

ABSTRACT

Title of dissertation: TOPICS IN HARMONIC ANALYSIS,
SPARSE REPRESENTATIONS, AND
DATA ANALYSIS

Weilin Li
Doctor of Philosophy, 2018

Dissertation directed by: Professor John J. Benedetto
Department of Mathematics

Professor Wojciech Czaja
Department of Mathematics

Classical harmonic analysis has traditionally focused on linear and invertible transformations. Motivated by modern applications, there is a growing interest in non-linear analysis and synthesis operators. This thesis encompasses applications of computational harmonic analysis, with a strong emphasis on time-frequency methods, to modern problems arising in deep learning, data analysis, imaging, and signal processing.

The first focus of this thesis deals with scattering transforms, which are particular realizations of convolutional neural networks. While the latter uses trained convolution kernels, scattering transforms use fixed ones, and this simplification allows mathematicians to develop a model of deep learning. Mallat originally introduced a wavelet scattering transform, but we study a complementary Fourier based version. We prove that the Fourier scattering transform enjoys properties

that make it an effective feature extractor for classification, and we also construct a rotationally invariant modification of this transform. We provide experimental evidence that shows its effectiveness at representing complicated spectral data.

The second focus of this thesis pertains to the mathematical foundations of super-resolution, which is concerned with the recovery of fine details from low-resolution observations. This imaging model can be mathematically formulated as an ill-posed inverse problem in the space of bounded complex measures. While the current theory primarily deals with the recovery of discrete measures with minimum separation greater than the Rayleigh length, we present alternative approaches. One direction exploits Beurling's results on minimal extrapolation to obtain a general theory that is pertinent to a wide class of measures, including those with geometric structure. Another approach is information theoretic and studies the min-max error for robust super-resolution of discrete measures below the Rayleigh length.

TOPICS IN HARMONIC ANALYSIS, SPARSE
REPRESENTATIONS, AND DATA ANALYSIS

by

Weilin Li

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2018

Advisory Committee:
Professor John J. Benedetto, Chair
Professor Wojciech Czaja, Co-Chair
Professor Rama Chellappa
Professor Tom Goldstein
Professor Kasso Okoudjou

© Copyright by
Weilin Li
2018

Dedication

To my parents, for their unwavering love and support.

Acknowledgments

First and foremost, I owe my gratitude to my two advisors, Professors John Benedetto and Wojciech Czaja. Over these five years, they provided me with valuable insights and suggestions, and gave me numerous opportunities. They introduced me to many interesting applications of harmonic analysis, and encouraged me to branch out and become interested in interdisciplinary fields. Much of my success can be attributed to their careful guidance, support, and wisdom.

Several other professors also significantly influenced my graduate career. Professors Kasso Okoudjou and Radu Balan were generous with their time, and were always enthusiastic to discuss mathematics. Kasso also allowed me to join him while he was in residence at the Mathematical Sciences Research Institute during the Spring 2017 semester, and my participation in that program was beneficial towards my development as a mathematician. Professors Matei Machedon, Maria Cameron, and Eitan Tadmor introduced me to several interesting topics and taught foundational classes during my first year. My collaborator, Professor Wenjing Liao, is a constant source of encouragement, and is a fantastic mentor and friend.

Many thanks to my friends and colleagues who have made my graduate career memorable: Dimitrios Ntoggas, Hamid Al-Saqban, Dana Botesteanu, Jacky Chong, Laura Iosip, Mark Magsino, Minsung Kim, Cheng Jie, Franck Ndjakou Njeunje, Yiran Li, Xin Dong, Ilya Kavalero, and Mike Pekala. Also many thanks to the former Norbert Wiener Center graduate students and postdocs who have given me valuable guidance: Matt Begué, Chae Clark, Alex Cloninger, Paul Kaprowski, Joey

Iverson, Ben Manning, James Murphy, Kara Yacoubou Djima, and Dongmian Zou.

Some experiments in this thesis involve radiological data. These datasets were collected and provided to us by Lance McLean and Thomas McCullough, both from the Department of Energy, and by John Bird, Kevin Kochersberger, and John Peterson, all three from Virginia Polytechnic Institute.

Table of Contents

Dedication	ii
Acknowledgements	iii
List of Figures	vii
List of Abbreviations	viii
1 Introduction	1
1.1 Harmonic Analysis	1
1.2 Modern Applications	4
1.3 Outline	7
2 Background on Scattering Transforms	9
2.1 Scattering Transforms	9
2.2 Uniform Covering Frames	12
2.3 Fourier Scattering Transform	17
3 Properties of Fourier Scattering	21
3.1 Energy Concentration	21
3.2 Fourier Scattering Properties	29
3.3 Comparison and Discussion	37
4 Rotational Fourier Scattering	43
4.1 Rotational UCF	43
4.2 Rotational Fourier Scattering	49
4.3 Rotational Invariance and Properties	54
4.4 Relationship to Directional Representations	58
5 Applications of Fourier Scattering	63
5.1 The Discretization Problem	63
5.2 Fast Fourier Scattering Transform	69
5.3 Image Feature Comparison	71
5.4 Anomaly Detection in Radiological Data	74

6	Background on Super-Resolution	80
6.1	Terminology	80
6.2	Convex Approach	84
6.3	Subspace Methods	88
7	Beurling Theory of Super-Resolution	91
7.1	Problem Statement	91
7.2	Some Functional Analysis Results	93
7.3	The Beurling Theory	100
7.4	Uniqueness and Non-uniqueness	106
7.5	Additional Results	112
7.6	Examples	116
7.7	Conclusion	126
8	Stable Super-Resolution Limit	129
8.1	Problem Statement	129
8.2	Characterizations	132
8.3	Sparse Lagrange Polynomials	135
8.4	Lower Bound using Polynomial Interpolation	140
8.5	Upper Bound and Uncertainty Principles	143
8.6	Related Work	146
	Bibliography	148

List of Figures

2.1	Scattering network	11
4.1	Frequency support of the rotational uniform covering frame	46
4.2	The geometry of the support sets	48
4.3	Rotational uniform covering frame elements	61
5.1	Fourier and Wavelet scattering transform comparison on Lena	72
5.2	Google Maps view of the experiment location	76
5.3	Gamma-ray spectrum of four examples	78
5.4	Results of the supervised Fourier scattering algorithm	79
7.1	The geometry of the Beurling theory for super-resolution	104
7.2	An illustration of the admissibility range for ε	108
8.1	Lagrange versus sparse Lagrange polynomials	141

List of Abbreviations

\mathbb{C}	complex numbers
\mathbb{N}	natural numbers
\mathbb{R}	real numbers
\mathbb{S}^1	unit circle
\mathbb{T}	torus group, identified with $[0, 1)$
\mathbb{Z}	integers
$*$	convolution of functions or measures on \mathbb{R}^d or \mathbb{T}^d
\times	product of two sets or measures
\circ	function composition
$A \setminus B$	set of elements in A but not in B
$A \leq B$	A is a subgroup of B
A/B	quotient of A by its subgroup G
I	identity map on \mathbb{R}^d
$[x]$	the greatest integer $n \geq x$
$\lfloor x \rfloor$	the smallest integer $n \leq x$
$ \cdot $	modulus, absolute value, or Lebesgue measure
$ \cdot _{\mathbb{T}}$	distance on \mathbb{T}
\mathbb{C}_S^N	set of S -sparse complex vectors of length N
$\ \cdot\ _p$	p -norm of a vector in \mathbb{C}^d , where $1 \leq p \leq \infty$
$B_R(c)$	the ball $\{x \in \mathbb{R}^d: \ x - c\ _2 \leq R\}$
$Q_R(c)$	the cube $\{x \in \mathbb{R}^d: \ x - c\ _{\infty} \leq R\}$
∇	gradient with respect to Euclidean coordinates
T_y	translation operator, $T_y f(x) = f(x - y)$
M_{ξ}	modulation operator, $M_{\xi} f(x) = e^{2\pi i \xi \cdot x} f(x)$
$L^p(X)$	functions f on X such that $ f ^p$ is Lebesgue integrable
$\ \cdot\ _{L^p}$	L^p norm
$M(X)$	bounded Radon measures on X
$\ \cdot\ _{\text{TV}}$	total variation norm
\widehat{f}	Fourier transform, $\widehat{f}(\xi) = \int_X f(x) e^{-2\pi i \xi \cdot x} dx$, where $X = \mathbb{R}^d$ or \mathbb{T}^d
$\widehat{\mu}$	Fourier transform, $\widehat{\mu}(\xi) = \int_X e^{-2\pi i \xi \cdot x} d\mu(x)$, where $X = \mathbb{R}^d$ or \mathbb{T}^d
δ_x	Dirac delta measure supported in $\{x\}$
$C(X)$	vector space of complex-valued continuous functions on X
$\ell^p(X)$	vector space of p -summable sequences indexed by X
$\ \cdot\ _{\ell^p}$	p -th norm of a sequence
$PW(\varepsilon, R)$	$f \in L^2(\mathbb{R}^d)$ such that $\ \widehat{f}\ _{L^2(Q_R(0))} \geq (1 - \varepsilon)\ f\ _{L^2}$
$\text{sign}(\cdot)$	the sign of a function or measure
$\text{supp}(\cdot)$	the support of a function or measure

Chapter 1: Introduction

1.1 Harmonic Analysis

Generally speaking, harmonic analysis is concerned with the representation of information (vectors, functions, sets, operators, etc.) as a collection of simpler pieces. Often, this decomposition is carried out using a transformation of the form,

$$T: X \rightarrow Y,$$

where X and Y are sets, and often times, are metric spaces, Hilbert spaces, etc. We often interpret X as the original or spatial domain, and Y as the more abstract coefficient or spectral domain.

