SIMPLE SQUARE SMOOTHING REGULARIZATION OPERATORS

LOTHAR REICHEL* AND QIANG YE†

Dedicated to Gérard Meurant on the occasion of his 60th birthday.

Abstract. Tikhonov regularization of linear discrete ill-posed problems often is applied with a finite difference regularization operator that approximates a low-order derivative. These operators generally are represented by banded rectangular matrices with fewer rows than columns. They therefore cannot be applied in iterative methods that are based on the Arnoldi process, which requires the regularization operator to be represented by a square matrix. This paper discusses two approaches to circumvent this difficulty: zero-padding of the rectangular matrices to make them square and extending the rectangular matrix to a square circulant. We also describe how to combine these operators by weighted averaging and with orthogonal projection. Applications to Arnoldi and Lanczos bidiagonalization-based Tikhonov regularization, as well as to truncated iteration with a range restricted minimal residual method, are presented.

Key words. ill-posed problem, regularization operator, Tikhonov regularization, truncated iteration.

1. Introduction. This paper is concerned with the computation of an approximate solution of linear systems of equations of the form

$$A\boldsymbol{x} = \boldsymbol{b}, \qquad A \in \mathbb{R}^{n\times n}, \qquad \boldsymbol{x}, \boldsymbol{b} \in \mathbb{R}^n, \tag{1.1}$$

with a matrix A of ill-determined rank. In particular, A is severely ill-conditioned and may be singular. This kind of linear systems of equations often are referred to as linear discrete ill-posed problems. They stem, e.g., from the discretization of ill-posed problems, such as Fredholm integral equations of the first kind with a smooth kernel. The system (1.1) is not required to be consistent.

Linear discrete ill-posed problems arise when one seeks to determine the cause of an observed effect. The latter is represented by the right-hand side $\boldsymbol{b}$, which in applications often is contaminated by an unknown measurement error $\boldsymbol{e} \in \mathbb{R}^n$, i.e.,

$$\boldsymbol{b} = \hat{\boldsymbol{b}} + \boldsymbol{e}, \tag{1.2}$$

where $\hat{\boldsymbol{b}}$ denotes the unknown error-free right-hand side vector associated with $\boldsymbol{b}$. We will refer to the error $\boldsymbol{e}$ as "noise."

The present paper discusses methods for the solution of (1.1) that are applicable when the norm of the noise,

$$\varepsilon := \|\boldsymbol{e}\|, \tag{1.3}$$

or an accurate approximation thereof, is available. However, the regularization operators considered in this paper also can be applied in methods that do not require knowledge of the error norm (1.3). Throughout this paper $\|\cdot\|$ denotes the Euclidean vector norm.

Introduce the linear system of equations

$$A\boldsymbol{x} = \hat{\boldsymbol{b}} \tag{1.4}$$

*Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA. E-mail: reichel@math.kent.edu.

†Department of Mathematics, University of Kentucky, Lexington, KY 40506, USA. E-mail: qye@ms.uky.edu. Supported in part by NSF under grant DMS-0411502.

associated with (1.1) with the unknown error-free right-hand side. The system (1.4) is assumed to be consistent and its solution of minimal Euclidean norm is denoted by $\hat{\boldsymbol{x}}$. We would like to determine an approximation of $\hat{\boldsymbol{x}}$ and seek to achieve this by computing an approximate solution of the available linear discrete ill-posed problem (1.1). Note that due to the severe ill-conditioning of A and the error $\boldsymbol{e}$ in the right-hand side $\boldsymbol{b}$, the least-squares solution of minimal Euclidean norm of (1.1) typically is not a meaningful approximation of $\hat{\boldsymbol{x}}$. In order to be able to obtain an accurate approximation of $\hat{\boldsymbol{x}}$, the system (1.1) has to be replaced by a nearby system, whose solution is less sensitive to perturbations in $\boldsymbol{b}$. This replacement is known as regularization.

One of the most popular regularization methods, known as Tikhonov regularization, replaces the linear system (1.1) by the minimization problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^n}\{\|A\boldsymbol{x}-\boldsymbol{b}\|^2+\mu\|L\boldsymbol{x}\|^2\}, \tag{1.5}$$

where the matrix $L\in\mathbb{R}^{k\times n}$, $k\leq n$, is referred to as the regularization operator and the scalar $\mu\geq 0$ as the regularization parameter; see, e.g., Engl et al. [14] and Hansen [19] for discussions on Tikhonov regularization.

Let $\mathcal{N}(M)$ and $\mathcal{R}(M)$ denote the null space and range of the matrix M, respectively. We assume throughout this paper that the matrices A and L satisfy

$$\mathcal{N}(A)\cap\mathcal{N}(L)=\{\mathbf{0}\}. \tag{1.6}$$

Then the Tikhonov minimization problem (1.5) has the unique solution

$$\boldsymbol{x}_\mu:=(A^TA+\mu L^TL)^{-1}A^T\boldsymbol{b} \tag{1.7}$$

for any $\mu>0$.

Our computational task is to determine a suitable positive value of the regularization parameter μ and the associated solution $\boldsymbol{x}_\mu$ of (1.5). We will determine μ with the aid of the discrepancy principle, which requires that the norm of the noise (1.3), or a fairly accurate estimate thereof, be available. This approach of determining μ generally calls for the evaluation of the norm of the residual error

$$\boldsymbol{r}_\mu:=\boldsymbol{b}-A\boldsymbol{x}_\mu \tag{1.8}$$

for several values of μ; see Section 3 for details.

Common choices of regularization operators for problems in one space-dimension are the identity matrix, as well as scaled finite difference approximations of a derivative, such as

$$L_1:=\frac{1}{2}\begin{bmatrix} 1 & -1 & & & 0\\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ 0 & & & 1 & -1 \end{bmatrix}\in\mathbb{R}^{(n-1)\times n}, \tag{1.9}$$

$$L_2:=\frac{1}{4}\begin{bmatrix} -1 & 2 & -1 & & & 0\\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ 0 & & & -1 & 2 & -1 \end{bmatrix}\in\mathbb{R}^{(n-2)\times n}, \tag{1.10}$$

and

$$L_3 := \frac{1}{8}\begin{bmatrix} -1 & 3 & -3 & 1 & & & 0 \\ & -1 & 3 & -3 & 1 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ 0 & & & -1 & 3 & -3 & 1 \end{bmatrix} \in \mathbb{R}^{(n-3)\times n}. \tag{1.11}$$

When the matrices A and L are of small to moderate size, the solution of (1.5) and the norm of the residual error (1.8) conveniently can be determined for several values of the regularization parameter with the aid of the generalized singular value decomposition (GSVD) of the matrix pair $\{A, L\}$. However, for large-scale problems, the computation of the GSVD is too expensive to be attractive. This paper is concerned with iterative solution methods that can be applied to the solution of large-scale problems.

Large-scale problems typically are transformed to standard form before solution. Let $L^\dagger \in \mathbb{R}^{n\times k}$ denote the Moore-Penrose pseudoinverse of the regularization operator L in (1.5). The A-weighted pseudoinverse of L is defined by

$$L_A^\dagger := \left(I - (A(I - L^\dagger L))^\dagger A\right) L^\dagger \in \mathbb{R}^{n\times k}; \tag{1.12}$$

see, e.g., Eldén [13] or Hansen [19, Section 2.3]. Introduce the matrix

$$\bar{A} := AL_A^\dagger \tag{1.13}$$

and the vectors

$$\boldsymbol{x}^{(0)} := \left(A(I - L^\dagger L)\right)^\dagger \boldsymbol{b}, \tag{1.14}$$

$$\bar{\boldsymbol{b}} := \boldsymbol{b} - A\boldsymbol{x}^{(0)}. \tag{1.15}$$

Let $\bar{\boldsymbol{x}} := L\boldsymbol{x}$. Then the Tikhonov minimization problem (1.5) can be expressed in standard form,

$$\min_{\bar{\boldsymbol{x}}\in\mathbb{R}^k} \{\|\bar{A}\bar{\boldsymbol{x}} - \bar{\boldsymbol{b}}\|^2 + \mu\|\bar{\boldsymbol{x}}\|^2\}. \tag{1.16}$$

The solution $\boldsymbol{x}_\mu$ of (1.5), given by (1.7), can be recovered from the solution $\bar{\boldsymbol{x}}_\mu$ of (1.16) according to

$$\boldsymbol{x}_\mu = L_A^\dagger \bar{\boldsymbol{x}}_\mu + \boldsymbol{x}^{(0)}, \tag{1.17}$$

see, e.g., [13] or [19, Section 2.3] for details. We note for future reference that if $\bar{\boldsymbol{x}}$ is any vector in $\mathbb{R}^k$ and $\boldsymbol{x} := L_A^\dagger \bar{\boldsymbol{x}} + \boldsymbol{x}^{(0)}$, then the associated residual vectors $\bar{\boldsymbol{r}} := \bar{\boldsymbol{b}} - \bar{A}\bar{\boldsymbol{x}}$ and $\boldsymbol{r} := \boldsymbol{b} - A\boldsymbol{x}$ satisfy

$$\|\bar{\boldsymbol{r}}\| = \|\boldsymbol{r}\|. \tag{1.18}$$

If L is a square nonsingular matrix, then $\boldsymbol{x}^{(0)} = \boldsymbol{0}$, and the above transformation amounts to $\bar{\boldsymbol{x}} := L\boldsymbol{x}$ and $\bar{A} := AL^{-1}$.

When the regularization operator L approximates a derivative, its pseudoinverse approximates an integral operator and therefore is a low-pass filter. This filter is applied to the computed solution $\bar{\boldsymbol{x}}_\mu$ of (1.16) when evaluating the associated solution $\boldsymbol{x}_\mu$ of (1.5); cf. (1.17). The smoothing achieved by filtering is beneficial when the

desired solution $\hat{\boldsymbol{x}}$ is smooth. Regularization operators of the form (1.9)-(1.11) often are referred to as smoothing operators.

