SOME MATRIX NEARNESS PROBLEMS SUGGESTED BY TIKHONOV REGULARIZATION

SILVIA NOSCHESE* AND LOTHAR REICHEL†

Abstract. The numerical solution of linear discrete ill-posed problems typically requires regularization, i.e., replacement of the available ill-conditioned problem by a nearby better conditioned one. The most popular regularization methods for problems of small to moderate size are Tikhonov regularization and truncated singular value decomposition (TSVD). By considering matrix nearness problems related to Tikhonov regularization, several novel regularization methods are derived. These methods share properties with both Tikhonov regularization and TSVD, and can give approximate solutions of higher quality than either one of these methods.

Key words. ill-posed problem, Tikhonov regularization, modified Tikhonov regularization, truncated singular value decomposition,

1. Introduction. Consider the computation of an approximate solution of the minimization problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \|A\boldsymbol{x}-\boldsymbol{b}\|, \tag{1.1}$$

where $\|\cdot\|$ denotes the Euclidean vector norm and $A \in \mathbb{R}^{m\times n}$ is a matrix whose singular values decay smoothly to zero without a significant gap. In particular, A may be singular. Minimization problems (1.1) with a matrix of this kind often are referred to as discrete ill-posed problems. They arise, for example, from the discretization of linear ill-posed problems, such as Fredholm integral equations of the first kind with a smooth kernel. We will for notational simplicity assume that $m \geq n$; however, the methods discussed also can be applied when $m < n$.

The data vector $\boldsymbol{b} \in \mathbb{R}^m$ in linear discrete ill-posed problems that arise in science and engineering typically is contaminated by an (unknown) error $\boldsymbol{e} \in \mathbb{R}^m$. We will refer to the error $\boldsymbol{e}$ as "noise." Let $\hat{\boldsymbol{b}} \in \mathbb{R}^m$ denote the (unknown) error-free vector associated with $\boldsymbol{b}$, i.e.,

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

The (unknown) linear system of equations with error-free right-hand side,

$$A\boldsymbol{x} = \hat{\boldsymbol{b}}, \tag{1.3}$$

is assumed to be consistent; however, we do not require the least-squares problem (1.1) to be consistent.

Let $A^\dagger$ denote the Moore–Penrose pseudoinverse of A. We are interested in computing an approximation of the solution $\hat{\boldsymbol{x}} = A^\dagger\hat{\boldsymbol{b}}$ of minimal Euclidean norm of the error-free linear system (1.3) by determining an approximate solution of the error-contaminated least-squares problem (1.1). Note that the solution of (1.1),

$$\boldsymbol{x} = A^\dagger\boldsymbol{b} = A^\dagger(\hat{\boldsymbol{b}}+\boldsymbol{e}) = \hat{\boldsymbol{x}} + A^\dagger\boldsymbol{e}, \tag{1.4}$$

*Dipartimento di Matematica "Guido Castelnuovo", SAPIENZA Università di Roma, P.le A. Moro, 2, I-00185 Roma, Italy. E-mail: noschese@mat.uniroma1.it. Research supported by a grant from SAPIENZA Università di Roma.

†Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA. E-mail: reichel@math.kent.edu. Research supported in part by NSF grant DMS-1115385.

typically is dominated by the propagated error $A^{\dagger}\boldsymbol{e}$ and then is meaningless.

Tikhonov regularization, in its simplest form, seeks to determine a useful approximation of $\hat{\boldsymbol{x}}$ by replacing the minimization problem (1.1) by the penalized least-squares problem

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

The scalar $\mu > 0$ is a regularization parameter. We are interested in developing modifications of this minimization problem by considering certain matrix nearness problems.

Solving (1.5) requires both the determination of a suitable value of $\mu > 0$ and the computation of the associated solution

$$\boldsymbol{x}_{\mu} = (A^TA+\mu^2I)^{-1}A^T\boldsymbol{b} \tag{1.6}$$

of (1.5). Throughout this paper the superscript ^{T} denotes transposition and I is the identity matrix of appropriate order. We will assume that a bound for the norm of the error-vector $\boldsymbol{e}$ is known. Then μ can be determined with the aid of the discrepancy principle; see below for details.

Another common regularization method for (1.1) is truncated singular value decomposition (TSVD). In this method the $n-k$ smallest singular values of A are set to zero and the minimal-norm solution of the resulting least-squares problem is computed. The truncation index k is a regularization parameter, which can be determined, e.g., with the discrepancy principle.

The TSVD method generally only dampens high frequencies in the computed solution, while Tikhonov regularization (1.5) dampens all frequencies. A modification of the Tikhonov minimization problem (1.5) that generally only dampens high frequencies has been described in [7]. This modification can be derived as the solution of a matrix nearness problem. It is the purpose of this paper to describe several matrix nearness problems that suggest modifications of the Tikhonov minimization problem (1.5). Some of these modifications perform particularly well for problems (1.1) in which the vector $\boldsymbol{b}$ is contaminated by colored noise dominated by high-frequency components.

This paper is organized as follows. Section 2 reviews TSVD and Tikhonov regularization, as well as the modified Tikhonov regularization method described in [7], and introduces new regularization methods suggested by certain matrix nearness problems. Section 3 presents a few computed examples, and Section 4 contains concluding remarks and discusses some extensions. In particular, the discussion of methods in this paper assumes the singular value decomposition (SVD) of the matrix A to be available. However, it is impractical to compute the SVD of large matrices. We comment in Section 4 on how the methods of this paper can be applied to the solution of large-scale least-squares problems (1.1).

2. Old and new regularization methods. We first discuss the SVD of A, then review regularization by the TSVD and Tikhonov methods, and finally describe several modifications of the Tikhonov minimization problem (1.5). The SVD of A is a factorization of the form

$$A = U\Sigma V^T, \tag{2.1}$$

where $U = [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m] \in \mathbb{R}^{m\times m}$ and $V = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n] \in \mathbb{R}^{n\times n}$ are orthogonal matrices, the superscript ^{T} denotes transposition, and

$$\Sigma = \mathrm{diag}[\sigma_1, \sigma_2, \ldots, \sigma_n] \in \mathbb{R}^{m\times n}$$

is a (possibly rectangular) diagonal matrix, whose diagonal entries $\sigma_j \geq 0$ are the singular values of A. They are ordered according to $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$.

Let A be of rank $\ell \geq 1$. Then (2.1) can be expressed as

$$A = \sum_{j=1}^{\ell} \sigma_j \boldsymbol{u}_j \boldsymbol{v}_j^T \tag{2.2}$$

with $\sigma_\ell > 0$. When the matrix A stems from the discretization of a compact operator, such as a Fredholm integral equation of the first kind with a smooth kernel, the vectors $\boldsymbol{v}_j$ and $\boldsymbol{u}_j$ represent discretizations of singular functions that are defined on the domains of the integral operator and its adjoint, respectively. These singular functions typically oscillate more with increasing index. The representation (2.2) then is a decomposition of A into rank-one matrices $\boldsymbol{u}_j \boldsymbol{v}_j^T$ that are discretizations of products of singular functions that oscillate more with increasing index j.

