

PROJECTED TIKHONOV REGULARIZATION OF LARGE-SCALE DISCRETE ILL-POSED PROBLEMS

DAVID R. MARTIN* AND LOTHAR REICHEL†

Abstract. The solution of linear discrete ill-posed problems is very sensitive to perturbations in the data. Confidence intervals for solution coordinates provide insight into the sensitivity. This paper presents an efficient method for computing confidence intervals for large-scale linear discrete ill-posed problems. The method is based on approximating the matrix in these problems by a partial singular value decomposition of low rank. We investigate how to choose the rank. Our analysis also yields novel approaches to the solution of linear discrete ill-posed problems with solution norm or residual norm constraints.

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

1. Introduction. We are concerned with the solution of least-squares problems

$$(1.1) \quad \min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|, \quad \mathbf{A} \in \mathbb{R}^{m \times n}, \quad \mathbf{b} \in \mathbb{R}^m,$$

with a large matrix, whose singular values “cluster” at the origin. Problems of this kind are referred to as linear discrete ill-posed problems; they arise, for instance, from the discretization of linear ill-posed problems, such as Fredholm integral equations of the first kind. We assume for notational simplicity that $m \geq n$; however, our analysis and the methods described can after minor modifications also be applied when $m < n$. The vector \mathbf{b} represents available data, which are assumed to be contaminated by an error $\mathbf{e} \in \mathbb{R}^m$ that can be modeled as white Gaussian noise. Thus,

$$(1.2) \quad \mathbf{b} = \mathbf{b}_{\text{true}} + \mathbf{e},$$

where \mathbf{b}_{true} represents an unavailable error-free vector associated with \mathbf{b} . The linear system of equations with the (unknown) error-free right-hand side,

$$(1.3) \quad \mathbf{A}\mathbf{x} = \mathbf{b}_{\text{true}},$$

is assumed to be consistent. We would like to determine its solution \mathbf{x}_{true} of minimal Euclidean norm by computing an approximate solution of the available least-squares problem (1.1). However, due to the clustering of the singular values of \mathbf{A} at the origin, the solution of (1.1) is very sensitive to the error \mathbf{e} in \mathbf{b} . Let A^\dagger denote the Moore-Penrose pseudoinverse of \mathbf{A} . Then the solution of (1.1),

$$(1.4) \quad A^\dagger \mathbf{b} = A^\dagger \mathbf{b}_{\text{true}} + A^\dagger \mathbf{e} = \mathbf{x}_{\text{true}} + A^\dagger \mathbf{e},$$

generally is dominated by the term $A^\dagger \mathbf{e}$. Therefore, straightforward solution of (1.1) typically does not yield a meaningful approximation of \mathbf{x}_{true} .

To gain insight into the sensitivity of the solution of (1.1) to the error \mathbf{e} , one can determine confidence intervals for the solution coordinates. This can be done by solving constrained minimization problems of the form

$$(1.5) \quad \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{w}^T \mathbf{x} \quad \text{subject to} \quad \|\mathbf{A}\mathbf{x} - \mathbf{b}\| \leq \varepsilon, \quad \|\mathbf{x} - \mathbf{d}\| \leq \delta.$$

*Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA. E-mail: dmarti49@kent.edu. Research supported in part by NSF grant DMS-1115385.

†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.

This minimization problem was first discussed by Eldén [5]. Throughout this paper $\|\cdot\|$ denotes the Euclidean vector norm or the associated induced matrix norm. The parameter $\varepsilon > 0$ is an a priori upper bound for $\|\mathbf{e}\|$, holding with some level of confidence. Such a bound is assumed to be available or computable; see Section 2 for comments on how this parameter can be determined from the distribution of \mathbf{e} . The vector \mathbf{d} is an a priori estimate of the solution, and the parameter $\delta > 0$ bounds the deviation of \mathbf{d} from \mathbf{x}_{true} . The vector \mathbf{w} is an arbitrary unit vector. We are particularly interested in the case $\mathbf{w} = \pm \mathbf{e}_j$, where $\mathbf{e}_j = [0, \dots, 0, 1, 0, \dots, 0]^T$ denotes the j th axis vector. Solving (1.5) with $\mathbf{w} = \pm \mathbf{e}_j$ yields a confidence interval for the j th coordinate of the solution \mathbf{x}_{true} of (1.3). We are interested in the situation when the matrix A is large and confidence intervals for many coordinates of \mathbf{x}_{true} are to be computed.

Eldén [5] presented an efficient solution method for minimization problems (1.5) of small size. This method transforms the problem into a nonlinear root-finding problem involving a sequence of quadratically constrained least-squares problems. We review Eldén’s solution method in Section 2. A minimization problem related to (1.5) was previously discussed by O’Leary and Rust [18]. More recently, Eldén et al. [6] described a method for the solution of large-scale problems (1.5) based on the LSTRS algorithm, which is designed to solve the quadratic minimization problems that arise in large-scale trust-region computations; see [19]. An approach to the solution of large-scale problems (1.5), in which bounds for certain matrix functionals are computed inexpensively by using the connection between the Lanczos process and Gauss quadrature, is described in [15].

The computational effort required by the available solution methods for large-scale problems (1.5) is proportional to the number of vectors \mathbf{w} for which the problem is to be solved. These methods therefore can be quite expensive to use when solutions for many vectors \mathbf{w} are desired, as is the case when confidence intervals for many coordinates of the solution \mathbf{x}_{true} of (1.3) are to be determined.

This paper presents a new solution scheme that is well suited for large-scale problems (1.5) when solutions for many vectors \mathbf{w} are desired. We propose to first approximate A by a matrix of low rank, using a partial singular value decomposition (SVD). A simple iterative method for computing such a decomposition is presented in Section 6. Using this low-rank approximation of A , a lower bound for the minimum (1.5) can be computed quite inexpensively for each vector \mathbf{w} ; see Section 2 for details. We first approximate A by a matrix of rank one. The rank is increased only when for some quadratic minimization problem a specified error level cannot be guaranteed, as described in Section 4.

The determination of a partial SVD of A also may be useful for the problem of computing an approximate solution of the least-squares problem (1.1) by the use of Tikhonov regularization, i.e., by solving a penalized least-squares problem of the form

$$(1.6) \quad \min_{\mathbf{x} \in \mathbb{R}^n} \{ \|A\mathbf{x} - \mathbf{b}\|^2 + \mu \|\mathbf{x}\|^2 \},$$

where $\mu > 0$ is known as the regularization parameter. The solution method for (1.5) described in Section 4 uses the connection between quadratically constrained minimization problems and Tikhonov regularization. We also consider the situation when the value of μ is determined by the discrepancy principle, and illustrate that a hybrid method combining low-rank approximation of A with Tikhonov regularization can give higher accuracy than straightforward solution of (1.1) by truncated SVD (TSVD). Bounds are provided for the norm of the difference between the exact solution

of (1.6) and the approximate solution determined by our hybrid method.

This paper is organized as follows. Section 2 reviews the approach by Eldén [5] for the solution of small-scale minimization problems (1.5) and describes how to apply the method to large-scale problems by first replacing a large matrix A by a partial SVD of low rank. In Section 3 we bound the error in certain functionals caused by replacing the matrix A by a partial SVD in Tikhonov regularization, and Section 4 uses this analysis to derive a criterion for choosing the rank of the partial SVD of A in the context of a constrained least-squares problem. This provides the theoretical foundation of our solution method for (1.5). Section 5 presents a similar analysis when the discrepancy principle is used to determine the parameter μ in Tikhonov regularization (1.6), and Section 6 describes a few computed examples. Concluding remarks can be found in Section 7.

2. Eldén’s method. We outline the method proposed by Eldén [5] for the solution of minimization problems of the form (1.5), after first commenting on the choice of the parameter ε . Ideally, this parameter would represent a fairly tight bound for the norm of the error \mathbf{e} in \mathbf{b} , i.e.,

$$(2.1) \quad \|\mathbf{e}\| \leq \varepsilon.$$

However, in applications such a bound might not be explicitly known. A typical assumption is that the error \mathbf{e} is random in nature. If available, knowledge of the distribution of the error can be used to determine a bound (2.1) holding with probability p , where $0 < p < 1$ is a confidence level specified by the user. For instance, if the coordinates of \mathbf{e} are independent, normally distributed random variables with mean 0 and variance σ^2 (i.e., \mathbf{b}_{true} has been corrupted by white Gaussian noise), then the random variable $\|\mathbf{e}\|^2/\sigma^2$ belongs to the χ^2 distribution with m degrees of freedom. We therefore choose ε so that

$$(2.2) \quad \int_0^{\varepsilon^2/\sigma^2} \chi_m^2(\rho) d\rho = p$$

holds, where χ_m^2 is the probability density function for the χ^2 distribution with m degrees of freedom. Assuming also that an upper bound δ for $\|\mathbf{x}_{\text{true}} - \mathbf{d}\|$ is available (for simplicity, holding with probability 1), solving (1.5) with $\mathbf{w} = \pm \mathbf{e}_j$ provides a confidence interval for the j th coordinate of \mathbf{x}_{true} , holding with probability p ; see [18, 5, 6] for further discussions on this topic.

Eldén [5] transformed the minimization problem (1.5) into an equivalent problem involving the solution of a nonlinear equation as follows. The minimum value $\theta := \mathbf{w}^T \mathbf{x}$ attained in (1.5) is the smallest zero of the nonlinear function

$$(2.3) \quad f(\theta) := L(\theta) - \varepsilon^2$$

in the interval $[\theta_*, \theta^*] := [\mathbf{w}^T \mathbf{d} - \delta, \mathbf{w}^T \mathbf{d} + \delta]$, where L is defined by

$$L(\theta) := \min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 \quad \text{subject to} \quad \|\mathbf{x} - \mathbf{d}\| = \delta, \quad \mathbf{w}^T \mathbf{x} = \theta.$$

This result assumes that both constraints in (1.5) are “active” at the solution, i.e., that equality holds in both of the inequalities; otherwise the analysis and solution of the problem are easier; see [5, 6] for details. Eldén determined the smallest zero

of the function (2.3) by iterating in θ with a zero-finder, after making the following simplifications. The substitution $\mathbf{x}' = \mathbf{x} - \mathbf{d}$ yields the equivalent problem

$$L(\theta) = \min_{\mathbf{x}' \in \mathbb{R}^n} \|\mathbf{A}\mathbf{x}' - \mathbf{b}'\|^2 \quad \text{subject to} \quad \|\mathbf{x}'\| = \delta, \quad \mathbf{w}^T \mathbf{x}' = \theta',$$

where

$$\mathbf{b}' := \mathbf{b} - \mathbf{A}\mathbf{d}, \quad \theta' := \theta - \mathbf{w}^T \mathbf{d}.$$

The linear constraint is eliminated with the substitution

$$\mathbf{x}' = \theta' \mathbf{w} + \mathbf{H}\mathbf{y},$$

where $Q = [\mathbf{w}, \mathbf{H}]$ is an orthogonal matrix, which can be chosen to be a Householder reflector. This yields the equivalent problem

$$(2.4) \quad L(\theta) = \min_{\mathbf{y} \in \mathbb{R}^{n-1}} \|\bar{\mathbf{A}}\mathbf{y} - \bar{\mathbf{b}}\|^2 \quad \text{subject to} \quad \|\mathbf{y}\| = \bar{\delta}.$$

where

$$\begin{aligned} \bar{\mathbf{A}} &:= \mathbf{A}\mathbf{H}, \\ \bar{\mathbf{b}} &:= \mathbf{b}' - \theta' \mathbf{A}\mathbf{w}, \\ \bar{\delta} &:= \sqrt{\delta^2 - (\theta')^2}. \end{aligned}$$

