

## Fractional regularization matrices for linear discrete ill-posed problems

Michiel E. Hochstenbach ·  
Silvia Noschese · Lothar Reichel

Received: date / Accepted: date

**Abstract** The numerical solution of linear discrete ill-posed problems typically requires regularization. Two of the most popular regularization methods are due to Tikhonov and Lavrentiev. These methods require the choice of a regularization matrix. Common choices include the identity matrix and finite difference approximations of a derivative operator. It is the purpose of the present paper to explore the use of fractional powers of the matrices  $A^T A$  (for Tikhonov regularization) and  $A$  (for Lavrentiev regularization) as regularization matrices, where  $A$  is the matrix that defines the linear discrete ill-posed problem. Both small and large-scale problems are considered.

**Keywords** Ill-posed problem · fractional Tikhonov regularization · fractional Lavrentiev regularization · fractional power regularization matrix

**Mathematics Subject Classification (2000)** 65F10 · 65F22 · 65R30

---

Michiel E. Hochstenbach  
Department of Mathematics and Computer Science  
Eindhoven University of Technology, PO Box 513, 5600 MB, The Netherlands  
[www.win.tue.nl/~hochsten](http://www.win.tue.nl/~hochsten). This author was supported by an NWO Vidi grant.

Silvia Noschese  
Dipartimento di Matematica “Guido Castelnuovo”  
SAPIENZA Università di Roma, Piazzale Aldo Moro 2, I-00185 Roma, Italy  
E-mail: [noschese@mat.uniroma1.it](mailto:noschese@mat.uniroma1.it)

Lothar Reichel  
Department of Mathematical Sciences  
Kent State University, Kent, OH 44242, USA  
E-mail: [reichel@math.kent.edu](mailto:reichel@math.kent.edu). This research was supported in part by NSF grant DMS-1115385.

## 1 Introduction

This paper discusses the solution of linear least-squares problems

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

with a nonsymmetric and possibly rectangular matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$  or with a symmetric positive semidefinite matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$ . Throughout this paper  $\|\cdot\|$  denotes the Euclidean vector norm or the associated induced matrix norm. The singular values of  $\mathbf{A}$  are assumed to “cluster” at the origin. In particular,  $\mathbf{A}$  is severely ill-conditioned and may be singular. This kind of least-squares problems often are referred to as 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 with a smooth kernel. The vector  $\mathbf{b}$  represents available data, which is contaminated by an error vector  $\mathbf{e}$ . This error may stem from measurement inaccuracies or discretization. Thus,

$$(1.2) \quad \mathbf{b} = \widehat{\mathbf{b}} + \mathbf{e},$$

where  $\widehat{\mathbf{b}}$  denotes the unknown error-free vector associated with  $\mathbf{b}$ . We will assume the unavailable error-free system

$$(1.3) \quad \mathbf{A}\mathbf{x} = \widehat{\mathbf{b}}$$

to be consistent and denote its solution of minimal Euclidean norm by  $\widehat{\mathbf{x}}$ . We would like to determine an approximation of  $\widehat{\mathbf{x}}$  by computing a suitable approximate solution of (1.1). Due to the ill-conditioning of the matrix  $\mathbf{A}$  and the error  $\mathbf{e}$  in  $\mathbf{b}$ , the solution of the least-squares problem (1.1) of minimal Euclidean norm is typically a poor approximation of  $\widehat{\mathbf{x}}$ .

Tikhonov regularization is a popular approach to determine an approximation of  $\widehat{\mathbf{x}}$ ; see, e.g., [3, 6, 8] for properties and applications. This method replaces the minimization problem (1.1) by a penalized least-squares problem of the form

$$(1.4) \quad \min_{\mathbf{x} \in \mathcal{N}(\mathbf{A})^\perp} \left\{ \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 + \mu \|\mathbf{M}\mathbf{x}\|^2 \right\},$$

where  $\mathcal{N}(\mathbf{A})^\perp$  denotes the orthogonal complement of the null space  $\mathcal{N}(\mathbf{A})$  of the matrix  $\mathbf{A}$ ,  $\mathbf{M} \in \mathbb{R}^{p \times n}$  is a regularization matrix, and the scalar  $\mu \geq 0$  a regularization parameter. The problem (1.4) has a unique solution  $\mathbf{x}_\mu$  for all nonnegative values of  $\mu$ . The choices of regularization matrix  $\mathbf{M}$  and value of the regularization parameter  $\mu$  determine how much  $\mathbf{x}_\mu$  differs from the desired solution  $\widehat{\mathbf{x}}$  of (1.3) and how sensitive  $\mathbf{x}_\mu$  is to the error  $\mathbf{e}$  in  $\mathbf{b}$ . The choice of a suitable regularization matrix can be challenging. We propose the use of a regularization matrix of the form

$$(1.5) \quad \mathbf{M} = (\mathbf{A}^T \mathbf{A})^{\beta/2},$$

where  $\beta \geq 0$  and  $\mathbf{A}^T$  denotes the transpose of  $\mathbf{A}$ .

The effect of the penalization term in (1.4) on the solution  $\mathbf{x}_\mu$  can be investigated with the aid of filter factors; see Section 2. We show that all filter factors increase with  $\beta$ . Thus, for fixed  $\mu > 0$ , the penalization term in (1.4) provides less penalization the larger  $\beta$  is. This is interesting because standard Tikhonov regularization, which corresponds to  $\beta = 0$ , generally oversmooths the computed solution; see Klann and Ramlau [12] for a discussion on the latter.

Tikhonov regularization allows the matrix  $A$  to be nonsymmetric or rectangular. For notational simplicity, we assume in our discussion on Tikhonov regularization that  $m \geq n$ ; only minor modifications are required when  $m < n$ .

Lavrentiev regularization can be applied when the matrix  $A$  is symmetric positive semidefinite. Many applications of this regularization method are reported in the literature, including to neural networks [4], magnetic resonance imaging [14], and image deblurring [15]. In its standard form, this regularization method replaces the least-squares problem (1.1) by the linear system of equations

$$(1.6) \quad (A + \mu I) \mathbf{x} = \mathbf{b}.$$

We propose to replace the regularization matrix  $I$  by the matrix

$$(1.7) \quad M = A^\beta$$

for some  $\beta > 0$ .

This paper is organized as follows. Section 2 discusses Tikhonov regularization with the regularization matrix (1.5). The relation to the fractional Tikhonov method presented in [10] is explored. Lavrentiev regularization with a fractional regularization matrix (1.7) is derived in Section 3, and we discuss the choices of  $\mu$  and  $\beta$  in Section 4. When the matrix  $A$  is nonsymmetric and of small to medium size, the regularization matrix (1.5) can be conveniently computed by first evaluating the singular valued decomposition of  $A$ . Similarly, when  $A$  is symmetric positive semidefinite, we can determine the matrix (1.7) by first evaluating the spectral factorization of  $A$ . However, these approaches to determining the matrices (1.5) and (1.7) are not attractive when  $A$  is large. We outline in Section 5 how to proceed in this situation. Computed examples that illustrate the performance of Lavrentiev regularization with a regularization matrix of the form (1.7) are presented in Section 6 and concluding remarks can be found in Section 7.

## 2 Fractional Tikhonov regularization

This section discusses Tikhonov regularization with the regularization matrix (1.5). A different derivation can be found in [10]. The discussion of the present paper sheds additional light on the method.

The normal equations associated with the Tikhonov minimization problem (1.4) with the regularization matrix (1.5) can be expressed as

$$(2.1) \quad ((A^T A)^{1-\beta} + \mu I) \mathbf{x} = (A^T A)^{-\beta} A^T \mathbf{b}, \quad \mathbf{x} \in \mathcal{N}(A)^\perp,$$

where  $(A^T A)^{-\beta}$  is defined with the aid of the Moore–Penrose generalized inverse when  $\beta > 0$ . Introduce the singular value decomposition (SVD)

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

where  $U = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m] \in \mathbb{R}^{m \times m}$  and  $V = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] \in \mathbb{R}^{n \times n}$  are orthogonal matrices, and  $\Sigma$  is the diagonal, possibly rectangular, matrix

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

whose nontrivial entries are the singular values ordered according to

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_n = 0.$$

The index  $r$  is the rank of  $A$ . We will throughout this paper assume the matrix  $A$  to be scaled so that

$$(2.2) \quad \|A\| < 1.$$

Using the SVD of  $A$ , the solution of (1.4) with  $M$  given by (1.5) can be expressed as

$$(2.3) \quad \mathbf{x}_{\mu, \beta} = \sum_{j=1}^r \varphi(\sigma_j, \beta) (\mathbf{u}_j^T \mathbf{b}) \mathbf{v}_j, \quad \varphi(\sigma, \beta) := \frac{\sigma^{1-2\beta}}{\sigma^{2-2\beta} + \mu}.$$

