Universität Bremen

Dissertation

Effective and Efficient Reconstruction Schemes for the Inverse Medium Problem in Scattering

in fulfillment of the requirements for the degree of
Doktor der Ingenieurwissenschaften (Dr.-Ing.)

submitted by

Diplom-Mathematiker (Dipl.-Math.)

Florian Bürgel[◇]

at

University of Bremen	Universität Bremen,
Department 3	Fachbereich 3
(Mathematics and Computer Science)	(Mathematik und Informatik)
Center for Industrial Mathematics	Zentrum für Technomathematik
Germany	Deutschland

in

April 2019.

Colloquium: August 19th, 2019.
Publication: December 2019.

Original supervisor: Prof. Dr. Armin Lechleiter.

Examiners:
1. Dr. Kamil S. Kazimierski, Berlin.
2. Prof. Dr. Thorsten Hohage, University of Göttingen.
3. Prof. Dr. Dr. h. c. Peter Maaß, University of Bremen.

[◇] born on June 10th, 1987 in Nienburg (Weser), Germany, email: fbuergel@uni-bremen.de.

Abstract This thesis challenges with the development of a computational framework facilitating the solution for the inverse medium problem in time-independent scattering in two- and three-dimensional setting. This includes three main application cases: the simulation of the scattered field for a given transmitter-receiver geometry; the generation of simulated data as well as the handling of real-world data; the reconstruction of the refractive index of a penetrable medium from several measured, scattered fields.

We focus on an effective and efficient reconstruction algorithm. Therefore we set up a variational reconstruction scheme. The underlying paradigm is to minimize the discrepancy between the predicted data based on the reconstructed refractive index and the given data while taking into account various structural a priori information via suitable penalty terms, which are designed to promote information expected in real-world environments. Finally, the scheme relies on a primal-dual algorithm. In addition, information about the obstacle's shape and position obtained by the factorization method can be used as a priori information to increase the overall effectiveness of the scheme.

An implementation is provided as MATLAB toolbox IPscatt. It is tailored to the needs of practitioners, e.g. a heuristic algorithm for an automatic, data-driven choice of the regularization parameters is available. The effectiveness and efficiency of the proposed approach are demonstrated for simulated as well as real-world data by comparisons with existing software packages.

Zusammenfassung Das Ziel dieser Arbeit ist die Entwicklung eines Softwarepakets, das die Lösung des inversen zeitunabhängigen Streuproblems für Medien in zwei- und dreidimensionaler Umgebung ermöglicht. Dies betrifft drei Hauptanwendungsfälle: die Simulation des gestreuten Feldes für eine gegebene Sender-Empfänger-Geometrie; die Generierung von simulierten Daten sowie den Umgang mit echten Daten; die Rekonstruktion des Brechungsindexes eines penetrablen Mediums aus mehreren gemessenen, gestreuten Feldern.

Der Fokus liegt auf einem effektiven und effizienten Rekonstruktionsalgorithmus. Dazu entwickeln wir ein variationelles Rekonstruktionsschema. Diesem liegt die Idee zu Grunde, die Diskrepanz zwischen den prognostizierten Daten, die auf dem rekonstruierten Brechungsindex basieren, und den gegebenen Daten zu minimieren und dabei verschiedene strukturelle a priori Informationen durch geeignete Strafterme zu berücksichtigen, die in echten Anwendungen zu erwartende Informationen bevorzugen. Dabei basiert das Verfahren auf einem Primal-Dual-Algorithmus. Zusätzlich können Informationen zu Gestalt und Position des Hindernisses, die man aus der Faktorisierungsmethode erhält, als a priori Information zur Effektivitätsverbesserung verwendet werden.

Eine Implementierung wird als MATLAB Toolbox IPscatt bereitgestellt. Diese ist auf die Bedürfnisse von Anwendern zugeschnitten, da sie beispielsweise einen heuristischen Algorithmus für eine automatische, datengesteuerte Wahl der Regularisierungsparameter enthält. Effektivität und Effizienz des vorgeschlagenen Ansatzes werden für simulierte und echte Daten durch Vergleiche mit bestehenden Softwarepaketen gezeigt.

Keywords Inverse Scattering Problem, Parameter Identification, Helmholtz Equation, Denoising, Sparsity Regularization, Total Variation Regularization, Primal-Dual Algorithm, Factorization Method, MATLAB Toolbox.

Preface

Publications Note that some ideas and figures have already been published. This refers to the article [BKL17], that was submitted in August 2016 and in a revised version in February 2017. Further, this refers to the article [BKL19], that was submitted in October 2017 and in a revised version in August 2018 and in April 2019. This publication also includes *supplementary material*, that contains *source code*, a *source code documentation* and a *user guide*.

In particular, the presentation of the direct scattering problem in Ch. 2 is based on [BKL17] and [BKL19]. Further, Ch. 3 is about the inverse scattering problem and contains material from [BKL17]. The description of the MATLAB toolbox IPscatt, see Ch. 4, and the numerical examples, see Ch. 5, are essentially from [BKL19]. Further, the reconstruction with a primal-dual algorithm in Ch. 7 was already presented in [BKL17].

In addition, the table in App. A containing all involved formulas of the direct scattering problem was already part of the *user guide* from [BKL19]. Further, the main parts of the *source code* were already published as *supplementary material* together with [BKL19]. In particular, that submission did not include: the heuristic algorithm for an automatic, data-driven choice of the regularization parameters; the modification of the reconstruction scheme to use the obstacle's shape and position obtained by the factorization method as a priori information to increase the overall effectiveness. A more detailed description about the origin of the *source code* is given in Ch. 1.

[BKL17] Florian Bürgel, Kamil S. Kazimierski, and Armin Lechleiter. A sparsity regularization and total variation based computational framework for the inverse medium problem in scattering. *Journal of Computational Physics*, 339:1–30, 2017. URL: https://doi.org/10.1016/j.jcp.2017.03.011.

[BKL19] Florian Bürgel, Kamil S. Kazimierski, and Armin Lechleiter. Algorithm 1001: IPscatt—A MATLAB Toolbox for the Inverse Medium Problem in Scattering. *ACM Transactions on Mathematical Software*, 45(4):45:1–45:20, 2019. URL: https://doi.org/10.1145/3328525.

Acknowledgments Without Armin Lechleiter's dedicated support as original supervisor this thesis would not exist. As long as possible he guided me with instinctive mathematical assurance and he willingly answered all my questions. I am incredibly grateful for everything I could learn from him—as a mathematician and as a human being.

I also thank my supervisor Kamil S. Kazimierski for his excellent and patient support via Skype. I suppose we have managed all technical problems Skype offers in the recent years. However, these calls were always imaginative and constructive. Despite the geographical distance we had target-oriented discussions and worked on manuscripts. In particular, he was able to teach me a little bit to write mathematical results as a story.

Also great thanks go to Thorsten Hohage, who agreed to read and rate my work. In particular, I would like to thank him for his useful comments after a talk in Göttingen.

I would also like to thank Peter Maaß for his offer to write a third review. This makes the constellation of two external examiners possible.

The funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) enabled this research and the presentation on conferences under grants Ka 3954/1-1 and Le 2499/2-1.

Furthermore, I thank every current and former member of the Working Group Inverse Problems for helpful discussions about mathematics (and other topics) and a pleasant stay in Bremen—namely, Thies Gerken, Zixian Jiang, Alexander Konschin, Frederieke Miesner, Stefan Peters, Marcel Rennoch, Tobias Rienmüller and Ruming Zhang. Further, thanks go to the work group's secretary Ebba Feldmann—I would have been lost in university's administration stuff without her—, to all room-mates in office for the harmonious time and to the IT experts of the Center for Industrial Mathematics—they kept things running.

Very often, I had to tell friends and family members that I have no time. Thanks for understanding, supporting and the sparse but relaxing common moments.

Contents

1 What is at Issue?

Scattering and absorption are interesting and familiar phenomena, e.g. the sunlight's decomposition in rainbow colors by a prism is a simple scattering example. Further, the bones on a X-ray photograph are very bright because they absorb the X-rays strongly. Apart from electromagnetic waves like light and X-rays, scattering phenomena occur with acoustic waves, e.g. when we speak. This describes a direct scattering problem. Note that it is common to use the term scattering for both physical effects scattering and absorption in general. Therefore the obstacle, that scatters or absorbs the waves, is just called scatterer.

Furthermore, pathologists look inside human bodies for medical diagnosis, but during lifetime it is preferable to examine this medium with non-destructive methods like computerized tomography. This is one example of an application of the inverse medium problem in scattering: the inside of the body, physically represented by the refractive index, is reconstructed from X-ray measurements from different angles. However, the solution does not depend continuously on the measurements. This is the reason why inverse scattering problems are both difficult and interesting.

Direct and Inverse Scattering Problem We focus on the electromagnetic and acoustic scattering based on the Helmholtz equation. This mathematical model of time-independent scattering is known in literature as the *direct scattering problem*, see e.g. [CK13]. Fast numerical solvers for that problem are described for example in [Vai00].

The task of identifying the refractive index of a penetrable medium from the measurements of the waves scattered from that medium is commonly referred to as the *inverse medium problem* or *inverse scattering problem*. In fact, as the refractive index equals one outside the obstacle, it is preferred to identify the contrast, that is the squared refractive index minus one. In particular, this task is encountered in remote-sensing and non-destructive testing applications, e.g. based on ultrasound measurements.

Difficulties of the Inverse Scattering Problem As a special case of a nonlinear, ill-posed (i.e. instable) parameter identification problem it is notoriously difficult to solve. One approach to solve this problem is, roughly speaking, to minimize the discrepancy between the reconstruction and the measured data while, at the same time, taking into account a priori information via a suitable penalty term. A penalty term is required to cope with the ill-posedness of the inverse problem, see [EHN96]. This minimization-based reconstruction approach is known in inverse problems as (nonlinear) *Tikhonov regularization* and is the underlying paradigm for further developments.

Furthermore, the discretization of the problem leads to a large system of equations, which makes the problem even more challenging to treat computationally. In particular, this holds true for problems modeling all three space dimensions.

Algorithm Classes Numerous algorithms solving inverse medium scattering problems are already established, all having specific advantages and disadvantages.

A first class of algorithms (including the proposed one) exploits Fréchet differentiability of the operator mapping the contrast onto the data. This allows to set up one of the various existing variants of regularized Newton-like inversion schemes, that numerically show local convergence, see [CK13, Hoh01]. This class includes popular techniques as Kleinman and van den Berg's CG-based modified gradient method or Gutman and Klibanov's simplified gradient method.

A second class of algorithms implements constructive uniqueness proofs for (features of) the contrast, see e.g. [dLS^{+}16], which also includes so-called quantitative methods, that merely identify spatial sets where the contrast differs from its background values, see [CK13, KG08, KS03].

Finally, a third class of algorithms relies on a high or low frequency assumption to linearize the inversion problem in the corresponding asymptotic regime (e.g. physical optics approximation or Born approximation). This allows to use particular linear inversion methods for inverse medium scattering problems but remains of limited use if the wavelength is in the range of the size of the scattering object.

Whenever an inversion algorithm for the inverse medium scattering problem relies on so-called *multi-static* data, see Fig. 2.1 below, then it is backed up by uniqueness results for the searched-for contrast, both in two and three dimensions, see [Häh96, Buk08].

Preliminary Works There are several preliminary works to the inverse scattering problem in a two-dimensional setting we would like to build upon.

The time-independent scattering of acoustic waves was examined in [Hoh01], in particular the *reconstruction* of the scattering obstacle's contrast from so-called *far field* measurements using the common quadratic penalty for regularization.

In [DDD04] it was shown that ℓ^p penalties with $1 \leq p \leq 2$ (instead of the usual quadratic) still regularizes *linear* inverse problems, such that a so-called *sparsity* constraint can taken into account, e.g. by a *soft-shrinkage operator*[◇]. Further, the usage of *wavelets* is described in this context. Considering sparse refractive indices with respect to pixel or wavelet basis, it is by now well-known that traditionally choosing the square of a Hilbert space norm yields worse results than choosing ℓ^p-norms for $p \in [1, 2)$ close to one, see [LKK13].

In [LKK13] the *thresholded, nonlinear Landweber*[†] scheme, which bases on the soft-shrinkage operator taking into account sparsity in pixel basis, is described for *nonlinear* inverse medium scattering problems for so-called *near field* data. Further, the existence and uniqueness of the nonlinear inverse scattering problem's solution is shown with exception of the most interesting case of $p = 1$, that enforces sparsity. In addition, the numerical examples do not consider complex-valued contrasts.

The application of sparsity promoting reconstruction methods to the inverse problem are also evaluated in [Str14] using the *linearized* model given by the Born approximation of the scattering process and taking into account sparsity in the wavelet basis.

[◇] The *soft-shrinkage operator* is called *shrinkage operator* in [DDD04].

[†] The *thresholded Landweber* is called *shrunk Landweber* in [LKK13].

At Present Including further a priori information of the contrast further improves the reconstruction quality. Coping with additional penalty terms (in addition to sparsity) requires to use a more general minimization algorithm than the thresholded, nonlinear Landweber used in [LKK13]. To this end, this thesis is motivated by developing an effective and efficient minimization-based reconstruction scheme to solve the inverse medium problem in scattering (including complex-valued contrasts) at a fixed frequency in a two- and three-dimensional setting.

A first approach is to extend the soft-shrinkage operator to additionally respect *physical bounds* of the real and imaginary part of the contrast, see Ch. 6.

In addition to enforce sparsity in the pixel basis and respect physical bounds via suitable penalty terms, we would like to enforce sharp edges. For the latter one we use the *total variation*, that is a well-known denoising method in digital image processing, see e.g. [CP11a]. These penalty terms take into account structural a priori information. They are designed to promote information expected in real-world environments and have a combined effectiveness. To the author's best knowledge here the first combination of sparsity promoting and total variation-based regularization is shown to jointly improve the reconstruction quality for the inverse medium problem. However, we need a more sophisticated method than the Landweber scheme to include all these penalty terms. In addition, the Landweber iteration is not an efficient method. Therefore in Ch. 7 we develop a variational reconstruction scheme that relies on the so-called primal-dual algorithm due to Pock, Bischof, Cremers and Chambolle, see [PCBC09, CP11a]. It turns out that this algorithm is efficient. Therefore we can effectively and efficiently solve the inverse medium problem in time-independent scattering in a two- as well as three-dimensional setting within an appropriate run time.

In addition, information about the obstacle's shape and position obtained by the factorization method can be used as a priori information to increase the overall effectiveness of the scheme, see Ch. 8. Furthermore, the factorization method facilitates a reconstruction scheme without any assumptions to the contrast.

In Ch. 5 we demonstrate the effectiveness and the efficiency of the developed reconstruction scheme on several, representative examples for simulated data as well as real-world data from Institute Fresnel. In particular, we compare the new scheme with two aforementioned methods: the thresholded, nonlinear Landweber regularization, see [LKK13], and the reconstruction scheme using the Born approximation, see [Str14].

An implementation of the proposed reconstruction scheme is provided as MATLAB toolbox IPscatt. The toolbox has three main application cases: simulation of the scattered field for a given transmitter-receiver geometry; the generation of simulated data as well as the handling of the real-world data from Institute Fresnel; the reconstruction of the contrast from several measured, scattered fields. The architecture and key features of IPscatt are presented in Ch. 4. Note as a preliminary remark that it does not matter, if simulated data, real-world data from Institute Fresnel or another real-world data set is used for the reconstruction of the contrast.

So, the ingredients of this thesis are mathematical physics, mathematics and programming.

Source Code Archeology Note that the toolbox that was used for numerical results in [LKK13] was written by Lechleiter and Kazimierski and is a predecessor of IPscatt. In particular, it was a great basis to solve the direct scattering problem and for handling of the real-world data from Institute Fresnel. Although in [LKK13, Sec. 5] it is mentioned that the imaginary component of the contrast is set to zero in each step, the provided code basis already included the use of the shrinkage separately to the real and imaginary part. However, the code basis was restricted to a two-dimensional setting and an experimental set-up with transmitters and receivers arranged on circles. Therefore the structure of the *source code* has been fundamentally changed to include all new features: in particular, more standard transmitter-receiver geometries, the extended version of the soft-shrinkage operator to improve the thresholded, nonlinear Landweber regularization, the extension to a three-dimensional setting and the implementation of the new, developed variational reconstruction scheme. Further, the already-mentioned usage of the information of the factorization method and an automatic, data-driven choice of the regularization parameters were implemented, where the implementation of the factorization method itself was provided by Armin Lechleiter. In addition, a *source code documentation* and a *user guide* were created to increase the usability. The latter contains an installation guide, a technical description and a set of hands-on best practice recommendations and guides, e.g. how to simulate data avoiding the "inverse crime". The *source code*, the *source code documentation* and the *user guide* are available as *supplementary material*[◇].

In addition, the source code from Lechleiter and Kazimierski also included a thresholded, nonlinear Landweber enforcing sparsity in wavelet basis instead of pixel basis. The results did not convince in comparison to sparsity in pixel basis and were not presented in [LKK13]. Based on this it was experimented with a weighting of the result of the discrete wavelet transform. This did also not result in better reconstructions. Therefore this was not a subject of ongoing research. However, the tested improvements are still part of IPscatt.

Furthermore, the developed variational reconstruction scheme was also implemented for sparsity in wavelet basis. Note that this version does not include the total variation-based regularization. Once again it did not outcompete the sparsity in pixel basis. Therefore the corresponding results are not part of this thesis. In addition, IPscatt provides to vary the exponent of the discrepancy term. For the exponents 1 and 2 this was done by the author. A generalization was implemented by Armin Lechleiter. However, the default choice of 2 as exponent coincides with the so-called Tikhonov functional and is recommended.

Applications However, both features (sparsity in wavelet basis and various exponents) were already used. The modular and flexible design of IPscatt facilitates its use. Indeed, a draft version of IPscatt was already modified by Lechleiter and Rennoch to tackle inverse scattering from so-called *anisotropic* penetrable media instead of *isotropic* penetrable media, see [LR17] or [Ren17, Ch. 3]. This required two changes: the replacement of the forward operator corresponding to the anisotropic Helmholtz equation and the computation and implementation of the adjoint of the forward operator's linearization. For the reconstruction the thresholded, nonlinear Landweber and the developed variational reconstruction scheme were used out of the box. In the theoretical part, they show that

---

[◇] Supplementary material is provided on an enclosed CD and on http://www.fbuergel.de/ipscatt/ in the latest version.

sparsity in wavelet basis regularizes the solution. This is more interesting than sparsity in pixel basis. Consequently, they use sparsity in wavelet basis in the numerical examples too. In addition, they tried out different exponents of the discrepancy term.

Analogously, the modular and flexible design of IPscatt facilitates to replace the used numerical solver from [Vai00] for the direct scattering problem. This requires two changes (if the proposed reconstruction scheme is used): the replacement of the forward operator and the forward operator's linearization (with the Fréchet derivative).[◇] This opens up the opportunity to replace the used numerical solver by more sophisticated algorithms, see e.g. [BGG17, Yin15]. In particular, this is interesting in the case of relatively high wave numbers. Note that IPscatt tackles the *single-frequency* inverse scattering problem in contrast to [BGG17], that considers the *multi-frequency* case.

Notation We denote standard Lebesgue spaces with integrable pth power on a domain D by $L^p(D)$. The corresponding Sobolev spaces are $W^{s,p}(D)$ for $s \geq 0$ and $p \in [1, \infty]$; for $p = 2$, we abbreviate $H^s(D) = W^{s,2}(D)$. The notation of inner products is $\langle \cdot, \cdot \rangle$. In Ch. 2 vector and array quantities are denoted by an underlined symbol. To select an element in such quantities we employ a subindex or, where it is beneficial for readability, a bracket, i.e. $\underline{q}_j$ and $\underline{q}(j)$ both denote the jth element of $\underline{q}$. Sometimes we use the notation $f \cdot g$ to point out the pointwise multiplication of functions. The notation $(f\cdot)$ is used to denote the operator of pointwise multiplication (with a function f). In the same vein, $\underline{f} \odot \underline{g}$ and $(\underline{f} \odot\ )$ are used for the element-wise multiplication of vector, matrix or array quantities as well as the related operator. Of course, for vectors the notation $(\underline{f} \odot\ )$ is the same as the diagonal operator $\mathrm{diag}(\underline{f})$. Further, except it is explicitly stated otherwise, I denotes the identity operator on an appropriate space.

[◇] Note that IPscatt was not specifically designed for these replacements. However, the existing parameter `seti.model`, that is used to distinguish between the Helmholtz equation in two and three dimensions (via `helmholtz2D` and `helmholtz3D`), helps with this task.

2 Direct Scattering Problem

In classical scattering theory are two basic problems: scattering of time-independent$^\diamond$ *acoustic* or *electromagnetic* waves by a *penetrable inhomogeneous medium of compact support* and by a *bounded impenetrable obstacle*, see [CK13, Sec. 1.1]. We will concentrate on the case of scattering of acoustic waves by penetrable inhomogeneous media in two and three dimensions.

The details of this kind of scattering are described in [CK13, Ch. 8]. Fast numerical solvers of the direct scattering problem are described in [Vai00]. This solver was used in [LKK13] with several improvements. A detailed and specific description of these improvements was published in [BKL17, Sec. 3], where the provided research notes to the toolbox, that was used for the numerical examples in [LKK13] and was written by Kazimierski and Lechleiter, were a good base. Therefore they are not repeated in detail here. However, a brief description is given to get familiar with the used notation, to refer to some formulas, point out the personal contribution and correct some mistakes.

In this section we describe the mathematical model of time-independent scattering, i.e. the *direct scattering problem*. It consists of three steps: initially, a *transmitter* propagates a field into a region containing the *obstacle*. Then that obstacle, which is expressed by the *contrast*

$$q := n^2 - 1 \quad \text{with } n \text{ as refractive index}, \tag{2.1}$$

scatters that field in that region. Finally, the scattered field is propagated towards *receivers*, which measure it.

In *mono-static* setting the related operator maps a contrast to the resulting near or far field operator. In the *multi-static* case the analogue is considered, i.e. the mapping of the contrast to the field operators for *all* transmitters, cf. Fig. 2.1. We denote in both cases the involved *forward operator* by $\mathcal{F}$ and concentrate on the multi-static case in our formulations. We will also discuss the Fréchet derivative of the forward operator and its adjoint. We are relatively brief regarding theoretical properties of any introduced quantity and refer to [BKL17, Secs. 2 and 3] as well as [CK13] for details. However, we offer the essential formulas of the direct scattering problem in the continuous and discretized form to get an overview.

This section contains the continuous formulation of the direct scattering problem and its discretization. A table of all involved formulas can also be found in App. A as Tab. A.1, where also a connection between mathematical formulas and their implementation is built.

Fundamental Terms We consider *time-independent* waves, particularly *point sources* and *plane waves*, illustrated in Fig. 2.2(a)/(b). Time-independent means that the frequency

---

$^\diamond$The term *time-harmonic* is also commonly used for *time-independent*.

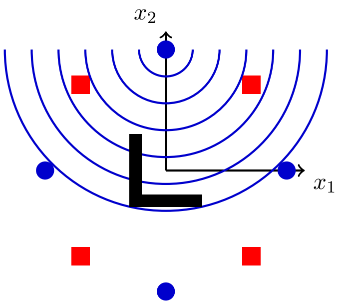
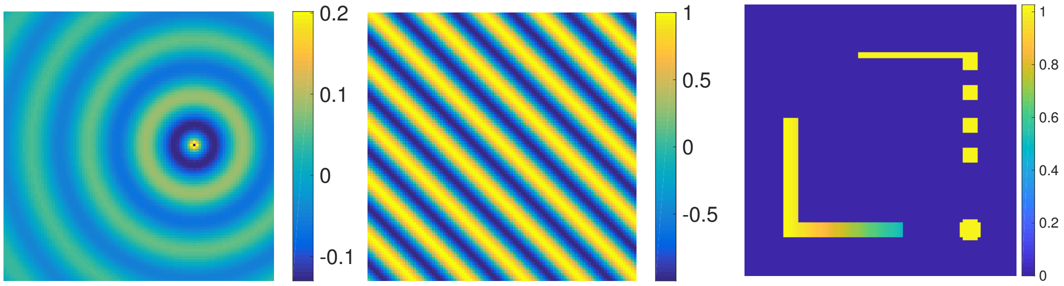


Figure 2.1: Experimental set-up for *multi-static* measurements: transmitters (blue filled circles) propagate incident fields one after the other while the receivers (red filled squares) measure the generated fields scattered by an obstacle.



(a) Point source.

(b) Plane wave.

(c) A contrast.

Figure 2.2: (a)/(b) Real parts of a point source and a plane wave. (c) A contrast: inhomogeneous obstacle and homogeneous surrounding medium (the background is zero).

is fixed. The time-dependence $\exp(-\mathrm{i}\omega t)$ can be easily discovered in the incident field of a *plane wave* $u^{\mathrm{i}}(x,t) = \exp(\mathrm{i}(kx \cdot \theta - \omega t))$, where θ is the direction of propagation, ω the angular frequency and $k = \omega/c_0 > 0$ the wave number (angular frequency divided by the speed of propagation). The direction of the plane wave in Fig. 2.2(b) is to the northeast. A *point source* is time-independent too and is shown in Fig. 2.2(a). We will consider point sources and plane waves as incident fields. As the scattering is time-independent we can omit the time dependency. Therefore this is called *time-independent* or *time-harmonic* scattering. The time-dependence is omitted from now on. We just consider $u^{\mathrm{i}}(x)$.

A time-independent wave can be characterized by wavelength λ, the frequency $f = c_0/\lambda$ with sound speed c_0 (in the background medium), the angular frequency $\omega = 2\pi f$ or the wave number $k = 2\pi/\lambda$.

A distinction is made between *penetrable* and *impenetrable* media. We consider only penetrable ones.

A medium can be *homogeneous*, e.g. paper, ceramic, or *inhomogeneous*, e.g. paper wrapped around ceramic. In general, we will consider inhomogeneous media. However, we always assume that the medium is homogeneous outside some sufficiently large sphere. For example, our inhomogeneous obstacle is paper wrapped around ceramic, but it is surrounded by a homogeneous host medium like air.

The *refractive index* n is a physical property describing the propagation of sound through the medium. The real part of the refractive index describes the scattering and the imaginary part the absorption. If there is no absorption the definition of n is

$n := c_0/c$, where c_0 is the sound speed in the homogeneous host medium and $c = c(x)$ the space-dependent sound speed in the inhomogeneous medium. It is usual to call the obstacle also *scatterer*—even if absorption occurs.

In this thesis the *obstacle* is an inhomogeneous medium $G \subset \mathbb{R}^d$, $d = 2, 3$, surrounded by a homogeneous medium. As the refractive index n equals one outside the obstacle, we prefer to consider the *contrast* $q := n^2 - 1$, see (2.1). An example is given in Fig. 2.2(c).

Throughout this thesis, we make the quiet assumption that the refractive index n is piece-wise continuous in $\mathbb{R}^d$ such that the contrast q in $\mathbb{R}^d$ has compact support and satisfies $1 + \mathrm{Re}(q) = \mathrm{Re}(n^2) > 0$ and $\mathrm{Im}(q) = \mathrm{Im}(n^2) \geq 0$, see [CK13, Sec. 8.2].

Further, a medium can be *isotropic*, i.e. a property is independent of the direction, or *anisotropic*, i.e. a property depends on the direction, e.g. wood. We will consider isotropic material.

2.1 What is a Direct Scattering Problem?

This section essentially follows [BKL19, Sec. 2] and therefore carefully distinguishes between the so-called region of interest D and the support of the contrast, $\overline{G}$, in contrast to [BKL17, Sec. 2].

Direct Scattering Problem The (classical) *Helmholtz equation* is the *homogeneous* differential equation

$$\Delta u^{\mathrm{i}}(x) + k^2 u^{\mathrm{i}}(x) = 0, \quad x \in \mathbb{R}^d, \tag{2.2}$$

and describes the propagation of an *incident* time-independent acoustic *wave* $u^{\mathrm{i}} : \mathbb{R}^d \to \mathbb{C}$ in dimension $d = 2, 3$ in a homogeneous *isotropic* medium for a constant wave number $k > 0$, see [CK13, Sec. 2.1] for details. Typical choices of u^{i} to solve the Helmholtz equation include plane waves and point sources.

In the model for the direct scattering problem we consider the scattering and absorption of time-independent acoustic waves by a penetrable and *isotropic*[◇] medium of compact support. Remember that the obstacle is an (inhomogeneous) medium $G \subset \mathbb{R}^d$, $d = 2, 3$, surrounded by a homogeneous medium. When the incident field interacts with an obstacle a *total field* u^{t} is generated.

If we denote by $q \colon \mathbb{R}^d \to \mathbb{C}$ the compactly supported contrast, that describes the obstacle, and by $\mathrm{supp}(q) = \overline{G}$ its support, the *total field* u^{t} solves the *inhomogeneous* Helmholtz equation[†]

$$\Delta u^{\mathrm{t}} + k^2 u^{\mathrm{t}} = -k^2 q u^{\mathrm{t}} \quad \text{in } \mathbb{R}^d. \tag{2.3}$$

[◇] Corresponding to (2.3), the differential equation for the *anisotropic* case is $\mathrm{div}((I+Q)\nabla u^t) + k^2 u^t = 0$ in $\mathbb{R}^d$, where the density is described by a matrix-valued material contrast parameter $Q \in \mathbb{C}^{d\times d}$. For the inverse acoustic scattering from anisotropic penetrable media see [LR17] or [Ren17, Ch. 3]. Electromagnetic scattering from anisotropic non-magnetic or magnetic media is described in [Ren17, Chs. 4 and 5].

[†] It is common to use the term *Helmholtz equation* for the homogeneous as well as the inhomogeneous differential equation because it is clear from the context which one is meant.

Further, the *scattered field*

$$u^{\mathrm{s}} = u^{\mathrm{t}} - u^{\mathrm{i}} \tag{2.4}$$

radiates away from the inhomogeneous medium, i.e. u^{s} satisfies *Sommerfeld's radiation condition* to ensure the radiation of energy

$$\lim_{|x|\to\infty} |x|^{(d-1)/2} \left(\frac{\partial}{\partial |x|} - \mathrm{i}k \right) u^{\mathrm{s}}(x) = 0, \tag{2.5}$$

uniformly in all directions $\hat{x} = x/|x| \in \mathbb{S} := \{y \in \mathbb{R}^d : |y| = 1\}$, where $|y| := \sqrt{y_1^2 + \ldots + y_d^2}$. (By abuse of notation, we do not explicitly denote the dimension of the sphere.)

The time-independent *direct scattering problem* is to find a scattered field u^{s} solving (2.3)–(2.5), or, equivalently, to find a radiating solution to the equation $\Delta u^{\mathrm{s}} + k^2(1+q)u^{\mathrm{s}} = -k^2\, q u^{\mathrm{i}}$. Note that the support of the right-hand side equals $\overline{G}$ (unless the incident field vanishes). (Clearly, the total field u^{t}, modeled by the solution to (2.3)–(2.5) is a quantity, which is accessible to physical measurements.)

Unique solvability of the scattering problem holds whenever the contrast q is smooth enough such that a unique continuation principle holds and, in particular, if q is bounded, see [CK13].

The direct scattering problem describes waves with fixed frequency that are interacting with a penetrable inhomogeneous medium, both in two and three dimensions. It models two physical settings: first, the scattering of acoustic waves and second, the scattering of electromagnetic waves in transverse magnetic (TM) polarization from some non-magnetic material, see e.g. [CK13].

Reformulation of the Direct Scattering Problem We denote by Φ the radiating *fundamental solution* of the Helmholtz equation,

$$\Phi(x) = \frac{\mathrm{i}}{4} H_0^{(1)}(k|x|) \quad \text{if } x \in \mathbb{R}^2 \setminus \{0\}, \quad \Phi(x) = \frac{1}{4\pi} \frac{\mathrm{e}^{\mathrm{i}k|x|}}{|x|} \quad \text{if } x \in \mathbb{R}^3 \setminus \{0\}, \tag{2.6}$$

where $H_0^{(1)}$ is the Hankel function of the first kind and order zero, see [AS65, Ch. 9]. It describes a radiating *spherical wave*. Typical models of the incident waves u^{i} are the real part of a *point source*, i.e. $u^{\mathrm{i}}(x) = \Phi(x - p)$ with origin in $p \in \mathbb{R}^d$, and the real part of a *plane wave*, i.e. $u^{\mathrm{i}}(x) = \mathrm{e}^{\mathrm{i}kx\cdot\theta}$ with θ as the direction of propagation.

Further, we define for the domain $D \supset \overline{G}$ the radiating *volume potential*

$$v(x) := (Vf)(x) := k^2 \int_D \Phi(x-y) f(y) \, \mathrm{d}y, \quad x \in \mathbb{R}^d. \tag{2.7}$$

The *volume potential operator* V defines the radiating solution $v = Vf$ of the Helmholtz equation $\Delta v + k^2\, v = -k^2\, f$ in $\mathbb{R}^d$, such that the direct scattering problem can be equivalently reformulated as follows: Find the scattered field u^{s} solving the so-called

Lippmann-Schwinger integral equation

$$u^{\mathrm{s}} - V(q \cdot u^{\mathrm{s}}) = V(q \cdot u^{\mathrm{i}}) \quad \text{for } x \in D. \tag{2.8}$$

Note that we use the notation $f \cdot g$ if we want to stress the pointwise multiplication. In general the set D is assumed to be only slightly larger than the support of the contrast, $\overline{G}$, and is called the *region of interest* because we know that the scattering takes place inside.

Further note in (2.8) that the scattered and incident fields, u^{s} and u^{i}, are scaled by k^2 in the volume potential operator. Therefore the linearized model given by the *Born approximation* $u^{\mathrm{s}}(q) \approx V(q \cdot u^{\mathrm{i}})$ holds for relatively small k, cf. [Bor26].

In addition, the Lippmann-Schwinger equation implicitly defines the *Lippmann-Schwinger solution operator* T_q mapping u^{i} to u^{t} on D via

$$T_q \colon L^2(D) \to L^2(D), \quad T_q := (I - V(q\,\cdot))^{-1}. \tag{2.9}$$

Incident Field and Scattered Field Next, we need to clarify the model for the generation of the incident field and the measurement of the scattered field. For the first part, we will present only the case of point sources. However, we remark that IPscatt can also handle the case of sources being plane waves. For point sources one assumes that the transmitters generating the field are located on some surface Γ_{i} in $\mathbb{R}^d \setminus \overline{G}$. Then, the so-called *single-layer potential*[◇], see [CK13, Sec. 3.1],

$$\mathrm{SL}_{\Gamma_{\mathrm{i}}\to D} \colon L^2(\Gamma_{\mathrm{i}}) \to L^2(D), \quad (\mathrm{SL}_{\Gamma_{\mathrm{i}}\to D}\, g)(x) := \int_{\Gamma_{\mathrm{i}}} \Phi(x-y) g(y)\, \mathrm{d}s(y), \quad x \in D \setminus \Gamma_{\mathrm{i}} \tag{2.10}$$

models the incident field via $u^{\mathrm{i}} = \mathrm{SL}_{\Gamma_{\mathrm{i}}\to D}\, g$ with density g.

Scattering Process As mentioned above, this field generated on Γ_{i} propagates into the region of interest D and interacts there with the contrast q via the Lippmann-Schwinger solution operator T_q. Next, the resulting fields u^{t} respectively u^{s} propagate out of D and can be measured.

It is known that every radiating solution u to the Helmholtz equation, i.e., in particular, the scattered field u^{s}, has the asymptotic behavior[†] of an outgoing spherical wave, see [CK13, Th. 2.6, Sec. 8.4] for 3D and [CK13, Sec. 3.4] for 2D,

$$u(x) = \frac{\exp(\mathrm{i}k|x|)}{|x|^{(d-1)/2}} \left(u_\infty(\hat{x}) + \mathcal{O}\left(\frac{1}{|x|}\right)\right) \tag{2.11}$$

as $|x| \to \infty$, uniformly in all directions $\hat{x} = x/|x| \in \mathbb{S}$. Therefore only the so-called *far field pattern*[‡] $u_\infty \colon \mathbb{S} \to \mathbb{C}$ is measured in direction $\hat{x}$ if the receivers are located far away[§]

[◇] In [BKL17, Sec. 3] it was mistakenly $x \in D$ instead of $x \in D \setminus \Gamma_{\mathrm{i}}$.

[†] Note the mistake in the scattered field's description by the far field in [BKL17, Sec. 2]. The factor $\exp(\mathrm{i}kr)/r$ with $r = |x|$ holds only in the three-dimensional case. In general, $\exp(\mathrm{i}kr)/r^{(d-1)/2}$ for $d = 2, 3$ includes the factor in (2.5).

[‡] The *far field pattern* is also known as *scattering amplitude*, see [CK13, Sec. 2.2].

[§] The receiver is called *far away* from the scattering object if the distance is several wavelengths, see [LKK13, Sec. 1].

from G. We use a notation with the dimension-dependent factor $\gamma = \exp(\mathrm{i}\pi/4)/\sqrt{8\pi k}$ in 2D and $\gamma = 1/(4\pi)$ in 3D in the far field pattern u_∞; otherwise such factor would be required in (2.11). For example, the translated fundamental solution $\Phi(x-y)$, i.e. the point source at $y \in \mathbb{R}^d$, behaves for large $x \in \mathbb{R}^d$ as a radiating spherical wave $\Phi(x)$ with the far field pattern

$$\Phi_\infty(\hat{x}, y) = \gamma \exp(-\mathrm{i}k\hat{x} \cdot y) \tag{2.12}$$

observed in direction $\hat{x} \in \mathbb{S}$, see e.g. [Lec08, Sec. I-2] or [DLP$^+$11, Sec. 4.1.2]. We will not do rely on far field measurements in the next chapters, but they will return in Ch. 8.

In contrast to the receivers far away: If the receivers are near G, it is advantageous to model the receivers as measuring the near field $u(x)$, i.e., in particular, the scattered field $u^{\mathrm{s}}(x)$. We remark that IPscatt can handle both cases, however, in what follows we only discuss the near field case for the sake of brevity. Let Γ_{s} in $\mathbb{R}^d \setminus \overline{G}$ be a surface on which the receivers are located. Then a field f given on D can be propagated onto Γ_{s} via the *solution-to-data operator*

$$V_{D\to\Gamma_{\mathrm{s}}} \colon L^2(D) \to L^2(\Gamma_{\mathrm{s}}), \quad (V_{D\to\Gamma_{\mathrm{s}}} f)(x) := k^2 \int_D \Phi(x-y) f(y) \, \mathrm{d}y, \quad x \in \Gamma_{\mathrm{s}}. \tag{2.13}$$

The data at a particular receiver on Γ_{s} of the resulting field is modeled as pointwise evaluation at the related position x. Altogether, the introduced operators can be used to model the whole scattering process via the *forward operator*

$$\mathcal{F}(q) := V_{D\to\Gamma_{\mathrm{s}}}(q\cdot) T_q \, \mathrm{SL}_{\Gamma_{\mathrm{i}}\to D} \, . \tag{2.14}$$

A sensible domain[◇] and range of the operator are given via $\mathcal{F} \colon L^2_{\mathrm{Im}\geq 0}(D) \to \mathrm{HS}$, where HS is the space of all Hilbert-Schmidt operators $\mathrm{HS}(L^2(\Gamma_{\mathrm{i}}), L^2(\Gamma_{\mathrm{s}}))$. This operator is nonlinear in q since the Lippmann-Schwinger solution operator T_q is so. However, the evaluation of $\mathcal{F}$ at q is a linear operator mapping a transmitter density g on Γ_{i} to the resulting field on Γ_{s}. Further, it is clear from the above considerations that the forward operator depends on the incident field type, the positions of the transmitters respectively the outgoing directions, the measurement model (near field or far field data) and the positions respectively directions of the receivers.

For the sake of completeness, we give the *Fréchet derivative of the forward operator* $\mathcal{F}$ because it is an important ingredient of many reconstruction and optimization schemes (if it is applied to the defect, i.e. the difference between the predicted data $\mathcal{F}(q)$, that bases on the reconstructed contrast q, and the given data),

$$\mathcal{F}'(q) \colon L^2(D) \to \mathrm{HS}, \quad \mathcal{F}'(q)[h]g = V_{D\to\Gamma_{\mathrm{s}}}(I + (q\cdot)T_q V)(h\cdot)T_q \, \mathrm{SL}_{\Gamma_{\mathrm{i}}\to D} \, g \tag{2.15}$$

for $g \in L^2(\Gamma_{\mathrm{i}})$.

Existence and Uniqueness of the Solution of the Lippmann-Schwinger Equation

In [LKK13, Th. 6] it was shown that the existence and unique solvability of the Lippmann-Schwinger integral equation (2.8) essentially holds for $\mathrm{Im}(q) \geq 0$ and $q \in L^p(D)$ with

---

[◇] In fact, a sensible domain is $L^p_{\mathrm{Im}\geq 0}$ (with $p > d/2$, $d = 2, 3$), see the existence and uniqueness below.

$p > d/2$ for dimensions $d = 2$ and 3. Further, the resulting scattered field u^{s} defines a radiating solution to the Helmholtz equation. (Further details are avoided for the sake of simplicity.) Therefore a sensible domain of the forward operator $\mathcal{F}$, see (2.14), is $L^p_{\mathrm{Im}\geq 0}$ with $p > d/2$ instead of the above mentioned $L^2_{\mathrm{Im}\geq 0}$.

2.2 Numerical Solution

For a numerical solution of the forward operator we essentially follow [Vai00]. We start with a rough survey to convey the basics. Afterwards we consider the steps corresponding to the periodization of the volume potential operator in detail because it ends in a correction of the Fourier coefficients in one case. The correction was already mentioned in [BKL17, Sec. 3.2]. However, here we will consider the whole computation of that case because this correction goes beyond the improvements from Kazimierski and Lechleiter. For the further steps of the forward problem, i.e. the discretization of the Lippmann–Schwinger equation and the implementation of the collocation scheme, see [BKL17, Secs. 3.3 and 3.4].

The basic idea of a fast numerical solution of the forward operator is to exploit the convolution structure of the volume potential operator V, see (2.7), consisting of the radiating fundamental solution Φ, see (2.6), and a function f, say, on B_R (ball with radius R) for some fixed radius $R > 0$. Therefore this operator can in Fourier space be represented as the element-wise product of the Fourier coefficients of both functions. This is the key for an efficient solution of the Lippmann-Schwinger equation (2.8) on an equidistant grid via the fast Fourier transform (FFT). This equidistant grid of the computational domain D_{2R}, defined in (2.16), is denoted by $\mathsf{CD} = \mathbb{C}^d_N$. More precisely, the *periodized volume potential operator* V_{2R}, defined in (2.19), allows to consider in the computational domain D_{2R} the *periodized Lippmann-Schwinger equation* corresponding to (2.8),

$$v - V_{2R}(q \cdot v) = V_{2R}(q \cdot u^{\mathrm{i}}) \quad \text{in } L^2(D_{2R}).$$

Fourier coefficients of the convolution kernel can be explicitly computed, see [Vai00]. Finally, we use the GMRES algorithm to solve the arising discrete linear system efficiently.

Computational Domain and Region of Interest Starting from radius R the *computational domain* (CD) is the smallest square/cube around the circle/ball with radius $2R$,

$$D_{2R} = [-2R, 2R)^d \subset \mathbb{R}^d, \tag{2.16}$$

see Fig. 2.3. The corresponding equidistant grid of D_{2R} is $\mathsf{CD} = \mathbb{C}^d_N$ with mesh size $h_N = 4R/N$. Following [Vai00] it is sufficient to evaluate the volume potential (2.7) on this computational domain D_{2R} instead of $\mathbb{R}^d$. Therefore we will define and consider the corresponding periodized volume potential operator V_{2R}, see (2.19).

IPscatt uses further improvements to evaluate the Lippmann-Schwinger equation (2.8) efficiently. One of these is the restriction of all unknowns and all data to the already-mentioned *region of interest* (ROI), that is the biggest square/cube inside the circle/ball

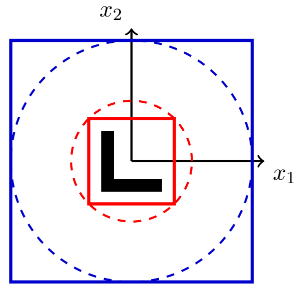


Figure 2.3: Areas: The square around the big circle with radius $2R$ is the computational domain (CD), $D_{2R} = [-2R, 2R)^2 \subset \mathbb{R}^2$. The square inside the small circle with radius R is the region of interest (ROI), $D = (-R/\sqrt{2}, R/\sqrt{2})^2$. The corner inside the region of interest represents an obstacle G.

with radius R, i.e.

$$D = (-R/\sqrt{2}, R/\sqrt{2})^d \tag{2.17}$$

(the definition in [BKL17, Sec. 3.1] was corrected$^\diamond$), see Fig. 2.3; the corresponding equidistant grid of D is $\mathsf{ROI} = \mathbb{C}^{N_\mathrm{D}}$ with $N_\mathrm{D} := \lfloor N/(2\sqrt{2}) \rfloor^d$; again, see [BKL17, Sec. 3.4] for details.

Actually, the region of interest should be the *mathematical sensible region*, i.e. B_R, however, the chosen square/cube $D \subset B_R$ as region of interest is easier to implement than a circle/ball and can be used as long as it is ensured that $\overline{G}$, i.e. the support of the contrast q, is inside D.

Periodization of the Volume Potential Operator In this paragraph we essentially follow [Vai00] considering the periodization of the volume potential operator V, see (2.7). Therefore we adapt the fundamental solution Φ of the Helmholtz equation, see (2.6), to define the volume potential via the periodized volume potential operator V_{2R} and consider the corresponding Fourier coefficients.

Remember that $\mathrm{supp}(q) \subset D$. This implies that $\mathrm{supp}(f) \subset D$ for $f = qu^\mathrm{i}$ or $f = qu^\mathrm{s}$. Plugging such f into the volume potential operator V from (2.7), we conclude that the source term f vanishes outside $B_R \supset D$. In addition, a close inspection of (2.7) reveals that to evaluate $Vf(x)$ the values of the fundamental solution Φ do only have to be known in the ball B_{2R}. This is the case since if the points x, y are in B_R the difference $x - y$ belongs to B_{2R}. Consequently, as in [Vai00], we define a modified *kernel*† Φ_{2R} by cutting off Φ outside of B_{2R} and multiplying it by k^2,

$$\Phi_{2R}(x) := \begin{cases} k^2\Phi(x) & \text{if } x \in \overline{B_{2R}}, \\ 0 & \text{if } x \in \overline{D_{2R}} \setminus \overline{B_{2R}}. \end{cases} \tag{2.18}$$

This allows us to define an equivalent *periodized volume potential operator* V_{2R} in the computational domain D_{2R}, see (2.16). Extending both Φ_{2R} and q from D_{2R} to $\mathbb{R}^d$ as $4R$-periodic functions with respect to $x_1, \ldots, x_d$, we define this periodized volume

---

$^\diamond$The defined $D = [-R/\sqrt{2}, R/\sqrt{2})^d$ in [BKL17, Sec. 3.1] is in $\overline{B_R}$ but not in the mathematical sensible region B_R. Therefore the open interval has to be used.

†In [BKL17, Sec. 3.2] a typo states B_{2R} instead of $\overline{B_{2R}}$.

potential operator

$$V_{2R}\colon L^2(D_{2R}) \to L^2(D_{2R}), \quad (V_{2R}f)(x) := \int_{D_{2R}} \Phi_{2R}(x-y)f(y)\,\mathrm{d}y, \quad x \in D_{2R}. \tag{2.19}$$

We use the same symbol for the original and extended versions since it is clear form the context which one is meant.

Further, note that the periodized volume potential operator V_{2R} replaces V for a numerical solution in the Lippmann-Schwinger solution operator T_q, see (2.9), and in the Fréchet derivative $\mathcal{F}'(q)$, see (2.15). Of course, in this case, the forward operator $\mathcal{F}$, see (2.14), is also based on the replaced Lippmann-Schwinger solution operator.

As discussed in the beginning of this paragraph, $(Vf)|_D = (V_{2R}f)|_D$ for all $f \in L^2(D)$ (and even for all $f \in L^2(B_R)$) if they are extended by zero to all of D_{2R}. Since, the operator V_{2R} is a $4R$-periodic convolution operator acting on functions in $L^2(D_{2R})$ it can be evaluated by means of the convolution theorem. To that end, we consider the complete orthogonal system in $L^2(D_{2R})$:

$$\varphi_j(x) := \frac{1}{(4R)^{d/2}} \exp\left(\frac{2\pi\mathrm{i}}{4R}\langle j, x\rangle\right), \quad j \in \mathbb{Z}^d\,. \tag{2.20}$$

Then for every $f \in L^2(D_{2R})$ the representation $f = \sum_{j\in\mathbb{Z}^d} \hat{f}(j)\varphi_j$ holds, where the Fourier coefficients $\hat{f}(j)$ are defined as $\int_{D_{2R}} f\,\overline{\varphi_j}\,\mathrm{d}x$. Further, the convolution theorem for that system shows that the Fourier coefficients of $V_{2R}f$ are given via

$$\widehat{V_{2R}f}(j) = (4R)^{d/2}\,\hat{\Phi}_{2R}(j)\,\hat{f}(j), \quad j \in \mathbb{Z}^d. \tag{2.21}$$

Further, the *Fourier coefficients* of the convolution kernel can be explicitly computed, see [Vai00, Hoh01]. They decay in j as $|\hat{\Phi}_{2R}(j)| \le C(1+|j|)^{-3/2}$ for $d=2$ and $|\hat{\Phi}_{2R}(j)| \le C(1+|j|)^{-2}$ for $d=3$, see Fig. 2.4. To omit repeating the factor $(4R)^{d/2}$ we define

$$\Psi_{2R} := (4R)^{d/2}\,\hat{\Phi}_{2R}. \tag{2.22}$$

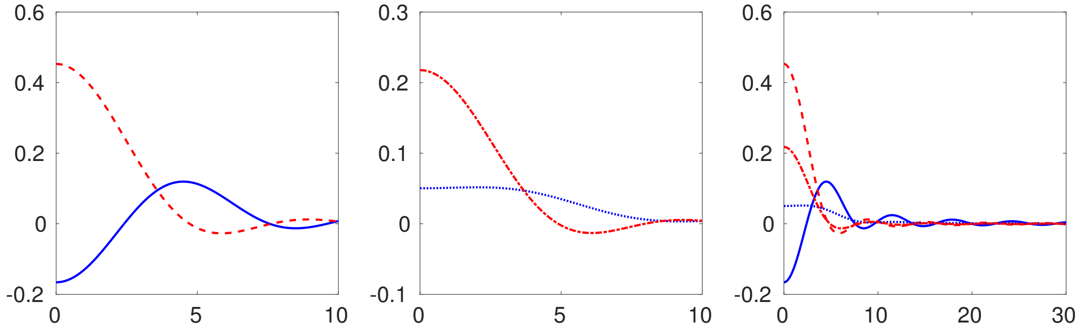
Further, we define $\kappa := 2Rk$ and $p_j := \pi|j|$. Then in the two-dimensional setting

$$\Psi_{2R}(j) = \begin{cases} \frac{\kappa^2}{p_j^2-\kappa^2}\left(1 + \frac{\mathrm{i}\pi}{2}\left[p_j J_1(p_j)H_0^{(1)}(\kappa) - \kappa J_0(p_j)H_1^{(1)}(\kappa)\right]\right) & \text{if } p_j \ne \kappa, \\ \frac{\mathrm{i}\pi\kappa^2}{4}\left[J_1(\kappa)H_1^{(1)}(\kappa) + J_0(\kappa)H_0^{(1)}(\kappa)\right] & \text{if } p_j = \kappa, \end{cases} \quad \text{for } d=2,$$

where J_ν and $H_\nu^{(1)}$ are the cylindrical Bessel and Hankel functions. Further, in three dimensions $\Psi_{2R}(0) = (1-\mathrm{i}\kappa)\exp(\mathrm{i}\kappa) - 1$ and

$$\Psi_{2R}(j) = \begin{cases} \frac{\kappa^2}{p_j^2-\kappa^2}\left(1 - \mathrm{e}^{\mathrm{i}\kappa}\left[\cos(p_j) - \frac{\mathrm{i}\kappa}{p_j}\sin(p_j)\right]\right) & \text{if } 0 \ne p_j \ne \kappa, \\ \frac{1}{2}\mathrm{e}^{\mathrm{i}\kappa}\left(\kappa\sin(\kappa) + \mathrm{i}[\kappa\cos(\kappa) - \sin(\kappa)]\right) & \text{if } p_j = \kappa, \end{cases} \quad \text{for } d=3.$$

Note the correction in the above formula for $d=3$ in the case of $p_j = \kappa$ of the value given in [Vai00, Hoh01]. Therefore the computation of this case is carried out in the next paragraph.



