Regularization and Numerical Solution of the Inverse Scattering Problem using Shearlet Frames

Gitta Kutyniok* Volker Mehrmann* Philipp Petersen*

July 28, 2014

Abstract

Regularization techniques for the numerical solution of nonlinear inverse scattering problems in two space dimensions are discussed. Assuming that the boundary of a scatterer is its most prominent feature, we exploit as model the class of cartoon-like functions. Since functions in this class are asymptotically optimally sparsely approximated by shearlet frames, we consider shearlets as a means for the regularization in a Tikhonov method. We examine both directly the nonlinear problem and a linearized problem obtained by the Born approximation technique. As problem classes we study the acoustic inverse scattering problem and the electromagnetic inverse scattering problem. We show that this approach introduces a sparse regularization for the nonlinear setting and we present a result describing the behavior of the local regularity of a scatterer under linearization, which shows that the linearization does not affect the sparsity of the problem. The analytical results are illustrated by numerical examples for the acoustic inverse scattering problem that highlight the effectiveness of this approach.

Keywords. Helmholtz equation, Inverse medium scattering, Regularization, Schrödinger equation, Shearlets, Sparse approximation.
AMS subject classification. 34L25, 35P25, 42C40, 42C15, 65J22, 65T60, 76B15, 78A46

1 Introduction

The scattering problem analyzes how incident waves, radiation, or particles, which are transmitted in a medium, are scattered at inhomogeneities of this medium. The associated inverse problem aims to determine characteristics of the inhomogeneities from the asymptotic behavior of such scattered waves. This problem appears in various flavors in different application areas, such as e.g. non-destructive testing, ultrasound tomography, and echolocation. For an overview of the problem and recent developments, we refer to the survey article [10].

Various numerical methods have been proposed for the solution of inverse scattering problems. A very common approach to solve a nonlinear inverse scattering problem are *fix-point iterations*, which produce a sequence of *linear* inverse scattering problems with solutions that converge, under some suitable assumptions, to a solution of the nonlinear problem. One such approximation technique is the *Born approximation*, see e.g. [3, 28]. However, one drawback of this class of approaches is the fact that it requires the solution of a linear inverse scattering problem in every iteration step, which is typically again an *ill-posed* problem that is hard

*Department of Mathematics, Technische Universität Berlin, 10623 Berlin, Germany; Email-Addresses: {kutyniok,mehrmann,petersen}@math.tu-berlin.de.

to solve in the presence of noisy data or data with linearization errors. A recently introduced different technique, see e.g. [26], tackles the nonlinear problem directly by minimizing a *Tikhonov functional* with a suitably chosen regularization term. The success of such an approach depends heavily on how properties of the solution are encoded in the regularization term. This, however, requires typically that a priori knowledge about characteristics of the solution is available.

We discuss both these approaches and combine them with a sparsity based methodology which makes use of representing the scatterer in a sparse way, as it has been suggested in several other areas of inverse problems. This methodology is based on the hypothesis that most types of data indeed admit a sparse approximation by a suitably chosen basis or frame, see Subsection 2.1, and today this is a well-accepted paradigm. Generally speaking, knowledge of a sparsifying basis or frame, appropriately applied, allows precise and stable reconstruction from very few and even noisy measurements. One prominent way to infuse such knowledge is by a regularization term such as in a Tikhonov functional. Indeed, in [26], it is assumed that the to-be-detected objects are sparse in the sense of small support, which is then encoded by using an L^p-norm for p close to 1 as regularization term, thereby promoting sparsity.

In this paper, we also aim to utilize sparsity to solve inverse scattering problems, but follow a different path. The key idea of our new approach is to generate a model for a large class of natural structures and an associated representation system, which provides asymptotically optimal sparse approximation of elements of this model class. We use this approach for both solution strategies for inverse scattering problems, and we focus on the *acoustic* and the *electromagnetic inverse scattering problem*.

1.1 Modeling of the Scatterer

Typically, a scatterer is a natural structure, which distinguishes itself from the surrounding medium by a change in density. In the 2D setting, this inhomogeneity can be regarded as a curve with, presumably, certain regularity properties. The interior as well as the exterior of this curve is usually assumed to be homogenous.

In the area of imaging sciences, the class of *cartoon-like functions* [14] is frequently used as model for images governed by anisotropic structures such as edges. Roughly speaking, a cartoon-like function is a compactly supported function which is a twice continuously differentiable function, apart from a piecewise C^2 discontinuity curve, see Definition 2.3 below. This cartoon-like model is well-suited for many inverse scattering problems, where the discontinuity curve models the boundary of a homogeneous domain. In some physical applications, one may debate this regularity of the curve as well as the homogeneity of the domains, but a certain smoothness on small pieces of the boundary seems a realistic scenario.

1.2 Directional Representation Systems

Having agreed on a model, one needs a suitably adapted representation system which ideally provides asymptotically optimal sparse approximations of cartoon-like functions in the sense of the decay of the L^2-error of best N-term approximation. Such a system can then be used for the regularization term of a Tikhonov functional.

The first (directional) representation system which achieved asymptotic optimality were *curvelets* introduced in [6]. In fact, in [7] curvelets are used to regularize linear ill-posed problems. This is done under the premise, that the solution of the inverse problem exhibits

edges, which tend to get smoothed out in a regularization procedure, while curvelets as a system adapted to edges overcomes this obstacle. However, on the practical side, curvelets suffer from the fact that often a faithful numerical realization of the associated transform is difficult.

In [17] shearlet systems were introduced, which similarly achieve the requested optimal sparse approximation rate [23], but in addition allow a unified treatment of the continuum and digital realm [25]. As curvelets, shearlets are mainly designed for image processing applications, in which they are also used for different inverse problems such as separation of morphologically distinct components [15, 24], recovery of missing data [16, 19], or reconstruction from the Radon transform [9]. Furthermore, in contrast to curvelets, compactly supported shearlet frames for high spatial localization are available [20], see [22] for a survey.

In view of this discussion, shearlet frames seem a good candidate as a regularizer for inverse scattering problems, and in fact this will be key to our new methodological approach.

1.3 Acoustic and Electromagnetic Inverse Scattering Problems

We focus on two inverse scattering problems, see e.g. [12], which are the *acoustic inverse scattering problem*, for which we introduce and study a strategy to directly solve the nonlinear problem, and the *electromagnetic inverse scattering problem*, for which we analyze the strategy to linearize the inverse scattering problem by means of the Born approximation.

The *acoustic inverse scattering problem* aims to reconstruct a contrast function which encodes the scatterer by emitting an acoustic wave and measuring the returning scattered waves. Common application areas are radar, sonar, and geophysical exploration, see e.g. [12] for a survey of applications.

The minimization of a suitable Tikhonov functional is a common approach to directly solve this nonlinear inverse problem. In [26] a sparsity-based regularization term is introduced which uses the L^p-norm with p close to 1 directly on the function to-be-recovered. This regularization scheme is very successful when the object under consideration has small support.

Following our methodological concept, and assuming that cartoon-like functions are an appropriate model for the scatterer, we instead choose as regularization term the ℓ_p-norm of the associated shearlet coefficient sequence with p larger or equal to 1. In Theorem 3.3 we prove convergence of this shearlet based regularization scheme. We also present numerical experiments that compare our approach to that of [26], see Subsection 5.2. These examples show convincing results, both in terms of the reconstruction error and the number of iterations. In particular, it is demonstrated that edges of the scatterer are recovered with high accuracy.

The *electromagnetic inverse scattering problem* aims to determine the shape of a scattering object from measurements of scattered incident electromagnetic plane waves. These problems appear, for instance, in applications such as medical imaging, where microwaves are used to detect leukemia, or non-destructive testing, where small cracks need to be detected inside of, for instance, metallic structures [11].

A prominent method to linearize the inverse scattering problem is by means of the Born approximation. Modeling the scatterer by cartoon-like functions, shearlets can be used again as a regularizer, provided that the transition from the nonlinear towards the linear problem does not influence the fact that the solution belongs to the class of cartoon-like functions. It has been shown in [29, 34] that certain singularities of the scatterer can still be found in the solution of the associated linearized problem. However, all these results require a global regularity of the scatterer to describe the regularity of the inverse Born approximation. On

the other hand, in the case of cartoon-like functions we have strong local but poor global regularity and therefore the results of [29, 34] can not be applied to our situation. To provide a theoretical basis for the application of shearlet frames, we prove that indeed the Born approximation to the electromagnetic Schrödinger equation gives rise to a scattering problem that exhibits sharp edges in the solution of the linearized problem. In particular, we show that the cartoon model is almost invariant under the linearization process, see Theorem 4.3 and Corollary 4.4. These results then provide the theoretical justification that shearlet systems can be used as regularization for the numerical solution of the associated linearized problems.

1.4 Outline of the Paper

The paper is organized as follows. The precise definition of shearlet systems, their frame properties, and their sparse approximation properties for cartoon-like functions is summarized in Section 2. Section 3 is devoted to the acoustic scattering problem. We first describe the direct and associated inverse problem, followed by the introduction of our new approach to regularize the inverse scattering problem using the shearlet transform in Subsection 3.3. We prove a convergence result in Theorem 3.3. The electromagnetic scattering problem is then introduced and studied in Section 4 with Theorem 4.3 being the main result on local regularity of the inverse Born approximation. Corollary 4.4 analyzes the situation of using the cartoon-like model as scatterer. Numerical experiments for the inverse acoustic scattering problem are provided in Section 5, highlighting the effectiveness of the shearlet-based regularization technique for the acoustic scattering problem as compared to other approaches.

2 Shearlet Systems

In this section we provide a precise definition of shearlet frames and recall their sparse approximation properties, see [22] for a survey on shearlets and [8] for a survey on frames.

2.1 Review of Frame Theory

Representation systems which are utilized for efficient encoding strategies often require a certain flexibility in their design, but should still lead to numerically stable algorithms. These desiderata are accommodated by the notion of a *frame*, which generalizes the notion of orthonormal bases by only requiring a norm equivalence between the Hilbert space norm of a vector and the ℓ_2-norm of the associated sequence of coefficients. To be more precise, given a Hilbert space $\mathcal{H}$ and an index set I, then a system $\{\varphi_i\}_{i\in I} \subset \mathcal{H}$, is called a *frame* for $\mathcal{H}$, if there exist constants $0 < \alpha_1 \leq \alpha_2 < \infty$ such that

$$\alpha_1 \|f\|^2 \leq \sum_{i\in I} |\langle f, \varphi_i\rangle|^2 \leq \alpha_2 \|f\|^2 \quad \text{for all } f \in \mathcal{H}.$$

The constants α_1, α_2 are referred to as the *lower* and *upper frame bound*, respectively. If $\alpha_1 = \alpha_2$ is possible, then the frame is called *tight*.

Frames provide a very good methodology for the analysis of functions and for efficient series expansions. For this, each frame $\Phi := \{\varphi_i\}_{i\in I} \subset \mathcal{H}$ is associated with three operators. The *analysis operator* T_Φ defined by

$$T_\Phi : \mathcal{H} \to \ell^2(I), \quad T_\Phi(f) = (\langle f, \varphi_i\rangle)_{i\in I},$$

decomposes a function into its *frame coefficients*, which typically allows the analysis of the original function. Second, the *synthesis operator* T_Φ^*, which is the adjoint of T_Φ, given by

$$T_\Phi^* : \ell^2(I) \to \mathcal{H}, \quad T_\Phi^*((c_i)_{i\in I}) = \sum_{i\in I} c_i \varphi_i,$$

allows to synthesize a function from the coefficients, and, third, the *frame operator* $S_\Phi := T_\Phi^* T_\Phi$, defined by

$$S_\Phi : \mathcal{H} \to \mathcal{H}, \quad S_\Phi(f) = \sum_{i\in I} \langle f, \varphi_i \rangle \, \varphi_i.$$

The operator S_φ, which can be shown to be self-adjoint and invertible, see e.g. [8], allows both a reconstruction of f given its frame coefficients and an expansion of f in terms of the frame elements, i.e.,

$$f = \sum_{i\in I} \langle f, \varphi_i \rangle \, S_\Phi^{-1} \varphi_i = \sum_{i\in I} \left\langle f, S_\Phi^{-1} \varphi_i \right\rangle \varphi_i \quad \text{for all } f \in \mathcal{H}.$$

Hence, although Φ does not constitute a basis, there exists a reconstruction formula using the system $\{\tilde{\varphi}_i\}_{i\in I} := \{S_\Phi^{-1}\varphi_i\}_{i\in I}$, which can actually be shown to also form a frame, the *so-called canonical dual frame*. In the case of a tight frame, the canonical dual frame is just a constant multiple of the original frame, which makes tightness a desirable property.

