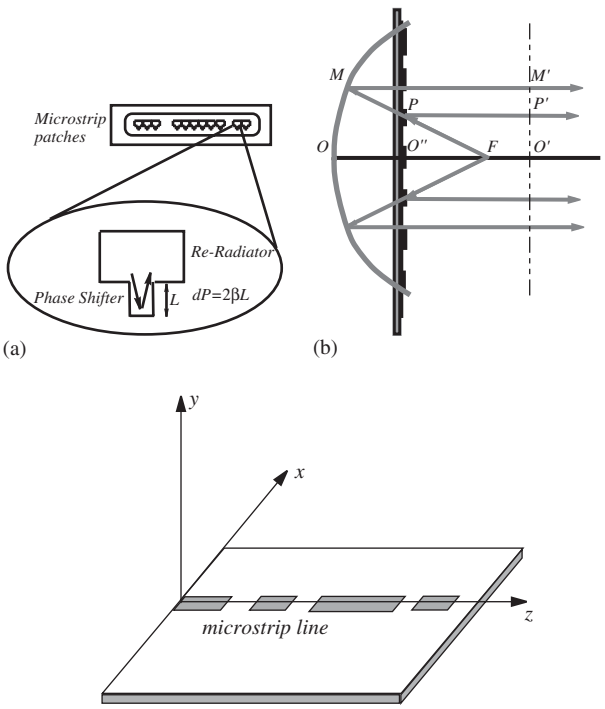

Figure 1. (Upper) (a) An MRA and patch element; and (b) the illustration of concept of the MRA comparing with a reflector antenna. (Lower) The geometry of a microstrip line.

$J_z(z) = \sum_n i_n \phi_n(z)$ and by applying inner products on both sides of the integral equations by test functions, ψ_m $(m \in \mathbb{N})$, this integral equation is reduced a linear system, $Zi = v$, where

$$[Z]_{m,n} = \int_{\mathscr{D}} \left(\int_{\mathscr{D}} \Gamma(z - z') \phi_n(z') \, \mathrm{d}z' \right) \psi_m(z) \, \mathrm{d}z \quad \text{and} \quad v_m = - \int_{\mathscr{D}} \int_{\mathscr{D}} E_z^i(z) \psi_m(z) \, \mathrm{d}z$$

Note that for the domain in Figure 1, $\mathscr{D}$ is a union of linear segments (not necessarily of same lengths) and therefore the linear system, $Zi = v$ is an LRE system. On the other hand if the domain $\mathscr{D}$ had been one segment, then this linear system would have been a Toeplitz system.

Example of collinear cracks

A simple integral equation to describe a crack located along a line (on the interval (a, b)) in an elastic body in a state of 2D plane strain is given by

$$k \int_a^b \frac{U(\xi)}{(x - \xi)^2} \, \mathrm{d}\xi = p$$

where k is a constant depending on material properties, $U(x)$ represents the crack opening displacement, and p represents the pressure applied to the boundary of the crack. A similar

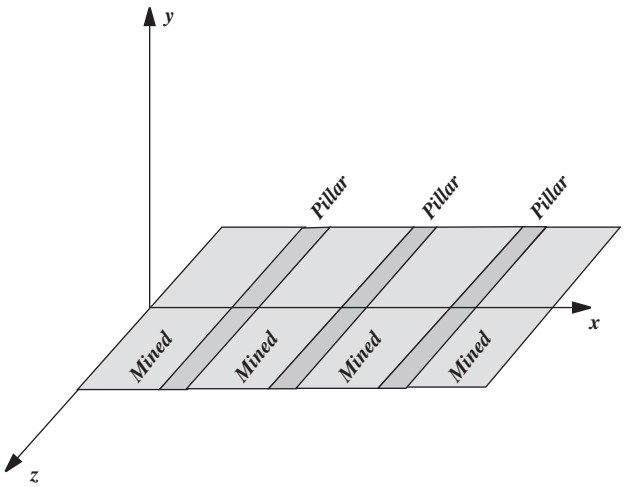

Figure 2. The 'rib-pillar' mining layout. This mine geometry can be accurately modelled by collinear segments in a state of plane strain (an LRE system).

integral equation can be used to model the closure of a tabular mining excavation, whose length in the out-of-plane direction is much larger than $b-a$ and in which the ambient stress in the rock prior to mining is given by $-p$ (see Reference [1] for details). A numerical approximation of this equation is obtained by partitioning the interval (a,b) into N subintervals of equal length and assuming a piecewise constant approximation to $U(\xi)$ on each subinterval. Finding the unknowns in this discrete approximation involve solving a symmetric Toeplitz system, $A_N x = b$, where $[A_N]_{i,j} = \tilde{k}/((i-j)^2 - \frac{1}{4})$, $\tilde{k}$ is a constant, $b_i = p$, and x_i is the approximation of $U(\xi)$ at the ith element of the partition.

Similarly, an integral equation to describe q interacting collinear cracks on the intervals $(a_1,b_1),\ldots,(a_q,b_q)$ under the same physical assumptions is given by

$$k\int_{\mathscr{D}} \frac{U(\xi)}{(x-\xi)^2}\,\mathrm{d}\xi = p$$

where $\mathscr{D}$ is the union of the intervals, $(a_1,b_1),\ldots,(a_q,b_q)$. The numerical model for this equation, obtained by applying the same procedure as in the single crack case, yields a LRE system, $A_p\tilde{x} = \tilde{b}$ where A_p consists of Toeplitz subblocks of A_N; and $\tilde{x}$ and $\tilde{b}$ are subvectors of x and b, respectively. In the mining context the LRE system represents the interaction of a sequence of coplanar tabular mining excavations in a state of plane strain. These coplanar mining excavations represent 'rib-pillar' mining layouts (see Figure 2) commonly used in the extraction of tabular ore bodies such as those found in the gold mining industry. Since, the extent of the mining excavations is much larger in the out-of-plane z direction, these mine geometries can be accurately modelled by collinear segments in a state of plane strain.

3. PROBLEM FORMULATION AND SOLUTION

Problem setting. In the previous section, we have seen that for every LRE system, we can associate a Toeplitz matrix representative of the kernel of the integral equation. This is a many-to-one association since many LRE systems having the same kernel but different geometries can be associated with the same Toeplitz system. In this paper, we consider a sequence of Toeplitz matrices, $\{A_N\}$, and study the sequence, $\{\mathscr{L}_N\}$, of sets of LRE coefficient matrices (A_p) that can be associated with each A_N. The sequence $\{A_N\}$ is assumed to satisfy

Assumption 1

1. $A_N = T_N^a$ (see *Notation*), formed from the N elements $a_0, \ldots, a_{N-1}$, of a given sequence $\{a_n\}$ in ℓ_1.
2. The sequence, $\{a_n\}$ is such that its generating function, given by $f(\theta) = \sum_{-\infty}^{\infty} a_k \mathrm{e}^{ik\theta}$, is real, symmetric, positive, and bounded away from 0; i.e. $M_a \triangleq \sum |a_k| < \infty$, $a_k = a_{-k}$ for all k in $\mathbb{Z}$, and there is a $\delta > 0$ such that $f(\theta) > \delta > 0$ for all θ in $[-\pi, \pi]$.

Remark
As a consequence of the generating function being real and positive, the matrices A_N, $N \in \mathbb{N}$ are symmetric and positive definite; and also the assumption $f(\theta) > \delta > 0$ guarantees an upper bound on the norms of A_N^{-1} which is necessary for the well-posedness of the recursive method to obtain the solution of the LRE system.

Preconditioned conjugate gradient method. We solve LRE systems using the preconditioned conjugate gradient method [12, 13]. In this method, a matrix (called the preconditioner) P_p is designed and the system $P_p A_p x = P_p b$ is solved instead of $A_p x = b$. Unlike in many other methods, here the convergence rate depends on the distribution of all eigenvalues of $P_p A_p$, and not exclusively on its extremal eigenvalues. Moreover, the PCGM convergence is fast when the eigenvalues are clustered and P_p are designed so as to achieve this property.

Proposed preconditioner. We prescribe a preconditioner for the coefficient matrix, A_p, of an LRE system in the following way. We first form matrices A_N and L_p as in previous section and then construct a $2N \times 2N$ circulant matrix C_N^a (see *Notation* for this construction). Since A_p is a principal submatrix of A_N, given by $L_p^{\mathrm{T}} A_N L_p$, it is also a principal submatrix of C_N^a; i.e. $A_p = \bar{L}_p^{\mathrm{T}} C_N^a \bar{L}_p$, where the *extracting matrix*, $\bar{L}_p$ is defined by $\bar{L}_p^{\mathrm{T}} = [0\ L_p^{\mathrm{T}}]$. Its structure is completely determined by, as well as determines, the geometry of the domain of the LRE system. The preconditioner, P_p is then defined by $P_p = \bar{L}_p^{\mathrm{T}} (C_N^a)^{-1} \bar{L}_p$. In the case of Toeplitz systems, L_p is equal to the $N \times N$ identity matrix and hence the corresponding matrix has a rank of N which is greater than any other LRE system associated with A_N. Hence, the name *lower rank extracted* matrices.

