

# Seismic wavefield redatuming with regularized multi-dimensional deconvolution

Nick Luiken<sup>1,\*</sup>, Tristan van Leeuwen<sup>1</sup>

<sup>1</sup> Department of Mathematics, Utrecht University

Email: n.a.luiken@uu.nl, t.vanleeuwen@uu.nl

\* corresponding author

March 25, 2020

## Abstract

In seismic imaging the aim is to obtain an image of the subsurface using reflection data. The reflection data are generated using sound waves and the sources and receivers are placed at the surface. The target zone, for example an oil or gas reservoir, lies relatively deep in the subsurface below several layers. The area above the target zone is called the overburden. This overburden will have an imprint on the image. Wavefield redatuming is an approach that removes the imprint of the overburden on the image by creating so-called virtual sources and receivers above the target zone. The virtual sources are obtained by determining the impulse response, or Green's function, in the subsurface. The impulse response is obtained by deconvolving all up- and downgoing wavefields at the desired location. In this paper, we pose this deconvolution problem as a constrained least-squares problem. We describe the constraints that are involved in the deconvolution and show that they are associated with orthogonal projection operators. We show different optimization strategies to solve the constrained least-squares problem and provide an explicit relation between them, showing that they are in a sense equivalent. We show that the constrained least-squares problem remains ill-posed and that additional regularization has to be provided. We show that Tikhonov regularization leads to improved resolution and a stable optimization procedure, but that we cannot estimate the correct regularization parameter using standard parameter selection methods. We also show that the constrained least-squares can be posed in such a way that additional nonlinear regularization is possible.

*Keywords*— seismic imaging, redatuming, multi-dimensional deconvolution, regularization, Tikhonov regularization, reciprocity, causality.

## 1 Introduction

In seismic imaging, one aims to obtain an image of the subsurface from reflection data. Here, an impulse source sends waves in to the subsurface and the reflected response is recorded by an array of receivers. Reconstructing an image from the reflected data is an inverse problem that has been studied extensively (for an extensive review see [27]). Wavefield redatuming is an inverse problem that appears in the same context and is often considered a pre-processing procedure. The goal of wavefield redatuming is to remove the effects of a part of the medium that is not of primary interest for imaging purposes (called the *overburden*), thereby making subsequent imaging of the *target zone* (e.g., an oil reservoir) easier. Redatuming transforms the response of the medium (overburden plus target zone) to the responses of the target zone only. This situation is depicted schematically in figure 1.

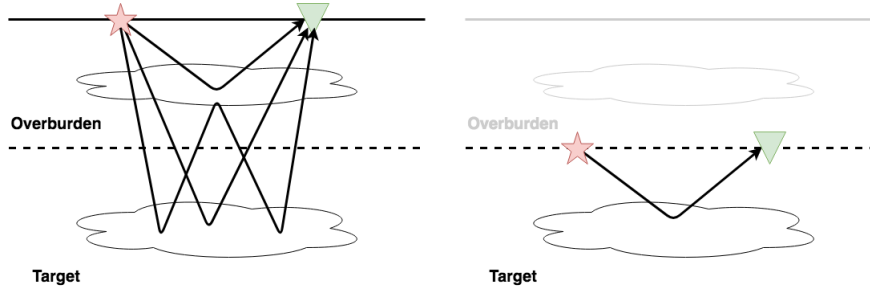


Figure 1: Schematic depiction of the redatuming procedure. A seismic survey consists of emitting waves in to the subsurface from an impulsive source (red star) and recording the reflected response (green triangle). The goal is to transform recorded reflection data that includes the response of both the overburden and target zone to the response of the target zone only.

For an extensive overview of the redatuming problem we refer to [2, 26, 36, 34, 29, 28, 30, 24]. A key ingredient in all redatuming methods is *multi-dimensional deconvolution*. Here, the data and unknown response are related through multi-dimensional convolution with a given kernel. This leads to a linear, ill-posed inverse problem

$$p(t, x, x') = \iint g(t - s, x, y) q(s, y, x') ds dy,$$

where  $p : \mathbb{R}^3 \rightarrow \mathbb{R}$  and  $q : \mathbb{R}^3 \rightarrow \mathbb{R}$  are given and  $g : \mathbb{R}^3 \rightarrow \mathbb{R}$  is the unknown impulse response. After discretisation this yields a linear system of matrix equations

$$QG = P,$$

where  $Q$  is a block Toeplitz matrix,  $P$  contains the measurements and  $G$  represents the impulse response. It should be noted here that generally, both  $Q$  and  $P$  are measured. Thus, the usual distinction between model and data does not hold as both  $Q$  and  $P$  are contaminated with noise. Nevertheless, we may attempt to solve the inverse problem by posing it as a regularized least-squares problem:

$$\min_G \|QG - P\|_F^2 + \lambda \|G\|_F^2, \quad (1)$$

where  $\|\cdot\|_F$  denotes the Frobenius norm and  $\lambda > 0$  is the regularization parameter. Because of the block-Toeplitz structure of  $Q$  this can be efficiently solved block-by-block in the Fourier domain. Typical difficulties that are encountered when solving such multi-dimensional deconvolution problems are illustrated in the following example.

**Example 1.1.** To illustrate the idea we show an example of source redatuming after the one in [26]. Here,  $q(t, r, s)$  is the transmitted (downgoing) wavefield generated by a point source at location  $(0, s)$  (indicated by red stars in figure 2 (a)) as recorded by the receivers (indicated by green triangles) at location  $(500, r)$  and  $p(t, r, s)$  is the reflected (upgoing) wavefield recorded at the same receivers. The impulse response  $g$  corresponds to a virtual experiment where sources are placed at  $z = 500$  (depicted in 2 (b)). The corresponding wavefields  $q$  and  $p$  for a source at  $(0, 0)$  as generated by a finite-difference modelling code are shown in figure 3 (a). The resulting estimate of  $g$  is depicted in figure 3 (b), alongside the true response (also generated by a finite-difference modelling code). Although the main features are reconstructed, we see some notable artefacts in the solution. In particular, we see non-physical events arriving before the first arrival. To counter such artefacts, we need a regularization method that takes in to account such prior knowledge of the underlying physics.

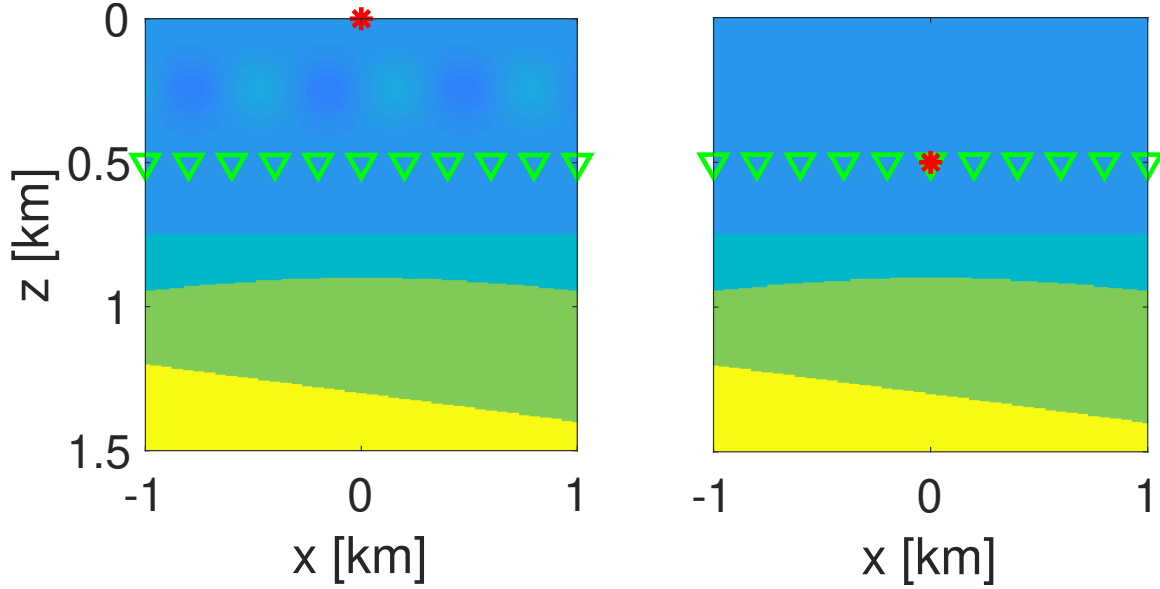


Figure 2: An example of source redatuming. The dashed red line indicates a wave traveling from source to receiver. The black line indicates a wave traveling the same path, but passing through the receiver after which it reflects and goes to another receiver. The difference between the two waves is the Green's function, shown in the figure on the right.

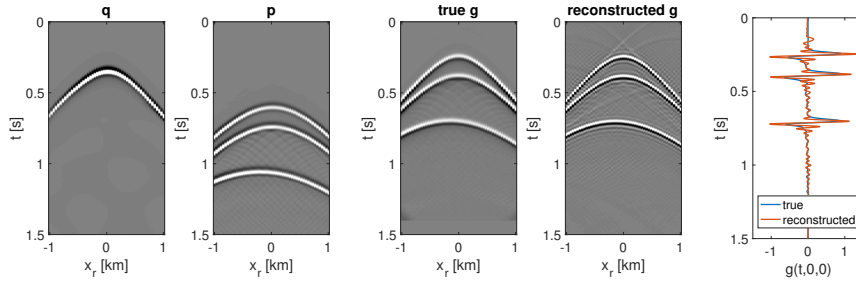


Figure 3: Wavefields  $q(t, r, s = 0)$ ,  $p(t, r, s = 0)$  associated with the subsurface in the example from figure (2) and the corresponding true and reconstructed response  $g(t, r, s = 0)$ .

