Received: 29 May 2017 | Revised: 26 October 2018 | Accepted: 15 November 2018
DOI: 10.1002/nla.2225

RESEARCH ARTICLE

WILEY

Regularizing properties of a class of matrices including the optimal and the superoptimal preconditioners

Stefano Cipolla[1] | Carmine Di Fiore[2] | Fabio Durastante[3] | Paolo Zellini[2]

[1]Department of Mathematics, University of Padova, Padua, Italy
[2]Department of Mathematics, University of Rome Tor Vergata, Rome, Italy
[3]Department of Computer Science, University of Pisa, Pisa, Italy

Correspondence
Stefano Cipolla, Department of Mathematics, University of Padova, via Trieste 33, 35121 Padua, Italy.
Email: cipolla@math.unipd.it

Funding information
Gruppo Nazionale per il Calcolo Scientifico dell'Istituto Nazionale di Alta Matematica (INdAM-GNCS); 2018 INdAM-GNCS project “Tecniche innovative per problemi di algebra lineare”; MIUR Excellence Department Project awarded to the Department of Mathematics, University of Rome Tor Vergata, Grant/Award Number: CUP E83C18000100006

Summary
In this work, given a positive definite matrix A, we introduce a class of matrices related to A, which is obtained by suitably combining projections of its powers onto algebras of matrices simultaneously diagonalized by a unitary transform. After a detailed theoretical study of some spectral properties of the matrices of this class, which suggests their use as regularizing preconditioners, we prove that such matrices can be cheaply computed when the matrix A has a Toeplitz structure. We provide numerical evidence of the advantages coming from the employment of the proposed preconditioners when used in regularizing procedures.

KEYWORDS
optimal preconditioning, regularizing preconditioners, superoptimal preconditioning

1 | INTRODUCTION

Given a matrix $A \in \mathbb{C}^{n\times n}$ and $\mathcal{L} := \mathrm{sd}\,U = \{Ud(\mathbf{z})U^* \;:\; \mathbf{z} \in \mathbb{C}^n\}$, with $d(\mathbf{z})$ being the diagonal matrix with entries z_i and U being a unitary fixed matrix, let us define the *optimal preconditioners* of A as $\mathcal{L}_A := \arg\min_{X\in\mathcal{L}} \|X - A\|_F$. Optimal circulant preconditioners (i.e., U is the Fourier matrix) have been introduced in the work of T. Chan[1] and studied in the works of R. Chan[2] and Tyrtyshnikov.[3] *Superoptimal circulant preconditioners* of A—obtained solving the problem $\min_{X\in\mathcal{L}} \|AX - I\|_F$—have been introduced in the work of Tyrtyshnikov[3] and studied in the works of Chan et al.[4] and Di Benedetto et al.[5] There is a huge amount of literature studying the properties of these preconditioners and both have been used in the context of several applications; see, for example, the works of Chan et al.[6] and Bertaccini et al.,[7] and the numerous references therein. In the work of Di Benedetto et al.,[5] it has been proved that the circulant superoptimal preconditioners for A = Toeplitz matrix provide a poor approximation of the original Toeplitz matrix when this is ill conditioned: If the considered Toeplitz matrix has a family of eigenvalues collapsing to zero, then the corresponding eigenvalues of the superoptimal preconditioner are well separated from zero; this peculiarity of the superoptimal preconditioners causes a slowdown of the preconditioned conjugate gradient (PCG) method when solving exactly the linear system. In contrast, the inability of the superoptimal preconditioner to accurately approximate the eigenvalues collapsing to zero is a particularly appreciated peculiarity in the field of image restoration and inverse problems where problems are typically ill conditioned and contaminated by noise[8]; in this context, *conjugate-like* iterations are used to obtain regularized solutions where the iterations count plays the role of a regularization parameter. As observed in other works,[5,9–11] the regularization power of the iterations is preserved by introducing a preconditioner such that the action of the preconditioned matrix is comparable with the action of the original matrix on the frequencies space where its eigenvalues collapse to zero. While in the earliest approach[9] this task has been pursued by means of *stabilization* procedures for the optimal circulant preconditioner, in the works of Di Benedetto et al.,[5,10] it has been observed that the superoptimal circulant preconditioner naturally exhibits the desired behavior on the "collapsing" frequencies, ensuring, in this case, that the noise in the data is not amplified by the ill-conditioned components. In the work of Estatico,[12] given a positive definite matrix A, the following class of regularizing preconditioners has been introduced with the aim to enforce the filtering capabilities on the ill-conditioning frequencies for the superoptimal preconditioner:

$$P_{(i)}(A) = \mathcal{L}_{A^2}^{i+1}\mathcal{L}_{A}^{-(2i+1)} \text{ for } i = 0, 1, \dots ;$$

the author shows, moreover, that this type of filtering preconditioners gives satisfactory performances even when the use of the optimal and the superoptimal preconditioners breaks down due to the severe ill conditioning of the problem and the large errors on the data.

In this paper, generalizing a relation intervening between the optimal and the superoptimal preconditioners,[13] we introduce a class of preconditioners of the form

$$P_{(i)}(A) := \mathcal{L}_{A^{2^i}}^{\frac{1}{2^{i}-1}}\mathcal{L}_{A}^{-1} \text{ for } i = 0, 1, \dots ,$$

which includes the optimal and the superoptimal preconditioners. We prove that the matrices in such a class share, in an enhanced form for increasing i, the peculiarities of the circulant superoptimal preconditioner, which are useful when exploited in regularization problems.[5,9,10,12] We prove, moreover, that the proposed preconditioners can be computed cheaply when the coefficient matrix of the linear system has a Toeplitz structure. Finally, we exhibit experimental results confirming the quality of the proposed preconditioners when employed as regularizing preconditioners.

2 | NOTATION

The eigenvalues of a Hermitian matrix $A \in \mathbb{C}^{n\times n}$ will be ordered in nonincreasing order, that is, $\lambda_1(A) \geq \cdots \geq \lambda_n(A)$. Given $A \in \mathbb{C}^{n\times n}$, $d(A)$ will indicate the diagonal matrix which has the same diagonal than A. Given $\mathbf{z} \in \mathbb{C}^n$, $d(\mathbf{z})$ will indicate the diagonal matrix whose diagonal entries are the elements of $\mathbf{z}$. Given $A \in \mathbb{C}^{n\times n}$, we will write $A > (\geq) 0$ iff A is Hermitian positive (semi)definite. Given two matrices $A, B \in \mathbb{C}^{n\times n}$ we will write $A \sim B$ if they are similar. We use the symbol $\kappa_2(A)$ for the condition number of the matrix A in 2-norm.

Given a unitary matrix $U \in \mathbb{C}^{n\times n}$ (i.e., $U^* = U^{-1}$), we define the space $\mathrm{sd}\,U$ of matrices simultaneously diagonalized by U setting

$$\mathcal{L} := \mathrm{sd}\,U = \{Ud(\mathbf{z})U^* \ : \ \mathbf{z} \in \mathbb{C}^n\}. \tag{1}$$

The space $\mathcal{L}$ is a closed subspace of $\mathbb{C}^{n\times n}$, which is a Hilbert space with respect to the inner product $(X, Y) = \sum_{i,j=1}^n \bar{x}_{ij}y_{ij}$. Notice that the norm induced by $(\cdot,\cdot)$ is the Frobenius norm $||X||_F = (\sum_{i,j=1}^n |x_{ij}|^2)^{\frac{1}{2}}$. Thus, by Pythagora's theorem (or equivalently by the Hilbert projection theorem), given a matrix $B \in \mathbb{C}^{n\times n}$, there exists a unique element $\mathcal{L}_B \in \mathcal{L}$ such that

$$||\mathcal{L}_B - B||_F \leq ||X - B||_F, \ \ \forall\, X \in \mathcal{L}, \tag{2}$$

or, equivalently, such that

$$(X, B - \mathcal{L}_B) = 0, \ \forall\, X \in \mathcal{L}.$$

For the sake of completeness, we recall in Lemma 1 few basic well-known results on the projection $\mathcal{L}_B$ of a matrix B onto a $\mathrm{sd}\,U$ subspace.

Lemma 1.

1. *$\mathcal{L}_B = Ud(\mathbf{z}_B)U^H$, where $[\mathbf{z}_B]_i = [U^*BU]_{ii}$, $i = 1, \dots, n$.*
2. *If $B = B^*$, then $\mathcal{L}_B = (\mathcal{L}_B)^*$ and $\min\lambda(B) \leq \lambda(\mathcal{L}_B) \leq \max\lambda(B)$, where $\lambda(X)$ denotes the generic eigenvalue of X. Therefore, $\mathcal{L}_B$ is Hermitian positive definite whenever B is Hermitian positive definite.*
3. $\mathrm{tr}(B) = \mathrm{tr}(\mathcal{L}_B)$, *and if B is Hermitian positive definite, then* $\det(B) \leq \det(\mathcal{L}_B)$.
4. *If $B \in \mathbb{R}^{n\times n}$, then $\mathcal{L}_B \in \mathbb{R}^{n\times n}$ whenever the conjugate of the space $\mathcal{L}$ is included in $\mathcal{L}$, that is, $\overline{\mathcal{L}} \subset \mathcal{L}$ ($\mathcal{L}$ is closed under conjugation).*

3 | MAIN RESULTS

In this section, we introduce and study a class of matrices parameterized by the natural numbers (see Definition 1). Theorem 2 represents the main result in this section and generalizes to this class of matrices an analogous result obtained in the work of Cheng et al.[13] involving the optimal and the superoptimal preconditioners. In order to prove Theorem 2, let us recall the following theorem by Ostrowski (theorem 4.5.9 in the work of Horn et al.[14]) and introduce a preliminary lemma.

