Tikhonov-Phillips Regularization with Operator Dependent Seminorms

Huckle Thomas and Sedlacek Matous
Technische Universität München, Fakultät für Informatik

December 2, 2011

Abstract. We focus on the solution of discrete ill-posed problems to recover the original information from blurred signals in the presence of Gaussian white noise more accurately. We derive seminorms for the Tikhonov-Phillips regularization based on the underlying blur operator H. In this way it is possible to improve the reconstruction using spectral information of H. Reconstructions of numerous discrete ill-posed model problems, arising both from realistic applications and examples generated on our own, demonstrate the effect of the presented approach.

Keywords: Seminorms, Tikhonov-Phillips regularization, Discrete ill-posed inverse problems

1. Introduction

For discrete ill-posed problems, as they arise in signal or image reconstruction, regularization techniques are important in order to recover the original information. We consider the discrete linear model problem

$$x \xrightarrow{blur} Hx \xrightarrow{noise} Hx + \eta = b \tag{1}$$

where $x \in \mathbb{R}^n$ is the original signal or image, $H \in \mathbb{R}^{n \times n}$ is the blur operator, $\eta \in \mathbb{R}^n$ is a vector representing the unknown perturbations such as noise or measurement errors, and $b \in \mathbb{R}^n$ is the observed signal and image, respectively. Our aim is to recover x as good as possible. Because H may be extremely ill-conditioned or even singular, and, because of the presence of the noise η, the direct solution of (1) will result in a useless reconstruction dominated by noise. Consequently, to avoid this and solve (1) mainly on the signal subspace a regularization technique has to be applied.

Based on a decomposition of H, like, for instance, the QR factorization or the singular value decomposition (SVD) (Golub and Van Loan, 1996), direct regularization methods can be seen as a spectral filter acting on the singular spectrum, diminishing the deterioration of the solution by noise. Within this class we focus on the classical Tikhonov-Phillips regularization (Phillips, 1962; Tikhonov, 1963) which can often be improved by including minimization in a seminorm. Usually, the seminorm is related to a smoothing operator like, for instance, the

Laplacian. Here, we introduce seminorms based on the operator H itself. In this way we restrict the regularization to the noise subspace leaving the signal information unchanged.

The outline of the paper is the following: In Section 2 we will have a closer look on the Tikhonov-Phillips regularization including smoothing norms. Subsequently, in Section 3, we present our approach to improve the reconstruction of regularization methods and derive seminorms depending on H. Section 4 contains numerical results using the proposed approaches for several test scenarios. We mainly focus on discrete ill-posed problems from the package REGULARIZATION TOOLS (Hansen, 1994), but we consider artificial problems constructed on our own as well. A conclusion with a short outlook closes the discussion in Section 6.

2. Tikhonov-Phillips Regularization Including Smoothing-Norms

One of the classical regularization methods is the Tikhonov regularization (Tikhonov, 1963) which solves

$$\min_x \left\{ \|Hx - b\|_2^2 + \alpha^2 \|x\|_2^2 \right\} \Leftrightarrow (H^T H + \alpha^2 I)x = H^T b \qquad (2)$$

instead of (1), for a fixed regularization parameter $\alpha \geq 0$. The weight α has to be chosen such that both minimization criterions yield the minimal value together: the computed solution x is as close as possible to the original problem and sufficiently regular. Following (Hanke and Hansen, 1993; Hansen, 2010; Hansen and Jensen, 2006), instead of using the 2-norm as a means to control the error in the regularized solution, another possibility is to use discrete smoothing-norms of the form $\|Lx\|_2$ to obtain regularity. With L being a discrete approximation to a derivative operator, the *standard form* problem (2) can be reformulated as Tikhonov-Phillips regularization in *general form* via

$$\min_x \left\{ \|Hx - b\|_2^2 + \alpha^2 \|Lx\|_2^2 \right\} \Leftrightarrow (H^T H + \alpha^2 L^T L)x = H^T b. \qquad (3)$$

Usually, the matrix L is an approximation to the first or second derivative operator, i.e.,

$$L_1 := \begin{pmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{pmatrix} \quad \textit{or} \quad L_2 := \begin{pmatrix} 1 & -2 & 1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & & 1 & -2 & 1 \end{pmatrix} \qquad (4)$$

with $L_1 \in \mathbb{R}^{(n-1)\times n}$ and $L_2 \in \mathbb{R}^{(n-2)\times n}$, respectively, not taking any boundary conditions into account. Consequently, rough oscillations caused by noisy components can be suppressed during the reconstruction and the regularized approximations will satisfy inherent smoothness properties. Therefore, for problems where the exact signal x is smooth, the solution of the general formulation (3), using a differential operator, will be smoother and thus a more accurate reconstruction.

If (3) is to be solved via spectral decomposition, a *generalized SVD* (GSVD) (Golub and Van Loan, 1996) of the matrix pair (H, L) has to be applied. Using filter factors similarly to the TSVD and thus truncating the GSVD to suppress the noise contribution is denoted as *truncated GSVD* (TGSVD) (Hansen, 1994; Hansen, 1998; Hansen, 2010) in literature. Note that for using smoothing preconditioning with L in an iterative regularization method like, for instance, CGLS, a transformation to standard form has to be applied (Hansen, 2010).

3. Operator Dependent Seminorms

We introduce a new approach for regularization including a seminorm by deriving operators which directly include H and thus use spectral information of the given problem. The modified regularization method allows larger values of the regularization parameter while improving the reconstruction. A regularizing seminorm should have the property that it leads to a regularizing effect for the noise subspace relative to small eigenvalues, but behaves like zero on the signal subspace relative to large eigenvalues. In the limiting case a seminorm should behave like 1 on the noise part providing much emphasis on the penalty term while it should tend to 0 for the signal part having most emphasis on the model fit. Using the original operator H, we can define such operator dependent seminorms $\|L_{k,*}x\|_2$, where $K_{k,*} := L_{k,*}^T L_{k,*}$ is positive semi-definite, in the following three different ways.