As for efficient expansions of a function $f \in \mathcal{H}$ in terms of Φ, although redundancy allows infinitely many coefficient sequences $(c_i)_{i\in I}$ such that

$$f = \sum_{i\in I} c_i \varphi_i,$$

we can identify with $(\langle f, \tilde{\varphi}_i \rangle)_{i\in I}$ one explicit coefficient sequence. However, this is typically by far not the 'best' possible in the sense of rapid decay in absolute value, since this sequence is only the smallest among all possible ones in ℓ^2-norm. Since one has better control over the sequence $(\langle f, \varphi_i \rangle)_{i\in I}$ of frame coefficients, it is often advantageous to instead consider the expansion

$$f = \sum_{i\in I} c_i \tilde{\varphi}_i$$

of f in terms of the canonical dual frame. The reason is that, if fast decay of the frame coefficients can be shown, then this form provides an efficient expansion of f.

2.2 Shearlet systems and Frame Properties

Shearlet systems are designed to asymptotically optimal encode geometric features. The key idea in shearlet systems are elements that are anisotropic and become more and more needlelike at fine scales. For this, we choose a *parabolic scaling matrix*, given by

$$A_j = \begin{bmatrix} 2^j & 0 \\ 0 & 2^{\frac{j}{2}} \end{bmatrix}, \quad j \in \mathbb{Z},$$

which ensures that the elements of the shearlet system have an essential support of size $2^{-j} \times 2^{-\frac{j}{2}}$ following the *parabolic scaling law* '$\mathit{width} \approx \mathit{length}^2$'. In fact, the choice of parabolic

scaling is oriented towards the fact that the governing feature of our model will be a piecewise C^2-discontinuity curve. In addition to translation, we now require a third parameter for changing the orientation of the shearlets. In contrast to curvelets [6] which are based on rotation, shearlets use *shearing matrices*

$$S_k = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}, \quad k \in \mathbb{Z},$$

which, by leaving the digital grid $\mathbb{Z}^2$ invariant, ensure the possibility of a faithful numerical realization. The formal definition of a shearlet system as it was defined in [20] is as follows.

Definition 2.1. *Let $\varphi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$, $c = [c_1, c_2]^T \in \mathbb{R}^2$ with $c_1, c_2 > 0$. Then the* (cone-adapted) shearlet system *is defined by*

$$\mathcal{SH}(\varphi, \psi, \tilde{\psi}, c) = \Phi(\varphi, c_1) \cup \Psi(\psi, c) \cup \tilde{\Psi}(\tilde{\psi}, c),$$

where

$$\begin{aligned}
\Phi(\psi, c_1) &= \left\{\varphi(\cdot - c_1 m) : m \in \mathbb{Z}^2\right\}, \\
\Psi(\psi, c) &= \left\{\psi_{j,k,m} = 2^{\frac{3j}{4}} \psi(S_k A_{2^j} \cdot - M_c m) : j \in \mathbb{N}_0, |k| \leq 2^{\lceil \frac{j}{2} \rceil}, m \in \mathbb{Z}^2\right\}, \\
\tilde{\Psi}(\tilde{\psi}, c) &= \left\{\tilde{\psi}_{j,k,m} = 2^{\frac{3j}{4}} \tilde{\psi}(S_k^T \tilde{A}_{2^j} \cdot - M_{\tilde{c}} m) : j \in \mathbb{N}_0, |k| \leq 2^{\lceil \frac{j}{2} \rceil}, m \in \mathbb{Z}^2\right\},
\end{aligned}$$

with $M_c := \begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix}$, $M_{\tilde{c}} = \begin{bmatrix} c_2 & 0 \\ 0 & c_1 \end{bmatrix}$, *and* $\tilde{A}_{2^j} = \operatorname{diag}(2^{\frac{j}{2}}, 2^j)$.

The following theorem shows that shearlets, and in particular compactly supported shearlets, form frames and gives theoretical estimates for the frame bounds.

Theorem 2.2 ([20])**.** *Let $\alpha > \gamma > 3$, and let $\varphi, \psi \in L^2(\mathbb{R}^2)$ be such that*

$$\begin{aligned}
|\widehat{\varphi}(\xi_1, \xi_2)| &\leq C_1 \min\{1, |\xi_1|^{-\gamma}\} \min\{1, |\xi_2|^{-\gamma}\} \text{ and} \\
|\widehat{\psi}(\xi_1, \xi_2)| &\leq C_2 \min\{1, |\xi_1|^{\alpha}\} \min\{1, |\xi_1|^{-\gamma}\} \min\{1, |\xi_2|^{-\gamma}\},
\end{aligned}$$

for some positive constants $C_1, C_2 < \infty$. Setting $\tilde{\psi}(x_1, x_2) := \psi(x_2, x_1)$, there exists a sampling vector $c = [c_1, c_2]^T \in \mathbb{R}^2$, $c_1, c_2 > 0$ such that the shearlet system $\mathcal{SH}(\varphi, \psi, \tilde{\psi}; c)$ forms a frame for $L^2(\mathbb{R}^2)$.

Faithful implementations of shearlet frames and the associated analysis operators are available at `www.ShearLab.org`, see also [25].

2.3 Sparse Approximation

Shearlets have well-analyzed approximation properties as can be seen, in particular, for cartoon-like functions as initially introduced in [14]. Denoting by $\chi_B \in L^2(\mathbb{R}^2)$ the characteristic function on a bounded, measurable set $B \subset \mathbb{R}^2$, we have the following definition.

Definition 2.3. *The class $\mathcal{E}^2(\mathbb{R}^2)$ of* cartoon-like functions *is the set of functions $f : \mathbb{R}^2 \to \mathbb{C}$ of the form*

$$f = f_0 + f_1 \chi_B,$$

where $B \subset [0,1]^2$ is a set with ∂B being a closed C^2-curve with bounded curvature and $f_i \in C^2(\mathbb{R}^2)$ are functions with support $\operatorname{supp} f_i \subset [0,1]^2$ *as well as $\|f_i\|_{C^2} \leq 1$ for $i = 0, 1$.*

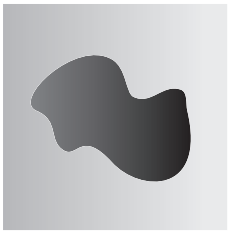

Figure 1: Example of a cartoon-like function.

For an illustration of a cartoon-like functions, see Figure 1. We measure the approximation quality of shearlets with respect to the cartoon model by the decay of the L^2-error of best N-term approximation. Recall that for a general representation system $\{\psi_i\}_{i\in I} \subset \mathcal{H}$ and $f \in \mathcal{H}$, the *best N-term approximation* is defined as

$$f_N = \underset{\substack{\Lambda\subset\mathbb{N},|\Lambda|=N,\\ \tilde{f}_N=\sum_{i\in\Lambda} c_i\psi_i}}{\arg\min} \|f - \tilde{f}_N\|.$$

In contrast to the situation of orthonormal bases, if $\{\psi_i\}_{i\in I}$ forms a frame or even a tight frame, it is not clear at all how the set Λ has to be chosen. Therefore, often the *best N-term approximation* is substituted by the N-term approximation using the N largest coefficients.

To be able to claim *asymptotic optimality* of a sparse approximation, one requires a benchmark result. In [14] it was shown that for an arbitrary representation system $\{\psi_i\}_{i\in I} \subset L^2(\mathbb{R}^2)$, the minimally achievable asymptotic approximation error for $f \in \mathcal{E}^2(\mathbb{R}^2)$ is

$$\|f - f_N\|_2^2 = O(N^{-2}) \quad \text{as } N \to \infty,$$

provided that only polynomial depth search is used to compute the approximation. Here, for a function f, the Landau symbol $O(f(a))$ describes the asymptotic convergence behavior as $a \to 0$ for the set of functions g such that $\limsup_{x\to a} \frac{g(x)}{f(x)} < \infty$.

Shearlets achieve this asymptotically optimal rate up to a log-factor as the following result shows.

Theorem 2.4 ([23])**.** *Let $\varphi, \psi, \tilde{\psi} \subset L^2(\mathbb{R}^2)$ be compactly supported, and assume that the shearlet system $\mathcal{SH}(\varphi, \psi, \tilde{\psi}, c)$ forms a frame for $L^2(\mathbb{R}^2)$. Furthermore, assume that, for all $\xi = [\xi_1, \xi_2]^T \in \mathbb{R}^2$, the function ψ satisfies*

$$\begin{aligned} |\widehat{\psi}(\xi)| &\leq C \min\{1, |\xi_1|^\delta\} \min\{1, |\xi_1|^{-\gamma}\} \min\{1, |\xi_2|^{-\gamma}\}, \\ \left|\frac{\partial}{\partial\xi_2}\widehat{\psi}(\xi)\right| &\leq |h(\xi_1)| \left(1 + \frac{|\xi_2|}{|\xi_1|}\right)^{-\gamma}, \end{aligned}$$

where $\delta > 6$, $\gamma \geq 3$, $h \in L^1(\mathbb{R})$ and C is a constant, and $\tilde{\psi}$ satisfies analogous conditions with the roles of ξ_1 and ξ_2 exchanged. Then $\mathcal{SH}(\varphi, \psi, \tilde{\psi}, c)$ provides an asymptotically optimal sparse approximation of $f \in \mathcal{E}^2(\mathbb{R}^2)$, i.e.,

$$\|f - f_N\|_2^2 = O(N^{-2} \cdot (\log N)^3) \quad \text{as } N \to \infty.$$

Theorem 2.4 indicates that shearlet systems provide a very good model for encoding the governing features of a scatterer.

3 The Acoustic Scattering Problem

In this section we focus on the first of the two scattering problems, which is the acoustic scattering problem. After briefly discussing the direct problem, we introduce the related inverse problem, which we approach using a Tikhonov type functional with regularization by ℓ_p minimization, with p close to or equal to 1, applied to the shearlet coefficients. For this setting, we will derive convergence results, which in fact even hold for a larger class of frames.

3.1 The Direct Problem

A very common model for the behavior of an acoustic wave $u : \mathbb{R}^2 \to \mathbb{C}$ in an inhomogeneous medium is the Helmholtz equation [12]. Given a *wave number* $k_0 > 0$ and a compactly supported *contrast function* $f \in L^2(\mathbb{R}^2)$, the Helmholtz equation has the form

$$\Delta u + k_0^2(1 - f)u = 0, \tag{1}$$

where the contrast function f models the inhomogeneity of the medium due the scatterer. In a typical situation one models f as a function which is smooth, apart from a model of the scatterer which is again assumed to be an essentially homogeneous medium, whose density is, however, significantly different from the surrounding medium. To model the boundary of the scatterer, a typical approach is to use a curve with a particular regularity, say C^2. Recalling the definition of a cartoon-like function in Definition 2.3, we suggest to use $\mathcal{E}^2(\mathbb{R}^2)$ as a model for the boundary. Furthermore, even if the boundary of the scatterer is only a piecewise C^2 curve, then the shearlets provide a very good model, since the sparse approximation results of shearlets as well as our analysis also hold in this more general situation. As further ingredient for the acoustic scattering problem, we introduce *incident waves* u^{inc}, which are solutions to the *homogenous Helmholtz equation*, i.e., (1) with $f \equiv 0$. A large class of such solutions take the form

$$u_d^{\text{inc}} : \mathbb{R}^2 \to \mathbb{C}, \quad u_d^{\text{inc}}(x) = e^{ik_0\langle x,d\rangle} \tag{2}$$

for some direction $d \in \mathbb{S}^1$. Then, for a given $f \in L^2(\mathbb{R}^2)$ and a solution u^{inc} to the homogeneous Helmholtz equation, every solution to (1) can be expressed as

$$u = u^s + u^{\text{inc}},$$

where u^s denotes the *scattered wave*. To obtain physically reasonable solutions we stipulate that the scattered wave obeys the *Sommerfeld radiation condition*, see e.g. [12],

$$\frac{\partial u^s}{\partial |x|} = ik_0 u^s(x) + O(|x|^{-\frac{1}{2}}) \text{ for } |x| \to \infty.$$

For a given $k_0 > 0$, and contrast function $f \in L^2(\mathbb{R}^2)$ with compact support and incident wave u^{inc}, the *acoustic scattering problem* then is to find $u \in H^2_{loc}(\mathbb{R}^2)$ such that

$$\begin{aligned} \Delta u + k_0^2(1 - f)u &= 0, \\ u &= u^s + u^{\text{inc}}, \\ \frac{\partial u^s}{\partial |x|} &= ik_0 u^s(x) + O(|x|^{-\frac{1}{2}}). \end{aligned}$$

To obtain an equivalent formulation, we introduce the *fundamental solution* G_{k_0} *to the Helmholtz equation*,

$$G_t(x,y) = \frac{i}{4} H_0^{(1)}(t|x-y|), \quad t > 0, x, y \in \mathbb{R}^2, \tag{3}$$

where $H_0^{(1)}$ is a *Hankel function*, see e.g. [1]. Letting B_R denoting the open ball of radius $R > 0$ centered at 0 and R chosen such that supp $f \subset B_R$, the *volume potential* is defined by

$$V(f)(x) := \int_{B_R} G_{k_0}(x,y) f(y)\, dy, \quad x \in \mathbb{R}^2.$$

Using this potential we can reformulate the acoustic scattering problem as the solution of the *Lippmann-Schwinger integral equation* given by

$$u^s(x) = -k_0^2 V(f(u^s - u^{\text{inc}})) \text{ in } B_R \tag{4}$$

for $f \in L^2(\mathbb{R}^2)$ with supp $f \subset B_R$. Any solution $u^s \in H^2_{loc}(B_R)$ of (4) indeed solves the acoustic scattering problem in B_R and can, by the unique continuation principle [18], be uniquely extended to a global solution of the acoustic scattering problem, see e.g. [12].