Theorem 1. *Let $A, S \in \mathbb{C}^{n\times n}$ with A Hermitian and S nonsingular. Let the eigenvalues of A, SAS^*, and SS^* be arranged in nondecreasing order. Let $\sigma_1 \geq \cdots \geq \sigma_n > 0$ be the singular values of S. For each $k = 1, \dots, n$, there is a positive real number $\theta_k \in [\sigma_n^2, \sigma_1^2]$ such that*

$$\lambda_k(SAS^*) = \theta_k \lambda_k(A).$$

Lemma 2. *For any Hermitian positive definite $A \in \mathbb{C}^{n\times n}$ and unitary $U \in \mathbb{C}^{n\times n}$, we have*

$$d(U^* A^{2^i} U) \geq d(U^* A^{2^{i-1}} U)^2, \quad i \in \mathbb{N} \setminus \{0\}. \tag{3}$$

Proof. By direct computation, exploiting the equality $U^* A^{2^i} U = U^* A^{2^{i-1}} U U^* A^{2^{i-1}} U$. □

Definition 1. For any $\mathcal{L} = \mathrm{sd}\, U$ and $i \in \mathbb{N}$, define

$$P_{(i)}(A) := \mathcal{L}_{A^{2^i}}^{\frac{1}{2^{i-1}}} \mathcal{L}_A^{-1}. \tag{4}$$

Observe that, choosing $i = 0$ in Equation (4), we obtain the optimal preconditioner introduced in the work of Chan,[1] and choosing instead $i = 1$, we obtain the superoptimal preconditioner introduced in the work of Tyrtyshnikov.[3] We will see in Theorem 2 that, for increasing i, it holds an analogous relation to that intervening between the eigenvalues of the matrix A when preconditioned with the optimal and superoptimal preconditioners.[13] We can consider for this reason $P_{(i)}(A)$ as a possible meaningful extension of the abovementioned preconditioners.

Remark 1. From Lemma 2, we have

$$d(U^* A^{2^i} U)^{\frac{1}{2^{i-1}}} \geq d(U^* A^{2^{i-1}} U)^{\frac{1}{2^{i-2}}}$$

or, equivalently,

$$P_{(i)}(A) \geq P_{(i-1)}(A) \text{ for } i \in \mathbb{N}\backslash\{0\},$$

and hence, for $i \in \mathbb{N}$, it holds

$$\max_k \lambda_k(P_{(i)}(A)) \geq \max_k \lambda_k(P_{(i-1)}(A)),$$
$$\min_k \lambda_k(P_{(i)}(A)) \geq \min_k \lambda_k(P_{(i-1)}(A)).$$

Observe moreover that, applying repeatedly Lemma 2, it follows that

$$d(U^* A^{2^i} U) \geq \cdots \geq d(U^* A U)^{2^i}$$

and, hence,

$$d(U^* A U)^{-1} d(U^* A^{2^i} U)^{\frac{1}{2^i}} \geq I_{n\times n}. \tag{5}$$

Theorem 2. *Given a Hermitian positive definite $A \in \mathbb{C}^{n\times n}$, for any $i \in \mathbb{N} \setminus \{0\}$, we have*

$$\lambda_k\left((P_{(i)}(A))^{-1} A\right) \leq \lambda_k\left((P_{(i-1)}(A))^{-1} A\right) \quad k = 1, \dots, n. \tag{6}$$

Proof.

$$\begin{aligned}
(P_{(i)}(A))^{-1} A &= (P_{(i)}(A))^{-1} P_{(i-1)}(A) (P_{(i-1)}(A))^{-1} A \\
&= \mathcal{L}_{A^{2^i}}^{-\frac{1}{2^{i-1}}} \mathcal{L}_{A^{2^{i-1}}}^{\frac{1}{2^{i-2}}} \mathcal{L}_{A^{2^{i-1}}}^{-\frac{1}{2^{i-2}}} \mathcal{L}_A A \\
&\sim d(U^* A^{2^i} U)^{-\frac{1}{2^{i-1}}} d(U^* A^{2^{i-1}} U)^{\frac{1}{2^{i-2}}} d(U^* A^{2^{i-1}} U)^{-\frac{1}{2^{i-2}}} d(U^* A U) U^* A U \\
&\sim DMD,
\end{aligned}$$

where

$$D = d(U^*A^{2^i}U)^{-\frac{1}{2^i}} d(U^*A^{2^{i-1}}U)^{\frac{1}{2^{i-1}}}$$

and

$$M = d(U^*A^{2^{i-1}}U)^{-\frac{1}{2^{i-1}}} d(U^*AU)^{\frac{1}{2}} U^*AUd(U^*AU)^{\frac{1}{2}} d(U^*A^{2^{i-1}}U)^{-\frac{1}{2^{i-1}}}.$$

From (3), we have $D_{ii} \in (0, 1]$, and hence, assertion follows from Theorem 1 where $S = D$ and $A = M$, observing that

$$M \sim (P_{(i-1)}(A))^{-1}A.$$

□

Corollary 1. *For every $k = 1, \dots, n$, there exist $\lambda_k^{\infty\downarrow}$ and $\lambda_k^{\infty\uparrow}$ such that*

$$\lim_{i\to\infty} \lambda_k\left((P_{(i)}(A))^{-1}A\right) = \lambda_k^{\infty\downarrow},$$
$$\lim_{i\to\infty} \lambda_k\left(P_{(i)}(A^{-1})A\right) = \lambda_k^{\infty\uparrow}.$$

Proof. Observe that, using Lemma 1, we have

$$\lambda\left(\mathcal{L}_{A^{2^i}}^{\frac{1}{2^{i-1}}}\right) \in \left[\lambda_n(A^{2^i})^{\frac{1}{2^{i-1}}}, \lambda_1(A^{2^i})^{\frac{1}{2^{i-1}}}\right] = \left[\lambda_n(A)^2, \lambda_1(A)^2\right]$$

and, hence,

$$\lambda(P_{(i)}(A)) \in \left[\frac{\lambda_n(A)^2}{\lambda_1(A)}, \frac{\lambda_1(A)^2}{\lambda_n(A)}\right]. \tag{7}$$

Assertion follows from the Bolzano–Weierstrass theorem, observing that

$$\lambda((P_{(i)}(A))^{-1}A) \in \left[\frac{\lambda_n(A)^2}{\lambda_1(A)^2}, \frac{\lambda_1(A)^2}{\lambda_n(A)^2}\right]$$

and observing that, from Theorem 2, $\lambda_k((P_{(i)}(A))^{-1}A)$ is a monotonically decreasing sequence for each $k = 1, \dots, n$.

For the second part, applying Theorem 2 to the inverse matrix A^{-1}, we obtain

$$\lambda_k\left((P_{(i)}(A^{-1}))^{-1}A^{-1}\right) \le \lambda_k\left((P_{(i-1)}(A^{-1}))^{-1}A^{-1}\right) \quad k = 1, \dots, n$$

and, hence,

$$\lambda_k\left(P_{(i)}(A^{-1})A\right) \ge \lambda_k\left(P_{(i-1)}(A^{-1})A\right) \quad k = 1, \dots, n. \tag{8}$$

Observe that (8) is now a monotonic increasing sequence and assertion follows as in the previous case. □

Remark 2. Using Proposition 1, Lemma 2, and analogous techniques to those in Theorem 2 and Corollary 1, it is possible to prove that also the sequences $\{\lambda_k(P_{(i)}(A)A)\}_{i\in\mathbb{N}}$ are monotonic increasing and convergent for every $k = 1, \dots, n$. Interestingly enough, using $P_{(i)}(A)$ and $(P_{(i)}(A))^{-1}$, it is possible to produce, in some cases, better clustering effects on A. To this extent, observe that

$$\min_{a,b\in\mathbb{R}} \left\| \left(aP_{(i)}(A) + b(P_{(i)}(A))^{-1}\right) A - I \right\|_F$$

has a nontrivial solution, namely, a, b are obtained solving, when possible, the following linear system:

$$\begin{bmatrix} \mathrm{tr}\left(A^2(P_{(i)}(A))^2\right) & \mathrm{tr}(A^2) \\ \mathrm{tr}(A^2) & \mathrm{tr}\left(A^2(P_{(i)}(A))^{-2}\right) \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \mathrm{tr}\left(A(P_{(i)}(A))\right) \\ \mathrm{tr}\left(A(P_{(i)}(A))^{-1}\right) \end{bmatrix}.$$

Theorem 3 gives a more accurate bound of $\lambda(P_{(i)}(A))$ if compared with the bound contained in (7). We need the following preliminary lemma.

Lemma 3. *If $A \in \mathbb{C}^{n\times n}$ is a Hermitian positive definite matrix, then, defining $M = \max_k A_{kk}$, we have*

$$|\Re A_{hk}| \le M \text{ and } |\Im A_{hk}| \le M \text{ for all } h, k \in \{1, \dots, n\}.$$

Proof. For $h \neq k$, choose the vectors $\mathbf{x}_1 = \mathbf{e}_h - \mathbf{e}_k$, $\mathbf{x}_2 = \mathbf{e}_h + \mathbf{e}_k$, $\mathbf{x}_3 = \imath\mathbf{e}_h - \mathbf{e}_k$, $\mathbf{x}_4 = \imath\mathbf{e}_h + \mathbf{e}_k$ and use the fact that $\mathbf{x}_l^*A\mathbf{x}_l$ for $l = 1, 2, 3, 4$ are strictly positive scalars. □

Theorem 3. *For every $i \in \mathbb{N}$, we have*

$$\sigma(P_{(i)}(A)) \subset [\beta_{\min}, C_{(i)}\beta_{\max}], \tag{9a}$$

$$\kappa_2(P_{(i)}(A)) \leq \min\left\{\kappa_2(A)^3, (2n-1)^{\frac{1}{2^{i-1}}}\kappa_2(A)^{\frac{1}{2^{i-2}}}\kappa_2(\mathcal{L}_A)\right\}, \tag{9b}$$

being $\beta_{\min} = \min_k (U^*AU)_{kk}$, $\beta_{\max} = \max_k (U^*AU)_{kk}$, and $C_{(i)} \geq 1$ a suitable constant.