We have prescribed P_p in terms of the circulant matrix, C_N^a, because circulant matrices are easy to invert. The linear equation with circulant matrix, $C_N^a x = b$, is equivalent to the discrete convolution, $c * x = b$, where c is the first column of C_N^a. This convolution equation can be solved easily by taking the fast Fourier transform ($\hat{x}_k = \hat{b}_k / \hat{c}_k$) and then obtaining x by applying the inverse fast Fourier transform. This requires only $O(N \log N)$ multiplications

and the operations can be done in parallel [26]. In a similar way, the number of computations in the multiplication of a vector by A_p can be reduced by using the fast Fourier transform (in a similar way as shown in Reference [17]). For the PCGM, if the dimension N of the system is large, the computational effort in the above algorithm is dominated by the preconditioner-residual product $P_p r_j$ and the matrix–vector product $A_p d_j$, where r_j and d_j are residual and conjugate direction vectors (see the PCGM algorithm in Reference [27]). For the symmetric LRE system considered in this paper *both* these matrix–vector products can be evaluated by using the diagonal representation of the circulant C_N^a with respect to the Fourier basis and by applying the appropriate extraction operators $\bar{L}_p$. These processes can be performed very efficiently because the proposed preconditioner product $P_p r_j$ only involves division by the same diagonal elements of the Fourier representation of C_N^a that are used for the matrix–vector product $A_p d_j$. In the following proposition, we summarize the relevant properties of circulant matrices.

Proposition 3.1

1. (a) The circulant matrix C_N^a is diagonalizable, i.e. $C_N^a = U_N \Lambda_N U_N^{\mathrm{T}}$ where $U_N = U_N^{\mathrm{T}} = U_N^{-1}$ and $[U_N]_{ij} = 1/\sqrt{2N}(\cos(ij\theta_N) + \sin(ij\theta_N))$ where $0 \leqslant i, j \leqslant 2N-1$, $\theta_N = \pi/N$.
   (b) $\Lambda_N = diag(\lambda_0^N, \ldots, \lambda_{2N-1}^N)$ with $\lambda_p^N = \sum_{k=-(N-1)}^{N} a_k \mathrm{e}^{ikp\theta_N}$, $0 \leqslant p \leqslant 2N-1$.
   (c) $\lambda_p^N = \lambda_{2N-p}^N$ for $0 < p \leqslant 2N-1$.
2. There exists an N_0 in $\mathbb{N}$ and an M_0 in $\mathbb{R}^+$ such that C_N^a is positive definite and $1/|\lambda_k^N| < M_0$ for all $N > N_0$ and k in $\mathbb{Z}$.
3. $(C_N^a)^{-1} = C_\xi^{\xi^N}$, where $(\xi^N)_p =: \xi_p^N = 1/2N \sum_{k=-(N-1)}^{N} 1/\lambda_k^N \mathrm{e}^{ipk\theta_N}$ for all $p \in \mathbb{Z}$.

Proof
These results can be easily verified by simple algebraic manipulations. □

Clustering of the preconditioned matrices. In this section, we shall show that the preconditioner that we proposed in the previous section achieves clustering of the eigenvalues. We define the clustering of the spectrum of the sequence of sets of matrices similar to the definition given for sequence of matrices in Reference [7] by

Definion 1
A sequence, $\{\mathscr{L}_N\}$ of sets of matrices is said to have spectra clustered around 1 if for any given $\varepsilon > 0$, there exist positive integers N_0 and N_1 such that for all $Q_p \in \mathscr{L}_N$, $N > N_0$, at most N_1 eigenvalues of the matrix $Q_p - I_p$ have absolute value larger than ε.

Relation to Fourier coefficients of $1/f$. Note that $\lambda_j^N = \sum_{k=-(N-1)}^{N} a_k e^{ijk\theta_N}$ is an approximation for $f(\theta)$ at $j\theta_N$; and $\xi_p^N = \theta_N/2\pi \sum_{k=-(N-1)}^{N} 1/\lambda_k^N e^{ipk\theta_N}$ is a Riemann sum approximation of the integral

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{f(\theta)} \mathrm{e}^{ij\theta} \, \mathrm{d}\theta$$

which is the kth Fourier coefficient of $1/f$. This suggests that the elements of P_p are approximations of the Fourier coefficients of $g(\theta) \triangleq 1/f(\theta)$.

Proposition 3.2

1. There exists a sequence $\{\gamma_k\} \in \ell_1(-\infty,\infty)$ with $\gamma_k = \gamma_{-k}$ such that $g(\theta) = \sum_{k=-\infty}^{\infty} \gamma_k e^{ik\theta}$ for all θ in $[-\pi,\pi]$.
2. $\lim_{N\to\infty} N\|\gamma^N - \xi^N\|^2 = 0$ if $\sum_{k=-(N-1)}^{N} |k^2 a_k^2| < \infty$.

Proof

(1) This is a direct consequence of Theorem 18.21 in Reference [28, p. 367–368].
(2) We define a vector h^N for each N by

$$h_j^N = g(j\theta_N) - \sum_{k=-(N-1)}^{N} \xi_k^N \cos(jk\theta_N)$$

for all $-(N-1) \leqslant j \leqslant N$. But the sum on the right-hand side of this equation is an approximation of $g(j\theta_N)$ and can be simplified as

$$\sum_{k=-(N-1)}^{N} \xi_k^N \cos(jk\theta_N) = \sum_{l=-(N-1)}^{N} \frac{1}{\lambda_l^N} \left(\frac{1}{2N} \sum_{k=-(N-1)}^{N} \cos(lk\theta_N)\cos(jk\theta_N) \right) = \frac{1}{\lambda_j^N}$$

for all $-(N-1) \leqslant j \leqslant N$. Now, $\lambda_j^N = \sum_{k=-(N-1)}^{N} a_k e^{ijk\theta_N}$ is an approximation of $f(j\theta_N)$ and if we define $r_N^a(\theta) \triangleq \sum_{k>N, k\leqslant -N} a_k e^{ik\theta}$ and denote $r_N^a(j\theta_N)$ by R_j^a, then an estimate of $h_j^N = g(j\theta_N) - 1/\lambda_j^N$ can be found as follows:

$$h_j^N = \underbrace{g(j\theta_N)}_{=\frac{1}{f(j\theta_N)}} - \frac{1}{\lambda_j^N} = \frac{\lambda_j^N - f(j\theta_N)}{\lambda_j^N f(j\theta_N)} = -\frac{R_j^a}{\lambda_j^N f(j\theta_N)} \;\Rightarrow\; |h_j^N| \leqslant \underbrace{M_0\|g\|_\infty}_{\triangleq M_2} |R_j^a| \tag{1}$$

This inequality can be rewritten as

$$|h_j^N| = \Bigg| \underbrace{\sum_{k=-(N-1)}^{N} \overbrace{(\gamma_k - \xi_k^N)}^{\triangleq \mu_k^N} \cos(jk\theta_N) + R_j^\gamma}_{x_j^N} \Bigg| \leqslant M_2 |R_j^a| \;\Rightarrow\; |x_j^N| \leqslant M_2 |R_j^a| + |R_j^\gamma| \tag{2}$$

where R_j^γ (and $r_N^\gamma(\theta)$) are defined in the same way as R_j^a (and $r_N^a(\theta)$). Note that the sequence $\{x_j^N\}$, $-(N-1) \leqslant j \leqslant N$ is a discrete Fourier series obtained from the sequence $\{\mu_k^N\}$, $-(N-1) \leqslant k \leqslant N$. Therefore, the coefficients of these two series satisfy the Parseval relationship [26]

$$\|\mu^N\|^2 = \frac{1}{2N}\|x^N\|^2 \tag{3}$$

We assume $\|f'\|_2^2 = \sum_{-\infty}^{\infty} |k^2 a_k^2| < \infty$. This implies that $\|g'(\theta)\|_2^2 = \|f'(\theta)/f^2(\theta)\|_2^2$ is bounded as $f(\theta)$ is bounded away from zero; i.e. $\|g'\|_2^2 = \sum_{-\infty}^{\infty} k^2\gamma_k^2 < \infty$. Let $\varepsilon > 0$. Then, there exists an N_0 in $\mathbb{Z}$ such that $\sum_N^\infty k^2\gamma_k^2 < \varepsilon$ and $\sum_N^\infty |k^2 a_k^2| \leqslant \varepsilon$ for all $N \geqslant N_0$.