2.1. Regularization by TSVD. The Moore–Penrose pseudoinverse of A is given by

$$A^\dagger = \sum_{j=1}^{\ell} \sigma_j^{-1} \boldsymbol{v}_j \boldsymbol{u}_j^T.$$

The difficulty of solving (1.1) without regularization stems from the fact that the matrix A has "tiny" positive singular values and the computation of the solution (1.4) of (1.1) involves division by these singular values. This results in severe propagation of the error $\boldsymbol{e}$ in $\boldsymbol{b}$ and of round-off errors introduced during the calculations of the computed approximate solution of (1.1).

Regularization by the TSVD method overcomes this difficulty by ignoring the tiny positive singular values of A. Introduce, for $1 \leq k \leq \ell$, the rank-k approximation of A,

$$A_k = \sum_{j=1}^{k} \sigma_j \boldsymbol{u}_j \boldsymbol{v}_j^T$$

with Moore–Penrose pseudoinverse

$$A_k^\dagger = \sum_{j=1}^{k} \sigma_j^{-1} \boldsymbol{v}_j \boldsymbol{u}_j^T.$$

The TSVD method yields approximate solutions of (1.1) of the form

$$\boldsymbol{x}_k = A_k^\dagger \boldsymbol{b} = \sum_{j=1}^{k} \frac{\boldsymbol{u}_j^T \boldsymbol{b}}{\sigma_j} \boldsymbol{v}_j, \qquad k = 1, 2, \ldots, \ell. \tag{2.3}$$

It is convenient to use the transformed quantities

$$\widetilde{\boldsymbol{x}}_k = V^T \boldsymbol{x}_k, \qquad \widetilde{\boldsymbol{b}} = [\widetilde{b}_1, \widetilde{b}_2, \ldots, \widetilde{b}_m]^T = U^T \boldsymbol{b}$$

in the computations. Thus, we compute

$$\widetilde{\boldsymbol{x}}_k = \left[\frac{\widetilde{b}_1}{\sigma_1}, \frac{\widetilde{b}_2}{\sigma_2}, \ldots, \frac{\widetilde{b}_k}{\sigma_k}, 0, \ldots, 0 \right]^T \tag{2.4}$$

for a suitable value of $1 \leq k \leq \ell$ and then determine the approximate solution $\boldsymbol{x}_k = V\widetilde{\boldsymbol{x}}_k$ of (1.1).

Let a bound for the norm of the error

$$\|\boldsymbol{e}\| \leq \varepsilon$$

in $\boldsymbol{b}$ be available. We then can determine a suitable truncation index k by the discrepancy principle, i.e., we choose k as small as possible so that

$$\|A\boldsymbol{x}_k - \boldsymbol{b}\| \leq \eta\varepsilon, \tag{2.5}$$

where $\eta \geq 1$ is a user-specified constant independent of ε. Thus, the truncation index $k = k_\varepsilon$ depends on ε and generally increases as ε decreases. A proof of the convergence of $\boldsymbol{x}_{k_\varepsilon}$ to $\hat{\boldsymbol{x}}$ as $\varepsilon \searrow 0$ in a Hilbert space setting is presented in [6]. It requires $\eta > 1$ in (2.5). In actual computations, we use the representation

$$\|A\boldsymbol{x}_k - \boldsymbol{b}\|^2 = \sum_{j=k+1}^{m} \widetilde{b}_j^2$$

to determine k_ε from (2.5). Further details on regularization by the TSVD method can be found in, e.g., [6, 8].

2.2. Standard Tikhonov regularization. Substituting (2.1), $\widetilde{\boldsymbol{x}} = V^T\boldsymbol{x}$, and $\widetilde{\boldsymbol{b}} = U^T\boldsymbol{b}$ into (1.5) yields the penalized least-squares problem

$$\min_{\widetilde{\boldsymbol{x}}\in\mathbb{R}^n} \{\|\Sigma\widetilde{\boldsymbol{x}} - \widetilde{\boldsymbol{b}}\|^2 + \mu^2\|\widetilde{\boldsymbol{x}}\|^2\}$$

with solution

$$\widetilde{\boldsymbol{x}}_\mu = (\Sigma^T\Sigma + \mu^2 I)^{-1}\Sigma^T\widetilde{\boldsymbol{b}} \tag{2.6}$$

for any $\mu > 0$. The associated solution of (1.5) is given by $\boldsymbol{x}_\mu = V\widetilde{\boldsymbol{x}}_\mu$. It satisfies

$$(A^TA + \mu^2 I)\boldsymbol{x}_\mu = A^T\boldsymbol{b}. \tag{2.7}$$

The discrepancy principle prescribes that the regularization parameter $\mu > 0$ be determined so that

$$\|A\boldsymbol{x}_\mu - \boldsymbol{b}\| = \eta\varepsilon, \tag{2.8}$$

or, equivalently, so that

$$\|\Sigma\widetilde{\boldsymbol{x}}_\mu - \widetilde{\boldsymbol{b}}\| = \eta\varepsilon, \tag{2.9}$$

where $\eta \geq 1$ is a user-chosen constant independent of ε. This nonlinear equation for μ can be solved, e.g., by Newton's method. Generally, μ decreases with ε. A proof of the convergence $\boldsymbol{x}_\mu \to \hat{\boldsymbol{x}}$ as $\varepsilon \searrow 0$ is provided in [6]. The proof is in a Hilbert space setting and requires that $\eta > 1$ in (2.8). All methods discussed in Subsections 2.3 and 2.5 use the value of μ determined by (2.8), i.e., $\mu > 0$ is for all methods chosen so that the solution $\boldsymbol{x}_\mu$ of (1.5) satisfies (2.8).

2.3. Modified Tikhonov regularization. It follows from (2.6) that Tikhonov regularization with $\mu > 0$ dampens all solution components $\boldsymbol{v}_j$ of $\boldsymbol{x}_\mu$. On the other hand, TSVD does not dampen any solution component that is not set to zero; cf. (2.4). It is well known that Tikhonov regularization may oversmooth the computed solution when the regularization parameter is determined by the discrepancy principle; see Hansen [8, §7.2]. A more recent discussion on the oversmoothing of the solution (1.6) obtained with Tikhonov regularization is provided by Klann and Ramlau [13].

In order to reduce the oversmoothing, it was suggested in [7] that the minimization problem (1.5) be replaced by

$$\min_{\boldsymbol{x}\in\mathbb{R}^n}\{\|A\boldsymbol{x}-\boldsymbol{b}\|^2+\|L_\mu\boldsymbol{x}\|^2\}, \tag{2.10}$$

where

$$L_\mu = D_\mu V^T \tag{2.11}$$

and

$$D_\mu^2 = \mathrm{diag}\left[\max\{\mu^2-\sigma_1^2,0\},\max\{\mu^2-\sigma_2^2,0\},\ldots,\max\{\mu^2-\sigma_n^2,0\}\right].$$

Thus, the elements of D_μ, and therefore of L_μ, are nonlinear functions of $\mu \geq 0$. Analogously to (2.6), one has

$$\widetilde{\boldsymbol{x}}_\mu = (\Sigma^T\Sigma + D_\mu^2)^{-1}\Sigma^T\widetilde{\boldsymbol{b}}. \tag{2.12}$$

We determine $\mu \geq 0$ so that the solution (1.6) of standard Tikhonov regularization (1.5) satisfies the discrepancy principle (2.8). If $\mu \geq \sigma_1$, then

$$\Sigma^T\Sigma + D_\mu^2 = \mu^2 I.$$

If, instead, $0 \leq \mu < \sigma_1$, then there is $1 \leq k \leq n$ such that $\sigma_k > \mu \geq \sigma_{k+1}$, where we define $\sigma_{n+1} = 0$ when $k = n$. These values of μ and k yield

$$\Sigma^T\Sigma + D_\mu^2 = \mathrm{diag}\left[\sigma_1^2,\sigma_2^2,\ldots,\sigma_k^2,\mu^2,\ldots,\mu^2\right] \in \mathbb{R}^{n\times n}.$$

