

Adaptive Sparsification Mechanisms in Signal Recovery

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

„Doctor rerum naturalium“

der Georg-August-Universität Göttingen

im Promotionsstudiengang Mathematical Sciences

der Georg-August University School of Science (GAUSS)

vorgelegt von

Jakob Alexander Geppert

aus Hamburg

Göttingen, 2021

Betreuungsausschuss:

Prof. Dr. Gerlind Plonka-Hoch
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Prof. David Russell Luke, PhD
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Prof. Dr. Thorsten Hohage
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Mitglieder der Prüfungskommission:

Referentin:

Prof. Dr. Gerlind Plonka-Hoch
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Korreferent:

Prof. David Russell Luke, PhD
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Dorothea Bahns
Mathematisches Institut
Georg-August-Universität Göttingen

Prof. Dr. Thorsten Hohage
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Prof. Dr. Dominic Schuhmacher
Institut für Mathematische Stochastik
Georg-August-Universität Göttingen

Jun.-Prof. Dr. Anne Wald
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Tag der mündlichen Prüfung: 22. März 2021

雨の富士

霧時雨

富士を見ぬ日ぞ

おもしんき

芭蕉

Mount Fuji in Rain

A day when Fuji is unseen
Veiled on a misty winter shower—
That day, too, is a joy.

BASHŌ

Original and Translation by [BMo2].

Acknowledgements

Shaping this thesis would not have been possible without the support of many in personal, professional and institutional ways. I would like to express my gratitude to all of them.

Especially, I would like to mention my friends and colleagues Hanna, Inge, Markus, Sina and Vlada for creating an environment for scientific exchange without neglecting the importance of socializing.

Naturally I am indebted to the underlying structures of the German Research Foundation supporting me both financially and structurally. Those are the CRC 755 under the leadership of Tim Salditt and the RTG 2088 under its speaker Gerlind Plonka-Hoch. In this context I particularly thank Tim Salditt, Russell Luke and Johannes Hagemann for numerous fruitful discussions.

Specifically I want to emphasize my thanks for the extensive help from the all the offices involved with my employment. To name a few, I thank Eva Hetzel, Diana Sieber and Carmen Barann.

Encouraged by the multifold input from the support programs and electrified by the conferences I attended, it was often my advisor Gerlind Plonka-Hoch who I would consult to direct my unsystematic thoughts into workable ideas. Her open-door mentality and her unique perception of research context were decisive in the shaping of the environment of my doctorate. I am also grateful for the enriching collaboration with my co-authors, most notably Felix Kraemer and Dominik Stöger.

In this context I also want to thank the family service of the university, in particular Renate Putschbach, for helping me to find the balance between research and family as well as guiding me through all the paperwork of the related matters, such as parental leave etc.

Being able to express my thoughts in an understandable way, especially those related to mathematics, is a skill that I deem most important – both as a researcher and as a responsible human being. Thence I want to thank Strunk & White for helping me become a better writer as well as Valentin Blomer for providing a basic concept for expressing mathematical thoughts.

At times when I would struggle with my research it often was the work of Randall Munroe and Jorge Cham which helped me put things into perspective. While this might have gone unseen by any other than myself, I consider this a nonetheless important contribution.

Keen interest in mathematics as a general concept is something that has accompanied me for almost as long as I can remember. In realizing this, I want to express my gratitude to everyone who has nurtured and furthered my interest in Mathematics. Naming but a few of them, I would especially like to thank my parents, Elke Riemann and Karl-Hermann Neeb.

All my endeavours, however, would not have been possible without the continuing and unquestioning support and love of my most wonderful family. I am grateful beyond expression for the home they have created and welcome me into every single day.

Contents

Acknowledgements	ix
Contents	xi
List of Figures	xiii
List of Symbols	xv
I. Introduction	1
II. Foundation	7
1. Frames, Bases and Transforms	9
1.1. Fourier Transform	10
1.2. Frames	11
Singular Value Decomposition	13
1.3. Wavelets	15
B-Spline Wavelets	17
1.4. Shearlets	17
Continuous Shearlets	18
Discrete Shearlets	20
1.5. Fresnel Transform	21
F-Spline Wavelets	22
1.6. Fresnel Transformed Shearlets	23
2. Convex Analysis	27
2.1. Differentiability and Continuity	30
2.2. Proximity Operators	31
2.3. Operator Splitting Algorithms	34
Reflections and Relaxation	36
3. Compressed Sensing	37
3.1. Lifting	38
3.2. Restricted Isometry Property	38
3.3. Peakiness	40

III. Feature Extraction	41
4. Nested Proximity Operators	43
4.1. Frame Shrinkage	44
Tailoring Operator Splitting Methods	46
Analyzing The Frame Soft Shrinkage Operator	48
Frame Soft Thresholding is a Proximity Operator	55
4.2. Proximity Operators Nested Into Injective Operators	60
Proximity Operators and Linear Operators	61
Frame Soft Shrinkage	64
5. Sparse Rank-One Matrix Recovery	65
5.1. Algorithms	68
Pursuit Algorithms	68
Sparse Power Factorization	69
Initialization	70
5.2. Assessing Locality	71
5.3. Adaptive Recovery	78
IV. Application: X-Ray Nanoscale Imaging	83
6. Physical Circumstances	85
6.1. Propagation of Light	85
Maxwell and Helmholtz Equations	86
Fresnel Propagation	87
6.2. Phase Retrieval	89
Exploitable Priors for Recovery	90
Noise	92
7. Algorithmic Approaches	93
7.1. Relaxed Douglas Rachford	93
Adjusting for Phase Retrieval	93
Extending the Proof of Looock	94
Outlook	95
7.2. Matrix Recovery	96
Alternative	97
7.3. Further Approaches	98
Transport of Intensity Equation	98
V. Conclusion	99
Bibliography	101
Curriculum Vitae	113

List of Figures

4.1. Visualization of H and Φ	50
6.1. Cell Image and Simulation	90
6.2. Recovery Priors	91
7.1. Demonstrating the Lifting Process for Fourier Magnitudes	97

List of Symbols

General

i	Imaginary unit $\sqrt{-1}$
id	Identity operator or identity matrix
\mathbb{R}_+	Non-negative real numbers, $[0, \infty)$
C^∞	Space of smooth functions
$\mathcal{CN}(\mu, \Gamma)$	Complex normal distribution
$\mathbb{E}[X]$,	Expected value of the random value X
$x^{(j)}$	Variable x in the j -th step of an iterative algorithm
$\text{GL}_n(\mathbb{R})$	General linear group, i.e. invertible matrices in $\mathbb{R}^{n \times n}$
$\delta_{j,k}$	Kronecker Delta, $\delta_{j,k} = 1$ if $j = k$ and $\delta_{j,k} = 0$ if $j \neq k$.

Matrices and Vectors

Let $A \in \mathbb{C}^{m \times n}$ and $x \in \mathbb{C}^n$.

x_i	i th entry
x_J	Projection onto the index set $J \subseteq \{1, \dots, n\}$, i.e. $[x_J]_i = x_i$ if $i \in J$ and $[x_J]_i = 0$ if $i \notin J$
x^*	Non-increasing rearrangement, i.e. $\{x_j\}_j = \{x_j^*\}_j$ and $ x_j^* \geq x_{j+1}^* $
$\text{sgn}(x)$	Signature vector, $\text{sgn}(x)_k = +1$ if $x_k > 0$, $\text{sgn}(x)_k = -1$ if $x_k < 0$, $\text{sgn}(x)_k = 0$ if $x_k = 0$
$\text{supp}(x)$	Support, $\text{supp}(x) := \{1 \leq j \leq n : x_j \neq 0\}$
$\ x\ _p$	ℓ^p -norm for $p \in \mathbb{N}$, $\ x\ _p := (\sum_{k=1}^n x_k ^p)^{1/p}$
$\ x\ _\infty$	Chebyshev distance, $\ x\ _\infty := \max_{1 \leq j \leq n} x_j $
$\ x\ _0$	So-called 0-norm, $\ x\ _0 := \#\text{supp}(x)$
A^T	Transposed matrix
A^*	Transposed and complex conjugated matrix
a_i	i th column of A for any $1 \leq i \leq n$
$a_{i,j}$	Entry in i th row and j th column of A for $1 \leq i \leq m$ and $1 \leq j \leq n$
$\ A\ _F$	Frobenius norm for square matrices A . $\ A\ _F^2 := \sum_{i,k=1}^n a_{i,j} ^2$
$\ A\ _2$	Spectral norm for square matrices A . $\ A\ _2 := \max\{\ Au\ _2 : \ u\ _2 = 1\}$

Functions

Let $f: X \rightarrow Y$ a function.

$\text{epi}(f)$	Epigraph, $\text{epi } f := \{(x, \xi) \in X \times Y : f(x) \leq \xi\}$ for a totally ordered space Y , e.g. $Y = \mathbb{R}$
$\Gamma(f)$	Graph, $\Gamma(f) := \{(x, \xi) \in X \times Y : f(x) = \xi\}$
$\mathcal{R}(f)$	Range, $\mathcal{R}(f) := f(X)$
$\mathcal{N}(f)$	Kernel or zero set, $\mathcal{N}(f) := \{x \in X : f(x) = 0\}$
$\text{supp}(f)$	Support, $\text{supp}(f) := \{x \in X : f(x) \neq 0\}$

List of Symbols

Sets

Let S a set.

ι_S	Indicator function $\iota_S(x) = 0$ if $x \in S$ and $\iota_S(x) = \infty$ else
\mathcal{P}_S	Projection onto the set
2^S	Power set

Lebesgue Spaces

\mathcal{L}^2	Hilbert space of square integrable functions and inner product $\langle f, g \rangle = \int f(x)\overline{g(x)} dx$
ℓ^p	Sequence space with norm $\ x\ _p = (\sum x_k ^p)^{1/p}$ for $p \in \mathbb{N}$
ℓ^∞	Sequence space of essentially bounded sequences with norm $\ x\ _\infty = \sup(x)$
ℓ^0	Pseudo-space with pseudo-norm $\ x\ _0 = \#\{j : x_j \neq 0\}$
$B(x, r, \ell^p)$	Closed ball around x with radius r and norm ℓ^p for $p \in \mathbb{N} \cup \{0, \infty\}$, $B(x, r, \ell^p) := \{t : \ x - t\ _p \leq r\}$

Transforms

Let $x \in \mathbb{C}^n$ and $f \in \mathcal{L}^2(\mathbb{R}^n)$.

ω_n	n -th root of unity, $\omega_n := \exp(-2\pi i/n)$
F_n	Fourier matrix
\hat{x}	Discrete Fourier transform, $\hat{x} := F_n x$
\hat{f}	Continuous Fourier transform
K_τ	Fresnel kernel for $\tau > 0$
\tilde{f}_τ	Fresnel transform, $\tilde{f}_\tau := f * K_\tau$

Hilbert Spaces

\mathcal{H}, \mathcal{K}	Generic Hilbert spaces
$\langle \cdot, \cdot \rangle$	Canonical inner product, unless specified otherwise
$\{e_j\}_j$	Canonical basis, if applicable. E.g. unit vectors in \mathbb{R}^n .
$\mathcal{B}(\mathcal{H}, \mathcal{K})$	Group of bounded linear operators from \mathcal{H} to \mathcal{K}
$\mathcal{U}(\mathcal{H})$	Group of of unitary operators
Φ^*	Adjoint of Φ

Part I.

Introduction

It may be one of mankind's most profound desires to “understand whatever; Binds the world's innermost core together”, [vGo3, ll. 382 – 383] to put it in the words of Goethe. This ambition has also for centuries been accompanied by beliefs that most – if not all – observable occurrences appertain to a deeper, more fundamental truth.¹ That in turn redounds to a redundancy in description of a class of objects that are similar to each other in a suitable manner.

This philosophical attitude has manifested itself in many ubiquitous phrases and expressions such as the *Sparsity of Effects Principle*, *Occam's Razor*, or the *Pareto Principle*. All these share the common belief that most information already is contained in a small fraction of the data that we actually observe, cf. e.g. [STW18, Sch14, New05]. Mathematically speaking, this translates into the persuasion that naturally occurring data allows for sparse encoding with respect to a suitable, possibly overcomplete, representation system, e.g. a frame, cf. e.g. [BP13].

The aim to find, assess and make sense of underlying structure has once more gained momentum in the last decades, not least alongside the emergent usage of neural networks, deep learning and, generally, the omnipresence of the buzzwords *Big Data* or *Data Science*, cf. e.g. [AÖ17, Sed18].

It has thusly become necessary for modelling schemes of all kinds to be able to incorporate sparsity as either a prior, a goal or a measure of success. Depending on the context one might also call these systems recovery or retrieval algorithms as they aim to extract the ground truth from the given observations. However, depending on the situation at hand and the inherent data structure, off-the-shelf procedures may fail to recognize the underlying patterns. Hence, there has been steady interest in the development of suitable adaptive algorithms.

New Results

In this thesis I consider two different adaptive recovery concepts in Part III and one application example in X-ray nanoscale imaging in Part IV.

¹This reasoning starts to drift into the direction of philosophy and religion or – more to the point – epistemology, and dates back millennia. A prominent example might be the discussion of the *Essence* or, literally, *Being – εἶναι* – of Plato in ancient Greek philosophy, cf. [RS08]. As this is clearly besides the topic of this thesis, I will not develop this thought any further.

The first concept is concerned with proximity operators intertwined between an injective linear operator T with bounded range and its pseudoinverse T^\dagger . Proximity operators are a versatile tool in convex analysis and are defined as

$$\text{prox}_f := \arg \min_{z \in \mathcal{H}} \left[f(z) + \frac{1}{2} \|z - \cdot\|^2 \right]. \quad (1)$$

Here, f is commonly a proper, lower semi-continuous, convex (plc) function and called the *prior* of the proximity operator. Proximity operators allow for several intuitions. One is to consider Equation (1) as a penalized least-squares problem, [GN20]. Another way would be to describe proximity operators as resolvents of subdifferentials, [BC11]. A prominent example in frame analysis is $\text{prox}_{\gamma \|\cdot\|_1}$ which coincides with the soft shrinkage operator \mathcal{S}_γ , [Bec17, GN20], and hence can be employed for sparsification. However, this will only lead to sparsity with respect to the canonical basis which in many applications is not helpful. Instead, one rather would like to have sparsity with respect to a suitable encoding, e.g. a frame, [Chr16, CK13, ZS14]. This motivates to nest the proximity operator into a frame encoding and decoding step, respectively. For an encoding T and prior f this reads

$$T^{-1} \text{prox}_f T,$$

which is impractical due to overcomplete frames being full-rank, but singular. Thence, we decided to use the Moore-Penrose pseudoinverse instead, yielding

$$T^\dagger \text{prox}_f T. \quad (2)$$

Clearly, this models the application of the specified proximity operator on the signal when encoded with respect to T . This immediately leads to the question whether this expression is a proximity operator itself. Besides direct applications, e.g. in customized denoising schemes, this property is crucial for the usability within standard algorithms, such as Douglas-Rachford or forward-backward splitting.

First, I prove that in finite dimensions this is indeed the case. The main theorem of that section, Theorem 4.10 states that for a frame $T \in \mathbb{R}^{L \times N}$ of full rank, the operator $T^\dagger \mathcal{S}_\gamma T$ is the proximity operator of a plc function. The key to this result is the introduction of a new inner product

$$\langle x, y \rangle_T := \langle Tx, Ty \rangle$$

that is weighted with the frame T . Due to this, we get a changed notion of (cyclical) monotonicity which through Rockafellar's *Cyclic Monotonicity Theorem*, Theorem 2.8, Minty's Theorem, Theorem 2.5, and Moreau's *Théorème des Fonctions Duales*, Corollary 2.15, carries over to the notion of proximity operators.

Afterwards we extend the result of Theorem 4.10 to arbitrary Hilbert spaces and injective operators with closed range, respectively, in Theorem 4.12. The main ingredient is again the weighted inner

product introduced in the previous section. This time, this is combined with an analysis of the Fréchet derivative in the aim to represent

$$T^\dagger \text{prox}_f T = \nabla \Psi,$$

i.e. writing the nested proximity operator as a Fréchet derivative with respect to the weighted inner product. Finally, in Theorem 4.13 we are moreover able to explicitly describe a plc function g subject to

$$T^\dagger \text{prox}_f T = \text{prox}_g \tag{3}$$

in terms of an infimal convolution.

Concludingly, we see that by proving that the expression $T^\dagger \text{prox}_f T$ is a proximity operator, we have created an adaptable tool that gives rise to adaptive algorithms when integrated, e.g., into larger recovery schemes. A brief example of doing so is given in Section 7.1 in the application in Part IV.

The second concept studied in this thesis concerns doubly sparse recovery from potentially noisy bilinear measurements. That is, we consider measurements of input vectors $u \in \mathbb{C}^{n_1}$ and $v \in \mathbb{C}^{n_2}$ of the form

$$b(u, v) + z$$

for a bilinear map $b: \mathbb{C}^{n_1} \times \mathbb{C}^{n_2} \rightarrow \mathbb{C}^m$ and a perturbation term z . Through lifting, this can be modelled as a linear problem of the outer product uv^* . In Chapter 5 we consider linear operators $\mathcal{A}: \mathbb{C}^{n_1 \times n_2} \rightarrow \mathbb{C}^m$. Expectably, we assume

$$\mathcal{A}(uv^*) = b(u, v).$$

More importantly, we will consider operators \mathcal{A} that satisfy an (s_1, s_2, r) -restricted isometry property (RIP), i.e.,

$$(1 - \delta) \|X\|_F^2 \leq \|\mathcal{A}(X)\|_2^2 \leq (1 + \delta) \|X\|_F^2$$

for all matrices $X \in \mathbb{C}^{n_1 \times n_2}$ of rank at most r with at most s_1 non-zero rows and at most s_2 non-zero columns. This condition can be understood as a measure of how close an operator \mathcal{A} is to unitarity. Most applications use the fact that classes of random matrices satisfy this property with high probability, e.g. Gaussians.

Recovering a column- and row-sparse rank-one matrix X from its RIP measurement $\mathcal{A}(X)$ is still a very hard problem due to the competing objectives. Also, bilinear inverse problems with sparsity constraints are still a relatively new field of study. We follow the approach of Lee et. al in [LWB18], which is able to guarantee successful recovery for signals whose singular vectors u and v satisfy $\|u\|_\infty \geq 0.78\|u\|_2$ and $\|v\|_\infty \geq 0.78\|v\|_2$. This does however imply that more than half of the mass already has to be contained in the largest entry of u and v , respectively, rendering applicability impossible for many settings in practice.

We solve this problem by providing a tradeoff between the peak to average power ratio and the number of required measurements in Theorem 5.13. For somewhat peaky signals with sparsity $\|u\|_0 = \|v\|_0 =: s$ this allows recovery with measurements in the order of $s^{3/2}$ as compared to arbitrary sparse signals with an established order of s^2 , cf. [LS15]. We also relax the stiff focus on the single-largest entry by introducing a new norm

$$\|x\|_{[k]} := \max_{\substack{I \subseteq \{1, \dots, n\} \\ |I|=k}} \left(\sum_{i \in I} |x_i|^2 \right)^{1/2}$$

for $x \in \mathbb{C}^n$ and $k \leq n$ which allows us to assess the fraction of the mass that is contained in the k largest entries of a given signal and scale the number of necessary measurements correspondingly.

Our new idea to introduce this twofold relaxation of requirements for guaranteed recovery provides us with an easily customizable algorithm that allows adaptive tailoring for the use case at hand.

The application part discusses applicability of the aforementioned recovery schemes in the context of phase retrieval. This is motivated by the phase loss problem inherent to diffractive imaging, e.g. X-ray nanoscale imaging. Firstly, I improve the approach by Looock, [LP14, Loo16], not least by providing a way to extend his proof to the case of arbitrary frames by the results of this work. Secondly, I analyze that phase retrieval, e.g. from Fourier magnitudes, can also be modelled as a sparse rank-one recovery problem. Unfortunately, albeit the Fourier transform being unitary, the phase loss destroys the length-preservation and does not even permit an RIP. In the sense of good scientific practice, and especially to counteract publication bias, I include this negative result here.

Structure

This thesis structured into five parts. Parts I and V cover the introduction and conclusion while Parts II to IV constitute the scientific part of this dissertation.

In Part II, I present a collection of concepts that I require to formulate and solve the questions of the following parts. Large portions of this part are cited from adequate literature as the definitions are well understood and most statements are direct consequences thereof. To this end, Chapter 1 is concerned with an introduction to Fourier analysis and frame theory as well as their extensions to wavelets, shearlets and the Fresnel transform. Followingly, Chapter 2 focusses on convex analysis and contains the notions of plc functions, proximity operators as well as extended notions of differentiability and continuity. This chapter also gives a very short prospect of operator splitting algorithms by the examples of forward-backward and Douglas-Rachford splitting. Lastly, Chapter 3 contains the compressed sensing concept of lifting along with the notions of the restricted isometry property and signal peakiness.

Most of my own contribution is located in Part III. Here I am concerned with addressing the task of adaptive sparsification mathematically and develop the new results stated above.

Chapter 4 considers nested proximity operators, i.e. expressions of the form $T^\dagger \text{prox}_f T$. I show that in suitable settings this again is a proximity operator. In Section 4.1, I consider the finite dimensional case and concentrate on soft shrinkage, i.e. in the notation of Equation (2), we let $f = \gamma \|\cdot\|_1$. Besides proving that frame soft shrinkage is a proximity operator, I also present some geometrical considerations aiming to provide heuristically graspable intuition, at least for small examples. In Section 4.2, we abandon this approach and instead aim for utmost generality. This also allows for description of the newly found prior g of Equation (3). As both sections of this chapter are based on individual manuscripts, [GP19, HNP⁺19], each will begin with a short introduction of its own to the specific focus of the respective paper.

Chapter 5 addresses sparse bilinear recovery which, through lifting, is modelled as doubly sparse rank one matrix recovery. Working with the Sparse Power Factorization algorithm of [LWB18] we develop relaxed sufficient recovery criteria. To this end, we analyze the proof ideas of Lee et al. and piecewisely improve or replace their technical lemmas and proofs in order to feed the improved criteria into their existing framework. As large portions of this chapter have been published in [GKS19], I have retained the introduction thereof at the beginning of this chapter for an appropriate embedding of this chapter into the corresponding context of compressed sensing.

In Part IV, I consider methodic approaches in the setting of X-ray nanoscale imaging. Here, Chapter 6 gives a brief outline on mathematical modelling of the propagation of light in the sense of an electromagnetic wave. Moreover, we see that phase loss is inherent to diffractive imaging and consequently look for exploitable priors to overcome the illposedness of the recovery problem. In Chapter 7, we follow two approaches of addressing phase retrieval, by a relaxed Douglas-Rachford iteration in the sense of [LP14] and via matrix recovery in the sense of [LYB13, CESV13].

Publications

Some parts of this thesis are based on the manuscripts [GP19, GKS19, HNP⁺19], of which I am a coauthor. In each case I will explicitly state the beginning of sections containing those excerpts. The content will be largely identical to the ones in the corresponding manuscripts except for, e.g., adaptations to fit the structure of this thesis or explanatory running text.

In this thesis, Section 4.1 is based on [GP19]. This manuscript evolved from the study of Stefan Loock's dissertation, [Loo16], upon which I noticed a flawed proof. When attempting to correct his proof, I found that by the new approach the assertion could be considerably generalized. Thus, Gerlind Plonka and I developed the idea of the adapted inner product that consequently gave rise to the core theorem of [GP19], i.e. Theorem 4.10. This manuscript has been submitted and is currently awaiting reviews.

Section 4.2 is based on [HNP⁺19]. The idea for this manuscript arose in the communication with Gabriele Steidl concerning the previous work, [GKS19]. Having brought along further researchers led this discussion into a new direction of more generality and less understanding by means of heuristics or geometry. Notwithstanding that, this correspondency amounted to [HNP⁺19]. After having finished this manuscript, some of the authors changed the focus of this work towards neural networks. Although carrying over portions of [HNP⁺19], this lead to a new work under a new title which has been published in [HHN⁺20]. Due to the not inconsiderable overlap, we decided not to pursue publication of [HNP⁺19].

Chapter 5 is based on [GKS19]. I have been working with the Sparse Power Factorization (SPF) algorithm since my Bachelor's thesis, [Gep15]. There, I showed applicability of different pursuit algorithms as part of SPF and, in correspondence with Kiryung Lee, the principal author of [LWB18], I found technical flaws in their early manuscript, [LWB13]. In my Master's thesis, [Gep17], I worked on generalized sufficient convergence criteria of SPF. Unfortunately, I encountered technical difficulties which led to an incorrect result. Under the lead of Dominik Stöger and together with Felix Krahmer, we were able to overcome these. Having presented and discussed the result on two conferences, [GKS17, SGK18], we published our manuscript as [GKS19].

Part II.

Foundation

This part lays the formal foundation for this thesis. More precisely, we will introduce the notations, definitions, notions and concepts that we will need later to formulate and address our main goals. Thereby, we especially provide detailed explanations to large parts of the symbols presented in the [List of Symbols](#).

First, in Chapter 1, we consider a suitable choice of bases, frames and transforms. To begin with, we fix the notation for the discrete and continuous Fourier transform in Section 1.1. While we will assume basic knowledge of the reader in this topic, there exist numerous variants of notation and normalization, prompting the necessity of this explicit introduction. Then, we introduce the concept of frames as an extension of bases in Section 1.2. In the process, we explicitly focus on the singular value decomposition of finite-dimensional frames, i.e. rectangular matrices. Following, we exemplify the notion of frames by two prominent representatives, wavelets and shearlets in Sections 1.3 and 1.4. In the latter case we will motivate the continuous version first and then come to terms with a discretization. Concludingly, we state the Fresnel transform in Section 1.5 and interlinkingly study its behavior with respect to preserving the structures of spline wavelets and, in Section 1.6, of shearlets.

Chapter 2 concerns itself with the introduction of some tools that one may associate with usage in convex analysis. This section covers three topics. First, Section 2.1 provides extension to the basic differentiability and continuity notions from introductory calculus, covering, e.g., Lipschitz continuity as well as Fréchet differentiability and subdifferentiability. Second, Section 2.2 introduces the proximity operator and states some properties that are directly derivable from its definition. Third, Section 2.3 is concerned with the operator splitting methods of forward-backward splitting and Douglas-Rachford splitting. Here, these are stated in a rather general manner, allowing, e.g., to derive a relaxed version based on averaged alternating reflections as a special case.

Lastly we introduce some basic concepts of compressed sensing in Chapter 3. In this context, we first shortly demonstrate the so-called lifting procedure which allows to reformulate a bilinear inverse problem as a linear problem of higher dimension in Section 3.1. Thereafter, we state the restricted isometry property in Section 3.2. Playing a key role in Chapter 5 in Part III, this property can be understood as a measure how close a given operator is to a unitary one. Finally, we discuss a quantification method for signal peakiness, that is a measure stating how much of a signal's mass is concentrated in few, large entries in Section 3.3.

1. Frames, Bases and Transforms

In this chapter we introduce the tools and notions of numerical Fourier analysis we will use in Parts III and IV.

As the name of Fourier analysis suggests, we will first introduce the Fourier transform itself in the discrete and continuous setting, Section 1.1. In this thesis we are not interested in Fourier series.

Following historical developments, one may base the wavelet transform on the Fourier transform. While we recognize this, we present the wavelet transform in Section 1.3 based on a multiresolution analysis. The wavelet construction in this thesis is knowingly not comprehensive, but will suffice for understanding of its interaction with the Fresnel transform at the end of this chapter.

As a further improvement of the wavelet transform, especially for two-dimensional, anisotropic signals, we will look at the shearlet transform in Section 1.4. Here, we follow [KL12] to provide a broader introduction by first considering the continuous transform being based on representation theory and then take a glimpse at discretization efforts undertaken by Kutyniok and others.

Both wavelets and shearlets give rise to representation systems, e.g. of $\mathcal{L}^2(\mathbb{R}^n)$. Depending on the choice of parameters, base functions etc., these systems can be complete or overcomplete. In the latter case, we call those systems *frames*. The corresponding theory is introduced in Section 1.2 prior to wavelets and shearlets in order to provide the necessary framework for an easier understanding.

Afterwards, we present the Fresnel kernel and Fresnel transform in Section 1.5. The Fresnel kernel is dominated by its exponential part that, neglecting constant terms in the argument, is of the form $\exp(\|x\|_2^2)$. It is hence straightforward to see a relationship to the Fourier transform. This observation is manifested later in Section 6.1 where we discuss usage of Fresnel and Fourier magnitudes as means of describing measurable intensities of electromagnetic waves. We then go on to apply the Fresnel transform to B-spline wavelets and shearlets. The former will result in a new kind of wavelets, so-called F-spline wavelets.

Finally, we pick up Lieblings idea of F-splines and try to adopt this concept for shearlets in Section 1.6. Here, we analyze which characteristics of shearlet functions are retained in the process of convolving with the Fresnel kernel and discuss whether this can yield a reasonable encoding scheme.

1.1. Fourier Transform

Based on n -th roots of unity, the Fourier basis lays the foundation to the discrete Fourier transform, cf. [Wonn1]. Note that as opposed to the usual convention in Fourier analysis, we will index vectors and matrices starting with 1 in order to stay consistent throughout this thesis.

1.1 Definition.

Let $n \in \mathbb{N}$. Denote ω_n the n -th root of unity, i.e.

$$\omega_n := \exp\left(\frac{-2\pi i}{n}\right).$$

We then define the *Fourier Matrix* $F_n \in \mathbb{C}^{n \times n}$ component-wise as

$$F_n := \left(\omega_n^{(j-1)(k-1)} \right)_{j,k=1}^n.$$

It is well known and easy to see that the columns of F_n provide an orthogonal basis of \mathbb{C}^n . More precisely, $n^{-1/2}F_n$ is unitary and commonly called *Discrete Fourier Basis*.

This lets us define the *Discrete Fourier Transform* (DFT) as the map $x \mapsto F_n x$. This definition easily yields the corresponding inversion theorem.

1.2 Theorem (Fourier Inversion Theorem).

We have

$$F_n^{-1} = \frac{1}{n} F_n^*.$$

We further introduce the continuous Fourier transform as an operator on \mathcal{L}^2 . We will assume some basic familiarity of the reader with Fourier analysis. Hence, we will not concern ourselves with the derivation of the Fourier transform or the reasoning on which spaces it is well defined. The interested reader finds a detailed introduction in [PPST18].

1.3 Definition (Cf. [Kre15]).

The *Continuous Fourier Transform* $\mathcal{F}: \mathcal{L}^2(\mathbb{R}^n) \rightarrow \mathcal{L}^2(\mathbb{R}^n)$ is defined as

$$\widehat{f}(\xi) := \mathcal{F}[f](\xi) = \int_{\mathbb{R}^n} f(x) \exp(-2\pi i \langle \xi, x \rangle) dx.$$

As special cases we introduce the lateral and spectral Fourier transform in \mathbb{R}^3 , which apply the continuous Fourier transform to only the first two or the last dimensions, respectively. This will be relevant later in modelling propagation of light in the direction of the third axis in Section 6.1.

1.4 Definition (Cf. [Kre15]).

Let $f \in \mathcal{L}^2(\mathbb{R}^3)$. We then define its *Spectral Fourier Transform* as

$$\mathcal{F}_{\text{sp}}[f](x_1, x_2, \omega) := \int_{\mathbb{R}} f(x_1, x_2, x_3) \exp(-2\pi i \omega x_3) dx_3$$

and its *Lateral Fourier Transform* as

$$\mathcal{F}_{\perp}[f](\omega_1, \omega_2, x_3) := \int_{\mathbb{R}^2} f(x_1, x_2, x_3) \exp(-2\pi i \langle \omega, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \rangle) d\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

The continuous Fourier transform also possesses a well known inversion theorem, cf. [Kre15, Goo05].

1.5 Theorem (Fourier Inversion Theorem).

The inverse of the continuous Fourier transform is given by

$$\mathcal{F}^{-1}[\widehat{f}](x) = f(x) = \int \widehat{f}(\xi) \exp(2\pi i \langle \xi, x \rangle) d\xi.$$

1.2. Frames

In numerous contexts, it is necessary to wisely choose an overcomplete representation system. Usually these are called *Frames*. The interested reader finds an ample introduction in [CKP13, Chr16]. From there, we repeat the following basics. Unless stated otherwise, all content of this Section 1.2 is taken from there.

We begin by stating the formal definition of a frame as an (over-)complete system.

1.6 Definition.

For a Hilbert space \mathcal{H} , a *Frame* is a sequence $\{f_k\}_{k=1}^{\infty} \subseteq \mathcal{H}$, if there exist constants $A, B > 0$ such that for all $f \in \mathcal{H}$

$$A\|f\|^2 \leq \sum_{k=1}^{\infty} |\langle f_k, f \rangle|^2 \leq B\|f\|^2.$$

We call A and B the *Frame Bounds* of $\{f_k\}$.

Having introduced the notion of frames, we further define the corresponding analysis and synthesis operators.

1.7 Definition.

Let $\{f_k\}$ a frame. Denote its *Analysis Operator* T and *Synthesis Operator* T^* by

$$T: \mathcal{H} \longrightarrow \ell^2 \qquad T^*: \ell^2 \longrightarrow \mathcal{H}$$

$$f \longmapsto (\langle f, f_k \rangle)_{k=1}^{\infty} \qquad \text{and} \qquad (a_k)_{k=1}^{\infty} \longmapsto \sum_{k=1}^{\infty} a_k f_k.$$

We call the product $T^*T: \mathcal{H} \longrightarrow \mathcal{H}$ the *Frame Operator* of $\{f_k\}$.

It is easy to see that the frame operator T^*T is invertible on \mathcal{H} , cf. [Chr16]. As a consequence, we can reconstruct f from its frame coefficients, i.e.

$$f = \sum_{k \in \mathbb{N}} \langle \cdot, f_k \rangle (T^*T)^{-1} f_k,$$

which readers familiar with functional analysis will easily recognize as the dual frame. For further interest in this topic, I recommend the reading of [Rud91].

1.8 Definition.

Let $\{f_k\}$ a frame with analysis operator T and synthesis operator T^* . Then its *Canonical Dual Frame* is defined as the sequence $\{(T^*T)^{-1} f_k\} \subseteq \mathcal{H}$.

In order to get an intuition of finite-dimensional frames, let us now consider a simple example. Let $\mathcal{H} = \mathbb{R}^n$ and $F \in \mathbb{R}^{m \times n}$, where $m \geq n$, a matrix of full rank. Then, clearly, the rows $\{f_1, \dots, f_m\} \subseteq \mathbb{R}^n$ of F form a frame. We obtain the frame coefficients in the form of a vector in \mathbb{R}^m by multiplication with F , i.e. the frame analysis of a vector $x \in \mathbb{R}^n$ can be written as

$$Tx = \begin{pmatrix} \langle x, f_1 \rangle \\ \langle x, f_2 \rangle \\ \vdots \\ \langle x, f_m \rangle \end{pmatrix} = Fx.$$

Analogously, the frame synthesis operator is realized as multiplication with F^* , i.e. for a coefficient vector $a \in \mathbb{R}^m$ we see

$$T^*a = \sum_{k=1}^m a_k f_k = \begin{pmatrix} | & | & & | \\ f_1 & f_2 & \dots & f_m \\ | & | & & | \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} = F^*a.$$

It is hence not uncommon to use the symbol T to denote both a frame and its analysis operator.

A special class of frames are so-called tight frames that are characterized by having a frame operator which is a multiple of the identity.

1.9 Definition.

A frame is called *Tight* if the frame bounds coincide, i.e. $A = B$. A tight frame is called *Parseval Frame* if $A = B = 1$.

We will now concern ourselves with the pseudoinverse operator for operators that are not directly invertible.

1.1 Definition.

Let \mathcal{H}, \mathcal{K} Hilbert spaces and $T: \mathcal{H} \rightarrow \mathcal{K}$ an injective, linear operator, e.g. a frame. Its *Moore-Penrose Inverse* or *Pseudoinverse*, T^\dagger , is defined by

$$T^\dagger := (T^*T)^{-1} T^*.$$

Clearly, for bijective operators the pseudoinverse coincides with the *real* inverse. If T is not bijective, the pseudoinverse acts as a left-inverse, i.e. $T^\dagger T = \text{id}$. When applied on the right hand side, we get an additional restriction to the range of T , that is, $TT^\dagger = P_{\mathcal{R}(T)}$, where $P_{\mathcal{R}(T)}$ is the projection onto the range of T , cf. [BC11].

A very important subclass of frames are the so-called *Riesz Bases*.

1.10 Definition (Cf. [Chr16]).

A frame $\{f_k\}_{k=1}^\infty \subset \mathcal{H}$ is called a *Riesz Basis* if there exist constants c and C , subject to $0 < c \leq C < \infty$, such that for all sequences $a := \{a_k\}_{k=1}^\infty \in \ell^2$ we have

$$c\|a\|_2^2 \leq \left\| \sum_{k=1}^\infty a_k f_k \right\|_2^2 \leq C\|a\|_2^2.$$

It is a standard result in functional analysis, cf. [Rud91], that if \mathcal{H} is separable, there exists another Riesz basis $\{\tilde{f}_k\}_{k=1}^\infty \subset \mathcal{H}$ which is dual to $\{f_k\}_{k=1}^\infty$ in the sense that

$$\langle f_j, \tilde{f}_k \rangle = \delta_{j,k}$$

for all $j, k \in \mathbb{N}$.

Singular Value Decomposition

This section introduces the concept of the so-called singular value decomposition (SVD) both over \mathbb{R} and \mathbb{C} . The SVD allows to decompose matrices of any kind, particularly frames, into a form that allows for better understanding of key properties such as rank, orthogonality or spectral norm. This is completely based on [GV96].