Proof. For the first part, to ease the notation, define

$$B := U^*AU \quad \text{and} \quad X_{(i)} := d(B)^{-2^{i-2}}(B^{2^{i-1}})d(B)^{-2^{i-2}}.$$

It can be easily checked that $P_{(i)}^{2^{i-1}}(A) = Ud(X_{(i)}d(B)^{2^{i-1}}X_{(i)})U^*$ and that

$$\beta_{\min}^{2^{i-1}}X_{(i)}^2 \leq X_{(i)}d(B)^{2^{i-1}}X_{(i)} \leq \beta_{\max}^{2^{i-1}}X_{(i)}^2. \tag{10}$$

From (10), it follows that

$$\beta_{\min} d\left(X_{(i)}^2\right)^{\frac{1}{2^{i-1}}} \leq d\left(X_{(i)}d(B)^{2^{i-1}}X_{(i)}\right)^{\frac{1}{2^{i-1}}} \leq \beta_{\max} d\left(X_{(i)}^2\right)^{\frac{1}{2^{i-1}}},$$

where $d(X_{(i)}d(B)^{2^{i-1}}X_{(i)})^{\frac{1}{2^{i-1}}}$ are the eigenvalues of $P_{(i)}(A)$. To complete the proof, define $M_{(i)} := \max_k (X_{(i)})_{kk}$. We have that

$$1 \leq \left(X_{(i)}^2\right)_{kk}^{\frac{1}{2^{i-1}}} \leq \left(M_{(i)}^2 + 2M_{(i)}^2(n-1)\right)^{\frac{1}{2^{i-1}}} \text{ for all } k \in 1, \ldots, n,$$

where the first inequality follows observing that $(X_{(i)}^2)_{kk} \geq (X_{(i)})_{kk}^2$ and (5), and the second inequality follows instead from Lemma 3. Define

$$C_{(i)} := \left(M_{(i)}^2(2n-1)\right)^{\frac{1}{2^{i-1}}}.$$

For the second part, let us bound the constant $M_{(i)}$. Observe that if we consider the 2-norm, we have for $i \in \mathbb{N}$ that

$$M_{(i)} \leq \rho(X_{(i)}) = \rho\left(d(B)^{-2^{i-1}}B^{2^{i-1}}\right) \leq \|d(B)^{-2^{i-1}}\|_2\|B^{2^{i-1}}\|_2 \Rightarrow (M_{(i)})^{\frac{1}{2^{i-1}}} \leq \|d(B)^{-2^{i-1}}\|_2^{\frac{1}{2^{i-1}}}\|B^{2^{i-1}}\|_2^{\frac{1}{2^{i-1}}} = \frac{\lambda_1(A)}{\beta_{\min}},$$

and hence,

$$1 \leq C_{(i)} \leq \left(\frac{\lambda_1(A)}{\beta_{\min}}\right)^{\frac{1}{2^{i-2}}}(2n-1)^{\frac{1}{2^{i-1}}}. \tag{11}$$

Observe finally that, from (11), we have $\lim_{i\to\infty} C_{(i)} = 1$. □

Before concluding this section, let us observe that an analogous bound to (9a) was derived in the work of Di Benedetto et al.[5] for the superoptimal preconditioner, and hence, Theorem 3 could be considered as an extension of Theorem 3.4 in the work of Di Benedetto et al.[5]

Observe, moreover, that (9b) could be particularly relevant if the matrix algebra $\mathcal{L} = \text{sd}\, U$ is chosen such that $\kappa_2(\mathcal{L}_A) << \kappa_2(A)$ (see the Appendix for a possible construction of such $\mathcal{L}$). In fact, once such $\mathcal{L}$ is available, the corresponding sequence of preconditioners $P_{(i)}(A) = \mathcal{L}_{A^{2^i}}^{\frac{1}{2^{i-1}}}\mathcal{L}_A^{-1}$ must be, by (9b), such that $\kappa_2(P_{(i)}(A)) \approx \kappa_2(\mathcal{L}_A)$ when $i \in \mathbb{N}$ is sufficiently large. Thus, it is possible to introduce and use preconditioners $P_{(i)}(A)$, which satisfy property (6)—a property that favors regularizing properties (see Section 5)— without resulting in a significant deterioration of the condition number.

Remark 3. It is possible to compute directly $\lim_{i\to\infty} P_{(i)}(A)$. To this end, let us define $\mathbf{u}_k := U\mathbf{e}_k$ for $k = 1, \ldots, n$, and write $\mathbf{u}_k = \sum_{j=1}^n \alpha_j^{(k)}\mathbf{v}_j$, where $A\mathbf{v}_j = \lambda_j(A)\mathbf{v}_j$. We have

$$\lim_{s\to+\infty} (\mathbf{u}_k A^s \mathbf{u}_k)^{\frac{1}{s}} = \lim_{s\to+\infty}\left(\sum_{j=1}^n \left(\alpha_j^{(k)}\right)^2 \lambda_j(A)^s\right)^{\frac{1}{s}} = \lambda_{(1)}(A),$$

and hence,

$$\lim_{i\to+\infty} P_{(i)}(A) = Ud(\mathbf{b})U^* \text{ where } \mathbf{b} = \left[\frac{\lambda_1(A)^2}{(U^*AU)_1}, \ldots, \frac{\lambda_1(A)^2}{(U^*AU)_n}\right].$$

4 | THE TOEPLITZ CASE

In this section, we will prove that, when T is a Toeplitz matrix and $C := \mathrm{sd}F$, where F is the Fourier matrix (i.e., C is the algebra of circulant matrices), the preconditioners $P_{(i)}(T)$ can be computed cheaply for moderate values of $i \in \mathbb{N}$; indeed, "when using superoptimal preconditioners in practice, one obviously should be assured that there is a way to compute them sufficiently quickly."[3]

4.1 | An inside to Toeplitz and Toeplitz-like structures

Let us start introducing some notations, definitions, and results. What follows in this subsection is entirely borrowed from the work of Kressner et al.,[15] and hence, we refer there for further details and proofs.

Definition 2. Define the *displacement* $\nabla_Z(A)$ *of* $A \in \mathbb{C}^{n\times n}$ as

$$\nabla_Z(A) := A - ZAZ^*, \quad Z = \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix},$$

and the rank of $\nabla_Z(A)$ as the *displacement rank* of A. Moreover, we define the Toeplitz-like matrices as those with small displacement rank.

Definition 3. For $r \geq \mathrm{rank}(\nabla_Z(A))$, we call a pair of matrices (G, B), where $G, B \in \mathbb{C}^{n\times r}$, generator for A if

$$\nabla_Z(A) := A - ZAZ^* = GB^*.$$

The matrix A can be reconstructed from generators as

$$A = \mathcal{T}(G, B) := \sum_{k=0}^{n-1} Z^k GB^* (Z^*)^k. \tag{12}$$

If we denote by $\mathbf{g}_j, \mathbf{b}_j \in \mathbb{C}^n$ respectively the columns of G, B for $j = 1, \dots, r$, we can rewrite (12) as

$$A = \mathcal{T}(G, B) = \sum_{k=1}^{r} L(\mathbf{g}_j) U\left(\mathbf{b}_j^*\right),$$

where $U(\mathbf{x})$ and $L(\mathbf{x})$ are the triangular Toeplitz matrices

$$L(\mathbf{x}) = \begin{pmatrix} x_0 & 0 & \cdots & 0 \\ x_1 & x_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ x_{n-1} & \cdots & x_1 & x_0 \end{pmatrix}, \quad U(\mathbf{x}) = \begin{pmatrix} x_0 & x_1 & \cdots & x_{n-1} \\ 0 & x_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & x_1 \\ 0 & \cdots & 0 & x_0 \end{pmatrix}.$$

Observe, moreover, that if $T \in \mathbb{C}^{n\times n}$ is the Toeplitz matrix $T_{ij} = t_{i-j}$ for $i, j \in \{0, \dots, n-1\}$, we have that $\mathrm{rank}(\nabla_Z(T)) = 2$ and a minimal generator is

$$G = \begin{pmatrix} t_0 & 1 \\ t_1 & 0 \\ \vdots & \vdots \\ t_{n-1} & 0 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & \bar{t}_{-1} \\ \vdots & \vdots \\ 0 & \bar{t}_{-(n-1)} \end{pmatrix}.$$

Lemma 4. *Let T be a Toeplitz matrix. Then, T^s is a Toeplitz-like matrix of displacement rank at most $2s$ for any integer $s \geq 1$. Letting (G, B) denote the generator for T, a sequence of (nonminimal) generators $(G_1, B_1), \dots, (G_s, B_s)$ for $T, T^2 \dots, T^s$ is given by*

$$G_1 = G, \quad G_{i+1} = \left[P_G^i G \;\; P_G^{i-1} G \;\; \cdots \;\; G \;\; -P_G \mathbf{e}_1 \;\; \cdots \;\; -P_G^i \mathbf{e}_1 \right],$$
$$B_1 = B, \quad B_{i+1} = \left[B \;\; P_B B \;\; \cdots \;\; P_B^i B \;\; P_B^i \mathbf{e}_1 \;\; \cdots \;\; P_B \mathbf{e}_1 \right],$$

for $i = 1, \dots, s-1$, where $P_G := (Z - I)T(Z - I)^{-1}$ and $P_B := (Z - I)T^*(Z - I)^{-1}$. Moreover,

$$\mathbf{e}_1 \in \mathrm{range}(G_1) \subset \cdots \subset \mathrm{range}(G_s) \text{ and } \mathbf{e}_1 \in \mathrm{range}(B_1) \subset \cdots \subset \mathrm{range}(B_s).$$