We will in the remainder of this section assume that $k \geq 1$. When $\mu > 0$, the above matrix is positive definite and the solution (2.12) exists and is unique. The corresponding approximate solution of (1.1) is given by $\boldsymbol{x}_\mu = V\widetilde{\boldsymbol{x}}_\mu$ and satisfies

$$(A^TA + L_\mu^TL_\mu)\boldsymbol{x} = A^T\boldsymbol{b}. \tag{2.13}$$

To avoid severe propagation of the error $\boldsymbol{e}$ in $\boldsymbol{b}$ into the solution of (2.13), the matrix $A^TA+L_\mu^TL_\mu$ should not be too ill-conditioned. This can be achieved by letting $\mu > 0$ be sufficiently large. We measure the conditioning of a matrix by its spectral condition number κ_2, which is defined as the ratio of the largest and smallest positive singular values of the matrix. For instance,

$$\kappa_2(A_k) = \frac{\sigma_1}{\sigma_k}, \qquad 1 \leq k \leq \ell,$$

$$\kappa_2(A^TA+\mu^2I) = \frac{\sigma_1^2+\mu^2}{\sigma_n^2+\mu^2}, \tag{2.14}$$

$$\kappa_2(A^TA+L_\mu^TL_\mu) = \frac{\sigma_1^2}{\mu^2}, \qquad \sigma_n \leq \mu < \sigma_1. \tag{2.15}$$

It is desirable that the matrix $L_\mu^T L_\mu$ be of small norm so that equation (2.13) is fairly close to the normal equations $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$ associated with (1.1), because this may help us determine an accurate approximation of $\hat{\boldsymbol{x}}$. Indeed, the matrix $L_\mu^T L_\mu$ can be shown to be the closest matrix to $A^T A$ in the Frobenius norm with the property that its smallest singular value is μ^2; see [7, Theorem 2.1 and Corollary 2.2]. We recall that the Frobenius norm of a matrix $M \in \mathbb{R}^{n\times n}$ is given by $\|M\|_F = \sqrt{\mathrm{trace}(M^T M)}$.

2.4. Filter factors. Properties of regularization methods can be studied with the aid of filter factors; see, e.g., Hansen [8] and Donatelli and Serra–Capizzano [5] for illustrations. The unregularized solution (1.4) can be expressed as

$$\boldsymbol{x} = \sum_{j=1}^{\ell} \frac{\boldsymbol{u}_j^T\boldsymbol{b}}{\sigma_j}\boldsymbol{v}_j.$$

The filter factors show how the components are modified by a regularization method. For instance, we can express the TSVD solution (2.3) as

$$\boldsymbol{x}_k = \sum_{j=1}^{\ell} \varphi_{k,j}^{(\mathrm{TSVD})} \frac{\boldsymbol{u}_j^T\boldsymbol{b}}{\sigma_j}\boldsymbol{v}_j$$

with the filter factors

$$\varphi_{k,j}^{(\mathrm{TSVD})} = \begin{cases} 1, & 1 \le j \le k, \\ 0, & k < j \le \ell. \end{cases}$$

Similarly, the Tikhonov solution of (2.7) can be written as

$$\boldsymbol{x}_\mu = \sum_{j=1}^{\ell} \varphi_{\mu,j}^{(\mathrm{Tikhonov})} \frac{\boldsymbol{u}_j^T\boldsymbol{b}}{\sigma_j}\boldsymbol{v}_j$$

with the filter factors

$$\varphi_{\mu,j}^{(\mathrm{Tikhonov})} = \frac{\sigma_j^2}{\sigma_j^2 + \mu^2}, \qquad 1 \le j \le \ell.$$

Let $\mu > 0$ and assume that k is such that $\sigma_k > \mu \ge \sigma_{k+1}$, where we define $\sigma_{n+1} = 0$ if $k = n$. The solution of the modified Tikhonov regularization method (2.10) can be expressed as

$$\boldsymbol{x}_\mu = \sum_{j=1}^{\ell} \varphi_{\mu,j} \frac{\boldsymbol{u}_j^T\boldsymbol{b}}{\sigma_j}\boldsymbol{v}_j$$

with the filter factors

$$\varphi_{\mu,j} = \begin{cases} 1, & 1 \le j \le k, \\ \dfrac{\sigma_j^2}{\mu^2}, & k < j \le \ell. \end{cases}$$

Thus, these filter factors are the same as $\varphi_{k,j}^{(\mathrm{TSVD})}$ for $1 \le j \le k$, and close to $\varphi_{\mu,j}^{(\mathrm{Tikhonov})}$ for $k < j \le \ell$.

2.5. New modified Tikhonov regularization methods. This section derives new modifications of Tikhonov regularization (1.5) by focusing on condition numbers. For all methods of this subsection, we determine $\mu \geq 0$ similarly as in Subsection 2.3, i.e., so that the solution (1.6) of (1.5) satisfies (2.8). Then k is chosen as a function of μ as described.

PROPOSITION 2.1. *Let L_μ be defined by (2.11) and assume that $\sigma_n \leq \mu \leq \sigma_1$. Then*

$$\max\{\kappa_2(A^TA + L_\mu^T L_\mu), \kappa_2(A^TA + \mu^2 I)\} \leq \kappa_2(A^TA). \tag{2.16}$$

Moreover,

$$\kappa_2(A^TA + L_\mu^T L_\mu) \leq \kappa_2(A^TA + \mu^2 I) \Leftrightarrow \mu^2 \geq \sigma_1\sigma_n. \tag{2.17}$$

Proof. The proofs of the inequalities (2.16) and (2.17) follow from (2.14) and (2.15). The requirement on μ^2 in (2.17) typically is satisfied for linear discrete ill-posed problems that arise in applications. □

We discuss Tikhonov regularization for several regularization matrices that are modifications of μI and yield condition numbers of the associated normal equations that are smaller than the condition number (2.14) of the matrix $A^TA + \mu^2 I$. We first consider the regularization matrix

$$L_{\mu,k} = D_{\mu,k} V^T \tag{2.18}$$

with

$$D_{\mu,k} = \operatorname{diag}\left[0, 0, \ldots, 0, \overbrace{\mu, \ldots, \mu}^{n-k}\right].$$

Given $\mu \geq 0$, the index $k = k_\mu$ is chosen so that the diagonal entries of

$$\Sigma^T\Sigma + D_{\mu,k}^2 = \operatorname{diag}\left[\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2, \sigma_{k+1}^2 + \mu^2, \ldots, \sigma_n^2 + \mu^2\right],$$

are non-increasing when the column index increases. Thus, the regularization matrix (2.18) leaves the largest k eigenvalues of A^TA invariant and shifts the remaining ones.

