Fall 2015:
Computational and Variational Methods for Inverse Problems

Georg Stadler
Courant Institute of Mathematical Sciences
New York University

Omar Ghattas
Jackson School of Geosciences
Department of Mechanical Engineering
Institute for Computational Engineering & Sciences
The University of Texas at Austin

CHAPTER 1

Introduction

The goal of these notes is to provide an overview of basic analysis, regularization and solution methods for inverse problem, with an emphasis on inverse problems that involve (partial) differential equations. As general references for inverse problem, and also as sources for these notes we refer to [1–3].

This section introduces basic definitions and naming conventions used throughout these notes. Based on an image deblurring problem, typical features of inverse problems are illustrated. While the deblurring problem does not involve a differential equation, it shares many features with problems in which the parameters and the measurements are linked through the solution of a partial differential equation. This makes it an illustrative introductionary example.

Let us introduce the notation used in these notes. We denote vectors with bold letters, e.g., $\boldsymbol{u} \in \mathbb{R}^n$. Components of these vectors are denoted using indices, i.e., $\boldsymbol{u} = (u_1, \dots, u_n)$ and vectors are understood as column vectors. The inner product between two vectors $\boldsymbol{u}, \boldsymbol{v}$ is defined by $\boldsymbol{u}^T\boldsymbol{v}$, and the norm of $\boldsymbol{u}$ by $\|\boldsymbol{u}\| := \sqrt{\boldsymbol{u}^T\boldsymbol{u}}$. Matrices are denoted using bold capital letters such as $\boldsymbol{A} \in \mathbb{R}^{n\times m}$. Real-valued functions and variables not further specified are denoted by lower case Latin letters such as g or h, and real scalars are usually denoted by Greek letters such as α or β.

1. Ill-posed problems

The basic setup of an inverse problem can be explained using the relation

$$d = F(p) + n, \tag{1}$$

where

- The variable p, which we want to reconstruct in the inverse problem is called the *parameter (field)*, the *image* or the *model*. The set of all parameters is called parameter (or image) space. In many of the applications we are interested in, p is a function or, after discretization, a vector in a high-dimensional space.
- The *forward (or direct) model* F maps p to the quantity we are able to measure. This forward mapping can be linear or nonlinear in p. Usually, F describes a physical theory, such as the propagation of waves, the diffusion of a substance, the absorption of rays when passing through an object, or fluid flow. In the cases we are

mainly interested in, F is given by an ordinary or a partial differential equation, in which case F often also contains an observation operator that restricts the solution of the differential equation to the quantity that can be measured, e.g., values at a part of the domain, at boundaries or points.

- The variable d denotes the *data* or the *measurements* we are able to make. Often, these measurements are a corrupted version of the outputs of the forward model. This error can be due to a model that does not fully describe the physical phenomenon, it can be intrinsic in the measurement process, or it can be due to roundoff error from a computer representation of the measurements. Measurement errors, denoted by n in (1) are usually not known, but statistical properties of n (such as the mean and the variance) are often available.

The *inverse problem* is to find the parameters p given the (noisy) measurements d, having knowledge over the forward operator F. A main difficulty is that in many applications of interest, inverse problems are not *well-posed* in the sense of Hadamard, who defined the inverse problem of solving

$$d = F(p)$$

to be well-posed, if the following properties are satisfied:

(1) *Existence:* For all data d (in an appropriate data space), there exists a parameter p of the problem (in an appropriate parameter space).
(2) *Uniqueness:* For all (suitable) data d, the solution p is unique.
(3) *Stability:* The solution depends continuously on the data, i.e., small changes in the data d result in small changes in the parameter p.

The problem (1) is called *ill-posed* if at least one of the above conditions is not satisfied. The main challenge in the numerical solution of inverse problems is the stability condition. If one wants to approximate a problem whose solution does not depend continuously on the data by a numerical method as one does for well-posed problems, one has to expect the method to become unstable.