We will proceed in two steps. First, we define the terms singular value decomposition, singular value and singular vectors by providing the formal definition. Second, we cite the vital theorem ensuring unconditional existence of this decomposition.

1.11 Definition.

Let $A \in \mathbb{C}^{m \times n}$. A decomposition of the form

$$A = U\Sigma V^*$$

where $U \in \mathbb{C}^{m \times n}$ and $V \in \mathbb{C}^{n \times n}$ are unitary and $\Sigma \in \mathbb{R}^{m \times n}$ is diagonal in the sense that $\sigma_{i,j} = 0$ for $i \neq j$, such that further $\sigma_{i,i} \geq \sigma_{j,j} \geq 0$ for $i \geq j$, is called the *Singular Value Decomposition* (SVD) of A .

The columns of U are called *Left Singular Vectors*, the columns of V are called *Right Singular Vectors*. The entries $\sigma_i := \sigma_{i,i}$ are called *Singular Values*.

This definition is indeed useful as the SVD always exists.

1.12 Theorem.

The SVD always exists. If A is real, so are U and V .

Due to the decreasing ordering of the singular values, Σ is uniquely determined. In general, however, U and V are not. To underline the usefulness of this decomposition, besides the numerical stability of the unitary matrices U and V , we will now list some commonly known properties of the SVD.

1.13 Proposition.

Let $A \in \mathbb{C}^{m \times n}$ with SVD $A = U\Sigma V^*$.

Then with $p := \min(m, n)$ and $r := \min\{1 \leq i \leq p : s_{i+1} = 0\}$

- a) $Av_i = \sigma_i u_i$ and $A^*u_i = \sigma_i v_i$ for any $i \in \{1, \dots, p\}$,
- b) $\text{rk}(A) = r$,
- c) the SVD can be shortened to

$$A = \sum_{i=1}^r \sigma_i u_i v_i^*,$$

- d) $\|A\|_2 = \sigma_1$ and $\|A\|_F^2 = \sigma_1^2 + \dots + \sigma_p^2$.

Based on c) in Proposition 1.13, there also exists a truncated version of the SVD that sizes down Σ to a square matrix. As a consequence, obviously, U has to be trimmed back, too. Yet this abridged decomposition still allows for meaningful representation of any matrix with more rows than columns.

1.14 Definition.

Let $A \in \mathbb{C}^{m \times n}$ where $m \geq n$ and consider its SVD $A = U\Sigma V^*$. Then also $A = U_1\Sigma_1V^*$ where $U_1 \in \mathbb{C}^{m \times n}$ and $\Sigma_1 \in \mathbb{C}^{n \times n}$ are the truncated submatrices of U and Σ , respectively. We call this decomposition *Thin Singular Value Decomposition* (Thin SVD).

1.3. Wavelets

The concept of the *Discrete Wavelet Transform* – or so-called *Wavelets* – as a form of encoding gained momentum in the 1990s. An accessible presentation can be found in e.g. [Nic17, Mal09]. While some degree of uncertainty will inevitably remain, the historical motivation of wavelets appears to be rather well understood. In [Nic17], Nickolas reports that wavelets were developed as an extension and improvement for Fourier series as they were only able to capture periodic signals. Wavelets on the other hand are designed to provide a family of functions that are scaled, dilated or translated with respect to the so-called *Mother Wavelet*. Pioneering work in this context is largely attributed to Chui and Daubechies, who independently developed the standard works [Dau92, Chu92]. In 2014, Plonka compiled an introductory lecture *Image and Geometry Processing I: Applied Fourier Analysis* based on [Mal09, LMR94, KK98], which we will closely follow. In this chapter, we will recurrently need the notion of Riesz bases from Definition 1.10.

In one dimension, wavelets allow great modelling for large classes of physically observable data, provided a smart choice of base wavelet. For the discrete wavelet transform on \mathbb{R} , we define the wavelet function as follows:

1.15 Definition.

A semi-orthogonal *Wavelet* is a function $\psi \in \mathcal{L}^2(\mathbb{R})$ with the property that the family of functions

$$\left\{ \psi_{j,k} := 2^{j/2} \psi(2^j \cdot -k) : j, k \in \mathbb{Z} \right\}$$

is a Riesz basis for $\mathcal{L}^2(\mathbb{R})$ such that further its dual Riesz basis

$$\left\{ \tilde{\psi}_{j,k} := 2^{j/2} \tilde{\psi}(2^j \cdot -k) : j, k \in \mathbb{Z} \right\}$$

satisfies

$$\left\langle \psi_{k_1}^{j_1}, \tilde{\psi}_{k_2}^{j_2} \right\rangle = \delta_{j_1, j_2} \delta_{k_1, k_2}$$

for $j_1, j_2, k_1, k_2 \in \mathbb{Z}$ and

$$\left\langle \psi_{k_1}^{j_1}, \psi_{k_2}^{j_2} \right\rangle = 0$$

for $j_1, j_2, k_1, k_2 \in \mathbb{Z}$, where $j_1 \neq j_2$.

In order to gain a more general understanding, it is advisable to include the notions of scaling

functions and of the multiresolution analysis. In fact, each of the following conditions has a few implications on its own, be it for theoretical understanding or use case applicability. For this thesis, however, we only aim to provide a shallow insight into this form of construction.

1.16 Definition.

Let $\{V_j\}_{j \in \mathbb{Z}}$ a sequence of closed subspaces of $\mathcal{L}^2(\mathbb{R})$ and $\phi \in V_0$ such that

(MRA 1) The spaces $\{V_j\}$ are nested, in the sense that $V_j \subset V_{j+1}$ for all $j \in \mathbb{Z}$,

(MRA 2) The set $\bigcup_{j \in \mathbb{Z}} V_j$ is dense in $\mathcal{L}^2(\mathbb{R})$,

(MRA 3) $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$,

(MRA 4) For all $j \in \mathbb{Z}$ we have $\phi \in V_0$ if and only if $\phi(2^j \cdot) \in V_j$,

(MRA 5) For all $k \in \mathbb{Z}$ we have $\phi \in V_0$ if and only if $\phi(\cdot - k) \in V_0$,

(MRA 6) The collection $\{\phi(\cdot - k) : k \in \mathbb{Z}\}$ is a Riesz basis for V_0 .

Then the pair $(\{V_j\}, \phi)$ is called *Multiresolution Analysis* and ϕ is called its *Scaling Function*.

Using any multiresolution analysis $(\{V_j\}, \phi)$ we can now construct a wavelet. By definition we have $\phi \in V_0$ and hence by (MRA 1) also $\phi \in V_1$. With (MRA 4) it follows that $\phi(\frac{1}{2} \cdot) \in V_0$. As any element in V_0 possesses a representation in the sense of (MRA 6), there exist coefficients $\{a_k\}_{k \in \mathbb{Z}}$ such that

$$\phi\left(\frac{\cdot}{2}\right) = \sum_{k \in \mathbb{Z}} a_k \phi(\cdot - k)$$

or

$$\phi = \sum_{k \in \mathbb{Z}} a_k \phi(2 \cdot - k). \tag{1.1}$$

We call Equation (1.1) the *Scaling Equation* of the multiresolution analysis. This now allows us to construct the corresponding wavelet.

1.17 Theorem.

Let $(\{V_j\}, \phi)$ a multiresolution analysis with an exponentially decaying scaling function ϕ satisfying Equation (1.1). Further, define a sequence $\{g_k\}_{k \in \mathbb{Z}}$ via

$$g_k := 2(-1)^{k+1} \int_{-\infty}^{\infty} \phi(2x + k - 1)\phi(x) \, dx$$

for all $k \in \mathbb{Z}$. Then we obtain a semi-orthogonal wavelet $\psi \in \mathcal{L}^2(\mathbb{R})$ defined as

$$\psi : x \mapsto \sum_{k \in \mathbb{Z}} g_k \phi(2x - k).$$

B-Spline Wavelets

An important class of wavelets can be constructed using *B-Splines*. Although the classical introduction of this topic is the one by Chui, [Chu92, CW91], here we will follow [SN96, LBU03].

First, we note that for $n \in \mathbb{N}_0$ the cardinal B-Splines β^n are defined in Fourier domain as

$$\widehat{\beta}^n : \xi \mapsto \left(\frac{1 - \exp(-2\pi i \xi)}{2\pi i \xi} \right)^{n+1}.$$

Note that in this form, β^n is the cardinal B-Spline of order $n + 1$, yet of degree n . As shown in [Chu92, Lio04], these satisfy the scaling equation, cf. Equation (1.1)

$$\beta^n = \sum_{k \in \mathbb{Z}} \frac{1}{2^n} \binom{n+1}{k} \beta^n(2 \cdot -k).$$

Thus, the cardinal B-splines generate a multiresolution analysis and subsequently give rise to a family of wavelets.

We will come back to this class of wavelets later in Section 1.5, where we discuss so-called F-splines, which will be constructed as the Fresnel transform of the B-splines mentioned here.

1.4. Shearlets

The shearlet transform is a powerful transform that – when discretized – gives rise to a versatile class of frames. These are of particular interest in the encoding of so-called *Cartoon-like Images*, i.e. piecewise constant functions with smooth boundaries. This is due to the defining feature of shearlets, the *shearing*, cf. Equation (1.2). Kutyniok and Labate motivate the necessity of this feature by the ubiquity of anisotropic features in data occurring in practice. In [KL12], they write

A fundamental property of virtually all data found in practical applications is that the relevant information which needs to be extracted or identified is sparse, i.e., data are typically highly correlated and the essential information lies on low-dimensional manifolds. [...] This is closely related to the observation that virtually all multivariate data are typically dominated by anisotropic features such as singularities on lower dimensional embedded manifolds.

For more details on manifolds and their properties, the interested reader is referred to [Lee13].

Recently, there has been renewed interest in topics such as constructing compactly supported shearlets [KKL12] or formulating the transform with the help of the Radon transform [BDDO17] or via a multiresolution analysis [CZ17].

Here, we present some common forms of shearlets that are of interest in the exemplary application

discussed in Part IV. The concept of shearlets is based on unitary representations. As I do not want to pursue any excursion into representation theory, I advise readers to the study of [Sug90]. This introduction paraphrases [KL12].

Continuous Shearlets

We introduce the *Parabolic Scaling Matrices* A_a for $a \in \mathbb{R}_+$ and *Shearing Matrices* S_s for $s \in \mathbb{R}$ by

$$A_a := \begin{bmatrix} a & 0 \\ 0 & \sqrt{a} \end{bmatrix} \quad S_s := \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}. \quad (1.2)$$

We further define two operators on $\mathcal{L}^2(\mathbb{R}^2)$. Parametrized with $M \in \text{GL}_2(\mathbb{R})$ and $t \in \mathbb{R}^2$, we define the *Dilation Operator* D_M and *Translation Operator* T_t via

$$\begin{aligned} D_M: \mathcal{L}^2(\mathbb{R}^2) &\longrightarrow \mathcal{L}^2(\mathbb{R}^2) & T_t: \mathcal{L}^2(\mathbb{R}^2) &\longrightarrow \mathcal{L}^2(\mathbb{R}^2) \\ \Phi &\longmapsto |\det(M)|^{1/2} \Phi(M^{-1}\cdot) & \text{and} & & \Phi &\longmapsto \Phi(\cdot - t). \end{aligned}$$

These let us now introduce the so-called *Shearlet Group*, a semi-direct product

$$\mathbb{S} := (\mathbb{R}_+ \times \mathbb{R}) \ltimes \mathbb{R}^2$$

characterized by

$$(a, s, t) \cdot (a, \tilde{s}, \tilde{t}) = (a\tilde{a}, s + \sqrt{a}\tilde{s}, t + S_s A_a \tilde{t}).$$

As a next step we introduce a unitary representation $\sigma: \mathbb{S} \longrightarrow \mathcal{U}(\mathcal{L}^2(\mathbb{R}^2))$ of \mathbb{S} , defined as

$$\sigma(a, s, t): \Phi \longmapsto T_t D_{A_a} D_{S_s} \Phi.$$

This now lets us characterize continuous shearlets.

1.18 Definition.

Let $\psi \in \mathcal{L}^2(\mathbb{R}^2)$.

We define its *Continuous Shearing System* by

$$\text{SH}(\psi) := \{\sigma(a, s, t)\psi : (a, s, t) \in \mathbb{S}\}.$$

When referring to this system, we call ψ its *Shearlet*. The corresponding *Continuous Shearlet Transform* is defined as a map onto \mathbb{S}^* , the subgroup of units of \mathbb{S} ,

$$\begin{aligned} \mathcal{SH}_\psi: \mathcal{L}^2(\mathbb{R}^2) &\longrightarrow \mathbb{S}^* \\ f &\longmapsto [(a, s, t) \mapsto \langle f, \sigma(a, s, t)\psi \rangle]. \end{aligned} \quad (1.3)$$

We will now take a look at three important classes of shearlets, namely *classical*, *admissible* and *separable* shearlets.

1.19 Definition.

A function is called *Classical Shearlet* if there exist functions $\psi_1, \psi_2 \in \mathcal{L}^2(\mathbb{R})$ such that for $\omega = (\omega_1, \omega_2) \in \mathbb{R}^2$,

$$\widehat{\psi}(\omega_1, \omega_2) = \widehat{\psi}_1(\omega_1) \widehat{\psi}_2\left(\frac{\omega_2}{\omega_1}\right),$$

where

- $\widehat{\psi}_1, \widehat{\psi}_2 \in C^\infty(\mathbb{R})$
- $\text{supp}(\widehat{\psi}_1) \subseteq [-\frac{1}{2}, -\frac{1}{16}] \cup [\frac{1}{16}, \frac{1}{2}]$ and $\text{supp}(\widehat{\psi}_2) \subseteq [-1, 1]$
- $\sum_{j \in \mathbb{Z}} |\widehat{\psi}_1(2^{-j}\omega)|^2 = 1$ for a.e. $\omega \in \mathbb{R}$
- $\sum_{k=-1}^1 |\widehat{\psi}_2(\omega + k)|^2 = 1$ for a.e. $\omega \in [-1, 1]$.

Despite these very specific conditions, classical shearlets provide an easily graspable geometrical interpretation. One may picture ψ_1 as a (discrete) wavelet and ψ_2 as a bump function. A simple yet suitable choice could be to select ψ_1 as a Lemarié–Meyer wavelet and ψ_2 as a spline.

Next, we present the notion of admissible shearlets.

1.20 Definition (4 in [KL12]).

A function $\psi \in \mathcal{L}^2(\mathbb{R}^2)$ is called *Admissible Shearlet* if

$$\int_{\mathbb{R}^2} \frac{|\widehat{\psi}(\xi_1, \xi_2)|^2}{\xi_1^2} d\xi < \infty. \quad (1.4)$$

With the help of the Plancherel theorem, one can show that admissible shearlets give rise to an isometric shearlet transform as soon as they satisfy the condition

$$\int_0^\infty \int_{-\infty}^\infty \frac{|\widehat{\psi}(\xi_1, \xi_2)|^2}{\xi_1^2} d\xi_2 d\xi_1 = \int_{-\infty}^0 \int_{-\infty}^\infty \frac{|\widehat{\psi}(\xi_1, \xi_2)|^2}{\xi_1^2} d\xi_2 d\xi_1 = 1. \quad (1.5)$$

It can also be shown, [DKM⁺08, KL12], that classical shearlets are especially admissible and satisfy the condition of Equation (1.5), i.e. yield an invertible shearlet transform.

Last, the separable shearlet transform, as considered in [Lim10, KKL12], aims at the construction of bandlimited shearlets. As Lim describes in [Lim10], separable shearlets are designed to have a rectangular essential support in Fourier domain rather than the isosceles trapezoids of common, e.g.

classical, shearlets. A complete classification of separable, bandlimited shearlets is far beyond the scope of this thesis. Instead we state a simplified, tractable approach yielding such shearlets.

1.21 Proposition (Cf. [KKL12]).

Let $K, L \in \mathbb{N}$ such that $L \geq 10$ and $\frac{3L}{2} \leq K \leq 3L - 2$. Let m_0 a trigonometric polynomial with real coefficients satisfying $m_0(0) = 1$ and

$$|m_0(x)|^2 = \cos(\pi x)^{2K} \sum_{n=0}^{L-1} \binom{K-1+n}{n} \sin(\pi x)^{2n}.$$

Further define $\phi \in \mathcal{L}^2(\mathbb{R})$ by

$$\widehat{\phi}(\xi) = \prod_{j=0}^{\infty} m_0(2^{-j}\xi).$$

Now define a band-pass filter m_1 by

$$|m_1(x)|^2 = |m_0(x + \frac{1}{2})|^2$$

for all $x \in \mathbb{R}$. Then the shearlet $\psi \in \mathcal{L}^2(\mathbb{R}^2)$ defined in Fourier domain by

$$\widehat{\psi}(\xi) = m_1(4\xi_1)\widehat{\phi}(\xi_1)\widehat{\phi}(2\xi_2), \tag{1.6}$$

for $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$ is compactly supported.

Discrete Shearlets

While continuous shearlets may be a versatile analysis tool, their handling is difficult in practice due to the cardinality of \mathbb{S} . Therefore, Kutyniok describes a way to discretize shearlets to create a frame. In that sense for a shearlet $\psi \in \mathcal{L}^2(\mathbb{R}^2)$ let

$$\psi_{j,k,m} = 2^{\frac{3}{4}j} \psi(S_k A_{2^j} \cdot -m)$$

and define

$$\text{SH}(\psi) := \{\psi_{j,k,m} : j, k \in \mathbb{Z}, m \in \mathbb{Z}^2\}.$$

Luckily, for classical shearlets this system already is a frame.

1.22 Proposition (2 in [KL12]).

Let $\psi \in \mathcal{L}^2(\mathbb{R}^2)$ a classical shearlet. Then SH is a Parseval frame for $\mathcal{L}^2(\mathbb{R}^2)$.

Shearlets also provide numerous opportunities for digitization whilst keeping intact core properties. To the interested reader I suggest the reading of [KLR16]. For our purpose, it is sufficient to know that there exist feasible algorithms that allow digital application of the shearlet frame.

1.5. Fresnel Transform

The Fresnel Transform is a parameterized unitary transform on \mathbb{R}^n . This transform will later become a crucial part in the *Fresnel Propagation* in Section 6.1 as an approximative model for the propagation of electromagnetic waves in vacuum. We will base this introduction on [LBU03, Lico4]. For the interested reader, I also recommend the reading of the latter.

We start by introducing the Fresnel kernel and, based on that, the Fresnel transform as the convolution therewith.

1.23 Definition.

Let $\tau > 0$. We define the *Fresnel Kernel* K_τ by letting

$$K_\tau: \mathbb{R}^n \longrightarrow \mathbb{C},$$

$$x \longmapsto \frac{1}{\tau^n} \exp\left(\frac{i\pi}{\tau^2} \|x\|_2^2\right).$$

We call the convolution with the Fresnel kernel the *Fresnel Transform* and write

$$\tilde{f}_\tau := f * K_\tau.$$

This definition is very approachable in the respect that, for instance, it exhibits a closed form in Fourier domain. That is,

$$\widehat{K}_\tau: \mathbb{R}^n \longrightarrow \mathbb{C},$$

$$\xi \longmapsto (i)^{n/2} \exp(-i\pi\tau^2 \|\xi\|_2^2).$$

Also, both the kernel K_τ and its Fourier transform \widehat{K}_τ are separable by dimension.

1.24 Remark.

When letting

$$k_\tau := \frac{1}{\tau} \exp(i\pi(\frac{1}{\tau}\cdot)^2)$$

and correspondingly

$$\widehat{k}_\tau := \sqrt{i} \exp(-i\pi(\tau\cdot)^2)$$

we see, e.g. in \mathbb{R}^2 , that

$$K_\tau(x_1, x_2) = k_\tau(x_1)k_\tau(x_2)$$

and

$$\widehat{K}_\tau(\omega_1, \omega_2) = \widehat{k}_\tau(\omega_1)\widehat{k}_\tau(\omega_2).$$

F-Spline Wavelets

In the early 2000s, Liebling considered B-Spline wavelets, cf. Section 1.3, and applied the Fresnel transform to the spline functions β^n . We denote $\tilde{\beta}_\tau^n = \beta^n * k_\tau$ as in Definition 1.23 and call this family *F-Splines*, cf. [LBU03, Lieo4].

Following [LBU03], we notice that F-splines do not satisfy the scaling equation, Equation (1.1), but instead

$$\tilde{\beta}_{\tau/2}^n = \sum_{k \in \mathbb{Z}} \frac{1}{2^n} \binom{n+1}{k} \tilde{\beta}_\tau^n(\cdot - k).$$

The authors go on to propose an alternate multiresolution analysis by letting

$$\tilde{V}_{j,\tau} := \text{span}_{k \in \mathbb{Z}} \left\{ \tilde{\beta}_\tau^n(2^{-j} \cdot - k) \right\} \cap \mathcal{L}^2(\mathbb{R})$$

satisfying

(MRA 1') The spaces $\{\tilde{V}_{j,\tau}\}$ are nested, in the sense that $\tilde{V}_{j,\tau} \subset \tilde{V}_{j-1,\tau}$ for all $j \in \mathbb{Z}$,

(MRA 2) The set $\bigcup_{j \in \mathbb{Z}} \tilde{V}_{j,\tau}$ is dense in $\mathcal{L}^2(\mathbb{R})$,

(MRA 3) $\bigcap_{j \in \mathbb{Z}} \tilde{V}_{j,\tau} = \{0\}$,

(MRA 4') For all $j \in \mathbb{Z}$ we have $f \in \tilde{V}_{j,\tau}$ if and only if $\tilde{f}_{\sqrt{3}\tau}(2 \cdot) \in \tilde{V}_{j-1,\tau}$,

(MRA 5) For all $k \in \mathbb{Z}$ we have $f \in \tilde{V}_{0,\tau}$ if and only if $f(\cdot - k) \in \tilde{V}_{0,\tau}$,

(MRA 6) The collection $\{\tilde{\beta}_\tau^n(\cdot - k) : k \in \mathbb{Z}\}$ is a Riesz basis for $\tilde{V}_{0,\tau}$.

Note that in this design the scaling spaces are in inverse nesting order as compared to Definition 1.16. This MRA still allows for a digitization of wavelets and the authors provide exemplary applications such as backpropagation of diffraction patterns.

1.6. Fresnel Transformed Shearlets

Generalizing the approach of Liebling based in [LBU03], i.e. applying the Fresnel transform on (B-spline) wavelets, we will now investigate the effects of consecutive application of Fresnel and shearlet transform. The main idea is that, if an image $I(x, y)$ can be sparsely represented by the shearlet transform w.r.t. ψ , its Fresnel transform $K_\tau * I$ should satisfy the same for an adequate counterpart. In this section, we investigate which kind of shearlets appear to be suitable.

In contrast to Liebling's success, however, we expect less convenient results due to the construction of shearlets themselves. While multi-dimensional wavelets are commonly constructed as the tensor product of their one-dimensional versions, [Oyv19], shearlets deliberately left this path in order to be able to address singularities along curves, [KL12]. Thus we expect this undertaking to be challenging at best.

In this sense, we will show that the Fresnel transform keeps the structure of admissible shearlets intact, Proposition 1.25, but not of classical shearlets, Proposition 1.26. It also maps band-limited shearlets onto band-limited shearlets, Proposition 1.27, and retains Equation (1.6) as one of the characteristic conditions for separable shearlets as introduced in Proposition 1.21, Proposition 1.28. We are, however, not able to show that the Fresnel transform of separable shearlets yields separable shearlets in an applicable manner.

To that end, recall the Fresnel transform of the previous section, Definition 1.23 and especially the notation

$$\tilde{f}_\tau := f * K_\tau$$

for the transformed function.

We start by considering admissible shearlets.

1.25 Proposition.

The Fresnel transform maps admissible shearlets on admissible shearlets.

Proof.

Assume that $\psi \in \mathcal{L}^2(\mathbb{R}^2)$ is a shearlet satisfying the condition in Equation (1.4). We will now investigate these conditions with regard to its Fresnel transformed counterpart, $\tilde{\psi}_\tau$. To that end, we calculate

$$|\mathcal{F}[\tilde{\psi}_\tau]|^2 = |\mathcal{F}[\psi * K_\tau]|^2 = |\widehat{\psi} \widehat{K}_\tau|^2 = |\widehat{\psi}|^2 |\widehat{K}_\tau|^2 = |\widehat{\psi}|^2 |\exp(-i\pi\tau^2 \|\cdot\|_2^2)|^2 = |\widehat{\psi}|^2.$$

So we find

$$\iint_{\mathbb{R}^2} \frac{|\mathcal{F}[\tilde{\psi}_\tau](\omega_1, \omega_2)|^2}{\omega_1^2} d\omega_2 d\omega_1 = \iint_{\mathbb{R}^2} \frac{|\widehat{\psi}(\omega_1, \omega_2)|^2}{\omega_1^2} d\omega_2 d\omega_1 < \infty.$$

Thus, the Fresnel transform does not change the properties of Equation (1.4), i.e. the admissibility assertion holds true for $\tilde{\psi}_\tau$. qed

Clearly, the same argument applies for Equation (1.5), i.e. admissible shearlets providing an isometric shearlet transform yield Fresnel transformed shearlets with the same property.

Now, we will analyse the structure of classical shearlets w.r.t. the Fresnel transform. Here, we will run into problems as the two-dimensional Fresnel transform is applied in a tensor-like manner, cf. Remark 1.24, while classical shearlets are decomposed as

$$\widehat{\psi}(\omega_1, \omega_2) = \widehat{\psi}_1(\omega_1) \widehat{\psi}_2\left(\frac{\omega_2}{\omega_1}\right),$$

cf. Definition 1.19. Expectably, this is incompatible.

1.26 Proposition.

The Fresnel transform does not preserve the structure of classical shearlets.

Proof.

Assume that $\psi \in \mathcal{L}^2(\mathbb{R}^2)$ is a shearlet satisfying the conditions of Definition 1.19 w.r.t. the functions ψ_1 and ψ_2 . To show that $\tilde{\psi}_\tau$ no longer is a classical shearlet, assume contrarily the existence of two functions ψ_1^τ and ψ_2^τ such that these would conform with the conditions of Definition 1.19, i.e.

$$\widehat{\psi}_1^\tau(\omega_1) \widehat{\psi}_2^\tau\left(\frac{\omega_2}{\omega_1}\right) = \mathcal{F}[\tilde{\psi}_\tau](\omega).$$

Note that we will not assume that $\psi_1^\tau = \psi_1 * K_\tau$ or $\psi_2^\tau = \psi_2 * K_\tau$.

Thence, with $\omega = (\omega_1, \omega_2) \in \mathbb{R}^2$,

$$\begin{aligned} \widehat{\psi}_1^\tau(\omega_1) \widehat{\psi}_2^\tau\left(\frac{\omega_2}{\omega_1}\right) &= \mathcal{F}[\tilde{\psi}_\tau](\omega) \\ &= \mathcal{F}[\psi * K_\tau](\omega) \\ &= \widehat{K}_\tau(\omega) \widehat{\psi}(\omega) \\ &= \widehat{k}_\tau(\omega_1) \widehat{k}_\tau(\omega_2) \widehat{\psi}_1(\omega_1) \widehat{\psi}_2\left(\frac{\omega_2}{\omega_1}\right) \\ &= \mathcal{F}[k_\tau * \psi_1](\omega_1) \mathcal{F}[\psi_2]\left(\frac{\omega_2}{\omega_1}\right) \int_{\mathbb{R}} k_\tau(t) \exp(-2\pi i t \omega_2) dt, \end{aligned}$$

where the integral is a function in ω_2 as compared to the desired ω_2/ω_1 , thereby rendering the desired decomposition impossible. qed

Next, we take a look the the Fresnel transform in conjunction with band-limited shearlets. Due to the closed form of the Fresnel kernel both in spatial and frequency domain, this property can easily be addressed.

1.27 Proposition.

The Fresnel transform maps band-limited shearlets onto band-limited shearlets.

Proof.

Let ψ be band-limited, i.e. assume existence of constants $a < b$ and $\alpha < \beta$ such that

$$\text{supp}(\widehat{\psi}) \subseteq [a, b] \times [\alpha, \beta].$$

Then

$$\text{supp}(\widehat{\tilde{\psi}_\tau}) = \text{supp}(\mathcal{F}[\psi * K_\tau]) = \text{supp}(\widehat{K_\tau} \cdot \widehat{\psi}) = \text{supp}(\widehat{K_\tau}) \cap \text{supp}(\widehat{\psi}) \subseteq \text{supp}(\widehat{\psi}).$$

So $\tilde{\psi}_\tau$ is band-limited as well. qed

Finally, we consider separable shearlets. These are an intricate topic in their own right, cf. [Lim10, KKL12], which is why we will restrict ourselves in this thesis to the analysis of the decomposition structure of Equation (1.6). This class of shearlets, however, is the most promising from a heuristic point of view as separability of the shearlet (in Fourier domain) may align well with the Fresnel kernel, which exhibits the same property.

1.28 Proposition.

The Fresnel transform maps shearlets generated from Proposition 1.21 onto functions satisfying Equation (1.6).

Proof.

Let ψ a compactly supported separable shearlet in the sense of Proposition 1.21, i.e. there exist a scaling function φ and a band-pass filter m_1 , such that for $\omega = (\omega_1, \omega_2) \in \mathbb{R}^2$,

$$\widehat{\psi}(\omega) = m_1(4\omega_1)\widehat{\varphi}(\omega_1)\widehat{\varphi}(2\omega_2).$$

Now let

$$\begin{aligned} m_\tau(4\omega_1) &:= 4 m_1(4\omega_1) \exp\left(-\frac{3}{4}i\pi\tau^2\omega_1^2\right) \\ \varphi_\tau &:= k_\tau(2\cdot) * \varphi. \end{aligned}$$

Then the Fresnel transform $\tilde{\psi}_\tau = \psi * K_\tau$ possesses a Fourier transform $\mathcal{F}\tilde{\psi}_\tau$ that can be decomposed as

$$\mathcal{F}\tilde{\psi}_\tau = \mathcal{F}[\psi * K_\tau](\omega)$$

1. *Frames, Bases and Transforms*

$$\begin{aligned}
&= \widehat{K}_\tau(\omega) \widehat{\psi}(\omega) \\
&= \widehat{k}_\tau(\omega_1) m_1(4\omega_1) \widehat{\varphi}(\omega_1) \widehat{k}_\tau(\omega_2) \widehat{\varphi}(2\omega_2) \\
&= m_1(4\omega_1) \widehat{k}_\tau(\omega_1) \widehat{\varphi}(\omega_1) 2\mathcal{F}[k_\tau(2\cdot)](2\omega_2) \widehat{\varphi}(2\omega_2) \\
&= 2m_1(4\omega_1) \mathcal{F}[k_\tau * \varphi](\omega_1) \mathcal{F}[k_\tau(2\cdot) * \varphi](2\omega_2) \\
&= 2m_1(4\omega_1) \mathcal{F}[k_\tau * \varphi](\omega_1) \widehat{\varphi}_\tau(2\omega_2) \\
&= 4m_1(4\omega_1) \exp\left(-\frac{3}{4}i\pi\tau^2\omega_1^2\right) \mathcal{F}[k_\tau(2\cdot) * \varphi](\omega_1) \widehat{\varphi}_\tau(2\omega_2) \\
&= m_\tau(4\omega_1) \widehat{\varphi}_\tau(\omega_1) \widehat{\varphi}_\tau(2\omega_2),
\end{aligned}$$

where the penultimate identity is due to

$$\begin{aligned}
\mathcal{F}[k_\tau * \varphi](\omega_1) &= \widehat{k}_\tau(\omega_1) \widehat{\varphi}(\omega_1) \\
&= \sqrt{i} \exp(-i\pi\tau^2\omega_1^2) \widehat{\varphi}(\omega_1) \\
&= \sqrt{i} \exp(-i\pi\tau^2\omega_1^2) \widehat{\varphi}(\omega_1) \\
&= \sqrt{i} \exp\left(-\frac{1}{4}i\pi\tau^2\omega_1^2\right) \widehat{\varphi}(\omega_1) \exp\left(-\frac{3}{4}i\pi\tau^2\omega_1^2\right) \\
&= \sqrt{i} \exp(-i\pi\tau^2(\frac{1}{2}\omega_1)^2) \widehat{\varphi}(\omega_1) \exp\left(-\frac{3}{4}i\pi\tau^2\omega_1^2\right) \\
&= \widehat{k}_\tau\left(\frac{\omega_1}{2}\right) \widehat{\varphi}(\omega_1) \exp\left(-\frac{3}{4}i\pi\tau^2\omega_1^2\right) \\
&= 2\mathcal{F}[k_\tau(2\cdot) * \varphi](\omega_1) \exp\left(-\frac{3}{4}i\pi\tau^2\omega_1^2\right).
\end{aligned}$$

So, indeed, the function $\tilde{\psi}_\tau$ can be decomposed in the same manner. qed

For achieving a decomposition suitable for digitization in the sense of the digital separable shearlet transform, cf. [KLZ12], we would need m_τ and φ_τ to behave sufficiently nice. However, numerical tests suggested this is not the case.

In summary, the characteristic feature of shearlets, i.e. the ability to follow singularities along (smooth) curves does not allow sensible propagation with respect to the paraxial Fresnel transform, as expected.

2. Convex Analysis

This chapter is concerned with introducing the concepts of convex analysis that we will use recurrently in Part III. In doing so, this chapter has three different, although related, focal points. Before coming to the first point, we address some introductory notions that are relevant in their own right but will not necessarily entail a cascade of successive results, at least not within the focus of this thesis. These notions include plc functions, the infimal convolution and firm nonexpansiveness. We also provide insight into the concept of monotonicity and its properties of cyclicity and maximality, correspondingly.

Our first focus will be on the extension of the notion of derivatives and continuity in Section 2.1. To this end, we introduce generalized concepts, e.g. the subdifferential, to create a notion of derivatives of functions that are not differentiable in the usual sense. This will become crucial when considering proximity operators in the subsequent section, since those are closely related, as we will demonstrate. We will also state the condition for Lipschitz continuity and derive the concept of nonexpansiveness.

Secondly, we focus on the proximity operator, which is the key to understanding Chapters 4 and 7. Most relevantly, we introduce the formal definition and show existence, uniqueness, and single-valuedness for plc priors. We also consider proximity operators as special cases of resolvents, thereby embedding this notion into a larger, more general context within convex analysis and operator theory. Moreover, we provide coherence to the preceding section by analyzing the proximity operator for e.g. nonexpansiveness.

Our third focus, in Section 2.3, will be on a choice of operator splitting algorithms, namely forward-backward splitting and Douglas-Rachford splitting. We state their schematic instructions and cite converge criteria from [BC11]. This will then allow us to consider a special case of Douglas-Rachford splitting, employing averaged alternating reflections. Following the works of Luke, cf. [BCL04, Luk05, LM20], we furthermore derive a relaxation of the Douglas-Rachford algorithm, which is known as DR λ .

We start by introducing the notion of plc functions. These will be essential in the context of proximity operators and consequently throughout Chapter 4. To this end, we state the requirements for *proper*, *lower semi-continuous* and *convex* functions and define plc functions as those that satisfy all three criteria.

2. Convex Analysis

2.1 Definition (Cf. [BC11]).

Let \mathcal{H} a Hilbert space and $f: \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$. Then f is called

- a) *Proper*, if $f(x) \neq -\infty$ for any x and $f(x_0) \neq +\infty$ for at least one x_0 .
- b) *Lower Semi-Continuous* at some point x , if for every net $(x_a)_{a \in A}$ in \mathbb{R}^d it holds that $x_a \rightarrow x$ implies $f(x) \leq \liminf f(x_a)$.
- c) *Convex*, if its epigraph is convex, i.e. for all $\alpha \in (0, 1)$ and arbitrary (x, ξ) and $(y, \eta) \in \text{epi } f$ we have that $(\alpha x + (1 - \alpha)y, \alpha\xi + (1 - \alpha)\eta) \in \text{epi } f$ as well.

If a function is proper, lower semi-continuous and convex, we also call that function to be *PLC*. We denote the set of plc functions from \mathcal{H} to $\mathbb{R} \cup \{+\infty\}$ by $\Gamma_0(\mathcal{H})$. We will omit the specification of \mathcal{H} whenever clear from context.

Next, we want to introduce the concept of the infimal convolution. This concept is much less common than the usual convolution, i.e. the integral convolution. These two kinds of convolution feature a noticeably close relationship, as their definition is identical when interchanging the integral operator for an infimum and the product for a sum.

2.2 Definition (Cf. [BC11]).

Let $f, g \in \Gamma_0(\mathcal{H})$. We define the *Infimal Convolution* $f \square g$ as

$$f \square g := \inf_{y \in \mathcal{H}} [f(y) + g(\cdot - y)].$$

The infimal convolution lives within the Γ_0 space of a given Hilbert space. Thus, as would be expected from any well-defined operation in functional calculus, it exhibits a closedness property.

2.3 Proposition (Cf. [Phe93, BGLW08]).

Let $f, g \in \Gamma_0$. Then also $f \square g \in \Gamma_0$.

We will now concern ourselves with the notion of monotonicity. To that end, we proceed by first recalling the classical notion of monotone functions and then expanding this to the more specific α -strong monotonicity.

2.1 Definition (20.1 and 22.1 in [BC11]).

Let $\alpha > 0$ and $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ an operator. A is called *Monotone* if we have

$$\langle x - y, u - v \rangle \geq 0$$

for all pairs $(x, u); (y, v) \in \Gamma(A)$. A is called *α -Strongly Monotone* if $A - \alpha \text{id}$ is monotone.