(a) 2D. (b) 3D. (c) Decay in 2D and 3D.

Figure 2.4: Real (blue solid and blue dotted) and imaginary part (red dashed and red dash-dotted) of Fourier coefficients $\hat{\Phi}_{2R}$ over $p_j = \pi|j| \geq 0$ with $R = 1$, $k = 1$, such that $\kappa = 2Rk = 2$. We see that the Fourier coefficients decay in j.

Fourier Coefficients' Computation The computation of the Fourier coefficients in three dimensions in the case of $p_j = \kappa$ is described in the following. If $0 \neq p_j \neq \kappa$, then

$$\Psi_{2R}(j) = \frac{\kappa^2}{p_j^2 - \kappa^2}\left(1 - \mathrm{e}^{\mathrm{i}\kappa}\left[\cos(p_j) - \frac{\mathrm{i}\kappa}{p_j}\sin(p_j)\right]\right) = \frac{\kappa^2 - \kappa^2\mathrm{e}^{\mathrm{i}\kappa}\cos(p_j) + \kappa^2\mathrm{e}^{\mathrm{i}\kappa}\frac{\mathrm{i}\kappa}{p_j}\sin(p_j)}{p_j^2 - \kappa^2}.$$

Numerator and denominator tend to 0 for $p_j \to \kappa$. Therefore L'Hospital's rule is used. This results in

$$\Psi_{2R}(j) = \frac{0 - \kappa^2\mathrm{e}^{\mathrm{i}\kappa}(-\sin(p_j)) + \kappa^2\mathrm{e}^{\mathrm{i}\kappa}\mathrm{i}\kappa\left[-\frac{1}{p_j^2}\sin(p_j) + \frac{1}{p_j}\cos(p_j)\right]}{2p_j}.$$

Further, with $p_j = \kappa$ it results in

$$\Psi_{2R}(j) = \frac{\kappa}{2}\mathrm{e}^{\mathrm{i}\kappa}\sin(\kappa) - \frac{1}{2}\mathrm{e}^{\mathrm{i}\kappa}\mathrm{i}\sin(\kappa) + \frac{\kappa}{2}\mathrm{e}^{\mathrm{i}\kappa}\mathrm{i}\cos(\kappa) = \frac{1}{2}\mathrm{e}^{\mathrm{i}\kappa}\left(\kappa\sin(\kappa) + \mathrm{i}[\kappa\cos(\kappa) - \sin(\kappa)]\right).$$

2.3 Discretization

Next, we discuss the discretization of the scattering problem, see (2.14). That concerns the discretization of the single-layer potential $\mathrm{SL}_{\Gamma_\mathrm{i}\to D}$, the volume potential operator V, the solution-to-data operator $V_{D\to\Gamma_\mathrm{s}}$, the Lippmann-Schwinger solution operator T_q and finally the forward operator $\mathcal{F}$. In addition, for practical usage we give the Fréchet derivative of the forward operator and its adjoint. This section follows [BKL19, Sec. 2.2].

As we have used the notation $f \cdot g$ to stress the pointwise multiplication in the continuous case we use $f \odot g$ in the discretized one. Where sensible underlined letters denote a discretized version of a continuous quantity.

Single-Layer Potential The *single-layer potential* $\mathrm{SL}_{\Gamma_\mathrm{i}\to D}$, see (2.10), for N_i source points at positions[◇] p_j is *discretized* via

$$\mathrm{SL}_{N_\mathrm{i},N_\mathrm{D}}\colon \mathbb{C}^{N_\mathrm{i}} \to \mathbb{C}^{N_\mathrm{D}}, \quad \mathrm{SL}_{N_\mathrm{i},N_\mathrm{D}} := \Phi_\mathrm{SL}\,(\omega^\mathrm{i} \odot \,),$$

where $(\Phi_\mathrm{SL})_{\ell,j} := \Phi(x_\ell - p_j)$ with $\ell = 1,\dots,N_\mathrm{D}$ and $j = 1,\dots,N_\mathrm{i}$ is the collocation of the transmitted field and $\omega_j^\mathrm{i} \approx \mathrm{d}s(p_j)$ are the approximations of the infinitesimal element on Γ_i.

One notices in the continuous formulation of the single-layer potential (2.10) the required absence of point sources inside the region of interest D because of the singularity of the radiating fundamental solution Φ, see (2.6), at the origin. This restriction is irrelevant in the discretized version as long as the transmitter is not near a grid point.[†]

Volume Potential Operator The *periodized volume potential operator* V_{2R}, see (2.19), is *discretized* via

$$V_{N_\mathrm{D}}\colon \mathbb{C}^{N_\mathrm{D}} \to \mathbb{C}^{N_\mathrm{D}}, \quad V_{N_\mathrm{D}} := \mathcal{R}_N\,\mathrm{FFT}_N^{-1}\,(\widehat{\Phi}_N \odot \,)\,\mathrm{FFT}_N\,\mathcal{E}_N,$$

where $\mathcal{E}_N\colon \mathbb{C}^{N_\mathrm{D}} \to \mathbb{C}_N^d$ and $\mathcal{R}_N\colon \mathbb{C}_N^d \to \mathbb{C}^{N_\mathrm{D}}$ are the discretizations of the extension and restriction operators from the region of interest to the computational domain (and vice versa). Further, $\mathrm{FFT}_N = $ `fft2` and $\mathrm{FFT}_N^{-1} = $ `ifft2` are the forward and inverse fast Fourier transforms with the common normalization constants as used in MATLAB, see [FJ05]. Finally, $\widehat{\Phi}_N$ denotes the Fourier coefficients of the fundamental solution stored in FFT convention, i.e. the constant term being at matrix position $(1,1)$.

Solution-to-Data Operator The *solution-to-data operator* $V_{D\to\Gamma_\mathrm{s}}$, see (2.13), is *discretized* via

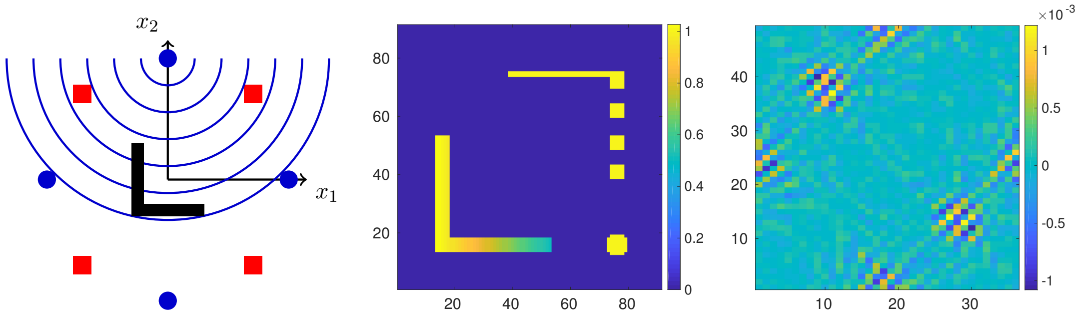
$$V_{N_\mathrm{D},N_\mathrm{s}}\colon \mathbb{C}^{N_\mathrm{D}} \to \mathbb{C}^{N_\mathrm{s}}, \quad V_{N_\mathrm{D},N_\mathrm{s}} := h_N^d\,k^2\,\Phi_V,$$

where the exact form of the kernel Φ_V depends on whether near or far field data is being measured and on the positions of the receivers x_ℓ with $\ell = 1,\dots,N_\mathrm{s}$ respectively their directions. Further, h_N^d is the approximation of the region of interest's infinitesimal element. One notices that there is no restriction for receivers' positions as long as they are not in $\overline{G}$.[‡]

[◇] Note that the notation p_j is slightly overloaded. In Sec. 2.2 we already defined $p_j := \pi|j|$. However, that was a temporary abbreviation. In this section p_j is a (more or less) arbitrary position.

[†] In the *source code* of IPscatt a stricter requirement is used than necessary: For the *transmitters* we simply require that they are outside the mathematical sensible region B_R to omit the implementation of a nearby condition. Remember that we have chosen the region of interest D as the biggest square inside the mathematical sensible region B_R. Hence, for the current choice of D it would be sufficient to check if the transmitters' positions are outside this square. However, omitting the hole mathematical sensible region prevents mistakes in the case of an adaption of the region of interest.

[‡] In the *source code* of IPscatt a stricter requirement is used than necessary: For the *receivers* we require that they are outside of the mathematical sensible region B_R because actually the true contrast q is unknown.



(a) Experimental set-up. (b) Contrast. (c) Measurements.

Figure 2.5: Illustration of the forward operator (contrast-to-measurement map in the case of multi-static measurements) mapping the discretized *contrast* $\underline{q} \in \mathbb{C}^{N_\mathrm{D}}$ onto the *measurements* (the *data*) $F_\mathrm{meas} \in \mathbb{C}^{N_\mathrm{s} \times N_\mathrm{i}}$. (a) The experimental set-up regardless of the number of transmitters and receivers (as in Fig. 2.1). (b) A discretized contrast in the region of interest, i.e. $\underline{q} \in \mathbb{C}^{N_\mathrm{D}}$. (c) The result of the contrast-to-measurement map in the case of $N_\mathrm{s} = 49$ receivers and $N_\mathrm{i} = 36$ transmitters roughly arranged as in (a) and the contrast in (b).

Forward Operator Finally, the *Lippmann-Schwinger solution operator* T_q, see (2.9), is *discretized* via

$$T_{\underline{q}} \colon \mathbb{C}^{N_\mathrm{D}} \to \mathbb{C}^{N_\mathrm{D}}, \quad T_{\underline{q}} := (I - V_{N_\mathrm{D}}(\underline{q} \odot\,))^{-1} \tag{2.23}$$

with $\underline{q} \in \mathbb{C}^{N_\mathrm{D}}$. Altogether, the *discretized forward operator*, see (2.14) for $\mathcal{F}$, is given via

$$\underline{\mathcal{F}} \colon \mathbb{C}^{N_\mathrm{D}} \to \mathbb{C}^{N_\mathrm{s} \times N_\mathrm{i}}, \quad \underline{\mathcal{F}}(\underline{q}) := V_{N_\mathrm{D},N_\mathrm{s}}(\underline{q} \odot\,) T_{\underline{q}}\, \mathrm{SL}_{N_\mathrm{i},N_\mathrm{D}}\,. \tag{2.24}$$

This forward operator $\underline{\mathcal{F}}$ is also called *contrast-to-measurement map* mapping the discretized *contrast* $\underline{q} \in \mathbb{C}^{N_\mathrm{D}}$ onto the *measurements* (the *data*) $F_\mathrm{meas} \in \mathbb{C}^{N_\mathrm{s} \times N_\mathrm{i}}$ that is the scattered (near or far) field at the receivers' positions (or the measurements direction in the case of the far field). This is illustrated in Fig. 2.5.

The resulting *discretized Fréchet derivative of the forward operator*, (2.15), is given via, see [BKL17, Sec. 3.6],

$$\underline{\mathcal{F}}'(\underline{q}) \colon \mathbb{C}^{N_\mathrm{D}} \to \mathbb{C}^{N_\mathrm{s} \times N_\mathrm{i}}, \quad \underline{\mathcal{F}}'(\underline{q})[\underline{h}] = A_{N_\mathrm{D},N_\mathrm{s}}(\underline{h} \odot\,) B_{N_\mathrm{D},N_\mathrm{i}}, \quad \underline{h} \in \mathbb{C}^{N_\mathrm{D}} \tag{2.25}$$

with auxiliary matrices

$$\begin{aligned} A_{N_\mathrm{D},N_\mathrm{s}} &= V_{N_\mathrm{D},N_\mathrm{s}}(I + (\underline{q} \odot\,) T_{\underline{q}} V_{N_\mathrm{D}}) \in \mathbb{C}^{N_\mathrm{s} \times N_\mathrm{D}}, \\ B_{N_\mathrm{D},N_\mathrm{i}} &= T_{\underline{q}}\, \mathrm{SL}_{N_\mathrm{i},N_\mathrm{D}} \in \mathbb{C}^{N_\mathrm{D} \times N_\mathrm{i}}. \end{aligned}$$

The *discretized adjoint of the derivative* is then given via, see [LKK13, App. E],

$$[\underline{\mathcal{F}}'(\underline{q})]^* \colon \mathbb{C}^{N_\mathrm{s} \times N_\mathrm{i}} \to \mathbb{C}^{N_\mathrm{D}}, \quad [\underline{\mathcal{F}}'(\underline{q})]^* \underline{H} = \sum_{j=1}^{N_\mathrm{s}} \sum_{\ell=1}^{N_\mathrm{i}} \underline{H}_{j,\ell}\, \overline{A_{N_\mathrm{D},N_\mathrm{s}}(j,\cdot)}\, \overline{B_{N_\mathrm{D},N_\mathrm{i}}(\cdot,\ell)} \tag{2.26}$$

for $\underline{H} \in \mathbb{C}^{N_\mathrm{s} \times N_\mathrm{i}}$. This adjoint applied to the defect $\mathcal{F}(q) - F^\delta_\mathrm{meas}$ is the derivative of the least-squares error of the forward operator[◇] and as already mentioned is an important ingredient of many reconstruction and optimization schemes.

Note that all spaces for the discretized operators are equipped with the usual non-weighted ℓ^2 norms. Moreover, up until this point all complex spaces involved in definitions of discretized operators were equipped with the standard inner product. However, for the definition of the full forward operator $\underline{\mathcal{F}}$ one could also consider weighted versions which mimic the behavior of the spaces involved in the definition of the operator $\mathcal{F}$. Therefore we follow [BKL17, Sec. 3.6] for the rest of this section.

The discretization of the norm in the domain space $L^p_{\mathrm{Im}\geq 0}$ yields the norm, related to the region of interest,

$$\|x\|_{\mathrm{roi},p} := \Big(h_N^d \sum_i |x_i|^p\Big)^{\frac{1}{p}}, \quad x \in \mathbb{C}^{N_\mathrm{D}},$$

for the space $\mathbb{C}^{N_\mathrm{D}}$. In particular for $p = 2$ the following inner product discretization

$$\langle x, y\rangle_{\mathrm{roi}} := h_N^d \sum_i x_i \overline{y_i}, \quad x, y \in \mathbb{C}^{N_\mathrm{D}}, \tag{2.27}$$

is sensible. For the space $\mathbb{C}^{N_\mathrm{s}\times N_\mathrm{i}}$ we propose the discretization of the inner product in $\mathrm{HS}(L^2(\Gamma_\mathrm{i}), L^2(\Gamma_\mathrm{s}))$ by the weighted Frobenius product

$$\langle A, B\rangle_{\mathrm{dis}} := \mathrm{trace}\big(B^*\,\omega^\mathrm{s}\,A\big) = \sum_{j=1}^{N_\mathrm{s}} \omega_j^\mathrm{s} \sum_{\ell=1}^{N_\mathrm{i}} A_{j\ell}\overline{B_{j\ell}}, \quad A, B \in \mathbb{C}^{N_\mathrm{s}\times N_\mathrm{i}}\,, \tag{2.28}$$

where the diagonal weight matrix $\omega^\mathrm{s} = \mathrm{diag}((\omega_j^\mathrm{s})_{j=1}^{N_\mathrm{s}})$ is due to the weights approximating the surface element on Γ_s. The letter combination "dis" is due to the fact that the induced norm $\|\cdot\|_{\mathrm{dis}}$ is used as discrepancy term in the reconstruction scheme discussed in the next chapter. Formally, the change of norms can be incorporated into the above definitions by introducing the embeddings $\iota_1 : (\mathbb{C}^{N_\mathrm{D}}, \|\cdot\|_{\mathrm{roi},p}) \to (\mathbb{C}^{N_\mathrm{D}}, \|\cdot\|_p)$ with $\iota_1(\underline{h}) := \underline{h}$ and $\iota_2 : (\mathbb{C}^{N_\mathrm{s}\times N_\mathrm{i}}, \|\cdot\|_\mathrm{F}) \to (\mathbb{C}^{N_\mathrm{s}\times N_\mathrm{i}}, \|\cdot\|_{\mathrm{dis}})$ with $\iota_2(\underline{H}) := \underline{H}$, where $\|\cdot\|_\mathrm{F}$ is the unweighted Frobenius norm. Clearly, prepending ι_2 and appending ι_1 to the definitions to the discretized forward operator will not change the computed result. The only changes will occur in the adjoint operators, since $\iota_1^*(\underline{g}) = \underline{g}/h_N^d$ and $\iota_2^*(\underline{G}) = \omega^\mathrm{s}\underline{G}$. For example the *discretized adjoint of the derivative* with respect to the new norms is given by

$$[\underline{\mathcal{F}}'(\underline{q})]^*\underline{H} = \sum_{j=1}^{N_\mathrm{s}} \frac{\omega_j^\mathrm{s}}{h_N^d} \sum_{\ell=1}^{N_\mathrm{i}} \underline{H}_{j,\ell}\, \overline{A_{N_\mathrm{D},N_\mathrm{s}}(j,\cdot)}\, \overline{B_{N_\mathrm{D},N_\mathrm{i}}(\cdot,\ell)} \quad \text{for } \underline{H} \in \mathbb{C}^{N_\mathrm{s}\times N_\mathrm{i}}. \tag{2.29}$$

[◇] This relation between the adjoint and the derivative is explained in Sec. 6.1.

3 Inverse Scattering Problem

This chapter describes inverse problems, in particular inverse scattering problems, which are challenging due to their intrinsic nonlinearity and ill-posedness. To anticipate the idea of an inverse scattering problem: For given near or far field measurements, we rely on the nonlinear forward operator mapping the contrast to the field measurements. We want to identify the contrast q of compact support, such that the forward operator $\mathcal{F}$ maps q to the measurements. Naturally, we tackle the inversion problem by seeking a contrast q such that $\mathcal{F}(q)$ matches data F^δ_{meas}, i.e. $\|\mathcal{F}(q) - F^\delta_{\text{meas}}\|$ is small in some appropriate norm. This chapter is a preparation of the Chs. 6, 7 and 8, that contain schemes to reconstruct the contrast. In addition, we consider some penalty terms and discuss their combined effect to the reconstruction result in Sec. 3.2, that is based on [BKL17, Sec. 4.3]. Furthermore, we look at some thoughts about numerical minimization techniques in Sec. 3.3 basing on [BKL17, Secs. 4.1 and 4.2] but adding an automatic parameter choice. Finally, in Sec. 3.4 we discuss the generating of synthetic data and the working with real-world data. This last section uses parts of [BKL17, Secs. 5 and 6].

Inverse Problem Before we consider the inverse scattering problem we will look at the definition of inverse problems in general. We have already defined a forward operator $\mathcal{F}$ in Ch. 2 to model physical reality. In general, such a forward operator is needed to define a so-called *mathematical model*, see e.g. [Rie03, Sec. 1.5], that is a function $\mathcal{F}\colon X \to Y$ from the set of causes (or parameters) X into the set of effects (or data) Y. Then the *direct problem* is to compute the effect from the cause, i.e. compute $\mathcal{F}(x)$ for a given $x \in X$. Further, the *inverse problem* is the reversed case, i.e. try to find a cause $x \in X$ from a given effect $y \in Y$ such that $\mathcal{F}(x) = y$.

The problem is called *well-posed* in the sense of Hadamard, see e.g. [Rie03, Sec. 1.5] or [Had23], if the following properties are valid: it exists a solution x to the equation $\mathcal{F}(x) = y$ for each $y \in Y$; this solution is unique; the inverse function $\mathcal{F}^{-1}\colon Y \to X$ is continuous, i.e. the solution x depends continuously on the data y. If one condition is violated, the problem is called *ill-posed*.

We cannot assume that a solution of the inverse problem exists. Therefore $x \in X$ should be chosen, such that the *discrepancy* (also called *residuum*), i.e. the *defect* between the *predicted data* $\mathcal{F}(x)$ and the *given data* y, is minimized, i.e. $\|\mathcal{F}(x) - y\|_Y \le \|\mathcal{F}(\varphi) - y\|_Y$ for all $\varphi \in X$, see e.g. [Rie03, Ch. 2]. This can be reformulated as an optimization problem, see [Rie03, Def. 9.5.1]:

$$\min_{x \in X} \|\mathcal{F}(x) - y\|_Y. \tag{3.1}$$

For the following remark we follow [Rie03, Sec. 1.5]. Usually, inverse problems are ill-posed because the inverse function is not continuous. This violation, in particular, is bad in the case of perturbed data because even a small noise level of measurements has a heavy influence on the reconstruction, i.e. the reconstruction error can be huge. It is possible to modify the topologies on X and Y such that continuity is enforced. However, the topologies are predetermined by the application.

It is well-known that in the case of an ill-posed problem the minimization of the discrepancy will in general not yield reasonable reconstructions, see e.g. [DDD04]. Therefore a *regularization* is needed to get a stable solution. To stabilize the process, a convex *penalty term* $\mathcal{P}$ (also called *regularization term*) is added to the discrepancy (3.1) such that the solution x depends continuously on the data y and the third condition of a well-posed problem is fulfilled. In literature, this general paradigm is called *variational regularization* or *Tikhonov regularization*. The resulting functional

$$q \mapsto \frac{1}{2}\|\mathcal{F}(q) - F^{\delta}_{\mathrm{meas}}\|^2_Y + \mathcal{P}(q) \tag{3.2}$$

is called *Tikhonov functional*.

In general, the penalty term is chosen such that a priori information on the solution x is taken into account. In Sec. 3.2 effective penalty terms and the related Tikhonov functional are presented, that are sensible in the context of inverse scattering problems.

Inverse Scattering Problems In fact, *inverse scattering problems* are commonly *ill-posed* because the inverse operator mapping the data (multi-static measurements) to the parameters (the obstacle's contrast) is not continuous (in any reasonable topology), see [DLP$^+$11, Sec. 4.2] and [CK13]. Therefore we assume some a priori information of the obstacle and include them by an appropriate penalty term. Furthermore, inverse scattering problems are *nonlinear* if the forward operator is nonlinear. Remember that our forward operator $\mathcal{F}$, presented in Ch. 2, is nonlinear.

Solving Nonlinear, Ill-Posed Inverse Scattering Problems There are three categories for algorithms solving nonlinear, ill-posed inverse scattering problems, see [LR06, Sec. 1]: noniterative algorithms based on global linearization; iterative solvers tackling the full nonlinear problem; direct methods.

The first and second category results in solving an optimization problem. We will tackle the full nonlinear problem with essentially two different algorithms: First, in Ch. 6 we will use the thresholded, nonlinear Landweber scheme that belongs to the iterative methods of the Newton type, see [Rie03, Sec. 7.5]. Second, in Ch. 7 we propose a variational reconstruction scheme in the class of Newton-like schemes.

Algorithms of the third category, direct methods, e.g. the linear sampling method from Colton and Kirsch, see [CK96], and the factorization method from Kirsch, see [Kir99], do only reconstruct the obstacle's shape instead of the obstacle's contrast. We will use the factorization method in Ch. 8 to have more a priori information for the proposed scheme in Ch. 7.

3.1 How can our Inverse Scattering Problem be Described?

The inverse scattering problem, i.e. the inverse problem to the direct scattering problem, which we considered in Sec. 2.1, is to look for the contrast q of the obstacle for known data (scattered fields) to known incident fields. As we focus on the *single-frequency* inverse scattering problem, i.e. the incident fields have the same fixed frequency, we omit the addition *single-frequency*. In contrast, the *multi-frequency* inverse scattering problem uses data from several fixed frequencies. We will formulate the continuous, the discretized and a simplified (single-frequency) inverse scattering problem.

Data To that end, the scattered field at the receivers' positions is called *data*. Further, simulated data with noise (or real-world data) is denoted by F^δ_{meas}, where $\delta > 0$ is the *relative noise level*, i.e. $\|\mathcal{F}(q) - F^\delta_{\text{meas}}\|_{\text{HS}} \leq \delta$, where HS is $\text{HS}(L^2(\Gamma_\text{i}), L^2(\Gamma_\text{s}))$, see Ch. 2, e.g. (2.14). Consequently, simulated data without noise is denoted by F_{meas}. For the sake of completeness, we mention that the synonyms *noise free data* and *exact data* for F_{meas} are used in the *user guide* and in the *source code*.

Continuous Inverse Scattering Problem For a description of the inverse scattering problem we recall the nonlinear forward operator $\mathcal{F}$, see (2.14) respectively (2.24). In the continuous version the operator $\mathcal{F}$ maps the contrast $q \in L^2_{\text{Im}\geq 0}(D)$ onto perturbed near field or far field data F^δ_{meas} in $\text{HS}(L^2(\Gamma_\text{i}), L^2(\Gamma_\text{s}))$. In the context of contrast reconstruction the task is to find a contrast q such that $\mathcal{F}(q)$ matches the data F^δ_{meas} as good as possible. Since the underlying problem is ill-posed, a sensible reconstruction can be computed by minimizing the following *Tikhonov functional*

$$q \mapsto \frac{1}{2}\|\mathcal{F}(q) - F^\delta_{\text{meas}}\|^2_{\text{HS}} + \mathcal{P}(q), \tag{3.3}$$

where the penalty term $\mathcal{P}(q)$ stabilizes the solution. This penalty term can consist of several summands. Each one is equipped with a factor, that is greater than zero. These factors are called *regularization parameters* and control the pay-off between the approximation in the discrepancy and the regularization term, [CP11a, Sec. 6.2.1].

Existence and Uniqueness of the Inverse Scattering Problem's Solution We will give a brief overview of statements belonging to existence and uniqueness of the inverse scattering problem's solution, i.e. the contrast q that fulfills $\mathcal{F}(q) \approx F^\delta_{\text{meas}}$ with relative noise level δ. They are restricted to near field data, see [LKK13]. Furthermore, it is assumed that $\text{Im}(q) \geq 0$. Further details are avoided for the sake of simplicity. The dimension of the problem is denoted by d.

If an exact solution of the noise free inverse scattering problem $\mathcal{F}(q) = F_{\text{meas}}$ is in $L^p(D)$ for $p > d/2$ with $d = 2$ or 3 and a family $\{F^\delta_{\text{meas}}\}_{\delta>0}$ satisfies $\|\mathcal{F}(q) - F^\delta_{\text{meas}}\|_{\text{HS}} \leq \delta$, where HS is $\text{HS}(L^2(\Gamma_\text{i}), L^2(\Gamma_\text{s}))$, see e.g. (2.14), then there exists a minimizer of $q \mapsto \frac{1}{2}\|\mathcal{F}(q) - F^\delta_{\text{meas}}\|^2_{\text{HS}} + \frac{\alpha}{p}\|q\|^p_{L^p(D)}$ on $L^p(D)$ for $p \in (1, \infty)$, where $\alpha = \alpha(\delta)$ appropriately, see [LKK13, Th. 10]. However, the most interesting case $p = 1$, that is formally not included in two dimensions, is also considered in numerical examples in [LKK13].

In the case of three dimensions, there is a proof for uniqueness for $p > 3/2$: in [LKK13, Th. 11] it was proofed for $p > 3$, the extension to $p > 3/2$ was proofed in [Ren17, Sec. 2.4].

Focus on Discretized Problem As we are focused on the implementation aspects, we only consider finite-dimensional spaces both for the searched-for contrast and the data (multi-static measurements). For simplicity, we omit to denote this discrete setting explicitly. We also remark that due to the shape of the region of interest the elements of the contrast q can be considered a vector or tuple ($\mathbb{C}^{N_\mathrm{D}}$) and at the same time a rectangular array ($\mathbb{C}^{N\times N}$ in 2D and $\mathbb{C}^{N\times N\times N}$ in 3D). To avoid overloading notation we will avoid to differentiate between these notations, whenever it is clear from the context which one is meant.

Forward Operator and his Spaces Therefore we do no longer need to differentiate in the notation between functions and their discretizations. To simplify notation in what follows, we no longer underline discrete quantities. In particular, from now on q shall refer to the discrete version of the contrast and $\mathcal{F}$, $\mathcal{F}'$ to the *discretized* forward operator and its derivative, see (2.24) and (2.25). The forward operator $\mathcal{F}$ is assumed to map between finite-dimensional spaces X and Y_dis with

$$\begin{array}{rcll} X & := & \mathbb{C}^{N_\mathrm{D}} & \text{equipped with the inner product (2.27) and} \\ Y_\mathrm{dis} & := & \mathbb{C}^{N_\mathrm{s}\times N_\mathrm{i}} & \text{equipped with the inner product (2.28).} \end{array} \tag{3.4}$$

The related norms are denoted as $\|\cdot\|_{\mathrm{roi},2}$ and $\|\cdot\|_\mathrm{dis}$. In both cases we identify the dual space of these Hilbert spaces with the spaces themselves. Sometimes, in cases where the inner product is not important, we will overload the notation and write $X = \mathbb{C}^{N_\mathrm{D}}$ and $Y_\mathrm{dis} = \mathbb{C}^{N_\mathrm{s}\times N_\mathrm{i}}$ to make the dimension of the space explicit.

Discretized Inverse Scattering Problem For a description of the discretized inverse scattering problem we recall the nonlinear forward operator $\mathcal{F}$, see (2.24). The operator $\mathcal{F}$ maps the contrast $q \in \mathbb{C}^{N_\mathrm{D}}$ onto perturbed near or far field data $F^\delta_\mathrm{meas} \in Y_\mathrm{dis} = \mathbb{C}^{N_\mathrm{s}\times N_\mathrm{i}}$, where N_s is the number of receivers and N_i the number of transmitters. In the context of contrast reconstruction we minimize, analogous to (3.3), the following *Tikhonov functional*,

$$q \mapsto \frac{1}{2}\|\mathcal{F}(q) - F^\delta_\mathrm{meas}\|^2_\mathrm{dis} + \mathcal{P}(q). \tag{3.5}$$

Simplified Inverse Scattering Problem For simplification we sometimes consider a *simplified Tikhonov functional*

$$q \mapsto \frac{1}{2}\|\mathcal{F}(q) - F^\delta_\mathrm{meas}\|^2_\mathrm{F} + \mathcal{P}(q). \tag{3.6}$$

with a Frobenius norm $\|\cdot\|_\mathrm{F}$. In particular, this simplification omits to respect the distances of the receivers to model physical reality. It is used in the next section to discuss appropriate penalty terms before presenting the full penalty terms.

3.2 Effective Penalty Terms

As already mentioned a penalty term is required to cope with the ill-posedness of the inverse problem, see [EHN96]. Choosing a combination of effective penalty terms is crucial for an appropriate reconstruction. We will consider penalty terms, which are designed to

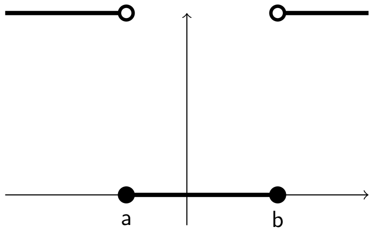


Figure 3.1: Indicator function $\delta_{[a,b]}$. The value of the function equals infinity if the input is outside the interval $[a, b]$ and it equals zero otherwise.

promote information expected in real-world environments. Some are well-known (sparsity and total variation) and one is unorthodox (physical bounds). However, the combination of sparsity and total variation is—to the author's best knowledge—new and results in a combined effectiveness. Furthermore, the resulting Tikhonov functional is presented.

First, we will consider simplified penalty terms fitting to the simplified Tikhonov functional (3.6), and second, the full penalty terms to (3.5). The combined effect of the chosen terms are demonstrated in Sec. 3.2.4.

3.2.1 Simplified Penalty Terms

In contrast to the next section we do not model physical reality to keep it simple. In addition, we use a simplified notation.

To take into account three kinds of a priori information about the contrast q, the ansatz for $\mathcal{P}(q)$ is the sum of the three following *simplified penalty terms*:

First, the contrast is assumed to be *sparse*, i.e. described by few non-zero coefficients, in the pixel basis. This is taken into account by a penalty term, weighted by a regularization parameter $\alpha > 0$, that is $\alpha \left\| q \right\|_1$, see e.g. [SKHK12, Sec. 1.5].

Second, the scattering object is supposed to have sharp edges (or is gradient-sparse). Thus, the *total variation* semi-norm is added as a penalty term using the notation $\beta \left\| \nabla q \right\|_1$, where $\beta > 0$ is an additional regularization parameter, see e.g. [CP11a, Sec. 6.2.1]. In general, for the restoration of images the total variation and not the L^2 norm is proper because total variation norms are essentially L^1 norms of derivatives and L^1 estimation procedures are more appropriate for the subject of image restoration, see [ROF92].

Third, an *indicator function* $\delta_{[a,b]}$ is defined: $\delta_{[a,b]}(q)$ equals infinity if one or more entries of q are outside a reasonable interval $[a, b]$ and it equals zero otherwise, see Fig. 3.1. This function is used to respect *physical bounds* of the real and imaginary part of the contrast by adding

$$\delta(q) = \delta_{[a,b]}(\mathrm{Re}(q)) + \delta_{[c,d]}(\mathrm{Im}(q)). \tag{3.7}$$

This forces the real and imaginary part of the contrast to be in $[a, b]$ and $[c, d]$. For example, if we know that the obstacle is a dielectricum we can take this into account by $\mathrm{Im}(q) = 0$. Altogether, the resulting penalty functional has the form $\mathcal{P}(q) = \alpha \left\| q \right\|_1 + \beta \left\| \nabla q \right\|_1 + \delta(q)$.

Enforce Sparsity We will consider why $\left\| q \right\|_1$ is used to enforce sparsity. The idea of *sparsity* is to express the solution by a few non-zero coefficients in an a priori known

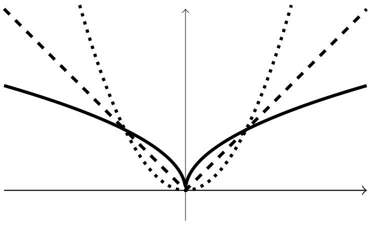


Figure 3.2: Plot of $f(x) = |x|^p$ with $p = 1/2$ (solid), $p = 1$ (dashed) and $p = 2$ (dotted).

basis. (We will focus on sparsity in the pixel basis.) To enforce that we have to make small coefficients expensive. Further, we have to minimize the Tikhonov functional (3.3) respectively (3.6). Therefore convex analysis, see Sec. 3.3.1, is helpful and finally, convexity is required.

We consider the penalty term $\|q\|_p^p$. If $p > 1$, e.g. $p = 2$, then we have a Hilbert space structure, but we do not enforce sparsity because small coefficients are not penalized; if $p < 1$, e.g. $p = 1/2$, then we enforce sparsity by penalizing small coefficients, but we get a non-convex problem, see [SKHK12, Sec. 1.5]. Finally, only the case $p = 1$ enforces sparsity and convexity. This is illustrated in Fig. 3.2 in one dimension.

The convergence analysis in [LKK13] for $q \mapsto \frac{1}{2}\|\mathcal{F}(q) - F^\delta_{\text{meas}}\|^2_{\text{HS}} + \frac{\alpha}{p}\|q\|^p_{L^p(D)}$ is for $p \in (1, \infty)$ as already mentioned in Sec. 3.1. The most interesting case $p = 1$, that enforces sparsity, is especially demanding and is still an open question.

Example A sparse contrast was already shown in Fig. 2.2(c). Essentially, this contrast has sharp edges as well, except for the gradual contrast ramp in the bottom. Of course, the total variation measures the size of a kind of gradient and a gradual contrast ramp results in a high total variation.

Remarks to Simplifications We have considered the simplified penalty terms. This referred to omit the related physical model and to use the simplified notation for the total variation semi-norm. Therefore we will present the full penalty terms in the next section, i.e. use $\|q\|_{\text{spa}}$ instead of $\|q\|_1$ as well as $\|\nabla q\|_{\text{tv}}$ instead of $\|\nabla q\|_1$ to incorporate the area or volume of a pixel/voxel. In addition, an exact definition of the total variation semi-norm will be given.

3.2.2 Full Penalty Terms

We will repeat the three already-mentioned assumptions to the contrast, but we model physical reality using the correct norms and define the total variation semi-norm precisely. The results are the *full penalty terms*.

We assume that the contrast array $q \in X = \mathbb{C}^{N_\text{D}}$ is sparse with respect to the pixel basis. This is enforced in the reconstructions by means of the penalty term, see [SKHK12, Sec. 1.5], $\alpha\,\|q\|_{\text{spa}}$, where

$$\|x\|_{\text{spa}} := h_N^d \sum_i \left(|\operatorname{Re} x_i| + |\operatorname{Im} x_i|\right), \quad x \in \mathbb{C}^{N_\text{D}},$$

and $\alpha > 0$ is a regularization parameter. (Remember that h_N is the mesh size, defined in Sec. 2.2, and h_N^d the area or volume of a pixel/voxel.)

In addition, we assume that the ground truth contrasts exhibit sharp edges. Therefore we introduce the *total variation* semi-norm, see [CP11a, Sec. 6.2.1]. In particular, in this paragraph the elements of X are interpreted as multi-dimensional arrays. A *discretization of the gradient* acting on square arrays is defined by

$$(\nabla a)_{i,j} = \begin{pmatrix} (\nabla a)_{i,j}^{(1)} \\ (\nabla a)_{i,j}^{(2)} \end{pmatrix} \quad \text{with } (\nabla a)_{i,j}^{(1)} := \begin{cases} \frac{a_{i+1,j}-a_{i,j}}{h} & \text{if } 1 \leq i < N, \\ 0 & \text{if } i = N, \end{cases}$$
$$\text{and } (\nabla a)_{i,j}^{(2)} := \begin{cases} \frac{a_{i,j+1}-a_{i,j}}{h} & \text{if } 1 \leq j < N, \\ 0 & \text{if } j = N, \end{cases} \tag{3.8}$$

where N is the size of the matrix in one direction and h is the mesh size of the underlying grid, cf. [CP11a, Sec. 6.1]. This operator can be implemented efficiently via shift operations[◇]. The three-dimensional case is straightforward to the two-dimensional case. Note that the above definition differs from the definition of the routine gradient in MATLAB, see [Mat15], when implementing the proposed algorithm. Next, the *total variation (TV)* semi-norm is defined in the case of two dimensions as

$$\|\nabla q\|_{\text{tv}} := h_N^2 \sum_{i,j} |(\nabla q)_{i,j}|, \quad \text{with} \quad |(\nabla q)_{i,j}| := \sqrt{|(\nabla q)_{i,j}^{(1)}|^2 + |(\nabla q)_{i,j}^{(2)}|^2}, \tag{3.9}$$

where the factor h_N^2 is due to the two-dimensional volume measure on the region of interest D. (The 3D case is analogous. Of course, h_N^2 has to be replaced by h_N^3.)

Note that, in general, $\|\nabla q\|_{\text{tv}}$ differs from the ℓ^1 norm of ∇q (if ∇q is written as a single vector). To promote sharp edges in the reconstruction the penalty $\beta \|\nabla q\|_{\text{tv}}$, where $\beta > 0$ is an additional regularization parameter, is included into the Tikhonov functional.

Finally, in general, characteristic *physical bounds* of the values for the contrast are known. As already mentioned, we incorporate this information using the indicator function $\delta_{[a,b]}$ by $\delta(q) := \delta_{[a,b]}(\text{Re}(q)) + \delta_{[c,d]}(\text{Im}(q))$.

Finally, the sum of the penalty terms is $\mathcal{P}(q) = \alpha \|q\|_{\text{spa}} + \beta \|\nabla q\|_{\text{tv}} + \delta(q)$.

3.2.3 Resulting Tikhonov Functional

Collecting the simplified respectively full penalty terms, see Secs. 3.2.1 and 3.2.2, the *simplified Tikhonov functional* (3.6) results in

$$\min_{q \in X} f(q), \quad f(q) := \underbrace{\tfrac{1}{2}\|\mathcal{F}(q) - F_{\text{meas}}^{\delta}\|_{\text{F}}^2}_{=:f_{\text{dis}}} + \underbrace{\alpha \|q\|_1}_{=:f_{\text{spa}}} + \underbrace{\beta \|\nabla q\|_1}_{=:f_{\text{tv}}} + \underbrace{\delta(q)}_{=:f_{\text{phy}}}, \tag{3.10}$$

and the *full Tikhonov functional* (3.5) yields the problem

$$\min_{q \in X} f(q), \quad f(q) := \underbrace{\tfrac{1}{2}\|\mathcal{F}(q) - F_{\text{meas}}^{\delta}\|_{\text{dis}}^2}_{=:f_{\text{dis}}} + \underbrace{\alpha \|q\|_{\text{spa}}}_{=:f_{\text{spa}}} + \underbrace{\beta \|\nabla q\|_{\text{tv}}}_{=:f_{\text{tv}}} + \underbrace{\delta(q)}_{=:f_{\text{phy}}}, \tag{3.11}$$

[◇] For an implementation in 2D and 3D circular shifting via the MATLAB routine circshift is useful.

where $\alpha, \beta > 0$ are regularization parameters. The summands are the discrepancy $f_{\mathrm{dis}}(q)$, the sparsity penalty $f_{\mathrm{spa}}(q)$, the total variation penalty $f_{\mathrm{tv}}(q)$ and the physical bounds $f_{\mathrm{phy}}(q)$. We take up these terms and symbols also in the *user guide* as well as the *source code*. Note that the Tikhonov functional $f(q)$ involves a nonlinear operator and will therefore in general fail to be convex.

We have defined f_{dis}, f_{spa}, f_{tv} and f_{phy} in both cases (simplified and full Tikhonov functional). In general, we consider for both

$$\min_q f(q), \quad f(q) := f_{\mathrm{dis}}(q) + f_{\mathrm{spa}}(q) + f_{\mathrm{tv}}(q) + f_{\mathrm{phy}}(q). \tag{3.12}$$

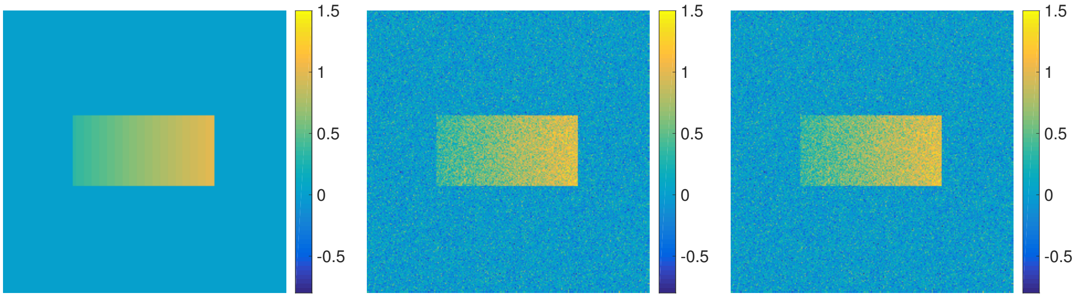
Alternative Discrepancy and Penalty Terms Of course, it is possible to use other discrepancy and penalty terms in (3.11). For example, for different noise statistics other discrepancy terms could be appropriate, e.g. unweighted Frobenius norm as in (3.10), Schatten norm etc. For the penalty, one could consider the sparsity of the contrast in another system, e.g. a wavelet or other advanced harmonic basis. For example, for wavelets this results in the penalty function $\alpha \|\mathcal{W}(q)\|_1$, where $\mathcal{W}$ is the forward wavelet operator. However, as already mentioned the corresponding reconstruction did not outcompete the sparsity in the pixel basis. For the sake of completeness, it should be mentioned that this reconstruction basing on a wavelet basis omits the total variation-based regularization. (Of course, it was omitted in the comparative computation basing on the pixel basis too.)

3.2.4 Motivation of the Combination of Sparsity and Total Variation

To show the combined effectiveness of these penalty terms, in particular sparsity and total variation—the effect of physical bounds is straightforward—, we replace the forward operator $\mathcal{F}$ and ignore the possibility to restrict the contrast by physical bounds. Then the simplified Tikhonov functional (3.10) is reduced to

$$\min_q f(q), \quad f(q) = \tfrac{1}{2}\|q - q^\delta\|_{\mathrm{F}}^2 + \alpha\, \|q\|_1 + \beta\, \|\nabla q\|_1.$$

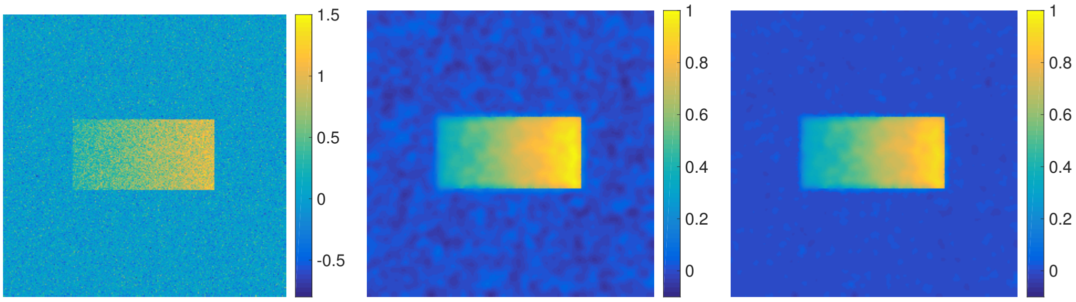
The combined effectiveness is shown in Fig. 3.3: Leaving out any penalty term as in (c) results in a not convincing reconstruction. We get a similar result if we use only sparsity, see (d). Using only total variation (TV) results in a reasonable reconstruction, see (e). The background artifacts essentially vanish if we additionally use sparsity, see (f), while the error decreases only slightly from 14% to 13%. Remember that the relative noise level is 20% such that a higher iteration number is not justified but emphasizes the effects of sparsity and total variation here. TV in general preserves edge locations and reduces the contrast. A detailed discussion of the effects of total variation is given in [SC03]. In (g) we use only TV and can observe the preservation of the edges and a loss of contrast in comparison to the exact contrast in (i). Furthermore, the high number of iteration results in step artifacts. Using both sparsity and total variation, see (h), has the same effects. In addition, we can observe no background artifacts and slightly improved edges as a result of the sparsity penalty term. The price of the sparsity penalty realized by a soft-shrinkage operator, defined in (3.13), is a slight shrinkage of the absolute values, i.e. a loss of contrast, that finally results in a higher error than without the sparsity penalty. In fact, the number of iteration is too high in (g) and (h).



(a) Exact contrast.

(b) Data with noise.

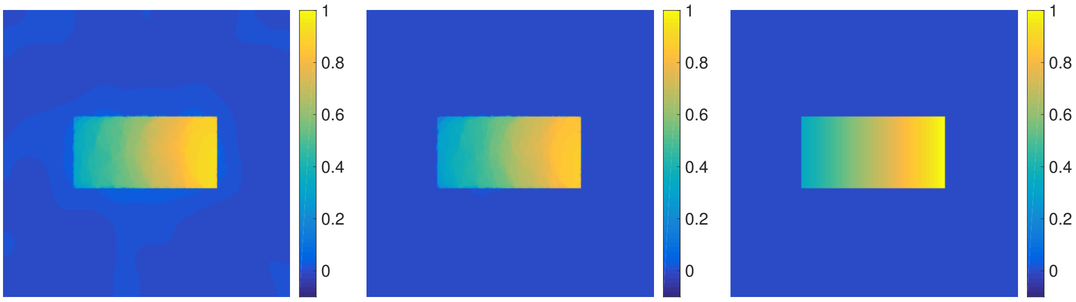
(c) Sparsity no, TV no, $N_{\text{in}} = 50$, err. 115%.



(d) Sparsity yes, TV no, $N_{\text{in}} = 50$, err. 95%.

(e) Sparsity no, TV yes, $N_{\text{in}} = 50$, err. 14%.

(f) Sparsity yes, TV yes, $N_{\text{in}} = 50$, err. 13%.



(g) Sparsity no, TV yes, $N_{\text{in}} = 1000$, err. 7%.

(h) Sparsity yes, TV yes, $N_{\text{in}} = 1000$, err. 12%.

(i) Exact contrast.

Figure 3.3: Simple example to demonstrate the combined effectiveness of sparsity and total variation. All figures show the real parts. Note that the imaginary part of the contrast is zero and essentially vanishes in reconstructions too. (a) and (i): Exact contrast (using different colorbars' limits). (b) Simulated data with noise level $\delta = 20\%$. The other figures (c)–(h) show the reconstructions. The usage of sparsity and total variation is controlled by the regularization parameters α or β. If we use them, we set $\alpha = 0.05$ for sparsity and $\beta = 0.5$ for total variation (TV); otherwise they are 0. The reconstructions are stopped after a specific number of inner iterations N_{in}. We give the relative reconstruction error. Note that we have changed the colorbars' limits in (e)–(i) in comparison to (a)–(d) because the better reconstructions.
In all examples it were chosen $N = 128$ discretization points for each dimension and a relative noise level $\delta = 20\%$. The examples use the same data with noise. As the replacement of $\mathcal{F}$ with I we tackle a linear problem. Therefore these examples are computed by the inner iterations of the developed variational reconstruction scheme, i.e. the iterations of the primal-dual algorithm, see Ch. 7.

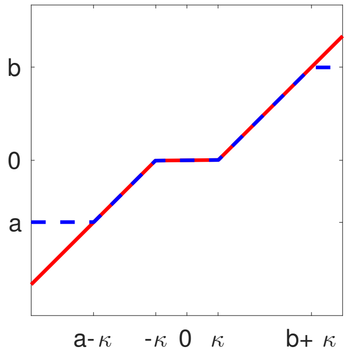


Figure 3.4: Soft-shrinkage (red continuous) and extended soft-shrinkage (blue dashed).

3.3 Numerical Minimization Techniques

After identifying reasonable penalty terms and their combined effectiveness, we are interested in minimizing the Tikhonov functional (3.12), i.e. $\min_q f(q)$ with $f(q) := f_{\mathrm{dis}}(q) + f_{\mathrm{spa}}(q) + f_{\mathrm{tv}}(q) + f_{\mathrm{phy}}(q)$. Therefore this section includes a rough overview of the numerical minimization techniques. The details are discussed in Chs. 6 and 7.

To take into account all penalty terms of (3.12) we will consider so-called *proximal splitting methods*, see [CP11b], that tackle convex optimization problems of the form $\min_{x\in\mathbb{R}^n}\{f_1(x) + \ldots + f_m(x)\}$ with convex functions $f_i\colon \mathbb{R}^n \to (-\infty, +\infty]$. The idea of splitting is to consider the functions f_i for $i = 1, \ldots, m$ individually for an easier implementation. They are called proximal because nonsmooth functions are involved via its so-called *proximal mapping*, defined in Lemma 3.3.2. So, nondifferentiable functions can be processed.

