Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Seminar for Applied Mathematics

Ridgelet Methods for Linear Transport Equations

P. Grohs and A. Obermeier

Research Report No. 2014-35
November 2014

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Funding SNF: 146356

Ridgelet Methods for Linear Transport Equations

Philipp Grohs, Axel Obermeier

November 17, 2014

Abstract

In this paper we present an overview of a novel method for the numerical solution of linear transport equations, which is based on ridgelets and has been introduced in [GO14, EGO14]. Such equations arise for instance in radiative transfer or in phase contrast imaging. Due to the fact that ridgelet systems are well adapted to the structure of linear transport operators, it can be shown that our scheme operates in optimal complexity, even if line singularities are present in the solution. After presenting the basic algorithm, we prove that certain operators are compressible, which is the key to obtain unconditional convergence results. Finally, we show some applications in radiative transport.

Contents

1 Introduction

In the past two decades, a wide range of multiscale systems have been introduced with lasting impact in many different fields, starting with wavelets [Dau92] and continuing with ridgelets [Can98], curvelets [CD05b, CD05a, CDDY06], shearlets [KLLW05, KL12], contourlets [DV05] etc. – the latter three of which fall into the framework of so-called "parabolic molecules" [GK14], while all of the mentioned systems are encompassed by the even broader framework of α-molecules [GKKS14].

These systems share the property that they are very well-adapted to representing certain classes of functions optimally (in the sense of the decay rate of the best N-term approximation) – functions with point singularities for wavelets, line singularities for ridgelets and curved singularities for parabolic molecules. Since these classes make up the fundamental phenomenological features of most images in an extremely diverse set of applications, it is perhaps not surprising, that many of the above-mentioned systems were originally investigated in view of their properties regarding image processing.

With a certain time-lag, it is becoming apparent that these systems are also very suitable for solving partial differential equations – again, wavelets were the first in this regard, for example leading to provably optimal solvers for elliptic equations [CDD01]. For differential equations with strong directional features – such as transport equations – it is intuitively clear that optimal solvers will need to take these features into account, however, the development of solvers based on directional systems is still in its infancy.

Following recent results [Gro12], that ridgelets permit the construction of simple diagonal preconditioners for linear transport equations which arise in collocation-type discretization methods for kinetic transport equations (such as radiative transport), we consider the papers [GO14, EGO14] to be a first step towards establishing directional representation systems as a useful tool for solving PDEs. The present paper proves results about the compressibility of certain operators that have been left open in [GO14].

Secondly, we present results from [EGO14], where we introduced an `FFT`-based implementation of a ridgelet transform: FFRT – **F**ast **F**inite **R**idgelet **T**ransform. We use this for the numerical solution of kinetic transport equations arising in radiative transport. Using the preconditioner from [Gro12] for linear transport equations together with a sparse discrete ordinates method similar to [GS11a], we construct a solver which mitigates the curse of dimensionality and which results in uniformly well-conditioned linear systems which can be solved efficiently with CG.

1.1 Radiative Transport Equation

The motivation for this work is the numerical solution of the following model equation, described by the radiative transport equation (RTE),

$$Au := \vec{s} \cdot \nabla u + \kappa\, u = f + \int_{\mathbb{S}^{d-1}} \sigma u \,\mathrm{d}\vec{s}'.$$

It is a steady state continuity equation describing the conservation of radiative intensity in an absorbing, emitting and scattering medium, see e.g. [Mod13]. We will, however, not treat the scattering operator in this paper, which can be incorporated through a variety of methods, one of which – the source iteration – we implemented in [EGO14]. Let us assume that the following quantities are known at all locations $\vec{x} \in \Omega \subset \mathbb{R}^d$ and for all directions $\vec{s} \in \mathbb{S}^{d-1} := \left\{\vec{s} \in \mathbb{R}^d \colon \|\vec{s}\|_2 = 1\right\}$:

- absorption coefficient $\kappa(\vec{x}, \vec{s}) \geq \kappa_0 > 0$
- source term $f(\vec{x}, \vec{s}) \in \mathbb{R}$

Then, the above equation allows us to find the unknown radiative intensity u as a function $\Omega \times \mathbb{S}^{d-1} \to \mathbb{R}$.

Although the RTE looks simple, standard numerical techniques for solving it do not perform well for a number of reasons, mainly:

- The transport term $s \cdot \nabla u$ leads to ill-conditioned systems of equations.
- Singularities in the input data may remain in the solution.
- With the dimension of the domain of u being 3 in 2-dimensional physical space and 5 in 3-dimensional space, the problem is fairly high-dimensional.

These issues make the accurate numerical solution of the RTE very costly or even impossible due to memory and compute power limitations of today's hardware.

1.2 Model Problem

We first simplify the problem by removing the scattering operator and fixing the transport direction $\vec{s}$. However, we will come back to the full problem once we have developed a solver for this easier problem, and then use solve the problem with scattering using a collocation approach, which can also be combined with sparse tensor methods to alleviate the curse of dimensionality.

Therefore, our starting point is the differential operator

$$A\colon\ H^{\vec{s}}(\mathbb{R}^d) \ni u \mapsto \vec{s} \cdot \nabla u(\vec{x}) + \kappa(\vec{x})u(\vec{x}) \in L^2(\mathbb{R}^d) \tag{1.1}$$

with fixed $\vec{s} \in \mathbb{S}^{d-1}$ and a function $\kappa \in L^\infty(\mathbb{R}^d)$ that satisfies $\kappa(\vec{x}) \geq \gamma > 0,\ \forall \vec{x} \in \mathbb{R}^d$. The space $H^{\vec{s}}$ is defined as follows.

Definition 1.1. Let $\vec{s} \in \mathbb{S}^{d-1}$, then we define the *anisotropic Sobolev space*

$$H^{\vec{s}}(\Omega) := \left\{ f \in L^2(\Omega) \colon (\vec{s} \cdot \nabla) f \in L^2(\Omega) \right\} \qquad \text{with norm} \quad \|f\|^2_{H^{\vec{s}}(\Omega)} := \|f\|^2_{L^2(\Omega)} + \|(\vec{s} \cdot \nabla) f\|^2_{L^2(\Omega)}.$$

This space is more easily characterised on the Fourier side,

$$H^{\vec{s}}(\widehat{\Omega}) := \left\{ \hat{f} \in L^2(\widehat{\Omega}) \colon \langle \vec{s} \cdot \vec{\xi} \rangle \hat{f}(\vec{\xi}) \in L^2(\widehat{\Omega}) \right\} \quad \text{with norm} \quad \left\|\hat{f}\right\|_{H^{\vec{s}}(\widehat{\Omega})} := \left\| \langle \vec{s} \cdot \vec{\xi} \rangle \hat{f} \right\|_{L^2(\widehat{\Omega})},$$

where $\langle x \rangle := \sqrt{1 + |x|^2}$ is the *regularised absolute value.*

To solve the equation $Au = f \in L^2(\mathbb{R}^d)$, we look for solutions by minimising the L^2-residual,

$$u_0 = \underset{v \in H^{\vec{s}}}{\operatorname{argmin}} \|Av - f\|_{L^2}. \tag{1.2}$$

With a variation-of-constants-argument (following [GS11b]), it is not difficult to show the following.

Proposition 1.2 ([GO14])**.** *For every $\ell \in \left(H^{\vec{s}}\right)'$ there exists a unique $u_0 \in H^{\vec{s}}$ which solves* (1.2)*. Moreover, the solution is characterised by the variational equation*

$$a(v, u_0) = \ell(v) \quad \text{for all } v \in H^{\vec{s}}, \quad \text{where} \quad a(v, u) := \langle Av, Au \rangle_{L^2}. \tag{1.3}$$

In particular, well-definedness holds for

$$\ell_f(v) := \langle Av, f \rangle_{L^2} \quad \text{with } f \in L^2(\mathbb{R}^d).$$

2 Discretization

In our paper we aim to solve (1.2) via solving a discretization of the linear system (1.3). Several ingredients are needed to render this approach efficient:

(i) Uniform well-conditionedness of the resulting infinite discrete linear system

(ii) Fast approximate matrix-vector multiplication for the discrete operator matrix

(iii) Efficient approximation of typical solutions

There exists several results which essentially state that, whenever (i), (ii) and (iii) are satisfied, then the linear system (1.3) can be solved in optimal computational complexity [CDD01, Ste03, DFR07].

2.1 Gelfand Frames

For a Gelfand triple $(\mathcal{H}, L^2(\mathbb{R}^d), \mathcal{H}')$, it can be shown that for an operator $F \colon \mathcal{H} \to \mathcal{H}'$ which is bounded and boundedly invertible, induces a symmetric bilinear form $a(v, u) := \langle v, Fu \rangle_{\mathcal{H} \times \mathcal{H}'}$ and is elliptic in the sense that $a(v, v) \sim \|v\|^2_{\mathcal{H}}$, the equation $Fu = f$ can be discretised (and preconditioned) with a Gelfand frame to $\mathbf{Fu} = \mathbf{f}$, such that the following result holds.