PROPOSITION 2.2. *Let L_μ and $L_{\mu,k}$ be defined by (2.11) and (2.18), respectively, and assume that $k = k_\mu$ in $L_{\mu,k}$ is chosen as described above. Then*

$$\kappa_2(A^TA + L_{\mu,k}^T L_{\mu,k}) = \frac{\sigma_1^2}{\sigma_n^2 + \mu^2}.$$

Therefore

$$\kappa_2(A^TA + L_{\mu,k}^T L_{\mu,k}) \leq \kappa_2(A^TA + \mu^2 I) \Leftrightarrow \mu^2 \neq 0 \tag{2.19}$$

and

$$\kappa_2(A^TA + L_{\mu,k}^T L_{\mu,k}) \leq \kappa_2(A^TA + L_\mu^T L_\mu), \tag{2.20}$$

where the latter inequality is strict if and only if A is of full rank. Moreover, for $k \geq 1$,

$$\|L_{\mu,k}\|_F < \|\mu I\|_F. \tag{2.21}$$

Proof. The proofs of (2.19) and (2.20) are immediate. The inequality (2.21) follows from the observation that

$$\|L_{\mu,k}\|_F^2 = \|D_{\mu,k}\|_F^2 = (n-k)\mu^2.$$

□

The filter factors for Tikhonov regularization with the regularization matrix (2.18) are given by

$$\varphi_{\mu,k,j} = \begin{cases} 1, & 1 \leq j \leq k, \\ \dfrac{\sigma_j^2}{\sigma_j^2+\mu^2}, & k < j \leq \ell. \end{cases} \tag{2.22}$$

Thus, these filter factors are the same as $\varphi_{k,j}^{(\mathrm{TSVD})}$ for $1 \leq j \leq k$, and the same as $\varphi_{\mu,j}^{(\mathrm{Tikhonov})}$ for $k < j \leq \ell$. However, the discrepancy principle applied to TSVD, cf. (2.5), may yield a different value of k.

We are lead to an alternative to the regularization matrix (2.18) when we instead of shifting the smallest eigenvalues of A^TA ignore them. Define the regularization matrix

$$L_k = D_k V^T \tag{2.23}$$

with

$$D_k^2 = \mathrm{diag}\left[0, 0, \ldots, 0, -\sigma_{k+1}^2, \ldots, -\sigma_n^2\right].$$

Then

$$\Sigma^T\Sigma + D_k^2 = \Sigma_k^T\Sigma_k = \mathrm{diag}\left[\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2, 0, 0, \ldots, 0\right].$$

PROPOSITION 2.3. *Let the regularization matrices L_μ, $L_{\mu,k}$, and L_k be defined by (2.11), (2.18), and (2.23), respectively. Then*

$$\kappa_2(A^TA + L_k^TL_k) = \kappa_2(A_k^TA_k) = \frac{\sigma_1^2}{\sigma_k^2}.$$

Therefore,

$$\kappa_2(A^TA + L_k^TL_k) \leq \kappa_2(A^TA + L_\mu^TL_\mu) \Leftrightarrow \mu \leq \sigma_k \tag{2.24}$$

and

$$\kappa_2(A^TA + L_k^TL_k) \leq \kappa_2(A^TA + L_{\mu,k}^TL_{\mu,k}) \Leftrightarrow \mu \leq \sqrt{\sigma_k^2 - \sigma_n^2}. \tag{2.25}$$

Moreover, if $\sigma_{k+1} \leq \mu < \sigma_k$, then

$$\|L_k\|_F \leq \|L_{\mu,k}\|_F. \tag{2.26}$$

Proof. The inequalities (2.24) and (2.25) are straightforward. Property (2.26) follows from

$$\|L_k\|_F^2 = \|D_k\|_F^2 = \sum_{\sigma_j^2 \leq \mu^2} \sigma_j^2 \leq (n-k)\mu^2 = \|D_{\mu,k}\|_F^2 = \|L_{\mu,k}\|_F^2.$$

□

The filter factors for Tikhonov regularization with the regularization matrix (2.23) are the same as $\varphi_{k,j}^{(\mathrm{TSVD})}$.

The observations at the end of Subsection 2.3 suggest that we seek to determine regularization matrices that give normal equations with the same condition number as $A^TA+\mu^2I$ but have smaller Frobenius norm than μI. Introduce the regularization matrix

$$\widetilde{L}_\mu = \widetilde{D}_\mu V^T \tag{2.27}$$

with

$$\widetilde{D}_\mu^2 = \frac{\mu^2}{\sigma_1^2+\mu^2}\,\mathrm{diag}\left[0, \sigma_1^2-\sigma_2^2, \dots, \sigma_1^2-\sigma_n^2\right].$$

Then

$$\Sigma^T\Sigma + \widetilde{D}_\mu^2 = \mathrm{diag}\left[\sigma_1^2, \frac{\sigma_1^2}{\sigma_1^2+\mu^2}(\sigma_2^2+\mu^2), \dots, \frac{\sigma_1^2}{\sigma_1^2+\mu^2}(\sigma_n^2+\mu^2)\right]. \tag{2.28}$$

Proposition 2.4. *Let $\widetilde{L}_\mu$ be given by (2.27). Then*

$$\kappa_2(A^TA+\widetilde{L}_\mu^T\widetilde{L}_\mu) = \kappa_2(A^TA+\mu^2I) \tag{2.29}$$

and

$$\|\widetilde{L}_\mu\|_F < \|\mu I\|_F. \tag{2.30}$$

Proof. The equality (2.29) follows from (2.28). The inequality (2.30) is a consequence of

$$\|\widetilde{L}_\mu\|_F^2 = \|\widetilde{D}_\mu\|_F^2 = \frac{\mu^2}{\sigma_1^2+\mu^2}\sum_{i=2}^n(\sigma_1^2-\sigma_i^2) < (n-1)\mu^2 < \|\mu I\|_F^2.$$

□

The filter factors for Tikhonov regularization with the regularization matrix (2.27) are given by

$$\widetilde{\varphi}_{\mu,j} = \frac{\sigma_j^2(\sigma_1^2+\mu^2)}{\sigma_1^2(\sigma_j^2+\mu^2)}, \qquad 1 \le j \le \ell.$$

Thus, these filter factors are the same as $\varphi_{k,j}^{(\mathrm{TSVD})}$ for $j=1$, and close to $\varphi_{\mu,j}^{(\mathrm{Tikhonov})}$ for $1<j\le\ell$. Specifically,

$$\widetilde{\varphi}_{\mu,j} = \frac{(\sigma_1^2+\mu^2)}{\sigma_1^2}\varphi_{\mu,j}^{(\mathrm{Tikhonov})}, \qquad 1<j\le\ell.$$

Another regularization matrix that also yields regularized normal equations with the same spectral condition number as $A^TA+\mu^2I$ is given by

$$\widetilde{L}_{\mu,k} = \widetilde{D}_{\mu,k}V^T \tag{2.31}$$

with

$$\widetilde{D}^2_{\mu,k} = \frac{\mu^2}{\sigma_1^2+\mu^2}\,\mathrm{diag}\left[0,\ldots,0,\sigma_1^2-\sigma_{k+1}^2,\ldots,\sigma_1^2-\sigma_n^2\right].$$

Then

$$\Sigma^T\Sigma+\widetilde{D}^2_{\mu,k} = \mathrm{diag}\left[\sigma_1^2,\ldots,\sigma_k^2,\frac{\sigma_1^2}{\sigma_1^2+\mu^2}(\sigma_{k+1}^2+\mu^2),\ldots,\frac{\sigma_1^2}{\sigma_1^2+\mu^2}(\sigma_n^2+\mu^2)\right]. \tag{2.32}$$

The index $k=k_\mu$ is chosen so that the diagonal entries of $\Sigma^T\Sigma+\widetilde{D}^2_{\mu,k}$ are nonincreasing. The following results are analogous to those of Proposition 2.4.