Soft-Shrinkage Operator In [LKK13] the minimization of $f_{\mathrm{dis}} + f_{\mathrm{spa}}$ was achieved by the *thresholded, nonlinear Landweber scheme*, discussed in Ch. 6, based on the so-called *soft-shrinkage operator*◇, that is defined for real-valued x and $\kappa > 0$ by, see e.g. [DDD04] or [CS05],

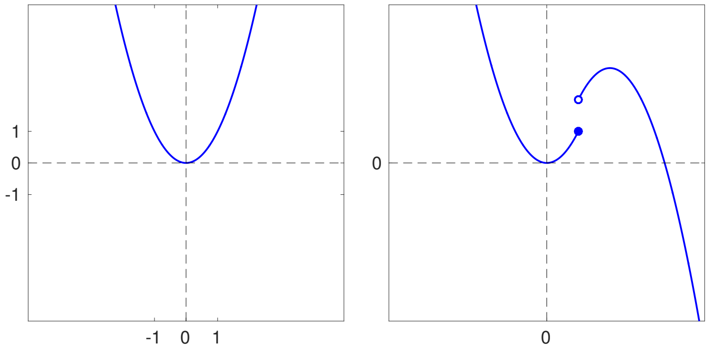
$$\mathbb{S}(x,\kappa) := \begin{cases} x+\kappa & \text{if } x \leq -\kappa, \\ 0 & \text{if } x \in (-\kappa, +\kappa), \\ x-\kappa & \text{if } x \geq \kappa \end{cases} \tag{3.13}$$

and depicted in Fig. 3.4 as red continuous. Of course, the element-wise application is assumed if x is a vector. The soft-shrinkage operator can be implemented by $\mathbb{S}(x,\kappa) = \mathrm{sign}(x)\max\{|x| - \kappa, 0\}$.

Extended Soft-Shrinkage Operator A result of this thesis is an extension of the soft-shrinkage operator, that additionally respects the penalty for physical bounds f_{phy} to minimize $f_{\mathrm{dis}} + f_{\mathrm{spa}} + f_{\mathrm{phy}}$. This extension is straightforward by cutting the soft-shrinkage

---

[◇] The *soft-shrinkage operator* is continuous in comparison to the so-called *hard-thresholding operator*, see e.g. [CS05, Ch. 4]. As already mentioned the *soft-shrinkage operator* is called *shrinkage operator* in [DDD04]. In addition, note that in [BKL17, Sec. 4.7] a typo states $\kappa \geq 0$ instead of $\kappa > 0$.



(a) Parabola.
(b) l.s.c.

Figure 3.5: (a) Parabola as an example of a convex function. (b) Example of a lower semi-continuous function. (If point and circle were changed, it would be upper semi-continuous.)

operator, see the blue dashed in Fig. 3.4. It receives the name *extended soft-shrinkage operator*.

Both, incorporating the sparsity penalty (via the soft-shrinkage operator) as well as the sparsity combined with physical bounds (via the extended soft-shrinkage operator) in the thresholded, nonlinear Landweber scheme are described in Ch. 6.

Primal-Dual Algorithm To bear in mind the total variation f_{tv} as well, i.e. completing the Tikhonov functional (3.12), a more sophisticated method is required. Therefore we will employ a so-called *primal-dual algorithm (PDA)*, see [PCBC09, CP11a], and thus have to deal with two problems. First, the forward operator $\mathcal{F}$ is nonlinear, whereas PDA requires a linear operator. Hence, a linearization is needed. Second, the contrast q is complex, but PDA requires real vector spaces. Therefore we need a transformation operator. The primal-dual algorithm and its application to our problem is discussed in Ch. 7.

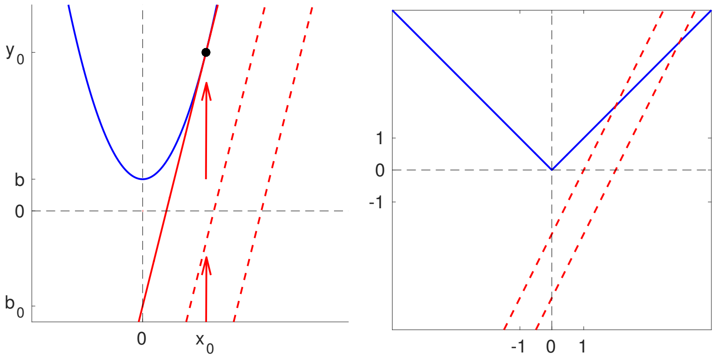
As a preparation, the next section provides some concepts of convex analysis, that are important for the mentioned numerical minimization techniques. The reader familiar with this topic may skip it.

3.3.1 Convex Analysis

As a preparation of Chs. 6 and 7 we will consider some concepts of convex analysis.

For the following definitions, see [SKHK12, Ch. 2] and [Roc67, Sec. 2], let X be a Banach space and $f\colon X \to \mathbb{R}\cup\{+\infty\}$ a *convex* function, i.e. $f(\lambda x+(1-\lambda)y) \leq \lambda f(x)+(1-\lambda)f(y)$ for all $x, y \in X$ and for all $\lambda \in [0,1]$, see e.g. Fig. 3.5(a). It is really essentially for the following definitions that we can order the codomain.

The convex function f is *lower semi-continuous (l.s.c.)* if the convex level set $\{x \in X : f(x) \leq c\}$ is closed for every $c \in \mathbb{R}$, see Fig. 3.5(b) for an example. The *effective domain* of the convex function f is $\text{dom}(f) := \{x \in X : f(x) < \infty\}$. The convex function f is *proper* if $\text{dom}(f) \neq \emptyset$. Of course, this is a useful requirement for a function to be minimized—otherwise we would try to minimize $f = \infty$.



(a) Fenchel conjugate on parabola. (b) Fenchel conjugate on absolute value function.

Figure 3.6: Geometric interpretation of the Fenchel conjugate. In (a) the Fenchel conjugate of the parabola with $f(x) = x^2 + 1$ at $x^* = 4$ is determined taking a straight line with slope $x^* = 4$ and shift it up until it touches the parabola; then the Fenchel conjugate is the negative y-intercept of this straight line, i.e. $f^*(4) = -b_0 = 3$. In (b) we look for the Fenchel conjugate of the absolute value function with $g(x) = |x|$ at $x^* = 2$; the straight line with slope $x^* = 2$ always cuts the function g; therefore the y-intercept has to be minus infinity; therefore the Fenchel conjugate is ∞; in general, $g^*(x^*) = 0$ for $|x^*| \leq 1$ and $g^*(x^*) = \infty$ for $|x^*| > 1$.

Fenchel Conjugate We still consider a convex function $f\colon X \to \mathbb{R} \cup \{+\infty\}$ with a Banach space X. The *Fenchel conjugate*$^\diamond$ is defined, see [Roc67], as†

$$f^*(x^*) := \sup_{x \in X} \left\{ \langle x, x^* \rangle - f(x) \right\}, \quad x^* \in X^*, \tag{3.14}$$

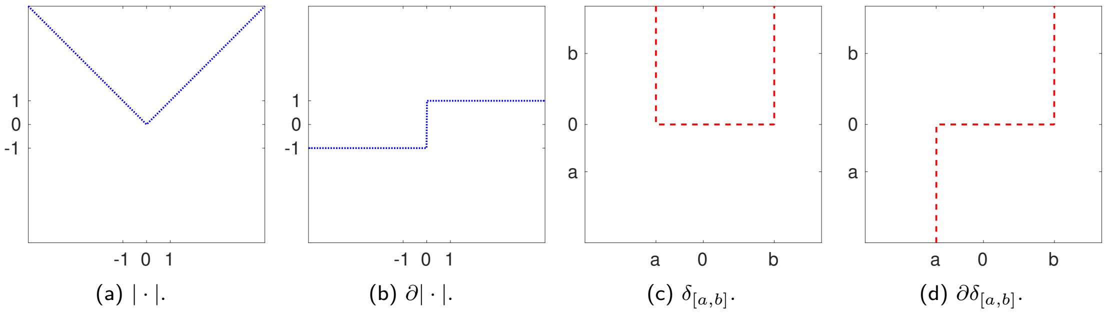
where X^* is the dual space of X. We will consider real finite-dimensional vector spaces with inner product. Therefore Riesz representation theorem delivers that they are isometrically isomorphic, i.e. $X \simeq X^*$. So, we identify X with its dual space, i.e. $X = X^*$. Actually, $\langle \cdot, \cdot \rangle$ in the definition is a dual pairing, but it is an inner product in the case of Hilbert spaces.

In a geometric interpretation the Fenchel conjugate is the negative part of the y-intercept of the tangent on the function f at the point, where f has the slope x^*. In general a tangential with slope x_0^* and y-intercept b_0 is denoted by $g(x) = x_0^* x + b_0$. Therefore the Fenchel conjugate $f^*(x_0^*)$ is $-b_0$, see [ZFM09]. This geometric interpretation is illustrated in Fig. 3.6(a)/(b).

Theorem 3.3.1 (Fenchel conjugate, see [RW98, Th. 11.1])**.** *If f on X is a proper, l.s.c. and convex function, then f^* on the dual space X^* is proper, l.s.c. and convex too. In a similar way the Fenchel biconjugate is defined via $f^{**}(x) = \sup_{x^* \in X^*} \{\langle x, x^* \rangle - f^*(x^*)\}$ for $x \in X$. It turns out that the Fenchel conjugation is a bijection between l.s.c. proper convex functions. In particular, for such functions $f = f^{**}$.*

[◇] Sometimes, in the literature the term *Fenchel dual* or *convex conjugate* is employed for this quantity. In addition, it is called Legendre-Fenchel transform, see [RW98]. In physics it is often referred to the Legendre transform, but note that the Legendre-Fenchel transform is in fact a generalization of the Legendre transform, see [Tou05, Remark 2]. The Legendre transform requires a convex and additionally smooth function f, see [ZFM09, Sec. 2].

[†] Note that the supremum is needed for uniqueness and is a natural choice in the definition because $-f(x)$ is concave.



(a) $|\cdot|$. (b) $\partial|\cdot|$. (c) $\delta_{[a,b]}$. (d) $\partial\delta_{[a,b]}$.

Figure 3.7: Functions and their subdifferentials: (a) Absolute value function. (b) The subdifferential of the absolute value function is the set-valued signum function. (c) Indicator function. (d) Subdifferential of the indicator function.

Subdifferential We still consider a convex function $f\colon X \to \mathbb{R} \cup \{+\infty\}$ with a Banach space X. An element x^* of the dual space X^* is a *subgradient* of f at $x \in X$ if $f(y) \geq f(x) + \langle x^*, y - x\rangle$ holds for all $y \in X$, see e.g. [SKHK12, Ch. 2] or [Roc67, Sec. 2]. The set of all subgradients of f at x is called the *subdifferential* of f at x. It is denoted by $\partial f(x)$.

The notions of subgradients and subdifferentials are a generalization of the gradient to (possibly) non-smooth convex functions. In particular, if X is finite-dimensional and f is continuously differentiable at x, then the subdifferential $\partial f(x)$ has exactly one element, namely the gradient $\nabla f(x)$. In fact, in that case we do not distinguish between the set and its sole element and write $\partial f(x) = \nabla f(x)$.

For an example of the subdifferential we consider the absolute value function $|\cdot|$ in $\mathbb{R}$, see Fig. 3.7(a). Of course, for the input $x > 0$ the slope is 1 and for $x < 0$ the slope is -1. There is no slope at $x = 0$, but if we try to draw a tangent its slope should be between -1 and 1, see Fig. 3.7(b). The definition of the subdifferential follows this idea.

For another example we consider the indicator function $\delta_{[a,b]}$, defined in Sec. 3.2.1, in one dimension, see Fig. 3.7(c). For the input $x < a$ and $x > b$ exists no subdifferential; for $x = a$ it is between $-\infty$ and 0, for $x = b$ it is between 0 and ∞; for x between a and b it is 0; see Fig. 3.7(d).

The subdifferentials of these two examples are

$$\partial|x| = \begin{cases} \{-1\} & \text{if } x < 0, \\ [-1,1] & \text{if } x = 0, \\ \{+1\} & \text{if } x > 0, \end{cases} \quad \text{and} \quad \partial\delta_{[a,b]}(x) = \begin{cases} (-\infty,0] & \text{if } x = a, \\ \{0\} & \text{if } x \in (a,b), \\ [0,\infty) & \text{if } x = b, \\ \emptyset & \text{otherwise.} \end{cases} \tag{3.15}$$

The first subdifferential will return as set-valued signum function sign in Sec. 6.1 and with the generalized notation $\operatorname{step}_{[-1,1]}$ in Sec. 6.2. The second subdifferential will return as $\operatorname{step}^{-1}_{[a,b]}$ in Sec. 6.2 as well. Indeed, $\operatorname{step}^{-1}_{[a,b]}$ is the inverse of $\operatorname{step}_{[a,b]}$.

Lemma 3.3.2 (Resolvent operator, proximal mapping, see [BL11, Lemma 6.134])**.** *Let X be a real Hilbert space and $G : X \to \mathbb{R} \cup \{+\infty\}$ a proper, convex and l.s.c. function.*

For all $\tau > 0$ and all $w \in X$ it holds that

$$(I + \tau\partial G)^{-1}(w) = \operatorname*{arg\,min}_{z\in X} \left\{ \frac{\|z-w\|_X^2}{2} + \tau G(z) \right\}.$$

Remark 3.3.3. In the literature sometimes a distinction is made between the left hand side operator $(I + \tau\partial G)^{-1}$ and the operator implicitly defined by the right hand side variational problem. The former is called *resolvent operator* while the latter is called *proximal mapping*. Since the assumptions of the lemma will always be fulfilled in our setting we will uniformly use the term proximal mapping. Note that the clue of this lemma is that the resolvent delivers the unique minimizer.

Indicator Functions The proximal mapping may be viewed as a generalization of the orthogonal projection. To see this, proximal mappings of indicator functions are considered. For a subset $C \subset X$ of a Hilbert space X we define an *indicator function*

$$\delta_C(w) := \begin{cases} 0 & \text{if } w \in C, \\ +\infty & \text{otherwise.} \end{cases}$$

Slightly overloading the notation we define for real-valued vectors w the element-wise indicator function on $[a, b]$,

$$\delta_{[a,b]}(w) := \sum_i \delta_{[a,b]}(w_i). \tag{3.16}$$

In other words, $\delta_{[a,b]}(w)$ is zero if every element w_i is in the interval $[a, b]$ and $\delta_{[a,b]}(w)$ is ∞ otherwise (as in the definition of $\delta_{[a,b]}$ in Sec. 3.2.1).

Remark 3.3.4 (Orthogonal Projection). We recall that the *orthogonal projection P_C* onto the closed convex set $C \subset X$ of a Hilbert space X, see [LMM12, Sec. 2.1], is defined as $P_C(w) = \frac{1}{2} \operatorname{arg\,min}_{z\in C} \|z-w\|_X^2$. Clearly, with the indicator function δ_C the projection operator P_C can also be written as an unconstrained minimization problem $P_C(w) = \operatorname{arg\,min}_{z\in X}\{\frac{1}{2}\|z-w\|_X^2 + \tau\delta_C(z)\}$ with arbitrary $\tau > 0$. However, this is exactly the definition of the proximal mapping for the indicator function.

3.3.2 When Should a Reconstruction Stop?

The *relative error* of the reconstruction is

$$\operatorname{err}(m) := \frac{\|q^{(m)} - q_{\text{exa}}\|_2}{\|q_{\text{exa}}\|_2},$$

where q_{exa} is the exact contrast, i.e. the ground truth. It would be desirable to stop the reconstruction scheme, when the relative error is minimal. However, the exact contrast q_{exa} is unknown in general. Therefore the discrepancy (defect) is consulted in the following discrepancy principle.

Definition 3.3.5 (Morozov's discrepancy principle, see [Rie03, Sec. 3.4])**.** Let $m = 1, \dots, N$ be the index of the iteration, then that iteration is stopped after iteration

$m = N$ if the *relative discrepancy*,

$$\mathrm{dis}(m) := \frac{\|\mathcal{F}(q^{(m)}) - F^\delta_{\mathrm{meas}}\|_{\mathrm{dis}}}{\|F^\delta_{\mathrm{meas}}\|_{\mathrm{dis}}}, \tag{3.17}$$

is for the first time less than a tolerance $\tau_{\mathrm{dis}} > 1$ times the relative noise level $\delta > 0$, i.e. it is stopped after N iterations

$$\begin{aligned} &\text{if} \quad \|\mathcal{F}(q^{(N)}) - F^\delta_{\mathrm{meas}}\|_{\mathrm{dis}} \le \tau_{\mathrm{dis}}\,\delta\,\|F^\delta_{\mathrm{meas}}\|_{\mathrm{dis}}, \\ &\text{whilst} \quad \|\mathcal{F}(q^{(m)}) - F^\delta_{\mathrm{meas}}\|_{\mathrm{dis}} > \tau_{\mathrm{dis}}\,\delta\,\|F^\delta_{\mathrm{meas}}\|_{\mathrm{dis}} \quad \text{for } m = 0, \ldots, N-1. \end{aligned} \tag{3.18}$$

3.3.3 Automatic Choice of Regularization Parameters?

A further interesting topic is an *automatic choice of the regularization parameters* α and β. Of course, this is helpful for any practitioner. In particular, as this choice is dependent on the relative noise level δ. However, it is difficult because we have to set two regularization parameters α and β. It is more simple to define β in dependence of α via a quotient ρ, whose choice will be described in Remark 3.3.7,

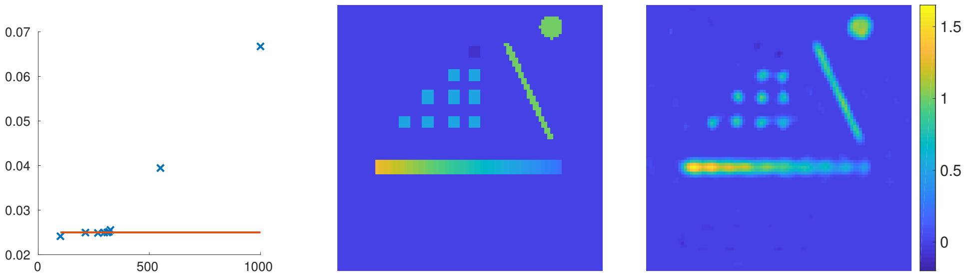
$$\beta(\alpha) = \rho\,\alpha \quad \text{e.g. with } \rho = 10^{-5}/500,$$

and look for a suitable $\alpha > 0$. We want to choose α as large as possible, such that the stabilizing penalty terms have a high effect; at the same time we like α to be small, such that the relative discrepancy is below the discrepancy principle threshold $\tau_{\mathrm{dis}}\,\delta$, see (3.18). Therefore the idea of the following algorithm is: first, find an α called a that fulfills the discrepancy principle; second, find a larger α called b that does not fulfill it; third, bisect the interval $[a, b]$. The whole heuristic algorithm for an automatic, data-driven parameter choice is described in Alg. 3.3.6. An example is given in Fig. 3.8.

Algorithm 3.3.6 (Automatic Choice of Regularization Parameters α and β)**.** In this description of the algorithm, the computation of the discrepancy dis(·) does mean to start the reconstruction process with α as well as corresponding $\beta(\alpha)$ and to stop it by the discrepancy principle or by reaching the maximal number of iterations. The following steps result in a reasonable parameter choice α_* for the regularization parameter α.

1. Initialization:
   a) Define a quotient ρ, e.g. $\rho = 10^{-5}/500$, therewith β depends of α by $\beta(\alpha) = \rho\,\alpha$.
   b) Define an initial guess of the regularization parameter α, e.g. $\alpha_1 = 10^4$.[◇]
   c) Set $i = 1$.
2. The aim of this step is to find suitable regularization parameters α stored as a and b such that the corresponding discrepancies dis(a) and dis(b) lie below, respectively above, the discrepancy principle threshold $\tau_{\mathrm{dis}}\,\delta$. Therefore we distinguish between two cases: If $\mathrm{dis}(\alpha_1) \le \tau_{\mathrm{dis}}\,\delta$, i.e. the discrepancy principle is fulfilled, then continue with step 2a. Otherwise go to step 2b.

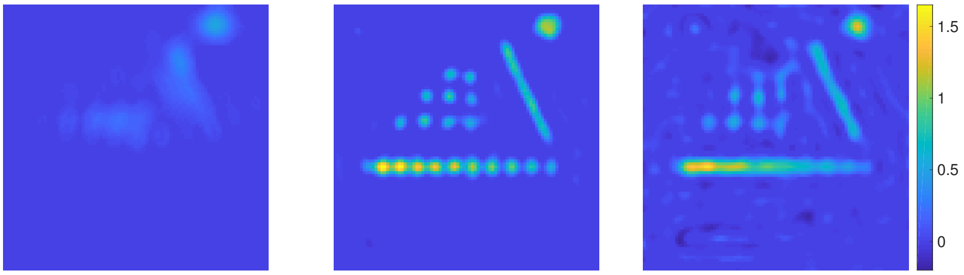
[◇] Of course, with regard to the run time it is advantageous to choose an initial guess α_1 near to the result α_*—if a reasonable guess is possible. In particular, it is advantageous with respect to the run time to fulfill that α_* is between α_1 and $10 \cdot \alpha_1$ or between $\alpha_1/10$ and α_1.



(a) Steps of parameter choice: rel. dis. over parameter α.

(b) Predefined contrast.

(c) Recon. with $\alpha = 310$.



(d) Recon. with $\alpha = 10^4$.

(e) Recon. with $\alpha = 1000$.

(f) Recon. with $\alpha = 100$.

Figure 3.8: Automatic choice of regularization parameters, see Alg. 3.3.6, using the proposed variational reconstruction scheme, see Alg. 7.5.1, and the *borehole* case as experimental set-up, that is described in detail in Ch. 5: (a) Steps of the automatic parameter choice showing the relative discrepancy dis over the regularization parameter α. (Remember that β is computed in dependence of α.) The red solid line is the discrepancy principle border $\tau_{\text{dis}}\,\delta$ with noise level $\delta = 0.01$ and tolerance $\tau_{\text{dis}} = 2.5$. (b)–(f) Real parts of contrasts.

(b) Predefined contrast. (This contrast is also used in Fig. 5.2(i)). (c) Reconstruction with the result of the automatic parameter choice, $\alpha = 310$. (This result is shown again in Fig. 5.2(k) in comparison to other reconstruction schemes.)

The steps of the automatic parameter choice: The automatic parameter choice starts with $\alpha = 10^4$. However, this value is omitted in (a). The corresponding reconstruction with $\alpha = 10^4$ is depicted in (d) and is as an example of overregularization. (Clearly, this is too sparse and an even higher value would end up in a contrast that is zero everywhere.) In keeping with this, the discrepancy principle was not fulfilled. Therefore a reconstruction is computed with $\alpha = 1000$. The reconstruction is slightly overregularized, see (e). The discrepancy principle is not fulfilled, see (a). Therefore a reconstruction is computed with $\alpha = 100$, see (f), that is slightly underregularized. Indeed, the discrepancy principle is fulfilled, see (a). Therefore an α between 100 and 1000 is searched by a kind of bisection method, again see (a) and, for the result, (c).

The whole automatic parameter choice stopped after 10 iterations and a run time of 2.3 h successfully because the criterion $(c - a)/a \leq 0.01$ was fulfilled, see step 3b. The result is $\alpha = 310.938$ (and $\beta = 6.219 \cdot 10^{-6}$) with relative discrepancy 0.025 and relative error 0.337. If rounded values are used it is recommended to round down to fulfill the discrepancy principle.

   a) Case $\mathrm{dis}(\alpha_1) \leq \tau_{\mathrm{dis}}\,\delta$: In this case we redefine $\alpha = \alpha \cdot 10$ until the discrepancy principle is *not* fulfilled, i.e. in detail:
      i. Redefine $\alpha_{i+1} := \alpha_i \cdot 10$. Set $i := i + 1$. Compute $\mathrm{dis}(\alpha_i)$.
      ii. If $\mathrm{dis}(\alpha_i) \leq \tau_{\mathrm{dis}}\,\delta$, go back to step 2(a)i. Otherwise continue.
      iii. Define $a := \alpha_{i-1}$ and $b := \alpha_i$ and go to step 3.
   b) Case $\mathrm{dis}(\alpha_1) > \tau_{\mathrm{dis}}\,\delta$: In this case we redefine $\alpha = \alpha/10$ until the discrepancy principle *is* fulfilled, i.e. in detail:
      i. Redefine $\alpha_{i+1} := \alpha_i/10$. Set $i := i + 1$. Compute $\mathrm{dis}(\alpha_i)$.
      ii. If $\mathrm{dis}(\alpha_i) > \tau_{\mathrm{dis}}\,\delta$, go back to step 2(b)i. Otherwise continue.
      iii. Define $a := \alpha_i$ and $b := \alpha_{i-1}$ and continue with step 3.
3. Use a kind of bisection method to find a suitable $\alpha_* \in [a, b]$, i.e. in detail:
   a) Define $c := a + (b - a)/2$ and compute $\mathrm{dis}(c)$.
   b) If $\mathrm{dis}(c) \leq \tau_{\mathrm{dis}}\,\delta$, then define $a := c$ and go back to step 3a. Otherwise define $b := c$ and go to step 3a again except for two cases: first, if $|\mathrm{dis}(c) - \mathrm{dis}(a)| \leq \delta/10$ and second, if $(c-a)/a \leq 0.01$. If one of these conditions is satisfied, then break the algorithm and choose $\alpha_* = a$ (because it is the largest α fulfilling the discrepancy principle).

The reader should keep in mind that the key point of this heuristic method is the discrepancy principle threshold $\tau_{\mathrm{dis}}\,\delta$ depending on our choice $\tau_{\mathrm{dis}} > 1$. Therefore a wise choice of τ_{dis} is important. Of course, a too-high τ_{dis} can lead to reconstructions that could be better. On the other hand, a too-small τ_{dis} can lead to a low discrepancy but a high reconstruction error as well. From this point of view τ_{dis} decides on the importance of the discrepancy.

Remark 3.3.7 (Remark to the Parameter Choice). Regarding the proposed variational reconstruction scheme, defined in Alg. 7.5.1, it is promising to set first $\alpha = 0$ and vary β such that the relative discrepancy is small but avoids a vanishing contrast. Second, it is recommended to vary α until the penalty terms are *balanced* (i.e. they have the same magnitude) in the case of a huge number (e.g. $5 \cdot 10^5$) of inner iterations at least in the last outer iteration step. In the case of $\delta = 0.01$ the choices $\beta = 10^{-5}$ and $\alpha = 3 \cdot 10^{-3}$ lead to balanced penalty terms in a very time-consuming process of several hours. Therefore it is recommended to stop the inner iteration earlier, which requires to increase α, e.g. up to 500, to exploit sparsity. It was derived empirically that the ratio $\alpha/\beta = 500/10^{-5}$ is reasonable, see Ch. 7. This choice was done using the second stopping strategy, see Sec. 7.5.

3.4 Synthetic Data and Real-World Data

In this section we focus on the generation of synthetic data, i.e. simulated data with noise, and the working with real-world data from Institute Fresnel. We will test the performance of the reconstruction algorithms of both kinds of data in Ch. 5.

3.4.1 Generating Synthetic Data

Synthetic data are generated by adding a specific relative noise level $\delta > 0$ to the exact data (noise free data) $F_{\mathrm{meas}} = \mathcal{F}(q) \in \mathbb{C}^{N_{\mathrm{s}} \times N_{\mathrm{i}}}$ for the contrast q. This results in perturbed data

$$F_{\mathrm{meas}}^{\delta} = F_{\mathrm{meas}} + \delta \, \frac{\|F_{\mathrm{meas}}\|_{\mathrm{dis}}}{\|N_{\mathrm{Re}} + \mathrm{i}N_{\mathrm{Im}}\|_{\mathrm{dis}}} (N_{\mathrm{Re}} + \mathrm{i}N_{\mathrm{Im}}), \tag{3.19}$$

where $N_{\mathrm{Re}}, N_{\mathrm{Im}} \in \mathbb{R}^{N_{\mathrm{s}} \times N_{\mathrm{i}}}$ are two real matrices with normally distributed entries. Consequently, the considered relative noise level is $\delta = \|F_{\mathrm{meas}}^{\delta} - F_{\mathrm{meas}}\|_{\mathrm{dis}} / \|F_{\mathrm{meas}}\|_{\mathrm{dis}}$. Note that this Gaussian prior can be interpreted as a ℓ^2 regularization term, see [Fig03, Sec. 3.1]. So, the Tikhonov functional (3.6) fits to the generated Gaussian noise. Further, note that it is a reasonable assumption in real-world environments that data is disturbed by Gaussian noise.

Avoiding the "Inverse Crime" The forward operator is just a physical model of the reality. Therefore the usage of the same forward operator for the generation and inversion of synthetic data leads to reconstructions that are "too good to be true“, see [KS05] or [MS12]. This problem is known in literature as "*inverse crime*", see [CK13]. To avoid the inverse crime for synthetic data, one should generate the synthetic data using a forward solver that has as little as possible to do with the solver employed in the inverse problem, see [CK13, Sec. 5.4]. Therefore we follow [MS12, Sec. 2.3.6] using a simple implementation of this philosophy: we *generate* synthetic data on a fine grid but *reconstruct* on a coarse grid not sharing common factors, e.g. the computational domain is discretized by $N = 1127$ points in each dimension in 2D ($N = 563$ in 3D) for computing synthetic data and by $N = 256$ for reconstructions (in 2D and 3D), as in [BKL17, Sec. 5].

3.4.2 Working with Real-World Data from Institute Fresnel

Clearly, the total field u^{t}, modeled by the solution to the direct scattering problem (2.3)–(2.5), is a quantity which is accessible to physical measurements. Here, we further assume that the incident field itself can also be made accessible, e.g. by reference measurements without the scattering object. Consequently, due to the equation $u^{\mathrm{s}} = u^{\mathrm{t}} - u^{\mathrm{i}}$, see (2.4), we assume in the sequel that (near or far field) measurements of the scattered field are available.

Institute Fresnel provides real-world data of the incident fields at the *receivers' positions*. For further computations we are actually interested in the corresponding incident field in a *region* instead of the incident field at the receivers' positions. This *matching* is essential to work with real-world data.

In this section is given a brief description of the experimental set-up for real-world two-dimensional data published by *Institute Fresnel* in a special issue of the journal *Inverse Problems*, see [BS01]. Further, we will consider how to match the corresponding incident fields following [Geh13, Sec. 6.2.2] and [BKL17, Sec. 6]. Note that the underlying idea of this matching came from Armin Lechleiter. Furthermore, we try out some alternative methods solving the underlying linear inverse problem, that is to get a vector containing coefficients. Further, we consider numerical experiments for a good choice of the approximation order and correct the positions of the obstacles manually. Finally, we estimate the relative errors of the data sets. For reconstruction results we refer to Ch. 5.

Institute Fresnel also published electromagnetic scattering data of three-dimensional targets for transverse magnetic (TM) and transverse electric (TE) polarization. Actually, the mathematical model of time-independent scattering can only deal with TM polarized data in the two-dimensional case, see Ch. 2.[◇]

Experimental Set-Up We sum up the main parameters of Institute Fresnel's data set. The part of the data set employed for numerical experiments consists of scattering of electromagnetic waves from long cylindrical objects in TM polarization. The transmitters and receivers are distributed on a circle around the obstacle. In particular, the data set contains the measurements of real and imaginary part of the total electric field $u^{\mathrm{t}}|_{\Gamma_{\mathrm{s}}}$ and the incident field $u^{\mathrm{i}}|_{\Gamma_{\mathrm{s}}}$ on a circle Γ_{s}. The 72 receivers are arranged on a circle with radius $760 \pm 3\,\mathrm{mm}$. There are 36 transmitters arranged on a circle with radius $720 \pm 3\,\mathrm{mm}$. These transmitters are modeled as point sources. For measurements of the incident field, no obstacle is present and the incident field merely illuminates the experimental set-up. Note that for metrological reasons some transmitter-receiver links are missing, i.e. the discretized forward operator $\mathcal{F}$, see (2.24), maps from $\mathbb{C}^{N_{\mathrm{D}}}$ to an appropriate subspace of $\mathbb{C}^{N_{\mathrm{s}} \times N_{\mathrm{i}}}$. A detailed description of the experimental set-up is given in [BKL17] and [BS01].

Matching the Incident Fields The forward operator discussed in Ch. 2 was designed for point sources or plane waves. However, the Fresnel data set was generated by real-world antennae. This necessitates to deduce the incident field on the region of interest D from the measurements $u^{\mathrm{i}}|_{\Gamma_{\mathrm{s}}}$ of that field at the receivers. In consequence, to apply reconstruction schemes to Institute Fresnel's data set we must change/match the functions employed in the single-layer potential operator (2.10). This *matching* is described in detail in [Geh13, Sec. 6.2.2] and [BKL17, Sec. 6]. It is done by computing the coefficients of the two-dimensional radiating series solutions to the Helmholtz equation, that fits best. The task, in particular, is to seek a coefficient vector $c = (c_{-\xi}, \dots, c_{+\xi})^{\top} \in \mathbb{C}^{(2\xi+1)\times 1}$, where ξ is the *approximation order*, such that

$$\mathcal{V}c \approx u^{\mathrm{i}}|_{\Gamma_{\mathrm{s}}} \tag{3.20}$$

for a matrix $\mathcal{V} \in \mathbb{C}^{N_{\mathrm{s}} \times (2\xi+1)}$ that contains two-dimensional radiating solutions to the Helmholtz equation in dependence of the receivers' positions and the approximation order. With this coefficient vector c the incident field on the region of interest, $u^{\mathrm{i}}|_{D}$, can be computed. Therefore a reasonable recovering of the coefficient vector c is important; also for the reconstruction of the contrast. For that reason, both a good choice of the approximation order ξ and a suitable method solving the approximation problem (3.20) are essential for a reasonable matching. We start with the discussion of the latter one.

A Suitable Method Solving the Approximation Problem The approximation problem (3.20) can be underdetermined as well as overdetermined depending on the size of the approximation order ξ. In general, the problem will be ill-conditioned. Therefore we tested three recovery/regularization methods tackling the approximation problem to get a suitable coefficient vector c.

[◇] In three dimensions we would have to take into account TM and TE polarized data, see [LAv09]. This would in any case require a numerical solver for the three-dimensional Maxwell's equations.

Freq.		Rel. err. of $\mathcal{V}c = u^{\mathrm{i}}\vert_{\Gamma_{\mathrm{s}}}$ in %	Rel. err. of data in %	
in GHz	Method, iterations	Mean	Real part	Imag. part
3	(M1)	1.3	14.9	15.3
3	(M2)	1.8	15.0	15.7
3	(M3), 10^5	1.6	14.9	15.7
3	(M3), 10^8	1.5	14.9	15.5
5	(M1)	3.4	20.1	22.4
5	(M2)	21.2	20.5	26.9
5	(M3), 10^5	15.2	19.1	23.1
5	(M3), 10^8	10.4	18.8	21.2

Table 3.1: Comparison of methods computing coefficient vector c using Institute Fresnel's data of a single dielectricum dielTM_dec8f.exp and approximation order $\xi = 10$. We give the relative error of the measured incident field $u^{\mathrm{i}}|_{\Gamma_{\mathrm{s}}}$ as well as the relative error of data, taking into account the manually corrected position of the target, see Sec. 3.4.2.

The first one, denoted as (M1) and already described in [Geh13, Sec. 6.2.2], computes the best-approximation to (3.20) in the case of an overdetermined system and the least-squares solution if it is underdetermined.[◇] The second method (M2) computes $c = (\gamma I + \mathcal{V}^*\mathcal{V})^{-1}(\mathcal{V}^* u^{\mathrm{i}}|_{\Gamma_{\mathrm{s}}})$ via *linear Tikhonov regularization* with a small $\gamma > 0$. The third method (M3) relies on the *Landweber iteration*[†] stopped after 10^5 respectively 10^8 iterations to solve (3.20). See e.g. [MS12] for more background on these regularization methods.

For Institute Fresnel's data set (3 GHz respectively 5 GHz data) and approximation order $\xi = 10$, the relative errors of the matched fields are listed in Tab. 3.1. Note that also the error of the real-world data is provided. This is possible since the corresponding "ground truth" about the scattering objects is known from [BS01]. Therefore it is possible to simulate the related data matrix. However, it must be noted that we take into account the manually corrected position of the target, see below, in particular Fig. 3.10. We display only the mean relative error for all incident fields because the minimal and maximal errors do not differ significantly. Obviously, the first method (M1) outperforms the other two methods. As the indicated errors for (M1) do not decrease by further increasing the value of ξ, we hence match all incident fields using (M1) with $\xi = 10$.

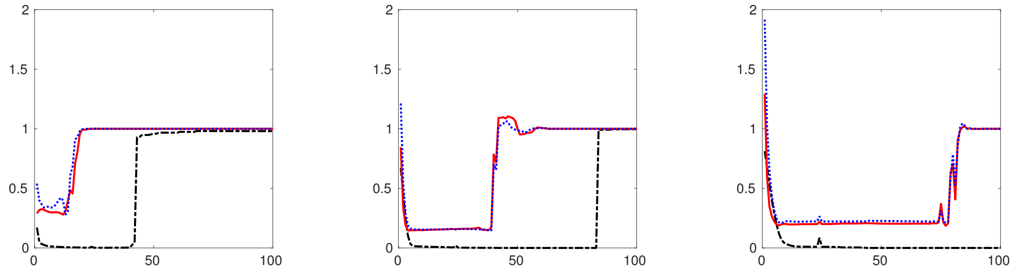
The Choice of the Approximation Order We have to justify the silent choice of the approximation order $\xi = 10$. Therefore we match the incident fields using different approximation orders and frequencies, see Fig. 3.9. We recognize that a too-large ξ is as bad as a too-small ξ. Furthermore, the choice of a larger ξ is possible in the case of a higher frequency. Finally, $\xi = 10$ is an appropriate choice.

Corrected Positions in the Contrast We consider the cases of single dielectric respectively two dielectrics as targets. In the experiment, the real part of the contrast of the target is $q = 2 \pm 0.3$. Since the targets are dielectric, the imaginary part vanishes, see

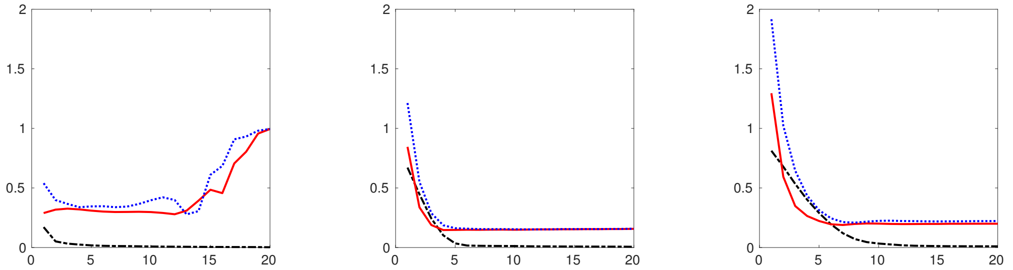
---

[◇] Of course, this describes the MATLAB routine mldivide, c = V\u.

[†] We will discuss the thresholded, nonlinear Landweber scheme in Ch. 6. However, in this case we consider the original, i.e. linear, Landweber iteration, $c^{[n+1]} = c^{[n]} - \omega\mathcal{V}^*(\mathcal{V}c^{[n]} - u^{\mathrm{i}}|_{\Gamma_{\mathrm{s}}})$ with a relaxation factor $0 < \omega < 2/\sigma_1^2$, where σ_1 is the largest singular value of $\mathcal{V}$.



(a) 1 GHz, $\xi = 1, \ldots, 100$. (b) 3 GHz, $\xi = 1, \ldots, 100$. (c) 5 GHz, $\xi = 1, \ldots, 100$.



(d) 1 GHz, $\xi = 1, \ldots, 20$. (e) 3 GHz, $\xi = 1, \ldots, 20$. (f) 5 GHz, $\xi = 1, \ldots, 20$.

Figure 3.9: Relative error of the measured incident field $u^{\mathrm{i}}|_{\Gamma_{\mathrm{s}}}$ (black dash-dotted) and the relative error of data (of the real part in red solid and of the imaginary part in blue dotted) for the manually corrected position of the single dielectric target, see Sec. 3.4.2, over the approximation order ξ for different frequencies. In every case, the related coefficient vectors c for the matched incident field were computed by best-approximation method (M1). (a)–(c) ξ from 1 to 100. (d)–(f) ξ from 1 to 20.

[BS01, Sec. 4.2]. Further, according to [BS01, Sec. 4.2] the radius of the filled dielectric cylinders is 15 mm, the position of the center of the single dielectric is left to the origin with a distance of 30 mm from the origin and the two dielectrics sit left and right to the origin (on the x_1-axis) with a distance of 45 mm from the origin.

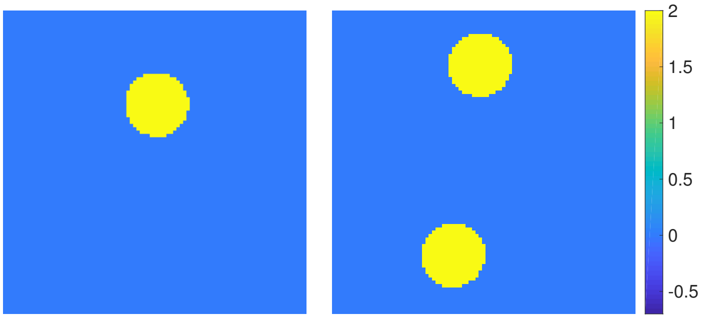
However, even rough reconstructions—we reconstruct in Ch. 5—show that the above-mentioned positions have to be slightly changed to describe the recorded data. The manually corrected positions are in Fig. 3.10. For a detailed description of the changes the reader is referred to [BKL17, Sec. 2]. The manually corrected positions are employed from now on.

Error Estimation To give an example, the manually corrected positions are used to estimate the relative data errors of the data from Institute Fresnel for different frequencies for the contrast with a single dielectric and another contrast with two dielectrics. The results are shown in Tab. 3.2.

The incident field matching is especially successful in the low frequency range. In particular, for frequencies 1–4 GHz the relative error is maximal 1.3%. For 5 and 6 GHz it is below 5%, while at the higher frequencies 7 and 8 GHz it exceeds 15%. In contrast, no clear trend is present in the relative data error.

This relative data error is an important ingredient in form of the relative noise level δ to stop the reconstruction iteration by Morozov's discrepancy principle (3.18), e.g. for a single dielectric at 3 GHz the noise level δ is set to the relative data error of 15%.

Reconstructions with different algorithms from real-world data from Institute Fresnel are presented in Ch. 5.



(a) One dielectric. (b) Two dielectrics.

Figure 3.10: The real part for the two "ground truth" contrast considered here as part of Institute Fresnel's data set. The positions were manually corrected.

	Rel. err. of $\mathcal{V}c = u^{\mathrm{i}}\|_{\Gamma_{\mathrm{s}}}$ in %	Rel. err. of data y in %	
Frequency	Mean	Real part	Imag. part
	Single dielectric (dielTM_dec8f.exp)		
1 GHz	0.9	29.8	39.7
2 GHz	0.6	18.0	14.9
3 GHz	1.3	14.9	15.3
4 GHz	0.8	25.2	19.7
5 GHz	3.4	20.1	22.4
6 GHz	4.2	39.7	32.7
7 GHz	16.4	22.2	28.1
8 GHz	18.1	44.3	39.1
	Two dielectrics (twodielTM_8f.exp)		
1 GHz	1.0	11.5	64.5
2 GHz	0.6	46.7	8.1
3 GHz	1.3	13.4	21.6
4 GHz	0.8	47.9	15.6
5 GHz	3.3	20.3	27.4
6 GHz	4.0	59.5	36.8
7 GHz	16.1	25.2	42.9
8 GHz	17.8	72.2	41.1

Table 3.2: Relative error of the measured incident field $u^{\mathrm{i}}|_{\Gamma_{\mathrm{s}}}$ and the relative error of data for the manually corrected positions of the targets, see Sec. 3.4.2, for different frequencies. In every case, the related coefficient vectors c for the matched incident field were computed by best-approximation method (M1) and approximation order $\xi = 10$.

4 MATLAB Toolbox IPscatt

There are—to the author's best knowledge—only two software packages dealing with the inverse medium problem: GLIDE, see [Str14], and sparseScatt (a predecessor of IPscatt), see [LKK13]. The main reconstruction algorithm of GLIDE is based on a linearization of the underlying problem, which leads to suboptimal reconstructions, e.g. for relatively high wave numbers. The sparseScatt package on the other hand is based on a very time-consuming algorithm. The latter algorithm is the so-called thresholded, nonlinear Landweber scheme and is discussed in Ch. 6. Further, both packages have no implementation available to the general research community.

The main motivation behind IPscatt was to close this gap and to provide a free, open-source implementation of a reconstruction scheme for the nonlinear inverse medium problem in time-independent scattering in two and three dimensions.

The variational regularization scheme, on which the reconstruction part of IPscatt is based, is developed in Ch. 7 and relies on the so-called primal-dual algorithm due to Pock, Bischof, Cremers and Chambolle, see [PCBC09, CP11a]. In Ch. 5 will be shown that the implementation of that scheme provided by IPscatt solves the underlying inverse medium problem efficiently and effectively.

To cope with the ill-posedness of the inverse problem we must add penalty terms, that are suited to a priori information about the contrast, see [EHN96]. IPscatt achieves this by employing terms in the underlying regularization scheme that enforce sparsity in the pixel basis and are based on total variation as well as physical bounds. These terms were selected among other alternatives based on their combined effectiveness as discussed in Sec. 3.2. However, if needed, all the necessary tools are provided within IPscatt to the user, who is interested in extending or altering the proposed reconstruction scheme. To the author's best knowledge, IPscatt is the first toolbox which provides such features and combines sparsity promoting and total variation-based regularization to jointly improve the reconstruction quality for the inverse medium problem.

Furthermore, a variety of options tailored to the needs of practitioners is provided by IPscatt. For example, the toolbox allows the simulation of the scattered near field as well as of the far field. Also, it provides methods for the modeling of the incident field via point sources or plane waves. Further, many common geometries of transmitters and receivers are included out of the box. In addition, an automatic, data-driven choice of the regularization parameters is implemented.

In this chapter an overview of the software framework IPscatt is given and its key features are presented. As already mentioned: For further information on the details of IPscatt the reader may refer to the *supplementary material*[◇] It consists of the *source code*, the *source code documentation* for advanced usage and an *user guide* for IPscatt.

The first two sections of this chapter essentially base on [BKL19, Secs. 1 and 2].

[◇] Supplementary material is provided on an enclosed CD and on http://www.fbuergel.de/ipscatt/ in the latest version.

4.1 Architecture and Key Features of IPscatt

A typical MATLAB toolbox is a collection of functions/routines. If possible, object-oriented paradigms are not used or even avoided. The design of IPscatt follows this common convention. Consequently, IPscatt is a collection of functions dealing with aspects of time-independent scattering.

Groups of IPscatt Each of the functions provided by IPscatt can be put into one of the following five groups: *Initialization*, *Experimental Set-Up*, *Data Readers*, *Scattering Simulation* and *Contrast Reconstruction*. Each of these groups can be considered a module of the IPscatt toolbox. The connections between these groups are described on the basis of the typical *use cases* of IPscatt in Fig. 4.1. In the following paragraphs we discuss the tasks of each group.

The state of the algorithms employed in IPscatt is stored in a “master” data structure. The construction and initialization of that structure is carried out by the functions of the *Initialization* group. Several convenience functions, e.g. for loading of presets are included.

The *Experimental Set-Up* group is responsible for setting up the simulation grid and the positioning of the transmitters and receivers. Of course, many common transmitter-receiver geometries, like rectangle geometry, are provided by IPscatt, see Fig. 4.2(a). However, custom, user-defined geometries are supported too, see Fig. 4.2(b).

Further, the *Experimental Set-Up* group provides routines for the generation of predefined contrasts. In two dimensions these are e.g. a corner, a cross, a ball and a combination of two corners and one ball. IPscatt also provides routines for the generation of rotated variants of these contrasts. In three dimensions a tripod, a cross, a ball and the edges of a cube are available, see Fig. 4.2(c). Further, contrasts fitting to real-world data from Institute Fresnel, see [BS01], are contained: a contrast with one dielectric cylinder and another contrast with two dielectric cylinders.

The routines of the *Data Readers* read the real-world data from Institute Fresnel. Further, these routines are able to match the incident field for that data as described in Sec. 3.4.2.

While the previous three groups are dealing with the preparation of the data structures necessary for the simulation and reconstruction, the fourth and fifth group, i.e. *Scattering Simulation* and *Contrast Reconstruction*, are the main power-horses of the toolbox.

Scattering Simulation provides the routines necessary for the simulation of the scattered fields at the receivers for a given contrast and a given transmitter-receiver geometry. This setting implicitly defines a nonlinear mapping of the contrast to the resulting fields. The details of this *forward mapping* were already discussed in Ch. 2.

Further, the *Scattering Simulation* provides routines for the evaluation of the (Fréchet) derivative of that mapping as well as the derivative of the least-squares error of the forward mapping. The latter is provided since it is an important ingredient of many reconstruction and optimization schemes, e.g. it is used in the thresholded, nonlinear Landweber scheme, that bases on the soft-shrinkage operator, see (3.13), and is discussed in Ch. 6.

One of the many applications of IPscatt is to be a building block for simulation and reconstruction schemes for time-independent scattering. However, some users may prefer to have a pre-packaged reconstruction scheme available. Therefore *Contrast Reconstruction* routines implement essentially three reconstruction schemes: the thresholded, nonlinear Landweber scheme, see Ch. 6, the developed reconstruction scheme using a primal-dual algorithm, see Ch. 7, and a modification of the latter one, that uses the factorization method's result as a prior information to increase the overall effectiveness, see Ch. 8. In it physically sensible contrasts are reconstructed from measured fields at the receivers. The developed reconstruction scheme is based on the assumption that such contrasts admit sparsity, total variation and physical bounds properties. It works in two- as well as three-dimensional setting, see Fig. 4.2(d). It were carried out extensive numerical experiments to tune the parameters of that scheme ensuring its effectiveness and efficiency.

Supplementary Material The interested user will find a detailed description of the provided routines in the *source code documentation* of IPscatt. The main cornerstones of the discretization approach were already presented in Ch. 2. The reconstruction schemes provided by IPscatt are discussed in Chs. 6–8. In particular, the developed variational reconstruction scheme is described in Ch. 7. In addition, a *user guide* of IPscatt is provided, where many typical use cases, like the loading of and the working with real-world data from Institute Fresnel or the matching of the incident fields from receivers' measurements, are described.