Corollary 2. *Let $T \in \mathbb{C}^{n\times n}$ be a Toeplitz matrix, then a set of generators for the monomial $T, T^2, \dots, T^s$ can be computed with $O(sn\log_2(n))$ operations.*

Proof. Applying $(Z - I)^{-1}$ to a vector amounts simply to computing the vector of its cumulative sums[15] and the application of $Z - I$ to a vector can be evaluated with $n - 1$ subtractions. The multiplication of a Toeplitz matrix by a vector can be performed in $O(n \log(n))$. □

4.2 | Projecting the power of Toeplitz matrices onto the Circulant Algebra

Let us start recalling theorem 5.1 and corollary 1 in the work of Tyrtyshnikov.[3]

Theorem 4. *Let $M = LR \in \mathbb{C}^{n\times n}$, where*

$$L = \begin{pmatrix} l_0 & 0 & \cdots & 0 \\ l_1 & l_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ l_{n-1} & \cdots & l_1 & l_0 \end{pmatrix}, \quad R = \begin{pmatrix} r_0 & r_1 & \cdots & r_{n-1} \\ 0 & r_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_1 \\ 0 & \cdots & 0 & r_0 \end{pmatrix}.$$

Then, defining

$$s_k(M) := \sum_{i-j=k \mod n} m_{ij} \text{ where } i, j, k \in \{0, \dots, n-1\},$$

we have

$$\begin{pmatrix} s_0 \\ s_1 \\ \vdots \\ s_{n-1} \end{pmatrix} = P \begin{pmatrix} r_{n-1} \\ r_{n-2} \\ \vdots \\ r_0 \end{pmatrix} + L \begin{pmatrix} 0 \\ r_{n-1} \\ \vdots \\ (n-1)r_1 \end{pmatrix},$$

where

$$P = \begin{pmatrix} l_{n-1} & 2l_{n-2} & \cdots & (n-1)l_1 & nl_0 \\ 0 & l_{n-1} & \ddots & \ddots & (n-1)l_1 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 2l_{n-2} \\ 0 & \cdots & \cdots & 0 & l_{n-1} \end{pmatrix}.$$

Corollary 3. *If $M = LR$ as in Theorem 4, then the values s_k for $0 \le k \le n-1$, can be computed in $O(n \log_2(n))$ operations.*

Theorem 5. *Given a Toeplitz matrix $T \in \mathbb{C}^{n\times n}$, then*

$$\mathcal{L}_{T^s} := \arg\min_{C\in\mathcal{C}} \|T^s - C\|_F \tag{13}$$

can be computed in $O(2sn \log_2(n))$ operations.

Proof. It is well known that any matrix $C \in \mathcal{C}$ can be written in the form

$$C = \sum_{k=0}^{n-1} c_j Q^j, \tag{14}$$

where

$$Q = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 0 & 0 \end{pmatrix},$$

and that $C_{T^s} = \sum_{j=0}^{n-1} \alpha_j Q^j$, where $\alpha_j = [B^{-1}\mathbf{c}]_j$, $B_{i,j} = (Q^i, Q^j)$ and $c_j = (Q^j, T^s)$ for $i, j = 0, \dots, n-1$.
Observe, moreover, that, from Lemma 4, we have

$$T^s = \sum_{i=1}^{2s} L\left(\mathbf{g}_i^s\right) U\left(\mathbf{b}_i^s\right) \tag{15}$$

being $\mathbf{g}_i^s$, $\mathbf{b}_i^s$ for $i = 1, \dots, 2s$, the columns of the generators (G_s, B_s) of T^s. The thesis follows from the linearity of the trace, from Theorem 4, and from observing that $s_k(M) = \mathrm{tr}(Q^k M)$ (direct computation). □

To conclude, let us point out that, using the representation of the optimal preconditioner from the works of Chan et al.[16,17] and Bini et al.,[18] and formula (15), the results presented in this section could be easily extended to Sine, Cosine, and Hartley's trigonometric matrix algebras (see other works[18–23] for a general overview on these algebras in the context of preconditioning and displacement theory).

5 | EXPERIMENTAL RESULTS

In the works of Di Benedetto et al.[5,10] and Estatico,[12] it has been proved that when $C = \mathrm{sd}F$ (F Fourier matrix) and A is a Toeplitz matrix, under suitable hypotheses on the generating function of T, the spectrum of $P_{(1)}(A)$ stays bounded from below. In these papers, this particular feature of the superoptimal preconditioner has been profitably exploited in order to use $P_{(1)}(A)$ as a regularizing preconditioner for A because it provides an approximation of the matrix, which ignores some "bad" frequencies corresponding to small eigenvalues. From Remark 1, we have that

$$P_{(i)}(A) \geq P_{(1)}(A) \text{ for } i \geq 2,$$

and hence, the same property holds for the class of preconditioners $P_{(i)}(A)$ presented in this paper whenever it holds for $P_{(1)}(A)$. In particular, using Theorem 2, we can infer that such class of preconditioners presents the same regularizing behavior of the superoptimal preconditioner.

In this section, some preliminary numerical experiences are carried on. Such experiments confirm that the preconditioners proposed in Section 3 could be suitably employed as regularizing preconditioners for the conjugate gradient (CG) and conjugate gradient least square (CGLS) methods.[5,9,10,12] To this end, we compare the methods and the preconditioners in terms of the relative restoration error (RRE) of the reconstructed signal and the number of iterations needed to produce such a solution. Specifically, we define $k_{m.e.}$ as the number of iterations at which the minimal RRE $\varepsilon := \|\mathbf{f} - \mathbf{x}_{k_{m.e.}}\|/\|\mathbf{f}\|$ is obtained, being $\mathbf{x}_{k_{m.e.}}$ and $\mathbf{f}$ respectively the reconstructed and the true signals.

Let us point out that, as stopping criterion for Krylov-type methods, when the noise η is of Gaussian type and its norm is known or well estimated, a somewhat canonical choice is represented by the *discrepancy principle*, that is, one stops the iterations for solving $A\mathbf{x} = \mathbf{g}$ as soon as

$$\|\mathbf{g} - A\mathbf{x}^{(k)}\| \leq \beta\|\boldsymbol{\eta}\|, \quad \beta > 1 \text{ and } \beta \approx 1.$$

We refer to the works of Dykes et al.,[11] Gazzola et al.,[24] and Morozov[25] for further discussion on this issue. However, in the following, we are going to focus on the best achievable RRE and on the number of iterations needed to compute this reconstruction because introducing the stopping criterion would make the comparison between the classic optimal/superoptimal preconditioners and the proposed ones harder to discern. Indeed, in any complete applicative setting, the discrepancy principle is easily built.

All the numerical experiments are performed on a laptop running Linux with 8-Gb memory and CPU Intel® Core™ i7-4710HQ CPU with a clock at 2.50 GHz. The code is written and executed in MATLAB R2018a. We used the numerical routines from the work of Nagy et al.[26] for both image treatment and iterative methods.

5.1 | Experiment 1

We focus on the solution of the system $A\mathbf{x} = \mathbf{g}$, where $A \in \mathbb{R}^{n\times n}$ is a symmetric severely ill-conditioned one-level Toeplitz matrix and the right-hand side is a smooth one-dimensional signal $\mathbf{g} \in \mathbb{R}^n$ contaminated by noise. We take the same test problem of Di Benedetto et al.,[10] Estatico,[12] and Eldén,[27] and we choose the matrix algebra $\mathcal{L} = C := \mathrm{sd}F$. The dimension of the system A is set as $n = 2^k$, $k = 7, \ldots, 11$, and the true signal $\mathbf{f}$ is the sum of two different "impulses"

$$(\mathbf{f})_j = f_j = 0.5k_{0.1}(x_j + 0.9) + k_{0.05}(x_j - 0.8), \quad x_j = -2 + \frac{4}{n+1}(j+1), \tag{16}$$

where the points x_j are equally distributed in $[-2, 2]$ and $k_\sigma(t)$ denotes the Gaussian distribution

$$k_\sigma(t) := \frac{1}{2\sqrt{\pi\sigma}}e^{-\frac{t^2}{4\sigma}}.$$

TABLE 1 Experiment 1. Numerical results for the reconstruction of the signal (16) with the PCG and PCGLS methods with preconditioners $P^{(i)}$