The classical example is the *Fourier transform* on \mathbb{R}^d , formally defined as

$$f \mapsto \widehat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i \xi \cdot x} dx, \quad \xi \in \mathbb{R}^d.$$

Similarly, the Fourier transform on \mathbb{T}^d is formally defined as

$$f \mapsto \widehat{f}(m) = \int_{\mathbb{T}^d} f(x) e^{-2\pi i m \cdot x} dx, \quad m \in \mathbb{Z}^d.$$

The Fourier transform is well-defined on many function spaces, including $X = L^1(\mathbb{R}^d)$ and $X = L^1(\mathbb{T}^d)$. Even for these classical examples, the image of L^1 under the Fourier transform is not completely well-understood. Thus, it is difficult to say

much about the transform domain. See the books and monographs [125, 6, 74] for further background and discussion.

The study of such transformations involves two components.

(a) The *analysis formulation* or *forward problem* is to describe how T acts on X .

What are the properties of the image $T(X)$?

(b) The *synthesis formulation* or the *backwards problem* is concerned with inverting the transformation T . What $x \in X$ can be reconstructed from $y = Tx$?

Answers to either question depend on the transformation, and classical harmonic analysis has typically dealt with linear operators. Again, using the classical Fourier transform as the model example, we know that it is a unitary operator on $L^2(\mathbb{R}^d)$, products become convolution, smoothness converts to decay, and preserves tensor products. These results are considered part of the analysis formulation. It is also known that certain functions f can be reconstructed from its *inverse Fourier transform*, formally defined as

$$f(x) = \int_{\mathbb{R}^d} \widehat{f}(\xi) e^{2\pi i x \cdot \xi} d\xi.$$

This is the synthesis formulation.

The interaction of harmonic analysis and its applications, both to pure and applied math (number theory and partial differential equations), the physical sciences (quantum mechanics), and engineering (image and signal processing), has motivated and led to the development of new harmonic analysis ideas. We name a few that are particularly relevant to this thesis.

- (a) The complex exponentials are not compactly supported, which means that the Fourier transform is a non-local operator, and aggregates information about f at a wide range of locations. To obtain localized frequency information, the *short-time Fourier transform* (STFT) localizes f at a particular time using a window function g and takes the Fourier transform of the product. The d -dimensional STFT is formally defined by the formula,

$$V_g f(x, \xi) = \int_{\mathbb{R}} f(y) \overline{g(x-y)} e^{-2\pi i y \cdot \xi} dy,$$

see [71, 77]. This is a time-frequency representation and $V_g f(x, \xi)$ contains information about f at location x and frequency ξ . It is a fundamental tool in audio-processing: the *spectrogram*, $|V_g f(x, \xi)|^2$, is a common visualization of audio data.

- (b) Both the classical Fourier transform and the STFT have difficulty with representing discontinuous functions using a limited number of coefficients. This is manifested in the Gibbs phenomenon for the Fourier transform and the Balian-Low theorem for Gabor systems [16, 12, 8, 9, 10]. Discontinuous functions arise frequently in time series and image processing, due to rapid jumps and edges, respectively. To circumvent these issues, the wavelet transform localizes f at a particular location using a dilation of a function ψ , and this results in a time-scale representation [46]. The one-dimensional *continuous wavelet transform* (CWT) is formally defined as,

$$W_\psi f(a, b) = \frac{1}{\sqrt{|a|}} \int_{\mathbb{R}} f(x) \psi\left(\frac{x-b}{a}\right) dx,$$

see [47, 105]. Dealing with discontinuities in higher dimensions has motivated the development of curvelets [28] and shearlets [80].

- (c) It is known that the complex exponentials, $\{x \mapsto e^{2\pi imx}\}_{m \in \mathbb{Z}}$, forms an orthonormal basis for $L^2(\mathbb{T})$. More generally, for any orthonormal basis $\{\phi_j\}_{j=1}^\infty$ of a separable Hilbert space \mathcal{H} , we have the formula,

$$f = \sum_{j=1}^{\infty} \langle f, \phi_j \rangle \phi_j,$$

where the sum converges in the Hilbert space norm, see [7] for this basic fact. This equation is in its synthesis formulation, as it describes how to reconstruct the original function from its basis representation $\{\langle f, \phi_j \rangle\}_{j=1}^\infty$.

When the STFT and CWT are discretized to produce a sequence of functions acting on f , it is usually not an orthonormal basis. Thus, it is not immediately clear if the function can even be reconstructed from its coefficient expansion. Understanding when and what types of sequences can be used to recover the original function led to the development of *frame theory*, which deals with non-orthogonal expansions [58].

1.2 Modern Applications

The examples presented in the previous section (the Fourier transform, STFT, CWT, and frame theory), all deal with linear analysis and synthesis transformations. However, recent advances in deep learning and sparse recovery have illustrated important advantages of non-linear methods over their linear counter-parts.

- (a) A *neural network* implements a hierarchical transformation, where typically, each operation is an affine map, non-linearity, pooling, or linear aggregation. More formally, it can be written as a function on \mathbb{R}^d , of the form,

$$x \in \mathbb{R}^d \mapsto \rho_N(A_N(\cdots \rho_2(A_2(\rho_1(A_1x + b_1)) + b_2)) + b_N),$$

where each A_j is a matrix, b_j is a vector, and ρ_j is a pooling operator and/or a non-linearity. The network structure imposes constraints on A_j . For example, in a *convolutional neural network*, the matrices A_j will have a circulant structure, and if the convolution filter is compactly supported, then the entries far away from the diagonal are zero. Typically, the network structure and the non-linear operators are fixed, while the non-zero entries of A_j and b_j are learned in a training process.

Neural networks provide state-of-the-art results in many classification tasks [86], but they are not fully understood from both a theoretical and computational point of view. The main conceptual challenge is to understand why neural networks work so well. This question deal with the analysis formulation of the problem, since after all, a neural network is a non-linear transformation from one Euclidean space to another and we are interested in the properties of the coefficient domain; this is commonly referred to as the *feature space* in the computer science and machine learning communities.

- (b) Linear algebra states that any vector in \mathbb{R}^N can be recovered from its inner products with some spanning set of vectors. However, in many applications, obtaining sufficiently many samples is impossible or expensive, and we only

have $M \ll N$ observations. If the unknown vector is *sparse*, or has few non-zero entries compared to N , then it still might be possible to recover it from only $M \ll N$ projections. One major breakthrough in this direction is *compressed sensing*, which formalized this intuition [31, 56]. One representative result from this body of work states that, if a $M \times N$ matrix Φ satisfies certain properties, such as the *restricted isometry property* (RIP) [32] or incoherence [25], then any sufficiently sparse vector can be recovered from $y = \Phi x$ by solving

$$\min_{\tilde{x} \in \mathbb{C}^N} \|\tilde{x}\|_1 \quad \text{such that} \quad y = \Phi \tilde{x}. \quad (\text{BP})$$

This optimization algorithm was introduced in [35] as *basis pursuit*.

This result is in the synthesis formulation, as it tells us how to recover a sparse vector from incomplete linear measurements. This result highlights the advantage of non-linear methods over linear recovery techniques, particularly when there is some known prior information, such as sparsity, about the underlying objects. Since the aftermath of compressed sensing, there is still much interest in extending and generalizing sparse recovery results [62]. We name two directions that are particularly relevant to this thesis. First, in many imaging applications, the sensing matrix Φ does not satisfy RIP or incoherence, primarily because it is not a random matrix [62]. Second, functions are defined on the continuous domain, and the discretization error incurred in approximating such functions can be large [65].

1.3 Outline

Chapter 2 discusses the motivation for Mallat’s wavelet scattering transform, from the viewpoint of encoding invariances and convolutional neural networks. It defines uniform covering frames and discusses their relationship to time-frequency analysis. These frames are used in the definition of the Fourier scattering transform.

Chapter 3 proves two important properties of the Fourier scattering transform: the energy decays exponentially in the depth of the network and the majority of the energy is concentrated along the path decreasing frequencies. The main theorems show that the Fourier scattering transform is bounded above and below, contracts sufficiently small translations and diffeomorphisms, and is non-expansive. These properties justify its use as a feature extractor for classification.