Letting $L^2(B_R)$ denote the square-integrable functions defined on B_R, which are in particular compactly supported, we now define the *solution operator* of the acoustic scattering problem by

$$\mathcal{S} : L^2(B_R) \times L^2(B_R) \to H^2_{loc}(B_R), \quad \mathcal{S}(f, u^{\text{inc}}) = u,$$

and the Lippmann-Schwinger equation (4) allows to compute this operator for a given scatterer f and incident wave u^{inc}.

3.2 The Inverse Problem

In the associated inverse problem, we assume that we know the incident wave u^{inc} as well as measurements of the scattered wave u^s and we aim to compute information about the scatterer f. Following [26], we model these measurements as $u^s|_{\Gamma_{\text{meas}}}$, where Γ_{meas} is the trace of a closed locally Lipschitz continuous curve with $\Gamma_{\text{meas}} \cap \overline{B_R} = \emptyset$.

In the case that we just have one incident wave u^{inc}, then the map $(f, u^{\text{inc}}) \mapsto u^s|_{\Gamma_{\text{meas}}}$, is called *mono-static contrast-to-measurement operator* in [26]. For multiple incident waves, and multi-static measurements, a closed set Γ_{inc} is introduced, which is again the trace of a closed locally Lipschitz curve enclosing B_R, such that $\Gamma_{\text{inc}} \cap \overline{B_R} = \emptyset$. The set Γ_{inc} serves to construct *single layer potentials*, which take the role of the incident waves. For $\varphi \in L^2(\Gamma_{\text{inc}})$, these single layer potentials are

$$SL_{\Gamma_{\text{inc}}}\varphi := \int_{\Gamma_{\text{inc}}} G_{k_0}(\cdot, y)\varphi(y)\, dy \in L^2(B_R),$$

see Figure 2 for an illustration. Let $L^p_{\text{Im}\geq 0}(B_R)$ denote the set of $L^p(B_R)$-functions with nonnegative imaginary part, and let $\text{HS}(\cdot,\cdot)$ denote the space of Hilbert Schmidt operators [35]. Then the *multi-static measurement operator* $\mathcal{N}$, which assigns each contrast function a Hilbert-Schmidt operator, maps a single layer potential to the associated solution of the acoustic scattering problem. Formally, $\mathcal{N}$ is defined by

$$\mathcal{N} : \ L^2_{\text{Im}\geq 0}(B_R) \to \text{HS}(L^2(\Gamma_{\text{inc}}), L^2(\Gamma_{\text{meas}})), \quad f \mapsto N_f,$$

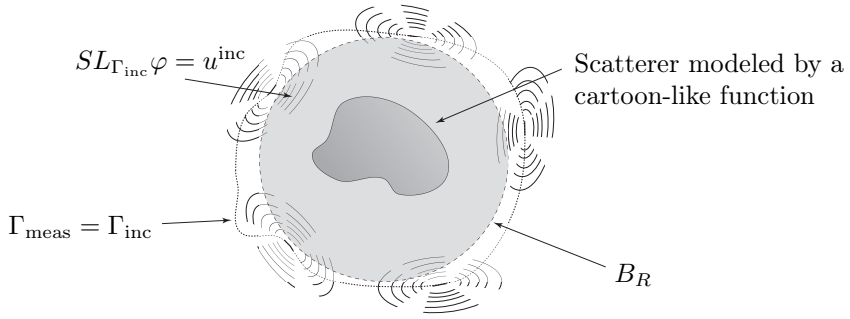

Figure 2: Model for the acoustic inverse scattering problem in which $\Gamma_{\text{meas}} = \Gamma_{\text{inc}}$.

where

$$N_f : L^2(\Gamma_{\text{inc}}) \to L^2(\Gamma_{\text{meas}}), \quad \varphi \mapsto \mathcal{S}(f, SL_{\Gamma_{\text{inc}}}\varphi).$$

Note that indeed $N_f \in \text{HS}(L^2(\Gamma_{\text{inc}}), L^2(\Gamma_{\text{meas}}))$, see [26], where it was also shown, even in a more general setting, that the operator $\mathcal{N}$ satisfies the following properties.

Theorem 3.1. *[26] The operator $\mathcal{N}$ is continuous, compact, and weakly sequentially closed from $L^2_{\text{Im}\geq 0}(B_R)$ into $\text{HS}(L^2(\Gamma_{\text{inc}}), L^2(\Gamma_{\text{meas}}))$.*

Since in realistic applications the signals always contain noise, we consider the *inverse acoustic scattering problem* with noise, which for a noise level $\varepsilon > 0$, and noisy measurements $N^\varepsilon_{\text{meas}} \in \text{HS}(L^2(\Gamma_{\text{inc}}), L^2(\Gamma_{\text{meas}}))$ satisfying

$$\|N^\varepsilon_{\text{meas}} - N_{f^\dagger}\|_{\text{HS}(L^2(\Gamma_{\text{inc}}),L^2(\Gamma_{\text{meas}}))} \leq \varepsilon, \tag{5}$$

is to recover the scatterer $f^\dagger$. This inverse problem is ill-posed and requires careful regularization which is discussed in the next subsection.

3.3 Regularization by Frames

A classical regularization approach to solve inverse problems is to minimize an appropriate Tikhonov functional. In [26] it was suggested to minimize

$$\tilde{\mathfrak{T}}^\varepsilon_\alpha(f) := \frac{1}{2}\|N_f - N^\varepsilon_{\text{meas}}\|^2_{\text{HS}(L^2(\Gamma_{\text{inc}}),L^2(\Gamma_{\text{meas}}))} + \frac{\alpha}{p}\|f\|^p_{L^p(B_R)}, \quad f \in L^p(B_R). \tag{6}$$

for fixed $p > 1$ and $\alpha > 0$. For p close to 1, the regularization term in this functional promotes sparsity in the representation of the scatterer.

Here we suggest a different regularization, which exploits that the scatterer f is modeled as a cartoon-like function $\mathcal{E}^2(\mathbb{R}^2)$ and that shearlet systems are used to obtain sparse approximations of the scatterer f. Let $\Phi := \mathcal{SH}(\varphi, \psi, \tilde{\psi}, c)$ be a shearlet frame satisfying the hypotheses of Theorem 2.4, and let T_Φ denote the analysis operator of the shearlet frame Φ, then the proof of Theorem 2.4 yields the decay behavior of the associated shearlet coefficients $T_\Phi(f)$. It has been shown in [26] that $T_\Phi(f)$ is in ℓ^p for every $p > \frac{2}{3}$.

We propose to regularize the acoustic inverse scattering problem by adapting the data fidelity term appropriately and by imposing a constraint on the ℓ^p-norm of the coefficient sequence

$T_\Phi(f)$. More precisely, for fixed $1 \le p \le 2$ and $\alpha > 0$, we consider the Tikhonov functional given by

$$\mathfrak{T}_\alpha^\varepsilon(f) := \frac{1}{2}\,\|\mathcal{N}(f) - N_{\text{meas}}^\varepsilon\|^2_{\text{HS}(L^2(\Gamma_{\text{inc}}),L^2(\Gamma_{\text{meas}}))} + \frac{\alpha}{p}\|T_\Phi(f)\|_{\ell^p}^p, \quad f \in L^2(B_R). \tag{7}$$

Note that the case of $p = 1$ is not excluded in our analysis in contrast to the situation in [26].

Having introduced the Tikhonov regularization, we now analyze the convergence of the minimization process. For this we are particularly interested in convergence to a *norm minimizing solution* $f^* \in L^2(B_R)$, i.e.,

$$\mathcal{N}(f^*) = N_f \quad \text{and} \quad \|T_\Phi(f^*)\|_p \le \|T_\Phi(f)\|_p \text{ for all } f \text{ such that } \mathcal{N}(f) = N_{f^\dagger}.$$

Such convergence properties have been extensively studied. Let us recall the following result of [33, Thm. 3.48], which was stated for orthonormal bases but the extension to frames is straightforward.

Theorem 3.2. *[33] Let $1 \le p, q < 2$, let U be a Hilbert space with frame $\Phi = (\varphi_i)_{i\in I}$, and let V be a Banach space. Let τ_V and τ_U be the weak topologies on V and U, respectively. Let $w = (w_i)_{i\in I}$ be a sequence of weights with $0 < w_{\min} \le w_i < \infty$ for all $i \in I$ and some constant $w_{\min}$. Let $\mathcal{R} : U \to [0,\infty]$ be defined as $\mathcal{R}(u) = \|wT_\Phi(u)\|_{\ell_p}^p$. Furthermore, let $\mathcal{D} := \mathcal{D}(F) \cap \mathcal{D}(\mathcal{R}) \neq \emptyset$ and assume that the operator $F : \mathcal{D}(F) \subset U \to V$ is weakly continuous and its domain is weakly sequentially closed.*

Suppose that (8) possesses a solution in $\mathcal{D}$, and that $\alpha : (0,\infty) \to (0,\infty)$ is a function which satisfies

$$\alpha(x) \to 0 \quad \textit{and} \quad \frac{x^q}{\alpha(x)} \to 0 \quad \textit{as } x \to 0.$$

Let $(\delta_n)_n \subseteq \mathbb{R}^+$ be a sequence of noise levels converging to 0 as $n \to \infty$, let $(v_n)_n \subseteq V$ be a sequence of noisy measurements deviating from the noise-free measurement v by at most $(\delta_n)_n$, i.e.,

$$\|v - v_n\|_V \le \delta_n,$$

and set $\alpha_n := \alpha(\delta_n)$, as regularization parameter. Then, for every sequence $(u_n)_n \subseteq U$ of elements minimizing $\mathfrak{T}_{\alpha_n,v_n}$, there exists a subsequence $(u_{n_j})_j$ of $(u_n)_n$ and a norm minimizing solution $u^* \in U$ such that

$$u_{n_j} \to u^* \quad \textit{as } j \to \infty, \textit{ with respect to } \tau_U.$$

Theorem 3.2 provides conditions for convergence with a general Tikhonov type functional

$$\mathfrak{T}_{\alpha,v}(u) := \|F(u) - v\|_V^q + \alpha\mathcal{R}(u), \qquad u \in U, \tag{8}$$

where $1 \le q < \infty$, $\alpha > 0$, U is a Hilbert space and V is a Banach space. Based on Theorem 3.2, using weights $w_i = 1$, we obtain a convergence result for the functional $\mathfrak{T}_\alpha^\varepsilon$ introduced in (7).

Theorem 3.3. *Consider the functional (7), let $1 \le p \le 2$ and $f^\dagger \in L^2_{\text{Im}\ge 0}(B_R)$. Furthermore, let $(\varepsilon_n)_n \subseteq \mathbb{R}^+$ be a sequence of noise levels converging to 0 as $n \to \infty$, and let $(N_{\text{meas}}^{\varepsilon_n})_{n\in\mathbb{N}} \subset \text{HS}(L^2(\Gamma_{\text{inc}}), L^2(\Gamma_{\text{meas}}))$ be an associated family of noisy measurements that obey (5). Finally, let $(\alpha_n)_n$ be a sequence of regularization parameters satisfying*

$$\alpha_n \to 0 \quad \textit{and} \quad \frac{\varepsilon_n^2}{\alpha_n} \to 0 \quad \textit{as } n \to \infty.$$

Then, for every n, there exists a minimizer $f^n \in L^2(B_R)$ of the functional $\mathfrak{T}_{\alpha_n}^{\varepsilon_n}$ from (7). In addition, for every such sequence $(f^n)_n \subset L^2(B_R)$, there exists a subsequence $(f^{n_l})_l$ and a norm minimizing solution $f^* \in L^2(B_R)$ such that

$$\|f^{n_l} - f^*\|_{L^2} \to 0 \quad \text{as } l \to \infty.$$

Proof. The proof is organized as follows. We first show that the assumptions of Theorem 3.2 are fulfilled, which yields weak convergence of subsequences $(f^{n_l})_l$ to a norm minimizing solution f^*. Afterwards we prove that in our setup we can even get strong convergence.