Now, as a further, related concept, we introduce cyclical monotonicity and describe the condition for maximality thereof. As Rockafellar shows in [Roc66], maximally cyclically monotone operators are indeed very common. For a plc function $f \in \Gamma_0(H)$ the subdifferential ∂f is a maximally cyclically monotone operator on $\mathcal{H} \times \mathcal{H}^*$, where \mathcal{H}^* denotes the dual space to \mathcal{H} . We will formally introduce the subdifferential in Definition 2.7 and formulate a theorem including this statement in Theorem 2.8.

2.4 Definition (Cf. 12.24 in [RW98]).

Denote $\langle \cdot, \cdot \rangle$ an inner product in \mathcal{H} . A mapping $H: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ is called *Cyclically Monotone* if for any $m \in \mathbb{N}$, $m \geq 2$ and any choice of points x_1, \dots, x_m and elements $y_i \in H(x_i)$ we have

$$\langle x_2 - x_1, y_1 \rangle + \langle x_3 - x_2, y_2 \rangle + \dots + \langle x_1 - x_m, y_m \rangle \leq 0.$$

We call H *Maximally Cyclically Monotone* if it is cyclically monotone and its graph cannot be enlarged without destroying this property.

The maximality condition in its presented form might appear rather abstract and of little use. Fortunately, maximal monotonicity can be characterized in a much more tangible way by Minty's Theorem.

2.5 Theorem (Minty's Theorem, 21.1 in [BC11]).

Let $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ be monotone. Then A is maximally monotone if and only if the operator $\text{id} + A$ has full range, i.e. $\mathcal{R}(\text{id} + A) = \mathcal{H}$.

At the end of this section we will now address the notion of firm nonexpansiveness and, based on that, cocoerciveness. We will later also present the definition of the weaker condition of (not necessarily firm) nonexpansiveness in Definition 2.10. This will be however based on Lipschitz continuity and is hence contextualized in Section 2.1. Differentiability and Continuity.

2.6 Definition (Cf. [HNP⁺19, BC11]).

Let \mathcal{H} a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and corresponding norm $\| \cdot \|$. An operator $G: \mathcal{H} \rightarrow \mathcal{H}$ is called *Firmly Nonexpansive* if for all $x, y \in \mathcal{H}$ the relation

$$\|Gx - Gy\|^2 \leq \langle x - y, Gx - Gy \rangle$$

is satisfied.

The operator G is said to be β -Cocoercive for some $\beta > 0$, if βG is firmly nonexpansive.

2.1. Differentiability and Continuity

In this section we provide terminology to extend the commonly known notions of differentiability and continuity, e.g. from basic calculus. To this end, we consider the subdifferential that provides a sense of derivative even for non-differentiable functions. This also allows us to state Rockafellar's Cyclic Monotonicity Theorem that we have anticipated since Definition 2.4. A second generalization of the classical derivative is the Fréchet derivative that we introduce in Definition 2.9. We conclude by presenting the concept of Lipschitz continuity, which is a stricter condition than basic continuity. This in turn leads to an easy characterization of nonexpansiveness, which adds the missing context to Definition 2.6.

We begin by quoting the definition of the subdifferential from [BC11].

2.7 Definition (16.1 in [BC11]).

Let \mathcal{H} a Hilbert space and $f: \mathcal{H} \rightarrow \mathbb{R} \cup \{\infty\}$ be proper. The *Subdifferential* of f is the set-valued operator

$$\begin{aligned} \partial f: \mathcal{H} &\rightarrow 2^{\mathcal{H}} \\ x &\mapsto \{u \in \mathcal{H} : \langle y - x, u \rangle + f(x) \leq f(y) \text{ for all } y \in \mathcal{H}\}. \end{aligned}$$

Let now $x \in \mathcal{H}$. Then f is *Subdifferentiable* at x if $\partial f(x) \neq \emptyset$.

We now have introduced, in Definitions 2.4 and 2.7, all required prerequisites to state Rockafellar's Cyclic Monotonicity Theorem. This will, eventually, allow us to investigate set-valued operators for maximal cyclical monotonicity.

2.8 Theorem (Rockafellar, 22.14 in [BC11]).

A set-valued mapping $H: \mathbb{R}^N \rightarrow 2^{\mathbb{R}^N}$ is the subdifferential of a function $\Phi \in \Gamma_0$ if and only if H is maximally cyclically monotone.

As a second generalized notion of derivation is given by the so-called Fréchet derivative. This concept also gives rise to an appropriate understanding of the gradient, which we will use later, e.g. in Section 4.2.

2.9 Definition (Cf. [FSG08, HNP⁺19]).

Let \mathcal{H} a Hilbert space and $f: \mathcal{H} \rightarrow \mathbb{R}$. The function f is called *Fréchet Differentiable* at $x \in \mathcal{H}$ if there exists a bounded linear operator $Df_x: \mathcal{H} \rightarrow \mathbb{R}$ such that

$$\lim_{\|h\|_{\mathcal{H}} \rightarrow 0} \frac{|f(x+h) - f(x) - Df_x h|}{\|h\|_{\mathcal{H}}} = 0.$$

For a Fréchet differentiable function $f: \mathcal{H} \rightarrow \mathbb{R}$, the gradient $\nabla f(x)$ at $x \in \mathcal{H}$ is defined as the vector satisfying

$$\langle \nabla f(x), h \rangle = Df_x h.$$

for all $h \in \mathcal{H}$.

If clear from context, we will omit the subscript of x and write Df .

As we have formulated this definition for arbitrary Hilbert spaces with a not further specified norm, the gradient and hence the Fréchet differentiability property fundamentally depend on the chosen inner product. The same holds true for the Lipschitz continuity between two arbitrary Hilbert spaces that we will introduce next.

2.10 Definition (Cf. [BC11]).

Let \mathcal{H}, \mathcal{K} two Hilbert spaces. A function $f: \mathcal{H} \rightarrow \mathcal{K}$ is called *L-Lipschitz Continuous* if there exists a constant $L \geq 0$ such that for any $x, y \in \mathcal{H}$ we have

$$\|f(x) - f(y)\|_{\mathcal{K}} \leq L\|x - y\|_{\mathcal{H}}.$$

A 1-Lipschitz continuous function is called *Nonexpansive*.

2.2. Proximity Operators

This section is dedicated to introducing the proximity operator and presents some basic results that follow immediately from its definition. We will consider the proximity operator for plc functions, that we have introduced in Definition 2.1, as this will ensure uniqueness and existence. Moreover, we will use this newly introduced tool to handily characterize maximal cyclical monotonicity and classify the proximity operator as the resolvent of the subdifferential of its prior.

First, we state the proximity operator and the inherent Moreau envelope.

2.11 Definition (Proximity Operator, cf. [BC11]).

Let $f \in \Gamma_0(\mathcal{H})$. We define its *Proximity Operator* (prox operator) as

$$\text{prox}_f := \arg \min_{z \in \mathcal{H}} [f(z) + \frac{1}{2}\|z - \cdot\|^2]$$

and its Moreau envelope $M_f: \mathcal{H} \rightarrow \mathbb{R}$ via

$$M_f := \inf_{z \in \mathcal{H}} [f(z) + \frac{1}{2}\|z - \cdot\|^2].$$

2. Convex Analysis

When discussing proximity operators prox_f or Moreau envelopes M_f , we will refer to f as their corresponding *Prior*. It is important to note that the proximity operator and the Moreau envelope depend crucially on the underlying space and the choice of the employed norm. Yet, this allows us to tailor these operators for specific use cases as we will demonstrate in Section 4.1.

There also exists an alternative characterization of the Moreau envelope as an infimal convolution, namely for $f \in \Gamma_0$ we see

$$M_f = f \square_{\frac{1}{2}} \|\cdot\|^2.$$

In relation to frame theory, cf. Chapter 1, the next example shows that some common shrinkage operators can be expressed as proximity operators. Reversely, this shows that for some priors the proximity operator exhibits a closed form which in general is rather rare.

2.12 Example (Cf. [Bec17, GN20]).

Considering the (pseudo-)norms of ℓ^0 , ℓ^1 and ℓ^2 as introduced in the [List of Symbols](#), we present their proximity operators in a closed form. We use a positive scaling coefficient $\lambda > 0$. In the first two cases, the proximity operator is best expressed componentwise; hence we will use some $x \in \mathbb{R}^n$ as a generic argument of the proximity operators.

Now, we see that hard shrinkage is the proximity operator of $\|\cdot\|_0$,

$$\left[\text{prox}_{\lambda \|\cdot\|_0}(x) \right]_j = \begin{cases} 0, & |x_j| < \sqrt{2\lambda}, \\ x_j, & |x_j| > \sqrt{2\lambda}, \\ 0 \text{ or } x_j, & |x_j| = \sqrt{2\lambda}, \end{cases}$$

soft shrinkage is the proximity operator of $\|\cdot\|_1$,

$$\left[\text{prox}_{\lambda \|\cdot\|_1}(x) \right]_j = \begin{cases} x_j - \lambda, & x_j \geq \lambda, \\ x_j + \lambda, & x_j \leq -\lambda, \\ 0, & \text{else,} \end{cases}$$

or the proximity operator of $\|\cdot\|_2$,

$$\text{prox}_{\lambda \|\cdot\|_2}(x) = \begin{cases} (1 - \lambda \|x\|_2^{-1}) x, & \|x\|_2 > \lambda, \\ 0, & \|x\|_2 \leq \lambda. \end{cases}$$

In the following, we will also refer to the soft shrinkage operator $\text{prox}_{\lambda \|\cdot\|_1}$ as \mathcal{S}_λ , as is more common in frame theory.

In order to embed the notion of proximity operators into some larger context in operator theory, we now consider the notion of so-called resolvents and subsequently prove that these form a superclass

to proximity operators.

2.13 Definition (23.1 and 23.10 in [BC11]).

Let $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$. We denote the *Resolvent* of A as

$$J_A := (\text{id} + A)^{-1}$$

and the *Reflected Resolvent* of A as

$$R_A := 2J_A - \text{id}.$$

The next proposition combines two core properties of proximity operators. First, and most importantly for unburdened usage of proximity operators, this states unconditional, unique existence and, notably, single-valuedness for plc functions. Second, this provides the relation between proximity operators, resolvents and subdifferentials, which are all pivotal notions in convex analysis.

2.14 Proposition (Cf. [ET87]).

The proximity operator prox_f of a plc function $f \in \Gamma_0$ exists, is uniquely defined, single-valued, and we have

$$\text{prox}_f = J_{\partial f}.$$

This proposition entails two useful results with respect to this thesis. First, we rewrite the relation between the proximity operator and the corresponding resolvent at a lower, i.e. element-wise, level in Moreau's *Théorème des Fonctions Duales*. Often, this is more useful in calculations than the operator-theoretic form of Proposition 2.14.

2.15 Corollary (Moreau, 4.a in [Mor65]).

Let $f \in \Gamma_0(\mathbb{R}^d)$ and $x, y \in \mathbb{R}^d$. Then $y \in \partial f(x)$ if and only if $x = \text{prox}_f(x + y)$.

Second, this yields cocoerciveness with coefficient 1, which is an instantaneous consequence of the monotonicity of the proximity operator.

2.16 Corollary.

Proximity operators are 1-cocoercive.

Proof.

Let $f \in \Gamma_0$ and $x, y \in \mathcal{H}$ be arbitrary. Let now $u := \text{prox}_f(x)$ and $v := \text{prox}_f(y)$. Then by Corollary 2.15, $x - u \in \partial f(x)$ and $y - v \in \partial f(y)$. As the subdifferential is monotone, we have

$$\langle (x - u) - (y - v), u - v \rangle \geq 0$$

which ends the proof.

qed

2. Convex Analysis

The end of this section will relate the proximity operator to the notions of the preceding section. We first show that a proximity operator is firmly nonexpansive and its Moreau envelope is Fréchet differentiable, cf. Definitions 2.6 and 2.9.

2.17 Theorem (12.27 and 12.29 in [BC11]).

Let $f \in \Gamma_0(\mathcal{H})$ and $\lambda > 0$. Then it holds that

- a) The operator $\text{prox}_{\lambda f}: \mathcal{H} \rightarrow \mathcal{H}$ is firmly nonexpansive,
- b) The function $M_{\lambda f}$ is (Fréchet) differentiable and has a Lipschitz-continuous gradient given by

$$\nabla M_{\lambda f}(x) = x - \text{prox}_{\lambda f}(x).$$

Furthermore, we relate a proximity operator to the Fréchet gradient of its Moreau envelope. This subtle relatedness will become beneficial in Section 4.2.

2.18 Corollary.

Let $\lambda > 0$ and f plc. For $\Phi := \frac{1}{2}\|\cdot\|^2 - M_{\lambda f}$ we see

$$\text{prox}_{\lambda f}(x) = \nabla \left(\frac{1}{2}\|\cdot\|^2 - M_{\lambda f}(x) \right) = \nabla \Phi(x).$$

Proof.

Considering that $\text{prox}_{\lambda f}$ is nonexpansive by Corollary 2.16 and that Φ is convex, the claim follows directly from Theorem 2.17 b). qed

Finally, it was shown by Moreau that also the following (reverse) statement holds true. This provides context in the sense of nonexpansiveness and the subdifferential, cf. Definitions 2.7 and 2.10.

2.19 Theorem (10.c in [Mor65]).

The operator $f: \mathcal{H} \rightarrow \mathcal{H}$ is a proximity operator if and only if it is nonexpansive and there exists a function $\psi \in \Gamma_0(\mathcal{H})$, such that for any $x \in \mathcal{H}$ we have $f(x) \in \partial\psi(x)$.

2.3. Operator Splitting Algorithms

This section is based on [BC11, Set10, Och18]. Here, we are looking for a possibly unique minimizer of the sum of two functions. As this problem can become arbitrarily complicated to solve explicitly, we will instead turn our attention to (numerical) algorithms. While there exist a myriad of algorithms with numerous assumptions and requirements, we will ex ante restrict ourselves to a small selection. In this thesis, we consider *Forward-Backward* and *Douglas-Rachford Splitting*. The latter

will also give rise to a relaxation, DR λ or RAAR, that has proven useful in phase retrieval problems, cf. [LP14, Lukos, LZ17]. Interestingly, these two algorithms can also be understood as two special forms of a more general, relaxable algorithm, the so-called *Generalized Forward-Backward Splitting*, cf. [RFP13]. In this context remember the notation of the kernel, $\mathcal{N}(f) := \{x : f(x) = 0\}$.

We will first consider the forward-backward splitting algorithm as stated by Bauschke and Combettes. This includes a convergence guarantee that fortunately is stated in extensive generality.

2.20 Proposition (25.9 in [BC11]).

Let $D \subseteq \mathcal{H}$ be nonempty and closed, let $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ be maximally monotone and such that $\text{dom}(A) \subseteq D$, let $B: D \rightarrow \mathcal{H}$, let $\alpha, \beta > 0$. Suppose that one of the following holds

- A is α -strongly monotone, B is β -cocoercive and $\gamma \in (0, 2\beta)$,
- $\alpha \leq \beta$, B is α -strongly monotone and β -Lipschitz continuous and $\gamma \in (0, 2\alpha/\beta^2)$.

Let $x^{(0)} \in D$ and update

$$\begin{aligned} y^{(n)} &:= x^{(n)} - \gamma Bx^{(n)} \\ x^{(n+1)} &:= J_{\gamma A}y^{(n)}. \end{aligned}$$

Then $(x^{(n)})$ converges linearly to the unique point in $\mathcal{N}(A + B)$.

Note that the assumptions directly ensure that $\mathcal{N}(A + B)$ is not empty.

When comparing distinct, yet akin, iterative algorithms, it can be useful to reformulate the iteration instructions in the form of a single line and to compare those. Capturing the characteristics of Proposition 2.20, the update step can be written in one line as

$$x^{(n+1)} = J_{\gamma A}(x^{(n)} - \gamma Bx^{(n)}). \quad (2.1)$$

We will now consider Douglas-Rachford splitting.

2.21 Theorem (25.6 in [BC11]).

Let A and B be maximally monotone operators from \mathcal{H} to $2^{\mathcal{H}}$ such that $\mathcal{N}(A + B) \neq \emptyset$, let (λ_n) be a sequence in $[0, 2]$ such that $\sum_{n \in \mathbb{N}} \lambda_n(2 - \lambda_n) = +\infty$, let $\gamma > 0$ and $x^{(0)} \in \mathcal{H}$. Update

$$\begin{aligned} y^{(n)} &:= J_{\gamma B}x^{(n)} \\ z^{(n)} &:= J_{\gamma A}(2y^{(n)} - x^{(n)}) \\ x^{(n+1)} &:= x_n + \lambda_n(z^{(n)} - y^{(n)}) \end{aligned}$$

Then there exists $x \in \text{Fix}(R_{\gamma A}R_{\gamma B})$ such that

2. Convex Analysis

- $J_\gamma Bx \in \mathcal{N}(A + B)$,
- $(x^{(n)})$ converges weakly to x ,
- both $(y^{(n)})$ and $(z^{(n)})$ converge weakly to $J_\gamma Bx$,
- if either A or B is uniformly monotone on every nonempty bounded subset of its domain, both $(y^{(n)})$ and $(z^{(n)})$ converge strongly to the unique point in $\mathcal{N}(A + B)$.

Again, the update step can be summarized in one line

$$x^{(n+1)} = x^{(n)} + \lambda_n (J_{\gamma A}(2J_{\gamma B}x^{(n)} - x^{(n)}) - J_{\gamma B}x^{(n)}). \quad (2.2)$$

While this looks decidedly different to Equation (2.1), there exists research on how to incorporate both Equations (2.1) and (2.2) into a larger, generalized class of iterative operator splitting algorithms. The interested reader is advised to the reading of [RFP13].

Reflections and Relaxation

We now discuss a special case of Douglas-Rachford, that is, the reflection off convex sets. Consider two convex sets S and M and denote their indicator functions ι_S and ι_M . Let now $\text{prox}_{\iota_S} = P_S$ the projector onto S and analogously for $\text{prox}_{\iota_M} = P_M$ onto M . Based on that we can now define the reflectors R_M and R_S of the sets by simply putting $R_M = 2P_M - \text{id}$ and $R_S = 2P_S - \text{id}$.

Plugging this into the Douglas-Rachford algorithm with $\lambda_n \equiv 1$, $\gamma = 1$ and $A = \partial\iota_S$, $B = \partial\iota_M$, the update instruction reads

$$x^{(n+1)} = \frac{1}{2}(R_S R_M + \text{id})x^{(n)}. \quad (2.3)$$

In many applications, especially when inexact data leads to empty intersections of constraint sets, we want to relax the applied algorithm to also address previously infeasible input and parameter regions. To the above formulation in Equation (2.3), there exists a well-studied relaxation, which is often referred to as DR λ , cf. [Luko5].¹ For a relaxation constant β rewrite the update rule as

$$x^{(n+1)} := \beta \frac{1}{2}(R_S R_M + \text{id})x^{(n)} + (1 - \beta)P_M x^{(n)}. \quad (2.4)$$

This can be extended similarly to the Douglas-Rachford algorithm of Theorem 2.21, yielding

$$x^{(n+1)} := \beta \left[x^{(n)} + \lambda_n J_{\gamma A}(2J_{\gamma B}x^{(n)} - x^{(n)}) \right] + (1 - \beta - \beta\lambda_n)J_{\gamma B}x^{(n)}.$$

Convergence results to this form can be found in [LM20]. For more reading on relaxations of the Douglas-Rachford splitting algorithm, I refer to [DP18].

¹Sometimes Equations (2.3) and (2.4) are called *(Relaxed) Alternating Averaged Reflections*, (R)AAR, [BCLo4, Luko5].

3. Compressed Sensing

Referencing [FR13, 岡藤伏+14], compressed sensing is widely understood as the mathematical discipline which aims at solving underdetermined linear systems. That is, as is general knowledge, finding solutions $x \in \mathbb{C}^n$ of the problem

$$Ax = y,$$

provided the measurement matrix $A \in \mathbb{C}^{m \times n}$ and the measurement $y \in \mathbb{C}^{m \times n}$, where $m < n$. Assuming that one solution exists, this problem possesses a non-trivial solution subspace, i.e. infinitely many solutions.

When addressing this problem with algorithmic ideas, one needs some additional priors, the exploitation of which often enables recovery and, depending on the setting, even uniqueness thereof. Commonly, sparsity is the key to either enabling algorithms or to making them run more stably or smoothly. Another focus of research is the suitable choice of a measurement operator that allows modelling of physically observable scenarios on the one hand, and provides sufficient conditions for theoretical or algorithmic solution of the problem on the other hand. One well-established option are random matrices, e.g. Gaussian matrices.

In this chapter we discuss three topics that fall under the sphere of compressed sensing. First, in Section 3.1, we present the idea of so-called lifting that reformulates a bilinear map as a linear map of higher dimension. This entails applicability of the entirety of concepts and algorithms of the much better understood linear problems.

Second, Section 3.2 gets back to the above mentioned choice of measurement operator. We present the so-called restricted isometry property (RIP) that allows the classification of the distance of operators to unitary ones. We also consider the previously stated Gaussian matrices and show that they indeed fulfill said RIP condition.

Third, we discuss the concentration of mass within a signal in Section 3.3. That is, we analyze how many large entries (in modulus) dominate the ℓ^2 -norm of the whole vector. Aiming for representation of this property, we present a norm that quantifies this behavior, which will become a crucial tool in Chapter 5.

3.1. Lifting

In some applications, one is facing not an underdetermined linear but bilinear problem. In order to be able to access the manifold of existing analysis and solution tools of linear problems, it is desirable to trace these problems back to linear ones. In this section, we present the so-called *Lifting* procedure that is able to achieve exactly this. The idea behind this scheme is to consider bilinear measurements of the vectors u and v as a linear measurement of their outer product uv^* . We introduce this concept based on [BBCE09, CSV12, LWB18].

Assume the noiseless measurement $b \in \mathbb{C}^m$ is provided as

$$b := B(u, v),$$

with respect to the bilinear map $B: \mathbb{C}^{n_1} \times \mathbb{C}^{n_2} \rightarrow \mathbb{C}^m$. We then can represent this bilinear map B by a linear map, which is a standard result in tensor theory. Here we consider the outer vector product as a tensor product, which yields the following reading of the theorem presented by Ryan, [Rya02].

3.1 Theorem (Cf. [Rya02]).

Let $B: \mathbb{C}^{n_1} \times \mathbb{C}^{n_2} \rightarrow \mathbb{C}^m$ be a bilinear map. Then there exists a uniquely defined linear operator $\mathcal{A}: \mathbb{C}^{n_1 \times n_2} \rightarrow \mathbb{C}^m$ that satisfies

$$B(u, v) = \mathcal{A}(uv^*)$$

for all pairs of vectors $u \in \mathbb{C}^{n_1}$ and $v \in \mathbb{C}^{n_2}$.

Much more relevant for algorithmic applicability is the fact that such a linear map \mathcal{A} is fully characterized by a unique set of (representation) matrices in a way that the k -th entry of $\mathcal{A}(X)$, denoted as $[\mathcal{A}(X)]_k$, is given by

$$[\mathcal{A}(X)]_k = \text{tr}(A_k^* X). \quad (3.1)$$

The right hand side of the equation above is easily recognizable as the Frobenius inner product of the representation matrix A_k and the argument X . Hence, the existence of these matrices is an immediate consequence of the Riesz representation theorem, [Rud91].

3.2. Restricted Isometry Property

Unitary operators are convenient in various contexts. In applications one often exploits, besides numerical stability, that these are length-preserving. The downside, however, is that the space of unitary operators is comparably small and many real-life applications are not representable by

unitary operators. Thus, for linear operators of the above form $\mathcal{A}: \mathbb{C}^{n_1 \times n_2} \rightarrow \mathbb{C}^m$ we would like to have a means of quantification to describe how close \mathcal{A} is to a unitary operator. This is often done via the so-called restricted isometry property.

3.2 Definition (Cf. [LWB18]).

A linear operator $\mathcal{A}: \mathbb{C}^{n_1 \times n_2} \rightarrow \mathbb{C}^m$ has the (s_1, s_2, r) -Restricted Isometry Property (RIP) with *Restricted Isometry Constant* δ if

$$(1 - \delta) \|X\|_F^2 \leq \|\mathcal{A}(X)\|_2^2 \leq (1 + \delta) \|X\|_F^2$$

for all matrices $X \in \mathbb{C}^{n_1 \times n_2}$ of rank at most r with at most s_1 non-zero rows and at most s_2 non-zero columns.

While there exists theory on constructing matrices satisfying a suitable RIP, cf. [Mix15], in most cases one defaults to using random matrices instead. While being an active research topic on its own, cf. [Ver12, KW11], we will here restrict ourselves to the definition of a Gaussian measurement operator in Definition 3.3 and its – highly probable – restricted isometry property in Proposition 3.4.

3.3 Definition.

An operator $\mathcal{A}: \mathbb{C}^{n_1 \times n_2} \rightarrow \mathbb{C}^m$ is called an *i.i.d. Gaussian Measurement Operator* if all the entries of the matrices $\{A_k\}_k$ as in Equation (3.1) are independent with distribution $\mathcal{CN}(0, \frac{1}{m})$.

That is, the entries of both $\text{re}(\{A_k\}_k)$ and $\text{im}(\{A_k\}_k)$ follow a real normal distribution $\mathcal{N}(0, \frac{1}{2m})$.

Gaussian measurement operators satisfy an RIP with high probability. One of the benefits of the following proposition is that it smoothly incorporates sparsity, which is integral to almost every compressed sensing application by adequately scaling the number of required measurements.

3.4 Proposition (Cf. Theorem III.7 in [LWB18]).

Let $\mathcal{A}: \mathbb{C}^{n_1 \times n_2} \rightarrow \mathbb{C}^m$ be an *i.i.d. Gaussian measurement operator*. There are absolute constants $c_1, c_2 > 0$, such that if

$$m \geq \frac{c_1}{\delta^2} r (s_1 + s_2) \log \left(\max \left\{ \frac{n_1}{s_1}, \frac{n_2}{s_2} \right\} \right),$$

for some $\delta > 0$, then \mathcal{A} has the (s_1, s_2, r) -restricted isometry property with restricted isometry constant δ with probability at least $1 - \exp(-c_2 m)$.

Scaling the number of measurements and thereby adapting the resulting RIP of the measurement operator is one of the core ideas of Chapter 5. In that context, we will revisit the above statement in a customized form in Theorem 5.4.

3.3. Peakiness

When recovering a distorted, e.g. noisy, signal from its linear measurements, sometimes there exist algorithms that excel at correctly identifying and recovering the largest entry (in modulus) of the desired signal whenever it is reasonably distinguished from the other elements. Often however, this comes at the cost of poor performance when signals are closer to a completely generic one, i.e. all entries having a modulus close to $1/\sqrt{n}$ for a signal with n non-zero entries.

In order to be able to analyze the performance of algorithms depending on the degree of genericity we will now present a norm to quantify the notion of peakiness. That is, a measure of how much mass is concentrated in a specified number of entries. In this context we will generally consider peakiness to be the counterpart of genericity. This is constructed as a prematurely cancelled calculation of the ℓ^2 -norm when rearranging the entries in non-increasing order.

3.5 Notation (Cf. [GKS19]).

Let $x \in \mathbb{C}^n$. We then define the following shorthand notation

$$\|x\|_{[k]} := \max_{\substack{I \subseteq \{1, \dots, n\} \\ |I|=k}} \left(\sum_{i \in I} |x_i|^2 \right)^{1/2}.$$

Using the notion of the non-increasing rearrangement, cf. [List of Symbols](#), the understanding of this notation as a shortened ℓ^2 -norm becomes easily visible, as

$$\|x\|_{[k]}^2 := \max_{\substack{I \subseteq \{1, \dots, n\} \\ |I|=k}} \sum_{i \in I} |x_i|^2 = \sum_{i=1}^k (x_i^*)^2 = \left\| (x_1^* \dots x_k^* 0 \dots 0)^T \right\|_2^2.$$

To furthermore quantify the relative content of information within a signal we consider the additional notion of the peak to average power ratio, i.e. the squared ratio between ℓ^2 and ℓ^∞ norm scaled by the size of the signal.

3.6 Definition (Cf. [NK03]).

Let $x \in \mathbb{C}^n \setminus \{0\}$. We call the quotient

$$n \left(\frac{\|x\|_\infty}{\|x\|_2} \right)^2$$

its *Peak to Average Power Ratio* (PAPR).

For further reading in the context of information density etc., I commend the interested reader to [CT06], where further suchlike tools can be found in terms of information theory.

Part III.

Feature Extraction

In this part we present two vastly different methods to promote sparsity in the signals at hand. Relying on the ubiquitous *sparsity of effects*, i.e. the observation that naturally occurring data can be sparsely represented with respect to a suitable encoding, we assert that enforcing sparsity in the right context will automatically yield emphasis on or enable the extraction of desired features.

In Chapter 4 we consider soft shrinkage as a method to promote sparsity. The soft shrinkage operator with parameter λ , \mathcal{S}_λ , can be understood in different ways. Firstly, as stated in Example 2.12, it is the proximity operator of the ℓ^1 -norm which is widely accepted as the continuous proxy for the *true* sparsity measure, i.e. the ℓ^0 -pseudonorm. Secondly, as a geometrical interpretation, \mathcal{S}_λ moves every value closer to zero by λ . Except for those that were smaller than λ in the beginning, which are set to zero instead. Both approaches yield the intuition that application of this operator favors sparse signals. This motivates the question, how this may be extended to enforce sparsity with respect to arbitrary complete systems, without losing the property of being a proximity operator. To this end, we analyze under what circumstances expressions of the form $T^\dagger \text{prox}_\Phi T$ are proximity operators, too.

In Section 4.1 we consider the above scenario with finite-dimensional frames T and the soft thresholding operator, i.e. $\Phi = \|\cdot\|_1$. This also includes a geometrical exemplification in order to gain a heuristic understanding of this nesting concept. In Section 4.2 we extend the range of operators to T being an injective operator with closed range between arbitrary Hilbert spaces and without posing restrictions on Φ . This, however, in turn leads to an unillustratable degree of abstractness.

Chapter 5 considers the aim of sparsification from a compressed sensing point of view. That is, in contrast to the above, we do not intend to denoise or correct a corrupted signal but instead aim to recover it from its measurements with respect to a known measurement operator. In this specific case, we consider a bilinear measurement by an operator satisfying a suitable restricted isometry property, e.g. a Gaussian. We then present the Sparse Power Factorization algorithm of [LWB18] in Section 5.1, which is able to address this problem in the case of particularly peaky signals. As this often does not match physically observable data, we extend the range of applicable parameters by providing a reverse relationship between the number of required measurements and the peakiness of the signals to be recovered.

While this part functions as the theoretical part and the core of this thesis, in Part IV the methods of both chapters will also be analyzed with regard to the applicability in an existing real world context. There, we will discuss an application of the schemes portrayed within this part, in the context of phase retrieval as considered in X-ray nanoscale imaging.

4. Nested Proximity Operators

As touched on in the preamble to this part, we aim at an emphasis of sparsity with the goal of feature accentuation. This chapter takes the approach of so-called soft shrinkage, or soft thresholding. This process is well studied, we exemplarily refer to [Don95].

Utilizing the ℓ^1 -norm as a convex proxy for sparsity is a common procedure, cf. e.g. [CWBo8, BJMO12]. Doing so does however often fall short of enforcing sparsity in a regard suitable to the application at hand. Instead, sparsity with respect to an encoding scheme, e.g. a frame, cf. Section 1.2, is desirable. This motivates the application of a composition of some sort of encoding followed by the shrinkage operator and an, ideally matching, decoding step. For a frame T and the soft shrinkage operator S_γ this could take the form

$$T^{-1}S_\gamma T$$

for $\gamma > 0$ and a matrix $T \in \mathbb{R}^{L \times N}$ with $L \geq N$ and full rank N . However, non-trivial frames are in general not directly invertible, which is why we will consider the pseudoinverse, cf. Definition 1.1, in lieu thereof. Thence, we obtain

$$T^\dagger S_\gamma T.$$

Analyzing expressions of this kind in finite dimensions will be the subject of Section 4.1. In general, this differs from intertwining the shrinkage operator with the frame analysis and synthesis operators due to the usage of the pseudoinverse. We study some geometrical properties of this expression and are able to show that this is indeed a proximity operator, cf. Section 2.2. This opens the variety of tools of convex analysis to be applied to this kind of expression, e.g. operator splitting algorithms, cf. Section 2.3.

Second, in Section 4.2, we expand the notion of frames to arbitrary injective operators and use the broader concept of proximity operators instead of restricting ourselves to soft shrinkage. That is, we consider compositions of the form

$$T^\dagger \text{prox}_\varphi T,$$

where φ is a plc function and T an injective operator with bounded range between arbitrary Hilbert spaces. Here we characterize the result of the application of linear transform, proximity operator and pseudoinverse transform. We are able to show that this is a proximity operator of a certain plc function. We further provide a closed form for said function in terms of an infimal convolution.

4.1. Frame Shrinkage

Note: The following portion of this thesis is predominantly taken from [GP19] of which I am a coauthor. Until the end of Section 4.1 the content is in large portions identical to the one in [GP19]. I have occasionally adapted the order, notation or, where adequate, wording and added further thoughts if considered reasonable. I will also at times leave out parts that have been referenced earlier within this thesis or are irrelevant with regard to the concern of this work. In these cases I have put references to the corresponding location whenever I believe this to be beneficial.

In this section, we will generally use T to represent an arbitrary frame and $S_\gamma := \text{prox}_{\gamma \|\cdot\|_1}$ the soft shrinkage operator introduced in Example 2.12. We will heavily rely on previously introduced concepts, most prominently the proximity operator, Definition 2.11, and the pseudoinverse, Definition 1.1. We will show that the commonly used frame shrinkage operator that maps a given vector $x \in \mathbb{R}^N$ onto the vector $T^\dagger S_\gamma T x$, is already a proximity operator, which can therefore be directly used in corresponding splitting algorithms.

Many reconstruction problems in signal and image processing are ill-posed and commonly solved by means of variational methods. In the finite-dimensional case, the ill-posed operator equation $Kx = f$ is often solved using a regularization approach,

$$\tilde{x} = \arg \min_{x \in \mathbb{R}^N} F(x) = \arg \min_{x \in \mathbb{R}^N} \left(\frac{1}{2} \|Kx - f\|_2^2 + \Phi(x) \right). \quad (4.1)$$

Here, $K: \mathbb{R}^N \rightarrow \mathbb{R}^M$ is a known linear or non-linear operator and $f \in \mathbb{R}^M$ represents the measured (noisy) data. For regularization, a plc functional $\Phi: \mathbb{R}^N \rightarrow \mathbb{R} \cup \{\infty\}$ as introduced in Definition 2.1 is employed that enforces desired properties of x such as regularity or sparsity. The observant reader will easily spot the familiarity of Equation (4.1) with the proximity operator of Definition 2.11.

In many applications, signals tend to be sparse with regard to a suitable encoding, and functionals of the form $\Phi = \gamma \|\cdot\|_1$ or $\Phi = \gamma \|T \cdot\|_1$ with $\gamma > 0$ and a linear transform $T \in \mathbb{R}^{L \times N}$ are frequently used.

While orthogonal transforms T have proven to be ideal with regard to numeric stability, they often fail to capture the underlying characteristics of many signal types. Therefore, in numerous applications one uses redundant representations that suit the purpose at hand, as e.g. redundant wavelet transforms, curvelets or shearlets for capturing non-isotropic features in 2D images, cf. [ELL08, KKL12, MP10] as well as Sections 1.3 and 1.4. In this case we have rectangular transform matrices T with $L > N$.

Within the last years, many computational algorithms have been proposed to solve the minimization problem in Equation (4.1), which are, for example, based on operator splitting methods, such as

forward-backward splitting (FBS) or Douglas-Rachford splitting (DRS), which we have introduced in Section 2.3. These approaches make use of the so-called *proximity operator* of Φ , cf. Definition 2.11. If we take the standard inner product and let $\Phi = \gamma \|\cdot\|_1$, we recall from Example 2.12 that we obtain the so-called soft shrinkage operator.

For $\Phi = \gamma \|T \cdot\|_1$ with T being an orthonormal matrix, it can be shown that the corresponding proximity operator is

$$\text{prox}_{\gamma \|T \cdot\|_1} = T^* S_\gamma T = T^{-1} S_\gamma T,$$

see Proposition 23.29 in [BC11]. This, however is no longer true if T is not orthogonal or if the rows of T do not even form a basis, i.e., if $T \in \mathbb{R}^{L \times N}$ for $L > N$, see e.g. [Ela06]. In this case, the proximity operator can no longer be represented in a closed form, and one has to make use of an iteration procedure to compute it.

Yet the question remains, what would happen if we replaced the exact proximity operator with the frame soft shrinkage operator $T^\dagger S_\gamma T$, also for rectangular matrices $T \in \mathbb{R}^{L \times N}$. For that investigation we replace $\text{prox}_{\gamma \|T \cdot\|_1}$ by $T^\dagger S_\gamma T$ in iteration procedures such as, e.g., forward-backward splitting or Douglas-Rachford splitting. Will these splitting algorithms still converge? And, if so, how will the regularization functional change?

A related question, but from a different point of view, has also been posed by Elad in [Ela06]. His main argument, to explain the good performance of methods that just employ the frame soft shrinkage operator, is based on the connection to the solution of basis pursuit denoising (BPDN) problems [CDS98]. He showed that the application of the frame soft shrinkage operator can be interpreted as the first iteration step of an iterative algorithm to solve the BPDN problem.