There are two conceptually very different approaches to solve ill-conditioned inverse problems:

- *Deterministic inversion* is usually based on regularization methods that help to overcome the difficulties due to the ill-posedness of inverse problems. These methods usually find a single parameter or image p, which solves (1) in an appropriate sense. In this class, we will mainly focus on this deterministic approach for the solution of inverse problems, discuss regularization methods, their influence on the reconstruction and numerical solution algorithms.
- *Bayesian inversion methods* compute a probability density for the parameter p rather than a single solution. This approach allows a

flexible integration of prior knowledge about p into the solution (i.e., the probability density function). Such a probabilistic approach is often preferable in practical problems since its solution also quantifies the uncertainties in the reconstruction. However, Bayesian inversion is very costly and sometimes infeasible, in particular for the large-dimensional problems, which arise as discretizations of inverse problems with partial differential equations.

In simple situations, connections between the Bayesian and the deterministic approach can be made. For instance, the choice of the prior in the probabilistic approach is closely related to regularization methods in the deterministic approach.

2. A deblurring problem

Let us consider a deblurring (or deconvolution) problem as illustrative example. Even though in this problem the parameters and the data are not connected through a differential equation, it shares several features with more complicated inverse problems that involve differential equations. For simplicity, we consider a one-dimensional blurring operator given by a Fredholm first kind integral equation. For a function $p : [0, 1] \to \mathbb{R}$, we consider the operator

$$F(p) = d : [0, 1] \to \mathbb{R} \tag{2}$$

defined by

$$d(x) = \int_0^1 k(x - x')p(x')\,dx' \quad \text{for } 0 \le x \le 1.$$

Here, the kernel $k(x)$ is given by

$$k(x) = C \exp(\frac{-x^2}{2\gamma^2}) \quad \text{with } C, \gamma > 0. \tag{3}$$

The forward problem is the following: Given the source function p and the kernel k, determine the blurred image d. The associated inverse problem is: Given the kernel k and the blurred image d, determine the original image p. To illustrate the ill-posedness of this inverse problem, consider a perturbation $\delta p(x) := \varepsilon \sin(2\pi\omega x)$ for p, where $\varepsilon > 0$ and $\omega = 1, 2, \ldots$. The corresponding perturbation for $F(p + \delta p)$ is

$$\delta d(x) = \varepsilon \int_0^1 k(x - x') \sin(2\pi\omega x')\,dx',$$

which converges to zero as $\omega \to \infty$[1]. Hence, the ratio between δp and δd[2] can become arbitrary large, which shows that the stability requirement for well-posedness cannot be satisfied.

Figure 2 illustrates the effect of the convolution operator. While the parameter function p contains jumps, the convolved data $\boldsymbol{Kp}$ is a smoothly

[1] This is a consequence of the Riemann Lebesgue lemma.

[2] The norms for δp and δd can either be the L^∞ or the L^2-norm on $(0, 1)$.

varying function. This smoothing effect of the convolution operator is particularly obvious in the interval $[0, 1/4]$, since the small wave length variations in p are averaged out in the convolved data. The amount of averaging depends on the width of the Gaussian, which is controlled by the value γ in (3).

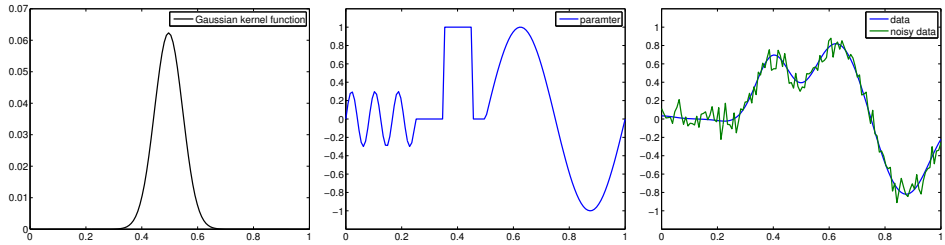

FIGURE 1. Gaussian kernel $k(x - 0.5)$ with $\gamma = 0.05, C = 1/(\gamma\sqrt{2\pi})$ as defined in (3) (left plot). A parameter function p (middle plot) and its convolution d (right plot). Shown are the exact convolution data in blue, and the noisy data in green. The discretization of the convolution operator uses $N = 128$ unknowns.