## 1.1 Approach and challenges

To include prior physical constraints in the reconstruction, we pose the inverse problem as a constrained least-squares problem:

$$\min_G \|QG - P\|_F^2 \quad \text{such that} \quad G \in \mathcal{A}, \quad (2)$$

where  $\mathcal{A}$  denotes a (convex) set of admissible solutions. Typical constraints include *causality*:  $g(t, s, r) = 0$  when  $t < \tau(s, r)$  for some given function  $\tau$ , and *reciprocity*:  $g(t, s, r) = g(t, r, s)$ . Note that even though  $p$  and  $q$  should obey this constraint as well, noise or modelling errors may cause the unregularized solution to violate this constraint. The difficulty in solving the inverse problem is that it is typically underdetermined and rank deficient. This is due to the fact that a seismic survey usually has less sources than receivers and the measurements are bandlimited. Therefore, further regularization besides the constraints is needed to

stabilize the solution.

## 1.2 Contribution

In this paper we pose multi-dimensional deconvolution as a constrained least-squares problem. In particular, we treat the source-redatuming problem with causality and reciprocity constraints. We show that these constraints are associated with orthogonal projection operators. We describe different optimization methods to incorporate the constraints in the optimization and show explicit relations between the methods. We show that the optimization methods are in a certain way equivalent, but that solving them numerically leads to different solutions. We show that even with incorporating the constraints, the problem exhibits semiconvergence, which means that the optimization scheme is still not stable. This means that the iterations have to be stopped at the appropriate point. Finally, we show that the addition of a Tikhonov penalty can further improve the reconstruction, but that standard parameter selection methods do not yield a good estimate for the regularization parameter. This makes the addition of a Tikhonov penalty impractical.

## 1.3 Outline

The paper is organized as follows. In section (2) we describe the MDD problem and set up the discretized linear system. In section (3) we describe how the symmetry constraint and the causality constraint can be incorporated in the optimization. In section (4) we compare different optimization strategies and detail the difficulties in solving the optimization problem. Finally, in section (5), we draw our conclusions and add a short discussion and outlook.

## 2 Source redatuming

We start from the scalar wave equation in  $\mathbb{R}^n$ :

$$(c(x)^{-2}\partial_t^2 - \nabla^2) u(t, x, x') = f(t)\delta(x - x'), \quad (3)$$

where  $c$  is the soundspeed in the medium, and  $f$  is the time-signature of the source. The wave equation is furnished with appropriate boundary and initial conditions to ensure causal, outward propagating solutions. We split the medium in two parts; the *overburden* where  $c(x) = c_0(x)$  and the *target zone*, where  $c(x) = c_1(x)$ . In essence, the inverse problem is as follows; given measurements of  $u$  at depth level  $\bar{z}$  we want to retrieve the impulse response of the target zone that excludes any effects from the overburden. In absence of horizontally propagating waves, we can split the solution to (3) in to an upgoing and downgoing part,  $u_-$  and  $u_+$  [4]. The upgoing and downgoing constituents can be obtained by solving a system of equations involving  $u$  and its vertical derivative [9, 35]. We now consider measurements of  $u$  at  $x_r = (\bar{z}, r)$  originating from a source at  $x_s = (z_0, s)$  and set  $q(t, r, s) = u_+(t, x_r, x_s)$  and  $p(t, r, s) = u_-(t, x_r, x_s)$ . These two quantities are related to the upgoing response at  $(\bar{z}, r)$  to a downward radiating source at  $(\bar{z}, s)$  via convolution:

$$p(t, r, s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(t - t', r, r') q(t, r', s) dt' dr'. \quad (4)$$

The goal is to recover  $g$  from (noisy) samples  $p_{ijk} := p(t_i, r_j, s_k) + \epsilon_{ijk}$  and  $q_{ijk} := q(t_i, r_j, s_k) + \delta_{ijk}$ , with  $\epsilon_{ijk}$  and  $\delta_{ijk}$  representing the noise terms. For more details regarding the derivation of this relation we refer to [36, 34, 24]. A concrete example illustrating the ill-posedness of the problem is given below.

### 2.1 Analysis for layered media

For horizontally layered media, the wavefields and impulse response can be expressed as

$$p(t, r, s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{p}(\omega, \xi) \exp(-i(\omega t + \xi(r - s))) d\omega d\xi,$$

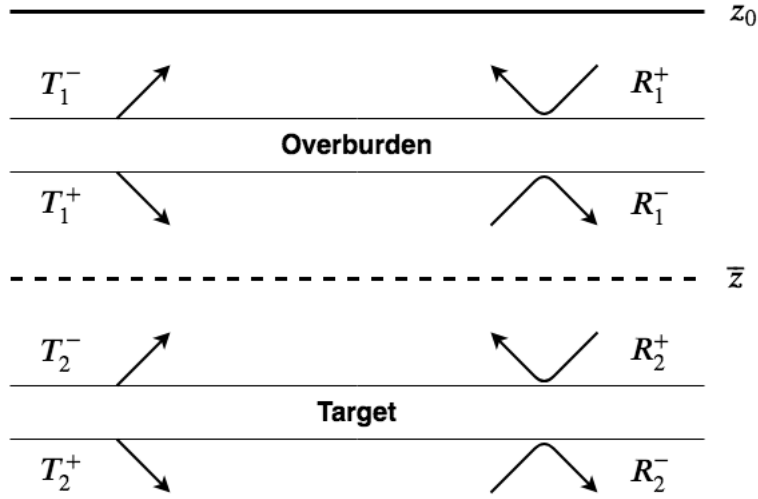


Figure 4: Schematic depiction of the up- and downgoing responses of the overburden and target. A superscript  $-$  indicates the responses of the medium to an upgoing plane wave, while a superscript  $+$  indicates the response to a downgoing plane wave.

and similary for  $q$  and  $g$ . The forward relation (4) can be expressed in terms of the Fourier transformed quantities as

$$\hat{p}(\omega, \xi) = \hat{q}(\omega, \xi) \cdot \hat{g}(\omega, \xi).$$

We can get an explicit expression for the up and downgoing solutions as follows. For a horizontally layered medium we can explicitly factorize the wave equation in the Fourier domain:

$$\hat{u}'_{\pm}(z) \mp k(z)\hat{u}_{\pm}(z) = 0,$$

with  $k(z) = \sqrt{c(z)^{-2}\omega^2 - |\xi|^2}$  where  $\omega$  is the temporal frequency and  $\xi$  the horizontal wavenumber. We can now think of the response of the overburden and target in terms of incoming plane waves. We denote the transmitted and reflected response of the overburden to an up/downgoing plane wave by  $T_0^{\pm}$ ,  $R_0^{\pm}$  and likewise we denote the responses of the target zone by  $T_1^{\pm}$ ,  $R_1^{\pm}$ . Figure 4 illustrates the situation. The downgoing response of the entire medium to a downgoing plane wave  $e^{ik_0 z}$  at  $\bar{z}$  can now be expressed as

$$\hat{q} = T_1^+ + R_1^- R_2^+ T_1^+ + \dots = (1 - R_1^- R_2^+)^{-1} T_1^+.$$

Likewise, the upgoing response at  $\bar{z}$  can be descibed by

$$\hat{p} = R_2^+ T_1^+ + R_2^+ R_1^- R_2^+ T_1^+ + \dots = (1 - R_1^- R_2^+)^{-1} R_2^+ T_1^+.$$

We immediately see that

$$\hat{p}/\hat{q} = R_2^+,$$

which is the impulse response of the target zone to a downgoing planewave measured at  $z = \bar{z}$ . The main cause of the ill-posedness of the inverse problem is the bandlimited nature of the measured responses; they only contain propagating modes for which  $|\xi| < \omega/c$ . In practice, the measured response contains the imprint of the source wavelet  $\hat{f}(\omega)$  and is further band-limited in  $\xi$  because it is measured with a finite array of receivers. The estimated response is then given by

$$\hat{g} = \frac{(\hat{q} + \hat{\epsilon})^* (\hat{p} + \hat{\delta})}{|\hat{q} + \hat{\epsilon}|^2 + \lambda},$$

where  $\hat{\epsilon}$  and  $\hat{\delta}$  represent the measurement noise and  $\lambda > 0$  is the regularization parameter (cf. (1)). This can be decomposed as

$$\hat{g} = \frac{|\hat{q} + \hat{\epsilon}|^2}{|\hat{q} + \hat{\epsilon}|^2 + \lambda} R_2^+ - \frac{(\hat{q} + \hat{\epsilon})^* \hat{\epsilon}}{|\hat{q} + \hat{\epsilon}|^2 + \lambda} R_2^+ + \frac{(\hat{q} + \hat{\epsilon})^* \delta}{|\hat{q} + \hat{\epsilon}|^2 + \lambda}.$$

Thus the error can be bounded as

$$|\hat{g} - R_2^+| \leq \frac{\lambda}{|\hat{q} + \hat{\epsilon}|^2 + \lambda} \cdot |R_2^+| + \frac{|\hat{q} + \hat{\epsilon}|}{|\hat{q} + \hat{\epsilon}|^2 + \lambda} \cdot (|\hat{\epsilon}| \cdot |R_2^+| + |\hat{\delta}|),$$

from which we recognize a bias and variance term. Notably, we see the regularizing effect that  $\lambda$  has on both sources of noise.