We would like to present another heuristic argument towards the use of the frame soft shrinkage operator. In phase retrieval, for example, one needs to address the inverse problem introduced by the loss of the phase information. In practice, this problem is often formulated as a feasibility problem. In our case we can define the measurement set

$$M := \{x : Kx = f\}$$

and a constraints set, as for example

$$C := \{x : \|x\|_2 \leq \|K^\dagger f\|_2, Tx \text{ is } s\text{-sparse}\}.$$

We are looking for some

$$x \in M \cap C. \tag{4.2}$$

It is quite likely that $M \cap C = \emptyset$, particularly for noisy data f . In this case we aim at finding $x \in C$ with shortest distance to M , or, more generally, x corresponding to the smallest sum of distances to C and M . The most simple iteration algorithms to tackle the feasibility problem of Equation (4.2)

4. Nested Proximity Operators

are alternating projection algorithms. We start with an arbitrary x_0 , apply the projection to M , then the projection to C , and then iterate. Another, more sophisticated algorithm is the relaxed version of Douglas-Rachford, DR λ of [Luk05]. As we have seen in Section 2.3, this will converge to the zeroset of the sum of two operators. In our case we can use this to model the distance to M and C respectively. The relaxation in turn is used to tackle the infeasible case $M \cap C = \emptyset$.

For linear K the projection onto M is well-defined, since M is an affine subspace of \mathbb{R}^N . The set C is bounded but usually no longer convex, and the projection cannot be obtained easily. For given $x \in C$ we could apply the frame shrinkage operator with a suitable γ such that $S_\gamma T x$ is s -sparse.

However, the frame soft shrinkage operator is not the orthogonal projector onto C , it is not even idempotent. But it obviously provides a vector which is somewhat close to C . In [LP14, Loo16], this approach has been used in connection with the DR λ algorithm for phase retrieval problems with very good numerical results.

Tailoring Operator Splitting Methods

To solve Equation (4.1), we observe that the first term $\|Kx - f\|_2^2$ of the functional F is differentiable. This may no longer hold true for the convex functional Φ . Therefore we apply the subdifferential as introduced in Definition 2.7. Note that the solution \tilde{x} of Equation (4.1) necessarily satisfies

$$0 \in \partial \left(\frac{1}{2} \|K\tilde{x} - f\|_2^2 + \Phi(\tilde{x}) \right) = K^*(K\tilde{x} - f) + \partial\Phi(\tilde{x}).$$

Multiplication with a constant $\lambda > 0$ and addition of \tilde{x} yields the equivalent statements

$$\begin{aligned} \tilde{x} - \lambda K^*(K\tilde{x} - f) &\in \tilde{x} + \lambda \partial\Phi(\tilde{x}) \\ (\text{id} - \lambda K^*K)\tilde{x} + \lambda K^*f &\in (\text{id} + \lambda \partial\Phi)^{-1}(\tilde{x}). \end{aligned}$$

Thus formally

$$\tilde{x} = (\text{id} + \lambda \partial\Phi)^{-1} [(\text{id} - \lambda K^*K)\tilde{x} + \lambda K^*f]. \quad (4.3)$$

The operator $(\text{id} + \lambda \partial\Phi)^{-1}$ is also called the resolvent of $\lambda \partial\Phi$, cf. Definition 2.13 and Proposition 2.14.

Indeed, we observe that $\frac{1}{2} \|y - x\|^2 + \lambda \Phi(y)$ is strictly convex and

$$\begin{aligned} y = \text{prox}_{\lambda\Phi}(x) &\iff 0 \in (y - x) + \lambda \partial\Phi(y) = -x + (\text{id} + \lambda \partial\Phi)(y) \\ &\iff y = (\text{id} + \lambda \partial\Phi)^{-1}(x). \end{aligned}$$

In particular,

$$x - y \in \lambda \partial\Phi(y) \iff y = \text{prox}_{\lambda\Phi}(x),$$

or equivalently,

$$t \in \lambda \partial\Phi(y) \iff y = \text{prox}_{\lambda\Phi}(y + t), \quad (4.4)$$

cf. Corollary 2.15. Now, Equation (4.3) already provides an iterative algorithm, the so-called forward-backward splitting iteration, see e.g. [LM79, BCII, Setio]:

4.1 Algorithm (Forward-Backward Splitting).

For an arbitrary starting vector $x^{(0)} \in \mathbb{R}^N$, iterate

$$\begin{aligned} y^{(j)} &:= (\text{id} - \lambda K^* K) x^{(j)} + \lambda K^* f \\ x^{(j+1)} &:= \text{prox}_{\lambda\Phi}(y^{(j)}). \end{aligned}$$

Note that Algorithm 4.1 is a special case of Proposition 2.20, where we select $A = \partial\Phi$ and $B = \partial(\frac{1}{2}\|K \cdot - f\|_2^2) = K^*(K \cdot - f)$.

This iteration converges for $\lambda \in (0, 2/\|K^*\|_2^2)$, see e.g. [Setio], since the map

$$x \mapsto \frac{1}{\|K^*\|_2^2} (K^* K x + K^* f)$$

is firmly nonexpansive, i.e., we have for arbitrary vectors x and y ,

$$\langle x - y, K^* K(x - y) \rangle = \|K(x - y)\|_2^2 \geq \frac{1}{\|K^*\|_2^2} \|K^* K(x - y)\|_2^2.$$

Let us consider a second operator splitting method. Observe that the fixed point relation in Equation (4.3) can be rewritten as

$$\tilde{x} + \lambda K^* K \tilde{x} \in \text{prox}_{\lambda\Phi} [(\text{id} - \lambda K^* K) \tilde{x} + \lambda K^* f] + \lambda K^* K \tilde{x}.$$

With $\tilde{t} := (\text{id} + \lambda K^* K) \tilde{x}$ it follows that

$$\tilde{t} \in \text{prox}_{\lambda\Phi}(2\tilde{x} - \tilde{t} + \lambda K^* f) + \tilde{t} - \tilde{x}.$$

This leads to the Douglas-Rachford iteration, which converges in a large parameter range, see [BCII, Setio].

4.2 Algorithm (Douglas-Rachford Splitting).

For $\lambda > 0$ and any starting vectors $x^{(0)}, t^{(0)} \in \mathbb{R}^N$, iterate

$$\begin{aligned} t^{(j+1)} &:= \text{prox}_{\lambda\Phi}(2x^{(j)} - t^{(j)} + \lambda K^* f) + t^{(j)} - x^{(j)} \\ x^{(j+1)} &:= (\text{id} + \lambda K^* K)^{-1} t^{(j+1)}. \end{aligned}$$

4. Nested Proximity Operators

Note that Algorithm 4.2 is again a special case, this time of Theorem 2.21 with $\lambda_n \equiv 1$ and $\gamma = \lambda$ and $B = K^*K$ and $A = \partial\Psi$ for $\Psi: x \mapsto \Phi(x) + K^*f$.

For the two considered algorithms we need to evaluate the proximity operator of the regularization functional $\lambda\Phi$.

Analyzing The Frame Soft Shrinkage Operator

Our goal is to show that for any frame matrix $T \in \mathbb{R}^{L \times N}$ with $L \geq N$ and full rank N the operator $T^\dagger S_\gamma T$ is the proximity operator of a plc functional. Remember that $S_\gamma = \text{prox}_{\gamma\|\cdot\|_1}$, cf. Example 2.12. In a first step we define the set-valued mapping

$$H: \mathbb{R}^N \longrightarrow 2^{\mathbb{R}^N}$$

$$x \longmapsto \{y : x = T^\dagger S_\gamma T(x + y)\}. \quad (4.5)$$

Then, according to Equation (4.4) and Corollary 2.15, we have to show that H is the subdifferential of a functional $\Phi \in \Gamma_0$.

Now, we will show that H in Equation (4.5) is well-defined, and we will study some properties of H . In a final step, we will then be able to show that indeed $H = \partial\Phi$ for some $\Phi \in \Gamma_0$.

In order to give first insight into the procedure, we start off with a toy example.

4.3 Lemma.

For $T = \begin{pmatrix} 1 \\ c \end{pmatrix}$ with $c \geq 1$ and $\gamma > 0$ we find for H in Equation (4.5) for $x \geq 0$

$$H(x) = \begin{cases} \gamma \left[-\frac{1}{c}, \frac{1}{c}\right] & x = 0, \\ \frac{\gamma}{c} + \frac{x}{c^2} & x \in \left(0, \frac{\gamma(c-1)c}{c^2+1}\right], \\ \gamma \left(\frac{1+c}{1+c^2}\right) & x > \frac{\gamma(c-1)c}{c^2+1}. \end{cases}$$

For $x < 0$ we have $H(x) = -H(-x)$. Then H is the subdifferential of the even function

$$\Phi(x) = \begin{cases} \frac{\gamma x}{c} + \frac{x^2}{2c^2} & x \in \left[0, \frac{\gamma(c-1)c}{c^2+1}\right], \\ \gamma \left(\frac{1+c}{1+c^2}\right) x & x > \frac{\gamma(c-1)c}{c^2+1}, \\ \Phi(-x) & x < 0. \end{cases}$$

Proof.

For $x = 0$, it follows from Equation (4.5) with $T^\dagger = \frac{1}{1+c^2}(1, c)$ that $y \in H(0)$ if and only if

$$\frac{1}{1+c^2}(1, c) S_\gamma \begin{pmatrix} 1 \\ c \end{pmatrix} y = 0.$$

This is only true if $S_\gamma(cy) = 0$, i.e. $y \in [-\frac{\gamma}{c}, \frac{\gamma}{c}]$.

For $x > 0$ we find

$$x = \frac{1}{1+c^2}(1, c) S_\gamma \begin{pmatrix} 1 \\ c \end{pmatrix} (y+x).$$

Since $x > 0$, we need that $S_\gamma T(x+y) > 0$. Thus $c(x+y) > \gamma$ as well as $x+y > 0$. We consider two cases.

If $x+y \leq \gamma$ and $c(x+y) > \gamma$, then

$$x = T^\dagger S_\gamma T(x+y) = \frac{1}{1+c^2}(1, c) \begin{pmatrix} 0 \\ c(x+y) - \gamma \end{pmatrix}$$

implies

$$y = \frac{x}{c^2} + \frac{\gamma}{c}.$$

Further, the condition $x+y = x + (\frac{x}{c^2} + \frac{\gamma}{c}) \leq \gamma$ yields $x \leq \frac{c(c-1)\gamma}{1+c^2}$.

Let now $x+y > \gamma$ and $c(x+y) > \gamma$, then

$$x = T^\dagger S_\gamma T(x+y) = \frac{1}{1+c^2}(1, c) \begin{pmatrix} x+y-\gamma \\ c(x+y)-\gamma \end{pmatrix} = x+y - \frac{\gamma(c+1)}{c^2+1}.$$

Thus, we find $y = \frac{\gamma(c+1)}{c^2+1}$, and $x+y > \gamma$ is true for $x > \frac{\gamma c(c-1)}{c^2+1}$. Similar considerations for $x < 0$ yield $H(-x) = -H(x)$. Integration yields Φ as asserted. **qed**

Considering an explicit example of Lemma 4.3, let $c = 2$ and $\gamma = \frac{5}{3}$. Then we find

$$H(x) = \begin{cases} [-\frac{5}{6}, \frac{5}{6}] & x = 0, \\ \frac{5}{6} + \frac{x}{4} & x \in (0, \frac{2}{3}], \\ 1 & x > \frac{2}{3}, \\ -H(-x) & x < 0, \end{cases} \quad \text{and} \quad \Phi(x) = \begin{cases} \frac{5x}{6} + \frac{x^2}{8} & x \in [0, \frac{2}{3}], \\ x - \frac{1}{18} & x > \frac{2}{3}, \\ \Phi(-x) & x < 0. \end{cases}$$

Thus Φ approximates the absolute value, see Figure 4.1.

4. Nested Proximity Operators

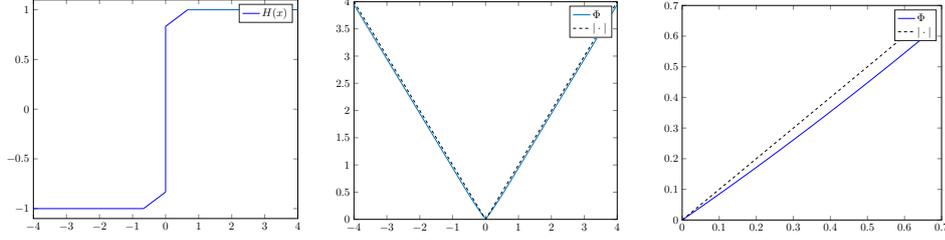


Figure 4.1.: Visualization of H and Φ .

Now we inspect the function H in Equation (4.5) and show first that for any fixed $\gamma > 0$ and each $x \in \mathbb{R}^N$ the set $H(x)$ is not empty.

4.4 Proposition.

Let $T \in \mathbb{R}^{L \times N}$ with $L \geq N$ have full rank N and let $\gamma > 0$. Then, for each $x \in \mathbb{R}^N$, we have $H(x) \neq \emptyset$. Further, for each $x \in \mathbb{R}^N$ we have $H(x) \subset \{y \in \mathbb{R}^N : \|y\|_\infty \leq \gamma \|T^\dagger\|_\infty\}$, i.e. the image of H is bounded.

Proof.

By Equation (4.5), we have $y \in H(x)$ if $x = T^\dagger S_\gamma T(x + y)$. With the substitution $t = x + y$, we get an equivalent fixed point representation, i.e. $t - x \in H(x)$ if and only if $t = x + (\text{id} - T^\dagger S_\gamma T)t$. This motivates to define a function

$$\begin{aligned} f_{x,T}: \mathbb{R}^N &\longrightarrow \mathbb{R}^N \\ t &\longmapsto x + (\text{id} - T^\dagger S_\gamma T)t. \end{aligned} \quad (4.6)$$

If now $f_{x,T}$ possesses a fixed point t , then $y = t - x$ is an element of $H(x)$. Hence, to prove that $H(x) \neq \emptyset$, we show that for each $x \in \mathbb{R}^N$ the function $f_{x,T}$ possesses at least one fixed point.

First, we consider $f_{x,P}(t) = x + (\text{id} - P^* S_\gamma P)t$, for some matrix $P \in \mathbb{R}^{L \times N}$ satisfying $P^* P = \text{id}$, i.e., $P^\dagger = P^*$. We define the closed ball

$$B(x, \gamma \|P^*\|_\infty, \ell^\infty) := \{t \in \mathbb{R}^N : \|x - t\|_\infty \leq \gamma \|P^*\|_\infty\}.$$

Then $f_{x,P}(t) \in B(x, \gamma \|P^*\|_\infty, \ell^\infty)$ for each $t \in \mathbb{R}^N$, since we have

$$\begin{aligned} \|x - f_{x,P}(t)\|_\infty &= \|(\text{id} - P^* S_\gamma P)t\|_\infty = \|P^*(\text{id} - S_\gamma)Pt\|_\infty \\ &\leq \|P^*\|_\infty \|(\text{id} - S_\gamma)Pt\|_\infty \leq \|P^*\|_\infty \sup_{s \in \mathbb{R}^L} \|(\text{id} - S_\gamma)s\|_\infty \leq \gamma \|P^*\|_\infty. \end{aligned}$$

Since S_γ is continuous, $f_{x,P}$ is also continuous, and it follows by Brouwer's fixed point theorem, cf. [Sek57] that there exists a fixed point of $f_{x,P}$ within $B(x, \gamma \|P^*\|_\infty, \ell^\infty)$.

Second, we recall that the matrix $T \in \mathbb{R}^{L \times N}$ possesses a (thin) singular value decomposition, cf. Definition 1.14,

$$T = PDQ^*,$$

where $P \in \mathbb{R}^{L \times N}$ has N orthogonal columns, i.e., $P^*P = \text{id}$, $D = \text{diag}(d_1, \dots, d_N) \in \mathbb{R}^{N \times N}$ contains the positive singular values of T , and $Q \in \mathbb{R}^{N \times N}$ is orthogonal. In particular, we have

$$T^\dagger = QD^{-1}P^*.$$

Using the singular value decomposition of T , we discover that the function $f_{DQ^*x,P}$, being defined analogously to $f_{x,T}$ – with DQ^*x instead of x and with the matrix P instead of T – satisfies

$$\begin{aligned} f_{DQ^*x,P}(t) &= DQ^*x + (\text{id} - P^*S_\gamma P)t \\ &= DQ^*x + DQ^*(QD^{-1})(\text{id} - P^*S_\gamma P)DQ^*(QD^{-1}t) \\ &= DQ^*x + DQ^*(\text{id} - T^\dagger S_\gamma T)(QD^{-1}t) \\ &= DQ^*(x + (\text{id} - T^\dagger S_\gamma T)(QD^{-1}t)) \\ &= DQ^*f_{x,T}(QD^{-1}t). \end{aligned}$$

Reversely,

$$f_{x,T}(t) = QD^{-1}f_{DQ^*x,P}(DQ^*t). \quad (4.7)$$

Since P satisfies $P^*P = \text{id}$, we already know from the first step that $f_{DQ^*x,P}$ possesses a fixed point \bar{x} . Now, Equation (4.7) implies

$$f_{x,T}(QD^{-1}\bar{x}) = QD^{-1}f_{DQ^*x,P}(DQ^*(QD^{-1}\bar{x})) = QD^{-1}\bar{x},$$

i.e., $f_{x,T}$ possesses at least the fixed point $QD^{-1}\bar{x}$.

Moreover, using $T^\dagger T = \text{id}$, we obtain for each fixed point $t = f_{x,T}(t)$,

$$\begin{aligned} \|x - t\|_\infty &= \|x - f_{x,T}(t)\|_\infty = \|(\text{id} - T^\dagger S_\gamma T)t\|_\infty \\ &= \|T^\dagger(\text{id} - S_\gamma)Tt\|_\infty \leq \|T^\dagger\|_\infty \|(\text{id} - S_\gamma)Tt\|_\infty \leq \|T^\dagger\|_\infty \gamma. \end{aligned}$$

Since $t - x \in H(x)$, the boundedness of the image of H follows. qed

Thus, we conclude that the mapping H in Equation (4.5) is well defined. Next we show that $H(0)$ is indeed not single-valued.

4.5 Proposition.

Let $T \in \mathbb{R}^{L \times N}$ with $L \geq N$ with full rank N . Further let $\gamma > 0$ and H as in Equation (4.5). Then $y \in H(0)$ if and only if $\|T y\|_\infty \leq \gamma$.

4. Nested Proximity Operators

Proof.

We recall that the pseudoinverse from Definition 1.1 and that $T^\dagger T = (T^* T)^{-1} T^* T = \text{id}$ as T has full rank N .

Now consider the soft shrinkage operator S_γ from Example 2.12 and let $\|T y\|_\infty \leq \gamma$. Then the definition of S_γ implies $S_\gamma T y = 0$ and hence also $T^\dagger S_\gamma T (y + 0) = 0$, that is, $y \in H(0)$. Next, let $y \in H(0)$, i.e.,

$$T^\dagger S_\gamma T y = 0. \quad (4.8)$$

We show that then $\|T y\|_\infty \leq \gamma$. We consider the components $[T y]_j$, for $1 \leq j \leq L$, and define three index sets I_1, I_2, I_3 that form a partition of $\{1, \dots, L\}$,

$$\begin{aligned} I_1 &:= \{1 \leq j \leq L : (T y)_j > \gamma\}, \\ I_2 &:= \{1 \leq j \leq L : (T y)_j < -\gamma\}, \\ I_3 &:= \{1 \leq j \leq L : (T y)_j \in [-\gamma, \gamma]\}. \end{aligned}$$

Suppose that $\|T y\|_\infty > \gamma$, which means that $I_1 \cup I_2 \neq \emptyset$. Then

$$S_\gamma T y = \sum_{j \in I_1} ([T y]_j - \gamma) e_j + \sum_{j \in I_2} ([T y]_j + \gamma) e_j, \quad (4.9)$$

where e_j denotes the j -th unit vector in \mathbb{R}^L . Now we combine Equations (4.8) and (4.9) to get

$$\begin{aligned} 0 &= T^\dagger S_\gamma T y = \sum_{j \in I_1} ([T y]_j - \gamma) T^\dagger e_j + \sum_{j \in I_2} ([T y]_j + \gamma) T^\dagger e_j \\ &= \sum_{j \in I_1} ([T y]_j - \gamma) \tilde{v}_j + \sum_{j \in I_2} ([T y]_j + \gamma) \tilde{v}_j, \end{aligned} \quad (4.10)$$

where $\tilde{v}_j := T^\dagger e_j$. In other words, the set $\{\tilde{v}_j : j \in I_1 \cup I_2\}$ is linearly dependent. At the same time, none of these vectors \tilde{v}_j vanish because $\tilde{v}_j = 0$ for $j \in I_1 \cup I_2$ leads to the contradiction

$$0 = |\langle \tilde{v}_j, (T^* T) y \rangle| = |\langle T^\dagger e_j, (T^* T) y \rangle| = |\langle (T^* T) T^\dagger e_j, y \rangle| = |\langle T^* e_j, y \rangle| = |[T y]_j| > \gamma.$$

Without loss of generality, assume $I_1 \neq \emptyset$ and choose $j_1 \in I_1$. Then $[T y]_{j_1} > \gamma$ and Equation (4.10) implies

$$\tilde{v}_{j_1} = \sum_{\substack{j \in I_1 \\ j \neq j_1}} \frac{\gamma - [T y]_j}{[T y]_{j_1} - \gamma} \tilde{v}_j + \sum_{j \in I_2} \frac{-\gamma - [T y]_j}{[T y]_{j_1} - \gamma} \tilde{v}_j. \quad (4.11)$$

A closer look at the coefficients shows that

$$\frac{\gamma - [T y]_j}{[T y]_{j_1} - \gamma} < 0 \text{ for } j \in I_1 \setminus \{j_1\} \quad \text{and} \quad \frac{-\gamma - [T y]_j}{[T y]_{j_1} - \gamma} > 0 \text{ for } j \in I_2.$$

Hence, we find with $T^*T\tilde{v}_j = T^*e_j$,

$$\begin{aligned}
[Ty]_{j_1} &= \langle Ty, e_{j_1} \rangle = \langle y, (T^*T)(T^*T)^{-1}T^*e_{j_1} \rangle = \langle y, (T^*T)\tilde{v}_{j_1} \rangle \\
&= \left\langle y, (T^*T) \left(\sum_{\substack{j \in I_1 \\ j \neq j_1}} \frac{\gamma - [Ty]_j}{[Ty]_{j_1} - \gamma} \tilde{v}_j + \sum_{j \in I_2} \frac{-\gamma - [Ty]_j}{[Ty]_{j_1} - \gamma} \tilde{v}_j \right) \right\rangle \\
&= \sum_{\substack{j \in I_1 \\ j \neq j_1}} \frac{\gamma - [Ty]_j}{[Ty]_{j_1} - \gamma} \langle y, (T^*T)\tilde{v}_j \rangle + \sum_{j \in I_2} \frac{-\gamma - [Ty]_j}{[Ty]_{j_1} - \gamma} \langle y, (T^*T)\tilde{v}_j \rangle \\
&= \sum_{\substack{j \in I_1 \\ j \neq j_1}} \underbrace{\frac{\gamma - [Ty]_j}{[Ty]_{j_1} - \gamma}}_{<0} \underbrace{\langle Ty, e_j \rangle}_{>\gamma} + \sum_{j \in I_2} \underbrace{\frac{-\gamma - [Ty]_j}{[Ty]_{j_1} - \gamma}}_{>0} \underbrace{\langle Ty, e_j \rangle}_{<-\gamma} < 0,
\end{aligned}$$

which contradicts the above assumption that $j_1 \in I_1$. Thus, $I_1 \cup I_2 = \emptyset$, i.e., all indices are located within I_3 , which readily shows that $\|Ty\|_\infty \leq \gamma$. qed

Further, if all components of Tx have a modulus greater than $\gamma(\|TT^\dagger\|_\infty + 1)$, we show that $H(x)$ consists of a single element.

4.6 Proposition.

Let $T \in \mathbb{R}^{L \times N}$ with $L \geq N$ with full rank N . Further let $\gamma > 0$ and

$$x \in \mathcal{U}_{T,\gamma} := \left\{ v \in \mathbb{R}^N : |(Tv)_j| > \gamma(\|TT^\dagger\|_\infty + 1) \forall j \in \{1, \dots, L\} \right\}.$$

Then $H(x)$ with H in Equation (4.5) contains a single element. Moreover, denote the set of all vectors in \mathbb{R}^L containing only components ± 1 by $\{u_l : l \in \{1, \dots, 2^L\}\}$. Then, on each set $S_l := \mathcal{U}_{T,\gamma} \cap \{x \in \mathbb{R}^N : \text{sgn}(Tx) = u_l\}$ the mapping H evaluates to the constant value $\gamma T^\dagger u_l$.

Proof.

By Proposition 4.4, $H(x)$ is not empty for each $x \in \mathbb{R}^N$, and each fixed point of $f_{x,T}$ in Equation (4.6) provides us with an element $y = t - x \in H(x)$. We show that $f_{x,T}$ possesses only one fixed point, if

$$\min_{j \in \{1, \dots, L\}} |(Tx)_j| > \gamma(\|TT^\dagger\|_\infty + 1).$$

Assume that $f_{x,T}$, with $x \in \mathcal{U}_{T,\gamma}$, possesses two fixed points t_1 and t_2 . Thus

$$t_1 = x + (\text{id} - T^\dagger S_\gamma T)t_1, \quad t_2 = x + (\text{id} - T^\dagger S_\gamma T)t_2,$$

implies

$$x = T^\dagger S_\gamma T t_1 = T^\dagger S_\gamma T t_2.$$

4. Nested Proximity Operators

Consequently,

$$T^\dagger S_\gamma T t_1 - T^\dagger S_\gamma T t_2 = 0. \quad (4.12)$$

Now, similarly as in the previous proof, we can show that $T^\dagger(S_\gamma T t_1 - S_\gamma T t_2) = 0$ implies $S_\gamma T t_1 - S_\gamma T t_2 = 0$. This time we consider the index sets

$$\begin{aligned} I_1 &= \{1 \leq j \leq L : [S_\gamma T t_1 - S_\gamma T t_2]_j > 0\}, \\ I_2 &= \{1 \leq j \leq L : [S_\gamma T t_1 - S_\gamma T t_2]_j = 0\}, \\ I_3 &= \{1 \leq j \leq L : [S_\gamma T t_1 - S_\gamma T t_2]_j < 0\}. \end{aligned}$$

We observe that $j \in I_1$ yields $[T t_1 - T t_2]_j > 0$ and similarly $j \in I_3$ yields $[T t_1 - T t_2]_j < 0$. We want to show that $I_1 \cup I_3 = \emptyset$.

The relation of Equation (4.12) implies

$$0 = \sum_{j \in I_1} [S_\gamma T t_1 - S_\gamma T t_2]_j T^\dagger e_j + \sum_{j \in I_3} [S_\gamma T t_1 - S_\gamma T t_2]_j T^\dagger e_j.$$

Now, suppose contrarily that w.l.o.g. $I_1 \neq \emptyset$, then for $j_1 \in I_1$ we find $\tilde{v}_{j_1} := T^\dagger e_{j_1}$ similarly as in Equation (4.11) of the form

$$\tilde{v}_{j_1} = \sum_{\substack{j \in I_1 \\ j \neq j_1}} \frac{-[S_\gamma T t_1 - S_\gamma T t_2]_j}{[S_\gamma T t_1 - S_\gamma T t_2]_{j_1}} \tilde{v}_j + \sum_{j \in I_3} \frac{-[S_\gamma T t_1 - S_\gamma T t_2]_j}{[S_\gamma T t_1 - S_\gamma T t_2]_{j_1}} \tilde{v}_j.$$

A closer look at the coefficients shows that

$$\frac{-[S_\gamma T t_1 - S_\gamma T t_2]_j}{[S_\gamma T t_1 - S_\gamma T t_2]_{j_1}} < 0$$

for $j \in I_1 \setminus \{j_1\}$ and

$$\frac{-[S_\gamma T t_1 - S_\gamma T t_2]_j}{[S_\gamma T t_1 - S_\gamma T t_2]_{j_1}} > 0$$

for $j \in I_3$.

Hence, we find with $T^* T \tilde{v}_j = T^* e_j$,

$$\begin{aligned} & [T(t_1 - t_2)]_{j_1} \\ &= \langle T(t_1 - t_2), e_{j_1} \rangle = \langle t_1 - t_2, T^* T \tilde{v}_{j_1} \rangle \\ &= \left\langle t_1 - t_2, (T^* T) \sum_{\substack{j \in I_1 \\ j \neq j_1}} \frac{-[S_\gamma T t_1 - S_\gamma T t_2]_j}{[S_\gamma T t_1 - S_\gamma T t_2]_{j_1}} \tilde{v}_j + \sum_{j \in I_3} \frac{-[S_\gamma T t_1 - S_\gamma T t_2]_j}{[S_\gamma T t_1 - S_\gamma T t_2]_{j_1}} \tilde{v}_j \right\rangle \end{aligned}$$

$$\begin{aligned}
 &= \sum_{\substack{j \in I_1 \\ j \neq j_1}} \frac{-[S_\gamma T t_1 - S_\gamma T t_2]_j}{[S_\gamma T t_1 - S_\gamma T t_2]_{j_1}} \langle t_1 - t_2, (T^* T) \tilde{v}_j \rangle + \sum_{j \in I_3} \frac{-[S_\gamma T t_1 - S_\gamma T t_2]_j}{[S_\gamma T t_1 - S_\gamma T t_2]_{j_1}} \langle t_1 - t_2, (T^* T) \tilde{v}_j \rangle \\
 &= \sum_{\substack{j \in I_1 \\ j \neq j_1}} \underbrace{\frac{-[S_\gamma T t_1 - S_\gamma T t_2]_j}{[S_\gamma T t_1 - S_\gamma T t_2]_{j_1}}}_{<0} \underbrace{\langle T(t_1 - t_2), e_j \rangle}_{>0} + \sum_{j \in I_3} \underbrace{\frac{-[S_\gamma T t_1 - S_\gamma T t_2]_j}{[S_\gamma T t_1 - S_\gamma T t_2]_{j_1}}}_{>0} \underbrace{\langle T(t_1 - t_2), e_j \rangle}_{<0} \\
 &< 0,
 \end{aligned}$$

which contradicts the above assumption that $j_1 \in I_1$. Therefore, $I_1 \cup I_3 = \emptyset$, and for each component it follows either that $[T t_1]_j = [T t_2]_j$ or $|[T t_1]_j| \leq \gamma$ and $|[T t_2]_j| \leq \gamma$. Further, observe that

$$\|T t_1 - T x\|_\infty = \|T t_1 - T T^\dagger S_\gamma T t_1\|_\infty = \|T T^\dagger (\text{id} - S_\gamma) T t_1\|_\infty \leq \|T T^\dagger\|_\infty \gamma,$$

and similarly for the fixed point t_2 . Therefore, $|[T x]_j| > \gamma(\|T^\dagger T\|_\infty + 1)$ for all $j \in \{1, \dots, L\}$ implies that

$$|[T t_1]_j| > |[T x]_j| - |[T t_1]_j - [T x]_j| > \gamma$$

for all $j \in \{1, \dots, L\}$. Thus, all components of $T t_1$ and $T t_2$ coincide, and we conclude $t_1 = t_2$.

Since $H(x)$ contains a single element for $x \in S_l$, it follows with $\text{sgn}(T x) = u_l$ that

$$x = T^\dagger S_\gamma T t_1 = T^\dagger (T t_1 - \gamma u_l) = t_1 - \gamma T^\dagger u_l$$

and thus $H(x) = t_1 - x = \gamma T^\dagger u_l$. qed

Our considerations show that H does in general not coincide with the subdifferential $\partial \|T \cdot\|_1 = T^* \text{sgn}(T \cdot)$. But similarly to H , we observe that $y \in \partial \|T \cdot\|_1(0)$ if and only if $\|T x\|_\infty \leq \gamma$, and $\partial \|T \cdot\|_1(x)$ is single-valued if

$$\min_{j \in \{1, \dots, L\}} |[T x]_j| > \gamma.$$

Frame Soft Thresholding is a Proximity Operator

Throughout this section, we again assume that $T \in \mathbb{R}^{L \times N}$ with $L > N$ has full rank N , $\gamma > 0$ and let $S_\gamma = \text{prox}_{\gamma \|\cdot\|_1}$ as before. We will now show that the set-valued function H in Equation (4.5) is the subdifferential of a plc function. To that end we introduce a weighted inner product

$$\langle x, y \rangle_T := \langle T x, T y \rangle \tag{4.13}$$

based on the canonical inner product $\langle \cdot, \cdot \rangle$ on the right hand side.

To prove that H is the subdifferential of a plc function we employ Rockafellar's Monotonicity

4. Nested Proximity Operators

Theorem, Theorem 2.8. That means, we have to show that H in Equation (4.5) is indeed maximally cyclically monotone. To this end, we need some preliminary lemmas, in the context of which we will use the notions of monotonicity, cyclical monotonicity, maximal monotonicity etc. as introduced in Chapter 2.

4.7 Lemma.

For $x \in \mathbb{R}^N$ and $y \in H(x)$ we have

$$Ty + (S_\gamma - \text{id})T(x + y) \in \mathcal{N}(T^\dagger) = \mathcal{N}(T^*),$$

and

$$y = T^\dagger(\text{id} - S_\gamma)T(x + y).$$

Proof.

Recall that $T^\dagger = (T^*T)^{-1}T^* \in \mathbb{R}^{N \times L}$ and $T^\dagger T = \text{id}$. Thus we have

$$\begin{aligned} y \in H(x) &\iff x = T^\dagger S_\gamma T(x + y) \\ &\iff T^\dagger T x = T^\dagger S_\gamma T(x + y) \\ &\iff \exists u \in \mathcal{N}(T^\dagger) : u + T x = S_\gamma T(x + y) \\ &\iff \exists u \in \mathcal{N}(T^\dagger) : u + T(x + y) - T y = S_\gamma T(x + y) \\ &\iff \exists u \in \mathcal{N}(T^\dagger) : u + T(x + y) - S_\gamma T(x + y) = T y \\ &\iff \exists u \in \mathcal{N}(T^\dagger) : u + (\text{id} - S_\gamma)T(x + y) = T y. \end{aligned} \quad (4.14)$$

In other words, $Ty + (S_\gamma - \text{id})T(x + y) \in \mathcal{N}(T^\dagger)$. Further, multiplying Equation (4.14) with T^\dagger , it follows that

$$T^\dagger T y = y = T^\dagger(\text{id} - S_\gamma)T(x + y),$$

which finishes this proof. qed

Next we consider a rather technical lemma. While not being interesting in its own right, it will become a core building block in proving cyclical monotonicity of H in Theorem 4.9.

4.8 Lemma.

Let $x_1, x_2 \in \mathbb{R}^N$ and $y_1 \in H(x_1), y_2 \in H(x_2)$. Further, let

$$z_1 := (\text{id} - S_\gamma)T(x_1 + y_1), \quad z_2 := (\text{id} - S_\gamma)T(x_2 + y_2).$$

Then

$$\langle S_\gamma T(x_1 + y_1), z_2 - z_1 \rangle \leq 0.$$

Proof.

From the definition of S_γ as the proximity operator of the ℓ^1 -norm in Example 2.12, it follows for $x \in \mathbb{R}$

$$(1 - S_\gamma)x = \begin{cases} x - (x - \gamma) = \gamma & x > \gamma, \\ x - (x + \gamma) = -\gamma & x < -\gamma, \\ x & |x| \leq \gamma, \end{cases}$$

and therefore for all $x \in \mathbb{R}^N$,

$$|[(\text{id} - S_\gamma)Tx]_j| \leq \gamma, \quad \text{for any } j \in \{1, \dots, L\}. \quad (4.15)$$

Let us consider the case $[T(x_1 + y_1)]_j \leq -\gamma$. Then $[(\text{id} - S_\gamma)T(x_1 + y_1)]_j = -\gamma$ and Equation (4.15) imply $[(\text{id} - S_\gamma)T(x_2 + y_2)]_j \geq -\gamma$, such that

$$\begin{aligned} [z_2 - z_1]_j &= [(\text{id} - S_\gamma)T(x_2 + y_2) - (\text{id} - S_\gamma)T(x_1 + y_1)]_j \\ &\geq -\gamma - [(\text{id} - S_\gamma)T(x_1 + y_1)]_j \\ &= -\gamma + \gamma = 0. \end{aligned}$$

In the case $[T(x_1 + y_1)]_j \geq \gamma$ we similarly conclude from $[(\text{id} - S_\gamma)T(x_1 + y_1)]_j = \gamma$ and Equation (4.15) that

$$\begin{aligned} [z_2 - z_1]_j &= [(\text{id} - S_\gamma)T(x_2 + y_2) - (\text{id} - S_\gamma)T(x_1 + y_1)]_j \\ &\leq \gamma - [(\text{id} - S_\gamma)T(x_1 + y_1)]_j \\ &= \gamma - \gamma = 0. \end{aligned}$$

Thus, in both cases we have $-\text{sgn}[z_2 - z_1]_j = \text{sgn}[S_\gamma T(x_1 + y_1)]_j$. Finally, for $|[T(x_1 + y_1)]_j| < \gamma$ it follows that $[S_\gamma T(x_1 + y_1)]_j = 0$. We therefore conclude

$$\langle S_\gamma T(x_1 + y_1), z_2 - z_1 \rangle = \sum_{j=1}^N [S_\gamma T(x_1 + y_1)]_j [z_2 - z_1]_j \leq 0.$$

and the assertion follows. qed

Now we can show cyclical monotonicity of H .

