Electronic Transactions on Numerical Analysis.
Volume 59, pp. 157–178, 2023.

ISSN 1068–9613.
DOI: 10.1553/etna_vol59s157

ETNA
Kent State University and
Johann Radon Institute (RICAM)

SYMMETRIZATION TECHNIQUES IN IMAGE DEBLURRING*

MARCO DONATELLI†, PAOLA FERRARI†, AND SILVIA GAZZOLA‡

Dedicated to Lothar Reichel on his seventieth birthday.

Abstract. This paper presents some preconditioning techniques that enhance the performance of iterative regularization methods applied to image deblurring problems determined by a wide variety of point spread functions (PSFs) and boundary conditions. We first consider the anti-identity preconditioner, which symmetrizes the coefficient matrix associated to problems with zero boundary conditions, allowing the use of MINRES as a regularization method. When considering more sophisticated boundary conditions and strongly nonsymmetric PSFs, we show that the anti-identity preconditioner improves the performance of GMRES. We then consider both stationary and iteration-dependent regularizing circulant preconditioners that, applied in connection with the anti-identity matrix and both standard and flexible Krylov subspaces, speed up the iterations. A theoretical result about the clustering of the eigenvalues of the preconditioned matrices is proved in a special case. Extensive numerical experiments show the effectiveness of the new preconditioning techniques, including when the deblurring of sparse images is considered.

Key words. inverse problems, regularization, preconditioning, Toeplitz matrices, Krylov subspace methods

AMS subject classifications. 65F08, 65F10, 65F22

1. Introduction.

We consider the restoration of blurred and noise-corrupted images in two-dimensional space. We assume that the point spread function (PSF) is spatially-invariant, therefore, the blurring phenomenon is modeled by a convolution of the form

$$b(\boldsymbol{s}) = [\mathcal{K}f](\boldsymbol{s}) + \xi(\boldsymbol{s}) = \int_{\mathbb{R}^2} h(\boldsymbol{s}-\boldsymbol{t})x(\boldsymbol{t})\mathrm{d}\boldsymbol{t} + \xi(\boldsymbol{s}), \quad \boldsymbol{s} \in \Omega \subset \mathbb{R}^2, \tag{1.1}$$

where b represents the observed (blurred and noisy) image, x the (unknown) exact image, h the PSF with compact support, and ξ the random noise. The real-valued nonnegative functions x and b determine the light intensity of the desired and available images, respectively. We assume that the PSF h, and thus the blurring phenomenon, is known.

Discretization of the above integral equation at equidistant nodes yields

$$B_{i,j} = \sum_{(k,\ell)\in\mathbb{Z}^2} h_{i-k,j-\ell}X_{k,\ell} + \Xi_{i,j}, \quad (i,j) \in \mathbb{Z}^2, \tag{1.2}$$

where the entries of the discrete images $B = [B_{i,j}]_{(i,j)\in\mathbb{Z}^2}$ and $X = [X_{k,\ell}]_{(k,\ell)\in\mathbb{Z}^2}$ represent the light intensity at each pixel and $\Xi = [\Xi_{i,j}]_{(i,j)\in\mathbb{Z}^2}$ models the noise-contamination at these pixels. Moreover, the observed image is available only in a finite region, the field of view (FOV), corresponding to $(i,j),(k,\ell) \in [1,n]^2$, which is assumed to be square only for notational simplicity. Therefore, when there are nonvanishing coefficients $h_{i,j}$ with $(i,j) \neq (0,0)$, the measured intensities near the boundary are affected by data outside the FOV, depending on the support of the PSF. Thus, the linear system of equations defined by (1.2) is underdetermined, since there are n^2 constraints, while the number of unknowns required to specify the equations is larger. A meaningful solution of this underdetermined system can be determined in several ways; see [1, 5]. In this paper we adopt a stack-ordered representation of

*Received December 12, 2022. Accepted September 5, 2023. Published online on October 5, 2023. Recommended by G. Rodriguez.

†Dipartimento di Scienza e Alta Tecnologia, Università dell'Insubria, Como, Italy ({paola.ferrari,marco.donatelli}@uninsubria.it).

‡Department of Mathematical Sciences, University of Bath, Bath, UK (S.Gazzola@bath.ac.uk).

ETNA
Kent State University and
Johann Radon Institute (RICAM)

the involved images and, starting from (1.2), by imposing boundary conditions we determine a linear system of equations with a square matrix,

$$A\boldsymbol{x} = \boldsymbol{b}, \quad A \in \mathbb{R}^{n^2 \times n^2}, \quad \boldsymbol{x}, \boldsymbol{b} \in \mathbb{R}^{n^2}, \tag{1.3}$$

where the x_j-values in (1.2) corresponding to pixels outside the FOV are assumed to be certain linear combinations of values inside the FOV; see [11, 26].

Since the singular values of the discrete convolution operator A gradually approach zero without a significant gap, A is ill-conditioned and may be numerically rank-deficient. The degree of ill-posedness depends on the decay of the PSF values: slower decay (i.e., wider PSF) implies more ill-posedness; see equations (2.1), (2.2), and the analysis in Section 2. A linear system of equations (1.3) with a matrix of this kind is commonly referred to as linear discrete ill-posed problem and requires regularization [24].

The structure of the matrix A depends on the boundary conditions. For instance, by using zero boundary conditions we obtain a Block Toeplitz with Toeplitz Blocks (BTTB) matrix, by using periodic boundary conditions we obtain a Block Circulant with Circulant Blocks (BCCB) matrix, while more sophisticated boundary conditions, such as reflective, antireflective, and synthetic, give rise to more complex matrix structures [11, 16]. Regardless of how complicated the structure of the matrix A is, matrix-vector product with A can always be computed in $O(n^2 \log(n))$ flops by padding the vector with the desired boundary conditions and applying the circular convolution by fast Fourier transforms (FFTs). This is the implementation of matrix-vector products in the Matlab toolbox *IR Tools* [18]. Therefore, for the restoration of large images, one generally resorts to iterative methods requiring only matrix-vectors product with A and, possibly, A^T, where the superscript T denotes the transposition, and do not require any factorization of A.

The adjoint of the convolution operator in (1.1) is the correlation operator

$$[\mathcal{K}^* x](\boldsymbol{s}) = \int_{\mathbb{R}^2} h(\boldsymbol{t} - \boldsymbol{s}) x(\boldsymbol{t}) \mathrm{d}\boldsymbol{t}, \tag{1.4}$$

where we have used the fact that h is real-valued. A discretization of (1.4) with the same boundary conditions used for (1.3) can be simply obtained from the PSF rotated by $180°$. The resulting matrix is denoted by A'. Therefore, matrix-vector products with A', i.e., the discretization of the adjoint operator, can be computed by rotating the PSF and then applying the same procedure described above for A. This is the common implementation of the matrix-vector products with the adjoint operator of A when zero or periodic boundary conditions are imposed. Unfortunately, the matrix A' could differ from A^T when the imposed boundary conditions are other than zero or periodic; see [11] for details. In such cases, using solvers like CGLS or LSQR with A^T replaced by A' lacks theoretical justification, which makes it natural to explore the performance of other iterative methods that do not require the adjoint operator. Some recent strategies are based on the preconditioned Arnoldi method and nonstationary iterations [5, 12, 13, 19].

In this paper, we compute a solution of (1.3) through iterative regularization methods, which should terminate when a desired approximation is obtained and before the noise starts to corrupt the solution, causing the restoration error to grow (this is the so-called *semi-convergence* phenomenon). For this reason, a reliable stopping criterion is crucial to compute a good reconstruction. On the one hand, preconditioning is usually applied to speed up the convergence of iterative methods replacing the linear system (1.3) by either

$$PA\boldsymbol{x} = P\boldsymbol{b} \quad \text{or} \quad AP\boldsymbol{z} = \boldsymbol{b},\ \boldsymbol{x} = P\boldsymbol{z}, \tag{1.5}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

where P is the preconditioner that could be applied to the left or right side of the matrix A. If iterations are stopped by the statistically-inspired discrepancy principle, right preconditioning is preferred because it does not modify the noise statistics; see [24, 25]. For discrete ill-posed problems, P must be chosen carefully by avoiding clustering of eigenvalues in the so-called noise subspace that exacerbate semi-convergence, since the signal components in this subspace are usually dominated by noise [22]. On the other hand, when the linear systems (1.5) are solved by a Krylov method, the approximate solution is computed in a different subspace and thus the choice of P affects the quality of the restored image rather than speeding up the convergence, or, possibly it achieves both [3, 9, 22]. In particular, to provide a good restoration, P should symmetrize the operator A and thus $P = A'$ is a favorable choice. This is the so-called reblurring strategy proposed in [11] and later studied in [13] in connection to Arnoldi methods. This approach has been further improved by adding a clustering of the eigenvalues in the signal space to obtain a fast convergence rate [9, 12].

Symmetrization of Toeplitz and BTTB linear systems arising from well-posed problems was recently explored by Pestana and Wathen in [33]. In detail, defining the anti-identity matrix $Y \in \mathbb{R}^{n^2 \times n^2}$ by

$$Y = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix},$$

the matrix YA is symmetric whenever A is persymmetric, i.e., $YA = A^T Y$, as in the case of Toeplitz and BTTB matrices. It follows that the linear system (1.3) can be replaced by the equivalent linear system

$$YA\boldsymbol{x} = Y\boldsymbol{b}, \tag{1.6}$$

which can be solved by methods developed for symmetric indefinite matrices, such as MINRES and MR-II, i.e., a variant of MINRES that is very popular for solving inverse problems. When A is a BTTB matrix, as in the case of zero Dirichlet boundary conditions, preconditioning the linear system (1.6) by BCCB matrices has been proposed and studied independently in [31] and [17, 27] proving the eigenvalue clustering at the two points -1 and 1. However, such a symmetrization strategy has never been explored for discrete ill-posed problems, where the preconditioner has to deal with the noise subspace.

In this paper, motivated by the importance of having an operator close to symmetric to generate the Krylov subspace in which to search for an approximate solution of a discrete ill-posed problem (see, e.g., [29]), we investigate the symmetrization technique (1.6) for image deblurring problems. More specifically, the contributions of this article are twofold. Firstly, we consider zero Dirichlet boundary conditions so that A is a BTTB matrix and we investigate the regularizing properties of MINRES, applied to the linear system (1.6). For the symmetrized linear system, we then define a regularizing preconditioner P for the matrix YA combining the analysis in [17] with the regularizing preconditioner used, e.g., in [9, 12]. We prove that the spectrum of the preconditioned matrix PYA is clustered at the three points $\{-1, 0, 1\}$. Preconditioned MR-II for deblurring astronomical images has been previously investigated in [23], but using a different symmetrization strategy for PSFs close to symmetric than our proposal. Differently from [23], our approach is also effective for strongly nonsymmetric PSFs such as the motion blur considered in the numerical results.