We remark that despite the complexity of the A-weighted pseudoinverse (1.12), matrix-vector products with the matrices $L_A^\dagger$ and $AL_A^\dagger$ can be evaluated quite inexpensively when L is a banded matrix with small bandwidth, a circulant, or an orthogonal projection; see Section 2 or [13], [19, Section 2.3], and [24] for further details. Commonly used iterative methods for the solution of Tikhonov minimization problems in standard form (1.16) are based on partial Lanczos bidiagonalization of the matrix (1.13); see, e.g., [3, 4, 7, 8, 15, 17, 22]. However, recent numerical results reported in [23] show that Tikhonov regularization based on the range restricted Arnoldi (RR-Arnoldi) process can be competitive with Tikhonov regularization based on Lanczos bidiagonalization. The RR-Arnoldi process is used to reduce $\bar{A}$ to a small upper Hessenberg matrix; see Section 3. Advantages of RR-Arnoldi-based Tikhonov regularization, compared with Lanczos bidiagonalization-based Tikhonov regularization, include:

i) Many problems require a smaller number of matrix-vector product evaluations. When the matrix A is large, these evaluations constitute the dominant computational work; see [23] for a comparison.
ii) The methods do not require the evaluation of matrix-vector products with A^T. They therefore are attractive to use for problems for which matrix-vector products with A^T are difficult to evaluate. This situation arises, e.g., when solving large nonlinear problems by Krylov subspace methods; see [11] for a discussion. It also arises when matrix-vector products are evaluated by multi-pole methods.
iii) The methods deliver more accurate approximations of the desired solution $\hat{\boldsymbol{x}}$ for some problems; see [23] for a few illustrations.

A difficulty with solution methods based on the RR-Arnoldi process is that they require the regularization operator L to be represented by a square matrix. In particular, the operators (1.9)-(1.11) cannot be applied. Several approaches to circumvent this difficulty have been discussed in the literature. Calvetti et al. [8, 9] propose to append or prepend rows to the regularization operators L_j to yield square and nonsingular regularization operators L; the nonsingularity simplifies the computations. For suitable extensions of the operators (1.9)-(1.11) this approach yields high accuracy. However, to achieve the most accurate approximations of $\hat{\boldsymbol{x}}$, the choice of extension often should depend on the behavior of the solution near the boundary. If this behavior is not known a priori, one can first solve the problem with $L := I$ to determine the boundary behavior of the solution, and then again with a suitably chosen square regularization operator L to achieve higher accuracy.

An alternative to solving the transformed Tikhonov minimization problem (1.16) is to apply an iterative method to the linear system of equations

$$\bar{A}\bar{\boldsymbol{x}} = \bar{\boldsymbol{b}}. \tag{1.19}$$

Regularization is achieved by terminating the iteration process sufficiently early. Denote the kth iterate generated by the iterative method by $\bar{\boldsymbol{x}}_k$, with $\bar{\boldsymbol{x}}_0 = \mathbf{0}$, and define the associated residual vector

$$\bar{\boldsymbol{r}}_k := \bar{\boldsymbol{b}} - \bar{A}\bar{\boldsymbol{x}}_k. \tag{1.20}$$

The iteration number, k, can be thought of as a regularization parameter. Application of the discrepancy principle to determine when to terminate the iterations requires

that the norm of the residual vector $\bar{\boldsymbol{r}}_k$ be available; see Section 3. In this paper, we discuss application of the range restricted GMRES (RR-GMRES) iterative method to the solution of (1.19). RR-GMRES is based on the RR-Arnoldi process; its use and comparison with standard GMRES are illustrated in [5]. RR-GMRES is better suited than standard GMRES for the computation of an approximate solution of linear discrete ill-posed problems, whose desired solution $\hat{\boldsymbol{x}}$ is smooth.

Hansen and Jensen [20] recently described how nonsquare regularization operators can be used in conjunction with RR-GMRES by augmenting the linear system (1.1) and multiplying the augmented system from the left-hand side by a matrix, which contains $(L_A^\dagger)^T$ as a submatrix. This approach is interesting, but it is not well suited for use with the discrepancy principle, because the norm of residual vector (1.20) is not explicitly available.

Morigi et al. [24] proposed to use orthogonal projections as regularization operators. These operators perform well for many linear discrete ill-problems, but they are not smoothing. They are well suited for problems, for which the desired solution $\hat{\boldsymbol{x}}$ is not smooth. The application of regularization operators without any particular structure is considered by Kilmer et al. [21].

The present paper discusses several approaches to determining square regularization operators from regularization operators of the form (1.9)-(1.11). We consider appending zero rows to obtain square regularization operators. Numerical examples illustrate the feasibility of this approach. Problems with a periodic solution may benefit from the use of circulant regularization operators obtained by augmenting suitable rows to regularization operators of the form (1.9)-(1.11). The spectral decomposition of circulant matrices is easily computable, and numerical examples illustrate that it may be beneficial to set certain eigenvalues to zero. The latter observation is believed to be new. Finally, we discuss how these regularization operators can be combined by weighted averaging or with an orthogonal projection to deal with problems with non-periodic solutions.

This paper is organized as follows. Section 2 discusses some properties of square zero-padded and circulant regularization operators obtained by modifying operators of the form (1.9)-(1.11). A weighted average of the two is proposed in order to circumvent the restrictions of each approach. We also describe how orthogonal projection regularization operators can be combined with other regularization operators. The RR-Arnoldi process, the discrepancy principle, and RR-GMRES are reviewed in Section 3. Computed examples are presented in Section 4, and Section 5 contains concluding remarks.

2. Regularization operators. We first discuss general properties of regularization operators $L \in \mathbb{R}^{k\times n}$, $k \leq n$, with a nontrivial null space and then consider square regularization operators related to operators of the form (1.9)-(1.11). This is followed by a discussion on orthogonal projection regularization operators and how they can be combined with square smoothing regularization operators.

Many regularization operators of interest have a nontrivial null space. Let

$$\begin{aligned} \boldsymbol{n}_1 &:= [1, 1, \ldots, 1, 1]^T, \\ \boldsymbol{n}_2 &:= [1, 2, \ldots, n-1, n]^T, \\ \boldsymbol{n}_3 &:= [1, 2^2, \ldots, (n-1)^2, n^2]^T. \end{aligned} \tag{2.1}$$

Then

$$\mathcal{N}(L_1) = \text{span}\{\boldsymbol{n}_1\},$$

$$\mathcal{N}(L_2) = \text{span}\{\boldsymbol{n}_1, \boldsymbol{n}_2\}, \tag{2.2}$$
$$\mathcal{N}(L_3) = \text{span}\{\boldsymbol{n}_1, \boldsymbol{n}_2, \boldsymbol{n}_3\}.$$

Let the columns of the matrix $U \in \mathbb{R}^{n\times\ell}$ form an orthonormal basis of $\mathcal{N}(L)$. Generally, the dimension, ℓ, of the null space of regularization operators of interest is fairly small; for the operators (1.9)-(1.11), we have $\dim\mathcal{N}(L_j) = j$. Introduce the QR-factorization

$$AU = QR, \tag{2.3}$$

where $Q \in \mathbb{R}^{n\times\ell}$ satisfies $Q^TQ = I$ and $R \in \mathbb{R}^{\ell\times\ell}$ is upper triangular.

PROPOSITION 2.1. *Let the columns of $U \in \mathbb{R}^{n\times\ell}$ form an orthonormal basis of the null space of the regularization operator $L \in \mathbb{R}^{k\times n}$, $k \leq n$. Thus, $\ell \geq n - k$. Let the matrices Q and R be determined by (2.3). Then*

$$L_A^\dagger = (I - UR^{-1}Q^TA)L^\dagger, \tag{2.4}$$
$$AL_A^\dagger = (I - QQ^T)AL^\dagger, \tag{2.5}$$
$$\bar{\boldsymbol{b}} = (I - QQ^T)\boldsymbol{b}. \tag{2.6}$$

Proof. The relation (2.4) follows by substituting $I - L^\dagger L = UU^T$ and the QR-factorization (2.3) into (1.12). Multiplication of (2.4) by A yields (2.5). Finally, relation (2.6) is a consequence of $A(A(I - L^\dagger L))^\dagger = QQ^T$. For further details; see [13] or [19, Section 2.3]. □

It is clear from (1.5) that the component of the solution $\boldsymbol{x}_\mu$ in $\mathcal{N}(L)$ is not affected by the value of the regularization parameter μ. The following result is a consequence of this observation.

PROPOSITION 2.2. ([24]) *Let the orthonormal columns of the matrix $U \in \mathbb{R}^{n\times\ell}$ span the null space of the regularization operator $L \in \mathbb{R}^{k\times n}$ and assume that (1.6) holds. Let $\mu > 0$ and let $\boldsymbol{x}_\mu$ be the unique solution (1.7) of the Tikhonov minimization problem (1.5). Consider the discrete ill-posed problem with modified right-hand side*

$$A\boldsymbol{x} = \boldsymbol{b}', \qquad \boldsymbol{b}' := \boldsymbol{b} + AU\boldsymbol{y},$$

for some $\boldsymbol{y} \in \mathbb{R}^\ell$. Then the unique solution of the associated Tikhonov minimization problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^n}\{\|A\boldsymbol{x} - \boldsymbol{b}'\|^2 + \mu\|L\boldsymbol{x}\|^2\}$$

is given by $\boldsymbol{x}'_\mu := \boldsymbol{x}_\mu + U\boldsymbol{y}$.

The above proposition suggests that the regularization operator L should be chosen so that known features of the desired solution $\hat{\boldsymbol{x}}$ can be represented by vectors in $\mathcal{N}(L)$. For instance, when $\hat{\boldsymbol{x}}$ is known to be monotonically increasing and smooth, $L := L_2$ or $L := L_3$ may be suitable regularization operators since their null spaces contain discretizations of linearly and quadratically increasing function; cf. (2.2).

2.1. Square regularization operators by zero-padding. We discuss the use of square regularization operators obtained by zero-padding matrices that represent scaled finite difference operators, such as (1.9)-(1.11). Let $L_{1,0}$ denote the $n\times n$ matrix obtained by appending a zero row-vector to the operator L_1 defined by (1.9). Then

$$L_{1,0}^\dagger = [L_1^\dagger, \boldsymbol{0}] \in \mathbb{R}^{n\times n}. \tag{2.7}$$

Substituting $L := L_{1,0}$ and $L^\dagger := L_{1,0}^\dagger$ into (1.12) yields a square A-weighted pseudoinverse $L_A^\dagger$. The matrix $AL_A^\dagger$ also is square and can be reduced to an upper Hessenberg matrix by the RR-Arnoldi process. Note that for any vector $\boldsymbol{v} \in \mathbb{R}^n$, $L_{1,0}^\dagger \boldsymbol{v}$ can be evaluated efficiently by first computing the QR-factorization of $L_{1,0}^T$ with the aid of Givens rotations, and then solving a least-squares problem. The least-squares solution of minimal-norm easily can be determined, because $\mathcal{N}(L_{1,0})$ is explicitly known; see [13], [19, Section 2.3] for details.