			CGLS		PCGLS		CG		PCG	
n	α	i	$k_{m.e.}$	ε	$k_{m.e.}$	ε	$k_{m.e.}$	ε	$k_{m.e.}$	ε
128	1.0e−01	0	7	6.10e−01	1	1.00e+00	1	1.00e+00	1	1.00e+00
		1			3	6.03e−01			1	1.00e+00
		2			3	6.43e−01			2	6.74e−01
		3			6	6.39e−01			3	6.78e−01
		4			11	6.34e−01			3	6.45e−01
		5			11	6.44e−01			4	6.27e−01
	1.0e−02	0	17	4.07e−01	1	1.00e+00	3	5.80e−01	1	1.00e+00
		1			12	4.04e−01			2	4.91e−01
		2			19	4.02e−01			4	4.46e−01
		3			26	3.92e−01			8	4.38e−01
		4			30	3.92e−01			14	4.32e−01
		5			36	3.96e−01			14	4.31e−01
	1.0e−03	0	25	3.36e−01	1	1.00e+00	7	4.24e−01	1	1.00e+00
		1			22	3.16e−01			4	3.82e−01
		2			27	3.02e−01			13	3.59e−01
		3			38	3.00e−01			18	3.64e−01
		4			46	2.98e−01			18	3.56e−01
		5			51	2.98e−01			18	3.53e−01
256	1.0e−01	0	3	7.41e−01	1	1.00e+00	1	1.00e+00	1	1.00e+00
		1			1	1.00e+00			1	1.00e+00
		2			2	7.24e−01			1	1.00e+00
		3			3	7.27e−01			2	7.47e−01
		4			5	7.34e−01			2	7.44e−01
		5			6	7.31e−01			2	7.49e−01
	1.0e−02	0	14	4.75e−01	1	1.00e+00	2	6.47e−01	1	1.00e+00
		1			3	8.63e−01			1	1.00e+00
		2			16	4.53e−01			3	5.38e−01
		3			16	4.62e−01			6	5.02e−01
		4			20	4.34e−01			10	4.77e−01
		5			24	4.53e−01			13	4.69e−01
	1.0e−03	0	18	3.20e−01	1	1.00e+00	6	4.40e−01	1	1.00e+00
		1			4	5.08e−01			2	6.34e−01
		2			18	3.22e−01			7	4.00e−01
		3			29	2.77e−01			15	3.28e−01
		4			45	2.62e−01			17	3.16e−01
		5			46	2.58e−01			17	2.97e−01
512	1.0e−01	0	3	6.36e−01	1	1.00e+00	1	1.00e+00	1	1.00e+00
		1			1	1.00e+00			1	1.00e+00
		2			1	1.00e+00			1	1.00e+00
		3			2	6.13e−01			4	6.83e−01
		4			3	6.35e−01			2	6.37e−01
		5			3	6.38e−01			2	6.37e−01
	1.0e−02	0	10	3.80e−01	1	1.00e+00	2	5.44e−01	1	1.00e+00
		1			1	1.00e+00			1	1.00e+00
		2			3	4.05e−01			2	7.44e−01
		3			9	3.70e−01			3	4.13e−01
		4			16	3.69e−01			6	3.93e−01
		5			18	3.66e−01			8	3.83e−01
	1.0e−03	0	19	2.19e−01	1	1.00e+00	11	2.91e−01	1	1.00e+00
		1			1	1.00e+00			1	1.00e+00
		2			6	3.03e−01			3	3.08e−01
		3			24	2.29e−01			17	2.24e−01
		4			31	2.13e−01			22	2.17e−01
		5			37	2.09e−01			24	2.00e−01

(Continues)

TABLE 1 (Continued)

			CGLS		PCGLS		CG		PCG	
n	α	i	$k_{m.e.}$	ε	$k_{m.e.}$	ε	$k_{m.e.}$	ε	$k_{m.e.}$	ε
1,024	1.0e−01	0	2	5.18e−01	1	1.00e+00	1	1.00e+00	1	1.00e+00
		1			1	1.00e+00			1	1.00e+00
		2			1	1.00e+00			1	1.00e+00
		3			2	6.07e−01			1	1.00e+00
		4			2	4.71e−01			2	7.04e−01
		5			2	4.91e−01			2	5.54e−01
	1.0e−02	0	5	1.93e−01	1	1.00e+00	2	4.34e−01	1	1.00e+00
		1			1	1.00e+00			1	1.00e+00
		2			2	6.20e−01			1	1.00e+00
		3			3	1.89e−01			2	2.41e−01
		4			6	1.82e−01			3	2.09e−01
		5			10	1.81e−01			4	2.02e−01
	1.0e−03	0	13	9.92e−02	1	1.00e+00	3	1.71e−01	1	1.00e+00
		1			1	1.00e+00			1	1.00e+00
		2			2	1.49e−01			2	2.52e−01
		3			10	1.09e−01			4	1.24e−01
		4			27	1.00e−01			6	1.08e−01
		5			32	9.70e−02			9	1.03e−01
2,048	1.0e−01	0	2	5.98e−01	1	1.00e+00	1	1.00e+00	1	1.00e+00
		1			1	1.00e+00			1	1.00e+00
		2			1	1.00e+00			1	1.00e+00
		3			1	1.00e+00			1	1.00e+00
		4			2	6.87e−01			1	1.00e+00
		5			2	5.68e−01			2	7.53e−01
	1.0e−02	0	3	9.50e−02	1	1.00e+00	2	5.92e−01	1	1.00e+00
		1			1	1.00e+00			1	1.00e+00
		2			1	1.00e+00			1	1.00e+00
		3			2	1.24e−01			2	3.62e−01
		4			2	7.14e−02			2	1.19e−01
		5			3	8.01e−02			2	8.69e−02
	1.0e−03	0	5	1.92e−02	1	1.00e+00	2	8.71e−02	1	1.00e+00
		1			1	1.00e+00			1	1.00e+00
		2			2	1.87e−01			2	5.30e−01
		3			2	1.75e−02			2	3.70e−02
		4			5	1.46e−02			4	1.80e−02
		5			7	1.39e−02			5	1.92e−02

Note. PCG = preconditioned conjugate gradient; PCGLS = preconditioned conjugate gradient and least squares; CGLS = conjugate gradient least square; CG = conjugate gradient.

The right-hand side vector $\mathbf{g}$ is the sum of $\overline{\mathbf{g}} = A\mathbf{f}$ and the noise component $\boldsymbol{\eta}$, that is, $\mathbf{g} = \overline{\mathbf{g}} + \boldsymbol{\eta}$, where $\boldsymbol{\eta}$ comes from a normal distribution with zero mean and deviation $\alpha\|\overline{\mathbf{g}}\|$. The matrix A is the symmetric real Toeplitz matrix

$$A_{r,s} = a_{r-s} = \begin{cases} \frac{4}{51} k_{0.15}(x_r - x_s) \text{ if } |r-s| \le b, & b \le n, \\ 0, & \text{otherwise}, \end{cases} \tag{17}$$

obtained from the discretization of a smooth Kernel from a Fredholm integral equation of the first kind; see the work of Eldén.[27] We recall that, as outlined in the work of Di Benedetto et al.,[10] the bandwidth amplitude b in (17) regulates the ill conditioning of the blurring operator; the wider the band, the higher the condition number. The results for this test case are collected in Table 1 when $b = 30$. The maximum number of iteration is set to 100; thus, according to the implementation in RestoreTools,[26] we obtain at most `maxit`+1 approximations counting the initial guess.

As Table 1 confirms, the introduction of the proposed preconditioners enhances the well-known regularization properties of *conjugate-like iterations*; indeed, for both PCG and preconditioned conjugate gradient and least squares (PCGLS),

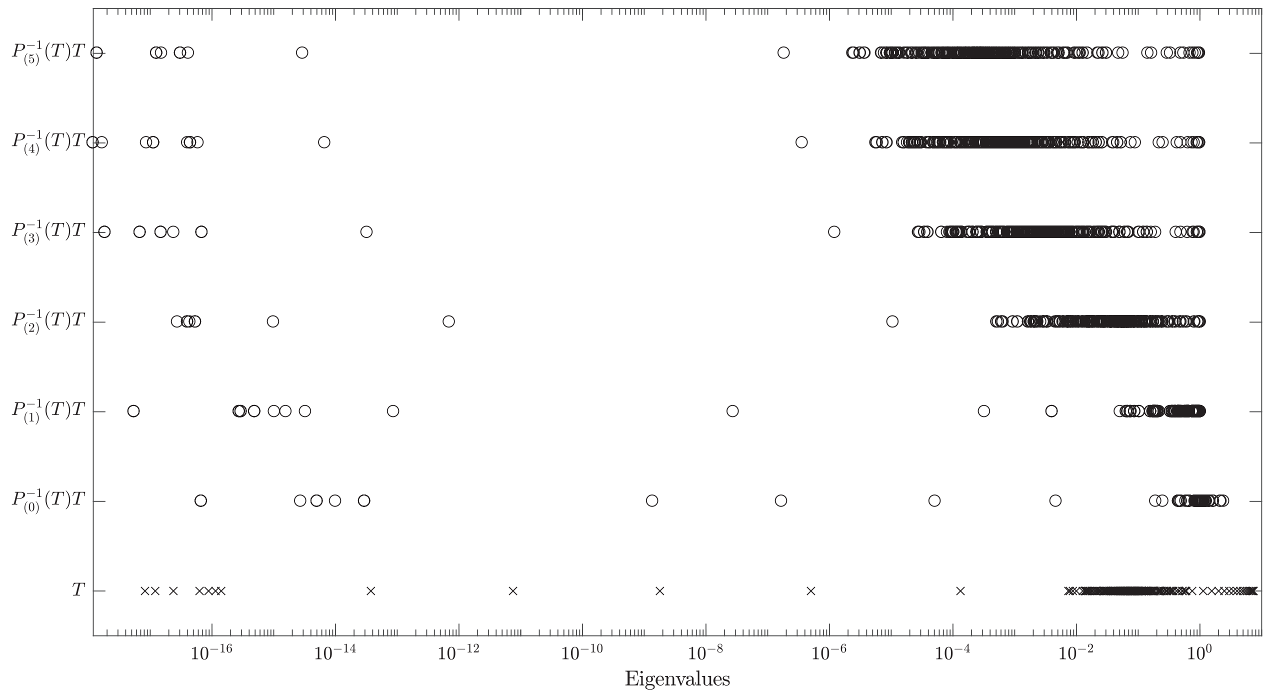

FIGURE 1 Experiment 1. Eigenvalues of the unpreconditioned and preconditioned matrices

they could slow down the number of iterations needed to obtain the best RRE solution for increasing value of i; at the same time, we observe that when this occurs, the quality of the reconstructed signal is improved when compared with the nonpreconditioned case. For the particular case of PCG, we observe that when $\alpha = 10^{-1}$ (i.e., in presence of high noise levels), a meaningful reconstruction can be achieved only choosing a sufficiently large i for $P_{(i)}(A)$. As a further confirmation of the theoretical analysis, in Figure 1, we depict both the original and the preconditioned spectra for several values of i; the plots show clearly the monotonic decreasing predicted in Equation (6).