4.9 Theorem.

The map H in Equation (4.5) is cyclically monotone with respect to the norm induced by $\langle \cdot, \cdot \rangle_T$.

4. Nested Proximity Operators

Proof.

Let $m \in \mathbb{N}$ and $m \geq 2$. Further, for $i \in \{1, \dots, m\}$ let $y_i \in H(x_i)$. As before we use a thin SVD of T , cf. Definition 1.14, i.e. $T = PDQ^T$ with $Q \in \mathbb{R}^{N \times N}$ orthogonal, $D \in \mathbb{R}^{N \times N}$ containing the N positive singular values of T , and $P \in \mathbb{R}^{L \times N}$ with $P^*P = \text{id}$. Further, let

$$\begin{aligned}\tilde{x}_i &:= DQ^T x_i, \\ \tilde{y}_i &:= DQ^T y_i, \\ z_i &:= (\text{id} - S_\gamma)T(x_i + y_i) = (\text{id} - S_\gamma)P(\tilde{x}_i + \tilde{y}_i), \\ u_i &:= Ty_i - z_i = P\tilde{y}_i - z_i.\end{aligned}\tag{4.16}$$

For simplicity we use the convention $x_{m+1} := x_1$ as well as $y_{m+1} := y_1$, and extend that similarly for $\tilde{x}_{m+1}, \tilde{y}_{m+1}, z_{m+1}$, and u_{m+1} . Then, by Lemma 4.7, $u_i \in \mathcal{N}(T^\dagger)$ and thus also $u_i \in \mathcal{N}(P^*)$, since $T^\dagger = QD^{-1}P^*$, where $QD^{-1} \in \mathbb{R}^{N \times N}$ is invertible. Further, with Equation (4.16) we can write

$$\tilde{y}_i = P^*P\tilde{y}_i = P^*(u_i + z_i) = P^*z_i.$$

We will show that

$$A := \sum_{i=1}^m \langle \tilde{x}_{i+1} - \tilde{x}_i, \tilde{y}_i \rangle \leq 0.$$

First, we observe that for all $i \in \{1, \dots, m\}$,

$$P\tilde{x}_i + u_i = P\tilde{x}_i + P\tilde{y}_i - z_i = P(\tilde{x}_i + \tilde{y}_i) - (\text{id} - S_\gamma)P(\tilde{x}_i + \tilde{y}_i) = S_\gamma P(\tilde{x}_i + \tilde{y}_i).\tag{4.17}$$

Using Equation (4.17), it follows

$$\begin{aligned}A &= \sum_{i=1}^m \langle \tilde{x}_{i+1} - \tilde{x}_i, P^*z_i \rangle = \sum_{i=1}^m \langle P(\tilde{x}_{i+1} - \tilde{x}_i), z_i \rangle \\ &= \sum_{i=1}^m \langle (P\tilde{x}_{i+1} + u_{i+1}) - (P\tilde{x}_i + u_i) - u_{i+1} + u_i, z_i \rangle \\ &= \sum_{i=1}^m \langle S_\gamma P(\tilde{x}_{i+1} + \tilde{y}_{i+1}), z_i \rangle - \sum_{i=1}^m \langle S_\gamma P(\tilde{x}_i + \tilde{y}_i), z_i \rangle + \sum_{i=1}^m \langle -u_{i+1} + u_i, z_i \rangle \\ &= \sum_{i=1}^m \langle S_\gamma P(\tilde{x}_{i+1} + \tilde{y}_{i+1}), z_i - z_{i+1} \rangle + \sum_{i=1}^m \langle -u_{i+1} + u_i, z_i \rangle.\end{aligned}$$

By Lemma 4.8, the first sum is not positive. Thus,

$$A \leq \sum_{i=1}^m \langle -u_{i+1} + u_i, z_i \rangle = \sum_{i=1}^m \langle -u_{i+1} + u_i, P\tilde{y}_i - u_i \rangle$$

$$= \sum_{i=1}^m \langle P^*(-u_{i+1} + u_i), \tilde{y}_i \rangle - \sum_{i=1}^m \langle u_i, u_i \rangle + \sum_{i=1}^m \langle u_{i+1}, u_i \rangle.$$

The first sum vanishes, since u_i and u_{i+1} are in $\mathcal{N}(P^*)$. Thus,

$$A \leq - \sum_{i=1}^m \|u_i\|_2^2 + \sum_{i=1}^m \langle u_{i+1}, u_i \rangle \leq - \sum_{i=1}^m \|u_i\|_2^2 + \frac{1}{2} \sum_{i=1}^m (\|u_i\|_2^2 + \|u_{i+1}\|_2^2) = 0.$$

Then we simply observe that, as P is orthogonal,

$$A = \sum_{i=1}^m \langle \tilde{x}_{i+1} - \tilde{x}_i, \tilde{y}_i \rangle = \sum_{i=1}^m \langle DQ^T(x_{i+1} - x_i), DQ^T y_i \rangle = \sum_{i=1}^m \langle x_{i+1} - x_i, y_i \rangle_T \leq 0.$$

Thus, the assertion of the theorem holds. qed

Now we can conclude the main theorem of this section.

4.10 Theorem.

Let $T \in \mathbb{R}^{L \times N}$ with $L \geq N$ have full rank N and let \mathbb{R}^N be equipped with the inner product $\langle \cdot, \cdot \rangle_T$ as defined in Equation (4.13). Then the operator $T^\dagger S_\gamma T$ is a proximity operator of a plc function Φ .

Proof.

First, by Equation (4.4), for any x and p we have $p - x \in \partial\Phi(x)$ if and only if $x = \text{prox}_\Phi(p)$. So, by Theorem 2.8, we need to prove that H in Equation (4.5) is maximally cyclically monotone.

As shown in Theorem 4.9, we already have that H is cyclically monotone. Further, by Proposition 4.4, H is bounded, i.e., for all $x \in \mathbb{R}^N$ we have that $y \in H(x)$ implies $\|y\|_\infty \leq \gamma \|T^*\|_\infty$. Therefore, we observe that the range of the operator $\text{id} + H$ is \mathbb{R}^N . By Minty's Theorem, Theorem 2.5, it follows that H is also maximally monotone. The assertion now follows from Theorem 2.8. qed

As a corollary, we can easily prove firm nonexpansiveness of the composition $T^\dagger S_\gamma T$. To this end, recall the notion of (firm) nonexpansiveness from Definitions 2.6 and 2.10.

4.11 Corollary.

Let $T \in \mathbb{R}^{L \times N}$ with $L \geq N$ have full rank N . Let $\|\cdot\|_T$ be the norm corresponding to the inner product in Equation (4.13). Then the operators $T^\dagger S_\gamma T$ and $\text{id} - T^\dagger S_\gamma T$ are firmly non-expansive, i.e., for all $x, y \in \mathbb{R}^N$ we have

$$\|T^\dagger S_\gamma T x - T^\dagger S_\gamma T y\|_T^2 + \|(\text{id} - T^\dagger S_\gamma T)x - (\text{id} - T^\dagger S_\gamma T)y\|_T^2 \leq \|x - y\|_T^2.$$

4. Nested Proximity Operators

Proof.

Observe that as $\mathcal{R}(S_\gamma T) \subseteq \mathcal{R}(T)$, it is easy to see that

$$\|T^\dagger S_\gamma T x - T^\dagger S_\gamma T y\|_T^2 = \|T T^\dagger (S_\gamma T x - S_\gamma T y)\|_2^2 = \|S_\gamma T x - S_\gamma T y\|_2^2.$$

Also,

$$\begin{aligned} & \|(\text{id} - T^\dagger S_\gamma T)x - (\text{id} - T^\dagger S_\gamma T)y\|_T^2 \\ &= \|T(\text{id} - T^\dagger S_\gamma T)x - T(\text{id} - T^\dagger S_\gamma T)y\|_2^2 \\ &= \|Tx - Ty - S_\gamma T x + S_\gamma T y\|_2^2 \\ &= \|x - y\|_T^2 + \|S_\gamma T x - S_\gamma T y\|_2^2 - 2\langle T(x - y), S_\gamma T x - S_\gamma T y \rangle \end{aligned}$$

To prove the assertion of the corollary we hence have to show that

$$\|S_\gamma T x - S_\gamma T y\|_2^2 - \langle T(x - y), S_\gamma T x - S_\gamma T y \rangle \leq 0.$$

But this assertion is obviously true since for each component $k \in \{1, \dots, L\}$, we have either $(S_\gamma T x - S_\gamma T y)_k = 0$ or $\text{sgn}(S_\gamma T x - S_\gamma T y)_k = \text{sgn}(T(x - y))_k$ and $|(S_\gamma T x - S_\gamma T y)_k| \leq |(T(x - y))_k|$. qed

4.2. Proximity Operators Nested Into Injective Operators

Note: The following portion of this thesis is predominantly taken from [HNP⁺19] of which I am a coauthor. Until the end of Section 4.2 the content is in large portions identical to the one in [HNP⁺19]. I have occasionally adapted the order, notation or, where adequate, wording and added further thoughts if considered reasonable. I will also at times leave out parts that have been referenced earlier within this thesis or are irrelevant with regard to the concern of this work. In these cases I have put references to the corresponding location whenever I believe this to be beneficial.

Above, we considered the question whether the frame soft shrinkage operator can itself be seen as a proximity operator, see Section 4.1. In this section, we prove that for any injective operator $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ with closed range and any proximity operator $\text{Prox}: \mathcal{K} \rightarrow \mathcal{K}$, the new operator $T^\dagger \text{Prox} T: \mathcal{H} \rightarrow \mathcal{H}$ is also a proximity operator on the linear space \mathcal{H} , but equipped with another inner product. This includes the above mentioned finite dimensional setting as a special case. In contrast to Section 4.1, we directly approach the question using a classical result of Moreau [Mor65].

There are several fields where our results may be of interest. Some different norms in the definition of the proximity operator were successfully used in variable metric algorithms, see [CPR13]. Recently, it was shown that many activation functions that appear in neural networks are indeed proximity functions [CP20]. Here, we concatenate proximity operators with a linear operator and

stay within the set of proximity operators in different Hilbert spaces. Finally, our findings may be of interest in connection with so-called Plug-and-Play algorithms [CWE17, SVW⁺16, TBDF18].

Proximity Operators and Linear Operators

We use notation and terminology as before. That is especially the notion of the proximity operator from Section 2.2, frame theory from Section 1.2 and the idea of nesting proximity operators as in Section 4.1.

Thanks to Corollary 2.18, we know that $\text{Prox}: \mathcal{H} \rightarrow \mathcal{H}$ is a proximity operator if and only if it is nonexpansive and the gradient of a convex, differentiable function $\Phi: \mathcal{H} \rightarrow \mathbb{R}$.

Let now \mathcal{H} and \mathcal{K} be real Hilbert spaces with inner products $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and $\langle \cdot, \cdot \rangle_{\mathcal{K}}$ and corresponding norms $\|\cdot\|_{\mathcal{H}}$ and $\|\cdot\|_{\mathcal{K}}$. By $\mathcal{B}(\mathcal{H}, \mathcal{K})$ we denote the space of bounded, linear operators from \mathcal{H} to \mathcal{K} with domain \mathcal{H} . In this section, we show that for any injective operator $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ with closed range $\mathcal{R}(T)$ and any proximity operator $\text{Prox}: \mathcal{K} \rightarrow \mathcal{K}$, the operator $T^\dagger \text{Prox} T: \mathcal{H} \rightarrow \mathcal{H}$ is itself a proximity operator on the linear space \mathcal{H} equipped with a suitable (equivalent) norm $\|\cdot\|_{\mathcal{H}_T}$.

For any injective $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ with closed range, recall its pseudoinverse T^\dagger from Definition 1.1. Such a map T also gives rise to an inner product in \mathcal{H} via $\langle x, y \rangle_{\mathcal{H}_T} = \langle Tx, Ty \rangle_{\mathcal{K}}$ and corresponding norm $\|x\|_{\mathcal{H}_T} = \|Tx\|_{\mathcal{K}}$ as before in Section 4.1, cf. Equation (4.13). In general, this defines a pre-Hilbert space. Since T has additionally closed range, the norms $\|\cdot\|_{\mathcal{H}}$ and $\|\cdot\|_{\mathcal{H}_T}$ are equivalent on \mathcal{H} due to

$$\|T\|_{\mathcal{B}(\mathcal{H}, \mathcal{K})}^{-1} \|x\|_{\mathcal{H}_T} \leq \|x\|_{\mathcal{H}} = \|T^\dagger Tx\|_{\mathcal{H}} \leq \|T^\dagger\|_{\mathcal{B}(\mathcal{K}, \mathcal{H})} \|x\|_{\mathcal{H}_T}$$

for all $x \in \mathcal{H}$. The norm equivalence also ensures the completeness of \mathcal{H} equipped with the new norm. To emphasize that we consider the linear space \mathcal{H} with this norm we write \mathcal{H}_T .

For a Fréchet differentiable function, cf. Definition 2.9, the gradient depends on the chosen norm through

$$\nabla_{\mathcal{H}_T} \Phi(x) = (T^* T)^{-1} \nabla_{\mathcal{H}} \Phi(x).$$

Now the desired result follows from the next theorem. Here, we refer to (firm) nonexpansiveness and convexity of functions as in Definitions 2.1, 2.6 and 2.10. We also use the pseudoinverse as introduced in Definition 1.1.

4.12 Theorem.

Let $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ be injective with closed range and $\text{Prox}: \mathcal{K} \rightarrow \mathcal{K}$ a proximity operator on \mathcal{K} . Then, the operator $T^\dagger \text{Prox} T: \mathcal{H}_T \rightarrow \mathcal{H}_T$ is a proximity operator.

4. Nested Proximity Operators

Proof.

In view of Theorems 2.17 and 2.19, it suffices to show that $T^\dagger \text{Prox} T$ is nonexpansive and that there exists a convex function $\Psi: \mathcal{H}_T \rightarrow \mathbb{R}$ with $T^\dagger \text{Prox} T = \nabla_{\mathcal{H}_T} \Psi$.

First, we show that $T^\dagger \text{Prox} T$ is firmly nonexpansive, and thus nonexpansive. To that end let $x, y \in \mathcal{H}$. Using that $TT^\dagger = P_{\mathcal{R}(T)}$, it follows

$$\begin{aligned} \|T^\dagger \text{Prox} Tx - T^\dagger \text{Prox} Ty\|_{\mathcal{H}_T}^2 &= \|TT^\dagger (\text{Prox} Tx - \text{Prox} Ty)\|_{\mathcal{K}}^2 \\ &\leq \|\text{Prox} Tx - \text{Prox} Ty\|_{\mathcal{K}}^2. \end{aligned} \quad (4.18)$$

Now, we obtain

$$\begin{aligned} \langle T^\dagger \text{Prox} Tx - T^\dagger \text{Prox} Ty, x - y \rangle_{\mathcal{H}_T} &= \langle TT^\dagger (\text{Prox} Tx - \text{Prox} Ty), Tx - Ty \rangle_{\mathcal{K}} \\ &= \langle T^* TT^\dagger (\text{Prox} Tx - \text{Prox} Ty), x - y \rangle_{\mathcal{H}} \\ &= \langle \text{Prox} Tx - \text{Prox} Ty, Tx - Ty \rangle_{\mathcal{K}}, \end{aligned}$$

and since Prox is firmly nonexpansive with respect to $\|\cdot\|_{\mathcal{K}}$ and by Equation (4.18) further

$$\begin{aligned} \langle T^\dagger \text{Prox} Tx - T^\dagger \text{Prox} Ty, x - y \rangle_{\mathcal{H}_T} &\geq \|\text{Prox} Tx - \text{Prox} Ty\|_{\mathcal{K}}^2 \\ &\geq \|T^\dagger \text{Prox} Tx - T^\dagger \text{Prox} Ty\|_{\mathcal{H}_T}^2. \end{aligned}$$

Thus, $T^\dagger \text{Prox} T$ is firmly nonexpansive.

Second, it remains to prove that there exists a convex function $\Psi: \mathcal{H}_T \rightarrow \mathbb{R}$ such that $\nabla_{\mathcal{H}_T} \Psi = T^\dagger \text{Prox} T$. Since Prox is a proximity operator, there exists $\Phi: \mathcal{H} \rightarrow \mathbb{R}$ with $\text{Prox} = \nabla_{\mathcal{K}} \Phi$. Then, a natural candidate is given by $\Psi = \Phi T$. Using the definition of the gradient and the chain rule, it holds for all $x, h \in \mathcal{H}$ that

$$\langle \nabla_{\mathcal{H}} \Psi(x), h \rangle_{\mathcal{H}} = D\Psi(x)h = D\Phi(Tx)Th = \langle \nabla_{\mathcal{K}} \Phi(Tx), Th \rangle_{\mathcal{K}} = \langle T^* \text{Prox} Tx, h \rangle_{\mathcal{H}}.$$

Hence, we conclude

$$\nabla_{\mathcal{H}_T} \Psi = (T^* T)^{-1} \nabla_{\mathcal{H}} \Psi(x) = (T^* T)^{-1} T^* \text{Prox} Tx = T^\dagger \text{Prox} Tx.$$

Finally, Ψ is convex since it is the concatenation of a convex function with a linear function. **qed**

For $\text{Prox} := \text{prox}_g$ with $g \in \Gamma_0(\mathcal{H})$, we are able to explicitly compute $f \in \Gamma_0(\mathcal{H})$ such that $T^\dagger \text{Prox} T = \text{prox}_f$ on \mathcal{H}_T .

We now present a formula to solve the expression $\text{prox}_f = T^\dagger \text{prox}_g T$ for f in terms of an infimal convolution, cf. Definition 2.2 and Proposition 2.3. Note however that despite the looks this is not necessarily a closed expression as the infimal convolution oftentimes is not computable explicitly.

4.13 Theorem.

Let $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ an injective operator with closed range and $\text{Prox} := \text{prox}_g$ for some $g \in \Gamma_0(\mathcal{K})$. Then $T^\dagger \text{prox}_g T: \mathcal{H}_T \rightarrow \mathcal{H}_T$ is the proximity operator on \mathcal{H}_T of $f \in \Gamma_0(\mathcal{H})$ given by

$$f(x) := g \square \left(\frac{1}{2} \|\cdot\|_{\mathcal{K}}^2 + \iota_{\mathcal{N}(T^*)} \right) (Tx). \quad (4.19)$$

Here $\mathcal{N}(T^*)$ denotes the kernel of T^* and $\iota_{\mathcal{N}(T^*)}$ its indicator function, cf. [List of Symbols](#).

For bijective $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ this expression simplifies to $f = gT$.

Proof.

Since $\mathcal{R}(T)$ is closed, we have the orthogonal decomposition $\mathcal{H} = \mathcal{R}(T) \oplus \mathcal{N}(T^*)$. Then, we get

$$\begin{aligned} T^\dagger \text{prox}_g(Tx) &= T^\dagger \arg \min_{z \in \mathcal{K}} \left\{ \frac{1}{2} \|z - Tx\|_{\mathcal{K}}^2 + g(z) \right\} \\ &= T^\dagger P_{\mathcal{R}(T)} \arg \min_{\substack{z = z_1 + z_2 \\ z_1 \in \mathcal{R}(T), z_2 \in \mathcal{N}(T^*)}} \left\{ \frac{1}{2} \|z_1 + z_2 - Tx\|_{\mathcal{K}}^2 + g(z_1 + z_2) \right\} \\ &= T^\dagger \arg \min_{z_1 \in \mathcal{R}(T)} \inf_{z_2 \in \mathcal{N}(T^*)} \left\{ \frac{1}{2} \|z_1 - Tx\|_{\mathcal{K}}^2 + \frac{1}{2} \|z_2\|_{\mathcal{K}}^2 + g(z_1 + z_2) \right\} \\ &= T^\dagger \arg \min_{z_1 \in \mathcal{R}(T)} \left\{ \frac{1}{2} \|z_1 - Tx\|_{\mathcal{K}}^2 + \inf_{z_2 \in \mathcal{N}(T^*)} \left\{ \frac{1}{2} \|z_2\|_{\mathcal{K}}^2 + g(z_1 + z_2) \right\} \right\} \\ &= T^\dagger T \arg \min_{y \in \mathcal{H}} \left\{ \frac{1}{2} \|Ty - Tx\|_{\mathcal{K}}^2 + \inf_{z_2 \in \mathcal{N}(T^*)} \left\{ \frac{1}{2} \|z_2\|_{\mathcal{K}}^2 + g(Ty + z_2) \right\} \right\} \\ &= \arg \min_{y \in \mathcal{H}} \left\{ \frac{1}{2} \|y - x\|_{\mathcal{H}_T}^2 + \inf_{z_2 \in \mathcal{N}(T^*)} \left\{ \frac{1}{2} \|z_2\|_{\mathcal{K}}^2 + g(Ty + z_2) \right\} \right\} \quad (4.20) \\ &= \arg \min_{y \in \mathcal{H}} \left\{ \frac{1}{2} \|y - x\|_{\mathcal{H}_T}^2 + g \square \left(\frac{1}{2} \|\cdot\|_{\mathcal{K}}^2 + \iota_{\mathcal{N}(T^*)} \right) (Ty) \right\}. \end{aligned}$$

Hence, we conclude that $T^\dagger \text{prox}_g T$ is the proximity operator on \mathcal{H}_T of f in Equation (4.19). In particular, we conclude for bijective T by Equation (4.20) that

$$T^\dagger \text{prox}_g(Tx) = \arg \min_{y \in \mathcal{H}} \left\{ \frac{1}{2} \|x - y\|_{\mathcal{H}_T}^2 + g(Ty) \right\}. \quad \text{qed}$$

Also note that in general f is a weaker regularizer than g , i.e., $f \leq g$. This is necessary since for the latter we would get, using the same reasoning as in Equation (4.20)

$$\arg \min_{y \in \mathcal{H}} \left\{ \frac{1}{2} \|x - y\|_T^2 + g(Ty) \right\} = T^\dagger \arg \min_{z \in \mathcal{K}} \left\{ \frac{1}{2} \|z - Tx\|_{\mathcal{K}}^2 + g(z) + \iota_{\mathcal{R}(T)}(z) \right\} \neq T^\dagger \text{prox}_g(Tx).$$

Frame Soft Shrinkage

Let us now investigate frame soft shrinkage as a special proximity operator and thereby finding a different characterization as compared to Section 4.1. Hence, let $\mathcal{K} = \ell^2$ be the Hilbert space of quadratic summable sequences with the usual ℓ^2 -norm and assume further that \mathcal{H} is separable.

Recalling frames from Section 1.2, note that T^\dagger is indeed the synthesis operator for the canonical dual frame of $\{f_k\}_{k \in \mathbb{N}}$. The relation between linear, bounded, injective operators of closed range and frame analysis operators reads as follows:

4.14 Proposition.

An operator $T \in \mathcal{B}(\mathcal{H}, \ell^2)$ is injective and has closed range if and only if it is the analysis operator of some frame of \mathcal{H} .

Proof.

If T is the analysis operator for a frame $\{x_k\}_{k \in \mathbb{N}}$, then T is bounded, injective and has closed range, see [Chr16]. Conversely, assume that $T \in \mathcal{B}(\mathcal{H}, \ell^2)$ is injective and that $\mathcal{R}(T)$ is closed. By the closed range theorem, cf. [MS79], it holds $\mathcal{R}(T^*) = H$. Let $\{\delta_k\}_{k \in \mathbb{N}}$ be the canonical basis of ℓ^2 and set $\{x_k\}_{k \in \mathbb{N}} := \{T^* \delta_k\}_{k \in \mathbb{N}}$. Since $\sum_{k \in \mathbb{N}} |\langle x, x_k \rangle_{\mathcal{H}}|^2 = \|Tx\|_2^2$, we conclude that $\{x_k\}_{k \in \mathbb{N}}$ is a frame of \mathcal{H} and T is the corresponding analysis operator. qed

The soft shrinkage operator S_λ on ℓ^2 (applied componentwise) is the proximity operator corresponding to the function $g := \lambda \|\cdot\|_1$ for some $\lambda > 0$, as shown in Example 2.12. Then we obtain as immediate consequence of Theorem 4.13 the following corollary.

4.15 Corollary.

Let $T: \mathcal{H} \rightarrow \ell^2$ be an analysis frame operator and $\text{Prox}: \ell^2 \rightarrow \ell^2$ an arbitrary proximity operator. Then $T^\dagger \text{Prox} T$ is itself a proximity operator on \mathcal{H} equipped with the norm $\|\cdot\|_{\mathcal{H}_T}$. In particular, it holds for $\text{Prox} := S_\lambda$, $\lambda > 0$ that

$$T^\dagger S_\lambda T(x) = \arg \min_{y \in \mathcal{H}} \{ \|x - y\|_{\mathcal{H}_T}^2 + f(y) \}$$

for

$$f(y) := \left[\lambda \|\cdot\|_1 \square \left(\frac{1}{2} \|\cdot\|_2^2 + \iota_{\mathcal{N}(T^*)} \right) \right] (Ty).$$

Finally, let us have a look at the finite dimensional setting with $\mathcal{H} := \mathbb{R}^d$, $\mathcal{K} := \mathbb{R}^n$, $n \geq d$. Then we have for any $T \in \mathbb{R}^{n \times d}$ with full rank d and an arbitrary proximity operator on \mathbb{R}^n that

$$T^\dagger \text{Prox} T(x) = \arg \min_{y \in \mathbb{R}^d} \left\{ \frac{1}{2} \|x - y\|_T^2 + f(y) \right\}$$

where

$$f(y) := \lambda \left[\|\cdot\|_1 \square \left(\frac{1}{2} \|\cdot\|_2^2 + \iota_{\mathcal{N}(T^\top)} \right) \right] (Ty).$$

5. Sparse Rank-One Matrix Recovery

In this chapter we are considering a different approach for extracting desired features. Set in a compressed sensing environment, we suppose a bilinear measurement setting, cf. Chapter 3. That is, we want to recover two sparse signals u, v from the corresponding function value $b(u, v)$ for some bilinear map b . With lifting, cf. Section 3.1, this becomes a sparse rank-one matrix recovery problem. While it is well known that recovering a simultaneously sparse and rank-one matrix is difficult, [AKR18], we propose an adaptive scheme, based on the *Sparse Power Factorization* algorithm by Lee et al., [LWB18].

Note: The following portion of this thesis is predominantly taken from [GKS19] of which I am a coauthor. Until the end of Chapter 5 the content is in large portions identical to the one in [GKS19]. I have occasionally adapted the order, notation or, where adequate, wording and added further thoughts if considered reasonable. I will also at times leave out parts that have been referenced earlier within this thesis or are irrelevant with regard to the concern of this work. In these cases I have put references to the corresponding location whenever I believe this to be beneficial.

Many measurement operations in signal and image processing as well as in communication follow a bilinear model. Namely, in addition to the measurements depending linearly on the unknown signal, certain parameters of the measurement procedure enter in a linear fashion. Hence, one cannot employ a linear model (for example, in connection compressed sensing techniques [CRT06]) unless one has an accurate estimate of these parameters.

When such estimates are not available or too expensive to obtain, there are certain asymmetric scenarios where one of the inputs can be recovered even though the other one is out of reach (e.g., [XLTk95, LKR18], this scenario is sometimes referred to as passive imaging). In most cases, however, the natural aim will be to recover both the signal and the parameters, that is, to solve the associated bilinear inverse problem. Even when some estimates of the parameters are available, such a unified approach will be preferred in many situations, especially when information is limited. Consequently, the study of bilinear inverse problems, including, but not limited to, the important problem of blind deconvolution, has been an active area of research for many years [Hay94].

Observe that bilinear maps admit a representation as a linear map in the rank one outer product of the unknown signal and the parameter vector. Thus, one can approach such problems using tools from the theory of low-rank recovery (see, e.g., [ARR14, LS17, JKS18]). Under sparsity assumptions however, the direct applicability of these approaches is limited, as two competing objectives arise:

5. Sparse Rank-One Matrix Recovery

one aims to simultaneously minimize rank and sparsity. That is the case when the signals and/or parameter vectors admit an approximate representation using only a small (but unknown) subset of an appropriate basis. For more details regarding the appearances of such assumptions in bilinear inverse problems, see [LS15]. As a consequence, the problem becomes considerably more difficult; Oymak et al., for example, have demonstrated that minimizing linear combinations of the nuclear norm (a standard convex proxy for the rank) and the ℓ^1 -norm (the corresponding quantity for sparsity) exhibits suboptimal scaling [OJF⁺15]. In fact it is not even clear if, without additional assumptions, efficient recovery is at all possible for a near-linear number of measurements (as it would be predicted identifiability considerations [KK17]).

Recently, a number of nonconvex algorithms for bilinear inverse problems have been proposed. For example, for such problems without sparsity constraints, several such algorithms have been analyzed for blind deconvolution and related problems [LLSW19, LS18] with near-optimal recovery guarantees. In contrast, our understanding of bilinear inverse problems with sparsity constraints is still in its infancy. Recently, several algorithms have been analyzed for sparse phase retrieval [BR18, Sol19] or blind deconvolution with sparsity constraints [QZEW20]. The recovery guarantees for these algorithms, however, are either suboptimal in the number of necessary measurements or only local convergence guarantees are available, i.e., one relies on the existence of a good initialization. A noteworthy exception are the two related papers [BR16, IVW17], where a two-stage approach for (sparsity) constrained bilinear inverse problems is proposed, which achieves recovery at near-optimal rate. However, the algorithm relies on a special nested structure of the measurements, which is not feasible for many practical applications.

In [LWB18] Lee, Wu, and Bresler introduced the *Sparse Power Factorization* (SPF) method together with a tractable initialization procedure based on alternating minimization. They also provide a first performance analysis of their method for random bilinear measurements in the sense that their lifted representation is a matrix with independent Gaussian entries. That is, they work with linear operators $\mathcal{A}: \mathbb{C}^{n_1 \times n_2} \rightarrow \mathbb{C}^m$ that admit a representation as

$$[\mathcal{A}(X)]_\ell = \text{tr}(A_\ell^* X)$$

for i.i.d. Gaussian matrices $A_\ell \in \mathbb{C}^{n_1 \times n_2}$ as considered in Section 3.1 in Equation (3.1).

For such measurements they show that with high probability, SPF converges locally to the right solution, i.e., one has convergence for initializations not too far from the signal to be recovered.

For signals that have a very large entry, they also devise a tractable initialization procedure – they call it thresholding initialization – such that one has global convergence to the right solution. Local convergence has also been shown for the multi-penalty approach *A-T-LAS*_{1,2} [FMN21], but to our knowledge, comparable global recovery guarantees are not available to date. This is why we focus on SPF in this chapter, using the results of [LWB18] as our starting point.

The precise condition for their guarantee to hold is that both (normalized) input signals need to be larger than some $c > \frac{1}{2}$ in supremum norm – more than one quarter of its mass needs to be located in just one entry, that is, the signals must have a very high peak to average power ratio, cf. Definition 3.6. Thusly motivated, we will develop an improved criterion that accepts arbitrary mass distribution within the signal while only requiring a mildly increased number of measurements.

Let $b \in \mathbb{C}^m$ be given by $b := B(u, v) + z$ where z is a perturbation term, e.g. noise. In the sense of Section 3.1, our goal will be to reconstruct u and v from linear measurements given by

$$b_\ell = \text{tr}(A_\ell^* u v^*)$$

At the core of the Sparse Power Factorization Algorithm, as introduced in [LWB18], are the linear operators $F: \mathbb{C}^{n_2} \rightarrow \mathbb{C}^{m \times n_1}$ and $G: \mathbb{C}^{n_1} \rightarrow \mathbb{C}^{m \times n_2}$ defined by

$$F(y) := \begin{pmatrix} y^* A_1^* \\ \vdots \\ y^* A_m^* \end{pmatrix}, \quad G(x) := \begin{pmatrix} x^* A_1 \\ \vdots \\ x^* A_m \end{pmatrix}. \quad (5.1)$$

A direct consequence of this definition is that for all $x \in \mathbb{C}^{n_1}$ and all $y \in \mathbb{C}^{n_2}$ we have

$$\mathcal{A}(x y^*) = [F(y)]x = \overline{[G(x)]y},$$

The preparation for our goal of relaxed, i.e. more general, performance criteria subdivides this chapter into three sections. We will first present the algorithmic framework of the Sparse Power Factorization (SPF) algorithm in Section 5.1. To this end, we first discuss pursuit algorithms. This class of algorithms provide (approximate) solutions to underdetermined linear inverse problems while adhering to a sparsity constraint. Afterwards we consider the SPF algorithm itself. Within each iteration, SPF requires a sparse solution to a system of linear equations and hence employs a pursuit algorithm. Moreover, as SPF will only exhibit local convergence, we address the crucial part of a sufficient initialization. We further consider the approach of SPF in the context of NP-hard sparse recovery problems, i.e. especially SparsePCA.

Then follows a rather technical section. The goal of Section 5.2 is to fit our improved recovery criteria onto the proof framework of [LWB18]. To do this, we first cite their recovery guarantee and some auxiliary lemmas. In a second step, we complement their approach by posing new lemmas with our relaxed criteria in mind to, eventually, achieve a seamless fit of both their contribution and our own in the final section.

Finally, Section 5.3 allows us to insert the framework buildup of the preceding section into our new, adaptive recovery criteria. In the end, we will discuss our result in comparison to competing concepts and in which cases this outperforms state-of-the-art algorithms.

5.1. Algorithms

In this section we present the *Sparse Power Factorization Algorithm* (SPF), developed by Lee, Wu and Bresler, [LWB18]. As their paper elaborately explains the composition of the algorithm, we will restrict ourselves to a brief introduction.

Here, we start with pursuit algorithms, which subsequently will be employed to solve sub-problems in the iterations of SPF. Secondly, we state the SPF algorithm in the form presented by Lee et al. in [LWB18]. Most importantly, we will look for a suitable initialization procedure to get close enough to the ground truth for SPF to perform satisfactorily.

Pursuit Algorithms

When using overcomplete representation systems, e.g. frames, cf. Section 1.2, one will at some point inevitably come to the problem of encoding a signal optimally with respect to said frame. While the meaning of the word *optimal* may depend on the larger context, often one wants to represent a signal with a fixed number s of coefficients with minimal error. This is well understood to be NP-hard, however. For a more in-depth discussion I refer the reading [Mal09] which provides a pleasant read on this topic. Especially the twelfth chapter may be of interest with respect to this very topic.

For this thesis it will suffice to understand a *Pursuit Algorithm* to be a procedure that aims at solving an underdetermined linear system

$$Ax = y,$$

for $x \in \mathbb{C}^n$, provided the measurement matrix $A \in \mathbb{C}^{m \times n}$ and measured data $y \in \mathbb{C}^m$, where $m < n$ while requiring solutions to possess a maximum number of non-zero coefficients, i.e. $\|x\|_0 \leq s$.

There exists an ample variety of different algorithms for this task. Many of these are based on a greedy approach to avoid the complexity imposed by NP-hardness. As discussed before in Chapter 4, the ℓ^1 -norm is an accessible proxy for sparsity and is hence employed in numerous algorithms. Others use novel ideas, such as a constantly updated weighted least-squares minimization.

Notable algorithms include Hard Thresholding Pursuit (HTP), [Fou11], Basis Pursuit (BP), [CDS98], (Compressive Sampling) Matching Pursuit (CoSaMP), [NT09, MZ93] or Iteratively Reweighted Least Squares (IRLS), [DDFG10].

In this thesis we will only use HTP in accordance with [LWB18]. There, the Hard Thresholding Pursuit Algorithm is defined as follows:

5.1 Algorithm (HTP(A, b, s)).

Input: Measurement matrix $A \in \mathbb{C}^{m \times n}$, measurement $b \in \mathbb{C}^m$, sparsity constraint $s \in \mathbb{N}$.

Output: $\hat{x} \in \mathbb{C}^n$.

- 1: $t \leftarrow 0$
- 2: **while** stop condition not satisfied **do**
- 3: $t \leftarrow t + 1$
- 4: $w = x^{(t-1)} + A^* (b - Ax^{(t-1)})$
- 5: $J \leftarrow \arg \max_{J \subset \{1, \dots, n\}, |J|=s} \|w_J\|_2$
- 6: $x^{(t)} \leftarrow \arg \min_{x: \text{supp}(x) \subset J} \|Ax - b\|_2$
- 7: **end while**
- 8: **return** $\hat{x} \leftarrow x^{(t)}$

Sparse Power Factorization

The idea of Sparse Power Factorization is to iteratively update estimates $u^{(t)}$ and $v^{(t)}$ for u and v in an alternating fashion. That is, in each iteration one keeps one of $v^{(t)}$ and $u^{(t)}$ fixed and updates the respective other by solving an (underdetermined) linear system. Solving each of these linear systems then amounts to solving a linear inverse problem with sparsity constraints, which is the topic of pursuit algorithms as discussed above.

In the context of my Bachelor's thesis, [Gep15], I have conducted research on interchanging HTP for different pursuit algorithms in the iterations of SPF. When using, e.g., CoSaMP, the required RIP of the measurement operator is slightly varied, i.e. $(4s_1, 4s_2, 2)$ -RIP for CoSaMP as compared to $(3s_1, 3s_2, 2)$ -RIP for HTP. Also the restricted isometry constant δ has a tighter bound for CoSaMP. Moreover, numerical testing suggests similar and, more importantly, acceptable behavior for a range of pursuit algorithms.