The RR-Arnoldi process may break down when applied to a singular matrix. The matrix (2.5) with $L^\dagger$ defined by (2.7) is singular because $L_A^\dagger \boldsymbol{e}_n = \boldsymbol{0}$; here and elsewhere in this paper $\boldsymbol{e}_j = [0, \ldots, 0, 1, 0, \ldots, 0]^T$ denotes the jth axis vector of appropriate dimension. Moreover, the matrix (2.5) also can be singular because A may be singular. A variant of the Arnoldi process that avoids breakdown when applied to a singular matrix is described in [27] and should be used in black-box software. However, in our experience, it is rare to encounter breakdown or near-breakdown of the RR-Arnoldi process before the computations are terminated by the discrepancy principle. Indeed, we have not encountered near-breakdown in any of our numerical experiments.

The regularization operators (1.10) and (1.11) can be extended to square regularization operators in a similar manner as L_1; the square operator $L_{j,0}$ is defined by appending j zero row-vectors to L_j.

2.2. Square regularization operators by circulant extension. Circulant regularization operators are well suited for use in problems with a periodic solution. We consider the following circulant extensions of the matrices (1.9) and (1.10),

$$C_1 := \frac{1}{2}\begin{bmatrix} 1 & -1 & & & 0 \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ -1 & & & & 1 \end{bmatrix} \in \mathbb{R}^{n\times n}$$

and

$$C_2 := \frac{1}{4}\begin{bmatrix} 2 & -1 & & & -1 \\ -1 & 2 & -1 & & \\ & -1 & \ddots & \ddots & \\ & & \ddots & & -1 \\ -1 & & & -1 & 2 \end{bmatrix} \in \mathbb{R}^{n\times n}. \tag{2.8}$$

Circulants have advantages of being normal matrices with a known spectral decomposition. Introduce the unitary "FFT-matrix" with the columns ordered according to increasing frequency,

$$W = [\boldsymbol{w}_0, \boldsymbol{w}_1, \boldsymbol{w}_{-1}, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_{(-1)^n\lfloor n/2\rfloor}] \in \mathbb{C}^{n\times n},$$
$$\boldsymbol{w}_k := \frac{1}{\sqrt{n}}[1, e^{-2\pi ik/n}, e^{-4\pi ik/n}, e^{-6\pi ik/n}, \ldots, e^{-2\pi i(n-1)k/n}]^T \in \mathbb{C}^n,$$

where $i := \sqrt{-1}$ and $\lfloor\alpha\rfloor$ denotes the integer part of $\alpha \geq 0$. The columns satisfy $\boldsymbol{w}_k = \bar{\boldsymbol{w}}_{-k}$, where the bar denotes complex conjugation. Each column $\boldsymbol{w}_k$ represents

rotations, with the frequency increasing with $|k|$. The $\boldsymbol{w}_k$ are eigenvectors for all circulants; see, e.g., [12]. In particular, the circulants C_j have the spectral decompositions

$$C_j = W\Lambda_j W^*, \qquad \Lambda_j = \mathrm{diag}[\lambda_0^{(j)}, \lambda_1^{(j)}, \lambda_{-1}^{(j)}, \lambda_2^{(j)}, \ldots, \lambda_{(-1)^n\lfloor n/2\rfloor}^{(j)}], \tag{2.9}$$

where the superscript * denotes transposition and complex conjugation. The eigenvalues satisfy

$$[\lambda_0^{(j)}, \lambda_1^{(j)}, \lambda_{-1}^{(j)}, \lambda_2^{(j)}, \ldots, \lambda_{(-1)^n\lfloor n/2\rfloor}^{(j)}]^T = \sqrt{n}W^* C_j \boldsymbol{e}_1$$

and therefore can be computed efficiently by the fast Fourier transform method.

PROPOSITION 2.3. *The spectral decomposition (2.9) expresses a circulant as the sum of n Hermitian rank-one circulants,*

$$C_j = \sum_k \lambda_k^{(j)} \boldsymbol{w}_k \boldsymbol{w}_k^*, \tag{2.10}$$

where summation is over the index set $\{0, 1, -1, 2, \ldots, (-1)^n\lfloor n/2\rfloor\}$ with n components.

Proof. Straightforward computations show that the rank-one matrices $\boldsymbol{w}_k\boldsymbol{w}_k^*$ are circulants. ∎

The eigenvalues of C_1 are distributed equidistantly on a circle of radius and center $1/2$, with $\lambda_0^{(1)} = 0$. All, or all but one, depending on the parity of n, of the remaining eigenvalues appear in complex conjugate pairs; we have $\lambda_k^{(1)} = \bar{\lambda}_{-k}^{(1)}$. The magnitude of the eigenvalues increases with increasing magnitude of their index. The null space is given by $\mathcal{N}(C_1) = \mathrm{span}\{\boldsymbol{n}_1\}$, where $\boldsymbol{n}_1$ is defined by (2.1).

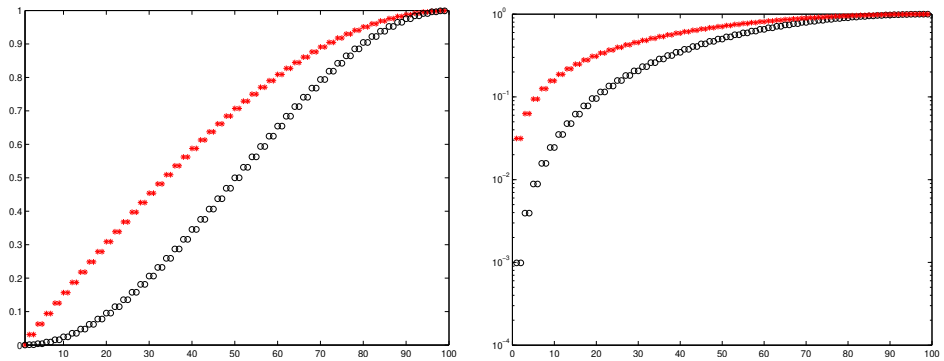


FIG. 2.1. *(a) Magnitude of the eigenvalues of the circulant $C_1 \in \mathbb{R}^{100\times 100}$ (red $*$) and eigenvalues of the circulant $C_2 \in \mathbb{R}^{100\times 100}$ (black o). The eigenvalues are ordered according to increasing index magnitude. (b) Logarithmic scale with the eigenvalues $\lambda_0^{(j)} = 0$, $j = 1, 2$, omitted.*

The matrix C_2 is the symmetric part of C_1. Its eigenvalues are the real part of the eigenvalues of C_1. They live in the interval $[0, 1]$ with $\lambda_0^{(2)} = 0$. All, or all but one, depending on the parity of n, of the remaining eigenvalues appear pairwise; we have $\lambda_k^{(2)} = \lambda_{-k}^{(2)}$. The eigenvalues increase with increasing magnitude of their index. Moreover, $\mathcal{N}(C_2) = \mathcal{N}(C_1)$. Figure 2.1(a) displays the magnitude of the eigenvalues of circulants C_1 and C_2 of order $n = 100$.

Since the $n \times n$ matrix C_2 is real and symmetric, it has a real orthonormal eigenvector basis, such as

$$\left\{ \boldsymbol{w}_0, \frac{1}{\sqrt{2}}(\boldsymbol{w}_1 + \boldsymbol{w}_{-1}), \frac{i}{\sqrt{2}}(\boldsymbol{w}_1 - \boldsymbol{w}_{-1}), \frac{1}{\sqrt{2}}(\boldsymbol{w}_2 + \boldsymbol{w}_{-2}), \dots \right\}.$$

For $n > 2$, real linear combinations of these basis-vectors do not yield a real circulant eigenvector matrix.

The Moore-Penrose pseudoinverse of C_j is given by

$$C_j^\dagger = W \Lambda_j^\dagger W^*, \quad \Lambda_j^\dagger = \mathrm{diag}[(\lambda_0^{(j)})^\dagger, (\lambda_1^{(j)})^\dagger, (\lambda_{-1}^{(j)})^\dagger, \dots, (\lambda_{(-1)^n \lfloor n/2 \rfloor}^{(j)})^\dagger],$$

where

$$(\lambda_k^{(j)})^\dagger := \begin{cases} 1/\lambda_k^{(j)}, & \text{if } \lambda_k^{(j)} \neq 0, \\ 0, & \text{if } \lambda_k^{(j)} = 0. \end{cases}$$

The smoothing effect of the circulant regularization operators C_j stems from the fact that the seminorm $\|C_j \boldsymbol{x}\|$ weighs high-frequency components of the vector $\boldsymbol{x}$, when expanded in terms of the eigenvectors $\boldsymbol{w}_k$, more than low-frequency components. Recall that the frequency as well as the magnitude of the eigenvalues increase with the magnitude of the eigenvector index. Figure 2.1(a) shows that C_1 weighs low-frequency components of $\boldsymbol{x}$ more than C_2.

Figure 2.1(b) differs from Figure 2.1(a) in the scaling of the vertical axis and in that the zero eigenvalues are not displayed. The figure provides better resolution for eigenvalues of small magnitude. In particular, Figure 2.1(b) shows that while both circulants C_1 and C_2 weigh high frequency components the most, they also provide some weighting of low frequencies. The latter may not be beneficial in situations when the desired solution is periodic with low frequency. We may want to modify the circulants C_j by setting one or more of its eigenvalues to zero and thereby avoid damping of the associated frequencies; cf. Proposition 2.2. The benefit of this is illustrated in Section 4. Here we show that the modified matrix obtained is a real circulant.