To show weak convergence, we apply Theorem 3.2 with $q = 2$, $V = \mathrm{HS}(L^2(\Gamma_{\mathrm{inc}}), L^2(\Gamma_{\mathrm{meas}}))$ and $U = L^2(B_R)$, and we choose τ_U, τ_V to be the respective weak topologies. Moreover, we use

$$\mathcal{D}(\mathcal{N}) \cap \mathcal{D}(R) = L^2(B_R) \cap T_\Phi^{-1}(\ell^p) = T_\Phi^{-1}(\ell^p) \neq \emptyset.$$

The continuity, compactness, and weakly sequentially closedness of $\mathcal{N}$ implies that $\mathcal{N}$ is weakly continuous. Theorem 3.2 then guarantees, that, for every n, there exists a minimizer $f^n \in L^2(B_R)$ of the functional $\mathfrak{T}_{\alpha_n}^{\varepsilon_n}$. In addition, for every such sequence $(f^n)_n \subset L^2(B_R)$, there exists a subsequence $(f^{n_l})_l$ that converges weakly to a norm minimizing solution $f^* \in L^2(B_R)$.

To prove strong convergence, we may assume w.l.o.g., that $(f^n)_n$ is the weakly converging subsequence and we aim to prove that $\|f^n - f^*\|_{L^2} \to 0$ for $n \to \infty$. Since f^n converges weakly to f^*, we also obtain that, for all $i \in I$, $\langle f^n - f^*, \varphi_i \rangle \to 0$ as $n \to \infty$. Furthermore, since f^* is a norm minimizing solution, we have that $\|T_\Phi f^*\|_p < \infty$ and, since $\|T_\Phi f^n\|_p^p \leq \mathfrak{T}_{\alpha_n}^{\varepsilon_n}(f^n) \leq \mathfrak{T}_{\alpha_n}^{\varepsilon_n}(f^*)$, we obtain that

$$\frac{\alpha}{p}\|T_\Phi f^n\|_p^p \leq \mathfrak{T}_{\alpha_n}^{\varepsilon_n}(f^*) = \frac{1}{2}\left\|\mathcal{N}(f^*) - N_{\mathrm{meas}}^{\varepsilon_n}\right\|_{\mathrm{HS}(L^2(\Gamma_{\mathrm{inc}}),L^2(\Gamma_{\mathrm{meas}}))}^2 + \frac{\alpha}{p}\|T_\Phi(f^*)\|_{\ell^p}^p.$$

Employing that $\|\mathcal{N}(f^*) - N_{\mathrm{meas}}^{\varepsilon_n}\|_{\mathrm{HS}(L^2(\Gamma_{\mathrm{inc}}),L^2(\Gamma_{\mathrm{meas}}))} \leq \varepsilon_n$ and $\frac{\varepsilon_n^2}{\alpha} \to 0$ it follows that $\|T_\Phi f^n\|_p^p$ is uniformly bounded in n. By the frame inequality and the fact that $p \leq 2$, we obtain that there exists constants $0 < C_1, C_2 < \infty$ such that

$$\|f^n - f^*\|_2 \leq C_1 \left(\sum_{i\in I} |\langle f^n - f^*, \varphi_i\rangle|^2\right)^{\frac{1}{2}} \leq C_2 \left(\sum_{i\in I} |\langle f^n - f^*, \varphi_i\rangle|^p\right)^{\frac{1}{p}}.$$

Using the boundedness of $\|T_\Phi f^n\|_p$ and $\|T_\Phi f^*\|_p$ and the dominated convergence theorem yields that for $n \to \infty$ the right hand side of the above expression vanishes. The strong convergence and hence the assertion follow. □

Note that Theorem 3.3 holds in more generality for arbitrary frames for $L^2(\mathbb{R}^2)$ with certain additional properties.

4 The Electromagnetic Scattering Problem

The second inverse problem that we consider is for *electromagnetic scattering*. After introducing the inverse problem, our goal will be to provide a theoretical basis for the application of shearlet frames. As before, we base our considerations on the premise that edges, i.e., curve-like singularities, are the governing features of the scatterer leading to the cartoon model as appropriate choice.

Once we step away from the nonlinear situation and introduce a linearization, then this argument may, however, not be valid anymore. It could be possible, that linearizing the inverse problem introduces a smoothing that erases all edge-like structures. Fortunately, it has been shown, see, e.g., [29, 34], that when using a linearization via the inverse Born approximation, certain singularities of the scatterer prevail. This gives a first indication, that methods involving shearlets can play a role again in a regularization of the inverse electromagnetic scattering problem. However, all results on the regularity of the inverse Born approximation in the literature describe the global regularity in the sense of weak differentiability. In the case of cartoon-like functions we though have strong local, but poor global regularity. To be able to exploit shearlets in the context of this problem, in this section, we will prove a local regularity result for the inverse Born approximation.

4.1 The Inverse Problem

The *magnetic Schrödinger equation*, for $f \in L^2(\mathbb{R}^2)$, a wave number $k > 0$, and with $u = u^s + u_d^{\text{inc}}$ as in (2), is given by

$$\begin{aligned} \Delta u + (f + k^2)u &= 0, \\ \lim_{r\to\infty} \left(\frac{\partial u^s(x)}{\partial r} - iku^s(x) \right) &= 0. \end{aligned}$$

For $\theta \in [0, \pi]$ and $\tau_\theta = (\cos(\theta), \sin(\theta))$, the associated *backscattering amplitude* is defined by

$$A(k, -\theta, \theta) = \int_{\mathbb{R}^2} e^{ik\langle \tau_\theta, y\rangle} f(y)u(y)\, dy,$$

and the inverse problem is to reconstruct the potential f from A.

Using $|\xi|\tau_\theta := \xi$ to denote polar coordinates, the *Born approximation* of the solution f is defined as the inverse Fourier transform of the function $\xi \mapsto A(k, -\theta, \theta)$, given by

$$f_B(x) = \int_{\mathbb{R}^2} e^{-i\langle \xi, x\rangle} A(\frac{|\xi|}{2}, \theta, -\theta)\, d\xi.$$

Applying the Lippmann-Schwinger integral equation (4) iteratively as in [29], yields that

$$f_B(x) = \sum_{j=1}^{m} q_j(x) + q_{m+1}^R(x), \tag{9}$$

with

$$\begin{aligned} q_j(x) &= \frac{1}{4\pi^2} \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} e^{i\langle \xi, x+\frac{y}{2}\rangle} f(y)(\mathcal{G}_{|\xi|})^j(e^{i\langle \xi, \cdot\rangle})(y)\, dy\, d\xi, \quad 1 \le j \le m, \\ q_{m+1}^R(x) &= \frac{1}{4\pi^2} \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} e^{i\langle \xi, x+\frac{y}{2}\rangle} f(y)(\mathcal{G}_{|\xi|})^{m+1}(u(\cdot, |\xi|, \hat{\xi}))(y)\, dy\, d\xi, \end{aligned} \tag{10}$$

where $\mathcal{G}_{|\xi|} : L^2_{loc}(\mathbb{R}^2) \to L^2_{loc}(\mathbb{R}^2)$ is the integral operator with kernel $G_{\frac{|\xi|}{2}}(x, y)f(y)$ as defined in (3).

4.2 Local Regularity of the inverse Born Approximation

To determine the local regularity of the Born approximation f_B from a given potential f, we invoke the representation (9) and observe that we can equally well examine the regularity of the functions $q_1, \dots, q_m$, and q_{m+1}^R. To analyze the regularity of the q_j's, we will make use of the *Radon transform* and the *Projection Slice Theorem*, see e.g. [27].

Definition 4.1. *Let $\theta \in [0, \pi)$. Then the* Radon transform $\mathcal{R}_\theta$ *of a function $f \in L^2(\mathbb{R}^2)$ along a ray $\Delta_{t,\theta} = \{x \in \mathbb{R}^2 : x_1 \cos(\theta) + x_2 \sin(\theta) = t\}$ for $t \in \mathbb{R}$ is defined as*

$$\mathcal{R}_\theta f(t) := \int_{\Delta_{t,\theta}} f(x)\, ds = \int\int f(x)\delta_0(\langle x, \tau_\theta\rangle - t)\, dx.$$

Theorem 4.2 (Projection Slice Theorem)**.** *Let $f \in L^2(\mathbb{R}^2)$. Then, for all $\theta \in [0, \pi)$ and $\xi \in \mathbb{R}$,*

$$\widehat{\mathcal{R}_\theta f}(\xi) = \widehat{f}(\xi\cos(\theta), \xi\sin(\theta)).$$

Here $\widehat{f}$ denotes the Fourier transform.

Before we state and prove the local regularity of the functions q_j, $j = 1, \dots, m$, we fix some notion. We will denote the Sobolev spaces of functions with s weak derivatives in $L^2(\mathbb{R}^2)$ by $H^s(\mathbb{R}^2)$ and the functions that are locally in $H^s(\mathbb{R}^2)$ by $H^s_{loc}(\mathbb{R}^2)$. Furthermore, $H^s(x)$ is the local Sobolev space of s times weakly differentiable functions with weak derivatives in $L^2(x)$, and $L^2(x)$ the space of distributions that are L^2 on a neighborhood of x, see [5]. Finally, we denote by $C^{k,\alpha}(\mathbb{R}^2)$ the k times differentiable functions with a Hölder continuous k-th derivative with Hölder coefficient α, writing $C^k(\mathbb{R}^2)$ if the k-th derivative is simply continuous.

Then we have the following regularity result.

Theorem 4.3. *Let $\varepsilon > 0$, let $s \in \mathbb{N}, s \geq 2$, and let, for some $x_0 \in \mathbb{R}^2$, $f \in L^2(\mathbb{R}^2) \cap H^{s+\varepsilon}(x_0)$ be compactly supported and real valued. Then the q_j defined in (10) satisfy $q_j \in H^s(x_0)$ for all $j \in \mathbb{N}$, and, in particular, $f_B \in H^s(x_0)$.*

Proof. We start by proving the local regularity of q_1. Let $x_0 \in \mathbb{R}^2$ be such that $f \in H^s_{\text{loc}}(x_0)$, then we aim to prove that

$$q_1(x) = \frac{1}{4\pi^2} \int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2} e^{i\left\langle \xi, x+\frac{y}{2}+\frac{z}{2}\right\rangle} f(y)\mathcal{G}_{|\xi|}(y,z)f(z)\, dz\, dy\, d\xi \in H^s(x_0).$$

For this, we introduce a smooth cutoff function φ supported in a neighborhood U_{x_0} of x_0, where f is $H^{s+\varepsilon}$ such that $\varphi \equiv 1$ on a strictly smaller neighborhood of x_0. For $x \in U_{x_0}$, the function φ is now used to decompose q_1 as

$$\begin{aligned} q_1(x) &= \frac{1}{4\pi^2} \int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2} e^{i\left\langle \xi, x+\frac{y}{2}+\frac{z}{2}\right\rangle} \varphi(y) f(y)\mathcal{G}_{|\xi|}(y,z)f(z)\, dz\, dy\, d\xi \\ &\quad + \frac{1}{4\pi^2} \int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2} e^{i\left\langle \xi, x+\frac{y}{2}+\frac{z}{2}\right\rangle} (1-\varphi)(y) f(y)\mathcal{G}_{|\xi|}(y,z)f(z)\, dz\, dy\, d\xi \\ &=: \mathcal{I}_1(x) + \mathcal{I}_2(x). \end{aligned}$$

We further decompose $\mathcal{I}_2$ as

$$\begin{aligned}\mathcal{I}_2 &= \frac{1}{4\pi^2}\int\limits_{|\xi|\geq 1}\int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^2} e^{i\left\langle \xi, x+\frac{y}{2}+\frac{z}{2}\right\rangle}(1-\varphi)(y)f(y)\mathcal{G}_{|\xi|}(y,z)\varphi(z)f(z)\,dz\,dy\,d\xi \\ &\quad + \frac{1}{4\pi^2}\int\limits_{|\xi|\geq 1}\int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^2} e^{i\left\langle \xi, x+\frac{y}{2}+\frac{z}{2}\right\rangle}(1-\varphi)(y)f(y)\mathcal{G}_{|\xi|}(y,z)(1-\varphi)(z)f(z)\,dz\,dy\,d\xi \\ &=: \mathcal{I}_{2,1}(x)+\mathcal{I}_{2,2}(x).\end{aligned}$$

and study each of the integrals $\mathcal{I}_1$, $\mathcal{I}_{2,1}$, and $\mathcal{I}_{2,2}$ separately.