Key Features Altogether, the *key features* of IPscatt can be summarized as follows:

- simulation of the time-independent scattering/the forward mapping of scattering,
- (Fréchet) derivative of the forward mapping,
- derivative of the least-squares error of the forward mapping,
- reconstruction schemes of the contrast from measured fields,
- predefined and user-defined transmitter-receiver geometries,
- predefined and user-defined contrasts in 2D and 3D,
- detailed *user guide* and *source code documentation.*

Public and Closed Source Code The enclosed *source code* can be divided into two parts: the *public source code* of IPscatt contains the same parts as in [BKL19]; additional code is denoted by *closed source code* of IPscatt. For the *public source code* a detailed reference of all functions is contained. In addition, the *user guide* only refers to the *public source code*, in particular the practical guides for using the library with code snippets as well as resulting figures. The *closed source code* can be recognized easily because it is in folders with the suffix “Closed”. It additionally contains two versions of the thresholded, nonlinear Landweber scheme: first, based on the soft-shrinkage operator, see (3.13) and [LKK13], and second, based on the developed extended soft-shrinkage operator, mentioned in Sec. 3.3 and defined in (6.7). The source code of the first version was adopted from sparseScatt, that is the MATLAB toolbox behind the numerical examples in [LKK13]. In addition, the *closed source code* contains the reconstruction method using the factorization method's result as described in Ch. 8 and an heuristic algorithm for an automatic, data-driven parameter choice of the regularization parameters as described in Sec. 3.3.3. However, a detailed discussion of the source code archeology was already given in Ch. 1. Another characteristic of the *closed source code* is the less detailed

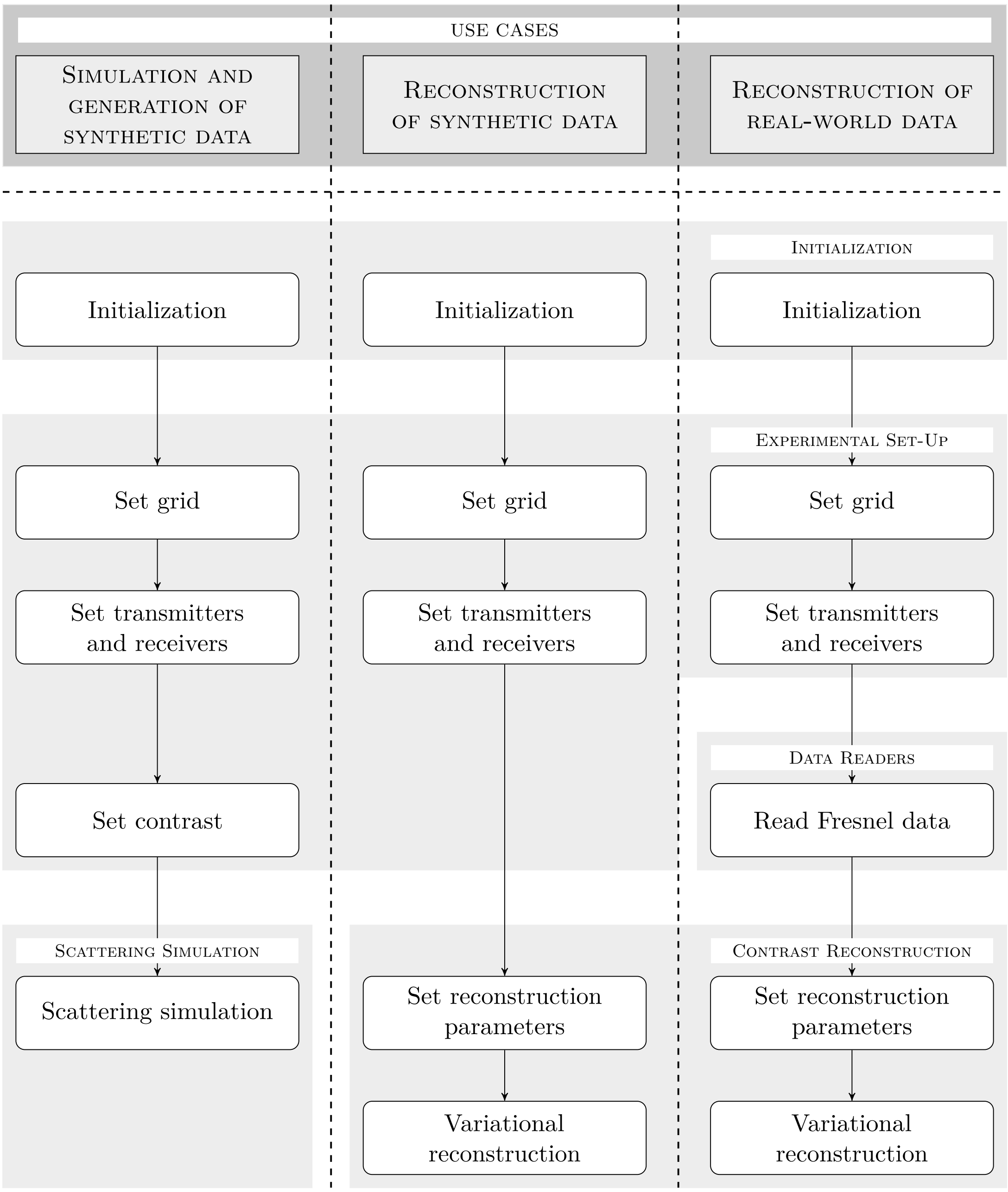
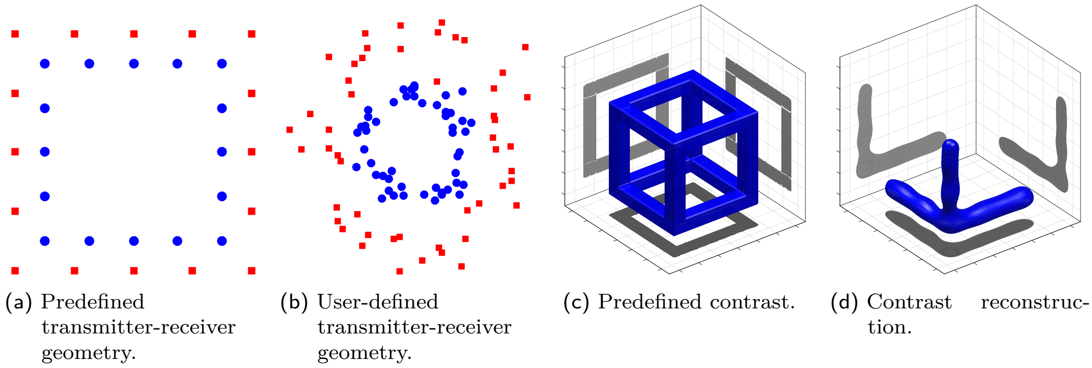


Figure 4.1: Typical use cases of IPscatt and their connection to the modules/function groups *Initialization*, *Experimental Set-Up*, *Data Readers*, *Scattering Simulation* and *Contrast Reconstruction*. Arrows indicate a typical data flow.


(a) Predefined transmitter-receiver geometry.

(b) User-defined transmitter-receiver geometry.

(c) Predefined contrast.

(d) Contrast reconstruction.

Figure 4.2: Some key features of IPscatt. (a) Standard geometries for transmitters/receivers are available out of the box. (b) User-defined positions for transmitters/receivers. (c) Predefined contrasts in 2D and 3D. (Not shown are user-defined contrasts, which can also be used in IPscatt.) (d) Contrast reconstructions in 2D and 3D.

documentation. This concerns the *source code documentation* as well as the *user guide*. In particular, the *user guide* is written for the *public source code*. Therefore it is restricted to the developed reconstruction scheme, that is proposed in Ch. 7.

4.2 Complexity

In this section we discuss the computational cost and main optimizations of the function groups *Scattering Simulation* and *Contrast Reconstruction*.

Scattering Simulation As can be seen in (2.14) and (2.24) the evaluation of the forward operator consists of the single-layer potential, the data-to-solution operator and the evaluation of the Lippmann-Schwinger solution operator. The latter is equivalent to the solution of the Lippmann-Schwinger equation.

The discretization of the first two operators (single-layer potential and data-to-solution operator) consists of a matrix multiplication and therefore their computational cost is negligible. However, the efficient solution of the Lippmann-Schwinger equation is more challenging.

Following [Vai00] this was already discussed in Sec. 2.2: The domain of the operator is periodized and the collocation method is applied. Therefore the convolution structure of the periodized volume potential operator V_{2R}, see (2.19), is exploited and Fourier coefficients are used. Further, using FFT lowers the number of floating-point operations from N^{2d} to $\mathcal{O}(N^d \log N)$, where N is the number of discretization points in each dimension d, see [Vai00]. In contrast to [Vai00] IPscatt stores the necessary Fourier coefficients such that unnecessary shift operations (and therefore memory operations) are avoided. As already mentioned in Sec. 2.3 appropriate restriction and extension operators are used to lower the memory footprint. In particular, all quantities outside of the operator V_{2R} are only stored within the region of interest D instead of the whole computational domain D_{2R}. Finally, IPscatt solves the underlying system of linear equations with the GMRES from Kelley, see [Kel95, Sec. 3.4] and [Kel02] for the source code. In numerical experiments this solver turned out to be superior to standard GMRES.

Contrast Reconstruction In the case of the *Contrast Reconstruction* using the developed reconstruction scheme, see Ch. 7, the forward operator must be evaluated and auxiliary matrices for the Jacobian matrix (Fréchet derivative) of the forward operator are required, see (2.25). These two preparations (forward operator and auxiliary matrices) for the inner iterations of the scheme are time consuming. We remark that each of the steps of the inner iteration, i.e. each of the steps in the primal-dual algorithm, see [CP11a], has a small computational effort. However, for a sufficiently good reconstruction a large number of these steps is necessary. Therefore the overall cost of these iterations is considerable.

As already mentioned, IPscatt provides the derivative of the least-squares error of the forward operator. For example, this is used in the thresholded, nonlinear Landweber scheme, see Ch. 6. Its computational cost is reasonable because it can be computed without evaluating the full derivative first. Further, IPscatt reduces the overall computational effort of reconstruction schemes using the forward operator as well as the derivative of the least-squares error by evaluating both in one routine.

4.3 First Steps with IPscatt

For a quick start with IPscatt two code snippets are given. However, see the *user guide* for more information.

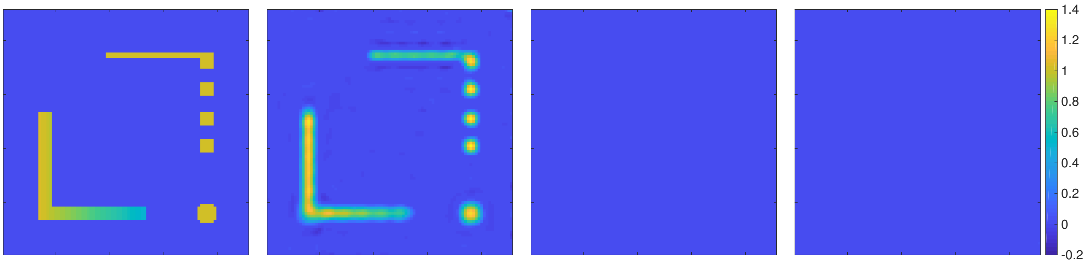
Quick Start and Typical Work Flow Input parameters can be set in the fields of the structure array$^\diamond$ seti, that is a mutable data structure. IPscatt takes care that all necessary fields are set in a consistent manner, e.g. if these fields are empty as in Lst. 4.1, default parameters will be set. In the routine to set the data, setData, this concerns e.g. the underlying grid, the positions of the transmitters and receivers, the type of the incident field as well as the predefined contrast of the obstacle. Furthermore, setData solves the direct scattering problem and adds Gaussian noise to the simulated data. In the end, the variational reconstruction starts by recon after reconstruction-dependent parameters were set in setRecon. By default the developed reconstruction scheme, see Ch. 7, is used.

```
init;
seti = struct; seti = setData(seti);
seti = setRecon(seti); seti = recon(seti);
```

Listing 4.1: Quick Start Example.

For a quick introduction to the toolbox IPscatt see the guides in the *user guide*. These guides will show how to: generate a grid; use some standard transmitter-receiver geometries for the experimental set-up; load predefined contrasts; set geometry and simulation; simulate data with noise; reconstruct in the case of two and three dimensional scattering.

$^\diamond$We follow MATLAB's nomenclature denoting an *associative array* as *structure array* and *return values* as *output arguments*.



(a) Exact, real part. (b) Recon., real part. (c) Exact, imag. part. (d) Recon., imag. part.

Figure 4.3: Real and imaginary part of the exact contrast in (a) and (c) as well as the reconstructed contrast from simulated data with 1% noise level after 11 outer iterations (stopped by the discrepancy principle) in (b) and (d) as generated by Lst. 4.2. The reconstruction error is 36%, the run time was 4 min.

Easy Usage of the Toolbox IPscatt The following Lst. 4.2 demonstrates the easy usage of IPscatt adapting some of the parameters, such that simulated data with noise in two dimensions is generated and the contrast is reconstructed. The result is given in Fig. 4.3.

```
init; % initialization

seti.contrast   = 'cornerBallSparseMod2D'; % contrast
seti.nCD        = 256; % number of discretization points for each dimension
seti.rCD        =  0.2; % computational domain of size [−0.2,0.2)^2 (rCD is in unit meter)
seti.k          =  250; % wave number (in unit 1/m)
seti.incNb      =    35; % transmitters (on a circle with radius 5 m)
seti.measNb =    35; % receivers (on a circle with radius 5 m)
seti.delta      =   0.01; % relative noise level
seti.pdaN       =      50; % number of inner iterations
seti.alpha      =     500; % sparsity regularization parameter
seti.beta       =   1E−5; % TV regularization parameter

seti = setData(seti);  % set experimental set−up, contrast,
                        % compute data, add noise
seti = setRecon(seti); % preparation of reconstruction
seti = recon(seti);     % reconstruction process
```

Listing 4.2: Easy usage of the toolbox adapting some parameters.

Necessary Discretization Points per Wavelength A rule of thumb is to *discretize a wavelength* by at least 10 points. Therefore an accurate choice of the computational domain's discretization is required. The computational domain D_{2R} with side length $4R$ was defined in (2.16). It is worth mentioning that it has side length $2r$ with $r = \text{seti.rCD}$ in the *source code*. It means that $r = 2R$. As in (2.16) the length of the infinitesimal element of CD is $h_N = 4R/N$. To compute the discretization points per wavelength, λ/h_N, we remember the wave number $k = 2\pi/\lambda$, see Ch. 2, and receive that

$$\frac{\lambda}{h_N} = \frac{2\pi/k}{4R/N} = \frac{2\pi/\text{seti.k}}{2\,\text{seti.rCD}/\text{seti.nCD}}$$

should be 10 or greater. In the 2D example in Fig. 4.3 we end up with $\lambda/h_N \approx 16$.

5 Comparative Numerical Examples

In this chapter we follow [BKL19, Sec. 3] and demonstrate the main applications of IPscatt using representative examples. Further, we compare IPscatt with two other software packages (sparseScatt and GLIDE) performing the same task, i.e. solving the inverse medium problem in scattering, and we discuss the advantages of IPscatt over these packages.

In this regard, we put the *Contrast Reconstruction* in the foreground because the essential differences among the three packages are in this group. Remember that several reconstruction schemes are available in IPscatt. However, for the comparative numerical examples in this chapter we only use the developed reconstruction scheme, that is proposed in Ch. 7. As mentioned in Ch. 4, *Contrast Reconstruction* is only one of the five function groups of IPscatt apart from *Initialization*, *Experimental Set-Up*, *Data Readers* and *Scattering Simulation*. However, we will choose the test cases such that all other function groups are implicitly included in at least one of the discussed examples.

5.1 Comparing Software Packages

We begin our discussion with a survey of the reconstruction methods implemented by the software packages sparseScatt and GLIDE and refer to [LKK13] respectively [Str14, Ch. 5] for details.

Package sparseScatt The package sparseScatt solves the nonlinear inverse medium problem in scattering using a thresholded, nonlinear Landweber scheme. This scheme is discussed in Ch. 6 and employs the so-called soft-shrinkage operator, see (3.13). In sparseScatt the contrast is assumed to be sparse with respect to the pixel basis. Therefore the underlying Tikhonov functional is given by

$$\min_{q\in X} \underbrace{\tfrac{1}{2}\|\mathcal{F}(q) - F^{\delta}_{\text{meas}}\|^2_{\text{dis}}}_{=:f_{\text{dis}}(q)} + \underbrace{\alpha\|q\|_{\text{spa}}}_{=:f_{\text{spa}}(q)} . \tag{5.1}$$

Package GLIDE The package GLIDE relies on the Born approximation on the physical model of scattering, defined in Sec. 2.1. The Born approximation can be considered as the linearization of the forward operator (2.14) in the origin, cf. [Bor26]. Further, for the reconstruction the contrast is assumed to be sparse. However, in contrast to IPscatt and sparseScatt the sparsity is assumed to be with respect to a wavelet basis, e.g. the Cohen-Daubechies-Feauveau-9/7 (CDF) wavelets. Finally, the related variational problem is solved via the *fast iterative shrinkage-thresholding algorithm for linear inverse problems* (FISTA), see [BT09].

Package IPscatt To the author's best knowledge, IPscatt is the first freely available software package providing an advanced contrast reconstruction scheme in two and three dimensions, providing data handling routines for real-world data from Institute Fresnel and finally, providing several transmitter-receiver geometries. As mentioned above, only the reconstruction scheme, that is proposed in Ch. 7, is used in IPscatt in this chapter.

5.2 Experimental Set-Ups

For our experiments we will consider the three transmitter-receiver geometries shown in Fig. 5.1, that we call *Fresnel* case, *borehole* case and *sphere* case.

Fresnel Case In the *Fresnel* case, we use the same experimental set-up as in [Str14, Sec. 5.5, Fig. 5.6], i.e. we consider scattering in two dimensions with a contrast consisting of a plane of two homogeneous, long cylindrical dielectrics, see Fig. 5.2(a)/(e). This is the same set-up as used in the *real-world* data set twodielTM_8f.exp from Institute Fresnel, see [BS01], and was already discussed in Sec. 3.4.2. There are 36 transmitters arranged on a circle with radius $720 \pm 3\,\mathrm{mm}$. Each of these point sources emits an electromagnetic wave in TM polarization with a frequency of $6\,\mathrm{GHz}$, i.e. wave number $k \approx 126\,\mathrm{m}^{-1}$. The 72 receivers are arranged on a circle with radius $760 \pm 3\,\mathrm{mm}$. The experimental set-up is depicted in Fig. 5.1(a). Note that for metrological reasons some transmitter-receiver links are missing, i.e. the discretized forward operator $\mathcal{F}$, see (2.24), maps from $\mathbb{C}^{N_\mathrm{D}}$ to an appropriate subspace of $\mathbb{C}^{N_\mathrm{s} \times N_\mathrm{i}}$. Therefore the minimization functional (3.11) was adapted accordingly.

Borehole Case In the *borehole* case, we consider a *synthetic* contrast in two dimensions consisting of a real and an imaginary part as depicted in Fig. 5.2(i)/(m). The contrast was constructed to be challenging as it consists of a gradual contrast ramp (the thick horizontal rectangle at the bottom of the real and imaginary part of the contrast) as well as fine features (a filled circle, small squares with different contrasts in real as well as imaginary part and a rotated thin line in the real part). The transmitter-receiver geometry consists of two boreholes $10\,\mathrm{m}$ apart and $20\,\mathrm{m}$ deep. Each borehole is assumed to be equipped with 18 transmitters and receivers, that are arranged in the configuration presented in Fig. 5.1(b). The transmitters are assumed to be point sources[◇] emitting with a wave number $k = 250\,\mathrm{m}^{-1}$, i.e. a frequency of $11.9\,\mathrm{GHz}$. Gaussian noise with noise level of 1% in real and imaginary part was employed to simulate data in this case.

Sphere Case In the *sphere* case, three-dimensional scattering is considered. The *synthetic* contrast is depicted in Fig. 5.3(a)/(e). The transmitter-receiver geometry consists of 31 transmitters as well as receivers that are distributed almost equidistantly on the surface of a sphere with a radius of $5\,\mathrm{m}$, see Fig. 5.1(c). To approximate an equidistant distribution the so-called *Fibonacci lattice*, see [Gon10, Sec. 3.1], was employed. The transmitters are point sources emitting with a wave number $k = 10\,\mathrm{m}^{-1}$, i.e. a frequency of $0.5\,\mathrm{GHz}$. Again 1% Gaussian noise was used in the data simulation.

[◇] Usually a pulse is used in a borehole setting. However, we indeed consider a continuous source as in the whole thesis.

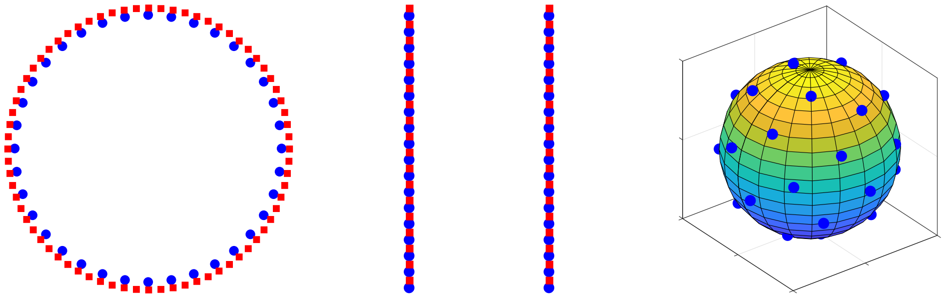

(a) *Fresnel* case. (b) *Borehole* case. (c) *Sphere* case.

Figure 5.1: Experimental set-ups for comparisons. Transmitters are blue filled circles and receivers red filled squares. (a) *Fresnel* case: transmitters and receivers are arranged on circles as in the experimental set-up of Institute Fresnel, see [BS01]. (b) *Borehole* case: transmitters and receivers are arranged on two boreholes. (c) *Sphere* case: transmitters (and receivers) are roughly equidistantly arranged on a sphere.

5.3 Results and Discussion

Results The results of the reconstructions in 2D are depicted in Fig. 5.2 and in 3D in Fig. 5.3. The contrast is always shown in the region of interest (ROI), which is a square approximately from $-0.07\,\mathrm{m}$ to $0.07\,\mathrm{m}$ in case of two dimensions and a cube from $-0.70\,\mathrm{m}$ to $0.70\,\mathrm{m}$ in three dimensions. Note that test cases are omitted if the package is not able to tackle them.

The results were obtained with the following set of regularization parameters: In case of IPscatt and sparseScatt we used the highest regularization parameter α, that fulfills the discrepancy principle (3.18), see Alg. 3.3.6 for an automatic choice. Further, in the case of IPscatt we generally use $\beta = \alpha \cdot 2 \cdot 10^{-8}$ as the second regularization parameter because the ratio $\alpha/\beta = 500/10^{-5}$ was derived empirically, cf. Ch. 7. In addition, we enforce the real part of the contrast to be in $[-1, 3]$ and the imaginary part to be in $[0, 3]$ by using physical bounds in IPscatt. However, an exception of the α-β-ratio is made in IPscatt for the case of *balanced penalty terms*—i.e. sparsity penalty and total variation penalty have the same scale. Balanced penalty terms are recommended when reconstruction quality is paramount and increased computational cost is disregarded.

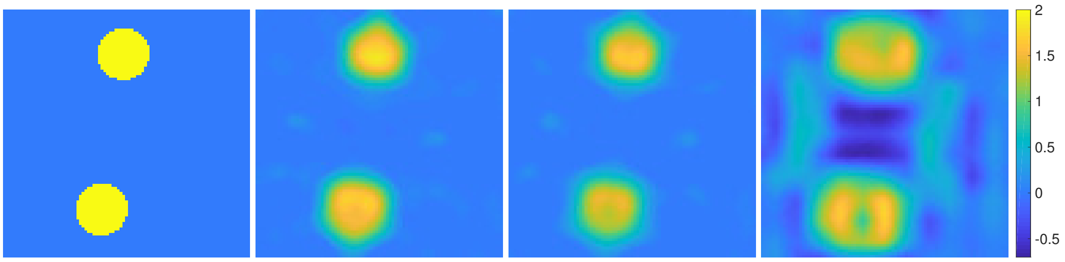
The reconstruction with sparseScatt turned out to be very sensitive to the choice of the step size of the thresholded, nonlinear Landweber scheme. If the value was chosen too small sparseScatt was very slow. If the value was chosen too big the reconstruction failed completely.[◇] After several experiments the value of 10^{-4} was deemed best for the *Fresnel* case, cf. Fig. 5.2(c)/(g), and the value of 1 was deemed best for the *borehole* case, cf. Fig. 5.2(l)/(p).

In the case of GLIDE the same parameter values were used as in [Str14, Sec. 5.5, Fig. 5.6].

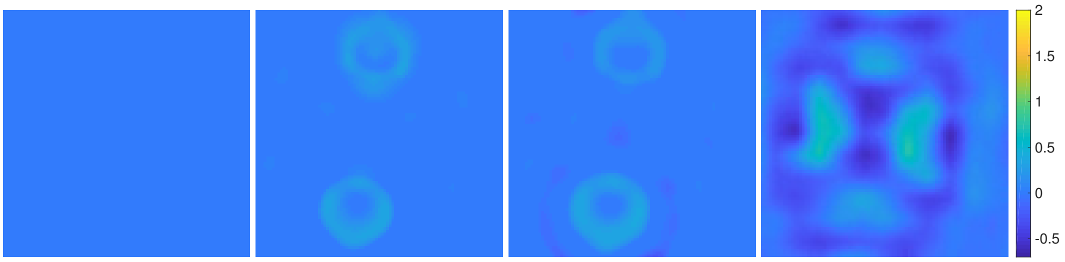
All computations were carried out on a workstation with an Intel(R) Core(TM) i7-3770 CPU with 3.40 GHz and 32 GByte RAM using MATLAB in version R2016b. Although

---

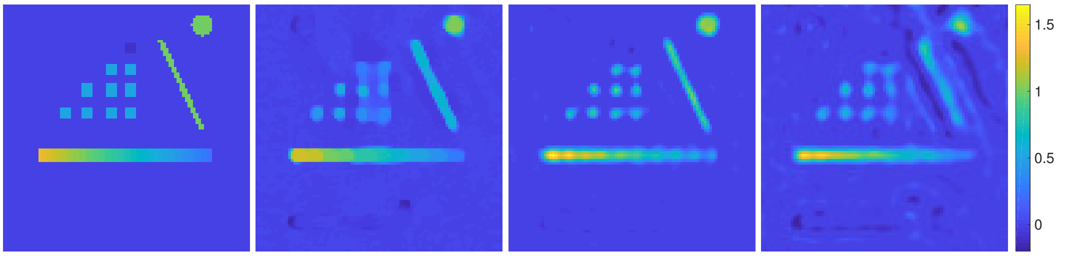
[◇] The step sizes in sparseScatt can be chosen automatically. This choice is done by Barzilai-Borwein rule, see [BB88], and redefined by Armijo rule, see e.g. [NW06]. (It is also mentioned in Sec. 6.1.) However, the choice frequently fails, such that a careful chosen constant step size is used for the comparative numerical examples.



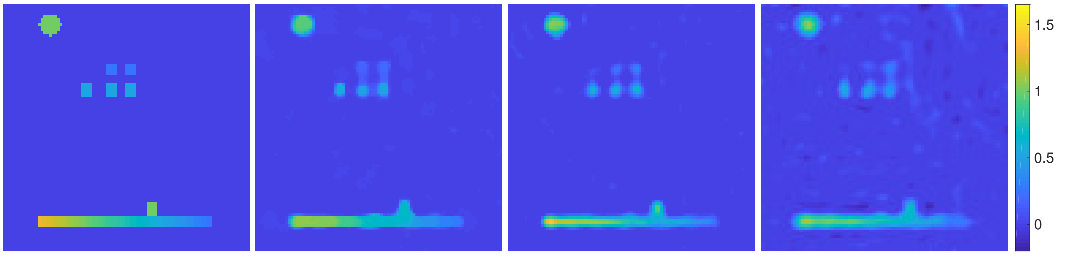
(a) Contrast of two dielectrics, real part. (b) Recon. with IPscatt, real part. (c) Recon. with sparseScatt, real part. (d) Recon. with GLIDE, real part.



(e) Contrast of two dielectrics, imag. part. (f) Recon. with IPscatt, imag. part. (g) Recon. with sparseScatt, imag. part. (h) Recon. with GLIDE, imag. part.



(i) Predefined contrast, real part. (j) Recon. with IPscatt, balanced penalty terms, real part. (k) Recon. with IPscatt, real part. (l) Recon. with sparseScatt, real part.

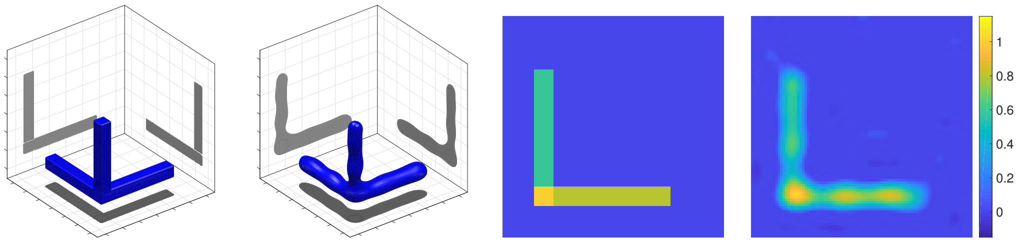


(m) Predefined contrast, imag. part. (n) Recon. with IPscatt, balanced penalty terms, imag. part. (o) Recon. with IPscatt, imag. part. (p) Recon. with sparseScatt, imag. part.

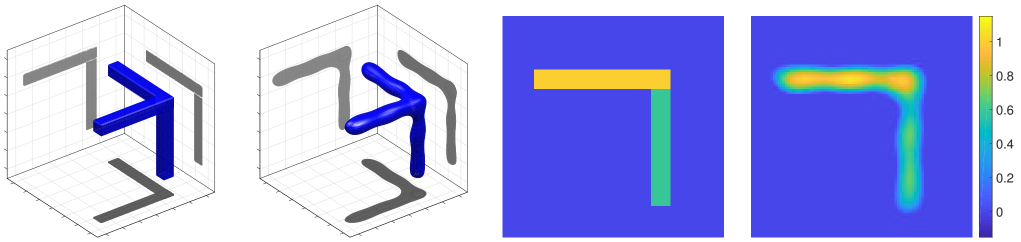
Figure 5.2: Comparison of IPscatt (with and without balanced penalty terms) with sparseScatt and GLIDE in two dimensions. *Fresnel* case (first and second row): exact contrast and reconstructions of two dielectrics from real-world data from Institute Fresnel with 6 GHz, real and imaginary part. The two dielectrics with a homogeneous contrast of $q = 2 \pm 0.3$ have a radius of 15 mm and are positioned 45 mm from the origin. *Borehole* case (third and fourth row): predefined contrast and reconstructions in real and imaginary part from simulated data with 1% noise level. The reconstructions with IPscatt and sparseScatt were stopped by discrepancy principle. The reconstruction with GLIDE stopped because the maximum number of iterations was reached.

In the *Fresnel* case (first and second row) the run times, relative discrepancies and relative errors were: with IPscatt (b)/(f) 1.3 min, dis. 0.216, err. 0.542 after 6 iterations with $\alpha = 10^6$, $\beta = 0.02$; with sparseScatt (c)/(g) 4.4 min, dis. 0.220, err. 0.552 after 50 iterations with $\alpha = 114$; and with GLIDE (d)/(h) 0.8 min.

In the *borehole* case (third and fourth row) the run times, relative discrepancies and relative errors were: with IPscatt using balanced penalty terms (j)/(n) 12.4 h, dis. 0.026, err. 0.333 after 64 iterations (with $5 \cdot 10^5$ inner iterations in the last one) with $\alpha = 2.04 \cdot 10^{-4}$, $\beta = 6 \cdot 10^{-7}$; with IPscatt (k)/(o) 16.5 min, dis. 0.025, err. 0.337 after 62 iterations with $\alpha = 310$, $\beta = 6.2 \cdot 10^{-6}$; and with sparseScatt (l)/(p) 5.1 h, dis. 0.015, err. 0.455 after 2319 iterations with $\alpha = 10^{-4}$.



(a) Predefined contrast, real part. (b) Recon. with IPscatt, real part. (c) Predefined contrast (plane), real part. (d) Recon. (plane) with IPscatt, real part.



(e) Predefined contrast, imag. part. (f) Recon. with IPscatt, imag. part. (g) Predefined contrast (plane), imag. part. (h) Recon. (plane) with IPscatt, imag. part.

Figure 5.3: *Sphere* case: reconstruction of two tripods in real and imaginary part with IPscatt from simulated data with 1% noise level. The arcs of the tripods have the contrast values 1.0, 0.8, 0.6 and 1.0i, 0.8i, 0.6i. The contrast on a sectional plane through the obstacle is shown in the second direction in the case of the real part and the first direction in the case of the imaginary part. Parameters for reconstruction were $\alpha = 1000$, $\beta = 2 \cdot 10^{-5}$. The reconstruction was stopped by the discrepancy principle after 15 iterations. The run time was 6.01 h, relative discrepancy 0.012 and relative error 0.464.

the run times should be taken with a pinch of salt because the computations rely on MATLAB, they give an impression of the performance of the algorithms.

The contrast and the reconstructions relating the *Fresnel* case of the experimental set-up are shown in Fig. 5.2(a)–(h). As mentioned above GLIDE relies on the Born approximation and therefore is essentially comparable to IPscatt respectively sparseScatt with only one outer loop. However, as can be seen in Fig. 5.2(d)/(h) the resulting reconstruction is reasonable. A better reconstruction is obtained with sparseScatt with a significantly higher computational cost, see Fig. 5.2(c)/(g). However, one notices that both reconstructions contain regions where the imaginary part of the contrast is negative, i.e. the obtained reconstructions are physically not sensible. Finally, since the reconstruction approach of IPscatt is based on physical constraints, it is able to avoid these artifacts, see Fig. 5.2(b)/(f). At the same time it has only slightly higher computational cost than GLIDE.

Next, the contrast and the reconstructions relating the *borehole* case are presented in Fig. 5.2(i)–(p). The result of IPscatt was accurate and fast again, see Fig. 5.2(k)/(o). As in the previous example, the reconstruction obtained by sparseScatt is reasonable as determined by visual inspection, see Fig. 5.2(l)/(p). It is not surprising that the same physical artifacts as were discussed in the *Fresnel* case are also encountered here. Further, we observe some artifacts around the rotated thin line. Moreover, sparseScatt was much slower than IPscatt. In addition, the contrast was reconstructed with IPscatt with the above mentioned balanced penalty terms, see Fig. 5.2(j)/(n). The thick horizontal rectangle is excellently recovered as well as is the gradual contrast ramp. However, the thickness of the thin diagonal line is overestimated by the reconstruction algorithm.

Finally, the contrast and the reconstruction relating the *sphere* case of the experimental set-ups are shown in Fig. 5.3. As neither sparseScatt nor GLIDE support reconstructions for scattering in three dimensions only the reconstruction obtained by IPscatt is shown. As can be seen it provides a precise reconstruction result in a reasonable time.

Discussion All three software packages have in common routines of the *Data Readers* group to deal with real-world data from Institute Fresnel and routines of the *Scattering Simulation* group. In the latter case, the only difference concerns the dimension: all can deal with 2D but only IPscatt with 3D. Therefore we focus on the groups *Experimental Set-Up* and *Contrast Reconstruction* in the discussion.

Out of the box the software package GLIDE is restricted to two-dimensional scattering problems and transmitter-receiver geometries on circles. GLIDE has these properties in common with sparseScatt. As the only package, amongst the three considered, GLIDE is restricted to small wave numbers because it solves the *linear* inverse medium problem instead of the *nonlinear* one. Consequently, in the case of high wave numbers GLIDE cannot compete with sparseScatt and IPscatt with regard to reconstruction quality. On the other hand, GLIDE outcompetes the compared packages with the low computational effort. Altogether, GLIDE is recommended only if the run time is the most important factor.

Out of the box the package sparseScatt is also restricted to two-dimensional scattering problems and transmitter-receiver geometries on circles. Further, among the three compared packages sparseScatt was consistently the computationally most expensive one. However, sparseScatt can deal with the *nonlinear* inverse medium problem and therefore delivers a better reconstruction than GLIDE but stays behind IPscatt.

The only disadvantage of IPscatt in comparison to GLIDE and sparseScatt is the slightly higher computational effort in comparison to GLIDE. Altogether one can summarize that IPscatt provides a better usability, reconstruction quality and run time properties than GLIDE or sparseScatt.

6 Reconstruction with the Thresholded, Nonlinear Landweber Scheme

As mentioned in Sec. 3.3 the soft-shrinkage operator is used to solve $\min_q f(q)$ with $f(q) := f_{\text{dis}}(q) + f_{\text{spa}}(q)$. In the first part of this chapter we will consider the corresponding thresholded, nonlinear Landweber scheme. In addition to the sparsity $f_{\text{spa}}(q)$ we will take into account physical bounds solving $\min_q f(q)$ with $f(q) := f_{\text{dis}}(q) + f_{\text{spa}}(q) + f_{\text{phy}}$. Therefore an extended soft-shrinkage operator is defined in the second part of this chapter. Finally, in the third part the use of wavelets in the thresholded, nonlinear Landweber scheme is discussed.

6.1 Soft-Shrinkage

We look for a numerical method to minimize the Tikhonov functional

$$f(q) := \tfrac{1}{2} \|\mathcal{F}(q) - F^\delta_{\text{meas}}\|^2_{\text{HS}(L^2(\Gamma_\text{i}),L^2(\Gamma_\text{s}))} + \alpha \|q\|_{L^1(D)} \quad \text{on } L^1(D).$$

In the discretized version we take into account the area/volume of a pixel/voxel, cf. (3.11),

$$\min_{q \in X} f(q), \quad f(q) := \underbrace{\tfrac{1}{2} \|\mathcal{F}(q) - F^\delta_{\text{meas}}\|^2_{\text{dis}}}_{=:f_{\text{dis}}} + \underbrace{\alpha \|q\|_{\text{spa}}}_{=:f_{\text{spa}}}. \tag{6.1}$$

For a numerical method to minimize the Tikhonov functional we follow [LKK13, Sec. 5] and, for the sake of simplicity, we only consider real-valued contrasts q. In this case $\|q\|_{\text{spa}} = h_N^d \|q\|_1$, see (3.4).

Note that the whole chapter considers only real-valued contrasts. However, the results can be extended to complex-valued contrasts. Therefore the soft-shrinkage operator is used separately to the real and imaginary part. As already mentioned in Ch. 1, the provided code basis to [LKK13] included this. The transformation of the complex problem is carried out as described in Sec. 7.3.

We will see that the *thresholded, nonlinear Landweber scheme* arises in a natural way from the first-order optimality conditions. We consider the derivation as a preparation of the analogous scheme with the extended soft-shrinkage operator. In addition, we will recycle results in the reconstruction with a primal-dual algorithm in Ch. 7.

We made sure that f_{spa} is convex, see Sec. 3.2.1. Therefore f is convex. Hence it is sufficient to look for a contrast q that fulfills $\partial f(q) = \partial f_{\text{dis}}(q) + \partial f_{\text{spa}}(q) \overset{!}{=} 0$. Therefore we will consider the subdifferentials of f_{dis} and f_{spa}.

Subdifferential of f_{dis} For the sake of simplicity, in this paragraph we consider the unweighted Frobenius norm $\|\cdot\|_{\text{F}}$ instead of $\|\cdot\|_{\text{dis}}$, cf. Sec. 2.3, i.e. we consider $f_{\text{dis,F}}(q) := \frac{1}{2}\|\mathcal{F}(q)-F^{\delta}_{\text{meas}}\|^2_{\text{F}}$ instead of $f_{\text{dis}}(q)$. Rewritten as an inner product this results in $\frac{1}{2}\langle\mathcal{F}(q)-F^{\delta}_{\text{meas}},\mathcal{F}(q)-F^{\delta}_{\text{meas}}\rangle_{\text{F}}$. As $\partial(\langle f,g\rangle) = \langle\partial f,g\rangle+\langle f,\partial g\rangle$ and $\partial\mathcal{F}(q)$ is the Fréchet derivative $\mathcal{F}'(q)$, it follows that $\partial f_{\text{dis,F}}(q) = \frac{1}{2}\langle\mathcal{F}'(q),\mathcal{F}(q)-F^{\delta}_{\text{meas}}\rangle_{\text{F}} + \frac{1}{2}\langle\mathcal{F}(q)-F^{\delta}_{\text{meas}},\mathcal{F}'(q)\rangle_{\text{F}}$. We employ the adjoint of the Fréchet derivative to change the position of the Fréchet derivative in both inner products. In addition, we use the symmetry by permuting the first and second argument of the second inner product. Then, we end up with $\partial f_{\text{dis,F}}(q) = \frac{1}{2}\langle I,[\mathcal{F}'(q)]^*[\mathcal{F}(q)-F^{\delta}_{\text{meas}}]\rangle_{\text{F}} + \frac{1}{2}\langle I,[\mathcal{F}'(q)]^*[\mathcal{F}(q)-F^{\delta}_{\text{meas}}]\rangle_{\text{F}} = [\mathcal{F}'(q)]^*[\mathcal{F}(q)-F^{\delta}_{\text{meas}}]$. Finally, we have shown that the derivative of the least-squares error of the forward operator is the adjoint of the Fréchet derivative applied to the defect $\mathcal{F}(q)-F^{\delta}_{\text{meas}}$. The discretized version of the Fréchet derivative was already given in Sec. 2.3 for the unweighted as well as the weighted norm, see (2.26) and (2.29).

Subdifferential of f_{spa} As mentioned above, we consider real-valued contrasts in this chapter. Therefore $f_{\text{spa}} = \alpha\|q\|_{\text{spa}} = \alpha h_N^d\|q\|_1$. The subdifferential of the absolute value function, see Fig. 3.7(a) or 6.1(a), is the set-valued signum function sign as already explained, see (3.15) and Fig. 3.7(b) or 6.1(b),

$$\operatorname{sign}(q(x)) = \begin{cases} \{-1\} & \text{if } q(x) < 0, \\ [-1,1] & \text{if } q(x) = 0, \\ \{1\} & \text{if } q(x) > 0, \end{cases} \qquad x \in D.$$

It is straightforward to handle the additional factors, regarding that $\partial(\kappa h) = \kappa\partial h$ for a factor $\kappa > 0$ and a convex function h, see Fig. 6.1(c).

Thresholded, nonlinear Landweber Finally, for $\partial f(q) = \partial f_{\text{dis}}(q)+\partial f_{\text{spa}}(q) \overset{!}{=} 0$ we end up with

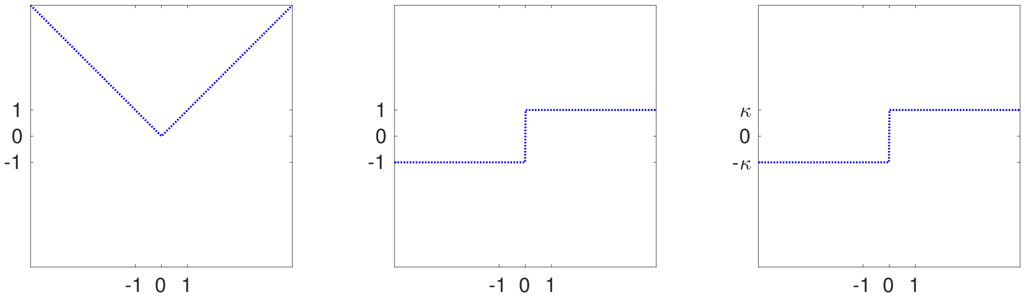
$$[\mathcal{F}'(q)]^*[\mathcal{F}(q)-F^{\delta}_{\text{meas}}] + \alpha\, h_N^d \operatorname{sign}(q) \overset{!}{=} 0.$$

We rearrange it by subtracting $[\mathcal{F}'(q)]^*[\mathcal{F}(q)-F^{\delta}_{\text{meas}}]$, multiplying $\tau > 0$ and adding q, hence, receive

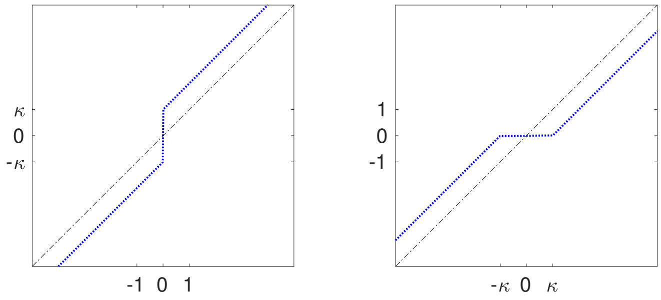
$$(I + \tau\,\alpha\, h_N^d \operatorname{sign})(q) = q - \tau[\mathcal{F}'(q)]^*[\mathcal{F}(q)-F^{\delta}_{\text{meas}}]. \tag{6.2}$$

The left hand side is $(I+\tau\partial f_{\text{spa}})(q)$, cf. Fig. 6.1(d). As we are interested in the contrast q we built the inverse, i.e. the proximal mapping $(I+\tau\partial f_{\text{spa}})^{-1}(q) = (I+\kappa\operatorname{sign})^{-1}(q)$ with $\kappa = \tau\,\alpha\,h_N^d$, cf. Fig. 6.1(e). This mapping is the well-defined *soft-shrinkage operator*, remember (3.13),

$$\mathbb{S}(q(x),\kappa) := \begin{cases} q(x)+\kappa & \text{if } q(x) \le -\kappa, \\ 0 & \text{if } q(x) \in (-\kappa,+\kappa), \\ q(x)-\kappa & \text{if } q(x) \ge \kappa, \end{cases}$$



(a) $|\cdot|$. (b) $\partial|\cdot| = \mathrm{sign}(\cdot)$. (c) $\kappa\partial|\cdot|$ with $\kappa > 0$.



(d) $I + \kappa\partial|\cdot|$. (e) $(I + \kappa\partial|\cdot|)^{-1}$.

Figure 6.1: (a) Absolute value function. (b) The subdifferential of the absolute value function is the set-valued signum function sign. (c) Subdifferential taking into account the factor κ. (d) Adding the identity. (e) The inverse is the soft-shrinkage operator.

for $\kappa > 0$ and $x \in D$, see [DDD04] or [CS05]. (Note that we will be interested in terms like $(I + \tau\partial f_{\mathrm{spa}})^{-1}(q)$ in the next chapter too.) Finally, we end up with the so-called *thresholded, nonlinear Landweber scheme*

$$q_{n+1} = \mathbb{S}\left(q_n - \tau_n[\mathcal{F}'(q)]^*[\mathcal{F}(q) - F^\delta_{\mathrm{meas}}], \tau_n\, \alpha\, h_N^d\right). \tag{6.3}$$

Thresholded, Nonlinear Landweber Scheme also for Non-Sparse Penalty Terms In fact, the thresholded, nonlinear Landweber scheme is described in [LKK13, Sec. 5] to minimize the Tikhonov functional $f(q) := \frac{1}{2}\|\mathcal{F}(q) - F^\delta_{\mathrm{meas}}\|^2_{\mathrm{HS}(L^2(\Gamma_\mathrm{i}),L^2(\Gamma_\mathrm{s}))} + \frac{\alpha}{p}\|q\|^p_{L^p(D)}$ on $L^p(D)$. This coincides[◇] with the just-mentioned case if $p = 1$. We follow [LKK13] for the rest of this paragraph. We denote the subdifferential of $\|q\|^p_{L^p(D)}$ by $J_p(q)$. For the duality mapping $J_p(\cdot)$ we refer to [SKHK12] and [Cio90]. If the Lebesgue index $p > 1$, it holds that $[J_p(q)](x) = |q(x)|^{p-1}\,\mathrm{sign}(q(x))$ for $x \in D$. Of course, $[J_1(q)](x) = \mathrm{sign}(q(x))$. Finally, we want to solve $[\mathcal{F}'(q)]^*[\mathcal{F}(q) - F^\delta_{\mathrm{meas}}] + \alpha\, J_p(q) = 0$. Rearranging results in $(I + \tau\,\alpha\, J_p)(q) = q - \tau[\mathcal{F}'(q)]^*[\mathcal{F}(q) - F^\delta_{\mathrm{meas}}]$. The inverse of the left hand side is $(I + \kappa J_p)^{-1}$ with $\kappa = \tau\,\alpha$. This mapping is well-defined and denoted by $\mathbb{S}_p(q(x), \kappa)$. Of course, for $p = 1$ it is the soft-shrinkage operator, i.e. $\mathbb{S}_1(q(x), \kappa) = \mathbb{S}(q(x), \kappa)$. Finally, we end up with the thresholded, nonlinear Landweber scheme $q_{n+1} = \mathbb{S}_p\left(q_n - \tau_n[\mathcal{F}'(q)]^*[\mathcal{F}(q) - F^\delta_{\mathrm{meas}}], \tau_n\,\alpha\right)$.

[◇] For the sake of simplicity we consider the continuous version and therefore omit the pixel area or voxel volume h_N^d of the discrete version.

Realization Notes The choice of step sizes τ_n of the thresholded, nonlinear Landweber iteration (6.3) was done in [LKK13] by the Barzilai-Borwein rule, see [BB88], and is redefined by the Armijo rule, see e.g. [NW06].[◇] As already mentioned in Sec. 4.2, the overall computational effort of reconstruction schemes using the forward operator as well as the derivative of the least-squares error is reduced by evaluating both in one routine; of course, the thresholded, nonlinear Landweber scheme is a use case.

Further Generalization We will consider a more general setting. However, for the sake of simplicity we only consider real-valued contrasts again. In contrast to (6.1) it is possible to allow Schatten norms $\| \cdot \|_{\mathrm{dis},P}$ for the discrepancy term, which are a generalization of the Hilbert-Schmidt norm, and as just mentioned a modified penalty term, that also includes non-sparse cases, i.e.[†]

$$\min_{q \in X} f(q), \quad f(q) := \tfrac{1}{P} \|\mathcal{F}(q) - F^{\delta}_{\mathrm{meas}}\|^{P}_{\mathrm{dis},P} + \tfrac{\alpha}{Q} \|q\|^{Q}_{\mathrm{spa},Q} \tag{6.4}$$

for $0 < P, Q < \infty$. Similar to (2.28) we define[‡] the weighted Schatten P-norm $\| \cdot \|_{\mathrm{dis},P}$ and $\| \cdot \|_{\mathrm{spa},Q}$ as

$$\|A\|_{\mathrm{dis},P} := \|\sigma\|_P \quad \text{and} \quad \|x\|_{\mathrm{spa},Q} := \left(h_N^d \sum_i |x_i|^Q \right)^{1/Q},$$

where the vector σ contains the singular values of $(\omega^{\mathrm{s}})^{1/P} A$ with weight matrix ω^{s} due to the weights approximating the surface element on Γ_{s}, see Sec. 2.3. Of course, for the choice of $P = 2$ and $Q = 1$ the Tikhonov functionals (6.4) and (6.1) coincide.

6.2 Extended Soft-Shrinkage

As mentioned in Sec. 3.3, we will use the extended soft-shrinkage operator to solve $\min_q f(q)$ with $f(q) := f_{\mathrm{dis}}(q) + f_{\mathrm{spa}}(q) + f_{\mathrm{phy}}(q)$.

For the sake of simplicity again we consider only real-valued contrasts. Then $g(q) := f_{\mathrm{spa}}(q) + f_{\mathrm{phy}}(q)$ equals $\alpha \|q\|_{\mathrm{spa}} + \delta_{[a,b]}(q)$, cf. Fig. 6.2(f), with the indicator function $\delta_{[a,b]}$ as defined in (3.16), see Fig. 6.2(a). Again considering first-order optimality conditions, we are interested in the subdifferential of $g(q) = f_{\mathrm{spa}}(q) + f_{\mathrm{phy}}(q)$ and finally the corresponding proximal mapping.