COROLLARY 2.4. *Let the matrix $\hat{C}_2$ be obtained by setting the p pairs of eigenvalues $\{\lambda_1^{(2)}, \lambda_{-1}^{(2)}\}, \{\lambda_2^{(2)}, \lambda_{-2}^{(2)}\}, \dots, \{\lambda_p^{(2)}, \lambda_{-p}^{(2)}\}$ in the decomposition (2.10) of C_2 to zero, where $p < n/2$. The matrix $\hat{C}_2$ so obtained is a real circulant, and so is its Moore-Penrose pseudoinverse $\hat{C}_2^\dagger$.*

Proof. Let $1 \leq k \leq p$. Then $\lambda_{-k}^{(2)} = \lambda_k^{(2)}$. Setting $\lambda_{-k}^{(2)}$ and $\lambda_k^{(2)}$ to zero for $1 \leq k \leq p$ yields the matrix

$$\hat{C}_2 := C_2 - \sum_{k=1}^{p} \left(\lambda_k^{(2)} \boldsymbol{w}_k \boldsymbol{w}_k^* + \lambda_{-k}^{(2)} \boldsymbol{w}_{-k} \boldsymbol{w}_{-k}^* \right) = C_2 - \sum_{k=1}^{p} \lambda_k^{(2)} \left(\boldsymbol{w}_k \boldsymbol{w}_k^* + \bar{\boldsymbol{w}}_k \bar{\boldsymbol{w}}_k^* \right).$$

The right-hand side shows that $\hat{C}_2$ is real. It is easy to verify that $\boldsymbol{w}_k \boldsymbol{w}_k^*$ is a circulant and so is $\bar{\boldsymbol{w}}_k \bar{\boldsymbol{w}}_k^*$. Therefore $\hat{C}_2$ is a circulant. It follows that $\hat{C}_2^\dagger$ also is a real circulant. ∎

We have discussed two circulant regularization operators. Other circulants can be investigated similarly and applied as regularization operators when appropriate.

2.3. Weighted average of circulant and zero-padded regularization operators. The null space of the regularization operator L_2, defined by (1.10), contains the discretization of linear functions; however, the circulant extension, C_2, does not. While the circulant C_2, and its modification $\hat{C}_2$ described in Corollary 2.4, are well suited for the solution of problems with a periodic solution, they damp non-periodic solution-components, such as linear growth. We are interested in deriving regularization operators that neither damp slowly oscillatory solution components nor linear growth. Namely, the null space of the regularization operator should contain discretizations of both linear and slowly oscillatory functions, or at least the complement of the orthogonal projection of these functions into the null space should be small. This section describes how regularization operators with this property can be defined.

First consider the regularization operator $L := D_\delta C_2$, where

$$D_\delta := \mathrm{diag}[\delta, 1, \ldots, 1, \delta], \tag{2.11}$$

and $\delta > 0$ is small. Then L is a weighted average of the the circulant operator C_2 and the zero-padded operator that is obtained by prepending and appending a zero row to L_2. Discrete linear functions, represented by $\alpha\boldsymbol{n}_1 + \beta\boldsymbol{n}_2$ for $\alpha, \beta \in \mathbb{R}$, are approximately in the null space of L, in the sense that they are hardly damped by L in the context of Tikhonov regularization. Unfortunately, this operator L is computationally difficult to use in large-scale problems as a formula for $L^\dagger$ that is inexpensive to evaluate is not obvious; see also Section 2.4 for further comments. However, $C_2^\dagger D_\delta^{-1}$ may be considered a rough approximation of $L^\dagger$. This leads us to consider the regularization operator

$$L := \left(C_2^\dagger D_\delta^{-1}\right)^\dagger,$$

which differs from $D_\delta C_2$, but has some similar desirable properties as we shall discuss below.

More generally, as in Section 2.2, we can set some small eigenvalues of C_2 to zero, and thereby avoid the damping of low frequency solution-components. Recall that the matrix $\hat{C}_2$ is obtained from C_2 by setting the p pairs of eigenvalues $\{\lambda_1^{(2)}, \lambda_{-1}^{(2)}\}, \{\lambda_2^{(2)}, \lambda_{-2}^{(2)}\}, \ldots, \{\lambda_p^{(2)}, \lambda_{-p}^{(2)}\}$ in the decomposition (2.10) to zero, where $0 \le p < n/2$ (with the convention that $\hat{C}_2 = C_2$ if $p = 0$.) We will use the regularization operator defined implicitly by

$$L := \left(\hat{C}_2^\dagger D_\delta^{-1}\right)^\dagger. \tag{2.12}$$

Note that only $L^\dagger = \hat{C}_2^\dagger D_\delta^{-1}$ is needed in computations and is readily available; L itself is not directly used. We show that the operator (2.12) inherits the low-pass property of $\hat{C}_2$ and hardly damps discrete linear functions $\alpha\boldsymbol{n}_1 + \beta\boldsymbol{n}_2$, $\alpha, \beta \in \mathbb{R}$.

THEOREM 2.5. *Let L be defined by (2.12) with D_δ given by (2.11). Then, for any $\boldsymbol{v} \in \mathbb{R}^n$,*

$$\|L\boldsymbol{v}\| \le \|D_\delta \hat{C}_2 \boldsymbol{v}\|.$$

In particular, $L\boldsymbol{w}_k = 0$ for $0 \le |k| \le p$. Thus, $\boldsymbol{n}_1 \in \mathcal{N}(L)$. Moreover,

$$\frac{\|L\boldsymbol{n}_2\|}{\|\boldsymbol{n}_2\|} \le \frac{\delta}{\sqrt{n}} + \lambda_{p+1}^{(2)}.$$

Proof. Let $\boldsymbol{u} := L\boldsymbol{v}$. Then $L^\dagger L\boldsymbol{v} = L^\dagger \boldsymbol{u} = \hat{C}_2^\dagger D_\delta^{-1}\boldsymbol{u}$. Therefore,

$$\hat{C}_2 L^\dagger L\boldsymbol{v} = \hat{C}_2\hat{C}_2^\dagger D_\delta^{-1}\boldsymbol{u} = D_\delta^{-1}\boldsymbol{u} - (I - \hat{C}_2\hat{C}_2^\dagger)D_\delta^{-1}\boldsymbol{u}.$$

Hence,

$$D_\delta\hat{C}_2 L^\dagger L\boldsymbol{v} = \boldsymbol{u} - D_\delta(I - \hat{C}_2\hat{C}_2^\dagger)D_\delta^{-1}\boldsymbol{u}.$$

Since $\mathcal{R}(L^\dagger) = \mathcal{R}(\hat{C}_2^\dagger) = \text{span}\{\boldsymbol{w}_{p+1}, \boldsymbol{w}_{-(p+1)}, \boldsymbol{w}_{p+2}, \boldsymbol{w}_{-(p+2)}, \ldots\}$, we have

$$I - L^\dagger L = \boldsymbol{w}_0\boldsymbol{w}_0^* + \sum_{k=1}^{p}\left(\boldsymbol{w}_k\boldsymbol{w}_k^* + \boldsymbol{w}_{-k}\boldsymbol{w}_{-k}^*\right).$$

It follows from this equation and $\hat{C}_2\boldsymbol{w}_k = 0$ for $0 \le |k| \le p$ that $D_\delta\hat{C}_2 L^\dagger L = D_\delta\hat{C}_2$. Writing $\boldsymbol{u}_0 := D_\delta(I - \hat{C}_2\hat{C}_2^\dagger)D_\delta^{-1}\boldsymbol{u}$, we have

$$\boldsymbol{u} - \boldsymbol{u}_0 = D_\delta\hat{C}_2\boldsymbol{v}.$$

Now, $L^\dagger\boldsymbol{u}_0 = \hat{C}_2^\dagger(I - \hat{C}_2\hat{C}_2^\dagger)D_\delta^{-1}\boldsymbol{u} = \boldsymbol{0}$, i.e., $\boldsymbol{u}_0 \in \mathcal{N}(L^\dagger)$. Therefore, $\boldsymbol{u}_0 \perp \boldsymbol{u} \in \mathcal{R}(L)$, and it follows that

$$\|\boldsymbol{u}\| \le \|\boldsymbol{u} - \boldsymbol{u}_0\| = \|D_\delta\hat{C}_2\boldsymbol{v}\|.$$

This inequality and $\hat{C}_2\boldsymbol{w}_k = \boldsymbol{0}$ show that $L\boldsymbol{w}_k = \boldsymbol{0}$ for $0 \le |k| \le p$. Furthermore, since $\|\hat{C}_2 - C_2\| \le \lambda_{p+1}^{(2)}$ and $C_2\boldsymbol{n}_2 = \frac{n}{4}[-1, 0, \ldots, 0, 1]^T$, we have

$$\|L\boldsymbol{n}_2\| \le \|D_\delta\hat{C}_2\boldsymbol{n}_2\| \le \|D_\delta C_2\boldsymbol{n}_2\| + \lambda_{p+1}^{(2)}\|\boldsymbol{n}_2\| \le \frac{\sqrt{2}n}{4}\delta + \lambda_{p+1}^{(2)}\|\boldsymbol{n}_2\|.$$

The bound for $\|L\boldsymbol{n}_2\|/\|\boldsymbol{n}_2\|$ now follows from $\|\boldsymbol{n}_2\| = \sqrt{n(n+1)(2n+1)/6} \ge \sqrt{n^3/3}$. ∎

2.4. Orthogonal projection regularization operators. Let the matrix $U \in \mathbb{R}^{n\times\ell}$, $\ell \ll n$, have orthonormal columns. Then the regularization operator

$$L := I - UU^T \tag{2.13}$$

is an orthogonal projector onto $\mathcal{R}^\perp(U)$. Application of orthogonal projection regularization operators is discussed in [24]. This section describes how to combine regularization by orthogonal projection and smoothing operators. This kind of regularization yields the smoothing properties of the latter and allows the choice of a sufficiently large null space to represent important features of the desired solution $\hat{\boldsymbol{x}}$.

While the regularization operator introduced in Section 2.3 has some of these properties, its null space only approximately contains discretizations of linear functions and it requires n to be large. In addition, there are situations where it might be desirable to include other solution components, such as a quadratic or exponential growth functions, in the null space. This can be achieved by combining an orthogonal projection regularization operator with a circulant smoothing regularization operator.

We first review relevant properties of orthogonal projection regularization operators and discuss the equivalence of these operators to a decomposition method

described in [2]. The latter provides a convenient way of implementing regularization with two regularization operators for large-scale problems.