**Example 2.1.** As a concrete example we consider a medium with three horizontal layers with sound speed  $c_i$  for  $z \in [z_i, z_{i+1})$ . The redatuming level is set at  $\bar{z} = (z_1 + z_2)/2$ . The corresponding transmission and reflection responses can then be explicitly expressed in terms of  $k_i = \sqrt{(\omega/c_i)^2 - \xi^2}$ :

$$T_1^+ = \frac{2k_0}{k_1 + k_0} e^{-ik_1 h_1/2}, \quad R_1^- = \frac{k_1 - k_0}{k_1 + k_0} e^{-ik_1 h_1}, \quad R_2^+ = \frac{k_1 - k_2}{k_1 + k_2} e^{-ik_1 h_1},$$

with  $h_1 = z_2 - z_1$ . Two typical examples of the corresponding spectra  $\hat{q}$ ,  $\hat{p}$  and  $\hat{g}$  are illustrated in figure 5. We see that when  $c_0 > c_1$ , part of the impulse response,  $\hat{g}$ , is in the null-space of  $\hat{q}$  – the modes for which  $|\xi| > \omega/c_0$ . A properly regularized inversion will thus at best give an estimate of  $\hat{g}$  that is restricted to the support of  $\hat{q}$  in the  $(\omega, \xi)$  domain.

## 2.2 Discretization

Assuming the signals are regularly sampled and the spatial samples are co-located, i.e.  $t_i = i \cdot \Delta t$  for  $i = 0 \dots n_t - 1$ ,  $r_j = j \cdot \Delta r$  for  $j = 0 \dots n_r - 1$  and  $r_k = k \cdot \Delta s$  for  $k = 0 \dots n_s - 1$ , we can represent the signals in terms of their samples as

$$p(t, r, s) = \sum_{ijk} p_{ijk} \operatorname{sinc}\left(\frac{t - t_i}{\Delta t}\right) \operatorname{sinc}\left(\frac{r - r_j}{\Delta r}\right) \operatorname{sinc}\left(\frac{s - s_k}{\Delta s}\right),$$

$$q(t, r, s) = \sum_{ijk} q_{ijk} \operatorname{sinc}\left(\frac{t - t_i}{\Delta t}\right) \operatorname{sinc}\left(\frac{r - r_j}{\Delta r}\right) \operatorname{sinc}\left(\frac{s - s_k}{\Delta s}\right),$$

and

$$g(t, r, s) = \sum_{ijk} g_{ijk} \operatorname{sinc}\left(\frac{t - t_i}{\Delta t}\right) \operatorname{sinc}\left(\frac{r - r_j}{\Delta r}\right) \operatorname{sinc}\left(\frac{s - s_k}{\Delta s}\right).$$

Using the orthogonality relations of the normalized sinc function, we can re-write this as a system of matrix-equations with a block-circulant structure<sup>1</sup>:

$$P = QG,$$

with

$$Q = \begin{pmatrix} Q_0 & Q_{n_t-1} & \dots & Q_1 \\ Q_1 & Q_0 & \dots & Q_2 \\ \vdots & \ddots & \ddots & \vdots \\ Q_{n_t-1} & \dots & Q_1 & Q_0 \end{pmatrix}, \quad G = \begin{pmatrix} G_0 \\ G_1 \\ \vdots \\ G_{n_r-1} \end{pmatrix}, \quad P = \begin{pmatrix} P_0 \\ P_1 \\ \vdots \\ P_{n_s-1} \end{pmatrix}.$$

Here,  $Q_i \in \mathbb{R}^{n_s \times n_r}$  is a matrix with elements  $q_{ijk}$ ,  $P$  is a block matrix with  $n_t$  blocks  $P_i \in \mathbb{R}^{n_s \times n_r}$  with elements  $p_{ijk}$  and  $G$  is a block matrix with  $n_t$  blocks  $G_i \in \mathbb{R}^{n_r \times n_r}$ . Since a circulant matrix diagonalizes under the Discrete Fourier Transform,  $F_{n_t}$  (with entries  $\exp(i \frac{2\pi ij}{n_t})$ ), we can express  $Q$  as

$$Q = (F_{n_t} \otimes I_{n_r \cdot n_s}) \hat{Q} (F_{n_t}^{-1} \otimes I_{n_r \cdot n_s}),$$

<sup>1</sup>assuming that we are looking for a solution that is periodic in time

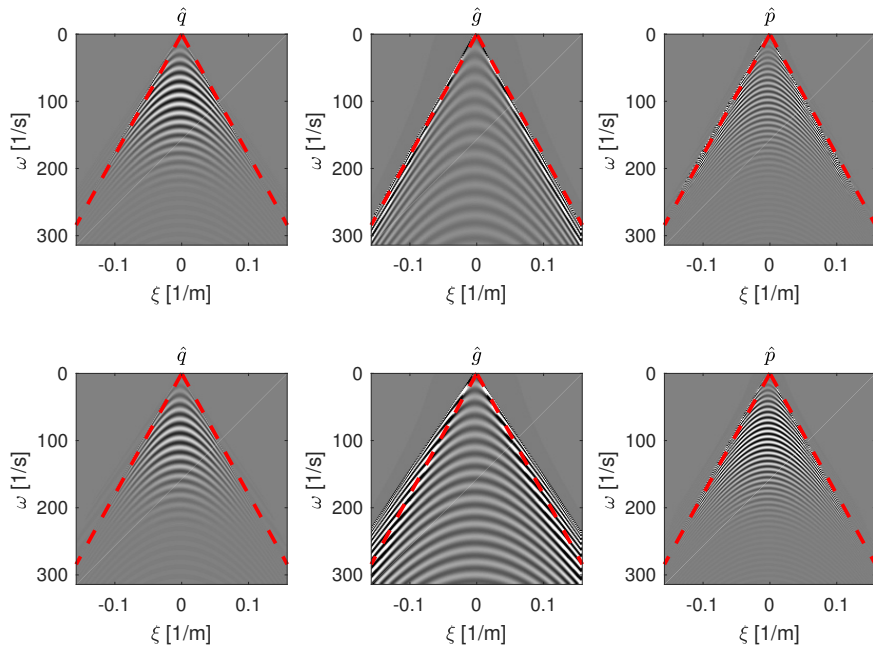


Figure 5: Spectrum of the responses  $\hat{q}$ ,  $\hat{g}$  and  $\hat{p}$  for a layered medium with  $z_0 = 0$  m,  $z_1 = 500$  m,  $z_2 = 1000$  m and  $c_0 = 1500$  m/s,  $c_1 = 1800$  m/s,  $c_2 = 2000$  m/s (top),  $c_0 = 1800$  m/s,  $c_1 = 1500$  m/s,  $c_2 = 2000$  m/s (bottom). In the top scenario, we see that the spectrum of  $\hat{q}$  has the same support as that of  $\hat{g}$ , making it in principle possible to retrieve  $\hat{g}$  completely from  $\hat{q}$ . In the bottom scenario, the spectrum of  $\hat{q}$  has a narrower support than that of  $\hat{g}$ , making it impossible to recover the complete spectrum of  $\hat{g}$  from  $\hat{p}$ .

where

$$\hat{Q} = \text{blockdiag} \left( \hat{Q}_0, \hat{Q}_1, \dots, \hat{Q}_{n_t} \right),$$

and

$$\hat{Q}_i = \sum_{j=0}^{n_t-1} \exp \left( i \frac{2\pi i j}{n_t} \right) Q_j.$$

This means we can decouple the system into  $n_t$  matrix equations

$$\hat{Q}_i \hat{G}_i = \hat{P}_i.$$

The constraints, however, may not decouple in this fashion. Moreover, it is not very attractive to have to estimate a separate regularization parameter for each frequency separately. We therefore stick with a time-domain formulation. Matrix-vector multiplication with  $Q$  are carried out in the frequency-domain, however, for computational efficiency.

### 3 Constrained least squares

In this section we describe how to solve the constrained least-squares problem (2) and describe two relevant constraints, causality and reciprocity, in detail. For each constraint we describe an orthogonal projection operator  $P_A : \mathbb{R}^{n_t \times n_r \times n_r} \rightarrow \mathbb{R}^{n_t \times n_r \times n_r}$  i.e., the operator that solves

$$P_A(G) = \arg \min_{G' \in \mathbb{R}^{n_t \times n_r \times n_r}} \|G' - G\|_F^2 \quad \text{s.t.} \quad G' \in \mathcal{A}.$$

Related to the projection operator is the penalty operator,  $L_A : \mathbb{R}^{n_t \times n_r \times n_r} \rightarrow \mathbb{R}^{n_t \times n_r \times n_r}$ , related to the projection as

$$P_A(G) + L_A(G) = G \quad \forall G \in \mathbb{R}^{n_t \times n_r \times n_r}.$$

#### 3.1 The constraints

The system has to satisfy two binding constraints. The first constraint is source-receiver reciprocity. This prior requires that a wave travels from source location to receiver location in the same time as a wave traveling from receiver location to source location. This means that the impulse response  $G$  has to satisfy  $g_{ijk} = g_{ikj}$ , for all  $i, j, k$ . The second constraint is causality which means that  $g_{ijk} = 0$  for  $i < \tau_{jk}$  for some given matrix  $\tau$ . Below, we formulate the projection and penalty operators for each constraint.