Lemma 2.1 ([DFR07, Lemma 4.1])**.** *The operator $\mathbf{F} : \ell^2(\Lambda) \to \ell^2(\Lambda)$ is bounded and boundedly invertible on its range. In particular, the system $\mathbf{Fu} = \mathbf{f}$ is a uniformly well-conditioned infinite linear system.*

2.2 Numerical Solution of the Discrete System

If we were able to compute with infinite vectors, at this point we could simply solve

$$\mathbf{F}\mathbf{u} = \mathbf{f}, \tag{2.1}$$

by using a standard iterative solver such as a damped Richardson iteration

$$\mathbf{u}^{(j+1)} = \mathbf{u}^{(j)} - \alpha\big(\mathbf{F}\mathbf{u}^{(j)} - \mathbf{f}\big), \quad \mathbf{u}^{(0)} = \mathbf{0}.$$

Due to the well-conditionedness of the matrix $\mathbf{F}$ ensured by Lemma 2.1 and the fact that the iterates stay in $\mathrm{ran}(\mathbf{F})$ in each step, it is easy to show that for appropriate damping α the sequence $\mathbf{u}^{(j)}$ converges geometrically to the sought solution $\mathbf{u}$ in the $\ell^2(\Lambda)$-norm, i.e.

$$\big\|\mathbf{u} - \mathbf{u}^{(j)}\big\|_{\ell^2(\Lambda)} \lesssim \rho^j$$

for some $\rho < 1$, depending on the spectral properties of the operator $\mathbf{F}$.

To deal with the fact that we can only compute with finite vectors, we consider how this is done for the by-now classical wavelet discretisations of elliptic PDEs. The approximative evaluation of the Richardson iteration utilises the following three procedures:

- $\mathbf{RHS}[\varepsilon, \mathbf{f}] \to \mathbf{f}_\varepsilon$: determines for $\mathbf{f} \in \ell^2(\Lambda)$ a finitely supported $\mathbf{f}_\varepsilon \in \ell^2(\Lambda)$ such that

$$\|\mathbf{f} - \mathbf{f}_\varepsilon\|_{\ell^2(\Lambda)} \le \varepsilon;$$

- $\mathbf{APPLY}[\varepsilon, \mathbf{A}, \mathbf{v}] \to \mathbf{v}_\varepsilon$: determines for $\mathbf{A} : \ell^2(\Lambda) \to \ell^2(\Lambda)$ and for a finitely supported $\mathbf{v} \in \ell^2(\Lambda)$ a finitely supported $\mathbf{v}_\varepsilon$ such that

$$\|\mathbf{A}\mathbf{v} - \mathbf{v}_\varepsilon\|_{\ell^2(\Lambda)} \le \varepsilon;$$

- $\mathbf{COARSE}[\varepsilon, \mathbf{u}] \to \mathbf{u}_\varepsilon$: determines for a finitely supported $\mathbf{u} \in \ell^2(\Lambda)$ a finitely supported $\mathbf{u}_\varepsilon \in \ell^2(\Lambda)$ with at most N nonzero coefficients, such that

$$\|\mathbf{u} - \mathbf{u}_\varepsilon\|_{\ell^2(\Lambda)} \le \varepsilon. \tag{2.2}$$

  Moreover, $N \lesssim N_{\min}$ holds, $N_{\min}$ being the minimal number of entries with (2.2).

We refer to [CDD01, Ste03, DFR07] for information on the numerical realization of these routines.

2.3 The Problem with $\ker(\mathbf{F})$

For the finite system, the iterates are no longer guaranteed to stay in $\mathrm{ran}(\mathbf{F})$ – in particular, errors in the kernel of the discretisation of $\mathbf{F}$ may accumulate. One possible remedy is to apply a projection $\mathbf{P}$ with $\ker(\mathbf{P}) = \ker(\mathbf{F})$ periodically, to eliminate these errors in the kernel.

Assuming the existence of numerical procedures as above and such a projection $\mathbf{P}$, we can formulate the numerical algorithm [Ste03] to solve the discrete linear system (2.1) up to accuracy $\varepsilon > 0$, given as Algorithm 2.2 below.

Conditional on the three routines above, we have thus formulated a feasible algorithm for the approximate solution of (2.1).

2.4 Compressibility

To achieve optimal convergence rates for our problem through the techniques introduced in [CDD01], a key ingredient is *compressibility* of the discretised operator equation. Such a property guarantees the existence of linear-time approximate matrix-vector multiplication algorithms **APPLY** which are used in the iterative solution of the operator equation, see [CDD01, Ste03] for more information.

Algorithm 2.2: Modified Inexact Damped Richardson Iteration

Data: $\varepsilon > 0$, $\mathbf{F}$, $\mathbf{f}$

Result: $\mathbf{u}_\varepsilon = \mathbf{modSOLVE}[\varepsilon, \mathbf{F}, \mathbf{P}, \mathbf{f}]$

Let $\theta < \frac{1}{3}$ and $K \in \mathbb{N}$ such that $3\rho^K \|\mathbf{P}\| < \theta$. $i := 0$, $\mathbf{u}^{(0)} := \mathbf{0}$, $\varepsilon_0 := \big\|\mathbf{P}\big\|\big\|\mathbf{F}\big|_{\mathrm{ran}(\mathbf{F})}^{-1}\big\|\|\mathbf{f}\|_{\ell^2(\Lambda)}$

while $\varepsilon_i > \varepsilon$ **do**

- $i := i + 1$;
- $\varepsilon_i := 3\rho^K \|\mathbf{P}\|\varepsilon_{i-1}/\theta$;
- $\mathbf{f}^{(i)} := \mathbf{RHS}[\theta\varepsilon_i/(6\alpha K\|\mathbf{P}\|), \mathbf{f}]$;
- $\mathbf{u}^{(i,0)} := \mathbf{u}^{(i-1)}$;
- **for** $j = 1, \dots, K$ **do**
  - $\mathbf{u}^{(i,j)} := \mathbf{u}^{(i,j-1)} - \alpha\big(\mathbf{APPLY}\big[\theta\varepsilon_i/(6\alpha K\|\mathbf{P}\|), \mathbf{F}, \mathbf{u}^{(i,j-1)}\big] - \mathbf{f}^{(i)}\big)$;
- $\mathbf{z}^{(i)} := \mathbf{APPLY}\big[\theta\varepsilon_i/3, \mathbf{P}, \mathbf{u}^{(i,K)}\big]$;
- $\mathbf{u}^{(i)} := \mathbf{COARSE}[(1-\theta)\varepsilon_i, \mathbf{z}^{(i)}]$;

$\mathbf{u}_\varepsilon := \mathbf{u}^{(i)}$;

Definition 2.3. A matrix $\mathbf{A}$ is called σ^*-*compressible* if for every $\sigma < \sigma^*$ and $k \in \mathbb{N}$ there exists a matrix $\mathbf{A}^{[k]}$ such that

(i) the matrix $\mathbf{A}^{[k]}$ has at most $\alpha_k 2^k$ non-zero entries in each column,

(ii) the following estimate holds

$$\big\|\mathbf{A} - \mathbf{A}^{[k]}\big\|_2 \le C_k,$$

in such a way that the sequences $(\alpha_k)_{k\in\mathbb{N}}$, $(C_k 2^{\sigma k})_{k\in\mathbb{N}}$ are both summable.

Definition 2.4 ([Ste03, Def. 3.9])**.** A vector $\mathbf{c} \in \ell^2$ is called σ^*-*optimal*, when for a suitable routine $\mathbf{RHS}$, for each $\sigma \in (0, \sigma^*)$ with $p := (\sigma + \frac{1}{2})^{-1}$, the following is valid for $\mathbf{c}_\varepsilon = \mathbf{RHS}[\varepsilon, \mathbf{c}]$:

1. $\#\mathrm{supp}\ \mathbf{c}_\varepsilon \lesssim \varepsilon^{-1/\sigma} |\mathbf{c}|_{\ell^p_w}^{1/\sigma}$
2. The number of arithmetic operations used to compute $\mathbf{c}_\varepsilon$ is at most a multiple of $\varepsilon^{-1/\sigma}|\mathbf{c}|_{\ell^p_w}^{1/\sigma}$.

We can now formulate the main result of [Ste03, Theorem 3.11].

Theorem 2.5 (Convergence of $\mathbf{modSOLVE}$)**.** *Assume that for some $\sigma^* > 0$, the matrices $\mathbf{F}$ and $\mathbf{P}$ are σ^*-compressible and that for some $\sigma \in (0, \sigma^*)$ and $p := (\sigma + \frac{1}{2})^{-1}$, the system $\mathbf{F}\mathbf{u} = \mathbf{f}$ has a solution $\mathbf{u} \in \ell^p_w(\Lambda)$. Moreover, assume that $\mathbf{f}$ is σ^*-optimal. Then for all $\varepsilon > 0$, $\mathbf{u}_\varepsilon := \mathbf{modSOLVE}[\varepsilon, \mathbf{F}, \mathbf{P}, \mathbf{f}]$ satisfies*

(I) $\#\mathrm{supp}\ \mathbf{u}_\varepsilon \lesssim \varepsilon^{-1/\sigma}|\mathbf{u}|_{\ell^p_w(\Lambda)}^{1/\sigma}$,

(II) the number of arithmetic operations to compute $\mathbf{u}_\varepsilon$ is at most a multiple of $\varepsilon^{-1/\sigma}|\mathbf{u}|_{\ell^p_w(\Lambda)}^{1/\sigma}$.