PROPOSITION 2.6. *Let the regularization operator $L \in \mathbb{R}^{n\times n}$ be defined by (2.13) and let the matrices Q and R be determined by (2.3). Then*

$$L_A^\dagger = I - UR^{-1}Q^TA, \tag{2.14}$$

$$AL_A^\dagger = (I - QQ^T)A. \tag{2.15}$$

Proof. A proof follows from Proposition 2.1 and straightforward computations. Details and related results can be found in [24]. □

For small to medium-sized problems, for which the GSVD can be applied, we may choose regularization operators of the form SL, where S is a smoothing regularization operator and L is an orthogonal projector (2.13) with a suitably chosen null space. However, products of regularization operators are difficult to handle in large-scale problems that are transformed to standard form (1.16). The reason for this is that one of the factors of $L_A^\dagger$ is the pseudoinverse of the regularization operator, cf. (2.4), and for most pairs $\{S, L\}$ of regularization operators of interest, the pseudoinverse $(SL)^\dagger$ is difficult to evaluate efficiently; see, e.g., [10, Section 1.4] and [16] for discussions on the pseudoinverse of a product of matrices. We therefore propose to apply pairs of regularization operators sequentially. This can be carried out in a conceptually simple manner by partitioning the linear system of equations (1.1) as follows. Introduce the orthogonal projectors

$$P_U := UU^T, \qquad P_U^\perp := I - UU^T, \qquad P_Q := QQ^T, \qquad P_Q^\perp := I - QQ^T,$$

and partition the vector $\boldsymbol{x} = P_U\boldsymbol{x} + P_U^\perp\boldsymbol{x}$ and the linear system of equations (1.1) according to

$$P_QAP_U\boldsymbol{x} + P_QAP_U^\perp\boldsymbol{x} = P_Q\boldsymbol{b}, \tag{2.16}$$

$$P_Q^\perp AP_U\boldsymbol{x} + P_Q^\perp AP_U^\perp\boldsymbol{x} = P_Q^\perp\boldsymbol{b}. \tag{2.17}$$

Application of partitioning to the solution of linear discrete ill-posed problems is discussed in [2].

THEOREM 2.7. *Let L be defined by (2.13) and let Q be given by (2.3). Then equation (2.16) is equivalent to (1.17), and equation (2.17) is equivalent to (1.19).*

Proof. We first establish the equivalence of (2.17) and (1.19). Note that

$$P_Q^\perp AP_U = 0. \tag{2.18}$$

It follows from (2.15) and (2.6) that

$$\bar{A} = P_Q^\perp A = P_Q^\perp AP_U^\perp, \qquad P_Q^\perp\boldsymbol{b} = \bar{\boldsymbol{b}}.$$

This shows the equivalence.

We turn to (2.16). Denote the computed approximate solution of (2.17) by $\widetilde{\boldsymbol{x}}$. In view of (2.18), the system of equations (2.16)-(2.17) is block upper triangular and $P_U\boldsymbol{x}$ satisfies

$$P_QAP_U\boldsymbol{x} = P_Q\boldsymbol{b} - P_QAP_U^\perp\widetilde{\boldsymbol{x}},$$

which can be written as

$$RU^T\boldsymbol{x} = Q^T\boldsymbol{b} - (Q^TA\widetilde{\boldsymbol{x}} - RU^T\widetilde{\boldsymbol{x}}),$$

i.e.,

$$U^T(\boldsymbol{x} - \widetilde{\boldsymbol{x}}) = R^{-1}Q^T(\boldsymbol{b} - A\widetilde{\boldsymbol{x}}).$$

We may choose

$$\boldsymbol{x} := \bar{\boldsymbol{x}} + UR^{-1}Q^T(\boldsymbol{b} - A\bar{\boldsymbol{x}}). \tag{2.19}$$

Note that (1.14) can be expressed as $\boldsymbol{x}^{(0)} = UR^{-1}Q^T\boldsymbol{b}$. Substituting this expression and (2.14) into $\boldsymbol{x} = L_A^\dagger\widetilde{\boldsymbol{x}} + \boldsymbol{x}^{(0)}$, which is analogous to (1.17), yields

$$\boldsymbol{x} = (I - UR^{-1}Q^TA)\widetilde{\boldsymbol{x}} + UR^{-1}Q^T\boldsymbol{b}.$$

This expression is equivalent to (2.19), which shows the theorem. □

Theorem 2.7 shows that partitioning is equivalent to the application of an orthogonal projection regularization operator, and that the partitioned system (2.16)-(2.17) is block upper triangular. We propose to apply an initial partitioning in situations when the null space of an otherwise suitable smoothing regularization operators is not large enough to be able to represent pertinent features of the desired solution $\hat{\boldsymbol{x}}$. Having chosen a suitable orthogonal projection regularization operator (2.13), and therefore the matrix U, we solve (2.17) using a smoothing regularization operator. Note that in view of (2.18) and $P_Q^\perp AP_U^\perp = P_Q^\perp A$, equation (2.17) simplifies to

$$P_Q^\perp A\boldsymbol{x} = P_Q^\perp\boldsymbol{b}. \tag{2.20}$$

For example, if $\hat{\boldsymbol{x}}$ has a linear growth function component, then the matrix U that defines the orthogonal projection regularization operator (2.13) should be chosen so that $\mathcal{R}(U) = \text{span}\{\boldsymbol{n}_1, \boldsymbol{n}_2\}$. We then may select the circulant

$$\hat{C}_2 := C_2 - \lambda_1^{(2)}(\boldsymbol{w}_1\boldsymbol{w}_1^* + \bar{\boldsymbol{w}}_1\bar{\boldsymbol{w}}_1^*),$$

whose null space contains slowly oscillating discrete functions, as smoothing regularization operator. Computations with this combination of regularization operators are presented in Example 4.5 in Section 4.

The following result shows that in order to determine the residual error norm associated with (1.1), it suffices to compute the norm of the residual error of the reduced system (2.20).

COROLLARY 2.8. *Let $\widetilde{\boldsymbol{x}} \in \mathbb{R}^n$ and define the residual vector*

$$\widetilde{\boldsymbol{r}} := P_Q^\perp\boldsymbol{b} - P_Q^\perp A\widetilde{\boldsymbol{x}}.$$

Let $\boldsymbol{x}$ be the associated approximate solution of (1.1) defined by (2.19) in the proof of Theorem 2.7, and let

$$\boldsymbol{r} := \boldsymbol{b} - A\boldsymbol{x}.$$

Then $\|\boldsymbol{r}\| = \|\widetilde{\boldsymbol{r}}\|$.

Proof. The result follows from (1.18) by using the equivalence of partitioning and regularization by an orthogonal projection operator. The corollary also can be

shown directly by using the partitioning (2.16)-(2.17); see [2] for details on the latter approach. □

We remark that instead of partitioning (1.1) and regularizing (2.20), we can carry out repeated transformations of the Tikhonov minimization problem (1.5) to standard form. The latter approach is more general. However, since we are not exploiting this generality, we will not pursue this approach further in the present paper. Note that the regularization technique of this section also can be used with nonsquare smoothing regularization operators.

3. Krylov subspace methods and the discrepancy principle. Application of $k < n$ steps of the range restricted Arnoldi process to the matrix $\bar{A} \in \mathbb{R}^{n\times n}$ in (1.16) or (1.19) with initial vector $\bar{\boldsymbol{b}}$ determines the decomposition

$$\bar{A}\bar{V}_k = \bar{V}_{k+1}\bar{H}_{k+1,k}, \tag{3.1}$$

where the orthonormal columns of the matrix $\bar{V}_{k+1} \in \mathbb{R}^{n\times(k+1)}$ form a basis of the Krylov subspace

$$\mathcal{K}_{k+1}(\bar{A}, \bar{\boldsymbol{b}}) := \text{span}\{\bar{A}\bar{\boldsymbol{b}}, \bar{A}^2\bar{\boldsymbol{b}}, \ldots, \bar{A}^{k+1}\bar{\boldsymbol{b}}\} \tag{3.2}$$

and the matrix $\bar{H}_{k+1,k} \in \mathbb{R}^{(k+1)\times k}$ is of upper Hessenberg form. The matrix $\bar{V}_k$ consists of the first k columns of $\bar{V}_{k+1}$.

The decomposition (3.1) is the basis for the Arnoldi-Tikhonov method for the solution of (1.16), as well as for the RR-GMRES iterative method for the solution of the linear systems of equations (1.19). We first outline the latter scheme.

The kth iterate, $\bar{\boldsymbol{x}}_k$, determined by RR-GMRES when applied to the solution of (1.19) satisfies

$$\|\bar{A}\bar{\boldsymbol{x}}_k - \bar{\boldsymbol{b}}\| = \min_{\bar{\boldsymbol{x}}\in\mathcal{K}_k(\bar{A},\bar{A}\bar{\boldsymbol{b}})} \|\bar{A}\bar{\boldsymbol{x}} - \bar{\boldsymbol{b}}\|, \qquad \bar{\boldsymbol{x}}_k \in \mathcal{K}_k(\bar{A}, \bar{A}\bar{\boldsymbol{b}}), \tag{3.3}$$

and is computed by first finding the solution $\bar{\boldsymbol{y}}_k$ of the minimization problem

$$\min_{\bar{\boldsymbol{y}}\in\mathbb{R}^k} \|\bar{H}_{k+1,k}\bar{\boldsymbol{y}} - \bar{V}_{k+1}^T\bar{\boldsymbol{b}}\|$$

and then setting $\bar{\boldsymbol{x}}_k := \bar{V}_k\bar{\boldsymbol{y}}_k$. The associated approximate solution of (1.1) is given by

$$\boldsymbol{x}_k := L_A^\dagger\bar{\boldsymbol{x}}_k + \boldsymbol{x}^{(0)};$$

cf. (1.17). The residual error $\boldsymbol{r}_k := \boldsymbol{b} - A\boldsymbol{x}_k$ associated with $\boldsymbol{x}_k$ can be expressed as

$$\boldsymbol{r}_k = \bar{\boldsymbol{b}} - \bar{A}\bar{\boldsymbol{x}}_k = \bar{V}_{k+1}(\bar{H}_{k+1,k}\bar{\boldsymbol{y}}_k - \bar{V}_{k+1}^T\bar{\boldsymbol{b}}) - (I - \bar{V}_{k+1}\bar{V}_{k+1}^T)\bar{\boldsymbol{b}},$$

and it follows that

$$\|\boldsymbol{r}_k\|^2 = (\boldsymbol{e}_{k+1}^T\bar{Q}_{k+1}^T\bar{V}_{k+1}^T\bar{\boldsymbol{b}})^2 + \|\bar{\boldsymbol{b}} - \bar{V}_{k+1}\bar{V}_{k+1}^T\bar{\boldsymbol{b}}\|^2,$$

where $\bar{Q}_{k+1} \in \mathbb{R}^{(k+1)\times(k+1)}$ is the orthogonal matrix in a QR-factorization of $\bar{H}_{k+1}$. The squared residual error norm $\|\boldsymbol{r}_k\|^2$ can be evaluated inexpensively by updating $\|\boldsymbol{r}_{k-1}\|^2$.