Next, we discretize the integral in (2) as needed for numerical computations. For that purpose we use split $[0, 1]$ into N intervals $[kh, (k + 1)h]$, $k = 0, N - 1$, where $h = 1/N$ and N is a large integer. Using midpoint quadrature, the discrete version of (2) becomes

$$\boldsymbol{d} = \boldsymbol{K}\boldsymbol{p} \tag{4}$$

with the vectors $\boldsymbol{d}$, $\boldsymbol{p}$ corresponding to the function values of d, p at the midpoints of the intervals, and a matrix $\boldsymbol{K} \in \mathbb{R}^{N\times N}$. The entries of $\boldsymbol{K}$ are given by

$$\boldsymbol{K}_{ij} = h\, C \exp\left(-\frac{((i-j)h)^2}{2\gamma^2}\right), \quad 1 \le i, j \le N.$$

Since the matrix $\boldsymbol{K}$ is invertible, for given (discrete) data $\boldsymbol{d}$ one may simply compute the image $\boldsymbol{p}$ as $\boldsymbol{p} = \boldsymbol{K}^{-1}\boldsymbol{d}$. However, $\boldsymbol{K}$ becomes increasingly ill-conditioned as N becomes large and small noise in $\boldsymbol{d}$ can result in large errors in $\boldsymbol{p}$, which is a consequence of the ill-posedness of the deconvolution problem[3].

2.1. SVD-based filtering. To analyze properties of the system (4) we include a noise vector $\boldsymbol{n}$, i.e., we consider

$$\boldsymbol{d} = \boldsymbol{K}\boldsymbol{p}_{\boldsymbol{true}} + \boldsymbol{n}. \tag{5}$$

[3]To be precise, the discrete system (4) is stable, but the stability constant grows as N becomes larger and (4) becomes a better approximation to the unstable continuous problem (2). The ill-posedness of this linear inverse problem is closely related to the fact that the matrix $\boldsymbol{K}$ is ill-conditioned.

Here, $\boldsymbol{p_{true}}$ are the true parameters we are trying to reconstruct from the data $\boldsymbol{d}$. The below analysis is based on the singular value decomposition (SVD) of the matrix $\boldsymbol{K}$. The SVD decomposition also allows to stabilize the inverse problem by filtering. While this regularization approach is illustrative and works well for moderate size inverse problems, it cannot be applied for large-scale inverse problems, where computing an explicit SVD decomposition is infeasible.

2.1.1. *Singular value decomposition.* For a real-valued matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, there exist orthogonal matrices

$$\boldsymbol{U} = [\boldsymbol{u}_1, \dots, \boldsymbol{u}_m] \in \mathbb{R}^{m\times m} \text{ and } \boldsymbol{V} = [\boldsymbol{v}_1, \dots, \boldsymbol{v}_n] \in \mathbb{R}^{n\times n} \tag{6}$$

such that

$$\boldsymbol{A} = \boldsymbol{U}\text{diag}(\sigma_1, \dots, \sigma_p)\boldsymbol{V}^T, \text{ with } p = \min\{m, n\}, \tag{7}$$

with $\sigma_1 \geq \sigma_2 \geq \dots \sigma_p \geq 0$. The σ_i are the singular values and the vectors $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ are the left and the right singular vectors, respectively.

If $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is symmetric and positive definite, the singular values are all positive, they coincide with the eigenvalues (i.e., $\lambda_i = \sigma_i$ for $1 \leq i \leq n$), and $\boldsymbol{U} = \boldsymbol{V}$. The columns $\boldsymbol{u}_j$ of $\boldsymbol{U}$ are then orthonormal eigenvectors of $\boldsymbol{A}$, i.e.,

$$\boldsymbol{A}\boldsymbol{u}_j = \lambda_j \boldsymbol{u}_j, \quad \boldsymbol{u}_i^T \boldsymbol{u}_j = \delta_{ij} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \qquad \text{for } 1 \leq i, j \leq n.$$

Moreover, the orthonormality of the matrix $\boldsymbol{U}$ implies that $\boldsymbol{U}^{-1} = \boldsymbol{U}^T$.