We have already seen in Fig. 6.1 the steps from the absolute value function to the corresponding proximal mapping, which turned out to be the soft-shrinkage operator. Likewise, the steps from the indicator function to the corresponding proximal mapping are illustrated in Fig. 6.2(a)–(e). Finally, the steps from the sum of absolute value function

---

[◇] In the *source code* of IPscatt the thresholded, nonlinear Landweber scheme can be used via `seti.invNo = 1`, see the *source code documentation* of `setInvType`. In fact, this starts the scheme using the extended soft-shrinkage operator, that is discussed in the next section. However, disabling the physical bounds results in the soft-shrinkage operator.

[†] Note the slightly changed notation: in the continuous case it was $\frac{\alpha}{p} \|q\|^{p}_{L^p(D)}$, now it is $\frac{\alpha}{Q} \|q\|^{Q}_{\mathrm{spa},Q}$ in the discretized case.

[‡] In the *source code* the corresponding functions to these norms are `normws` and `normroi`. The corresponding parameters are $P =$ `seti.pNorm` (default: 2) and $Q =$ `seti.qNorm` (default: 1).

and indicator function are visualized in Fig. 6.2(f)–(j). Analogous to this, in the following we consider the individual steps for the function $g(q) = f_{\text{spa}}(q) + f_{\text{phy}}(q)$.

Remembering $\partial f_{\text{spa}} = \alpha\partial\|q\|_{\text{spa}} = \alpha h_N^d \partial\|q\|_1 = \alpha h_N^d \operatorname{sign}(q)$ with the set-valued signum function sign, see (3.15) and Fig. 3.7(c) or 6.1(b), as the subdifferential of the absolute value function. The subdifferential of $f_{\text{phy}} = \delta_{[a,b]}$ was already explained too, see (3.15) and Fig. 3.7(d) or 6.2(b). For the sum of these subdifferentials cf. Fig. 6.2(g) and 6.2(h). The left hand side of (6.2) is in this case $(I + \tau\partial g)(q)$, cf. Fig. 6.2(i), and we are finally interested in its inverse, cf. Fig. 6.2(j), that is the proximal mapping.

To this end, the auxiliary, set-valued step function $\operatorname{step}_{[a,b]}$ and their inverse $\operatorname{step}_{[a,b]}^{-1}$, both illustrated in Fig. 6.3, defined by

$$\operatorname{step}_{[a,b]}(x) := \begin{cases} \{a\} & \text{if } x \leq 0, \\ [a,b] & \text{if } x = 0, \\ \{b\} & \text{if } x \geq 0, \end{cases} \quad \text{and} \quad \operatorname{step}_{[a,b]}^{-1}(x) := \begin{cases} (-\infty, 0] & \text{if } x = a, \\ \{0\} & \text{if } x \in (a,b), \\ [0, \infty) & \text{if } x = b, \\ \emptyset & \text{otherwise,} \end{cases} \tag{6.5}$$

turn out to be useful. Of course, the element-wise application is assumed if x is a vector. Note that $\operatorname{step}_{[-1,1]}(q) = \operatorname{sign}(q) = \partial\|q\|_1$ and $\operatorname{step}_{[a,b]}^{-1}(q) = \partial\delta_{[a,b]}(q)$.

Therefore $(I + \tau\partial g)(q) = (I + \tau\,\partial f_{\text{spa}} + \tau\,\partial f_{\text{phy}})(q) = q + \kappa \operatorname{step}_{[-1,1]}(q) + \operatorname{step}_{[a,b]}^{-1}(q)$ with $\kappa = \tau\,\alpha\,h_N^d$, which is visualized in Fig. 6.2(i) for one dimension.

As can be seen in Fig. 6.2(j) the inverse of $I + \tau\partial g$ exists. To express this reflection on the identity we remember the soft-shrinkage operator (3.13) and introduce the *interval projection operator* defined for real-valued x and $a \geq b$ by

$$\mathcal{I}_{[a,b]}(x) := \begin{cases} a & \text{if } x < a, \\ x & \text{if } x \in [a,b], \\ b & \text{if } x > b. \end{cases} \tag{6.6}$$

Of course, the element-wise application is assumed if x is a vector. In any case the interval projection can be implemented by $\mathcal{I}_{[a,b]}(x) = \max\{a, \min\{b, x\}\}$.

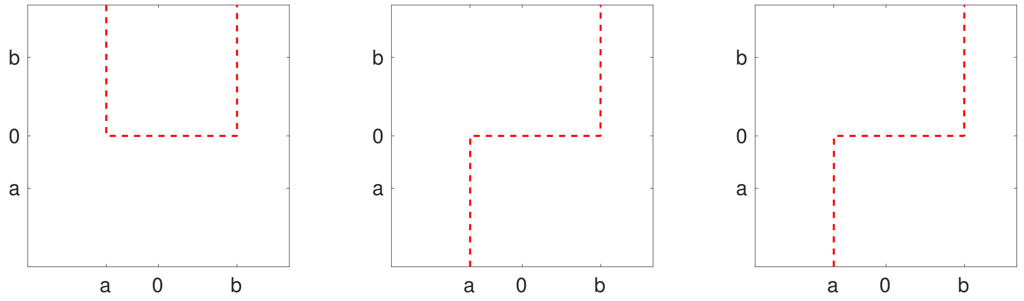
Then one can write the proximal mapping of $f_{\text{spa}} + f_{\text{phy}}$ as *extended soft-shrinkage operator*, i.e.

$$(I + \tau\partial(f_{\text{spa}} + f_{\text{phy}}))^{-1}(q) = (I + \tau\alpha h_N^d \partial\|\cdot\|_1 + \delta_{[a,b]}(\cdot))^{-1}(q) = \mathcal{I}_{[a,b]}(\mathbb{S}(q, \tau\alpha h_N^d)). \tag{6.7}$$

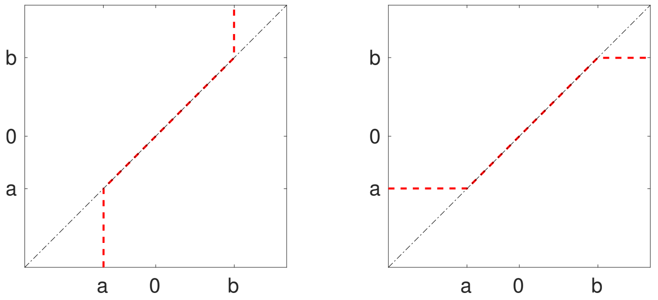
The implementation of $\mathcal{I}_{[a,b]}\mathbb{S}(z,\kappa)$ by $\max\{a, \min\{b, \operatorname{sign}(z)\, \max\{|z| - \kappa, 0\}\}\}$ follows from the implementation notes to $\mathbb{S}$, see (3.13), and $\mathcal{I}_{[a,b]}$, see (6.6).

Analogous to (6.3) we end up with an *extended thresholded, nonlinear Landweber scheme*

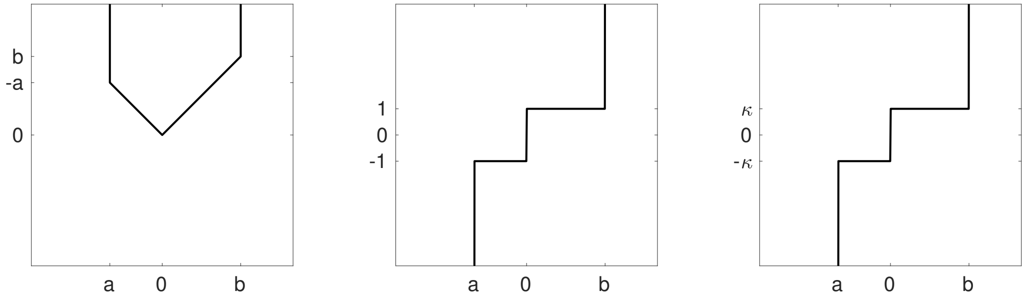
$$q_{n+1} = \mathcal{I}_{[a,b]}\left(\mathbb{S}\left(q_n - \tau_n[\mathcal{F}'(q)]^*[\mathcal{F}(q) - F_{\text{meas}}^\delta], \tau_n\,\alpha\,h_N^d\right)\right).$$



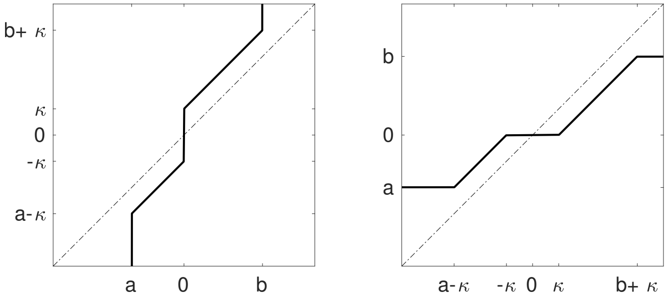
(a) $\delta_{[a,b]}$. (b) $\partial\delta_{[a,b]} = \operatorname{step}^{-1}_{[a,b]}$. (c) $\kappa\partial\delta_{[a,b]}(\cdot) = \operatorname{step}^{-1}_{[a,b]}$ with $\kappa > 0$.



(d) $(I + \kappa\partial\delta_{[a,b]})$. (e) $(I + \kappa\partial\delta_{[a,b]})^{-1}$.



(f) $|\cdot| + \delta_{[a,b]}(\cdot)$. (g) $\partial(|\cdot| + \delta_{[a,b]}(\cdot))$. (h) $\kappa\partial(|\cdot| + \delta_{[a,b]}$ with $\kappa > 0$.



(i) $I + \kappa\partial(|\cdot| + \delta_{[a,b]}(\cdot)) = I + \kappa \operatorname{step}_{[-1,1]} + \operatorname{step}^{-1}_{[a,b]}$. (j) $(I + \kappa\partial(|\cdot| + \delta_{[a,b]}(\cdot)))^{-1}$.

Figure 6.2: (a)–(e) From the indicator function to the corresponding proximal mapping: (a) Indicator function. (b) The subdifferential of the indicator function. (c) Subdifferential taking into account the factor κ. (Keep in mind that $\kappa \operatorname{step}^{-1}_{[a,b]}(x) = \operatorname{step}^{-1}_{[a,b]}(x)$.) (d) Adding the identity. (e) The inverse is the proximal mapping.
(f)–(j) From the sum of absolute value function and indicator function to the corresponding proximal mapping: (f) Sum of absolute value function and indicator function. (g) The subdifferential. (h) Subdifferential taking into account the factor κ. (i) Adding the identity. (j) The inverse of $I + \kappa\partial(|\cdot| + \delta_{[a,b]})$ is the proximal mapping of $|\cdot| + \delta_{[a,b]}$. In this case it is the extended soft-shrinkage operator.

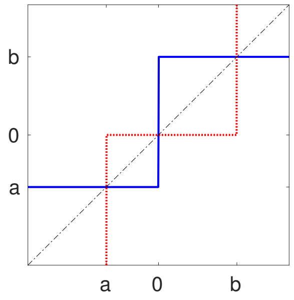


Figure 6.3: Illustration of $\mathrm{step}_{[a,b]}$ (blue solid) and $\mathrm{step}_{[a,b]}^{-1}$ (red dotted).

6.3 Soft-Shrinkage with Wavelets

In many applications the usage of *sparsity* in a *wavelet* basis instead of the pixel basis, see [DDD04], is a successful concept. As already mentioned in Ch. 1 the source code from Lechleiter and Kazimierski to [LKK13] includes the thresholded, nonlinear Landweber scheme corresponding to the Tikhonov functional, that enforces sparsity in a wavelet basis. However, the thresholded, nonlinear Landweber scheme enforcing sparsity in a wavelet basis does not outcompete the same method enforcing sparsity in the pixel basis.

There are promising reconstruction results with sparsity in a wavelet basis in the case of the linearized model given by the Born approximation, see [Str14, Ch. 5] or the results of the software package GLIDE in Ch. 5.

Therefore based on [LKK13] it was experimented with a weighting of the result of the discrete wavelet transform. Starting with the computation of white noise the underlying idea was to give high levels, i.e. low frequencies, a low weight and to give low levels, i.e. high frequencies, a high weight. Note that this did also not result in better reconstructions. Therefore this was not a subject of ongoing research. However, the tested improvements are still part of IPscatt and are briefly introduced.

Tikhonov Functional To enforce sparsity in a wavelet basis we introduce the *discrete wavelet transform*[◇] $\mathcal{W}$, its inverse $\mathcal{W}^{-1}$ as well as weights w for the wavelet coefficients and adapt (6.4),

$$\min_{q \in X} f(q), \quad f(q) := \frac{1}{P}\|\mathcal{F}(q) - F^{\delta}_{\text{meas}}\|^{P}_{\text{dis},P} + \frac{\alpha}{Q}\|w \odot \mathcal{W}(q)\|^{Q}_{\text{spa},Q}. \tag{6.8}$$

The corresponding thresholded, nonlinear Landweber scheme enforces sparsity in a wavelet basis.[†]

Computation of the Weights For the computation of the weights w we consider a random matrix $\mathcal{R}$, that has the same size as the grid of the region of interest, with

---

[◇] As already mentioned in Ch. 1 the discrete wavelet transform and its inverse were already part of the source code to [LKK13]. This includes the Haar wavelet, the LeGall-5/3 wavelet and the Cohen-Daubechies-Feauveau-9/7 (CDF) wavelet.

[†] In the *source code* of IPscatt the thresholded, nonlinear Landweber scheme enforcing sparsity in wavelet basis can be used via seti.invNo = 2, see the *source code documentation* of setInvType. In fact this uses the extended soft-shrinkage operator. Therefore it also includes physical bounds.

normally distributed entries. Then the *white noise* matrix is $\mathcal{N} = \frac{1}{n}\sum_{i=1}^{n} |\mathcal{W}(\mathcal{R})|^2$, where n is the number of samples and $|\cdot|$ is meant element-wise for once. The *weight matrix* can be defined by $w = 1/\mathcal{N}$, where an element-wise division is meant. In addition, this weight matrix can be modified: to give high levels a low weight we multiply the entry with a small factor; to give low levels a high weight we multiply the entry with a high factor.

Further Generalization In a more general setting we allow another exponent for the discrepancy term than in the norm$^\diamond$, cf. (6.8),

$$\min_{q\in X} f(q), \quad f(q) := \frac{1}{p}\|\mathcal{F}(q) - F^{\delta}_{\text{meas}}\|^{p}_{\text{dis},P} + \frac{\alpha}{Q}\|w \odot \mathcal{W}(q)\|^{Q}_{\text{spa},Q}. \tag{6.9}$$

If we use thresholded, nonlinear Landweber scheme, $p = P$ is required, but Q can be adapted. This adaption of Q was already a part of the code basis provided by Kazimierski and Lechleiter. It basis on the soft-shrinkage operator that can also be used for non-sparse penalty terms, is denoted by $\mathbb{S}_Q$ and was defined in Sec. 6.1.[†] However, another method allows a choice $p \neq P$ as well, see Sec. 7.7.

[◇] In the *source code* are the parameters defined as $p =$ `seti.p`, $P =$ `seti.pNorm` and $Q =$ `seti.qNorm`.

[†] Note that the corresponding soft-shrinkage operator in Sec. 6.1 is denoted by $\mathbb{S}_p$ instead of $\mathbb{S}_Q$ as the notation of the norm differs.

7 Reconstruction with a Primal-Dual Algorithm

In the last chapter we minimized $f_{\text{dis}} + f_{\text{spa}} + f_{\text{phy}}$. As mentioned in Sec. 3.3 it is difficult to additionally take into account the total variation penalty f_{tv}, i.e. minimize (3.11). However, the combined effect with the sparsity penalty f_{spa}, see Sec. 3.2, is a motivation. We will use a sophisticated algorithm—the primal-dual algorithm (PDA) by Pock, Bischof, Cremers and Chambolle, see [PCBC09] or [CP11a, Alg. 1]. However, we have to deal with two problems: First, the operator $\mathcal{F}$ is nonlinear, but PDA requires a linear operator. Thus we need a linearization. Second, the contrast q is complex-valued, but PDA requires real vector spaces. Therefore we need a transformation operator T and its inverse T^{-1} mapping between $\mathbb{C}$ and $\mathbb{R} \times \mathbb{R}$.

We will consider the proposed variational reconstruction scheme in detail.[◇] In particular, this includes the computation of the occurring proximal mappings. In addition, we discuss the influence of parameter choices by numerical examples.

Note that this chapter essentially bases on the already published results in [BKL17, Secs. 4–6]. However, it should be pointed out that the just-mentioned transformation operator was not developed in the context of this thesis: as already mentioned in Ch. 1, the provided code basis from Kazimierski and Lechleiter already included the use of the shrinkage separately to the real and imaginary part and therefore the concept of this transformation operator.

7.1 Linearization of the Problem

Initially, it follows a brief overview of the concepts including the linearization of the problem. We will enter into the details in the next sections.

Application of the Primal-Dual Algorithm on Inverse Scattering Problems We want to connect the Tikhonov functional (3.11) with the so-called *primal problem* $\min_q\{F(Kq) + G(q)\}$, where F and G are convex, lower semi-continuous (l.s.c.) functions, G is proper in addition, and K is a continuous linear operator, see Sec. 7.2 for all assumptions. Hence, we can solve the inverse scattering problem with the mentioned primal-dual algorithm (PDA). A possible assignment is

$$f(q) := \underbrace{\tfrac{1}{2}\|\mathcal{F}(q) - F^{\delta}_{\text{meas}}\|^2_{\text{dis}} + \beta\,\|\nabla q\|_{\text{tv}}}_{=:F(K(q))} + \underbrace{\alpha\,\|q\|_{\text{spa}} + \delta(q)}_{=:G(q)}\,.$$

Remember that $\delta(q) = \delta_{[a,b]}(\mathrm{Re}(q)) + \delta_{[c,d]}(\mathrm{Im}(q))$, see Sec. 3.2.2.

[◇] In the *source code* of IPscatt this reconstruction scheme is used by default. It corresponds to the case `seti.invNo = 6`, see the *source code documentation* of setInvType.

We recall that the forward operator $\mathcal{F}$ is nonlinear. Hence, $K(q) := (\mathcal{F}(q), \beta\,\nabla q)$ is nonlinear. (Clearly, we have to define $F(y_1, y_2) := \frac{1}{2}\|y_1 - F^{\delta}_{\text{meas}}\|^2_{\text{dis}} + \|y_2\|_{\text{tv}}$.) However, the primal-dual algorithm requires a linear operator K. Therefore we consider the Tikhonov functional with $q + h$ instead of q

$$f(q+h) := \tfrac{1}{2}\|\mathcal{F}(q+h) - F^{\delta}_{\text{meas}}\|^2_{\text{dis}} + \beta\,\|\nabla(q+h)\|_{\text{tv}} + \alpha\,\|q+h\|_{\text{spa}} + \delta(q+h).$$

To get rid of the nonlinearity of K, we use the linearization $\mathcal{F}(q+h) \approx \mathcal{F}(q) + \mathcal{F}'(q)[h]$ with the Fréchet derivative. Then the extended Tikhonov functional is

$$f_q(h) := \underbrace{\tfrac{1}{2}\|\mathcal{F}'(q)[h] + \mathcal{F}(q) - F^{\delta}_{\text{meas}}\|^2_{\text{dis}} + \beta\,\|\nabla(q+h)\|_{\text{tv}}}_{=:F(Kh)} + \underbrace{\alpha\,\|q+h\|_{\text{spa}} + \delta(q+h)}_{=:G(h)}. \tag{7.1}$$

We denote this minimization with respect to h as post-linearization problem. The assignment of F and G is defined analogously to the previous. However, they are defined for a fixed contrast q, i.e. $F(y_1, y_2) := \frac{1}{2}\|y_1 + \mathcal{F}(q) - F^{\delta}_{\text{meas}}\|^2_{\text{dis}} + \|y_2 + \beta\,\nabla(q)\|_{\text{tv}}$ and $Kh := (\mathcal{F}'(q)[h], \beta\,\nabla h)$. In particular, this K is linear.

Proposed Variational Reconstruction Scheme We continue with an overview of the steps that the *proposed variational reconstruction scheme* processes. Stopping criteria for the two parts of the reconstruction algorithm are presented elsewhere: a stopping criterion for the *outer iteration* is Morozov's discrepancy principle (3.18); stopping criteria for the *inner iteration* are presented in Sec. 7.5. A full pseudo code of the proposed variational reconstruction scheme with all technical details like the corresponding proximal mappings is given in Sec. 7.5. The scheme can be summarized as follows:

1. Initialization: Set the contrast $q = 0$.
2. Minimize $f_q(h)$ with respect to h, i.e. (7.1), with the primal-dual algorithm, that is given in [CP11a]. We call this iteration an *inner iteration*. If a stopping criterion is fulfilled, then go to step 3.
3. Update $q := q + h$. We call this linearization step an *outer iteration*. If a stopping criterion is fulfilled, then the variational reconstruction scheme ends.

7.2 The Primal-Dual Algorithm and its Derivation

In this section we consider the primal-dual algorithm (PDA) from [CP11a, Alg. 1]. Furthermore, we discuss its derivation for a better understanding. For this section it is useful to remember the introduction to the convex analysis in Sec. 3.3.1. In particular, we need the definition of a proper, lower semi-continuous (l.s.c.) and convex function, the Fenchel conjugate and the subdifferential.

Assumptions for the Primal-Dual Algorithn Let the two finite-dimensional real vector spaces X, Y be equipped with inner products $\langle\cdot,\cdot\rangle_X$ as well as $\langle\cdot,\cdot\rangle_Y$ and associated norms $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$. Let $K\colon X \to Y$ be a continuous linear operator with operator norm $\|K\| := \max_{0 \neq x \in X} \|Kx\|_Y / \|x\|_X$. Let $G\colon X \to [0, \infty]$ be proper, l.s.c. and convex and $F\colon X \to [0, \infty)$ l.s.c. and convex. Of course, F is then also proper because the codomain does not include infinity.

Remember that the Fenchel conjugate $F^*: X \to [0, \infty]$ is proper, l.s.c. and convex too because of Th. 3.3.1.

Primal and Primal-Dual Problem First, we want to solve the (nonlinear) *primal problem*, see [CP11a, Alg. 1],

$$\min_{x \in X} \{F(Kx) + G(x)\}. \tag{7.2}$$

Second, we want to connect the Tikhonov functional (3.11) with this primal problem to solve it by the primal-dual algorithm later on. Therefore we need a reformulation of (3.11). However, we start with a reformulation of the *primal problem* into a *dual problem* inserting some interims steps into the short derivation, that is given in [BL11, Sec. 6.2.4]. This reformulation is done via reformulations as primal-dual problem and dual-primal problem.

Note that the asterisk * is an ambiguous notation as it is used for the Fenchel conjugates F^* and G^* of F and G, for the adjoint K^* of K and for the dual space Y^* of Y.

To reformulate the primal problem into the primal-dual problem we will consider the term $F(Kx)$, set $z := Kx$ and remember that the Fenchel conjugate for f on Y is $f^*(z) := \max_{y \in Y^*}\{\langle z, y\rangle - f(y)\}$ for $z \in Y^* \simeq Y$, cf. (3.14). We use real finite-dimensional vector spaces with inner products. Therefore we can identify Y^* with Y as well as X^* with X again, see Sec. 3.3.1. Furthermore, we assume that the supremum is reached. Therefore we use the maximum instead of the supremum.

In fact, for reformulation we need the assumption that at least one solution $(\hat{x}, \hat{y})$ of the primal problem exists because this implies the so-called *strong duality*, i.e. the primal problem equals the dual problem, see Th. 7.2.1 below.

Remember $f = f^{**}$ from Th. 3.3.1. If we set $f = F^*$ the result is $F(Kx) = F(z) = (F^*)^*(z) = \max_{y \in Y^*}\{\langle z, y\rangle - F^*(y)\} = \max_{y \in Y^*}\{\langle Kx, y\rangle - F^*(y)\}$. Replacing $F(Kx)$ in the primal problem with the last term results in the *primal-dual problem*, see [CP11a, Sec. 2], which is an example of the saddle-point problem,

$$\min_{x \in X} \max_{y \in Y^*} \{\langle Kx, y\rangle - F^*(y) + G(x)\}. \tag{7.3}$$

Dual-Primal and Dual Problem We remember the primal-dual problem (7.3), switch minimum and maximum, that will be allowed by Th. 7.2.1, and use the adjoint operator K^* by $\langle Kx, y\rangle = \langle x, K^* y\rangle$ to obtain the *dual-primal problem*

$$\max_{y \in Y^*} \min_{x \in X} \{\langle x, K^* y\rangle - F^*(y) + G(x)\}. \tag{7.4}$$

Next, we will derive the dual problem. The Fenchel conjugate of G is by definition, cf. (3.14), $G^*(z) = \max_{x \in X}\{\langle z, x\rangle - G(x)\}$. We set $z := -K^* y$. As we consider real vector spaces, the inner product is symmetric, such that $\langle -K^* y, x\rangle = -\langle x, K^* y\rangle$. Then we receive $G^*(-K^* y) = \max_{x \in X}\{-\langle x, K^* y\rangle - G(x)\} = \max_{x \in X} -f(x)$ for $f(x) := \langle x, K^* y\rangle + G(x)$. If f is a convex function, then $-f(x)$ is concave and $\max\{-f(x)\} = -\min f(x)$. Therefore $G^*(-K^* y) = -\min_{x \in X} f(x)$. Multiplying both sides by (-1) and inserting $f(x)$ again ends up with $-G^*(-K^* y) = \min_{x \in X} f(x) = \min_{x \in X}\{\langle x, K^* y\rangle + G(x)\}$. We put this inside the dual-primal problem (7.4) to receive $\max_{y\in Y^*}\{-G^*(-K^*y) - F^*(y)\}$ and end up with the *dual problem*

$$\max_{y\in Y^*} -\{G^*(-K^*y) + F^*(y)\}. \tag{7.5}$$

The statement of the next theorem, the Fenchel-Rockafellar duality, is the equality of the primal and dual problem, see (7.2) and (7.5). Therefore it delivers the assumptions to switch minimum and maximum, see [BL11, Satz 6.68]. The equality is called *strong duality*$^\diamond$, see [Roc67, Sec. 5].

Theorem 7.2.1 (Fenchel-Rockafellar duality)**.** *Let* $F\colon Y \to \mathbb{R}\cup\{+\infty\}$ *and* $G\colon X \to \mathbb{R}\cup\{+\infty\}$ *be proper, l.s.c. and convex on real Banach spaces* X *and* Y*. Let* $K\colon X \to Y$ *be linear and bounded. We assume that the primal problem* (7.2) *has at least one solution* $\hat{x} \in X$*. If there is a point* $x_0 \in X$ *such that* $F(Kx_0) < \infty$*,* $G(x_0) < \infty$ *and* F *is continuous in* Kx_0*, then the primal problem* (7.2) *equals the dual problem* (7.5)*, i.e.*

$$\min_{x\in X}\{F(Kx) + G(x)\} = \max_{y\in Y^*} -\{G^*(-K^*y) + F^*(y)\}.$$

The next corollary shows the connection between strong duality and two extremality conditions. We will derive these conditions afterwards, see (7.6) and (7.7).

Corollary 7.2.2 (Fenchel-Rockafellar optimality system, see [BL11, Korollar 6.70])**.** *Let be* $F\colon X \to \mathbb{R}\cup\{+\infty\}$ *and* $G\colon Y \to \mathbb{R}\cup\{+\infty\}$ *proper, l.s.c. and convex. If strong duality is given, i.e. the primal problem equals the dual problem, i.e.*

$$\min_{x\in X}\{F(Kx) + G(x)\} = \max_{y\in Y^*} -\{G^*(-K^*y) + F^*(y)\},$$

then the pair $(\hat{x}, \hat{y}) \in X \times Y^*$ *solves the primal-dual problem if and only if the following* extremality conditions *are fulfilled*

$$K\hat{x} \in \partial F^*(\hat{y}) \quad \textit{and} \quad -(K^*\hat{y}) \in \partial G(\hat{x}).$$

Note that we assumed the existence of a solution during the derivation before. There is a theorem, see [Roc67, Th. 9], that strong duality holds for Kuhn-Tucker functions $\mathcal{K}(x,y) := \langle Kx, y\rangle - F^*(y) + G(x)$ with a saddle-point.

Extremality Conditions Both extremality conditions, that we will derive, are given in [Roc67, Sec. 8]. If we want to solve the primal problem (7.2) the necessary condition is that $0 \in \partial\{F(Kx) + G(x)\}$. The sufficient condition is uninteresting because F and G are convex. We consider the primal-dual problem (7.3) with a fixed $\hat{x}$ instead of x, $y = \arg\max_{y\in Y^*}\{\langle K\hat{x}, y\rangle - F^*(y) + G(\hat{x})\}$. The subdifferential with respect to y is $\partial_y\left\{\langle K\hat{x}, y\rangle - F^*(y) + G(\hat{x})\right\} = K\hat{x} - \partial F^*(y) + 0$. Remember that the subdifferential may be a set, such that the necessary condition is that 0 is an element of this set. The reformulation of this extremality condition results in the first extremality condition,

$$K\hat{x} \in \partial F^*(\hat{y}). \tag{7.6}$$

$^\diamond$There is also a *weak duality*. This is described e.g. in [Roc67, Sec. 6].

For the second extremality condition we consider the dual-primal problem (7.4) with a fixed $\hat{y}$ instead of y as $x = \arg\min_{x \in X}\{\langle x, K^*\hat{y}\rangle - F^*(\hat{y}) + G(x)\}$. The subdifferential with respect to x is $K^*\hat{y} + \partial G(x)$. This results in the second extremality condition,

$$-K^*\hat{y} \in \partial G(\hat{x}). \tag{7.7}$$

Fixed-Point Iterations As the PDA will base on fixed-point iterations, we derive those from the extremality conditions. Multiplying both sides of the first extremality condition (7.6) with $\sigma > 0$ and adding $\hat{y}$ we obtain

$$\hat{y} + \sigma K\hat{x} \in \hat{y} + \sigma\partial F^*(\hat{y}) = (I + \sigma\partial F^*)(\hat{y}) \quad \text{and} \quad \hat{y} = (I + \sigma\partial F^*)^{-1}(\hat{y} + \sigma K\hat{x}).$$

For the right equation remember that F is proper, convex and l.s.c. and therefore F^* has the same properties, such that Lemma 3.3.2 (about the equality of resolvent operator and proximal mapping) can indeed be applied on F^*. Again, the clue of this lemma is that the resolvent $(I + \sigma\partial F^*)^{-1}$ delivers the unique minimizer. Therefore we have, in fact, equality. Finally, we end up with a fixed-point iteration from the first extremality condition,

$$y^{n+1} = (I + \sigma\partial F^*)^{-1}(y^n + \sigma Kx^n). \tag{7.8}$$

The second extremality condition (7.7) also results in a fixed-point iteration. Multiplying $\tau > 0$ and adding $\hat{x}$ results in $\hat{x} - \tau K^*\hat{y} \in \hat{x} + \tau\partial G(\hat{x}) = (I + \tau\partial G)(\hat{x})$. Remember that G is proper, l.s.c. and convex by definition. Therefore Lemma 3.3.2 can be applied on G and we receive $\hat{x} = (I + \tau\partial G)^{-1}(\hat{x} - \tau K^*\hat{y})$. We end up with the second fixed-point iteration from the second extremality condition,

$$x^{n+1} = (I + \tau\partial G)^{-1}(x^n - \tau K^*y^{n+1}). \tag{7.9}$$

Primal-Dual Algorithm (PDA) With the two fixed-point iterations (7.8), (7.9) and an over-relaxation step (to speed up the scheme) we have all ingredients to formulate the *primal-dual algorithm (PDA)*.

Algorithm 7.2.3 (Primal-dual algorithm for convex problems, see [CP11a, Alg. 1])**.**

- Initialization: Choose the *primal step size* $\tau > 0$, the *dual step size* $\sigma > 0$, initial vectors $(x^0, y^0) \in X \times Y$ (usually $x^0 = 0$, $y^0 = 0$) and set $\bar{x}^0 = x^0$.
- Iterations ($n \geq 0$): Update $x^n, y^n, \bar{x}^n$ as follows:

$$\begin{aligned} &1.\quad y^{n+1} = (I + \sigma\partial F^*)^{-1}(y^n + \sigma K\bar{x}^n),\\ &2.\quad x^{n+1} = (I + \tau\partial G)^{-1}(x^n - \tau K^*y^{n+1}),\\ &3.\quad \bar{x}^{n+1} = 2x^{n+1} - x^n. \end{aligned} \tag{7.10}$$

Note that in general the third step is $\bar{x}^{n+1} = x^{n+1} + \theta(x^{n+1} - x^n)$ for a $\theta \in [0, 1]$. For the sake of simplicity we choose $\theta = 1$. The challenge is to compute the resolvents $(I + \sigma\partial F^*)^{-1}$ and $(I + \tau\partial G)^{-1}$. For our problem (7.1) this is done in Secs. 7.4.1 and 7.4.2.

Theorem 7.2.4 (Convergence of PDA)**.** *We assume that the primal-dual problem* (7.3) *has a saddle-point* $(\hat{x}, \hat{y})$ *and the* step sizes τ*,* σ *are such that* $\tau\sigma\|K\|^2 < 1$*. Then there exists a saddle-point* (x^*, y^*)*, such that* $x^n \to x^*$ *and* $y^n \to y^*$ *as* $n \to \infty$*.*

A proof of Th. 7.2.4 for real vector spaces can be found in [CP11a, Th. 1] and essentially consists of estimates. Note that real vector spaces are indeed crucial for Alg. 7.2.3 as the symmetry of the inner product is exploited in the proof and both subgradient and Fenchel conjugate are real-valued concepts. We will discuss a related minimization scheme for complex-valued functionals in Sec. 7.3.

Note that the extension of the primal-dual algorithm to Banach spaces has been considered in [HH14], which also indicates numerical examples for one-dimensional deconvolution and two-dimensional phase reconstruction.

7.3 Transformation of the Complex Problem

In order to construct the linear operator K in the primal-dual algorithm (PDA), see Alg. 7.2.3, as an operator between real vector spaces we transform complex-valued quantities by means of the identification of $\mathbb{C}$ as $\mathbb{R} \times \mathbb{R}$.

Transformation of Complex into Real Vectors Formally, the transport between complex and real-valued form is carried out by the *transformations*:

$$\begin{array}{rclrcl} \mathrm{T} & : & \mathbb{C} \to \mathbb{R} \times \mathbb{R}, & \mathrm{T}(x) & := & (\mathrm{Re}(x), \mathrm{Im}(x)), \\ \mathrm{T}^{-1} & : & \mathbb{R} \times \mathbb{R} \to \mathbb{C}, & \mathrm{T}^{-1}(y) & = & y^{\mathrm{Re}} + \mathrm{i}\, y^{\mathrm{Im}} \quad \text{where} \quad y = (y^{\mathrm{Re}}, y^{\mathrm{Im}}). \end{array} \tag{7.11}$$

It is easy to keep in mind

$$\begin{pmatrix} a_1 + \mathrm{i}b_1 \\ \vdots \\ a_n + \mathrm{i}b_n \end{pmatrix} \underset{\mathrm{T}^{-1}}{\overset{\mathrm{T}}{\rightleftarrows}} \Big(a_1, \dots, a_n, b_1, \dots, b_n\Big)^\top .$$

Of course, as usual the above pair extends by the element-wise application to vectors, matrices and arrays. For example, $\mathbb{C}^n$ is identified with $\mathbb{R}^n \times \mathbb{R}^n$, where the first $\mathbb{R}^n$ is responsible for the real parts and the second $\mathbb{R}^n$ for the imaginary part. In the above definitions the set of complex numbers is considered to be a vector space over the field of real numbers and not (as is usually the case) over the complex numbers. This setting is necessary to ensure that the operators are linear. Formally this can only be the case if the related operator maps between vector spaces over the same field.

Reformulated Tikhonov Problem In the next step we formulate the post-linearization problem (7.1) in real vector spaces. For the remainder of this section we assume that the linearization point q as well as the minimization variable h are already in the transported real vector space. Due to (3.4), this transported version of X is given by the space$^\diamond$

$$X_{\mathbb{R}} := \mathbb{R}^{N_{\mathrm{D}}} \times \mathbb{R}^{N_{\mathrm{D}}}. \tag{7.12}$$

$^\diamond$In [BKL17, Sec. 4.5] a typo states the transported version of $X_{\mathbb{R}}$ instead of X.

First, we define the auxiliary quantities

$$\begin{aligned} K_{\mathrm{dis}} &:= \mathrm{T}[\mathcal{F}'(\mathrm{T}^{-1}q)]\mathrm{T}^{-1}, & v_{\mathrm{dis}} &:= \mathrm{T}(\mathcal{F}(\mathrm{T}^{-1}q) - F^{\delta}_{\mathrm{meas}}), \\ K_{\mathrm{tv}} &:= \beta\, \mathrm{T}\nabla\mathrm{T}^{-1}, & v_{\mathrm{tv}} &:= \beta\, \mathrm{T}\nabla\mathrm{T}^{-1}q. \end{aligned} \tag{7.13}$$

Second, transporting the functionals of (7.1) into the real vector space $X_{\mathbb{R}} := \mathbb{R}^{N_{\mathrm{D}}} \times \mathbb{R}^{N_{\mathrm{D}}}$ yields the

$$\begin{aligned} \text{discrepancy (of post-linearization problem)} \quad & f_{\mathrm{dis}}(h) := \tfrac{1}{2}\|K_{\mathrm{dis}}h + v_{\mathrm{dis}}\|^2_{\mathrm{dis},\mathbb{R}}, \\ \text{the sparsity penalty} \quad & f_{\mathrm{spa}}(h) := \alpha\|h+q\|_{\mathrm{spa},\mathbb{R}}, \\ \text{the total variation penalty} \quad & f_{\mathrm{tv}}(h) := \|K_{\mathrm{tv}}h + v_{\mathrm{tv}}\|_{\mathrm{tv},\mathbb{R}}, \\ \text{and the penalty for physical bounds} \quad & f_{\mathrm{phy}}(h) := \delta(h+q). \end{aligned} \tag{7.14}$$

Roughly, the norms with the subindex $\mathbb{R}$ are real-valued identifications of their complex counterparts, i.e. $\|x\| = \|Tx\|_{\mathbb{R}}$. Of course, in the next steps these terms are defined explicitly.

To equip the space $X_{\mathbb{R}}$, see (7.12), with an inner product we rely on the original inner product of X, cf. (3.4) and (2.27),

$$\langle x, y\rangle_{\mathrm{roi},\mathbb{R}} := \mathrm{Re}\langle \mathrm{T}^{-1}x, \mathrm{T}^{-1}y\rangle_{\mathrm{roi}} = h_N^d \sum_i \left(x_i^{\mathrm{Re}} y_i^{\mathrm{Re}} + x_i^{\mathrm{Im}} y_i^{\mathrm{Im}}\right), \quad x, y \in X_{\mathbb{R}}. \tag{7.15}$$

Of course, this is done so to ensure that the topology on the original and transported space is the same. This, in particular, ensures that the derivatives do not change. Next, we define

$$\|x\|_{\mathrm{spa},\mathbb{R}} := h_N^d \sum_i \left(|x_i^{\mathrm{Re}}| + |x_i^{\mathrm{Im}}|\right), \quad x = (x^{\mathrm{Re}}, x^{\mathrm{Im}}) \in X_{\mathbb{R}} \tag{7.16}$$

$$\text{and} \quad \delta(x) := \delta_{[a,b]}(x^{\mathrm{Re}}) + \delta_{[c,d]}(x^{\mathrm{Im}}), \quad x = (x^{\mathrm{Re}}, x^{\mathrm{Im}}) \in X_{\mathbb{R}}.$$

In the next step we define $\|\cdot\|_{\mathrm{tv},\mathbb{R}}$. For the sake of simplicity we discuss the 2D case, since the changes necessary for the 3D case are straightforward. First, we recall that in 2D the $\|\cdot\|_{\mathrm{tv}}$ norm operates on the gradients of vectors from X, i.e. the space $X \times X$. Consequently, the norm $\|\cdot\|_{\mathrm{tv},\mathbb{R}}$ must operate on

$$Y_{\mathrm{tv},\mathbb{R}} := X_{\mathbb{R}} \times X_{\mathbb{R}}.$$

This results in the definition, cf. (3.9),

$$\|a\|_{\mathrm{tv},\mathbb{R}} := h_N^2 \sum_i |(a_i^{(1),\mathrm{Re}}, a_i^{(1),\mathrm{Im}}, a_i^{(2),\mathrm{Re}}, a_i^{(2),\mathrm{Im}})|, \quad \text{(2D case)}, \tag{7.17}$$

where of course $((a^{(1),\mathrm{Re}}, a^{(1),\mathrm{Im}}), (a^{(2),\mathrm{Re}}, a^{(2),\mathrm{Im}})) \in Y_{\mathrm{tv},\mathbb{R}}$ and

$$|(a_i^{(1),\mathrm{Re}}, a_i^{(1),\mathrm{Im}}, a_i^{(2),\mathrm{Re}}, a_i^{(2),\mathrm{Im}})|^2 := |a_i^{(1),\mathrm{Re}}|^2 + |a_i^{(1),\mathrm{Im}}|^2 + |a_i^{(2),\mathrm{Re}}|^2 + |a_i^{(2),\mathrm{Im}}|^2. \tag{7.18}$$

In the same way as for the space X, we receive for Y_{dis}

$$Y_{\text{dis},\mathbb{R}} := \mathbb{R}^{N_{\text{s}}\times N_{\text{i}}} \times \mathbb{R}^{N_{\text{s}}\times N_{\text{i}}}$$

and the inner product on $Y_{\mathbb{R}}$, cf. (2.28),

$$\langle x, y\rangle_{\text{dis},\mathbb{R}} := \operatorname{Re}\langle \mathrm{T}^{-1}x, \mathrm{T}^{-1}y\rangle_{\text{dis}}, \quad x, y \in Y_{\text{dis},\mathbb{R}}. \tag{7.19}$$

All the above elementary but indispensable work allows us to identify the three main ingredients of the primal problem (7.2), i.e. the operator K (with its domain and codomain space) and the functionals F and G. The remainder of the section is devoted to that task.

Operator K and its Adjoint We begin with the definition of the related linear operator K. In our approach it shall be given by

$$K := (K_{\text{dis}}, K_{\text{tv}}),$$

where K_{dis} and K_{tv} are as in (7.13). All operators involved in the definition of K are $\mathbb{R}$-linear. Therefore the operator K is $\mathbb{R}$-linear too. The proofs are straightforward. Of course, some operators are even $\mathbb{C}$-linear which does not bother us here. (Note that T^{-1} is $\mathbb{C}$-linear but T only $\mathbb{R}$-linear.)

For the sake of the proposed algorithm we also need the exact form of the adjoint of K. Therefore we need to discuss closely the domain space and the codomain space of K together with their related inner products/norms. After the considerations of the previous paragraph it is clear that the canonical choice for the domain space is the space $X_{\mathbb{R}}$. The canonical choice for the codomain space is the product space $Y := Y_{\text{dis}} \times Y_{\text{tv}}$, since K_{dis} maps into Y_{dis} and K_{tv} maps into Y_{tv}. Of course, instead of the penalty $\|\cdot\|_{\text{tv},\mathbb{R}}$ the space Y_{tv} is equipped with the norm generated by the inner product

$$\langle x, y\rangle_{\text{tv},\mathbb{R}} := \langle x^{(1)}, y^{(1)}\rangle_{\text{roi},\mathbb{R}} + \langle x^{(2)}, y^{(2)}\rangle_{\text{roi},\mathbb{R}}, \quad x, y \in Y_{\text{tv},\mathbb{R}}. \tag{7.20}$$

The reader should be aware that the norm induced by that inner product is not the functional $\|\cdot\|_{\text{tv},\mathbb{R}}$ in (7.17). The indices “tv” were chosen in both cases to indicate on which objects both quantities act.

In a similar way to (7.20) the inner product on Y is generated by

$$\langle (y_{\text{dis}}, y_{\text{tv}}), (z_{\text{dis}}, z_{\text{tv}})\rangle_Y := \langle y_{\text{dis}}, z_{\text{dis}}\rangle_{\text{dis},\mathbb{R}} + \langle y_{\text{tv}}, z_{\text{tv}}\rangle_{\text{tv},\mathbb{R}} \tag{7.21}$$

with the inner products (7.19) and (7.20). Altogether

$$K : X_{\mathbb{R}} \to Y_{\text{dis},\mathbb{R}} \times Y_{\text{tv},\mathbb{R}}.$$

The domain and codomain space are both Hilbert spaces and we identify their duals with the spaces themselves. Since

$$\langle Kh, y\rangle_Y = \langle (K_{\text{dis}}h, K_{\text{tv}}h), (y_{\text{dis}}, y_{\text{tv}})\rangle_Y = \langle h, K_{\text{dis}}^{*}\, y_{\text{dis}}\rangle_{\text{roi},\mathbb{R}} + \langle h, K_{\text{tv}}^{*}\, y_{\text{tv}}\rangle_{\text{roi},\mathbb{R}},$$

the adjoint of K is given by

$$K^* : Y_{\mathrm{dis},\mathbb{R}} \times Y_{\mathrm{tv},\mathbb{R}} \to X_{\mathbb{R}}, \quad K^*(y_{\mathrm{dis}}, y_{\mathrm{tv}}) = K^*_{\mathrm{dis}}(y_{\mathrm{dis}}) + K^*_{\mathrm{tv}}(y_{\mathrm{tv}})\,. \tag{7.22}$$

To determine the explicit form of K^*_{dis} we connect the definition of $K_{\mathrm{dis}} = \mathrm{T}[\mathcal{F}'(\mathrm{T}^{-1}q)]\mathrm{T}^{-1}$, see (7.13), with the efficient evaluation of $[\mathcal{F}'(q)]^*$, see (2.25) and (2.29). For K^*_{tv}, on the other hand, we have to consider the *adjoint of the discrete gradient operator* which is the negative *discrete divergence operator* because of $\langle \nabla a, b\rangle_Y = \langle a, -\operatorname{div} b\rangle_X$, see [CP11a, Sec. 6.1]. For the sake of brevity we consider the form of the operator div for the 2D case only, the 3D holds mutatis mutandis. For arrays of size N (with underlying grid size h) the 2D version of div is given by[◇]

$$\operatorname{div} b = (\operatorname{div} b)^{(1)} + (\operatorname{div} b)^{(2)},$$

where

$$(\operatorname{div} b)^{(1)}_{i,j} := \begin{cases} \frac{b^{(1)}_{1,j}}{h} & \text{if } i = 1, \\ \frac{b^{(1)}_{i} - b^{(1)}_{i-1,j}}{h} & \text{if } 1 < i < N, \\ -\frac{b_{N-1,j}}{h} & \text{if } i = N, \end{cases} \qquad (\operatorname{div} b)^{(2)}_{i,j} := \begin{cases} \frac{b^{(2)}_{i,1}}{h} & \text{if } j = 1, \\ \frac{b^{(2)}_{i} - b^{(2)}_{i,j-1}}{h} & \text{if } 1 < j < N, \\ -\frac{b_{i,N-1}}{h} & \text{if } j = N. \end{cases}$$

Remark 7.3.1 (Compatibility of adjoint K^*). We defined $\mathcal{F}'(q)$ and its adjoint $[\mathcal{F}'(q)]^*$ in a complex vector space, see (2.25) and (2.29), but consider K in real vector space. We show the compatibility of the adjoint K^* before and after identification of $\mathbb{C}$ with $\mathbb{R} \times \mathbb{R}$ via the operator T. Therefore we consider the operator $K_{\mathbb{C}} := (K_{\mathrm{dis},\mathbb{C}}, K_{\mathrm{tv},\mathbb{C}})$ with $K_{\mathrm{dis},\mathbb{C}} = \mathcal{F}'(q)$ and $K_{\mathrm{tv}} = \beta\nabla$. (Note that this operator is $\mathbb{C}$-linear in comparison to the corresponding K, that is only $\mathbb{R}$-linear.) For this remark we will consider in general $K_{\mathbb{C}} = \operatorname{Re} K_{\mathbb{C}} + \mathrm{i} \operatorname{Im} K_{\mathbb{C}}$ and the complex vector $q = \operatorname{Re} q + \mathrm{i} \operatorname{Im} q$:

$$\begin{aligned} K_{\mathbb{C}}\, q &= (\operatorname{Re} K_{\mathbb{C}} + \mathrm{i} \operatorname{Im} K_{\mathbb{C}})(\operatorname{Re} q + \mathrm{i} \operatorname{Im} q) \\ &= \operatorname{Re} K_{\mathbb{C}} \operatorname{Re} q + \mathrm{i} \operatorname{Im} K_{\mathbb{C}} \operatorname{Re} q + \mathrm{i} \operatorname{Re} K_{\mathbb{C}} \operatorname{Im} q - \operatorname{Im} K_{\mathbb{C}} \operatorname{Im} q, \\ \mathrm{T}[K_{\mathbb{C}}\, q] &= \underbrace{\begin{pmatrix} \operatorname{Re} K_{\mathbb{C}} & -\operatorname{Im} K_{\mathbb{C}} \\ \operatorname{Im} K_{\mathbb{C}} & \operatorname{Re} K_{\mathbb{C}} \end{pmatrix}}_{=:K_{\mathbb{R}}} \begin{pmatrix} \operatorname{Re} q \\ \operatorname{Im} q \end{pmatrix} = K_{\mathbb{R}}[\mathrm{T}q]. \end{aligned}$$

[◇] First, note that the definition in [BKL17, Sec. 4.5] was corrected in the cases $1 < i, j < N$ by using a subtraction instead of an addition. Second, note that the discretization of the gradient in one dimension is $(\nabla a)_i = (a_{i+1} - a_i)/h$ if $1 \le i < N$ and $(\nabla a)_i = 0$ if $i = N$, cf. the 2D case (3.8). A suspicion for the corresponding adjoint $\nabla^* = -\operatorname{div}$ results in $(\operatorname{div} b)_i = (b_{i+1} - b_i)/h$ if $1 \le i < N$ and $(\operatorname{div} b)_i = 0$ if $i = N$. However, a rigorous computation in the discretized form shows that a shift is required, i.e. $(\operatorname{div} b)_i = b_1$ if $i = 1$, $(\operatorname{div} b)_i = (b_i - b_{i-1})/h$ if $1 < i < N$ and $(\operatorname{div} b)_i = 0$ if $i = N$. The two-dimensional case is analogous. To prove the claim in 1D we consider the discrete gradient operator in 1D, cf. (3.8), $\nabla a = ([a_2 - a_1]/h, \ldots, [a_N - a_{N-1}]/h, 0)^T$. Then $\langle \nabla a, b\rangle = \sum_{i=1}^{N} (\nabla a)_i b_i = \frac{1}{h}\sum_{i=1}^{N-1}(a_{i+1} - a_i) b_i = \frac{1}{h}[(a_2 - a_1)b_1 + (a_3 - a_2)b_2 + \ldots + (a_{N-1} - a_{N-2})b_{N-2} + (a_N - a_{N-1})b_{N-1}]$. This equals $\frac{1}{h}[a_1(-b_1) + a_2(b_1 - b_2) + \ldots + a_{N-1}(b_{N-2} - b_{N-1}) + a_N b_{N-1}]$, which is the same as $\sum_{i=1}^{N} a_i(\nabla^* b)_i$ with $(\nabla^* b)_i = -(\operatorname{div} b)_i$, where $(\operatorname{div} b)_i$ is defined as above including the shift. Finally, we have a discretized form of $\langle \nabla a, b\rangle = \langle a, -\operatorname{div} b\rangle$.