To apply the techniques of Sections 3 and 4 to the problem (2.4), we would like to be able to cheaply determine a partial SVD for each matrix $\bar{\mathbf{A}}$ from one for \mathbf{A} . However, in general a partial SVD for a product of two matrices cannot be determined from a partial SVD of each factor in a simple way. We therefore propose to remove the matrix \mathbf{H} from (2.4), and consider instead the function

$$(2.5) \quad \tilde{L}(\theta) := \min_{\mathbf{y} \in \mathbb{R}^n} \|\mathbf{A}\mathbf{y} - \bar{\mathbf{b}}\|^2 \quad \text{subject to} \quad \|\mathbf{y}\| = \bar{\delta}.$$

Notice that the minimization problem (2.4) can be obtained by appending the minimization problem (2.5) with the additional constraint that \mathbf{y} be in the range of \mathbf{H} . Since the residual $\|\mathbf{A}\mathbf{y} - \bar{\mathbf{b}}\|$ is minimized over a larger subset when solving (2.5), it follows that for a given value of θ , the minimum attained in (2.5) cannot exceed the minimum attained in (2.4); that is, $\tilde{L}(\theta) \leq L(\theta)$ for all values of θ . Recall that the function L was shown to be continuous and convex in [6]. Moreover, as described in [6], in the case of interest (when equality holds in both constraints for the \mathbf{x} that minimizes (1.5)), we have $L(\theta_*) - \varepsilon^2 > 0$. Thus, in the case of interest, L is decreasing on the interval $(\theta_*, \theta_{\text{zero}})$, where θ_{zero} denotes the smallest zero of (2.3), or equivalently the minimum value attained in (1.5). Assuming in addition that $\tilde{L}(\theta_*) - \varepsilon^2 > 0$ (which holds in all examples discussed in Section 6), since \tilde{L} is continuous and $\tilde{L} \leq L$, the minimal zero $\tilde{\theta}_{\text{zero}}$ of

$$(2.6) \quad \tilde{f}(\theta) := \tilde{L}(\theta) - \varepsilon^2$$

is a lower bound for θ_{zero} . Thus, in order to be able to utilize a partial SVD of \mathbf{A} , we compute a lower bound for the minimum attained in (1.5). In the context of computing confidence intervals for components of a linear discrete ill-posed problem

(1.1), elimination of H (that is, use of $\tilde{\theta}_{\text{zero}}$ in place of θ_{zero}) results in potentially looser bounds. We illustrate in Section 6 that for many problems, the elimination of H only has a small impact on the quality of the computed bounds.

When $\tilde{L}(\theta_*) - \varepsilon^2 > 0$, since \tilde{L} is continuous and convex (as shown for L in [6]), assuming the existence of at least one zero of the function (2.6), application of the secant method with initial approximate solutions θ_0 and θ_1 such that $\theta_* = \theta_0 < \theta_1 \ll \theta^*$ produces a sequence of iterates θ_i , $i = 2, 3, \dots$, that converge monotonically to the smallest zero of \tilde{f} , assuming that \tilde{L} is evaluated exactly. Iteration is terminated when an approximation of the smallest zero has been determined with sufficient accuracy, i.e., as soon as

$$(2.7) \quad \left| \tilde{f}(\theta_i) \right| < \tau \varepsilon^2$$

holds for a user-specified parameter τ . We used $\tau = 0.1$ in the computed examples reported in Section 6.

Eldén [5] relies on complete factorization of A , affording very efficient solution of the constrained least-squares problems (2.4). We are interested in the case when A is so large that complete factorization is not feasible, and therefore propose to solve each problem (2.4) arising in the solution of (1.5) by first replacing A by a partial SVD of low rank. A simple iterative method for computing such a decomposition is presented in Section 6. Other methods for computing a partial SVD are described in, e.g., [1, 2, 12]. All these methods use A and A^T only for the evaluation of matrix-vector products. The solution of (1.5) for different vectors \mathbf{w} does not require further evaluations of matrix-vector products with A and A^T if a partial SVD of sufficiently high rank is available.

3. Truncation of Tikhonov regularization. The SVD of A is a factorization

$$A = U \Sigma V^T,$$

where $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ are orthogonal matrices, and Σ is the, possibly rectangular, diagonal matrix

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

The diagonal entries σ_j , known as the singular values of A , are nonnegative and decreasing as a function of their index. The rank- ℓ partial SVD of A is defined as

$$(3.1) \quad A_\ell := U_\ell \Sigma_\ell V_\ell^T,$$

where Σ_ℓ is the $\ell \times \ell$ leading principal submatrix of Σ , and $U_\ell \in \mathbb{R}^{m \times \ell}$ and $V_\ell \in \mathbb{R}^{n \times \ell}$ consist of the first ℓ columns of the matrices U and V , respectively. In our applications, ℓ is fairly small, and much smaller than m and n .

The normal equations associated with the Tikhonov minimization problem (1.6) are given by

$$(3.2) \quad (A^T A + \mu I) \mathbf{x} = A^T \mathbf{b},$$

where I denotes the identity matrix. Let, for $\mu > 0$, the vector \mathbf{x}_μ be the unique solution of (3.2) (and of (1.6)), and introduce the functions

$$(3.3) \quad \varphi(\mu) := \|\mathbf{x}_\mu\|^2,$$

$$(3.4) \quad \psi(\mu) := \|A\mathbf{x}_\mu - \mathbf{b}\|^2.$$

Replacing the matrix A in (1.6) by its rank- ℓ approximant A_ℓ gives, analogously to (3.2),

$$(3.5) \quad \left((A_\ell)^T A_\ell + \mu I \right) \mathbf{x} = (A_\ell)^T \mathbf{b}.$$

We denote the solution of (3.5) for $\mu > 0$ by $\mathbf{x}_\mu^{(\ell)}$ and introduce the functions

$$\varphi_\ell(\mu) := \|\mathbf{x}_\mu^{(\ell)}\|^2 \quad \text{and} \quad \psi_\ell(\mu) := \|A_\ell \mathbf{x}_\mu^{(\ell)} - \mathbf{b}\|^2.$$

Defining $f_\mu(t) := t/(t + \mu)^2$ and inserting the SVD of A into (3.2) yields

$$(3.6) \quad \begin{aligned} \varphi(\mu) &= (A^T \mathbf{b})^T (A^T A + \mu I)^{-2} (A^T \mathbf{b}) \\ &= (U^T \mathbf{b})^T \Sigma (\Sigma^T \Sigma + \mu I)^{-2} \Sigma^T (U^T \mathbf{b}) \\ &= \sum_{i=1}^n f_\mu(\sigma_i^2) (\mathbf{u}_i^T \mathbf{b})^2. \end{aligned}$$

Similarly, we obtain

$$(3.7) \quad \varphi_\ell(\mu) = \sum_{i=1}^{\ell} f_\mu(\sigma_i^2) (\mathbf{u}_i^T \mathbf{b})^2.$$

For $t > 0$ and fixed $\mu > 0$, we have $f'_\mu(t) = (\mu - t)/(t + \mu)^3$. Therefore, the maximum value of f_μ occurs at the point $(\mu, 1/(4\mu))$. Moreover, f_μ is increasing for $t < \mu$ and decreasing for $t > \mu$. It follows that for $\ell + 1 \leq i \leq n$, we have

$$f_\mu(\sigma_i^2) \leq f_{\mu,\ell} := \max_{0 \leq x \leq \sigma_\ell^2} f_\mu(x) = \begin{cases} \frac{1}{4\mu}, & \sigma_\ell^2 \geq \mu, \\ f_\mu(\sigma_\ell^2), & \sigma_\ell^2 < \mu. \end{cases}$$

Hence, the error $\varphi(\mu) - \varphi_\ell(\mu)$ satisfies

$$(3.8) \quad \begin{aligned} \varphi(\mu) - \varphi_\ell(\mu) &= \sum_{i=\ell+1}^n f_\mu(\sigma_i^2) (\mathbf{u}_i^T \mathbf{b})^2 \\ &\leq f_{\mu,\ell} \left(\|\mathbf{b}\|^2 - \sum_{i=1}^{\ell} (\mathbf{u}_i^T \mathbf{b})^2 \right) =: E_\varphi^{(\ell)}(\mu). \end{aligned}$$

By equations (3.6) and (3.7), we have $\varphi_\ell(\mu) \leq \varphi(\mu)$, so that

$$(3.9) \quad \varphi_\ell^-(\mu) := \varphi_\ell(\mu) \leq \varphi(\mu) \leq \varphi_\ell(\mu) + E_\varphi^{(\ell)}(\mu) =: \varphi_\ell^+(\mu).$$

THEOREM 3.1. *Let φ_ℓ^- and φ_ℓ^+ be as defined in (3.9). Then*

$$\varphi_\ell^-(\mu) \leq \varphi_{\ell+1}^-(\mu) \leq \varphi(\mu) \leq \varphi_{\ell+1}^+(\mu) \leq \varphi_\ell^+(\mu)$$

holds for all $\mu > 0$ and $\ell \geq 1$.

Proof. It is clear from the representations (3.6) and (3.7) that

$$(3.10) \quad \varphi_\ell^-(\mu) \leq \varphi_{\ell+1}^-(\mu) \leq \varphi(\mu)$$

holds for all $\ell \geq 1$ and $\mu > 0$, where both inequalities are generically strict. Writing

$$\varphi_\ell^+(\mu) = \sum_{i=1}^{\ell} f_\mu(\sigma_i^2) (\mathbf{u}_i^T \mathbf{b})^2 + f_{\mu,\ell} \left(\|\mathbf{b}\|^2 - \sum_{i=1}^{\ell} (\mathbf{u}_i^T \mathbf{b})^2 \right),$$

the difference $\varphi_{\ell+1}^+(\mu) - \varphi_\ell^+(\mu)$ satisfies

$$\begin{aligned} \varphi_{\ell+1}^+(\mu) - \varphi_\ell^+(\mu) &= (f_\mu(\sigma_{\ell+1}^2) - f_{\mu,\ell}) (\mathbf{u}_{\ell+1}^T \mathbf{b})^2 \\ &\quad + (f_{\mu,\ell+1} - f_{\mu,\ell}) \left(\|\mathbf{b}\|^2 - \sum_{i=1}^{\ell+1} (\mathbf{u}_i^T \mathbf{b})^2 \right) \leq 0, \end{aligned}$$

so that, analogous to (3.10),

$$(3.11) \quad \varphi(\mu) \leq \varphi_{\ell+1}^+(\mu) \leq \varphi_\ell^+(\mu)$$

holds for all $\ell \geq 1$ and $\mu > 0$, where both inequalities are generically strict. \square

A representation for ψ analogous to (3.6) is given by

$$\begin{aligned} (3.12) \quad \psi(\mu) &= (A(A^T A + \mu I)^{-1} A^T \mathbf{b} - \mathbf{b})^T (A(A^T A + \mu I)^{-1} A^T \mathbf{b} - \mathbf{b}) \\ &= \mu^2 \mathbf{b}^T (A A^T + \mu I)^{-2} \mathbf{b} \\ &= \mu^2 (U^T \mathbf{b})^T (\Sigma \Sigma^T + \mu I)^{-2} (U^T \mathbf{b}) \\ &= \sum_{i=1}^n \frac{\mu^2}{(\sigma_i^2 + \mu)^2} (\mathbf{u}_i^T \mathbf{b})^2 + \left(\|\mathbf{b}\|^2 - \sum_{i=1}^n (\mathbf{u}_i^T \mathbf{b})^2 \right) \\ &= \|\mathbf{b}\|^2 + \sum_{i=1}^n \left(\frac{\mu^2}{(\sigma_i^2 + \mu)^2} - 1 \right) (\mathbf{u}_i^T \mathbf{b})^2, \end{aligned}$$

where we have used the identity

$$(3.13) \quad I - A(A^T A + \mu I)^{-1} A^T = \mu(A A^T + \mu I)^{-1},$$

which can be shown, e.g., by multiplying (3.13) by $A A^T + \mu I$. Similarly, we have

$$(3.14) \quad \psi_\ell(\mu) = \|\mathbf{b}\|^2 + \sum_{i=1}^{\ell} \left(\frac{\mu^2}{(\sigma_i^2 + \mu)^2} - 1 \right) (\mathbf{u}_i^T \mathbf{b})^2$$

and, therefore,

$$\begin{aligned} \psi_\ell(\mu) - \psi(\mu) &= \sum_{i=\ell+1}^n \left(1 - \frac{\mu^2}{(\sigma_i^2 + \mu)^2} \right) (\mathbf{u}_i^T \mathbf{b})^2 \\ &\leq \left(1 - \frac{\mu^2}{(\sigma_\ell^2 + \mu)^2} \right) \left(\|\mathbf{b}\|^2 - \sum_{i=1}^{\ell} (\mathbf{u}_i^T \mathbf{b})^2 \right) =: E_\psi^{(\ell)}(\mu). \end{aligned}$$

This bound easily can be evaluated when the partial SVD of A (3.1) is available. The sharper bound that would be obtained by replacing σ_ℓ by $\sigma_{\ell+1}$ in $E_\psi^{(\ell)}(\mu)$ cannot be computed without increasing ℓ in (3.1). We therefore do not use the latter.