Regularity of $\mathcal{I}_1$: We use a representation from [29, Lem. 1.1], which yields, for a function $\nu_2 \in C^\infty(\mathbb{R}^2)$ and $x \in U_{x_0}$,

$$\begin{aligned}\mathcal{I}_1(x) &= \int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^2} \frac{\mathcal{F}(f)(\xi)\mathcal{F}(\varphi f)(\eta)}{\langle \eta,\xi\rangle - i0} e^{i\langle x,\xi+\eta\rangle}\,d\eta\,d\xi \\ &= \int\limits_{|\xi|\geq 1}\int\limits_{\mathbb{R}^2} \frac{\mathcal{F}(f)(\xi)\mathcal{F}(\varphi f)(\eta)}{\langle \eta,\xi\rangle - i0} e^{i\langle x,\xi+\eta\rangle}\,d\eta\,d\xi + \nu_2(x), \qquad (11)\end{aligned}$$

where

$$(\langle \xi,\eta\rangle - i0)^{-1} = \mathsf{p.v.}(\langle \xi,\eta\rangle)^{-1} - \pi i\delta_0(\langle \xi,\eta\rangle), \qquad (12)$$

with $\mathsf{p.v.}$ denoting the Cauchy principal value as stated in [29, Lem. 1.2]. Furthermore, we omit the small values of ξ in the integration, since they only contribute to the smooth part of $\mathcal{I}_1$ in (11). To show the local regularity of (11) as a function of x, we introduce the symbol

$$a(x,\xi) := \int\limits_{\mathbb{R}^2} \frac{\mathcal{F}(\varphi f)(\eta)}{\langle \eta,\xi\rangle - i0} e^{i\langle x,\eta\rangle} d\eta, \quad (x,\xi)\in\mathbb{R}^2\times\mathbb{R}^2.$$

It follows from [5, Thm. 1.3], that if the function

$$g_\xi : U_{x_0} \to \mathbb{C}, \quad g_\xi(x) := a(x,\xi), \quad \xi\in\mathbb{R}^2,$$

is in $H^s(x)$ for some $s\in\mathbb{N}$, $s>1$ and if $\|g_\xi\|_{H^s(x_0)}$ is independent of ξ, then the corresponding pseudo-differential operator $a(x,D)$ is a bounded operator from $H^s(x_0)$ to $H^s(x_0)$. By definition, we have

$$a(x,D)(f) = \mathcal{I}_1(x) - \nu_2(x). \qquad (13)$$

Thus it remains to prove that, for any $|\xi|\geq 1$, we have $g_\xi \in H^s(x_0)$ and $\|g_\xi\|_{H^s(x_0)}$ is independent of ξ. By (12), we obtain

$$a(x,\xi) = \mathsf{p.v.}\int\limits_{\mathbb{R}^2} \frac{\mathcal{F}(\varphi f)(\eta)}{\langle \eta,\xi\rangle} e^{i\langle x,\eta\rangle}\,d\eta - \pi i \int\limits_{\langle \eta,\xi\rangle = 0} \mathcal{F}(\varphi f)(\eta) e^{i\langle x,\eta\rangle}\,d\eta. \qquad (14)$$

Since $\varphi f \in H^{s+\varepsilon}(\mathbb{R}^2)$ is compactly supported, it follows that $\mathcal{F}(\varphi f)\in C^\infty(\mathbb{R}^2)$ and

$$\int_{\mathbb{R}^2} (1+|\xi|^2)^{s+\varepsilon}|\mathcal{F}(\varphi f)(\eta)|^2 d\eta < \infty.$$

Passing to polar coordinates yields

$$\int_0^{2\pi} \int_{\mathbb{R}^+} r(1+|r|^2)^{s+\varepsilon} |\mathcal{F}(\varphi f)(r\tau_\theta)|^2 \, dr \, d\theta < \infty,$$

and by the smoothness of

$$\theta \mapsto \int_{\mathbb{R}} r(1+|r|^2)^{s+\varepsilon} |\mathcal{F}(\varphi f)(r\tau_\theta)|^2 dr, \tag{15}$$

the terms $\int_{\mathbb{R}} |r|(1+|r|^2)^{s+\varepsilon} |\mathcal{F}(\varphi f)(r\tau_\theta)|^2 dr$ are uniformly bounded with respect to θ. This implies the desired regularity of the second term of (14) as a function of x.

We continue with the first term of (14), i.e., with

$$\mathsf{p.v.} \int_{\mathbb{R}^2} \frac{\mathcal{F}(\varphi f)(\eta)}{\langle \eta, \xi \rangle} e^{i\langle x, \eta \rangle} \, d\eta = \mathsf{p.v.} \int_0^\pi \frac{1}{\langle \tau_\theta, \xi \rangle} \int_{\mathbb{R}} \mathcal{F}(\varphi f)(r\tau_\theta) e^{ir\langle x, \tau_\theta \rangle} \, dr \, d\theta$$

By substitution, w.l.o.g. we may assume that $\xi = (1, 0)$. Application of Theorem 4.2 shows that (4.2) equals

$$\mathsf{p.v.} \int_0^\pi \frac{1}{\langle \tau_\theta, \xi \rangle} \mathcal{R}_\theta(\varphi f)(\langle x, \tau_\theta \rangle) \, d\theta = \lim_{\varepsilon \downarrow 0} \int_{[0, \frac{\pi}{2} - \varepsilon] \cup [\frac{\pi}{2} + \varepsilon, \pi]} \frac{1}{\cos(\theta)} \mathcal{R}_\theta(\varphi f)(\langle \cdot, \tau_\theta \rangle) \, d\theta.$$

For a given sequence $\varepsilon_n \to 0$ as $n \to \infty$, we set $E_n := [0, \frac{\pi}{2} - \varepsilon_n] \cup [\frac{\pi}{2} + \varepsilon_n, \pi]$ and show that the sequence

$$\left(\int_{E_n} \frac{1}{\cos(\theta)} \mathcal{R}_\theta(\varphi f)(\langle \cdot, \tau_\theta \rangle) \, d\theta \right)_{n \in \mathbb{N}} \tag{16}$$

is a Cauchy sequence in $H^s(x_0)$.

We first observe that by Theorem 4.2, the finiteness of the integral in (15) implies that

$$\xi \mapsto |\xi|^{s+\frac{1}{2}+\varepsilon} \cdot \widehat{\mathcal{R}_\theta(\varphi f)}(\xi) \in L^2(\mathbb{R}).$$

This yields $\mathcal{R}_\theta(\varphi f) \in H^{s+\frac{1}{2}+\varepsilon}(\mathbb{R})$. By the Sobolev embedding theorem, [2, Thm 5.4] we have $H^{s+\frac{1}{2}+\varepsilon}(\mathbb{R}) \hookrightarrow C^{s,\varepsilon}(\mathbb{R})$, which implies that $\mathcal{R}_\theta(\varphi f)(\langle \cdot, \tau_\theta \rangle) \in C^{s,\varepsilon}(\mathbb{R}^2)$. Hence for a multi-index γ with $|\gamma| \leq s$, taking the γth derivative of each element of (16) yields

$$\left(\int_{\mathcal{E}_n} \frac{1}{\cos(\theta)} D^\gamma (\mathcal{R}_\theta(\varphi f))(\langle \cdot, \tau_\theta \rangle) d\theta \right)_{n \in \mathbb{N}}. \tag{17}$$

To prove that this sequence is a Cauchy sequence in $L^2(x_0)$ we show that $\theta \mapsto D^\gamma \mathcal{R}_\theta(\varphi f)(\langle \cdot, \tau_\theta \rangle)$ is Hölder continuous on $[0, \pi)$. In fact, by the chain rule and Theorem 4.2, for $\theta, \theta' \in [0, \pi)$, we have

$$\begin{aligned} &|D^\gamma(\mathcal{R}_\theta(\varphi f))(\langle \cdot, \tau_\theta \rangle) - D^\gamma(\mathcal{R}_{\theta'}(\varphi f))(\langle \cdot, \tau_{\theta'} \rangle)| \\ &\quad \leq \; C \cdot \Big(\Big| (\mathcal{R}_{\theta'}(\varphi f))^{(|\gamma|)}(\langle \cdot, \tau_{\theta'} \rangle) - (\mathcal{R}_{\theta'}(\varphi f))^{(|\gamma|)}(\langle \cdot, \tau_\theta \rangle) \Big| \\ &\qquad\quad + \Big| (\mathcal{R}_{\theta'}(\varphi f))^{(|\gamma|)}(\langle \cdot, \tau_\theta \rangle) - (\mathcal{R}_\theta(\varphi f))^{(|\gamma|)}(\langle \cdot, \tau_\theta \rangle) \Big| \Big). \end{aligned} \tag{18}$$

Since $\mathcal{R}_\theta(\varphi f) \in C^{s,\varepsilon}(\mathbb{R})$, we obtain that the first term of (18) is bounded by $C_0|\tau_\theta - \tau_{\theta'}|^\alpha$ for some $\alpha \leq \varepsilon$ and $C_0 > 0$. Hence also the local $L^2(x_0)$-norm of the first term is bounded by $C_1|\tau_\theta - \tau_{\theta'}|^\alpha$ with a possibly different constant C_1. In the sequel, C_ν, $\nu \in \mathbb{N}$ will always denote a positive constant.

To estimate the $L^2(x_0)$-norm of the second term of (18), it suffices to show

$$\|(\mathcal{R}_\theta(\varphi f))^{(|\gamma|)} - (\mathcal{R}_{\theta'}(\varphi f))^{(|\gamma|)}\|_{L^2(\mathbb{R})} \leq C_2|\tau_\theta - \tau_{\theta'}|^\alpha \quad \text{for some } 0 < \alpha < 1/2. \tag{19}$$

Using the Plancherel identity [27] and Theorem 4.2, we obtain

$$\begin{aligned}
&\frac{1}{|\tau_\theta - \tau_{\theta'}|^\alpha} \left\| \left((\mathcal{R}_\theta(\varphi f))^{(|\gamma|)} - (\mathcal{R}_{\theta'}(\varphi f))^{(|\gamma|)} \right) \right\|_{L^2(\mathbb{R})} \\
&= \frac{1}{2\pi} \left\| \frac{(i\,\cdot)^{|\gamma|}}{|\tau_\theta - \tau_{\theta'}|^\alpha} \left(\mathcal{F}(\varphi f)(\cdot\,\tau_\theta) - \mathcal{F}(\varphi f)(\cdot\,\tau_{\theta'}) \right) \right\|_{L^2(\mathbb{R})} \\
&\leq \frac{1}{2\pi} \left\| (i\,\cdot)^{|\gamma|+\alpha} \frac{\mathcal{F}(\varphi f)(\cdot\,\tau_\theta) - \mathcal{F}(\varphi f)(\cdot\,\tau_{\theta'})}{|\,\cdot\,\tau_\theta - \cdot\,\tau_{\theta'}|^\alpha} \right\|_{L^2(\mathbb{R})}
\end{aligned} \tag{20}$$

For $|\tau_\theta - \tau_{\theta'}|$ small enough, we now pick any multiindex ρ with $|\rho| = |\gamma|$ and $|\tau_\theta^\rho|, |\tau_{\theta'}^\rho| \geq \left(\frac{1}{2}\right)^{|\gamma|}$. Thus, by (20),

$$\begin{aligned}
&\frac{1}{|\tau_\theta - \tau_{\theta'}|^\alpha} \left\| \left((\mathcal{R}_\theta(\varphi f))^{(|\gamma|)} - (\mathcal{R}_{\theta'}(\varphi f))^{(|\gamma|)} \right) \right\|_{L^2(\mathbb{R})} \\
&\leq \frac{1}{2\pi} \left\| (i\,\cdot)^{\alpha} \frac{\tau_\theta^{-\rho}\mathcal{F}(D^\rho(\varphi f))(\cdot\,\tau_\theta) - \tau_{\theta'}^{-\rho}\mathcal{F}(D^\rho(\varphi f))(\cdot\,\tau_{\theta'})}{|\,\cdot\,\tau_\theta - \cdot\,\tau_{\theta'}|^\alpha} \right\|_{L^2(\mathbb{R})} \\
&\leq \frac{1}{2\pi} \left\| (i\,\cdot)^{\alpha} \tau_\theta^{-\rho} \frac{\mathcal{F}(D^\rho(\varphi f))(\cdot\,\tau_\theta) - \mathcal{F}(D^\rho(\varphi f))(\cdot\,\tau_{\theta'})}{|\,\cdot\,\tau_\theta - \cdot\,\tau_{\theta'}|^\alpha} \right\|_{L^2(\mathbb{R})} \\
&\qquad + \frac{1}{2\pi} \left\| (i\,\cdot)^{\alpha} \frac{\tau_\theta^{-\rho}\mathcal{F}(D^\rho(\varphi f))(\cdot\,\tau_{\theta'}) - \tau_{\theta'}^{-\rho}\mathcal{F}(D^\rho(\varphi f))(\cdot\,\tau_{\theta'})}{|\,\cdot\,\tau_\theta - \cdot\,\tau_{\theta'}|^\alpha} \right\|_{L^2(\mathbb{R})} \\
&=: \mathcal{M}_1 + \mathcal{M}_2.
\end{aligned} \tag{21}$$