Chapter 4 is concerned with the problem of incorporating rotational invariance into a time-frequency scattering transform. We construct uniform covering frames that are partially generated by rotation and modulation. When incorporated into a suitable network structure, the resulting rotational Fourier scattering transform is also rotationally invariant, in addition to the other properties. We also discuss the connection between this material and recent developments on directional harmonic analysis.

Chapter 5 contains numerical results and applications of the Fourier scattering transform. We show how to discretize the uniform covering frame elements to obtain digital versions, and this allows us to define the fast Fourier scattering transform. We apply this to supervised anomaly detection in radiological data.

Chapter 6 discusses the motivations and applications of super-resolution. The spectral extrapolation version can be viewed as an ill-posed inverse problems in the space of measures. Here, the measure encodes the object and the given information consists of low-frequency and possibly noisy Fourier samples. Depending on the type of measure, typically assumed to be discrete, we discuss important algorithms and theoretical results.

Chapter 7 develops the Beurling theory of super-resolution. We examine the total variation minimization method, and we describe the solutions to this optimization technique in terms of a set that depends on the given data. This set provides valuable information and almost characterizes the contrasting behaviors of the solutions. There is pathological phenomenon associated with some situations. We provide numerous examples to illustrate the applications of this theory.

Chapter 8 studies the recovery of discrete measures whose separation is below the Rayleigh length. There are only several algorithms that can handle this situation, but their exact limitations are unclear. We study the min-max error under a sparsity constraint on the measure, and we obtain the exact dependence of the min-max error, up to a constant depending only on the sparsity. This is done by obtaining a sharp lower bound on certain restricted Fourier matrices.

Chapter 2: Background on Scattering Transforms

2.1 Scattering Transforms

Introduced by LeCun [98], a *convolutional neural network* (CNN) is a composition of a finite number of transformations, where each transformation is one of three types: a convolution against a filter bank, a non-linearity, and an averaging. CNNs approximate functions through an adaptive and iterative learning process and have been extremely successful for classifying data [98, 86, 94]. Since they have complex architectures and their parameters are learned through “black-box” optimization schemes, such as stochastic gradient descent, training is computationally expensive and there is no widely accepted rigorous theory that explains their remarkable success.

Recently, Mallat [106] provided an intriguing example of a predetermined convolutional neural network with formal mathematical guarantees. His *wavelet scattering transform* propagates the input information through multiple iterations of the wavelet transform and the complex modulus, and finishes the process with a local averaging. It is typically used as a *feature extractor*, which is a transformation that organizes the input data into a particular form, while simultaneously discarding irrelevant information. It has been shown that the wavelet scattering transform

provides an effective representation of hand-digit recognition [23], texture images [124], audio and music analysis [1], and molecular classification [61].

Scattering transforms were originally introduced by Mallat [106] in the context of wavelets. Since then, more general operators have also been referred to as scattering transforms, so there is not an agreed upon criteria for what is considered a scattering transform. The definition that we employ is more general than the original formulation found in [106], but it is more restrictive than some recent formulations, see [36, 135] for variations on this theme.

A scattering transform depends on an underlying sequence of complex-valued functions on \mathbb{R}^d ,

$$\Phi = \{\varphi\} \cup \{\psi_\lambda\}_{\lambda \in \Lambda},$$

where Λ is a countable index set. The network is created by forming a tree structure from the index set Λ and associating each element of the tree with a corresponding operator. Indeed, let $\Lambda^0 = \emptyset$, and for integers $k \geq 1$, let

$$\Lambda^k = \underbrace{\Lambda \times \Lambda \times \cdots \times \Lambda}_{k\text{-times}}.$$

Then, each $\lambda \in \Lambda^k$ is associated with the *scattering propagator* $U[\lambda]$, formally defined as

$$U[\lambda](f) = \begin{cases} f & \text{if } \lambda \in \Lambda^0, \\ |f * \psi_\lambda| & \text{if } \lambda \in \Lambda, \\ U[\lambda_k]U[\lambda_{k-1}] \cdots U[\lambda_1]f & \text{if } \lambda = (\lambda_1, \lambda_2, \dots, \lambda_k) \in \Lambda^k. \end{cases}$$

Strictly speaking, it does not make sense to write $\lambda \in \Lambda^0 = \emptyset$, but we use this convention for convenience. Associated with Φ is the *scattering transform* \mathcal{S}_Φ , which

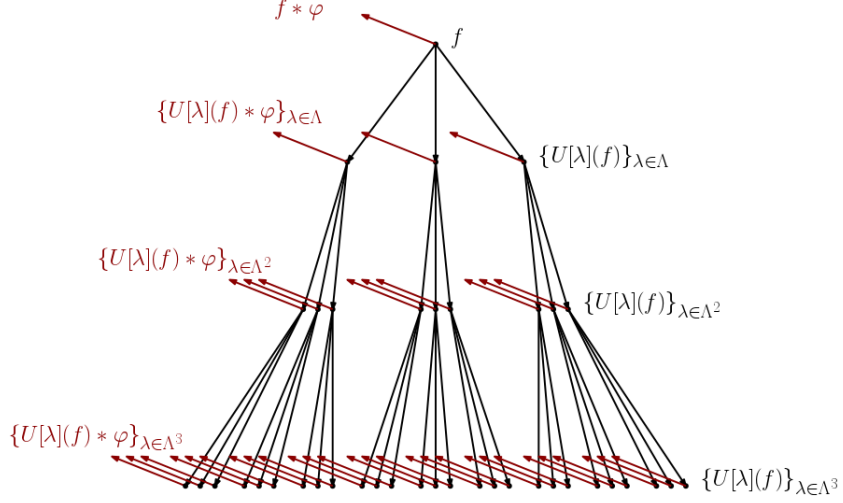


Figure 2.1: The network structure of the scattering transform. The black dots represent $U[\lambda](f)$ and can be computed iteratively. The scattering coefficients are calculated by convolving each $U[\lambda](f)$ with φ .

is formally defined as the sequence of functions,

$$\mathcal{S}_\Phi(f) = \{U[\lambda]f * \varphi : \lambda \in \Lambda^k, k = 0, 1, \dots\}.$$

We call the sequence,

$$\{U[\lambda]f * \varphi : \lambda \in \Lambda^k\},$$

the k -th order scattering coefficients of f . We mention that, in general, the scattering transform is not invertible due to the loss of the phase factor in each layer. See Figure 2.1, which demonstrates the tree network associated with the scattering transform.

Mallat originally used a specific wavelet frame \mathcal{W} in the above scattering transform and we call the resulting operator the *wavelet scattering transform* $\mathcal{S}_\mathcal{W}$. Let J be an integer and let G be a finite group of rotations on \mathbb{R}^d together with

reflection about the origin. Consider the wavelet frame,

$$\mathcal{W} = \mathcal{W}(J, G) = \{\varphi\} \cup \{\psi_{2^j, r} : j > -J, r \in G\},$$

where $\varphi(x) = 2^{-dJ}\varphi_0(2^{-J}x)$ is the wavelet corresponding to the coarsest scale 2^J , and $\psi_{2^j, r}(x) = 2^{dj}\psi(2^j r^{-1}x)$ is a detail wavelet of scale 2^{-j} and localization r . Here, we have followed Mallat's notation in [106] where the dilations of φ_0 and ψ are inversely related. The index set of \mathcal{W} is the countably infinite set

$$\Lambda = \{(2^j, r) : j > -J, r \in G\}.$$

Scattering transforms are used to help classify complicated data. Many types of digital data live in a high-dimensional Euclidean space, but we cannot accurately describe the properties of this subset (its geometry, distribution, or dimension). However, a reasonable assumption is that, for certain types, such as audio and image, a data point's classification and information content are invariant under small distortions. Assuming that this hypothesis is correct, effective feature extractor should contract small perturbations. To this end, Mallat proved that the wavelet scattering transform contracts small translations and diffeomorphisms, see [106] for the precise statements.

2.2 Uniform Covering Frames

While Mallat's results are impressive, there are several reasons to consider an alternative case where a non-wavelet frame is used for scattering.

(a) Since neural networks were originally inspired by the structure of the brain, it

makes sense to mimic the visual system of mammals when designing a feature extractor for image classification. The ground-breaking work of Daugman [49, 50] demonstrated that simple cells in the mammalian visual cortex are modeled by modulations of a fixed 2-dimensional Gaussian. In other words, this is a Gabor system with a Gaussian window. We remark that modern neural networks also incorporate ideas that are not strictly biologically motivated.

- (b) The authors of [99] observed that the learned filters (the experimentally “optimal” filters) in Hinton’s algorithm for learning *deep belief networks* [87] are localized, oriented, band-pass filters, which resemble Gabor functions. Strictly speaking, this set of functions is not a Gabor system since it is not derived from a single generating function, but it is not a wavelet system either.
- (c) The use of Gabor frames for classification is not unprecedented. The STFT has been used as a feature extractor for various image classification problems [82, 92, 2]. These papers pre-dated Mallat’s work on scattering transforms and did not use Gabor functions in a multi-layer decomposition. We are not aware of any prior work that combines multi-layer neural networks with Gabor functions.