It follows from (3.12) and (3.14) that $\psi_\ell(\mu) \geq \psi(\mu)$. Therefore,

$$(3.15) \quad \psi_\ell^-(\mu) := \psi_\ell(\mu) - E_\psi^{(\ell)}(\mu) \leq \psi(\mu) \leq \psi_\ell(\mu) =: \psi_\ell^+(\mu).$$

Analogously to Theorem 3.1, we have the following result.

THEOREM 3.2. *Let ψ_ℓ^- and ψ_ℓ^+ be as defined in (3.15). Then*

$$\psi_\ell^-(\mu) \leq \psi_{\ell+1}^-(\mu) \leq \psi(\mu) \leq \psi_{\ell+1}^+(\mu) \leq \psi_\ell^+(\mu)$$

holds for all $\mu > 0$ and $\ell \geq 1$.

Proof. It follows from the representations (3.12) and (3.14) that

$$(3.16) \quad \psi(\mu) \leq \psi_{\ell+1}^+(\mu) \leq \psi_\ell^+(\mu)$$

holds for all $\ell \geq 1$ and $\mu > 0$, and similarly one can show that

$$(3.17) \quad \psi_\ell^-(\mu) \leq \psi_{\ell+1}^-(\mu) \leq \psi(\mu).$$

Again, all inequalities in (3.16) and (3.17) are generically strict. \square

4. Constrained minimization. We are interested in the solution of constrained least-squares problems of the form

$$(4.1) \quad \min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 \quad \text{subject to} \quad \|\mathbf{x}\| = \delta,$$

because they arise in the evaluation of $\tilde{L}(\theta)$ when $\mathbf{b} = \bar{\mathbf{b}}$; cf. (2.5). It is easy to show, e.g., by using a Lagrange multiplier, that the Tikhonov minimization problem (1.6) for a suitable finite value of the regularization parameter $\mu = \mu_\delta > 0$ yields the solution \mathbf{x}_{μ_δ} of (4.1). Therefore, the constrained least-squares problem (4.1) is equivalent to the nonlinear equation

$$(4.2) \quad \varphi(\mu) = \delta^2,$$

where φ is defined by (3.3). We also note that the minimization problem (4.1) is closely related to the trust-region subproblem that arises in optimization methods; see, e.g., [16, 19] for discussions of the latter.

The function φ is expensive to evaluate when the matrix A is large. Therefore, we instead calculate φ_ℓ^- for one or a few small values of ℓ . We begin with $\ell = 1$, and increase ℓ only when necessary. In this section, the primary goal when solving (4.1) is to evaluate $\psi(\mu_\delta)$, since this is precisely the value of $\tilde{L}(\theta)$ in (2.5) when $\mathbf{b} = \bar{\mathbf{b}}$. We seek to bracket the solution μ_δ of (4.2), thereby bracketing $\psi(\mu_\delta)$. This can be accomplished as follows. We first determine approximations of the solutions μ_ℓ^- and μ_ℓ^+ , respectively, of the equations

$$(4.3) \quad \varphi_\ell^-(\mu) = \delta^2,$$

$$(4.4) \quad \varphi_\ell^+(\mu) = \delta^2.$$

From (3.7), (3.8), and (3.9), it follows that both $\varphi_\ell^-(\mu)$ and $\varphi_\ell^+(\mu)$ are decreasing functions of $\mu > 0$. Therefore, since $\varphi_\ell^- \leq \varphi \leq \varphi_\ell^+$, we conclude that

$$(4.5) \quad \mu_\ell^- \leq \mu_\delta \leq \mu_\ell^+.$$

Moreover, since $\varphi_\ell^-(\mu)$ and $\varphi_\ell^+(\mu)$ are continuous for $\mu > 0$, the equations (4.3)-(4.4) have unique bounded solutions whenever the inequalities

$$(4.6) \quad 0 = \lim_{\mu \rightarrow \infty} \varphi_\ell^-(\mu) < \delta^2 < \lim_{\mu \rightarrow 0} \varphi_\ell^-(\mu) = \varphi_\ell^-(0),$$

$$(4.7) \quad 0 = \lim_{\mu \rightarrow \infty} \varphi_\ell^+(\mu) < \delta^2 < \lim_{\mu \rightarrow 0} \varphi_\ell^+(\mu) = \infty$$

hold. We begin by ensuring that the last inequality of (4.6) is valid by increasing ℓ if necessary. The inequality

$$(4.8) \quad \|A^\dagger \mathbf{b}\| > \delta.$$

may be assumed to hold, because we will choose δ to be of the order $\|\mathbf{x}_{\text{true}}\|$, while $\|A^\dagger \mathbf{b}\|$ typically is much larger; see the discussion following (1.4). Note that $\varphi(0) = \|A^\dagger \mathbf{b}\|^2$. Since $\varphi_\ell^-(0) = \varphi_\ell(0) \rightarrow \varphi(0)$ as ℓ increases, by (4.8) the last inequality of (4.6) holds for all sufficiently large ℓ . The inequalities (4.7) hold whenever $0 < \delta < \infty$.

Since φ_ℓ^- is decreasing and convex, application of Newton's method or the zero-finder discussed by Golub and von Matt [8, equation (74)] with an initial iterate $\mu_{\ell,0}^- < \mu_\ell^-$ will produce a sequence of iterates $\mu_{\ell,j}^-$, $j = 1, 2, 3, \dots$, that converge to μ_ℓ^- and satisfy

$$(4.9) \quad \mu_{\ell,0}^- \leq \mu_{\ell,1}^- \leq \dots \leq \mu_\ell^-.$$

We terminate the iterations as soon as for some $q = q_\ell$ the inequality

$$(4.10) \quad \frac{\varphi_\ell^-(\mu_{\ell,q}^-) - \delta^2}{\delta^2} < \nu$$

holds for a user-specified parameter $0 < \nu \ll 1$. In the examples reported in Section 6, we let $\nu = 1 \times 10^{-5}$ and use the initial iterate $\mu_{\ell_0,0}^- := 0$, where ℓ_0 is the first value of ℓ for which both (4.6) and (4.7) hold. For $\ell > \ell_0$, we define $\mu_{\ell,0}^- := \mu_{\ell-1,q}^-$. This ensures that $\mu_{\ell,0}^- \leq \mu_\ell^-$ by (3.10).

We turn to the solution of equation (4.4) and employ the monotonically and quadratically convergent zero-finder described by Golub and von Matt [8, equations (75)-(78)] with an initial iterate $\mu_{\ell,0}^+ > \mu_\ell^+$ to obtain a sequence of iterates $\mu_{\ell,j}^+$, $j = 1, 2, 3, \dots$, that converge to μ_ℓ^+ and satisfy

$$(4.11) \quad \mu_\ell^+ \leq \dots \leq \mu_{\ell,1}^+ \leq \mu_{\ell,0}^+.$$

Iteration is terminated as soon as for some q the inequality

$$(4.12) \quad \frac{\delta^2 - \varphi_\ell^+(\mu_{\ell,q}^+)}{\delta^2} < \nu$$

holds. We choose an arbitrary initial iterate $\mu_{\ell_0,0}^+$ such that $\varphi_\ell^+(\mu_{\ell_0,0}^+) < \delta^2$. For $\ell > \ell_0$, we define $\mu_{\ell,0}^+ := \mu_{\ell-1,q}^+$. Then $\mu_{\ell,0}^+ \geq \mu_\ell^+$ by (3.11).

It is clear from the representation (3.12) that the function ψ is increasing for $\mu > 0$. Therefore, it follows from (4.5) that $\psi(\mu_\ell^-) \leq \psi(\mu_\delta) \leq \psi(\mu_\ell^+)$. Moreover, by (4.9) and (4.11), we have

$$\psi_\ell^-(\mu_{\ell,j}^-) \leq \psi(\mu_{\ell,j}^-) \leq \psi(\mu_\delta) \leq \psi(\mu_{\ell,j}^+) \leq \psi_\ell^+(\mu_{\ell,j}^+).$$

Letting γ represent the maximum allowable relative error in the computed value of $\psi(\mu_\delta)$ (i.e., of $\tilde{L}(\theta)$), we require

$$(4.13) \quad \frac{\psi_\ell^+(\mu_{\ell,q}^+) - \psi_\ell^-(\mu_{\ell,q}^-)}{2\psi_\ell^-(\mu_{\ell,q}^-)} < \gamma.$$

We accept the approximation $\psi(\mu_\delta) \approx (\psi_\ell^-(\mu_{\ell,q}^-) + \psi_\ell^+(\mu_{\ell,q}^+))/2$ when the inequality (4.13) holds; otherwise we increase ℓ and repeat the above process. The inequality (4.13) trivially holds for $\ell = n$, at least for sufficiently small ν . In the case of interest, i.e., when the matrix A and vector \mathbf{b} have the properties described following (1.1), the inequality (4.13) typically holds for some $\ell \ll n$ for reasonable values of γ . We use $\gamma = 0.01$ in the examples reported in Section 6. The following algorithm summarizes the computations.

ALGORITHM 4.1 (Solution of the minimization problem (1.5)).

Transform the problem (1.5) to (2.5);

$\ell := \ell_0 - 1;$

while (4.10) *does not hold do*

$\ell := \ell + 1;$ *Solve equation (4.3) for $\mu_{\ell,q}^-$;*

end while

$\ell := \ell - 1;$

while (4.12) *does not hold do*

$\ell := \ell + 1;$ *Solve equation (4.4) for $\mu_{\ell,q}^+$;*

end while

$\ell := \ell - 1;$

while (4.13) *does not hold do*

$\ell := \ell + 1;$ *Solve equations (4.3) and (4.4) for $\mu_{\ell,q}^-$ and $\mu_{\ell,q}^+$, respectively;*

end while □

The average $(\psi_\ell^-(\mu_{\ell,q}^-) + \psi_\ell^+(\mu_{\ell,q}^+))/2$ furnished by the algorithm is our computed approximation of (2.5).