This implies

$$\|r_N^\gamma\|_2^2 \leqslant 4\sum_N^\infty \gamma_k^2 \leqslant \frac{4}{N^2}\sum_N^\infty k^2\gamma_k^2 \leqslant \frac{4\varepsilon^2}{N^2}$$

and $\|r_N'\|_2^2 = \sum_N^\infty k^2\gamma_k^2 \leqslant \varepsilon^2$ for all $N \geqslant N_0$. Therefore, by using the Sobolev inequality (see Reference [29, A. 12, p. 496]), we have an upper bound on its infinity norm by

$$\|r_N\|_\infty = c_1\|r_N^\gamma\|_2^{1/2}(\|r_N^\gamma\|_2 + \|(r_N^\gamma)'\|_2)^{1/2} \leqslant \frac{c_2\varepsilon}{\sqrt{N}}$$

where c_1 and c_2 are constants. This implies that $|R_j^\gamma| = |r_N^\gamma(j\theta_N)| \leqslant c_2\varepsilon/\sqrt{N}$. Similarly, we can show that there exists $\tilde{c}_2$ in $\mathbb{R}$ such that $|R_j^a| \leqslant \tilde{c}_2\varepsilon/\sqrt{N}$. Therefore, using inequality 2, we have $|x_j^N| \leqslant M_2|R_j^a| + |R_j^\gamma| \leqslant M_3\varepsilon/\sqrt{N}$, where $M_3 = M_2\tilde{c}_2 + c_2$. Using this estimate in Equation (3), we have

$$\|\gamma^N - \xi^N\|^2 = \|\mu^N\|^2 = \frac{1}{2N}\|x^N\|^2 < \frac{M_3^2\varepsilon^2}{N}$$

for all $N \geqslant N_0$ and, therefore,

$$\lim_{N\to\infty} N\|\gamma^N - \xi^N\|^2 = 0 \qquad \square$$

Remark
It should be noted that $f \in C^q$, $q \geqslant 1$ implies $\sum_{k=-\infty}^\infty k^2|a_k|^2 < \infty$. This implies that

$$\lim_{N\to\infty} N\|\gamma^N - \xi^N\|^2 = 0 \quad \text{if } f \in C^q,\ q \geqslant 1$$

Clustering of the spectrum of LRE matrices. In this section, we define a class of LRE systems and show the clustering properties of the corresponding preconditioned LRE matrices. We define a sequence of sets of LRE matrices associated with the sequence of Toeplitz matrices $\{A_N\}$ in the following way. For every $\varepsilon > 0$, let $N_0(\varepsilon)$ and $N_1(\varepsilon)$ be such that $\sum_{N_0}^\infty ka_k^2 \leqslant \varepsilon$, $\sum_{N_0}^\infty k\gamma_k^2 \leqslant \varepsilon$ and $N\|\gamma^N - \xi^N\|^2 \leqslant \varepsilon$ for all $N > N_1(\varepsilon)$ (this is possible by Proposition 3.2). Then, to every A_N for $N > N_1(\varepsilon)$, we denote a set of LRE matrices by $\mathscr{L}_N^A(\varepsilon)$ whose elements have the form given by $A_p = L_p^{\mathrm{T}} A_N L_p$, where

1. L_p has the structure given by

$$L_p = \begin{pmatrix} I_{p_0} & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ 0 & I_{p_1} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{matrix} \}r_0 \\ \}r_1 \\ \}r_2 \\ \vdots \end{matrix}$$

   and has n_p block-columns (with $\sum_{i=0}^{n_p-1} p_i = p$) and n_r block rows. In this matrix, the ith block-column has only one identity matrix (I_{p_i}) (with all other entries in this block-column being 0); and every alternate block-row is a zero block.
2. $r_i > N_0(\varepsilon)$ for all $0 \leqslant i \leqslant n_r - 1$.

To each $A_p = L_p^{\mathrm{T}} T_N^a L_p \in \mathscr{L}_N^A(\varepsilon)$, we associate a preconditioner as described earlier in this paper; i.e. $P_p = \bar{L}_p^{\mathrm{T}}(C_N^a)^{-1}\bar{L}_p = L_p^{\mathrm{T}} T_\xi^{\xi^N} L_p$. The sequence of sets of preconditioned LRE matrices can be now defined by

$$\mathscr{L}_N(\varepsilon) = \{P_p A_p \text{ such that } A_p = L_p^{\mathrm{T}} T_N^a L_p \in \mathscr{L}_N^A(\varepsilon) \text{ and } P_p = L_p^{\mathrm{T}} T_N^{\xi^N} L_p\}$$

These definitions being given, we present the following proposition.

Proposition 3.3
If $f \in C^q$, $q \geqslant 1$ or if $\sum_{k=-\infty}^{\infty} |k^2 a_k^2| < \infty$, and under the Assumptions 1, for every $\varepsilon > 0$, there exist N_0 and N_1 in $\mathbb{N}$ such that

$$\|I - P_p A_p - D_p\|_{\mathrm{F}} \leqslant \varepsilon \quad \text{for all } P_p A_p \in \mathscr{L}_N(\varepsilon) \quad \text{and} \quad N \geqslant N_1$$

where D_p is a block diagonal matrix which has at most $2n_p$ non-zero $N_0 \times N_0$ blocks.

Before proceeding with the proof of this proposition, we first present a lemma which will be used in it,

Lemma 3.1
Let $\{\alpha_k\}$ be a sequence of real numbers such that $\sum_{k=1}^{\infty} |\alpha_k| < \infty$, and $\{H_N\}$ be a sequence of infinite-dimensional Hankel matrices given by

$$H_N = \begin{pmatrix} \alpha_N & \alpha_{N+1} & \cdots \\ \alpha_{N+1} & \iddots & \iddots \\ \vdots & \iddots & \iddots \end{pmatrix}$$

then

1. $\lim_{N\to\infty} \|H_N\| = 0$.
2. $\lim_{N\to\infty} \|H_N\|_{\mathrm{F}} = 0$ if $\|H_0\|_{\mathrm{F}} < \infty$.

Proof
See Appendix A for the proof. □

Proof of Proposition 3.3
Let $\varepsilon > 0$ and $N_0(\varepsilon)$ and $N_1(\varepsilon)$ be such that $\sum_{N_0}^{\infty} k a_k^2 \leqslant \varepsilon$, $\sum_{N_0}^{\infty} k \gamma_k^2 \leqslant \varepsilon$ and $N\|\gamma^N - \xi^N\|^2 \leqslant \varepsilon$ for all $N > N_1$ (this is possible by Proposition 3.2). Let $A_p \in \mathscr{L}_N^A(\varepsilon)$ for some $N > N_1$ and $P_p = L_p^{\mathrm{T}} T_N^{\xi^N} L_p$ be its preconditioner. We consider the product $P_p A_p$ and study its spectral properties by studying the matrix, $L_p^{\mathrm{T}} T_N^{\xi^N} T_N^a L_p - P_p A_p$. Note that from Proposition A.1, we have that $P_N A_N = T_N^{\xi^N} T_N^a = I + \bar{D}$ where $\bar{D}$ has at most 2 non-zero $N_0 \times N_0$ blocks. Also $L_p^{\mathrm{T}} T_N^{\xi^N} T_N^a L_p$

$-\underbrace{L_p^{\mathrm{T}} T_N^{\xi^N} L_p}_{P_p} \underbrace{L_p^{\mathrm{T}} T_N^a L_p}_{A_p}$ can be rewritten as $L_p^{\mathrm{T}} T_N^{\xi^N} \tilde{L}_p \tilde{L}_p^{\mathrm{T}} T_N^a L_p$ where

$$\tilde{L}_p = \begin{pmatrix} 0 & 0 & \cdots \\ I_{q_o} & 0 & \cdots \\ 0 & 0 & \cdots \\ \vdots & \vdots & \vdots \end{pmatrix} \begin{matrix} \}r_o \\ \}r_1 \\ \}r_2 \\ \vdots \end{matrix}$$

is such that $L_p L_p^{\mathrm{T}} + \tilde{L}_p \tilde{L}_p^{\mathrm{T}} = I$. This implies that

$$I + L_p^{\mathrm{T}} \bar{D} L_p - P_p A_p = L_p^{\mathrm{T}} T_N^{\xi^N} \tilde{L}_p \tilde{L}_p^{\mathrm{T}} T_N^a L_p \tag{4}$$