5.2 | Experiment 2

Let us consider now the case of a two-dimensional signal, for example, an image. In this experiment, we use the classic satellite image as $\mathbf{f}$ (see Figure 2a).

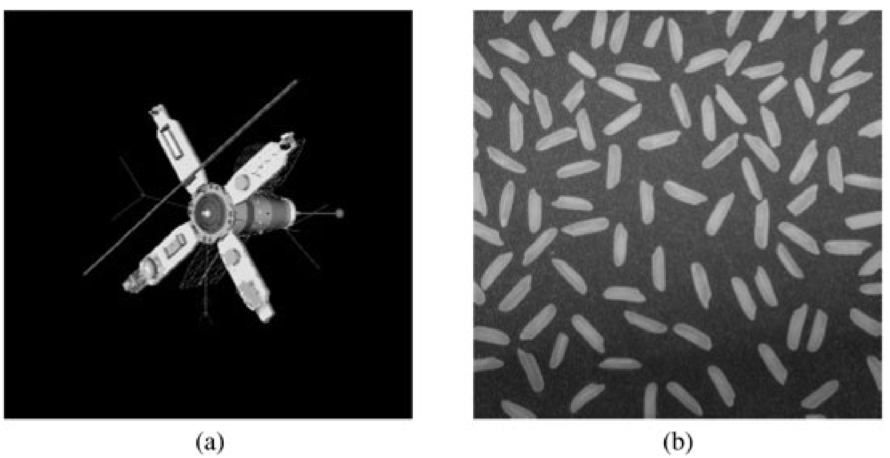

FIGURE 2 Test images for Experiments 1 and 2. (a) Satellite image (Experiment 2). (b) Rice image (Experiment 3)

TABLE 2 Experiment 2. Numerical results for the reconstruction of the image in Figure 2a using the PCG algorithm preconditioned by $P_{(i)}(A)$ α are reported

(a) $b = 15$											
α	1.0e−2			1.0e−3			1.0e−4			1.0e−5	
i	$k_{m.e.}$	ε		$k_{m.e.}$	ε		$k_{m.e.}$	ε	i	$k_{m.e.}$	ε
CG	27	3.9472e−01	–	162	2.6248e−01	–	199	2.3313e−01	–	196	2.3375e−01
0	1	1.4027e+01	0	38	4.7603e+00	0	51	9.4918e−01	0	51	5.3902e−01
1	3	5.1717e+00	1	3	6.8361e−01	1	18	2.4195e−01	1	97	1.2568e−01
2	8	5.2065e−01	2	20	3.4986e−01	2	148	2.0211e−01	2	150	1.9402e−01
3	35	3.9575e−01	3	197	3.0047e−01	3	197	2.9413e−01	3	197	2.9702e−01
CGLS	93	3.5429e−01		201	2.4638e−01		201	2.4372e−01		201	2.4356e−01
(b) $b = 30$											
α	1.0e−2			1.0e−3			1.0e−4			1.0e−5	
i	$k_{m.e.}$	ε		$k_{m.e.}$	ε	i	$k_{m.e.}$	ε		$k_{m.e.}$	ε
CG	48	5.1567e−01	–	144	3.3095e−01	–	201	2.8453e−01	–	197	2.8579e−01
0	1	4.7565e+01	0	1	4.7565e+01	0	2	1.4818e+01	0	2	1.2249e+01
1	4	1.1243e+01	1	12	2.3054e+00	1	41	5.4332e−01	1	100	2.2965e−01
2	26	8.9217e−01	2	79	3.9355e−01	2	146	2.9436e−01	2	146	2.9138e−01
3	120	6.1670e−01	3	201	3.9353e−01	3	185	3.9266e−01	3	201	3.8743e−01
CGLS	106	6.0539e−01		201	3.0084e−01		201	2.9504e−01		201	2.9555e−01

Note. Results for different band amplitudes b and level of noise. PCG = preconditioned conjugate gradient; CG = conjugate gradient; CGLS = conjugate gradient least square.

We build the two-level blur operator T as $T = A \otimes A$ with A from (17) and different band amplitudes b. As observed in the work of Di Benedetto et al.,[10] this case represents a model of a bounded point spread function (PSF), that is, we are dealing with the regularization of an out-of-focus image. We select the Block–Circulant with Circulant Blocks algebra $\mathcal{L} = \mathrm{sd}(F \otimes F)$, with F the Fourier matrix of suitable size. The results for this test case are collected in Table 2. To ensure a fair comparison, we fix the number of both CGLS and CG iteration to 200; thus, according to the implementation in RestoreTools,[26] we obtain at most `maxit+1` approximations counting the initial guess.

We observe that, in these settings, the standard CG usually recovers better reconstructions than the CGLS algorithm. Moreover, in the behavior of PCG when preconditioned with $P_{(i)}(A)$, we retrace the analysis obtained in Theorem 2; in particular, for higher level of noise ($\alpha = 10^{-2}, 10^{-3}$), considering $i > 1$ for the preconditioner $P_{(i)}(A)$, it is possible to obtain RREs of comparable order to the best RRE (nonpreconditioned case), but obtained with fewer iterations. An identical behavior is observed also with the PCGLS algorithm, but because it achieves worse reconstruction error than the PCG, we have reported in Table 2 only the data of the CGLS (see Figure 3).

Finally, we want to stress that, using higher values of i in $P_{(i)}(A)$, we succeed in having a satisfactory reconstruction also at noise levels for which the employment of both the optimal $P_{(0)}(A)$ and the superoptimal $P_{(1)}(A)$ preconditioners breaks down when used in a regularization framework.

5.3 | Experiment 3

In the final experiment, we use as $\mathbf{f}$, the image in Figure 2b. Here, we consider a regularization problem obtained by convolving an image $\mathbf{f}$ with a Gaussian PSF with standard deviation $\sigma = 5.5$ and centered in the middle of the image. To complete the construction of the operator A, *reflexive* boundary conditions have been selected.[28,29] The maximum number of iterations for this experiment is set to $(i + 1)50$, where i is again the order of the preconditioner $P_{(i)}(A)$; thus, according to the implementation in RestoreTools,[26] we obtain at most `maxit+1` approximations counting the initial guess.

Observe that, with this choice for the boundary conditions, the resulting structured matrix A is indeed the sum of matrices with structures Block–Toeplitz with Toeplitz–Blocks and Block–Hankel with Hankel–Blocks. Thus, it is diagonalizable in the algebra $\mathcal{L} = \mathrm{sd}\,C$, being C the orthogonal matrix associated with the trigonometric transform DCT–III; see the works of Chan et al.[29] and Serra-Capizzano.[30] In this case, instead of directly solving the linear system (prone to severe ill conditioning), we build the preconditioner for the iterative solver by using a Kronecker approximation of A as $A \approx T \otimes T$ with T being a symmetric Toeplitz matrix, that is, we use the preconditioner $P_{(i)}(T \otimes T)$. The results obtained in this case, collected in Table 3 and Figure 4, are completely analogous to the results obtained in Experiment 2.

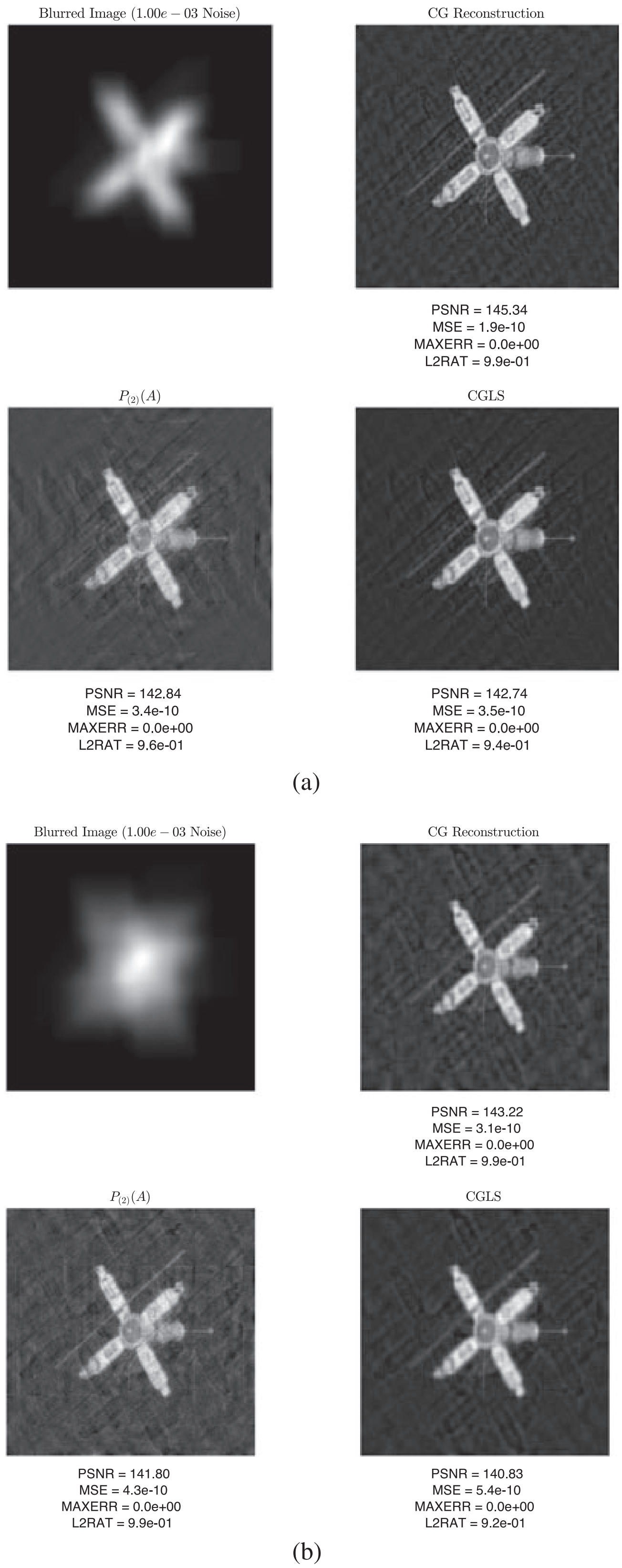

FIGURE 3 Experiment 2. Reconstructed solutions with conjugate gradient (CG), conjugate gradient least square (CGLS), and preconditioned conjugate gradient (PCG), where $P_{(i)}$ is such that the same order of best relative restoration error in CG and CGLS is achieved