PROPOSITION 2.5. *Let the matrix $\widetilde{L}_{\mu,k}$ be defined by (2.31) with the index $k=k_\mu$ chosen as indicated above. Then*

$$\kappa_2(A^TA+\widetilde{L}^T_{\mu,k}\widetilde{L}_{\mu,k}) = \kappa_2(A^TA+\mu^2 I) \tag{2.33}$$

and

$$\|\widetilde{L}_{\mu,k}\|_F^2 < (n-k)\mu^2 < \|\mu I\|_F^2. \tag{2.34}$$

Proof. Property (2.33) is a consequence of (2.32), and (2.34) follows from the choice of k, i.e., $\sigma_{k+1}\leq\mu<\sigma_k$.

Indeed, the squared Frobenius norm of the regularization matrix defined by $\widetilde{L}_{\mu,k}$ in (2.31) is less than or equal to that of the one defined by $\widetilde{L}_\mu$ in (2.27), i.e.

$$\|\widetilde{L}_{\mu,k}\|_F^2 = \|\widetilde{D}_{\mu,k}\|_F^2 = \frac{\mu^2}{\sigma_1^2+\mu^2}\sum_{i=k+1}^{n}(\sigma_1^2-\sigma_i^2) < (n-k)\mu^2 < \|\mu I\|_F^2.$$

□

We next compare the regularization matrices (2.11) and (2.31).

PROPOSITION 2.6. *Let L_μ and $\widetilde{L}_{\mu,k}$ be given by (2.11) and (2.31), respectively. Assume that k is such that*

$$\sigma_k > \frac{\mu^2}{\sigma_1} \geq \sigma_{k+1}. \tag{2.35}$$

Then

$$\|\widetilde{L}_{\mu,k}\|_F \leq \|L_\mu\|_F.$$

Proof. For any $j>k$, one has $\sigma_1\sigma_j\leq\mu^2$. Therefore,

$$\frac{\mu^2}{\sigma_1^2+\mu^2}(\sigma_1^2-\sigma_j^2) \leq \mu^2-\sigma_j^2,$$

and it follows that

$$\|\widetilde{L}_{\mu,k}\|_F^2 = \frac{\mu^2}{\sigma_1^2+\mu^2}\sum_{j=k+1}^{n}(\sigma_1^2-\sigma_j^2) \leq \sum_{j=k+1}^{n}(\mu^2-\sigma_j^2).$$

Assuming $\mu < \sigma_1$, so that $\mu^2 < \mu\sigma_1$, we obtain

$$\sum_{j=k+1}^{n} (\mu^2 - \sigma_j^2) \leq \sum_{\sigma_j^2 < \mu^2} (\mu^2 - \sigma_j^2),$$

which concludes the proof. □

Note that the parameter k such that (2.35) is satisfied may differ from the parameter $\tilde{k}$ such that $\sigma_{\tilde{k}} > \mu \geq \sigma_{\tilde{k}+1}$. Specifically, $k \geq \tilde{k}$.

We also can establish the relations

$$\|\widetilde{L}_{\mu,k}\|_F < \|L_{\mu,k}\|_F, \qquad \|\widetilde{L}_{\mu,k}\|_F \leq \|\widetilde{L}_{\mu}\|_F,$$

where the latter inequality is strict if $\sigma_1 > \sigma_k$. Thus, the regularization matrix $\widetilde{L}_{\mu,k}$ yields normal equations with the same condition number as the regularization matrix μI, but is of smaller norm than this and several other regularization matrices considered. We therefore expect $\widetilde{L}_{\mu,k}$ to often yield more accurate approximations of the desired solution $\hat{\boldsymbol{x}}$ than the other regularization matrices discussed above. That this is, indeed, the case is illustrated in Section 3.

The filter factors for Tikhonov regularization with the regularization matrix (2.31) are given by

$$\widetilde{\varphi}_{\mu,k,j} = \begin{cases} 1, & 1 \leq j \leq k, \\ \dfrac{\sigma_j^2(\sigma_1^2 + \mu^2)}{\sigma_1^2(\sigma_j^2 + \mu^2)}, & k < j \leq \ell, \end{cases}$$

i.e., they are same as $\varphi_{k,j}^{(\mathrm{TSVD})}$ for $1 \leq j \leq k$, and are close to $\varphi_{\mu,j}^{(\mathrm{Tikhonov})}$ for $k < j \leq \ell$.

The above analysis suggests that we introduce a parameter θ that allows us to interpolate between the regularization matrices (2.18) and (2.31). Thus, define for $0 \leq \theta \leq 1$ the regularization matrices

$$L_{\mu,k}(\theta) = D_{\mu,k}(\theta) V^T \tag{2.36}$$

with

$$D_{\mu,k}^2(\theta) = \frac{\mu^2}{\sigma_1^2 + \theta\mu^2} \operatorname{diag}\left[0, \ldots, 0, \sigma_1^2 - \theta\sigma_{k+1}^2, \ldots, \sigma_1^2 - \theta\sigma_n^2\right].$$

Then

$$\Sigma^T\Sigma + D_{\mu,k}^2(\theta) = \operatorname{diag}\left[\sigma_1^2, \ldots, \sigma_k^2, \frac{\sigma_1^2}{\sigma_1^2 + \theta\mu^2}(\sigma_{k+1}^2 + \mu^2), \ldots, \frac{\sigma_1^2}{\sigma_1^2 + \theta\mu^2}(\sigma_n^2 + \mu^2)\right]$$

from which it follows that

$$\begin{aligned} \kappa_2(A^TA + L_{\mu,k}(\theta)^T L_{\mu,k}(\theta)) &= (1-\theta)\kappa_2(A^TA + L_{\mu,k}^T L_{\mu,k}) + \theta\kappa_2(A^TA + \widetilde{L}_{\mu,k}^T \widetilde{L}_{\mu,k}) \\ &= \frac{\sigma_1^2 + \theta\mu^2}{\sigma_n^2 + \mu^2}. \end{aligned}$$

Moreover,

$$\|L_{\mu,k}(\theta)\|_F^2 = \frac{\mu^2}{\sigma_1^2 + \theta\mu^2} \sum_{i=k+1}^{n} (\sigma_1^2 - \theta\sigma_i^2).$$

Hence, the norm $\|L_{\mu,k}(\theta)\|_F^2$ is a nonincreasing function of θ, whereas the condition number $\kappa_2(A^TA + L_{\mu,k}(\theta)^T L_{\mu,k}(\theta))$ is an increasing function of θ.

The filter factors for Tikhonov regularization with the regularization matrix (2.36) are given by

$$\varphi_{\mu,k,j}(\theta) = (1-\theta)\varphi_{\mu,k,j} + \theta\widetilde{\varphi}_{\mu,k,j} = \begin{cases} 1, & 1 \le j \le k, \\ \dfrac{\sigma_j^2(\sigma_1^2 + \theta\mu^2)}{\sigma_1^2(\sigma_j^2 + \mu^2)}, & k < j \le \ell, \end{cases}$$

where $\varphi_{\mu,k,j}$ is defined by (2.22). Thus, the filter factors $\varphi_{\mu,k,j}(\theta)$ agree with $\varphi_{k,j}^{(\mathrm{TSVD})}$ for $1 \le j \le k$, and are close to $\varphi_{\mu,j}^{(\mathrm{Tikhonov})}$ for $k < j \le \ell$.