Since φf has compact support, also $D^\rho(\varphi f)$ is compactly supported. Hence, its Fourier transform is Hölder continuous and obeys

$$\frac{\mathcal{F}(D^\rho(\varphi f))(r\tau_\theta) - \mathcal{F}(D^\rho(\varphi f))(r\tau_{\theta'})}{|r\tau_\theta - r\tau_{\theta'}|^\alpha} < h(r\tau_\theta), \quad \text{for all } r \in \mathbb{R},\, \theta \in [0,\pi),$$

for a function $h \in L^2(\mathbb{R}^2)$. Thus, the first term $\mathcal{M}_1$ in (21) is bounded, if $\alpha < \frac{1}{2}$. To estimate the second term $\mathcal{M}_2$ in (21), we observe, that

$$\frac{1}{\tau_\theta^\rho} - \frac{1}{\tau_{\theta'}^\rho} = \frac{\tau_\theta^\rho - \tau_{\theta'}^\rho}{\tau_\theta^\rho \tau_{\theta'}^\rho} \leq C_3|\tau_\theta - \tau_{\theta'}|.$$

Thus, the $\mathcal{M}_2$ is bounded by $C_4|\tau_\theta - \tau_{\theta'}|^{1-\alpha}$, and we have proved (19). Using the estimates for the two terms in (18), yields that

$$\|D^\gamma(\mathcal{R}_\theta(\varphi f))(\langle\cdot,\tau_\theta\rangle) - D^\gamma(\mathcal{R}_{\theta'}(\varphi f))(\langle\cdot,\tau_{\theta'}\rangle)\|_{H^s(x_0)} < C_5|\tau_\theta - \tau_{\theta'}|^\alpha.$$

Returning to the sequence in (17), for $m > n$, we have the estimate

$$\begin{aligned}
&\left\| \int_{\frac{\pi}{2}-\varepsilon_n}^{\frac{\pi}{2}-\varepsilon_m} \frac{1}{\cos(\theta)} (\mathcal{R}_\theta(\varphi f))(\langle \cdot, \tau_\theta \rangle) d\theta + \int_{\frac{\pi}{2}+\varepsilon_m}^{\frac{\pi}{2}+\varepsilon_n} \frac{1}{\cos(\theta)} (\mathcal{R}_\theta(\varphi f))(\langle \cdot, \tau_\theta \rangle) d\theta \right\|_{H^s(x_0)} \\
&= \left\| \int_{\frac{\pi}{2}-\varepsilon_n}^{\frac{\pi}{2}-\varepsilon_m} \frac{1}{\cos(\theta)} (\mathcal{R}_\theta(\varphi f))(\langle \cdot, \tau_\theta \rangle) d\theta - \int_{\frac{\pi}{2}-\varepsilon_n}^{\frac{\pi}{2}-\varepsilon_m} \frac{1}{\cos(\theta)} (\mathcal{R}_{\pi-\theta}(\varphi f))(\langle \cdot, \tau_{\pi-\theta} \rangle) d\theta \right\|_{H^s(x_0)} \\
&= \int_{\frac{\pi}{2}-\varepsilon_n}^{\frac{\pi}{2}-\varepsilon_m} \frac{1}{\cos(\theta)} \left\| (\mathcal{R}_\theta(\varphi f))(\langle \cdot, \tau_\theta \rangle) - (\mathcal{R}_{\pi-\theta}(\varphi f))(\langle \cdot, \tau_{\pi-\theta} \rangle) \right\|_{H^s(x_0)} d\theta \\
&\leq C_5 \int_{\varepsilon_m}^{\varepsilon_n} \frac{|\tau_\theta - \tau_{\pi-\theta}|^\alpha}{\cos(\theta)} d\theta \leq C_6 \int_{\varepsilon_m}^{\varepsilon_n} \frac{|\pi - 2\theta|^\alpha}{\cos(\theta)} d\theta. \qquad (22)
\end{aligned}$$

Since $\int_0^\pi |\pi - 2\theta|^\alpha / \cos(\theta)\, d\theta < \infty$, (22) converges to 0 for $m > n$ and as $n \to \infty$ and hence (16) is a Cauchy sequence in $H^s(x_0)$, which implies the required regularity of g_ξ. Thus, $a(\cdot, D)(f)$ is $s-$times weakly differentiable and using (13), we obtain the required differentiability of $\mathcal{I}_1$.

Regularity of $\mathcal{I}_{2,1}$: The proof of local regularity of $\mathcal{I}_1$ can be applied in a similar way to also prove local regularity of $\mathcal{I}_{2,1}$.

Regularity of $\mathcal{I}_{2,2}$: Using the same argument as for $\mathcal{I}_1(x)$, we obtain

$$\begin{aligned}
\mathcal{I}_{2,2}(x) &= \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \frac{\mathcal{F}((1-\varphi)f)(\xi) \mathcal{F}((1-\varphi)f)(\eta)}{\langle \eta, \xi \rangle - i0} e^{i\langle x, \xi+\eta \rangle}\, d\eta\, d\xi \\
&= \int_{|\xi| \geq 1} \int_{\mathbb{R}^2} \frac{\mathcal{F}((1-\varphi)f)(\xi) \mathcal{F}((1-\varphi)f)(\eta)}{\langle \eta, \xi \rangle - i0} e^{i\langle x, \xi+\eta \rangle}\, d\eta\, d\xi + \nu_3(x), \quad x \in U_{x_0},
\end{aligned}$$

where $\nu_3 \in C^\infty(\mathbb{R}^2)$. In this case the approach as for $\mathcal{I}_1(x)$ is not applicable anymore, since $(1-\varphi)f$ is not globally s-times differentiable and, consequently, its Radon transform does not need to be as well. Thus, it is not possible to construct a pseudo-differential operator, which is bounded from $H^s(x_0)$ to $H^s(x_0)$.

However, since $(1-\varphi)f \in L^2(\mathbb{R}^2)$, using the argument of (22) with $s = 0$ and considering the Radon transform $\mathcal{R}_\theta((1-\varphi)f)$ instead of $\mathcal{R}_\theta(\varphi f)$ yields that with the symbol b being defined as

$$b(x, \xi) := \int_{\mathbb{R}^2} \frac{\mathcal{F}((1-\varphi)f)(\eta)}{\langle \eta, \xi \rangle - i0} e^{i\langle x, \eta \rangle} d\eta, \quad (x, \xi) \in \mathbb{R}^2 \times \mathbb{R}^2,$$

the function

$$h_\xi : U_{x_0} \to \mathbb{C}, \quad h_\xi(x) := b(x, \xi), \quad \xi \in \mathbb{R}^2,$$

is an $L^2(\mathbb{R}^2)$ function with $\|h_\xi\|_{L^2(\mathbb{R}^2)}$ independent of ξ.

Approximating b via a sequence

$$b^M(x, \xi) := \int_{|\eta| \leq M} \frac{\mathcal{F}((1-\varphi)f)(\eta)}{\langle \eta, \xi \rangle - i0} e^{i\langle x, \eta \rangle} d\eta, \quad \text{for } M \in \mathbb{N},$$

we have that for fixed ξ, the function $x \mapsto b^M(x,\xi)$ is C^∞ on a neighborhood of x_0. Hence, $b^M(x,D)$ is a bounded operator from $H^t(x_0)$ to $H^t(x_0)$ for all $t \geq 0$. In particular, since $(1-\varphi)f \equiv 0$ on a neighborhood of x_0, we obtain

$$b^M(x,D)((1-\varphi)f) \equiv 0 \quad \text{on a neighborhood of } x_0,$$

which can be chosen to be the same for all M. Then

$$\|b^M(\cdot,\xi) - b(\cdot,\xi)\|_{L^2(\mathbb{R}^2)} \to 0 \quad \text{as } M \to \infty \text{ uniformly in } \xi,$$

and hence

$$\|b^M(x,D)((1-\varphi)f) - b(x,D)((1-\varphi)f)\|_{L^2(\mathbb{R}^2)} \to 0 \quad \text{as } M \to \infty.$$

In particular, since $b^M(x,D)((1-\varphi)f) = 0$ on a neighborhood of x_0, also $b(x,D)((1-\varphi)f) = 0$ on a neighborhood of x_0. Since $\mathcal{I}_2$ equals $b(x,D)((1-\varphi)f)$ up to a smooth function, this yields the claimed regularity of $\mathcal{I}_{2,2}$.

Combining all the terms $\mathcal{I}_1$, $\mathcal{I}_{2,1}$, and $\mathcal{I}_{2,2}$ finishes the proof that $q_1 \in H^s(x_0)$. For the functions $q_2, \ldots, q_m$, using a similar computation as in the proof of the main theorem of [34], we obtain

$$q_{j+1}(x) = \frac{1}{4\pi^2} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} e^{i\left\langle \xi, x + \frac{y}{2} + \frac{z}{2} \right\rangle} f(y) \mathcal{G}_{|\xi|}(y,z) q_j(z)\, dz\, dy\, d\xi, \quad x \in \mathbb{R}^2, 1 \leq j \leq m-1.$$

Now we can apply similar arguments as in the proof for $q_1 \in H^s(x_0)$, in particular, splitting f and q_j into two parts and estimating the resulting terms in the same fashion as before.

Finally, to show that $f_B \in H^s(x_0)$, the decomposition (9) indicates that it remains to analyze the regularity of the function q^R_{m+1}. It has been shown in [29, Prop. 4.1], that $q^R_{m+1} \in H^t(\mathbb{R}^2)$ for all $t < (m+1/2)/2 - 1$. Hence choosing m large enough yields the final claim. □

Remark 4.1. Observe that in Theorem (4.3) we locally lose an ε in the derivative for arbitrarily small $\varepsilon > 0$, when going from $f \in L^2(\mathbb{R}^2) \cap H^{s+\varepsilon}(x_0)$ to $f_B \in H^s(x_0)$. Certainly, one might ask whether this is in fact necessary. The examination of $\mathcal{I}_{2,2}$ in the proof of Theorem (4.3) suggests that the regularity of f_B depends only on the term

$$\int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} e^{i\left\langle \xi, x + \frac{y}{2} + \frac{z}{2} \right\rangle} \varphi(y) f(y) \mathcal{G}_{|\xi|}(y,z) \varphi(z) f(z)\, dz\, dy\, d\xi.$$

A careful review of the methods of [29] and [34] seems to indicate that this term should be even smoother than $\varphi(y)f(y)$. Hence we believe that Theorem 4.3 can be improved in the sense that locally the regularity of the Born approximation f_B is higher than the regularity of the contrast function f.

We now turn to the question of how the Born approximation affects the regularity of a function f that is modeled as a cartoon-like function. It would certainly be highly desirable that f_B is again a cartoon-like function, and we show next that this is indeed almost the case when posing some weak additional conditions to f.

The proof will use both the known results that the inverse Born approximation does not introduce a global smoothing, see [29, 32, 34], as well as Theorem 4.3, which proves that

locally the smoothness does not decrease. The key point will be that for a scatterer, which is smooth except for some singularity curve, this curve will still be present in the inverse Born approximation. For this result, we introduce the notion of a neighborhood $N_\delta(X)$ of a subset $X \subset \mathbb{R}^2$ defined by $N_\delta(X) := \{x \in \mathbb{R}^2 : \inf_{y\in X} \|x - y\|_2 < \delta\}$, where $\delta > 0$.

Corollary 4.4. *Let $\varepsilon > 0$, let $f_0, f_1 \in H^{3+\varepsilon}(\mathbb{R}^2)$ be compactly supported, let B be some compact domain with piecewise C^2 boundary ∂B, and set*

$$f = f_0 + f_1\chi_B.$$

Then, for every $\delta > 0$, there exist $\tilde{f}_0^\delta, \tilde{f}_1^\delta \in H^3(\mathbb{R}^2)$ with compact support, $h^\delta \in H^r(\mathbb{R}^2)$ for every $r < \frac{1}{2}$ with $\operatorname{supp} h^\delta \subset N_\delta(\partial B)$, and $\nu^\delta \in C^\infty(\mathbb{R}^2)$ such that the inverse Born approximation f_B of f can be written as

$$f_B = \tilde{f}_0^\delta + \tilde{f}_1^\delta\chi_D + h^\delta + \nu^\delta.$$

In particular, f_B is a cartoon-like function up to a C^∞ function and an arbitrarily well localized correction term at the boundary.