3.1. Seminorm $\|L_{k,\sigma}x\|_2$

Let σ_{max} be the maximum singular value supposed to be available from $H \overset{SVD}{=} U\Sigma V^T$ or being an approximation resulting from a couple of Arnoldi iterations. Then we set for $k \in \{1, 2, 3, \ldots\}$

$$K_{k,\sigma} = L_{k,\sigma}^T L_{k,\sigma} := \left(I - \frac{H^T H}{\sigma_{max}^2}\right)^k \qquad with \qquad 0 \leq K_{k,\sigma} \leq 1.$$

$K_{k,\sigma}$ is a polynomial in $H^T H$ of the general form $p_k(x) = (1 - x)^k$ with $0 \leq x \leq 1$. Via the power k, it is possible to steer the "L"-shape of $p_k(x)$ which is illustrated in Figure 1.a. For small x, corresponding to singular vectors relative to noise, the function $p_k(x)$ tends to 1, for large k approaching the ordinate. For $x > 0$ and near 1, representing the signal subspace, $p_k(x)$ is close to zero with a zero of order k at $x = 1$. Including $K_{k,\sigma}$, we can modify the standard form Tikhonov regularization (2) to the general representation

$$\min_x \left\{ \|Hx - b\|_2^2 + \alpha^2 \|L_{k,\sigma} x\|_2^2 \right\} \Leftrightarrow \left[H^T H + \alpha^2 \left(I - \frac{H^T H}{\sigma_{max}^2} \right)^k \right] x = H^T b. \qquad (5)$$

Especially for $k = 1$, (5) can be written in the form

$$\left[H^T H + \alpha^2 \left(I - \frac{H^T H}{\sigma_{max}^2} \right) \right] x = \left[\left(1 - \frac{\alpha^2}{\sigma_{max}^2} \right) H^T H + \alpha^2 I \right] x = H^T b,$$

equivalent to a standard Tikhonov regularization for a slightly modified problem.

3.2. Seminorm $\|L_{k,r,\lambda} x\|_2$

For symmetric indefinite operators H, where the smallest and largest eigenvalue satisfy $\lambda_{min} \ll 0$ and $\lambda_{max} \gg 0$, respectively, it can be helpful to use a seminorm $\|L_{k,r,\lambda} x\|_2$ defined by the matrix

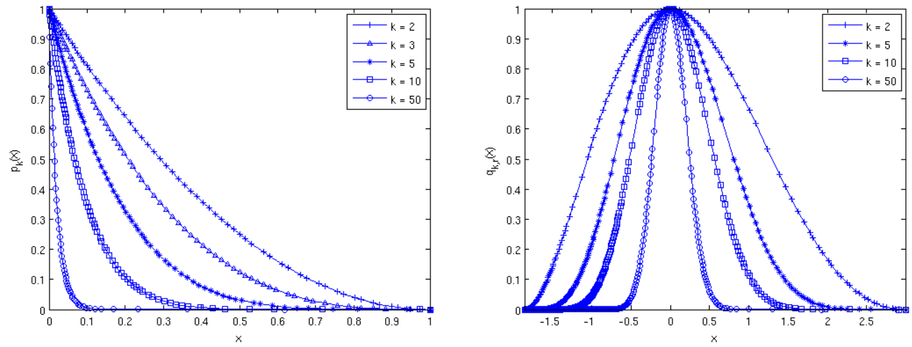
$$K_{k,r,\lambda} = L_{k,r,\lambda}^T L_{k,r,\lambda} := \left[\left(I - \frac{H}{\lambda_{min}} \right) \left(I - \frac{H}{\lambda_{max}} \right) \right]^k \left[I + \frac{k}{r} \left(\frac{1}{\lambda_{min}} + \frac{1}{\lambda_{max}} \right) H \right]^r.$$

$K_{k,r,\lambda}$ can be analyzed via the polynomial $q_{k,r}(x) =$

$$\left[\left(1 - \frac{x}{\lambda_{min}} \right) \left(1 - \frac{x}{\lambda_{max}} \right) \right]^k \left[1 + \frac{k}{r} \left(\frac{1}{\lambda_{min}} + \frac{1}{\lambda_{max}} \right) x \right]^r. \qquad (6)$$

Again k should be chosen out of $\{1, 2, 3, \ldots\}$ and $r \in \{0, 1, 2, 3, \ldots\}$. This definition makes sense only if the maximum or the minimum eigenvalue is not close to zero. The additional third term of $K_{k,r,\lambda}$, i.e., $([\cdot]^r)$, ensures that the regularization is effective near zero. Furthermore, there are conditions on k, r, λ_{min}, and λ_{max} in order to avoid negative or very large values, and to avoid additional maxima of $q_{k,r}(x)$ in the interval $[\lambda_{min}, \lambda_{max}]$. It turns out that the above requests can be satisfied with the condition

$$r := \lceil (s-1)k \rceil, \qquad \textit{with} \qquad s := \max \left(\left| \frac{\lambda_{min}}{\lambda_{max}} \right|, \left| \frac{\lambda_{max}}{\lambda_{min}} \right| \right). \qquad (7)$$



a. Polynomial $p_k(x) = (1-x)^k$. b. Polynomial $q_{k,r}(x)$ (6).

Figure 1. Polynomials for increasing values of k. Figure 1.b shows $q_{k,r}(x)$ for the model problem `shaw` from (Hansen, 1994) where all eigenvalues are located in $[\lambda_{min}, \lambda_{max}] = [-1.8567, 2.9933]$.

Figure 1.b shows $q_{k,r}(x)$ for increasing values of k for the spectrum of H resulting from the model problem `shaw` from (Hansen, 1994). The shape of $q_{k,r}(x)$ (6) slightly resembles a Gaussian bell curve. For increasing values of k it falls towards the abscissa tending to the δ-function.