Numerical examples in the following section show the regularization matrices $L_{\mu,k}(1) = \widetilde{L}_{\mu,k}$ and $L_{\mu,k}(0) = L_{\mu,k}$ to yield the most accurate approximations of $\hat{\boldsymbol{x}}$. The former matrix has the smallest Frobenius norm and the latter yields normal equations for Tikhonov regularization with the smallest condition number.

3. Computed examples. The calculations of this section were carried out using MATLAB with relative accuracy $2.2 \cdot 10^{-16}$. Most of the examples are obtained by discretizing Fredholm integral equations of the first kind

$$\int_a^b h(s,t)x(t)\,dt = g(s), \qquad c \le s \le d, \tag{3.1}$$

with a smooth kernel h. The discretizations are carried out by Galerkin or Nyström methods and yield linear discrete ill-posed problems (1.1). MATLAB functions in Regularization Tools [9] determine discretizations $A \in \mathbb{R}^{m\times n}$ of the integral operators and scaled discrete approximations $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ of the solution x of (3.1). In all examples, we let $m = n = 200$. The performance of the regularization matrices discussed in this paper is illustrated when the error $\boldsymbol{e}$ in $\boldsymbol{b}$ is white Gaussian noise or colored noise. We begin with the former.

3.1. Tests with white noise. In the experiments of this subsection the error vector $\boldsymbol{e} \in \mathbb{R}^m$ has normally distributed random entries with zero mean. The vector is scaled to yield a specified noise level $\|\boldsymbol{e}\|/\|\hat{\boldsymbol{b}}\|$ and added to the error-free data vector $\hat{\boldsymbol{b}} := A\hat{\boldsymbol{x}}$ to obtain the vector $\boldsymbol{b}$ in (1.1); cf. (1.2). In particular, $\|\boldsymbol{e}\|$ is available and we can apply the discrepancy principle with $\varepsilon = \|\boldsymbol{e}\|$ to determine the regularization parameter μ in Tikhonov regularization and the truncation index k in TSVD. The parameter η in (2.5) and (2.9) is set to one.

The computed approximation of $\hat{\boldsymbol{x}}$ is denoted by $\boldsymbol{x}_{\mathrm{comp}}$. We are interested in the relative error $\|\boldsymbol{x}_{\mathrm{comp}} - \hat{\boldsymbol{x}}\|/\|\hat{\boldsymbol{x}}\|$ in the computed solutions determined by Tikhonov regularization with the different regularization matrices described, and by TSVD. The difference $\boldsymbol{x}_{\mathrm{comp}} - \hat{\boldsymbol{x}}$ depends on the entries of the error vector $\boldsymbol{e}$. We report for every example the average of the relative errors in $\boldsymbol{x}_{\mathrm{comp}}$ over 1000 runs for each noise level.

Example 3.1. We first consider the problem phillips from [9]. Let

$$\phi(t) = \begin{cases} 1 + \cos(\frac{\pi t}{3}), & |t| < 3, \\ 0, & |t| \ge 3, \end{cases}$$

and $a = c = -6$, $b = d = 6$. The kernel, right-hand side function, and solution of the integral equation (3.1) are given by

$$h(s,t) = \phi(s-t), \quad x(t) = \phi(t), \quad g(s) = (6 - |s|)\left(1 + \frac{1}{2}\cos\left(\frac{\pi s}{3}\right)\right) + \frac{9}{2\pi}\sin\left(\frac{\pi|s|}{3}\right).$$

Noise level	Tikhonov regularization			TSVD
%	L in (2.11)	$L = \mu I$	L in (2.18)	
10.0	$6.70 \cdot 10^{-2}$	$6.83 \cdot 10^{-2}$	$\mathbf{6.32} \cdot 10^{-2}$	$7.86 \cdot 10^{-2}$
1.0	$2.72 \cdot 10^{-2}$	$2.62 \cdot 10^{-2}$	$2.62 \cdot 10^{-2}$	$\mathbf{2.57} \cdot 10^{-2}$
0.5	$2.17 \cdot 10^{-2}$	$2.08 \cdot 10^{-2}$	$\mathbf{2.07} \cdot 10^{-2}$	$2.47 \cdot 10^{-2}$
0.1	$1.08 \cdot 10^{-2}$	$1.11 \cdot 10^{-2}$	$\mathbf{1.03} \cdot 10^{-2}$	$1.23 \cdot 10^{-2}$

TABLE 3.1

Example 3.1: Average relative errors in the computed solutions for the phillips *test problem for several noise levels.*

Table 3.1 displays the averages of the relative errors in the computed solutions over 1000 runs for each noise level. The smallest average relative error is for each noise level marked in boldface. Tikhonov regularization with the regularization matrix (2.18) is seen to yield the same or smaller average errors as Tikhonov regularization with the regularization matrices (2.11) and μI. The only average error that is smaller than for Tikhonov regularization with the matrix (2.18) is obtained for 1% noise by the TSVD method. We conclude that the regularization matrix (2.18) yields competitive results and, in particular, determines more accurate approximations of $\hat{\boldsymbol{x}}$ than standard Tikhonov regularization (1.5). □

Noise level	Tikhonov regularization			TSVD
%	L in (2.11)	$L = \mu I$	L in (2.18)	
10.0	$\mathbf{1.69} \cdot 10^{-1}$	$1.76 \cdot 10^{-1}$	$1.70 \cdot 10^{-1}$	$1.86 \cdot 10^{-1}$
1.0	$\mathbf{1.02} \cdot 10^{-1}$	$1.13 \cdot 10^{-1}$	$1.11 \cdot 10^{-1}$	$1.30 \cdot 10^{-1}$
0.5	$\mathbf{6.76} \cdot 10^{-2}$	$8.35 \cdot 10^{-2}$	$7.53 \cdot 10^{-2}$	$7.86 \cdot 10^{-2}$
0.1	$4.83 \cdot 10^{-2}$	$5.03 \cdot 10^{-2}$	$\mathbf{4.80} \cdot 10^{-2}$	$4.83 \cdot 10^{-2}$

TABLE 3.2

Example 3.2: Average relative errors in the computed solutions for the shaw *test problem for several noise levels.*

Example 3.2. The test problem shaw from [9] is an integral equation (3.1) with kernel and solution

$$h(s,t) = (\cos(s) + \cos(t))^2 \left(\frac{\sin(u)}{u}\right)^2, \quad u = \pi(\sin(s) + \sin(t)),$$
$$x(t) = 2\exp\left(-6\left(t - \frac{4}{5}\right)^2\right) + \exp\left(-2\left(t + \frac{1}{2}\right)^2\right),$$

and parameters $a = c = -\pi/2$, $b = d = \pi/2$. Table 3.2 is analogous to Table 3.1; it displays the averages of the relative errors in the computed solutions over 1000 runs for each noise level. The regularization parameter μ for Tikhonov regularization and the truncation index k for TSVD are determined with the aid of the discrepancy principle. The smallest entry in each row is in boldface. The regularization matrices (2.11) and (2.18) can be seen to perform the best.