Motivated by these considerations, we would like to incorporate Gabor-like functions into the scattering transform and establish theoretical properties of the resulting operator. We first recall some definitions, as well as discuss the history behind our viewpoint.

Recall that a *semi-discrete frame* for $L^2(\mathbb{R}^d)$ is a sequence of functions on \mathbb{R}^d , $\{\phi_j\}_{j \in \mathcal{J}} \subseteq L^2(\mathbb{R}^d)$, indexed by a possibly uncountable index set \mathcal{J} , with the

property that there exist $0 < A \leq B < \infty$ such that for all $f \in L^2(\mathbb{R}^d)$,

$$A\|f\|_{L^2}^2 \leq \sum_{j \in \mathcal{J}} \|f * \phi_j\|_{L^2}^2 \leq B\|f\|_{L^2}^2.$$

The constants A and B are the lower and upper frame bounds, respectively; when $A = B = 1$, the sequence is called a *Parseval frame*. We call $\{f * \phi_j\}_{j \in \mathcal{J}}$ the semi-discrete frame coefficients of f . The frame condition allows us to reconstruct any $f \in L^2(\mathbb{R}^d)$ from its frame coefficients using the Calderón-like formula: In the special case of a Parseval frame, then the reconstruction formula becomes,

$$f = \sum_{j \in \mathcal{J}} f * \phi_j * \bar{\phi}_j.$$

Semi-discrete Gabor frames naturally arise in the context of the STFT and audio processing. Let $\phi_j = M_{\xi_j} \bar{g}$ for some fixed function g , called the *window function*, and let $\{\xi_j\}_{j \in \mathcal{J}} \subseteq \mathbb{R}^d$ is a discrete set. Suppose we sample the STFT at the frequencies $\{\xi_j\}_{j \in \mathcal{J}}$. A simple calculation shows that

$$\begin{aligned} V_g f(x, \xi_j) &= \int_{\mathbb{R}^d} f(y) \overline{g(x-y)} e^{-2\pi i y \cdot \xi_j} dy \\ &= \int_{\mathbb{R}^d} f(y) (M_{\xi_j} \bar{g})(x-y) e^{-2\pi i x \cdot \xi_j} dy \\ &= e^{-2\pi i x \cdot \xi_j} (f * (M_{\xi_j} \bar{g}))(x), \end{aligned}$$

The phase factor $e^{-2\pi i x \cdot \xi_j}$ is harmless, whereas the sequence of functions ϕ_j is more important in determining the invertibility of this operator. In fact, if $\{M_{\xi_j} \bar{g}\}_{j \in \mathcal{J}}$ is a semi-discrete frame, then the original function f can be recovered in a stable way from its frequency sampled STFT by slightly modifying the Calderón formula. There is a simple method for verifying whether $\{M_{\xi_j} \bar{g}\}_{j \in \mathcal{J}}$ is a frame, but the exact frame bounds are hard to compute, see [48].

We typically select a smooth and fast decaying window, such as a Schwartz function, so that the frame coefficients capture jointly localized in space and frequency information. Gaussians are the unique minimizers of the Heisenberg uncertainty principle, and this motivated Gabor to introduce what are now called Gabor systems [72]. Generalizing this type of idea, we have the following concept.

Definition 2.2.1. Let \mathcal{P} be a countably infinite index set. A *uniform covering frame* is a sequence of functions,

$$\mathcal{F} = \{f_0\} \cup \{f_p\}_{p \in \mathcal{P}},$$

satisfying the following assumptions:

- (a) Assumptions on f_0 and f_p . Let $f_0 \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d) \cap C^1(\mathbb{R}^d)$ such that $\widehat{f_0}$ is supported in a neighborhood of the origin and $|\widehat{f_0}(0)| = 1$. For each $p \in \mathcal{P}$, let $f_p \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ such that $\text{supp}(\widehat{f_p})$ is compact and connected.
- (b) Uniform covering property. For any $R > 0$, there exists an integer $N > 0$ such that for each $p \in \mathcal{P}$, the set $\text{supp}(\widehat{f_p})$ can be covered by N cubes of side length $2R$.
- (c) Frame condition. Assume that for all $\xi \in \mathbb{R}^d$,

$$|\widehat{f_0}(\xi)|^2 + \sum_{p \in \mathcal{P}} |\widehat{f_p}(\xi)|^2 = 1. \quad (2.1)$$

This implies \mathcal{F} is a semi-discrete Parseval frame for $L^2(\mathbb{R}^d)$: For all $f \in L^2(\mathbb{R}^d)$,

$$\|f * f_0\|_{L^2}^2 + \sum_{p \in \mathcal{P}} \|f * f_p\|_{L^2}^2 = \|f\|_{L^2}^2.$$

The uniform covering property can be thought of as a size and a shape constraint on the sets $\{\text{supp}(\widehat{f}_p)\}_{p \in \mathcal{P}}$. The covering property is a size constraint because it implies $\sup_{p \in \mathcal{P}} |\text{supp}(\widehat{f}_p)| < \infty$, where $|S|$ denotes the Lebesgue measure of the set S . It is also a shape constraint because the number of cubes of a fixed side length required to cover the unit cube is much less than the number required to cover an elongated rectangular prism of unit volume.

While each f_p is defined in the Fourier domain, it is straightforward to imagine how each one appears in the spatial domain. Indeed, each f_p is supported near the origin, and it oscillates in a certain direction and at some frequency, depending only on p . Although f_p is not compactly supported due to the uncertainty principle, it can still have excellent spatial localization. For example, we can choose f_p to be a Schwartz function. If this is the case, the uniform covering property implies that each member of the family of functions $\{f_p\}_{p \in \mathcal{P}}$ is concentrated in a set that has approximately the same size and shape as the rest.

It is not surprising that a variety of Gabor frames are UCFs, as shown in the following proposition.

Proposition 2.2.2. *Let $g \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d) \cap C^1(\mathbb{R}^d)$ be such that $\text{supp}(\widehat{g})$ is compact and connected, $|\widehat{g}(0)| = 1$, and $\sum_{m \in \mathbb{Z}^d} |\widehat{g}(\xi - m)|^2 = 1$ for all $\xi \in \mathbb{R}^d$. Let $A: \mathbb{R}^d \rightarrow \mathbb{R}^d$ be an invertible linear transformation, $f_0(x) = |\det A| g(Ax)$, $\mathcal{P} = A^t \mathbb{Z}^d \setminus \{0\}$, and $f_p(x) = e^{2\pi i p \cdot x} f_0(x)$ for each $p \in \mathcal{P}$. Then, $\mathcal{F} = \{f_0\} \cup \{f_p\}_{p \in \mathcal{P}}$ is a Gabor frame, as well as a uniform covering frame.*

Proof. Since $\text{supp}(\widehat{f}_p)$ is a translation of the connected and compact set $\text{supp}(\widehat{f}_0)$,

the uniform covering property automatically holds. Let $A^{-t} = (A^{-1})^t$. For all $\xi \in \mathbb{R}^d$,

$$\begin{aligned} |\widehat{f}_0(\xi)|^2 + \sum_{p \in \mathcal{P}} |\widehat{f}_p(\xi)|^2 &= |\widehat{g}(A^{-t}\xi)|^2 + \sum_{p \in \mathcal{P}} |\widehat{g}(A^{-t}(\xi - p))|^2 \\ &= \sum_{m \in \mathbb{Z}^d} |\widehat{g}(A^{-t}\xi - m)|^2 = 1. \end{aligned}$$

□

A semi-discrete Gabor frame covers the frequency space uniformly by translating a fixed set, while a UCF covers the frequency domain by sets of approximately equal size and shape, which is a more general approach. In contrast, wavelets are partially generated by dilations of a single function, so a wavelet frame covers the frequency space non-uniformly by dilating a fixed function. Hence, due to the uniform covering property, no wavelet frame can be UCF. This is the major difference between our setting and that of Mallat's. Examples of wavelet frames include standard wavelets [47, 105], curvelets [28], shearlets [79, 80, 39, 40], composite wavelets [81], and α -molecules [78].

2.3 Fourier Scattering Transform

In contrast to Mallat's wavelet approach, we use a uniform covering frame

$$\mathcal{F} = \{f_0\} \cup \{f_p\}_{p \in \mathcal{P}}$$

in the scattering transform. Slightly abusing notation, we associate each multi-index $p \in \mathcal{P}^k$ with the scattering propagator $U[p]$, defined as

$$U[p]f = \begin{cases} f & \text{if } p \in \mathcal{P}^0, \\ |f * f_p| & \text{if } p \in \mathcal{P}, \\ U[p_k]U[p_{k-1}] \cdots U[p_1]f & \text{if } p = (p_1, p_2, \dots, p_k) \in \mathcal{P}^k. \end{cases}$$

Definition 2.3.1. The *Fourier scattering transform* $\mathcal{S}_{\mathcal{F}}$ is formally defined as

$$\mathcal{S}_{\mathcal{F}}(f) = \{U[p]f * f_0 : p \in \mathcal{P}^k, k = 0, 1, \dots\}.$$

Since the index set \mathcal{P} is just an abstract set, the frame identity (2.1), which is a partition of unity statement, does not immediately provide any information on how the partitioning is structured. However, there is a canonical way in which the tiling has to be done.