The coefficients  $\varphi(\sigma_j, \beta)$  are commonly referred to as filter factors. Generally, right singular vectors  $\mathbf{v}_j$  with a small index  $j$  represent discretizations of constant or slowly oscillating functions, while right singular vectors  $\mathbf{v}_j$  with a large index represent discretizations of rapidly oscillating functions. The latter vectors often are associated with “tiny” positive singular values. When the desired solution  $\hat{\mathbf{x}}$  is smooth, the right singular vectors  $\mathbf{v}_j$  with  $j$  large should be damped in the representation (2.3) of  $\mathbf{x}_{\mu, \beta}$ . This can be achieved by requiring the filter factors  $\varphi(\sigma_j, \beta)$  to be small for large  $j$  and all  $\mu > 0$ , and leads to the demand that the function  $\varphi$  should satisfy

$$(2.4) \quad \lim_{\sigma \searrow 0} \varphi(\sigma, \beta) = 0$$

for any  $\mu > 0$ . We therefore should choose

$$(2.5) \quad \beta < \frac{1}{2}.$$

Tikhonov regularization problems of the form (2.1) also can be derived from the penalized least-squares problem

$$(2.6) \quad \min_{\mathbf{x} \in \mathcal{N}(A)^\perp} \{ \|A\mathbf{x} - \mathbf{b}\|_W^2 + \mu \|\mathbf{x}\|^2 \},$$

where  $\|\mathbf{x}\|_W = (\mathbf{x}^T W \mathbf{x})^{1/2}$  is a seminorm with  $W = (AA^T)^{(\alpha-1)/2}$ . This approach to Tikhonov regularization is described in [10]. We define  $W$  with the aid of the Moore–Penrose pseudoinverse of  $AA^T$  when  $\alpha < 1$ .

**Proposition 2.1** *The penalized least-squares problems (1.4) and (2.6) are equivalent when  $\alpha = 1 - 2\beta$ .*

*Proof* The normal equations associated with (2.6) are given by (see [10])

$$(2.7) \quad ((A^T A)^{(\alpha+1)/2} + \mu I) \mathbf{x} = (A^T A)^{(\alpha-1)/2} A^T \mathbf{b}, \quad \mathbf{x} \in \mathcal{N}(A)^\perp.$$

Comparing with (2.1) shows the proposition.  $\square$

We note that with  $\alpha = 1 - 2\beta$  the requirements  $\beta \geq 0$  and (2.5) yield the bounds

$$(2.8) \quad 0 < \alpha \leq 1.$$

The Tikhonov solution (2.3) can be expressed in terms of  $\alpha$  as follows,

$$(2.9) \quad \mathbf{x}_{\mu, (1-\alpha)/2} = \sum_{j=1}^r \tilde{\varphi}(\sigma_j, \alpha) (\mathbf{u}_j^T \mathbf{b}) \mathbf{v}_j, \quad \tilde{\varphi}(\sigma, \alpha) = \frac{\sigma^\alpha}{\sigma^{\alpha+1} + \mu}.$$

With the scaling (2.2), the filter factors in (2.3) satisfy

$$(2.10) \quad \varphi(\sigma_j, \beta_1) > \varphi(\sigma_j, \beta_2) > \varphi(\sigma_j, 0) \quad \text{for } \beta_1 > \beta_2 > 0$$

when  $\mu > 0$  and  $\sigma_j > 0$ , where  $\beta = 0$  corresponds to standard Tikhonov regularization with  $M = I$ . The inequalities (2.10) show that all components of the computed solution (2.3) are damped less the larger  $\beta \geq 0$  is. Moreover, the filter factors converge to zero slower as  $j$  increases the larger  $\beta$  is. A positive value of  $\beta$  may yield a more accurate approximation of  $\hat{\mathbf{x}}$  than  $\beta = 0$ , because standard Tikhonov regularization may oversmooth the solution; see [12] for an analysis of the latter. Numerical examples in [10] illustrate that it can be beneficial to let  $\beta > 0$ .

### 3 Fractional Lavrentiev regularization

The matrix  $A \in \mathbb{R}^{n \times n}$  is in this and the following sections assumed to be symmetric and positive semidefinite, and scaled so that (2.2) holds. Let  $P_{\mathcal{R}(A)}$  denote the orthogonal projector onto  $\mathcal{R}(A)$ , the range of  $A$ , and recall that  $\mathcal{R}(A) = \mathcal{N}(A)^\perp$ . The Lavrentiev equation (1.6) is not guaranteed to be consistent for  $\mu = 0$ . Therefore we propose to replace (1.6) by the equation

$$(3.1) \quad (A + \mu I) \mathbf{x} = P_{\mathcal{R}(A)} \mathbf{b}, \quad \mathbf{x} \in \mathcal{R}(A),$$

which is consistent for all  $\mu \geq 0$ .

The matrix  $A$  can be expressed as  $A = H^2$ , where  $H$  is a symmetric and positive semidefinite matrix. It can easily be verified that  $P_{\mathcal{R}(A)} = HH^+$ , where  $H^+$  denotes the Moore–Penrose pseudo-inverse of  $H$ . Introduce the vector  $\mathbf{d} = H^+ \mathbf{b}$ . Then (3.1) can be written as

$$(H^2 + \mu I) \mathbf{x} = H \mathbf{d}, \quad \mathbf{x} \in \mathcal{R}(H).$$

These are the normal equations associated with the Tikhonov minimization problem

$$\min_{\mathbf{x} \in \mathcal{R}(H)} \{ \|H\mathbf{x} - \mathbf{d}\|^2 + \mu \|\mathbf{x}\|^2 \}.$$

The good performance reported in [10] of fractional Tikhonov regularization suggests that it may be advantageous to replace the above minimization problem by

$$\min_{\mathbf{x} \in \mathcal{R}(H)} \{ \|H\mathbf{x} - \mathbf{d}\|_W^2 + \mu \|\mathbf{x}\|^2 \},$$

where  $W = H^{\gamma-1}$  for some  $\gamma > 0$ ; cf. (2.6). When  $\gamma < 1$ ,  $H^{\gamma-1}$  is defined with the aid of the Moore–Penrose pseudoinverse of  $H$ . The associated normal equations are given by (cf. the proof of Proposition 2.1)

$$(3.2) \quad (H^{\gamma+1} + \mu I) \mathbf{x} = H^{\gamma-1} H \mathbf{d}, \quad \mathbf{x} \in \mathcal{R}(H),$$

which we can express as

$$(A^{(\gamma+1)/2} + \mu I) \mathbf{x} = A^{(\gamma-1)/2} P_{\mathcal{R}(A)} \mathbf{b}, \quad \mathbf{x} \in \mathcal{R}(A).$$

To obtain filter factors of the same simple form as (2.9), we let  $\alpha = \frac{1}{2}(\gamma - 1)$ . This yields

$$(3.3) \quad (A^{\alpha+1} + \mu I) \mathbf{x} = A^\alpha P_{\mathcal{R}(A)} \mathbf{b}, \quad \mathbf{x} \in \mathcal{R}(A).$$

These equations can be obtained from the normal equations (2.7) when the matrix  $A$  is symmetric and  $\alpha > 0$ . This suggests that it may be interesting to consider equation (3.3) for  $\alpha$ -values in the interval (2.8). The projection  $P_{\mathcal{R}(A)}$  is relevant only when  $\alpha = 0$ , since  $A^\alpha P_{\mathcal{R}(A)} = A^\alpha$  for  $\alpha > 0$ . The value  $\alpha = 0$  corresponds to standard Lavrentiev regularization (1.6) with the extra projection  $P_{\mathcal{R}(A)}$ . For future reference, we define the vector

$$\mathbf{b}_{\mathcal{R}(A)} = P_{\mathcal{R}(A)} \mathbf{b}.$$

We refer to (3.3) as the fractional Lavrentiev method. Note that equation (3.2), and therefore (3.3), also can be derived by forming the normal equations associated with the minimization problem

$$\min_{\mathbf{x} \in \mathcal{R}(H)} \{ \|H\mathbf{x} - \mathbf{d}\|^2 + \mu \|M\mathbf{x}\|^2 \}$$

with  $M = H^{(1-\gamma)/2}$ .

We turn to the filter factors associated with the fractional Lavrentiev method (3.3). Introduce the spectral factorization

$$(3.4) \quad A = U \Lambda U^T,$$

where  $U = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n] \in \mathbb{R}^{n \times n}$  is orthogonal and

$$\Lambda = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_n] \in \mathbb{R}^{n \times n}.$$

We order the eigenvalues according to

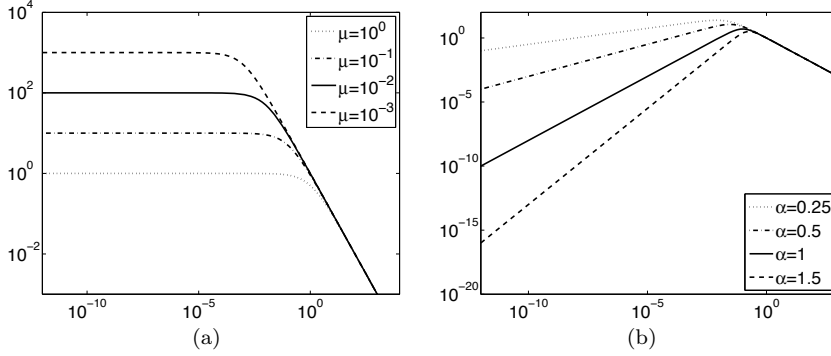
$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_n = 0,$$

where the index  $r$  is the rank of  $A$ . Using the spectral factorization (3.4), the solution of (3.3) can be written as

$$\mathbf{x}_{\mu,\alpha} = \sum_{j=1}^r \widehat{\varphi}(\lambda_j, \alpha) (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)}) \mathbf{u}_j, \quad \widehat{\varphi}(\lambda, \alpha) = \frac{\lambda^\alpha}{\lambda^{\alpha+1} + \mu}.$$

The filter factors  $\widehat{\varphi}(\lambda_j, \alpha)$  decrease as  $\alpha$  increases when the matrix  $A$  is scaled so that (2.2) holds.