Table 3.3 compares the performance of the methods when the optimal values of the regularization parameter μ in Tikhonov regularization is used, i.e., we use the values that give the most accurate approximations of $\hat{\boldsymbol{x}}$. These values of μ are generally not available when solving discrete ill-posed problems. Nevertheless, it is interesting to see how the regularization matrices would perform if the optimal values of μ were available. The table shows, in increasing order, the average relative errors over 1000 runs in the computed approximate solutions determined by Tikhonov regularization for the noise level 0.1%. All the modifications (2.11), (2.18), and (2.31) give approximate solutions of higher quality than $L = \mu I$. For the sake of completeness, we also report the average of the relative errors in the computed solutions obtained with TSVD when the truncation index k is chosen to give the most accurate approximation of $\hat{\boldsymbol{x}}$. It is $4.4777146 \cdot 10^{-2}$, which is slightly larger than the average errors reported in Table 3.3. □

L in (2.18)	L in (2.31)	L in (2.11)	$L = \mu I$
$4.3750446 \cdot 10^{-2}$	$4.3750452 \cdot 10^{-2}$	$4.3855830 \cdot 10^{-2}$	$4.4713012 \cdot 10^{-2}$

Table 3.3

Example 3.2: Average relative errors in the computed solutions for the `shaw` *test problem for noise level* 0.1% *with optimal regularization parameters* μ *and* k*.*

Example 3.3. Consider the problem `heat` from [9]. It is a discretization of a Volterra integral equation of the first kind on the interval $[0, 1]$ with a convolution kernel. Table 3.4 shows the average relative errors in the computed solutions determined by Tikhonov regularization and TSVD over 1000 runs for each noise level. The regularization matrices (2.11) and (2.18) are seen to yield the smallest average relative errors. □

Noise level	Tikhonov regularization			TSVD
%	L in (2.11)	$L = \mu I$	L in (2.18)	
10.0	$2.61 \cdot 10^{-1}$	$2.88 \cdot 10^{-1}$	$\mathbf{2.59} \cdot 10^{-1}$	$3.04 \cdot 10^{-1}$
1.0	$9.95 \cdot 10^{-2}$	$1.08 \cdot 10^{-1}$	$\mathbf{9.78} \cdot 10^{-2}$	$1.20 \cdot 10^{-1}$
0.5	$\mathbf{7.17} \cdot 10^{-2}$	$7.75 \cdot 10^{-2}$	$7.21 \cdot 10^{-2}$	$9.67 \cdot 10^{-2}$
0.1	$3.50 \cdot 10^{-2}$	$3.67 \cdot 10^{-2}$	$\mathbf{3.43} \cdot 10^{-2}$	$4.61 \cdot 10^{-2}$

Table 3.4

Example 3.3: Average relative errors in the computed solutions for the `heat` *test problem for several noise levels.*

3.2. Tests with colored noise. In this subsection, we consider noise whose power density increases with the frequency, i.e., the noise has more energy in the high frequencies than white Gaussian noise. This kind of noise is known as "colored noise" and is sometimes referred to as "violet noise"; see, e.g., Hansen [10] for a discussion of colored noise in discrete ill-posed problems. Let U be the orthogonal matrix of left singular vectors of the matrix A in (1.1). Hansen [10, p. 74] generates colored noise with the MATLAB command

(3.2)
```
e=U*(logspace(-alpha,0,200)'.*(U'*randn(200,1)));
```

Here `randn(200,1)` yields a vector in $\mathbb{R}^{200}$ with normally distributed random entries and the parameter $\alpha =$ `alpha` determines how much the energy in the high frequencies dominate; they dominate more the larger $\alpha > 0$. We add the vector $\boldsymbol{e} =$ `e` to the noise-free data vector $\hat{\boldsymbol{b}}$ to obtain the noise-contaminated data vector $\boldsymbol{b}$; cf. (1.2). When the covariance matrix for the noise is known, then its Cholesky factorization can be used to prewhitening the noise; see [10, p. 76]. We assume the covariance matrix not to be available and would like to illustrate how the methods considered in this paper perform in this situation. The vector $\boldsymbol{e}$ is scaled to yield a specified noise level $\|\boldsymbol{e}\|/\|\hat{\boldsymbol{b}}\|$ and we use the discrepancy principle to determine the regularization parameters in Tikhonov regularization and TSVD with $\eta = 1$ in (2.5) and (2.9). We also will replace the matrix U in (3.2) by other orthogonal matrices.

Example 3.4. Consider the integral equation of the first kind (3.1) with the kernel and right-hand side function given by

$$h(s,t) = \begin{cases} s\,(t-1), & s < t, \\ t\,(s-1), & s \geq t, \end{cases}$$

and

$$g(s) = \begin{cases} (4\,s^3 - 3\,s)/24, & s < 0.5, \\ (-4\,s^3 + 12\,s^2 - 9\,s + 1)/24, & s \geq 0.5. \end{cases}$$

We use the MATLAB function deriv2 from [9] to determine a discretization $A \in \mathbb{R}^{200\times 200}$ of the integral operator, and a scaled discrete approximation $\hat{\boldsymbol{x}}$ of the solution

$$x(t) = \begin{cases} t, & t < 0.5, \\ 1-t, & t \geq 0.5. \end{cases}$$

We compute the noise-free data vector $\hat{\boldsymbol{b}} := A\hat{\boldsymbol{x}}$ to which we add the noise-vector $\boldsymbol{e}$. The latter is generated by (3.2) with $\alpha = 1$ followed by scaling.

Table 3.5 displays the averages of the relative errors in the computed solutions over 1000 runs for each noise level. Tikhonov regularization with the regularization matrix (2.18) is seen to yield the smallest average errors for all noise levels. Table 3.6 is obtained by replacing the orthogonal matrix U of left singular vectors in (3.2) by a random orthogonal matrix, and for the results of Table 3.7 this matrix is replaced by the orthogonal cosine transform matrix. The regularization matrix (2.18) is seen to perform well in each one of these tables. □

Noise level	Tikhonov regularization			TSVD
%	L in (2.11)	$L = \mu I$	L in (2.18)	
1.0	$2.31 \cdot 10^{-2}$	$2.25 \cdot 10^{-2}$	$\mathbf{2.16} \cdot 10^{-1}$	$2.34 \cdot 10^{-1}$
0.5	$1.81 \cdot 10^{-2}$	$1.76 \cdot 10^{-2}$	$\mathbf{1.72} \cdot 10^{-2}$	$1.81 \cdot 10^{-2}$
0.1	$1.01 \cdot 10^{-2}$	$9.86 \cdot 10^{-3}$	$\mathbf{9.62} \cdot 10^{-3}$	$1.03 \cdot 10^{-2}$

TABLE 3.5

Example 3.4: Average relative errors in the computed solutions for the deriv2 *test problem for several noise levels. Moderate violet noise ($\alpha = 1$).*

Noise level	Tikhonov regularization			TSVD
%	L in (2.11)	$L = \mu I$	L in (2.18)	
1.0	$3.90 \cdot 10^{-2}$	$3.78 \cdot 10^{-2}$	$\mathbf{3.65} \cdot 10^{-1}$	$4.07 \cdot 10^{-1}$
0.5	$3.05 \cdot 10^{-2}$	$2.96 \cdot 10^{-2}$	$\mathbf{2.92} \cdot 10^{-2}$	$3.02 \cdot 10^{-2}$
0.1	$1.65 \cdot 10^{-2}$	$1.62 \cdot 10^{-3}$	$\mathbf{1.56} \cdot 10^{-3}$	$1.74 \cdot 10^{-2}$