We first prove the following properties of $L_p^{\mathrm{T}} T_N^{\xi^N} \tilde{L}_p$ and $\tilde{L}_p^{\mathrm{T}} T_N^a L_p$ which we shall use to study the spectrum of $P_p A_p$,

1. $L_p^{\mathrm{T}} T_N^{\xi^N} \tilde{L}_p = L_p^{\mathrm{T}} D_N^\gamma \tilde{L}_p + E_N^\gamma$ where D_N^γ is a block tridiagonal matrix given by

$$D_N^\gamma \triangleq \begin{pmatrix} 0 & R_{01}^\gamma & 0 & \cdots & \cdots & 0 \\ L_{10}^\gamma & 0 & R_{12}^\gamma & \ddots & \ddots & \vdots \\ 0 & \ddots & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & R_{n_r-2,n_r-1}^\gamma \\ 0 & \cdots & \cdots & 0 & L_{n_r-1,n_r-2}^\gamma & 0 \end{pmatrix}$$

   where $R_{ij}^\gamma = D_R^\gamma(N_0, r_i, r_j)$ and $L_{ij}^\gamma = D_L^\gamma(N_0, r_i, r_j)$ (see *Notation*); and $\|E_N^\gamma\|_{\mathrm{F}} \leqslant 2n_p(n_r - n_p)\varepsilon$.
2. $\tilde{L}_p^{\mathrm{T}} T_N^a L_p = \tilde{L}_p^{\mathrm{T}} D_N^a L_p + E_N^a$ where $\|E_N^a\|_{\mathrm{F}} \leqslant 2n_p(n_r - n_p)\varepsilon$ and D_N^a is defined in the same way as D_N^γ.
3. $D_N^\gamma D_N^a$ is a block diagonal matrix with only $2n_r - 2$ non-zero $N_0 \times N_0$ blocks.

(1) and (2) Consider the product $L_p^{\mathrm{T}} T_N^{\xi^N} \tilde{L}_p$. It is independent of the n_r diagonal blocks ($r_i \times r_i$ blocks, $0 \leqslant i \leqslant n_r - 1$) in $T_N^{\xi^N}$ since L_p and $\tilde{L}_p$ are submatrices of the same permutation matrix and satisfy $L_p^{\mathrm{T}} \tilde{L}_p = 0$ (which is easily verified using their structure). Therefore these diagonal blocks in $T_N^{\xi^N}$ can be replaced by zeros and still the product remains unchanged.

Therefore this product can be rewritten as

$$L_p^{\mathrm{T}} \underbrace{\begin{pmatrix} 0 & \tilde{R}_0^{\xi^N} & & & & \\ \tilde{L}_1^{\xi^N} & 0 & \tilde{R}_1^{\xi^N} & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & 0 & \tilde{R}_{n_r-2}^{\xi^N} \\ & & & & \tilde{L}_{n_r-1}^{\xi^N} & 0 \end{pmatrix}}_{\triangleq \tilde{D}_N^{\xi^N}} \tilde{L}_p$$

where $\tilde{D}_N^{\xi^N}$ is obtained by substituting the diagonal blocks (the n_r $r_i \times r_i$ blocks) in $T_N^{\xi^N}$ by zero blocks. Also note that $\tilde{R}_i^{\xi^N}$ and $\tilde{L}_i^{\xi^N}$ are submatrices of $J_N H_N^{\xi^N}$ and $H_N^{\xi^N} J_N$, respectively. From our choice of N_1, we have $\|J_N H_N^{\xi^N} - J_N H_N^{\gamma}\|_{\mathrm{F}} \leqslant \varepsilon$ which implies $\|\tilde{R}_i^{\xi^N} - \tilde{R}_i^{\gamma}\|_{\mathrm{F}} \leqslant \varepsilon$. Therefore

$$\|\tilde{R}_i^{\xi^N} - (R_{ij}^{\gamma}\ 0)\|_{\mathrm{F}} \leqslant \|\tilde{R}_i^{\xi^N} - \tilde{R}_i^{\gamma}\|_{\mathrm{F}} + \|\tilde{R}_i^{\gamma} - (R_{ij}^{\gamma}\ 0)\|_{\mathrm{F}} \leqslant 2\varepsilon$$

where the zero block is of the appropriate size. Similarly, we have $\|\tilde{L}_i^{\xi^N} - (0\ L_{ij}^{\gamma})\|_{\mathrm{F}} \leqslant 2\varepsilon$. Therefore

$$\| \underbrace{L_p^{\mathrm{T}} T_N^{\xi^N} \tilde{L}_p - L_p^{\mathrm{T}} D_N^{\gamma} \tilde{L}_p}_{\triangleq E_N^{\gamma}} \|_{\mathrm{F}} = \|L_p^{\mathrm{T}} (\tilde{D}_N^{\xi^N} - D_N^{\gamma}) \tilde{L}_p\|_{\mathrm{F}} \leqslant 2 n_p (n_r - n_p) \varepsilon$$

(2) This can be proved in the same way as (1).

(3) Note that the matrices R_{ij}^{γ} and R_{ij}^{a} have only a lower left non-zero $N_0 \times N_0$ block; and the matrices L_{ij}^{γ} and L_{ij}^{a} have only a top right non-zero $N_0 \times N_0$ block. This structure implies that the products $R_{ij}^{\gamma} R_{i'j'}^{a} = 0$ and $L_{ij}^{\gamma} L_{i'j'}^{a} = 0$; and the products $R_{ij}^{\gamma} L_{i'j'}^{a}$ and $L_{ij}^{\gamma} R_{i'j'}^{a}$ are block diagonal with only one non-zero $N_0 \times N_0$ block for all $0 \leqslant i, j, i', j' \leqslant n_r - 1$. This implies that the product $D_N^{\gamma} D_N^{a}$ is a block diagonal matrix ($= \mathrm{diag}(R_{01}^{\gamma} L_{10}^{a}, L_{10}^{\gamma} R_{01}^{a} + R_{12}^{\gamma} L_{21}^{a}, \ldots)$) with at most $2n_r - 2$ non-zero $N_0 \times N_0$ blocks. Therefore $L_p^{\mathrm{T}} D_N^{\gamma} D_N^{a} L_p$ is a block diagonal matrix with at most $2n_p - 2$ non-zero $N_0 \times N_0$ blocks.

Also, note that from the structures of D_N^{a} and D_N^{γ}, we have $L_p^{\mathrm{T}} D_N^{\gamma} L_p = 0$ and $L_p^{\mathrm{T}} D_N^{a} L_p = 0$ and therefore $L_p^{\mathrm{T}} D_N^{\gamma} L_p L_p^{\mathrm{T}} D_N^{a} L_p = 0$. Now Equation (4) can be further simplified as

$$\begin{aligned} I + L_p^{\mathrm{T}} \bar{D} L_p - P_p A_p &= L_p^{\mathrm{T}} T_N^{\xi^N} \tilde{L}_p \tilde{L}_p^{\mathrm{T}} T_N^{a} L_p = (L_p^{\mathrm{T}} D_N^{\gamma} \tilde{L}_p + E_N^{\gamma})(\tilde{L}_p^{\mathrm{T}} D_N^{a} L_p + E_N^{a}) \\ &= L_p^{\mathrm{T}} D_N^{\gamma} \tilde{L}_p \tilde{L}_p^{\mathrm{T}} D_N^{a} L_p + \overbrace{L_p^{\mathrm{T}} D_N^{\gamma} L_p L_p^{\mathrm{T}} D_N^{a} L_p}^{=0} + E_N^{\gamma} (\tilde{L}_p^{\mathrm{T}} D_N^{a} L_p + E_N^{a}) + (L_p^{\mathrm{T}} D_N^{\gamma} \tilde{L}_p) E_N^{a} \\ &= L_p^{\mathrm{T}} \underbrace{D_N^{\gamma} D_N^{a}}_{\triangleq \tilde{D}} L_p + E_N^{\gamma} (\tilde{L}_p^{\mathrm{T}} D_N^{a} L_p + E_N^{a}) + (L_p^{\mathrm{T}} D_N^{\gamma} \tilde{L}_p) E_N^{a} \end{aligned}$$

From the structure of D_N^a and extracting matrices $\tilde{L}_p$ and L_p, we have that $\tilde{L}_p^{\mathrm{T}} D_N^a L_p$ has at most $2n_p(n_r - n_p)$ non-zero blocks each of which is a submatrix of C_N^a. Since $\|C_N^a\| \leqslant \sum_{k=-\infty}^{\infty} |a_k| = M_a < \infty$ (from Proposition 3.1(1)), we have $\|\tilde{L}_p^{\mathrm{T}} D_N^a L_p\| \leqslant 2n_p(n_r - n_p)M_a$. Similarly, there exists an $M_\gamma > 0$ (since g is bounded away from 0), we obtain $\|\tilde{L}_p^{\mathrm{T}} D_N^\gamma L_p\| \leqslant 2n_p(n_r - n_p)M_\gamma$. Therefore,

$$\|I - P_p A_p - \underbrace{L_p^{\mathrm{T}}(\tilde{D} - \bar{D})L_p}_{\triangleq D_p}\|_{\mathrm{F}} \leqslant M\varepsilon$$

where $M \triangleq (2n_p(n_r - n_p))^2(M_a + 1 + M_\gamma)$ and D_p is a matrix with at most $2n_p$ non-zero $N_0 \times N_0$ blocks. As $\varepsilon > 0$ and $N > N_1$ were chosen arbitrarily, we have proved the proposition. □

In Proposition 3.3, we showed that $I - P_p A_p$ can be approximated by a matrix which has at least $p - 2n_p N_0$ zero eigenvalues. We exploit the positive definiteness of A_p and Theorem 2 in Reference [30] to prove the following proposition.