Figure 3.1 displays the filter functions  $\lambda \mapsto \widehat{\varphi}(\lambda, \alpha)$  for standard Lavrentiev regularization (a) and for fractional Lavrentiev regularization (b). The former filter functions do not satisfy (2.4). We therefore expect the computed solutions obtained with the fractional Lavrentiev method to often be better approximations of the desired solution  $\widehat{\mathbf{x}}$  than the approximate solutions determined by standard Lavrentiev regularization.



**Fig. 3.1** (a) The filter factors  $(\lambda + \mu)^{-1}$  as function of  $\lambda$  for standard Lavrentiev for different values of  $\mu$ . (b) The filter factors  $\frac{\lambda^\alpha}{\lambda^{\alpha+1} + \mu}$  as function of  $\lambda$  for fractional Lavrentiev for  $\mu = 0.01$  and different values of  $\alpha$ .

#### 4 Choosing $\mu$ and $\alpha$ in fractional Lavrentiev regularization

The choice of  $\mu$  and  $\alpha$  for fractional Tikhonov regularization (2.6) is discussed in [10]. This section presents analogous results for the fractional Lavrentiev method. We consider two popular choices for the regularization parameter  $\mu$ : the discrepancy principle and the imposition of a solution norm constraint.

We first investigate the dependence of the solution  $\mathbf{x}_{\mu,\alpha}$  of (3.3) on the parameters  $\mu$  and  $\alpha$ . Substituting the spectral factorization (3.4) into (3.3) yields

$$(\Lambda^{\alpha+1} + \mu I) \mathbf{y} = \Lambda^\alpha U^T \mathbf{b}_{\mathcal{R}(A)}.$$

Denote the solution by  $\mathbf{y}_{\mu,\alpha}$ . Then  $\mathbf{x}_{\mu,\alpha} = U\mathbf{y}_{\mu,\alpha}$  solves (3.3), and we have

$$\|\mathbf{x}_{\mu,\alpha}\|^2 = \|\mathbf{y}_{\mu,\alpha}\|^2 = \sum_{j=1}^r \frac{\lambda_j^{2\alpha}}{(\lambda_j^{\alpha+1} + \mu)^2} (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2,$$

where  $r$  is the rank of  $A$ . Thus,

$$(4.1) \quad \frac{\partial}{\partial \mu} \|\mathbf{x}_{\mu,\alpha}\|^2 = -2 \sum_{j=1}^r \frac{\lambda_j^{2\alpha}}{(\lambda_j^{\alpha+1} + \mu)^3} (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2.$$

Clearly,  $\mu \mapsto \|\mathbf{x}_{\mu,\alpha}\|^2$  is a monotonically decreasing function. Similarly,

$$\frac{\partial}{\partial \alpha} \|\mathbf{x}_{\mu,\alpha}\|^2 = 2\mu \sum_{j=1}^r \frac{\log(\lambda_j) \lambda_j^{-\alpha}}{(\lambda_j + \mu \lambda_j^{-\alpha})^3} (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2.$$

Due to (2.2), we have  $\log(\lambda_j) < 0$  and it follows that  $\alpha \mapsto \|\mathbf{x}_{\mu,\alpha}\|^2$  is monotonically decreasing.

The choice of the regularization parameter  $\mu$  depends on the amount of error  $\mathbf{e}$  in  $\mathbf{b}$ . Consider for the moment standard Tikhonov regularization when  $A$  is symmetric. Then  $\alpha = 1$  in (2.7) and the normal equations become

$$(4.2) \quad (A^2 + \mu I)\mathbf{x} = A\mathbf{b}.$$

Generally, the larger  $\|\mathbf{e}\|$ , the larger  $\mu$  should be in order for the solution  $\mathbf{x}_\mu$  of (3.3) to be a suitable approximation of the desired solution  $\hat{\mathbf{x}}$ ; see, e.g., Proposition 4.1 below. However, it follows from (4.1) that increasing  $\mu$  decreases the norm of the solution  $\mathbf{x}_{\mu,1}$  of (4.2). Therefore, the computed solution may be of significantly smaller norm than the desired solution  $\hat{\mathbf{x}}$ . This difficulty can be remedied by choosing  $\alpha < 1$ , because this increases the norm of the computed solution.

Consider the situation when a fairly accurate bound for the error in  $\mathbf{b}$ ,

$$\|\mathbf{e}\| \leq \varepsilon,$$

is available. Then we can apply the discrepancy principle to determine a suitable value of the regularization parameter  $\mu$ . Let  $\alpha$  and  $0 < s < 1$  be fixed, define

$$(4.3) \quad \delta = \varepsilon^s,$$

and determine  $\mu > 0$  so that

$$(4.4) \quad \|\mathbf{b}_{\mathcal{R}(A)} - A\mathbf{x}_{\mu,\alpha}\| = \delta.$$

We refer to this approach to determine  $\mu$  as the discrepancy principle. It is shown in [15] that  $\mathbf{x}_{\mu,0} \rightarrow \hat{\mathbf{x}}$  as  $\varepsilon \searrow 0$  for this choice of  $\mu$ . A different choice of  $\mu$  is proposed by Groetsch and Cuacamene [7] for  $\alpha = 0$ , who also show that  $\mathbf{x}_{\mu,0}$  is not guaranteed to converge to  $\hat{\mathbf{x}}$  in a Hilbert space setting when  $\varepsilon \searrow 0$  and  $s = 1$ . Computed examples in [15] show that the choice of  $s$  can affect the



quality of the computed approximation  $\mathbf{x}_{\mu,\alpha}$  of  $\hat{\mathbf{x}}$  significantly when  $\alpha = 0$ . We will illustrate the dependence of  $\|\mathbf{x}_{\mu,\alpha} - \hat{\mathbf{x}}\|$  on  $s$  and  $\alpha$  in Section 6.

Results analogous to those shown in the remainder of this section have been established for Tikhonov regularization in [10], but many details of the proofs below and in [10] differ. We therefore provide complete proofs for the convenience of the reader. It follows from

$$\mathbf{b}_{\mathcal{R}(A)} - A\mathbf{x}_{\mu,\alpha} = \sum_{j=1}^r \left(1 - \lambda_j \frac{\lambda_j^\alpha}{\lambda_j^{\alpha+1} + \mu}\right) \mathbf{u}_j (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)}) + \sum_{j=r+1}^n \mathbf{u}_j (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})$$

that

$$(4.5) \quad \|\mathbf{b}_{\mathcal{R}(A)} - A\mathbf{x}_{\mu,\alpha}\|^2 = \sum_{j=1}^r \left(1 - \frac{\lambda_j^{\alpha+1}}{\lambda_j^{\alpha+1} + \mu}\right)^2 (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2 + \sum_{j=r+1}^n (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2.$$

Substituting  $\nu = \mu^{-1}$  into (4.5) shows that the solution of (4.4) for  $\mu > 0$  is equivalent to the computation of the positive zero of the function

$$(4.6) \quad F_\alpha(\nu) = \sum_{j=1}^r (\nu \lambda_j^{\alpha+1} + 1)^{-2} (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2 + \sum_{j=r+1}^n (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2 - \delta^2.$$

We are in a position to show how  $\mu$ , such that  $\mathbf{x}_{\mu,\alpha}$  satisfies (4.4) for fixed  $\alpha > 0$ , depends on  $\delta$ .

**Proposition 4.1** *Let  $\mu = \mu(\delta) > 0$  be such that  $\mathbf{x}_{\mu,\alpha}$  satisfies (4.4) for fixed  $\alpha > 0$ . Then  $d\mu/d\delta > 0$ .*

*Proof* Consider  $\nu(\delta) = 1/\mu(\delta)$ . It follows from (4.6) that the inverse function satisfies

$$\delta(\nu)^2 = \sum_{j=1}^r (\nu \lambda_j^{\alpha+1} + 1)^{-2} (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2 + \sum_{j=r+1}^n (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2.$$

Differentiating with respect to  $\nu$  yields

$$2\delta(\nu)\delta'(\nu) = -2 \sum_{j=1}^r \frac{\lambda_j^{\alpha+1}}{(\nu \lambda_j^{\alpha+1} + 1)^3} (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2.$$

It follows that  $\delta'(\nu) < 0$ . Consequently,  $\nu'(\delta) < 0$  and  $\mu'(\delta) > 0$ . □

We discuss Newton's method for computing the positive zero of the function (4.6). However, other zero-finders also can be used.