Let η be a constant strictly larger than one. The iterate $\boldsymbol{x}_k$ is said to satisfy the discrepancy principle if

$$\|\boldsymbol{r}_k\| \le \eta\varepsilon, \tag{3.4}$$

where ε is the norm of the noise (1.3) or an estimate thereof. If the available estimate is known to be accurate, then we choose η close to unity. It follows from (3.3) and $\mathcal{K}_{k-1}(\bar{A}, \bar{A}\bar{\boldsymbol{b}}) \subset \mathcal{K}_k(\bar{A}, \bar{A}\bar{\boldsymbol{b}})$ that $\|\boldsymbol{r}_{k-1}\| \leq \|\boldsymbol{r}_k\|$ for $k = 1, 2, 3, \ldots$, where $\boldsymbol{r}_0 := \boldsymbol{b} - A\boldsymbol{x}^{(0)}$. We terminate the iterations as soon as an iterate $\boldsymbol{x}_{\hat{k}}$ has been determined that satisfies the discrepancy principle. Thus, $\boldsymbol{r}_{\hat{k}}$ is the first residual vector, such that (3.4) holds. An analysis of this stopping rule for standard GMRES is provided in [6]; the analysis there carries over to RR-GMRES. The evaluation of $\boldsymbol{x}_{\hat{k}}$ requires $\hat{k}+\ell+1$ matrix-vector product evaluations with the matrix A, with ℓ of which required to compute the QR-factorization (2.3). For large-scale problems, the matrix-vector product evaluations dominate the computational work. We therefore tabulate this number in the computed examples of Section 4.

We turn to Tikhonov regularization. Let $\hat{k}$ be defined as above. In the numerical examples of Section 4, we carry out k Arnoldi steps, where $k := \hat{k}$ or $k := \hat{k} + 1$. Substituting the decomposition (3.1) into (1.16) yields a minimization problem of fairly small size, whose solution, $\bar{\boldsymbol{x}}_{k,\mu}$, we compute. Let $\boldsymbol{x}_{k,\mu} := L_A^\dagger \bar{\boldsymbol{x}}_{k,\mu} + \boldsymbol{x}^{(0)}$. We choose μ so that, analogously to (3.4), the associated residual error $\boldsymbol{r}_{k,\mu} := \boldsymbol{b} - A\boldsymbol{x}_{k,\mu}$ satisfies $\|\boldsymbol{r}_{k,\mu}\| = \eta\varepsilon$; see [23] for further details.

4. Computed examples. We illustrate the performance of the regularization operators discussed with some numerical examples. The "noise-vector" $\boldsymbol{e}$ has in all examples normally distributed pseudorandom entries with mean zero, and is normalized to correspond to a chosen noise-level

$$\nu := \frac{\|\boldsymbol{e}\|}{\|\hat{\boldsymbol{b}}\|}.$$

Here $\hat{\boldsymbol{b}}$ denotes the noise-free right-hand side vector in (1.4). We let $\eta := 1.01$ in (3.4) in all examples. The computations are carried out in MATLAB with about 16 significant decimal digits. We reorthogonalize the columns of the matrix $\bar{V}_{k+1}$ in the decomposition (3.1). Since typically the dimension $k+1$ of the Krylov subspaces (3.2) used is fairly small, reorthogonalization is inexpensive. Reorthogonalization may reduce the size of the Krylov subspace required for determining an approximate solution that satisfies the discrepancy principle; see [26, Example 4.1] for an illustration with a symmetric matrix.

reg. oper.	# iterations k	# mat.-vec. prod.	$\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\|/\|\hat{\boldsymbol{x}}\|$
I	3	4	$1.6 \cdot 10^{-3}$
$L_{1,0}$	2	4	$2.2 \cdot 10^{-4}$
$L_{2,0}$	1	4	$2.0 \cdot 10^{-4}$
$L_{3,0}$	0	3	$1.7 \cdot 10^{-4}$

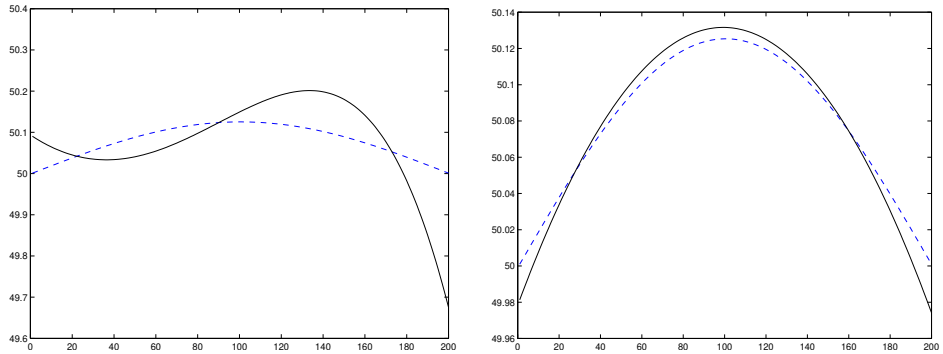
TABLE 4.1

Example 4.1: Relative error in approximate solutions $\boldsymbol{x}_k$ determined by truncated iteration with RR-GMRES.

Example 4.1. The Fredholm integral equation of the first kind,

$$\int_0^{\pi/2} \kappa(\sigma,\tau)x(\sigma)d\sigma = b(\tau), \qquad 0 \leq \tau \leq \pi, \tag{4.1}$$

with $\kappa(\sigma,\tau) := \exp(\sigma\cos(\tau))$, $b(\tau) := 2\sinh(\tau)/\tau$, and solution $x(\tau) := \sin(\tau)$ is discussed by Baart [1]. We use the MATLAB code baart from [18] to discretize (4.1)



(a) (b)

FIG. 4.1. *Example 4.1: Computed approximate solutions $\boldsymbol{x}_k$ determined by RR-GMRES using the discrepancy principle. The dashed curves show the vector $\hat{\boldsymbol{x}}$; the continuous curves display in (a) the iterate $\boldsymbol{x}_3$ determined without regularization operator ($L := I$), and in (b) the iterate $\boldsymbol{x}_0$ determined with the regularization operator $L := L_{3,0}$.*

by a Galerkin method with 200 orthonormal box functions as test and trial functions. The code produces the matrix $A \in \mathbb{R}^{200\times 200}$ and a scaled discrete approximation of $x(\tau)$. Adding $50\boldsymbol{n}_1$ to the latter yields the vector $\hat{\boldsymbol{x}} \in \mathbb{R}^{200}$ with which we compute the noise-free right-hand side $\hat{\boldsymbol{b}} := A\hat{\boldsymbol{x}}$.

Let the entries of the error vector $\boldsymbol{e} \in \mathbb{R}^{200}$ be normally distributed with zero mean, and be normalized to yield the noise-level $\nu = 5 \cdot 10^{-5}$. This corresponds to an absolute error of $\|\boldsymbol{e}\| = 9.8 \cdot 10^{-2}$. The right-hand side $\boldsymbol{b}$ in the system (1.1) is obtained from (1.2).

Table 4.1 displays results obtained with RR-GMRES for several regularization operators and Figure 4.1 shows two computed solutions. The continuous curve in Figure 4.1(a) displays the computed solution obtained without explicit use of a regularization operator ($L := I$); the continuous curve in Figure 4.1(b) depicts the corresponding solution determined with $L := L_{3,0}$. The iterations are terminated by the discrepancy principle (3.4). The dashed curves in Figures 4.1(a) and (b) show the solution $\hat{\boldsymbol{x}}$ of the noise-free problem (1.4).

This example shows RR-GMRES without explicit use of a regularization operator to perform poorly. The reason for this is that the desired solution is a small relative perturbation of the vector $50\boldsymbol{n}_1$. We note that this vector lives in $\mathcal{N}(L_{j,0})$, $1 \leq j \leq 3$, and therefore does not affect the Krylov subspaces for $\bar{A}$ when one of the regularization operators $L_{j,0}$ is applied. When $L := L_{3,0}$, the norm of the initial residual $\boldsymbol{r}_0 := \boldsymbol{b} - A\boldsymbol{x}^{(0)}$ satisfies the discrepancy principle (3.4) and no iterations are carried out. □

method	# iterations k	# mat.-vec. prod.	$\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\|$
Arnoldi-Tikhonov	9	10	$1.5 \cdot 10^{-2}$
Arnoldi-Tikhonov +1	10	11	$1.2 \cdot 10^{-2}$
LBD-Tikhonov	11	22	$1.2 \cdot 10^{-2}$

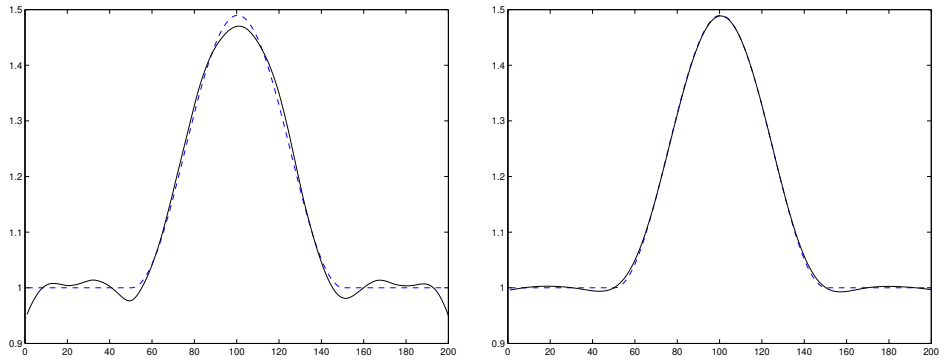
TABLE 4.2

Example 4.2: Errors in approximate solutions of a modification of (4.2) determined by several Tikhonov regularization methods with regularization operator $L := I$.

method	# iterations k	# mat.-vec. prod.	$\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\|$
Arnoldi-Tikhonov	6	8	$5.7 \cdot 10^{-3}$
Arnoldi-Tikhonov +1	7	9	$2.9 \cdot 10^{-3}$
LBD-Tikhonov	7	15	$3.6 \cdot 10^{-3}$

TABLE 4.3

Example 4.2: Errors in approximate solutions of a modification of (4.2) determined by several Tikhonov regularization methods with regularization operator $L_{1,0}$.