3.3. Seminorm $\|L_{k,\tau}x\|_2$

A third seminorm can be defined for the general nonsymmetric case via

$$K_{k,\tau} := \left[\left(I - \frac{H^T}{\tau}\right)\left(I - \frac{H}{\tau}\right)\right]^k,$$

where we have to make sure that $K_{k,\tau} \geq 0$ in order to have a related seminorm $\|L_{k,\tau}x\|_2$ with $K_{k,\tau} = L_{k,\tau}^T L_{k,\tau}$. Furthermore, τ has to be chosen in such a way that for the signal subspace the regularization is turned off, while for the noise subspace it is turned on. Heuristically, we can model the signal and noise subspace by probing vectors, like, for example, $e_S = n^{-\frac{1}{2}}(1,1,...,1)^T$ and $e_N = n^{-\frac{1}{2}}(1,-1,1,-1,...)^T$, respectively, with τ chosen such that

$$\|K_{k,\tau}e_S\|_2 \ll 1 \qquad \textit{and} \qquad \|K_{k,\tau}e_N - e_N\|_2 \ll 1. \tag{8}$$

In other words, we suggest to choose τ such that the seminorm acts as 0 on the signal subspace and like the identity I on the noise subspace.

Compared to derivative operators L, like, for instance, L_1 or L_2 (4), $L_{k,\sigma}$, $L_{k,r,\lambda}$, and $L_{k,\tau}$ are not restricted to smooth right-hand sides. We provide a comparison between the introduced seminorms in Section 4.3. To obtain higher degrees in the polynomials it is possible to use the scaling and squaring in the matrix power. Thus, degrees of 2^k will only require $\log_2(2^k) = k$ multiplications.

4. Numerical Results

We focus on ill-posed problems from Regularization Tools (Hansen, 1994). Hence, we mainly examine Fredholm integral equations of the first kind. The Matlab (The MathWorks Inc., 2010) functions in (Hansen, 1994) provide discretizations H of the integral operators using Galerking or quadrature methods and scaled discrete approximations of the solution. We affect our right-hand sides with Gaussian white noise of different order, and we perform all computations on normalized values. Note that $\xi \in \mathbb{R}$ refers to the order of the white noise. Our direct regularization method is the Tikhonov-Phillips regularization.

Concerning the regularization parameter α we illustrate the maximum obtainable improvement for a perfect estimator. Therefore, we compute it via the Matlab function `fminbnd` which attempts to find a local minimizer in a given interval. We only consider reconstructions as reasonable where $\alpha \in [0, 10^3]$. Thus we fade out reconstructions for which `fminbnd` hits the interval boundary ($\alpha = 10^3$) and highlight them via †. We refine the search by setting the termination tolerance to $\texttt{TolX} = 10^{-9}$. Note that `fminbnd` is based on golden section search and parabolic interpolation and may provide only a local solution (The MathWorks Inc., 2010), i.e., the estimation may still not be optimal for certain problems. Hence, we measure the quality of the reconstructions via the relative reconstruction error (RRE) $\|\tilde{x} - x\|_2/\|x\|_2$, where $\tilde{x}$ denotes the reconstruction and x the exact signal. All our test problems have a fixed problem size $n = 300$. We emphasize reconstructions from standard form Tikhonov regularization with values marked in bold style.

4.1. The general case: test problems `baart` and `heat`

In our first examples we focus on the general nonsymmetric case and thus on the seminorms $\|L_{k,\sigma}x\|_2$ and $\|L_{k,\tau=\sigma_{max}}x\|_2$. For the test problem `baart` we obtained degradation with smoothing-norms using derivative operators according to (4) and therefore do not use them as a comparison in our results. In contrast to this, smoothing-norms yield highly improved reconstructions for the test problem `heat` which models the inverse heat equation. Here, we choose $\texttt{kappa} = 5$. In both examples we affect the right-hand side with white noise of order 10%, 5%, 1%, and 0.1%. Besides the results in the Tables II and III, we illustrate two additional informations for the two noise levels 1% and 5% in Figures 2 and 3. While fixing the polynomial degree at $k = 16$ we plot the regularization parameter α versus the RRE and show the associated reconstructions for the optimal regularization parameter. Note that any choice of α will provide an improved reconstruction using the seminorms, as their curves remain below the one corresponding to standard form Tikhonov-Phillips regularization.

Among most of all our experiments we obtain improved results using the seminorm $\|L_{k,\tau}x\|_2$ when the heuristic conditions (8) are satisfied. However, this is not a necessary condition. See for instance the `heat` example for $\xi = 5\%$, where improvement is gained for the selected measures according to Table I.

Table I. Regularization property for seminorm $\|L_{k,\tau=\sigma_{max}}x\|_2$ according to (8) for test problem `heat` with $\xi = 5\%$.

Norm \ Degree	$k = 4$	$k = 8$	$k = 16$
$\|K_{k,\tau=\sigma_{max}}e_S\|_2$	0.0257	0.0224	0.0294
$\|K_{k,\tau\sigma_{max}}e_N - e_N\|_2$	0.2762	0.6284	1.6517

4.2. The symmetric indefinite case: test problem `shaw`

We illustrate the behaviour of the seminorms $\|L_{k,r,\lambda}x\|_2$ and $\|L_{k,\sigma}x\|_2$ for the test problem `shaw` where the spectrum of H satisfies the necessary conditions $\lambda_{min} = -1.86 \ll 0$ and $\lambda_{max} = 2.99 \gg 0$. See the characteristics of the polynomial $q_{k,r}(x)$ for this problem in Figure 1.b. The scenario setting is the same as in Section 4.1 except that we fix the polynomial degree at $k = 32$ in Figure 4. Thus, we obtain better distinguishable plots. For large noise levels the seminorm $\|L_{k,r,\lambda}x\|_2$ yields improved results while it does not degrade the solution for small noise levels. In comparison, the seminorm $\|L_{k,\sigma}x\|_2$ yields similar behaviour with slightly weaker improvement.