Furthermore, $\|\mathbf{P}\mathbf{u} - \mathbf{u}_\varepsilon\|_{\ell^2(\Lambda)} \le \varepsilon$ and so, the recovered approximation ($D_{\mathcal{H}}^{-1}$ undoes the preconditioning, and G_Φ^* is the frame synthesis operator) satisfies $\big\|u - G_\Phi^* D_{\mathcal{H}}^{-1}\mathbf{u}_\varepsilon\big\|_{\mathcal{H}} \lesssim \varepsilon$.

One crucial assumption in Theorem 2.5 is that $\mathbf{P}$ be σ^*-compressible, because for the most obvious choice – an orthogonal projection – compressibility cannot be verified with current mathematical technology except in trivial cases [Ste03, DFR07]. However, as we will demonstrate for ridgelets below, other projections can be carefully constructed hat are in fact compressible.

The line of attack to solve the operator equation (1.2) is now clear: We have to construct a Gelfand frame Φ for the Gelfand triple $\big(H^{\vec{s}}, L^2, (H^{\vec{s}})'\big)$ and show that the resulting matrices $\mathbf{F}$ and $\mathbf{P}$ are compressible.

2.5 Ridgelet Frames

To this end, in [Gro12], a Parseval frame $\Phi = (\varphi_\lambda)_{\lambda\in\Lambda}$ of ridgelets for $\mathcal{H} = H^{\vec{s}}$ was constructed. The key to the construction is a certain set of functions $\psi_{j,\ell} \in L^2(\mathbb{R}^d)$, which form a partition of unity in the frequency domain, i.e.

$$(\psi_{j,\ell})_{j\in\mathbb{N}_0,\,\ell\in\{0,\dots,L_j\}} \quad \text{such that} \quad \sum_{j=0}^{\infty}\sum_{\ell=0}^{L_j} \hat{\psi}_{j,\ell}^2 = 1. \tag{2.3}$$

The support $P_{j,\ell} := \operatorname{supp} \hat{\psi}_{j,\ell}$ is explicitly given from the construction in (hyper-)spherical coordinates

$$P_{j,\ell} = \left\{(r,\vec{\theta}) \colon r \in [2^{j-1}, 2^{j+1}], \operatorname{dist}_{\mathbb{S}^{d-1}}(\vec{\theta}, \vec{s}_{j,\ell}) \le \alpha_j := 2^{-j+1}\right\} \tag{2.4}$$

Definition 2.6. Using (2.3), a tight frame for $L^2(\mathbb{R}^d)$ is defined by

$$\varphi_{j,\ell,\vec{k}} = 2^{-\frac{j}{2}} T_{U_{j,\ell}\vec{k}}\, \psi_{j,\ell}, \quad j\in\mathbb{N}_0,\, \ell\in\{0,\dots,L_j\},\, \vec{k}\in\mathbb{Z}^d,$$

with T the translation operator, $T_{\vec{y}} f(\cdot) := f(\cdot - \vec{y})$, and $U_{j,\ell} := R_{j,\ell}^{-1} D_{2^{-j}}$. The rotation $R_{j,\ell}$ takes $\vec{s}_{j,\ell}$ into the first canonical unit vector $\vec{e}_1$ and D_a scales the first element by a. Whenever possible, we will subsume the indices of φ by $\lambda = (j,\ell,\vec{k})$.

Remark 2.7. In $d > 3$ dimensions, rotations $R_{\vec{s}}$ turning $\vec{s}$ into $\vec{e}_1$ are no longer necessarily unique. However, for fixed $R_{\vec{s}}$, it is always possible [GO14, Lem. A.4] to choose $R_{\vec{s}'}$ such that the following Lipschitz conditions still holds,

$$\|R_{\vec{s}} - R_{\vec{s}'}\| \lesssim \operatorname{dist}_{\mathbb{S}^{d-1}}(\vec{s}, \vec{s}'). \tag{2.5}$$

To prove compressibility of $\mathbf{F}$ and $\mathbf{P}$, we will need the following assumption on the partition-of-unity. This is not an undue restriction, as we show in [GO14, Lem. B.1] that this can be satisfied by constructing $\hat{\psi}_{j,\ell}$ in a certain way, but still leaving many possibilities to choose the window functions in question.

Assumption 2.8. *The $\hat{\psi}_{j,\ell}$ are constructed in such a way, that* for any *rotation $R_{j,\ell}$ (taking $\vec{s}_{j,\ell}$ to $\vec{e}_1$), the pullbacks under the transformation $U_{j,\ell}^{-\top} = R_{j,\ell}^{-1} D_{2^j}$, have bounded derivatives* independently *of j and ℓ, i.e.*

$$\text{for } \hat{\psi}_{(j,\ell)}(\vec{\eta}) := \hat{\psi}_{j,\ell}(U_{j,\ell}^{-\top}\vec{\eta}) \colon \qquad \left\|\hat{\psi}_{(j,\ell)}\right\|_{\mathcal{C}^n} \le \beta_n, \qquad \forall\, n \le N.$$

With the ridgelet frame Φ in hand we go on to show that Φ is indeed a Gelfand frame for the Gelfand triple $\left(H^{\vec{s}}, L^2(\mathbb{R}^d), (H^{\vec{s}})'\right)$. First, we need to find suitable sequence spaces $\mathcal{H}_d$. To this end we introduce the diagonal preconditioning matrix

$$\mathbf{W}_{\lambda,\lambda'} = \begin{cases} 0, & \lambda \ne \lambda', \\ w(\lambda) := 1 + 2^j|\vec{s}\cdot\vec{s}_{j,\ell}|, & \lambda = \lambda', \end{cases} \tag{2.6}$$

and define the weighted ℓ^2-spaces

$$\mathcal{H}_d := \ell^2_{\mathbf{W}}(\Lambda) := \{\mathbf{c} \in \ell^2(\Lambda) :\ \|\mathbf{Wc}\|_{\ell^2(\Lambda)} < \infty\}$$

and the corresponding isomorphisms

$$D_{\mathcal{H}_d} \colon \begin{cases} \mathcal{H}_d & \to & \ell^2(\Lambda), \\ \mathbf{c} & \mapsto & \mathbf{Wc}, \end{cases} \qquad \text{and} \qquad D^*_{\mathcal{H}_d} \colon \begin{cases} \ell^2(\Lambda) & \to & \mathcal{H}'_d = \ell^2_{\mathbf{W}^{-1}}(\Lambda), \\ \mathbf{c} & \mapsto & \mathbf{Wc}. \end{cases}$$

Finally, we have the frame operators

$$G^*_\Phi \colon \begin{cases} \mathcal{H}_d & \to & \mathcal{H}, \\ \mathbf{c} & \mapsto & \Phi\mathbf{c}, \end{cases} \qquad \text{and} \qquad G_\Phi \colon \begin{cases} \mathcal{H}' & \to & \mathcal{H}'_d, \\ f & \mapsto & \langle \Phi, f\rangle_{\mathcal{H}'\times\mathcal{H}}, \end{cases}$$

which are bounded.

Then, as desired, we have the following result:

Theorem 2.9 ([GO14]). *The ridgelet frame Φ as constructed above constitutes a Gelfand frame for the Gelfand triple $\big(H^{\vec{s}}, L^2(\mathbb{R}^d), (H^{\vec{s}})'\big)$.*

Together with Lemma 2.1, this yields the following.

Theorem 2.10 ([GO14]). *With Φ the ridgelet system and A the differential operator defined above, consider the (infinite) matrix*

$$\mathbf{F} := \mathbf{W}^{-1}\langle A\,\Phi, A\,\Phi\rangle_{L^2}\mathbf{W}^{-1}.$$

Then the operator $\mathbf{F} : \ell^2(\Lambda) \to \ell^2(\Lambda)$ is bounded as well as boundedly invertible on its range $\operatorname{ran}(\mathbf{F}) = \operatorname{ran}\big(D_{\mathcal{H}_d}^{-1}G_\Phi\big)$. *Furthermore,* $\ker(\mathbf{F}) = \ker(G_\Phi^* D_{\mathcal{H}_d}^{-1})$.

Proposition 2.11. *The matrix*

$$\mathbf{P} := \mathbf{W}\langle \Phi, \Phi\rangle_{L^2}\mathbf{W}^{-1}$$

is a projection and satisfies $\ker(\mathbf{P}) = \ker(\mathbf{F})$.