(a) (b)

FIG. 4.2. *Example 4.2: Approximate solutions $\boldsymbol{x}_{k,\mu}$ computed by the Arnoldi-Tikhonov method with the regularization parameter μ determined by the discrepancy principle. The dashed curves show the vector $\hat{\boldsymbol{x}}$; the continuous curves display in (a) the approximate solution $\boldsymbol{x}_{10,\mu}$ determined with regularization operator $L := I$ and in (b) the approximate solution $\boldsymbol{x}_{7,\mu}$ determined with $L := L_{1,0}$.*

Example 4.2. Consider the Fredholm integral equation of the first kind

$$\int_{-6}^{6} \kappa(\tau,\sigma)x(\sigma)d\sigma = g(\tau), \qquad -6 \leq \tau \leq 6, \tag{4.2}$$

with kernel and solution given by

$$\kappa(\tau,\sigma) := x(\tau-\sigma),$$
$$x(\sigma) := \begin{cases} 1+\cos(\frac{\pi}{3}\sigma), & \text{if } |\sigma| < 3, \\ 0, & \text{otherwise.} \end{cases}$$

The right-hand side $g(\tau)$ is defined by (4.2). This integral equation is discussed by Phillips [25]. The MATLAB code phillips in [18] determines a discretization by a Galerkin method with orthonormal box functions. A discretization of a scaled solution also is provided. Let the matrix $A \in \mathbb{R}^{200\times 200}$ determined by phillips represent the discretized integral operator, and let $\hat{\boldsymbol{x}}$ be the sum of the scaled discrete solution provided by phillips and the vector $\boldsymbol{n}_1$. The noise-free right-hand side is given by $\hat{\boldsymbol{b}} := A\hat{\boldsymbol{x}}$. A "noise-vector" $\boldsymbol{e}$ similar to the one in Example 4.1 and scaled to correspond to a noise-level of $1 \cdot 10^{-3}$ is added to $\hat{\boldsymbol{b}}$ to yield the right-hand side $\boldsymbol{b}$ of the linear system (1.1).

Tables 4.2 and 4.3 report the performance of several Tikhonov regularization methods with the regularization operators $L := I$ and $L := L_{1,0}$, respectively. The latter regularization operator gives more accurate results for the present problem. For many ill-posed problems the Arnoldi-Tikhonov method yields higher accuracy by

applying one more step of the Arnoldi process than the $\hat{k}$ steps necessary to satisfy the discrepancy principle; see the last paragraph of Section 3. This approach is denoted by Arnoldi-Tikhonov +1 in the tables. LBD-Tikhonov is the Lanczos bidiagonalization-based Tikhonov regularization method described in [7]. Lanczos bidiagonalization is implemented with reorthogonalization. We remark that the operators $L_{1,0}$ and L_1 are equivalent when applied in LBD-Tikhonov. Figure 4.2 displays the most accurate computed solutions in the Tables 4.2 and 4.3 for $L := I$ and $L := L_{1,0}$.

The tables show the Arnoldi-Tikhonov method to yield about the same accuracy as LBD-Tikhonov and to require fewer matrix-vector product evaluations. We therefore omit graphs for the approximate solutions determined by LBD-Tikhonov. □

reg. oper.	# iterations k	# mat.-vec. prod.	$\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\|/\|\hat{\boldsymbol{x}}\|$
I	2	3	$1.3 \cdot 10^{-1}$
$L_{3,0}$	0	3	$5.0 \cdot 10^{-2}$

TABLE 4.4

Example 4.3: Relative error in approximate solutions $\boldsymbol{x}_k$ determined by truncated iteration with RR-GMRES.

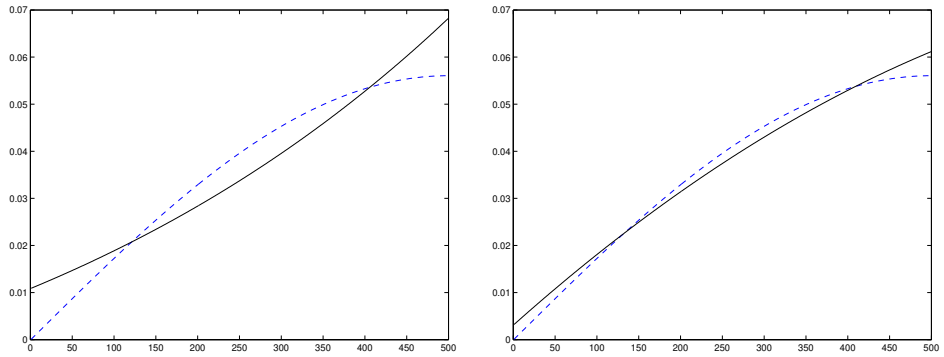


FIG. 4.3. *Example 4.3: Computed approximate solutions $\boldsymbol{x}_k$ determined by RR-GMRES using the discrepancy principle. The dashed curves show the vector $\hat{\boldsymbol{x}}$; the continuous curves display in (a) the iterate $\boldsymbol{x}_2$ determined without regularization operator ($L := I$), and in (b) the iterate $\boldsymbol{x}_0$ determined with the regularization operator $L := L_{3,0}$.*

Example 4.3. RR-GMRES would in Example 4.1 be able to determine a quite accurate approximation of $\hat{\boldsymbol{x}}$ without the use of the regularization operator $L_{1,0}$, if before iterative solution $50A\boldsymbol{n}_1$ is subtracted from the right-hand side $\boldsymbol{b}$, and after iterative solution the vector $50\boldsymbol{n}_1$ is added to the approximate solution determined by RR-GMRES. However, it is not always obvious how a problem can be modified in order for RR-GMRES to be able to achieve high accuracy without applying a regularization operator. This is illustrated by the present example, for which subtracting a multiple of $A\boldsymbol{n}_1$ before iterative solution with RR-GMRES, and adding the same multiple of $\boldsymbol{n}_1$ after iterative solution, does not yield an accurate approximation of $\hat{\boldsymbol{x}}$.

Consider

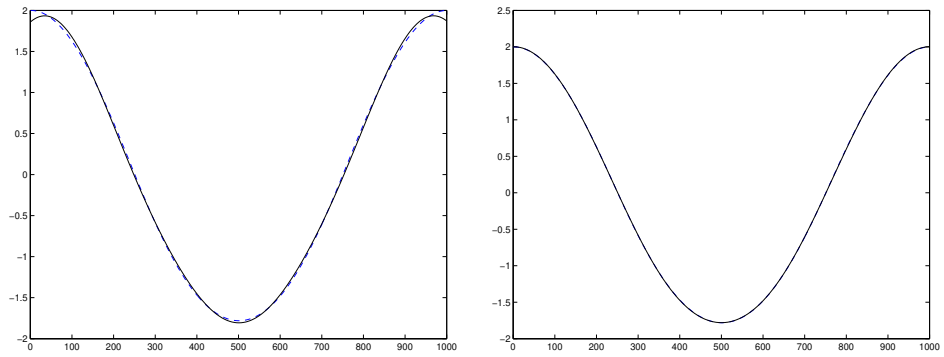
$$\int_0^{\pi/4} \kappa(\sigma,\tau)x(\sigma)d\sigma = b(\tau), \qquad 0 \leq \tau \leq \frac{\pi}{2},$$

where the kernel and the right-hand side are the same as in (4.1). We discretize in a similar fashion as in Example 4.1 to obtain the matrix $A \in \mathbb{R}^{500\times 500}$ and the noise-free right-hand side $\hat{\boldsymbol{b}} \in \mathbb{R}^{500}$. Adding a "noise-vector" $\boldsymbol{e}$ to $\hat{\boldsymbol{b}}$ yields the right-hand side of (1.1). The vector $\boldsymbol{e}$ is scaled to correspond to the noise-level $\nu = 2 \cdot 10^{-3}$. We solve the system (1.1) by RR-GMRES. Table 4.4 reports results for the cases when no regularization operator is used ($L := I$) and when the regularization operator $L := L_{3,0}$ is applied. Figure 4.3 displays the computed solutions. The operator $L_{3,0}$ is seen to improve the quality of the computed solution. □

regularization operator	# iterations k	# mat.-vec. prod.	$\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\|$
-	5	6	$5.7 \cdot 10^{-2}$
$L_{2,0}$	5	8	$2.4 \cdot 10^{-2}$
C_2	3	5	$5.7 \cdot 10^{-3}$
$\hat{C}_2$	1	5	$2.4 \cdot 10^{-3}$

TABLE 4.5

Example 4.4: Errors in approximate solutions of computed by RR-GMRES without and with several regularization operators. The operator $L_{2,0}$ is defined by zero-padding the operator (1.10), C_2 by (2.8), and $\hat{C}_2$ by setting the 2 smallest nonvanishing eigenvalues of C_2 to zero.



(a) (b)

FIG. 4.4. *Example 4.4: Computed approximate solutions $\boldsymbol{x}_k$ by RR-GMRES using the simplified discrepancy principle. The dashed curves depict the vector $\hat{\boldsymbol{x}}$; the continuous curves display in (a) the iterate $\boldsymbol{x}_5$ determined with the regularization operator $L := L_{2,0}$, and in (b) the iterate $\boldsymbol{x}_1$ computed with the regularization operator $L := \hat{C}_2$ defined by setting the two smallest nonvanishing eigenvalues of C_2 to zero.*

Example 4.4. We modify the integral equation of Example 4.2 in two ways: the integral equation is discretized with finer resolution to obtain the matrix $A \in \mathbb{R}^{1000\times 1000}$, and instead of adding a discretization of the function 1 to the scaled discrete solution determined by the MATLAB code phillips, we add a discretization of the function $2\cos(\pi(1+\frac{\sigma}{6}))$. This defines $\hat{\boldsymbol{x}}$. The noise-free right-hand side is given by $\hat{\boldsymbol{b}} := A\hat{\boldsymbol{x}}$. A "noise-vector" $\boldsymbol{e}$ with normally distributed entries with mean zero, and scaled to correspond to the noise-level $1 \cdot 10^{-2}$ is added to $\hat{\boldsymbol{b}}$ to yield the right-hand side $\boldsymbol{b}$ of the linear system (1.1).