Proposition 2.3.2. *Let $\mathcal{F} = \{f_0\} \cup \{f_p\}_{p \in \mathcal{P}}$ be a uniform covering frame. There exist a constant $C = C_{\text{tiling}} > 0$ and subsets $\{\mathcal{P}[m]\}_{m=1}^{\infty} \subseteq \mathcal{P}$ such that for all integers $m \geq 1$,*

$$|\widehat{f}_0(\xi)|^2 + \sum_{p \in \mathcal{P}[m]} |\widehat{f}_p(\xi)|^2 = \begin{cases} 1 & \text{if } \xi \in \overline{Q_{C_1 m}(0)}, \\ 0 & \text{if } \xi \notin Q_{C_1(m+1)}(0). \end{cases} \quad (2.2)$$

Proof. For any set $S \subseteq \mathbb{R}^d$, let $\text{diam}(S) = \sup_{x, y \in S} |x - y|$ be the diameter of S .

Define

$$C = C_{\text{tiling}} = \max \left(\text{diam}(\text{supp}(\widehat{f}_0)), \sup_{p \in \mathcal{P}} \text{diam}(\text{supp}(\widehat{f}_p)) \right).$$

Note that C is finite because of the uniform covering property. Indeed, the diameter of $\text{supp}(\widehat{f}_0)$ is finite since \widehat{f}_0 is supported in a compact set containing the origin. Fix

$R > 0$, and by assumption, the closed and connected set $\text{supp}(\widehat{f}_p)$ can be covered by N cubes of side length $2R$. Then, the diameter of $\text{supp}(\widehat{f}_p)$ is bounded by $2NR$.

For integers $m \geq 1$, we define

$$\mathcal{P}[m] = \{p \in \mathcal{P} : \text{supp}(\widehat{f}_p) \subseteq \overline{Q_{C(m+1)}(0)}\}.$$

By definition, of $\mathcal{P}[m]$, we have

$$|\widehat{f}_0(\xi)|^2 + \sum_{p \in \mathcal{P}[m]} |\widehat{f}_p(\xi)|^2 = 0 \quad \text{if } \xi \notin Q_{C(m+1)}(0).$$

To complete the proof, we prove (2.2) by contradiction. Suppose there exists $\xi_0 \in Q_{Cm}(0)$ such that

$$|\widehat{f}_0(\xi_0)|^2 + \sum_{p \in \mathcal{P}[m]} |\widehat{f}_p(\xi_0)|^2 < 1.$$

By the frame condition (2.1), there exists $q \in \mathcal{P}$ such that $|\widehat{f}_q(\xi_0)| > 0$. Then, $\xi_0 \in \text{supp}(\widehat{f}_q)$ and by definition of $C > 0$, we have

$$\text{supp}(\widehat{f}_q) \subseteq \overline{Q_C(\xi_0)} \subseteq \overline{Q_{C(m+1)}(0)}.$$

This shows that $q \in \mathcal{P}[m]$, which contradicts the definition of $\mathcal{P}[m]$. \square

Remark 2.3.3. Suppose \mathcal{F} is a Gabor frame satisfying Proposition 2.2.2 for $A = aI$, where I is the identity transformation on \mathbb{R}^d and $a > 0$. By definition, we have $\mathcal{P} = a\mathbb{Z}^d \setminus \{0\}$. We can determine the family of sets $\{\mathcal{P}[m] : m \geq 1\}$ satisfying Proposition 2.3.2. Let $C_{\text{tiling}} = a$ and

$$\mathcal{P}[m] = \{p \in \mathcal{P} : \|p\|_\infty \leq m\}.$$

For all integers $m \geq 1$, we have

$$|\widehat{f}_0(\xi)|^2 + \sum_{p \in \mathcal{P}[m]} |\widehat{f}_p(\xi)|^2 = \begin{cases} 1 & \text{if } \xi \in \overline{Q_{am}(0)}, \\ 0 & \text{if } \xi \notin Q_{a(m+1)}(0). \end{cases}$$

A scattering transform corresponds to a network with an infinite number of layers and infinitely many nodes per layer, but when used in practice, it must be truncated. To truncate $\mathcal{S}_{\mathcal{F}}$, we keep terms up to a certain depth K and terms belonging to appropriate finite subsets of \mathcal{P}^k , for $k = 1, 2, \dots, K$. Recall that in Proposition 2.3.2, we established the existence of the family of sets $\{\mathcal{P}[m]\}_{m \geq 1}$, which are defined in terms of a constant $C_{\text{tiling}} > 0$. Similar to before, we create a tree from this collection of sets. For integers $M, K \geq 1$, we define the discrete set

$$\mathcal{P}[M]^K = \underbrace{\mathcal{P}[M] \times \mathcal{P}[M] \times \dots \times \mathcal{P}[M]}_{K\text{-times}} \subseteq \mathcal{P}^K.$$

Again, we use the convention that $\mathcal{P}[M]^0 = \emptyset$.

Definition 2.3.4. The *truncated Fourier scattering transform* $\mathcal{S}_{\mathcal{F}}[M, K]$ is formally defined as

$$\mathcal{S}_{\mathcal{F}}[M, K](f) = \{U[p]f * f_0 : p \in \mathcal{P}[M]^k, k = 0, 1, \dots, K\}.$$

Chapter 3: Properties of Fourier Scattering

The material in this chapter contains results from the paper [41], which was written by the author and Wojciech Czaja.

3.1 Energy Concentration

Since $\mathcal{S}_{\mathcal{W}}$ and $\mathcal{S}_{\mathcal{F}}$ have the same network structure, it is natural to ask whether $\mathcal{S}_{\mathcal{F}}$ satisfies the same broad mathematical properties. This is not immediately clear because wavelets and Gabor functions are qualitatively and mathematically different, see [46] for a comparison and discussion. Despite their differences, we prove that $\mathcal{S}_{\mathcal{F}}$ satisfies all the same properties of $\mathcal{S}_{\mathcal{W}}$. However, our proof techniques are very different from those of Mallat's. For example, he used scaling and almost orthogonality arguments to exploit the dyadic structure of wavelets, while we use tiling arguments to take advantage of the uniform covering property.

We need several preliminary results. For all $k \geq 0$ and $p \in \mathcal{P}^k$, it immediately follows from the frame property (2.1) that $U[p]: L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ is bounded with operator norm satisfying $\|U[p]\|_{L^2 \rightarrow L^2} \leq 1$. The following proposition contains some additional results that follow from the frame property. Mallat proved these for his wavelet frame \mathcal{W} in [106], but the arguments only rely on the frame identity (2.1),

so we omit their proofs.

Proposition 3.1.1. *Let $\mathcal{F} = \{f_0\} \cup \{f_p\}_{p \in \mathcal{P}}$ be a uniform covering frame. For any $f, g \in L^2(\mathbb{R}^d)$ and integers $K \geq 0$, we have*

$$\sum_{p \in \mathcal{P}^{K+1}} \|U[p]f\|_{L^2}^2 + \sum_{k=0}^K \sum_{p \in \mathcal{P}^k} \|U[p]f * f_0\|_{L^2}^2 = \|f\|_{L^2}^2,$$

and

$$\sum_{k=0}^K \sum_{p \in \mathcal{P}^k} \|U[p]f * f_0 - U[p]g * f_0\|_{L^2}^2 \leq \|f - g\|_{L^2}^2.$$

The first identity of Proposition 3.1.1 implies that $\mathcal{S}_{\mathcal{F}}: L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d; \ell^2(\mathbb{Z}))$

is bounded with operator norm satisfying $\|\mathcal{S}_{\mathcal{F}}\|_{L^2 \rightarrow L^2 \ell^2} \leq 1$. Indeed, we have

$$\|\mathcal{S}_{\mathcal{F}}(f)\|_{L^2 \ell^2}^2 = \lim_{K \rightarrow \infty} \sum_{k=0}^K \sum_{p \in \mathcal{P}^k} \|U[p]f * f_0\|_{L^2}^2 \leq \|f\|_{L^2}^2.$$

The following proposition is a basic result on positive definite functions. For any integer $k \geq 1$ and dimension d , Wendland [133] constructed a compactly supported, radial, and positive-definite $C^{2k}(\mathbb{R}^d)$ function. These functions are essentially anti-derivatives of positive-definite polynomial splines.