**Proposition 4.2** *Newton's method applied to the computation of the positive zero of  $F_\alpha$  with initial iterate  $\nu_0 = 0$  converges quadratically and monotonically.*

*Proof* The quadratic convergence is a consequence of the analyticity of  $F_\alpha(\nu)$  in a neighborhood of the positive real axis in the complex plane. The monotonic convergence follows from the fact that for every fixed  $\alpha > 0$  and  $\nu \geq 0$ , the function  $F_\alpha$  satisfies  $F'_\alpha(\nu) < 0$  and  $F''_\alpha(\nu) > 0$ .  $\square$

Let  $\alpha > 0$  and determine  $\mu = \mu(\alpha)$  so that  $\mathbf{x}_{\mu,\alpha}$  satisfies (4.4) with  $\delta$  given by (4.3) for some fixed  $0 < s < 1$ . The following result shows that the solution  $\mathbf{x}_{\mu,1}$  of (4.2) locally minimizes the solution norm.

**Proposition 4.3** *Let for  $\alpha > 0$  the regularization parameter  $\mu = \mu(\alpha)$  be such that  $\mathbf{x}_{\mu,\alpha}$  satisfies (4.4). Then there is an open real interval  $\Omega$  containing unity such that  $\argmin_{\alpha \in \Omega} \|\mathbf{x}_{\mu(\alpha),\alpha}\| = 1$ .*

*Proof* The equation  $F_\alpha(\nu) = 0$  can be expressed as

$$(4.7) \quad \sum_{j=1}^r \frac{\mu^2}{(\lambda_j^{\alpha+1} + \mu)^2} (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2 = \delta^2 - \sum_{j=r+1}^n (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2.$$

We may consider  $\mu = \mu(\alpha)$  a function of  $\alpha$ . Implicit differentiation of (4.7) with respect to  $\alpha$  yields

$$(4.8) \quad 2\mu \sum_{j=1}^r \frac{\lambda_j^{\alpha+1} (\mu' - \mu \log(\lambda_j))}{(\lambda_j^{\alpha+1} + \mu)^3} (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2 = 0,$$

which, since  $\mu > 0$ , implies that

$$(4.9) \quad \sum_{j=1}^r \xi_j (\mu' - \mu \log(\lambda_j)) = 0,$$

where

$$(4.10) \quad \xi_j = \frac{\lambda_j^{\alpha+1}}{(\lambda_j^{\alpha+1} + \mu)^3} (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2.$$

Introduce the function

$$G(\alpha) = \|\mathbf{x}_{\mu(\alpha),\alpha}\|^2 = \sum_{j=1}^r \frac{\lambda_j^{2\alpha}}{(\lambda_j^{\alpha+1} + \mu)^2} (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2.$$

Then

$$\begin{aligned} G'(\alpha) &= \sum_{j=1}^r \frac{2\lambda_j^{2\alpha} \log(\lambda_j) (\lambda_j^{\alpha+1} + \mu) - 2\lambda_j^{2\alpha} (\lambda_j^{\alpha+1} \log(\lambda_j) + \mu')}{(\lambda_j^{\alpha+1} + \mu)^3} (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2 \\ &= \sum_{j=1}^r \frac{2\lambda_j^{2\alpha} (\log(\lambda_j) \mu - \mu')}{(\lambda_j^{\alpha+1} + \mu)^3} (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2 \\ &= 2 \sum_{j=1}^r \xi_j \lambda_j^{2\alpha} (\mu \log(\lambda_j) - \mu'). \end{aligned}$$

It follows from (4.9) that  $G'(1) = 0$ . Moreover, differentiating (4.9) yields

$$(4.11) \quad \sum_{j=1}^r \left\{ \xi'_j(\mu' - \mu \log(\lambda_j)) + \xi_j(\mu'' - \mu' \log(\lambda_j)) \right\} = 0.$$

Since

$$G''(\alpha) = 2 \sum_{j=1}^r \lambda_j^{2\alpha} \left\{ (\xi'_j + \xi_j \log(\lambda_j))(\mu \log(\lambda_j) - \mu') + \xi_j(\mu' \log(\lambda_j) - \mu'') \right\},$$

we obtain, in view of (4.11), that

$$G''(1) = 2 \sum_{j=1}^r \xi_j \log(\lambda_j)(\mu \log(\lambda_j) - \mu').$$

The above sum can be determined by multiplying the terms in (4.9) by the positive weights  $-\log(\lambda_j)$ ; the largest weights multiply the largest terms. Therefore,  $G''(1) > 0$ . By continuity,  $G''(\alpha)$  is positive in a neighborhood  $\Omega$  of  $\alpha = 1$ . Thus,  $G(\alpha)$  has a local minimum at  $\alpha = 1$ .  $\square$

For some linear discrete ill-posed problems (1.1) an estimate  $\Delta$  of the norm of the desired solution  $\hat{\mathbf{x}}$  may be known. Then it might be desirable to require the computed solution  $\mathbf{x}_{\mu,\alpha}$  to be of the same norm, i.e.,

$$(4.12) \quad \Delta = \|\mathbf{x}_{\mu,\alpha}\|.$$

This type of problems is discussed in [2, 11, 13, 17]. The following result shows that the Tikhonov solution  $\mathbf{x}_{\mu,1}$  of (4.2) locally minimizes the solution norm.

**Proposition 4.4** *Let, for  $\alpha > 0$ , the regularization parameter  $\mu = \mu(\alpha)$  be such that  $\mathbf{x}_{\mu,\alpha}$  satisfies (4.12). Then there is an open real interval  $\Omega$  containing unity, such that  $\operatorname{argmin}_{\alpha \in \Omega} \|\mathbf{b}_{\mathcal{R}(A)} - A\mathbf{x}_{\mu(\alpha),\alpha}\| = 1$ .*

*Proof* This result is shown in a similar fashion as Proposition 4.3. Differentiating the squares of the right-hand and left-hand sides of (4.12) with respect to  $\alpha$ , keeping in mind that  $\mu = \mu(\alpha)$ , gives analogously to (4.9) the equation

$$(4.13) \quad \sum_{j=1}^r \zeta_j (\mu \log(\lambda_j) - \mu') = 0,$$

where  $\zeta_j = \xi_j \lambda_j^{2\alpha}$  and  $\xi_j$  is defined by (4.10). Introduce the function

$$H(\alpha) = \|\mathbf{b}_{\mathcal{R}(A)} - A\mathbf{x}_{\mu(\alpha),\alpha}\|^2 = \sum_{j=1}^r \frac{\mu^2}{(\lambda_j^{\alpha+1} + \mu)^2} (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2 + \sum_{j=r+1}^n (\mathbf{u}_j^T \mathbf{b}_{\mathcal{R}(A)})^2;$$

cf. (4.5). Using (4.8) now yields

$$H'(\alpha) = 2\mu \sum_{j=1}^r \zeta_j \lambda_j^{1-\alpha} (\mu' - \mu \log(\lambda_j))$$

and it follows from (4.13) that  $H'(1) = 0$ . Differentiating the representation

$$H'(\alpha) = 2\mu \sum_{j=1}^r \xi_j (\mu' - \mu \log(\lambda_j))$$

gives

$$(4.14) \quad \begin{aligned} H''(\alpha) = & 2\mu' \sum_{j=1}^r \xi_j (\mu' - \mu \log(\lambda_j)) \\ & + 2\mu \sum_{j=1}^r [\xi_j' (\mu' - \mu \log(\lambda_j)) + \xi_j (\mu'' - \mu' \log(\lambda_j))], \end{aligned}$$

and differentiating (4.13) yields

$$(4.15) \quad \sum_{j=1}^r \zeta_j' (\mu \log(\lambda_j) - \mu') + \zeta_j (\mu' \log(\lambda_j) - \mu'') = 0.$$

Let  $\alpha = 1$ . Then  $\zeta_j = \xi_j$  for all  $j$ . Using this property when substituting (4.15) into (4.14), we then obtain, in view of (4.13), that

$$(4.16) \quad H''(1) = 2\mu \sum_{j=1}^r (\xi_j' - \zeta_j') (\mu' - \mu \log(\lambda_j)).$$

It follows from  $\xi_j = \zeta_j \lambda_j^{1-\alpha}$  that, for  $\alpha = 1$ ,  $\xi_j' = \zeta_j' - \zeta_j \log(\lambda_j)$ . Substituting the latter expression into (4.16) yields

$$H''(1) = -2\mu \sum_{j=1}^r \zeta_j (\mu' - \mu \log(\lambda_j)) \log(\lambda_j).$$

Comparing this sum with (4.13) shows that  $H''(1) > 0$ , similarly as the analogous result for  $G''(1)$  in the proof of Proposition 4.3. We conclude similarly as in the proof of Proposition 4.3 that  $H$  is convex in a neighborhood  $\Omega$  of  $\alpha = 1$  and therefore has a local minimum at  $\alpha = 1$ .  $\square$

Propositions 4.3 and 4.4 show the choice  $\alpha = 1$  to be quite natural. By Proposition 4.3 this choice minimizes  $\|\mathbf{x}_{\mu(\alpha), \alpha}\|$  locally when the residual norm  $\|\mathbf{b}_{\mathcal{R}(A)} - A\mathbf{x}_{\mu(\alpha), \alpha}\|$  is specified and by Propositions 4.4 the residual norm has a local minimum for  $\alpha = 1$  when  $\|\mathbf{x}_{\mu(\alpha), \alpha}\|$  is specified. We remark that the value of  $\delta$  used in Proposition 4.3 does not have to be defined by (4.3) and, similarly, the value of  $\Delta$  in Proposition 4.4 does not have to be close to  $\|\hat{\mathbf{x}}\|$ . Nevertheless, numerical examples of Section 6 show  $0 < \alpha < 1$  to yield more accurate approximations of  $\hat{\mathbf{x}}$  than  $\alpha = 1$ .