5. Application of the discrepancy principle. This section is concerned with the computation of an approximate solution of (1.4) using the discrepancy principle to determine a suitable value of the regularization parameter in Tikhonov regularization. The discrepancy principle can be applied when a bound (2.1) of the error in the data \mathbf{b} is available or can be estimated, e.g., by using (2.2). The discrepancy principle prescribes that the regularization parameter $\mu > 0$ in (1.6) be determined so that

$$(5.1) \quad \psi(\mu) = \alpha^2 \varepsilon^2,$$

where ψ is defined by (3.4) and $\alpha > 1$ is a constant independent of ε . Thus, we require the solution \mathbf{x}_μ of the Tikhonov minimization problem (1.6) to satisfy $\|A\mathbf{x}_\mu - \mathbf{b}\| = \alpha\varepsilon$. Let μ_ε denote the solution of (5.1). One can show that $\mathbf{x}_{\mu_\varepsilon} \rightarrow \mathbf{x}_{\text{true}}$ as $\varepsilon \searrow 0$; see, e.g., [7, 14] for proofs in Hilbert space settings. The parameter α can be chosen arbitrarily close to unity in these proofs. We set $\alpha = 1$ in the computed examples as well as in the remainder of this section. Perhaps a more intuitive motivation for the discrepancy principle is the following. By premise, \mathbf{x}_{true} satisfies $\|A\mathbf{x}_{\text{true}} - \mathbf{b}\| = \|\mathbf{b}_{\text{true}} - \mathbf{b}\| = \|\mathbf{e}\| \leq \varepsilon$. The discrepancy principle dictates use of the largest value of μ such that $\|A\mathbf{x}_\mu - \mathbf{b}\| \leq \varepsilon$ or, equivalently, the solution of the best conditioned problem (3.2) for which this inequality holds.

The function ψ is expensive to evaluate when the matrix A is large. Similarly as in the previous section, we therefore replace A by a partial singular value decomposition A_ℓ of low rank ℓ . Hence, we apply the discrepancy principle to the Tikhonov minimization problem (1.6) with the matrix A replaced by A_ℓ , i.e., we solve (3.5). We refer to this as the Truncated Tikhonov (TT) method. The TT method provides an alternative to the approximate solution of (1.1) by TSVD [7, 10], as well as to the approximate solution of the Tikhonov minimization problem (1.6) by partial Lanczos bidiagonalization of A ; see, e.g., [4, 9]. The TT method can be cheaper to apply than the latter scheme when the Tikhonov minimization problem (1.6) has to be solved for many data vectors \mathbf{b} . Moreover, the TT method yields, on the average, somewhat more accurate approximations of \mathbf{x}_{true} than the TSVD method. This is illustrated in Section 6. Further comments on the TSVD can be found at the end of this section.

We approximate the function ψ by ψ_ℓ^+ and ψ_ℓ^- similarly as in Section 4, and start with $\ell = 1$. The rank ℓ of A_ℓ is increased only if necessary. We seek to bracket the solution μ_ε of (5.1) in order to be able to bound the error in the computed approximation of the solution $\mathbf{x}_{\mu_\varepsilon}$ of (1.6) with $\mu = \mu_\varepsilon$. Note that differently from φ , the function ψ is not convex. This potentially makes the computation of the root of (5.1) complicated, because root-finders for non-convex functions generally have to be safeguarded, while this may not be required for root-finders for decreasing convex functions. For instance, Newton's method does not have to be safeguarded when applied to such a function with a suitable initial iterate. We therefore make the substitution $\lambda := 1/\mu$ and introduce

$$\begin{aligned}\Psi(\lambda) &:= \psi(1/\lambda) = \|\mathbf{b}\|^2 + \sum_{i=1}^n \left(\frac{1}{(\lambda\sigma_i^2 + 1)^2} - 1 \right) (\mathbf{u}_i^T \mathbf{b})^2, \\ \Psi_\ell^+(\lambda) &:= \psi_\ell^+(1/\lambda) = \|\mathbf{b}\|^2 + \sum_{i=1}^\ell \left(\frac{1}{(\lambda\sigma_i^2 + 1)^2} - 1 \right) (\mathbf{u}_i^T \mathbf{b})^2, \\ \Psi_\ell^-(\lambda) &:= \psi_\ell^-(1/\lambda) = \|\mathbf{b}\|^2 + \sum_{i=1}^\ell \left(\frac{1}{(\lambda\sigma_i^2 + 1)^2} - 1 \right) (\mathbf{u}_i^T \mathbf{b})^2 \\ &\quad - \left(1 - \frac{1}{(\lambda\sigma_\ell^2 + 1)^2} \right) \left(\|\mathbf{b}\|^2 - \sum_{i=1}^\ell (\mathbf{u}_i^T \mathbf{b})^2 \right) \\ &= \frac{1}{(\lambda\sigma_\ell^2 + 1)^2} \left(\|\mathbf{b}\|^2 - \sum_{i=1}^\ell (\mathbf{u}_i^T \mathbf{b})^2 \right) + \sum_{i=1}^\ell \frac{(\mathbf{u}_i^T \mathbf{b})^2}{(\lambda\sigma_i^2 + 1)^2}.\end{aligned}$$

All three of these functions are decreasing and convex. We first approximate the solutions λ_ℓ^- and λ_ℓ^+ , respectively, of the equations

$$\begin{aligned}\Psi_\ell^-(\lambda) &= \varepsilon^2, \\ \Psi_\ell^+(\lambda) &= \varepsilon^2,\end{aligned}$$

which have unique bounded solutions whenever the inequalities

$$(5.2) \quad 0 = \lim_{\lambda \rightarrow \infty} \Psi_\ell^-(\lambda) < \varepsilon^2 < \Psi_\ell^-(0) = \|\mathbf{b}\|^2,$$

$$(5.3) \quad \|\mathbf{b}\|^2 - \sum_{i=1}^\ell (\mathbf{u}_i^T \mathbf{b})^2 = \lim_{\lambda \rightarrow \infty} \Psi_\ell^+(\lambda) < \varepsilon^2 < \Psi_\ell^+(0) = \|\mathbf{b}\|^2$$

hold. We ensure that the inequalities (5.2) and (5.3) hold by increasing ℓ if needed. One may reasonably assume that $\|\mathbf{e}\| < \|\mathbf{b}\|$, in which case $\varepsilon^2 < \|\mathbf{b}\|^2$ holds and the inequalities (5.2) are valid for all $\varepsilon > 0$ sufficiently close to $\|\mathbf{e}\|$. Let r denote the rank of A . By assumption, \mathbf{b}_{true} is in the range of A , or equivalently in the span of the first r left singular vectors \mathbf{u}_i of A . Using (1.2), we obtain for ℓ sufficiently large that

$$\begin{aligned} \sqrt{\|\mathbf{b}\|^2 - \sum_{i=1}^{\ell} (\mathbf{u}_i^T \mathbf{b})^2} &= \|(I - U_{\ell} U_{\ell}^T) \mathbf{b}\| \\ &\leq \|(I - U_{\ell} U_{\ell}^T) \mathbf{b}_{\text{true}}\| + \|(I - U_{\ell} U_{\ell}^T) \mathbf{e}\| < \|\mathbf{e}\|, \end{aligned}$$

where the last inequality holds with probability 1. To see this, first note that $\|(I - U_{\ell} U_{\ell}^T) \mathbf{b}_{\text{true}}\|$ vanishes for $\ell \geq r$. Moreover, since \mathbf{e} is assumed to represent white Gaussian noise, we have

$$\mathbf{u}_i^T \mathbf{e} \neq 0, \quad 1 \leq i \leq \ell,$$

with probability 1. Therefore, when $\varepsilon > 0$ is sufficiently close to $\|\mathbf{e}\|$, there exists (with probability 1) some $\ell \leq r$ for which the inequality (5.3) holds. In fact, the smallest ℓ for which (5.3) holds is the number of singular triplets $(\sigma_i, \mathbf{u}_i, \mathbf{v}_i)$ required by the TSVD when determining the truncation index by the discrepancy principle. This method is discussed at the end of this section. As illustrated in Section 6, for the problems of interest described following (1.1), the left inequality (5.3) typically holds already for some $\ell \ll r$ when ε is close to $\|\mathbf{e}\|$.

Because $\Psi_{\ell}^{-} \leq \Psi \leq \Psi_{\ell}^{+}$ and all three functions are decreasing for $\lambda > 0$, it follows that $\lambda_{\ell}^{-} \leq \lambda_{\varepsilon} \leq \lambda_{\ell}^{+}$, where λ_{ε} solves $\Psi(\lambda_{\varepsilon}) = \varepsilon^2$; equivalently, $\lambda_{\varepsilon} = 1/\mu_{\varepsilon}$. In view of that Ψ_{ℓ}^{-} is convex, application of Newton's method or the zero-finder discussed by Golub and von Matt [8, equation (74)] with any initial iterate $\lambda_{\ell,0}^{-} < \lambda_{\ell}^{-}$ will produce a sequence of iterates $\lambda_{\ell,j}^{-}$, $j = 1, 2, 3, \dots$, that converge to λ_{ℓ}^{-} and satisfy

$$\lambda_{\ell,0}^{-} \leq \lambda_{\ell,1}^{-} \leq \dots \leq \lambda_{\ell}^{-}.$$

We terminate the iterations as soon as for some $q = q_{\ell}$ the inequality

$$\frac{\Psi_{\ell}^{-}(\lambda_{\ell,q}^{-}) - \varepsilon^2}{\varepsilon^2} < \nu$$

holds, where $0 < \nu \ll 1$ is a user-specified parameter. This yields the upper bound $\hat{\mu}_{\ell}^{+} := \lambda_{\ell,q}^{-}$ for μ_{ε} . The initial iterate is chosen as $\lambda_{\ell,0}^{-} := 0$, where ℓ_0 is the smallest value of ℓ such that both (5.2) and (5.3) are satisfied. For $\ell > \ell_0$, we let $\lambda_{\ell,0}^{-} := \lambda_{\ell-1,q}^{-}$. This ensures that $\lambda_{\ell,0}^{-} \leq \lambda_{\ell}^{-}$ by (3.17).

Next we employ the monotonically and quadratically convergent zero-finder described by Golub and von Matt [8, equations (75)-(78)] with an initial iterate $\lambda_{\ell,0}^{+} > \lambda_{\ell}^{+}$ to obtain a sequence $\lambda_{\ell,j}^{+}$, $j = 1, 2, 3, \dots$, that converges to λ_{ℓ}^{+} and satisfies

$$\lambda_{\ell}^{+} \leq \dots \leq \lambda_{\ell,1}^{+} \leq \lambda_{\ell,0}^{+}.$$

Iteration is terminated as soon as the inequality

$$\frac{\varepsilon^2 - \Psi_{\ell}^{+}(\lambda_{\ell,q}^{+})}{\varepsilon^2} < \nu$$

holds for some q , and we let $\hat{\mu}_\ell^- := \lambda_{\ell,q}^+$ (as this is a lower bound for μ_ε). The initial iterate $\lambda_{\ell_0,0}^+$ is chosen so that $\Psi_\ell^+(\lambda_{\ell_0,0}^+) < \varepsilon^2$ holds. For $\ell > \ell_0$, we let $\lambda_{\ell,0}^+ := \lambda_{\ell-1,q}^+$, which ensures that $\lambda_{\ell,0}^+ \geq \lambda_\ell^+$ by (3.16).