Proof. Let f_0, f_1, B be as assumed. For a fixed $\delta > 0$, choose $\varphi_1, \varphi_2, \varphi_3 \in C^\infty(\mathbb{R}^2)$ such that $\varphi_1 + \varphi_2 + \varphi_3 \equiv 1$, $\varphi_i \geq 0$ for $i = 1, 2, 3$, and

$$\begin{aligned} \varphi_1 &\equiv 1 \text{ on } N_{\frac{\delta}{2}}(\partial B), \quad \operatorname{supp} \varphi_1 \subset N_\delta(\partial B), \\ \varphi_2 &\equiv 1 \text{ on } (\operatorname{supp} f_0 \cup \operatorname{supp} f_1) \setminus N_\delta(\partial B), \quad \operatorname{supp} \varphi_2 \subset N_\delta(\operatorname{supp} f_0 \cup \operatorname{supp} f_1) \setminus N_{\frac{\delta}{2}}(\partial B). \end{aligned}$$

By Theorem 4.3, it follows that $\varphi_2 f_B \in H^3(\mathbb{R}^2)$ and $\varphi_3 f_B \in C^\infty(\mathbb{R}^2)$. Then (9) implies that

$$\varphi_i f_B = \varphi_i f + \varphi_i q_1^R, \qquad \text{for } i = 1, 2, 3,$$

and thus $\varphi_2 q_1^R \in H^3(\mathbb{R}^2)$ and $\varphi_3 q_1^R \in C^\infty(\mathbb{R}^2)$.

Defining

$$\tilde{f}_0^\delta := f_0 + \varphi_2 q_1^R, \quad \tilde{f}_1^\delta := f_1, \quad h^\delta := \varphi_1 q_1^R, \quad \text{and} \quad \nu^\delta := \varphi_3 q_1^R,$$

then the Sobolev embedding theorem [2], implies that $\tilde{f}_0^\delta, \tilde{f}_1^\delta \in C^2(\mathbb{R}^2)$. Then [29, Prop. 4.1] implies that $q_1^R \in H^r(\mathbb{R}^2)$ for all $r < \frac{1}{2}$, and hence $h^\delta \in H^r(\mathbb{R}^2)$, and $\operatorname{supp} h^\delta \subset N_\delta(\partial B)$ follows by construction. The function ν^δ is C^∞, since $\varphi_3 q_1^R \in C^\infty(\mathbb{R}^2)$. Thus the main assertion is proved and the 'in particular' part follows immediately. □

As highlighted before in [7, 9, 30] as well as in [21], there exist various numerical approaches to enhance the solution of linear inverse problems under the premise of sparsity in the shearlet expansion. With the local regularity of f_B established, we now have the full repertoire of methods from shearlet theory at hand to enhance the solution of the inverse Born approximation for the electromagnetic Schrödinger equation.

5 Numerical Methods for the Acoustic Inverse Scattering Problem

In this section, we will analyze numerical approaches to solve the acoustic scattering problem of Section 3. After discussing an algorithmic realization of our approach (7), we briefly present the other numerical methods that we compare with, followed by a detailed description of the numerical experiments. It will turn out that our new method is advantageous to the other methods in the situation that the scatterer is a body consisting of a more or less homogeneous medium, whose density is significantly different from the surrounding medium.

5.1 The New Algorithmic Approach

As suggested in Subsection 3.3, we aim to solve the minimization problem, compare also (7),

$$\min_{f\in L^2(B_R)} \left(\frac{1}{2} \|\mathcal{N}(f) - N^{\varepsilon}_{\text{meas}}\|^2_{\text{HS}(L^2(\Gamma_{\text{inc}}),L^2(\Gamma_{\text{meas}}))} + \frac{\alpha}{p} \|T_\Phi(f)\|^p_{\ell^p} \right),$$

where Φ is a shearlet frame for $L^2(\mathbb{R}^2)$, and $\tilde{\Phi}$ denotes the associated canonical dual frame. Employing the sign function, the mapping

$$J_q : B_R \to \mathbb{C}, \quad x \mapsto [J_p(q)](x) := |q(x)|^{p-1}\text{sign}(q(x)),$$

and the operator $\mathcal{S}_{\alpha\mu,p} := (I + \alpha\mu J_p)^{-1}$, it has been shown in [26] that the solution via the standard Tikhonov functional (6) can be obtained as the limit of the *Landweber iteration*

$$f_{n+1} = \mathcal{S}_{\alpha\mu,p}\left[f_n - \mu_n[\mathcal{N}'(f_n)]^*(\mathcal{N}(f_n) - N^{\varepsilon}_{\text{meas}})\right] \quad \text{for } (\mu_n)_n \subset \mathbb{R}^+. \tag{23}$$

By [31], the solution to (7) with a frame based regularization term can be computed as a limit of the iteration

$$f_{n+1} = T_{\tilde{\Phi}^*}\mathcal{S}_{\alpha\mu,p}\left[T_\Phi(f_n - \mu_n[\mathcal{N}'(f_n)]^*(\mathcal{N}(f_n) - N^{\varepsilon}_{\text{meas}}))\right] \quad \text{for } (\mu_n)_n \subset \mathbb{R}^+, \tag{24}$$

see [26] for an explicit construction of $[\mathcal{N}']^*$.

Since the ℓ_1-norm promotes sparsity, in our experiments we will choose $p = 1$. In this case, $\mathcal{S}_{\alpha\mu,1}$ is the soft-thresholding operator, which for a scalar ω, is defined as

$$\mathcal{S}_{\alpha,1}(\omega) = \begin{cases} \omega - \alpha, & \text{if } \omega \geq \alpha, \\ 0, & \text{if } |\omega| < \alpha, \\ \omega + \alpha, & \text{if } \omega \leq \alpha, \end{cases}$$

with element-wise application for sequences.

The general setup of the numerical experiments, whose results will be described in Subsection 5.2, follows that for similar experiments presented in [26]. We chose the stepsize μ_n according to the Barzilai-Borwein rule [4], and stop the iteration in accordance to the standard discrepancy principle with parameter $\tau = 1.6$, i.e., when

$$\|\mathcal{N}(f_n) - N^{\varepsilon}_{\text{meas}}\|_{\text{HS}(L^2(\Gamma_i),L^2(\Gamma_m))} \leq \tau\varepsilon, \tag{25}$$

with ε being a fixed parameter chosen according to the noise level.

In each step of (24), one shearlet decomposition and reconstruction needs to be performed for which *Shearlab* [25] is used. In all experiments, a discretization of the domain with a 512×512 grid is used and the standard subsampled shearlet system of Shearlab using 5 scales is chosen.

We select as domain $[-1, 1]^2$, and let the scatterer be supported in B_R with $R = 0.75$. We then pick 32 transmitter-receiver pairs, equidistributed on the circle of radius 0.9. Thus 64 Lippmann-Schwinger equations need to be solved in every step, 32 for the evaluation of $\mathcal{N}(f_n)$ for the different single layer potentials, and 32 for the evaluation of $[\mathcal{N}'(f_n)]^*$.

We then solve these equations with a simple, and admittedly slower, method than [26], by discretizing the Lippmann-Schwinger equation with respect to a wavelet system, in our case a Daubechies 6 Wavelet, [13]. We solve the resulting linear system using a GMRES iteration without preconditioning. The results in Subsection 5.2 show that even with this simple

$\mathbf{k_0 = 40}$	Regularization method	Noise level	Relative error	#iterations
1.	L^1 Tikhonov	0.01	0.4095	19
2.	Shearlets	0.01	**0.3914**	15
3.	No Penalty	0.01	0.4580	11
4.	L^1 Tikhonov	0.005	0.2689	35
5.	Shearlets	0.005	**0.2100**	31
6.	No Penalty	0.005	0.2736	28
7.	L^1 Tikhonov	0.002	0.1664	91
8.	Shearlets	0.002	**0.1095**	63
9.	No Penalty	0.002	0.1719	67
10.	L^1 Tikhonov	0.001	0.1132	217
11.	Shearlets	0.001	**0.0635**	123
12.	No Penalty	0.001	0.1189	185

Table 1: Numerical results of the three regularization methods for different noise levels with the scatterer chosen as in Figure 3.

approach the advantage of the shearlet regularization over other regularization methods can be observed. The increased runtime per step does not affect the overall runtime significantly, since we require less iterations. However, using the method of [26] to solve the Lippmann-Schwinger equations should provide a significant further speed-up in our algorithm, when aiming for higher numerical efficiency and not only for the accuracy of the reconstruction.

As scatterers f, we consider prototypes of a cartoon-like functions, as depicted in Figures 3 and 4.

5.2 Comparison Results

We compare our approach with two other approaches, first the method introduced in [26], which is based on the assumption that the scatterer is itself sparse and hence an L^1 regularization is used, solving

$$\min_{f\in L^2(B_R)} \left(\frac{1}{2} \left\| N_f - N^{\varepsilon}_{\text{meas}} \right\|^2_{\text{HS}(L^2(\Gamma_{\text{inc}}),L^2(\Gamma_{\text{meas}}))} + \frac{\alpha}{p} \|f\|^1_{L^1(B_R)} \right)$$

via the Landweber iteration (23), and second

$$\min_{f\in L^2(B_R)} \left\| N_f - N^{\varepsilon}_{\text{meas}} \right\|_{\text{HS}(L^2(\Gamma_{\text{inc}}),L^2(\Gamma_{\text{meas}}))},$$

which does not contain a regularization term, hence does not exploit sparsity in any way. We stop the iteration when the discrepancy principle is achieved.

In the first set of experiments we choose a wave number of $k_0 = 40$ and compute reconstructions with our approach, see Subsection 5.1. The different noise levels we impose are described in Table 1 and Figure 3. In Table 1, for each of the three regularization methods, we provide the relative error measured in the discrete L^2-norm as well as the number of iterations until (25) is achieved for different noise levels.

The shearlet scheme shows the best performance both visually and with respect to the relative error. Interestingly, it also requires the least number of iterations. The inferior performance of the L^1 regularization from [26] is due to the fact that the scatterer is not sparse itself in the sense of having a relatively small support. Certainly, if no penalty term

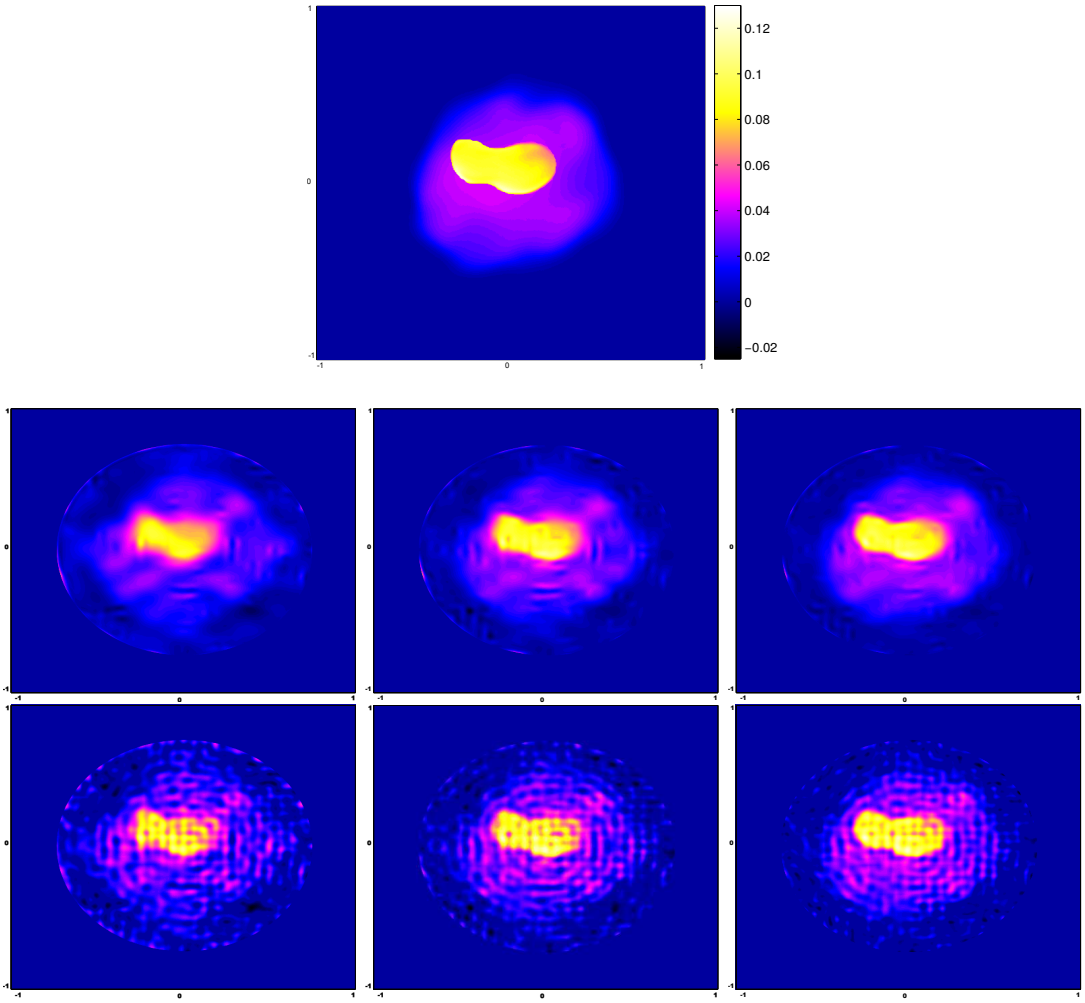

Figure 3: **Top:** Scatterer. **Second Row:** Reconstructed scatterers using the shearlet regularization, noise levels from left to right: $\varepsilon = 0.005, 0.002, 0.001$. **Third Row:** Reconstructed scatterers using the L^1 regularization, noise levels from left to right: $\varepsilon = 0.005, 0.002, 0.001$.

is used, then the solution is not at all adapted to the specific structure and expectedly, the performance is worse.