Nonetheless, in this thesis we will restrict ourselves to Hard Thresholding Pursuit (HTP), cf. Algorithm 5.1, as the authors of [LWB18] have done for their analysis, too. In order to alternately update the left and right singular vectors $u^{(t)}$ and $v^{(t)}$ we need to consider the application of \mathcal{A} with one fixed singular vector, i.e. the application of $\mathcal{A}(u \cdot *)$ and $\mathcal{A}(\cdot v^*)$. This can be handily realized with the help of the functions F and G of Equation (5.1) as

$$\mathcal{A}(u \cdot *) = \overline{G(u)} \quad \text{and} \quad \mathcal{A}(\cdot v^*) = F(y).$$

In the case of no desired sparsity in either singular vector, the algorithm will default to common least-squares minimization in the respective update step.

With this, the Sparse Power Factorization Algorithm reads as follows:

5. Sparse Rank-One Matrix Recovery

5.2 Algorithm (Algorithm 1 in [LWB18]).

Input: Operator \mathcal{A} , Measurement b , Sparsity Constraints s_1, s_2 , Initialization v_{init} .

Output: Estimate \widehat{X} .

```

1:  $t \leftarrow 0$ 
2: while stop condition not satisfied do
3:    $t \leftarrow t + 1$ 
4:    $v^{(t-1)} \leftarrow \frac{v^{(t-1)}}{\|v^{(t-1)}\|_2}$ 
5:   if  $s_1 < n_1$  then
6:      $u^{(t)} \leftarrow \text{HTP}(\mathcal{F}(v^{(t-1)}), b, s_1)$ 
7:   else
8:      $u^{(t)} \leftarrow \arg \min_x \|b - [\mathcal{F}(v^{(t-1)})]x\|_2^2$ 
9:   end if
10:   $u^{(t)} \leftarrow \frac{u^{(t)}}{\|u^{(t)}\|_2}$ 
11:  if  $s_2 < n_2$  then
12:     $v^{(t)} \leftarrow \text{HTP}(\mathcal{G}(u^{(t)}), \bar{b}, s_2)$ 
13:  else
14:     $v^{(t)} \leftarrow \arg \min_b \|\bar{b} - [\mathcal{G}(u^{(t)})]b\|_2^2$ 
15:  end if
16: end while
17: return  $\widehat{X} \leftarrow u^{(t)}(v^{(t)})^*$ 

```

Initialization

As for many other non-convex algorithms (e.g. [JNS13, CLS15]), the convergence properties of Sparse Power Factorization depend crucially on the choice of the starting point. In [JNS13, CLS15] the starting point is chosen via a spectral initialization. That is, one chooses the leading left- and right-singular vectors of $\mathcal{A}^*(y)$ as the starting point. The general idea behind this approach is that, if the map \mathcal{A} is random, under appropriate assumptions, the leading eigenvectors of the expected value, $\mathbb{E}[\mathcal{A}(y)]$ are given by u and v . Indeed under the assumptions in this chapter, we have that $\mathbb{E}[\mathcal{A}^* \mathcal{A}] = \text{id}$. If $z = 0$, this implies that $\mathbb{E}[\mathcal{A}(y)] = \mathbb{E}[\mathcal{A}^* \mathcal{A}(uv^*)] = uv^*$. However, spectral initialization requires that the number of measurements is at the order of $\max\{n_1, n_2\}$, as otherwise the matrix $\mathcal{A}^*(y)$ will in general not concentrate around its expectation. Hence, spectral initialization will in general not be optimal as it does not take into account the sparsity of the vectors u and v . One way to incorporate the sparsity assumption would be to solve the Sparse Principal Component Analysis (SparsePCA) problem.

$$\begin{aligned}
& \max \quad \text{Re}(\tilde{u}^* \mathcal{A}^*(y) \tilde{v}) \\
& \text{subject to} \quad \|\tilde{u}\|_0 \leq s_1, \|\tilde{u}\|_2 = 1 \\
& \quad \quad \quad \|\tilde{v}\|_0 \leq s_2, \|\tilde{v}\|_2 = 1,
\end{aligned} \tag{5.2}$$

where $\|\cdot\|_0$ denotes the number of non-zero entries. As it was shown in [LWB18, Proposition III.4], Algorithm 5.2, if initialized by a solution of Equation (5.2) is able to recover the solution u and v from a number of measurements at the order of $(s_1 + s_2) \max\left\{\frac{s_1}{n_1}, \frac{s_2}{n_2}\right\}$. However, the SparsePCA problem has been shown to be NP-hard [TP14]. Nevertheless, in the last fifteen years there has been a lot of research on the SparsePCA problem and, in particular, on tractable (i.e., polynomial-time) algorithms, which yield good approximations to the true solution. Several computationally tractable algorithms have been proposed for solving Equation (5.2), e.g., thresholdings algorithms [Mar13], a general version of the power method [JNRS10] and semidefinite programs [dBE08]. From the statistical perspective, a particular emphasis has been put for computationally efficient or at least tractable algorithms on the analysis of the single spike model, [AW09, KNV15, DM16]. These approaches, however, require that the number of samples scales with the square of the number of non-zero entries of the signal to estimate (up to log-factors). This raised the question whether there are fundamental barriers preventing the SparsePCA problem to be solved in polynomial time at a sampling rate close to the information theoretic limit. Indeed, it has been shown that an algorithm achieving this would also allow for an algorithm which solves the k -clique problem in polynomial time [BR13, WBS16]. However, a widely believed conjecture in theoretical computer science states that this is not the case, which indicates that this approach will not be suited for initializing bilinear recovery problems either.

Here, we will analyze the following initialization algorithm, which is the one proposed in [LWB18]. For a set $J_1 \subset \{1, \dots, n_1\}$, respectively $J_2 \subset \{1, \dots, n_2\}$ in the following we will denote by Π_{J_1} , respectively Π_{J_2} the matrix, which projects a vector onto the components which belong to J_1 , respectively J_2 . We further will use the short-hand $\Pi_{J_2}^\perp = \Pi_{\{1, \dots, n_2\} \setminus J_2}$.

5.3 Algorithm (Algorithm 3 in [LWB18]).

Input: Operator \mathcal{A} , Measurement b , Sparsity Constraints s_1, s_2 ,

Output: Initial guess v_{init} for $v \in \mathbb{C}^{n_2}$.

- 1: For all $i \in \{1, \dots, n_1\}$ let ξ_i be the ℓ^2 -norm of the best s_2 -sparse approximation of the i th row of the matrix $\mathcal{A}^*(b) \in \mathbb{C}^{n_1 \times n_2}$.
- 2: Let $\hat{J}_1 \subset \{1, \dots, n_1\}$ be the set of the s_2 largest elements in $\{\xi_1; \xi_2; \dots; \xi_{n_1}\}$
- 3: Choose \hat{J}_2 to contain the indices of the s_2 columns of $\Pi_{\hat{J}_1} \mathcal{A}^*(b)$ largest in ℓ^2 -norm, i.e.,

$$\hat{J}_2 := \arg \max_{J \subset \{1, \dots, n_2\}, |J|=s_2} \|\Pi_{\hat{J}_1} [\mathcal{A}^*(b)] \Pi_J\|_{\text{F}}. \quad (5.3)$$

- 4: **return** v_{init} , the leading right singular vector of $\Pi_{\hat{J}_1} [\mathcal{A}^*(b)] \Pi_{\hat{J}_2}$.

5.2. Assessing Locality

As SPF will at best converge locally, successful recovery crucially depends on its initialization. In this sense, this section prepares toolwork to analyze the initialization for the performance guaran-

5. Sparse Rank-One Matrix Recovery

tees in Section 5.3. To that end, we check, amongst other things, which entries the initialization finds correctly and assess the angle between the correct left singular vector v and the result of the initialization v_{init} . Large parts of this section will be identical or close to [LWB18] and are stated here for completeness of the chain of arguments. Some lemmas will be of a more technical nature, but will become relevant pieces in Section 5.3 nonetheless.

In the following, we will work with the model of Equation (3.1), i.e., we observe

$$\text{tr}(A_\ell^* uv^*) + z_\ell$$

where $u \in \mathbb{C}^{n_1}$ is s_1 -sparse, $v \in \mathbb{C}^{n_2}$ is s_2 -sparse, and $z_\ell \in \mathbb{C}^m$ is noise. As in [LWB18], ν will quantify the Noise-to-Signal Ratio by

$$\nu(z) := \frac{\|z\|_2}{\|\mathcal{A}(uv^*)\|_2}.$$

In our results, we will make no assumptions on the noise z except that $\nu(z)$ needs to be smaller than a certain threshold. Our results will hold for adversarial noise, i.e., the recovery guarantees will hold uniformly for all $z \in \mathbb{C}^m$, such that $\nu(z)$ satisfies this threshold. For our analysis, \mathcal{A} will be a Gaussian linear operator, that is, all the entries of the matrices A_1, \dots, A_m are independent with distribution $\mathcal{CN}(0, \frac{1}{m})$. Here, a complex-valued random variable X has distribution $\mathcal{CN}(0, \frac{1}{m})$ if its real and complex part are independent Gaussians with expectation 0 and variance $\sqrt{\frac{\sigma}{2}}$.

In [LWB18], the authors derived that Algorithm 5.2, if initialized by Algorithm 5.3, is able to recover both u and v (up to scale ambiguity), if both u and v belong to a certain restricted class of signals. More precisely, they proved the following result. This extends the RIP of Gaussian matrices, as discussed in Proposition 3.4, by the adequate recovery guarantee.

5.4 Theorem (Cf. III.7 and III.10 in [LWB18]).

Assume that $\mathcal{A}: \mathbb{C}^{n_1 \times n_2} \rightarrow \mathbb{C}^m$ is a Gaussian linear operator as described before. Let $b = \mathcal{A}(uv^*) + z$, where u is s_1 -sparse and v is s_2 -sparse. Suppose that $\|u\|_\infty \geq 0.78\|u\|_2$, $\|v\|_\infty \geq 0.78\|v\|_2$, and that the noise level satisfies $\nu(z) \leq 0.04$. Then, with probability exceeding $1 - \exp(-c_1 m)$, the output of the Algorithm 5.2, initialized by Algorithm 5.3, converges linearly to uv^* provided that

$$m \geq c_2 (s_1 + s_2) \log \left(\max \left\{ \frac{n_1}{s_1}, \frac{n_2}{s_2} \right\} \right),$$

where $c_1, c_2 > 0$ are absolute constants.

Note that in order to apply Theorem 5.4 to signals u and v , one needs to require that more than half of the mass of u and v is located in one single entry. This is a severe restriction, which can be prohibitive for many applications. Our goal in the following will be to considerably relax this assumption by slightly increasing the amount of required measurements. We will relax this

assumption in two different ways: On the one hand we will show that one can replace 0.78 by an arbitrary small constant that will then show up in the number of measurements. On the other hand we generalize the result to the case that a significant portion of mass of u is concentrated on a small number of entries k , rather than just one of them.

Now, we prepare the necessary tools for the main goal of this chapter, Theorem 5.13. As in [LWB18, Lemma VIII.7] we will need the following quantity, which depends on δ and ν .

$$\omega_{\text{sup}} := \sup \left\{ \omega \in [0, \frac{\pi}{2}) : \omega \geq \arcsin (C_{\delta} [\delta \tan(\omega) + (1 + \delta) \nu \sec(\omega)]) \right\} \quad (5.4)$$

Here, the constant C_{δ} is given by the expression

$$C_{\delta} = 1.1 \frac{\sqrt{\frac{2}{1-\delta^2}} + \frac{1}{1-\delta}}{1 - \sqrt{\frac{2}{1-\delta^2}} \delta},$$

as it can be seen by an inspection of the proof of Lemma VIII.1 in [LWB18]. The precise value of C_{δ} will not be important in the following, we will only use that $2 \leq C_{\delta} \leq 5$ for $\delta \leq 0.04$.

A simple estimate for ω_{sup} is given by the following lemma.

5.5 Lemma.

Assume that $\delta \leq 0.04$ and $\nu \leq 0.04$. Then it holds that

$$\frac{1}{2} \leq \sin(\omega_{\text{sup}}) \leq 1.$$

Proof.

We observe that, in order to show the claim, it is enough to verify that $\omega = \arcsin(\frac{1}{2})$ fulfills the inequality in Equation (5.4). Indeed, using $\cos(\omega) = \sqrt{\frac{3}{4}}$ and $C_{\delta} \leq 5$ we obtain that

$$C_{\delta} \left[\delta \tan \left(\arcsin \left(\frac{1}{2} \right) \right) + (1 + \delta) \nu \sec \left(\arcsin \left(\frac{1}{2} \right) \right) \right] = C_{\delta} \left[0.04 \frac{1/2}{\sqrt{3/4}} + \frac{1.04 \cdot 0.04}{\sqrt{3/4}} \right] \leq \frac{1}{2}.$$

qed

The quantity ω_{sup} controls the maximal angle between the initialization v_{init} and the ground truth v such that the Sparse Power Factorization is guaranteed to converge as captured by the following theorem.

5.6 Theorem (Theorem III.9 in [LWB18]).

Assume that

- 1) \mathcal{A} has the $(3s_1, 3s_2, 2)$ -RIP with isometry constant $\delta \leq 0.08$,
- 2) $\nu \leq 0.08$,
- 3) the initialization v_{init} satisfies $\sin(\angle(v_{\text{init}}, v)) < \sin(\omega_{\text{sup}})$.

Then the iterates $\{x^{(t)}\}_{t \in \mathbb{N}}$ generated by Algorithm 5.2, initialized via Algorithm 5.3, satisfy

$$\limsup_{t \rightarrow \infty} \frac{\|x^{(t)} - uv^*\|_{\text{F}}}{\|uv^*\|_{\text{F}}} \leq 8.3\nu.$$

Furthermore, the convergence is linear, i.e., for all $t \gtrsim \log(\frac{1}{\varepsilon})$ we have that

$$\frac{\|X_t - uv^*\|_{\text{F}}}{\|uv^*\|_{\text{F}}} \leq 8.3\nu + \varepsilon. \quad (5.5)$$

Thus, it remains to verify that the initialization satisfies $\sin(\angle(v_{\text{init}}, v)) < \sin(\omega_{\text{sup}})$. The following lemma gives an upper bound on $\sin(\angle(v_{\text{init}}, v))$.

5.7 Lemma (Lemma VIII.10 in [LWB18]).

Assume that the $(3s_1, 3s_2, 2)$ -restricted isometry property holds for some constant $\delta > 0$. Furthermore, assume that $\|u\|_2 = \|v\|_2 = 1$. Let $\widehat{\mathcal{J}}_1 \subseteq \{1, \dots, n_1\}$ and $\widehat{\mathcal{J}}_2 \subseteq \{1, \dots, n_2\}$ denote the output resulting from Algorithm 5.3. Denote by v_{init} the leading right singular vector of $\Pi_{\widehat{\mathcal{J}}_1}[\mathcal{A}^*(b)]\Pi_{\widehat{\mathcal{J}}_2}$. Then it holds that

$$\sin(\angle(v_{\text{init}}, v)) \leq \frac{\|\Pi_{\widehat{\mathcal{J}}_1} u\|_2 \|\Pi_{\widehat{\mathcal{J}}_2}^\perp v\|_2 + (\delta + \nu + \delta\nu)}{\|\Pi_{\widehat{\mathcal{J}}_1} u\|_2 - (\delta + \nu + \delta\nu)}. \quad (5.6)$$

The next lemma shows that initialization according to Algorithm 5.3 successfully finds non-zero components whose absolute value is above $2(\delta + \nu + \delta\nu)$.

5.8 Lemma (Lemma VIII.12 in [LWB18]).

Assume that u and v , are s_1 -sparse, respectively s_2 -sparse, and satisfy $\|u\|_2 = \|v\|_2 = 1$ and assume that the measurement operator \mathcal{A} satisfies the $(3s_1, 3s_2, 2)$ -RIP with constant δ . Recall that $\widehat{\mathcal{J}}_1 \subset \{1, \dots, n_1\}$ is the support estimate for v_{init} given by the initialization Algorithm 5.3. Define

$$\widetilde{\mathcal{J}}_1 := \{1 \leq j \leq n_1 : |u_j| \geq 2(\delta + \nu + \delta\nu)\}. \quad (5.7)$$

Then we have that $\widetilde{\mathcal{J}}_1 \subset \widehat{\mathcal{J}}_1$.

We now present a lemma that can be understood as a case-specific adaptation of the restricted isometry property. The assertion states that the tensor operator $\Pi_{\tilde{J}_1}(\mathcal{A}^* \mathcal{A} - I) \otimes \Pi_{\tilde{J}_2}$ is bounded in terms of δ , the isometry constant of \mathcal{A} .

5.9 Lemma (Lemma A.2 and Lemma A.3 in [LWB18]).

Assume that the $(3s_1, 3s_2, 2)$ -restricted isometry property is fulfilled for some restricted isometry constant $\delta > 0$. Assume that the cardinality of $\tilde{J}_1 \subseteq \{1, \dots, n_1\}$, respectively $\tilde{J}_2 \subseteq \{1, \dots, n_2\}$ is at most $2s_1$, respectively $2s_2$. Then, whenever $u \in \mathbb{C}^{n_1}$ is at most $2s_1$ -sparse and $v \in \mathbb{C}^{n_2}$ is at most $2s_2$ -sparse, we have that

$$\|\Pi_{\tilde{J}_1}[(\mathcal{A}^* \mathcal{A} - I)(uv^*)]\Pi_{\tilde{J}_2}\|_2 \leq \delta \|uv^*\|_F.$$

Furthermore for all $z \in \mathbb{C}^n$ and for all $\tilde{J}_1 \subseteq \{1, \dots, n_1\}$, respectively $\tilde{J}_2 \subseteq \{1, \dots, n_2\}$, with cardinality at most s_1 , respectively s_2 , we have that

$$\|\Pi_{\tilde{J}_1}[\mathcal{A}^*(z)]\Pi_{\tilde{J}_2}\|_2 \leq \sqrt{1 + \delta} \|z\|_2.$$

Following the series of lemmas quoted from [LWB18], the next statement provides a relation between the mass of u and v in the index sets \hat{J}_1 and \hat{J}_2 , respectively, on the one hand and the index set \tilde{J}_1 together with the largest entry of v on the other hand.

5.10 Lemma.

Assume that \mathcal{A} has the $(3s_1, 3s_2, 2)$ -restricted isometry property with isometry constant $\delta > 0$ and let u and v as in Lemma 5.8. Let \tilde{J}_1 be defined as in Equation (5.7). Then, it holds that

$$\|\Pi_{\tilde{J}_1} u\|_2 \|\Pi_{\hat{J}_2} v\|_2 \geq \|\Pi_{\tilde{J}_1} u\|_2 \|v\|_\infty - 2(\delta + \nu + \delta\nu).$$

Proof.

Recall that $b = \mathcal{A}(X) + z$ and define k_1 and k_2 by

$$\begin{aligned} k_1 &:= \arg \max_{k \in \{1, \dots, n_2\}} |v_k| \\ k_2 &:= \arg \max_{k \in \{1, \dots, n_2\}} \|\Pi_{\hat{J}_1}[\mathcal{A}^*(b)]\Pi_{\{k\}}\|_F. \end{aligned} \tag{5.8}$$

The starting point of our proof is the observation that

$$\|\Pi_{\tilde{J}_1}[\mathcal{A}^*(b)]\Pi_{\{k_2\}}\|_F \geq \|\Pi_{\tilde{J}_1}[\mathcal{A}^*(b)]\Pi_{\{k_1\}}\|_F \geq \|\Pi_{\tilde{J}_1}[\mathcal{A}^*(b)]\Pi_{\{k_1\}}\|_F, \tag{5.9}$$

where the first inequality is due to the definition of k_2 and the second one follows from Lemma 5.8, entailing $\tilde{J}_1 \subset \hat{J}_1$. The right hand side of the inequality chain can be estimated from below by

5. Sparse Rank-One Matrix Recovery

$$\begin{aligned}
& \|\Pi_{\widehat{\mathcal{J}}_1}[\mathcal{A}^*(b)]\Pi_{\{k_1\}}\|_{\text{F}} \\
& \geq \|\Pi_{\widehat{\mathcal{J}}_1}uv^*\Pi_{\{k_1\}}\|_{\text{F}} - \|\Pi_{\widehat{\mathcal{J}}_1}[(\mathcal{A}^*\mathcal{A} - I)(uv^*)]\Pi_{\{k_1\}}\|_{\text{F}} - \|\Pi_{\widehat{\mathcal{J}}_1}\mathcal{A}^*(z)\Pi_{\{k_1\}}\|_{\text{F}} \\
& \geq \|\Pi_{\widehat{\mathcal{J}}_1}uv^*\Pi_{\{k_1\}}\|_{\text{F}} - \left(\delta\|uv^*\|_F + \sqrt{1 + \delta}\|z\|_2\right) \\
& \geq \|\Pi_{\widehat{\mathcal{J}}_1}u\|_2\|v\|_{\infty} - (\delta + \nu + \delta\nu).
\end{aligned} \tag{5.10}$$

In the first inequality we used $b = \mathcal{A}(uv^*) + z$ and the triangle inequality. The second inequality follows from Lemma 5.9. The last line follows from $\|uv^*\|_F = 1$ and $\|z\|_2 = \nu$. Next, we will estimate the left hand side of Equation (5.9) by

$$\begin{aligned}
& \|\Pi_{\widehat{\mathcal{J}}_1}[\mathcal{A}^*(b)]\Pi_{\{k_2\}}\|_{\text{F}} \\
& \leq \|\Pi_{\widehat{\mathcal{J}}_1}uv^*\Pi_{\{k_2\}}\|_{\text{F}} + \left(\delta\|uv^*\|_F + \sqrt{1 + \delta}\|z\|_2\right) \\
& \leq \|\Pi_{\widehat{\mathcal{J}}_1}u\|_2\|\Pi_{\{k_2\}}v\|_2 + (\delta + \nu + \delta\nu) \\
& \leq \|\Pi_{\widehat{\mathcal{J}}_1}u\|_2\|\Pi_{\widehat{\mathcal{J}}_2}v\|_2 + (\delta + \nu + \delta\nu).
\end{aligned} \tag{5.11}$$

The first two lines are obtained by an analogous reasoning as for Equation (5.10). The last line is due to $\{k_2\} \subset \widehat{\mathcal{J}}_2$, which is a consequence of the definition of $\widehat{\mathcal{J}}_2$ in Equation (5.3) and the definition of $\{k_2\}$ in Equation (5.8). We finish the proof by combining the inequality chains of Equations (5.9) to (5.11). **qed**

Lemma 5.10 is actually a slight generalization of what has been shown in [LWB18, p. 1685]. For completeness we have included a proof, which closely follows the proof in [LWB18].

We will now piece together these ingredients to obtain a sufficient condition; in the remainder of this section we will then show that this condition holds in our measurement setup. First note that in order to apply Theorem 5.6 we need to check that

$$\sin(\angle(v_{\text{init}}, v)) < \sin(\omega_{\text{sup}})$$

is satisfied. By Lemma 5.7 it is sufficient to show that the right hand side of Equation (5.6) is strictly smaller than $\sin(\omega_{\text{sup}})$. Combining this with the equality $\|\Pi_{\widehat{\mathcal{J}}_2}^\perp v\|_2 = \sqrt{1 - \|\Pi_{\widehat{\mathcal{J}}_2}v\|_2^2}$ we obtain the sufficient condition

$$\|\Pi_{\widehat{\mathcal{J}}_1}u\|_2\sqrt{1 - \|\Pi_{\widehat{\mathcal{J}}_2}v\|_2^2} < \sin(\omega_{\text{sup}}) \left(\|\Pi_{\widehat{\mathcal{J}}_1}u\|_2 - (\delta + \nu + \delta\nu) \right) - (\delta + \nu + \delta\nu).$$

Further manipulations yield that this is equivalent to

$$\begin{aligned} \|\Pi_{\hat{\mathcal{J}}_1} u\|_2^2 &< \left(\sin(\omega_{\text{sup}}) \|\Pi_{\hat{\mathcal{J}}_1} u\|_2 - (1 + \sin(\omega_{\text{sup}}))(\delta + \nu + \delta\nu) \right)^2 \\ &+ \|\Pi_{\hat{\mathcal{J}}_1} u\|_2^2 \|\Pi_{\hat{\mathcal{J}}_2} v\|_2^2. \end{aligned} \quad (5.12)$$

Hence, in the following our goal will be to verify Equation (5.12). We already noticed that the angle ω_{sup} measures how much the vector v_{init} given by the initialization has to be aligned with the ground truth v in order for the Sparse Power Factorization to converge. Consequently, it is natural to expect that the smaller the constant δ and the noise-to-signal ratio ν , the less the initialization vector has to be aligned with the ground truth, i.e., the larger ω_{sup} can be. This fact is captured by the following lemma.

5.11 Lemma.

Let $\delta \leq 0.04$ and $\nu \leq 0.04$. Then it holds that

$$\sin(\omega_{\text{sup}}) \geq 1 - C_\delta^2 (\delta + 2\delta\nu + 2\nu)^2.$$

Proof.

We will first show that it holds that

$$\sin(\omega_{\text{sup}}) = C_\delta [\delta \tan(\omega_{\text{sup}}) + (1 + \delta)\nu \sec(\omega_{\text{sup}})]. \quad (5.13)$$

For that, we define the set

$$\Omega := \{\omega \in [0, \pi/2) : \sin(\omega) \geq C_\delta [\delta \tan(\omega) + (1 + \delta)\nu \sec(\omega)]\}$$

and note that $\omega_{\text{sup}} = \sup(\Omega)$. Note that from Lemma 5.5 and its proof, it follows that Ω is nonempty. We observe that

$$\lim_{\omega \nearrow \pi/2} C_\delta [\delta \tan(\omega) + (1 + \delta)\nu \sec(\omega)] = +\infty.$$

Hence, Ω is a compact set. In particular, ω_{sup} is contained in this set. Now assume by contradiction that

$$\sin(\omega_{\text{sup}}) > C_\delta [\delta \tan(\omega_{\text{sup}}) + (1 + \delta)\nu \sec(\omega_{\text{sup}})].$$

Because of Section 5.2 and as the sine is bounded in $[0, \pi/2)$, this implies by continuity that there must be an $\tilde{\omega} \in [0, \pi/2)$ such that $\tilde{\omega} > \omega_{\text{sup}}$ and

$$\sin(\tilde{\omega}) = C_\delta [\delta \tan(\tilde{\omega}) + (1 + \delta)\nu \sec(\tilde{\omega})].$$

5. Sparse Rank-One Matrix Recovery

Hence, $\tilde{\omega} \in \Omega$ which is a contradiction to the definition of ω_{sup} . Hence, we have proven Equation (5.13). Using trigonometric identities, we obtain that

$$\sin(\omega_{\text{sup}}) = C_\delta \left[\delta \frac{\sin(\omega_{\text{sup}})}{\sqrt{1 - \sin(\omega_{\text{sup}})^2}} + (1 + \delta) \nu \frac{1}{\sqrt{1 - \sin(\omega_{\text{sup}})^2}} \right].$$

Lemma 5.5 implies that

$$\sin(\omega_{\text{sup}}) \leq \frac{\sin(\omega_{\text{sup}})}{\sqrt{1 - \sin(\omega_{\text{sup}})^2}} C_\delta (\delta + 2(1 + \delta)\nu).$$

Rearranging terms yields that

$$\sin(\omega_{\text{sup}}) \geq \sqrt{1 - C_\delta^2 (\delta + 2\delta\nu + 2\nu)^2}.$$

The claim follows then using the fact that $\sqrt{x} \geq x$ for all $x \in [0, 1]$. qed

5.3. Adaptive Recovery

With the preparation completed, we can now prove our desired result, enabling recovery in various peakiness settings by trading off peakiness against the number of required measurements.

5.12 Proposition.

There are absolute constants $c_1, c_2, c_3 > 0$ such that if

$$m \geq c_1 \delta^{-2} (s_1 + s_2) \log \left(\max \left\{ \frac{n_1}{s_1}, \frac{n_2}{s_2} \right\} \right), \quad (5.14)$$

for some $0 < \delta < 0.01$, then with probability at least $1 - \exp(-c_2 m)$ the following statement holds uniformly for all s_1 -sparse $u \in \mathbb{C}^{n_1}$, s_2 -sparse $v \in \mathbb{C}^{n_2}$ and $z \in \mathbb{C}^m$ such that $\|u\|_2 = \|v\|_2 = 1$ and $\nu(z) \leq 0.01$:

Let the measurements be given by $b = \mathcal{A}(uv^*) + z$ for \mathcal{A} Gaussian as above and let \tilde{J}_1 be defined by

$$\tilde{J}_1 := \{j \in \{1, \dots, n_1\} : |u_j| \geq M_{\delta, \nu}\}, \quad (5.15)$$

where

$$M_{\delta, \nu} := 2(\delta + \nu + \delta\nu).$$

Then, whenever

$$\|\Pi_{\tilde{\mathcal{J}}_1} u\|_2 \|v\|_\infty > c_3 \sqrt{M_{\delta,\nu}}, \quad (5.16)$$

the iterates $\{x^{(t)}\}_{t \in \mathbb{N}}$ generated by Algorithm 5.2 initialized via Algorithm 5.3, satisfy

$$\limsup_{t \rightarrow \infty} \|x^{(t)} - uv^*\|_F \leq 8.3\nu.$$

Furthermore, the convergence is linear in the sense of Equation (5.5).

Proof.

Assumption of Equation (5.14) and Theorem 5.4 yield that with probability at least $1 - \exp(-cm)$ the $(3s_1, 3s_2, 2)$ -restricted isometry property holds with constant δ . For the remainder of the proof, we will consider the event that the restricted isometry property holds for such δ .

We obtain

$$\|\Pi_{\tilde{\mathcal{J}}_1} u\|_2 \|v\|_\infty \geq \left(\sqrt{C_\delta^2 + 1} + 1 \right) \sqrt{M_{\delta,\nu}}$$

from $2 \leq C_\delta \leq 5$ and by choosing the constant c_3 in Equation (5.16) large enough. Combining this with Lemma 5.10 we obtain that

$$\begin{aligned} \|\Pi_{\hat{\mathcal{J}}_1} u\|_2 \|\Pi_{\hat{\mathcal{J}}_2} v\|_2 &\geq \|\Pi_{\tilde{\mathcal{J}}_1} u\|_2 \|v\|_\infty - M_{\delta,\nu}. \\ &> \sqrt{(C_\delta^2 + 1) M_{\delta,\nu}}, \end{aligned} \quad (5.17)$$

where we used that $\sqrt{x} \geq x$ for all $x \in [0, 1]$. This yields a lower bound for the second summand of the right hand side of Equation (5.12). To bound the first summand we estimate

$$\begin{aligned} &\sin(\omega_{\text{sup}}) \|\Pi_{\hat{\mathcal{J}}_1} u\| - (\sin(\omega_{\text{sup}}) + 1) (\delta + \nu + \delta\nu) \\ &\geq \left(1 - C_\delta^2 (\delta + 2\nu + 2\delta\nu)^2 \right) \|\Pi_{\hat{\mathcal{J}}_1} u\| - 2(\delta + \nu + \delta\nu) \\ &\geq \|\Pi_{\hat{\mathcal{J}}_1} u\| - C_\delta^2 (\delta + 2\nu + 2\delta\nu)^2 - 2(\delta + \nu + \delta\nu) \\ &\geq \|\Pi_{\hat{\mathcal{J}}_1} u\| - (C_\delta^2 + 1) M_{\delta,\nu} \\ &\geq 0. \end{aligned} \quad (5.18)$$

In the first line we used Lemma 5.11 and the fact that $\sin(\omega_{\text{sup}}) \leq 1$. The second line is due to $\|\Pi_{\hat{\mathcal{J}}_1} u\| \leq 1$ and the third inequality is due to $\delta \geq 0, \nu \geq 0$. In order to verify the last inequality, it is enough to observe that due to Lemma 5.8 and due to Equation (5.16) with c_3 large enough

$$\|\Pi_{\hat{\mathcal{J}}_1} u\| \geq \|\Pi_{\tilde{\mathcal{J}}_1} u\|_2 \geq \|\Pi_{\tilde{\mathcal{J}}_1} u\|_2 \|v\|_\infty \geq (C_\delta^2 + 1) M_{\delta,\nu},$$

where the last inequality uses that $C_\delta \leq 5$ and $0 \leq \delta, \nu \leq 0.01$. Hence, by squaring Equation (5.18)

5. Sparse Rank-One Matrix Recovery

we obtain that

$$\begin{aligned}
& \left(\sin(\omega_{\text{sup}}) \|\Pi_{\hat{J}_1} u\| - (\sin(\omega_{\text{sup}}) + 1) (\delta + \nu + \delta\nu) \right)^2 \\
& \geq \left(\|\Pi_{\hat{J}_1} u\| - \frac{1}{2} (C_\delta^2 + 1) M_{\delta,\nu} \right)^2 \\
& \geq \|\Pi_{\hat{J}_1} u\|^2 - (C_\delta^2 + 1) M_{\delta,\nu} \|\Pi_{\hat{J}_1} u\| \\
& \geq \|\Pi_{\hat{J}_1} u\|^2 - (C_\delta^2 + 1) M_{\delta,\nu},
\end{aligned}$$

where in the last line we again used that $\|\Pi_{\hat{J}_1} u\| \leq 1$. Together with Equation (5.17) this yields Equation (5.12), as desired. qed

We will now state the main result of this chapter, Theorem 5.13. For that, we recall the $\|\cdot\|_{[k]}$ norm from Notation 3.5. Our main requirement on the vector u will be that a significant amount of its mass is located in the largest k entries, i.e., that $\frac{\|u\|_{[k]}}{\|u\|_2}$ is large enough.

5.13 Theorem.

Let $k \in \{1, \dots, n_1\}$ and $0 < \xi < 1, 0 < \mu < 1$. Then, there are absolute constants $C_1, C_2, C_3 > 0$ such that if, with \mathcal{A} as in Proposition 5.12,

$$m \geq C_1 \max \left\{ \frac{1}{\xi^4 \mu^4}, \frac{k}{\xi^2} \right\} (s_1 + s_2) \log \left(\max \left\{ \frac{n_1}{s_1}, \frac{n_2}{s_2} \right\} \right),$$

then with probability at least $1 - \exp(-C_2 m)$ the following holds.

For all s_1 -sparse $u \in \mathbb{C}^{n_1}$ with $\|u\|_{[k]} \geq \xi \|u\|_2$, all s_2 -sparse $v \in \mathbb{C}^{n_2}$ with $\|v\|_\infty \geq \mu \|v\|_2$, and all $z \in \mathbb{C}^m$ with $\nu(z) \leq C_3 \min \left\{ \xi^2 \mu^2, \frac{\xi}{\sqrt{k}} \right\}$ the iterates $\{x^{(t)}\}_{t \in \mathbb{N}}$ generated by applying Algorithm 5.2, initialized by Algorithm 5.3, satisfy

$$\limsup_{t \rightarrow \infty} \frac{\|x^{(t)} - uv^*\|_F}{\|uv^*\|_F} \leq 8.3\nu.$$

Furthermore, the convergence is linear in the sense of Equation (5.5).

Proof.

We will prove this result by applying Proposition 5.12 with

$$\delta = \min \left\{ \frac{\xi}{6\sqrt{2k}}, \frac{\xi^2 \mu^2}{8c_3^2} \right\}.$$

Let $u \in \mathbb{C}^{n_1}$ s_1 -sparse, $v \in \mathbb{C}^{n_2}$ s_2 -sparse and $z \in \mathbb{C}^m$ such that the assumptions of this theorem are satisfied. Without loss of generality we may assume in the following that $\|u\|_2 = \|v\|_2 = 1$.

First, we note that invoking $\delta, \nu < 0.01$ and potentially decreasing the size of C_3 we have that

$$2(\delta + \nu(z) + \delta\nu(z)) < 2(\delta + 2\nu(z)) \leq \frac{\xi}{\sqrt{2k}}.$$

Hence, we obtain that

$$\check{J}_1 := \left\{ j \in \{1, \dots, n_1\} : |u_j| \geq \frac{\xi}{\sqrt{2k}} \right\} \subset \tilde{J}_1, \quad (5.19)$$

where \tilde{J}_1 is the set defined in Equation (5.15).

Note that

$$\sum_{i \in \{1, \dots, k\} \setminus \check{J}_1} (u_i^*)^2 < \sum_{i \in \{1, \dots, k\} \setminus \check{J}_1} \frac{\xi^2}{2k} \leq \frac{\xi^2}{2},$$

where in the first inequality we have used that $u_i^* < \frac{\xi}{\sqrt{2k}}$ for all $i \in \{1, \dots, k\} \setminus \check{J}_1$. By the assumption $\|u\|_{[k]} \geq \xi$ this yields that

$$\sum_{i \in \{1, \dots, k\} \cap \check{J}_1} (u_i^*)^2 \geq \frac{\xi^2}{2},$$

which in turn implies that $\|\Pi_{\check{J}_1} u\|_2 \geq \frac{\xi}{\sqrt{2}}$. By the inclusion $\check{J}_1 \subset \tilde{J}_1$ of Equation (5.19) we obtain that $\|\Pi_{\tilde{J}_1} u\|_2 \geq \frac{\xi}{\sqrt{2}}$. Hence, using the assumption $\|v\|_\infty \geq \mu$, our choice of δ , the assumption on the noise level $\nu(z)$ and potentially again decreasing the value of the constant C_3 we obtain that

$$\|\Pi_{\tilde{J}_1} u\|_2 \|v\|_\infty \geq \frac{\xi\mu}{\sqrt{2}} \geq c_3 \sqrt{M_{\delta, \nu}}.$$

This shows that Equation (5.16) is satisfied. Hence, we can apply Proposition 5.12 and by inserting our choice of δ into Equation (5.14), so choosing the constant C_1 large enough, we obtain the main result. qed

This result gives rise to a number of special cases.