TABLE 3.6

Example 3.4: Average relative errors in the computed solutions for the deriv2 *test problem for several noise levels. U in (3.2) is an orthogonal random matrix. Violet noise ($\alpha = 2$).*

Example 3.5. Consider again the test problem heat from [9]. Tables 3.8 and 3.9 are analogous to Tables 3.5 and 3.6, respectively. Tikhonov regularization with the regularization matrix (2.18) is seen to perform well. □

Noise level	Tikhonov regularization			TSVD
%	L in (2.11)	$L=\mu I$	L in (2.18)	
1.0	$2.32 \cdot 10^{-2}$	$2.25 \cdot 10^{-2}$	$\mathbf{2.16} \cdot 10^{-1}$	$2.34 \cdot 10^{-1}$
0.5	$1.81 \cdot 10^{-2}$	$1.76 \cdot 10^{-2}$	$\mathbf{1.72} \cdot 10^{-2}$	$1.80 \cdot 10^{-2}$
0.1	$1.02 \cdot 10^{-2}$	$9.90 \cdot 10^{-3}$	$\mathbf{9.64} \cdot 10^{-3}$	$1.03 \cdot 10^{-2}$

TABLE 3.7

Example 3.4: Average relative errors in the computed solutions for the deriv2 *test problem for several noise levels. U in (3.2) is a orthogonal cosine transform matrix. Moderate violet noise ($\alpha = 1$).*

Noise level	Tikhonov regularization			TSVD
%	L in (2.11)	$L=\mu I$	L in (2.18)-(2.31)	
1.0	$5.78 \cdot 10^{-2}$	$5.92 \cdot 10^{-2}$	$\mathbf{5.40} \cdot 10^{-2}$	$6.76 \cdot 10^{-2}$
0.5	$4.34 \cdot 10^{-2}$	$4.36 \cdot 10^{-2}$	$\mathbf{4.21} \cdot 10^{-2}$	$4.95 \cdot 10^{-2}$
0.1	$2.48 \cdot 10^{-2}$	$\mathbf{2.29} \cdot 10^{-2}$	$2.30 \cdot 10^{-2}$	$2.34 \cdot 10^{-2}$

TABLE 3.8

Example 3.5: Average relative errors in the computed solutions for the heat *test problem for several noise levels. Moderate violet noise ($\alpha = 1$).*

Noise level	Tikhonov regularization			TSVD
%	L in (2.11)	$L=\mu I$	L in (2.18)-(2.31)	
1.0	$9.76 \cdot 10^{-2}$	$1.06 \cdot 10^{-1}$	$\mathbf{9.67} \cdot 10^{-2}$	$1.18 \cdot 10^{-1}$
0.5	$\mathbf{7.14} \cdot 10^{-2}$	$7.73 \cdot 10^{-1}$	$7.18 \cdot 10^{-2}$	$9.68 \cdot 10^{-1}$
0.1	$3.50 \cdot 10^{-2}$	$3.70 \cdot 10^{-2}$	$\mathbf{3.44} \cdot 10^{-2}$	$4.61 \cdot 10^{-2}$

TABLE 3.9

Example 3.5: Average relative errors in the computed solutions for the heat *test problem for several noise levels. The matrix U in (3.2) is an orthogonal random matrix. Violet noise ($\alpha = 2$).*

4. Conclusion and extension. Tikhonov regularization suggests several matrix nearness problems for determining regularization matrices. Regularization matrices so defined can give approximate solutions of higher quality than both Tikhonov regularization (1.5) with regularization matrix μI and the TSVD method. The computational effort is dominated by the computation of the SVD (2.1) of the given matrix A in (1.1) and, consequently, is essentially the same for all methods considered in this paper. The new regularization matrices are attractive both when the noise $\boldsymbol{e}$ is white Gaussian or violet.

For ease of description of the methods, we assumed the SVD of A to be available. This requirement can be removed. A least-squares problem (1.1) with a matrix too large to compute its SVD can be reduced to a small problem by a Krylov subspace method. The methods of the present paper can be applied to the reduced problem so obtained. Reduction methods include partial Golub–Kahan bidiagonalization and partial Arnoldi decomposition; see, e.g., [1, 4, 14, 17] for illustrations of application of these reduction methods.

We also note that the methods of this paper can be applied to Tikhonov regularization problems (1.5) with a more general regularization matrix than μI by first transforming the more general problem to the form (1.5). Transformation methods are discussed in [8, Sections 2.3.1 and 2.3.2] and [16].

We used the discrepancy principle to determine the amount of regularization in all computed examples. However, the regularization methods described also can be applied in conjunction with parameter choice rules that do not require a bound for $\|\boldsymbol{e}\|$ to be known. Many such parameter choice rules are discussed and analyzed in [2, 3, 8, 10, 11, 12, 15] and in references therein.

Acknowledgement. We would like to thank a referee for comments that improved the presentation.

REFERENCES

[1] Å. Björck, *A bidiagonalization algorithm for solving large and sparse ill-posed systems of linear equations*, BIT, 18 (1988), pp. 659–670.
[2] C. Brezinski, M. Redivo–Zaglia, G. Rodriguez, and S. Seatzu, *Extrapolation techniques for ill-conditioned linear systems*, Numer. Math., 81 (1998), pp. 1–29.
[3] C. Brezinski, G. Rodriguez, and S. Seatzu, *Error estimates for the regularization of least squares problems*, Numer. Algorithms, 51 (2009), pp. 61–76.
[4] D. Calvetti, G. H. Golub and L. Reichel, *Estimation of the L-curve via Lanczos bidiagonalization*, BIT, 39 (1999), pp. 603–619.
[5] M. Donatelli and S. Serra–Capizzano, *Filter factor analysis of an iterative multilevel regularization method*, Electron. Trans. Numer. Anal., 29 (2008), pp. 163–177.
[6] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.
[7] M. Fuhry and L. Reichel, *A new Tikhonov regularization method*, Numer. Algorithms, 59 (2012), pp. 433–445.
[8] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems*, SIAM, Philadelphia, 1998.
[9] P. C. Hansen, *Regularization tools version 4.0 for Matlab 7.3*, Numer. Algorithms, 46 (2007), pp. 189–194.
[10] P. C. Hansen, *Discrete Inverse Problems: Insight and Algorithms*, SIAM, Philadelphia, 2010.
[11] S. Kindermann, *Convergence analysis of minimization-based noise level-free parameter choice rules for linear ill-posed problems*, Electron. Trans. Numer. Anal., 38 (2011), pp. 233–257.
[12] S. Kindermann, *Discretization independent convergence rates for noise level-free parameter choice rules for the regularization of ill-conditioned problems*, Electron. Trans. Numer. Anal., 40 (2013), pp. 58–81.
[13] E. Klann and R. Ramlau, *Regularization by fractional filter methods and data smoothing*, Inverse Problems, 24 (2008), 025018.
[14] A. Neuman, L. Reichel, and H. Sadok, *Implementations of range restricted iterative methods for linear discrete ill-posed problems*, Linear Algebra Appl., 436 (2012), pp. 3974–3990.
[15] L. Reichel and G. Rodriguez, *Old and new parameter choice rules for discrete ill-posed problems*, Numer. Algorithms, 63 (2013), pp. 65–87.
[16] L. Reichel and Q. Ye, *Simple square smoothing regularization operators*, Electron. Trans. Numer. Anal., 33 (2009), pp. 63–83.
[17] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.