TABLE 3 Experiment 3. Numerical results for the reconstruction of the image in Figure 2b using the PCGLS and PCG algorithms preconditioned by $P_{(i)}(A)$

α	1.0e−1			1.0e−2			1.0e−3			1.0e−4	
i	$k_{m.e.}$	ε		$k_{m.e.}$	ε		$k_{m.e.}$	ε		$k_{m.e.}$	ε
CG	1	2.4819e−01	–	3	1.7395e−01	–	8	1.3197e−01	–	22	1.1380e−01
0	1	2.4819e−01	0	1	2.2880e−01	0	1	2.2858e−01	0	1	2.2858e−01
1	1	2.4819e−01	1	2	1.6280e−01	1	6	1.2431e−01	1	15	1.1177e−01
2	2	2.0664e−01	2	7	1.3832e−01	2	26	1.1808e−01	2	85	1.0885e−01
3	4	2.0395e−01	3	17	1.3574e−01	3	66	1.1730e−01	3	201	1.0908e−01
4	5	2.0237e−01	4	28	1.3528e−01	4	113	1.1730e−01	4	251	1.1071e−01
5	7	2.0158e−01	5	36	1.3528e−01	5	154	1.1736e−01	5	301	1.1137e−01
CGLS	13	1.9905e−01	–	83	1.3159e−01	–	201	1.1770e−01	–	201	1.1749e−01
0	1	2.4819e−01	0	1	2.2880e−01	0	1	2.2858e−01	0	1	2.2858e−01
1	5	2.3048e−01	1	12	1.4245e−01	1	76	1.2102e−01	1	101	1.1474e−01
2	6	2.0374e−01	2	78	1.3766e−01	2	151	1.2630e−01	2	151	1.2624e−01
3	21	2.0212e−01	3	201	1.3863e−01	3	201	1.3612e−01	3	201	1.3610e−01
4	51	2.0106e−01	4	251	1.4483e−01	4	251	1.4372e−01	4	251	1.4370e−01
5	83	2.0080e−01	5	301	1.4755e−01	5	301	1.4657e−01	5	301	1.4655e−01

Note. Results for different level of noise α are reported. PCGLS = preconditioned conjugate gradient and least squares; PCG = preconditioned conjugate gradient; CG = conjugate gradient; CGLS = conjugate gradient least square.

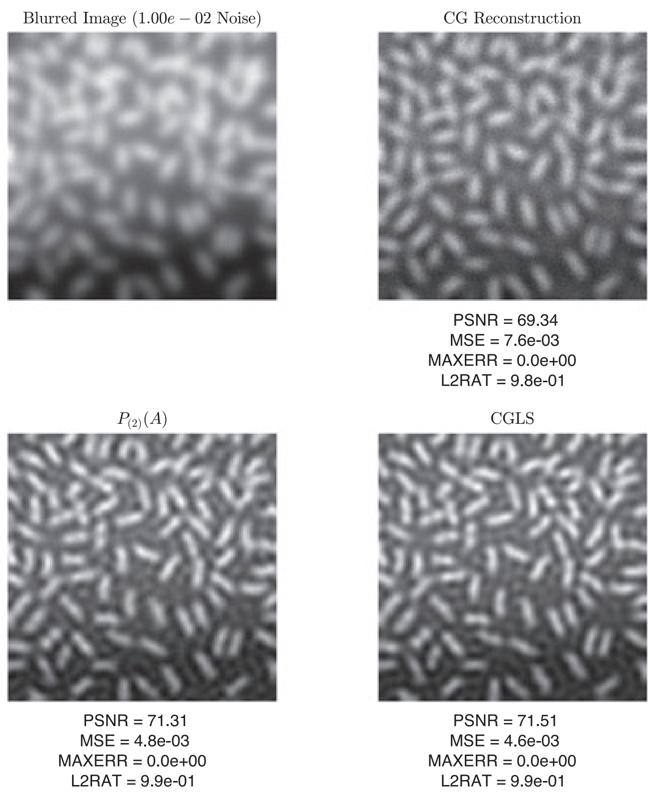

FIGURE 4 Experiment 3. Reconstructed solutions with conjugate gradient (CG), conjugate gradient least square (CGLS), and preconditioned conjugate gradient (PCG), where $P_{(i)}$ is the one that reaches an error of the same order with respect to the CGLS and a better reconstruction with respect to the CG algorithm

6 | CONCLUSIONS

In this paper, we have introduced a new class of preconditioners including the optimal and superoptimal preconditioners. We have carried on a detailed analysis and rigorously proved the regularizing properties of the new preconditioners. All the results have been confirmed by extensive numerical examples, which show that when the value of i increases for the preconditioner $P_{(i)}(A)$, a higher regularization level and better filtering capabilities for the noise space are obtained.

The class of the regularizing preconditioners proposed in this paper could be suitably employed even when critical conditions are registered, that is, high noise level or excessive ill conditioning. Indeed, even in these unfavorable conditions, satisfactory reconstruction performances can be obtained.

ACKNOWLEDGEMENTS

The authors would like to thank two anonymous referees for their constructive comments that have improved both the readability and the quality of this paper. The work of Stefano Cipolla and Carmine Di Fiore was partially supported by Gruppo Nazionale per il Calcolo Scientifico dell'Istituto Nazionale di Alta Matematica (INdAM-GNCS). The work of Fabio Durastante was partially supported by the 2018 INdAM-GNCS project "Tecniche innovative per problemi di algebra lineare". The work of Carmine Di Fiore and Paolo Zellini was partially supported by the MIUR Excellence Department Project awarded to the Department of Mathematics, University of Rome Tor Vergata (CUP E83C18000100006).

ORCID

Stefano Cipolla https://orcid.org/0000-0002-8000-4719
Fabio Durastante https://orcid.org/0000-0002-1412-8289

REFERENCES

1. Chan TF. An optimal circulant preconditioner for Toeplitz systems. SIAM J Sci Statist Comput. 1988;9(4):766–771. Available from: https://doi.org/10.1137/0909051
2. Chan RH. Circulant preconditioners for Hermitian Toeplitz systems. SIAM J Matrix Anal Appl. 1989;10(4):542–550. Available from: https://doi.org/10.1137/0610039
3. Tyrtyshnikov EE. Optimal and superoptimal circulant preconditioners. SIAM J Matrix Anal Appl. 1992;13(2):459–473. Available from: https://doi.org/10.1137/0613030
4. Chan RH, Jin XQ, Yeung MC. The spectra of super-optimal circulant preconditioned Toeplitz systems. SIAM J Numer Anal. 1991;28(3):871–879. Available from: https://doi.org/10.1137/0728046
5. Di Benedetto F, Capizzano SS. A note on the superoptimal matrix algebra operators. Linear Multilinear Algebra. 2002;50(4):343–372. Available from: https://doi.org/10.1080/03081080210000049301
6. Chan RHF, Jin XQ. An introduction to iterative Toeplitz solvers. Philadelphia, PA: Society for Industrial and Applied Mathematics; 2007. Fundamentals of algorithms, vol. 5. Available from: https://doi.org/10.1137/1.9780898718850
7. Bertaccini D, Durastante F. Iterative methods and preconditioning for large and sparse linear systems with applications. In: Monographs and research notes in mathematics. Boca Raton, FL: CRC Press; 2018.
8. Hansen PC, Nagy JG, O'Leary DP. Deblurring images: Matrices, spectra, and filtering. Philadelphia, PA: Society for Industrial and Applied Mathematics; 2006. Fundamentals of algorithms, vol. 3. Available from: https://doi.org/10.1137/1.9780898718874
9. Hanke M, Nagy JG, Plemmons RJ. Preconditioned iterative regularization. Numer Linear Algebra. 1993:141–163.
10. Di Benedetto F, Estatico C, Capizzano SS. Superoptimal preconditioned conjugate gradient iteration for image deblurring. SIAM J Sci Comput. 2005;26(3):1012–1035. Available from: https://doi.org/10.1137/S1064827503421653
11. Dykes L, Noschese S, Reichel L. Circulant preconditioners for discrete ill-posed Toeplitz systems. Numer Algorithms. 2017;75(2):477–490. Available from: https://doi.org/10.1007/s11075-016-0205-9
12. Estatico C. A class of filtering superoptimal preconditioners for highly ill-conditioned linear systems. BIT Numer Math. 2002;42(4):753–778. Available from: https://doi.org/10.1023/A:1021948319714
13. Cheng CM, Jin XQ, Vong SW, Wang W. A note on spectra of optimal and superoptimal preconditioned matrices. Linear Algebra Appl. 2007;422(2–3):482–485. Available from: https://doi.org/10.1016/j.laa.2006.11.006
14. Horn RA, Johnson CR, Matrix analysis. 2nd ed. Cambridge, UK: Cambridge University Press; 2013.
15. Kressner D, Luce R. Fast computation of the matrix exponential for a Toeplitz matrix. SIAM J Matrix Anal Appl. 2018;39(1):23–47. Available from: https://doi.org/10.1137/16M1083633
16. Chan RH, Ng MK, Wong CK. Sine transform based preconditioners for symmetric Toeplitz systems. Linear Algebra Appl. 1996;232:237–259. Available from: https://doi.org/10.1016/0024-3795(94)00049-2
17. Chan RH, Ching WK, Wong CK. Optimal trigonometric preconditioners for elliptic and queueing problems. Southeast Asian Bull Math. 1996;20(3):117–124. Conference on Scientific Computation'94 (Shatin, Hong Kong, 1994).
18. Bini D, Favati P. On a matrix algebra related to the discrete Hartley transform. SIAM J Matrix Anal Appl. 1993;14(2):500–507. Available from: https://doi.org/10.1137/0614035
19. Bozzo E, Di Fiore C. On the use of certain matrix algebras associated with discrete trigonometric transforms in matrix displacement decomposition. SIAM J Matrix Anal Appl. 1995;16(1):312–326. Available from: https://doi.org/10.1137/S0895479893245103