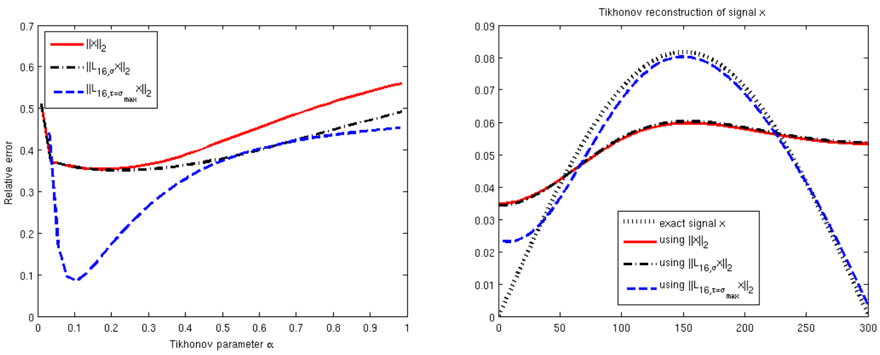
4.3. Further problems from Regularization Tools

Finally, we provide results on further examples from Regularization Tools to give an overall summary. For some model problems we illustrate detailed results for $\|L_{k,\sigma}x\|_2$, $\|L_{k,\tau}x\|_2$ and the noise levels 10%, 5%, and 1% in the Table V. For model problems with smooth right-hand side we additionally provide results for the smoothing-norms (4).

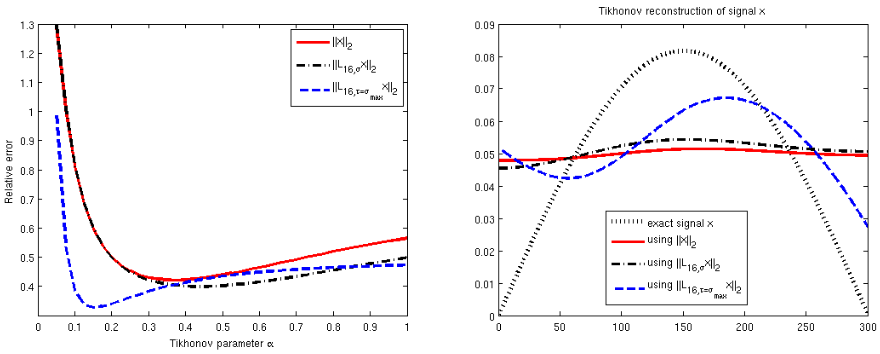
In our last experiment, we are interested in the behaviour using different right-hand sides for each model problem from Table VI. Be-

Table II. RRE using Tikhonov regularization with optimal α to compute the reconstruction $\tilde{x}$ of test problem baart.

Noise / Norm	$\xi = 10\%$	$\xi = 5\%$	$\xi = 1\%$	$\xi = 0.1\%$
$\|\|Ix\|\|_2$	**0.4638**	**0.4211**	**0.3535**	**0.1275**
$\|\|L_{1,\sigma}x\|\|_2$	0.4631	0.4201	0.3534	0.1275
$\|\|L_{4,\sigma}x\|\|_2$	0.4562	0.4153	0.3528	0.1275
$\|\|L_{8,\sigma}x\|\|_2$	0.4476	0.4095	0.3522	0.1275
$\|\|L_{16,\sigma}x\|\|_2$	0.4320	0.3999	0.3510	0.1275
$\|\|L_{32,\sigma}x\|\|_2$	0.4071	0.3826	0.3491	0.1275
$\|\|L_{64,\sigma}x\|\|_2$	0.3764	0.3629	0.3469	0.1275
$\|\|L_{1,\tau=\sigma_{max}}x\|\|_2$	0.4890	0.4839	0.3825	0.1401
$\|\|L_{4,\tau=\sigma_{max}}x\|\|_2$	0.4887	0.4887	0.3119	0.1186
$\|\|L_{8,\tau=\sigma_{max}}x\|\|_2$	0.4928	0.3744	0.1269	0.0639
$\|\|L_{16,\tau=\sigma_{max}}x\|\|_2$	0.4928	0.3272	0.0837	0.0631
$\|\|L_{32,\tau=\sigma_{max}}x\|\|_2$	0.4928	0.4895	0.0595	0.0646
$\|\|L_{64,\tau=\sigma_{max}}x\|\|_2$	†	†	†	†



a. RRE depending on α for $\xi = 1\%$. b. Optimal reconstructions for $\xi = 1\%$.

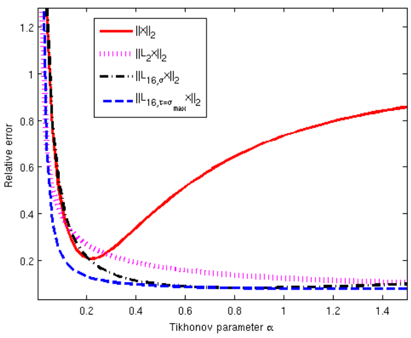


c. RRE depending on α for $\xi = 5\%$. d. Optimal reconstructions for $\xi = 5\%$.

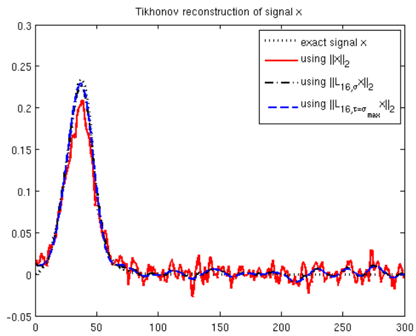
Figure 2. Reconstruction of the test problem baart using the different seminorms of degree $k = 16$. a,c illustrate the RRE depending on α for the noise levels $\xi = 1\%$ and $\xi = 5\%$. b,d show the corresponding optimally reconstructed signals.