##### 3.1.1 Causality

The set of causal solutions is given by

$$\mathcal{C} = \{G \in \mathbb{R}^{n_t \times n_r \times n_r} \mid g_{ijk} = 0 \text{ for } 1 \leq i < \tau_{jk}, 1 \leq j \leq n_r, 1 \leq k \leq n_r\},$$

where  $\tau_{jk} > 1$  are given. The corresponding projection and penalty operators are given by

$$P_{\mathcal{C}}(G)_{ijk} = \begin{cases} g_{ijk} & \text{if } i \geq \tau_{jk} \\ 0 & \text{otherwise} \end{cases}$$

$$L_{\mathcal{C}}(G)_{ijk} = \begin{cases} g_{ijk} & \text{if } i < \tau_{jk} \\ 0 & \text{otherwise} \end{cases}$$



### 3.1.2 Reciprocity

The set of solutions satisfying reciprocity is given by

$$\mathcal{R} = \{G \in \mathbb{R}^{n_t \times n_r \times n_r} \mid g_{ijk} = g_{ikj} \text{ for } 1 \leq i \leq n_t, 1 \leq j \leq n_r, 1 \leq k \leq n_r\}.$$

The dimension of  $\mathcal{R}$  is  $n_t \cdot n_r \cdot n_r(n_r + 1)/2$  and orthogonal basis for  $\mathcal{R}$  and its complement can be easily constructed. An example for  $n_t = 1$ ,  $n_r = 3$  is shown in figure 6. The corresponding projection and penalty operators are given by

$$P_{\mathcal{R}}(G)_{ijk} = \frac{1}{2} (g_{ijk} + g_{ikj}),$$

$$L_{\mathcal{R}}(G)_{ijk} = \frac{1}{2} (g_{ijk} - g_{ikj}).$$

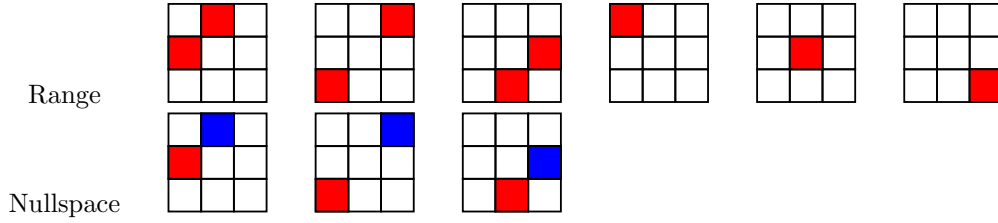


Figure 6: Basis elements for the symmetric and anti-symmetric  $3 \times 3$  matrices. Blue is -1 and red is 1.

### 3.1.3 Intersection

To include both constraints we need the projection and penalty operators for the set  $\mathcal{R} \cap \mathcal{C}$ . In general, one cannot naively project on to the intersection by concatenating the individual projection operators. In this particular setting, however, the constraints are consistent with each other and causality will not violate reciprocity and vice versa. This requires that the time window  $\tau$  is symmetric (i.e., satisfies reciprocity). The orthogonal projection on to  $\mathcal{R} \cap \mathcal{C}$  is thus simply given by

$$P_{\mathcal{R} \cap \mathcal{C}}(G)_{ijk} = \begin{cases} \frac{1}{2}(g_{ijk} + g_{ikj}) & \text{if } i \geq \tau_{jk} \\ 0 & \text{otherwise} \end{cases},$$

and the corresponding penalty operator by

$$L_{\mathcal{R} \cap \mathcal{C}}(G)_{ijk} = \begin{cases} \frac{1}{2}(g_{ijk} - g_{ikj}) & \text{if } i \geq \tau_{jk} \\ g_{ijk} & \text{otherwise} \end{cases},$$

An example of the orthogonal basis of  $\mathcal{R} \cap \mathcal{C}$  and its complement for  $n_t = 1$ ,  $n_r = 3$  is shown in figure 7.

## 3.2 Solving the constrained least-squares problem

Before continuing we first express (2) in a more convenient form

$$\min_{\mathbf{g}} \|(I \otimes Q)\mathbf{g} - \mathbf{p}\|_2^2 \quad \text{s.t.} \quad L\mathbf{g} = \mathbf{0}, \quad (5)$$

where  $\mathbf{g} = \text{vec}(G) \in \mathbb{R}^n$ ,  $\mathbf{p} = \text{vec}(P) \in \mathbb{R}^m$  with  $n = n_t \cdot n_r^2$  and  $m = n_t \cdot n_r \cdot n_s$ , and  $L \in \mathbb{R}^{n \times n}$  is the penalty operator corresponding to the admissible set  $\mathcal{A}$ . In the remainder of the paper we will refer to  $I \otimes Q$  as  $Q$  for ease of notation. Given an orthogonal basis,  $A$ , for  $\mathcal{A}$  we have  $P = AA^T$  and  $L = I - AA^T$ .

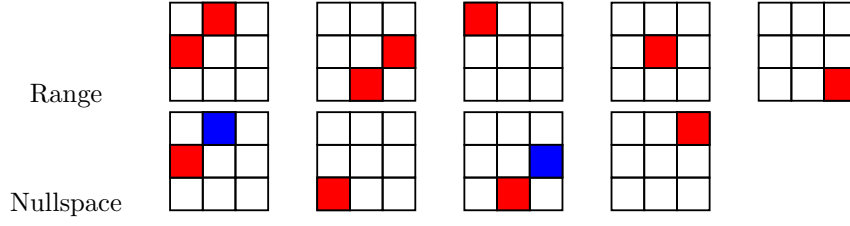


Figure 7: Basis elements for the symmetric matrices with  $G_{31} = G_{13} = 0$ . Note that the second element for the symmetric and anti-symmetric matrices have been replaced by basis elements making the corresponding entries 0.

The solution to (5) can be expressed as  $\mathbf{g} = A\mathbf{y}$  where  $\mathbf{y}$  solves

$$\min_{\mathbf{y}} \|QA\mathbf{y} - \mathbf{p}\|_2^2.$$

Hence, the (minimum-norm) solution to (5) is given by

$$\mathbf{g} = A \left( A^T Q^T Q A \right)^\dagger A^T Q^T \mathbf{p}. \quad (6)$$

We note that in practical applications,  $Q$ ,  $P$  and  $L$  are never formed explicitly; their action is computed on-the-fly in a matrix-free fashion. Next, we discuss three approaches for finding a solution to (5).

### 3.2.1 All-at-once

The first-order Karush-Kuhn-Tucker optimality conditions corresponding to (5) lead to a saddle-point problem

$$\begin{pmatrix} 0 & Q^T & L \\ Q & -I & 0 \\ L & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{g} \\ \mathbf{r} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{p} \\ \mathbf{0} \end{pmatrix}, \quad (7)$$

from which the residual  $\mathbf{r}$  can be eliminated to yield

$$\begin{pmatrix} Q^T Q & L \\ L & 0 \end{pmatrix} \begin{pmatrix} \mathbf{g} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} Q^T \mathbf{p} \\ \mathbf{0} \end{pmatrix}. \quad (8)$$

This system of equations can be readily solved with an iterative method like MINRES to yield the desired solution [23]. Applying the substitution  $\mathbf{g} = A\mathbf{y}$  the system reduces to

$$Q^T Q A \mathbf{y} + L \boldsymbol{\mu} = Q^T \mathbf{p}. \quad (9)$$

Projecting onto  $A^T$  yields  $A^T Q^T Q A \mathbf{y} = A^T Q^T \mathbf{p}$ , whose solution is given by  $\mathbf{y} = (A^T Q^T Q A)^\dagger A^T Q^T \mathbf{p}$ , and hence  $\mathbf{g} = A (A^T Q^T Q A)^\dagger A^T Q^T \mathbf{p}$ , which coincides with (6).

Note that we do not need to explicitly form the matrix in order to do so; we can easily compute matrix-vector multiplications with the system matrix on-the-fly.

### 3.2.2 Right preconditioning

We can explicitly eliminate the constraint in (5) via a substitution  $\mathbf{g} = P\tilde{\mathbf{g}}$ :

$$\min_{\tilde{\mathbf{g}}} \|QP\tilde{\mathbf{g}} - \mathbf{p}\|_2^2. \quad (10)$$

The resulting least-squares problem may not have a unique solution due to overlap of the null-spaces of  $Q$  and  $P$ . A minimum norm solution can be readily obtained using LSQR. To see that this is indeed equivalent to (5), we note that the minimum norm solution to (10) is given by

$$\mathbf{g} = P \left( PQ^T QP \right)^\dagger PQ^T \mathbf{p},$$

which is indeed equivalent to (6). To show this, use that  $(AHA^T)^\dagger = AH^\dagger A^T$  where  $A^T A = I$  and  $H$  is symmetric.

### 3.2.3 Quadratic penalty

We may incorporate the constraints via a quadratic penalty and solve

$$\min_{\mathbf{g}} \|Q\mathbf{g} - \mathbf{p}\|_2^2 + \rho \|L\mathbf{g}\|_2^2. \quad (11)$$

We can solve this in a straightforward fashion using LSQR by introducing the augmented matrix  $[Q^T, \sqrt{\rho}L]^T$ . In [21] and [10] this method is used to analyze symmetric solutions to matrix equations with  $\rho = 1$ . For ill-posed inverse problems, the choice of  $\rho$  may be slightly more involved. If the system admits a symmetric solution then this approach works for any  $\rho$ . However, if  $\mathbf{p}$  and possibly also  $\mathbf{g}$  are perturbed, the solution may no longer admit a symmetric solution. In this case, we have to find the best symmetric solution and in this case there is a trade-off between symmetry and data misfit. To ensure a symmetric solution, we have to choose  $\rho$  large enough. Consider the following example.

**Example 3.1.** Let  $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ ,  $X = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$  and  $AX = B = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ . One can consider the matrices  $A$  and  $B$  as perturbed such that the solution is now  $X$ , which is not symmetric. The solution using (8) or (10) is  $X = \begin{bmatrix} 1/2 & 1/3 & 1/2 \\ 1/3 & 1 & 1/3 \\ 1/2 & 1/3 & 1/2 \end{bmatrix}$ , and  $A^\dagger B = \begin{bmatrix} 1/2 & 1/2 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 1/2 & 1/2 \end{bmatrix}$ . Plugging in  $\rho = 1$  in (11) gives the solution  $\begin{bmatrix} 1/2 & 3/8 & 1/2 \\ 1/4 & 1 & 3/8 \\ 1/2 & 1/4 & 1/2 \end{bmatrix}$ , which is not symmetric. Plugging in  $\rho = 10000$  gives the desired result. Increasing the value of  $\rho$  will gradually make the solution more symmetric.