## 5 Large-scale problems

The computations described in the previous section use the spectral factorization (3.4) of the symmetric positive definite matrix  $A \in \mathbb{R}^{n \times n}$ . When this matrix is large, the evaluation of its spectral factorization might be too expensive to be attractive. In this situation, we may use the symmetric Lanczos process to reduce  $A$  to a small matrix. Application of  $\ell$  steps of the symmetric Lanczos process to  $A$  with initial vector  $\mathbf{b}/\|\mathbf{b}\|$  yields the Lanczos decomposition

$$(5.1) \quad AW_\ell = W_\ell T_\ell + \mathbf{f}_\ell \mathbf{e}_\ell^T,$$

where the orthonormal columns of the matrix  $W_\ell \in \mathbb{R}^{n \times \ell}$  form a basis for the Krylov subspace  $\text{span}\{\mathbf{b}, A\mathbf{b}, \dots, A^{\ell-1}\mathbf{b}\}$  with  $W_\ell \mathbf{e}_1 = \mathbf{b}/\|\mathbf{b}\|$ ,  $T_\ell \in \mathbb{R}^{\ell \times \ell}$  is symmetric and tridiagonal, and  $\mathbf{f}_\ell \in \mathbb{R}^n$  satisfies  $W_\ell^T \mathbf{f}_\ell = \mathbf{0}$ . Here  $\mathbf{e}_j$  denotes the  $j$ th axis vector; see, e.g., [5, Chapter 9] for details on the symmetric Lanczos process. We may assume that  $\ell$  is chosen small enough so that a decomposition of the form (5.1) with the stated properties exists. The matrix  $T_\ell$  is an orthogonal projection of  $A$  onto  $\mathcal{R}(W_\ell)$ . Therefore its eigenvalues live in the convex hull of the spectrum of  $A$  and it follows that  $T_\ell$  is positive semidefinite.

We determine an approximate solution of (1.1) in  $\mathcal{R}(W_\ell)$  by a Galerkin method. This leads to the reduced problem

$$(5.2) \quad \min_{\mathbf{z} \in \mathbb{R}^\ell} \|T_\ell \mathbf{z} - \mathbf{e}_1\| \|\mathbf{b}\|,$$

which can be solved by the fractional Lavrentiev method. The so computed approximate solution  $\mathbf{z}_{\mu, \alpha}$  of (5.2) gives the approximate solution  $\mathbf{x}_{\mu, \alpha, \ell} = W_\ell \mathbf{z}_{\mu, \alpha}$  of (1.1). Since  $\ell$  typically can be chosen quite small, the spectral factorization of  $T_\ell$  can be readily evaluated.

We remark that in order for the fractional Lavrentiev method to be applicable, it suffices that  $T_\ell$  is symmetric and positive semidefinite. However, typically the matrix  $T_\ell$  determined by the symmetric Lanczos process is positive definite, because generally the vector  $\mathbf{f}_\ell$  in (5.1) is nonvanishing.

**Proposition 5.1** *Let the matrices  $A$  and  $T_\ell$ , and the vector  $\mathbf{f}_\ell$ , be those in the Lanczos decomposition (5.1). Assume that  $A$  is symmetric positive semidefinite and that  $\mathbf{f}_\ell \neq \mathbf{0}$ . Then  $T_\ell$  is positive definite and  $\mathcal{N}(A) \cap \mathcal{R}(W_\ell) = \emptyset$ .*

*Proof* Since  $\mathbf{f}_\ell \neq \mathbf{0}$  another Lanczos step can be carried out, which yields the symmetric positive semidefinite tridiagonal matrix  $T_{\ell+1} \in \mathbb{R}^{(\ell+1) \times (\ell+1)}$ . This matrix has nonvanishing subdiagonal entries. The eigenvalues of  $T_\ell$  strictly interlace those of  $T_{\ell+1}$ ; see, e.g., [9, Theorem 1]. It follows that the smallest eigenvalue of  $T_\ell$  is positive.

Assume that  $\mathcal{N}(A) \cap \mathcal{R}(W_\ell) \neq \emptyset$ . Then there is a vector  $\mathbf{y} \in \mathbb{R}^\ell \setminus \{\mathbf{0}\}$  such that  $AW_\ell \mathbf{y} = \mathbf{0}$ . We obtain from (5.1) that

$$\mathbf{0} = AW_\ell \mathbf{y} = W_\ell T_\ell \mathbf{y} + \mathbf{f}_\ell \mathbf{e}_\ell^T \mathbf{y}.$$

It follows that  $\mathbf{e}_\ell^T \mathbf{y} \neq 0$  and therefore  $\mathbf{f}_\ell \in \mathcal{R}(W_\ell)$ . This is a contradiction.  $\square$

When  $A \in \mathbb{R}^{m \times n}$  is a large nonsymmetric matrix, it can be reduced to a smaller one by partial Golub–Kahan bidiagonalization. Application of  $\ell$  steps of Golub–Kahan bidiagonalization to  $A$  with initial vector  $\mathbf{b}/\|\mathbf{b}\|$  yields decompositions that determine a Tikhonov minimization problem analogous to (2.1) with  $\ell \times \ell$  matrices. Generally,  $\ell$  can be chosen fairly small and the reduced problem can be solved as described in Section 2. Further details can be found in [10].

## 6 Computed examples

The examples of this section illustrate the performance of the fractional Lavrentiev method when the regularization parameter  $\mu = \mu(\alpha)$  is determined by the discrepancy principle (4.4). All computations were carried out in MATLAB with machine epsilon  $2.2 \cdot 10^{-16}$ .

**Example 6.1.** We consider the Fredholm integral equation of the first kind

$$(6.1) \quad \int_0^\pi \exp(u \cos(v)) x(v) dv = 2 \frac{\sinh(u)}{u}, \quad 0 \leq u \leq \frac{\pi}{2}$$

discussed by Baart [1]. We first discretize (6.1) by a Galerkin method using the function `baart` from the MATLAB package Regularization Tools by Hansen [8] to determine the matrix  $B \in \mathbb{R}^{100 \times 100}$  and a scaled discrete approximation  $\hat{\mathbf{x}} \in \mathbb{R}^{100}$  of the solution  $x(v) = \sin(v)$  of (6.1). Fractional Lavrentiev regularization is applied to the symmetric positive semidefinite matrix  $A = BB^T$  scaled to have norm 0.5. The matrix is of ill-determined rank. In fact, the MATLAB function `cond` gives the condition number  $\kappa(B) = 5 \cdot 10^{18}$ , where  $\kappa(B) := \|B\| \|B^{-1}\|$ . This yields  $\kappa(A) = \kappa(B)^2 = 2 \cdot 10^{37}$ . In particular,  $A$  is numerically singular. Let  $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$ . An error vector  $\mathbf{e}$  with normally distributed random entries with zero mean is scaled to correspond to a chosen noise level  $\epsilon = \|\mathbf{e}\|/\|\hat{\mathbf{b}}\|$  and added to  $\hat{\mathbf{b}}$ ; cf. (1.2).

Table 6.1 reports relative errors  $\|\mathbf{x}_{\mu,\alpha} - \hat{\mathbf{x}}\|/\|\hat{\mathbf{x}}\|$  for several noise levels  $\epsilon$ . The regularization parameter  $\mu$  is determined by the discrepancy principle (4.4) for powers  $s \in \{0.5, 0.6, \dots, 1\}$  in (4.3). The column labeled “Lavrentiev” reports the relative errors obtained for  $\alpha = 0$ , and the column “Fractional Lavrentiev” displays the relative errors achieved for the optimal value  $\alpha^*$ . Specifically, for each  $s$  and  $\epsilon$ , the column labeled  $\alpha^*$  shows the value of the parameter  $\alpha \in \{0, 0.1, 0.2, \dots, 1\}$  that gives the most accurate approximation  $\mathbf{x}_{\mu,\alpha}$  of  $\hat{\mathbf{x}}$  when  $\mu$  is determined by the discrepancy principle. The table shows the relative error  $\|\mathbf{x}_{\mu,\alpha} - \hat{\mathbf{x}}\|/\|\hat{\mathbf{x}}\|$  for standard Lavrentiev to increase with  $s$  for  $s$  close to unity. In particular,  $s = 1$  yields the largest errors. Fractional Lavrentiev gives the smallest errors for  $s = 1$ . The optimal value  $\alpha^*$  is seen to increase with  $s$  when  $s$  is close to one.