Using the properties of the SVD and denoting by $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n)$ the matrix with eigenvalues on the diagonal, we can find the inverse of $\boldsymbol{K}$ as $\boldsymbol{K}^{-1} = \boldsymbol{U}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}^T$ and obtain from (5) that

$$\boldsymbol{K}^{-1}\boldsymbol{d} = \boldsymbol{U}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}^T\boldsymbol{d} = \sum_{i=1}^{N} \lambda_i^{-1}(\boldsymbol{u}_i^T\boldsymbol{d})\boldsymbol{u}_i = \boldsymbol{p}_{true} + \sum_{i=1}^{N} \lambda_i^{-1}(\boldsymbol{u}_i^T\boldsymbol{n})\boldsymbol{u}_i. \tag{8}$$

From (8), it can be seen that instability arises for small eigenvalues λ_i, since (8) involves terms weighted by λ_i^{-1}. The eigenvalues for $\boldsymbol{K} \in \mathbb{R}^{128\times 128}$ are shown in Figure 2.1.1. For the convolution matrix $\boldsymbol{K}$, as well as for many inverse problems with PDEs, large eigenvalues λ_i correspond to smooth eigenfunctions, and small eigenvalues correspond to oscillatory eigenfunctions, as can be seen in Figure 3. Thus, from (8) it follows that oscillatory components cannot reliably be reconstructed from noisy data (i.e., when $\boldsymbol{n} \neq 0$) since they correspond to small eigenvalues. Often the noise introduced by round off error is large enough to render the explicit inversion (8) useless.

2.1.2. *Truncated SVD and Tikhonov filtering.* As a remedy to the above described problems, one can employ *filter methods*, which remove or dampen the terms corresponding to the small eigenvalues in (8). Filter functions

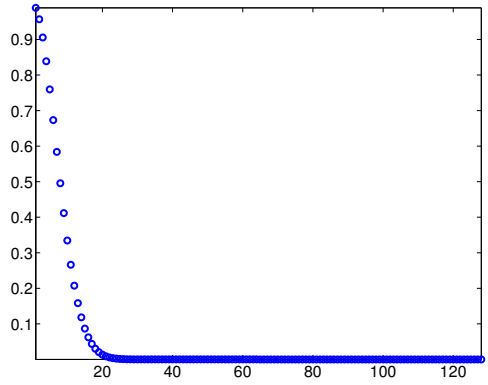

FIGURE 2. Eigenvalues of the discrete convolution operator $\boldsymbol{K}$ with $N = 128$. All eigenvalues are positive, the largest being 1 and the smallest being about 10^{-20}.

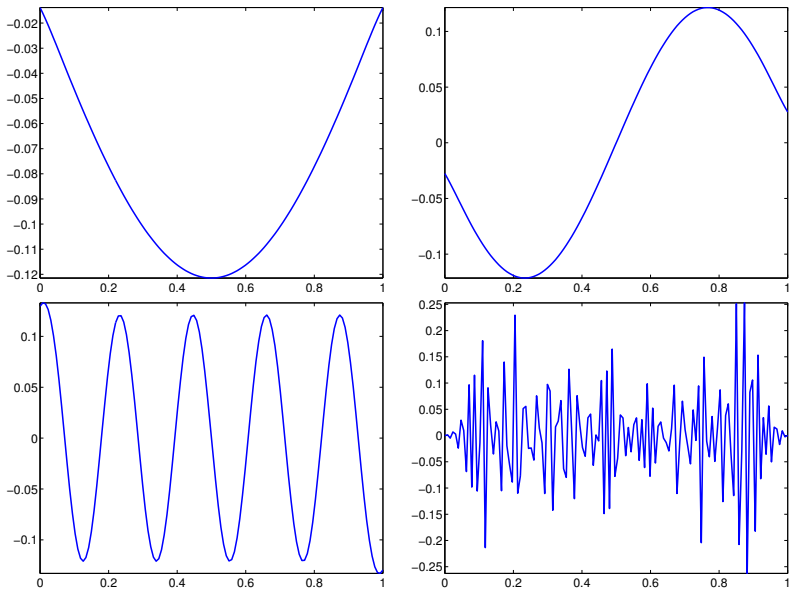

FIGURE 3. Eigenvectors corresponding to the largest (top left), second largest (top right), 10th largest (bottom left) and 100th largest (bottom right) eigenvalue.