Finally, we show for a complex vector w that

$$\begin{aligned}\langle K_{\mathbb{C}}\, q, w\rangle_{\mathbb{C}} &= \sum_i (K_{\mathbb{C}}\, q)_i\, \overline{w}_i = \sum_i \sum_j K_{\mathbb{C}_{i,j}}\, q_j\, \overline{w}_i \\ &= \sum_j \sum_i q_j\, K_{\mathbb{C}_{i,j}}\, \overline{w}_i = \sum_j \sum_i q_j\, \overline{K^*_{\mathbb{C}_{j,i}}}\, \overline{w}_i = \sum_j q_j\, (\overline{K^*_{\mathbb{C}}\, w})_j = \langle q, K^*_{\mathbb{C}}\, w\rangle_{\mathbb{C}}.\end{aligned}$$

Note the use of $\overline{K^*_{\mathbb{C}_{j,i}}} = \overline{\overline{K_{\mathbb{C}_{i,j}}}} = K_{\mathbb{C}_{i,j}}$ in the last equation.

Functionals F and G Remember (7.14), in the next step we are ready to identify the functionals F and G of the primal problem (7.2) as

$$F(Kh) = f_{\text{dis}}(h) + f_{\text{tv}}(h), \quad G(h) = f_{\text{spa}}(h) + f_{\text{phy}}(h). \tag{7.23}$$

Consequently, F splits into two terms

$$\begin{aligned}&F(y_{\text{dis}}, y_{\text{tv}}) := F_{\text{dis}}(y_{\text{dis}}) + F_{\text{tv}}(y_{\text{tv}}) \\ &\quad \text{with} \quad F_{\text{dis}}(y) := \tfrac{1}{2}\|y + v_{\text{dis}}\|^2_{\text{dis},\mathbb{R}} \quad \text{and} \quad F_{\text{tv}}(y) := \|y + v_{\text{tv}}\|_{\text{tv},\mathbb{R}},\end{aligned} \tag{7.24}$$

where $y_{\text{dis}} \in Y_{\text{dis},\mathbb{R}}$, $y_{\text{tv}} \in Y_{\text{tv},\mathbb{R}}$ and $v_{\text{tv}}, v_{\text{dis}}$ are as in (7.13). Note that this assignment of F and G is analogous to (7.1).

The above identification of K, F, G is not the only one possible. However, in particular, it turned out to be numerically favorable in our experiments. To implement the primal-dual algorithm, see Alg. 7.2.3, we need the explicit form of the proximal mappings $(I + \sigma\partial F^*)^{-1}$ and $(I + \tau\partial G)^{-1}$ for the proposed functionals F and G. The next two sections are devoted to that subject. For the latter proximal mapping we essentially can refer back to the extended soft-shrinkage operator (6.7).

7.4 Proximal Mappings

7.4.1 The Proximal Mapping of F^*

In this section we determine the form of the *proximal mapping of* F^* necessary for the primal-dual algorithm, see Alg. 7.2.3. First, we remark that both the Fenchel conjugate F^* and the subdifferential ∂F^* are well-defined, since we consider finite-dimensional real vector spaces. Next, we determine the explicit form of the Fenchel conjugate. The form $F(y_{\text{dis}}, y_{\text{tv}}) = F_{\text{dis}}(y_{\text{dis}}) + F_{\text{tv}}(y_{\text{tv}})$ of the functional F, see (7.24), implies that $F^*(y_{\text{dis}}, y_{\text{tv}}) = F^*_{\text{dis}}(y_{\text{dis}}) + F^*_{\text{tv}}(y_{\text{tv}})$ where still $y_{\text{dis}} \in Y_{\text{dis},\mathbb{R}}$ and $y_{\text{tv}} \in Y_{\text{tv},\mathbb{R}}$. This results from the following lemma.

Lemma 7.4.1 (Sum of Fenchel Conjugates, [RW98, Prop. 11.22]). *Let Y_1 and Y_2 be finite-dimensional real Hilbert spaces, with $Y_1^* = Y_1$, $Y_2^* = Y_2$, further, $Y := Y_1 \times Y_2$ be equipped with the inner product $\langle x, y\rangle_Y = \langle x_1, y_1\rangle_{Y_1} + \langle x_2, y_2\rangle_{Y_2}$ where $x = (x_1, x_2) \in Y$, $y = (y_1, y_2) \in Y$ and, finally, F_1, F_2 be convex functionals. If $F(y) = F(y_1, y_2) := F_1(y_1) + F_2(y_2)$ for all $y \in Y$, then $F^*(y) = F^*(y_1, y_2) = F_1^*(y_1) + F_2^*(y_2)$ for all $y \in Y$.*

Proof. Recalling the definition of the Fenchel conjugate and

$$\begin{aligned} F^*(y_1,y_2) &= \sup\{\langle (y_1,y_2),(x_1,x_2)\rangle_{Y_1\times Y_2} - F(y_1,y_2) : (x_1,x_2)\in Y_1\times Y_2\} \\ &= \sup\{\langle y_1,x_1\rangle_{Y_1} + \langle y_2,x_2\rangle_{Y_2} - F_1(x_1) - F_2(x_2) : x_1\in Y_1, x_2\in Y_2\} \\ &= \sup\{\langle y_1,x_1\rangle_{Y_1} - F_1(x_1) : x_1\in Y_1] + \sup[\langle y_2,x_2\rangle_{Y_2} - F_2(x_2) : x_2\in Y_2\} \\ &= F_1^*(y_1) + F_2^*(y_2) \end{aligned}$$

proves the claim. ☐

Remark 7.4.2. This element-wise split implies that the proximal mapping splits too,

$$(I+\sigma\partial F^*)^{-1}(y_1,y_2) = ((I+\partial F_1^*)^{-1}(y_1),(I+\partial F_2^*)^{-1}(y_2)).$$

Proximal Mapping of F^* Therefore the decomposition of F^* together with the definition of Y implies that the *proximal mapping of F^** is

$$(I+\sigma\partial F^*)^{-1}(y_{\text{dis}},y_{\text{tv}}) = ((I+\partial F_{\text{dis}}^*)^{-1}(y_{\text{dis}}),(I+\partial F_{\text{tv}}^*)^{-1}(y_{\text{tv}})).$$

Consequently, we need to determine the proximal mappings of F_{dis}^* and F_{tv}^*, see (7.26) and (7.28) for the results.[◇]

Further, one recalls that for shift function $f(y) = g(y+b)$ one has $f^*(y) = g^*(y) - \langle b, y\rangle$, see [BL06, Tab. 3.2]. Therefore[†]

$$f(y) = g(y+b) \quad \text{implies} \quad (I+\sigma\partial f^*)^{-1}(y) = (I+\sigma\partial g^*)^{-1}(y+\sigma b)\,. \tag{7.25}$$

Proximal Mapping of F_{dis}^* To compute the *proximal mapping of F_{dis}^**, see (7.24) for F_{dis}, we have $(I+\sigma\partial F_{\text{dis}}^*)^{-1}(y) = (I+\sigma\partial(\frac{1}{2}\|\cdot\|_{\text{dis},\mathbb{R}}^2)^*)^{-1}(y+\sigma v_{\text{dis}})$. For Hilbert norms the functional $\frac{1}{2}\|\cdot\|^2$ has as the Fenchel conjugate the same term but with dual norm. Therefore the derivative of the Fenchel conjugate is the identity in the dual space. In our case this means that $(\frac{1}{2}\|\cdot\|_{\text{dis},\mathbb{R}}^2)^* = \frac{1}{2}\|\cdot\|_{\text{dis},\mathbb{R}}^2$ and therefore $\partial(\frac{1}{2}\|\cdot\|_{\text{dis},\mathbb{R}}^2)^* = I$, since $Y_{\text{dis},\mathbb{R}}$ was identified with its dual. Therefore the proximal mapping is[‡]

$$(I+\sigma\partial F_{\text{dis}}^*)^{-1}(y) = \frac{y+\sigma v_{\text{dis}}}{1+\sigma}, \tag{7.26}$$

with v_{dis} as in (7.13).

Proximal Mapping of F_{tv}^* For the *proximal mapping of F_{tv}^**, see (7.24) for F_{tv}, we can again invoke (7.25) which gives[§]

$$(I+\sigma\partial F_{\text{tv}}^*)^{-1}(y) = (I+\sigma\partial(\|\cdot\|_{\text{tv},\mathbb{R}})^*)^{-1}(y+\sigma v_{\text{tv}}).$$

[◇] In [BKL17, Sec. 4.6] some typos state the proximal mappings of F_{dis} and F_{tv} instead of F_{dis}^* and F_{tv}^*.

[†] For the derivation we consider $x = (I+\sigma\partial f^*)^{-1}(y)$ and invert it, i.e. $(I+\sigma\partial f^*)(x) = y$. As $\sigma\partial f^*(x) = \sigma\partial g^*(x) - \sigma b$ this results in $x + \sigma\partial g^*(x) - \sigma b = y$. The proximal mapping $(I+\sigma\partial g^*)^{-1}$ applied to $(I+\sigma\partial g^*)(x) = y + \sigma b$ finally results in $x = (I+\sigma\partial g^*)^{-1}(y+\sigma b)$.

[‡] Of course, inserting of $\partial(\frac{1}{2}\|\cdot\|_{\text{dis},\mathbb{R}}^2)^* = I$ in $(I+\sigma\partial F_{\text{dis}}^*)^{-1}(y)$ results in $(I+\sigma I)^{-1}(y+\sigma v_{\text{dis}}) = (y+\sigma v_{\text{dis}})/(1+\sigma)$.

[§] In [BKL17, Sec. 4.6] a typo states v_{dis} instead of v_{tv}.

Next, we remember $y \in Y_{\mathrm{tv},\mathbb{R}}$, the notation in (7.17), and define tuples $y_i \in Y_i$ via $Y_{\mathrm{tv},\mathbb{R}} = Y_1 \times \ldots \times Y_{N_\mathrm{D}}$ and

$$y_i := (y_i^{(1),\mathrm{Re}}, y_i^{(1),\mathrm{Im}}, y_i^{(2),\mathrm{Re}}, y_i^{(2),\mathrm{Im}}), \quad \text{(2D case)}, \tag{7.27}$$

with the straightforward modification for the 3D case. Then, the functional $\|\cdot\|_{\mathrm{tv},\mathbb{R}}$, see (7.17), splits into a sum of terms of the form $h_N^d |y_i|$ with $y_i \in Y_i$ and norm $|\cdot|$ as defined in (7.18). Consequently, Remark 7.4.2 applies in extension and implies that the proximal mapping of $(\|\cdot\|_{\mathrm{tv},\mathbb{R}})^*$ can be discussed for the tuples separately.

Remembering that Y is equipped with an inner product, see Sec. 7.2, we have to look for the Fenchel conjugate to the induced norm by the inner product corresponding to Y_i, i.e. using (7.20) and (7.15),

$$\langle x_i, y_i \rangle_{Y_i} = h_N^d \langle x_i, y_i \rangle_2.$$

The induced norm is $\|x\|_{Y_i} = (h_N^d)^{1/2}|x|$, where x means a tuple x_i.

This means, we look for the Fenchel conjugate of $h_N^d \|x\|_2 = (h_N^d)^{1/2}(h_N^d)^{1/2}|x| = (h_N^d)^{1/2}\|x\|_{Y_i} = a\,\|\cdot\|$ with $a = (h_N^d)^{1/2}$ and $\|\cdot\| = \|\cdot\|_{Y_i}$ corresponding to the inner product $\langle \cdot, \cdot \rangle_{Y_i}$.[◇]

To this end, we will see that the proximal mapping[†] of $(a\,\|\cdot\|)^*$ is the orthogonal projection onto the ball with radius a onto the pre-dual norm of $\|\cdot\|$, if the dual pairing between the pre-dual and the norm itself is the same as the inner product used for the orthogonal projection and the inner product used in the definition of the Fenchel conjugate (3.14). To see this, $\|\cdot\|_*$ denotes the pre-dual norm. Then $a\,\|x\| = \sup_{\|x^*\|_* \le 1} \langle a x^*, x \rangle = \sup_{\|x^*\|_* \le a} \langle x^*, x \rangle = \sup\{\langle x^*, x \rangle - \delta_C(x^*)\}$, where $C = \{x^* : \|x^*\|_* \le a\}$.[‡] A close inspection of (3.14) reveals that δ_C is the Fenchel conjugate of $(a\,\|\cdot\|)$. Remark 3.3.4 then yields the claim for the proximal mapping of $(a\,\|\cdot\|)^*$. Of course, this projection can be explicitly expressed by $x/\max\{1, \|x\|_*/a\}$, respectively componentwise

$$x_i / \max\{1, \|x\|_*/a\}.$$

In our case the inner product is as mentioned $\langle x_i, y_i \rangle_{Y_i} = h_N^d \langle x_i, y_i \rangle_2$, see above. Clearly, since the norm in the term $(a\|\cdot\|)$ is exactly the norm induced by the inner product, its pre-dual is the norm itself, i.e.

$$\|x\|_*/a = \|x\|_{Y_i}/(h_N^d)^{1/2} = (h_N^d)^{1/2}|x|/(h_N^d)^{1/2} = |x|.$$

Together with the previous considerations this yields for the proximal mapping of F_{tv}^* componentwise[§]

$$((I + \sigma F_{\mathrm{tv}}^*)^{-1}(y))_i^{(k),\ell} = \frac{y_i^{(k),\ell} + \sigma (v_{\mathrm{tv}})_i^{(k),\ell}}{\max\{1, |y_i + \sigma (v_{\mathrm{tv}})_i|\}}, \tag{7.28}$$

[◇] The notation in [BKL17] was slightly adapted.

[†] In [BKL17, Sec. 4.6] a typo states *proximal mapping* $(a\,\|\cdot\|)$ instead of *proximal mapping of* $(a\,\|\cdot\|)^*$.

[‡] The norms in [BKL17] were corrected. In one norm was missing that all norms are pre-dual norms.

[§] Remembering $(I + \sigma\partial F_{\mathrm{tv}}^*)^{-1}(y) = (I + \sigma\partial(\|\cdot\|_{\mathrm{tv},\mathbb{R}})^*)^{-1}(y + \sigma v_{\mathrm{tv}})$ and $x_i/\max\{1, |x|\}$ as the result of the proximal mapping of $(\|\cdot\|_{\mathrm{tv},\mathbb{R}})^*$ for one tuple x indeed yields the result.

where the norm $|\cdot|$ is as defined in (7.18) and v_{tv} as in (7.13), the indexing is in parallel to (7.27), $k = 1, \dots, d$, and ℓ is Re or Im.

7.4.2 The Proximal Mapping of G

After having computed the proximal mapping for F^* we now direct our attention to the *proximal mapping of* G, which is needed for the second step of the primal-dual algorithm, see Alg. 7.2.3. To this end, one observes that G separates in an element-wise sum with terms of the form $\alpha h_N^d |x^{\mathrm{Re}} + v^{\mathrm{Re}}| + \delta_{[a,b]}(x^{\mathrm{Re}} + v^{\mathrm{Re}})$ and $\alpha h_N^d |x^{\mathrm{Im}} + v^{\mathrm{Im}}| + \delta_{[c,d]}(x^{\mathrm{Im}} + v^{\mathrm{Im}})$, where $x^{\mathrm{Re}}, x^{\mathrm{Im}}, v^{\mathrm{Re}}, v^{\mathrm{Im}}$ are real-valued, see (7.23) and (7.16). By the same argumentation that led to Lemma 7.4.1 and Remark 7.4.2 the proximal mapping of G splits into the element-wise proximal mappings of the terms of the previous two forms. Since both term kinds are structurally similar we consider the first one only.

We observe, complementary to (7.25), if $f(x) = g(x+v)$ then[◇] $(I + \tau\partial f)^{-1}(x) = (I + \tau\partial g)^{-1}(x+v) - v$. Therefore we consider the proximal mapping of $g(x) := \alpha h_N^d |x| + \delta_{[a,b]}(x)$ as in Sec. 6.2, i.e. $(I + \tau\partial g)^{-1}(x) = \mathcal{I}_{[a,b]}(\mathbb{S}(x, \tau\alpha h_N^d))$, see (6.7) and cf. Fig. 6.2(f)–(j).

Accommodating for the just discussed shift $+v$ in the original terms, it results in the following formula for the proximal mapping of G:

$$\begin{aligned} ((I + \tau\partial G)^{-1}(x))^{\mathrm{Re}} &= \mathcal{I}_{[a,b]}(\mathbb{S}(x^{\mathrm{Re}} + v^{\mathrm{Re}}, \tau\alpha h_N^d)) - v^{\mathrm{Re}}, \\ ((I + \tau\partial G)^{-1}(x))^{\mathrm{Im}} &= \mathcal{I}_{[c,d]}(\mathbb{S}(x^{\mathrm{Im}} + v^{\mathrm{Im}}, \tau\alpha h_N^d)) - v^{\mathrm{Im}}. \end{aligned} \tag{7.29}$$

7.5 Proposed Variational Reconstruction Scheme

Now we have all ingredients to describe the *proposed variational reconstruction scheme*, that uses various penalty terms (sparsity, total variation and physical bounds).

An idea of this scheme was already given in Sec. 7.1. Again, the proposed scheme then iteratively interleaves the linearization and variational step. The idea is to improve the current iterate $q^{(m)} \in X$ by a step $h^{(m)} \in X$ that minimizes the functional (7.1) for $q = q^{(m)}$ (*inner iteration*) and update $q^{(m+1)} := q^{(m)} + h^{(m)}$ (*outer iteration*). This successively computes the linearization of $\mathcal{F}$ at $q^{(m+1)}$. Since the primal-dual algorithm (PDA) of (7.1) is formulated for real vector spaces, we transform the linearized problem into a minimization problem on a real vector space, i.e. we seek for q in $X_{\mathbb{R}} := \mathbb{R}^{N_{\mathrm{D}}} \times \mathbb{R}^{N_{\mathrm{D}}}$ instead of q in $X := \mathbb{C}^{N_{\mathrm{D}}}$, as is discussed in Sec. 7.3. In the following we will look at the complete inversion scheme as pseudo code with all technical details.

Algorithm 7.5.1 (Proposed variational reconstruction scheme).

1. Initialization of the outer iteration:
   set outer iteration number $m = 0$ and initial contrast $q^{(0)} = 0 \in X_{\mathbb{R}} := \mathbb{R}^{N_{\mathrm{D}}} \times \mathbb{R}^{N_{\mathrm{D}}}$.
2. *Inner iteration* solving the linearized minimization problem (7.1) employing primal-dual algorithm (PDA), see Alg. 7.2.3:
   a) Initialization of the inner iteration: Set inner iteration number $n = 0$. Remember $\mathcal{F}(q^{(m)})$ in (2.24). Compute $A_{N_{\mathrm{D}},N_{\mathrm{s}}}$ and $B_{N_{\mathrm{D}},N_{\mathrm{i}}}$ for fast computation of

---

[◇] For the derivation we consider $y = (I + \tau\partial f)^{-1}(x)$ and invert it, i.e. $(I + \tau\partial f)(y) = x$. As $\tau\partial f(x) = \tau\partial g(x+v)$ this results in $y + \tau\partial g(y+v) = x$. The proximal mapping $(I + \tau\partial g)^{-1}$ applied to $(I + \tau\partial g)(y+v) = x + v$ finally results in $y = (I + \tau\partial g)^{-1}(x+v) - v$.

$\mathcal{F}'(q^{(m)})[h]$ and $[\mathcal{F}'(q^{(m)})]^*H$, see (2.25) and (2.29). Choose step sizes $\tau, \sigma > 0$ as in (7.34). Set $\bar{x}^{[0]} = x^{[0]} = 0 \in X_{\mathbb{R}}$, $y_{\text{dis}}^{[0]} = 0 \in Y_{\text{dis},\mathbb{R}} = \mathbb{R}^{N_\text{s} \times N_\text{i}} \times \mathbb{R}^{N_\text{s} \times N_\text{i}}$, and $y_{\text{tv}}^{[0]} = 0 \in Y_{\text{tv},\mathbb{R}} = X_{\mathbb{R}} \times X_{\mathbb{R}}$ (in 2D case) as initial vectors. That means, we consider a tuple $((y^{(1),\text{Re}}, y^{(1),\text{Im}}), (y^{(2),\text{Re}}, y^{(2),\text{Im}})) \in Y_{\text{tv},\mathbb{R}}$. For the definitions of v_{dis}, v_{tv}, K_{dis}, K_{tv} see (7.13) in combination with (7.11) and for their adjoints K_{dis}^* and K_{tv}^* see (7.22).

b) *Dual step* $(\bar{x}^{[n]}, y_{\text{dis}}^{[n]}, y_{\text{tv}}^{[n]}) \rightarrow (\bar{x}^{[n]}, y_{\text{dis}}^{[n+1]}, y_{\text{tv}}^{[n+1]})$ with *dual step size* $\sigma > 0$:

$$\begin{aligned} w_{\text{y,dis}} &:= y_{\text{dis}}^{[n]} + \sigma K_{\text{dis}}(\bar{x}^{[n]}), \\ y_{\text{dis}}^{[n+1]} &= (I + \sigma F_{\text{dis}}^*)^{-1}(w_{\text{y,dis}}) = \frac{w_{\text{y,dis}} + \sigma v_{\text{dis}}}{1+\sigma}, \qquad (7.30) \\ w_{\text{y,tv}} &:= y_{\text{tv}}^{[n]} + \sigma K_{\text{tv}}(\bar{x}^{[n]}), \\ z &:= w_{\text{y,tv}} + \sigma v_{\text{tv}}, \\ (y_{\text{tv}}^{[n+1]})_i^{(k),\ell} &= ((I + \sigma F_{\text{tv}}^*)^{-1}(w_{\text{y,tv}}))_i^{(k),\ell} = \frac{z_i^{(k),\ell}}{\max\{1, |z_i|\}} \qquad (7.31) \end{aligned}$$

with $k = 1, \dots, d$, ℓ is Re or Im, and $|z_i|$ as defined in (7.18).

c) *Primal step* $(x^{[n]}, y_{\text{dis}}^{[n+1]}, y_{\text{tv}}^{[n+1]}) \rightarrow (x^{[n+1]}, y_{\text{dis}}^{[n+1]}, y_{\text{tv}}^{[n+1]})$ with *primal step size* $\tau > 0$:

$$w_{\text{x}} = x^{[n]} - \tau(K_{\text{dis}}^*(y_{\text{dis}}^{[n+1]}) + K_{\text{tv}}^*(y_{\text{tv}}^{[n+1]})), \qquad (7.32)$$

$$\kappa := \tau\alpha\, h_N^d, \quad N_{\mathbb{R}} := \{1, \dots, N_{\text{D}}\}, \quad N_{\mathbb{C}} := \{N_{\text{D}} + 1, \dots, 2N_{\text{D}}\},$$

$$x_j^{[n+1]} = (I + \tau G)^{-1}(w_{\text{x},j}) = -q_j + \begin{cases} \mathcal{I}_{[a,b]}\mathbb{S}(w_{\text{x},j} + q_j, \kappa) & \text{if } j \in N_{\mathbb{R}}, \\ \mathcal{I}_{[c,d]}\mathbb{S}(w_{\text{x},j} + q_j, \kappa) & \text{if } j \in N_{\mathbb{C}}. \end{cases} \qquad (7.33)$$

Remember the implementation of the extended soft-shrinkage operator[◇] denoted by $\mathcal{I}_{[a,b]}\mathbb{S}(z, \kappa)$ via $\max\{a, \min\{b, \operatorname{sign}(z)\, \max\{|z| - \kappa, 0\}\}\}$, see (6.7).

d) *Extrapolation step*: $\bar{x}^{[n+1]} = 2x^{[n+1]} - x^{[n]}$.

e) Stop the inner iteration by tolerance-based rules, which will be presented below as outer or inner tolerance principle, and set

$$h^{(m)} := x^{[n+1]}.$$

Otherwise set $n := n + 1$ and go to step 2b again.

3. *Outer iteration step*:
   a) $q^{(m+1)} = q^{(m)} + h^{(m)}$ and $m := m + 1$.
   b) Stop the outer iteration by the discrepancy principle, see (3.18). Otherwise go to step 2 again.

Parameter Choices and a Note to Convergence Theory Of course, the proposed reconstruction scheme depends on various parameters as, e.g. regularization and stopping parameters. All convergence theory for noise level tending to zero of such an iterative,

---

[◇] In [BKL17, Sec. 4.4] a typo states $\mathbb{S}_\kappa(z)$ instead of $\mathbb{S}(z, \kappa)$.

Newton-like algorithm seems to require nonlinearity conditions that are not yet verified. There is hence no parameter choice rule that theoretically guarantees convergence. Practically, suitable parameters, that were determined in exhaustive numerical experiments, are suggested. Further, in the case of regularization parameters an automatic, data-driven parameter choice is available, see Alg. 3.3.6.

Norm Estimation and Choice of Step Sizes The inner iteration of the reconstruction scheme proposed in Alg. 7.5.1, i.e. the primal-dual algorithm, see Alg. 7.2.3, converges if the *step sizes* τ and σ are chosen such that $\tau\sigma\|K\|^2 < 1$, see Th. 7.2.4. The norm of K is given by $\|K\| = \max_{0\neq x\in X} \frac{\|Kx\|_Y}{\|x\|_{\mathrm{roi},2}}$, where the Hilbert norm on Y is as induced by the inner product (7.21), see (7.13), and $\|\cdot\|_{\mathrm{roi},2}$ is induced by (3.4). The notation $\|\cdot\|_{\cdot,2}$ was chosen to emphasize that the norm is induced by an inner product. In particular, $\|Kx\|_Y = \sqrt{\|Kx\|^2_{\mathrm{dis},2} + \|Kx\|^2_{\mathrm{tv},2}}$, where $\|\cdot\|_{\mathrm{dis},2}$ is induced by (7.19) and $\|\cdot\|_{\mathrm{tv},2}$ is induced by (7.20). (As already mentioned, the latter one is not the same as $\|\cdot\|_{\mathrm{tv},\mathbb{R}}$, defined in (7.17).) It is possible to estimate the norm of K by breaking it down to the norm of the involved operators. However, this rough estimate turned out to be too inefficient for our purposes. Of course, one can proceed to reformulate the problem in spaces with standard inner products, which then enables to employ power iteration for the estimation of the biggest eigenvalue. For our purposes it sufficed to test the operator with a sample of vectors of two kinds, i.e. random vectors and vectors of the form $(1, \dots, 1, 0, \dots, 0)$. (The use of all unit basis vectors was omitted because they deliver a much too low estimate.) This coarse estimate was then multiplied by a safety factor, in general equal to 2. In the numerical experiments it was observed that Alg. 7.5.1 diverged sometimes evidently[◇] if $\|K\|^2$ was estimated too low.

In the numerical experiments the choice of step sizes

$$\sigma = \tau \quad \text{both slightly smaller than} \quad \|K\|^{-1} \tag{7.34}$$

yielded the best results. As a less severe condition than $\tau\sigma\|K\|^2 < 1$ is sufficient to guarantee convergence of Alg. 7.2.3, an *adaptive step size choice* proposed in [BH15, Sec. 2.2] is possible. Therefore new parameters $\vartheta \in (0, 1)$ as well as $\eta > 0$ are defined,

$$x_{\mathrm{d}} := x^{(n)} - x^{(n-1)}, \quad \chi := \frac{\|x_{\mathrm{d}}\|_{\mathrm{roi},2}}{\|Kx_{\mathrm{d}}\|_Y}, \quad \mathbb{P}(\chi) := \begin{cases} \chi & \text{if } \chi \leq \vartheta\sigma\tau, \\ \sqrt{\vartheta\sigma\tau} & \text{if } \vartheta\sigma\tau < \chi \leq \sigma\tau, \\ \sqrt{\sigma\tau} & \text{if } \sigma\tau < \chi, \end{cases}$$

are set and *adaptive step sizes* are defined by $\sigma = \mathbb{P}(\chi)\eta$ and $\tau = \mathbb{P}(\chi)/\eta$. Note that η balances the primal and dual step sizes, τ respectively σ; e.g. the choice $\eta = 1$ follows [BH15, Sec. 2.2].

This choice was tested in numerical examples. However, the requirement for activation was not satisfied.

[◇] In fact, the discrepancy as well as the whole Tikhonov functional (3.11) change significantly after each inner iteration. In particular, after a few inner iterations it essentially alternates between two values.

Stopping Strategies Recall that the proposed variational reconstruction scheme, see Alg. 7.5.1, consists of outer iterations, responsible for re-linearization of the forward problem, as well as inner iterations, responsible for update/minimization step within a linearized problem. In general, for the strict analysis of this approach one needs to assume that an infinite number of re-linearizations (outer iteration) and an infinite number of minimization steps (inner iteration) within each linearized minimization problem has been carried out. Clearly, this approach is computationally not feasible. To provide an efficient reconstruction scheme the number of outer and inner iteration steps has to be restricted. A fixed number is only one possibility. To that end, we discuss appropriate *stopping strategies* for the outer and the inner iteration.

Stopping the Outer Iteration Following the best practice for inverse problems it is recommended to stop the outer iteration by Morozov's discrepancy principle, see (3.18) with $N = N_{\text{out}}$. In the numerical examples, satisfactory results were achieved for the tolerance choice $\tau_{\text{dis}} = 2.5$.

Stopping the Inner Iteration In analogy to the previous paragraph, let the index of the inner iterations with m-th outer iteration be given by $n = 1, \ldots, N_{\text{in}}^{(m)}$.

A straightforward stopping for the inner iteration is given after a fixed number of inner iterations. In the numerical experiments the choice $N_{\text{in}}^{(m)} = 50$ for all outer iterations turned out to be sufficient to robustly and reasonably update the linearized problem. However, two more sophisticated, heuristic strategies are presented in the following. Both ensure that only a few inner iterations are executed inside the first outer iterations and allow a high number in later iterations. A comparison to the constant iteration number is also given in Fig. 7.2 in Sec. 7.6.

For the two strategies we define the *linearized relative discrepancy*

$$d_\ell(q, h) := \|\mathcal{F}'(q)[h] + \mathcal{F}(q) - F_{\text{meas}}^{\delta}\|_{\text{dis}} / \|F_{\text{meas}}^{\delta}\|_{\text{dis}}$$

and the *non-linearized relative discrepancy*, in accordance with (3.17),

$$d(q + h) := \|\mathcal{F}(q + h) - F_{\text{meas}}^{\delta}\|_{\text{dis}} / \|F_{\text{meas}}^{\delta}\|_{\text{dis}}.$$

Stopping Strategy 1 for the Inner Iteration The basic idea of the first strategy is to compare the quantities $d_\ell(q, h)$ and $d(q + h)$ each time the inner iteration terminates. If they are similar, the number of inner iterations is increased, otherwise it is decreased.

So, it is assumed that an a priori choice for the number of inner iterations in the m-th outer iteration step was made. We denote this choice by $N_{\text{in}}^{(m)}$. This choice is a posteriori evaluated, i.e. *after* the computation of the inner iterations. We consider the choice "good" if the difference between the linearized and original problem gives approximately the same difference. To quantify this we consider the linearized discrepancy $\text{dis}_{\text{lin}} := d_\ell(q^{(m)}, h^{(m)})$ and the non-linearized discrepancy $\text{dis}_{\text{nonlin}} := d(q^{(m)} + h^{(m)})$, where $q^{(m)}, h^{(m)}$ are the results of the inner iteration. To keep the notation simple, the iteration index for both discrepancies is omitted. Then, the number of inner iteration steps is considered "good" if the quotient

$$\mathrm{dis}_{\mathrm{rel}} := \mathrm{dis}_{\mathrm{lin}}/\mathrm{dis}_{\mathrm{nonlin}}$$

is approximately 1. Only in that case a higher number of inner iterations "pays off" with respect to the reconstruction quality. Consequently, the number of inner iterations is then updated via

$$N_{\mathrm{in}}^{(m+1)} = \begin{cases} \lceil \mu_{\uparrow} N_{\mathrm{in}}^{(m)} \rceil & \text{if } \mathrm{dis}_{\mathrm{rel}} \in (1-\tau_{\mathrm{out}}, 1+\tau_{\mathrm{out}}), \\ \lfloor \mu_{\downarrow} N_{\mathrm{in}}^{(m)} \rfloor & \text{otherwise} \end{cases}$$

with $\mu_{\uparrow} = 1 + \sqrt{1/m}\,\ln(m)$ and $\mu_{\downarrow} = (\min\{1/\mathrm{dis}_{\mathrm{rel}}, \mathrm{dis}_{\mathrm{rel}}\})^2$. In the experiments the outer tolerance τ_{out} was set to 0.05. Since the above update tends to be too optimistic in the first outer steps, a conservative start choice of $N_{\mathrm{in}}^{(1)} = 1$ is recommended. The number of inner iterations is additionally limited to 250 for all outer iteration steps.

Stopping Strategy 2 for the Inner Iteration The basic idea of the second stopping strategy is to compare the quantities $d_{\ell}(q,h)$ and $d(q)$ *inside* the inner iteration. If they are similar, the inner iteration is terminated. Note that it would be interesting to compare $d_{\ell}(q,h)$ and $d(q+h)$ directly, however, this involves an additional evaluation of the forward operator $\mathcal{F}$, which was deemed too expensive for the purpose of a stopping strategy.

So, the second strategy controls the number of necessary inner iteration steps from within the inner iteration. The idea follows the inexact stopping rule for a Newton-like method, cf. [Rie03, Sec. 7.5.3]. Similar to the previous strategy, the idea is to compute the quotient $\mathrm{dis}_{\mathrm{lin}}/\mathrm{dis}_{\mathrm{nonlin}}$ after each inner iteration. To keep the computational effort moderate, however, the non-linearized discrepancy is fixed within the inner iteration, i.e. $\mathrm{dis}_{\mathrm{nonlin}} := d(q^{(m)} + h^{(m)})$, where $q^{(m)}, h^{(m)}$ are again the results of the last outer iteration step. The linearized discrepancy is then reevaluated after every inner iteration step, i.e. $\mathrm{dis}_{\mathrm{lin}}^{(n)} := d_{\ell}(q^{(m)}, h^{(n)})$, where $h^{(n)}$ is the current iterate within the inner iteration. The inner iteration is stopped if $\mathrm{dis}_{\mathrm{lin}}^{(n)}/\mathrm{dis}_{\mathrm{nonlin}} < \Theta_m$ for some inner tolerance $\Theta_m \in (0,1]$. The inner tolerances Θ_m are computed via the following algorithm.

Algorithm 7.5.2 (Compute inner tolerances Θ_m, cf. [Rie03, Sec. 7.5.3])**.** Set $\Theta_{\mathrm{start}} \in (0,1)$, $\Theta_{\max} \in (\Theta_{\mathrm{start}}, 1)$, and $\gamma_{\mathrm{tol}} \in (0,1]$. For $m = 1,2$ set auxiliary tolerances $\tilde{\Theta}_m = \Theta_{\mathrm{start}}$ and compute, for $m \geq 3$,

$$\tilde{\Theta}_m = \begin{cases} 1 - (1-\Theta_{m-1})\, N_{\mathrm{in}}^{(m-2)}/N_{\mathrm{in}}^{(m-1)} & \text{if } N_{\mathrm{in}}^{(m-1)} \geq N_{\mathrm{in}}^{(m-2)}, \\ \gamma_{\mathrm{tol}}\, \Theta_{m-1} & \text{otherwise}, \end{cases}$$

where $N_{\mathrm{in}}^{(m-2)}$ and $N_{\mathrm{in}}^{(m-1)}$ are the number of steps in the inner iteration for the two previous outer iteration steps. Finally, set

$$\Theta_m = \Theta_{\max} \max\left\{\tau\, \delta/\mathrm{dis}_{\mathrm{nonlin}}^{(m)}, \tilde{\Theta}_m\right\}, \quad m \in \mathbb{N}.$$

As previously, since inner iterations should contribute to the overall reconstruction,

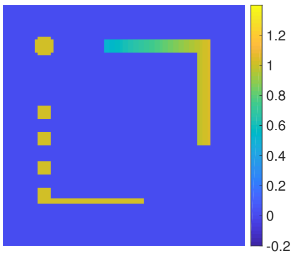


Figure 7.1: Test phantom for 2D synthetic data (corner with a gradual contrast ramp and a ball as well as a broken corner with $q = 1.0$).

the maximal number of inner iterations is capped to 250. Further, in the experiments the parameters were set to $\Theta_{\text{start}} = 0.925$, $\Theta_{\max} = 0.95$ and $\gamma_{\text{tol}} = 0.90$.[◇]

7.6 Numerical Examples

The performance of the proposed variational reconstruction scheme, see Alg. 7.5.1, was already presented in Ch. 5. In this section we focus on the influence of parameter choices and on test series to the suggested stopping strategies. For that we will consider reconstructions from simulated data with noise in two and three dimensions as well as from real-world data in two dimensions from Institute Fresnel, see [BS01].

Regularization Parameters Unless otherwise stated, the *regularization parameters* of the Tikhonov functional (3.11) were set to $\alpha = 500$ and $\beta = 10^{-5}$. These parameters were determined manually. A high regularization parameter β influences the operator norm $\|K\|$ (otherwise the discrepancy dominates) and therefore step sizes τ and σ will decrease, which results in a slow iteration. Therefore the following procedure is recommended: initially, set $\alpha = 0$ and find a suitable β and finally, find a convenient α. Of course, the proposed automatic choice of the regularization parameters, see Alg. 3.3.6, is recommended if a reasonable ratio α/β is known.

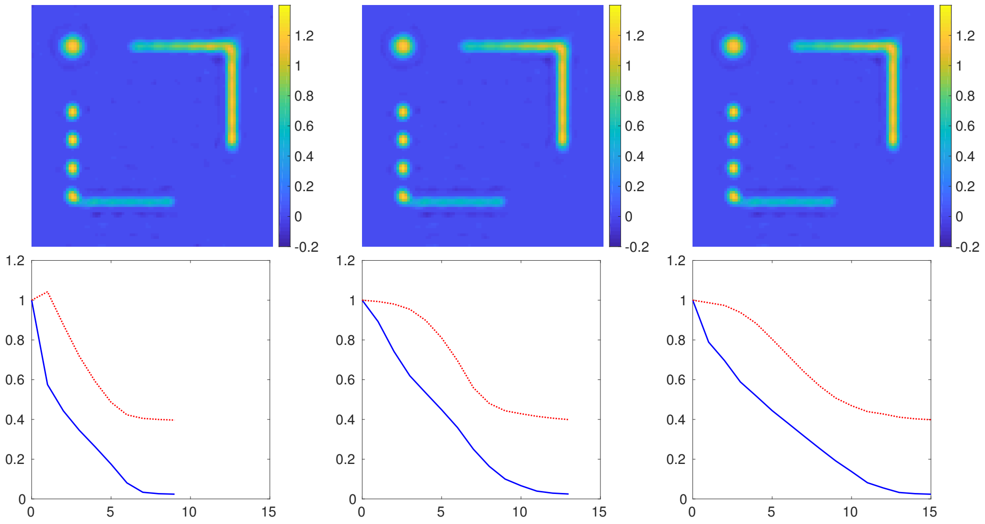
7.6.1 Numerical Examples for Synthetic Data in 2D

Setting and Parameters The wave number was set to $k = 250\,\mathrm{m}^{-1}$. Consequently, we consider fields with wavelength of about $0.025\,\mathrm{m}$. Further, we fix 35 identical source/receiver points for incident point sources distributed evenly on a circle with radius $5\,\mathrm{m}$ centered in the origin. (Near field data is considered for the measurement.) The radius R determining the region of interest D is set to $R = 0.1, \mathrm{m}$ and the computational domain D_{2R} is discretized by $N = 256$ points in each dimension. Synthetic data is generated using Gaussian noise. The relative noise level is set to $\delta = 0.01$. We set physical bounds assuming that the contrast q satisfies $-1 \leq \mathrm{Re}(q) \leq 3$ and $0 \leq \mathrm{Im}(q) \leq 3$.

As the test phantom in 2D the scattering object depicted in Fig. 7.1 is used. The imaginary part of the phantom vanishes. In what follows we consider the influence of

---

[◇] These parameters turned out to be more appropriate than the proposed ones in [Rie03, Sec. 7.5.3]—at least for that application case.



(a) Constant. (b) Stop. strat. 1. (c) Stop. strat. 2.

Figure 7.2: Effect of the stopping strategies for the inner iterations. Upper row: real part of the reconstructed synthetic contrast. Lower row: relative discrepancy (blue solid) and error (red dotted) as functions of the outer iteration number (x-axis). (The outer iteration was stopped by Morozov's discrepancy principle with $\tau_{\text{dis}} = 2.5$.) Run times, relative discrepancies and relative errors were in (a) 2.9 min, dis. 0.024, err. 0.397, (b) 3.5 min, dis. 0.025, err. 0.399, and (c) 4.2 min, dis. 0.024, err. 0.399.

several parameters of the reconstruction algorithm on the computational cost and the reconstruction quality as measured by relative discrepancy and relative error.

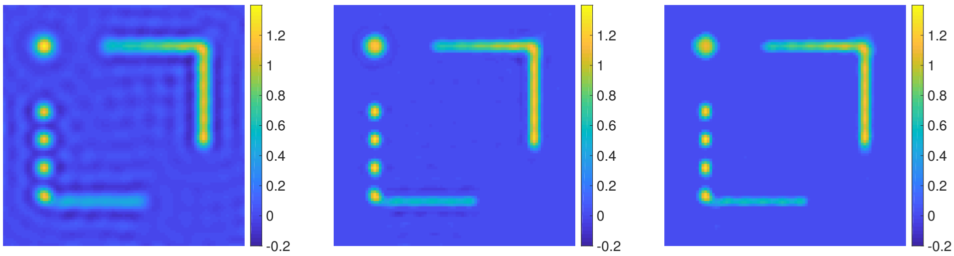
Reconstruction of the Obstacle in the Imaginary Part In the case of wave number◇ $k = 250\,\text{m}^{-1}$ it was observable that the obstacle appears in the imaginary part (i.e. physical absorption) after the first outer iteration. (This behavior decreases for smaller wavelengths.) Then the obstacle appears successively in the real part. Finally, the imaginary part vanishes as expected. However, the same behavior is delivered by the thresholded, nonlinear Landweber scheme, that was discussed in Ch. 6.

Influence of the Inner Iteration Stopping Strategies The influence of the inner iteration stopping strategies is presented in Fig. 7.2. Visually clearly, all three stopping criteria provide similarly satisfactory results. Quantitative description of this similarity is provided by the graphs below.

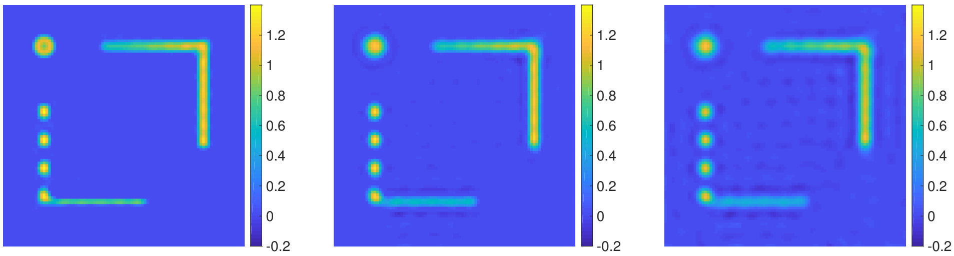
Influence of Parameter Choices α and β The next step is the study of the influence of the sparsity regularization parameter α and the TV regularization parameter β on the computational cost and the reconstruction quality. It is well known that, all other things being equal, the reconstructions will exhibit the typical sparsity or TV artifacts if the related other regularization is switched off by setting the according parameter to zero. Further, numerical experiments show an interesting dependency between the choice of α and β and related "good" choice of the parameter τ_{dis} in Morozov's discrepancy

---

◇In [BKL17, Secs. 5 and 6] the units are often missing: lengths are given in m; wave numbers in m^{-1}.



(a) $\alpha = 0$, $\tau_{\text{dis}} = 1.5$. (b) $\alpha = 500$, $\tau_{\text{dis}} = 2.5$. (c) $\alpha = 2500$, $\tau_{\text{dis}} = 6.0$.



(d) $\beta = 0$, $\tau_{\text{dis}} = 1.5$. (e) $\beta = 1 \cdot 10^{-5}$, $\tau_{\text{dis}} = 2.5$. (f) $\beta = 5 \cdot 10^{-5}$, $\tau_{\text{dis}} = 6.0$.

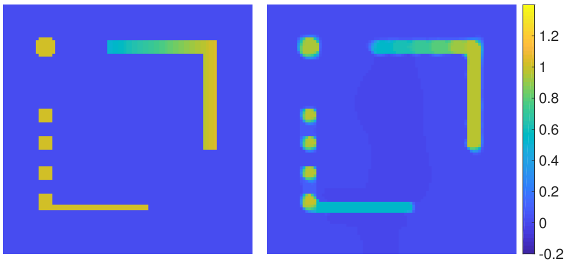
Figure 7.3: Influence of the sparsity regularization parameter α and the TV regularization parameter β. The upper row deals with changes in α, the lower row with changes in β. If not given explicitly the other parameter is given as $\alpha = 500$, $\beta = 10^{-5}$. In all cases Morozov's discrepancy principle was employed for stopping the outer iteration and stopping strategy 2 for the inner iteration. (Of course, the Subfigures (b) and (e) are the same and were duplicated for sake of easier in-row comparison.) Run times, relative discrepancies and relative errors were in (b) and (e) 4.0 min, dis. 0.024, err. 0.422, in (a) 4.1 min, dis. 0.015, err. 0.463, (c) 2.6 min, dis. 0.056, err. 0.429, (d) 7.1 min, dis. 0.015, err. 0.391 and (f) 2.5 min, dis. 0.059, err. 0.448.

principle. In particular, a sensible value for τ_{dis} increases/decreases as the values of α and β increase/decrease. This is depicted in Fig. 7.3.

Further, we notice that the sensible value for τ_{dis} also decreases to 1.5 if $\alpha = 0$ (instead of the default value $\alpha = 500$). However, even with this adaptation the reconstruction contains obvious "ripples" in the background, see Fig. 7.3(a). As mentioned above these artifacts are expected in this underregularization case. On the other hand, using $\alpha = 2500$ in Fig. 7.3(c), the discrepancy parameter had to be increased to $\tau_{\text{dis}} = 6.0$ to stop early enough to receive an artifact-free reconstruction.

On the other hand, the choice of $\beta = 0$ leads to the reconstruction in Fig. 7.3(d). In comparison to the reconstruction in Fig. 7.3(e) one can observe slightly less artifacts in the background but a hole in the ball. The result of setting $\beta = 5 \cdot 10^{-5}$ is presented in Fig. 7.3(f). One can observe a hint of "ripples" in the background. Note that it was necessary to set the discrepancy principle parameter to $\tau_{\text{dis}} = 6.0$.

Balanced Penalty Terms We expect visible effects of the sparsity promoting and TV penalty term, while these penalty terms have the same scale, i.e. they are *balanced*. The previous choices of regularization parameters as in Figs. 7.2 and 7.3 lead into highly unbalanced penalty terms of scale $f_{\text{spa}} = 1$ and $f_{\text{tv}} = 10^{-5}$. In particular, for a visible sparsity effect during balanced penalty terms, the primal-dual algorithm needs a high iteration number at least in the last outer step. This is demonstrated in Fig. 7.4 choosing $\beta = 10^{-5}$ (as before) and $\alpha = 3 \cdot 10^{-3}$. For the first 11 outer iterations were employed 50 inner iterations and for the last outer step there were $5 \cdot 10^5$ inner iterations, which



(a) Test phantom. (b) Reconstruction.

Figure 7.4: Visible effects of the sparsity promoting and TV penalty term while these penalty terms have the same scale, i.e. they are balanced. (a) Test phantom as in Fig. 7.1. (b) Balanced penalty terms are achieved with sparsity regularization parameter $\alpha = 3 \cdot 10^{-3}$ and TV regularization parameter $\beta = 10^{-5}$. Run times, relative discrepancies and relative errors were 10.8 h, dis. 0.058, err. 0.320.

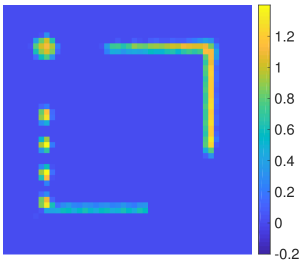


Figure 7.5: Reconstruction with grid scaling (halved number of discretization points in each dimension for the inversion process).

needs huge computational cost. We see a good reconstruction with the correct shape and an artifact-free and essentially fitting contrast. The main niggles are a general loss of contrast, typical step-artifacts of TV penalty in the part of the corner with the gradual contrast ramp, the overestimated thickness of the thin line and its very low contrast as well as the circle-like look of the squares. However, in this case penalty terms are balanced: $f_{\text{spa}} = 4.2 \cdot 10^{-6}$ and $f_{\text{tv}} = 4.1 \cdot 10^{-6}$. The discrepancy was $f_{\text{dis}} = 2.4 \cdot 10^{-7}$.

To get an efficient algorithm we relinquish the balance of the penalty terms and increase the regularization parameter α to 500 as in the examples before. This enables the utilization of sparsity with a much lower number of inner iterations.

Grid Scaling To save computational cost the inversion process for the contrast q can be carried out on a coarser *grid* than the one used for the forward mapping.[◇] To that end, the contrast is scaled down for the inversion process and scaled up for the direct scattering problem. For fast downscaling the MATLAB function imresize is employed with interpolation method nearest in 2D. In the case of 3D imresize is applied on the slices of the grid. To analyze the effects of this approach we want to compare the result with the example from Fig. 7.2(a). Therefore the number of outer iterations was fixed to 9 and a fixed number of 50 inner iterations was used. The related reconstruction with halved number of discretization points in each dimension for the inversion process is depicted in Fig. 7.5. Altogether, a slight speed up can be reported (about 15%), but the reconstruction quality significantly decreased.

[◇] Remember the necessary discretization points per wavelength to compute the wave fields, see Sec. 4.3.

7.6.2 Numerical Examples for Synthetic Data in 3D

We have already considered a numerical example for a reconstruction from synthetic data in 3D in Ch. 5. However, in this section we investigate the influence of the different stopping strategies for the inner iteration. Furthermore, the phantom is a "cube-like" object, see Fig. 7.6(a). To the author's best knowledge, this is a challenging obstacle.

The experimental set-up is the same as in the *sphere* case in Sec. 5.2. So, near field data is used with 31 transmitters and receivers distributed roughly uniformly on a sphere with Fibonacci lattice, see [Gon10, Sec. 3.1], with radius 5 m around the target. The wave number is set to $k = 10\,\mathrm{m}^{-1}$, i.e. a frequency of 0.5 GHz. The radius determining the computational domain D_{2R} is set to $R = 1\,\mathrm{m}$. The computational domain is discretized by $N = 256$ points in every dimension.[◇]

As in the 2D case we use the relative noise level $\delta = 0.01$, physical bounds $-1 \leq \mathrm{Re}(q) \leq 3$ and $0 \leq \mathrm{Im}(q) \leq 3$. The regularization parameters were determined automatically by Alg. 3.3.6. The parameter $\tau_{\mathrm{dis}} = 1.5$ was set to stop the outer iteration by the discrepancy principle. For the inner stopping strategies the same default settings as in the 2D case were employed.

The results are depicted in Fig. 7.6. In particular, we focus on the dependency of the quality and efficiency of the reconstruction scheme with respect to the stopping strategy in the inner iteration. Run times, relative discrepancies and relative errors are given in the caption. The only significant difference was the run times.