We now discretize (6.1) more accurately with the function `baart` and obtain the matrix  $B \in \mathbb{R}^{1000 \times 1000}$  and a scaled discrete approximation  $\hat{\mathbf{x}} \in \mathbb{R}^{1000}$  of the solution of (6.1). Table 6.2 is analogous to Table 6.1 for this larger problem.

Noise level %	$s$	Lavrentiev	Fractional Lavrentiev $\alpha^*$	
5	0.5	$6.08 \cdot 10^{-1}$	0.1	$6.06 \cdot 10^{-1}$
	0.6	$5.93 \cdot 10^{-1}$	0.1	$5.75 \cdot 10^{-1}$
	0.7	$6.11 \cdot 10^{-1}$	0.1	$5.59 \cdot 10^{-1}$
	0.8	$6.86 \cdot 10^{-1}$	0.2	$5.52 \cdot 10^{-1}$
	0.9	$9.09 \cdot 10^{-1}$	0.5	$5.47 \cdot 10^{-1}$
	1.0	$2.94 \cdot 10^0$	0.8	$5.44 \cdot 10^{-1}$
1	0.5	$5.30 \cdot 10^{-1}$	0	$5.30 \cdot 10^{-1}$
	0.6	$5.13 \cdot 10^{-1}$	0	$5.13 \cdot 10^{-1}$
	0.7	$5.14 \cdot 10^{-1}$	0.1	$4.98 \cdot 10^{-1}$
	0.8	$5.67 \cdot 10^{-1}$	0.1	$4.82 \cdot 10^{-1}$
	0.9	$7.46 \cdot 10^{-1}$	0.1	$4.70 \cdot 10^{-1}$
	1.0	$4.13 \cdot 10^0$	0.2	$4.61 \cdot 10^{-1}$
0.1	0.5	$4.83 \cdot 10^{-1}$	0	$4.83 \cdot 10^{-1}$
	0.6	$4.62 \cdot 10^{-1}$	0	$4.62 \cdot 10^{-1}$
	0.7	$4.49 \cdot 10^{-1}$	0	$4.49 \cdot 10^{-1}$
	0.8	$4.55 \cdot 10^{-1}$	0.1	$4.37 \cdot 10^{-1}$
	0.9	$5.48 \cdot 10^{-1}$	0.1	$4.11 \cdot 10^{-1}$
	1.0	$5.48 \cdot 10^{-1}$	0.1	$4.11 \cdot 10^{-1}$

**Table 6.1** Example 6.1: Discretization of the integral equation (6.1). Relative errors of the approximate solutions. Dimension  $n = 100$ .

Noise level %	$s$	Lavrentiev	Fractional Lavrentiev $\alpha^*$	
5	0.5	$5.99 \cdot 10^{-1}$	0	$5.99 \cdot 10^{-1}$
	0.6	$5.78 \cdot 10^{-1}$	0.1	$5.63 \cdot 10^{-1}$
	0.7	$5.91 \cdot 10^{-1}$	0.1	$5.41 \cdot 10^{-1}$
	0.8	$6.65 \cdot 10^{-1}$	0.1	$5.25 \cdot 10^{-1}$
	0.9	$9.22 \cdot 10^{-1}$	0.1	$5.08 \cdot 10^{-1}$
	1.0	$2.10 \cdot 10^1$	0.2	$4.61 \cdot 10^{-1}$
1	0.5	$5.27 \cdot 10^{-1}$	0	$5.27 \cdot 10^{-1}$
	0.6	$5.09 \cdot 10^{-1}$	0	$5.09 \cdot 10^{-1}$
	0.7	$5.06 \cdot 10^{-1}$	0.1	$4.94 \cdot 10^{-1}$
	0.8	$5.52 \cdot 10^{-1}$	0.1	$4.47 \cdot 10^{-1}$
	0.9	$7.29 \cdot 10^{-1}$	0.1	$4.50 \cdot 10^{-1}$
	1.0	$4.19 \cdot 10^0$	0.1	$3.17 \cdot 10^{-1}$
0.1	0.5	$4.85 \cdot 10^{-1}$	0	$4.85 \cdot 10^{-1}$
	0.6	$4.66 \cdot 10^{-1}$	0	$4.66 \cdot 10^{-1}$
	0.7	$4.56 \cdot 10^{-1}$	0	$4.56 \cdot 10^{-1}$
	0.8	$4.69 \cdot 10^{-1}$	0.1	$4.44 \cdot 10^{-1}$
	0.9	$5.68 \cdot 10^{-1}$	0.1	$4.24 \cdot 10^{-1}$
	1.0	$5.68 \cdot 10^{-1}$	0.1	$4.24 \cdot 10^{-1}$

**Table 6.2** Example 6.1: Discretization of the integral equation (6.1). Relative errors of the approximate solutions. Dimension  $n = 1000$ .

Fractional Lavrentiev regularization can be seen to perform similarly as in Table 6.1.  $\square$

**Example 6.2.** The Fredholm integral equation of the first kind

$$(6.2) \quad \int_{-6}^6 \kappa(u, v) x(v) dv = g(u), \quad -6 \leq u \leq 6,$$

with kernel  $\kappa$  and right-hand side  $g$  given by

$$\begin{aligned} \kappa(u, v) &= \phi(u - v), \\ g(u) &= (6 - |u|)(1 + \tfrac{1}{2} \cos(\tfrac{\pi u}{3})) + \tfrac{9}{2\pi} \sin(\tfrac{\pi |u|}{3}), \end{aligned}$$

where

$$\phi(y) = \begin{cases} 1 + \cos(\tfrac{\pi y}{3}), & |y| < 3, \\ 0, & |y| \geq 3, \end{cases}$$

is discussed by Phillips [16]. We discretize this integral equation by a Galerkin method with orthonormal box functions as test and trial functions using the function `phillips` from [8]. This function yields  $A \in \mathbb{R}^{100 \times 100}$  and a scaled discrete approximation  $\hat{\mathbf{x}} \in \mathbb{R}^{100}$  of the solution  $x(v) = \phi(v)$  of (6.2). The matrix  $A$  is scaled to have norm 0.5, and is symmetric indefinite and of ill-determined rank. Its condition number is  $\kappa(A) = \|A\| \|A^{-1}\| = 2.64 \cdot 10^6$ . The error-free right-hand side vector is computed as  $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$ . The error vector  $\mathbf{e}$  has normally distributed entries with zero mean; they are scaled to correspond to specified noise levels. The vector  $\mathbf{b}$  in (1.1) is determined by (1.2).

Application of  $\ell = 5$  steps of the symmetric Lanczos process to  $A$  with initial vector  $\mathbf{b}/\|\mathbf{b}\|$  gives the decomposition (5.1) with a symmetric positive definite matrix  $T_5$ . We apply the fractional Lavrentiev method to the reduced problem (5.2).

Table 6.3 is analogous to Table 6.1 and reports relative errors  $\|\mathbf{x}_{\mu,0,5} - \hat{\mathbf{x}}\|/\|\hat{\mathbf{x}}\|$  and  $\|\mathbf{x}_{\mu,\alpha^*,5} - \hat{\mathbf{x}}\|/\|\hat{\mathbf{x}}\|$  for several values of  $s$  and  $\epsilon$ . Similarly as in Table 6.1, the error for standard Lavrentiev grows and the error for fractional Lavrentiev decreases as  $s$  increases. The value  $\alpha^*$  can be seen to be an increasing function of  $s$ .

We turn to a finer discretization of the integral equation (6.2) using 1000 orthonormal test and trial functions. The function `phillips` gives  $A \in \mathbb{R}^{1000 \times 1000}$  and a scaled discrete approximation  $\hat{\mathbf{x}} \in \mathbb{R}^{1000}$  of the solution of (6.2). The results are displayed in Table 6.4. They are seen to be similar to those of Table 6.3.  $\square$

**Example 6.3.** Consider the integral equation of the first kind

$$(6.3) \quad \int_0^1 \kappa(u, v) x(v) dv = g(u), \quad 0 \leq u \leq 1,$$

where the kernel and the right-hand side function  $g$  are given by

$$\kappa(u, v) = \begin{cases} u(v-1), & u < v, \\ v(u-1), & u \geq v, \end{cases} \quad g(u) = \frac{u^3 - u}{6}.$$

The function `deriv2` from [8] determines a discretization of (6.3) with a symmetric negative definite matrix  $A \in \mathbb{R}^{100 \times 100}$  and a scaled discrete approximation