$\omega(\lambda_i)$ for $i = 1, \ldots, n$ are employed by modifying (8) as follows:

$$\boldsymbol{p} \approx \sum_{i=1}^{N} \omega(\lambda_i^2)\lambda_i^{-1}(\boldsymbol{u}_i^T \boldsymbol{d})\boldsymbol{u}_i. \tag{9}$$

A popular choice for a family of filter function is

$$\bar{\omega}_\alpha(\lambda^2) := \begin{cases} 1 & \text{if } \lambda^2 \geq \alpha, \\ 0 & \text{else,} \end{cases} \tag{10}$$

where $\alpha > 0$ is a regularization parameter. Using this filter, (9) simplifies to

$$\boldsymbol{p}_{\text{TSVD}} = \sum_{\substack{i=1 \\ \lambda_i^2 \geq \alpha}}^{N} \lambda_i^{-1} (\boldsymbol{u}_i^T \boldsymbol{d}) \boldsymbol{u}_i, \tag{11}$$

that is, all terms corresponding to eigenvalues that are smaller than the square root of α are dropped from the sum. Due to this truncation, this method is known as *truncated singular value decomposition (TSVD)*. The parameter α controls where the sum is truncated and must be adjusted according to the noise level, see the discussion in Section 2.2.

An alternative family of filter functions in (9) is

$$\tilde{\omega}_\alpha(\lambda^2) = \frac{\lambda^2}{\lambda^2 + \alpha}, \tag{12}$$

where, again, α is a regularization parameter. Note that $\tilde{\omega}_\alpha(\lambda^2)$ is close to one when $\lambda^2 \gg \alpha$, and it is close to zero when $\lambda^2 \ll \alpha$. Thus, this filter strongly dampens terms corresponding to the small eigenvalues of $\boldsymbol{K}$, while letting the terms for large eigenvalues (almost) unchanged. It is called *Tikhonov filter* and uses a smoothed version of the truncation filter function (10). It results in the filtered image

$$\boldsymbol{p}_{\text{TIK}} = \sum_{i=1}^{N} \frac{\lambda_i}{\lambda_i^2 + \alpha} (\boldsymbol{u}_i^T \boldsymbol{d}) \boldsymbol{u}_i. \tag{13}$$

An advantage of the Tikhonov filter compared to the TSVD filter is that it can be computed without explicit knowledge of the SVD of $\boldsymbol{K}$. This is due to the fact that $\boldsymbol{p}_{\text{TIK}}$ can also be found as solution of the minimization problem

$$\min_{\boldsymbol{p}} \frac{1}{2} \|\boldsymbol{K}\boldsymbol{p} - \boldsymbol{d}\|^2 + \frac{\alpha}{2} \|\boldsymbol{p}\|^2, \tag{14}$$

where $\|\cdot\|$ denoted the Euklidian vector norm in $\mathbb{R}^N$. To show that the minimizer of (14) is $\boldsymbol{p}_{\text{TIK}}$, one uses that the minimizer is also characterized by the normal equations

$$\boldsymbol{p}_{\text{TIK}} = \left(\boldsymbol{K}^T \boldsymbol{K} + \alpha \boldsymbol{I}\right)^{-1} \boldsymbol{K}^T \boldsymbol{d}, \tag{15}$$

and uses the SVD decomposition of $\boldsymbol{K}$ in (15). The question arises if the regularization parameter α can be chosen such that the filtered solutions converge as the noise level goes to zero. For TSVD and Tikhonov filtering for the deblurring problem, this question is answered next.