Peaky Signals. In [LWB18] the authors discuss recovery guarantees for signals u and v with $\frac{\|u\|_\infty}{\|u\|_2}$ and $\frac{\|v\|_\infty}{\|v\|_2}$, both bounded below by an absolute constant $\mu \approx 0.78$. The case $k = 1$ of our theorem yields a direct improvement of this result in the sense that μ can be chosen arbitrarily small with the number of required measurements only increasing by a factor of order μ^{-8} . Hence, even when this constant decays logarithmically in the dimension, the required number of measurements will only increase by logarithmic factors.

5. Sparse Rank-One Matrix Recovery

Signals With Multiple Large Entries. When one of the input signals has multiple large entries, using the $\|\cdot\|_{[k]}$ norm improves upon the resulting guarantee as compared to the scenario just discussed. As an example, assume that $s_1 = s_2 = s$, that u and v are normalized with

$$\|v\|_\infty \geq c_1 s^{-1/8},$$

and that $k = c_2 s^{1/2}$ of the entries of u are of absolute value at least $c_3 s^{-1/4}$. Then $\|u\|_{[k]} \geq \sqrt{c_2} c_3$. Using Theorem 5.13 we obtain that the vectors u and v can be recovered if the number of measurements is on the order of $s^{3/2}$, thus below the order of s^2 that has been established for arbitrary sparse signals in [LS15] (cf. next item). In contrast, applying Theorem 5.13 with $k = 1$ would yield that the number of measurements would have to be on the order of $s^{5/2}$, which is worse than the state-of-the-art.

Arbitrary Sparse Signals. Applying Theorem 5.13 to non-peaky signals yields suboptimal results. Indeed, let $u \in \mathbb{C}^{n_1}$ s_1 -sparse and $v \in \mathbb{C}^{n_2}$ s_2 -sparse be generic vectors. Observe that

$$\|v\|_\infty \asymp \frac{1}{\sqrt{s_2}} \|v\|_2.$$

Consequently, Theorem 5.13 applied with $\xi = 1$, $k = s_1$, and $\mu = \frac{1}{\sqrt{s_2}}$ yields that with high probability a generic s_1 -sparse u and a generic s_2 -sparse v can be recovered from $y = \mathcal{A}(uv^*) + z$, if the number of measurements satisfies

$$m \geq C \max\{s_1; s_2^2\} (s_1 + s_2) \log \left(\max \left\{ \frac{n_1}{s_1}, \frac{n_2}{s_2} \right\} \right),$$

and if the noise level ν is on the order of $\mathcal{O} \left(\max \left\{ \frac{1}{s_2}; \frac{1}{\sqrt{s_1}} \right\} \right)$. Previous results (see, e.g., [LS15]), in contrast, require $m \geq C \max\{s_1^2; s_2^2\} \log \left(\max \left\{ \frac{n_1}{s_1}, \frac{n_2}{s_2} \right\} \right)$ samples.

The peakiness assumptions in Theorem 5.13 may seem arbitrary at first sight, but in certain applications they are reasonable. Namely, when u is the signal transmitted via a wireless channel and v is the unknown vector of channel parameters it is natural to assume that v has a large entry, as the direct path will always carry most of the energy. The signal u can be modified by the sender, so some large entries can be artificially introduced. In this regard, being able to consider multiple entries of comparable size is of advantage as adding a single very large entry will result in a dramatic increase of the peak to average power ratio.

Part IV.

Application: X-Ray Nanoscale Imaging

The topic of *Phase Retrieval* concerns itself with the inversion of phase loss, i.e. the mapping

$$\begin{aligned}\mathbb{C}^{m \times n} &\longrightarrow \mathbb{R}_+^{m \times n} \\ (z_{i,j})_{i,j} &\longmapsto (|z_{i,j}|^2)_{i,j}\end{aligned}$$

which maps any given input entrywise to its squared modulus. While in the real case this leads to loss of the sign, in the complex plane we lose an entire dimension. Thus, generally speaking, this problem is highly ill-posed and any inversion attempts have to make further assumptions in order to become feasible. Global uniqueness of recovery is often impossible, but in many applications one is satisfied with the remaining ambiguities, cf. [BP15, Bei17].

Based on my collaboration with X-ray physicists, this part focusses on reconstructing images created from nanoscale X-ray imaging. In this context, we will restrict ourselves to the following setup. We assume a monochromatic light source which illuminates a thin, so-called pure phase object. By passing through said object, the electromagnetic wave will only shift its phase but will stay short of reducing the light intensity in a significant way. Hence, the light behind the object possesses unchanged energy but the phase shift enables successive interference. The thusly modified wave front is then measured by one or multiple detector planes which register the intensities at given coordinates.

Bearing this setup in mind, we split this part into two chapters.

First, in Chapter 6 we provide a very brief glance into the physical principles of light as an electromagnetic wave. We then use this understanding to form a simple mathematical model of light propagation. More in relevance to this thesis, we also discuss what kind of images we would like to propagate and what kind of priors for inversion schemes this will yield.

Second, in Chapter 7 we discuss algorithmic concepts that could potentially ensure successful recovery. Based upon Part III, we apply a relaxed Douglas-Rachford scheme with exploitation of frame sparsity in Section 7.1 and discuss applicability of phase retrieval through lifting and matrix recovery by Sparse Power Factorization in Section 7.2.

6. Physical Circumstances

In this chapter, we provide the context for phase retrieval in X-ray nanoscale imaging. For the diffraction imaging setup presented in the preamble of this part, we first have to understand the physical properties that govern the propagation and measurement of light – in our case in the form of X-rays. Bearing in mind that this is a work of mathematics and in the interest of brevity we will allow for some abridgement of the presentation, as well as simplification of the underlying matter. For a more extensive treatise I recommend the reading of [Kre15], especially the first two chapters.

The first part of this chapter, Section 6.1, is concerned with an – admittedly compendious – summary of how physicists describe the propagation of electromagnetic waves mathematically. To that end, we state Maxwell’s differential equations that describe wave propagation in vacuum and sketch how to derive the idea of Fresnel propagation from that.

In the second part, Section 6.2, we discuss the problem of phase loss and the resulting illposedness of the recovery problem. This leads to the necessity of additional a priori knowledge that manifests itself in mathematically exploitable priors for reconstruction methods. We further briefly touch on the subject of noise, as distortions obviously will make reconstruction much harder to do, yet are unavoidable in practice.

6.1. Propagation of Light

The physical understanding of the propagation of light – as an electromagnetic wave – can be described by a set of PDEs, known as *Maxwell’s Equations*. There exist numerous approaches that provide (approximate) solutions to those equations that still exhibit an approachable propagation operator. While this is an interesting topic in its own right, this part of the dissertation is focussed rather on enhancing existing recovery schemes by applying results of the preceding sections. Hence, we will touch this rather briefly in this thesis. The interested reader finds a comprehensive presentation in [Loo16, Kre15, Goo05], of which this chapter serves as a summarizing excerpt. For putting this topic into the larger context of physics generally the obvious reading recommendation would be the standard work *The Feynman Lectures on Physics*, [FLSi0]. We will first present a mathematical description for electromagnetic waves and then derive approximative forward propagation operators based on the Fresnel kernel.

Maxwell and Helmholtz Equations

In order to derive and describe the propagation of light by physical terms, we need to fix some notations first.

We begin by introducing the magnetic and electrical fields and their inherent elementary physical constants.

6.1 Notation (Magnetic and Electrical Fields).

We denote the *Electrical Field* by E and the *Magnetic Induction* by B . Both are dependent on a three dimensional coordinate vector $r \in \mathbb{R}^3$ and a time value $t \in \mathbb{R}$.

We further denote ε_0 the *Electric Field Constant* and μ_0 the *Magnetic Field Constant*. With these we define the *Vacuum Wave Velocity* $c := 1/\sqrt{\varepsilon_0\mu_0}$.

Next, we shortly state some derivational concepts for vector operators. We will not concern ourselves with these mathematically but only use them to state Maxwell's equations in Theorem 6.3. For the interested reader, I recommend the reading of [Mat98] for a detailed view on vector calculus, and especially chapter three for the topics of this section.

6.2 Notation (Derivational Operators).

Let $\Phi: \mathbb{R}^3 \rightarrow \mathbb{C}$ a partially differentiable function. We then denote its *Gradient* by $\nabla\Phi$, its *Divergence* by $\operatorname{div}\Phi$ and its *Curl* by $\operatorname{curl}\Phi$.

We also use the shorthand $\partial_t := \partial/\partial t$ for the derivative with respect to the time coordinate.

This now allows us to state the so-called *Maxwell's Equations*. These are a set of coupled partial differential equations that are able to describe the spatial distribution of electromagnetic waves. We will settle for the vacuum case here, because, although the introduction of matter will distinctly change many physical aspects of the wave propagation, the phase retrieval problem we are interested in solving will remain unchanged by that.

6.3 Theorem.

Denoting spatial coordinates by $r \in \mathbb{R}^3$ and the time by $t \in \mathbb{R}$, in vacuum electromagnetic waves satisfy this set of equations

$$\begin{aligned}\operatorname{div} E(r, t) &= 0, \\ \operatorname{div} B(r, t) &= 0, \\ \operatorname{curl} E(r, t) + \partial_t B(r, t) &= 0, \\ \operatorname{curl} B(r, t) - \varepsilon_0\mu_0\partial_t E(r, t) &= 0.\end{aligned}$$

From these equations, one can derive two so-called *Independent Wave Equations*

$$\begin{aligned}(\varepsilon_0\mu_0\partial_t^2 - \nabla^2)E(r, t) &= 0, \\(\varepsilon_0\mu_0\partial_t^2 - \nabla^2)B(r, t) &= 0.\end{aligned}$$

Seeing that this forces both magnetic induction and electric field to be of similar form, this motivates the so-called scalar wave equation or *D'Alembert Equation*

$$\left(\frac{1}{c^2}\partial_t^2 - \nabla^2\right)\Psi(r, t) = 0.$$

A special solution of this is the so-called *Plane Wave* with wavelength λ ,

$$\Psi_{\text{PW}}(r, t) = \exp(i\langle k, r \rangle - \omega t), \quad (6.1)$$

with *Wave Vector* $k \in \mathbb{R}^3$ and $\|k\|_2 = 2\pi/\lambda$ such that $\|k\|_2 = \omega/c$.

We now assume that we have such a wave Ψ . This can be decomposed using the spectral Fourier transform as in Definition 1.4¹

$$\Psi(r, t) = \int_0^\infty \psi_\omega(r) \exp(-2\pi i\omega t) d\omega.$$

for frequency weights ψ_ω .

Combining the above equations yields

$$\int_0^\infty \left[\left(\nabla^2 + \frac{\omega^2}{c^2} \right) \psi_\omega(r) \right] \exp(-2\pi i\omega t) d\omega = 0.$$

For general functions ψ_ω this means

$$\left(\nabla^2 + \|k\|_2^2 \right) \psi_\omega(r) = 0, \quad (6.2)$$

which is known as the *Helmholtz Equation*.

Fresnel Propagation

Now consider a solution ψ of Equation (6.2), known in a plane perpendicular to the optical axis at position z . Further denote $\mathbb{R}^3 \ni r =: (r_\perp, z) \in \mathbb{R}^2 \times \mathbb{R}$. We then decompose this into lateral

¹Apparently, in physical application the negative path of integration is often left out to better mimic physical observations as there are no *negative frequencies* in real-world descriptions. For more background, we refer to [Pago06, Kre15].

6. Physical Circumstances

Fourier components, cf. Definition 1.4,

$$\psi(r_{\perp}, z) = \int_{\mathbb{R}^2} \widehat{\psi}(k_{\perp}, z) \exp(2\pi i \langle k_{\perp}, r_{\perp} \rangle) dk_{\perp}. \quad (6.3)$$

In combination with the Helmholtz equation, this yields a second-order differential equation in terms of z for $\widehat{\psi}(k_{\perp}, z)$. This problem has the general solution

$$\widehat{\psi}(k_{\perp}, z_1) = \widehat{\psi}(k_{\perp}, z_0) \exp\left(i(z_1 - z_0) \sqrt{\|k\|_2^2 - \|k_{\perp}\|_2^2}\right),$$

for wave vector k as before.

Plugging this into Equation (6.3), we see

$$\psi(r_{\perp}, z_1) = \int_{\mathbb{R}^2} \widehat{\psi}(k_{\perp}, z_0) \exp\left(i(z_1 - z_0) \sqrt{\|k\|_2^2 - \|k_{\perp}\|_2^2}\right) \exp(2\pi i \langle k_{\perp}, r_{\perp} \rangle) dk_{\perp}.$$

This is a propagator from z_0 to z_1 . For simplicity, let us assume that $z_0 = 0$. Then with $z = z_1$ we get

$$\psi(r_{\perp}, z) = \int_{\mathbb{R}^2} \widehat{\psi}(k_{\perp}, 0) \exp\left(iz \sqrt{\|k\|_2^2 - \|k_{\perp}\|_2^2}\right) \exp(2\pi i \langle k_{\perp}, r_{\perp} \rangle) dk_{\perp}.$$

In the case of X-ray imaging, we can assume that $\|k\|_2^2 \gg \|k_{\perp}\|_2^2$ which enables a Taylor expansion

$$\sqrt{\|k\|_2^2 - \|k_{\perp}\|_2^2} = \|k\|_2 - \frac{\|k_{\perp}\|_2^2}{2\|k\|_2} + \mathcal{O}(\|k_{\perp}\|_2^4).$$

Omission of the remainder leads to

$$\begin{aligned} \psi(r_{\perp}, z) &= \int_{\mathbb{R}^2} \widehat{\psi}(k_{\perp}, 0) \exp\left(iz \sqrt{\|k\|_2^2 - \|k_{\perp}\|_2^2}\right) \exp(2\pi i \langle k_{\perp}, r_{\perp} \rangle) dk_{\perp} \\ &\approx \int_{\mathbb{R}^2} \widehat{\psi}(k_{\perp}, 0) \exp\left(iz \left(\|k\|_2 - \frac{\|k_{\perp}\|_2^2}{2\|k\|_2}\right)\right) \exp(2\pi i \langle k_{\perp}, r_{\perp} \rangle) dk_{\perp} \\ &= \exp(i\|k\|_2 z) \int_{\mathbb{R}^2} \widehat{\psi}(k_{\perp}, 0) \exp\left(-iz \frac{\|k_{\perp}\|_2^2}{2\|k\|_2}\right) \exp(2\pi i \langle k_{\perp}, r_{\perp} \rangle) dk_{\perp} \\ &= -i \exp(i\|k\|_2 z) \int_{\mathbb{R}^2} \widehat{\psi}(k_{\perp}, 0) i \exp\left(-iz \frac{1}{\|k\|_2} \frac{\|k_{\perp}\|_2^2}{2}\right) \exp(2\pi i \langle k_{\perp}, r_{\perp} \rangle) dk_{\perp} \\ &= -i \exp(i\|k\|_2 z) \int_{\mathbb{R}^2} \widehat{\psi}(k_{\perp}, 0) i \exp\left(-iz \frac{\lambda}{2\pi} \frac{\|k_{\perp}\|_2^2}{2}\right) \exp(2\pi i \langle k_{\perp}, r_{\perp} \rangle) dk_{\perp} \end{aligned}$$

$$\begin{aligned}
&= -i \exp(i\|k\|_2 z) \int_{\mathbb{R}^2} \widehat{\psi}(k_\perp, 0) i \exp\left(-i\pi \frac{\lambda z}{4\pi^2} \|k_\perp\|_2^2\right) \exp(2\pi i \langle k_\perp, r_\perp \rangle) dk_\perp \\
&= -i \exp(i\|k\|_2 z) \int_{\mathbb{R}^2} \widehat{\psi}(k_\perp, 0) \widehat{K}_{\frac{\sqrt{\lambda z}}{2\pi}}(k_\perp) \exp(2\pi i \langle k_\perp, r_\perp \rangle) dk_\perp \\
&= -i \exp(i\|k\|_2 z) \left[\psi(\cdot, 0) * K_{\frac{\sqrt{\lambda z}}{2\pi}} \right](r_\perp)
\end{aligned}$$

with $K_\tau = \frac{1}{\tau^n} \exp\left(\frac{i\pi}{\tau^2} \|\cdot\|_2^2\right)$ as in Section 1.5 and we recall that $\|k\|_2 = 2\pi/\lambda$.

6.4 Definition.

We call the operator

$$\mathcal{D}_{\lambda, z}: \psi(\cdot, 0) \mapsto -i \exp(i\|k\|_2 z) \left[\psi(\cdot, 0) * K_{\frac{\sqrt{\lambda z}}{2\pi}} \right]$$

the *Fresnel Propagation Operator* by frequency λ onto distance z .

Often, the frequency is fixed for a given context, and we hence refer to the operator as \mathcal{D}_z .

From a mathematical point of view, the Fresnel propagation is a handy approximation as it consists of a convolution and a multiplication with a unitary factor with respect to the ℓ^2 -norm. Also the kernel exhibits a closed form both in spatial and frequency domain.

As a consequence, in many applications based on the propagation of measurable light, considered as an electromagnetic wave, the recovery problem is modelled as reconstructing a signal X from its Fresnel convolution $|K_\tau * X|^2$ or even its Fourier magnitudes $|\mathcal{F}(X)|^2$.

6.2. Phase Retrieval

While the Fresnel propagation operator \mathcal{D}_z as a convolution operator is comparatively manageable from a mathematical point of view, problems arise when this step is followed by phase loss. When modelling light as a complex wave $\psi: \mathbb{R} \rightarrow \mathbb{C}$, losing its phase information means a reduction on the non-negative real line $|\psi|^2: \mathbb{R} \rightarrow \mathbb{R}_+$. While in one dimension the ambiguities have been fully understood, [BP15], in higher dimensions this is very much incomprehensible.

Nanoscale imaging, e.g. of living cells, however is clearly two-dimensional. Hence we imperatively need additional priors to be integrated into recovery schemes. The realization of some of these recovery methods will be discussed in Chapter 7. Here we follow Looock in his presentation of [LP14] where he detailedly compares the results of the integration of different priors into his recovery scheme, which is based on DR λ , cf. Sections 2.3 and 7.1.

Next, we give a short reiteration on noise as inherent to nanoscale imaging. Due to the desired

6. Physical Circumstances



Figure 6.1.: Cell Image and Simulation²

low exposure of living material to X-rays, the imaging regime becomes low illuminated. When considering non-destructive imaging, in the biological sense, the power of the light source will be effectively reduced to a degree where single deviant photons will become measurable. This phenomenon is referred to as shot noise.

Exploitable Priors for Recovery

As this obviously makes the recovery problem ill-posed, we need to consider additional priors on the original wave to make inversion feasible. In our context of nanoscale X-ray imaging, there are a number of commonly used assumptions.

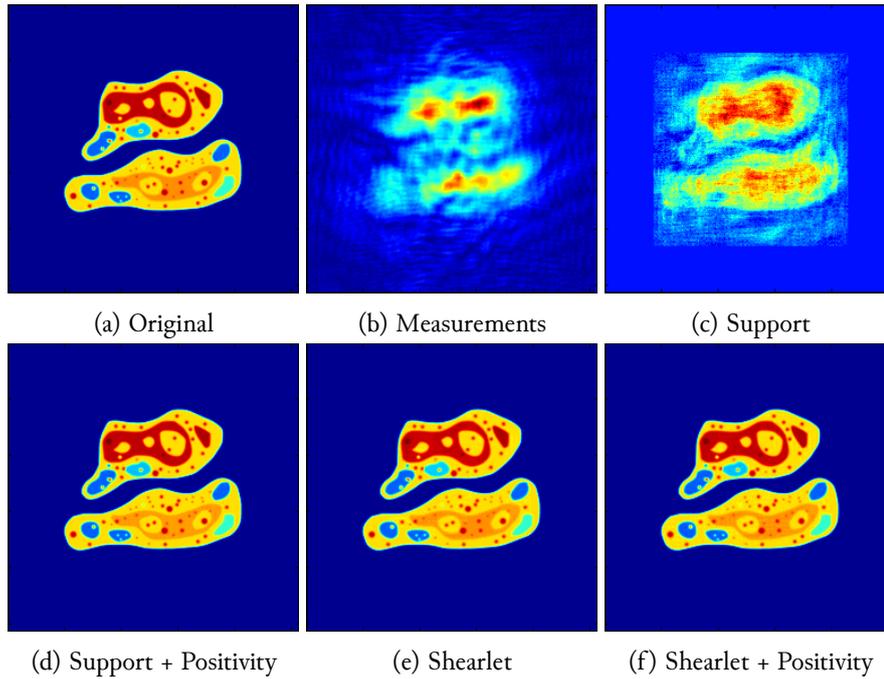
For the sake of simplicity of this part, let us exemplarily consider a test image, representing a living cell. In Figure 6.1 we see a cell cross section alongside a more organized, simulated version. From this, a number of characteristic properties comes to mind. The image of a single intact cell clearly is compact and often one can describe a rectangular area in the interior of which the cell is located. Further, different regions appear to be properly and smoothly bounded while their interiors are almost constant. From a computational perspective, it can also be desirable to scale the matrix encoding the 2D image such that all values lie in the interval $[0, 1]$ without altering any other properties.

This motivates to include the priors of compactness, support constraints, cartoon-likeness and positivity. Looock has done a comparison of these in [LP14], see Figure 6.2. He performed experiments without imposing noise with the objective to analyze the priors of a rectangular support set, positivity, shearlet sparsity. For the numerical details, we quote [LP14],³

The parameters for the Fresnel transform are $\lambda = 0.1 \text{ nm}$, $d = 100 \text{ mm}$, the pixel size is

²Images courtesy of Tim Salditt, Institute for X-ray Physics, Georg-August-Universität Göttingen

³I have adapted the references to point to the bibliography of this thesis as well as change units to SI standards.

Figure 6.2.: Recovery Priors⁴

$dx = 100$ nm and the images are 256×256 pixels in size. These parameters correspond to coherent imaging experiments using hard x-rays, see [KBS13]. In all cases, we apply [DR λ] with $N = 250$ iterations.

Figure 6.2a displays the synthetic cell image considered before and Figure 6.2b represents Looock’s simulated measurements of the magnitudes of the Fresnel transform. Figure 6.2c shows the recovered image with an inexact support constraint which clearly is insufficiently recovering the original image. Adding positivity as a constraint leads to the reconstruction in Figure 6.2d, which appears to be acceptable to the human eye. However, when using a shearlet sparsity constraint in Figure 6.2e, the obtained solution is very close to the original image. Still, adding the positivity constraint here as well improves the recovery further, as shown in Figure 6.2f. While differences between the images in the second row are hardly perceptible by the human eye, in [LP14] Looock supports the above findings by analyzing the distinct behavior of the corresponding methods with respect to the convergence rate of the reconstruction error.

This serves as a good example on how additional priors can be crucial to successful recovery without having to occupy ourselves too much with the underlying algorithmic scheme. We once more come back to the approach taken by Looock in Section 7.1 where we will look into the mathematical part of the phase retrieval process.

⁴Images courtesy of Stefan Looock, [LP14]

Noise

In application, one often is concerned about the power of the light source, especially in the X-ray regime, as too strong illumination might alter or even destroy a biological sample. Naturally, the process of measuring diffraction patterns with low illumination makes the obtained intensities dependent on the photon density per pixel, [KVS⁺14].

This accumulates to a difference from the theoretically modelled propagation and is considered a distortion. This phenomenon is commonly referred to as *Shot Noise* and can be modelled as Poisson noise due to its dependency on the signal intensity. There exists research on how to quantitatively predict such behavior or respectively reduce its impact, cf. [KVS⁺14, PMBBG20].

This can be counteracted when properties of the original data are known, as discussed above. Additionally to including these into the algorithm, there also exists the possibility to add a post-processing step to straighten out some further perturbations. This can be especially useful if there exists the tendency for disturbances that might be irrecognizable from the viewpoint of the algorithm but still discernible for the use case at hand.

When the end result is to be judged by a human eye, assessment by the Haar-PSI metric can provide insight into the amount of perceived discrepancy, [RBK⁺18]. With regards to content, one might consider Savitzky-Golay filters for smoothing, [KJK15], or promotion of cartoon-likeness, [RGM⁺19], to further close in on a perceivably acceptable result.

One might also want to consider to address the problem of denoising by application of learned dictionaries. Both the method of dictionary learning for image representation [AEB06] and application of learned dictionaries for image separation [LHX14, ID15] are well studied. There also exists research on improving this on piecewise stationary signals, [WZY12]. However, patterns of Fresnel-propagated cartoon-like images no longer adhere to easily describable patterns. Thus, as denoising in the measurement plane becomes infeasible, there is no noticeable benefit in the object plane as compared to, e.g., shearlet sparsity, as far as we are concerned.

7. Algorithmic Approaches

In this chapter we discuss two approaches for phase retrieval problems in the fashion of Chapter 6 by making use of the achievements of Part III.

First, we use the approach of [LP16], that is, alternating between a measurement set M and a suitable constraint set C with the help of an iterative procedure in Section 7.1. Here, we use a relaxed version of the Douglas-Rachford algorithm and improve the recovery scheme by upgrading from projecting onto a constraint *set* to a more involved constraining *proximity operator*.

Second, we model phase retrieval from Fourier magnitudes through lifting as a rank-one sparse matrix recovery problem in Section 7.2. We then analyze if that problem can be fed into Sparse Power Factorization to be solved.

7.1. Relaxed Douglas Rachford

In this section we investigate phase retrieval by means of a relaxed Douglas-Rachford algorithm (DR λ) by Luke, [Luko5, LM20]. Note that in earlier publications, e.g. [Luko5], this algorithm has been referred to as *Relaxed Averaged Alternating Reflections* (RAAR).

Adjusting for Phase Retrieval

The idea of applying an iterative projection, or reflection, scheme in phase-retrieval problems can be shortly summarized in the sense of [LP16, Loo16].

First, recall the propagation of light of Section 6.1 and the phase loss inherent to magnitudinal measurements. In the most simple adaptation this gives rise to a discrete recovery scheme. Denote $x_0 \in \mathbb{C}^d$ the vectorized illuminated object and $U \in \mathbb{C}^{d \times d}$ a unitary propagation operator, e.g. a digitized version of the Fresnel propagator, cf. Definition 6.4. Thus, the wavefront at the measurement plane can be described as Ux_0 . This leaves us with the task of modelling the information loss due to magnitudinal measurements. To that end, we denote by

$$|\cdot|_{\circ} : \mathbb{C}^d \longrightarrow \mathbb{R}_+^d$$

the pointwise modulus map and let $m := |Ux_0|_{\circ} \in \mathbb{R}_+^d$ the vectorized measured data.

7. Algorithmic Approaches

With that in mind we define the set of all possible specimens resulting in the obtained measurements

$$M := \{x \in \mathbb{C}^d : |Ux|_o = m\}.$$

As we have seen in Section 6.2, this is highly ambiguous, which is hardly surprising, given the model-inherent information loss.

Now Looock suggests that for application of the relaxed Douglas-Rachford algorithm introduced earlier, cf. Section 2.3, we choose the projection onto M as the first operator, i.e. letting $J_{\gamma B} := P_M$. For the other resolvent and in order to reduce the ambiguity introduced by M , Looock offers two choices of $J_{\gamma A}$. A simple choice would be to define a constraint set incorporating, realness, positivity and a support constraint, e.g.

$$C := \{x \in \mathbb{R}_+^d : \text{supp}(x) \subseteq D\},$$

for some bounded region D and then choose $J_{\gamma A}$ as the projection onto C .

As shown in Figure 6.2, a support constraint set defined as an rectangular region, is rather unsatisfactory and can be improved perspicuously. In a first step, one might consider to instead enforce sparsity. This means, replacing $P_C = \text{prox}_{\iota_C}$ by $\text{prox}_{\gamma \|\cdot\|_1}$, where the ℓ^1 -norm acts as a convex proxy for $\|\cdot\|_0$, which is neither a norm nor convex. As we have seen in Example 2.12, this yields the soft shrinkage operator, promoting sparsity as desired. However, as discussed in Chapter 4, naturally occurring data only tends to be sparse with respect to a suitable frame T rather than with respect to the canonical basis, cf., e.g., [Loo16, Sal19]. For example, in applications of medical imaging as considered by Looock, cf. [LP14] and Figure 6.2, shearlet sparsity may be desired as it promotes cartoon-like functions as discussed in Section 1.4. So, in a further improvement we change $J_{\gamma A}$ once more from $\text{prox}_{\gamma \|\cdot\|_1} = S_\gamma$ to $T^\dagger S_\gamma T$.

Extending the Proof of Looock

Now, to allow frame soft shrinkage, i.e. $T^\dagger S_\gamma T$, as the operator $J_{\gamma A}$ in the iterations of relaxed Douglas-Rachford, we need to show that there exists a resolvent of that form. Commonly this problem is approached by showing that frame soft shrinkage is a proximity operator. As a first step into this direction, Looock provides a proof of this goal for Parseval frames, cf. Definition 1.9.

7.1 Theorem (4.4.6 in [Loo16]).

Let $T \in \mathbb{R}^{n \times d}$, where $n \geq d$, be a tight frame, i.e. $T^*T = \nu \text{id}$ for some $\nu > 0$. Then the operator $T^\dagger S_\gamma T$ is the proximity operator of a plc function.

Note that due to a small detail in the proof this actually only holds true for Parseval frames, i.e. if $\nu = 1$. Honoring his approach, we will sketch his proof here.

Proof Sketch.

First, define a set-valued map

$$H: \mathbb{R}^d \longrightarrow 2^{\mathbb{R}^d}$$

$$x \longmapsto \{y : \nu T^* S_\gamma T(x + y) = x\}.$$

By lengthy, manual computation it can be shown that H is maximally cyclically monotone. Application of Theorem 2.8 yields existence of some $\Phi \in \Gamma_0(\mathbb{R}^d)$ such that $H = \partial\Phi$. Then by Corollary 2.15, we see that for $x, y \in \mathbb{R}^d$ we have $y \in H(x)$ if and only if

$$x = \text{prox}_\Phi(x + y).$$

Now the definition of H clearly states that $y \in H(x)$ if and only if

$$x = \nu T^* S_\gamma T(x + y).$$

Combining these two statements provides

$$T^\dagger S_\gamma T = \nu^{-1} \text{prox}_\Phi.$$

Now, for $\nu = 1$, the expression $\nu^{-1} \text{prox}_\Phi$ is clearly proximity operator itself without further assumptions. qed

Now, in [GP19], we were able to improve this idea and in Theorem 4.10 proved that the operator $T^\dagger S_\gamma T$ is in fact a proximity operator for any frame T by an adaptive choice of norm.

Outlook

Whilst not a topic of this thesis, one may also want to make the algorithm even more robust, especially with respect to noise. In the context of other recovery methods, this is done frequently by incorporating multiple measurement distances, cf. [WTB10, RS13]. While, e.g. Waller uses multiple distances in order to approximate higher order derivatives, we might consider an extended iteration scheme that projects onto all measuring planes in a cyclic manner. Serendipitously, Luke has recently extended the two-step DR λ algorithm to a multi-level cyclic procedure, [LM20].

This allows for example to iterate through multiple measurement sets M_1, \dots, M_n or constraint sets for achieving more stability when consistency has been lost because of noise. The topic of cyclic Douglas-Rachford schemes is also still actively investigated, cf. [BT13, BT14, Tam16, ACG18], so more adaptations for applications in phase-retrieval may be anticipated.

7.2. Matrix Recovery

In Chapter 5 we presented the SPF algorithm that can recover a set of doubly sparse vectors from their bilinear measurements. Being able to present sparse preimages of a bilinear operator is a powerful tool as generally this problem is really hard without further assumptions, [AKR18]. Fortunately, this tool provides numerous applications, one of the most prominent being multichannel blind deconvolution, cf. [LYB13].

We will now show that the same idea of lifting, cf. Section 3.1, can also be used to model the recovery from magnitude measurements, e.g. Fourier magnitudes. So, in order to formalize our goal we formulate our problem.

7.2 Problem.

Given a set of vectors $\{a_k\}_k \in \mathbb{C}^n$ and measurements $|\langle a_k, x \rangle|^2$, recover $x \in \mathbb{C}^n$.

With the help of lifting we reformulate our problem to a matrix recovery. To that end let $A_k := a_k a_k^* \in \mathbb{C}^{n \times n}$ and define

$$\begin{aligned} \mathcal{A}_k: \mathbb{C}^{n \times n} &\longrightarrow \mathbb{C} \\ X &\longmapsto \text{tr}(A_k X). \end{aligned}$$

Observe that now $|\langle a_k, x \rangle|^2 = \text{tr}(x^* a_k a_k^* x) = \text{tr}(A_k X)$ for $X = x x^*$. Now we reformulate the problem by applying the reasoning of Section 2.2 in [CESV13].

7.3 Problem.

Given a measurement $\mathcal{A}_k(X)$, find X subject to $\text{rk}(X) = 1$ and X positive semidefinite.

For the case of recovering from Fourier measurements, we recall the discrete Fourier transform of Section 1.1. For a signal $f = (f_j)_{j=1}^n \in \mathbb{C}^n$, the Fourier transform reads $\hat{f} = (\hat{f}_k)_{k=1}^n := F_n f$ for the Fourier matrix F_n . That is

$$\hat{f}_k = \sum_{j=1}^n f_j \exp(-2\pi i(j-1)(k-1)/n).$$

So, with $\omega_n := \exp(-2\pi i/n)$ we obtain

$$\hat{f}_k = \sum_{j=1}^n f_j \omega_n^{(j-1)(k-1)} = \left\langle f; \left(\omega_n^{(j-1)(k-1)} \right)_{j=1}^n \right\rangle.$$

To match our problem above we let $a_k = \left(\omega_n^{(j-1)(k-1)} \right)_{j=1}^n$ and $A_k = a_k a_k^*$. Then using $\mathcal{A}_k(X) = \text{tr}(A_k X)$ this extends to

$$\mathcal{A}: \mathbb{C}^{n \times n} \longrightarrow \mathbb{C}^n$$

$$X \longmapsto \begin{pmatrix} \text{tr}(A_1 X) \\ \text{tr}(A_2 X) \\ \vdots \\ \text{tr}(A_n X) \end{pmatrix}.$$

This allows us to contrast the original and lifted formulation as demonstrated in Figure 7.1.

Property	Original	Lifted
Signal	$x = (x_j)_{j=1}^n \in \mathbb{C}^n$	$X = xx^* = (x_j \bar{x}_k)_{j,k=1}^n \in \mathbb{C}^{n \times n}$
Measurement Functionals	$\langle a_k, \cdot \rangle$ for $a_k \in \mathbb{C}^n$	$\text{tr}(A_k \cdot)$ for $A_k \in \mathbb{C}^{n \times n}$
Measurements	$(\langle a_k, x \rangle)_{k=1}^n \in \mathbb{C}^n$	$(\text{tr}(A_k X))_{k=1}^n \in \mathbb{C}^n$

Figure 7.1.: Demonstrating the Lifting Process for Fourier Magnitudes

To now be able for SPF to recover X , we need \mathcal{A} to satisfy the conditions of Theorem 5.6 as, clearly, \mathcal{A} is not a Gaussian. This especially requires \mathcal{A} to fulfill a restricted isometry property, cf. Definition 3.2. Hence, we now investigate the relation of $\|\mathcal{A}(X)\|_2^2$ and $\|X\|_{\mathbb{F}}^2$ as their quotient is required to be bounded within $[1 - \delta, 1 + \delta]$ in order to satisfy the RIP of appropriate rank with constant δ . On the one hand, we see with $X = xx^*$ and $\hat{x} = (\hat{x}_j)_{j=1}^n := F_n x$,

$$\|\mathcal{A}(X)\|_2^2 = \sum_{j=1}^n |\text{tr}(A_j X)|^2 = \sum_{j=1}^n \left| |\langle a_j, x \rangle|^2 \right|^2 = \sum_{j=1}^n |\langle a_j, x \rangle|^4 = \sum_{j=1}^n |\hat{x}_j|^4 = \|\hat{x}\|_4^4$$

which in fact is the ℓ^4 -norm. This is troubling inasmuch as the Fourier matrix F_n is unitary, i.e., has ℓ^2 -norm one. This, however, will allow no generalization whatsoever for ℓ^4 .

On the other hand, we find

$$\|X\|_{\mathbb{F}}^2 = \sum_{j=1}^n \sum_{k=1}^n |X_{j,k}|^2 = \sum_{j=1}^n \sum_{k=1}^n (|x_j \bar{x}_k|)^2 = \sum_{j=1}^n \sum_{k=1}^n |x_j|^2 |x_k|^2 = \|x\|_2^2 \|x\|_2^2 = \|x\|_2^4.$$

Unfortunately, it is easy to see that the quotient $\|F_n x\|_4 / \|x\|_2$ is unbounded in general, making the RIP assumption unsatisfiable.

Alternative

As we have just seen, phase retrieval from Fourier magnitudes cannot be modelled as a problem satisfying any restricted isometry property. Clearly, this is due to the phase loss step which allows

components of the signal to amplify in an unbounded fashion.

There exist other approaches for realizing phase retrieval through matrix recovery without having to rely on an algorithm depending on an RIP. For example, in [CESV13] the authors detailedly present a recovery scheme using convex programming.

7.3. Further Approaches

Phase retrieval is a very active field of research for a multitude of different applications. Besides lenseless imaging considered in this thesis, this includes e.g. ptychography, [Sol19], or crystallography [Fri16]. Expectably, this also has motivated numerous recovery schemes.

One very noteworthy concept is that of F-splines, introduced in Section 1.5. They are unique in the sense that the encoding of signals with respect to F-splines allows for both backpropagation and denoising at the same time. When the unpropagated signal allows a sparse representation with respect to the unpropagated spline wavelet, so will their propagated counterparts. Hence, enforcing sparsity successfully within the diffraction patterns, will yield a cleaned up signal without having to iterate back and forth between object and measurement plane.