Table III. RRE using Tikhonov regularization with optimal α to compute the reconstruction $\tilde{x}$ of test problem `heat`.

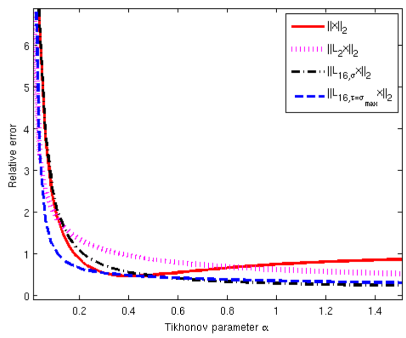
Noise / Norm	$\xi = 10\%$	$\xi = 5\%$	$\xi = 1\%$	$\xi = 0.1\%$
$\|Ix\|_2$	**0.6041**	**0.4554**	**0.2042**	**0.0540**
$\|L_1 x\|_2$	0.3789	0.2416	0.0869	0.0197
$\|L_2 x\|_2$	0.3545	0.2138	0.0771	0.0172
$\|L_{1,\sigma} x\|_2$	0.5141	0.3847	0.1726	0.0462
$\|L_{4,\sigma} x\|_2$	0.4261	0.2986	0.1270	0.0343
$\|L_{8,\sigma} x\|_2$	0.3789	0.2456	0.0986	0.0269
$\|L_{16,\sigma} x\|_2$	0.3373	0.2080	0.0812	0.0211
$\|L_{32,\sigma} x\|_2$	0.3207	0.2013	0.0785	0.0179
$\|L_{1,\tau=\sigma_{max}} x\|_2$	0.4661	0.3386	0.1483	0.0400
$\|L_{4,\tau=\sigma_{max}} x\|_2$	0.3671	0.2358	0.0923	0.0251
$\|L_{8,\tau=\sigma_{max}} x\|_2$	0.3388	0.2083	0.0780	0.0197
$\|L_{16,\tau=\sigma_{max}} x\|_2$	0.3266	0.2100	0.0769	0.0172
$\|L_{32,\tau=\sigma_{max}} x\|_2$	†	†	0.0789	0.0171



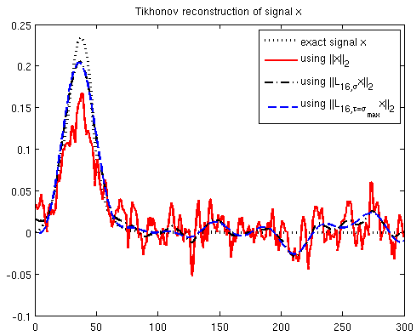
a. RRE depending on α for $\xi = 1\%$.



b. Optimal reconstructions for $\xi = 1\%$.



c. RRE depending on α for $\xi = 5\%$.

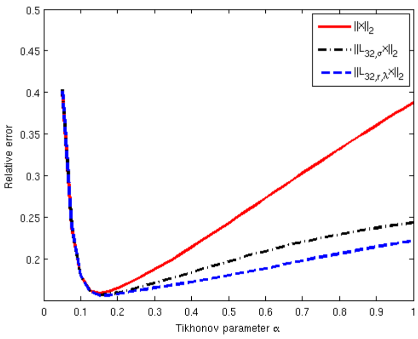


d. Optimal reconstructions for $\xi = 5\%$.

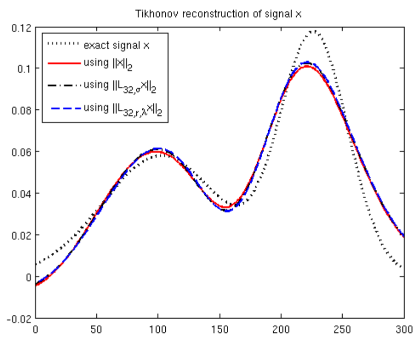
Figure 3. Reconstruction of the test problem `heat` using the different seminorms of degree $k = 16$. a,c illustrate the RRE depending on α for the noise levels $\xi = 1\%$ and $\xi = 5\%$. b,d show the corresponding optimally reconstructed signals.

Table IV. RRE using Tikhonov regularization with optimal α to compute the reconstruction $\tilde{x}$ of test problem `shaw`.

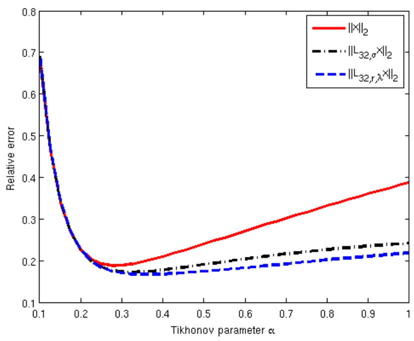
Norm \ Noise	$\xi = 10\%$	$\xi = 5\%$	$\xi = 1\%$	$\xi = 0.1\%$
$\|Ix\|_2$	**0.2195**	**0.1878**	**0.1584**	**0.0990**
$\|L_{1,\sigma}x\|_2$	0.2182	0.1868	0.1582	0.0990
$\|L_{4,\sigma}x\|_2$	0.2125	0.1836	0.1578	0.0990
$\|L_{8,\sigma}x\|_2$	0.2064	0.1805	0.1574	0.0990
$\|L_{16,\sigma}x\|_2$	0.1989	0.1766	0.1569	0.0990
$\|L_{32,\sigma}x\|_2$	†	0.1721	0.1564	0.0990
$\|L_{64,\sigma}x\|_2$	0.1829	0.1672	0.1558	0.0990
$\|L_{1,r,\lambda}x\|_2$	0.2180	0.1862	0.1581	0.0990
$\|L_{4,r,\lambda}x\|_2$	0.2069	0.1806	0.1574	0.0990
$\|L_{8,r,\lambda}x\|_2$	0.1985	0.1762	0.1569	0.0990
$\|L_{16,r,\lambda}x\|_2$	0.1900	0.1714	0.1563	0.0990
$\|L_{32,r,\lambda}x\|_2$	0.1817	0.1665	0.1558	0.0990
$\|L_{64,r,\lambda}x\|_2$	0.1793	0.1652	0.1556	0.0990