2.1.3. *A deterministic error analysis.* We consider filtered solutions of (5) denoted by

$$\boldsymbol{K}_\alpha^{-1}\boldsymbol{d} := \boldsymbol{p}_\alpha = \sum_{i=1}^{N} \omega(\lambda_i^2)\lambda_i^{-1}(\boldsymbol{u}_i^T\boldsymbol{d})\boldsymbol{u}_i, \tag{16}$$

where $\boldsymbol{K}_\alpha$ denotes the filtered convolution matrix corresponding either to TSVD or Tikhonov filtering with filter parameter α. Depending on the choice of α, an error $\boldsymbol{e}_\alpha$ in the reconstruction is committed, namely

$$\boldsymbol{e}_\alpha := \boldsymbol{p}_\alpha - \boldsymbol{p}_{\text{true}} = \boldsymbol{K}_\alpha^{-1}(\boldsymbol{K}\boldsymbol{p}_{\text{true}} + \boldsymbol{n}) - \boldsymbol{p}_{\text{true}} =: \boldsymbol{e}_\alpha^{\text{trunc}} + \boldsymbol{e}_\alpha^{\text{noise}}, \tag{17}$$

where the *truncation error due to the regularization* $\boldsymbol{e}_\alpha^{\text{trunc}}$ and the *noise amplification error* $\boldsymbol{e}_\alpha^{\text{noise}}$ are defined as

$$\boldsymbol{e}_\alpha^{\text{trunc}} = \boldsymbol{K}_\alpha^{-1}\boldsymbol{K}\boldsymbol{p}_{\text{true}} - \boldsymbol{p}_{\text{true}} = \sum_{i=1}^{N} \left(\omega_\alpha(\lambda_i^2) - 1\right)(\boldsymbol{u}_i^T\boldsymbol{p}_{\text{true}})\boldsymbol{u}_i,$$

$$\boldsymbol{e}_\alpha^{\text{noise}} = \boldsymbol{K}_\alpha^{-1}\boldsymbol{n} = \sum_{i=1}^{N} \omega_\alpha(\lambda_i^2)\lambda_i^{-1}(\boldsymbol{u}_i^T\boldsymbol{n})\boldsymbol{u}_i.$$

Next, we show that for Tikhonov filtering and the TSVD, the parameter α can be chosen such that both errors converge to zero as the noise level $\delta := \|\boldsymbol{n}\|$ goes to zero. We first estimate the truncation error:

- By definition of the filter weight functions (10) and (12), it follows that $\omega_\alpha(\lambda^2) \to 1$ as $\alpha \to 0$. This immediately implies that

$$\boldsymbol{e}_\alpha^{\text{trunc}} \to 0 \quad \text{as } \alpha \to 0. \tag{19}$$

- Next we study the noise amplification factor. By using the explicit form of the filter functions, it can be verified that

$$\omega_\alpha(\lambda^2)\lambda^{-1} \le \frac{1}{\sqrt{\alpha}} \tag{20}$$

for both, TSVD and Tikhonov filtering. Using the orthonormality of $\boldsymbol{U}$, this implies that

$$\|\boldsymbol{e}_\alpha^{\text{noise}}\| \le \frac{1}{\sqrt{\alpha}}\|\sum_{i=1}^{N}(\boldsymbol{u}_i^T\boldsymbol{n})\boldsymbol{u}_i\| = \frac{\delta}{\sqrt{\alpha}}.$$

Thus, if we choose the filter parameter as $\alpha := \delta^p$ with $p < 2$ we obtain

$$\boldsymbol{e}_\alpha^{\text{noise}} \to 0 \quad \text{as } \delta \to 0. \tag{21}$$

Combining the requirements for the truncation and the noise amplification error, the choice $\alpha := \delta^p$ with $0 < p < 2$ guarantees that $\boldsymbol{e}_\alpha \to 0$ as the noise level $\delta \to 0$. This means that the TSVD and Tikhonov filters, together with the above choice for the regularization parameter are *convergent*. A significant amount of research in inverse problems deals with the computation of

rates for this convergence in linear and nonlinear inverse problems. Next, we consider practical methods for choosing the regularization parameters.