Transport of Intensity Equation

Another popular modelling approach is the so-called *Transport of Intensity Equation* (TIE). Without going into too much detail, we will describe this approximation of the propagation of electromagnetic waves based on [KBS13, Kre15].

The TIE is given as

$$\nabla_{\perp} (I \nabla_{\perp} \phi) = -\|k\|_2 \frac{\partial I}{\partial z},$$

where I denotes the wave's intensity, ϕ its phase and k a wave vector as before, e.g. in Equation (6.1), and ∇_{\perp} represents the gradient w.r.t. x and y direction. Thus, the TIE creates a relation between the signal's phase and intensity distributions.

In practice, the derivative in z -direction is approximated by a discrete derivative between (at least) two close measurement planes. As this obviously is extremely sensitive to noise, applications often consider an increased number of measuring distances. One such approach is used by e.g. [ZXX⁺12, ZCYA13, JCD⁺14].

Indeed numerical examples suggest that an increased number of measurements results in a subjectively better reconstructed image, as expected. It is, however, still an open research question how to choose the number and distances of the measurement planes and whether to choose them randomly or deterministically, cf. e.g. [JCD⁺14] for exponential spacing.

Part V.

Conclusion

In this thesis, I considered two approaches for feature extraction through promotion and enforcement of sparsity. Additionally, I presented an exemplary application setting in the context of diffractive imaging, resulting in phase retrieval problems.

Chapter 4 chose the way of nested proximity operators, i.e. operators of the form

$$T^\dagger \operatorname{prox}_f T$$

for an (overcomplete) representation system T and some plc prior f .

In Section 4.1, we showed that in finite dimensions, i.e. for a frame $T \in \mathbb{R}^{L \times N}$ with $L \geq N$ and full rank N and $f = \gamma \|\cdot\|_1$, i.e. $\operatorname{prox}_f = \mathcal{S}_\gamma$, the *Frame Soft Shrinkage* operator $T^\dagger \mathcal{S}_\gamma T$ is indeed a proximity operator. Key to this result was the consideration of a new, equivalent inner product defined via

$$\langle x, y \rangle_T := \langle Tx, Ty \rangle.$$

We also considered a toy example in Lemma 4.3 that allowed visualization, cf. Figure 4.1. This does however in turn raise the question, in which settings and for which parametrizations there exists some form of visual representation. Also, it would be interesting to know whether this is able to yield some form of helpful heuristic, i.e. an intuition on what to expect from the design of such an operator.

In Section 4.2, we advanced this idea to a much more general setting. Here we showed that the nested proximity operator $T^\dagger \operatorname{prox}_f T$ is again a proximity operator for any linear, injective operator T with closed range between arbitrary Hilbert spaces. This statement posed no further restrictions on f as long as prox_f exists, e.g. if f is plc. Again the key ingredient to this was the weighted inner product as above. Here, we also provided a formula for a prior g subject to $\operatorname{prox}_g = T^\dagger \operatorname{prox}_f T$, namely

$$g: x \mapsto f \square \left(\frac{1}{2} \|\cdot\|_{\mathcal{K}}^2 + \iota_{\mathcal{N}(T^*)} \right) (Tx).$$

While it does indeed look good to express g in terms of what looks like a closed expression, this in fact is not. It is not clear at all, in which cases there does really exist a closed form for g . In analogy to the proximity operator, which also only rarely possesses an explicit form, this can be bothersome in practice. It would be nice to gain an understanding, at least heuristically, on how this formula can practically be used and whether this result really is applicable in actual settings.

Chapter 5 was concerned with the Sparse Power Factorization algorithm (SPF). Here, we presented a relaxed sufficient criterion for successful recovery of both singular vectors u and v from their (noisy) bilinear measurement. Most notably, we relaxed the requirement on the mass distribution within the signal so that we no longer need to require half of the mass to be concentrated in a single entry. We were able to realize this by introducing a new norm

$$\|x\|_{[k]} := \max_{\substack{I \subseteq \{1, \dots, n\} \\ |I|=k}} \left(\sum_{i \in I} |x_i|^2 \right)^{1/2}$$

that, for $x \in \mathbb{C}^n$ and some $k \leq n$, quantifies the mass concentrated in the k biggest entries. This allowed us to formulate a tradeoff between the number of required measurements and the peakiness of the ground truth.

However, this result still leaves some questions unanswered, as we wrote in [GKS19]. For instance, our formula for the number of required measurements demands a very tight bound on the Noise-to-Signal Ratio ν , which in practice may be unsatisfiable. Also, we only considered adversarial noise and it might be reasonable to consider noise that adheres to a given distribution, e.g. classically Gaussian noise or Poisson noise as considered in Section 6.2. We also would like to broaden the range of applicability towards settings that do not use random matrices or not even RIP operators – which has exactly been the case in Section 7.2. Research in this direction may have progressed in recent years, cf. applications for blind deconvolution, e.g., [LJ15, LLJB17], but to my concern is far from resolving this issue.

Finally, in Chapter 7 we considered both above approaches in a phase retrieval setting from diffractive imaging. As already reported by Looock in [LP14, Loo16], the approach of Section 7.1 based on DR λ yields good results and exhibits acceptable numerical properties, e.g. noise robustness. Yet we would also like to know more about the dependency on the choice of number and placement of measurement planes. We also are interested to understand if the new result of [GP19], allowing sparsification with respect to arbitrary frames, could yield new recovery ideas. Also, we would be thrilled to reiterate the approach of Section 7.2 for an improved version of SPF that does no longer require a (strict) RIP.

Bibliography

- [ACG18] Francisco Javier Aragón Artacho, Yair Censor, and Aviv Gibali. The cyclic Douglas-Rachford algorithm with r -sets-Douglas-Rachford operators. *Optimization Methods and Software*, 34(4):875–889, August 2018.
- [AEB06] Michal Aharon, Michael Elad, and Alfred Bruckstein. K -SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation. *IEEE Transactions on Signal Processing*, 54(11):4311–4322, November 2006.
- [AKR18] Ali Ahmed, Felix Kraher, and Justin Romberg. Empirical Chaos Processes and Blind Deconvolution. *arXiv.org*, 2018.
- [AÖ17] Jonas Adler and Ozan Öktem. Solving ill-posed inverse problems using iterative deep neural networks. *Inverse Problems*, 33(12):124007, November 2017.
- [ARR14] Ali Ahmed, Benjamin Recht, and Justin Romberg. Blind Deconvolution Using Convex Programming. *IEEE Transactions on Information Theory*, 60(3):1711–1732, March 2014.
- [AW09] Arash A. Amini and Martin J. Wainwright. High-dimensional analysis of semidefinite relaxations for sparse principal components. *The Annals of Statistics*, 37(5B):2877–2921, October 2009.
- [BBCE09] Radu Balan, Bernhard G. Bodmann, Peter G. Casazza, and Dan Edidin. Painless Reconstruction from Magnitudes of Frame Coefficients. *Journal of Fourier Analysis and Applications*, 15(4):488–501, March 2009.
- [BC11] Heinz H. Bauschke and Patrick Louis Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer New York, 2011.
- [BCL04] Heinz H. Bauschke, Patrick Luis Combettes, and David Russell Luke. Finding best approximation pairs relative to two closed convex sets in Hilbert spaces. *Journal of Approximation Theory*, 127(2):178–192, April 2004.
- [BDDO17] Francesca Bartolucci, Filippo De Mari, Ernesto De Vito, and Francesca Odone. Shearlets as multi-scale Radon transform. In *2017 International Conference on Sampling Theory and Applications (SampTA)*, pages 625–629, July 2017.
- [Bec17] Amir Beck. Chapter 6: The Proximal Operator. In *First-Order Methods in Optimization*, pages 129–177. Society for Industrial and Applied Mathematics, October 2017.

Bibliography

- [Bei17] Robert Beinert. Ambiguities in one-dimensional phase retrieval from magnitudes of a linear canonical transform. *ZAMM*, 97(9):1078–1082, April 2017.
- [BGLW08] Heinz H. Bauschke, Rafal Goebel, Yves Lucet, and Xianfu Wang. The Proximal Average: Basic Theory. *SIAM Journal on Optimization*, 19(2):766–785, January 2008.
- [BJMO12] Francis Bach, Rodolphe Jenatton, Julien Mairal, and Guillaume Obozinski. Structured Sparsity through Convex Optimization. *Statistical Science*, 27(4):450–468, November 2012.
- [BM02] Matsuo Bashō and Asatarō Miyamori. *Classic Haiku: An Anthology of Poems by Bashō and His Followers*. Mineola, N.Y.: Dover, 2002.
- [BP13] Waheed U. Bajwa and Ali Pezeshki. Finite Frames for Sparse Signal Processing. In *Finite Frames*, pages 303–335. Birkhäuser Boston, 2013.
- [BP15] Robert Beinert and Gerlind Plonka. Ambiguities in One-Dimensional Discrete Phase Retrieval from Fourier Magnitudes. *Journal of Fourier Analysis and Applications*, 21(6):1169–1198, April 2015.
- [BR13] Quentin Berthet and Philippe Rigollet. Optimal detection of sparse principal components in high dimension. *The Annals of Statistics*, 41(4):1780–1815, August 2013.
- [BR16] Sohail Bahmani and Justin Romberg. Near-optimal estimation of simultaneously sparse and low-rank matrices from nested linear measurements. *Information and Inference*, 5(3):331–351, May 2016.
- [BR18] Sohail Bahmani and Justin Romberg. Solving Equations of Random Convex Functions via Anchored Regression. *Foundations of Computational Mathematics*, 19(4):813–841, October 2018.
- [BT13] Jonathan Michael Borwein and Matthew Kyle Tam. A Cyclic Douglas-Rachford Iteration Scheme. *Journal of Optimization Theory and Applications*, 160(1):1–29, August 2013.
- [BT14] Jonathan Michael Borwein and Matthew Kyle Tam. The Cyclic Douglas-Rachford Method for Inconsistent Feasibility Problems. *Journal of nonlinear and convex analysis*, 16(4):537–584, 2014.
- [CDS98] Scott Shaobing Chen, David Leigh Donoho, and Michael Alan Saunders. Atomic Decomposition by Basis Pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, January 1998.
- [CESV13] Emmanuel Jean Candès, Yonina C. Eldar, Thomas Strohmer, and Vladislav Voroninski. Phase Retrieval via Matrix Completion. *SIAM Journal on Imaging Sciences*, 6(1):199–225, January 2013.
- [Chr16] Ole Christensen. *An Introduction to Frames and Riesz Bases*. Springer International Publishing, 2016.

- [Chu92] Charles Kam-Tai Chui. *Introduction to Wavelets*. Elsevier Science, 1992.
- [CK13] Peter G. Casazza and Gitta Kutyniok, editors. *Finite Frames*. Birkhäuser Boston, 2013.
- [CKP13] Peter G. Casazza, Gitta Kutyniok, and Friedrich Philipp. Introduction to Finite Frame Theory. In *Finite Frames*, pages 1–53. Birkhäuser Boston, 2013.
- [CLS15] Emmanuel Jean Candès, Xiaodong Li, and Mahdi Soltanolkotabi. Phase Retrieval via Wirtinger Flow: Theory and Algorithms. *IEEE Transactions on Information Theory*, 61(4):1985–2007, April 2015.
- [CP20] Patrick Luis Combettes and Jean-Christophe Pesquet. Deep Neural Network Structures Solving Variational Inequalities. *Set-Valued and Variational Analysis*, 28(3):491–518, February 2020.
- [CPR13] Emilie Chouzenoux, Jean-Christophe Pesquet, and Audrey Repetti. Variable Metric Forward-Backward Algorithm for Minimizing the Sum of a Differentiable Function and a Convex Function. *Journal of Optimization Theory and Applications*, 162(1):107–132, November 2013.
- [CRT06] Emmanuel Jean Candès, Justin Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, February 2006.
- [CSV12] Emmanuel Jean Candès, Thomas Strohmer, and Vladislav Voroninski. PhaseLift: Exact and Stable Signal Recovery from Magnitude Measurements via Convex Programming. *Communications on Pure and Applied Mathematics*, 66(8):1241–1274, November 2012.
- [CT06] Thomas Merrill Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley, 2nd edition, 2006.
- [CW91] Charles Kam-Tai Chui and Jian-Zhong Wang. A cardinal spline approach to wavelets. *Proceedings of the American Mathematical Society*, 113(3):785–785, March 1991.
- [CWB08] Emmanuel Jean Candès, Michael B. Wakin, and Stephen P. Boyd. Enhancing Sparsity by Reweighted ℓ_1 Minimization. *Journal of Fourier Analysis and Applications*, 14(5-6):877–905, October 2008.
- [CWE17] Stanley H. Chan, Xiran Wang, and Omar A. Elgendy. Plug-and-Play ADMM for Image Restoration: Fixed-Point Convergence and Applications. *IEEE Transactions on Computational Imaging*, 3(1):84–98, March 2017.
- [CZ17] Zhihua Che and Xiaosheng Zhuang. Digital affine shear filter banks with 2-layer structure. In *2017 International Conference on Sampling Theory and Applications (SampTA)*, pages 575–579, July 2017.
- [Dau92] Ingrid Daubechies. *Ten lectures on wavelets*. Philadelphia, PA: SIAM, 1992.

Bibliography

- [dBE08] Alexandre d’Aspremont, Francis Bach, and Laurent El Ghaoui. Optimal Solutions for Sparse Principal Component Analysis. *Journal of Machine Learning Research*, 9:1269–1294, July 2008.
- [DDFG10] Ingrid Daubechies, Ronald DeVore, Massimo Fornasier, and C. Sinan Güntürk. Iteratively reweighted least squares minimization for sparse recovery. *Communications on Pure and Applied Mathematics*, 63(1):1–38, January 2010.
- [DKM⁺08] Stephan Dahlke, Gitta Kutyniok, Peter Maaß, Chen Sagiv, Hans-Georg Stark, and Gerd Teschke. The Uncertainty Principle Associated With the Continuous Shearlet Transform. *International Journal of Wavelets, Multiresolution and Information Processing*, 06(02):157–181, March 2008.
- [DM16] Yash Deshpande and Andrea Montanari. Sparse PCA via covariance thresholding. *The Journal of Machine Learning Research*, 17(1):4913–4953, January 2016.
- [Don95] David Leigh Donoho. De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3):613–627, May 1995.
- [DP18] Minh Ngoc Dao and Hung M. Phan. Linear convergence of the generalized Douglas-Rachford algorithm for feasibility problems. *Journal of Global Optimization*, 72(3):443–474, April 2018.
- [Elao6] Michael Elad. Why simple shrinkage is still relevant for redundant representations? *IEEE Transactions on Information Theory*, 52(12):5559–5569, December 2006.
- [ELL08] Glenn R. Easley, Demetrio Labate, and Wang-Q Lim. Sparse directional image representations using the discrete shearlet transform. *Applied and Computational Harmonic Analysis*, 25(1):25–46, 2008.
- [ET87] Ivar Ekeland and Roger Témam. *Convex Analysis and Variational Problems*. SIAM, 1987.
- [FLS10] Richard Phillips Feynman, Robert Benjamin Leighton, and Matthew Sands. *The Feynman lectures on physics; New millennium ed.* Basic Books, 2010. Originally published 1963 – 1965.
- [FMN21] Massimo Fornasier, Johannes Maly, and Valeriya Naumova. Robust recovery of low-rank matrices with non-orthogonal sparse decomposition from incomplete measurements. *Applied Mathematics and Computation*, 392:125702, March 2021.
- [Fou11] Simon Foucart. Hard Thresholding Pursuit: An Algorithm for Compressive Sensing. *SIAM Journal on Numerical Analysis*, 49(6):2543–2563, January 2011.
- [FR13] Simon Foucart and Holger Rauhut. *A mathematical introduction to compressive sensing*. New York, NY: Birkhäuser/Springer, 2013.
- [Fri16] Gero Friesecke. Die Schönheit der molekularen Architektur von Viren und ihre Entschlüsselung durch Beugungsmuster. *Mitteilungen der Deutschen Mathematiker-Vereinigung*, 24(4):208–213, 2016.

- [FSGo8] Béla A. Frigyik, Santosh Srivastava, and Maya R. Gupta. An Introduction to Functional Derivatives. techreport UWEETR-2008-0001, Department of Electrical Engineering, University of Washington, July 2008.
- [Gep15] Jakob Alexander Geppert. Different Pursuit Algorithms in the Iterations of Sparse Power Factorisation. Bachelor's thesis, Georg-August-Universität Göttingen, September 2015.
- [Gep17] Jakob Alexander Geppert. Refined Performance Guarantees for Sparse Power Factorization. Master's thesis, Georg-August-Universität Göttingen, January 2017.
- [GKS17] Jakob Alexander Geppert, Felix Kraemer, and Dominik Stöger. Refined performance guarantees for Sparse Power Factorization. In *2017 International Conference on Sampling Theory and Applications (SampTA)*. IEEE, July 2017.
- [GKS19] Jakob Alexander Geppert, Felix Kraemer, and Dominik Stöger. Sparse Power Factorization: Balancing peakiness and sample complexity. *Advances in Computational Mathematics*, 45(3):1711–1728, June 2019.
- [GN20] Rémi Gribonval and Mila Nikolova. A Characterization of Proximity Operators. *Journal of Mathematical Imaging and Vision*, 62(6-7):773–789, March 2020.
- [Goo05] Joseph W. Goodman. *Introduction to Fourier optics*. Roberts, 3rd edition, 2005.
- [GP19] Jakob Alexander Geppert and Gerlind Plonka. Frame Soft Shrinkage Operators are Proximity Operators. *arXiv.org*, 2019.
- [GV96] Gene Howard Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, 1996.
- [Hay94] Simon Haykin. *Blind Deconvolution*. Prentice Hall, 1994.
- [HHN⁺20] Marzieh Hasannasab, Johannes Hertrich, Sebastian Neumayer, Gerlind Plonka, Simon Setzer, and Gabriele Steidl. Parseval Proximal Neural Networks. *Journal of Fourier Analysis and Applications*, 26(4), July 2020.
- [HNP⁺19] Marzieh Hasannasab, Sebastian Neumayer, Gerlind Plonka, Simon Setzer, Gabriele Steidl, and Jakob Alexander Geppert. Frame Soft Shrinkage as Proximity Operator. *arXiv.org*, 2019.
- [ID15] Paul Irofti and Bogdan Dumitrescu. Overcomplete Dictionary Design: The Impact of the Sparse Representation Algorithm. In *2015 20th International Conference on Control Systems and Computer Science*, pages 901–908, May 2015.
- [IWW17] Mark Iwen, Aditya Viswanathan, and Yang Wang. Robust sparse phase retrieval made easy. *Applied and Computational Harmonic Analysis*, 42(1):135–142, January 2017.
- [JCD⁺14] Zhong Jingshan, Rene Andre Claus, Justin Dauwels, Lei Tian, and Laura Waller. Transport of Intensity phase imaging by intensity spectrum fitting of exponentially spaced defocus planes. *Optics Express*, 22(9):10661, April 2014.

Bibliography

- [JKS18] Peter Jung, Felix Krahmer, and Dominik Stoger. Blind Demixing and Deconvolution at Near-Optimal Rate. *IEEE Transactions on Information Theory*, 64(2):704–727, February 2018.
- [JNRS10] Michel Journée, Yurii Nesterov, Peter Richtárik, and Rodolphe Sepulchre. Generalized Power Method for Sparse Principal Component Analysis. *Journal of Machine Learning Research*, 11:517–553, February 2010.
- [JNS13] Prateek Jain, Praneeth Netrapalli, and Sujay Sanghavi. Low-rank matrix completion using alternating minimization. In *Proceedings of the 45th annual ACM symposium on Symposium on theory of computing - STOC '13*. ACM Press, 2013.
- [KBS13] Martin Krenkel, Matthias Bartels, and Tim Salditt. Transport of intensity phase reconstruction to solve the twin image problem in holographic x-ray imaging. *Optics Express*, 21(2):2220–2235, January 2013.
- [KJK15] Jeong-Hwan Kim, Gyeo-Wun Jeung, and Kyeong-Seop Kim. Image Smoothing by Pseudo-2D Savitzky-Golay Filter. In Xiaodong Zhuang, editor, *Recent Advantages of Computer Engineering*, Seoul, September 2015. 14th International Conference on Applications of Computer Engineering.
- [KK98] Karl-Dirk Kammeyer and Kristian Kroschel. *Digitale Signalverarbeitung*. Vieweg+Teubner Verlag, 1998.
- [KK17] Michael Kech and Felix Krahmer. Optimal Injectivity Conditions for Bilinear Inverse Problems with Applications to Identifiability of Deconvolution Problems. *SIAM Journal on Applied Algebra and Geometry*, 1(1):20–37, January 2017.
- [KKL12] Pisamai Kittipoom, Gitta Kutyniok, and Wang-Q Lim. Construction of Compactly Supported Shearlet Frames. *Constructive Approximation*, 35(1):21–72, 2012.
- [KL12] Gitta Kutyniok and Demetrio Labate. *Introduction to Shearlets*, pages 1–38. Birkhäuser Boston, Boston, 2012.
- [KLR16] Gitta Kutyniok, Wang-Q Lim, and Rafael Reisenhofer. ShearLab 3D: Faithful Digital Shearlet Transforms based on Compactly Supported Shearlets. *ACM Transactions on Mathematical Software*, 42(1):42, 2016.
- [KLZ12] Gitta Kutyniok, Wang-Q Lim, and Xiaosheng Zhuang. *Digital Shearlet Transforms*, pages 239–282. Birkhäuser Boston, Boston, 2012.
- [KNV15] Robert Krauthgamer, Boaz Nadler, and Dan Vilenchik. Do semidefinite relaxations solve sparse PCA up to the information limit? *The Annals of Statistics*, 43(3):1300–1322, June 2015.
- [Kre15] Martin Krenkel. *Cone-beam x-ray phase-contrast tomography for the observation of single cells in whole organs*. PhD thesis, Georg-August University School of Science, October 2015.

- [KVS⁺14] Ahmed Kirmani, Dheera Venkatraman, Dongeek Shin, Andrea Colaço, Franco N. C. Wong, Jeffrey H. Shapiro, and Vivek K Goyal. First-Photon Imaging. *Science*, 343(6166):58–61, 2014.
- [KW11] Felix Krahmer and Rachel Ward. New and Improved Johnson-Lindenstrauss Embeddings via the Restricted Isometry Property. *SIAM Journal on Mathematical Analysis*, 43(3):1269–1281, January 2011.
- [LBU03] Michael Liebling, Thierry Blu, and Michael Unser. Fresnelets: New Multiresolution Wavelet Bases for Digital Holography. *IEEE Transactions on Image Processing*, 12(1):29–43, January 2003.
- [Lee13] John Marshall Lee. *Introduction to smooth manifolds*. New York, NY: Springer, 2nd revised edition, 2013.
- [LHX14] Shuaiqi Liu, Shaohai Hu, and Yang Xiao. Image separation using wavelet-complex shearlet dictionary. *Journal of Systems Engineering and Electronics*, 25(2):314–321, April 2014.
- [Lio04] Michael Liebling. On Fresnelets, Interference Fringes, and Digital Holography. *EPFL*, page 151, 2004. Sélectionné pour le "Prix EPFL de doctorats 2004" avec mention spéciale - "EPFL doctorate award 2004" distinction nominee.
- [Lim10] Wang-Q Lim. The Discrete Shearlet Transform: A New Directional Transform and Compactly Supported Shearlet Frames. *IEEE Transactions on Image Processing*, 19(5):1166–1180, May 2010.
- [LJ15] Kiryung Lee and Marius Junge. RIP-like Properties in Subsampled Blind Deconvolution. *arXiv.org*, 2015.
- [LKR18] Kiryung Lee, Felix Krahmer, and Justin Romberg. Spectral Methods for Passive Imaging: Nonasymptotic Performance and Robustness. *SIAM Journal on Imaging Sciences*, 11(3):2110–2164, January 2018.
- [LLJB17] Kiryung Lee, Yanjun Li, Marius Junge, and Yoram Bresler. Blind Recovery of Sparse Signals From Subsampled Convolution. *IEEE Transactions on Information Theory*, 63(2):802–821, February 2017.
- [LLSW19] Xiaodong Li, Shuyang Ling, Thomas Strohmer, and Ke Wei. Rapid, robust, and reliable blind deconvolution via nonconvex optimization. *Applied and Computational Harmonic Analysis*, 47(3):893–934, November 2019.
- [LM79] Pierre-Louis Lions and B. Mercier. Splitting Algorithms for the Sum of Two Nonlinear Operators. *SIAM Journal on Numerical Analysis*, 16(6):964–979, December 1979.
- [LM20] David Russell Luke and Anna-Lena Martins. Convergence Analysis of the Relaxed Douglas-Rachford Algorithm. *SIAM Journal on Optimization*, 30(1):542–584, January 2020.

Bibliography

- [LMR94] Alfred Karl Louis, Peter Maaß, and Andreas Rieder. *Wavelets*. Vieweg+Teubner Verlag, 1994.
- [Loo16] Stefan Loock. *Phase Retrieval with Sparsity Constraints*. PhD thesis, Georg-August University School of Science, July 2016.
- [LP14] Stefan Loock and Gerlind Plonka. Phase retrieval for Fresnel measurements using a shearlet sparsity constraint. *Inverse Problems*, 30(5):055005, 2014.
- [LP16] Stefan Loock and Gerlind Plonka. Iterative Phase Retrieval with Sparsity Constraints. *Proceedings in Applied Mathematics and Mechanics*, 16(1):835–836, 2016.
- [LS15] Shuyang Ling and Thomas Strohmer. Self-calibration and biconvex compressive sensing. *Inverse Problems*, 31(11):115002, September 2015.
- [LS17] Shuyang Ling and Thomas Strohmer. Blind Deconvolution Meets Blind Demixing: Algorithms and Performance Bounds. *IEEE Transactions on Information Theory*, 63(7):4497–4520, July 2017.
- [LS18] Shuyang Ling and Thomas Strohmer. Regularized gradient descent: a non-convex recipe for fast joint blind deconvolution and demixing. *Information and Inference: A Journal of the IMA*, 8(1):1–49, March 2018.
- [Luk05] David Russell Luke. Relaxed averaged alternating reflections for diffraction imaging. *Inverse Problems*, 21(1):37, 2005.
- [LWB13] Kiryung Lee, Yihong Wu, and Yoram Bresler. Near Optimal Compressed Sensing of Sparse Rank-One Matrices via Sparse Power Factorization. *arXiv.org*, 2013.
- [LWB18] Kiryung Lee, Yihong Wu, and Yoram Bresler. Near-Optimal Compressed Sensing of a Class of Sparse Low-Rank Matrices Via Sparse Power Factorization. *IEEE Transactions on Information Theory*, 64(3):1666–1698, March 2018.
- [LYB13] Kiryung Lee, Elad Yarkony, and Yoram Bresler. Subsampled Multichannel Blind Deconvolution by Sparse Power Factorization. Technical report, University of Illinois at Urbana-Champaign, September 2013.
- [LZ17] Ji Li and Tie Zhou. On relaxed averaged alternating reflections (RAAR) algorithm for phase retrieval with structured illumination. *Inverse Problems*, 33(2):025012, January 2017.
- [Mai13] Zongming Ma. Sparse principal component analysis and iterative thresholding. *The Annals of Statistics*, 41(2):772–801, April 2013.
- [Mal09] Stéphane Georges Mallat. *A Wavelet Tour of Signal Processing*. Elsevier, 2009.
- [Mat98] Paul Charles Matthews. *Vector Calculus*. Springer London, 1998.
- [Mix15] Dustin G. Mixon. Explicit Matrices with the Restricted Isometry Property: Breaking the Square-Root Bottleneck. In *Compressed Sensing and its Applications*, pages 389–417. Springer International Publishing, 2015.

- [Mor65] Jean Jacques Moreau. Proximité et dualité dans un espace hilbertien. *Bulletin de la Société Mathématique de France*, 79:273–299, 1965.
- [MP10] Jianwei Ma and Gerlind Plonka. The Curvelet Transform. *IEEE Signal Processing Magazine*, 27(2):118–133, March 2010.
- [MS79] Reinhard Mennicken and Boris Sagraloff. On Banach’s closed range theorem. *Archiv der Mathematik*, 33(1):461–465, December 1979.
- [MZ93] Stéphane Georges Mallat and Zhifeng Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12):3397–3415, 1993.
- [New05] Mark E. J. Newman. Power laws, Pareto distributions and Zipf’s law. *Contemporary Physics*, 46(5):323–351, September 2005.
- [Nic17] Peter Nickolas. *Wavelets. A student guide.*, volume 24. Cambridge: Cambridge University Press, 2017.
- [NK03] Michael P. Norton and Denis G. Karczub. *Fundamentals of Noise and Vibration Analysis for Engineers*. Cambridge University Press, 2003.
- [NT09] Deanna Needell and Joel Aaron Tropp. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, 26(3):301–321, May 2009.
- [Och18] Peter Ochs. Accelerations of Forward-Backward Splitting. In *Computational Methods for Inverse Problems in Imaging*, Como, June 2018. INDAM.
- [OJF⁺15] Samet Oymak, Amin Jalali, Maryam Fazel, Yonina C. Eldar, and Babak Hassibi. Simultaneously Structured Models With Application to Sparse and Low-Rank Matrices. *IEEE Transactions on Information Theory*, 61(5):2886–2908, May 2015.
- [Øyv19] Ryan Øyvind. *Linear Algebra, Signal Processing, and Wavelets – A Unified Approach*, chapter Using Tensor Products to Apply Wavelets to Images, pages 317–342. Springer, Cham, 2019.
- [Pag06] David Paganin. *Coherent X-Ray Optics*. Oxford University Press, January 2006.
- [Phe93] Robert Ralph Phelps. *Convex Functions, Monotone Operators and Differentiability*. Springer, 1993.
- [PMBBG20] Minxu Peng, John Murray-Bruce, Karl K. Berggren, and Vivek K Goyal. Source shot noise mitigation in focused ion beam microscopy by time-resolved measurement. *Ultramicroscopy*, 211:112948, April 2020.
- [PPST18] Gerlind Plonka, Daniel Potts, Gabriele Steidl, and Manfred Tasche. *Numerical Fourier Analysis*. Springer International Publishing, 2018.
- [QZEW20] Qing Qu, Yuqian Zhang, Yonina C. Eldar, and John Wright. Convolutional Phase Retrieval via Gradient Descent. *IEEE Transactions on Information Theory*, 66(3):1785–1821, March 2020.

Bibliography

- [RBK⁺18] Rafael Reisenhofer, Sebastian Bosse, Gitta Kutyniok, Thomas Wiegand, Rafael Reisenhofer, Sebastian Bosse, Gitta Kutyniok, and Thomas Wiegand. A Haar Wavelet-Based Perceptual Similarity Index for Image Quality Assessment. *Signal Processing: Image Communication*, 61:33–43, 2018.
- [RFP13] Hugo Raguét, Jalal Fadili, and Gabriel Peyré. A Generalized Forward-Backward Splitting. *SIAM Journal on Imaging Sciences*, 6(3):1199–1226, January 2013.
- [RGM⁺19] Robin Richter, Carsten Gottschlich, Lucas Mentch, Duy H. Thai, and Stephan F. Huckemann. Smudge Noise for Quality Estimation of Fingerprints and its Validation. *IEEE Transactions on Information Forensics and Security*, 14(8):1963–1974, August 2019.
- [Roc66] Ralph Tyrrell Rockafellar. Characterization of the subdifferentials of convex functions. *Pacific Journal of Mathematics*, 17(3):497–510, June 1966.
- [RS08] Pauliina Remes and Juha Sihvola, editors. *Ancient Philosophy of the Self*. Springer Netherlands, 2008.
- [RS13] Anna-Lena Robisch and Tim Salditt. Phase retrieval for object and probe using a series of defocus near-field images. *Optics Express*, 21(20):23345–23357, October 2013.
- [Rud91] Walter Rudin. *Functional Analysis*. New York, NY: McGraw-Hill, 2nd edition, 1991.
- [RW98] Ralph Tyrrell Rockafellar and Roger Jean-Baptiste Wets. *Variational analysis*. Berlin: Springer, 1998.
- [Rya02] Raymond A. Ryan. *Introduction to Tensor Products of Banach Spaces*. Springer London, 2002.
- [Sal19] Palina Salanevich. Stability of Phase Retrieval Problem. In *2019 International Conference on Sampling Theory and Applications (SampTA)*, 2019.
- [Sch14] Jonathan Schaffer. What Not to Multiply Without Necessity. *Australasian Journal of Philosophy*, 93(4):644–664, December 2014.
- [Sed18] Soraya Sedkaoui. *Data Analytics and Big Data*. John Wiley & Sons, Inc., May 2018.
- [Sek57] Takejiro Seki. An elementary proof of Brouwer’s fixed point theorem. *Toboku Mathematical Journal*, 9(2):105–109, 1957.
- [Set10] Simon Setzer. Operator Splittings, Bregman Methods and Frame Shrinkage in Image Processing. *International Journal of Computer Vision*, 92(3):265–280, July 2010.
- [SGK18] Dominik Stöger, Jakob Alexander Geppert, and Felix Krahmer. Sparse Power Factorization With Refined Peakiness Conditions. In *2018 IEEE Statistical Signal Processing Workshop (SSP)*. IEEE, June 2018.
- [SN96] Gilbert Strang and Truong Quang Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, Wellesley, MA, 2nd edition, 1996.

- [Sol19] Mahdi Soltanolkotabi. Structured Signal Recovery From Quadratic Measurements: Breaking Sample Complexity Barriers via Nonconvex Optimization. *IEEE Transactions on Information Theory*, 65(4):2374–2400, April 2019.
- [STW18] Hayden Schaeffer, Giang Tran, and Rachel Ward. Extracting Sparse High-Dimensional Dynamics from Limited Data. *SIAM Journal on Applied Mathematics*, 78(6):3279–3295, January 2018.
- [Sug90] Mitsuo Sugiura. *Unitary representations and harmonic analysis. An introduction.*, volume 44. Amsterdam etc.: North-Holland; Tokyo: Kodansha Ltd., 2nd edition, 1990.
- [SVW⁺16] Suhas Sreehari, Singanallur V. Venkatakrishnan, Brendt Wohlberg, Gregory T. Buzzard, Lawrence F. Drummy, Jeffrey P. Simmons, and Charles A. Bouman. Plug-and-Play Priors for Bright Field Electron Tomography and Sparse Interpolation. *IEEE Transactions on Computational Imaging*, 2016.
- [Tam16] Matthew Kyle Tam. Iterative Projection and Reflection Methods: Theory and Practice. *Bulletin of the Australian Mathematical Society*, 94(1):175–176, May 2016.
- [TBDF18] Afonso M. Teodoro, José M. Bioucas-Dias, and Mário A. T. Figueiredo. Scene-Adapted Plug-and-Play Algorithm with Guaranteed Convergence: Applications to Data Fusion in Imaging. *arXiv.org*, 2018.
- [TP14] Andreas M. Tillmann and Marc Emanuel Pfetsch. The Computational Complexity of the Restricted Isometry Property, the Nullspace Property, and Related Concepts in Compressed Sensing. *IEEE Transactions on Information Theory*, 60(2):1248–1259, February 2014.
- [Ver12] Roman Vershynin. Introduction to the non-asymptotic analysis of random matrices. In Yonina C. Eldar and Gitta Kutyniok, editors, *Compressed Sensing: Theory and Applications*, pages 210–268. Cambridge University Press, 2012.
- [vGo3] Johann Wolfgang von Goethe. *Faust Parts I & II*. Poetry In Translation, 2003. Translated by Anthony S. Kline.
- [WBS16] Tengyao Wang, Quentin Berthet, and Richard John Samworth. Statistical and computational trade-offs in estimation of sparse principal components. *The Annals of Statistics*, 44(5):1896–1930, October 2016.
- [Won11] Man Wah Won. *Discrete Fourier Analysis*. Springer-Verlag GmbH, 2011.
- [WTB10] Laura Waller, Lei Tian, and George Barbastathis. Transport of Intensity imaging with higher order derivatives. *Optics Express*, 18(12):12552, May 2010.
- [WZY12] Tian Jing Wang, Bao Yu Zheng, and Zhen Yang. An overcomplete dictionary design algorithm for sparse representation of piecewise stationary signals. In *2012 18th Asia-Pacific Conference on Communications (APCC)*, pages 427–430, October 2012.

Bibliography

- [XLTK95] Guanghan Xu, Hui Liu, Lang Tong, and Thomas Kailath. A least-squares approach to blind channel identification. *IEEE Transactions on Signal Processing*, 43(12):2982–2993, 1995.
- [ZCYA13] Chao Zuo, Qian Chen, Yingjie Yu, and Anand Asundi. Transport-of-intensity phase imaging using Savitzky-Golay differentiation filter - theory and applications. *Optical Express*, 21(5):5346–5362, March 2013.
- [ZS14] Ahmed I. Zayed and Gerhard Schmeisser, editors. *New Perspectives on Approximation and Sampling Theory*. Springer International Publishing, 2014.
- [ZXX⁺12] Shiling Zheng, Bindang Xue, Wenfang Xue, Xiangzhi Bai, and Fugen Zhou. Transport of intensity phase imaging from multiple noisy intensities measured in unequally-spaced planes. *Optics Express*, 20(2):972, January 2012.
- [岡藤伏⁺14] 岡田知久, 藤本晃司, 伏見育崇, 山本憲, and 富樫かおり. 医用画像における圧縮センシングの現状と課題. 数理解析研究所講究録, 1928:89–91, December 2014.