Let γ represent the maximum allowable relative error in the computed approximation of $\mathbf{x}_{\mu_\varepsilon}$. Defining

$$\mu_\varepsilon^{(\ell)} := \frac{\hat{\mu}_\ell^- + \hat{\mu}_\ell^+}{2},$$

we seek an upper bound for the error $\|\mathbf{x}_{\mu_\varepsilon^{(\ell)}}^{(\ell)} - \mathbf{x}_{\mu_\varepsilon}\|$. This can be accomplished via

$$(5.4) \quad \left\| \mathbf{x}_{\mu_\varepsilon^{(\ell)}}^{(\ell)} - \mathbf{x}_{\mu_\varepsilon} \right\| \leq \left\| \mathbf{x}_{\mu_\varepsilon^{(\ell)}}^{(\ell)} - \mathbf{x}_{\mu_\varepsilon}^{(\ell)} \right\| + \left\| \mathbf{x}_{\mu_\varepsilon}^{(\ell)} - \mathbf{x}_{\mu_\varepsilon} \right\|,$$

using the bound

$$(5.5) \quad \left\| \mathbf{x}_{\mu_\varepsilon^{(\ell)}}^{(\ell)} - \mathbf{x}_{\mu_\varepsilon}^{(\ell)} \right\| \leq \max \left\{ \left\| \mathbf{x}_{\hat{\mu}_\ell^+}^{(\ell)} - \mathbf{x}_{\mu_\varepsilon}^{(\ell)} \right\|, \left\| \mathbf{x}_{\mu_\varepsilon}^{(\ell)} - \mathbf{x}_{\hat{\mu}_\ell^-}^{(\ell)} \right\| \right\}$$

and, following (3.8),

$$(5.6) \quad \begin{aligned} \left\| \mathbf{x}_{\mu_\varepsilon}^{(\ell)} - \mathbf{x}_{\mu_\varepsilon} \right\|^2 &= \varphi(\mu_\varepsilon) - \varphi_\ell(\mu_\varepsilon) \\ &= \sum_{i=\ell+1}^n f_{\mu_\varepsilon}(\sigma_i^2) (\mathbf{u}_i^T \mathbf{b})^2 \\ &\leq \sum_{i=\ell+1}^n f_{\hat{\mu}_\ell^-}(\sigma_i^2) (\mathbf{u}_i^T \mathbf{b})^2 \\ &\leq f_{\hat{\mu}_\ell^-, \ell} \left(\|\mathbf{b}\|^2 - \sum_{i=1}^{\ell} (\mathbf{u}_i^T \mathbf{b})^2 \right). \end{aligned}$$

When the bounds (5.4), (5.5), and (5.6) together imply the last inequality of

$$\frac{\left\| \mathbf{x}_{\mu_\varepsilon^{(\ell)}}^{(\ell)} - \mathbf{x}_{\mu_\varepsilon} \right\|}{\left\| \mathbf{x}_{\mu_\varepsilon} \right\|} \leq \frac{\left\| \mathbf{x}_{\mu_\varepsilon^{(\ell)}}^{(\ell)} - \mathbf{x}_{\mu_\varepsilon} \right\|}{\left\| \mathbf{x}_{\hat{\mu}_\ell^+} \right\|} \leq \gamma,$$

we accept $\mathbf{x}_{\mu_\varepsilon^{(\ell)}}^{(\ell)}$ as our approximation of $\mathbf{x}_{\mu_\varepsilon}$.

The TT method described above applies the discrepancy principle to Tikhonov regularization by using a low-rank partial SVD of A . This method is an alternative to applying the discrepancy principle in conjunction with the TSVD [7, 10]. The latter method seeks an approximate solution of (1.1) of the form

$$\mathbf{x}_\ell := \sum_{i=1}^{\ell} \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{v}_i,$$

where ℓ is chosen as small as possible so that $\|A\mathbf{x}_\ell - \mathbf{b}\| < \varepsilon$ holds. We will refer to this method as “TSVD discrepancy” for conciseness. The numerical examples of Section 6 demonstrate that for some problems, the TT method yields a more accurate approximation of \mathbf{x}_{true} than TSVD discrepancy.

We applied the zero-finders described by Golub and von Matt [8] in this and the previous sections. Zero-finders also are discussed by Moré and Sorensen [16] and Morozov [17, Section 26]. When ℓ is not large, which is the situation of interest, the computational effort required by the zero-finder is insignificant. We therefore do not dwell on the choice of zero-finder further.

This section focused on the use of the discrepancy principle for determining a value of the regularization parameter, because the discrepancy principle often yields a suitable value. However, partial singular value decomposition of the matrix A also can be used in conjunction with other methods for determining the regularization parameter, such as the L-curve criterion, generalized cross validation, and the quasi-optimality principle; see [7, 10, 13] for discussions on properties of these methods.

6. Numerical examples. The computed examples of this section illustrate the performance of the methods described in Sections 4 and 5. All numerical experiments were conducted in MATLAB R2011a with $\epsilon_{\text{machine}} = 2.22 \times 10^{-16}$. We consider 6 linear discrete ill-posed problems that arise from the discretization of Fredholm integral equations of the first kind with a smooth kernel. All problems were generated by the functions `baart`, `shaw`, `deriv2`, `phillips`, `foxgood`, and `heat` from the MATLAB package *Regularization Tools* by Hansen [11]. Each function was used to determine a matrix $A \in \mathbb{R}^{1024 \times 1024}$ and the desired solution \mathbf{x}_{true} . We let $\mathbf{b}_{\text{true}} := A\mathbf{x}_{\text{true}}$ and generated an “error vector” \mathbf{e} with normally distributed random entries with zero mean, scaled to correspond to a prescribed relative error $\|\mathbf{e}\|/\|\mathbf{b}_{\text{true}}\|$. The contaminated data vector \mathbf{b} was determined using (1.2).

Each partial SVD was computed with the aid of a simple algorithm based on partial Lanczos bidiagonalization. Application of k steps of partial Lanczos bidiagonalization to the matrix A with random initial unit vector $\mathbf{p}_1 \in \mathbb{R}^n$ yields the decompositions

$$(6.1) \quad \begin{aligned} AP_k &= Q_k B_k, \\ A^T Q_k &= P_k B_k^T + \mathbf{r}_k \mathbf{e}_k^T, \end{aligned}$$

where $P_k \in \mathbb{R}^{n \times k}$ satisfies $P_k^T P_k = I$, $Q_k \in \mathbb{R}^{m \times k}$ satisfies $Q_k^T Q_k = I$, $P_k \mathbf{e}_1 = \mathbf{p}_1$, $P_k^T \mathbf{r}_k = 0$, and $B_k \in \mathbb{R}^{k \times k}$ is upper bidiagonal. Reorthogonalization was used when computing the columns of P_k and Q_k . We assume for notational simplicity that no breakdown occurs, i.e., that all diagonal and superdiagonal entries of B_k are non-zero; see, e.g., [3] for further discussion on Lanczos bidiagonalization, and [1, 2] for more details on the application of Lanczos bidiagonalization to the computation of singular triplets. From (6.1) it follows that

$$\begin{aligned} A^T A P_k &= P_k B_k^T B_k + \text{scalar} \cdot \mathbf{r}_k \mathbf{e}_k^T &\approx P_k B_k^T B_k, \\ A A^T Q_k &= Q_k B_k B_k^T + A \mathbf{r}_k \mathbf{e}_k^T &\approx Q_k B_k B_k^T. \end{aligned}$$

Let

$$B_k = \hat{U}_k \hat{\Sigma}_k \hat{V}_k^T$$

be an SVD of B_k , where $\hat{\Sigma}_k = \text{diag}[\hat{\sigma}_1^{(k)}, \dots, \hat{\sigma}_k^{(k)}]$. We use $\tilde{\sigma}_i^{(k)} := \hat{\sigma}_i^{(k)}$, $1 \leq i \leq \ell$, as approximations of the first ℓ singular values of A , provided that they do not change much with k ; see below. The corresponding approximate left and right singular vectors $\tilde{\mathbf{u}}_i^{(k)}$ and $\tilde{\mathbf{v}}_i^{(k)}$ of A are given by the i th columns of $Q_k \hat{U}_k$ and $P_k \hat{V}_k$, respectively, for $i = 1, 2, \dots, \ell$. If approximations of ℓ singular triplets $(\sigma_i, \mathbf{u}_i, \mathbf{v}_i)$ are required, then k

is increased until (for some $k \geq \ell$) the relative change in each of the ℓ first approximate singular values $\tilde{\sigma}_i^{(k)}$ of A satisfies

$$\frac{|\tilde{\sigma}_i^{(k)} - \tilde{\sigma}_i^{(k-1)}|}{\tilde{\sigma}_i^{(k)}} < \text{tol}, \quad 1 \leq i \leq \ell.$$

For all examples in this paper, we use $\text{tol} = 1 \times 10^{-6}$.

This way of computing singular triplets is well-suited for problems in which the number, ℓ , of desired triplets is increased incrementally, as in the methods of Sections 4 and 5; see Table 6.1 for details regarding the computational expense and accuracy. In order for the above method to be successful, the components $\mathbf{v}_i^T \mathbf{p}_1$, $1 \leq i \leq k$, of the initial vector \mathbf{p}_1 must be sufficiently large in magnitude. In practice, random initial vectors \mathbf{p}_1 for which the algorithm performs poorly are exceedingly rare. Each result of Table 6.1 was computed using a different random initial vector.

| Problem | ℓ | MVP | MRE |
|----------|--------|-----|------------|
| baart | 5 | 13 | 0 |
| shaw | 10 | 25 | $5.03e-11$ |
| deriv2 | 43 | 123 | $4.59e-07$ |
| phillips | 16 | 53 | $1.36e-11$ |
| foxgood | 5 | 15 | $2.44e-08$ |
| heat | 49 | 121 | $1.52e-09$ |

TABLE 6.1

For each rank- ℓ partial SVD reported in Table 6.7, the number of required matrix-vector products (MVP) with either A or A^T and the maximum relative error (MRE) over all computed singular values for the algorithm described above.

6.1. Confidence intervals and constrained minimization. We begin with a numerical example that exhibits the performance of the method described in Section 2 in the context of computing confidence intervals for components of the solution \mathbf{x}_{true} of (1.1). Each problem (2.5) that arises when solving (2.6) in order to compute a lower bound for the minimum attained in (1.5) is solved using the method described in Section 4.

For each one of the 6 discrete ill-posed problems considered, bounds are computed for 16 equally spaced coordinates of the solution, by computing a lower bound for the minimum attained in (1.5), via solution of (2.6), with $\mathbf{d} = \mathbf{0}$, $\delta = \|\mathbf{x}_{\text{true}}\|$, $\varepsilon = \|\mathbf{e}\|$, and $\mathbf{w} = \pm \mathbf{e}_{64i}$ for $i = 1, 2, \dots, 16$. In each case, the error is scaled to satisfy $\|\mathbf{e}\|/\|\mathbf{b}_{\text{true}}\| = 1 \times 10^{-3}$. We used the auxiliary parameters $\tau = 10^{-3}$ from (2.7), $\gamma = 10^{-4}$, and $\nu = 1 \times 10^{-7}$. Computed bounds are displayed in Figures 6.1-6.6 and documented in Table 6.2. See Table 6.3 for details regarding cost and accuracy when the quadrature-based method described in [15] is applied to the same problems, with $\tau = 10^{-3}$ and the auxiliary parameters (in the notation of that paper) $\nu = 1 \times 10^{-2}$ and $\gamma = 1 \times 10^{-4}$. Note that the parameter ν is used to denote different quantities in the algorithms of [15] and in this paper. Graphs on the right show the same bounds as their counterparts on the left, only with a smaller viewing window. In Tables 6.2 and 6.3, L -eval refers to the average number of evaluations of L or \tilde{L} , respectively, required to solve (2.3) or (2.6), respectively, for each vector \mathbf{w} ; Δ^- and Δ^+ denote, respectively, the maximum error (over all 16 coordinates) of the computed lower bound and the maximum error of the computed upper bound; ℓ refers to the number of singular values required to compute all bounds in Table 6.2, or to the average number of Lanczos bidiagonalization steps per minimization (1.5) in Table

6.3; MVP refers to the number of matrix-vector product evaluations with either A or A^T required to compute all bounds.