Below, we analyze this in more detail using the standard form transformation after [11], see also [18]. First, we split the solution in two parts;

$$\mathbf{g} = \mathbf{g}_A + \mathbf{g}_B,$$

where  $\mathbf{g}_A \in \mathcal{A}$  represents the admissible part with  $L\mathbf{g}_A = 0$  and  $\mathbf{g}_B \in \mathcal{B}$  is the remainder. Here, we take  $\mathcal{B}$  to be the  $Q$ -orthogonal complement of  $\mathcal{A}$ , meaning that  $\mathbf{g}_A^T Q^T Q \mathbf{g}_B = 0$ . We introduce the corresponding oblique projection operators  $L_Q$  and  $P_Q$ . Given an orthonormal basis,  $A$ , for  $\mathcal{A}$ , these are explicitly given by

$$P_Q = A(QA)^\dagger Q, \quad L_Q = I - P_Q.$$

The problem then decomposes in two parts

$$\mathbf{g}_A = A \mathbf{g}_A, \quad \mathbf{g}_A = \arg \min_{\mathbf{g}} \|QA\mathbf{g} - \mathbf{p}_A\|_2^2, \quad (12)$$

$$\mathbf{g}_B = L \mathbf{g}_B, \quad \mathbf{g}_B = \arg \min_{\mathbf{g}} \|QL_Q L^\dagger \mathbf{g} - \mathbf{p}_B\|_2^2 + \rho \|\mathbf{g}\|_2^2, \quad (13)$$

where  $\mathbf{p}_A \in \mathcal{R}(QP_Q)$  and  $\mathbf{p}_B \in \mathcal{R}(QL_Q)$ . First note that we may replace  $\mathbf{p}_A$  in (12) by  $\mathbf{p}$  without changing the solution. The solution to (12) is given by  $\mathbf{g} = (A^T Q^T Q A)^\dagger A^T Q^T \mathbf{p}$  and hence  $\mathbf{g}_A = A(A^T Q^T Q A)^\dagger A^T Q^T \mathbf{p}$ . Thus,  $\mathbf{g}_A$  coincides with the solution of (10). If the system  $Q\mathbf{g} = \mathbf{p}$  has an

admissible solution we have  $\mathbf{p}_B = 0$  and any non-zero value of  $\rho$  will suffice to suppress non-admissible solutions. The same argument holds, when all non-admissible solutions are in the null-space of  $Q$ . If, like in our example, there is no symmetric solution, then  $\rho$  has to be chosen large enough to make  $\mathbf{g}_B = 0$ . This means that  $\rho > \sigma_1(Q L_Q L^\dagger)$ , where  $\sigma_1(\cdot)$  denotes the largest singular value. In practice, we can not calculate this singular value due to the size of the system. Moreover, the singular values of  $Q L_Q L^\dagger$  are not related to the singular values of  $Q$  and  $L_Q$  or  $L$ . Even an iterative scheme like the power method does not work, because we can not compute the matrix  $L_Q$ , nor are matrix-vector multiplications available.

## 4 Numerical experiments

Our numerical experiments are carried on wavefields generated for the subsurface model presented in [12]. We show the subsurface model in figure (8). We discuss the deconvolution problem for two different source to receiver sampling ratios, namely 1:1 and 1:4. The 1:4 source to receiver ratio is most realistic, but we have added the other setup to investigate the effects of undersampling. For our experiments we always have 151 receivers spread 20 meters apart and we have 512 time samples. Depending on the source to receiver ratio, we either have 151 or 38 sources.

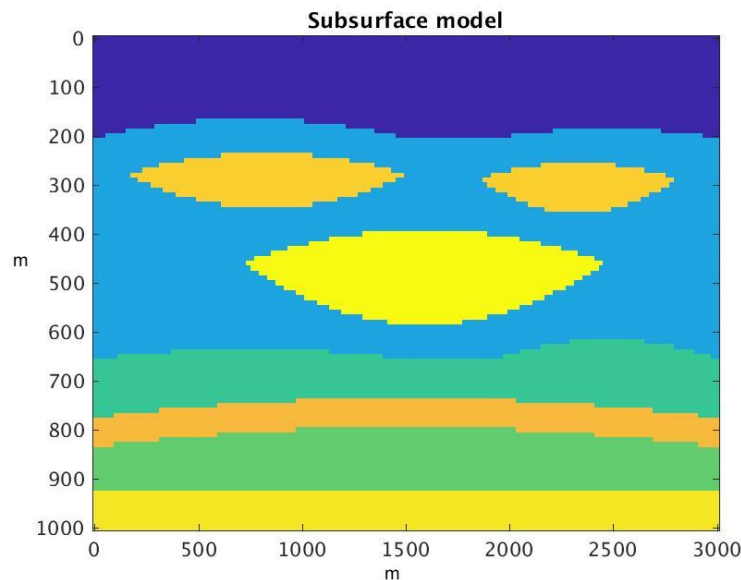


Figure 8: The subsurface model. The sources and receivers are redatumed at 680 m.

### 4.1 The wavefields and the impulse response

In figures (9) and (10) we show the wavefields and the impulse response that has to be inverted for, both for the center source. Both wavefields contain modeling errors that have to be accounted for. We clearly see some ringing artefacts and the wavefield  $P$  clearly contains some modeling errors near the boundary. We also see that the wavefields for the 1:4 sampling clearly suffers from limited source sampling. The data have been modeled by the Full Wavefield Modeling scheme (FWMod) [3]. We benchmark our results against the impulse response obtained in [12], which is obtained by nonlinear inversion. We will consider this the true impulse response. For further details we refer to [12].

Firstly, it is important to investigate whether the problem is actually ill-posed. It is clear that the system

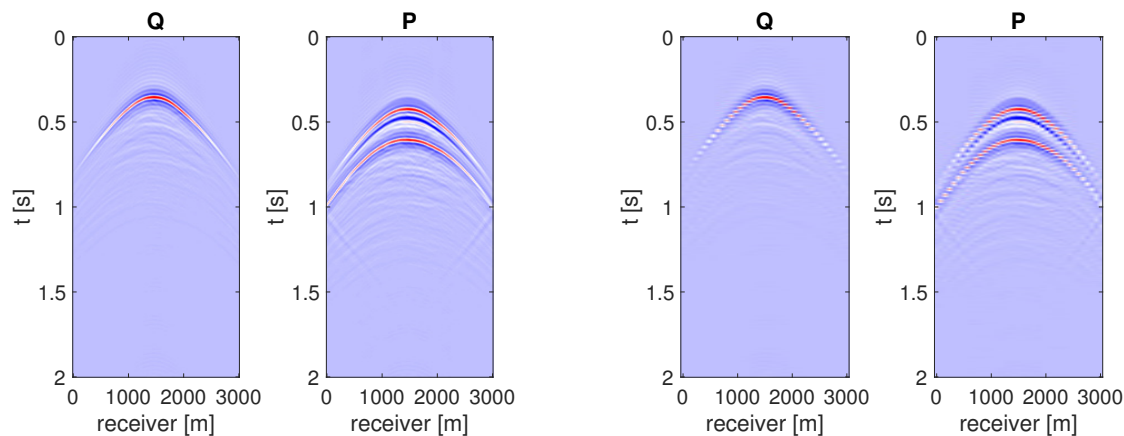


Figure 9: The wavefields for the different sampling scenarios.

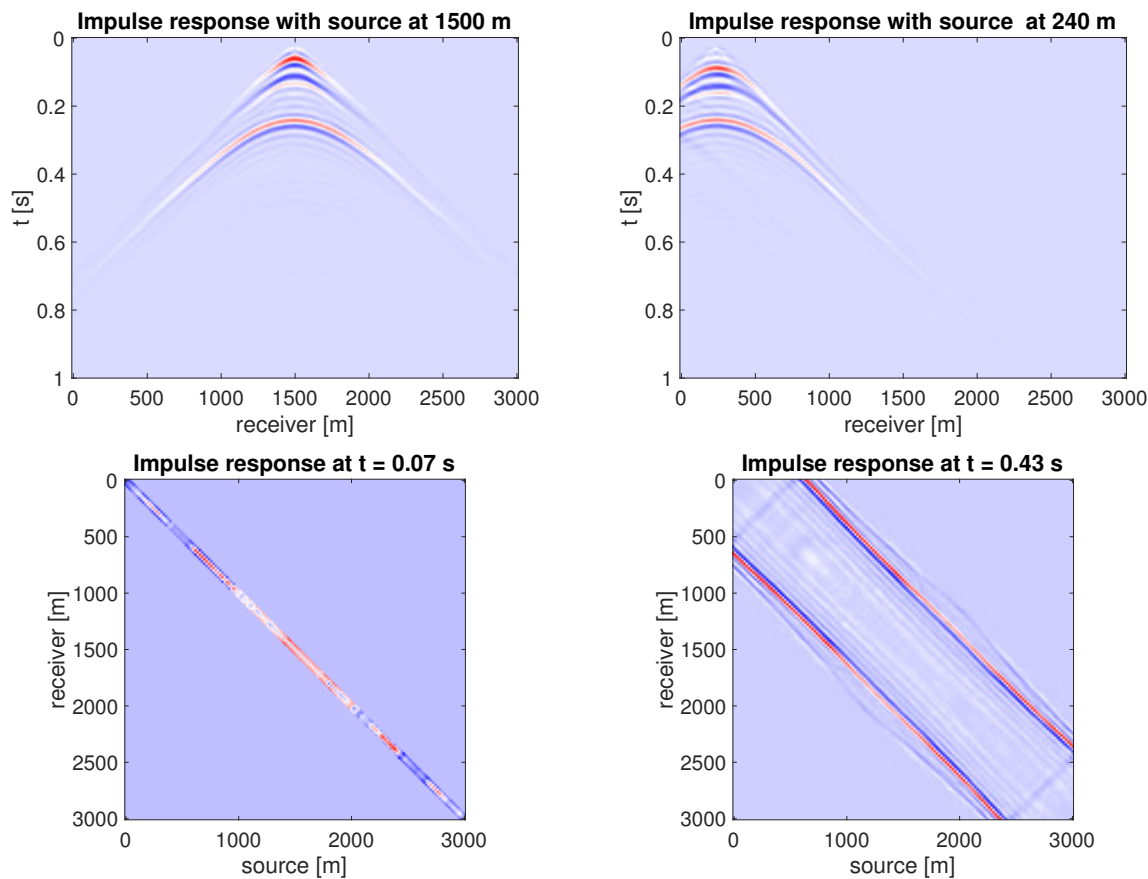


Figure 10: The impulse response in two different domains.