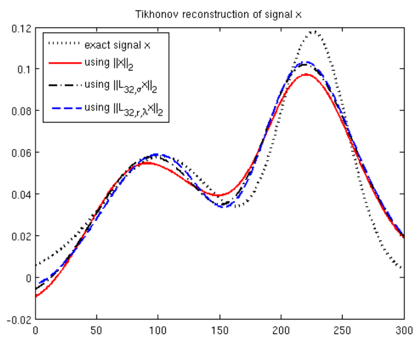
a. RRE depending on α for $\xi = 1\%$.



b. Optimal reconstructions for $\xi = 1\%$.



c. RRE depending on α for $\xi = 5\%$.



d. Optimal reconstructions for $\xi = 5\%$.

Figure 4. Reconstruction of the test problem `shaw` using the different seminorms of degree $k = 32$. a,c illustrate the RRE depending on α for the noise levels $\xi = 1\%$ and $\xi = 5\%$. b,d show the corresponding optimally reconstructed signals.

Table V. RRE of problems from (Hansen, 1994) using Tikhonov regularization.

Problem	Noise / Norm	$\xi = 10\%$	$\xi = 5\%$	$\xi = 1\%$
gravity(n,3,0,1,$\frac{1}{4}$)	$\|\|Ix\|\|_2$	**0.2243**	**0.1832**	**0.1185**
	$\|\|L_{1,\sigma}x\|\|_2$	0.2074	0.1735	0.1179
	$\|\|L_{4,\sigma}x\|\|_2$	0.1863	0.1611	0.1173
	$\|\|L_{32,\sigma}x\|\|_2$	0.1419	0.1342	0.1193
	$\|\|L_{1,\tau=\sigma_{max}}x\|\|_2$	0.1845	0.1601	0.1179
	$\|\|L_{4,\tau=\sigma_{max}}x\|\|_2$	0.1512	0.1408	0.1201
	$\|\|L_{8,\tau=\sigma_{max}}x\|\|_2$	0.1414	0.1340	†
	$\|\|L_{16,\tau=\sigma_{max}}x\|\|_2$	†	0.1525	0.1219
foxgood(n)	$\|\|Ix\|\|_2$	**0.1097**	**0.0727**	**0.0276**
	$\|\|L_{1,\sigma}x\|\|_2$	0.1141	0.0744	0.0276
	$\|\|L_{4,\sigma}x\|\|_2$	0.1123	0.0728	0.0268
	$\|\|L_{16,\sigma}x\|\|_2$	0.1057	0.0668	0.0237
	$\|\|L_{64,\sigma}x\|\|_2$	0.0867	0.0488	0.0145
	$\|\|L_{1,\tau=\sigma_{max}}x\|\|_2$	0.1242	0.0834	0.0324
	$\|\|L_{8,\tau=\sigma_{max}}x\|\|_2$	0.2117	0.1640	0.0792
	$\|\|L_{32,\tau=\sigma_{max}}x\|\|_2$	0.3062	0.2988	0.2964
deriv2(n,3)	$\|\|Ix\|\|_2$	**0.1882**	**0.1483**	**0.0983**
	$\|\|L_{1,\sigma}x\|\|_2$	0.1226	0.1209	0.1203
	$\|\|L_{4,\sigma}x\|\|_2$	0.1217	0.1179	0.0951
	$\|\|L_{8,\sigma}x\|\|_2$	0.1218	0.1182	0.0943
	$\|\|L_{32,\sigma}x\|\|_2$	0.1225	0.1205	0.0897
	$\|\|L_{1,\tau=\sigma_{max}}x\|\|_2$	0.2758	0.2059	0.1135
	$\|\|L_{4,\tau=\sigma_{max}}x\|\|_2$	0.8962	0.7636	0.3852
	$\|\|L_{8,\tau=\sigma_{max}}x\|\|_2$	†	1.0000	0.9977
i_laplace(n,2)	$\|\|Ix\|\|_2$	**0.8297**	**0.8206**	**0.8116**
	$\|\|L_1x\|\|_2$	0.1771	0.1461	0.1248
	$\|\|L_2x\|\|_2$	0.1491	0.1163	0.0954
	$\|\|L_{1,\sigma}x\|\|_2$	0.8306	0.8208	0.8116
	$\|\|L_{4,\sigma}x\|\|_2$	0.8306	0.8207	0.8116
	$\|\|L_{32,\sigma}x\|\|_2$	0.8308	0.8200	0.8114
	$\|\|L_{1,\tau=\sigma_{max}}x\|\|_2$	0.8976	0.8882	0.8820
	$\|\|L_{4,\tau=\sigma_{max}}x\|\|_2$	0.9222	0.9222	0.9222
	$\|\|L_{32,\tau=\sigma_{max}}x\|\|_2$	0.9900	0.9898	0.9915
phillips(n)	$\|\|Ix\|\|_2$	**0.2610**	**0.1787**	**0.0811**
	$\|\|L_{1,\sigma}x\|\|_2$	0.2221	0.1566	0.0785
	$\|\|L_{4,\sigma}x\|\|_2$	0.1918	0.1399	0.0781
	$\|\|L_{32,\sigma}x\|\|_2$	0.1954	0.1358	0.1097
	$\|\|L_{1,\tau=\lambda_{max}}x\|\|_2$	0.1955	0.1417	0.0770
	$\|\|L_{4,\tau=\lambda_{max}}x\|\|_2$	0.1931	0.1358	0.0740
	$\|\|L_{32,\tau=\lambda_{max}}x\|\|_2$	†	0.2333	0.0533
wing(n,$\frac{1}{3}$,$\frac{2}{3}$)	$\|\|Ix\|\|_2$	**0.6165**	**0.6107**	**0.6048**
	$\|\|L_{1,\sigma}x\|\|_2$	0.6165	0.6107	0.6048
	$\|\|L_{4,\sigma}x\|\|_2$	0.6163	0.6106	0.6048
	$\|\|L_{32,\sigma}x\|\|_2$	0.6143	0.6094	0.6045
	$\|\|L_{4,\tau=\lambda_{max}}x\|\|_2$	0.6637	0.6485	0.6333
	$\|\|L_{16,\tau=\lambda_{max}}x\|\|_2$	0.5956	0.5838	0.5739
	$\|\|L_{32,\tau=\lambda_{max}}x\|\|_2$	0.8388	0.8343	0.8312