The columns labeled $\text{cond}(A)$ and $\text{rank}(A)$ report the condition number and rank, respectively, of each matrix A as computed by the MATLAB functions `cond` and `rank`. In Figures 6.1-6.6, “a priori u.b.” and “a priori l.b.” denote the crude upper and lower bounds for each coordinate, $+\delta$ and $-\delta$, respectively, that would be obtained by removing the first constraint in (1.5). It is clear that the use of both constraints when solving (1.5) provides much tighter bounds. We remark that, as evidenced by the fairly large number of singular triplets (and matrix-vector products) required when applying our method to `deriv2`, the method may be expensive for some problems. Nevertheless, even for such problems, our method may be cheaper than other available methods when computing a large number of confidence intervals.

| Problem | $\text{cond}(A)$ | $\text{rank}(A)$ | $\ A^\dagger b\ $ | ℓ | MVP | $L\text{-eval}$ | $\ \mathbf{x}_{\text{true}}\ $ | Δ^- | Δ^+ |
|----------|------------------|------------------|-------------------|--------|-----|-----------------|--------------------------------|------------|------------|
| baart | $1.34e+019$ | 13 | $1.10e+09$ | 6 | 49 | 30 | $1.25e+0$ | $1.95e-1$ | $2.02e-1$ |
| shaw | $2.11e+020$ | 20 | $3.66e+10$ | 11 | 59 | 29 | $3.19e+1$ | $3.13e+0$ | $2.92e+0$ |
| deriv2 | $1.27e+006$ | 1024 | $2.85e+03$ | 258 | 721 | 29 | $5.77e-1$ | $1.35e-1$ | $1.07e-1$ |
| phillips | $2.90e+010$ | 1024 | $2.01e+07$ | 60 | 187 | 28 | $3.00e+0$ | $2.25e-1$ | $2.19e-1$ |
| foxgood | $3.13e+020$ | 30 | $2.59e+10$ | 7 | 53 | 29 | $1.85e+1$ | $1.15e+0$ | $1.14e+0$ |
| heat | $1.96e+237$ | 587 | $3.66e+10$ | 89 | 239 | 29 | $7.88e+0$ | $1.12e+0$ | $1.04e+0$ |

TABLE 6.2

Numerical results when computing confidence intervals for 6 discrete ill-posed problems, obtained by computing a lower bound for the minimum attained in (1.5) by solving (2.6) with the method of Section 4, using $\delta = \|\mathbf{x}_{\text{true}}\|$, $\varepsilon = \|\mathbf{e}\|$, and $\mathbf{w} = \pm \mathbf{e}_{64i}$, $i = 1, \dots, 16$.

| Problem | $\text{cond}(A)$ | $\text{rank}(A)$ | $\ A^\dagger b\ $ | ℓ | MVP | $L\text{-eval}$ | $\ \mathbf{x}_{\text{true}}\ $ | Δ^- | Δ^+ |
|----------|------------------|------------------|-------------------|--------|------|-----------------|--------------------------------|------------|------------|
| baart | $1.34e+019$ | 13 | $1.10e+09$ | 5 | 992 | 20 | $1.25e+0$ | $1.92e-1$ | $2.02e-1$ |
| shaw | $2.11e+020$ | 20 | $3.66e+10$ | 9 | 1760 | 20 | $3.19e+1$ | $2.43e+0$ | $2.42e+0$ |
| deriv2 | $1.27e+006$ | 1024 | $2.85e+03$ | 20 | 3872 | 20 | $5.77e-1$ | $1.35e-1$ | $1.05e-1$ |
| phillips | $2.90e+010$ | 1024 | $2.01e+07$ | 13 | 2528 | 19 | $3.00e+0$ | $6.13e-2$ | $6.13e-2$ |
| foxgood | $3.13e+020$ | 30 | $2.59e+10$ | 4 | 800 | 20 | $1.85e+1$ | $5.57e-1$ | $5.58e-1$ |
| heat | $1.96e+237$ | 587 | $3.66e+10$ | 27 | 5216 | 19 | $7.88e+0$ | $5.01e-1$ | $5.01e-1$ |

TABLE 6.3

Numerical results when computing confidence intervals for 6 discrete ill-posed problems, obtained by solving (1.5) via the method of [15], using $\delta = \|\mathbf{x}_{\text{true}}\|$, $\varepsilon = \|\mathbf{e}\|$, and $\mathbf{w} = \pm \mathbf{e}_{64i}$, $i = 1, \dots, 16$.

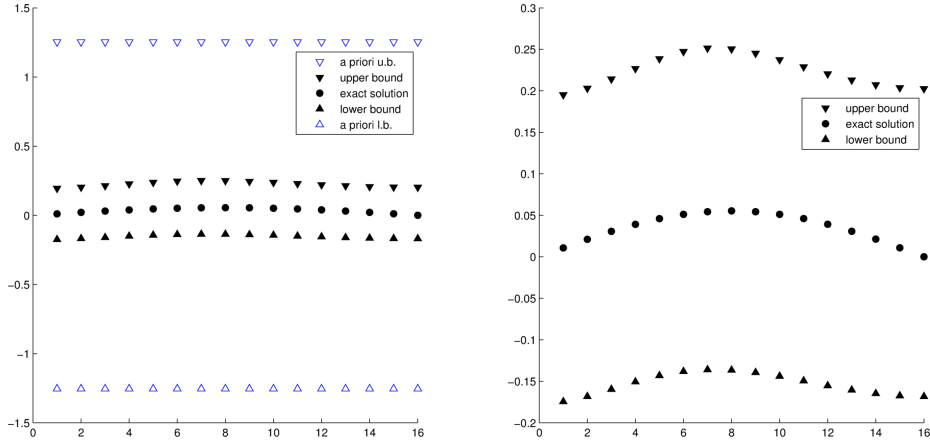


FIG. 6.1. Coordinates $64k, k = 1, \dots, 16$ of \mathbf{x}_{true} , along with the upper and lower bounds computed by solving (2.6) with the method of Section 4, with $A \in \mathbb{R}^{1024 \times 1024}$ generated by `baart`. Here, $\delta = \|\mathbf{x}_{\text{true}}\| = 1.25 \times 10^1$ and $\varepsilon = \|\mathbf{e}\| = 2.90 \times 10^{-3}$.

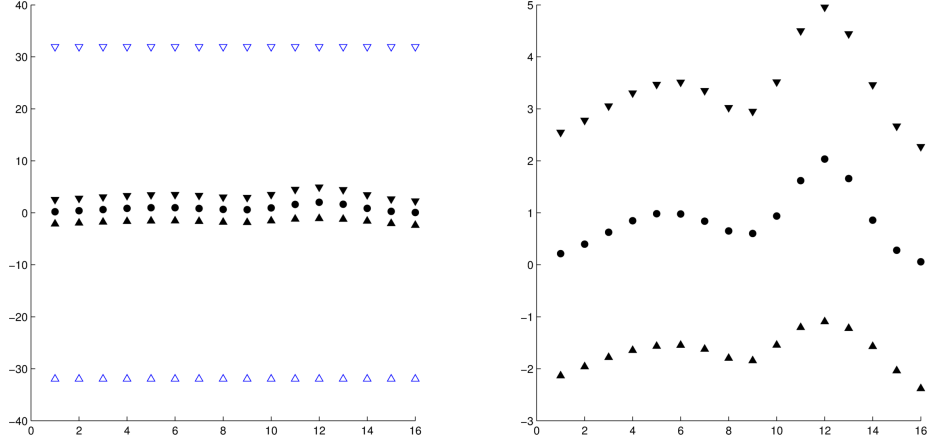


FIG. 6.2. Coordinates $64k, k = 1, \dots, 16$ of \mathbf{x}_{true} , along with the upper and lower bounds computed by solving (2.6) with the method of Section 4, with $A \in \mathbb{R}^{1024 \times 1024}$ generated by shaw. Here, $\delta = \|\mathbf{x}_{\text{true}}\| = 3.19 \times 10^1$ and $\varepsilon = \|\mathbf{e}\| = 7.46 \times 10^{-2}$.

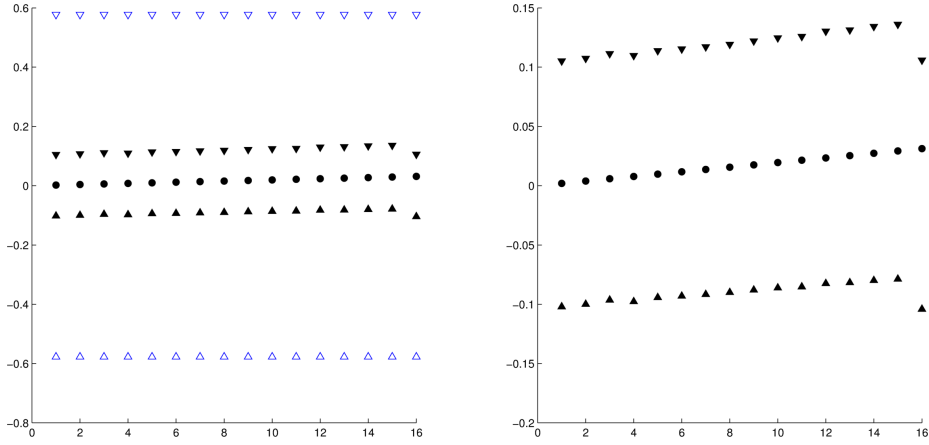


FIG. 6.3. Coordinates $64k, k = 1, \dots, 16$ of \mathbf{x}_{true} , along with the upper and lower bounds computed by solving (2.6) with the method of Section 4, with $A \in \mathbb{R}^{1024 \times 1024}$ generated by deriv2. Here, $\delta = \|\mathbf{x}_{\text{true}}\| = 5.77 \times 10^{-1}$ and $\varepsilon = \|\mathbf{e}\| = 4.60 \times 10^{-5}$.

6.2. Discrepancy principle. We now illustrate the performance of the TT method described in Section 5 and compare it to the TSVD discrepancy method outlined at the end of that section. We first apply the methods to the 6 linear discrete ill-posed problems mentioned above with matrices $A \in \mathbb{R}^{1024 \times 1024}$, using $\varepsilon = \|\mathbf{e}\|$ and the auxiliary parameters $\gamma = 1 \times 10^{-2}$ and $\nu = 1 \times 10^{-5}$. The relative error $\sigma := \|\mathbf{e}\|/\|\mathbf{b}_{\text{true}}\|$ assumes the values 1×10^{-k} for $k = 1, 2, 3$. Tables 6.4 and 6.5 report the average error in the approximate solutions computed with the TT and TSVD discrepancy methods, which we denote by

$$\mathbf{x}_{\text{TT}} := \mathbf{x}_{\mu_\varepsilon^{(\ell_{\text{TT}})}}^{(\ell_{\text{TT}})} \quad \text{and} \quad \mathbf{x}_{\text{TSVD}} := \sum_{i=1}^{\ell_{\text{TSVD}}} \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{v}_i,$$

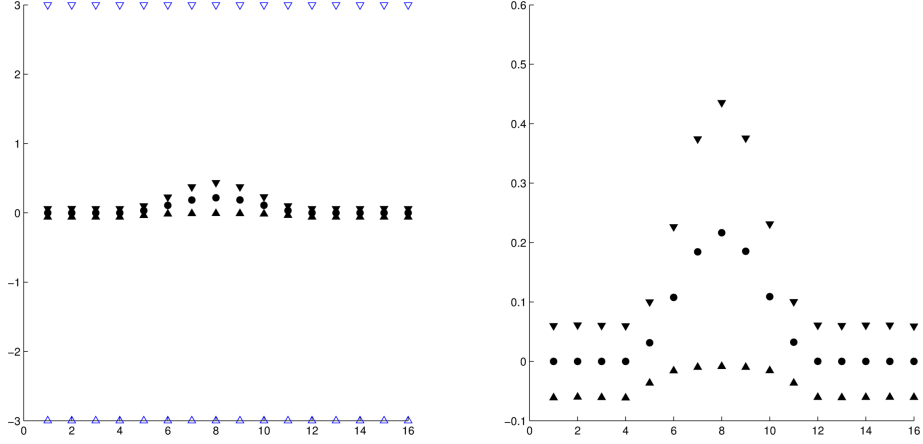


FIG. 6.4. Coordinates $64k, k = 1, \dots, 16$ of \mathbf{x}_{true} , along with the upper and lower bounds computed by solving (2.6) with the method of Section 4, with $A \in \mathbb{R}^{1024 \times 1024}$ generated by phillips. Here, $\delta = \|\mathbf{x}_{\text{true}}\| = 3.00 \times 10^0$ and $\varepsilon = \|\mathbf{e}\| = 1.53 \times 10^{-2}$.