Proof. We begin by writing

$$\mathbf{P}\colon \begin{cases} \mathcal{H} & \to & \ell^2(\Lambda), \\ f & \mapsto & D_{\mathcal{H}_d}G_\Phi G_\Phi^* D_{\mathcal{H}_d}^{-1} f. \end{cases}$$

Then

$$\mathbf{P}\mathbf{P}f = D_{\mathcal{H}_d}G_\Phi G_\Phi^* D_{\mathcal{H}_d}^{-1}D_{\mathcal{H}_d}G_\Phi G_\Phi^* D_{\mathcal{H}_d}^{-1}f = D_{\mathcal{H}_d}G_\Phi G_\Phi^* G_\Phi G_\Phi^* D_{\mathcal{H}_d}^{-1}f = \mathbf{P}f,$$

due to the fact that we have a Parseval frame – i.e. $G_\Phi^* G_\Phi g = g$. Finally, because $D_{\mathcal{H}_d}G_\Phi$ is injective due to the frame property, we have that $\ker(\mathbf{P}) = \ker(G_\Phi^* D_{\mathcal{H}_d}^{-1})$, which matches $\ker(\mathbf{F})$ by Theorem 2.10. □

2.6 Compressibility of F

The main property left to verify is the compressibility of $\mathbf{F}$ and $\mathbf{P}$ – however, compressibility is difficult to verify directly in general. Instead we use the following notion of sparsity for a (possible bi-infinite) matrix $\mathbf{A}$:

Definition 2.12. Let $p > 0$. A matrix $\mathbf{A} = \big(a_{\lambda,\lambda'}\big)_{\lambda\in\Lambda,\lambda'\in\Lambda'}$ is called *p-sparse* if

$$\|\mathbf{A}\|_{\ell^p(\Lambda)\to\ell^p(\Lambda)} := \max\bigg(\sup_{\lambda'\in\Lambda'}\sum_{\lambda\in\Lambda}|a_{\lambda,\lambda'}|^p, \sup_{\lambda\in\Lambda}\sum_{\lambda'\in\Lambda'}|a_{\lambda,\lambda'}|^p\bigg)^{\frac{1}{p}} < \infty. \tag{2.7}$$

Then, as a consequence of Schur's test, one can show:

Proposition 2.13. *Assume that $\mathbf{A}$ is p-sparse for $0 < p < 1$. Then $\mathbf{A}$ is $\frac{1}{2}\big(\frac{1}{p}-1\big)$-compressible.*

Theorem 2.14 ([GO14]). *We consider the frame $\Phi = (\varphi_\lambda)_{\lambda\in\Lambda}$, satisfying Assumption 2.8 with $N = 2n > d$ for some $n \in \mathbb{N}$, and choose $p \in \mathbb{R}$ such that $1 \geq p > \frac{d}{2n}$. For the operator A from* (1.1)*, we require that the absorption coefficient κ has a decomposition $\kappa = \gamma + \kappa_0$ with constant $\gamma > 0$, and κ_0 satisfying $\kappa_0, \hat{\kappa}_0 \in L_\infty(\mathbb{R}^d)$. Finally, we demand the existence of $r_0, c_0 > 0$, such that the decay condition*

$$\big|\hat{\kappa}_0(\vec{\xi})\big| \leq \frac{c_0}{|\vec{\xi}|^q} \qquad \forall\,\vec{\xi}\in\mathbb{R}^d\colon |\vec{\xi}| \geq r_0,$$

is fulfilled for a fixed $q > 2d + 2n + \frac{3}{2} + \frac{d-1}{p}$. Then the stiffness matrix for the operator A (with appropriate preconditioning, see (2.6)*) is p-sparse in this frame – in other words,*

$$\Big\|\mathbf{W}^{-1}\langle A\,\Phi, A\,\Phi\rangle_{L^2}\mathbf{W}^{-1}\Big\|_{\ell^p(\Lambda)\to\ell^p(\Lambda)} < \infty. \tag{2.8}$$

Remark 2.15. As we have seen in Proposition 2.13, the smaller p, the better the compressibility. The theorem is formulated in a way that p is chosen according to the restrictions imposed by d and n – however, since it is possible to construct window functions of arbitrary smoothness (and thus arbitrarily smooth $\hat{\psi}_{j,\ell}$), the limiting factor for p then becomes the decay rate of $\hat{\kappa}_0$. In the case that $\hat{\kappa}_0$ decays faster than any polynomial (say, exponentially), arbitrarily small p can be achieved (for infinitely smooth $\hat{\psi}_{j,\ell}$) – of course at the cost of exploding constants.

Instead of relying on unverified assumptions about the compressibility of the orthogonal projection, we can very sparsity (and thus compressibility) for the projection $\mathbf{P}$ defined above directly.

Theorem 2.16. *Again, let $\Phi = (\varphi_\lambda)_{\lambda\in\Lambda}$ satisfy Assumption 2.8 with $N = 2n > d$ for some $n \in \mathbb{N}$, and choose $p \in \mathbb{R}$ such that $1 \geq p > \frac{d}{2n}$. Then the projection $\mathbf{P}$ is p-sparse in this frame – in other words,*

$$\big\|\mathbf{P}\big\|_{\ell^p(\Lambda)\to\ell^p(\Lambda)} = \Big\|\mathbf{W}\langle\Phi,\Phi\rangle_{L^2}\mathbf{W}^{-1}\Big\|_{\ell^p(\Lambda)\to\ell^p(\Lambda)} < \infty.$$

Before we come to the proof, we collect two auxiliary technical results.

Proposition 2.17 ([GO14, Cor. C.3])**.** *For functions $f, g \in \mathcal{C}^{2n}$, we have the following estimate for the Laplacian in d dimensions,*

$$\big|\big[\Delta^n(fg)\big](\vec{\eta})\big| \leq (4d)^n |f(\vec{\eta})|_{\mathcal{C}^{2n}}\, |g(\vec{\eta})|_{\mathcal{C}^{2n}} \leq (4d)^n \|f\|_{\mathcal{C}^{2n}}\|g\|_{\mathcal{C}^{2n}}, \tag{2.9}$$

where $|f(\vec{\eta})|_{\mathcal{C}^{2n}} = \max_{0\leq r\leq 2n} |f^{(r)}(\vec{\eta})|$ is the maximum of all derivatives up to order $2n$ of f at $\vec{\eta}$.

Proposition 2.18 ([GO14, Prop. A.3])**.** *For fixed j' and ℓ', the number of $P_{j,\ell}$ on scale j that can intersect $P_{j',\ell'}$ is bounded as follows*

$$\#\Big\{\ell \in \{0,\dots,L\} \colon P_{j,\ell} \cap P_{j',\ell'} \neq \emptyset\Big\} \lesssim 2^{|j-j'|(d-1)}. \tag{2.10}$$

Proof of Theorem 2.16. Due to symmetry, we are able to express (2.8) without taking the maximum (cf. (2.7)),

$$\big\|\mathbf{P}\big\|^p_{\ell^p(\Lambda)\to\ell^p(\Lambda)} = \sup_{\lambda'\in\Lambda} \sum_{\lambda\in\Lambda} \Big|\frac{w(\lambda)}{w(\lambda')}\langle\varphi_\lambda,\varphi_{\lambda'}\rangle_{L^2}\Big|^p = \sup_{\lambda'\in\Lambda} \sum_{j\in\mathbb{N}_0}\sum_{\ell=0}^{L_j}\sum_{\vec{k}\in\mathbb{Z}^d} \Big|\frac{w(\lambda)}{w(\lambda')}\big\langle\varphi_{j,\ell,\vec{k}},\varphi_{j',\ell',\vec{k}'}\big\rangle_{L^2}\Big|^p < \infty. \tag{2.11}$$

Step 1 – Transforming the integral: Recalling the definition of the φ_λ, we compute

$$\begin{aligned}
\mathbf{P}_{\lambda,\lambda'} &= \frac{w(\lambda)}{w(\lambda')}\langle\varphi_\lambda,\varphi_{\lambda'}\rangle_{L^2} = \frac{w(\lambda)}{w(\lambda')}\langle\hat{\varphi}_\lambda,\hat{\varphi}_{\lambda'}\rangle_{L^2} \\
&= \quad 2^{-\frac{j+j'}{2}}\frac{w(\lambda)}{w(\lambda')}\int \overline{\mathcal{F}\big(\psi_{j,\ell}(\vec{x}-U_{j,\ell}\vec{k})\big)}\mathcal{F}\big(\psi_{j',\ell'}(\vec{x}-U_{j',\ell'}\vec{k}')\big)\,\mathrm{d}\vec{\xi} \\
&= c_{\mathcal{F}} 2^{-\frac{j+j'}{2}}\frac{w(\lambda)}{w(\lambda')}\int \hat{\psi}_{j,\ell}(\vec{\xi})\hat{\psi}_{j',\ell'}(\vec{\xi})\exp\big(2\pi\mathrm{i}\vec{\xi}\cdot(U_{j,\ell}\vec{k}-U_{j',\ell'}\vec{k}')\big)\,\mathrm{d}\vec{\xi}
\end{aligned}$$

where $c_{\mathcal{F}} = (2\pi\mathrm{i})^2$.