20. Di Fiore C, Zellini P. Matrix algebras in optimal preconditioning. Linear Algebra Appl. 2001;335:1–54. Available from: https://doi.org/10.1016/S0024-3795(00)00137-3
21. Di Fiore C, Zellini P. Matrix decompositions using displacement rank and classes of commutative matrix algebras. Linear Algebra Appl. 1995;229:49–99. Available from: https://doi.org/10.1016/0024-3795(93)00347-3
22. Di Benedetto F, Capizzano SS. Optimal multilevel matrix algebra operators. Linear Multilinear Algebra. 2000;48(1):35–66. Available from: https://doi.org/10.1080/03081080008818658
23. Bortoletti A, Di Fiore C. On a set of matrix algebras related to discrete Hartley-type transforms. Linear Algebra Appl. 2003;366:65–85. Available from: https://doi.org/10.1016/S0024-3795(02)00500-1. Special Issue on Structured Matrices: Analysis Algorithms and Applications (Cortona, Italy, 2000).
24. Gazzola S, Novati P, Russo MR. On Krylov projection methods and Tikhonov regularization. Electron Trans Numer Anal. 2015;44:83–123.
25. Morozov V. On the solution of functional equations by the method of regularization. Soviet Math Dokl. 1966;7:414–417.
26. Nagy J, Berisha S, Chung J, Palmer K, Perrone L, Wright W. RestoreTools: An object Oriented Matlab Package for Image Restoration. 2012. Available from: http://www.mathcs.emory.edu/~nagy/RestoreTools/
27. Eldén L. An algorithm for the regularization of ill-conditioned, banded least squares problems. SIAM J Sci Statist Comput. 1984;5(1):237–254. Available from: https://doi.org/10.1137/0905017
28. Ng MK, Chan RH, Tang WC. A fast algorithm for deblurring models with Neumann boundary conditions. SIAM J Sci Comput. 1999;21(3):851–866. Available from: https://doi.org/10.1137/S1064827598341384
29. Chan TF, Chan R, Wong CK. Cosine transform based preconditioners for total variation minimization problems in image processing. In: Proceedings of the IMACS Conference on Iterative Methods; 1995; Blagoevgrad, Bulgaria. Piscataway, NJ: IMACS Secretariat, Department of Computer Science, Rutgers University; p. 311–329. IMACS Series in Computational and Applied Mathematics.
30. Serra-Capizzano S. A note on antireflective boundary conditions and fast deblurring models. SIAM J Sci Comput. 2003;25(4):1307–1325.

How to cite this article: Cipolla S, Di Fiore C, Durastante F, Zellini P. Regularizing properties of a class of matrices including the optimal and the superoptimal preconditioners. *Numer Linear Algebra Appl.* 2018;e2225. https://doi.org/10.1002/nla.2225

APPENDIX

Given A positive definite, for any $\mathcal{L} = \mathrm{sd}\, U$, we have that $\kappa_2(\mathcal{L}_A) \le \kappa_2(A)$. In this Appendix, for a real positive definite matrix A, we show that, performing no more than $O(n^2)$ flops, one can define a matrix algebra $\mathcal{L} = \mathrm{sd}\, U$ such that $\kappa_2(\mathcal{L}_A)$ is as small as desired. It can be easily proved that, given

$$A = \begin{bmatrix} a & c \\ c & b \end{bmatrix}, \quad a, b, c \in \mathbb{R}, \text{ with } a < b, \text{ and any } z \in (a, b),$$

there exist $\alpha, \beta \in \mathbb{R}$ such that $\alpha^2 + \beta^2 = 1$ and

$$\begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \begin{bmatrix} a & c \\ c & b \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} = \begin{bmatrix} a_z & c_z \\ c_z & b_z \end{bmatrix}, \tag{A1}$$

being $|c_z| > |c|$ and a_z, b_z such that

$$a < a_z \le b_z < b, \; a_z - a = b - b_z,$$

$$a_z = z \text{ if } z \in \left(a, \frac{a+b}{2}\right] \text{ or } b_z = z \text{ if } z \in \left[\frac{a+b}{2}, b\right);$$

note that $a_{a+t} = a_{b-t}, b_{a+t} = b_{b-t}, c_{a+t} = c_{b-t}, \forall t \in (0, b-a)$. An analogous result can be stated if $a > b$.

Thus, if $A = \begin{bmatrix} a & c \\ c & b \end{bmatrix}$, $a, b, c \in \mathbb{R}$, $a \ne b$, then, by a Givens similarity transformation, that is, a transformation as in (A1), we can cluster the diagonal entries a and b of A as much as we want, maintaining their order, or, equivalently, given any $z \in (\min\{a, b\}, \max\{a, b\})$, we can cluster the diagonal entries a and b of A to equate to z the one of them nearer to z.

Let A be an arbitrary $n \times n$ symmetric matrix with real entries. Assume that $[A]_{ii} \ne [A]_{jj}$ for some $i \ne j$ (i.e., we suppose that the equalities $[A]_{11} = [A]_{22} = \dots = [A]_{nn}$ are not all verified).

Set $A_0 = A$ and, for $k = 0, 1, \dots$, define the $n \times n$ matrix A_{k+1} from the $n \times n$ matrix A_k as follows.

1. Choose i, j such that $[A_k]_{ii} < \frac{\mathrm{tr}(A)}{n} < [A_k]_{jj}$.

2. If $\frac{\mathrm{tr}(A)}{n} - [A_k]_{ii} \leq [A_k]_{jj} - \frac{\mathrm{tr}(A)}{n}$, then introduce the $n \times n$ Givens matrix G_{k+1} satisfying

$$\left[G_{k+1}^T A_k G_{k+1}\right]_{ii} = [A_k]_{ii} + \left(\frac{\mathrm{tr}(A)}{n} - [A_k]_{ii}\right) \text{ and } \left[G_{k+1}^T A_k G_{k+1}\right]_{jj} = [A_k]_{jj} - \left(\frac{\mathrm{tr}(A)}{n} - [A_k]_{ii}\right);$$

otherwise, introduce the $n \times n$ Givens matrix G_{k+1} such that

$$\left[G_{k+1}^T A_k G_{k+1}\right]_{jj} = [A_k]_{jj} - \left([A_k]_{jj} - \frac{\mathrm{tr}(A)}{n}\right) \text{ and } \left[G_{k+1}^T A_k G_{k+1}\right]_{ii} = [A_k]_{ii} + \left([A_k]_{jj} - \frac{\mathrm{tr}(A)}{n}\right).$$

After $\hat{k}$ steps (with $\hat{k} \leq n-1-d$, being $d \geq 0$ the number of diagonal entries of A equal to $\frac{\mathrm{tr}(A)}{n}$), this procedure yields a matrix

$$G_{\hat{k}}^T \cdots G_2^T G_1^T A G_1 G_2 \cdots G_{\hat{k}} = \begin{bmatrix} \frac{\mathrm{tr}(A)}{n} & * & \cdot & * \\ * & \frac{\mathrm{tr}(A)}{n} & \cdot & \cdot \\ \cdot & \cdot & \ddots & * \\ * & \cdot & * & \frac{\mathrm{tr}(A)}{n} \end{bmatrix}$$

with diagonal entries all equal to $\frac{\mathrm{tr}(A)}{n}$.

If $U_k = G_1 G_2 \cdots G_k$ and $\mathcal{L}_k = \mathrm{sd}\, U_k$, then the eigenvalues of $(\mathcal{L}_k)_A = U_k \mathrm{diag}([U_k^T A U_k]_{ss}) U_k^T$ satisfy the following inequalities:

$$\min_s \left[U_k^T A U_k\right]_{ss} \leq \min_s \left[U_{k+1}^T A U_{k+1}\right]_{ss}, \; \max_s \left[U_{k+1}^T A U_{k+1}\right]_{ss} \leq \max_s \left[U_k^T A U_k\right]_{ss},$$

$$l_{k+1} \leq l_k := \max_s \left[U_k^T A U_k\right]_{ss} - \min_s \left[U_k^T A U_k\right]_{ss},$$

$$l_{\hat{k}} = 0, \quad (\mathcal{L}_{\hat{k}})_A = \frac{\mathrm{tr}(A)}{n} I.$$

In particular, if A is positive definite, we have

$$\kappa_2\left((\mathcal{L}_{\hat{k}})_A\right) = 1 \leq \kappa_2\left((\mathcal{L}_{k+1})_A\right) \leq \kappa_2\left((\mathcal{L}_k)_A\right) \leq \kappa_2\left((\mathcal{L}_0)_A\right) \leq \kappa_2(A),$$

where $\mathcal{L}_0 = \mathrm{sd}\, I$.

Of course, if in step (1) we require that $[A_k]_{ii}$ is the smallest among the $[A_k]_{ss}$ smaller than $\frac{\mathrm{tr}(A)}{n}$ and $[A_k]_{jj}$ is the greatest among the $[A_k]_{ss}$ greater than $\frac{\mathrm{tr}(A)}{n}$, then

1. the sequence l_k decreases as fast as possible, and
2. if A is positive definite, the sequences l_k and $\kappa_2((\mathcal{L}_k)_A)$ decrease as fast as possible.