Proposition 3.1.2. *There exists a non-negative function $\phi: \mathbb{R}^d \rightarrow \mathbb{R}$, such that $\widehat{\phi}$ is continuous, decreasing along each Euclidean coordinate, and $\text{supp}(\widehat{\phi}) = \overline{Q_1(0)}$.*

Proof. For $j = 1, 2, \dots, d$, let $\phi_j: \mathbb{R} \rightarrow \mathbb{R}$ be defined by its one-dimensional Fourier transform,

$$\widehat{\phi}_j(\xi_j) = (1 - |\xi_j|) \mathbb{1}_{[0,1]}(|\xi_j|).$$

Here, $\mathbb{1}_S$ is the characteristic function of the set S and for a positive number x , $[x]$ stands for the integer n satisfying $n \leq x < n + 1$. Note that each ϕ_j is non-negative

because $\widehat{\phi}_j$ is a univariate positive-definite function, see [133]. Then, let $\phi: \mathbb{R}^d \rightarrow \mathbb{R}$ be the function,

$$\phi(x) = \phi_1(x_1)\phi_2(x_2)\cdots\phi_d(x_d).$$

By construction, ϕ satisfies the desired properties. \square

The following result is the crucial exponential decay of energy estimate, and from here onwards, we let C_{decay} denote the constant that appears in the proposition.

Proposition 3.1.3. *Let $\mathcal{F} = \{f_0\} \cup \{f_p\}_{p \in \mathcal{P}}$ be a uniform covering frame. There exists a constant $C_{\text{decay}} \in (0, 1)$ depending only on \mathcal{F} , such that for all $f \in L^2(\mathbb{R}^d)$ and integers $K \geq 1$,*

$$C_{\text{decay}}^{K-1} \|f * f_0\|_{L^2}^2 + \sum_{p \in \mathcal{P}^K} \|U[p]f\|_{L^2}^2 \leq C_{\text{decay}}^{K-1} \|f\|_{L^2}^2.$$

Proof. By assumption, \widehat{f}_0 is continuous, supported in a neighborhood of the origin, and $|\widehat{f}_0(0)| = 1$. Then, by appropriately scaling the function discussed in Proposition 3.1.2, there exists a non-negative ϕ such that $\widehat{\phi}$ is continuous, decreasing along each Euclidean coordinate, $|\widehat{\phi}(0)| > 0$, and $|\widehat{\phi}(\xi)| \leq |\widehat{f}_0(\xi)|$ for all $\xi \in \mathbb{R}^d$. Then, there exist constants $R = R_\phi > 0$ and $C = C_\phi \in (0, 1)$, such that $|\widehat{\phi}(\xi)|^2 \geq C$ for all $\xi \in Q_R(0)$. By the uniform covering property, there exists an integer $N = N_R > 0$, such that for all $p \in \mathcal{P}$, there exist $\{\xi_{p,n} \in \mathbb{R}^d: n = 1, 2, \dots, N\}$ such that

$$\text{supp}(\widehat{f}_p) \subseteq \bigcup_{n=1}^N Q_R(\xi_{p,n}).$$

Let $1 \leq k \leq K$ and $q \in \mathcal{P}^{k-1}$. By Plancherel's formula and the above inclusion,

$$\begin{aligned} \|U[q]f * f_p\|_{L^2}^2 &= \int_{\mathbb{R}^d} |\widehat{U[q]f}(\xi)|^2 |\widehat{f_p}(\xi)|^2 d\xi \\ &\leq \sum_{n=1}^N \int_{Q_R(\xi_{p,n})} |\widehat{U[q]f}(\xi)|^2 |\widehat{f_p}(\xi)|^2 d\xi. \end{aligned}$$

Since $|\widehat{\phi}(\xi - \xi_{p,n})|^2 \geq C$ for all $\xi \in Q_R(\xi_{p,n})$, we have

$$\begin{aligned} &\sum_{n=1}^N \int_{Q_R(\xi_{p,n})} |\widehat{U[q]f}(\xi)|^2 |\widehat{f_p}(\xi)|^2 d\xi \\ &\leq \frac{1}{C} \sum_{n=1}^N \int_{Q_R(\xi_{p,n})} |\widehat{U[q]f}(\xi)|^2 |\widehat{f_p}(\xi)|^2 |\widehat{\phi}(\xi - \xi_{p,n})|^2 d\xi. \end{aligned}$$

By Plancherel's formula, we have

$$\sum_{n=1}^N \int_{Q_R(\xi_{p,n})} |\widehat{U[q]f}(\xi)|^2 |\widehat{f_p}(\xi)|^2 |\widehat{\phi}(\xi - \xi_{p,n})|^2 d\xi \leq \sum_{n=1}^N \|U[q]f * f_p * M_{\xi_{p,n}}\phi\|_{L^2}^2.$$

Using that $\phi \geq 0$ and triangle inequality, we have

$$\sum_{n=1}^N \|U[q]f * f_p * M_{\xi_{p,n}}\phi\|_{L^2}^2 \leq \sum_{n=1}^N \| |U[q]f * f_p| * \phi \|_{L^2}^2.$$

Observe that the terms inside the summation on the right hand side do not depend on the index n . Using Plancherel's formula and that $|\widehat{\phi}(\xi)| \leq |\widehat{f_0}(\xi)|$ for all $\xi \in \mathbb{R}^d$, we have

$$\| |U[q]f * f_p| * \phi \|_{L^2}^2 \leq \| |U[q]f * f_p| * f_0 \|_{L^2}^2.$$

Combining the previous inequalities and rearranging the result, we obtain

$$\| |U[q]f * f_p| * f_0 \|_{L^2}^2 \geq \frac{C}{N} \|U[q]f * f_p\|_{L^2}^2.$$

The strength of this inequality is that C and N are independent of $q \in \mathcal{P}^{k-1}$ and $p \in \mathcal{P}$. Then, summing this inequality over all $p \in \mathcal{P}$ and $q \in \mathcal{P}^{k-1}$, we see that

$$\sum_{p \in \mathcal{P}^k} \|U[p]f * f_0\|_{L^2}^2 \geq \frac{C}{N} \sum_{p \in \mathcal{P}^k} \|U[p]f\|_{L^2}^2.$$

Applying the frame identity (2.1) to the left hand side and setting $C_{\text{decay}} = 1 - C/N$, we have

$$\sum_{p \in \mathcal{P}^{k+1}} \|U[p]f\|_{L^2}^2 \leq C_{\text{decay}} \sum_{p \in \mathcal{P}^k} \|U[p]f\|_{L^2}^2.$$

Iterating the above inequality, we obtain

$$\sum_{p \in \mathcal{P}^K} \|U[p]f\|_{L^2}^2 \leq C_{\text{decay}}^{K-1} \sum_{p \in \mathcal{P}} \|U[p]f\|_{L^2}^2 = C_{\text{decay}}^{K-1} \|f\|_{L^2}^2 - C_{\text{decay}}^{K-1} \|f * f_0\|_{L^2}^2.$$

□

Remark 3.1.4. The key step in the proof of Proposition 3.1.3 is to obtain an inequality of the form,

$$\| |f * g| * f_0 \|_{L^2} \geq C \|f * g\|_{L^2},$$

for some constant $C > 0$ independent of $f, g \in L^2(\mathbb{R}^d)$. It is straightforward to obtain a lower bound where the constant depends on f and p . Indeed, suppose $f * g \neq 0$. Since $|f * g|$ is continuous, we have

$$(|f * g|)^\wedge(0) = \int_{\mathbb{R}^d} |(f * g)(x)| dx = \|f * g\|_{L^1} > 0.$$

By continuity of $|f * g|$, the above inequality, and the assumption that $|\widehat{f_0}(0)| = 1$, we can find a sufficiently small neighborhood $V = V_{f_0, f, g}$ of the origin and constants $C_{f_0}, C_{f, g} > 0$, such that $|\widehat{f_0}(\xi)| \geq C_{f_0}$ and $|(|f * g|)^\wedge(\xi)| \geq C_{f, g} \|f * g\|_{L^1}$ for all $\xi \in V$. Then,

$$\| |f * g| * f_0 \|_{L^2}^2 \geq \int_V |(|f * g|)^\wedge(\xi)|^2 |\widehat{f_0}(\xi)|^2 d\xi \geq C_{f_0} C_{f, g} \|f * g\|_{L^1} |V|,$$

where $|V|$ is the Lebesgue measure of V . However, both $C_{f, g}$ and $|V|$ depend on f and g , and $L^1(\mathbb{R}^d)$ and $L^2(\mathbb{R}^d)$ norms are not equivalent. Thus, this inequality is not

very useful for our purposes. However, this naive reasoning suggests that a more sophisticated covering argument, such as the one given in the proof of Proposition 3.1.3, could work.

Remark 3.1.5. We have several comments about Proposition 3.1.3.

(a) Since C_{decay} describes the rate of decay of $\sum_{p \in \mathcal{P}^k} \|U[p]f\|_{L^2}^2$, it is of interest to determine the optimal (smallest) value C_{decay} for which Proposition 3.1.3 holds. Minimizing C_{decay} is equivalent to maximizing the ratio C/N , where these constants were defined in the proof. Since N is related to the optimal covering by cubes of side length $2R$ and C is the minimum of $|\widehat{\phi}|^2$ on $Q_R(0)$, both C and N decrease as R increases.