Noise level %	$s$	Lavrentiev	Fractional Lavrentiev $\alpha^*$	
5	0.5	$3.00 \cdot 10^{-1}$	0.1	$2.59 \cdot 10^{-1}$
	0.6	$3.31 \cdot 10^{-1}$	0.2	$2.15 \cdot 10^{-1}$
	0.7	$4.27 \cdot 10^{-1}$	0.3	$1.79 \cdot 10^{-1}$
	0.8	$5.97 \cdot 10^{-1}$	0.3	$1.48 \cdot 10^{-1}$
	0.9	$9.33 \cdot 10^{-1}$	0.4	$1.22 \cdot 10^{-1}$
	1.0	$2.24 \cdot 10^0$	0.6	$9.32 \cdot 10^{-2}$
1	0.5	$1.44 \cdot 10^{-1}$	0.1	$1.24 \cdot 10^{-1}$
	0.6	$1.73 \cdot 10^{-1}$	0.2	$9.31 \cdot 10^{-2}$
	0.7	$3.59 \cdot 10^{-1}$	0.3	$7.07 \cdot 10^{-2}$
	0.8	$4.04 \cdot 10^{-1}$	0.4	$5.55 \cdot 10^{-2}$
	0.9	$6.56 \cdot 10^{-1}$	0.5	$4.42 \cdot 10^{-2}$
	1.0	$1.77 \cdot 10^0$	0.7	$3.16 \cdot 10^{-2}$
0.1	0.5	$2.65 \cdot 10^{-2}$	0.2	$5.32 \cdot 10^{-2}$
	0.6	$1.79 \cdot 10^{-1}$	0.4	$4.02 \cdot 10^{-2}$
	0.7	$4.60 \cdot 10^{-1}$	0.6	$2.91 \cdot 10^{-2}$
	0.8	$4.60 \cdot 10^{-1}$	0.6	$2.91 \cdot 10^{-2}$
	0.9	$2.31 \cdot 10^{-1}$	0.7	$2.55 \cdot 10^{-2}$
	1.0	$2.31 \cdot 10^{-1}$	0.7	$2.55 \cdot 10^{-2}$

**Table 6.3** Example 6.2: Discretization of the integral equation (6.2). Relative errors of the approximate solutions. Dimension  $n = 100$ .

Noise level %	$s$	Lavrentiev	Fractional Lavrentiev $\alpha^*$	
5	0.5	$3.03 \cdot 10^{-1}$	0.1	$2.46 \cdot 10^{-1}$
	0.6	$3.44 \cdot 10^{-1}$	0.2	$2.00 \cdot 10^{-1}$
	0.7	$4.57 \cdot 10^{-1}$	0.2	$1.58 \cdot 10^{-1}$
	0.8	$6.72 \cdot 10^{-1}$	0.3	$1.22 \cdot 10^{-1}$
	0.9	$1.10 \cdot 10^0$	0.4	$9.07 \cdot 10^{-2}$
	1.0	$7.39 \cdot 10^0$	0.6	$3.52 \cdot 10^{-2}$
1	0.5	$1.51 \cdot 10^{-1}$	0.1	$1.25 \cdot 10^{-1}$
	0.6	$1.86 \cdot 10^{-1}$	0.2	$9.30 \cdot 10^{-2}$
	0.7	$2.84 \cdot 10^{-1}$	0.3	$7.00 \cdot 10^{-2}$
	0.8	$4.49 \cdot 10^{-1}$	0.4	$5.46 \cdot 10^{-2}$
	0.9	$9.06 \cdot 10^{-1}$	0.5	$3.91 \cdot 10^{-2}$
	1.0	$2.36 \cdot 10^0$	0.6	$2.94 \cdot 10^{-2}$
0.1	0.5	$7.65 \cdot 10^{-2}$	0.1	$5.36 \cdot 10^{-2}$
	0.6	$1.79 \cdot 10^{-1}$	0.4	$4.05 \cdot 10^{-2}$
	0.7	$4.51 \cdot 10^{-1}$	0.6	$2.89 \cdot 10^{-2}$
	0.8	$4.51 \cdot 10^{-1}$	0.6	$2.89 \cdot 10^{-2}$
	0.9	$2.27 \cdot 10^{-1}$	0.7	$2.54 \cdot 10^{-2}$
	1.0	$2.27 \cdot 10^{-1}$	0.7	$2.54 \cdot 10^{-2}$

**Table 6.4** Example 6.2: Discretization of the integral equation (6.2). Relative errors of the approximate solutions. Dimension  $n = 1000$ .

$\hat{\mathbf{x}} \in \mathbb{R}^{100}$  of the solution  $x(v) = v$ . Fractional Lavrentiev regularization is applied to the matrix  $-A$  scaled to have norm 0.5. Let  $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$ . An error vector  $\mathbf{e}$  with normally distributed random entries with zero mean, scaled to corre-

spond to a chosen noise level  $\epsilon = \|e\|/\|\hat{\mathbf{b}}\|$ , is added to  $\hat{\mathbf{b}}$  to yield the vector  $\mathbf{b}$  in (1.1).

Noise level %	$s$	Lavrentiev	Fractional Lavrentiev $\alpha^*$	
5	0.5	$6.55 \cdot 10^{-1}$	0	$6.55 \cdot 10^{-1}$
	0.6	$5.42 \cdot 10^{-1}$	0	$5.42 \cdot 10^{-1}$
	0.7	$4.82 \cdot 10^{-1}$	0	$4.82 \cdot 10^{-1}$
	0.8	$5.01 \cdot 10^{-1}$	0.1	$4.38 \cdot 10^{-1}$
	0.9	$7.74 \cdot 10^{-1}$	0.2	$4.01 \cdot 10^{-1}$
	1.0	$1.77 \cdot 10^0$	0.4	$3.63 \cdot 10^{-1}$
1	0.5	$4.93 \cdot 10^{-1}$	0	$4.93 \cdot 10^{-1}$
	0.6	$4.15 \cdot 10^{-1}$	0	$4.15 \cdot 10^{-1}$
	0.7	$3.69 \cdot 10^{-1}$	0	$3.69 \cdot 10^{-1}$
	0.8	$3.86 \cdot 10^{-1}$	0.1	$3.40 \cdot 10^{-1}$
	0.9	$5.72 \cdot 10^{-1}$	0.2	$3.16 \cdot 10^{-1}$
	1.0	$1.69 \cdot 10^{-1}$	0.4	$2.92 \cdot 10^{-1}$
0.1	0.5	$3.48 \cdot 10^{-1}$	0	$3.48 \cdot 10^{-1}$
	0.6	$2.79 \cdot 10^{-1}$	0	$2.79 \cdot 10^{-1}$
	0.7	$2.30 \cdot 10^{-1}$	0	$2.30 \cdot 10^{-1}$
	0.8	$2.42 \cdot 10^{-1}$	0.1	$2.14 \cdot 10^{-1}$
	0.9	$3.46 \cdot 10^{-1}$	0.2	$2.06 \cdot 10^{-1}$
	1.0	$3.46 \cdot 10^{-1}$	0.2	$2.06 \cdot 10^{-1}$

**Table 6.5** Example 6.3: Discretization of the integral equation (6.3). Relative errors of the approximate solution. Dimension  $n = 100$ .

We show in Table 6.5 relative errors for several values of  $s$  and  $\epsilon$  for  $\alpha = 0$  and  $\alpha = \alpha^*$ . Fractional Lavrentiev can be seen to give the smallest error when  $s = 1$ .

We also apply the function `deriv2` to determine a symmetric negative definite matrix  $A \in \mathbb{R}^{1000 \times 1000}$  and a scaled discrete approximation  $\hat{\mathbf{x}} \in \mathbb{R}^{1000}$  of the solution of (6.3). The results are displayed in Table 6.6. They are similar to those of Table 6.5.  $\square$

**Example 6.4.** The function `wing` from [8] determines a discretization of the Fredholm integral equation of the first kind

$$(6.4) \quad \int_0^1 t \exp(-st^2) x(t) dt = \frac{1}{2s} \left( \exp\left(-\frac{s}{9}\right) - \exp\left(-\frac{4s}{9}\right) \right), \quad 0 < s < 1,$$

with the discontinuous solution

$$x(t) = \begin{cases} 1, & \frac{1}{3} < t < \frac{2}{3}, \\ 0, & \text{elsewhere.} \end{cases}$$

This integral equation is discussed in [18]. We compute the matrix  $B \in \mathbb{R}^{100 \times 100}$  with `wing` and form  $A = B^T B$ . The matrix  $B$  is numerically singular and, therefore, so is  $A$ . The function `wing` also yields a scaled discretization

Noise level %	$s$	Lavrentiev	Fractional Lavrentiev $\alpha^*$	
5	0.5	$6.69 \cdot 10^{-1}$	0	$6.69 \cdot 10^{-1}$
	0.6	$5.56 \cdot 10^{-1}$	0	$5.56 \cdot 10^{-1}$
	0.7	$5.08 \cdot 10^{-1}$	0.1	$5.00 \cdot 10^{-1}$
	0.8	$5.31 \cdot 10^{-1}$	0.1	$4.46 \cdot 10^{-1}$
	0.9	$7.61 \cdot 10^{-1}$	0.2	$4.05 \cdot 10^{-1}$
	1.0	$6.48 \cdot 10^0$	0.4	$3.19 \cdot 10^{-1}$
1	0.5	$4.89 \cdot 10^{-1}$	0	$4.89 \cdot 10^{-1}$
	0.6	$4.09 \cdot 10^{-1}$	0	$4.09 \cdot 10^{-1}$
	0.7	$3.63 \cdot 10^{-1}$	0	$3.63 \cdot 10^{-1}$
	0.8	$3.96 \cdot 10^{-1}$	0.1	$3.20 \cdot 10^{-1}$
	0.9	$6.05 \cdot 10^{-1}$	0.2	$2.95 \cdot 10^{-1}$
	1.0	$3.84 \cdot 10^0$	0.4	$2.44 \cdot 10^{-1}$
0.1	0.5	$3.48 \cdot 10^{-1}$	0	$3.48 \cdot 10^{-1}$
	0.6	$2.79 \cdot 10^{-1}$	0	$2.79 \cdot 10^{-1}$
	0.7	$2.32 \cdot 10^{-1}$	0	$2.32 \cdot 10^{-1}$
	0.8	$2.61 \cdot 10^{-1}$	0.1	$2.08 \cdot 10^{-1}$
	0.9	$4.12 \cdot 10^{-1}$	0.2	$1.99 \cdot 10^{-1}$
	1.0	$4.12 \cdot 10^{-1}$	0.2	$1.99 \cdot 10^{-1}$

**Table 6.6** Example 6.3: Discretization of the integral equation (6.3). Relative errors of the approximate solutions. Dimension  $n = 1000$ .