Table VI. Test problems from (Hansen, 1994) used in Figure 5.

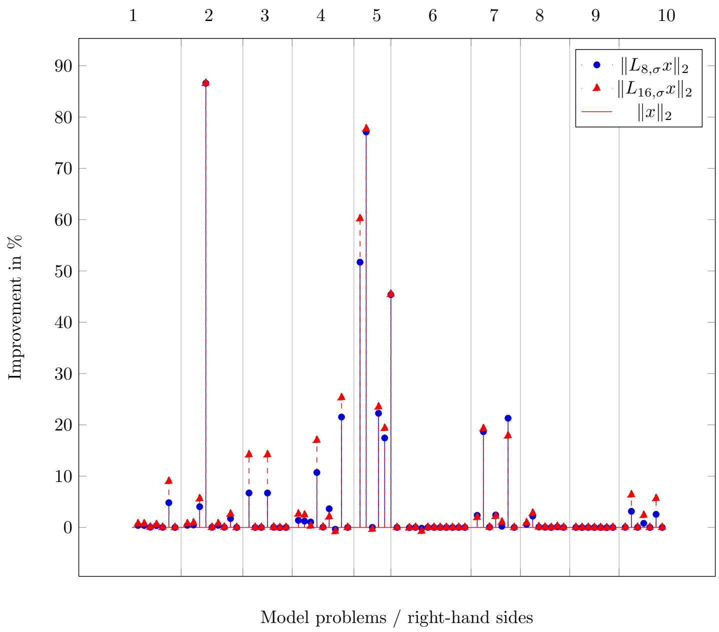
No.	Problem	No.	Problem
1	`baart(n)`	6	`i_laplace(n,[1,2,3,4])`
2	`deriv2(n,[1,2,3])`	7	`phillips(n)`
3	`foxgood(n)`	8	`shaw(n)`
4	`gravity(n,[1,2],0,1,0.5)`	9	`spikes(n,2)`
	`gravity(n,3,0,1,0.25)`	10	`wing(n,1/3,2/3)`
5	`heat(n,5)`		

sides those available from (Hansen, 1994) (x_1) we additionally use the following right-hand sides:

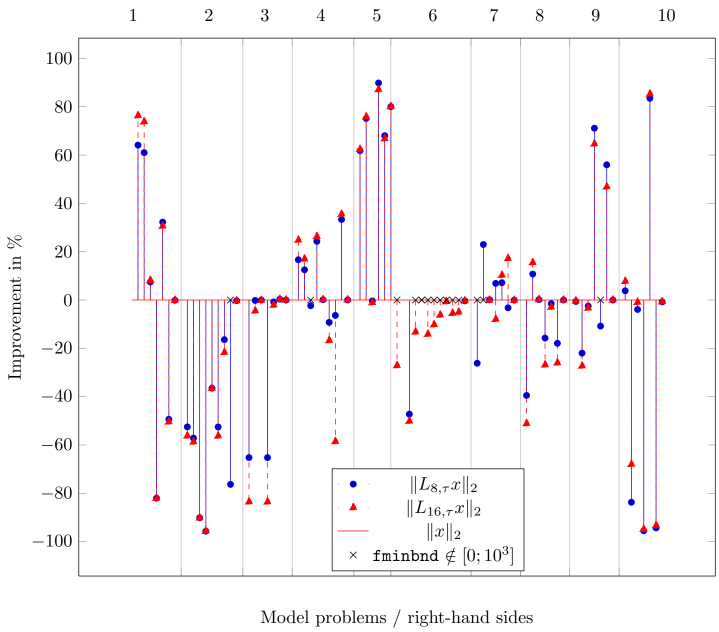
$$\begin{aligned}
(x_2)_i &:= \sin\left(\frac{i\pi}{n}\right), \\
(x_3)_i &:= \sin\left(\frac{i\pi}{n}\right),\ (x_3)_{i=100,\ i=200,\ i=300} = 10\cdot\sin\left(\frac{i\pi}{n}\right), \\
(x_4)_i &:= i, \\
(x_5)_i &:= \frac{\left(i-\lfloor\frac{n}{2}\rfloor\right)^2}{\left(\lceil\frac{n}{2}\rceil\right)^2}, \\
(x_6)_i &:= \frac{i}{n}+\sin\left(\frac{\pi(i-1)}{n}\right), \\
(x_7)_i &:= (0,\ldots,0)^T,\ (x_7)_{i=100,\ i=200,\ i=300} = 5,
\end{aligned}$$

where $i = 1,\ldots,n$. We plot the improvement obtained from the two seminorms $\|L_{k,\sigma}x\|_2$ and $\|L_{k,\tau}x\|_2$ for $\xi = 1\%$ in Figure 5. For $\xi = 10\%$ we observed similar behaviour for both seminorms. As REGULARIZATION TOOLS provides several right-hand sides for some model problems, like, e.g., `deriv2` or `gravity` (see Table VI), the overall number varies in the plots. We choose $\tau = \sigma_{max}$ and $\tau = \lambda_{max}$ but only plot the better result of both, even when the conditions (8) are violated. We summarize the results seminorm dependently:

- $\|L_{k,\sigma}x\|_2$: this seminorm provides the most robust behaviour, i.e., yields improvement for the most examples and operator spectra while it does not destroy the reconstruction in the remaining problems. Nevertheless, the improvement might be small for certain test problems or smaller compared to the other seminorms, like, for example, for the `baart` or the `shaw` scenario. Note that even for the indefinite spectrum of the latter problem we get an improvement compared to Tikhonov regularization in standard form.



a. Seminorm $\|L_{k,\sigma}x\|_2$.



b. Seminorm $\|L_{k,\tau}x\|_2$.