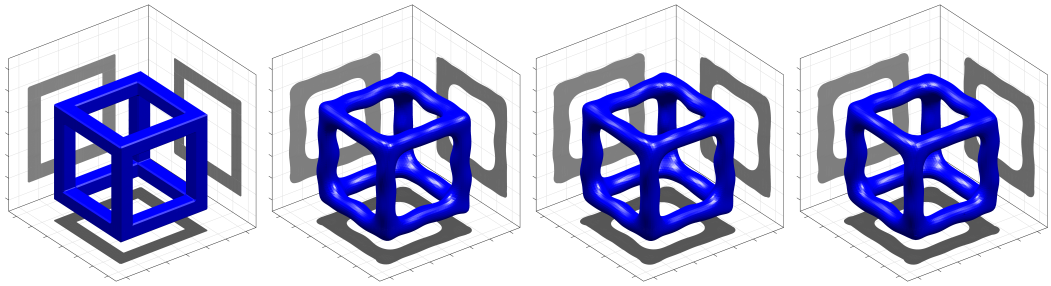
7.6.3 Numerical Examples for Real-World Data in 2D

In this section the proposed variational reconstruction algorithm is tested with real-world two-dimensional data published by Institute Fresnel. This data was already presented in Sec. 3.4.2. Further, the case of two dielectrics with 6 GHz has been considered before in Fig. 5.2. However, different stopping strategies are tested again for this data. Further, in addition to the contrast with two dielectrics another contrast with one dielectric is considered, see Fig. 3.10. Moreover, both obstacles are tested with 3 GHz and 5 GHz.

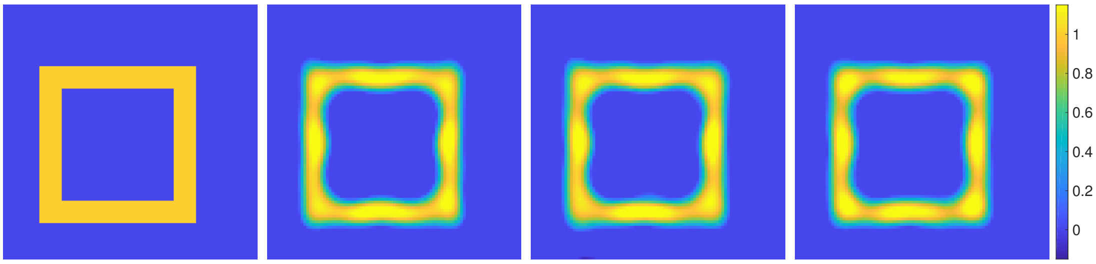
For the reconstruction, the grid size is $N = 256$ in every dimension. The regularization parameters are empirically derived to be $\alpha = 500$ and $\beta = 10^{-5}$ and the physical bounds are $-1 \leq \mathrm{Re}(q) \leq 3$ and $0 \leq \mathrm{Im}(q) \leq 1$.

Note that to stop the outer iteration by Morozov's discrepancy principle, see (3.18), one needs to know an estimate of the noise level δ. Based on the relative data error in Tab. 3.2 the values in Tab. 7.1 are chosen as noise levels to stop the outer iteration by the discrepancy principle with $\tau_{\mathrm{dis}} = 1.6$, see (3.18). In Fig. 7.7 the reconstructions for the above-mentioned targets and frequencies are presented for the case of stopping the inner iteration by stopping strategy 1. The related quantitative evaluation can be found in Tab. 7.1. The reconstructions for the constant stopping criterion with $N_{\mathrm{in}} = 10$ and stopping strategy 2 were very similar, qualitatively (visually) as well as quantitatively (run times, relative discrepancies, relative errors). Therefore they are omitted here.

[◇] A similar setting was also used in [BKL17, Sec 5.2] considering two other obstacles. The *latitude-longitude lattice*, see e.g. [Gon10], instead of the Fibonacci lattice was used there to distribute the transmitters/receivers. As in that case only 32 transmitters/receivers were set although 50 were requested, the setting is similar. However, there the computational domain was mistakenly denoted by D.

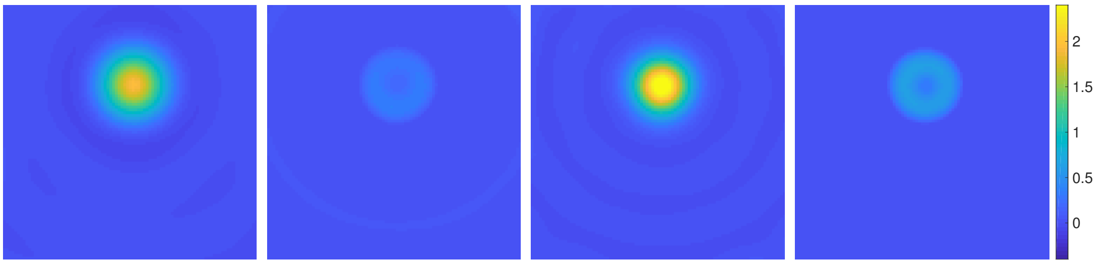


(a) Predef. contrast. (b) Const. (c) Stop. strategy 1. (d) Stop. strategy 2.

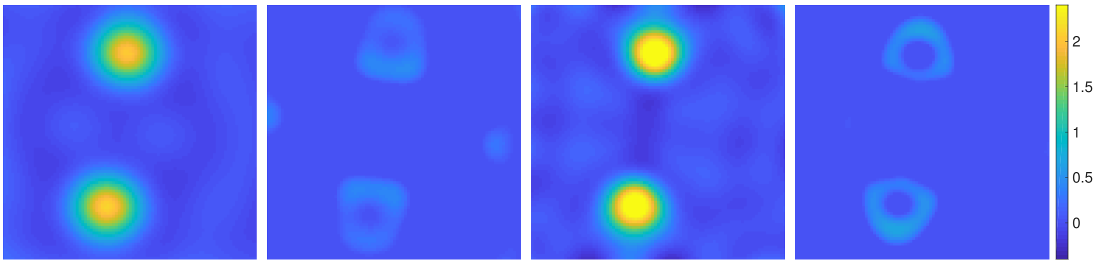


(e) Predef. contrast (plane). (f) Const. (plane). (g) Stop. strategy 1 (plane). (h) Stop. strategy 2 (plane).

Figure 7.6: Influence of the stopping strategies for the inner iterations. In all cases the real part of predefined contrast/reconstruction for the "cube-like" phantom with contrast $q = 1.0$ is depicted. (The obstacle has no imaginary part. The reconstruction vanishes in the imaginary part. Therefore the imaginary part is not shown.) In all cases the employed parameters were $\delta = 0.01$, $\alpha = 8031.25$ and $\beta = \alpha \cdot 10^{-5}/500$, $\tau_{\text{dis}} = 1.5$. The regularization parameters α and β were determined automatically by Alg. 3.3.6. For visualization isosurface data with isosurface value 0.3 was computed, see (a)–(d). The contrast on a sectional plane through the obstacle is presented in (e)–(h). Run times, relative discrepancies and relative errors were in (b)/(f) 3.3 h, dis. 0.014, err. 0.421, (c)/(g) 3.5 h, dis. 0.014, err. 0.428, (d)/(h) 5.3 h, dis. 0.015, err. 0.428.



(a) 3 GHz, real part. (b) 3 GHz, imag. part. (c) 5 GHz, real part. (d) 5 GHz, imag. part.



(e) 3 GHz, real part. (f) 3 GHz, imag. part. (g) 5 GHz, real part. (h) 5 GHz, imag. part.

Figure 7.7: Reconstructions of the real-world data set of Institute Fresnel for the one dielectric case (top row) and two dielectric case (bottom row) for two frequencies. In both cases the parameters described in the text were employed. The related "ground truth" are given in Fig. 3.10.

Target	Freq.	Noise level δ	Inner stop. strat.	N_{out}	Run time	Rel. dis.	Rel. err.
single dielectric	3 GHz	15%	1	7	1.3 min	0.239	0.547
single dielectric	5 GHz	20%	1	7	1.4 min	0.267	0.564
two dielectrics	3 GHz	15%	1	7	1.4 min	0.234	0.541
two dielectrics	5 GHz	25%	1	7	1.6 min	0.245	0.513

Table 7.1: Quantitative performance of the reconstruction scheme for Institute Fresnel's data set in two dimensions published by Institute Fresnel with our choice of noise levels. The noise level was estimated by the method described in text.

7.7 Reconstruction Using Wavelets

In analogy to Sec. 6.3 we are interested in a modification of the proposed variational reconstruction scheme, see Alg. 7.5.1, using *sparsity* in *wavelet* basis instead of pixel basis. We will consider two cases. However, as mentioned in Ch. 1 the total variation-based regularization is omitted in both cases.

We remember the generalization of the Tikhonov functional in (6.9), that has another exponent for the discrepancy term than in its norm. Taking into account physical bounds and the linearization as in (7.1) we end up with

$$f_q(h) := \underbrace{\tfrac{1}{p}\|\mathcal{F}'(q)[h]+\mathcal{F}(q)-F^{\delta}_{\text{meas}}\|^{p}_{\text{dis},P} + \frac{\alpha}{Q}\|w\odot\mathcal{W}(q+h)\|^{Q}_{\text{spa},Q}}_{=:F(Kh)} + \underbrace{\delta(q+h)}_{=:G}. \tag{7.35}$$

For the sake of simplicity we avoid the transformation of the complex-valued quantities in the notation. However, this is analogously respected as in Sec. 7.3.

Now, the two cases we will consider use the exponents $p = 1$ and $p = 2$ for the discrepancy term.[◇] Furthermore we limit ourselves to $P = 2$ to stay in Hilbert spaces. In addition, these two cases are restricted to $Q = 1$. However, as already mentioned in Ch. 1 Armin Lechleiter implemented a generalization, that provides more flexibility for the choice of P and Q but is restricted to $p = P$. It is likewise the generalization in Sec. 6.3 and uses a similar proximal mapping (apart from a shift resulting from the linearization, that was taken into account in (7.35)).[†] However, in the following we will consider the two restricted cases.

We will not look at the whole computations to modify the proposed variational reconstruction scheme because the techniques are similar to Secs. 7.3 and 7.4. However, the first difference is the assignment to F and G, defined in (7.35), as the sparsity penalty term is in F and not in G. Then,

$$\begin{aligned} K_{\text{dis}} &:= [\mathcal{F}'(q)], & v_{\text{dis}} &:= \mathcal{F}(q) - F^{\delta}_{\text{meas}}, \\ K_{\text{spa}} &:= \alpha\, w \odot \mathcal{W}, & v_{\text{spa}} &:= \alpha\, w \odot \mathcal{W}q. \end{aligned} \tag{7.36}$$

Of course, we used $w \odot \mathcal{W}(q+h) = w \odot \mathcal{W}q + w \odot \mathcal{W}h$. As already mentioned we omit the transformation of the complex-valued quantities in the notation. However, in the

---

[◇] In the *source code* of IPscatt the corresponding proximal mappings to them are used via seti.invNo = 4 and 5, see the *source code documentation* of setInvType.

[†] This generalization can be used via seti.invNo = 7 in IPscatt.

following we assume that the auxiliary quantities K_{dis}, v_{dis}, K_{spa} and v_{spa} are transformed analogously to (7.13). The definition of F (ignoring that the norms are in real vector spaces) is then given by $F(y_1, y_2) := \frac{1}{p}\|y_1 + v_{\text{dis}}\|_{\text{dis}}^p + \|y_2 + v_{\text{spa}}\|_{\text{spa}}$. The first summand is defined as F_{dis} and the second as F_{spa}.

In analogy to Sec. 7.3: The hidden operators T and T^{-1} are $\mathbb{R}$-linear. Both K_{dis} and K_{spa} are $\mathbb{R}$-linear. Finally, $K := (K_{\text{dis}}, K_{\text{spa}})$ is $\mathbb{R}$-linear.

Again, we are interested in the Fenchel conjugates of F_{dis} and F_{spa} to compute the proximal mappings of F_{dis}^* and F_{spa}^*. Then, we can compute the proximal mapping of F^*. In addition, we are interested in the proximal mapping of G. Finally, we have all ingredients to modify the proposed variational reconstruction scheme to minimize (7.35).

Proximal Mappings for $p = 1$ For $p = 1$ this computations end up with the proximal mapping of F_{dis}^* componentwise

$$((I + \sigma\,\partial F_{\text{dis}}^*)^{-1}(y))_i = \frac{y_i + \sigma\,(v_{\text{dis}})_i}{\max\{1, |y_i + \sigma\,(v_{\text{dis}})_i|\}}.$$

Note that it is similar to (7.28). Further, the proximal mapping of F_{spa}^* results in

$$(I + \sigma\,\partial F_{\text{spa}}^*)^{-1}(y) = \mathcal{I}_{[-h_N^d, h_N^d]}(y + \sigma\,v_{\text{spa}}). \tag{7.37}$$

As this proximal mapping is not similar to already treated ones, the computation will be considered below in Remark 7.7.1. Then, the proximal mapping of F^* can be computed using Remark 7.4.2. For the proximal mapping of G we receive

$$(I + \sigma\,\partial G)^{-1}(y) = \mathcal{I}_{[a,b]}(y + q) - q.$$

Remark 7.7.1 (The Proximal Mapping of F_{spa}^*). As $\|\cdot\|_{\text{spa}}$ is a sum, F_{spa} is a sum and its Fenchel conjugate splits into a sum too, see Lemma 7.4.1. Further, the proximal mapping of F_{spa}^* splits too, see Remark 7.4.2. Therefore it is sufficient to consider the proximal mapping of the Fenchel conjugate of $f(y) = h_N^d\,|y + v|$, where y is an entry of the vector y_2 and v is an entry of the vector v_{spa} in this remark. The three ingredients to compute the Fenchel conjugate of $f(y)$ are the following: First, the Fenchel conjugate of $f_1(y) = \rho\,g_1(y)$ for $\rho > 0$ is $f_1^*(y) = \rho\,g_1^*(y/a)$, see [BL06, Tab. 3.2]. Second, the Fenchel conjugate for shift function $f_2(y) = g_2(y+b)$ is $f_2^*(y) = g_2^*(y) - \langle b, y\rangle$, see [BL06, Tab. 3.2]. (This was already mentioned in Sec. 7.4.1.) Third, the Fenchel conjugate of $f_3(y) = |y|$ is $f_3^*(y) = 0$ for $|y| \leq 1$ and $f_3^*(y) = \infty$ for $|y| > 1$, see Fig. 3.6(b) in Sec. 3.3.1.

We compute the Fenchel conjugate of $f(y) = h_N^d\,|y + v|$ step by step using the just-mentioned three ingredients: First, the Fenchel conjugate of $f(y) = \rho\,g_1(y)$ with $\rho = h_N^d$ and $g_1(y) = |y+v|$ is $f^*(y) = \rho\,g_1^*(y/\rho)$. Second, the Fenchel conjugate of $g_1(y) = g_2(y+v)$ with $g_2(y) = |y|$ is $g_1^*(y) = g_2^*(y) - \langle v, y\rangle$. Third, the Fenchel conjugate of g_2 is $g_2^*(y) = 0$ for $|y| \leq 1$ and $g_2^*(y) = \infty$ for $|y| > 1$. Then we can insert $g_2^*(y)$ into $g_1^*(y)$ and use $g_1^*(y/\rho)$ in $f^*(y)$. This results in

$$f^*(y) = \rho\left(-\langle v, y/\rho\rangle + \begin{cases} 0 & \text{if } |y/\rho| \leq 1, \\ \infty & \text{if } |y/\rho| > 1, \end{cases}\right) = -\langle v, y\rangle + \begin{cases} 0 & \text{if } |y/\rho| \leq 1, \\ \infty & \text{if } |y/\rho| > 1. \end{cases}$$

We note that the last summand is the indicator function $\delta_{[-\rho,\rho]}(y)$, that was defined in Sec. 3.2.1. We noticed in Sec. 6.2 that the subdifferential of this indicator function is the inverse of the auxiliary step function, $\mathrm{step}^{-1}_{[-\rho,\rho]}$, see (6.5). Therefore the subdifferential of f^* is $\partial f^*(y) = -v + \mathrm{step}^{-1}_{[-\rho,\rho]}(y)$.

For the derivation of the proximal mapping of f^* we consider $x = (I + \sigma\,\partial f^*)^{-1}(y)$ and invert it, i.e. $x + \sigma\,\partial f^*(x) = y$. As $\sigma\,\mathrm{step}^{-1}_{[-\rho,\rho]} = \mathrm{step}^{-1}_{[-\rho,\rho]}$, the insertion of ∂f^* results in $y + \sigma\,v = x + \mathrm{step}^{-1}_{[-\rho,\rho]}(x)$. We noticed in Sec. 6.2 that the proximal mapping $(I + \mathrm{step}^{-1}_{[-\rho,\rho]})^{-1}$ is the interval projection operator $\mathcal{I}_{[-\rho,\rho]}$, defined in (6.6). Therefore $x = \mathcal{I}_{[-\rho,\rho]}(y + \sigma\,v)$ with $\rho = h_N^d$.

Finally, we have shown that the proximal mapping $(I+\sigma\,\partial f^*)^{-1}(y)$ is $\mathcal{I}_{[-h_N^d,h_N^d]}(y+\sigma\,v)$. Therefore (7.37) holds.

Proximal Mappings for $p = 2$ Of course, the proximal mappings of F^*_{spa} and G in the case of $p = 2$ are the same as in the case of $p = 1$. As to be expected, the proximal mapping of F^*_{dis} is (7.26).

As mentioned in Ch. 1 the sparsity in the wavelet basis did not outcompete the sparsity in the pixel basis. Therefore it was not a topic for further research.

8 Using the Factorization Method to Generate Further A Priori Information

The *factorization method* for inverse problems, see e.g. [Kir98, Kir99, Lec06], [KG08, Secs. 1.4, 4.3 and 4.4], [DLP$^+$11, Sec. 4.2] or [CK13, Sec. 5.6], delivers information about the obstacle's shape and position. This method is restricted to plane waves and far field data. However, the factorization method is fast and works without a prior knowledge. The main idea of this chapter is to use this a priori information to set the pixels to zero that are (most likely) outside this shape. We will realize that by area-based physical bounds. Therefore we will adapt the proposed variational reconstruction scheme, see Alg. 7.5.1.[◇] A further reduction of background artifacts can be expected as result. Clearly, the main motivation is to increase the reconstruction quality using the additional a priori information. Moreover, we will see that the use of the factorization method facilitates to set up a variational reconstruction scheme without any assumptions about the contrast. Furthermore, we will discuss the influence of the physical bounds on reconstructions (with and without the use of the factorization method) to show their importance.

One at least theoretical disadvantage of the factorization method is that it requires the far field patterns for *all* angles of measurements and *all* incident directions, see e.g. [DLP$^+$11, Sec. 4.2.1]. In practice it is sufficient to use enough angles of measurements and incident directions. However, the limitation for far field data remains in force.

This chapter is divided into three parts: a brief presentation of the factorization method; the interpretation of the factorization method's result and the necessary adaption of the proposed variational reconstruction scheme; numerical examples.

8.1 The Factorization Method

For a brief presentation of the factorization method we follow [Lec08, Sec. I-3] for the two-dimensional case and [DLP$^+$11, Sec. 4.2.1] for the three-dimensional case.

This method is a member of the sampling method's class and determines the shape of the obstacle from far field measurements. Hence, this also solves an inverse scattering problem. Therefore it is ill-posed in the sense of Hadamard, see at the beginning of Ch. 3, and, as already mentioned, a regularization is needed to get a stable solution.

Until here we were interested in the physical properties of the obstacle. More precisely, we reconstructed the contrast. Of course, this is more than just the shape and is computationally more expensive. Therefore the low computational effort of the factorization method is not surprising.

[◇] In the *source code* of IPscatt this adapted version can be used via seti.invNo = 8, see the *source code documentation* of setInvType.

Incident Plane Wave We follow [DLP$^+$11, Sec. 4.2.1] and consider an incident *plane wave*. In this chapter it is denoted by $u^{\mathrm{i}}(x,\theta) = \exp(\mathrm{i}kx \cdot \theta)$ for the observation position $x \in \mathbb{R}^d$ and the incident direction $\theta \in \mathbb{S}$.[◇]

Far Field Pattern As described in Sec. 2.1, cf. (2.11), far away the scattered field $u^{\mathrm{s}}(\cdot,\theta)$ behaves like an outgoing spherical wave with *far field pattern* $u^{\mathrm{s}}_\infty(\cdot,\theta)$. In this far field pattern $u_\infty(\hat{x},\theta)$ is $\hat{x}$ the angle of observation and θ the incident direction.[†]

For the contrast q we make the following assumptions hereafter in this section.

Assumption 8.1.1 ([Lec08, Assumption I.3] for 2D, [DLP$^+$11, Assumption 4.2.1] for 3D)**.** The contrast $q \in L^\infty(\mathbb{R}^d)$ has $\operatorname{supp}(q) = \overline{G}$, where G is a Lipschitz domain in $\mathbb{R}^d$ with connected complement. Furthermore, $\operatorname{Re}(q) \geq c$ for a $c > 0$ and $\operatorname{Im}(q) \geq 0$ almost everywhere in G.

The *far field operator* is defined as

$$F\colon L^2(\mathbb{S}) \to L^2(\mathbb{S}), \quad g \mapsto \int_{\mathbb{S}} g(\theta) u_\infty(\cdot,\theta)\,\mathrm{d}s(\theta). \tag{8.1}$$

This integral operator has a continuous (even analytical) kernel, see [CK13]. Therefore it is a compact operator on $L^2(\mathbb{S})$. Further, it maps the density g to the far field corresponding to the incident *Herglotz wave function*

$$v_g(x) = \int_{\mathbb{S}} g(\theta) u^{\mathrm{i}}(x,\theta)\,\mathrm{d}s(\theta) = \int_{\mathbb{S}} g(\theta) \mathrm{e}^{\mathrm{i}kx\cdot\theta}\,\mathrm{d}s(\theta), \quad x \in \mathbb{R}^d. \tag{8.2}$$

The far field operator F is known if the far field pattern $u_\infty(\hat{x},\theta)$ is known for all directions $\hat{x}$ of observation and all incident directions θ. Herglotz wave functions are entire solutions to the Helmholtz equation, see [CK13, Def. 3.18]. Finally, the problem to determine the shape of the obstacle is to determine $\overline{G}$ for a given F.

The factorization of the far field operator F is used to reconstruct G. The first proof thereof was given in [CK96]. To illustrate the concept we define two operators.

First, we define the *Herglotz operator*[‡],

$$H\colon L^2(\mathbb{S}) \to L^2(G), \quad g \mapsto \int_{\mathbb{S}} g(\theta) \mathrm{e}^{\mathrm{i}k\theta\cdot x}\,\mathrm{d}s(\theta), \quad x \in G.$$

This operator is a restriction of the Herglotz wave function v_g, see (8.2), to G.

We remember that the far field operator F, see (8.1), maps the density g to the far field, which corresponds to the incident Herglotz wave function v_g. Of course, the incident field v_g has the far field pattern $v_\infty = Fg$.

[◇] The notation $u^{\mathrm{i}}(\cdot,\cdot)$ is slightly overloaded because at the beginning of Ch. 2 we introduced the plane wave $u^{\mathrm{i}}(x,t) = \exp(\mathrm{i}(kx\cdot\theta - \omega t))$ with time-dependence $\exp(-\mathrm{i}\omega t)$. However, we omitted the time-dependence afterwards.

[†] Note that the incident direction is included in the notation of the far field pattern in comparison to (2.11).

[‡] In this case we do not follow [Lec08, Sec. I-3] for the 2D case because it includes the additional factor $\sqrt{|q|}$ in the definition. Instead of that we use the definition in [CK13, Sec. 3.3].

Second, we define the *solution operator*[◇]

$$S\colon L^2(G) \to L^2(\mathbb{S}), \quad f \mapsto v_\infty, \tag{8.3}$$

where v denotes the solution of the time-independent scattering problem, i.e. the radiating solution of the Helmholtz equation $\Delta v + k^2(1+q)v = -k^2 qf$ in $\mathbb{R}^d$, cf. Sec. 2.1[†] and v_∞ denotes the corresponding far field.

By construction it holds that

$$F = SH. \tag{8.4}$$

The following Lemma shows that the image of S can be used to characterize the shape of the obstacle.

Lemma 8.1.2 ([Lec08, Lemma I.5(c)] for 2D, [DLP$^+$11, Lemma 4.2.2] for 3D)**.** *The far field $\Phi_\infty(\cdot, z)$ of a point source $\Phi(\cdot - z)$ located at $z \in \mathbb{R}^d$, see* (2.6)*, belongs to the image of S if and only if $z \in G$.*

A proof is given in [Lec08, Lemma I.5(c)] and [DLP$^+$11, Lemma 4.2.2].

Note that we do not know S because G is unknown. We know the far field operator $F = SH$. However, in general, as the Herglotz operator H is compact, the image of F is different from the image of S. The clue of the method is to get a link between these two images.

Note further that the linear problem $Fg_z = \Phi_\infty(\cdot, z)$ in $L^2(\mathbb{S})$ is ill-posed because of the compactness of F, see [Rie03, Sec. 7.3].

The next step is the *factorization* of $F\colon L^2(\mathbb{S}) \to L^2(\mathbb{S})$ by $F = \gamma H^*TH$ with $H\colon L^2(\mathbb{S}) \to L^2(G)$, the middle operator $T\colon L^2(G) \to L^2(G)$, which is defined in the next theorem, and finally $H^*\colon L^2(G) \to L^2(\mathbb{S})$. The dimension-dependent factor γ was already defined in Sec. 2.1 and is repeated in the next theorem. Again this factor depends on the definition of the far field pattern u_∞.

Theorem 8.1.3 ([Lec08, Th. 1.4] for 2D, [DLP$^+$11, Th. 4.2.3] for 3D)**.** *The far field operator can be factored as*

$$F = \gamma H^*TH,$$

where γ is again the dimension-dependent factor $\gamma = \exp(\mathrm{i}\pi/4)/\sqrt{8\pi k}$ in 2D and $\gamma = 1/(4\pi)$ in 3D and $T\colon L^2(G) \to L^2(G)$ is defined by $Tf = k^2 q(f + v|_G)$ with $v \in H^1_{\mathrm{loc}}(\mathbb{R}^d)$ as radiating solution of the Helmholtz equation

$$\Delta v + k^2(1+q)v = -k^2 qf \quad \text{in } \mathbb{R}^d. \tag{8.5}$$

Proof. The proof follows [Lec08, Th. I.4] for 2D and [DLP$^+$11, Th. 4.2.3] for 3D. We consider the proof as it leads to adjoint of the Herglotz operator.

The unique solution $v \in H^1_{\mathrm{loc}}(\mathbb{R}^d)$ to the Helmholtz equation (8.5) and Sommerfeld's

---

[◇] The *solution operator* is called *data-to-pattern operator* in [Lec08, Sec. I-3].

[†] In Sec. 2.1 this equation is written down with $f = u^{\mathrm{i}}$.

radiation condition (2.5) is given by the volume potential[◇] see (2.7), see [DLP⁺11, Th. 4.1.4],

$$v = V(q(f + v|_G)). \tag{8.6}$$

Remember that the far field pattern of the fundamental solution $\Phi(x-y)$, i.e. the point source at $y \in \mathbb{R}^d$, is $\Phi_\infty(\hat{x}, y) := \gamma \exp(-\mathrm{i}k\hat{x}\cdot y)$ for $\hat{x} \in \mathbb{S}, y \in \mathbb{R}^d$, see (2.12). That is the reason why the adjoint of the Herglotz operator

$$H^*\colon L^2(G) \to L^2(\mathbb{S}), \quad f \mapsto \int_G f(x)\mathrm{e}^{-\mathrm{i}k\theta\cdot x}\,\mathrm{d}x, \quad x \in \mathbb{R}^d$$

maps $f \in L^2(G)$ to the far field pattern of the volume potential $V(f)$. More precisely, $\gamma H^* f = (Vf)_\infty / k^2$.[†]

Replacing f by Tf results in

$$\gamma H^* T f = \Big(V(q(f + v|_G))\Big)_\infty .$$

If we compare the right-hand side of this equation with (8.6), we see that $\gamma H^* T f$ is the far field pattern of the radiating solution v, see (8.6), i.e. $\gamma H^* T = S$ remembering (8.3). As $F = SH$, see (8.4), the factorization of F exists and is $F = \gamma H^* T H$. ☐

The factorization method, that is used for imaging, is constructed by this factorization.

Lemma 8.1.4 ([Lec08, Lemma I.5] for 2D, [DLP⁺11, Lemma 4.2.4] for 3D)**.** *We assume that the contrast q satisfies Assumption 8.1.1. Then the following statements hold.*

- *(a) The Herglotz operator $H\colon L^2(\mathbb{S}) \to L^2(G)$ is compact and injective.*
- *(b) The middle operator $T\colon L^2(G) \to L^2(G)$ is injective and has a natural splitting $T = T_0 + T_1$ into a sum of a coercive[‡] operator T_0, $T_0 f := k^2 q f$ and a compact operator T_1, $T_1 f = k^2 q v|_G$, where v is the radiating solution of* (8.5)*. Moreover, T is an isomorphism and*

$$\operatorname{Im}\left(\int_G \bar{f} T f\,\mathrm{d}x\right) \geq 0, \quad f \in L^2(G). \tag{8.7}$$

The corresponding proof is given in [Lec08, Lemma I.5] for 2D and in [DLP⁺11, Lemma 4.2.4] for 3D.

Note that the real and imaginary parts of a bounded linear operator T on a Hilbert space over $\mathbb{C}$ are the self-adjoint operators, see [KG08, Sec. 1.4],

$$\operatorname{Re}(T) := \frac{1}{2}(T + T^*), \quad \operatorname{Im}(T) := \frac{1}{2\mathrm{i}}(T - T^*), \tag{8.8}$$

[◇] This definition of the volume potential operator V already includes the factor k^2 in comparison to the definition in [Lec08, Ch. I] and [DLP⁺11, Ch. 4].

[†] The division by k^2 is necessary because the factor k^2 is already included in the volume potential operator V.

[‡] By definition, an operator $T : L^2(G) \to L^2(G)$ is coercive if a $c > 0$ exists such that $\operatorname{Re}\left(\int_G \bar{f} T f\,\mathrm{d}x\right) > c\|f\|^2_{L^2(G)}$ for all $f \in L^2(G)$.

where T^* is the adjoint operator of T. These definitions are in accordance with the corresponding definition for complex numbers, see [Lec08, Sec. I-3]. That does mean that $T = \mathrm{Re}T + \mathrm{i}\,\mathrm{Im}T$, see [Con85, Sec. 2.2]. This does not mean that in the numerical implementation the real part only contains real numbers and the imaginary part only imaginary ones for a matrix T. For the operator T, $\mathrm{Re}(T)$ is the self-adjoint part of T and $\mathrm{i}\,\mathrm{Im}(T)$ the non-self-adjoint part of T, see [Lec08, Sec. I-3].[◇]

In this context a slightly modified setting is used. We consider a Gelfand triple $X \subset U \subset X^*$, where U is a Hilbert space, X is a reflexive Banach space with dual X^* for the inner product of U and the embeddings are injective as well as dense. The real and imaginary part of the bounded operator $T\colon X \to X^*$ are defined as in the last paragraph, see [DLP$^+$11, Sec. 4.2.1].

Although the factorization is $F = \gamma\, H^*TH$, the factor γ is omitted in the following theorem. With the factor the same statement holds for $F_\sharp = |\mathrm{Re}(\gamma^{-1}F)| + \mathrm{Im}(\gamma^{-1}F)$, cf. [Lec06, Th. I.6]. However, the following theorem claiming an image identity for factorizations is the fundamental functional analytical result for the factorization method, see [Lec08, Sec. I-4].

Theorem 8.1.5 ([Lec08, Th. I.7] for 2D, [DLP$^+$11, Th. 4.2.5] for 3D)**.** *Let $X \subset U \subset X^*$ be a Gelfand triple with Hilbert space U and reflexive Banach space X. Let V be a second Hilbert space and $F\colon V \to V$, $H\colon V \to X$ and $T\colon X \to X^*$ linear and bounded operators with*

$$F = H^*TH.$$

Further assumptions are:

(a) H is compact and injective.

(b) The form of $\mathrm{Re}(T)$ is $\mathrm{Re}(T) = T_0 + T_1$, where T_0 is a coercive operator and $T_1\colon X \to X^*$ is a compact one.

(c) $\mathrm{Im}(T)$ is positive semidefinite[†] on X, i.e. $\langle \mathrm{Im}(T)\phi, \phi\rangle \geq 0$ for all $\phi \in X$.

(d) T is injective.

Then

$$F_\sharp = |\mathrm{Re}(F)| + \mathrm{Im}(F)$$

is positive definite and the images of $H^*\colon X^* \to V$ and $F_\sharp^{1/2}\colon V \to V$ coincide, i.e. $\mathrm{image}(H^*) = \mathrm{image}(F_\sharp^{1/2})$.

Note that the notation $|A|$ follows the definition in [KG08, Sec. 2.2],

$$\sqrt{|A|} = (A^*A)^{1/4}. \tag{8.9}$$

In particular, this notation does not mean the absolute values of the corresponding matrix in the discrete case.

[◇] Note the correction of the non-self-adjoint part: In [Lec08, Sec. I-3] $\mathrm{Im}(T)$ was described as non-self-adjoint part. However, $\mathrm{Im}(T)$ is self-adjoint and $\mathrm{i}\,\mathrm{Im}(T)$ is non-self-adjoint.

[†] This is also called *nonnegative* instead of *positive semidefinite*.

A proof of the theorem is given in [Lec08, Th. I.7] for 2D and in [DLP$^+$11, Th. 4.2.5] for 3D.

Remark 8.1.6 (The Meaning and Implementation of $|\mathrm{Re}(F)|$)**.** In this paragraph we discuss the meaning of the operator $|\mathrm{Re}(F)|$ to build an understanding of the implementation. In particular, we discuss why $|\mathrm{Re}(F)|$ is implemented as $V\mathsf{abs}(D)V^*$ for a unitary matrix V and a diagonal matrix D containing the eigenvalues of $\mathrm{Re}(F)$.

Therefore we define the operator $A := \mathrm{Re}(F)$, that is self-adjoint as mentioned above. As we are interested in the numerical implementation we consider the discrete case. The corresponding matrix $A := \mathrm{Re}(F) \in \mathbb{C}^{m\times m}$ is self-adjoint, i.e. $\langle Ay, x\rangle = \langle y, Ax\rangle$ for all $x, y \in \mathbb{C}^m$. Hence, A is hermitian, i.e. $A = A^*$ with $A^* := \bar{A}^T$.[◇] Therefore A is normal, i.e. $AA^* = A^*A$. Then, the spectral theorem delivers the existence of a unitary matrix $V \in \mathbb{C}^{m\times m}$, i.e. $V^* = V^{-1}$, with $V^{-1}AV = D$ with a diagonal matrix D containing the eigenvalues of A on the diagonal. Hence, $A = VDV^{-1} = VDV^*$.

We remember that $|A| = \sqrt{A^*A}$, see (8.9). In the following we insert $A^* = A = VDV^*$ and use the unitary property of V, i.e. $V^* = V^{-1}$. Then $|A| = \sqrt{A^*A} = \sqrt{VDV^*VDV^*} = \sqrt{VDDV^*}$. If a matrix $B = V\mathcal{D}V^{-1}$ with a diagonal matrix $\mathcal{D}$, then $\sqrt{B} = V\sqrt{\mathcal{D}}V^{-1}$ because $(V\sqrt{\mathcal{D}}V^{-1})^2 = V\sqrt{\mathcal{D}}V^{-1}V\sqrt{\mathcal{D}}V^{-1} = V\sqrt{\mathcal{D}}\sqrt{\mathcal{D}}V^{-1} = V\mathcal{D}V^{-1} = B$. Using this statement with the diagonal matrix $\mathcal{D} = DD$ results in $|A| = \sqrt{VDDV^*} = V\sqrt{DD}V^*$. Remember that D contains the eigenvalues of the self-adjoint operator $\mathrm{Re}(F)$ on the diagonal. As the eigenvalues of a self-adjoint operator are real, see e.g. [Con85, Ch. 2, Prop. 5.8], it holds $D = D^*$. Therefore $\sqrt{DD} = \sqrt{D^*D} = |D|$, see (8.9). Finally, we end up with $|A| = V|D|V^*$.

For the numerical implementation we use $|D| = \sqrt{D^2} = \mathsf{abs}(D)$ with the absolute values of each entry. Finally, we have shown that $|A| = V\mathsf{abs}(D)V^*$.

We consider Picard's criterion in the next theorem. The theorem afterwards gives the link between the field operator F and the obstacle G.[†]

Theorem 8.1.7 (Picard's criterion, [Rie03, Satz 2.3.7])**.** *We consider a compact operator* $A\colon X \to Y$ *with singular system* $\{(\lambda_j; v_j, \psi_j)\}$*. An element* $g \in \overline{\mathrm{image}(A)}$ *is in* $\mathrm{image}(A)$ *if and only if the following series converges,*

$$\sum_{j=1}^{\infty} \frac{|\langle g, \psi_j\rangle_Y|^2}{\lambda_j^2}.$$

A proof is given in [Rie03, Satz 2.3.7].

Theorem 8.1.8 ([Lec08, Th. I.6] for 2D, [DLP$^+$11, Th. 4.2.6] for 3D)**.** *We consider an orthonormal basis* $(\lambda_j, \psi_j)_{j\in\mathbb{N}}$ *of the self-adjoint compact operator* $F_\sharp := |\mathrm{Re}(\gamma^{-1}F)| + \mathrm{Im}(\gamma^{-1}F)$*. A point* $y \in \mathbb{R}^d$ *belongs to* G *if and only if*

$$\sum_{j=1}^{\infty} \frac{\left|\langle \Phi_\infty(\cdot, y), \psi_j\rangle_{L^2(\mathbb{S})}\right|^2}{\lambda_j} < \infty. \tag{8.10}$$

[◇] More precisely, a complex matrix is hermitian if and only if it is self-adjoint.

[†] Remember that $\overline{G} = \mathrm{supp}(q)$.

As this theorem is important we consider the proof from [Lec08, Th. I.6] for 2D and [DLP$^+$11, Th. 4.2.6] in 3D.

Proof. The proof essentially consists of an application of Th. 8.1.5 to the factorization $\gamma F = H^*TH$ with $V = L^2(\mathbb{S})$, $X = X^* = L^2(G)$. The assumptions of Th. 8.1.5 were verified in Lemma 8.1.4. Note that $\mathrm{Im}(T)$ is positive semidefinite because (8.7) implies that $\mathrm{Im}(\int_G \bar{f}Tf \,\mathrm{d}x) = \int_G \bar{f}\mathrm{Im}(T)f \,\mathrm{d}x$. So, Th. 8.1.5 delivers that $\mathrm{image}(H^*) = \mathrm{image}(F_\sharp^{1/2})$.

We remember the factorization $F = \gamma\, H^*TH$ from Th. 8.1.3 and $F = SH$, see (8.4). This implies that $\gamma\, H^*T = S$. The inverse of T exists because T is injective by Lemma 8.1.4. Therefore $\gamma\, H^* = ST^{-1}$ and $\mathrm{image}(T^{-1}) = L^2(G)$. Hence, $\mathrm{image}(H^*) = \mathrm{image}(ST^{-1}) = \mathrm{image}(S)$. Finally, we conclude that

$$\mathrm{image}(F_\sharp^{1/2}) = \mathrm{image}(S).$$

As (λ_j, ψ_j) is an eigensystem of $F_\sharp$, $(\lambda_j^{1/2}, \psi_j)$ is an eigensystem of $F_\sharp^{1/2}$. Finally, with Picard's criterion, see Th. 8.1.7, we conclude that

$$f \in L^2(\mathbb{S}) \text{ belongs to } \mathrm{image}(S) \quad \Leftrightarrow \quad \sum_{j=1}^{\infty} \frac{|\langle f, \psi_j\rangle_{L^2(\mathbb{S})}|^2}{\lambda_j} < \infty.$$

The claim of the theorem yields by Lemma 8.1.2. □

Numerical Implementation of the Factorization Method For the numerical implementation of the factorization method we follow [Lec08, Sec. I-3] and [DLP$^+$11, Sec. 4.2.1]. The far field $\Phi_\infty(\cdot, y)$ of the translated fundamental solution $\Phi(\cdot - y)$ is given by $\hat{x} \mapsto \gamma \exp(-\mathrm{i}k\hat{x} \cdot y)$, see (2.12). For a numerical implementation we plot the inverse of the value of the series in (8.10) on a grid. Of course, we have to truncate this series. We expect to get large values if y is inside G (i.e. inside the obstacle) and small values if it is outside. Unfortunately, this is not quite the case as the division by small eigenvalues λ_j leads to an ill-posed problem. For an approximation of λ_j by λ_j^δ with $|\lambda_j^\delta - \lambda_j| \le \delta|\lambda_j|$, the error in $|\langle \Phi_\infty(\cdot, y), \psi_j\rangle_{L^2(\mathbb{S})}|^2/\lambda_j^\delta$ is in general much larger than $\delta|\lambda_j|$. Furthermore, the singular vectors will have a numerical error, i.e. $\|\psi_j - \psi_j^\delta\| < \delta\|\psi_j\|$. Therefore it is necessary to regularize the series (8.10). One method is the Tikhonov regularization, i.e. to plot

$$z \mapsto \left(\sum_{j=1}^{\infty} \frac{\lambda_j}{(\lambda_j + \varepsilon)^2} \left| \langle \Phi_\infty(\cdot, y), \psi_j\rangle_{L^2(\mathbb{S})} \right|^2 \right)^{-1}, \tag{8.11}$$

where ε is a well-chosen regularization parameter. Of course, the eigensystem (λ_j, ψ_j) has to be replaced by numerical approximations $(\lambda_j^\delta, \psi_j^\delta)$. As the exponential decay of the singular values, keep in mind to trust only a restricted number of eigenvalues. Further, note that the factorization method's result separates the obstacle from the background but does not distinguish whether the contrast is different from zero in the real or imaginary part. This is important for the interpretation of the result.

8.2 Factorization Method's Result and its Interpretation

We want to use the factorization method to generate further a priori information to improve the proposed variational reconstruction scheme Alg. 7.5.1. Therefore we discuss: first, how to interpret the result such that the obstacle is separated from the background and second, the adaption of Alg. 7.5.1 to take into account this additional a priori information.

8.2.1 Interpretation of the Factorization Method's result

The generation of a priori information using the factorization method's result consists of three steps: initially, the factorization method itself is the evaluation of (8.11) for far field data. Then, this result has to be interpreted such that the obstacle is separated from the background. Finally, it is recommended to enlarge the suspected obstacle for safety reasons by adding some pixels in the neighborhood of this suspected obstacle. In the numerical examples the results of these steps are depicted in Fig. 8.1. The factorization method itself works for two as well as three dimensions. However, the presented separation corresponds to the two-dimensional setting.

Parameters in the Factorization Method In the factorization method we have to make two choices: the number of used eigenvalues to truncate the series in (8.11) and the regularization parameter ε.

The first insight will be that the number of eigenvalues is not greater than the number of receivers. The second one will be that the number of receivers, N_s, has to equal the number of transmitters, N_i. Both is justified by the following: As $F_\sharp$ is self-adjoint, its discretization is a square matrix. Then, it follows by the definition of $F_\sharp$ and (8.8) that the far field operator F is represented by a square matrix with the same dimensions as $F_\sharp$. In the implementation F is simply the far field data F^δ_meas, that is a matrix $\mathbb{C}^{N_\mathrm{s}\times N_\mathrm{i}}$, where N_s is the number of receivers and N_i the number of transmitters. Therefore we conclude that the number of receivers and transmitters is the same. Further, we remember that (λ_j, ψ_j) is an eigensystem to $F_\sharp$. Therefore we have not more than $N_\mathrm{s} = N_\mathrm{i}$ eigenvalues.

In the numerical experiments it was preferable to use all available eigenvalues: decreasing the number of used eigenvalues may result in a better reconstruction of the obstacle's shape but also in artifacts at the corners of the region of interest.

In [DLP$^+$11, Sec. 4.2.1] the choice of the regularization parameter ε as two times δ is suggested. However, in the numerical experiments presented in this thesis, $\varepsilon = \delta$ is a recommendable choice. A lower regularization parameter, e.g. $\varepsilon = \delta/2$, results in a less precise obstacle shape. A higher one, e.g. $\varepsilon = 2\delta$, produces artifacts in the corners of the region of interest, which might be misinterpreted as part of the obstacle. For even higher regularization parameters the artifacts in the corners grow.

Separation of the Obstacle from the Background To separate the suspected obstacle from the suspected background in the region of interest D we define a *mask* $\mathcal{M}$ (logical matrix) of size $N_R \times N_R$ labeling the suspected obstacle by 1 and the suspected background by 0. We denote the resulting matrix of (8.11) in the region of interest as Z and its mean value as m, then

$$\mathcal{M}_{i,j} := \begin{cases} 1 & \text{if } Z_{i,j} \geq \sigma m \text{ for } i,j = 1,\ldots,N_R, \\ 0 & \text{otherwise,} \end{cases} \tag{8.12}$$

where σ is a *separation factor* for the mask, e.g. we choose $\sigma = 1.25$.

Enlarge the Suspected Obstacle To avoid cutting off parts of the obstacle we enlarge the suspected obstacle by adding some pixels in the neighborhood of it. Therefore we define a new *mask* $\mathcal{N}$ in the region of interest D by

$$\begin{aligned} \mathcal{N}_{k,\ell} = 1 \text{ for } k = i-n,\ldots,i+n, \quad \ell = j-n,\ldots,j+n \\ \text{if } \mathcal{M}_{i,j} = 1 \text{ for } i,j = n+1,\ldots,N_R-(n+1). \end{aligned} \tag{8.13}$$

We have to restrict i and j to stay inside ROI. Moreover, n defines the number of pixels away from the original pixel by $n = \max\{1, \mathsf{round}(N_R \cdot \rho/100)\}$ with the *enlargement factor* ρ expressing the percentage of ROI, e.g. we choose $\rho = 2$. Finally, the mask $\mathcal{N}$ separates the region of interest D in the suspected obstacle and the suspected background part. In the following we simply use the descriptions obstacle and background although the suspected parts are intended.

8.2.2 Adaption of the Proposed Variational Reconstruction Scheme

The physical bounds must be zero in this background part. This additional a priori information gained through the factorization method's result can be used in the proposed variational reconstruction scheme, Alg. 7.5.1, after a slight modification.

We will use the mask $\mathcal{N}$, see (8.13), to separate the contrast $q \in \mathbb{C}^{N_R \times N_R}$. As described in Sec. 3.1 the contrast is considered as a vector $q \in \mathbb{C}^{N_\mathrm{D}}$. The obstacle part of the contrast q is denoted by $\check{q}$ and the background part by $\mathring{q}$. Therefore we rewrite the mask $\mathcal{N}$ as a vector denoted by $\tilde{\mathcal{N}}$. Then the indices of the vector q are separated into

$$\check{J} := \{i \in \{1,\ldots,N_\mathrm{D}\} : \tilde{\mathcal{N}}(i) = 1\}, \quad \mathring{J} := \{i \in \{1,\ldots,N_\mathrm{D}\} : \tilde{\mathcal{N}}(i) = 0\}, \tag{8.14}$$

where $\check{J}$ are the indices of the obstacle part and $\mathring{J}$ the indices of the background part. From now on, we split the contrast $q \in \mathbb{C}^{N_\mathrm{D}}$ in the obstacle part $\check{q}$ and the background part $\mathring{q}$ by $\check{q} = q(\check{J})$ and $\mathring{q} = q(\mathring{J})$.

As already mentioned the factorization method cannot differ between the real and imaginary part of an obstacle. However, Alg. 7.5.1 needs to consider them separately. Hence, we remember the splitting of the contrast into real and imaginary part, see Sec. 7.3, such that $q \in \mathbb{R}^{2N_\mathrm{D}}$, where the first N_D entries contain the real part and the last N_D the imaginary part. Then, the indices of $q \in \mathbb{R}^{2N_\mathrm{D}}$ can also be separated into the obstacle and background part. Finally, we end up with

$$\check{N}_\mathbb{R} = \check{J}, \quad \mathring{N}_\mathbb{R} = \mathring{J}, \quad \check{N}_\mathbb{C} = \check{J} + N_\mathrm{D}, \quad \mathring{N}_\mathbb{C} = \mathring{J} + N_\mathrm{D},$$

where the addition is meant for each element of the set. Of course, $\check{N}_\mathbb{R} \dot\cup \mathring{N}_\mathbb{R} \dot\cup \check{N}_\mathbb{C} \dot\cup \mathring{N}_\mathbb{C} = \check{J} \dot\cup \mathring{J} = \{1,\ldots,2N_\mathrm{D}\}$.

Remember the full Tikhonov functional (3.11)

$$\min_{q\in X} f(q), \quad f(q) := \underbrace{\tfrac{1}{2}\|\mathcal{F}(q) - F^{\delta}_{\mathrm{meas}}\|^2_{\mathrm{dis}}}_{=:f_{\mathrm{dis}}} + \underbrace{\alpha\|q\|_{\mathrm{spa}}}_{=:f_{\mathrm{spa}}} + \underbrace{\beta\|\nabla q\|_{\mathrm{tv}}}_{=:f_{\mathrm{tv}}} + \underbrace{\delta(q)}_{=:f_{\mathrm{phy}}}, \tag{8.15}$$

where $\alpha, \beta \geq 0$[◇] are regularization parameters and the (convex) functional for the physical bounds (3.7) is

$$\delta(q) := \delta_{[a,b]}(\mathrm{Re}(q)) + \delta_{[c,d]}(\mathrm{Im}(q)), \tag{8.16}$$

where $a \leq b$ and $c \leq d$ describe physical bounds for the real and imaginary part.

Taking into account that the background part of the contrast is zero, the functional for the physical bounds (8.16) changes to

$$\delta(q) := \delta_{[a,b]}(\mathrm{Re}\,\check{q}) + \delta_{[c,d]}(\mathrm{Im}\,\check{q}) + \delta_{[0,0]}(\mathrm{Re}\,\mathring{q}) + \delta_{[0,0]}(\mathrm{Im}\,\mathring{q}).$$

Accordingly, we have to adapt the proposed variational reconstruction scheme, Alg. 7.5.1, in particular (7.33), which is

$$x_j^{[n+1]} = (I + \tau G)^{-1}(w_{\mathrm{x},j}) = -q_j + \begin{cases} \mathcal{I}_{[a,b]}\mathbb{S}(w_{\mathrm{x},j} + q_j, \kappa) & \text{if } j \in N_{\mathbb{R}}, \\ \mathcal{I}_{[c,d]}\mathbb{S}(w_{\mathrm{x},j} + q_j, \kappa) & \text{if } j \in N_{\mathbb{C}}, \end{cases}$$

where $\mathcal{I}$ denotes the interval projection operator, see (6.6), and $\mathbb{S}$ the soft-shrinkage operator, see (3.13). We replace (7.33) by

$$x_j^{[n+1]} = (I + \tau G)^{-1}(w_{\mathrm{x},j}) = -q_j + \begin{cases} \mathcal{I}_{[a,b]}\mathbb{S}(w_{\mathrm{x},j} + q_j, \kappa) & \text{if } j \in \check{N}_{\mathbb{R}}, \\ \mathcal{I}_{[0,0]}\mathbb{S}(w_{\mathrm{x},j} + q_j, \kappa) & \text{if } j \in \mathring{N}_{\mathbb{R}}, \\ \mathcal{I}_{[c,d]}\mathbb{S}(w_{\mathrm{x},j} + q_j, \kappa) & \text{if } j \in \check{N}_{\mathbb{C}}, \\ \mathcal{I}_{[0,0]}\mathbb{S}(w_{\mathrm{x},j} + q_j, \kappa) & \text{if } j \in \mathring{N}_{\mathbb{C}}. \end{cases}$$

Indeed, the background parts are set to zero because $\mathcal{I}_{[0,0]}(x) = 0$.