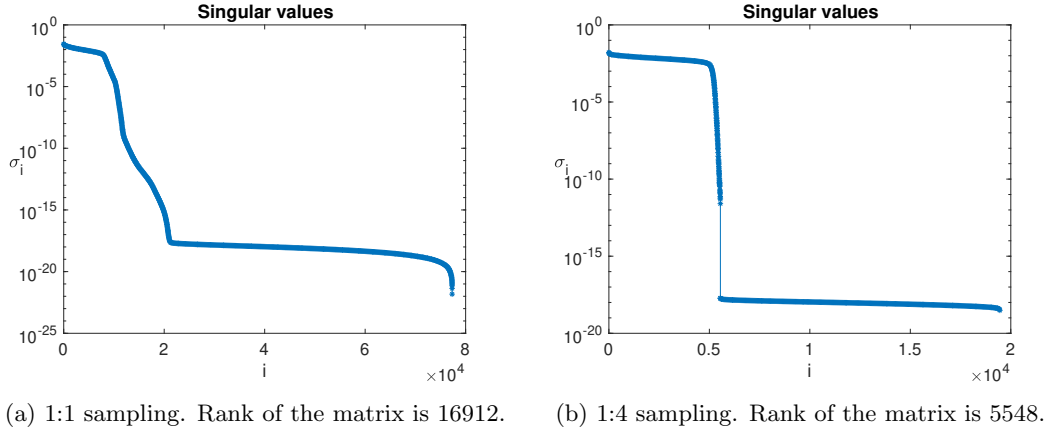


Figure 11: Singular values for different sampling scenarios.

for the 1:4 source to receiver ratio is ill-posed, due to the fact that they lead to underdetermined systems. The 1:1 source to receiver ratio leads to a square systems which may be well-posed. It turns out it is not, as can be seen from the singular values, which we show in figure (11a). From figure (11a) we see that for the 1:1 source to receiver sampling the system is rank deficient, which means that additional regularization is required. As we have stated before, it is not given that the symmetry constraint and the time window will be sufficient prior information to regularize the problem. From figure (11b) we see that even in the under-sampled case the problem is rank deficient. However, in this case the very small singular values correspond to frequencies that have very little impact anyway.

	1:4 sampling	1:1 sampling
Preconditioned	0.17	0.14
Tikhonov	0.17	0.14
KKT	0.13	0.10

Table 1: Relative error for three different optimization strategies.

## 4.2 Results for different optimization strategies

In this section we discuss the results for the different optimization strategies for the two different sampling scenarios. We compare the results to the true solution using the relative error, given by

$$\text{err} = \frac{\|\mathbf{g} - \mathbf{g}_{\text{true}}\|_2}{\|\mathbf{g}_{\text{true}}\|_2}.$$

We present the errors in table (1). We solve the Tikhonov system and the preconditioned system using LSQR. The KKT system is solved using MINRES. The Tikhonov approach and the preconditioned system have almost identical solutions but the KKT solution is slightly better. Interestingly, although the wavefields for the 1:4 sampling ratio contains much less information, the error does not increase much. The reported results are optimal in the sense that we have chosen the amount of iterations that gives the lowest error, by comparing the solution with the true solution.

## 4.3 Semiconvergence

The projection constraints do not stabilize the solution and semiconvergence may still be observed. In figure (12) we show the semiconvergence for the three different optimization strategies for the 1:4 sampling ratio. The Tikhonov and KKT approach show interesting convergence behavior where the error goes down in stages. The convergence is much slower than for the preconditioned system, but the area around the minimizer is flat as opposed to the minimizer of the preconditioned system. The Tikhonov approach requires half the iterations of the KKT system. Due to the slower convergence and the plateau the solution of the Tikhonov approach and the KKT system seem more stable.

## 4.4 Additional Tikhonov regularization

In order to stabilize the solution, and possibly even improve accuracy, we can use an additional Tikhonov penalty. This leads to the following systems:

$$\min_{\mathbf{g}} \|QP_{ST} \mathbf{g} - \mathbf{p}\|_2^2 + \lambda \|\mathbf{g}\|_2^2 \quad (14)$$

$$\min_{\mathbf{g}} \|Q \mathbf{g} - \mathbf{p}\|_2^2 + \|L_{ST} \mathbf{g}\|_2^2 + \lambda \|\mathbf{g}\|_2^2 \quad (15)$$

$$\begin{bmatrix} Q^T Q + \lambda I & L_{ST} \\ L_{ST} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{g} \\ \mathbf{z} \end{bmatrix} = \begin{bmatrix} Q^T \mathbf{p} \\ 0 \end{bmatrix} \quad (16)$$

To check whether this approach is effective we first determine the optimal  $\lambda_{\text{opt}}$ , given by

$$\lambda_{\text{opt}} = \arg \min_{\lambda} \|\mathbf{g}_{\lambda} - \mathbf{g}_{\text{true}}\|,$$

where  $\mathbf{g}_{\lambda}$  is a solution a given system for a given  $\lambda$ . The optimal  $\lambda$  and the associated error are reported in table (2). The errors decrease for both sampling scenarios and all three optimization strategies, although the improvement for the KKT system is not as big as for the other two. In figure (13) we show the stabilizing effect of Tikhonov regularization, which is in line with the results shown in [16] and [8]. Interestingly, the optimal  $\lambda$  is not necessarily the one that leads to a stable solution. The Tikhonov system and the KKT system show similar behavior.

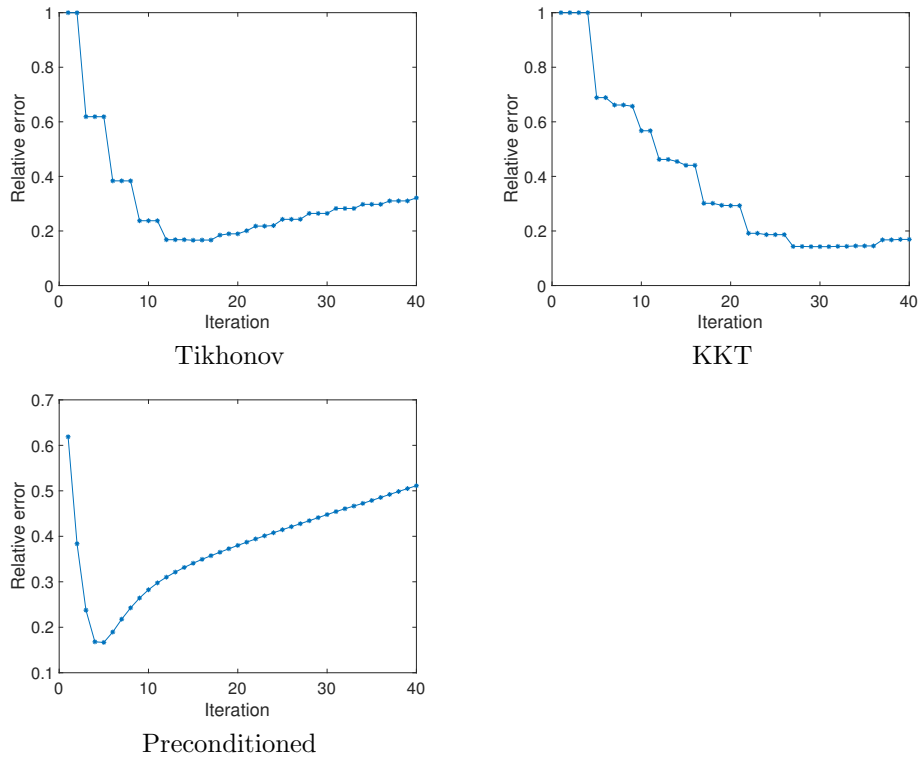


Figure 12: Semiconvergence for three different optimization strategies for the 1:4 sampling ratio.