The second contribution of this paper is to heuristically extend such approach to generic boundary conditions and to consider more sophisticated regularization methods, such as projection methods that enforce sparsity in the computed solution [7]; such investigations

ETNA
Kent State University and
Johann Radon Institute (RICAM)

are supported by extensive numerical tests. In this case, since the matrix A might not be persymmetric, YA is not symmetric even though it is close to being symmetric and MINRES is replaced by GMRES [6]. Since the regularizing preconditioner P depends on a parameter, nonstationary preconditioning is explored together with flexible GMRES to avoid the a priori estimation of such parameter, as proposed in [10]. When enforcing some sparsity into the computed solution is appropriate, we adopt efficient algorithms for 1-norm regularization based on iteratively reweighted least squares, which formally handle the inverted weights as iteration-dependent preconditioners that modify the approximation subspace within methods based on the flexible Golub-Kahan decomposition (such as FLSQR [20]) or methods based on the flexible Arnoldi decomposition (such as FGMRES [19]). This approach results in FLSQR and FGMRES methods, respectively, where two iteration-dependent preconditioners (P and the inverted weights) are sequentially applied at each iteration; to the best of our knowledge, a regularizing preconditioner P has never been applied to FLSQR before.

This paper is organized as follows. Section 2 provides some background material on the links between boundary conditions for image deblurring problems and structured matrices appearing in the linear system formulation (1.3), including their associated spectral decompositions. Section 3 describes the circulant preconditioners considered in this paper and the spectral analysis of the preconditioned matrices in the zero boundary condition case. Section 4 summarizes the iterative regularization methods considered in this paper and specifies the strategies adopted to precondition them. In particular, we discuss iteration-dependent circulant preconditioners and consider a diagonal preconditioner for enforcing sparsity in the computed solution. Section 5 displays the results of four different test problems, which show the performance of the new preconditioners applied in different settings. In particular, Sections 5.1 and 5.2 illustrate the efficiency of the symmetrization strategy for zero boundary conditions with BTTB matrices; Section 5.3 demonstrates performance of the approximate symmetrization strategy for reflective boundary conditions; Section 5.4 provides a numerical study of sparse image reconstruction to show the effects of symmetrization and preconditioning within flexible Krylov subspace methods that enforce sparsity into the solution. In the latter section, a discussion on the combination of the circulant and sparsity-enforcing preconditioners is also presented. Section 6 presents some conclusions and outlines some possible extensions to the present work.

2. Boundary conditions and structured matrices. Let $h_{i,j}$ be the entries of the PSF, with $i, j \in \mathbb{Z}$, where $h_{0,0}$ is the designated central entry, e.g., the entry of maximum intensity. Because the PSF has compact support, $h_{i,j} = 0$ for $|i|$ or $|j|$ large enough, in particular when $h_{i,j}$ is outside the FOV, i.e., when $\min\{|i|, |j|\} \geq n$, since it depends on the position of the central coefficient $h_{0,0}$, which, in the worst case, may be at a corner of the FOV. Given the coefficients $h_{i,j}$, it is possible to associate the PSF with the so-called generating function $f : [-\pi, \pi]^2 \to \mathbb{C}$ defined by

$$f(\vartheta_1, \vartheta_2) = \sum_{i,j=-n+1}^{n-1} h_{i,j} \mathrm{e}^{\hat{\imath}(i\vartheta_1 + j\vartheta_2)}, \quad \hat{\imath}^2 = -1. \tag{2.1}$$

Note that $h_{i,j}$ are the Fourier coefficients of the function f.

The structure of the matrix $A \in \mathbb{R}^{n^2 \times n^2}$ in (1.3) depends on the coefficients $h_{i,j}$ and the imposed boundary conditions. As we already stated in the introduction, we represent the images as vectors in $\mathbb{R}^{n^2}$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

When the exact image has a black background like, e.g., in astronomical imaging, zero boundary conditions are to be preferred. In this case

$$A = \begin{bmatrix} T_0 & T_{-1} & \cdots & T_{-n+1} \\ T_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & T_{-1} \\ T_{n-1} & \cdots & T_1 & T_0 \end{bmatrix}_{n^2 \times n^2},$$

with

$$T_k = \begin{bmatrix} h_{k,0} & h_{k,-1} & \cdots & h_{k,-n+1} \\ h_{k,1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & h_{k,-1} \\ h_{k,n-1} & \cdots & h_{k,1} & h_{k,0} \end{bmatrix}_{n \times n}, \quad k = -n+1, \ldots, n-1.$$

In this case A is a BTTB matrix and we use the notation

$$A = \mathcal{T}_n(f),$$

where f is the generating function defined in (2.1). Note that $\{\mathcal{T}_n(f)\}_n$ denotes the sequence of BTTB matrices generated by f as n varies. If the PSF is not quadrantally symmetric, i.e., symmetric in both horizontal and vertical directions, then A is not symmetric. On the other hand, it is always persymmetric independently of the PSF.

By imposing boundary conditions different from the zero Dirichlet ones, the matrix A is no longer BTTB, but small norm and/or small rank corrections are added to $\mathcal{T}_n(f)$ depending on the support and the decay rate of the PSF. Among the various kinds of boundary conditions, the periodic ones are computationally attractive, since the resulting matrix A is BCCB and can be diagonalized by FFTs [26]. The BCCB matrix A associated to a PSF, and thus its generating function f defined in (2.1), will be denoted by $A = \mathcal{C}_n(f)$. The main property of such matrix is its spectral decomposition in terms of the Fourier matrix. Let F_1 be the discrete Fourier matrix defined as $[F_1]_{i,j} = \frac{1}{\sqrt{n}} \mathrm{e}^{-\hat{\imath}\frac{2\pi ij}{n}}$, for $i, j = 0, \ldots, n-1$. The two-dimensional Fourier matrix is defined by the tensor product $F_2 = F_1 \otimes F_1$ and matrix-vector products with F_2 can be computed in $O(n^2 \log(n))$ using the FFT algorithm. The eigenvalues λ_j, with $j = 1, \ldots, n^2$, of $\mathcal{C}_n(f)$ can be computed by applying F_2 to the first column of $\mathcal{C}_n(f)$, which is obtained by stacking a proper permutation of the PSF; see [26] for details. Therefore, the matrix A can be factorized as

$$A = \mathcal{C}_n(f) = F_2^H \Lambda F_2,$$

where Λ is the diagonal matrix of the eigenvalues λ_j, with $j = 1, \ldots, n^2$. Note that $\mathcal{C}_n(f)$ is the Strang preconditioner of $\mathcal{T}_n(f)$ and the n^2 eigenvalues of $\mathcal{C}_n(f)$ can also be written as

$$\lambda_{i+jn+1} = f\left(\frac{2\pi i}{n}, \frac{2\pi j}{n}\right), \quad i, j = 0, \ldots, n-1, \tag{2.2}$$

where f is the generating function (2.1); see [8].

The inverse problem (1.1) is well-known to be ill-posed and the associated linear system (1.3) is ill-conditioned; see [14, 24]. In particular, one can clearly see from equations (2.1) and (2.2) that, in the case of periodic boundary conditions, the ill-conditioning of problem (1.3)

ETNA
Kent State University and
Johann Radon Institute (RICAM)

is related to the decay of the PSF values: slower decay, i.e., more h_{ij} coefficients different from zero, implies a more oscillating generating function, which is reflected in the associated circulant matrix having a few dominant singular modes that capture the most significant variations, while other modes are suppressed due to the oscillatory behavior. As a result, singular values associated with these suppressed modes decay rapidly. Therefore, the matrix is severely ill-conditioned and (1.3) is very sensitive to noise. Because of the ill-conditioning regularization should be applied when computing an approximation of the original image. For instance, Tikhonov regularization defines a nearby well-posed problem that can potentially be solved analytically. Similarly, when approximately solving (1.3) by means of an iterative regularization method, the preconditioner P has to be well-posed and, hence, a strategy to employ BCCB matrices as regularizing preconditioners is linked to Tikhonov regularization.

The solution of the linear system (1.3) by Tikhonov regularization gives rise to the minimization problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^{n^2}} \|A\boldsymbol{x}-\boldsymbol{b}\|^2+\alpha\|\boldsymbol{x}\|^2,$$

where $\|\cdot\|$ denotes the Euclidean norm and $\alpha>0$ is a regularization parameter that balances the regularization term $\|\boldsymbol{x}\|^2$ and the data fitting term $\|A\boldsymbol{x}-\boldsymbol{b}\|^2$. This minimization problem has the unique solution

$$\boldsymbol{x}_\alpha=(A^TA+\alpha I)^{-1}A^T\boldsymbol{b}.$$

In the case of periodic boundary conditions, since $A=\mathcal{C}_n(f)$, the Tikhonov solution $\boldsymbol{x}_\alpha$ can be computed by applying three FFTs as

$$\boldsymbol{x}_\alpha=\mathcal{C}_n(p_\alpha)\boldsymbol{b},$$

where

$$p_\alpha(\vartheta_1,\vartheta_2)=\frac{\overline{f(\vartheta_1,\vartheta_2)}}{|f(\vartheta_1,\vartheta_2)|^2+\alpha} \tag{2.3}$$

and thus the eigenvalues of $\mathcal{C}_n(p_\alpha)$ are

$$\frac{\overline{\lambda_j}}{|\lambda_j|^2+\alpha},\quad j=1,\dots,n^2.$$

In the next section we explain how to exploit these considerations to develop regularizing preconditioners for symmetrized BTTB matrices.

3. Regularizing preconditioner for symmetrized BTTB matrices. Starting from the seminal paper [22], regularizing preconditioners have been largely investigated to speed up the convergence of iterative regularization methods without spoiling the computed reconstructions [3, 9, 13]. The idea behind these approaches is to cluster the eigenvalues in the signal subspace without modifying the noise subspace. Moreover, the linear system with the preconditioned matrix has to be solved with a computational cost comparable to the matrix-vector product with the matrix A. Regarding the first requirement, it can be shown that it is fulfilled by the circulant Tikhonov operator $\mathcal{C}_n(p_\alpha)$, where p_α is defined in (2.3). A discussion on this is given in Remark 3.4. On the other hand, the class of BCCB matrices is the most efficient one among the possible classes of matrices that arise from space-invariant image deblurring, as far as the computational cost of matrix-vector product is concerned. Therefore, $\mathcal{C}_n(p_\alpha)$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

is a common choice for the preconditioning of A, independently of the imposed boundary conditions.