The transformation $U_{j,\ell}$ modifying $\vec{k}$ in the exponential function makes summing $\vec{k}$ difficult, and therefore, we will transform all the integral by $\vec{\xi} = U_{j,\ell}^{-\top}\vec{\eta}$ – introducing a factor 2^j from the determinant of the Jacobian and yielding the exponent

$$2\pi\mathrm{i}\vec{\eta}\cdot(\vec{k}-U_{j,\ell}^{-1}U_{j',\ell'}\vec{k}') = 2\pi\mathrm{i}\vec{\eta}\cdot(\vec{k}-U^{j,\ell}_{j',\ell'}\vec{k}'), \quad \text{where} \quad U^{j,\ell}_{j',\ell'} := U_{j,\ell}^{-1}U_{j',\ell'}.$$

We want to use the representation of $\hat{\psi}_{j,\ell}$ from Assumption 2.8, which holds for arbitrary rotations $\widetilde{R}_{j',\ell'}$ taking $\vec{s}_{j',\ell'}$ to $\vec{e}_1$. We choose $\widetilde{R}_{j',\ell'}$ in $\widetilde{U}_{j',\ell'} := \widetilde{R}^{-1}_{j',\ell'} D_{2^{-j}}$ in such a way that (2.5) holds for $\vec{s} = \vec{s}_{j,\ell}$ and $\vec{s}' = \vec{s}_{j',\ell'}$, and unsurprisingly, we set $\widetilde{U}^{j,\ell}_{j',\ell'} := U^{-1}_{j,\ell} \widetilde{U}_{j',\ell'}$. Thus,

$$\mathbf{P}_{\lambda,\lambda'} = c_{\mathcal{F}} 2^{\frac{j-j'}{2}} \frac{w(\lambda)}{w(\lambda')} \int \underbrace{\hat{\psi}_{(j,\ell)}(\vec{\eta}) \hat{\psi}_{(j',\ell')}(\widetilde{U}^{\top}_{j',\ell'} U^{-\top}_{j,\ell} \vec{\eta})}_{=:h_{\lambda,\lambda'}(\vec{\eta})} \exp\big(2\pi \mathrm{i} \vec{\eta} \cdot (\vec{k} - U^{j,\ell}_{j',\ell'} \vec{k}')\big) \, \mathrm{d}\vec{\eta}. \tag{2.12}$$

It should be noted that h-terms does not depend on $\vec{k}, \vec{k}'$ – however, we have chosen this notation for reasons of notational brevity.

We now have to show that the sum of (2.12) over all parameters in (2.11) is finite – which we will do for $\vec{k}$ first, then for ℓ and finally for j.

Step 2 – Integration by parts: Even though the exponent is purely imaginary, we cannot estimate the exponential function by one, as we would then sum constants in $\vec{k}$. However, a simple calculation shows $\Delta_{\vec{\eta}} \exp(2\pi \mathrm{i} \vec{\eta} \cdot \vec{y}) = -(2\pi)^2 |\vec{y}|^2 \exp(2\pi \mathrm{i} \vec{\eta} \cdot \vec{y})$, which entails

$$\Delta_{\vec{\eta}} \exp\big(2\pi \mathrm{i} \vec{\eta} \cdot (\vec{k} - U^{j,\ell}_{j',\ell'} \vec{k}')\big) = -(2\pi)^2 \big|\vec{k} - U^{j,\ell}_{j',\ell'} \vec{k}'\big|^2 \exp\big(2\pi \mathrm{i} \vec{\eta} \cdot (\vec{k} - U^{j,\ell}_{j',\ell'} \vec{k}')\big).$$

Applying Green's second identity iteratively (boundary terms disappear due to the compact support of $h_{\lambda,\lambda'}$), we will use this to generate a denominator of sufficient power to be summed over all $\vec{k} \in \mathbb{Z}^d$ – on the other hand, this forces us to estimate the derivatives of the remaining factors of the integrands. All differential operators will be with respect to $\vec{\eta}$, which we will not indicate anymore in the following.

Thus, for $\vec{k} \neq U^{j,\ell}_{j',\ell'} \vec{k}'$,

$$\mathbf{P}_{\lambda,\lambda'} = c_{\mathcal{F}} 2^{\frac{j-j'}{2}} \frac{w(\lambda)}{w(\lambda')} \frac{(-1)^n (2\pi)^{-2n}}{\big|\vec{k} - U^{j,\ell}_{j',\ell'} \vec{k}'\big|^{2n}} \int \Delta^n \big(h_{\lambda,\lambda'}(\vec{\eta})\big) \exp\big(2\pi \mathrm{i} \vec{\eta} \cdot (\vec{k} - U^{j,\ell}_{j',\ell'} \vec{k}')\big) \, \mathrm{d}\vec{\eta}. \tag{2.13}$$

Step 3 – Estimating the Derivatives: Using (2.9) and Assumption 2.8, we see

$$\big\|\Delta^n h_{\lambda,\lambda'}\big\| \leq (4d)^n \big\|\hat{\psi}_{(j,\ell)}\big\|_{\mathcal{C}^{2n}} \big\|\hat{\psi}_{(j',\ell')}\big((\widetilde{U}^{j,\ell}_{j',\ell'})^{\top} \cdot\big)\big\|_{\mathcal{C}^{2n}} \leq (4d)^n \beta^2_{2n} \big\|\widetilde{U}^{j,\ell}_{j',\ell'}\big\|^{2n}$$

which allows us to estimate (2.13), while remembering to keep the support information of terms we estimate away:

$$|\mathbf{P}_{\lambda,\lambda'}| \lesssim \frac{w(\lambda)}{w(\lambda')} \frac{2^{\frac{j-j'}{2}}}{\big|\vec{k} - U^{j,\ell}_{j',\ell'} \vec{k}'\big|^{2n}} \int_{U^{\top}_{j,\ell}(P_{j,\ell} \cap P_{j',\ell'})} \big\|\widetilde{U}^{j,\ell}_{j',\ell'}\big\|^{2n} \, \mathrm{d}\vec{\eta} \tag{2.14}$$

Step 4 – Estimating the transformation: We begin by considering the matrix $R_{j,\ell} \widetilde{R}^{-1}_{j',\ell'}$. Denoting the identity by $\mathbb{I}$, we exploit the orthogonality of the $R_{j,\ell}$ and (2.5) to yield

$$\big\|R_{j,\ell} \widetilde{R}^{-1}_{j',\ell'} - \mathbb{I}\big\| = \big\|\widetilde{R}_{j',\ell'} - R_{j,\ell}\big\| \lesssim \mathrm{dist}_{\mathbb{S}^{d-1}}(\vec{s}_{j,\ell}, \vec{s}_{j',\ell'}).$$

Due the necessary proximity of $\vec{s}_{j,\ell}$ and $\vec{s}_{j',\ell'}$ (for the intersection $P_{j,\ell} \cap P_{j',\ell'}$ to be non-empty),

$$\mathrm{dist}_{\mathbb{S}^{d-1}}(\vec{s}_{j,\ell}, \vec{s}_{j',\ell'}) \leq \alpha_j + \alpha_{j'} = 2^{-j+1} + 2^{-j'+1}, \tag{2.15}$$

we can finish the estimate,

$$\begin{aligned} \big\|\widetilde{U}^{j,\ell}_{j',\ell'}\big\| &= \big\|D_{2^j} R_{j,\ell} \widetilde{R}^{-1}_{j',\ell'} D_{2^{-j'}}\big\| = \big\|D_{2^{j-j'}} + D_{2^j}(R_{j,\ell} \widetilde{R}^{-1}_{j',\ell'} - \mathbb{I}) D_{2^{-j'}}\big\| \\ &\lesssim \max(2^{j-j'}, 1) + 2^j \mathrm{dist}_{\mathbb{S}^{d-1}}(\vec{s}_{j,\ell}, \vec{s}_{j',\ell'}) \lesssim 2^{|j-j'|}. \end{aligned}$$

Step 5 – Estimating the weights: The factor $\frac{w(\lambda)}{w(\lambda')}$ can be easily estimated, again due to (2.15). Inserting the definition, we have

$$\frac{1+2^j|\vec{s}\cdot\vec{s}_{j,\ell}|}{1+2^{j'}|\vec{s}\cdot\vec{s}_{j',\ell'}|} \leq \frac{1+2^j|\vec{s}\cdot\vec{s}_{j',\ell'}|+2^j|\vec{s}\cdot(\vec{s}_{j,\ell}-\vec{s}_{j',\ell'})|}{1+2^{j'}|\vec{s}\cdot\vec{s}_{j',\ell'}|} \leq 1+2^{j-j'}+2^j \operatorname{dist}_{\mathbb{S}^{d-1}}(\vec{s}_{j,\ell},\vec{s}_{j',\ell'}) \lesssim 2^{|j-j'|},$$

because $\left|\vec{s}-\vec{s}'\right| \leq \operatorname{dist}_{\mathbb{S}^{d-1}}(\vec{s},\vec{s}')$ for $\vec{s},\vec{s}' \in \mathbb{S}^{d-1}$.

Step 6 – Estimating the integral: For $j \geq 1$, the transformation $U_{j,\ell}^{\top}$ takes the "frequency tiles" $P_{j,\ell}$ back into a bounded set around the origin,

$$U_{j,\ell}^{\top}P_{j,\ell} \subseteq \left[\frac{1}{2}\cos(\alpha_j),2\right] \times \mathcal{P}_{(\operatorname{span}\{\vec{e}_1\})^{\perp}}\left(B_{\mathbb{R}^d}(0,4)\right) \subseteq B_{\mathbb{R}^d}(0,5).$$

This can be seen from (2.4), as $P_{j,\ell}$ is contained in the intersection between a spherical shell (between radii 2^{j-1} and 2^{j+1}) and a cone around $\vec{s}_{j,\ell}$ with opening angle $\alpha_j = 2^{-j+1}$. The rotation in $U_{j,\ell}^{\top} = D_{2^{-j}}R_{j,\ell}$ brings the axis of this cone into $\vec{e}_1$. We see that the smallest value of η_1 for $\vec{\eta} \in U_{j,\ell}^{\top}P_{j,\ell}$ is $2^{-j}2^{j-1}\cos(\alpha_j) = \frac{1}{2}\cos(\alpha_j) > \frac{1}{4}$ since $\alpha_j = 2^{-j+1} \leq 1 < \frac{\pi}{3}$ for $j \geq 1$.