Proposition 3.4
If $f \in C^q, q \geqslant 1$ or if $\sum_{k=-\infty}^{\infty} |k^2 a_k^2| < \infty$, and under the Assumptions 1, for every $\varepsilon > 0$, there exist N_0 and N_1 in $\mathbb{N}$ such that there are at least $p - 2n_p N_0$ eigenvalues α_j, of $P_p A_p$ satisfying $\sum(\alpha_j - 1)^2 \leqslant 4\varepsilon^2$ for all $P_p A_p \in \mathscr{L}_N(\varepsilon)$ and $N > N_1$; i.e. the spectrum of the sequence of sets of preconditioned LRE matrices, $\{\mathscr{L}_N(\varepsilon)\}$ is clustered around 1 for every $\varepsilon > 0$.

Proof
Let $\tilde{\varepsilon} > 0$, and $P_p A_p \in \mathscr{L}_N(\tilde{\varepsilon})$, N_0, N_1 and D_p (using Proposition 3.3) be such that

$$\|P_p A_p - \underbrace{(I - D_P)}_{\triangleq \Upsilon}\|_{\mathrm{F}} \leqslant \tilde{\varepsilon}$$

First, we show that $A_p \in \mathscr{L}_N^A(\tilde{\varepsilon})$ is positive definite for some arbitrarily chosen A_p and $\tilde{\varepsilon} > 0$ and then study the spectrum of the symmetric operator $A_p^{1/2} P_p A_p^{1/2}$ which has the same spectrum as $P_p A_p$.
From Proposition 3.1(1), for every N in $\mathbb{N}$ $\|C_N^a\| = \max_j |\lambda_j^N = \sum_{k=-(N-1)}^{N} a_k \mathrm{e}^{ikj\theta_N}| \leqslant \sum_{k=-\infty}^{\infty} |a_k|$ $= M_a < \infty$. Therefore $\|A_p\| = \|\bar{L}_p^{\mathrm{T}} C_N^a \bar{L}_p\| \leqslant \|C_N^a\| \leqslant M_a$, and since A_p is symmetric, $\|(A_p)^{1/2}\|$ $= \|A_p\|^{1/2} \leqslant M_a^{1/2}$. Note that this implies $M_a^{1/2}$ is an upper bound on $\|(A_p)^{1/2}\|$ which is independent of the geometry of the domain and the index N.

From Proposition 3.1(2), there exists an M in $\mathbb{N}$ such that $C_N^a > 0$ for all $N > M$ which implies $x^{\mathrm{T}} C_N^a x > 0$ for all x in $\mathbb{R}^{2N}$, which in turn implies that $y^{\mathrm{T}} A_p y > 0$ for all y in $\mathbb{R}^p$. Therefore, $A_p > 0$ for all $N > M$. Also, there exists an M_0 in $\mathbb{R}$ such that $\lambda_{\min}(A_p) y^{\mathrm{T}} y > (1/M_0) y^{\mathrm{T}} y$. This implies $\|A_p^{-1}\| \leqslant M_0$. Since A_p^{-1} is symmetric, $\|(A_p)^{-1/2}\| = \|A_p^{-1}\|^{1/2} \leqslant M_0^{1/2}$.

If we define $\tilde{\mathscr{T}} \triangleq (A_p)^{1/2} P_p (A_p)^{1/2}$, $\tilde{\Upsilon} \triangleq (A_p)^{1/2} \Upsilon (A_p)^{-1/2}$ and $\varepsilon = (M_0 M_a)^{1/2} \tilde{\varepsilon}$, then we have

$$\|\tilde{\mathscr{T}} - \tilde{\Upsilon}\| \leqslant \|(A_p)^{1/2}\| \|(A_p)^{-1/2}\| \, \|P_p A_p - \Upsilon\|_{\mathrm{F}} \leqslant \varepsilon$$

Using Theorem 2 of Kahan [30], we have $\sum_j(\alpha_j-\eta_j)^2\leqslant 4\|\tilde{\mathscr{T}}-\tilde{\Upsilon}\|_{\mathrm{F}}^2$ where α_j are the eigenvalues of $\tilde{\mathscr{T}}$ and η_j are the real parts of eigenvalues of $\tilde{\Upsilon}$. Therefore $\sum(\alpha_j-1)^2\leqslant 4\varepsilon^2$ for at least $p-2n_pN_0$ values of j. But the eigenvalues of P_pA_p are equal to the eigenvalues of $\tilde{\mathscr{T}}$ (see Theorem 1.3.20 in p. 53 of Horn and Johnson [31]). As $\tilde{\varepsilon}$ is chosen arbitrarily, we have for every $\varepsilon>0$, there exist N_0 and N_1 in $\mathbb{N}$ such that there are at least $N-2n_pN_0$ eigenvalues α_j, of P_pA_p satisfying $\sum(\alpha_j-1)^2\leqslant 4\varepsilon^2$ for all $N>N_1$. Also, since P_pA_p was chosen arbitrarily from $\mathscr{L}_N(\tilde{\varepsilon})$, the sequence of sets of matrices $\{\mathscr{L}_N(\varepsilon)\}$ has clustered spectra around 1.

Now, we show that minimum eigenvalue of P_pA_p is bounded away from zero. It follows from $1/M_0\leqslant\lambda_i(C_a^N)\leqslant M_a$ for all i. This implies $\lambda_{\min}(A_p)\geqslant 1/M_0$ and $\lambda_{\min}(P_p)\geqslant 1/M_a$. Therefore $x^{\mathrm{T}}P_p^{T/2}A_pP_p^{1/2}x\geqslant(1/M_0)x^{\mathrm{T}}P_px\geqslant(1/M_0M_a)x^{\mathrm{T}}x$ for all x which implies $\lambda_{\min}(P_pA_p)=\lambda_{\min}(P_p^{T/2}A_pP_p^{1/2})\geqslant 1/M_0M_a>0$. Hence, the eigenvalues of the preconditioned matrices are bounded away from zero. □

Remark

It should be observed that we assumed a smoothness condition on the generating function f ($\sum k^2|a_k|^2$) to prove the above clustering result. If we relax this condition, i.e. if we assume only absolute summability, then we can show that the sequence of eigenvalues $\{\lambda_k^N(P_pA_p)\}$ and the constant sequence $\{1\}$ are *equally distributed* (see Reference [32] for the definition). This result follows from Proposition 2.2 and Theorems 2.1 from Reference [33] and 3.1 from [32]. However, we cannot guarantee that the sequences are *strongly equally distributed* and therefore cannot guarantee the clustering results proven here.

4. SIMULATION RESULTS

We have shown in previous sections that the preconditioners, P_p, which are extracted from the inverse of the circulant matrix C_N^a yield spectra of $\{P_pA_p\}$ that are clustered around 1. This is desirable from a computational point of view as circulant matrices are easy to invert and the PCGM converges more rapidly if the eigenvalues are clustered (see Reference [34]).

In Figure 3, we provide simulation results for LRE systems with different kernels but defined on the same domain. Each of these LRE systems, $A_px=b$, is described by a 51×51 coefficient matrix, A_p which is a principal submatrix of a corresponding 64×64 Toeplitz matrix, A; i.e. $A_p=L_p^{\mathrm{T}}AL_p$. Since, the extracting matrix, L_p, is completely characterized by the domain of integration (see Section 2), it is the same for both these LRE systems. This domain consists of 3 line segments and the structure of L_p (as laid out in Section 1) is completely specified by the dimensions, $r_0=17$, $r_1=7$, $r_2=17$, $r_3=6$ and $r_4=21$. The kernels of the integral equations are specified by a different A in each of these LRE systems. In all simulations, we assumed that the given vector, $b=[1\ 1\ \cdots]^{\mathrm{T}}$. All the computations were done in MATLAB.