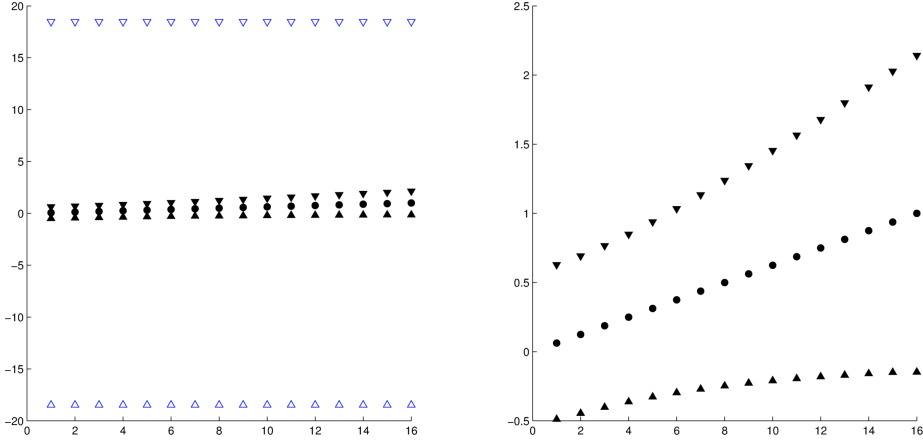


FIG. 6.5. Coordinates $64k, k = 1, \dots, 16$ of \mathbf{x}_{true} , along with the upper and lower bounds computed by solving (2.6) with the method of Section 4, with $A \in \mathbb{R}^{1024 \times 1024}$ generated by foxgood. Here, $\delta = \|\mathbf{x}_{\text{true}}\| = 1.85 \times 10^1$ and $\varepsilon = \|\mathbf{e}\| = 1.43 \times 10^{-2}$.

respectively, for each problem and error level. The tables also display the average number of singular values required for the computations. All averages are taken over 10^4 properly scaled error vectors $\|\mathbf{e}\|$ with independent, normally distributed entries with zero mean. Table 6.6 displays the entry-wise ratio of Tables 6.4 and 6.5. For instance, the (5,2) entry of Table 6.6 shows that for the problem foxgood with $\|\mathbf{e}\|/\|\mathbf{b}_{\text{true}}\| = 1 \times 10^{-2}$, the TT method produces (on average) only about 69% as much error in the computed approximations of \mathbf{x}_{true} as TSVD discrepancy. Similarly, the (5,5) entry of the same table displays that the TT method requires (on average) about 2.37 times as many singular values as TSVD discrepancy, although glancing at Tables 6.4 and 6.5 reveals that the average number of singular values required by the two methods differs only by about 3. Overall, the TT method determines more accurate approximations of \mathbf{x}_{true} than TSVD discrepancy for 14 of the 18 problems considered. It should be noted that, due to the fairly large number of singular triplets

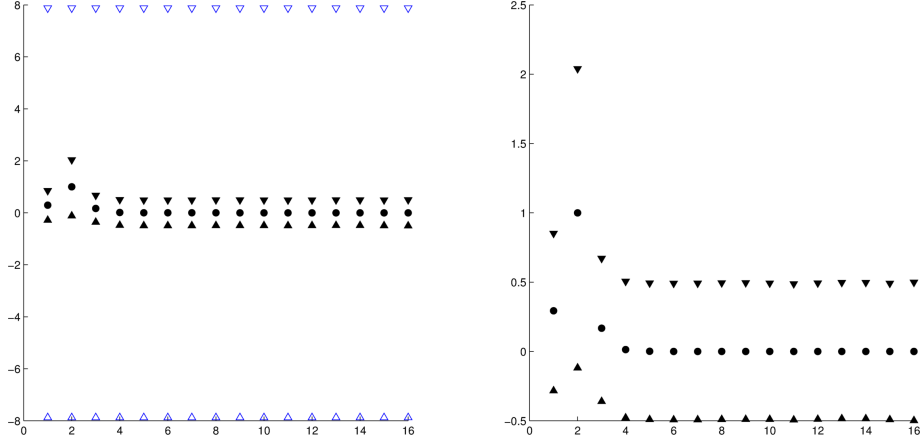


FIG. 6.6. Coordinates $64k, k = 1, \dots, 16$ of \mathbf{x}_{true} , along with the upper and lower bounds computed by solving (2.6) with the method of Section 4, with $A \in \mathbb{R}^{1024 \times 1024}$ generated by heat. Here, $\delta = \|\mathbf{x}_{\text{true}}\| = 7.88 \times 10^0$ and $\varepsilon = \|\mathbf{e}\| = 1.50 \times 10^{-3}$.

required when applying the TT method to heat and deriv2, the method might not be attractive to use for these problems.

In addition to the average results provided in Tables 6.4–6.6, we provide results for a single, representative problem instance for each ill-posed problem with relative error $\sigma = 1 \times 10^{-2}$ but the same auxiliary parameters as above. Table 6.7 compares the true solution \mathbf{x}_{true} , the TT solution \mathbf{x}_{TT} , and the discrepancy principle solution $\mathbf{x}_{\mu_\varepsilon}$ obtained working with the SVD of A (computed using `[U,S,V]=svd(A)` in MATLAB) and solving (5.1) within machine (relative) precision. In all cases, the relative error $\|\mathbf{x}_{\text{TT}} - \mathbf{x}_{\mu_\varepsilon}\| / \|\mathbf{x}_{\mu_\varepsilon}\|$ is smaller than $\sigma = 1 \times 10^{-2}$, as required. Moreover, notice that in all cases, the error $\|\mathbf{x}_{\text{TT}} - \mathbf{x}_{\mu_\varepsilon}\|$ caused by approximation of A by a partial SVD, within Tikhonov regularization, is smaller than the error $\|\mathbf{x}_{\text{true}} - \mathbf{x}_{\mu_\varepsilon}\|$ obtained when using the discrepancy principle with Tikhonov regularization. The Figures 6.7–6.9 display for some of the problem considered the TT approximation, the TSVD discrepancy approximation, and the exact solution.

| Problem | $\ \mathbf{x}_{\text{TT}} - \mathbf{x}_{\text{true}}\ $ | | | ℓ_{TT} | | |
|----------|---|--------------------|--------------------|--------------------|--------------------|--------------------|
| | $\sigma = 10^{-1}$ | $\sigma = 10^{-2}$ | $\sigma = 10^{-3}$ | $\sigma = 10^{-1}$ | $\sigma = 10^{-2}$ | $\sigma = 10^{-3}$ |
| baart | 0.28 | 0.20 | 0.15 | 4.98 | 5.17 | 5.89 |
| shaw | 5.05 | 2.52 | 1.50 | 9.57 | 9.99 | 10.0 |
| deriv2 | 0.18 | 0.13 | 0.09 | 32.0 | 46.8 | 69.3 |
| phillips | 0.14 | 0.06 | 0.03 | 17.6 | 17.5 | 16.7 |
| foxgood | 0.89 | 0.41 | 0.16 | 5.47 | 5.50 | 5.73 |
| heat | 1.62 | 0.58 | 0.20 | 46.8 | 50.1 | 51.6 |

TABLE 6.4

On the left, the average error (over 10^4 experiments) in the TT discrepancy principle solution \mathbf{x}_{TT} for relative error levels of $\sigma = 1 \times 10^{-k}$, $k = 1, 2, 3$. On the right, the average number of singular values ℓ_{TT} required by the TT method for the corresponding relative error levels.

We finish our comparison of the two methods by presenting a problem for which the TT method produces a much more accurate approximation of \mathbf{x}_{true} than TSVD discrepancy. Consider the problem **baart**. The matrix $A \in \mathbb{R}^{1024 \times 1024}$ has computed rank 13 and $\|\mathbf{x}_{\text{true}}\| = 1.25$. Let \mathbf{e} be the projection of a random vector with normally distributed entries onto the orthogonal complement of the space spanned by the first

| Problem | $\ \mathbf{x}_{\text{TSVD}} - \mathbf{x}_{\text{true}}\ $ | | | ℓ_{TSVD} | | |
|----------|---|--------------------|--------------------|----------------------|--------------------|--------------------|
| | $\sigma = 10^{-1}$ | $\sigma = 10^{-2}$ | $\sigma = 10^{-3}$ | $\sigma = 10^{-1}$ | $\sigma = 10^{-2}$ | $\sigma = 10^{-3}$ |
| baart | 0.27 | 0.22 | 0.16 | 2.87 | 3.20 | 4.03 |
| shaw | 5.82 | 2.30 | 1.54 | 4.19 | 6.74 | 7.03 |
| deriv2 | 0.20 | 0.14 | 0.09 | 5.41 | 11.5 | 24.5 |
| phillips | 0.10 | 0.07 | 0.03 | 6.98 | 7.39 | 11.1 |
| foxgood | 0.99 | 0.60 | 0.17 | 2.03 | 2.32 | 3.05 |
| heat | 1.96 | 0.72 | 0.23 | 12.0 | 19.2 | 27.6 |

TABLE 6.5

On the left, the average error (over 10^4 experiments) in the TSVD discrepancy solution \mathbf{x}_{TSVD} , for relative error levels $\sigma = 1 \times 10^{-k}$, $k = 1, 2, 3$. On the right, the average number of singular values ℓ_{TSVD} required by TSVD discrepancy for the corresponding relative error levels.

| Problem | $\ \mathbf{x}_{\text{TT}} - \mathbf{x}_{\text{true}}\ / \ \mathbf{x}_{\text{TSVD}} - \mathbf{x}_{\text{true}}\ $ | | | $\ell_{\text{TT}} / \ell_{\text{TSVD}}$ | | |
|----------|---|--------------------|--------------------|---|--------------------|--------------------|
| | $\sigma = 10^{-1}$ | $\sigma = 10^{-2}$ | $\sigma = 10^{-3}$ | $\sigma = 10^{-1}$ | $\sigma = 10^{-2}$ | $\sigma = 10^{-3}$ |
| baart | 1.02 | 0.89 | 0.95 | 1.74 | 1.61 | 1.46 |
| shaw | 0.87 | 1.10 | 0.98 | 2.28 | 1.48 | 1.42 |
| deriv2 | 0.93 | 0.94 | 0.94 | 5.91 | 4.08 | 2.83 |
| phillips | 1.34 | 0.82 | 1.01 | 2.51 | 2.36 | 1.51 |
| foxgood | 0.90 | 0.69 | 0.97 | 2.69 | 2.37 | 1.88 |
| heat | 0.83 | 0.80 | 0.87 | 3.89 | 2.61 | 1.87 |

TABLE 6.6

The entry-wise ratio of Tables 6.4 and 6.5.

| Problem | $\ A^\dagger \mathbf{b}\ $ | ℓ | $\ \mathbf{x}_{\mu_\varepsilon}\ $ | $\ \mathbf{x}_{\text{true}} - \mathbf{x}_{\mu_\varepsilon}\ / \ \mathbf{x}_{\mu_\varepsilon}\ $ | $\ \mathbf{x}_{\text{TT}} - \mathbf{x}_{\mu_\varepsilon}\ / \ \mathbf{x}_{\mu_\varepsilon}\ $ |
|----------|----------------------------|--------|------------------------------------|--|--|
| baart | $1.10e+09$ | 5 | $1.23e+0$ | $1.69e-1$ | $2.67e-4$ |
| shaw | $3.66e+10$ | 10 | $3.16e+1$ | $1.10e-1$ | $1.07e-5$ |
| deriv2 | $2.85e+03$ | 43 | $5.54e-1$ | $2.33e-1$ | $1.04e-2$ |
| phillips | $2.01e+07$ | 16 | $3.00e+0$ | $1.40e-2$ | $3.90e-3$ |
| foxgood | $2.59e+10$ | 5 | $1.84e+1$ | $3.23e-2$ | $2.05e-4$ |
| heat | $3.66e+10$ | 49 | $7.77e+0$ | $8.01e-2$ | $2.89e-3$ |

TABLE 6.7

Selected results for a single problem instance for 6 discrete ill-posed problems, with $\varepsilon = \|\mathbf{e}\|$, $\sigma = 1 \times 10^{-2}$, $\gamma = 1 \times 10^{-2}$, $\nu = 1 \times 10^{-5}$.