We introduce the distribution of the eigenvalues that will be useful to study the clustering of the eigenvalues of the preconditioned matrices. By the Szegö-Tilli theorem [34], a distribution relation holds for the eigenvalues of the matrix sequence $\{\mathcal{T}_n(g)\}_n$ for a real-valued $g \in L^1([-\pi,\pi]^2)$

(3.1)
$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{j=1}^{n^2} F[\lambda_j(\mathcal{T}_n(g))] = \frac{1}{(2\pi)^2} \int_{[-\pi,\pi]^2} F(g(\theta_1,\theta_2))\,\mathrm{d}\theta_1\,\mathrm{d}\theta_2,\ \forall F \in C_c(\mathbb{R},\mathbb{C}),$$

where $C_c(\mathbb{R},\mathbb{C})$ denotes the space of continuous functions $F : \mathbb{R} \to \mathbb{C}$. The function g is called the *symbol* of the Toeplitz family and we write

$$\{\mathcal{T}_n(g)\}_n \sim_\lambda g.$$

The informal meaning behind the above distribution result is the following. If g is continuous and n is large enough, then the spectrum of $\mathcal{T}_n(g)$ "behaves" like a uniform sampling of g over $[-\pi,\pi]^2$. It follows that $\mathcal{C}_n(g)$ is a common preconditioner for $\mathcal{T}_n(g)$ due to its eigenvalue distribution, cf. equation (2.2).

For completeness, we add here the general definition of eigenvalue distribution given for spectral symbols that are (possibly) matrix-valued. This is a case that we will encounter in the next sections.

DEFINITION 3.1 (Eigenvalue distribution of a sequence of matrices). *Let $\{A_n\}_n$ be a sequence of matrices with A_n of size d_n, and let $f : D \subset \mathbb{R}^k \to \mathbb{C}^{r\times r}$ be a measurable matrix-valued function defined on a set D with $0 < \mu_k(D) < \infty$.*

We say that $\{A_n\}_n$ has an (asymptotic) eigenvalue (or spectral) distribution described by f, and we write $\{A_n\}_n \sim_\lambda f$, if

$$\lim_{n\to\infty} \frac{1}{d_n} \sum_{j=1}^{d_n} F\left(\lambda_j\left(A_n\right)\right) = \frac{1}{\mu_k(D)} \int_D \frac{\sum_{i=1}^{r} F\left(\lambda_i(f(\mathbf{x}))\right)}{r}\,\mathrm{d}\mathbf{x}, \quad \forall F \in C_c(\mathbb{C}).$$

In this case the function f is referred to as the eigenvalue (or spectral) symbol of $\{A_n\}_n$.

When we impose accurate boundary conditions, such as reflective or antireflective, the coefficient matrix is a BTTB matrix, up to a "small" correction. It turns out that, independently of the imposed boundary conditions, the matrices associated to the PSF have the same symbol f as their Toeplitz part.

Since the PSF performs an average of neighboring pixels, we have

$$\sum_{i,j=-n+1}^{n-1} h_{i,j} = 1, \quad h_{i,j} \geq 0.$$

Therefore, the function $|f|$ has maximum at the origin and then decays, not necessarily uniformly, reaching the minimum in $[\pi,\pi]$. Indeed, it is well-known that deconvolution problems of the form (1.1) are ill-posed problems with an ill-conditioned subspace in the high frequencies, i.e., around the point $[\pi,\pi]$, see [23]. For instance, for the PSF used in Example 5.1, plotted in the left panel of Fig. 3.1, the right panel of Fig. 3.1 depicts the behavior of $|f|$ as described above. Therefore, the generating function f satisfies the assumption of the following theorem, which we will prove in Appendix A.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

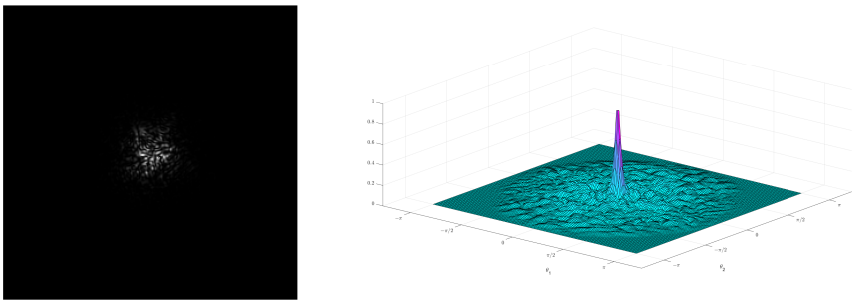

FIG. 3.1. *PSF of speckle blur (left) and associated symbol (right).*

THEOREM 3.2. *Let ε and τ be positive values such that $\varepsilon \in (0,1)$ and $\tau \in (0,\pi)$. Let $f \in L^1([-\pi,\pi]^2)$ be a bivariate function with real Fourier coefficients, periodically extended to the whole real plane, and such that*

$$\begin{cases} |f(\vartheta_1,\vartheta_2)| > \varepsilon, & \text{if } \left|\vartheta_1^2+\vartheta_2^2\right| < \tau, \\ |f(\vartheta_1,\vartheta_2)| \le \varepsilon, & \text{otherwise.} \end{cases} \tag{3.2}$$

Define

$$g_\tau(\vartheta_1,\vartheta_2) = \begin{cases} |f(\vartheta_1,\vartheta_2)|, & \text{if } \left|\vartheta_1^2+\vartheta_2^2\right| < \tau, \\ 1, & \text{otherwise.} \end{cases} \tag{3.3}$$

Then,

$$\left\{\mathcal{C}_n(g_\tau)^{-1} Y \mathcal{T}_n(f)\right\}_n \sim_\lambda \psi,$$

where

$$\psi(\vartheta_1,\vartheta_2) = \begin{cases} \frac{1}{|f|}\begin{bmatrix} 0 & f(\vartheta_1,\vartheta_2) \\ \overline{f(\vartheta_1,\vartheta_2)} & 0 \end{bmatrix}, & \text{if } \left|\vartheta_1^2+\vartheta_2^2\right| < \tau, \\ \begin{bmatrix} 0 & f(\vartheta_1,\vartheta_2) \\ \overline{f(\vartheta_1,\vartheta_2)} & 0 \end{bmatrix}, & \text{otherwise.} \end{cases} \tag{3.4}$$

Proof. See Appendix A. □

The latter theorem provides information on the spectrum of the preconditioned matrix, which can be exploited to study the behavior of $\mathcal{C}_n(g_\tau)$ as a regularizing preconditioner, as we will see in the next remark. Before going into details, we would like to point out how $\mathcal{C}_n(g_\tau)$ is constructed in practice. The PSF h is known from the application and it uniquely defines the matrix $\mathcal{T}_n(f)$ and, hence, its generating function f. The function f is then used to compute g_τ for a suitable choice of τ as in equation (3.3) and, finally, the preconditioner $\mathcal{C}_n(g_\tau)$ is determined.

REMARK 3.3. On the set where $|f| < \varepsilon$, we have $\psi = \begin{bmatrix} 0 & f \\ \overline{f} & 0 \end{bmatrix}$, which has eigenvalue functions $\pm|f|$ with image contained in $(-\varepsilon,\varepsilon)$. If $|f| > \varepsilon$, we have $\psi = \frac{1}{|f|}\begin{bmatrix} 0 & f \\ \overline{f} & 0 \end{bmatrix}$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

with eigenvalue functions ± 1. Hence, according to [2, Remarks 2.12–2.13], the spectral distribution (3.4) suggests that there are at most $o(n^2)$ eigenvalues of $\mathcal{C}_n(g_\tau)^{-1}Y\mathcal{T}_n(f)$ outside the three clusters at $\lambda = 1$, $\lambda = -1$, and $\lambda \in [-\varepsilon, \varepsilon]$, whose cardinality depends on ε and τ, provided that n is large enough. Note that ε and τ are not independent: for each ε we find a τ such that (3.2) holds. If ε tends to 0, τ tends to π, that is, the set where $|f| < \varepsilon$ has the empty set as a limit, making the eigenvalues of the preconditioned matrix clustered at ± 1. On the other hand, if ε is chosen greater than the maximum of $|f|$, the preconditioner has no effect at all. So, it is important to choose ε, and consequently τ, accurately, so that the preconditioner has a significant clustering effect without amplifying the noise.

The next remark explains how the preconditioner $\mathcal{C}_n\,(g_\tau)$ is related to the preconditioner $\mathcal{C}_n(p_\alpha)$, with p_α is defined in (2.3). To be more precise, we consider the absolute value of $\mathcal{C}_n(p_\alpha)$ as a preconditioner for symmetrized BTTB matrices and we explain why this is a regularizing preconditioner.

REMARK 3.4. As we have already stated, a common choice for the preconditioning of A in the image deblurring context is $\mathcal{C}_n(p_\alpha)$, where p_α is defined in (2.3). In [17, 31, 33] it was shown that, under proper assumptions, if $\mathcal{C}_n(p_\alpha)$ is such that the singular values of $\mathcal{C}_n(p_\alpha)A$ are clustered, then the absolute value of $\mathcal{C}_n(p_\alpha)$ is such that the eigenvalues of $|\mathcal{C}_n(p_\alpha)|YA$ are clustered. We recall that $|\mathcal{C}_n(p_\alpha)|$ is the circulant matrix whose eigenvalues are the modulus of the eigenvalues of $\mathcal{C}_n(p_\alpha)$. For these reasons, we consider the circulant matrix $\mathcal{C}_n(|p_\alpha|)$ with

$$|p_\alpha|(\vartheta_1, \vartheta_2) = \frac{|f(\vartheta_1, \vartheta_2)|}{|f(\vartheta_1, \vartheta_2)|^2 + \alpha}.$$

In general, the function g_τ is discontinuous and this guarantees a sharp subdivision between eigenvalue clusters. For $|p_\alpha|$ this is not true since it is a smooth low-pass filter. This implies that $\mathcal{C}_n(|p_\alpha|)$ is less sensitive than $\mathcal{C}_n(g_\tau)$ to the choice of the threshold parameter α and τ, respectively, which is related to ε.