2.2. Choice of the regularization parameter. As seen above, the choice of the regularization parameter α in either TSVD or Tikhonov filtering is important. If α is small, the computation of the parameter $\boldsymbol{p}$ is unstable as in the case without filtering. On the other hand, if α is too large, information is lost in the filtered image. In this section we discuss methods to choose appropriate filter (or regularization) parameters. Both methods are *a posteriori* parameter choice methods, i.e., they require the solution of several regularized inverse problems to find an appropriate value for α. While the L-curve criterion, which is presented first, does not require knowledge of the noise level, the discrepancy principle presented afterwards is based on an estimate of $\delta := \|\boldsymbol{n}\|$.

2.2.1. *The L-curve criterion.* Choosing the filter parameter using the L-curve criterion requires the solution of inverse problems for a sequence of regularization parameters α. Then, for each α, the norm of the data misfit (also called residual) $\|\boldsymbol{K}\boldsymbol{p}_\alpha - \boldsymbol{d}\|$ is plotted against the norm of the regularization term $\|\boldsymbol{p}_\alpha\|$ in a log-log plot. This curve usually is found to be L-shaped and thus has an "elbow," i.e. a point of greatest curvature. The L-curve criterion chooses the regularization parameter corresponding to that point; see the left plot in Figure 4 for an illustration. The idea behind the L-curve criterion is that this choice for the regularization parameter is a good compromise between fitting the data and controlling the stability of the parameters. A smaller α, which correspond to points to the left of the optimal value in Figure 4, only leads to a slightly better data fit while significantly increasing the norm of the parameters. Conversely, a larger α, corresponding to points to the right of the optimal value, slightly decrease the norm of the solution, but they increase the data misfit significantly. Proving convergence for this parameter choice method is problematic and cannot be shown in all cases.

2.2.2. *The discrepancy principle.* The *discrepancy principle*, due to Morozov, chooses the regularization parameter to be the largest value of α such that the norm of the misfit is bounded by the noise level in the data, i.e.,

$$\|\boldsymbol{K}\boldsymbol{p}_\alpha - \boldsymbol{d}\| \leq \delta, \tag{22}$$

where δ is the noise level. Here, $\boldsymbol{p}_\alpha$ denotes the parameter found either using a TSVD filter or Tikhonov regularization with parameter α. This choice aims to avoid overfitting of the data, i.e., fitting the noise. The criterion is illustrated in the right plot in Figure 4. Convergence results and rates for the parameter when determined by the Morozov criterion as the noise level goes to zero are available.

Next, we prove that for the discretized deblurring problem such a regularization parameter α always exists provided the noise level is less than the

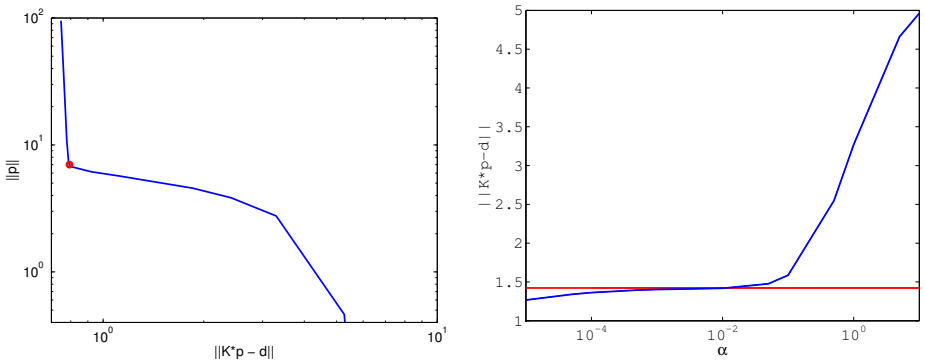

FIGURE 4. Choosing the regularization parameter α: The red dot on the L-curve (left plot), which corresponds to the point with largest curvature, yields the optimal filter/regularization parameter according to the L-curve criterion. For the discrepancy criterion (right plot), the optimal parameter corresponds to the intersection of the data misfit curve with the red line indicating the noise level.