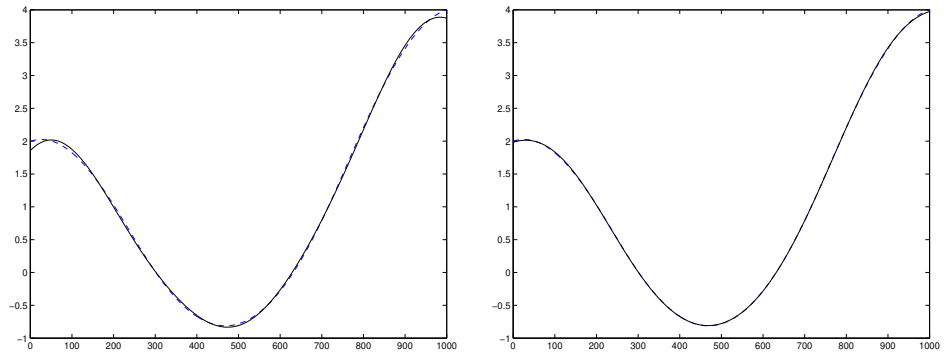
Table 4.5 reports the performance of RR-GMRES without and with several regularization operators. The table shows that the regularization operator $\hat{C}_2$ obtained from (2.8) by setting the 2 smallest nonvanishing eigenvalues, $\lambda_1^{(2)}$ and $\lambda_{-1}^{(2)}$, to zero

yields the best approximation of the desired solution $\hat{\boldsymbol{x}}$. In particular, the regularization operator $\hat{C}_2$ yields higher accuracy than $L_{2,0}$. Table 4.5 also displays the number of iterations and the number of matrix-vector product evaluations required. Figure 4.4 shows $\hat{\boldsymbol{x}}$ (dashed curves) and the computed approximate solutions determined with the regularization operators $L_{2,0}$ and $\hat{C}_2$ (continuous curves). The latter curves differ primarily at their endpoints. □

regularization operator	# iterations k	# mat.-vec. prod.	$\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\|$
-	8	9	$6.0 \cdot 10^{-2}$
$L_{2,0}$	5	8	$1.8 \cdot 10^{-2}$
C_2	7	9	$1.1 \cdot 10^{-1}$
$\hat{C}_2$	3	7	$1.1 \cdot 10^{-1}$
$(\hat{C}_2^\dagger D_\delta^{-1})^\dagger$	2	6	$5.6 \cdot 10^{-3}$
OP and C_2	3	7	$4.3 \cdot 10^{-3}$
OP	4	7	$1.0 \cdot 10^{-2}$

TABLE 4.6

Example 4.5: Errors in approximate solutions computed by RR-GMRES without and with several regularization operators. The operator $L_{2,0}$ is defined by zero-padding the operator (1.10), C_2 by (2.8), and $\hat{C}_2$ by setting the 2 smallest nonvanishing eigenvalues of C_2 to zero. The diagonal matrix D_δ has $\delta = 1 \cdot 10^{-8}$. OP stands for orthogonal projection onto a subspace not containing discretized linear functions, i.e., $\mathcal{R}(U) = span\{\boldsymbol{n}_1, \boldsymbol{n}_2\}$ in (2.13).



(a) (b)

FIG. 4.5. *Example 4.5: Approximate solutions $\boldsymbol{x}_k$ computed by RR-GMRES using the simplified discrepancy principle. The dashed curves depict the vector $\hat{\boldsymbol{x}}$; the continuous curves display in (a) the iterate $\boldsymbol{x}_5$ determined with regularization operator $L := L_{2,0}$, and in (b) the iterate $\boldsymbol{x}_3$ determined by initial orthogonal projection onto the complement of the discretized linear functions, and then solving the projected problem by RR-GMRES with regularization operator C_2.*

Example 4.5. The matrix in this example is the same as in Example 4.4, and we add a discretization of the linear function $\ell(\sigma) := 1 + \sigma/6$ to the vector $\hat{\boldsymbol{x}}$. The noise-free and noise-contaminated right-hand sides are determined similarly as in Example 4.4; the noise-level is $1 \cdot 10^{-2}$.

Table 4.6 reports the performance of RR-GMRES without and with several regularization operators. Because of the oscillatory behavior of the desired solution $\hat{\boldsymbol{x}}$, the regularization operator $L_{2,0}$ does not perform well, and due to the linear term in $\hat{\boldsymbol{x}}$, neither do the operators C_2 and $\hat{C}_2$. The latter operator is the same as in Example 4.4. We therefore consider the approaches of Section 2.3 and 2.4, which yield

clearly better results. For the weighted average regularization operator L in (2.12), the matrix D_δ is given by (2.11) with $\delta := 1 \cdot 10^{-8}$. In the method of Section 2.4, we first carry out an orthogonal projection onto the complement of discrete linear functions. Hence, we let $U \in \mathbb{R}^{1000\times 2}$ in (2.13) be such that $\mathcal{R}(U) = \text{span}\{\boldsymbol{n}_1, \boldsymbol{n}_2\}$. This projection is in Table 4.6 referred to as "OP". RR-GMRES is applied to the projected equations (2.20) with the regularization operators C_2 or without further regularization. Results for the latter approach are displayed in the last line of the table. Table 4.6 shows orthogonal projection followed by regularization with C_2 to give the best approximation of $\hat{\boldsymbol{x}}$. The computed solution is shown in Figure 4.5(b). Also the regularization operator $(\hat{C}_2^\dagger D_\delta^{-1})^\dagger$ is seen to determine an accurate approximation of $\hat{\boldsymbol{x}}$. Table 4.6 also illustrates that the smoothing regularization operator $L_{2,0}$ yields a better approximation than the orthogonal projection regularization operator (OP) with the same null space. This depends on that the desired solution $\hat{\boldsymbol{x}}$ is smooth and $L_{2,0}$ is smoothing, but the orthogonal projector is not. Figure 4.5(a) displays the solution determined with $L_{2,0}$. This computed solution differs from $\hat{\boldsymbol{x}}$ the most at the endpoints. □

5. Conclusion. We have presented several square extensions of some standard regularization operators based on finite difference discretization of derivatives. The numerical examples illustrate that the square regularization operators discussed here can improve the quality of the computed approximate solution determined by Arnoldi-based Tikhonov regularization and minimal residual methods. Moreover, they are quite simple to implement.

REFERENCES

[1] M. L. Baart, *The use of auto-correlation for pseudo-rank determination in noisy ill-conditioned least-squares problems*, IMA J. Numer. Anal., 2 (1982), pp. 241–247.

[2] J. Baglama and L. Reichel, *Decomposition methods for large linear discrete ill-posed problems*, J. Comput. Appl. Math., 198 (2007), pp. 332–342.

[3] Å. Björck, *A bidiagonalization algorithm for solving large sparse ill-posed systems of linear equations*, BIT, 28 (1988), pp. 659–670.

[4] D. Calvetti, G. H. Golub, and L. Reichel, *Estimation of the L-curve via Lanczos bidiagonalization*, BIT, 39 (1999), pp. 603–619.

[5] 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.

[6] D. Calvetti, B. Lewis, and L. Reichel, *On the regularizing properties of the GMRES method*, Numer. Math., 91 (2002), pp. 605–625.

[7] D. Calvetti and L. Reichel, *Tikhonov regularization of large linear problems*, BIT, 43 (2003), pp. 263–283.

[8] D. Calvetti, L. Reichel, and A. Shuibi, *L-curve and curvature bounds for Tikhonov regularization*, Numer. Algorithms, 35 (2004), pp. 301–314.

[9] D. Calvetti, L. Reichel, and A. Shuibi, *Invertible smoothing preconditioners for linear discrete ill-posed problems*, Appl. Numer. Math., 54 (2005), pp. 135–149.

[10] S. L. Campbell and C. D. Meyer, *Generalized Inverses of Linear Transformations*, Dover, Mineola, 1991.

[11] T. F. Chan and K. R. Jackson, *Nonlinearly preconditioned Krylov subspace methods for discrete Newton algorithms*, SIAM J. Sci. Statist. Comput., 5 (1984), pp. 533–542.

[12] P. J. Davis, *Circulant Matrices*, Chelsea, New York, 1994.

[13] L. Eldén, *A weighted pseudoinverse, generalized singular values, and constrained least squares problems*, BIT, 22 (1982), pp. 487–501.

[14] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.

[15] G. H. Golub and U. von Matt, *Tikhonov regularization for large scale problems*, in Workshop on Scientific Computing, eds. G. H. Golub, S. H. Lui, F. Luk, and R. Plemmons, Springer, New York, 1997, pp. 3–26.

[16] T. N. E. Greville, *Note on the generalized inverse of a matrix product*, SIAM Rev., 8 (1966), pp. 518–521.

[17] M. Hanke and P. C. Hansen, *Regularization methods for large-scale problems*, Surv. Math. Ind., 3 (1993), pp. 253–315.

[18] P. C. Hansen, *Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems*, Numer. Algorithms, 6 (1994), pp. 1–35. Software is available in Netlib at the web site `http://www.netlib.org`

[19] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems*, SIAM, Philadelphia, 1998.

[20] P. C. Hansen and T. K. Jensen, *Smoothing norm preconditioning for regularizing minimum residual methods*, SIAM J. Matrix Anal. Appl., 29 (2006), pp. 1–14.

[21] M. E. Kilmer, P. C. Hansen, and M. I. Espanol, *A projection-based approach to general-form Tikhonov regularization*, SIAM J. Sci. Comput., 29 (2007), pp. 315–330.

[22] M. E. Kilmer and D. P. O'Leary, *Choosing regularization parameters in iterative methods for ill-posed problems*, SIAM J. Matrix Anal. Appl., 22 (2001), pp. 1204–1221.

[23] B. Lewis and L. Reichel, *Arnoldi-Tikhonov regularization methods*, J. Comput. Appl. Math., to appear.

[24] S. Morigi, L. Reichel, and F. Sgallari, *Orthogonal projection regularization operators*, Numer. Algorithms, 44 (2007), pp. 99–114.

[25] D. L. Phillips, *A technique for the numerical solution of certain integral equations of the first kind*, J. ACM, 9 (1962), pp. 84–97.

[26] L. Reichel, H. Sadok, and A. Shyshkov, *Greedy Tikhonov regularization for large linear ill-posed problems*, Int. J. Comput. Math., 84 (2007), pp. 1151–1166.

[27] L. Reichel and Q. Ye, *Breakdown-free GMRES for singular systems*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 1001–1021.