To prove that $\mathcal{C}_n(|p_\alpha|)$ is a regularizing preconditioner, suppose for simplicity that we take $\alpha = \varepsilon$ and study $|p_\varepsilon|$. With the considerations that we make in Appendix A, we can state that

$$\{\mathcal{C}_n(|p_\varepsilon|)Y\mathcal{T}_n(f)\}_n \sim_\lambda |p_\varepsilon| \begin{bmatrix} 0 & f \\ \overline{f} & 0 \end{bmatrix}, \tag{3.5}$$

which has eigenvalue functions

$$\frac{|f|}{|f|^2 + \varepsilon}(\pm|f|) = \pm\frac{|f|^2}{|f|^2 + \varepsilon}.$$

Note that $\frac{|f|^2}{|f|^2+\varepsilon} < 1$, and when $|f|$ is much greater than ε the eigenvalues are close to 1, which means a speed up of the convergence in the signal subspace.

On the other hand, if $|f| \leq \varepsilon$, we have

$$\frac{|f|^2}{|f|^2 + \varepsilon} \leq \frac{\varepsilon^2}{|f|^2 + \varepsilon} \leq \frac{\varepsilon^2}{\varepsilon} = \varepsilon,$$

which means that the small eigenvalues are not amplified by the preconditioner.

In the numerical examples we will use $|p_\alpha|$ instead of g_τ since it shows an overall greater robustness and better performance.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

4. Iterative regularization methods. The purpose of the present section is to briefly introduce the iterative regularization methods that we use in the numerical experiments section; more details about all the solvers considered in this section can be found in the recent review paper [7]. As mentioned in Section 1, iterative methods used to reconstruct blurred and noisy images are affected by *semiconvergence*, i.e., they first reduce the error and then diverge from the exact solution. When possible, we will use the discrepancy principle as stopping criterion for the iterations, that is, we stop the iterations as soon as the norm of the residual vector is less than a tolerance times the noise level, i.e., the norm of the unknown noise vector affecting the right-hand-side vector $\boldsymbol{b}$ in (1.3).

As we anticipated, in the case where A is a BTTB matrix, we want to study the behavior of MINRES applied to the linear system (1.6), which has a symmetric coefficient matrix. The regularization properties of the MINRES method were proven in [29, 30]. When using MINRES, a preconditioning strategy needs to be applied symmetrically (i.e., both on the left and on the right), that is, the preconditioned symmetrized linear system becomes

$$P^{1/2}YAP^{1/2}\boldsymbol{z} = P^{1/2}Y\boldsymbol{b},\ \boldsymbol{x} = P^{1/2}\boldsymbol{z}.$$

Observe that in this case the preconditioner P needs to be positive definite to ensure the existence of $P^{1/2}$. Regarding the computational cost of this approach, note that the preconditioned MINRES algorithm can be reformulated in such a way that only one matrix-vector product with the circulant preconditioner is performed per iteration (see [21]), while the product with the anti-identity preconditioner has a negligible cost, being a permutation. When applying right preconditioning the residual of the preconditioned system is the same as that of the non-preconditioned one and the discrepancy principle can be applied without a significant additional computational cost. This is not true for left preconditioning. Therefore, in the symmetric preconditioning case, we will comment on the best reconstruction achieved by the considered methods and their stability. Other stopping criteria can be chosen, but the study of their behavior is beyond the purpose of the present paper.

While theoretical results guarantee a regularizing behavior of MINRES, the success of GMRES as regularization method is problem dependent. Although some theory has been developed [6], it often happens that the GMRES is not effective when the coefficient matrix A is not close to normal [29]. In the case where we consider reflective boundary conditions, if the PSF is not quadrantally symmetric, the matrix A is neither symmetric nor persymmetric. However, YA is close to being symmetric even if A is highly non-symmetric. In this case, we expect that GMRES applied to the system (1.6) over-performs GMRES applied to the original system (1.3).

The LSQR method requires A^T, and replacing it with A' is easy to implement but is not theoretically sound, as explained in the introduction. We stress that one iteration of LSQR costs a matrix-vector product with matrix A^T more than one iteration of MINRES and GMRES, so this needs to be taken into account when analyzing the convergence speed in terms of iteration number.

In order to speed up the convergence of the iterative methods listed above, we apply the preconditioning strategy analyzed in Section 3. More precisely, we use $\mathcal{C}_n(p_\alpha)$ as a (right) preconditioner for LSQR type methods, which we apply to systems with A and $\mathcal{C}_n(|p_\alpha|)$ for GMRES type methods, which we apply to systems with YA, on the right. Moreover, adopting FLSQR and FGMRES instead of LSQR and GMRES, respectively, we can consider the iteration-dependent circulant matrices $\mathcal{C}_n(p_{\alpha_k})$ and $\mathcal{C}_n(|p_{\alpha_k}|)$ as preconditioners, where we choose the parameter as the geometric sequence $\alpha_k = \alpha_0 q^k$, where k is the iteration counter, while $\alpha_0 = 0.1$ and $q = 0.8$ are chosen following the recommendations in [12]. In the rest of the paper, when we denote a preconditioner by P, we mean a circulant preconditioner, whose exact formulation will be clear from the context.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

For enforcing sparsity in the computed solution, we apply the iteration-dependent preconditioner studied in [19, 20], which here we simply denote by W. More specifically, at the kth iteration of the considered methods we have

$$W = \text{diag}(|\boldsymbol{x}_{k-1}|^{1/2}),$$

where $\boldsymbol{x}_{k-1}$ is the solution computed at the previous, i.e., the $(k-1)$th, iteration, and where absolute value and exponentiation are applied component-wise. Such preconditioner stems from the iteratively reweighted least squares (IRLS) method applied to the problem

$$\min_{\boldsymbol{x}} \|A\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \lambda\|\boldsymbol{x}\|_1, \tag{4.1}$$

whereby, assuming that the matrix W is periodically updated with the most recent approximation of the solution of (4.1) and invertible*, the regularization term $\lambda\|\boldsymbol{x}\|_1$ is first approximated by the squared 2-norm term $\lambda\|W^{-1}\boldsymbol{x}\|_2^2$ and then undergoes a transformation to standard form (change of variable), leading to the problem

$$\min_{\boldsymbol{y}} \|AW\boldsymbol{y} - \boldsymbol{b}\|_2^2 + \lambda\|\boldsymbol{y}\|_2^2, \quad \boldsymbol{y} = W^{-1}\boldsymbol{x}. \tag{4.2}$$

Following common practice, we apply either FGMRES or FLSQR to the above problem, after dropping the regularization term. That is, in the FGMRES case, starting from $\boldsymbol{v}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|_2$, at the mth iteration we update the partial matrix factorization

$$AZ_m = V_{m+1}H_m, \quad \text{where} \quad \begin{array}{lcl} Z_m & = & [W_1\boldsymbol{v}_1, ..., W_m\boldsymbol{v}_m] \quad \in \mathbb{R}^{n^2\times m}, \\ V_{m+1} & = & [\boldsymbol{v}_1, ..., \boldsymbol{v}_{m+1}] \quad \in \mathbb{R}^{n^2\times(m+1)}. \end{array} \tag{4.3}$$

Here we have used the notation $W_i = \text{diag}(|\boldsymbol{x}_{i-1}|^{1/2})$, with $i = 1, \ldots, m$; the columns of Z_m span the subspace for the mth FGMRES approximate solution, V_{m+1} has orthonormal columns, and $H_m \in \mathbb{R}^{(m+1)\times m}$ is upper Hessenberg. The problem solved at the mth FGMRES iteration reads

$$\begin{aligned} \boldsymbol{x}_m &= \arg \min_{\boldsymbol{x}\in\text{range}(Z_m)} \|A\boldsymbol{x} - \boldsymbol{b}\|_2 \\ &= Z_m \left(\arg \min_{\boldsymbol{s}\in\mathbb{R}^m} \|AZ_m\boldsymbol{s} - \boldsymbol{b}\|_2 \right) = Z_m \left(\arg \min_{\boldsymbol{s}\in\mathbb{R}^m} \|H_m\boldsymbol{s} - \|\boldsymbol{b}\|_2\boldsymbol{e}_1\|_2 \right), \end{aligned}$$

where we have exploited the properties of the matrices appearing in (4.3) and $\boldsymbol{e}_1$ is the first canonical basis vector of $\mathbb{R}^{m+1}$. A similar procedure can be adopted for FLSQR. Such methods allow updates of W as the iterations proceed and, therefore, are more efficient than classical IRLS schemes that update W only after each instance of problem (4.2) is fully solved (which typically requires a number of iterations of a solver such as CGLS). The iteration-dependent matrix W modifies the approximation subspace for the FGMRES and FLSQR solution in such a way that sparsity is naturally enforced within its basis vectors. We refer to [20] for more details about these approaches; a numerical illustration is also provided in Section 5.4. An iteration-depended regularizing circulant preconditioner P can be used to speed-up the convergence of the sparsity-enforcing FGMRES and FLSQR methods just described; such an approach will be detailed and numerically investigated in Section 5.4, as well.

*The matrix W^{-1} is required to state the original reweighted problem but, in practice, after the change of variables has been applied, the application of the preconditioner requires only a matrix-vector product with the matrix W. Therefore, we do not implement any shifting strategy to guarantee the invertibility of W when $\boldsymbol{x}_{k-1}$ has null entries.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

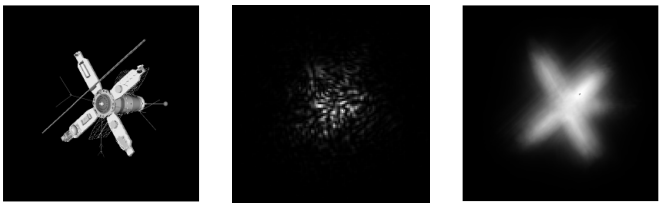

FIG. 5.1. *Satellite test problem. From left to right: exact image, PSF cropped to* $[h_{i,j}]_{i,j=-32}^{32}$*, and blurred image. Note that* $h_{i,j} < 10^{-3}$ *when* $\min\{|i|, |j|\} > 32$.

5. Numerical Examples. In this section we present four examples. The Satellite test problem in Section 5.1 and the Phantom test problem in 5.2 are aimed at discussing the efficiency of the symmetrization strategy in the case of zero boundary conditions, namely, when the matrix A is BTTB. The Cameraman test problem in Section 5.3 is an example with reflective boundary conditions, which shows the performance of the "symmetrization" strategy on matrices that are close to being symmetric. Section 5.4 is devoted to the numerical study of the reconstruction of a sparse image with reflective boundary conditions, showing the effect of the "symmetrization" strategy and preconditioning within flexible Krylov subspace methods.