We also conducted a second set of experiments with a different wave number, i.e., $k_0 = 30$ and display the results in Table 2. Furthermore, the reconstruction error is depicted in Figure 4, where we observe that the shearlet regularization produces satisfying results. Most importantly, the singularity curve of the scatterer, which is the most prominent feature of the cartoon model, is obtained with decent precision.

The reason for the superior performance of the regularization by the shearlet transform is also visible in Figure 4. All three methods handle the singularity curve fairly well, although, naturally, the error is the largest, at points where the singularity is most pronounced, i.e., the upper and lower right corners as well as the middle of the left edge of the centered square. Away from the singularities, the shearlet regularization yields a far better approximation than the other two approaches, since it is designed to deal very well with smooth regions.

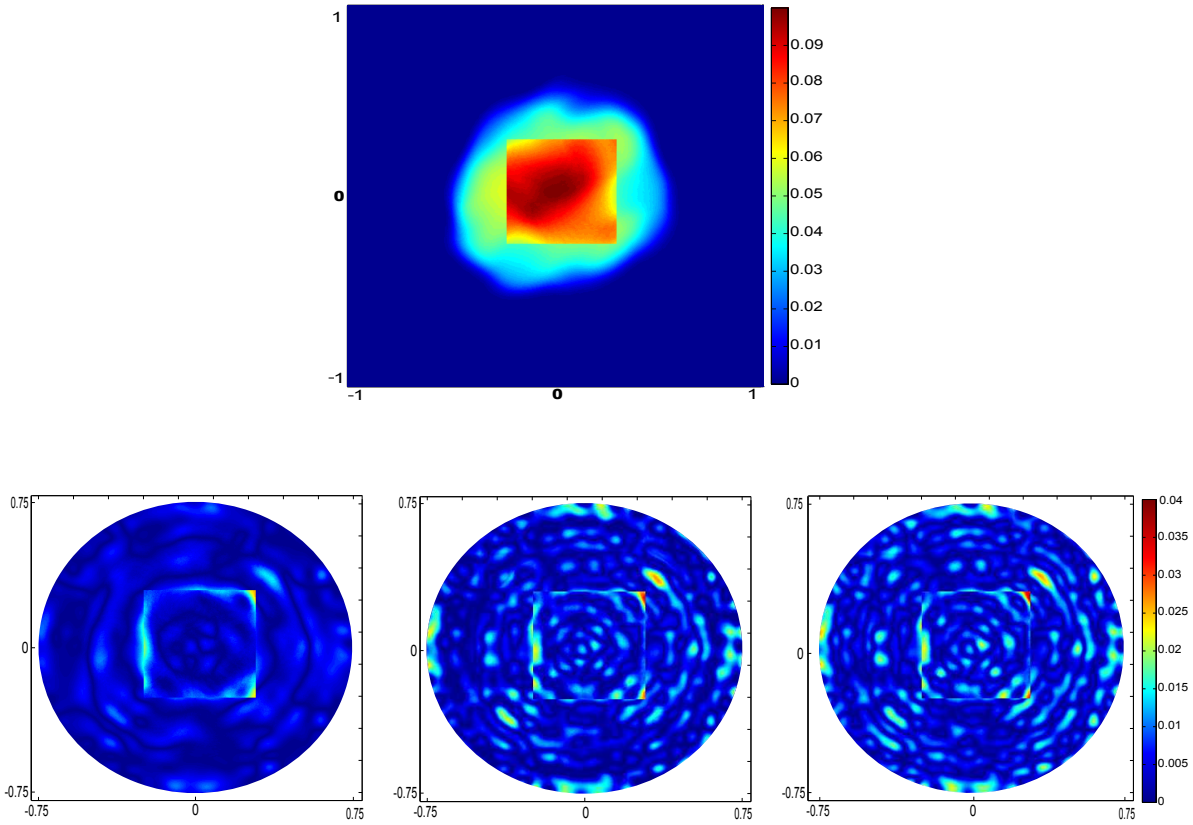

Figure 4: **Top:** Square in front of a smooth background, **Second Row:** Error of the reconstruction using the shearlet regularization, L^1 Tikhonov Regularization, and without penalty term, $\varepsilon = 0.002$.

Acknowledgements

The first author acknowledges support by the Einstein Foundation Berlin, by the Einstein Center for Mathematics Berlin (ECMath), by Deutsche Forschungsgemeinschaft (DFG) Grant KU 1446/14, by the DFG Collaborative Research Center TRR 109 "Discretization in Geometry and Dynamics", and by the DFG Research Center MATHEON "Mathematics for key technologies" in Berlin. The second author acknowledges support by the DFG Research Center MATHEON "Mathematics for key technologies" in Berlin. The third author thanks the DFG Collaborative Research Center TRR 109 "Discretization in Geometry and Dynamics" for its support.

References

[1] Milton Abramowitz and Irene A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Dover Publications Inc., New York, 1992. Reprint of the 1972 edition.

[2] Robert A. Adams. *Sobolev spaces*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.

Table 2:

$\mathbf{k_0 = 30}$	Regularization method	Noise level	Relative error	#iterations
1.	L^1 Tikhonov	0.01	0.2786	8
2.	Shearlets	0.01	**0.2286**	8
3.	No Penalty	0.01	0.2548	7
4.	L^1 Tikhonov	0.005	0.1572	12
5.	Shearlets	0.005	**0.1135**	12
6.	No Penalty	0.005	0.1357	12
7.	L^1 Tikhonov	0.002	0.0929	21
8.	Shearlets	0.002	**0.0738**	20
9.	No Penalty	0.002	0.0911	21
10.	L^1 Tikhonov	0.001	0.0644	51
11.	Shearlets	0.001	**0.0545**	41
12.	No Penalty	0.001	0.0597	56

Table 2: Numerical results of the three regularization methods for different noise levels with the scatterer chosen as in 4.

[3] Gang Bao and Faouzi Triki. Error estimates for the recursive linearization of inverse medium problems. *J. Comput. Math.*, 28(6):725–744, 2010.

[4] Jonathan Barzilai and Jonathan M. Borwein. Two-point step size gradient methods. *IMA J. Numer. Anal.*, 8(1):141–148, 1988.

[5] Michael Beals and Michael Reed. Microlocal regularity theorems for nonsmooth pseudodifferential operators and applications to nonlinear problems. *Trans. Amer. Math. Soc.*, 285(1):159–184, 1984.

[6] Emmanuel Candès and David L. Donoho. New tight frames of curvelets and optimal representation of objects with C^2 singularities. *Comm. Pure. Appl. Math.*, 57(2):219–266, 2004.

[7] Emmanuel J. Candès and David L. Donoho. Recovering edges in ill-posed inverse problems: optimality of curvelet frames. *Ann. Statist.*, 30(3):784–842, 2002. Dedicated to the memory of Lucien Le Cam.

[8] Ole Christensen. *An introduction to frames and Riesz bases*. Birkhäuser, Basel, 2003.

[9] Flavia Colonna, Glenn Easley, Kanghui Guo, and Demetrio Labate. Radon transform inversion using the shearlet representation. *Appl. Comput. Harmon. Anal.*, 29(2):232–250, 2010.

[10] David Colton, Joe Coyle, and Peter Monk. Recent developments in inverse acoustic scattering theory. *SIAM Rev.*, 42(3):369–414, 2000.

[11] David Colton, Houssem Haddar, and Michele Piana. The linear sampling method in inverse electromagnetic scattering theory. *Inverse Problems*, 19(6):S105–S137, 2003. Special section on imaging.

[12] David Colton and Rainer Kress. *Inverse acoustic and electromagnetic scattering theory*, volume 93 of *Applied Mathematical Sciences*. Springer, New York, third edition, 2013.

[13] Ingrid Daubechies. *Ten lectures on wavelets*, volume 61. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[14] David Donoho. Sparse components of images and optimal atomic decompositions. *Constr. Approx.*, 17(3):353–382, 2001.

[15] David Donoho and Gitta Kutyniok. Microlocal analysis of the geometric separation problem. *Comm. Pure Appl. Math.*, 66(1):1–47, 2013.

[16] Martin Genzel and Gitta Kutyniok. Asymptotical analysis of inpainting via universal shearlet systems. *preprint*, 2014.

[17] Kanghui Guo, Gitta Kutyniok, and Demetrio Labate. Sparse multidimensional representations using anisotropic dilation and shear operators. In *Wavelets and splines: Athens 2005*, Mod. Methods Math., pages 189–201. Nashboro Press, Brentwood, TN, 2006.

[18] David Jerison and Carlos E. Kenig. Unique continuation and absence of positive eigenvalues for Schrödinger operators. *Ann. of Math. (2)*, 121(3):463–494, 1985. With an appendix by E. M. Stein.

[19] Emily J. King, Gitta Kutyniok, and Xiaosheng Zhuang. Analysis of inpainting via clustered sparsity and microlocal analysis. *J. Math. Imaging Vis.*, 48:205–234, 2014.

[20] Pisamai Kittipoom, Gitta Kutyniok, and Wang-Q Lim. Construction of compactly supported shearlet frames. *Constr. Approx.*, 35(1):21–72, 2012.

[21] Gitta Kutyniok and Demetrio Labate. In *Shearlets*, Appl. Numer. Harmon. Anal., pages 1–38. Birkhäuser/Springer, New York, 2012.

[22] Gitta Kutyniok and Demetrio Labate. *Shearlets: Multiscale analysis for multivariate data*, chapter Introduction to shearlets, pages 1–38. Birkhäuser Boston, 2012.

[23] Gitta Kutyniok and Wang-Q Lim. Compactly supported shearlets are optimally sparse. *J. Approx. Theory*, 163(11):1564–1589, 2011.

[24] Gitta Kutyniok and Wang-Q Lim. Image separation using wavelets and shearlets. In *Curves and surfaces*, volume 6920 of *Lecture Notes in Comput. Sci.*, pages 416–430. Springer, Heidelberg, 2012.

[25] Gitta Kutyniok, Wang-Q Lim, and Rafael Reisenhofer. Shearlab: Faithful digital shearlet transform with compactly supported shearlets. *preprint*, 2014.

[26] Armin Lechleiter, Kamil S Kazimierski, and Mirza Karamehmedovic. Tikhonov regularization in L^p applied to inverse medium scattering. *Inverse Problems*, 29(7), 2013.

[27] Stéphane Mallat. *A wavelet tour of signal processing: The sparse way*. Elsevier/Academic Press, Amsterdam, third edition, 2009.

[28] Shari Moskow and John C. Schotland. Convergence and stability of the inverse scattering series for diffuse waves. *Inverse Problems*, 24(6):065005, 16, 2008.

[29] Petri Ola, Lassi Päivärinta, and Valeri Serov. Recovering singularities from backscattering in two dimensions. *Comm. Partial Differential Equations*, 26(3-4):697–715, 2001.

[30] Vishal M. Patel, Glenn R. Easley, and Dennis M. Healy, Jr. Shearlet-based deconvolution. *IEEE Trans. Image Process.*, 18(12):2673–2685, 2009.

[31] Ronny Ramlau and Gerd Teschke. A Tikhonov-based projection iteration for nonlinear ill-posed problems with sparsity constraints. *Numer. Math.*, 104(2):177–203, 2006.

[32] Juan M. Reyes. Inverse backscattering for the Schrödinger equation in 2d. *Inverse Problems*, 23(2):625, 2007.

[33] Otmar Scherzer, Markus Grasmair, Harald Grossauer, Markus Haltmeier, and Frank Lenzen. *Variational methods in imaging*, volume 167 of *Applied Mathematical Sciences*. Springer, New York, 2009.

[34] Valery Serov. Inverse backscattering Born approximation for a two-dimensional magnetic Schrödinger operator. *Inverse Problems*, 29(7):075015, 9, 2013.

[35] Joachim Weidmann. *Linear operators in Hilbert spaces*, volume 68 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1980.