In Figure 3 (Upper), the kernel, A is generated by the function, $f(\theta)=\theta^4+1$. In Reference [7], this function has been used as a test generating function to compare different algorithms to solve Toeplitz systems. Here, we use it to study the LRE systems. In (a), we plot the number of eigenvalues of the preconditioned matrix, P_pA_p, in a ball of radius r centred at 1 vs the radius, r. We observe that a majority (32 out of 51) of the eigenvalues are clustered around 1 (within a radius of 10^{-4}). This clustering of the eigenvalues is exploited by the PCGM and the rapid convergence of the PCGM can be observed in (b). We also see that

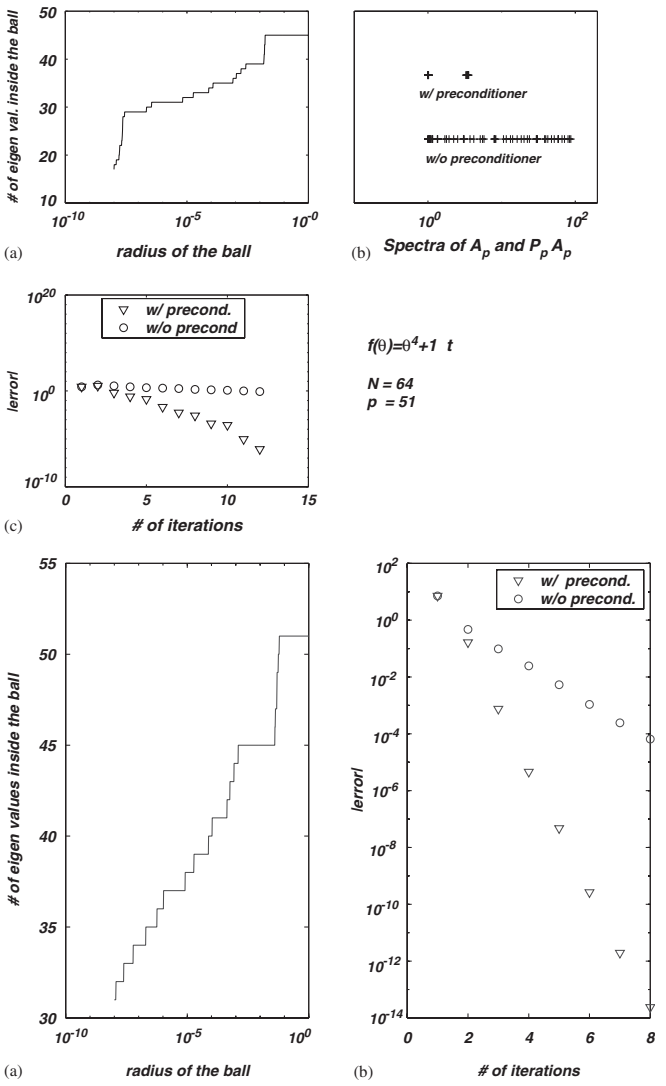

Figure 3. (Upper) The test problem ($N = 64$, $p = 51$): (a) plot of number of eigenvalues (of P_pA_p) within a ball around 1 vs the radius of the ball; (b) clustering of eigenvalues of P_pA_p compared to that of A_p; and (c) comparison of the convergence rates of PCGM between the preconditioned and non-preconditioned cases. (Lower) The three cracks problem ($N = 64$, $p = 51$): (a) plot of number of eigenvalues (of P_pA_p) within a ball around 1 versus the radius of the ball; and (b) comparison of the convergence rates of PCGM between the preconditioned and non-preconditioned cases.

the remaining eigenvalues are not scattered but are in fact clustered about one other point. This secondary clustering, although not captured directly by the analysis presented here, is exploited by the PCGM which achieves good convergence rates as shown in (c).

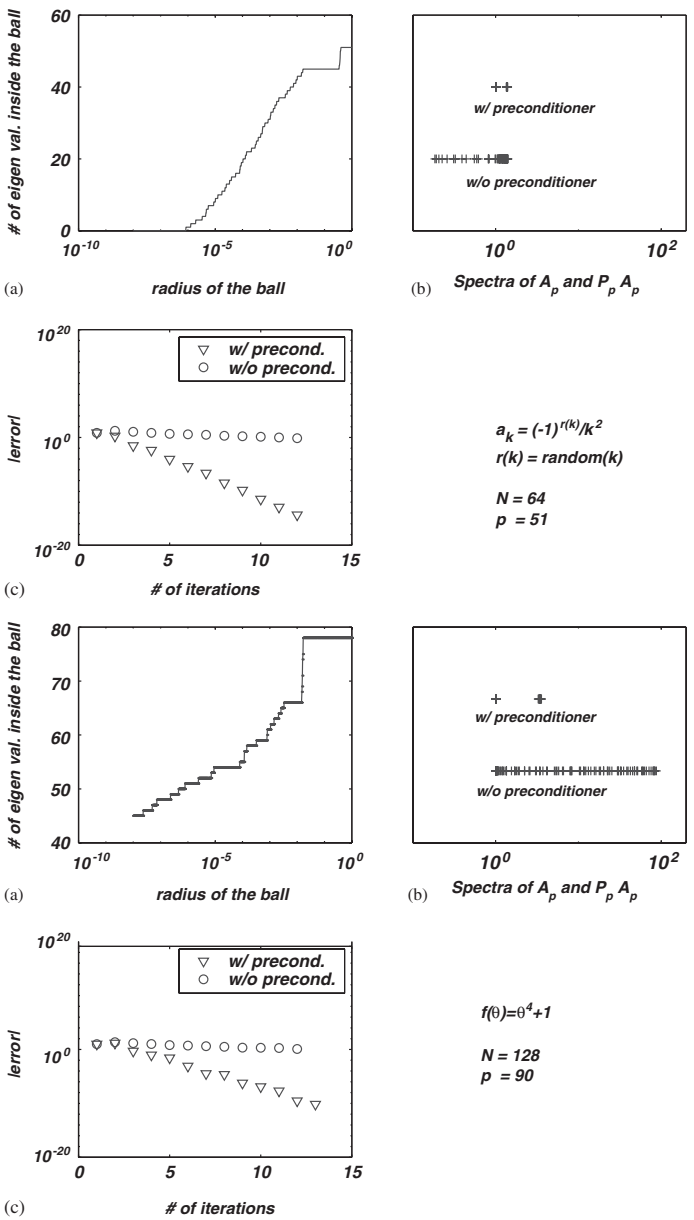

Figure 4. (Upper) LRE system with the generating sequence $\{(-1)^{\mathrm{rand}(k)}/k^2\}$ ($N = 64$, $p = 51$)), (Lower) The test problem ($N = 128$, $p = 90$): (a) plot of number of eigenvalues (of P_pA_p) within a ball around 1 versus the radius of the ball; (b) clustering of eigenvalues of P_pA_p compared to that of A_p; and (c) Comparison of the convergence rates of PCGM between the preconditioned and non-preconditioned cases.

In Figure 3 (Lower), the kernel, A is generated by the function $f(\theta) = 2\pi \sin(|\theta|/2)$ whose Fourier coefficients form the sequence $\{a_n\} = \{-1/(n^2 - \frac{1}{4})\}$. This problem represents the interaction of three cracks (see the mining example in Section 1). In this case too, we observe (see plot (a)) that a majority (40 out of 51) of the eigenvalues are clustered around 1 (within a radius of 10^{-4}). This clustering of the eigenvalues is exploited by the PCGM and the rapid convergence of the PCGM can be observed in (b). In fact, the solution within a tolerance of 10^{-14} is achieved in just 8 iterations. Here, it should be noted that $f(\theta)$ is not bounded away from 0, in fact $f(0) = 0$. Equivalently, in terms of its Fourier coefficients, $\Rightarrow a_0 + 2\sum_{n=1}^{\infty} a_n = 0$. Even though this function does not satisfy the hypotheses that we have assumed, we can explain the convergence of the PCGM for this case by noting that any finite truncation of the Fourier series expansion of $f(\theta)$ is always bounded away from 0. More precisely, in this sequence, $a_0 = 4$ is positive and all the other terms are negative and from the above equation their absolute sum is equal to $a_0/2$. This implies that any finite truncation of $f(\theta)$ given by $f_M(\theta) = \sum_{k=-(M-1)}^{M} a_k e^{ik\theta}$ is bounded away from 0; i.e. $f_M(\theta) > \delta_M \triangleq \sum_{|k|>M+1} |a_k| > 0$. For large enough M ($M \gg N$), the analysis presented in this paper can be shown to be applicable to this generating function.