The largest extent perpendicular to $\vec{e}_1$ can be calculated as

$$2^{j+1}\cos\alpha_j \sin\alpha_j = 2^j \sin 2\alpha_j \leq 2^j \cdot 2\alpha_j = 4,$$

which is what we wanted. Therefore,

$$\int_{U_{j,\ell}^{\top}(P_{j,\ell}\cap P_{j',\ell'})} \mathrm{d}\vec{\eta} \lesssim 1. \tag{2.16}$$

However, we will keep the integrals for now, as the support conveniently encodes the condition that j and j', resp. ℓ and ℓ' have to be close. Applying the other estimates to (2.14), we arrive at

$$|\mathbf{P}_{\lambda,\lambda'}| \lesssim 2^{|j-j'|(2n+\frac{3}{2})}\left|\vec{k}-U_{j',\ell'}^{j,\ell}\vec{k}'\right|^{-2n}\int_{U_{j,\ell}^{\top}(P_{j,\ell}\cap P_{j',\ell'})} \mathrm{d}\vec{\eta}$$

Step 7 – Summing $\vec{k}$: Thus far, we have omitted the case $\vec{k} = U_{j',\ell'}^{j,\ell}\vec{k}'$ – in fact, to sum over $\vec{k}$, we need treat even more elements differently. In order to estimate the term $\left|\vec{k}-U_{j',\ell'}^{j,\ell}\vec{k}'\right|$, we choose $K_{j',\ell'}^{j,\ell}\vec{k}' \in \mathbb{Z}^d$ as a (possibly non-unique) closest lattice element to $U_{j',\ell'}^{j,\ell}\vec{k}'$ (for example by rounding every component to the nearest integer), which may be interpreted as a projection of $U_{j',\ell'}^{j,\ell}\vec{k}'$ onto the lattice $\mathbb{Z}^d$. Then $\left|K_{j',\ell'}^{j,\ell}\vec{k}'-U_{j',\ell'}^{j,\ell}\vec{k}'\right| \leq \frac{\sqrt{d}}{2}$, and if we restrict $\vec{k} \in \mathbb{Z}^d$ such that $\left|\vec{k}-K_{j',\ell'}^{j,\ell}\vec{k}'\right| \geq \sqrt{d}$, it holds that

$$\left|\vec{k}-U_{j',\ell'}^{j,\ell}\vec{k}'\right| \geq \left|\vec{k}-K_{j',\ell'}^{j,\ell}\vec{k}'\right| - \frac{\sqrt{d}}{2} \geq \frac{1}{2}\left|\vec{k}-K_{j',\ell'}^{j,\ell}\vec{k}'\right|. \tag{2.17}$$

For $\vec{k} \in \mathbb{Z}^d$ such that $\left|\vec{k}-K_{j',\ell'}^{j,\ell}\vec{k}'\right| < \sqrt{d}$, we retrace the derivation of all above estimates without the partial integration, which, in effect, only eliminates the divisor $\left|\vec{k}-U_{j',\ell'}^{j,\ell}\vec{k}'\right|^{2n}$ (and reduces the implicit constants). Summing up, this means that

$$|\mathbf{P}_{\lambda,\lambda'}| \lesssim 2^{|j-j'|(2n+\frac{3}{2})}\left|\vec{k}-U_{j',\ell'}^{j,\ell}\vec{k}'\right|^{-2n}\int_{U_{j,\ell}^{\top}(P_{j,\ell}\cap P_{j',\ell'})} \mathrm{d}\vec{\eta} =: Z_{\lambda,\lambda'}$$

for $\left|\vec{k}-K_{j',\ell'}^{j,\ell}\vec{k}'\right| \geq \sqrt{d}$, and similarly for $\left|\vec{k}-K_{j',\ell'}^{j,\ell}\vec{k}'\right| < \sqrt{d}$,

$$|\mathbf{P}_{\lambda,\lambda'}| \lesssim 2^{|j-j'|(2n+\frac{3}{2})}\int_{U_{j,\ell}^{\top}(P_{j,\ell}\cap P_{j',\ell'})} \mathrm{d}\vec{\eta} = Z_{\lambda,\lambda'}$$

Note that the different cases for $\vec{k}$ are incorporated in the definition of the Z-terms.

The intention now is to prove (2.11) by showing

$$\sup_{\lambda'\in\Lambda}\sum_{\lambda\in\Lambda}\left|\mathbf{P}_{\lambda,\lambda'}\right|^p \lesssim \sup_{\lambda'\in\Lambda}\sum_{\lambda\in\Lambda}\left(Z_{\lambda,\lambda'}\right)^p < \infty.$$

We begin by summing $\vec{k}$, which crucially requires the condition $p > \frac{d}{2n}$,

$$\sum_{\substack{\vec{k}\in\mathbb{Z}^d\\ \left|\vec{k}-K^{j,\ell}_{j',\ell'}\vec{k}'\right|\geq\sqrt{d}}} \frac{1}{\left|\vec{k}-U^{j,\ell}_{j',\ell'}\vec{k}'\right|^{2np}} \overset{(2.17)}{\leq} \sum_{\substack{\vec{k}\in\mathbb{Z}^d\\ \left|\vec{k}-K^{j,\ell}_{j',\ell'}\vec{k}'\right|\geq\sqrt{d}}} \frac{2^{2np}}{\left|\vec{k}-K^{j,\ell}_{j',\ell'}\vec{k}'\right|^{2np}} = \sum_{\substack{\vec{k}\in\mathbb{Z}^d\\ \left|\vec{k}\right|\geq\sqrt{d}}} \frac{2^{2np}}{\left|\vec{k}\right|^{2np}} =: G_{d,2np} < \infty.$$

The remaining sum over $\vec{k}\in\mathbb{Z}^d : \left|\vec{k}-K^{j,\ell}_{j',\ell'}\vec{k}'\right| < \sqrt{d}$ has at most $\mathcal{O}(\sqrt{d}^d)$ terms. Taken together, this implies

$$\sum_{\lambda\in\Lambda}\left|\mathbf{P}_{\lambda,\lambda'}\right|^p \lesssim \sum_{j\in\mathbb{N}_0}\sum_{\ell=0}^{L_j} 2^{|j-j'|(2n+\frac{3}{2})p}\left(\int_{U^\top_{j,\ell}(P_{j,\ell}\cap P_{j',\ell'})}\mathrm{d}\vec{\eta}\right)^p.$$

Step 8 – Summing ℓ and j: From (2.10), we know how many intersections in ℓ are maximally possible on scale j. Therefore

$$\sum_{\lambda\in\Lambda}\left|\mathbf{P}_{\lambda,\lambda'}\right|^p \lesssim \sum_{j\in\mathbb{N}_0} 2^{|j-j'|(2np+\frac{3p}{2}+d-1)}\left(\int_{U^\top_{j,\ell}(P_{j,\ell}\cap P_{j',\ell'})}\mathrm{d}\vec{\eta}\right)^p.$$

Finally, because $|j-j'|\leq 1$ (and using (2.16)), we have bounded the left-hand side independently of λ', and therefore taking the sup doesn't change anything and the proof is finished. □

3 Main results

The results so far allow us to formulate the following corollary to Theorem 2.9, which, in essence, states that the complexity of **modSOLVE** is *linear* with respect to the number of relevant coefficients of the discretisation.

Corollary 3.1. *Assume that* $\mathbf{f}$ *is* σ^**-optimal (compare Definition 2.4) and that the system* $\mathbf{Fu}=\mathbf{f}$ *has a solution* $\mathbf{u}\in\ell^p_w(\Lambda)$ *for* $\sigma\in(0,\sigma^*)$ *and* $p:=(\sigma+\frac{1}{2})^{-1}$*. Then the solution* $\mathbf{u}_\varepsilon := \mathbf{modSOLVE}[\varepsilon,\mathbf{F},\mathbf{P},\mathbf{f}]$ *of the ridgelet-based solver recovers this approximation rate – i.e.*

$$\#\mathrm{supp}\ \mathbf{u}_\varepsilon \lesssim \varepsilon^{-1/\sigma}|\mathbf{u}|^{1/\sigma}_{\ell^p_w(\Lambda)},$$

and the number of arithmetic operations is at most a multiple of $\varepsilon^{-1/\sigma}|\mathbf{u}|^{1/\sigma}_{\ell^p_w(\Lambda)}$*.*

Finally, the last assumption – that the discretisation of typical solutions are in $\ell^p_w(\Lambda)$ – is also satisfied by the ridgelet discretisation. The proof of this theorem is based on arguments of [Can01] and is the subject of an upcoming paper [GO15].