## 4.5 Parameter selection methods

In order to make the regularization algorithm useful in practice we have to provide a parameter selection rule. There are two types of parameter selection rules: methods that require an estimate of the noise level and methods that do not. The latter are sometimes referred to as heuristic methods. Here, noise is identically, independently distributed white noise. Such noise is not present in our data. We are dealing with noise, but the noise comes from modeling errors. Therefore, we have to rely on heuristic parameter selection methods to give an estimate of the regularization parameter. Parameter selection rules are designed for the preconditioned system and the Tikhonov system, but not for the KKT system. The Tikhonov system has to be modified, by defining the operator  $\begin{bmatrix} Q \\ L_{ST} \end{bmatrix}$  and the data  $\begin{bmatrix} \mathbf{p} \\ 0 \end{bmatrix}$ . We will use the Lanczos process to obtain a low dimensional subspace that is invariant with respect to  $\lambda$  and which allows us to evaluate the parameter selection methods efficiently. We briefly describe some heuristic parameter selection methods. For clarity of presentation we change notation, and now assume that we are solving a system

$$\mathbf{g}_\lambda := \min_{\mathbf{g}} \|\mathbf{Q} \mathbf{g} - \mathbf{p}\|_2^2 + \lambda \|\mathbf{g}\|_2^2.$$

We define

$$\mathbf{r}_\lambda := \|\mathbf{Q} \mathbf{g}_\lambda - \mathbf{p}\|_2.$$



	1:1 sampling		1:4 sampling	
	Relative error	$\lambda_{\text{opt}}$	Relative error	$\lambda_{\text{opt}}$
Preconditioned	0.08	$1.1 \cdot 10^{-5}$	0.12	$3.1 \cdot 10^{-6}$
Tikhonov	0.08	$1.1 \cdot 10^{-5}$	0.12	$3.1 \cdot 10^{-6}$
KKT	0.08	$1 \cdot 10^{-5}$	0.12	$2.6 \cdot 10^{-6}$

Table 2: Reconstruction for three different optimization strategies using additional Tikhonov regularization.

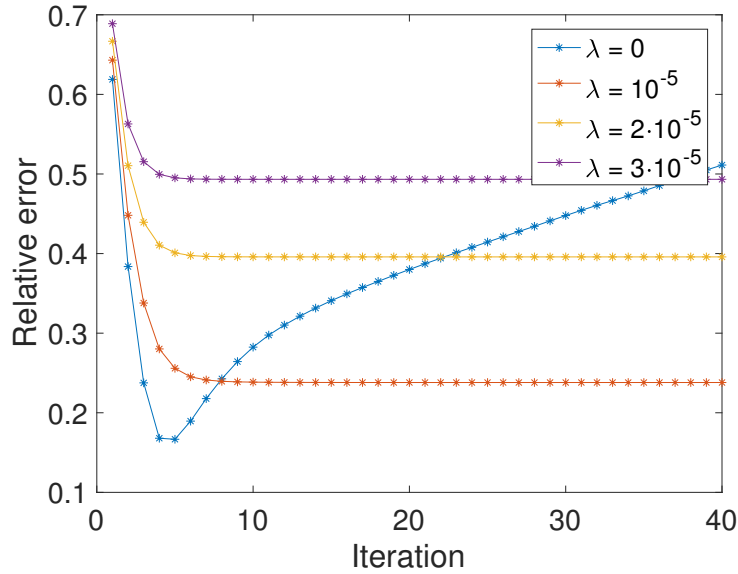


Figure 13: Reconstruction for different  $\lambda$ .

#### 4.5.1 Reginska's rule

Reginska's rule [25] is a variant of the L-curve [17], also known as the Pareto curve. It estimates the optimal  $\lambda$  as the minimizer of

$$\lambda_{\text{Reginska}} = \min_{\lambda} \|\mathbf{g}_{\lambda}\|^2 \|\mathbf{r}_{\lambda}\|^2.$$

We choose Reginska's rule over the L-curve because it is easier to evaluate. A relation between the  $\lambda$  estimated by Reginska's rule and the  $\lambda$  estimated by the L-curve can be found in [25].

#### GCV

The Generalized Cross Validation [33], [13] is a parameter selection methods that estimates the optimal  $\lambda$  as the minimizer of

$$\lambda_{\text{GCV}} = \min_{\lambda} \frac{\|\mathbf{r}_{\lambda}\|^2}{\text{tr}(Q(Q^T Q + \lambda I)^{-1} Q^T)}.$$

The denominator of the GCV can be seen as a measure for the degrees of freedom of the system. If  $A$  is large it is costly to evaluate the trace. Therefore, it has been proposed in [13] to use a randomized trace

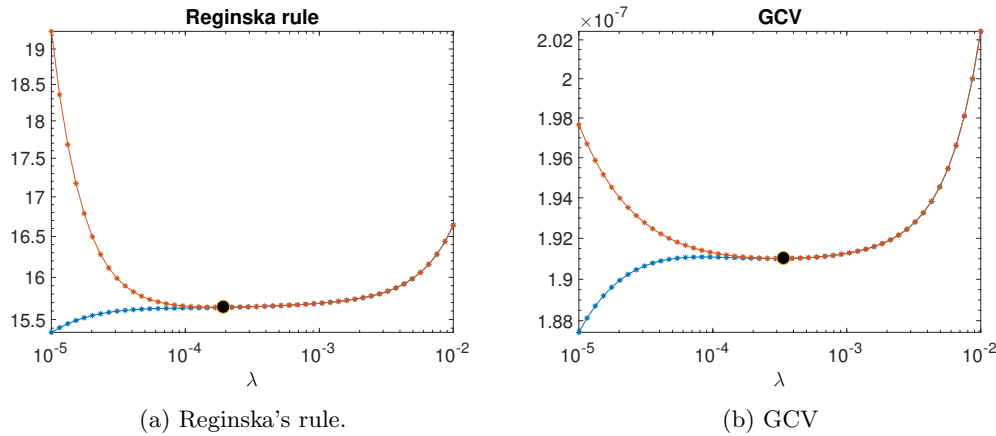


Figure 14: Lower and upper bounds for Reginska's rule and GCV for a standard problem. The black dot indicates the minimizer which is an estimate for  $\lambda$ .

estimator instead. The trace is estimated by

$$\mathbf{u}^T Q(Q^T Q + \lambda I)^{-1} Q^T \mathbf{u},$$

where  $\mathbf{u}$  is a vector whose entries are drawn from the Rademacher distribution. For more information on randomized trace estimation we refer to [20].

#### 4.5.2 Lower and upper bounds

Using the Lanczos bidiagonalization process, which is also the basis for LSQR, we can obtain lower and upper bounds for these parameter selection methods. The upper and lower bounds are calculated using a low dimensional approximation to  $Q$ , which makes them cheap to evaluate. For details we refer to [14], [15], [5], [6]. The difference between the upper and lower bound indicates how close we are to the true value of the parameter selection method. To show what the lower and upper bounds should look like we show them for a standard problem in figure (14). We can see from figure (15) that the parameter selection methods fail to give an approximation to the regularization parameter. Although Tikhonov regularization does improve the accuracy of the reconstruction, we can not estimate the correct regularization parameter.

### 4.6 Nonlinear regularization

Instead of using Tikhonov regularization we could also use additional nonlinear regularization. This excludes the use of the KKT system, as it can only deal with linear (in)equality constraints and we can not use it in combination with nonlinear regularizers. However, the systems (10) and (11) allow for additional nonlinear regularization such that the constraints are still satisfied. The system (10) is preferred, as we have shown that the solutions to the two systems are equal and it is easier to use. We see two possible nonlinear regularizers. The first is an  $\ell_1$  penalty on the impulse response in the curvelet domain, which has shown to be a good basis to represent seismic data [7], [19]. Another possibility is low rank minimization. Seismic data is shown to have low rank in the midpoint-offset domain [1], [22], which can be enforced by penalizing the nuclear norm of the impulse response in the midpoint-offset domain.

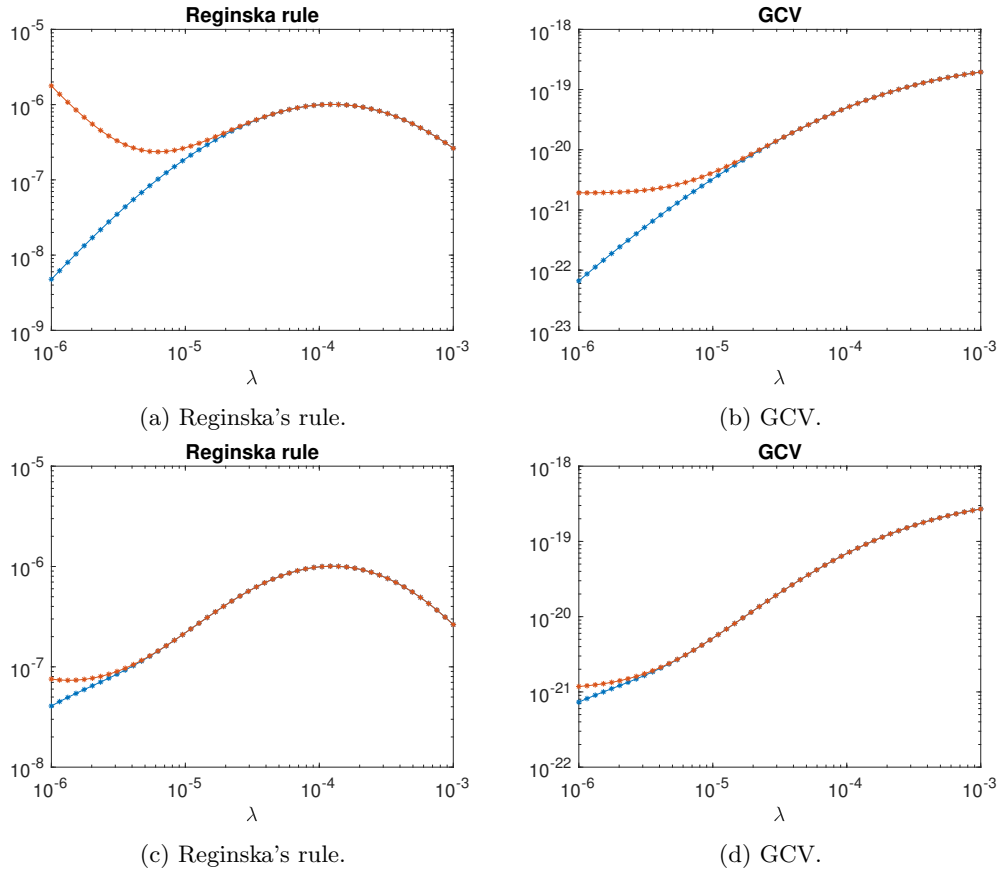


Figure 15: Both parameter selection methods fail to have a minimizer estimating the optimal  $\lambda$  for the 1:4 source sampling. The red line shows the upper bound and the blue line shows the lower bound. The top row is for the Tikhonov system and the bottom row for the preconditioned system.