These convergence trends were found to persist on other simulation performed with other generating sequences as in Figure 4 (Upper). Here the generating sequence is $\{(-1)^{\text{rand}(k)}/k^2\}$, where $\text{rand}(k) \in \mathbb{N}$ is obtained by truncating a random real number, r ($0 < r < 1000$) generated by using the 'rand' function in MATLAB.

We also simulated some LRE systems with different sizes and domains. We also simulated the LRE system with the kernel generated by the function $f(\theta) = \theta^4 + 1$ for a geometry where L_p is specified by $r_0 = 17$, $r_1 = 7$, $r_2 = 17$, $r_3 = 6$, $r_4 = 21$, $r_5 = 10$, $r_6 = 10$, $r_7 = 8$, $r_8 = 15$, $r_9 = 7$ and $r_4 = 21$. Here A_p is a 90×90 matrix and the corresponding A is a 128×128 matrix. The results are shown in Figure 4 (Lower). We again observe similar clustering and convergence properties and also that the spectrum clusters around more than one point.

The clustering, convergence and robustness properties were found to persist in many other simulations (which we do not present here) that we did by changing the kernels, domain geometries and matrix sizes. In this paper, the analytical results concerning the clustering of the preconditioned matrices are presented for a specific class of generating functions; however, the simulations show that these algorithms converge quickly for various generating functions that are not limited to the assumptions made in our analysis.

5. CONCLUSIONS

In this paper, we have introduced and analysed preconditioners (P_p) in PCGM for the efficient solution of lower rank extracted systems (LRES), $A_p x = b$. The elements of the preconditioners are shown to approximate the Fourier coefficients of the reciprocal of the generating function associated with the LRE system. Under fairly mild assumptions on the generating function, $f(\theta)$ or alternatively on the generating sequence $\{a_N\}$ these properties are exploited in order to prove clustering of the eigenvalues of the matrices $P_p A_p$. Also, these systems are shown to be subsystems of Toeplitz systems, $A_N x = b$. For LRES, the PCGM converges to a specified tolerance in $O(N \log N)$ operations where N is the size of A_N. To study the preconditioner, P_p, many simulations of LRES with different kernels, sizes and domains have been presented.

Simulation results corroborate the theoretical findings regarding clustering of the spectra of preconditioned matrices and the associated convergence rates. In particular, the majority of the eigenvalues of P_pA_p fall in the vicinity of 1. They also indicate that the remaining eigenvalues are not scattered but are in fact clustered about two other points. This secondary clustering, although not captured directly by the analysis presented here, can be exploited by the PCGM. In addition, the simulations demonstrate that the algorithm is robust in that it still yields significant clustering even for Toeplitz matrices derived from sequences which did not satisfy the restrictions imposed by the hypotheses of the propositions presented. This indicates that theoretical results established in this paper might be proved under more relaxed conditions.

APPENDIX A

Lemma A.1

$I - P_NA_N$ can be written as a product of two symmetric Hankel matrices, that is, $I - P_NA_N = H_N^{\xi^N} H_N^a$.

Let $\tilde{L}_N = (\mathbf{I}\ \mathbf{0})^{\mathrm{T}}$ be a $2N \times N$ matrix. Then $L_NL_N^{\mathrm{T}} = I - \tilde{L}_N\tilde{L}_N^{\mathrm{T}}$. Therefore,

$$\begin{aligned} I - P_NA_N &= \underbrace{L_N^{\mathrm{T}}(C_N^a)^{-1}C_N^aL_N}_{=I} - \underbrace{L_N^{\mathrm{T}}(C_N^a)^{-1}L_N}_{=P_N}\underbrace{L_N^{\mathrm{T}}C_N^aL_N}_{=A_N} \\ &= L_N^{\mathrm{T}}(C_N^a)^{-1}(I - L_NL_N^{\mathrm{T}})C_N^aL_N = L_N^{\mathrm{T}}C_{\xi}^{\xi^N}\tilde{L}_N\tilde{L}_N^{\mathrm{T}}C_N^aL_N \end{aligned} \tag{A1}$$

Now $\tilde{L}_N^{\mathrm{T}}C_N^aL_N$ picks up the $N \times N$ submatrix obtained by deleting the first N columns and last N rows of the circulant matrix C_N^a. Similarly, $L_N^{\mathrm{T}}(C_N^a)^{-1}\tilde{L}_N$ is obtained from $C_{\xi}^{\xi^N}$. Therefore

$$I - P_NT_N^a = \begin{pmatrix} \xi_N^N & \xi_{N-1}^N & \cdots & \xi_2^N & \xi_1^N \\ \xi_{N-1}^N & \xi_N^N & \xi_{N-1}^N & \cdots & \xi_2^N \\ \xi_{N-2}^N & \xi_{N-1}^N & \xi_N^N & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \xi_1^N & \cdots & \cdots & \xi_{N-1}^N & \xi_N^N \end{pmatrix} \underbrace{J_NJ_N}_{=I} \begin{pmatrix} a_N & a_{N-1} & \cdots & a_2 & a_1 \\ a_{N-1} & a_N & a_{N-1} & \cdots & a_2 \\ a_{N-2} & a_{N-1} & a_N & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_1 & \cdots & \cdots & x_1 & a_N \end{pmatrix} = H_{\xi}^{\xi^N} H_N^a \qquad \square$$

Proof of Lemma 3.1

(1) From Nehari's theorem, we have

$$\|H_N\| = \inf_{v_0, v_{-1}, \ldots} \sup_{\theta} \left| \sum_{-\infty}^{\infty} v_j e^{ij\theta} \right|$$

where $v_j, j \geqslant 1$ are the elements of the first row of H_N. Therefore,

$$\lim_{N\to\infty} \|H_N\| \leqslant \lim_{N\to\infty} \sup_{\theta} \left| \sum_{N}^{\infty} \alpha_j e^{ij\theta} \right| \leqslant \lim_{N\to\infty} \sum_{N}^{\infty} |\alpha_j| = 0$$

(2)

$$\|H_0\|_F < \infty \Rightarrow \sum_{j=0}^{\infty} (j+1)\alpha_j^2 < \infty \Rightarrow \|H_N\|_F^2 = \sum_{N}^{\infty} j\alpha_{N+j}^2 \leqslant \sum_{N}^{\infty} (N+j)\alpha_{N+j}^2 \leqslant \sum_{2N}^{\infty} j'\alpha_{j'}^2 \to 0$$
as $N \to \infty$

Therefore, $\lim_{N\to\infty} \|H_N\|_F = 0$. □

A.1. Clustering of the spectrum of Toeplitz matrices

Here, we present a proposition that shows that in the case of Toeplitz matrices, the product $I - P_N A_N$ can be approximated by a block diagonal matrix with a large 0 block. This case has been analysed [14] and the proposition presented here is very similar to Lemma 7 in Reference [14]. We still present this case, because our analysis of these systems is different from the one given in Reference [14] and the concepts used in this section are used in dealing with LRE systems.

The interpretation that ξ_p^N is an approximation of γ_p is very useful as the analysis of the spectral properties of $P_N A_N$ can be estimated by studying the spectral properties of $T_N^{\gamma} T_N^a$, which is relatively easier to study. To study the spectrum of $T_N^{\gamma} T_N^a$, we use the structure of $H_N^{\gamma} H_N^a = I - T_N^{\gamma} T_N^a$ (see Lemma A.1).

Proposition A.1
If f is in $C^q, q \geqslant 1$ or if $\sum_{k=-\infty}^{\infty} |k^2 a_k^2| < \infty$, then for $\varepsilon > 0$ there exist N_0 and N_1 in $\mathbb{N}$ such that

$$\left\| H_{\xi^N}^N H_N^a - D_N \right\|_F \leqslant \varepsilon \quad \text{for all } N \geqslant N_1$$

where D_N is a block diagonal matrix with only two non-zero $N_0 \times N_0$ blocks.