(b) Let us examine why the argument proving Proposition 3.1.3 fails for the wavelet case. Recall that $\psi_{2^j, r}$ has frequency scale 2^j whereas φ has frequency scale 2^{-J} . Using the same argument as in the proof of Proposition 3.1.3, we obtain: For each $j > -J$, there exists an integer $N_j > 0$ and $C_j \in (0, 1)$, such that for all $k \geq 1$, $p \in \Lambda^k$, $r \in G$, and $f \in L^2(\mathbb{R}^d)$,

$$\| |U[p]f * \psi_{2^j, r}| * \varphi_{2^j} \|_{L^2}^2 \geq \frac{C_j}{N_j} \|U[p]f * \psi_{2^j, r}\|_{L^2}^2.$$

The measure of $\text{supp}(\widehat{\psi_{2^j, r}})$ is proportional to 2^j , and N_j is the number of cubes required to cover this set with cubes of side length bounded by a constant multiple of 2^{-J} . Hence, $\lim_{j \rightarrow \infty} N_j = \infty$ and this inequality is not meaningful for large j .

(c) Rather interestingly, numerical experiments in [106, page 1345] have suggested

that the exponential decay of energy described in the Proposition 3.1.3 holds for the wavelet case. Mallat conjectured that there exists $C \in (0, 1)$ such that

$$\sum_{\lambda \in \Lambda^K} \|U[\lambda]f\|_{L^2}^2 \leq C^{K-1} \|f\|_{L^2}^2,$$

for all $f \in L^2(\mathbb{R}^d)$, and $K \geq 1$. Determining whether this property holds for a given wavelet frame is of interest in the scattering community.

Proposition 3.1.3 proved that the energy decreases exponentially quickly in the depth, which provides excellent control over the deep layers of the network. For width truncation, we need to show that the energy of f is concentrated on a small subset of the coefficients. We are concerned with bounding terms of the form, $\| |f * f_p| * f_q \|_{L^2}$, where $p \in \mathcal{P}[M]$ and $q \in \mathcal{P}[M]^c$. These are the terms that are thrown away due to truncation, and since f_p has lower frequencies than f_q , we expect them to be small. Proposition 3.1.6 shows that this intuition holds: most of the energy is concentrated along the frequency decaying paths.

Proposition 3.1.6. *Let $\mathcal{F} = \{f_0\} \cup \{f_p\}_{p \in \mathcal{P}}$ be a uniform covering frame. For any integer $M \geq 1$, there exists $C_M \in (0, 1)$ such that $C_M \rightarrow 1$ as $M \rightarrow \infty$ and for all integers $k \geq 1$, $p \in \mathcal{P}^k$, and $f \in L^2(\mathbb{R}^d)$,*

$$\|U[p]f * f_0\|_{L^2}^2 + \sum_{q \in \mathcal{P}[M]} \|U[p]f * f_q\|_{L^2}^2 \geq C_M \|U[p]f\|_{L^2}^2.$$

Proof. By Proposition 3.1.2, there exists a non-negative function ϕ , such that $\widehat{\phi}$ is continuous, decreasing along each Euclidean coordinate, $\text{supp}(\widehat{\phi}) = \overline{Q_1(0)}$, and $|\widehat{\phi}(0)| = 1$. Define ϕ_M by its Fourier transform, $\widehat{\phi}_M(\xi) = \widehat{\phi}(C^{-1}M^{-1}\xi)$, where $C =$

C_{tiling} . By definition of $\mathcal{P}[M]$, for all $\xi \in \mathbb{R}^d$,

$$|\widehat{\phi}_M(\xi)|^2 \leq |\widehat{f}_0(\xi)|^2 + \sum_{q \in \mathcal{P}[M]} |\widehat{f}_q(\xi)|^2.$$

Plancherel's formula and this inequality imply

$$\|U[p]f * f_0\|_{L^2}^2 + \sum_{q \in \mathcal{P}[M]} \|U[p]f * f_q\|_{L^2}^2 = \| |U[p']f * f_s| * \phi_M \|_{L^2}^2,$$

where $p = (p', s)$. By definition of C_{tiling} , there exists $\xi_s \in \mathbb{R}^d$ such that $\text{supp}(\widehat{f}_s) \subseteq \overline{Q_{C/2}(\xi_s)}$. Applying triangle inequality to the right hand side of the previous inequality and using that $\phi \geq 0$, we have

$$\begin{aligned} \| |U[p']f * f_s| * \phi_M \|_{L^2}^2 &\geq \|U[p']f * f_s * M_{\xi_s} \phi_M\|_{L^2}^2 \\ &= \int_{\mathbb{R}^d} |\widehat{U[p']f}(\xi)|^2 |\widehat{f}_s(\xi)|^2 |\widehat{\phi}_M(\xi - \xi_s)|^2 d\xi. \end{aligned}$$

Since $\widehat{\phi}$ is decreasing along each Euclidean coordinate and the inclusion $\text{supp}(\widehat{f}_s) \subseteq \overline{Q_{C/2}(\xi_s)}$, we have

$$\begin{aligned} C_M &= \inf_{\xi \in \text{supp}(\widehat{f}_s)} |\widehat{\phi}_M(\xi - \xi_s)|^2 \\ &\geq \inf_{\xi \in \overline{Q_{C/2}(0)}} |\widehat{\phi}_M(\xi)|^2 \\ &= |\widehat{\phi}(2^{-1}M^{-1}, 2^{-1}M^{-1}, \dots, 2^{-1}M^{-1})|^2 > 0. \end{aligned}$$

Applying Plancherel's formula yields

$$\begin{aligned} \int_{\mathbb{R}^d} |\widehat{U[p']f}(\xi)|^2 |\widehat{f}_s(\xi)|^2 |\widehat{\phi}_M(\xi - \xi_s)|^2 d\xi &\geq C_M \|U[p']f * f_s\|_{L^2}^2 \\ &= C_M \|U[p]f\|_{L^2}^2. \end{aligned}$$

We have the trivial inequality $C_M \leq 1$, and observe that

$$\begin{aligned} \liminf_{M \rightarrow \infty} C_M &\geq \liminf_{M \rightarrow \infty} |\widehat{\phi}(2^{-1}M^{-1}, 2^{-1}M^{-1}, \dots, 2^{-1}M^{-1})|^2 \\ &= |\widehat{\phi}(0)|^2 \\ &= 1. \end{aligned}$$

□

Remark 3.1.7. This argument fails for wavelet frames. Indeed, we made use of the uniform tiling property in Proposition 2.3.2 and that \widehat{f}_p is supported in a cube of side length C_{tiling} , independent of $p \in \mathcal{P}$.

3.2 Fourier Scattering Properties

We are ready to prove our first main theorem, which shows that $\mathcal{S}_{\mathcal{F}}$ satisfies several desirable properties as a feature extractor.

Theorem 3.2.1. *Let $\mathcal{F} = \{f_0\} \cup \{f_p\}_{p \in \mathcal{P}}$ be a uniform covering frame.*

(a) Energy conservation: For all $f \in L^2(\mathbb{R}^d)$,

$$\|\mathcal{S}_{\mathcal{F}}(f)\|_{L^2\ell^2} = \|f\|_{L^2}.$$

(b) Non-expansiveness: For all $f, g \in L^2(\mathbb{R}^d)$,

$$\|\mathcal{S}_{\mathcal{F}}(f) - \mathcal{S}_{\mathcal{F}}(g)\|_{L^2\ell^2} \leq \|f - g\|_{L^2}.$$

(c) Translation contraction: There exists $C > 0$ depending only on \mathcal{F} , such that for all $f \in L^2(\mathbb{R}^d)$ and $y \in \mathbb{R}^d$,

$$\|\mathcal{S}_{\mathcal{F}}(T_y f) - \mathcal{S}_{\mathcal{F}}(f)\|_{L^2\ell^2} \leq C|y| \|\nabla f_0\|_{L^1} \|f\|_{L^2}.$$

(d) Additive diffeomorphisms contraction: Let $\varepsilon \in [0, 1)$ and $R > 0$. There exists a universal constant $C > 0$, such that for all $f \in PW(\varepsilon, R)$, and all $\tau \in C^1(\mathbb{R}^d; \mathbb{R}^d)$ with $\|\nabla\tau\|_{L^\infty} \leq 1/(2d)$,

$$\|\mathcal{S}_{\mathcal{F}}(f \circ (I - \tau)) - \mathcal{S}_{\mathcal{F}}(f)\|_{L^2\ell^2} \leq C(R\|\tau\|_{L^\infty} + \varepsilon)\|f\|_{L^2}.$$

Proof.