In all examples, the blurred and noisy image is given by

$$\boldsymbol{b} = A\boldsymbol{x} + \frac{\boldsymbol{\xi}}{\|\boldsymbol{\xi}\|}\sigma\|A\boldsymbol{x}\|,$$

where $\boldsymbol{x}$ is the exact image, $\boldsymbol{\xi}$ is a random Gaussian white noise vector, and σ is the noise level. When available, we use the blurring functions and the implementation of the iterative methods included in *IR Tools* [18]. We highlight that in *IR Tools* the LSQR-based methods are implemented using a Kronecker product approximation of A^T instead of replacing it with A'; see [32] for details.

In order to evaluate the quality of the deblurring processes, we use two different metrics: Relative Restoration Error (RRE) and Peak Signal-to-Noise Ratio (PSNR). The RRE measures the norm of the difference between the reconstructed image $\tilde{\boldsymbol{x}}$ and the original image $\boldsymbol{x}$ relative to the norm of the original image

$$\text{RRE}(\tilde{\boldsymbol{x}}) = \frac{\|\tilde{\boldsymbol{x}} - \boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2}, \quad \boldsymbol{x}, \tilde{\boldsymbol{x}} \in \mathbb{R}^{n^2},$$

lower RRE values indicate better deblurring performance. Instead, the PSNR quantifies image quality by comparing the peak signal, that is, the maximum pixel value of the image $\max(\boldsymbol{x})$, to the noise, calculated from the mean squared error between the original image and the deblurred image

$$\text{PSNR}(\tilde{\boldsymbol{x}}) = 20 \cdot \log_{10}\left(\frac{n^2 \max(\boldsymbol{x})}{\|\tilde{\boldsymbol{x}} - \boldsymbol{x}\|_2}\right), \quad \boldsymbol{x}, \tilde{\boldsymbol{x}} \in \mathbb{R}^{n^2},$$

higher PSNR values indicate better deblurring performance.

5.1. Satellite with zero boundary conditions. The satellite image is blurred with a medium level speckle blur, which simulates spatially invariant blurring caused by atmospheric turbulence; a noise of level $\sigma = 0.05$ is added. Fig. 5.1 shows the exact image, the PSF, and the blurred image, all of size 256×256 pixels. In Fig. 5.2, we report the error behavior of (preconditioned) MINRES applied to the symmetrized system (1.6) and of (preconditioned)

ETNA
Kent State University and
Johann Radon Institute (RICAM)

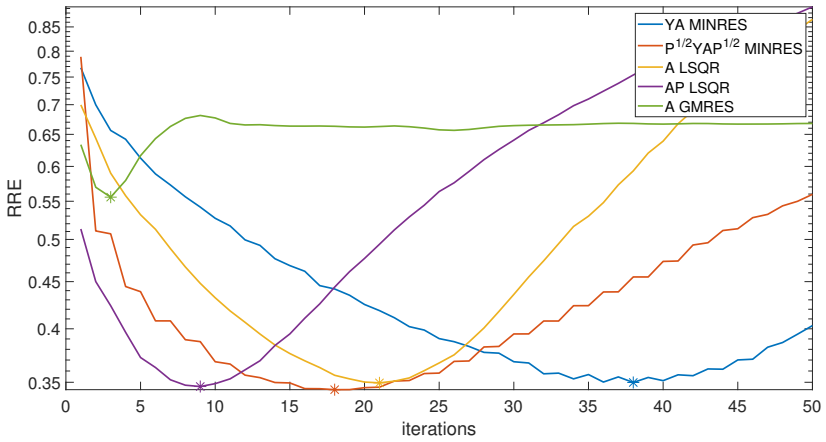

FIG. 5.2. *Satellite test problem. Comparison between the error behaviors of (preconditioned) MINRES applied to the symmetrized system (1.6) and of (preconditioned) LSQR and GMRES applied to the non-symmetrized linear system. The asterisks mark the iterations giving the best reconstruction in terms of RRE.*

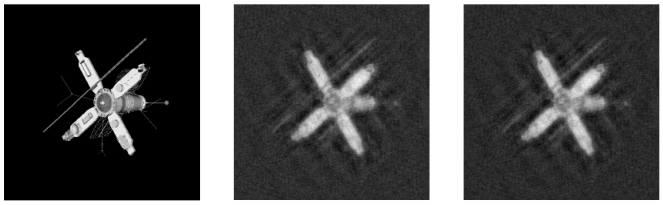

FIG. 5.3. *Satellite test problem. Exact image (left), best preconditioned MINRES reconstruction (center), best preconditioned LSQR reconstruction (right).*

LSQR and GMRES applied to the non-symmetrized linear system. The regularizing parameter α needed in the construction of the circulant preconditioners $\mathcal{C}_n(p_\alpha)$ and $\mathcal{C}_n(|p_\alpha|)$ with p_α defined in (2.3) was manually tuned, i.e., we chose the $\alpha \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}\}$ which numerically proved to be the best in terms of balancing the convergence speed and the quality of the reconstruction, which is $\alpha = 10^{-2}$. MINRES is slower to reach the minimizer than LSQR in terms of iteration number, however one iteration of MINRES is computationally less costly, not requiring the multiplication by A^T. Moreover, we highlight that preconditioned MINRES is more robust than preconditioned LSQR: for the latter, the iteration should be stopped between 5 and 13 to obtain an RRE that differs from the minimum RRE achieved less than 0.03, while for preconditioned MINRES the range of possible stopping iteration numbers for achieving an analogous result is between 10 and 27, which is a wider range. GMRES applied to the non-symmetrized linear system does not provide a reconstruction of high quality when compared to the other methods. In Fig. 5.3 we compare the best reconstruction by preconditioned MINRES and LSQR for this difficult deblurring problem. We can clearly see that, although both restored images are corrupted by some ringing artifacts, they visually look very similar; indeed, according to Fig. 5.2, the two methods achieve a comparable relative errors, although preconditioned MINRES does so more efficiently.

5.2. Phantom with zero boundary conditions. The Phantom test problem analyzes the modified Shepp-Logan phantom blurred with a unidirectional motion, which uniformly causes

ETNA
Kent State University and
Johann Radon Institute (RICAM)

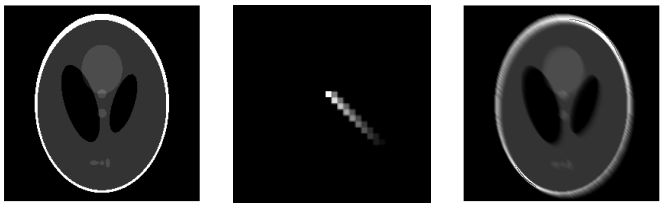

FIG. 5.4. *Phantom test problem. From left to right: exact image, PSF cropped to* $[h_{i,j}]_{i,j=-16}^{16}$*, and blurred image. Note that the PSF is exactly zero in the pixels where light is not visible.*

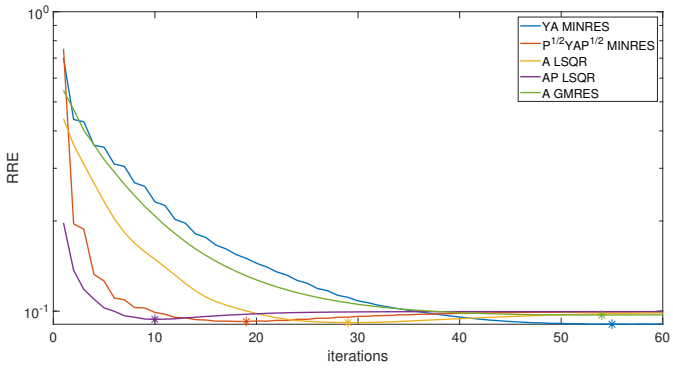

FIG. 5.5. *Phantom test problem. Comparison between the error behaviors of (preconditioned) MINRES applied to the symmetrized system* (1.6) *and of (preconditioned) LSQR and GMRES applied to the non-symmetrized linear system. The asterisks mark the iterations giving the best reconstruction in terms of RRE.*

pixels' displacement towards the right bottom direction. The PSF can be seen in Fig. 5.4, together with the exact image and the blurred image. A noise of level $\sigma = 0.01$ is applied. In Fig. 5.5, we report the error behavior of (preconditioned) MINRES applied to the symmetrized system (1.6) and of (preconditioned) LSQR and GMRES applied to the non-symmetrized linear system. Also in this case, the regularizing parameter α needed in the construction of the preconditioner $\mathcal{C}_n(p_\alpha)$ was manually chosen equal to 10^{-2}. This example confirms that preconditioned MINRES is more stable than preconditioned LSQR, since its semi-convergence is slower. We remark again that this does not translate into a higher computational cost of the overall method, because the cost of a single iteration needs to be taken into account. In the Phantom case, GMRES applied to the non-symmetrized linear system performs better than in the Satellite case, but it is still over-performed by the other methods. In Fig. 5.6 we compare the best reconstruction by preconditioned MINRES and LSQR.

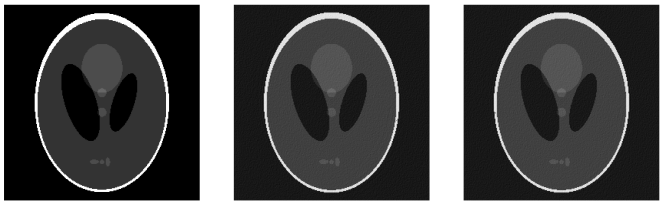

FIG. 5.6. *Phantom test problem. Exact image (left), best preconditioned MINRES reconstruction (center), best preconditioned LSQR reconstruction (right).*

ETNA
Kent State University and
Johann Radon Institute (RICAM)

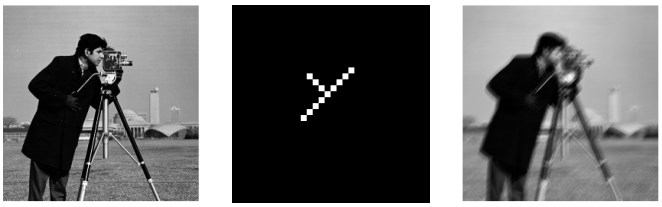

FIG. 5.7. *Cameraman test problem. From left to right: exact image, PSF cropped to* $[h_{i,j}]_{i,j=-16}^{16}$*, and blurred image. Note that the PSF is exactly zero in the pixels where light is not visible.*