Figure 5. Percental improvement using the seminorms $\|L_{k,\sigma}x\|_2$ and $\|L_{k,\tau}x\|_2$ for the polynomial degrees $k = 8$, $k = 16$ and the noise level $\xi = 1\%$. Refer to Table VI for the used model problems. The improvement is measured with reference to standard form Tikhonov-Phillips regularization.

- $\|L_{k,r,\lambda}x\|_2$: in case that H is symmetric, the more the indefinite eigenspectrum is separated from zero, and the better $-\lambda_{min} \approx \lambda_{max}$, the higher the improvement should be when applying this seminorm. See the `shaw` test problem in Section 4.2, where the usage of $\|L_{k,r,\lambda}x\|_2$ leads to best results. As we always choose r according to (7) the operator $K_{k,r,\lambda}$ is well-defined. On the other hand, the usage of $\|L_{k,r,\lambda}x\|_2$ will eventually destroy the reconstruction if $\lambda_{min} \approx 0$ or $\lambda_{max} \approx 0$.

- $\|L_{k,\tau}x\|_2$: for certain general operators, this seminorm may yield better improvements than $\|L_{k,\sigma}x\|_2$. However, one has to choose τ with caution. For nonsymmetric H, τ has to be chosen such that $K_{k,\tau} \geq 0$ and that it has a regularizing effect only on the noise subspace. For example, we achieved best results using $K_{k,\tau=\sigma_{max}}$ for the problems `baart` and `heat`. Here, $K_{k,\tau=\sigma_{max}} \geq 0$. For `heat` both the conditions $\|K_{k,\tau=\sigma_{max}}e_S\|_2 \ll 1$ and $\|K_{k,\tau=\sigma_{max}}e_N - e_N\|_2 \ll 1$ are often satisfied. Though, for `baart` these conditions are not satisfied for larger polynomial degress, we still observe much improvement. This reflects the heuristic choice of modeling the subspaces by e_S and e_N. For symmetric H, like, for example, in the model problems `phillips` or `gravity 3`, $K_{k,\tau=\lambda_{max}}$ is well-defined and yields improved results too.

- Smoothing norm $\|Lx\|_2$: for model problems containing a smooth right-hand side this well-known seminorm provides highly improved reconstructions. E.g., while reconstructing the `i_laplace(n,2)` or `i_laplace(n,4)` model problems, none of the spectral dependent seminorms provide improved results. Here, a smoothing-norm is very effective.

5. Conclusion

We considered the impact of incorporating spectral information of H and therewith use operator dependent seminorms in the Tikhonov-Phillips regularization. Depending on the definiteness and the location of the spectral values, an appropriate seminorm, corresponding to a polynomial with certain properties, will have a regularizing effect on the noise subspace but no action on the signal subspace. The improvement of the solution can be enhanced by increasing values of the polynomial degree. Here, the computation should be based on the scaling and squaring according to the matrix power to keep it efficient. As a brief summary we point out:

- The seminorm $\|L_{k,\sigma}x\|_2$ yields robust behaviour meaning that it will not destroy the solution if not appropriate for an underlying problem.
- For indefinite operators satisfying $\lambda_{min} \ll 0$ and $\lambda_{max} \gg 0$ the seminorm $\|L_{k,r,\lambda}x\|_2$ is an alternative to obtain improved reconstructions.
- If τ is chosen carefully to satisfy the mentioned conditions, $\|L_{k,\tau}x\|_2$ will produce improved results for certain problems as well.
- If the signal is known to be smooth, smoothing norms will yield distinct improvement.

Among several of our test problems we observed that the larger the noise level, the more the improvement will be when applying the seminorms. Note that operator dependent seminorms rely on the assumption that a priori knowledge of the extremal spectral values is available. If not, however, approximations can be obtained via a couple of Arnoldi iterations.

References

G. H. Golub and C. F. Van Loan. Matrix Computations. The Johns Hopkins University Press, 3rd ed., 1996.

M. Hanke and P. C. Hansen. Regularization Methods for Large-Scale Problems. *Surv. Math. Ind.*, 3:253–315, 1993.

P. C. Hansen. Regularization Tools: A Matlab Package for Analysis and Solution of Discrete Ill-Posed Problems. *Numer. Algorithms*, 6(1–2):1–35, 1994.

P. C. Hansen. Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1998.

P. C. Hansen. Discrete Inverse Problems: Insight and Algorithms. SIAM, 1st ed., 2010.

P. C. Hansen and T. K. Jensen. Smoothing-Norm Preconditioning for Regularizing Minimum-Residual Methods. *SIAM J. Matrix Anal. Appl.*, 29:1–14, 2006.

Matlab. version 7.11.0 (R2010b). The MathWorks Inc., Natick, Massachusetts, 2010.

D. L. Phillips. A Technique for the Numerical Solution of Certain Integral Equations of the First Kind. *J. Assoc. Comput. Mach.*, ACM, 9:84–97, 1962.

A. N. Tikhonov. Solution of Incorrectly Formulated Problems and the Regularization method. *Sov. Math. Dokl.*, 4:1035–1038, 1963, English translation of *Dokl. Akad. Nauk. SSSR* 151:501–504, 1963.