## 5 Conclusion, discussion and outlook

In this paper we have discussed deconvolving two wavefields to obtain the impulse response in the context of redatuming. We have discussed two constraints that the impulse response has to satisfy and we have shown that they are associated with orthogonal projection operators and a related penalty operator. We have shown three different optimization methods that incorporate the constraints and have shown that they are equivalent in a certain sense. Incorporating the constraints as a penalty using a generalized Tikhonov approach should be avoided, as it is inferior to the preconditioned system in terms of computational time and simplicity. The KKT system was superior in our numerical experiments. However, it is computationally more expensive and it may be difficult or impossible to incorporate additional regularization. Lastly, we have shown that the constraints for the impulse response do not have a stabilizing effect on the solution and that additional regularization is necessary.

We have shown that Tikhonov regularization can be used to improve the reconstruction. However, we have seen that the parameter selection methods Reginska's rule and GCV are not able to estimate the optimal  $\lambda$ . Moreover, we have to use heuristic parameter selection methods because the noise level is not known and, perhaps more importantly, the noise consists of modeling errors and is not Gaussian and independently and identically distributed.

To solve the MDD problem we have applied standard techniques and theory from the inverse problem literature. The theory is developed for linear systems where the data have been generated by a known forward model. For the MDD problem we deal with two datasets in the form of wavefields that are related by a convolution with the impulse response. Now, both datasets contain noise and a strict model and data separation is artificial. An approach like Total Least Squares (TLS) [31], where both model and data are assumed to be noisy, seems more natural. Specifically, Restricted Total Least Squares [32], where equality constraints can be taken into account, seems the most natural approach for the MDD problem. However, the solution to this problem is given by the Restricted Singular Value Decomposition, which can not be computed for the large matrices arising in MDD problems.

Finally, we could add a regularization filter that filters out the non-recoverable modes. As shown in our example, see figure (5), the support of  $\hat{q}$  determines how much of  $\hat{g}$  can be recovered. We can stabilize the reconstruction by restricting the support of  $\hat{g}$  to the support of  $\hat{q}$ . Anything outside of this support can not be recovered and can be considered as noise.

## 6 Acknowledgements

The first author is sponsored by the DELPHI consortium and would like to thank the sponsors for their support. The second author is financially supported by the Netherlands Organization for Scientific Research (NWO) as part of research programme 613.009.032. The authors would like to thank Dr. Eric Verschuur and Aayush Garg for describing to them the MDD problem and providing them with the dataset. The authors would like to thank Dr. Ajinkya Kadu for pointing them to the KKT system and for fruitful discussions.

## References

- [1] A. Aravkin, R. Kumar, H. Mansour, B. Recht, and F.J. Herrmann. Fast methods for denoising matrix completion formulations, with applications to robust seismic data interpolation. *SIAM Journal on Scientific Computing*, 36:237–266, 2014.
- [2] A.J. Berkhout. *Seismic Migration. Imaging of Acoustic Energy by Wave Field Extrapolation*. Elsevier, Elsevier, 1982.
- [3] A.J. Berkhout. Review paper: An outlook on the future of seismic imaging, part i: forward and reverse modelling. *Geophysical Prospecting*, 62(5):911–930, 2014.
- [4] Christiaan C. Stolk. A pseudodifferential equation with damping for one-way wave propagation in inhomogeneous acoustic media. *Wave Motion*, 40(2):111–121, aug 2004.
- [5] D. Calvetti, G. H. Golub, and L. Reichel. Estimation of the l-curve via lanczos bidiagonalization. *BIT Numerical Mathematics*, 39(4):603–619, Dec 1999.
- [6] D. Calvetti, G. Spaletta, L. Reichel, and F. Sgallari. An l-ribbon for large underdetermined linear discrete ill-posed problems. *Numerical Algorithms*, 25:89–107, 2000.
- [7] Emmanuel J. Cands and Laurent Demanet. The curvelet representation of wave propagators is optimally sparse, 2004.
- [8] Julianne Chung, James G. Nagy, and Dianne P. O’Leary. A weighted gcv method for lanczos hybrid regularization. *Electronic Transactions on Numerical Analysis*, 28:149 – 167, 2008/// 2008.
- [9] Jon F. Claerbout. TOWARD A UNIFIED THEORY OF REFLECTOR MAPPING. *GEOPHYSICS*, 36(3):467–481, jun 1971.
- [10] F.J.H. Don. On the symmetric solutions of a linear matrix equation. *Linear Algebra and its Applications*, 93:1–7.
- [11] Lars Eldén. A weighted pseudoinverse, generalized singular values, and constrained least squares problems. *BIT*, 15:487–502, 1982.

- [12] A. Garg, S. Sharma, and D.J. Verschuur. Reservoir elastic parameters estimation from surface seismic data using jmi-res: A full-wavefield approach. *80th EAGE Conference and Exhibition 2018*, 2018.
- [13] Gene H. Golub, Michael Heath, and Grace Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21(2):215–223, 1979.
- [14] Gene H. Golub and Urs von Matt. Generalized cross-validation for large scale problems. *J. Comput. Graph. Stat*, 6:1–34, 1995.
- [15] Gene H. Golub and Urs von Matt. Tikhonov regularization for large scale problems, 1997.
- [16] Martin Hanke. On lanczos based methods for the regularization of discrete ill-posed problems. *BIT Numerical Mathematics*, 41(5):1008–1018, Dec 2001.
- [17] Per Christian Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review*, 34:561–580, 1992.
- [18] Per Christian Hansen. Oblique projections and standard-form transformations for discrete inverse problems. *Numerical Linear Algebra With Applications*, 20:250–258, 2013.
- [19] F.J. Herrmann and G. Hennenfent. Non-parametric seismic data recovery with curvelet frames. *Geophysical Journal International*, 173:233–248.
- [20] M.F. Hutchinson. A stochastic estimator for the trace of the influence matrix for laplacian smoothing splines. *Communications in Statistics, Simulation and Computation*, 19:433–450, 1990.
- [21] Eric Chu King-wah. Symmetric solutions of linear matrix equations by matrix decompositions. *Linear Algebra and its Applications*, 119:35–50, 1989.
- [22] Rajiv Kumar, Curt Da Silva, Okan Akalin, Aleksandr Y. Aravkin, Hassan Mansour, Benjamin Recht, and Felix J. Herrmann. Efficient matrix completion for seismic data reconstruction. *Geophysics*, 80:97–114, 2015.
- [23] C.C. Paige and M.A. Saunders. Solution of sparse indefinite systems of linear equations. *SIAM J. Numerical Analysis*, 12:617–629, 1975.
- [24] Thomas Planès, Roel Snieder, and Satyan Singh. Model-based redatuming of seismic data: An inverse-filter approach. *GEOPHYSICS*, 83(2):1–13, 2 2018.
- [25] T. Reginska. A regularization parameter in discrete ill-posed problems. *SIAM Journal for Scientific Computing*, 17, 1996.
- [26] Gerard T. Schuster and Min Zhou. A theoretical overview of model-based and correlation-based redatuming methods. *Geophysics*, 71(4):130–110, 07 2006.
- [27] W W Symes. The seismic reflection inverse problem. *Inverse Problems*, 25(12):123008, dec 2009.
- [28] J. van der Neut, M. Tatanova, J. Thorbecke, E. Slob, and K. Wapenaar. Controlled source interferometric redatuming by crosscorrelation and multidimensional deconvolution in elastic media. *Geophysics*, 76(4), 2011.
- [29] J. van der Neut, M. Tatanova, J. Thorbecke, E. Slob, and K. Wapenaar. Deghosting, demultiple, and deblurring in controlled-source seismic interferometry. *International Journal of Geophysics*, 2011.
- [30] Joost van der Neut and Felix J. Herrmann. Interferometric redatuming by sparse inversion. *Geophysical Journal International*, 192:666–670, 02 2013.
- [31] Sabine van Huffel and Joos Vandewalle. *The total least squares problem: computational aspects and analysis*, volume 9. SIAM, 1991.
- [32] Sabine Van Huffel and Hongyuan Zha. The restricted total least squares problem: Formulation, algorithm, and properties. *SIAM Journal of Matrix Analysis and Applications*, 12(2):292–309, 2012.
- [33] G. Wahba. *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics, Philadelphia, 1990.

- [34] C.P.A Wapenaar. A representation for greens function retrieval by multidimensional deconvolution. *The Journal of the Acoustical Society of America*, 128, 2010.
- [35] Kees Wapenaar. Reciprocity properties of oneway propagators. *GEOPHYSICS*, 63(5):1795–1798, sep 1998.
- [36] Kees Wapenaar, Evert Slob, and Roel Snieder. Seismic and electromagnetic controlled-source interferometry in dissipative media. *Geophysical Prospecting*, 56:419434, 2008.