Adapt the Initial Guess Another idea to further adapt the proposed variational reconstruction scheme, Alg. 7.5.1, is to use the a priori information about the suspected obstacle's shape for a suitable guess of the initial contrast. (Currently, this guess is 0.) Therefore we modify step 1 in Alg. 7.5.1, setting the entries of the initial contrast $q^{(0)}$ to 1 in the real and imaginary part if the obstacle is suspected there and 0 otherwise. However, the results of this approach were not convincing. Therefore this feature was not used in the following numerical examples.

8.3 Numerical Examples

In this section we consider numerical examples corresponding to the feature that uses the factorization method's result as a priori information for the reconstruction.

[◇] Of course, equality is only useful to switch off the corresponding penalty term.

Setting and Parameters For these examples we essentially use the default parameters of **IPscatt** for two dimensions, which are described in Sec. 7.6.1. The main difference to the default experimental set-up is the use of plane waves instead of point sources as incident waves and the use of the far field data instead of the near field data. Further, we increase the discrepancy tolerance τ_{dis} from 2.5 to 3.5 to stop the outer iteration by the discrepancy principle. Moreover, we decrease the fixed number of inner iterations to $N_{\text{in}} = 25$ (from 50). The regularization parameters $\alpha = 88750$ and $\beta = \alpha \cdot 10^{-5}/500$ were determined automatically by the corresponding heuristic algorithm, see Alg. 3.3.6.

For the factorization method and its interpretation we set the parameters as described in Sec. 8.2.1, i.e. the regularization parameter $\varepsilon = \delta$, the separation factor $\sigma = 1.25$ and the enlargement factor $\rho = 2$.

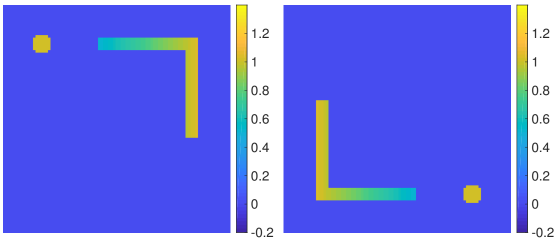
Interpretation of Factorization Method's Result As described in Sec. 8.2.1 the interpretation of the factorization method's result consists of three steps. They are depicted in Fig. 8.1. It is worth mentioning that a huge advantage of the factorization method is its low computational effort. Indeed, the factorization method cannot distinguish between an obstacle in the real and imaginary part, see Fig. 8.1(c). Further, it can be observed that the obstacle in the imaginary part is slightly more difficult to detect (and therefore to distinguish from the background) than the real part although they have the same contrast values. This is also reflected in the mask in Fig. 8.1(d) and in the increased mask in Fig. 8.1(e).

The main risk is to suspect background at an obstacle's place. In particular, this is challenging if the obstacle is not homogeneous. It requires to choose carefully the separation factor σ and the enlargement factor ρ.

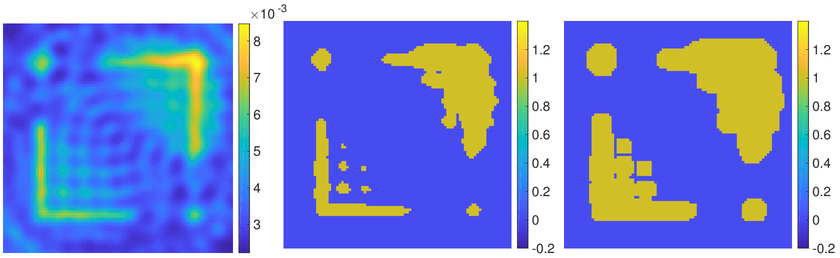
Reconstruction with and without the Factorization Method's Result The effect of this a priori information on the reconstruction using the proposed variational reconstruction scheme Alg. 7.5.1 with the modification described in Sec. 8.2.2 is shown in Fig. 8.2. It is compared with Alg. 7.5.1 without these modifications. As expected the additional a priori information improve the reconstruction result eliminating some background artifacts. Moreover, it improves the reconstruction quality of the obstacle itself. The a posteriori application of the used mask to the contrast in Fig. 8.2(c)/(f) would not have the same effect as the used modification of the reconstruction scheme. In particular, this is visible in the reconstruction of the gradual contrast ramp in Fig. 8.2(b) in comparison to Fig. 8.2(c).

Both reconstruction schemes stop in this example after the same number of outer iterations by the discrepancy principle. Therefore the run times do not differ significantly. However, in some numerical experiments the modified reconstruction scheme stopped earlier and was therefore faster—the reconstruction quality was still better.

A Priori Information from the Data Another interesting point is that the factorization method's result delivers a priori information about the contrast that only bases on the given data itself. In particular, this a priori information does not impose any conditions to the contrast like the sparsity promoting or total variation-based regularization or physical bounds—all these penalty terms are ingredients in the variational reconstruction scheme

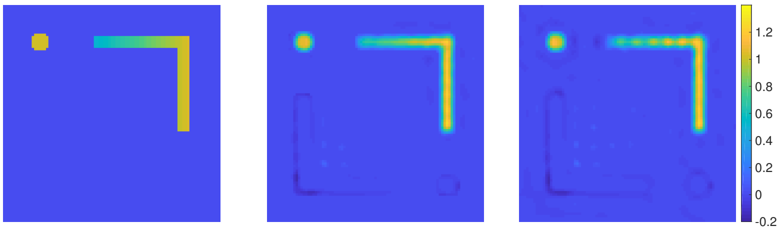


(a) Predefined contrast, real part. (b) Predefined contrast, imag. part.

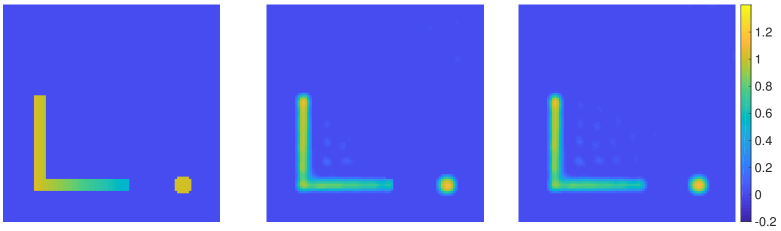


(c) Result of the factorization method, see (8.11). (d) Mask. (e) Increased Mask.

Figure 8.1: (a)/(b) Real and imaginary part of predefined contrast. (c) Result of the factorization method, see (8.11). The run time was 0.15 s. (d) Interpretation of the factorization method's result as mask. (e) Increased mask.



(a) Predefined contrast, real part. (b) Recon. with fact. method, real part. (c) Recon. without fact. method, real part.

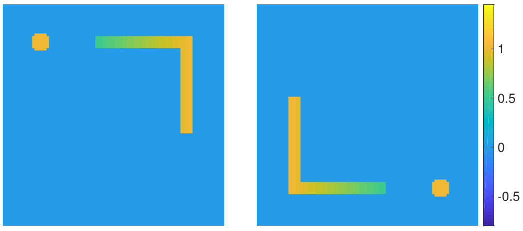


(d) Predefined contrast, imag. part. (e) Recon. with fact. method, imag. part. (f) Recon. without fact. method, imag. part.

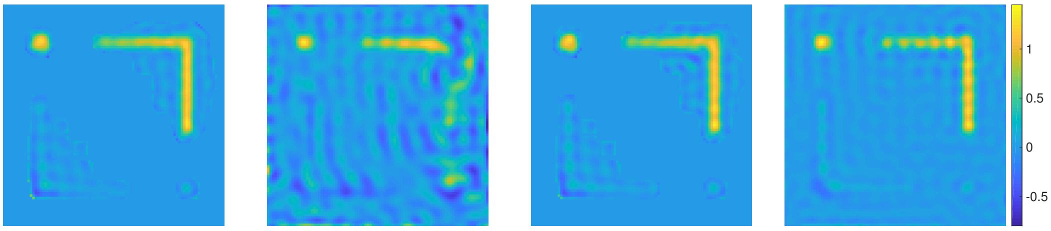
Figure 8.2: Reconstructions with and without the factorization method's result: (a)/(d) Real and imaginary part of predefined contrast. (Of course, these subfigures are the same as Fig. 8.1(a)/(b) and were duplicated for sake of easier comparison.) (b)/(e) Reconstruction with factorization method's result as a priori information. (c)/(f) Reconstruction without factorization method's result. Both reconstructions stopped after 10 outer iterations (by the discrepancy principle). Run times, relative discrepancies and relative errors were in (b)/(e) 2.1 min, 0.034, 0.313 and in (c)/(f) 2.0 min, 0.033, 0.326.

described in Alg. 7.5.1. Therefore we can construct a reconstruction scheme without any assumptions about the contrast[◇], at least for far field data: We use the factorization method's result and employ Alg. 7.5.1 with the modification described in Sec. 8.2.2, but we switch off the sparsity promoting and total variation-based regularization and the physical bounds. In fact, this results in a reasonable reconstruction, see Fig. 8.3(c)/(g). The same scheme without using the factorization method's result does not contain any regularization. As the inverse medium problem in scattering is ill-posed this, as expected, does not result in a reasonable reconstruction, see Fig. 8.3(d)/(h). Moreover, the same two schemes can be employed using physical bounds, see Fig. 8.3(e)/(i) and Fig. 8.3(f)/(j). The effect in the case with the factorization method is as expected: There are less background artifacts in Fig. 8.3(e)/(i) compared to Fig. 8.3(c)/(g). In particular, this is visible in the imaginary part of the contrast. The effect in the case without the factorization method is considerably larger: indeed, the use of physical bounds results in an upgradable but reasonable reconstruction, see Fig. 8.3(f)/(j), whereas switching them off results in the opposite, see Fig. 8.3(d)/(h). Hence, in this case the use of physical bounds is more than an improvement: it makes the difference between a reasonable reconstruction and a failure.

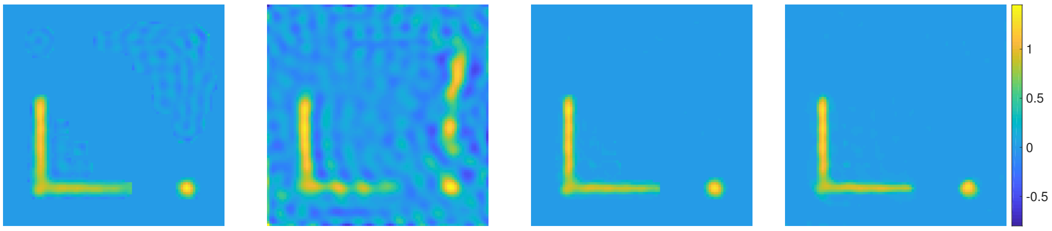
[◇] It should be mentioned that the separation of the obstacle from the background, i.e. the mask $\mathcal{M}$, see (8.12), needs enough background pixels to work reasonable. Clearly, it benefits from a sparse contrast.



(a) Predefined contrast, real part. (b) Predefined contrast, imag. part.



(c) Recon. with fact. method, no phys. bounds, real part. (d) Recon. without fact. method, no phys. bounds, real part. (e) Recon. with fact. method, with phys. bounds, real part. (f) Recon. without fact. method, with phys. bounds, real part.



(g) Recon. with fact. method, no phys. bounds, imag. part. (h) Recon. without fact. method, no phys. bounds, imag. part. (i) Recon. with fact. method, with phys. bounds, imag. part. (j) Recon. without fact. method, with phys. bounds, imag. part.

Figure 8.3: Reconstructions with or without the factorization method's result but without the sparsity promoting and total variation-based regularization: (a)/(b) Real and imaginary part of predefined contrast. (These contrasts are the same as in Fig. 8.1(a)/(b). However, they are repeated with extended limits of the colorbar for sake of easier comparison.) (c)–(j) Reconstructions with or without factorization method's result as a priori information and with or without the use of physical bounds. In all reconstructions the regularization parameters α and β are set to zero, i.e. sparsity promoting and total variation-based regularization are not used. Hence, if physical bounds are not used, then there are no a priori information used (except for the perhaps used factorization method's result).

Run times, number of outer iterations, relative discrepancies and relative errors were in (c)/(g) 2.7 min, 11, 0.024, 0.379, in (d)/(h) 3.6 min, 15, 0.192, 1.067, in (e)/(i) 1.8 min, 8, 0.029, 0.375 and in (f)/(j) 2.5 min, 11, 0.031, 0.405. In the case of (d)/(h) the reconstruction scheme was stopped manually after 15 outer iterations. All other reconstructions stopped by the discrepancy principle. Note that a higher number of outer iterations in (d)/(h) results in a slightly decreased discrepancy but in an even worse reconstruction (indeed, the error increases).

9 Summary and Outlook

The direct scattering problem and its numerical solution are already known. However, we briefly discussed them in Ch. 2 as they are required to tackle the inverse medium problem in scattering. The first step towards an effective and efficient reconstruction scheme for the inverse problem is to work out effective penalty terms, which, in particular, promote information expected in real-world environments. This was done in Ch. 3 and resulted in three penalty terms: sparsity promoting, total variation-based regularization, and physical bounds. As the quality of the reconstruction depends on a suitable choice of the involved regularization parameters, a heuristic algorithm for an automatic, data-driven parameter choice is helpful for any practitioner. Therefore such a routine was developed in Ch. 3. The already-known thresholded, nonlinear Landweber scheme takes into account sparsity regularization and was repeated in Ch. 6. An extension that takes into account physical bounds was developed in Ch. 6. As one penalty term, namely the total variation-based regularization, was still missing and the thresholded, nonlinear Landweber scheme is definitely not efficient, another scheme to reconstruct the contrast was developed in Ch. 7. This proposed variational reconstruction scheme, see Alg. 7.5.1, relies on the primal-dual algorithm, see [CP11a] and includes all three penalty terms. This contrast reconstruction is one of the main power-horses of the MATLAB toolbox IPscatt, that was briefly presented in Ch. 4 by discussing its main features and application cases. Numerical examples of the inverse medium problem in time-independent scattering were given in Ch. 5. The effectiveness and efficiency of IPscatt was shown by a comparison with two other software packages. The comparisons were carried out for synthetic as well as real-world data in two dimensions. Furthermore, it was demonstrated that IPscatt is efficient enough to deal with the inverse scattering problem in a three-dimensional setting. Indeed, the key to its effectiveness is the combination of the penalty terms and the key to its efficiency the utilization of the primal-dual algorithm. In Ch. 8 we briefly discussed the already-known factorization method, that delivers information about the obstacle's shape and position, and combined it with the proposed variational reconstruction scheme to increase the overall effectiveness of the reconstruction scheme. Further, we came across a reconstruction scheme without assumptions to the contrast (at least for far field data).

The modular and flexible design of IPscatt facilitates its use in a wide range of applications. For example, the ability to define arbitrary transmitter and receiver geometries is interesting for users interested in the simulation of scattering with extremely few transmitters and receivers. Together with the sophisticated variational reconstruction scheme it allows to study the limit cases of reconstructability in the inverse medium problem. For example, in the case of a setting as in Fig. 4.3 it turned out that at least 30 transmitters and receivers are required for a similar reasonable result. In addition, the influence of systematic errors besides the statistical ones is interesting, in particular for real-world data, and can be studied with IPscatt.

An idea to further improve the reconstruction result is to use data from several fixed frequencies instead of one frequency. In particular, this is interesting as the Fresnel data includes several frequencies. Further, if the homogeneity of an obstacle is known a priori as in the *Fresnel* case, see Sec. 5.2, this should be exploited in the reconstruction scheme. For example, a mask as for the interpretation of the factorization method's result, see Sec. 8.2.1, could be used to separate the suspected obstacle from the background. Then, the value of the obstacle's contrast should be the mean of the contrast values in the suspected obstacle. In addition, an automatic choice of at least one parameter in the factorization method, e.g. the separation factor σ, see Sec. 8.2.1, would be helpful for any practitioner. It should be effective to switch off the sparsity promoting and total variation-based regularization by setting $\alpha = \beta = 0$ and control the discrepancy by the separation factor—similar to the automatic parameter choice, see Alg. 3.3.6. Probably, this is easier with enabled physical bounds. However, it might be possible without them.

Furthermore, in many practical applications the phase of the (near or far) field is not available (at least not in a sufficient accuracy), e.g. see [MH17, ZZ18]. Therefore it is interesting to reconstruct the contrast of the obstacle (or at least its shape) from the intensity, i.e. the squared modulus, of the near or far field pattern. As the phase information is not available, this data is called *phaseless*. This kind of inverse scattering problem is also called *phase retrieval problem*. It would be interesting to give IPscatt an attempt at reconstruction from phaseless near field data $|u^{\mathrm{s}}|^2$ (instead of u^{s}) or phaseless far field data $|u_\infty|^2$ (instead of u_∞).

In addition, the replacement of the used numerical solver from [Vai00] for the direct scattering problem by more sophisticated algorithms, see e.g. [BGG17, Yin15], might be useful to deal with relatively high wave numbers. This was already mentioned in Ch. 1. A further step is to change the physical model and try to reconstruct the contrast. As mentioned in Ch. 1 the forward operator corresponding to the *anisotropic* Helmholtz equation was already implemented by Lechleiter and Rennoch in a draft version of IPscatt, see [LR17] or [Ren17, Ch. 3]. It is natural to reimplement it in the current release. In particular, this is interesting because sparsity regularization in wavelet basis and physical bounds were used in the draft version but not the total variation-based regularization. It is to be expected that the combined effect of all three penalty terms of the proposed variational reconstruction scheme, Alg. 7.5.1, namely the sparsity in the pixel basis, the total variation-based regularization and the physical bounds, increase the reconstruction quality. The presented automatic choice of the regularization parameters, Alg. 3.3.6, will also be helpful to improve the results. Furthermore, an implementation of the full Maxwell equations as forward operator would extend the application cases of IPscatt.

In addition, there are still two interesting open questions regarding the scattering simulation's and the reconstruction scheme's theory. The first question is about the convergence of the collocation method. In the case of smooth contrasts it was proven in [Vai00]. However, in the case considered here, i.e. piecewise smooth contrasts, the convergence was usually given in our numerical experiments, but its proof remains elusive. The second question is about the convergence of the underlying variational reconstruction scheme. Some of its aspects like the usage of the primal-dual algorithm are justified by [PCBC09, CP11a], but its connection to the properties of the outer iterations is yet to be entirely understood. However, as above, although the proof is still a matter of current research, the proposed reconstruction scheme turned out to be reliable in comprehensive numerical experiments in 2D and 3D, in particular, in the case of real-world data in 2D.

A Scattering Simulation

Scattering simulation is one of the main components of IPscatt. Hence, we summarize the basic formulas of the direct scattering problem in Tab. A.1 essentially in the chronological order of scattering. Without the claim of completeness we give them in continuous and discretized form as well as source code to build a connection between mathematical formulas and their implementation.

Note that the table was developed on the basis of the detailed description of the forward problem in [BKL17, Sec. 3]. As already mentioned in Ch. 2, [BKL17, Sec. 3] essentially bases on [LKK13] and research notes to the toolbox to [LKK13]. Further, it uses [CK13, Ch. 8] and [Vai00] as sources.

Single-layer potential for source points

C	$\mathrm{SL}_{\Gamma_\mathrm{i}\to D}\colon L^2(\Gamma_\mathrm{i}) \to L^2(D)$,	$(\mathrm{SL}_{\Gamma_\mathrm{i}\to D}\, g)(x) := \int_{\Gamma_\mathrm{i}} \Phi(x-y)g(y)\,\mathrm{d}s(y), \quad x \in D \setminus \Gamma_\mathrm{i}.$	(A.1)
D	$\mathrm{SL}_{N_\mathrm{i},N_D} : \mathbb{C}^{N_\mathrm{i}} \to \mathbb{C}^{N_\mathrm{D}}$,	$(\mathrm{SL}_{N_\mathrm{i},N_D}\, \underline{g}_{N_\mathrm{i}})(\ell) := \sum_{j=1}^{N_\mathrm{i}} \omega_j^\mathrm{i} u_j^\mathrm{i}(x_\ell)\underline{g}_{N_\mathrm{i}}(j)$ with $u_j^\mathrm{i}(x) = \Phi(x-p_j)$, $j = 1,\dots,N_\mathrm{i}$, in the case of source points at p_j and approximations $\omega_j^\mathrm{i} =$ `seti.dSInc` of the infinitesimal element on Γ_i.	
S	SL: incPnts → ROI,	`SL = dSInc(j)*incField(:,j)` with `dSInc`$(j) = \omega_j^\mathrm{i}$ and `incField` $= u_j^\mathrm{i}(x)\underline{g}_{N_\mathrm{i}}$ in `mimo.m`.	
C	$\mathrm{SL}_{\mathbb{S}\to D}$	The single-layer potential in the case of plane waves.	

Volume potential operator◇

C	$V_{2R}\colon L^2(D_{2R}) \to L^2(D_{2R})$,	$(V_{2R}f)(x) := \int_{D_{2R}} \Phi_{2R}(x-y)f(y)\,\mathrm{d}y, \quad x \in D_{2R}.$	(A.2)
D	$V_{N_\mathrm{D}} : \mathbb{C}^{N_\mathrm{D}} \to \mathbb{C}^{N_\mathrm{D}}$,	$V_{N_\mathrm{D}}\underline{f}_{N_\mathrm{D}} := \mathcal{R}_N \mathrm{FFT}_N^{-1}(\widehat{\Phi}_N \odot\,)\mathrm{FFT}_N \mathcal{E}_N \underline{f}_{N_\mathrm{D}}.$ It is restricted to ROI and shifted because of the FFT.	

S	V: ROI → ROI,	`V = seti.k^2.*helmholtz2Dr2r` with `helmholtz2Dr2r` computing
		`restrictCDtoROI(ifft2(reshape(seti.kHat,seti.nCD,seti.nCD).*...`
		`fft2(extendROItoCD(fROI,seti.ROImask))),seti.ROImask).`
			(A.3)

◇More precisely, we consider the *periodized* volume potential operator in this table.

Additional formulas for the volume potential operator

$\mathcal{E}\colon L^2(D)\to L^2(D_{2R})$,	$\boxed{\text{C}}$: $\mathcal{E}$ extends, $\mathcal{R}$ restricts and $\mathcal{E}^*=\mathcal{R}$.	
$\mathcal{R}\colon L^2(D_{2R})\to L^2(D)$	$\boxed{\text{D}}$: $\mathcal{E}_N\colon \mathbb{C}^{N_\mathrm{D}}\to\mathbb{C}^d_N$ and $\mathcal{R}_N\colon\mathbb{C}^d_N\to\mathbb{C}^{N_\mathrm{D}}$.	
	$\boxed{\text{S}}$: extendROItoCD: ROI → CD, restrictCDtoROI: CD → ROI.	
FFT_N, FFT_N^{-1}	$\boxed{\text{D}}$ and $\boxed{\text{S}}$: $\mathrm{FFT}_N=$ fft2, $\mathrm{FFT}_N^{-1}=$ ifft2.	
$\Phi_{2R},\widehat{\Phi}_N$, seti.kHat	$\boxed{\text{D}}$ and $\boxed{\text{S}}$: The kernel $\Phi_{2R}(x)$ in the computational domain is $k^2\Phi(x)$ if $x\in\overline{B_{2R}}$ and 0 if $x\in\overline{D_{2R}}\setminus\overline{B_{2R}}$. The *shifted Fourier coefficients* $\widehat{\Phi}_N=S_N^{-1}([(4R)^{d/2}\,\hat{\Phi}_{2R}(j)]_{j\in\mathbb{Z}^d_N})$ of Φ_N, where S_N^{-1} represents the shifting, are defined as seti.kHat in setKernel.m.	(A.4)

Application of the volume potential operator V_{2R} computing the scattered field u^s

	One incident field:	The routine solveLippmannSchwinger(Vq, f, seti) solves (A.12) computing v such that v − Vq(v) = f. Therefore a GMRES is used. For contrast qROI and incident field uIncROI, the scattered field is uScattROI = solveLippmannSchwinger(@(x) V(qROI.*x), V(qROI.*uIncROI), seti).	(A.5)
$\boxed{\text{S}}$	uScattROI: ROI → ROI		
	Multi-static:		
$\boxed{\text{S}}$	FFqROI	FFqROI(:,j) = uScattROI for each transmitter $j=1,\dots,N_\mathrm{i}$ in mimo.m.	(A.6)

Adjoint of the volume potential operator

$\boxed{\text{C}}$	$V^*_{2R}\colon L^2(D_{2R})\to L^2(D_{2R})$,	$(V_{2R}f)^*(x):=\displaystyle\int_{D_{2R}}\overline{\Phi_{2R}(x-y)}f(y)\,\mathrm{d}y,\quad x\in D_{2R}.$
$\boxed{\text{D}}$	$V^*_{N_\mathrm{D}}\colon\mathbb{C}^{N_\mathrm{D}}\to\mathbb{C}^{N_\mathrm{D}}$,	$V^*_{N_\mathrm{D}}\underline{f}_{N_\mathrm{D}}:=\mathcal{R}_N\mathrm{FFT}_N^{-1}(\overline{\widehat{\Phi}_N}\odot)\mathrm{FFT}_N\mathcal{E}_N\underline{f}_{N_\mathrm{D}}.$
$\boxed{\text{S}}$	VStar: ROI → ROI,	VStar = seti.k^2.*helmholtz2Dr2rAdjoint, cf. (A.3) using seti.kHat'.

Solution-to-data operator (for near field data)

$\boxed{\text{C}}$	$V_{D\to\Gamma_\mathrm{s}}\colon L^2(D)\to L^2(\Gamma_\mathrm{s})$	$(V_{D\to\Gamma_\mathrm{s}}f)(x):=k^2\displaystyle\int_D\Phi(x-y)f(y)\,\mathrm{d}y,\quad x\in\Gamma_\mathrm{s}.$	(A.7)
$\boxed{\text{D}}$	$V_{N_\mathrm{D},N_\mathrm{s}}\colon\mathbb{C}^{N_\mathrm{D}}\to\mathbb{C}^{N_\mathrm{s}}$	$(V_{N_\mathrm{D},N_\mathrm{s}}\underline{f}_{N_\mathrm{D}})(\ell):=h_N^dk^2\sum_{j\in\mathbb{Z}^d_N}\Phi(x_\ell-y_j^{(N)})\underline{f}_{N_\mathrm{D}}(j)$ with points $y_j^{(N)}$ in ROI, the position x_ℓ of the ℓth receiver and the area/volume h_N^d of the ROI's infinitesimal element (pixel/voxel).	(A.8)
$\boxed{\text{S}}$	uScattRX: ROI → measPnts	uScattRX = seti.k^2.*helmholtz2Dr2data(fROI, seti) in simo.m, where helmholtz2Dr2data yields (seti.measKer*fROI)*seti.dV with seti.dV = h_N^d, fROI = $\underline{f}_{N_\mathrm{D}}$ and seti.measKer = $\Phi(x_\ell-y_j^{(N)})$ for near field data defined in setMeasKer.m.	(A.9)

Adjoint of the solution-to-data operator (for near field data)

C	$V^*_{D\to\Gamma_\mathrm{s}}\colon L^2(\Gamma_\mathrm{s}) \to L^2(D)$	$(V^*_{D\to\Gamma_\mathrm{s}}f)(x) := k^2 \int_{\Gamma_\mathrm{s}} \overline{\Phi(x-y)}f(y)\,\mathrm{d}y, \quad x \in D.$	
D	$V^*_{N_\mathrm{D},N_\mathrm{s}}\colon \mathbb{C}^{N_\mathrm{s}} \to \mathbb{C}^{N_\mathrm{D}},$	$(V^*_{N_\mathrm{D},N_\mathrm{s}}\underline{f}_{N_\mathrm{s}})(j) := \omega^\mathrm{s}_j\, k^2 \sum_{\ell=1,\dots,N_\mathrm{s}} \overline{\Phi(x_\ell - y^{(N)}_j)}\underline{f}_{N_\mathrm{s}}(\ell).$ ω^s_j = seti.dSMeas: approximations of infinitesimal element of Γ_s.	(A.10)
S	VGStar: measPnts → ROI,	VGStar = seti.k^2.*helmholtz2Dr2dataAdjoint under the use of (seti.measKer')*(f.*seti.dSMeas) in helmholtz2Dr2dataAdjoint.m.	

Solution-to-data operator (for far field data)

C	$V_{D\to\mathbb{S}}$	Operator as in (A.7), (A.9), but `seti.measKer` $= \gamma\exp(-\mathrm{i}k\langle y, \theta_\ell\rangle)$, where y are points in ROI, θ_ℓ is the direction of the ℓth receiver and $\gamma = \exp(\mathrm{i}\pi/4)/\sqrt{8\pi k}$ if $x \in \mathbb{R}^2$ and $\gamma = 1/(4\pi)$ if $x \in \mathbb{R}^3$.	(A.11)
S			

Lippmann-Schwinger solution operator

C	$T_q\colon L^2(D) \to L^2(D)$	$T_q := (I - V_{2R}(q\,\cdot))^{-1}$	(A.12)
D	$T_{\underline{q}}\colon \mathbb{C}^{N_\mathrm{D}} \to \mathbb{C}^{N_\mathrm{D}}$	$T_{\underline{q}} := (I - V_{N_\mathrm{D}}(\underline{q}\odot\,))^{-1}$ with $\underline{q} \in \mathbb{C}^{N_\mathrm{D}}$.	
C	$T^*_q\colon L^2(D) \to L^2(D)$	$T^*_q := (I - (V_{2R}(q\,\cdot))^*)^{-1}$	
D	$T^*_{\underline{q}}\colon \mathbb{C}^{N_\mathrm{D}} \to \mathbb{C}^{N_\mathrm{D}}$	$T^*_{\underline{q}} := (I - (V_{N_\mathrm{D}}(\underline{q}\odot\,))^*)^{-1}$ with $\underline{q} \in \mathbb{C}^{N_\mathrm{D}}$	
S	TStar: ROI → ROI	TStar = @(f) solveLippmannSchwinger(VqStar,f,seti) under the use of VqStar = @(x) conj(qROI).*VStar(x). Remember that solveLippmannSchwinger computes v such that v − VqStar(v) = f.	

Forward operator (multi-static contrast-to-measurement operator)

C	$\mathcal{F}\colon L^p_{\mathrm{Im}\geq 0}(D) \to \mathrm{HS}$	$\mathcal{F}(q) := V_{D\to\Gamma_\mathrm{s}}(q\cdot)T_q\,\mathrm{SL}_{\Gamma_\mathrm{i}\to D}.$	(A.13)
		Note that HS is the space of all Hilbert-Schmidt operators $\mathrm{HS}(L^2(\Gamma_\mathrm{i}), L^2(\Gamma_\mathrm{s}))$.	
D	$\underline{\mathcal{F}}\colon \mathbb{C}^{N_\mathrm{D}} \to \mathbb{C}^{N_\mathrm{s}\times N_\mathrm{i}}$	$\underline{\mathcal{F}}(\underline{q}) := V_{N_\mathrm{D},N_\mathrm{s}}(\underline{q}\odot\,)T_{\underline{q}}\,\mathrm{SL}_{N_\mathrm{i},N_\mathrm{D}}.$	
	One incident field:		
S	uScattRX: ROI → $\mathbb{C}^{N_\mathrm{s}}$	Compute [uScattRX, uScattROI] = S(SL) in mimo.m, where S = @(s) simo(qROI, s, seti) in intOpsFuncs.m. The task of simo.m is to compute u^s = uScattROI as in (A.5), fROI = QU(uIncROI+uScattROI) $= q \odot (u^\mathrm{i} + u^\mathrm{s})$ and uScattRX as in (A.9).	(A.14)
	Multi-static:		
S	FFq: ROI → $\mathbb{C}^{N_\mathrm{s}\times N_\mathrm{i}}$	FFq(:,j) = uScattRX for each transmitter $j = 1, \dots, N_\mathrm{i}$ in mimo.m.	(A.15)

FRÉCHET DERIVATIVE OF THE FORWARD OPERATOR			
C	$\mathcal{F}'(q)\colon L^2(D) \to \mathrm{HS}$, (cf. (A.13) for HS)	$\mathcal{F}'(q)[h]g = V_{D\to\Gamma_\mathrm{s}}(I + (q\cdot)T_q V_{2R})(h\cdot)T_q\,\mathrm{SL}_{\Gamma_\mathrm{i}\to D}\,g, \quad g \in L^2(\Gamma_\mathrm{i})$. In finite-dimensional spaces the derivative $\mathcal{F}'(q)$ is represented by the Jacobian matrix denoted by $\underline{\mathcal{F}}'(\underline{q})$, see [BKL17, Sec. 3.6]:	(A.16)
D	$\underline{\mathcal{F}}'(\underline{q})\colon \mathbb{C}^{N_\mathrm{D}} \to \mathbb{C}^{N_\mathrm{s}\times N_\mathrm{i}}$	$\underline{\mathcal{F}}'(\underline{q})[\underline{h}] = A_{N_\mathrm{D},N_\mathrm{s}}(\underline{h}\odot)B_{N_\mathrm{D},N_\mathrm{i}}$ for $\underline{h} \in \mathbb{C}^{N_\mathrm{D}}$ with $A_{N_\mathrm{D},N_\mathrm{s}} = V_{N_\mathrm{D},N_\mathrm{s}}\big(I + (\underline{q}\odot)T_{\underline{q}}V_{N_\mathrm{D}}\big) \in \mathbb{C}^{N_\mathrm{s}\times N_\mathrm{D}}$, $B_{N_\mathrm{D},N_\mathrm{i}} = T_{\underline{q}}\,\mathrm{SL}_{N_\mathrm{i},N_\mathrm{D}} \in \mathbb{C}^{N_\mathrm{D}\times N_\mathrm{i}}$.	(A.17)
S	DFFq	[JA,JB] = derivative(seti,qROI) with JA = $A_{N_\mathrm{D},N_\mathrm{s}}$ and JB = $B_{N_\mathrm{D},N_\mathrm{i}}$ as well as qROI = $\underline{q}$, such that DFFq = @(h) JA*diag(h)*JB computes $\underline{\mathcal{F}}'(\underline{q})[h]$ by DFFqh = DFFq(h).	
FRÉCHET DERIVATIVE'S ADJOINT OF THE FORWARD OPERATOR			
D	$[\underline{\mathcal{F}}'(\underline{q})]^*\colon \mathbb{C}^{N_\mathrm{s}\times N_\mathrm{i}} \to \mathbb{C}^{N_\mathrm{D}}$	$[\underline{\mathcal{F}}'(\underline{q})]^*\underline{H} = \sum_{j=1}^{N_\mathrm{s}}\sum_{\ell=1}^{N_\mathrm{i}} \underline{H}_{j,\ell}\,\overline{A_{N_\mathrm{D},N_\mathrm{s}}(j,\cdot)}\,\overline{B_{N_\mathrm{D},N_\mathrm{i}}(\cdot,\ell)}$ for $\underline{H} \in \mathbb{C}^{N_\mathrm{s}\times N_\mathrm{i}}$. Since discretizations of the domain space $L^p_{\mathrm{Im}\geq 0}$ and the inner product in HS, see (A.13), take into account weights for physical reasons, there is an additional factor $\omega^\mathrm{s}_j/h^d_N$, see (A.8) and (A.10).	(A.18)
S	ADFFq	[ADFFq,seti] = adjOfDer(seti,qROI,FmeasDelta) where ADFFq = $[\underline{\mathcal{F}}'(\underline{q})]^*[\underline{\mathcal{F}}(\underline{q}) - F^\delta_\mathrm{meas}]$ with qROI = $\underline{q}$ and FmeasDelta = F^δ_meas. Note that the adjoint applied to the defect results in the derivative of the least-squares error of the forward operator.	

Table A.1: The table contains the basic formulas of the direct scattering problem: the single-layer potential $\mathrm{SL}_{\Gamma_\mathrm{i}\to D}$, the volume potential operator V, the solution-to-data operator $V_{D\to\Gamma_\mathrm{s}}$ and the Lippmann-Schwinger solution operator T_q. In addition, for practical usage their adjoints are given as well as the forward operator $\mathcal{F}$. Important ingredients of many reconstruction and optimization schemes are in the table too: the Fréchet derivative $\mathcal{F}'(q)$ and its adjoint $[\mathcal{F}'(q)]^*$.
Without the claim of completeness they are presented in the order of continuous and discretized formulas and the source code highlighted by the symbols C, D and S. If we want to stress the pointwise multiplication we use the notation $f \cdot g$ in the continuous case and $f \odot g$ in the discretized one. We underline a symbol to emphasize the discretization.

Bibliography

[AS65] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions*, volume 44 of *Dover books on intermediate and advanced mathematics*. Dover Publications, New York, unabridged and unaltered republication of the 1964 edition, 1965.

[BB88] Jonathan Barzilai and Jonathan M. Borwein. Two-Point Step Size Gradient Methods. *IMA Journal of Numerical Analysis*, 8(1):141–148, 1988. URL: https://doi.org/10.1093/imanum/8.1.141.

[BGG17] Carlos Borges, Adrianna Gillman, and Leslie Greengard. High Resolution Inverse Scattering in Two Dimensions Using Recursive Linearization. *SIAM Journal on Imaging Sciences*, 10(2):641–664, 2017. URL: https://doi.org/10.1137/16M1093562.

[BH15] Kristian Bredies and Martin Holler. A TGV-Based Framework for Variational Image Decompression, Zooming, and Reconstruction. Part II: Numerics. *SIAM Journal on Imaging Sciences*, 8(4):2851–2886, 2015. URL: https://doi.org/10.1137/15M1023877.

[BKL17] Florian Bürgel, Kamil S. Kazimierski, and Armin Lechleiter. A sparsity regularization and total variation based computational framework for the inverse medium problem in scattering. *Journal of Computational Physics*, 339:1–30, 2017. URL: https://doi.org/10.1016/j.jcp.2017.03.011.

[BKL19] Florian Bürgel, Kamil S. Kazimierski, and Armin Lechleiter. Algorithm 1001: IPscatt—A MATLAB Toolbox for the Inverse Medium Problem in Scattering. *ACM Transactions on Mathematical Software*, 45(4):45:1–45:20, 2019. URL: https://doi.org/10.1145/3328525.

[BL06] Jonathan Borwein and Adrian Lewis. *Convex Analysis and Nonlinear Optimization*. CMS Books in Mathematics. Springer, New York, second edition, 2006. URL: https://doi.org/10.1007/978-0-387-31256-9.

[BL11] Kristian Bredies and Dirk Lorenz. *Mathematische Bildverarbeitung: Einführung in Grundlagen und moderne Theorie*. Vieweg+Teubner, 2011. URL: https://doi.org/10.1007/978-3-8348-9814-2.

[Bor26] Max Born. Quantenmechanik der Stoßvorgänge. *Zeitschrift für Physik*, 38(11):803–827, 1926. URL: https://doi.org/10.1007/BF01397184.

[BS01] Kamal Belkebir and Marc Saillard. Special section: Testing inversion algorithms against experimental data. *Inverse Problems*, 17(6):1565–1571, 2001. URL: https://doi.org/10.1088/0266-5611/17/6/301.

[BT09] Amir Beck and Marc Teboulle. A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009. URL: https://doi.org/10.1137/080716542.

[Buk08] Alexander L. Bukhgeim. Recovering a potential from Cauchy data in the two-dimensional case. *Journal of Inverse and Ill-posed Problems*, 16(1):19–33, 2008. URL: https://doi.org/10.1515/jiip.2008.002.

[Cio90] Ioana Cioranescu. *Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems.* Springer Netherlands, 1990. URL: https://doi.org/10.1007/978-94-009-2121-4.

[CK96] David Colton and Andreas Kirsch. A simple method for solving inverse scattering problems in the resonance region. *Inverse Problems*, 12(4):383, 1996. URL: https://doi.org/10.1088/0266-5611/12/4/003.

[CK13] David Colton and Rainer Kress. *Inverse Acoustic and Electromagnetic Scattering Theory.* Springer, New York, 2013. URL: https://doi.org/10.1007/978-1-4614-4942-3.

[Con85] John B. Conway. *A Course in Functional Analysis*, volume 96 of *Graduate Texts in Mathematics.* Springer, New York, 1985. URL: https://doi.org/10.1007/978-1-4757-3828-5.

[CP11a] Antonin Chambolle and Thomas Pock. A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011. URL: https://doi.org/10.1007/s10851-010-0251-1.

[CP11b] Patrick L. Combettes and Jean-Christophe Pesquet. Proximal Splitting Methods in Signal Processing. In Heinz H. Bauschke, Regina S. Burachik, Patrick L. Combettes, Veit Elser, D. Russell Luke, and Henry Wolkowicz, editors, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pages 185–212. Springer, New York, 2011. URL: https://doi.org/10.1007/978-1-4419-9569-8_10.

[CS05] Tony F. Chan and Jianhong (Jackie) Shen. *Image Processing and Analysis.* Society for Industrial and Applied Mathematics, Philadelphia, 2005. URL: https://doi.org/10.1137/1.9780898717877.

[DDD04] Ingrid Daubechies, Michel Defrise, and Christine De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57(11):1413–1457, 2004. URL: https://doi.org/10.1002/cpa.20042.

[DLP$^+$11] Willy Dörfler, Armin Lechleiter, Michael Plum, Guido Schneider, and Christian Wieners. *Photonic Crystals: Mathematical Analysis and Numerical Approximation*, volume 42 of *Oberwolfach Seminars.* Birkhäuser, Basel, 2011. URL: https://doi.org/10.1007/978-3-0348-0113-3.

[dLS$^+$16] Maarten de Hoop, Matti Lassas, Matteo Santacesaria, Samuli Siltanen, and Janne P. Tamminen. Positive-energy D-bar method for acoustic tomography: a computational study. *Inverse Problems*, 32(2):025003, 2016. URL: https://doi.org/10.1088/0266-5611/32/2/025003.

[EHN96] Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of Inverse Problems.* Kluwer Academic Publishers, Dordrecht, Netherlands, 1996.

[Fig03] Mário A. T. Figueiredo. Adaptive sparseness for supervised learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9):1150–1159, 2003. URL: https://doi.org/10.1109/TPAMI.2003.1227989.

[FJ05] Matteo Frigo and Steven G. Johnson. The Design and Implementation of FFTW3. *Proceedings of the IEEE*, 93(2):216–231, 2005. Special issue on “Program Generation, Optimization, and Platform Adaptation”. URL: https://doi.org/10.1109/jproc.2004.840301.

[Geh13] Matthias Gehre. *Rapid Uncertainty Quantification for Nonlinear Inverse Problems.* PhD thesis, Universität Bremen, 2013. Retrieved from https://nbn-resolving.de/urn:nbn:de:gbv:46-00103519-10.

[Gon10] Álvaro González. Measurement of Areas on a Sphere Using Fibonacci and Latitude–Longitude Lattices. *Mathematical Geosciences*, 42(1):49, 2010. URL: https://doi.org/10.1007/s11004-009-9257-x.

[Had23] Jacques Hadamard. *Lectures on Cauchy's Problem in Linear Partial Differential Equations.* Yale University Press, New Haven, 1923.

[Häh96] Peter Hähner. A Periodic Faddeev-Type Solution Operator. *Journal of Differential Equations*, 128:300–308, 1996. URL: https://doi.org/10.1006/jdeq.1996.0096.

[HH14] Thorsten Hohage and Carolin Homann. A Generalization of the Chambolle-Pock Algorithm to Banach Spaces with Applications to Inverse Problems. arXiv:1412.0126, 2014. URL: https://arxiv.org/abs/1412.0126.

[Hoh01] Thorsten Hohage. On the numerical solution of a three-dimensional inverse medium scattering problem. *Inverse Problems*, 17(6):1743, 2001. URL: https://doi.org/10.1088/0266-5611/17/6/314.

[Kel95] C. Tim Kelley. *Iterative Methods for Linear and Nonlinear Equations.* Society for Industrial and Applied Mathematics, 1995. URL: https://doi.org/10.1137/1.9781611970944.

[Kel02] C. Tim Kelley. Iterative Methods for Linear and Nonlinear Equations. Companion Software, 2002. https://de.mathworks.com/matlabcentral/fileexchange/2198-iterative-methods-for-linear-and-nonlinear-equations/content/kelley/gmres.m. Accessed: March 2017.

[KG08] Andreas Kirsch and Natalia Grinberg. *The Factorization Method for Inverse Problems*, volume 36 of *Oxford Lecture Series in Mathematics and its Applications.* Oxford University Press, 2008. URL: https://doi.org/10.1093/acprof:oso/9780199213535.001.0001.

[Kir98] Andreas Kirsch. Characterization of the shape of a scattering obstacle using the spectral data of the far field operator. *Inverse Problems*, 14(6):1489, 1998. URL: https://doi.org/10.1088/0266-5611/14/6/009.

[Kir99] Andreas Kirsch. Factorization of the far-field operator for the inhomogeneous medium case and an application in inverse scattering theory. *Inverse Problems*, 15(2):413, 1999. URL: https://doi.org/10.1088/0266-5611/15/2/005.

[KS03] Steven Kusiak and John Sylvester. The scattering support. *Communications on Pure and Applied Mathematics*, 56(11):1525–1548, 2003. URL: https://doi.org/10.1002/cpa.3038.

[KS05] Jari Kaipio and Erkki Somersalo. *Statistical and Computational Inverse Problems*, volume 160 of *Applied Mathematical Sciences.* Springer, New York, 2005. URL: https://doi.org/10.1007/b138659.

[LAv09] Maokun Li, Aria Abubakar, and Peter M. van den Berg. Application of the multiplicative regularized contrast source inversion method on 3D experimental Fresnel data. *Inverse Problems*, 25(2):024006, 2009. URL: https://doi.org/10.1088/0266-5611/25/2/024006.

[Lec06] Armin Lechleiter. A regularization technique for the factorization method. *Inverse Problems*, 22(5):1605, 2006. URL: https://doi.org/10.1088/0266-5611/22/5/006.

[Lec08] Armin Lechleiter. *Factorization Methods for Photonics and Rough Surfaces*. PhD thesis, Universität Karlsruhe, 2008. URL: https://doi.org/10.5445/KSP/1000008735.

[LKK13] Armin Lechleiter, Kamil S. Kazimierski, and Mirza Karamehmedović. Tikhonov regularization in L^p applied to inverse medium scattering. *Inverse Problems*, 29(7):075003, 2013. URL: https://doi.org/10.1088/0266-5611/29/7/075003.

[LMM12] Dirk A. Lorenz, Peter Maaß, and Pham Q. Muoi. Gradient descent methods based on quadratic approximations of Tikhonov functionals with sparsity constraints: Theory and numerical comparison of stepsize rules. *Electronic Transactions on Numerical Analysis*, 39:437–463, 2012.

[LR06] Armin Lechleiter and Andreas Rieder. Newton regularizations for impedance tomography: a numerical study. *Inverse Problems*, 22(6):1967, 2006. URL: https://doi.org/10.1088/0266-5611/22/6/004.

[LR17] Armin Lechleiter and Marcel Rennoch. Non-linear Tikhonov Regularization in Banach Spaces for Inverse Scattering from Anisotropic Penetrable Media. *Inverse Problems and Imaging*, 11(1):151–176, 2017. URL: https://doi.org/10.3934/ipi.2017008.

[Mat15] MathWorks. MATLAB Documentation. https://de.mathworks.com/help/matlab/, 2015. Accessed: November 2015.

[MH17] Simon Maretzke and Thorsten Hohage. Stability Estimates for Linearized Near-Field Phase Retrieval in X-ray Phase Contrast Imaging. *SIAM Journal on Applied Mathematics*, 77(2):384–408, 2017. URL: https://doi.org/10.1137/16M1086170.

[MS12] Jennifer L. Mueller and Samuli Siltanen. *Linear and Nonlinear Inverse Problems with Practical Applications.* Computational Science & Engineering. SIAM, Philadelphia, 2012. URL: https://doi.org/10.1137/1.9781611972344.

[NW06] Jorge Nocedal and Stephen Wright. *Numerical Optimization.* Springer Series in Operations Research and Financial Engineering. Springer, New York, second edition, 2006. URL: https://doi.org/10.1007/978-0-387-40065-5.

[PCBC09] Thomas Pock, Daniel Cremers, Horst Bischof, and Antonin Chambolle. An algorithm for minimizing the Mumford-Shah functional. In *2009 IEEE 12th International Conference on Computer Vision*, pages 1133–1140, 2009. URL: https://doi.org/10.1109/ICCV.2009.5459348.

[Ren17] Marcel Rennoch. *Regularization Methods in Banach Spaces Applied to Inverse Medium Scattering Problems.* PhD thesis, Universität Bremen, 2017. Retrieved from https://nbn-resolving.de/urn:nbn:de:gbv:46-00105983-18.

[Rie03] Andreas Rieder. *Keine Probleme mit Inversen Problemen.* Vieweg, Wiesbaden, 2003. URL: https://doi.org/10.1007/978-3-322-80234-7.

[Roc67] R. Tyrrell Rockafellar. Duality and stability in extremum problems involving convex functions. *Pacific Journal of Mathematics*, 21(1):167–187, 1967. URL: https://doi.org/10.2140/pjm.1967.21.167.

[ROF92] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992. URL: https://doi.org/10.1016/0167-2789(92)90242-F.

[RW98] R. Tyrrell Rockafellar and Roger J.-B. Wets. *Variational Analysis*, volume 317 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, Heidelberg, 1998. URL: https://doi.org/10.1007/978-3-642-02431-3.

[SC03] David Strong and Tony Chan. Edge-preserving and scale-dependent properties of total variation regularization. *Inverse Problems*, 19(6):S165, 2003. URL: https://doi.org/10.1088/0266-5611/19/6/059.

[SKHK12] Thomas Schuster, Barbara Kaltenbacher, Bernd Hofmann, and Kamil S. Kazimierski. *Regularization Methods in Banach Spaces*, volume 10 of *Radon Series on Computational and Applied Mathematics*. De Gruyter, Berlin, 2012. URL: https://doi.org/10.1515/9783110255720.

[Str14] Robin Strehlow. *Regularization of the inverse medium problem: on nonstandard methods for sparse reconstruction.* PhD thesis, Universität Bremen, 2014. Retrieved from https://nbn-resolving.de/urn:nbn:de:gbv:46-00104187-19.

[Tou05] Hugo Touchette. Legendre-fenchel transforms in a nutshell. http://appliedmaths.sun.ac.za/~htouchette/archive/notes/lfth2.pdf, 2005. Accessed: October 2018.

[Vai00] Gennadi Vainikko. Fast Solvers of the Lippmann-Schwinger Equation. In Robert P. Gilbert, Joji Kajiwara, and Yongzhi S. Xu, editors, *Direct and Inverse Problems of Mathematical Physics*, pages 423–440. Springer, Boston, 2000. URL: https://doi.org/10.1007/978-1-4757-3214-6_25.

[Yin15] Lexing Ying. Sparsifying Preconditioner for the Lippmann-Schwinger Equation. *Multiscale Modeling & Simulation*, 13(2):644–660, 2015. URL: https://doi.org/10.1137/140985147.

[ZFM09] Royce K. P. Zia, Redish Edward F., and Susan R. McKay. Making sense of the Legendre transform. *American Journal of Physics*, 77(7):614–622, 2009. URL: https://doi.org/10.1119/1.3119512.

[ZZ18] Bo Zhang and Haiwen Zhang. Fast imaging of scattering obstacles from phaseless far-field measurements at a fixed frequency. *Inverse Problems*, 34(10):104005, 2018. URL: https://doi.org/10.1088/1361-6420/aad81f.

Index