Noise level %	$s$	Lavrentiev	Fractional Lavrentiev $\alpha^*$	
5	0.5	$8.88 \cdot 10^{-1}$	0	$8.88 \cdot 10^{-1}$
	0.6	$8.48 \cdot 10^{-1}$	0.1	$8.47 \cdot 10^{-1}$
	0.7	$8.39 \cdot 10^{-1}$	0.1	$8.30 \cdot 10^{-1}$
	0.8	$8.53 \cdot 10^{-1}$	0.1	$8.23 \cdot 10^{-1}$
	0.9	$9.27 \cdot 10^{-1}$	0.2	$8.19 \cdot 10^{-1}$
	1.0	$2.93 \cdot 10^0$	0.3	$8.11 \cdot 10^{-1}$
1	0.5	$8.34 \cdot 10^{-1}$	0	$8.34 \cdot 10^{-1}$
	0.6	$8.24 \cdot 10^{-1}$	0	$8.24 \cdot 10^{-1}$
	0.7	$8.22 \cdot 10^{-1}$	0	$8.22 \cdot 10^{-1}$
	0.8	$8.27 \cdot 10^{-1}$	0.1	$8.22 \cdot 10^{-1}$
	0.9	$8.72 \cdot 10^{-1}$	0.1	$8.20 \cdot 10^{-1}$
	1.0	$2.85 \cdot 10^0$	0.1	$8.07 \cdot 10^{-1}$
0.1	0.5	$8.21 \cdot 10^{-2}$	0	$8.21 \cdot 10^{-1}$
	0.6	$8.17 \cdot 10^{-1}$	0	$8.17 \cdot 10^{-1}$
	0.7	$8.12 \cdot 10^{-1}$	0	$8.12 \cdot 10^{-1}$
	0.8	$8.08 \cdot 10^{-1}$	0	$8.08 \cdot 10^{-1}$
	0.9	$8.32 \cdot 10^{-1}$	0.1	$8.04 \cdot 10^{-1}$
	1.0	$8.32 \cdot 10^{-1}$	0.1	$8.04 \cdot 10^{-1}$

**Table 6.7** Example 6.4: Discretization of the integral equation (6.4). Relative errors of the approximate solutions. Dimension:  $n = 100$ .

of the solution  $\hat{\mathbf{x}}$  from which we compute  $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$ . The error-contaminated vector  $\mathbf{b}$  in (1.1) is determined similarly as in the previous examples.

Table 6.7 displays the computed results. In particular, the table shows that also when the desired solution  $\hat{\mathbf{x}}$  is discontinuous, it is advantageous to let  $\alpha$  be

positive and smaller than unity. Table 6.8 is analogous to Table 6.7; it displays computed results for a finer discretization of (6.4) with  $A \in \mathbb{R}^{1000 \times 1000}$ .  $\square$

Noise level %	$s$	Lavrentiev	Fractional Lavrentiev $\alpha^*$	
5	0.5	$8.83 \cdot 10^{-1}$	0	$8.83 \cdot 10^{-1}$
	0.6	$8.50 \cdot 10^{-1}$	0	$8.50 \cdot 10^{-1}$
	0.7	$8.42 \cdot 10^{-1}$	0.1	$8.35 \cdot 10^{-1}$
	0.8	$8.54 \cdot 10^{-1}$	0.1	$8.28 \cdot 10^{-1}$
	0.9	$9.26 \cdot 10^{-1}$	0.1	$8.25 \cdot 10^{-1}$
	1.0	$1.41 \cdot 10^1$	0.2	$8.12 \cdot 10^{-1}$
1	0.5	$8.37 \cdot 10^{-1}$	0	$8.37 \cdot 10^{-1}$
	0.6	$8.28 \cdot 10^{-1}$	0	$8.28 \cdot 10^{-1}$
	0.7	$8.26 \cdot 10^{-1}$	0.1	$8.25 \cdot 10^{-1}$
	0.8	$8.32 \cdot 10^{-1}$	0.1	$8.22 \cdot 10^{-1}$
	0.9	$8.82 \cdot 10^{-1}$	0.1	$8.16 \cdot 10^{-1}$
	1.0	$2.91 \cdot 10^0$	0.1	$7.79 \cdot 10^{-1}$
0.1	0.5	$8.25 \cdot 10^{-1}$	0	$8.25 \cdot 10^{-1}$
	0.6	$8.22 \cdot 10^{-1}$	0	$8.22 \cdot 10^{-1}$
	0.7	$8.20 \cdot 10^{-1}$	0	$8.20 \cdot 10^{-1}$
	0.8	$8.21 \cdot 10^{-1}$	0.1	$8.17 \cdot 10^{-1}$
	0.9	$8.64 \cdot 10^{-1}$	0.1	$8.14 \cdot 10^{-1}$
	1.0	$8.64 \cdot 10^{-1}$	0.1	$8.14 \cdot 10^{-1}$

**Table 6.8** Example 6.4: Discretization of the integral equation (6.4). Relative errors of the approximate solutions. Dimension  $n = 1000$ .

## 7 Conclusion

We have presented a fractional Lavrentiev method that uses a fractional power of the matrix  $A$  as regularization matrix. Computed examples show this method to perform better than the standard Lavrentiev method. The fractional Lavrentiev method performs best for the parameter  $s = 1$ ; the standard Lavrentiev method performs poorly for this choice of  $s$ . The value of the parameter  $\alpha^*$  is in all examples strictly smaller than unity. This shows that the fractional Lavrentiev method yields more accurate approximations of the desired solution  $\hat{\mathbf{x}}$  than standard Tikhonov regularization (4.2).

## References

1. M. L. Baart, The use of auto-correlation for pseudo-rank determination in noisy ill-conditioned linear least-squares problems, IMA J. Numer. Anal., 2 (1982), pp. 241–247.
2. D. Calvetti and L. Reichel, Tikhonov regularization with a solution constraint, SIAM J. Sci. Comput., 26 (2004), pp. 224–239.
3. H. W. Engl, M. Hanke, and A. Neubauer, Regularization of Inverse Problems, Kluwer, Dordrecht, 1996.

4. F. Girosi, M. Jones, and T. Poggio, Regularization theory and neural network architecture, *Neural Computing*, 7 (1995), pp. 219–269.
5. G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, 1996.
6. C. W. Groetsch, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Pitman, Boston, 1984.
7. C. W. Groetsch and J. Guacamene, Arcangeli's method for Fredholm equations of the first kind, *Proc. Amer. Math. Soc.*, 99 (1987), pp. 256–260.
8. P. C. Hansen, Regularization tools version 4.0 for Matlab 7.3, *Numer. Algorithms*, 46 (2007), pp. 189–194.
9. R. O. Hill, Jr., and B. N. Parlett, Refined interlacing properties, *SIAM J. Matrix Anal. Appl.*, 13 (1992), pp. 239–247.
10. M. E. Hochstenbach and L. Reichel, Fractional Tikhonov regularization for linear discrete ill-posed problems, *BIT*, 51 (2011), pp. 197–215.
11. M. E. Hochstenbach, N. McNinch, and L. Reichel, Discrete ill-posed least-squares problems with a solution norm constraint, *Linear Algebra Appl.*, 436 (2012), pp. 3801–3818.
12. E. Klann and R. Ramlau, Regularization by fractional filter methods and data smoothing, *Inverse Problems*, 24 (2008), 025018.
13. J. Lampe, M. Rojas, D. C. Sorensen, and H. Voss, Accelerating the LSTRS algorithm, *SIAM J. Sci. Comput.*, 33 (2011), pp. 175–194.
14. Z. P. Liang and P. C. Lauterbur, An efficient method for dynamic magnetic resonance imaging, *IEEE Trans. Med. Imag.*, 13 (1994), pp. 677–686.
15. S. Morigi, L. Reichel, and F. Sgallari, An iterative Lavrentiev regularization method, *BIT*, 46 (2006), pp. 589–606.
16. D. L. Phillips, A technique for the numerical solution of certain integral equations of the first kind, *J. ACM*, 9 (1962), pp. 84–97.
17. M. Rojas and D. C. Sorensen, A trust-region approach to regularization of large-scale discrete forms of ill-posed problems, *SIAM J. Sci. Comput.*, 23 (2002), pp. 1842–1860.
18. G. M. Wing, *A Primer on Integral Equations of the First Kind*, SIAM, Philadelphia, 1991.