Theorem 3.2. *For a function* $u\in L^2(\mathbb{R}^d)$ *such that* u, $\vec{s}\cdot\nabla u\in H^t(\mathbb{R}^d)$ *apart from discontinuities across hyperplanes containing* $\vec{s}$*. Then* $\mathbf{W}\langle\Phi,u\rangle_{L^2}\in\ell^p_w$*, the weak* ℓ^p*-space with* $p=(\frac{t}{d}+\frac{1}{2})^{-1}$*. This is the best possible approximation rate for functions in* $H^t(\mathbb{R}^d)$ *(even without singularities!).*

3.1 Conclusion

The bottom line is that the presented construction "sparsifies" *both* the system matrix as well as typical solutions of transport problems (in the sense of compressibility and N-term approximations, respectively), which makes it the ideal candidate for the development of fast algorithms, as underscored also by the results of Corollary 3.1.

4 Applications

4.1 Discrete Ordinates Method

As outlined in the introduction, we will now construct an algorithm to solve the full transport problem by collocation in different directions (where we can use our uni-directional solver)[1]. Therefore, we consider

$$\vec{s} \cdot \nabla u + \kappa\, u = f, \tag{4.1}$$

but this time we let $\vec{s} \in \mathbb{S}^1$ also be an independent variable such that $u, \kappa, f : \Omega \times \mathbb{S}^1 \to \mathbb{R}$. The *discrete ordinates method* (DOM) as outlined in [GS11a, Section 2] solves this problem in the following way:

- Choose some directions $\{\vec{s}_i\}_{i=1}^{N_s} \subset \mathbb{S}^1, N_s \in \mathbb{N}$.
- Solve (4.1) for these fixed directions, yielding the one-directional solutions $u'_i(x, y)$.
- Interpolate the $(\vec{s}_i, u'_i)$ to get a solution u' for the full domain $\Omega \times \mathbb{S}^1$.

The approximation error $\|u - u'\|_{L^2(\Omega\times\mathbb{S}^1)}$ depends on the quality of the interpolation – which is determined by N_s and the angular smoothness of u – on the one hand, and the error in the uni-directional solver on the other hand. Balancing these two errors gives a certain base b for choosing $N_s \sim b^J$, where, for example, $b = 2^{\frac{k}{3}}$ if the right-hand side $f \in H^k$ and linear interpolation is used (with u being sufficiently smooth in angle). The outlined discrete ordinates method – where the uni-directional solver goes up to scale J – has complexity $\mathcal{O}(N_s 4^J) = \mathcal{O}\big((4b)^J\big)$, which quickly becomes very expensive.

4.2 Sparse Discrete Ordinates Method

In order to mitigate the scaling problem of the (full) discrete ordinates method, the *sparse discrete ordinates method* (SDOM) was developed in [GS11a, Section 4]. The main idea is that the highest resolution in radius does not need the highest resolution in angle, and vice versa. Under suitable smoothness assumptions, matching decreasing resolution in one with increasing resolution in the other allows to lower the complexity of the solver to $\mathcal{O}\big(\max(4, b)^J\big)$, while losing only a logarithmic factor in accuracy compared to the full DOM (with the highest radial and angular resolutions used for the SDOM), see [GS11a, Lem. 4.3] or [EGO14].

A graphical representation of the SDOM is given in Figure 4.1 for $J = 3$ and $b = 4$ – the interpolations between the green arrows are added, while the interpolation between the red arrows is subtracted. In the upper row, the lengths of the arrows represent the number of scales that were used whereas their number and directions indicate how many and which directions are used for angular interpolation. The bottom row shows in which detail spaces the functions obtained in this way live (the J-arrow denotes increasing frame size, the S-arrow denotes increasing number of angular interpolation points).

4.3 Source Iterations

Finally, we are able to tackle the complete RTE including the scattering term:

$$\vec{s} \cdot \nabla u + \kappa\, u = f + \int_{\mathbb{S}^1} \sigma u \mathrm{d}\vec{s}'$$

This problem can be solved using the *source iteration* method, which is:

- Set $u'^{(0)}(x, y, \vec{s}) = 0$.
- For $t = 1, ..., T$, solve
$$\vec{s} \cdot \nabla u'^{(t)} + \kappa\, u'^{(t)} = f + \int_{\mathbb{S}^1} \sigma u'^{(t-1)} \mathrm{d}\vec{s}'$$
  using e.g. the DOM or SDOM based on the basic ridgelet RTE solver.

[1] A detailed account can be found in [EGO14].

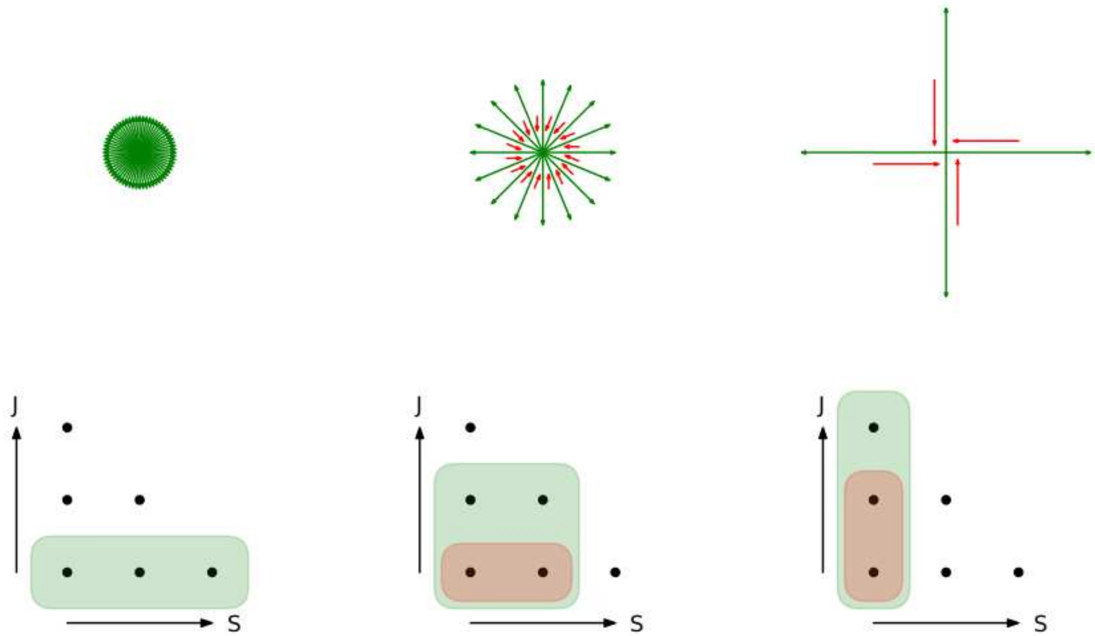


Figure 4.1: Illustration of the SDOM

The idea of the source iterations is that the $u^{(t)}$ will converge to the true solution u for large enough t – which is what we observe numerically.

Since three-dimensional functions are difficult to visualize, we will only look at the *incident radiation*

$$G[u](x,y) := \int_{\mathbb{S}^1} u(x,y,\vec{s})\,\mathrm{d}\vec{s}.$$

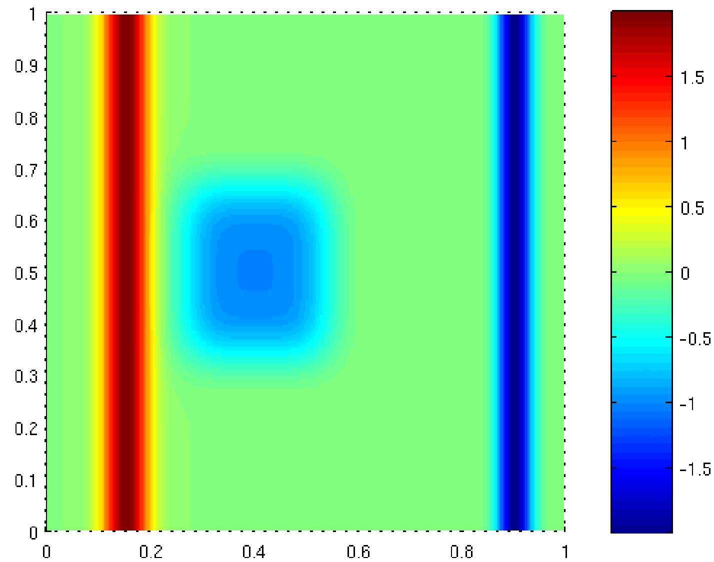
Figure 4.2 shows the incident radiation of solutions of to the complete radiative transfer problem with sources and sinks as in (a) – details below. Subfigures (b), (c) and (d) show the incident radiation for three different value of the scattering coefficient σ (0, 0.2 and 0.5, respectively). In Figure 4.2a, the f and κ are illustrated: The red line on the left shows the shape of the source term

$$f(x,y,\varphi) = \mathrm{e}^{-500\,(x-0.15)^2 - 10\min\{\varphi, 2\pi-\varphi\}^2}$$

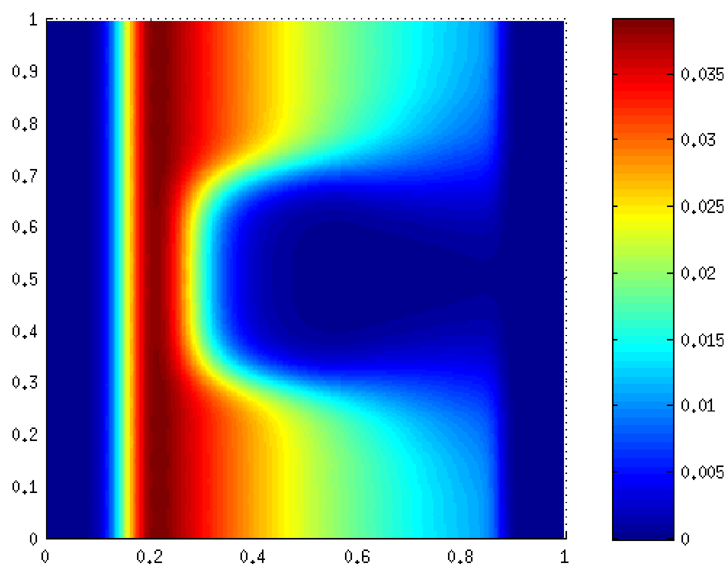
for some constant φ. The light blue area in the middle represents the obstacle which corresponds to the second term in

$$\kappa(x,y) = 2 + 18\,\mathrm{e}^{-2000\,(x-0.4)^4 - 1000\,(y-0.5)^4} + 98\,\mathrm{e}^{-900\,(x-0.9)^2}.$$

The last term in κ, shown in dark blue in Figure 4.2a, was introduced in order to avoid that radiation flows across the y-boundary.