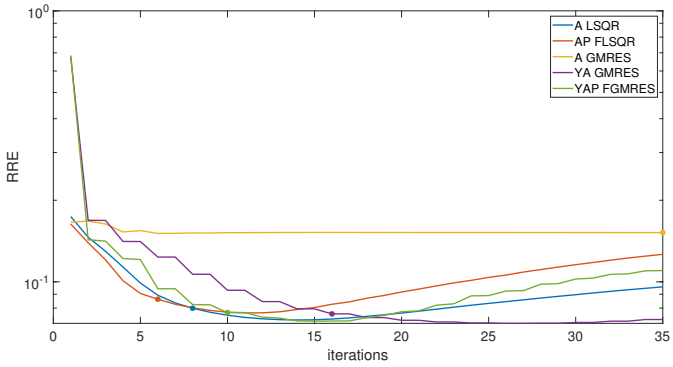

FIG. 5.8. *Cameraman test problem. Comparison between the error behaviors of the GMRES and FGMRES applied to the symmetrized system* (1.6) *and of the LSQR, FLSQR, and GMRES applied to the non-symmetrized linear system. The dots mark the iterations satisfying the discrepancy principle stopping criterion.*

5.3. Cameraman with reflective boundary conditions. In this case, we analyze the reconstruction of the cameraman image contaminated by motion blur in two directions and by a noise of level $\sigma = 0.01$. The PSF can be seen in Fig. 5.7, together with the exact image and the blurred image, all of size 256×256 pixels. In Fig. 5.8, we report the error behavior of GMRES and FGMRES applied to the symmetrized system (1.6) and of LSQR, FLSQR, and GMRES applied to the non-symmetrized linear system. The iteration-dependent preconditioners for FGMRES and FLSQR are, respectively, $\mathcal{C}_n(|p_{\alpha_k}|)$ and $\mathcal{C}_n(p_{\alpha_k})$, with $\alpha_k = 0.1 \cdot (0.8)^k$. The dots in Fig. 5.8 show the iteration for which the discrepancy principle is fulfilled. Table 5.1 shows the RRE and PSNR values with the corresponding iteration numbers and computational times for the restoration with minimum RRE and for the restoration determined when terminating the iterations with the discrepancy principle. The error behavior is in accordance with the computational times, since one iteration of LSQR costs approximately as two iterations of GMRES. We see that for FGMRES the circulant preconditioning strategy accelerates the semi-convergence, while for FLSQR it fails to do so. Of course, the behavior of the preconditioner strictly depends on the choice of α_k, for which more sophisticated strategies can be adopted, but this is beyond the scope of this work. In Fig. 5.9 we compare the best reconstruction by FGMRES and FLSQR and we notice that FLSQR produces some artifacts on the boundary, which are not present in the FGMRES reconstruction.

5.4. Edges with reflective boundary conditions. In this case, we analyze the reconstruction of a sparse image containing the edges of geometrical figures contaminated by a severe shake blur and a noise of level $\sigma = 0.1$. Fig. 5.10 shows the exact image, the PSF, and

ETNA
Kent State University and
Johann Radon Institute (RICAM)

FIG. 5.9. *Cameraman test problem. Exact image (left), preconditioned GMRES reconstruction (center), preconditioned LSQR reconstruction (right) given by the discrepancy principle.*

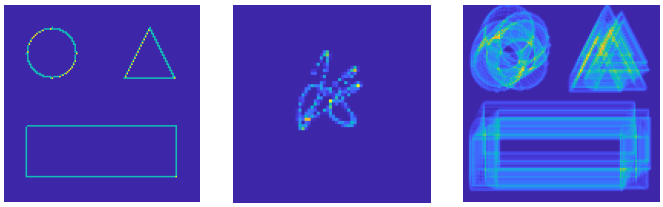

FIG. 5.10. *Edges test problem. From left to right: exact image, PSF cropped to* $[h_{i,j}]_{i,j=-32}^{32}$*, and blurred image.*

the blurred image, all of size 128×128 pixels, in a colormap that better emphasises the light intensity of the grayscale images.

In Fig. 5.11, we report the error behavior of FGMRES applied to the symmetrized system (1.6) and of FLSQR and FGMRES applied to the original linear system (1.3). We consider two different iteration-dependent preconditioners for FGMRES and FLSQR. The W preconditioner implements the re-weighting strategy from [20] used to enforce sparsity in the solution, as explained in Section 4. Then, we combine the latter strategy with the circulant preconditioning technique of $\mathcal{C}_n(|p_{\alpha_k}|)$ and $\mathcal{C}_n(p_{\alpha_k})$, with $\alpha_k = 0.1 \cdot (0.8)^k$. To our knowledge, this is the first time that the combination of these two preconditioners has been explored within flexible regularizing Krylov methods. Between the two options of preconditioners WP and PW, we choose WP since it is the choice that enforces sparsity within the basis vectors of the approximation subspace for the solution. In support of this statement, in Fig. 5.12 and Fig. 5.13 we show the first four, the eighth, and the twelfth basis vectors of the flexible Krylov subspaces generated when applying FGMRES with the two different preconditioning strategies. We can clearly see that, when preconditioning with WP, since the sparsity-enforcing preconditioner W is applied on the top of the regularizing preconditioner P, the sparsity pattern of the original image is consistently maintained across the different basis vectors, resulting in a restored image where the nonzero parts are well defined. The same is not true when preconditioning with PW: we can clearly see that the first basis vectors are not sparse and, as a consequence, many spurious nonzero pixels appear in the final reconstruction.

Table 5.1 shows the RRE and PSNR values with the corresponding iteration numbers and computational times for the restoration with minimum RRE and for the restoration determined when terminating the iterations with the discrepancy principle.

This test highlights that the sparse preconditioning technique from [20], which fails in combination with FGMRES when the matrix A is highly non-symmetric as in this case, is instead a valid choice when applied to the matrix YA. The YAW FGMRES is slightly better than the AW LSQR method in terms of RRE and PSNR of the discrepancy principle reconstruction and it reaches the stopping criterion in comparable computational times. Moreover, regarding the combination with the circulant preconditioner, we see that with this choice

ETNA
Kent State University and
Johann Radon Institute (RICAM)

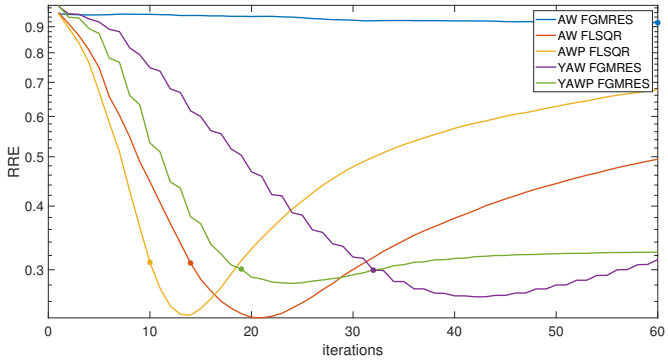

FIG. 5.11. *Edges test problem. Comparison between the error behaviors of FGMRES applied to the symmetrized system (1.6) and of FLSQR and FGMRES applied to the non-symmetrized linear system. The dots mark the iterations satisfying the discrepancy principle stopping criterion.*

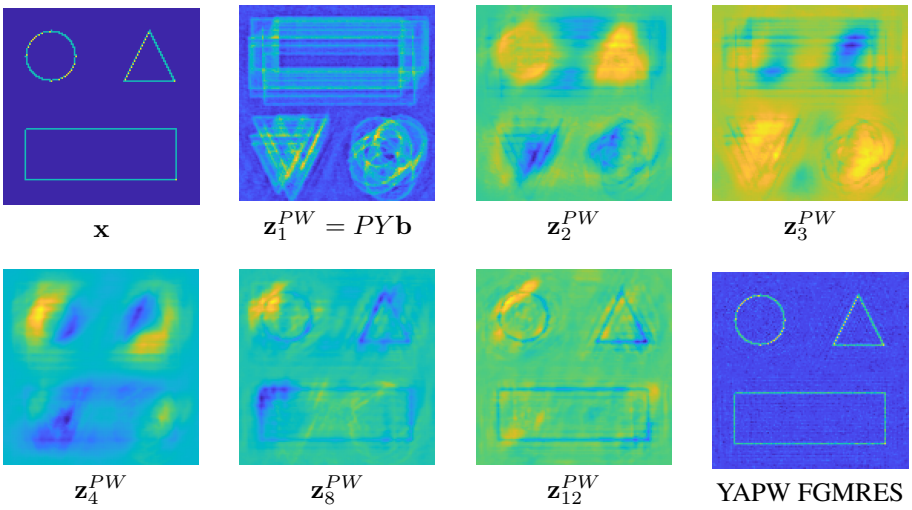

FIG. 5.12. *Edges test problem. The true image* $\mathbf{x}$*, the basis images number 1, 2, 3, 4, 8, and 12 of the flexible Krylov subspace, and the reconstruction returned by the discrepancy principle when applying FGMRES to with a* PW *iteration-dependent right preconditioning strategy. In the first iteration, the preconditioning matrix* W *is set to be the identity matrix.*

of α_k the semi-convergence of both FGMRES and FLSQR is accelerated, but what is most remarkable is that for FGMRES the convergence also becomes more stable.

Finally, we see in Fig. 5.14 that the reconstructions of AW LSQR and YAW FGMRES are visually comparable.