Proof
This proof exploits the Hankel structure of the matrices H_N^a and H_N^{γ}. We show that these matrices are close to block diagonal matrices (in the Frobenius norm) and then infer the same property for their product. We rewrite H_N^a as

$$H_N^a = \underbrace{\begin{pmatrix} D_1^{N_0} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & J_{N_0} D_1^{N_0} J_{N_0} \end{pmatrix}}_{\triangleq Z_N^a} + H_1 + J_N H_1 J_N \tag{A2}$$

where

$$D_1^{N_0} = \begin{pmatrix} a_1 & a_2 & \cdots & a_{N_0} \\ a_2 & \iddots & \iddots & 0 \\ \vdots & \iddots & \iddots & \vdots \\ a_{N_0} & 0 & \cdots & 0 \end{pmatrix}$$

and

$$H_1 = \begin{pmatrix} 0 & \cdots & 0 & a_{N_0+1} & \cdots & a_N \\ \vdots & \iddots & \iddots & \iddots & \iddots & 0 \\ 0 & \iddots & \iddots & \iddots & \iddots & \vdots \\ a_{N_0+1} & \iddots & \iddots & \iddots & \iddots & \vdots \\ \vdots & \iddots & \iddots & \iddots & \iddots & \vdots \\ a_N & 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

Let $\tilde{\varepsilon} > 0$. Note that $f \in C^q$ with $q > 1$ and/or $\sum_k |k^2 a_k^2| < \infty$ implies $\sum_k k^2 a_k^2 < \infty$ and $\sum_k k^2 \gamma_k^2 < \infty$. Therefore there exist constants such that $\|H_N^a\|_F \leqslant M_a$ and $\|H_N^\gamma\|_F \leqslant M_\gamma$; and N_0, N_1 in $\mathbb{N}$ such that $\sum_{N_0+1}^{\infty} k a_k^2 \leqslant \tilde{\varepsilon}^2$, $\sum_{N_0+1}^{\infty} k \gamma_k^2 \leqslant \tilde{\varepsilon}^2$ and $N \|\gamma^N - \xi^N\|^2 \leqslant \tilde{\varepsilon}^2$ (from Proposition 3.2) for all $N > N_1$. This implies $\|H_1\|_F \leqslant \tilde{\varepsilon}$ and $\|J_N H_1 J_N\|_F \leqslant \tilde{\varepsilon}$. Therefore

$$\| \underbrace{H_N^a - Z_N^a}_{\triangleq E_N^a} \|_F \leqslant 2\tilde{\varepsilon} \quad \text{and similarly} \quad \| \underbrace{H_N^\gamma - Z_N^\gamma}_{\triangleq E_N^\gamma} \|_F \leqslant 2\tilde{\varepsilon}$$

for all $N > N_1$, where Z_N^γ is defined in the same way as Z_N^a (which is defined in Equation (A2). Therefore, if M_a and M_γ represent the upper bounds on $\sum_{k=-\infty}^{\infty} |a_k|$ and $\sum_{k=-\infty}^{\infty} |\gamma_k|$, respectively,

$$\begin{aligned} \|H_N^\gamma H_N^a - \overbrace{Z_N^\gamma Z_N^a}^{D_N}\|_F &\leqslant \|(Z_N^\gamma + E_N^\gamma)(Z_N^a + E_N^a) - Z_N^\gamma Z_N^a\|_F \\ &\leqslant \|Z_N^\gamma\|_F \|E_N^a\|_F + \|E_N^\gamma\|_F \|H_N^a\|_F \leqslant 2(M_\gamma + M_a)\tilde{\varepsilon} \\ \Rightarrow \|H_N^{\xi^N} H_N^a - D_N\|_F &\leqslant \|H_N^{\xi^N} - H_N^\gamma\|_F \|H_N^a\| + \|H_N^\gamma H_N^a - D_N\|_F \\ &\leqslant 2(M_\gamma + M_a)\tilde{\varepsilon} + M_a\tilde{\varepsilon} \triangleq \varepsilon \end{aligned} \tag{A3}$$

As $\varepsilon > 0$ can be chosen arbitrarily, we have proved the proposition. □

ACKNOWLEDGEMENT

The second author wishes to acknowledge the support of the National Science and Engineering Research Council of Canada.

REFERENCES

1. Peirce AP, Spottiswoode S, Napier JAL. The spectral boundary element method: a new window on boundary elements in rock mechanics. *International Journal of Rock Mechanics Mineral Sciences and Geomechanical Abstracts* 1992; **29**(4):379–400.
2. Zhuang Y. On the use of hybrid full-wave cg-fft and spectrally preconditioned cg-fft methods for analyzing large microstrip antenna arrays. *Ph.D. Thesis*, Mc Master University, 7, August 1995.
3. Akhiezer NI. *The Classical Moment Problem* (1st edn). Oliver and Boyd: Edinburgh, 1965.
4. Kailath T. A view of three decades of linear filtering theory. *IEEE Transactions on Information Theory* 1974; **IT-20**:145–181.
5. Willsky AS. *Digital Signal Processing and Control and Estimation Theory* (1st edn). MIT Press: Cambridge, 1979.
6. Haykin S. *Adaptive Filter Theory*. Prentice-Hall: Englewood Cliffs, NJ, 1991.
7. Chan RH, Ng MK. Conjugate gradient methods for Toeplitz systems. *SIAM Review* 1996; **38**(3):427–482.
8. Levinson N. The Wiener RMS error criterion in filter design and prediction. *Journal of Mathematics Physics* 1946; **25**:261–278.
9. Baxter G. Polynomials defined by a difference system. *Journal of Mathematical Analysis and Applications* 1961; **2**:223–263.
10. Trench W. An algorithm for inversion of finite Toeplitz matrices. *SIAM Journal of Applied Mathematics* 1964; **12**:515–522.
11. Zohar S. The solution of Toeplitz set of linear equations. *Journal of Associations for Computational Mathematics* 1974; **21**:272–276.
12. Golub G, Vanloan C. *Matrix Computations* (3rd edn). The Johns Hopkins University Press: London and Baltimore, 1989.
13. Quarteroni A, Sacco R, Fausto S. *Numerical Mathematics* (1st edn). Springer: Berlin, 2000.
14. Chan RH, Ng KP. Toeplitz preconditioners for Hermitian Toeplitz systems. *Linear Algebra and its Applications* 1993; **190**:181–208.
15. Strang G. A proposal for Toeplitz matrix calculations. *Studies in Applied Mathematics* 1986; **74**:171–176.
16. Chan T. An optimal circulant preconditioner for Toeplitz systems. *SIAM Journal of Numerical Analysis* 1993; **30**:1193–1207.
17. Chan R. Circulant preconditioners for Hermitian Toeplitz systems. *SIAM Journal of Matrix Analysis and Applications* 1989; **10**:542–550.
18. Ku T, Kuo C. Design and analysis of Toeplitz preconditioners. *IEEE Transactions on Signal Processing* 1992; **40**:129–141.
19. Tyrtyshnikov E. Optimal and super optimal circulant preconditioners. *SIAM Journal of Matrix Analysis and Applications* 1992; **13**:459–473.
20. Huckle T. Circulant and skew circulant matrices for solving Toeplitz matrix problems. *SIAM Journal of Matrix Analysis and Applications* 1992; **13**:767–777.
21. Hanke M, Nagy J. Toeplitz-approximate inverse preconditioner for banded toeplitz matrices. *Numerical Algorithms* 1994; **7**:183–199.
22. Nagy JG, Plemmons RJ, Torgersen TC. Iterative image restoration using approximate inverse preconditioning. *IEEE Transactions on Image Processing* 1996; **5**(7):1151–1161.
23. Chan T, Olkin J. Circulant preconditioners for Toeplitz-block matrices. *Numerical Algorithms* 1994; **6**:89–101.
24. Chan R, Nagy J, Plemmons R. Circulant preconditioned Toeplitz least squares iterations. *SIAM Journal of Matrix Analysis and Applications* 1994; 15:80–97.
25. King RWP, Harrison CW Jr. *Antennas and Waves: A Modern Approach* (1st edn). MIT Press: Cambridge, MA, 1969.
26. Oppenheim AV, Schafer RW. *Discrete-Time Signal Processing* (1st edn). Prentice-Hall: Englewood Cliffs, NJ, 1989.
27. Chan RH, Strang G. Toeplitz equations by conjugate gradients with circulant preconditioner. *SIAM Journal of Science Statistics and Computers* 1989; **10**:104–119.
28. Rudin W. *Real & Complex Analysis. Mathematics Series* (3rd edn). McGraw-Hill: New York, 1987.
29. Canuto C, Hussaini MY, Quarteroni A, Zang TA. *Spectral Methods in Fluid Dynamics* (1st edn). Springer: New York, 1988.
30. Kahan W. Spectra of nearly Hermitian matrices. *Proceedings of American Mathematical Society* 1975; **48**(1):11–17.
31. Horn RA, Johnson CR. *Matrix Analysis* (1st edn). Cambridge University Press: Cambridge, 1990.
32. Tyrtyshnikov E. A unifying approach to some old and new theorems on distribution and clustering. *Linear Algebra and its Applications* 1996; **232**:1–43.
33. Tyrtyshnikov E. Influence of matrix operations on the distribution of eigenvalues and singular values of Toeplitz matrices. *Linear Algebra and its Applications* 1994; **207**:225–249.
34. Luenberger DG. *Linear and Nonlinear Programming* (2nd edn). Addison-Wesley: Reading, MA, 1984.