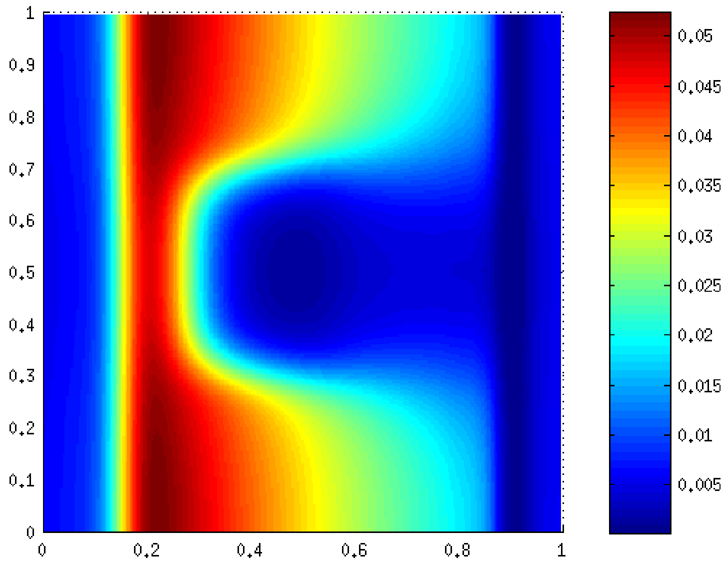


(a) The parameters of the problem

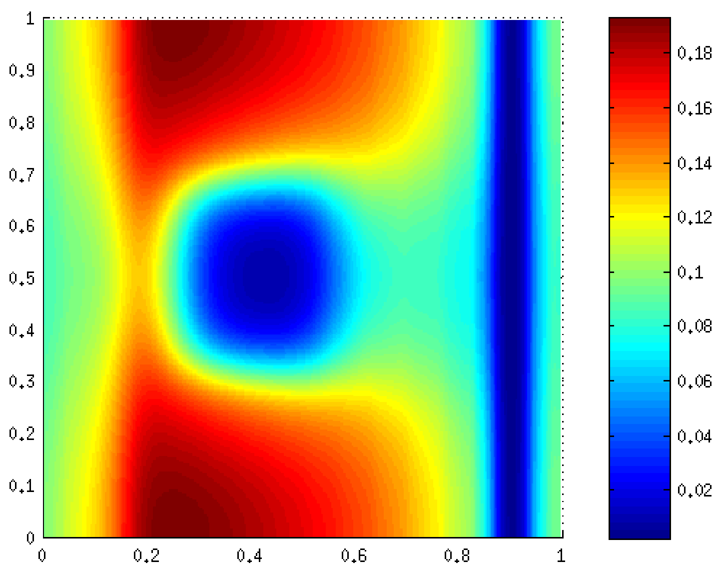


(b) Solution without scattering

Figure 4.2: (Continued on next page)



(c) Solution with $\sigma = 0.2$ after 10 source iteration steps



(d) Solution with $\sigma = 0.5$ after 40 source iteration steps

Figure 4.2: Scattering of radiation around an obstacle for different values of the scattering coefficient σ

References

[Can98] E. Candès. Ridgelets: Theory and applications. PhD thesis, Stanford University, 1998.

[Can01] E. Candès. Ridgelets and the representation of mutilated Sobolev functions. *SIAM J. Math. Anal.*, 33(2):347–368, 2001.

[CD05a] E. Candès and D.L. Donoho. Continuous curvelet transform: I. Resolution of the Wavefront Set. *Appl. Comput. Harmon. Anal.*, 19(2):162–197, 2005.

[CD05b] E. Candès and D.L. Donoho. Continuous curvelet transform: II. Discretization and frames. *Appl. Comput. Harmon. Anal.*, 19(2):198–222, 2005.

[CDD01] A. Cohen, W. Dahmen, and R. DeVore. Adaptive wavelet methods for elliptic operator equations: convergence rates. *Math. Comp.*, 70(233):27–75, 2001.

[CDDY06] E. Candès, L. Demanet, D.L. Donoho, and L. Ying. Fast discrete curvelet transforms. *Mult. Model. Simul.*, 5:861–899, 2006.

[Dau92] I. Daubechies. *Ten lectures on wavelets*, volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[DFR07] S. Dahlke, M. Fornasier, and T. Raasch. Adaptive frame methods for elliptic operator equations. *Adv. Comput. Math.*, 27(1):27–63, 2007.

[DV05] M.N. Do and M. Vetterli. The contourlet transform: an efficient directional multiresolution image representation. *IEEE Trans. Image Proc.*, 14:2091–2106, 2005.

[EGO14] S. Etter, P. Grohs, and A. Obermeier. Ffrt – a fast finite ridgelet transform for radiative transport. *submitted*, 2014.

[GK14] P. Grohs and G. Kutyniok. Parabolic molecules. *Foundations of Computational Mathematics*, 14(2):299–337, 2014.

[GKKS14] P. Grohs, S. Keiper, G. Kutyniok, and M. Schäfer. α-molecules. *Submitted*, 2014. Preprint available as a SAM Report (2014), ETH Zürich, `http://www.sam.math.ethz.ch/sam_reports/index.php?id=2014-16`.

[GO14] P. Grohs and A. Obermeier. Optimal adaptive ridgelet schemes for linear transport equations. *Submitted*, 2014.

[GO15] P. Grohs and A. Obermeier. On the approximation of functions with line singularities by ridgelets. *In preparation*, 2015.

[Gro12] P. Grohs. Ridgelet-type frame decompositions for Sobolev spaces related to linear transport. *J. Fourier Anal. Appl.*, 18(2):309–325, 2012.

[GS11a] K. Grella and Ch. Schwab. Sparse discrete ordinates method in radiative transfer. *Comput. Methods Appl. Math.*, 11(3):305–326, 2011.

[GS11b] P. Grohs and Ch. Schwab. Sparse twisted tensor frame discretization of parametric transport operators. *SAM REPORT?*, 2011.

[KL12] G. Kutyniok and D. Labate. *Shearlets: Multiscale Analysis for Multivariate Data*, chapter Introduction to Shearlets, pages 1–38. Birkhäuser, 2012.

[KLLW05] G. Kutyniok, D. Labate, W.-Q Lim, and G. Weiss. Sparse multidimensional representation using shearlets. *Wavelets XI(San Diego, CA), SPIE Proc.*, 5914:254–262, 2005.

[Mod13] M.F. Modest. *Radiative heat transfer*. Academic press, 2013.

[Ste03] R. Stevenson. Adaptive solution of operator equations using wavelet frames. *SIAM J. Numer. Anal.*, 41(3):1074–1100, 2003.

Recent Research Reports

Nr.	Authors/Title
2014-25	R.N. Gantner and Ch. Schwab Computational Higher Order Quasi-Monte Carlo Integration
2014-26	C. Schillings and Ch. Schwab Scaling Limits in Computational Bayesian Inversion
2014-27	R. Hiptmair and A. Paganini Shape optimization by pursuing diffeomorphisms
2014-28	D. Ray and P. Chandrashekar and U. Fjordholm and S. Mishra Entropy stable schemes on two-dimensional unstructured grids
2014-29	H. Rauhut and Ch. Schwab Compressive sensing Petrov-Galerkin approximation of high-dimensional parametric operator equations
2014-30	M. Hansen A new embedding result for Kondratiev spaces and application to adaptive approximation of elliptic PDEs
2014-31	F. Mueller and Ch. Schwab Finite elements with mesh refinement for elastic wave propagation in polygons
2014-32	R. Casagrande and C. Winkelmann and R. Hiptmair and J. Ostrowski DG Treatment of Non-Conforming Interfaces in 3D Curl-Curl Problems
2014-33	U. Fjordholm and R. Kappeli and S. Mishra and E. Tadmor Construction of approximate entropy measure valued solutions for systems of conservation laws.
2014-34	S. Lanthaler and S. Mishra Computation of measure valued solutions for the incompressible Euler equations.