6. Conclusions and future work. This paper described a couple of preconditioning strategies to be used when solving image deblurring problems through iterative regularization methods. Some theoretical insight and substantial numerical experimentation showed that using such preconditioners within the MINRES, GMRES, FGMRES and FLSQR methods and with a variety of PSFs and boundary conditions leads to better reconstructions, which are computed more efficiently and often in a more stable way with respect to their unpreconditioned counterparts.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

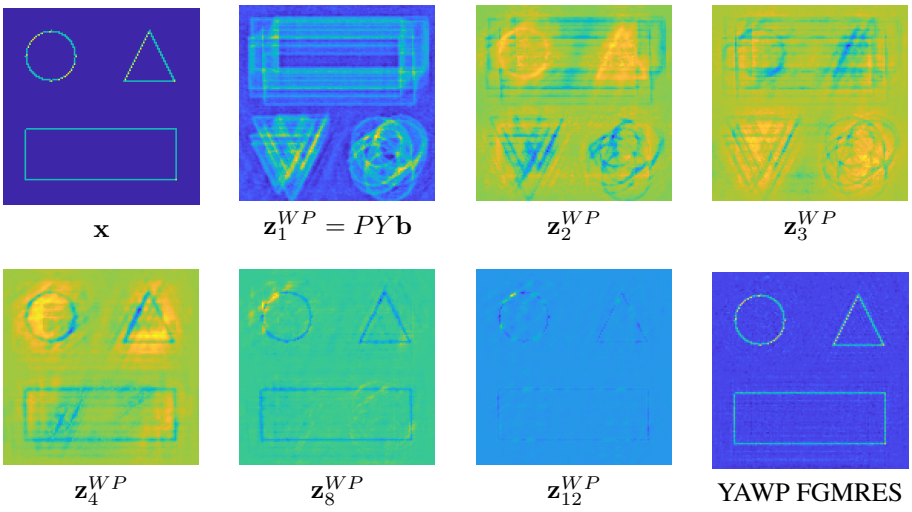

FIG. 5.13. *Edges test problem. The true image* $\mathbf{x}$*, the basis images number 1, 2, 3, 4, 8, and 12 of the flexible Krylov subspace, and the reconstruction returned by the discrepancy principle when applying FGMRES with a* WP *iteration-dependent right preconditioning strategy. In the first iteration, the preconditioning matrix* W *is set to be the identity matrix.*

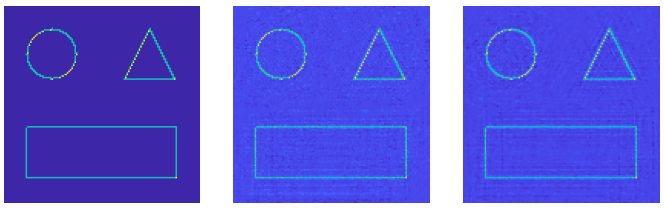

FIG. 5.14. *Edges test problem. Exact image (left),* YAW *FGMRES discrepancy principle reconstruction (center), and* AW *FLSQR discrepancy principle reconstruction (right).*

There are a number of extensions to the present work. First of all, as hinted in Section 4, one may consider different adaptive ways of setting the regularization parameter α within the regularizing circulant preconditioner $\mathcal{C}_n(|p_\alpha|)$, including strategies based on the discrepancy principle and generalized cross-validation. Secondly, although this paper focused on purely iterative regularization methods, all the Krylov subspace methods considered here can be used in a hybrid fashion, i.e., combining projection onto Krylov subspaces and Tikhonov regularization; see [7]. An interesting new research avenue may explore the performance of the preconditioners considered here in a hybrid setting. Finally, when employing Krylov methods for the computation of sparse solutions, there exists popular alternatives to the flexible methods considered here, which are based on generalized Krylov subspaces; see [4,28]. Future work can focus on the numerical study and analysis of preconditioning techniques based on regularizing circulant preconditioners applied to such methods.

Acknowledgments. The work of the first and second authors was partially supported by the Gruppo Nazionale per il Calcolo Scientifico (GNCS).

Appendix A. Proof of Theorem 3.2. The proof of Theorem 3.2 is based on a distribution result in [31], which in turn exploits the concept of Generalized Locally Toeplitz (GLT) sequences. The formal definition of the GLT class and the derivation of their properties require

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 5.1
Cameraman and Edges test problems. RRE and PSNR values with the corresponding iteration numbers and computational times for the restoration with minimum RRE and for the restoration determined when terminating the iterations with the discrepancy principle.

Example	Method	*Best Reconstruction*			*Discrepancy Principle*			
		RRE	PSNR	iter	RRE	PSNR	iter	time (s)
Cameraman	A LSQR	0.0723	28.3279	14	0.0799	27.4657	8	1.3023
	AP FLSQR	0.0768	27.8058	11	0.0863	26.7965	6	1.3191
	A GMRES	0.1509	21.9410	6	-	-	-	-
	YA GMRES	0.0703	28.5773	26	0.0762	27.8791	16	1.4398
	YAP FGMRES	0.0716	28.4157	15	0.0770	27.7841	10	1.2180
Edges	AW FLSQR	0.2414	32.7000	21	0.3093	30.5465	14	0.8622
	APW FLSQR	0.2443	32.5954	14	0.3103	30.5201	10	0.8258
	AW FGMRES	0.9166	21.1105	60	-	-	-	-
	YAW FGMRES	0.2655	31.8737	42	0.2994	30.8293	32	0.9335
	YAPW FGMRES	0.2821	31.3466	24	0.3010	30.7835	19	0.8536

rather technical tools. We refer the reader to [2] for a discussion on the GLT theory in its general multilevel block form.

The crucial feature of a GLT sequence $\{A_n\}_n$ is that we can associate each sequence to a function $\tilde{f} : [0,1]^k \times [-\pi,\pi]^k \to \mathbb{C}^{s\times s}$, called GLT symbol. We denote this relation with $\{A_n\}_n \sim_{\text{GLT}} \tilde{f}$. If all the matrices of the sequence are Hermitian, then $\tilde{f}$ is the eigenvalue symbol of $\{A_n\}_n$ in a sense analogous to formula (3.1). Every k-level Toeplitz sequence generated by a function $f \in L^1([-\pi,\pi]^k)$ is a GLT sequence and its symbol is $\tilde{f}(\cdot,\boldsymbol{\vartheta}) = f(\boldsymbol{\vartheta})$. For image deblurring problems $k = 2$ and a 2-level Toeplitz matrix is a BTTB matrix.

GLT sequences constitute a $*$-algebra of matrix-sequences to which multilevel Toeplitz matrix-sequences with Lebesgue integrable generating functions belong. The sequence obtained via algebraic operations on a finite set of given GLT sequences is still a GLT sequence and its symbol is obtained by performing the same algebraic manipulations on the corresponding symbols of the input GLT sequences.

In [31], the authors make use of the GLT theory to discover the spectral distribution of $\{Y\mathcal{T}_n(g)\}_n$ and of the preconditioned sequence $\{P_n^{-1}Y\mathcal{T}_n(g)\}_n$, under specific assumptions on P_n. For the proof of Theorem 3.2, we use part of the theory in [31], from which we report here a definition and two theorems.

DEFINITION A.1 ([31]). *We introduce the following matrices*

- $\Pi_n = \Pi_{n_1} \otimes \Pi_{n_2} \otimes \cdots \otimes \Pi_{n_d}$ *with* Π_{n_k}, n_k *even, such that its* j *th column* π_j, $j = 1,\ldots,n_k$, *is*

$$\pi_j = \begin{cases} e_{2j-1}, & j = 1,\ldots,n_k/2, \\ e_{2(j-n_k/2)}, & j = n_k/2+1,\ldots,n_k, \end{cases}$$

  where $e_j, j = 1,\ldots,n_k$, *is the* j *th column of the identity matrix of dimension* n_k.
- $U_n = U_{n_1} \otimes U_{n_2} \otimes \cdots \otimes U_{n_d}$ *with* U_{n_k} *such that*

$$U_{n_k} = \begin{bmatrix} Y_{\lceil n_k/2 \rceil} & \\ & I_{\lfloor n_k/2 \rfloor} \end{bmatrix}.$$

THEOREM A.2 ([31, Remark 4]). *Assume that* $n = (n_1,\ldots,n_d)$ *with* $n_k = 2m_k$, $m_k \in \mathbb{N}$. *Then,*

$$\left\{\Pi_n U_n \mathcal{T}_n(f) U_n \Pi_n^T\right\}_n \sim_{\text{GLT}} \begin{bmatrix} f^* & 0 \\ 0 & f \end{bmatrix}.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

THEOREM A.3 ([31, Theorem 3.3]). *Let* $\{\mathcal{T}_n(f)\}_n, \mathcal{T}_n(f) \in \mathbb{R}^{d_n \times d_n}$ *with* $n_k = 2m_k$ *and* $m_k \in \mathbb{N}$, *be the multilevel Toeplitz sequence associated with* $f \in L^1\left([-\pi,\pi]^d\right)$, *let* $\{Y_n\mathcal{T}_n(f)\}_n$ *be the corresponding sequence of flipped multilevel Toeplitz matrices, and let* $\{\mathcal{P}_n\}_n$ *be a sequence of Hermitian positive definite matrices such that* $\{\mathcal{P}_n\}_n \sim_{\text{GLT}} h$, *and* $\left\{\Pi_n U_n \mathcal{P}_n U_n \Pi_n^T\right\}_n \sim_{\text{GLT}} h \cdot I_2$ *with* $h : [-\pi,\pi]^d \to \mathbb{C}$ *and* $h \neq 0$ *a.e. Then,*

$$\left\{\mathcal{P}_n^{-1} Y_n \mathcal{T}_n(f)\right\}_n \sim_\lambda \left(h^{-1} g, [-\pi,\pi]^d\right).$$

We are now ready to prove Theorem 3.2, which we rewrite here to improve readability.

THEOREM A.4. *Let* ε *and* τ *be positive values such that* $\varepsilon \in (0,1)$ *and* $\tau \in (0,\pi)$. *Let* $f \in L^1([-\pi,\pi]^2)$ *be a bivariate function with real Fourier coefficients, periodically extended to the whole real plane, and such that*

$$\begin{cases} |f(\vartheta_1,\vartheta_2)| > \varepsilon, & \text{if } \left|\vartheta_1^2 + \vartheta_2^2\right| < \tau, \\ |f(\vartheta_1,\vartheta_2)| \le \varepsilon, & \text{otherwise.} \end{cases}$$

Define

$$g_\tau(\vartheta_1,\vartheta_2) = \begin{cases} |f(\vartheta_1,\vartheta_2)|, & \text{if } \left|\vartheta_1^2 + \vartheta_2^2\right| < \tau, \\ 1, & \text{otherwise.} \end{cases}$$

Then,

$$\left\{\mathcal{C}_n(g_\tau)^{-1} Y \mathcal{T}_n(f)\right\}_n \sim_\lambda \psi, \tag{A.1}$$

where

$$\psi(\vartheta_1,\vartheta_2) = \begin{cases} \frac{1}{|f|}\begin{bmatrix} 0 & f(\vartheta_1,\vartheta_2) \\ \overline{f(\vartheta_1,\vartheta_2)} & 0 \end{bmatrix}, & \text{if } \left|\vartheta_1^2 + \vartheta_2^2\right| < \tau, \\ \begin{bmatrix} 0 & f(\vartheta_1,\vartheta_2) \\ \overline{f(\vartheta_1,\vartheta_2)} & 0 \end{bmatrix}, & \text{otherwise.} \end{cases}$$

Proof. Since g_τ is a strictly positive function, the matrices $\mathcal{C}_n(g_\tau)$ are Hermitian positive definite for all n. According to Theorem A.3, the spectral distribution (A.1) holds under the assumptions

(i) $\{\mathcal{C}_n(g_\tau)\}_n \sim_{\text{GLT}} g_\tau$;
(ii) $\{\Pi_n U_n \mathcal{C}_n(g_\tau) U_n \Pi_n^T\}_n \sim_{\text{GLT}} g_\tau I_2$, where Π_n and U_n are the permutation matrices reported in Definition A.1.