norm of the data, i.e., $\delta < \|\boldsymbol{d}\|$. For that purpose we define the function

$$D(\alpha) := \|\boldsymbol{K}\boldsymbol{p}_\alpha - \boldsymbol{d}\| \tag{23}$$

Using the form of the Tikhonov-regularized parameter $\boldsymbol{p}_\alpha$ as given in (13) and the eigenvalue decomposition $\boldsymbol{K} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T$, one obtains

$$\boldsymbol{K}\boldsymbol{p}_\alpha - \boldsymbol{d} = \sum_{i=1}^{N} \left(\frac{\lambda_i^2}{\lambda_i^2 + \alpha} - 1 \right) (\boldsymbol{u}_i^T \boldsymbol{d}) \boldsymbol{u}_i,$$

and due to the orthonormality of $\boldsymbol{U}$ this implies

$$D^2(\alpha) = \sum_{i=1}^{N} \left(\frac{\lambda_i^2}{\lambda_i^2 + \alpha} - 1 \right)^2 (\boldsymbol{u}_i^T \boldsymbol{d})^2. \tag{24}$$

This shows that $D(\alpha)$ is continuous in α, that $D(0) = 0$ and that D is monotonically increasing. Moreover, $D(\alpha) \to \|\boldsymbol{d}\|$ as $\alpha \to \infty$. Thus, provided $\delta < \|\boldsymbol{d}\|$, there exists an α such that $D(\alpha) = \delta$, as desired.

Note that a similar argument does not work for TSVD filtering, since the function D is not continuous in α for the filter function (10). Thus, in general, the optimal α according to the discrepancy principle will satisfy (22) with a strict inequality.

2.3. Variational regularization methods. As discussed above, the Tikhonov filtered solution can also be found through the solution of an optimization problem. This has the advantage that no explicit SVD is required. Moreover, such an optimization approach allows more flexibility in the choice of norms for the misfit and the regularization, as is often desired in

variational inverse problems, i.e., inverse problems that involve differential equations. To illustrate this flexibility we consider the following generalization of the optimization problem (14):

$$\min_{\boldsymbol{p}} \frac{1}{2}\|\boldsymbol{K}\boldsymbol{p} - \boldsymbol{d}\|^2 + \mathcal{R}(\boldsymbol{p}), \tag{25}$$

with a regularization function $\mathcal{R} : \mathbb{R}^N \to \mathbb{R}$. Above we have discussed the choice of $\mathcal{R}(\boldsymbol{p}) = \alpha/2\|\boldsymbol{p}\|^2$. Alternative choice are the squared difference operator

$$\mathcal{R}_2(\boldsymbol{p}) = \frac{\alpha}{2}\sum_{i=1}^{N-1}(p_{i+1} - p_i)^2, \quad \text{with } \boldsymbol{p} = (p_1, \ldots, p_N), \tag{26}$$

which is closely related to the squared gradient if $\boldsymbol{p}$ corresponds to the discretization of a function. The choice (26) favors the parameter $\boldsymbol{p}$ that has small differences between its components. If the vector $\boldsymbol{p}$ originates from a discretized parameter function, (26) expresses a preference for smooth parameter functions. An alternative choice is to replace the sum of squares by a sum of absolute values

$$\mathcal{R}_1(\boldsymbol{p}) = \alpha\sum_{i=1}^{N-1}|p_{i+1} - p_i|. \tag{27}$$

Similar to (26), the regularization (27) favors small differences. However, compared to (26) it puts less emphasis on large values in the sum. For discretized functions, the choice (27) corresponds to *total variation regularization*, which is a popular regularization for inverse problems, in particular in imaging. Note that $\mathcal{R}_2$ corresponds to the squared Euclidian (also called ℓ_2-norm) of the differences, while $\mathcal{R}_1$ corresponds to the ℓ_1 norm of the differences. Since $\mathcal{R}_1$ is not differentiable due to the absolute value $|\cdot|$, which makes computing derivatives of (25), as required by numerical optimization methods, challenging.