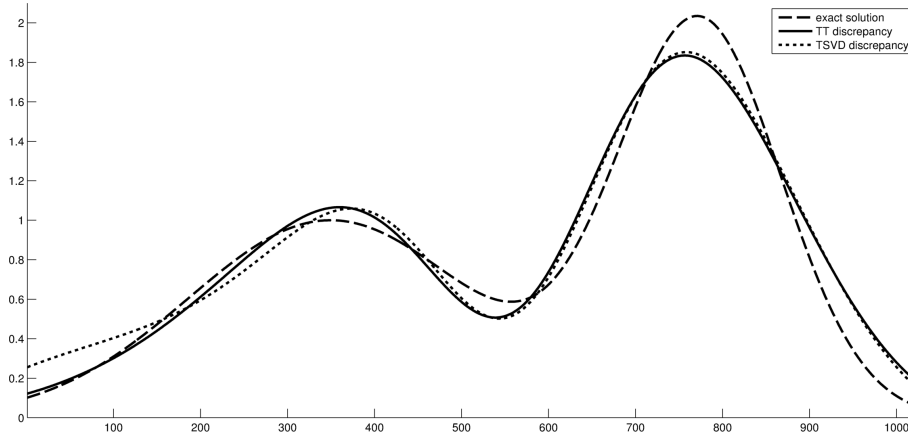


FIG. 6.7. For an instance of the problem shaw, the true solution \mathbf{x}_{true} , the TT solution \mathbf{x}_{TT} , and the TSVD solution \mathbf{x}_{TSVD} , using $\varepsilon = \|\mathbf{e}\|$, $\sigma = 1 \times 10^{-2}$, $\gamma = 1 \times 10^{-2}$, $\nu = 1 \times 10^{-5}$. Here, $\|\mathbf{x}_{\text{TT}} - \mathbf{x}_{\text{true}}\| / \|\mathbf{x}_{\text{TSVD}} - \mathbf{x}_{\text{true}}\| = 0.965$, and accepted iterations were $\ell_{\text{TT}} = 10$ and $\ell_{\text{TSVD}} = 6$.

5 left singular vectors \mathbf{u}_i , scaled to satisfy $\|\mathbf{e}\| / \|\mathbf{b}_{\text{true}}\| = 10^{-1}$. Using $\gamma = 10^{-1}$ and $\nu = 10^{-10}$, and carrying out 10^4 experiments with different error vectors \mathbf{e} , the TT

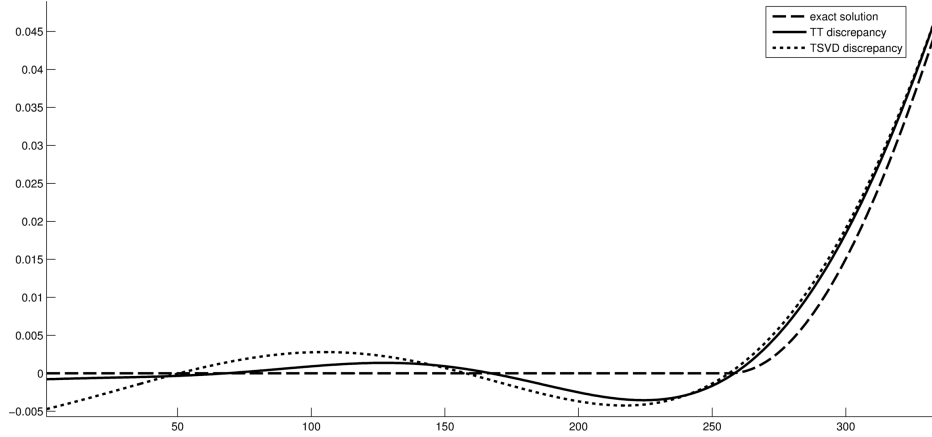


FIG. 6.8. For an instance of the problem *phillips*, the true solution \mathbf{x}_{true} , the TT solution \mathbf{x}_{TT} , and the TSVD solution \mathbf{x}_{TSVD} , using $\varepsilon = \|\mathbf{e}\|$, $\sigma = 1 \times 10^{-2}$, $\gamma = 1 \times 10^{-2}$, $\nu = 1 \times 10^{-5}$. Here, $\|\mathbf{x}_{\text{TT}} - \mathbf{x}_{\text{true}}\| / \|\mathbf{x}_{\text{TSVD}} - \mathbf{x}_{\text{true}}\| = 0.637$, and accepted iterations were $\ell_{\text{TT}} = 16$ and $\ell_{\text{TSVD}} = 7$. Only the left part of the true and computed solutions are displayed.

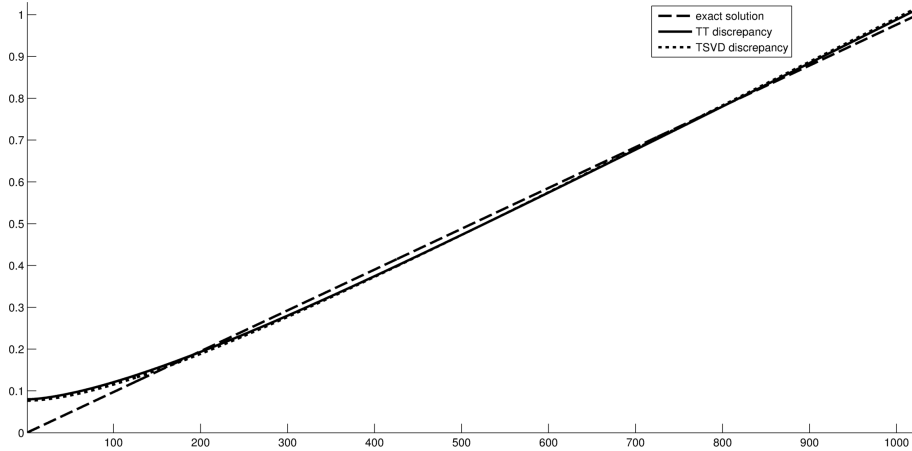


FIG. 6.9. For an instance of the problem *foxgood*, the true solution \mathbf{x}_{true} , the TT solution \mathbf{x}_{TT} , and the TSVD solution \mathbf{x}_{TSVD} , using $\varepsilon = \|\mathbf{e}\|$, $\sigma = 1 \times 10^{-2}$, $\gamma = 1 \times 10^{-2}$, $\nu = 1 \times 10^{-5}$. Here, $\|\mathbf{x}_{\text{TT}} - \mathbf{x}_{\text{true}}\| / \|\mathbf{x}_{\text{TSVD}} - \mathbf{x}_{\text{true}}\| = 1.030$, and accepted iterations were $\ell_{\text{TT}} = 5$ and $\ell_{\text{TSVD}} = 2$.

method gave computed approximations of \mathbf{x}_{true} with an average error 0.1406, while TSVD discrepancy determined approximate solutions with an average error 404.9. The TT method required on average 7.1 singular triplets, while TSVD discrepancy needed on average 5.1 singular triplets. Thus, TT produced on average much more accurate approximations of \mathbf{x}_{true} . We remark that in this problem TSVD discrepancy performs very poorly, rather than TT performing exceptionally well. Error vectors that yield dramatic differences in the performance of the TT and TSVD discrepancy methods may occur in application, but only with low probability.

7. Conclusion. The first part of Section 6 shows that, when sufficient knowledge of the distribution of \mathbf{e} and the norm of \mathbf{x}_{true} is available, the methods described in Sections 2 and 4 can be used to compute lower and upper bounds for each coordinate

of \mathbf{x}_{true} . Unlike previously available methods for solving large-scale problems (1.5), our method of computing a lower bound for the minimum attained in (1.5) affords the use of a single reduction, a partial SVD, for all confidence intervals. This offers a significant reduction in computational cost when confidence intervals for many solution components are to be determined. Our numerical examples demonstrate that for many problems, the solution of (2.6) does not deviate significantly from that of (2.3); in such cases, the present method yields bounds that are similar in quality to those produced by available methods.

For some problems and error levels, TT offers significant improvement in accuracy over TSVD discrepancy, with only a moderate increase in expense. Thus, for some problems, TT offers a valuable second opinion, and a slightly more robust alternative than the leading method.

REFERENCES

- [1] J. Baglama and L. Reichel, *Augmented implicitly restarted Lanczos bidiagonalization methods*, SIAM J. Sci. Comput., 27 (2005), pp. 19–42.
- [2] J. Baglama and L. Reichel, *Restarted block Lanczos bidiagonalization methods*, Numer. Algorithms, 43 (2006), pp. 251–272.
- [3] Å. Björck, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, 1996.
- [4] D. Calvetti and L. Reichel, *Tikhonov regularization of large linear problems*, BIT Numer. Math., 43 (2003), pp. 263–283.
- [5] L. Eldén, *Algorithm for the computation of functionals defined on the solution of discrete ill-posed problems*, BIT, 30 (1990), pp. 466–483.
- [6] L. Eldén, P. C. Hansen, and M. Rojas, *Minimization of linear functionals defined on solutions of large-scale discrete ill-posed problems*, BIT Numer. Math., 45 (2005), pp. 329–340.
- [7] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.
- [8] G. H. Golub and U. von Matt, *Quadratically constrained least squares and quadratic problems*, Numer. Math., 59 (1991), pp. 561–580.
- [9] 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.
- [10] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems*, SIAM, Philadelphia, 1998.
- [11] P. C. Hansen, *Regularization tools version 4.0 for Matlab 7.3*, Numer. Algorithms, 46 (2007), pp. 189–194.
- [12] M. E. Hochstenbach, *A Jacobi-Davidson type SVD method*, SIAM J. Sci. Comput., 23 (2001), pp. 606–628.
- [13] 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.
- [14] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, Springer, New York, 1996.
- [15] D. R. Martin and L. Reichel, *Minimization of functionals on the solution of a large-scale discrete ill-posed problem*, BIT Numer. Math., in press.
- [16] J. J. Moré and D. C. Sorensen, *Computing a trust region step*, SIAM J. Sci. Stat. Comput., 4 (1983), pp. 553–572.
- [17] V. A. Morozov, *Methods for Solving Incorrectly Posed Problems*, Springer, New York, 1984.
- [18] D. P. O’Leary and B. W. Rust, *Confidence intervals for inequality-constrained least squares problems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 473–489.
- [19] M. Rojas, S. A. Santos, and D. C. Sorensen, *A new matrix-free algorithm for the large-scale trust-region subproblem*, SIAM J. Optim., 11 (2000), pp. 611–646.