In order to prove (i) and (ii), we state the following results

- $\{\mathcal{T}_n(g_\tau)\}_n \sim_{\text{GLT}} g_\tau$, which is a GLT property;
- $\{\Pi_n U_n \mathcal{T}_n(g_\tau) U_n \Pi_n^T\}_n \sim_{\text{GLT}} g_\tau I_2$, which holds thanks to Theorem A.2;
- $\{\mathcal{C}_n(g_\tau) - \mathcal{T}_n(g_\tau)\}_n \sim_{\text{GLT}} 0$; see Remark A.5.

Combining the latter statements with the $*$-algebra property of GLT sequences, we deduce that (i) and (ii) hold. Hence, $\left\{\mathcal{C}_n(g_\tau)^{-1} Y \mathcal{T}_n(f)\right\}_n \sim_\lambda \psi$ and the proof is complete. ∎

REMARK A.5. For brevity's sake, in Section 2 we stated that $\mathcal{C}_n(f)$ is the Strang preconditioner of $\mathcal{T}_n(f)$. This is true if f is a trigonometric polynomial for n large enough,

ETNA
Kent State University and
Johann Radon Institute (RICAM)

which is the case of the functions f related to the blurring operators that we are considering. The Strang preconditioner might not be a feasible choice for general functions $f \in L^1$; see [15] for details. However, the Strang preconditioner that we consider for g_τ is related to the Strang preconditioner constructed from trigonometric polynomial. Indeed, this is the procedure to construct the preconditioner:

1. from the trigonometric polynomial f find the eigenvalues of the Strang preconditioner $\mathcal{C}_n(f)$;
2. set to 1 the eigenvalues that are equal to or less than ε;
3. construct the circulant matrix with the eigenvalues computed in Step 2.

So, the eigenvalues of the circulant preconditioner are actually a uniform sampling of the function g_τ and hence the distribution $\{\mathcal{C}_n(g_\tau) - \mathcal{T}_n(g_\tau)\}_n \sim_{\text{GLT}} 0$ holds, where the notation $\mathcal{C}_n(g_\tau)$ is an abuse of notation since, in general, $\mathcal{C}_n(g_\tau)$ is neither the Strang preconditioner nor the circulant matrix generated by g_τ, but it well-approximates both in the case where the Fourier coefficients of g_τ decay rapidly.

REMARK A.6. Thanks to the GLT theory, it is also possible to analyze the spectral behavior of the preconditioned matrices $\mathcal{C}_n\left(|p_\alpha|\right) Y \mathcal{T}_n(f)$. In the case where $|p_\alpha|$ does not vanish at any points, it is possible to mimic the proof of Theorem 3.2 replacing g_τ with the reciprocal of $|p_\alpha|$. Otherwise, in order to deduce the spectral distribution in formula (3.5), one needs to study the matrix-sequence

$$\{\mathcal{C}_n\left(|p_\alpha|\right) Y \mathcal{T}_n(f) - \mathcal{C}_n(g_\tau)^{-1} Y \mathcal{T}_n(f)\}_n,$$

which is zero distributed in the GLT sense for a proper choice of α in relation to τ. The full exposition of the proof for the vanishing case is omitted in this paper, since such a detailed elaboration would require extensive additional background material, diverging from the core focus of this study.

REFERENCES

[1] M. S. ALMEIDA AND M. FIGUEIREDO, *Deconvolving images with unknown boundaries using the alternating direction method of multipliers*, IEEE Trans. Image Process., 22 (2013), pp. 3074–3086.

[2] G. BARBARINO, C. GARONI, AND S. SERRA-CAPIZZANO, *Block generalized locally Toeplitz sequences: theory and applications in the multidimensional case*, Electron. Trans. Numer. Anal., 53 (2020), pp. 113–216. https://etna.ricam.oeaw.ac.at/vol.53.2020/pp113-216.dir/pp113-216.pdf

[3] P. BRIANZI, F. DI BENEDETTO, AND C. ESTATICO, *Improvement of space-invariant image deblurring by preconditioned landweber iterations*, SIAM J. Sci. Comput., 30 (2008), pp. 1430–1458.

[4] A. BUCCINI AND L. REICHEL, *An ℓ_2-ℓ_q regularization method for large discrete ill-posed problems*, J. Sci. Comput., 78 (2019), pp. 1526–1549.

[5] Y. CAI, M. DONATELLI, D. BIANCHI, AND T.-Z. HUANG, *Regularization preconditioners for frame-based image deblurring with reduced boundary artifacts*, SIAM J. Sci. Comput., 38 (2016), pp. B164–B189.

[6] D. CALVETTI, B. LEWIS, AND L. REICHEL, *On the choice of subspace for iterative methods for linear discrete ill-posed problems*, Int. J. Appl. Math. Comput. Sci., 11 (2001), pp. 1069–1092.

[7] J. CHUNG AND S. GAZZOLA, *Computational methods for large-scale inverse problems: a survey on hybrid projection methods*. Preprint on arXiv, 2022. https://arxiv.org/abs/2105.07221v2.

[8] P. J. DAVIS, *Circulant Matrices*, Wiley, New York, 1979.

[9] P. DELL'ACQUA, M. DONATELLI, AND C. ESTATICO, *Preconditioners for image restoration by reblurring techniques*, J. Comput. Appl. Math., 272 (2014), pp. 313–333.

[10] P. DELL'ACQUA, M. DONATELLI, AND L. REICHEL, *Non-stationary structure-preserving preconditioning for image restoration*, in Computational Methods for Inverse Problems in Imaging, M. Donatelli and S. Serra-Capizzano, eds., Springer, Cham, 2019, pp. 51–75.

[11] M. DONATELLI, C. ESTATICO, A. MARTINELLI, AND S. SERRA-CAPIZZANO, *Improved image deblurring with anti-reflective boundary conditions and re-blurring*, Inverse Problems, 22 (2006), pp. 2035–2053.

[12] M. DONATELLI AND M. HANKE, *Fast nonstationary preconditioned iterative methods for ill-posed problems, with application to image deblurring*, Inverse Problems, 29 (2013), Art. 095008 (16 pages).

[13] M. DONATELLI, D. MARTIN, AND L. REICHEL, *Arnoldi methods for image deblurring with anti-reflective boundary conditions*, Appl. Math. Comput., 253 (2015), pp. 135–150.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[14] H. W. ENGL, M. HANKE, AND A. NEUBAUER, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.
[15] C. ESTATICO AND S. SERRA-CAPIZZANO, *Superoptimal approximation for unbounded symbols*, Linear Algebra Appl., 428 (2008), pp. 564–585.
[16] Y. W. D. FAN AND J. G. NAGY, *Synthetic boundary conditions for image deblurring*, Linear Algebra Appl., 434 (2011), pp. 2244–2268.
[17] P. FERRARI, I. FURCI, S. HON, M. A. MURSALEEN, AND S. SERRA-CAPIZZANO, *The eigenvalue distribution of special 2-by-2 block matrix-sequences with applications to the case of symmetrized Toeplitz structures*, SIAM J. Matrix Anal. Appl., 40 (2019), pp. 1066–1086.
[18] S. GAZZOLA, P. C. HANSEN, AND J. G. NAGY, *IR Tools: a MATLAB package of iterative regularization methods and large-scale test problems*, Numer. Algorithms, 81 (2019), pp. 773–811.
[19] S. GAZZOLA AND J. G. NAGY, *Generalized Arnoldi–Tikhonov method for sparse reconstruction*, SIAM J. Sci. Comput., 36 (2014), pp. B225–B247.
[20] S. GAZZOLA, J. G. NAGY, AND M. SABATÉ LANDMAN, *Iteratively reweighted FGMRES and FLSQR for sparse reconstruction*, SIAM J. Sci. Comput., 43 (2021), pp. S47–S69.
[21] A. GREENBAUM, *Iterative Methods for Solving Linear Systems*, SIAM, Philadelphia, 1997.
[22] M. HANKE, J. NAGY, AND R. PLEMMONS, *Preconditioned iterative regularization for ill—posed problems*, in Numerical Linear Algebra, L. Reichel, A. Ruttan and R.S. Varga, eds., De Gruyter, Berlin, 1993, pp. 141–164.
[23] M. HANKE AND J. G. NAGY, *Restoration of atmospherically blurred images by symmetric indefinite conjugate gradient techniques*, Inverse Problems, 12 (1996), pp. 157–173.
[24] P. C. HANSEN, *Rank-Deficient and Discrete Ill-Posed Problems*, SIAM, Philadelphia, 1998.
[25] ———, *Discrete Inverse Problems*, SIAM, Philadelphia, 2010.
[26] P. C. HANSEN, J. G. NAGY, AND D. P. O'LEARY, *Deblurring Images: Matrices, Spectra, and Filtering*, SIAM, Philadelphia, 2006.
[27] S. HON, M. A. MURSALEEN, AND S. SERRA-CAPIZZANO, *A note on the spectral distribution of symmetrized Toeplitz sequences*, Linear Algebra Appl., 579 (2019), pp. 32–50.
[28] G. HUANG, A. LANZA, S. MORIGI, L. REICHEL, AND F. SGALLARI, *Majorization–minimization generalized Krylov subspace methods for ℓ_p-ℓ_q optimization applied to image restoration*, BIT, 57 (2017), pp. 351–378.
[29] T. K. JENSEN AND P. C. HANSEN, *Iterative regularization with minimum-residual methods*, BIT, 47 (2007), pp. 103–120.
[30] M. KILMER AND G. W. STEWART, *Iterative regularization and MINRES*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 613–628.
[31] M. MAZZA AND J. PESTANA, *The asymptotic spectrum of flipped multilevel Toeplitz matrices and of certain preconditionings*, SIAM J. Matrix Anal. Appl., 42 (2021), pp. 1319–1336.
[32] J. G. NAGY, M. K. NG, AND L. PERRONE, *Kronecker product approximations for image restoration with reflexive boundary conditions*, SIAM J. Matrix Anal. Appl., 25 (2003), pp. 829–841.
[33] J. PESTANA AND A. J. WATHEN, *A preconditioned MINRES method for nonsymmetric Toeplitz matrices*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 273–288.
[34] P. TILLI, *A note on the spectral distribution of Toeplitz matrices*, Linear Multilinear Algebra, 45 (1998), pp. 147–159.