(a) Using Propositions 3.1.1 and 3.1.3, we obtain

$$\begin{aligned} \|\mathcal{S}_{\mathcal{F}}(f)\|_{L^2\ell^2}^2 &= \lim_{K \rightarrow \infty} \sum_{k=0}^K \sum_{p \in \mathcal{P}^k} \|U[p]f * f_0\|_{L^2}^2 \\ &= \|f\|_{L^2}^2 - \lim_{K \rightarrow \infty} \sum_{p \in \mathcal{P}^K} \|U[p]f\|_{L^2}^2 = \|f\|_{L^2}^2. \end{aligned}$$

(b) Using Proposition 3.1.1, we obtain

$$\begin{aligned} \|\mathcal{S}_{\mathcal{F}}(f) - \mathcal{S}_{\mathcal{F}}(g)\|_{L^2\ell^2}^2 &= \lim_{K \rightarrow \infty} \sum_{k=0}^K \sum_{p \in \mathcal{P}^k} \|U[p]f * f_0 - U[p]g * f_0\|_{L^2}^2 \\ &\leq \|f - g\|_{L^2}^2. \end{aligned}$$

(c) By definition, we have

$$\|\mathcal{S}_{\mathcal{F}}(T_y f) - \mathcal{S}_{\mathcal{F}}(f)\|_{L^2\ell^2}^2 = \sum_{k=0}^{\infty} \sum_{p \in \mathcal{P}^k} \|U[p](T_y f) * f_0 - U[p]f * f_0\|_{L^2}^2.$$

Since translation commutes with convolution and the complex modulus, for all

$k \geq 0$ and $p \in \mathcal{P}^k$, we have

$$U[p](T_y f) * f_0 = T_y(U[p]f) * f_0 = U[p]f * T_y f_0.$$

This fact, combined with Young's inequality, yields

$$\begin{aligned} \|U[p](T_y f) * f_0 - U[p]f * f_0\|_{L^2} &= \|U[p]f * (T_y f_0 - f_0)\|_{L^2} \\ &\leq \|T_y f_0 - f_0\|_{L^1} \|U[p]f\|_{L^2}. \end{aligned}$$

Then, we have

$$\|\mathcal{S}_{\mathcal{F}}(T_y f) - \mathcal{S}_{\mathcal{F}}(f)\|_{L^2 \ell^2} \leq \|T_y f_0 - f_0\|_{L^1} \left(\sum_{k=0}^{\infty} \sum_{p \in \mathcal{P}^k} \|U[p]f\|_{L^2}^2 \right)^{1/2}. \quad (3.1)$$

We first bound the summation in (3.1). Using Proposition 3.1.3, we have

$$\begin{aligned} \sum_{k=0}^{\infty} \sum_{p \in \mathcal{P}^k} \|U[p]f\|_{L^2}^2 &\leq \|f\|_{L^2}^2 + \sum_{k=1}^{\infty} C_{\text{decay}}^{k-1} \|f\|_{L^2}^2 \\ &= \left(1 + \frac{1}{1 - C_{\text{decay}}}\right) \|f\|_{L^2}^2. \end{aligned}$$

To bound the L^1 term in (3.1), we use the fundamental theorem of calculus, which is justified by the assumption that $f_0 \in C^1(\mathbb{R}^d)$. Then, we obtain

$$\int_{\mathbb{R}^d} |f_0(x-y) - f_0(x)| dx = \int_{\mathbb{R}^d} \left| \int_0^1 \nabla f_0(x-ty) \cdot y dt \right| dx \leq |y| \|\nabla f_0\|_{L^1}.$$

- (d) Let ϕ be a Schwartz function such that $\widehat{\phi}$ is real-valued, supported in $\overline{Q_2(0)}$, and $\widehat{\phi}(\xi) = 1$ for all $\xi \in Q_1(0)$. Let $\phi_R(x) = R^d \phi(Rx)$, and let $f_R = f * \phi_R$. By the non-expansiveness property and triangle inequality

$$\begin{aligned} &\|\mathcal{S}_{\mathcal{F}}(f \circ (I - \tau)) - \mathcal{S}_{\mathcal{F}}(f)\|_{L^2 \ell^2} \\ &\leq \|f - f_R\|_{L^2} + \|f_R \circ (I - \tau) - f_R\|_{L^2} + \|f_R \circ (I - \tau) - T_{\tau} f\|_{L^2}. \end{aligned} \quad (3.2)$$

The bound for the first term in (3.2) follows by assumption

$$\|f - f_R\|_{L^2} \leq \varepsilon \|f\|_{L^2}.$$

To bound the second term in (3.2), we make the change of variable $u = x - \tau(x)$ and note that

$$\left| \frac{\partial u}{\partial x} \right| = |\det(I - \nabla \tau(x))| \geq (1 - d \|\nabla \tau\|_{L^\infty}) \geq \frac{1}{2},$$

see [22]. Then, we have

$$\begin{aligned}
\|f_R \circ (I - \tau) - f \circ (I - \tau)\|_{L^2}^2 &= \int_{\mathbb{R}^d} |f_R(x - \tau(x)) - f(x - \tau(x))|^2 dx \\
&\leq 2\|f - f_R\|_{L^2}^2 \\
&\leq 2\varepsilon^2\|f\|_{L^2}^2.
\end{aligned}$$

It remains to bound the third term of (3.2), and we use the argument proved in [134, Proposition 5]. We have

$$\begin{aligned}
(f_R \circ (I - \tau))(x) - f_R(x) &= (f * \phi_R)(x - \tau(x)) - (f * \phi_R)(x) \\
&= \int_{\mathbb{R}^d} (\phi_R(x - \tau(x) - y) - \phi_R(x - y))f(y) dy.
\end{aligned}$$

The above can be interpreted as an integral kernel operator acting on $f \in L^2(\mathbb{R}^d)$ with kernel

$$k(x, y) = \phi_R(x - \tau(x) - y) - \phi_R(x - y).$$

The proof is completed by verifying that this kernel satisfies the assumptions of Schur's lemma with the appropriate bounds. By the fundamental theorem of calculus, we have

$$\begin{aligned}
|k(x, y)| &= \left| \int_0^1 \nabla \phi_R(x - t\tau(x) - y) \cdot \tau(x) dt \right| \\
&\leq \|\tau\|_{L^\infty} \int_0^1 |\nabla \phi_R(x - t\tau(x) - y)| dt.
\end{aligned}$$

(i) For each $x \in \mathbb{R}^d$, we have

$$\begin{aligned}
\int_{\mathbb{R}^d} |k(x, y)| dy &\leq \|\tau\|_{L^\infty} \int_0^1 \int_{\mathbb{R}^d} |\nabla \phi_R(x - t\tau(x) - y)| dy dt \\
&= R\|\nabla \phi\|_{L^1}\|\tau\|_{L^\infty}.
\end{aligned}$$

(ii) For each $y \in \mathbb{R}^d$, we have

$$\int_{\mathbb{R}^d} |k(x, y)| dx \leq \|\tau\|_{L^\infty} \int_0^1 \int_{\mathbb{R}^d} |\nabla \phi_R(x - t\tau(x) - y)| dx dt.$$

For fixed $y \in \mathbb{R}^d$ and $t \in [0, 1]$, we make the change of variables $v = x - t\tau(x) - y$ and note that

$$\left| \frac{\partial v}{\partial x} \right| = |\det(I - t\nabla\tau(x))| \geq (1 - td\|\nabla\tau\|_{L^\infty}) \geq \frac{1}{2}.$$

Thus, for all $y \in \mathbb{R}^d$,

$$\int_{\mathbb{R}^d} |k(x, y)| dx \leq 2\|\tau\|_{L^\infty} \int_0^1 \int_{\mathbb{R}^d} |\nabla \phi_R(v)| dv dt = 2R\|\nabla\phi\|_{L^1}\|\tau\|_{L^\infty}.$$

By Schur's lemma, we conclude that

$$\|T_\tau f - f\|_{L^2} \leq 2R\|\nabla\phi\|_{L^1}\|\tau\|_{L^\infty}\|f\|_{L^2}.$$

Combining the above results, we obtain the inequality

$$\|\mathcal{S}_{\mathcal{F}}(T_\tau f) - \mathcal{S}_{\mathcal{F}}(f)\|_{L^2 \ell^2} \leq (2R\|\nabla\phi\|_{L^1}\|\tau\|_{L^\infty} + \varepsilon + \sqrt{2\varepsilon})\|f\|_{L^2}.$$

Set $C = \max(2\|\nabla\phi\|_{L^1}, 1 + \sqrt{2})$, which completes the proof.

□

We are ready to prove our second main theorem, which deals with the truncated Fourier scattering transform. We first provide some motivation for our choice of truncation and the assumptions that we require below. At this point, it is not clear whether $\mathcal{S}_{\mathcal{F}}[M, K]$ is non-trivial for appropriate M and K . For example, let us focus our attention on the first layer. Say we fix M and compute a finite number of first-order coefficients,

$$\{|f * f_p| * f_0 : p \in \mathcal{P}[M]\}.$$

Observe that there exists a non-trivial $f \in L^2(\mathbb{R}^d)$ such that $f * f_p = 0$ for all $p \in \mathcal{P}[M]$. This is already problematic, since it shows that this truncated operator has non-trivial kernel and consequently, is not bounded from below. This shows that, in order to truncate just the first layer of coefficients, we need an additional assumption on $f \in L^2(\mathbb{R}^d)$. The most natural assumption is that f is (ε, R) band-limited, and then M can be chosen appropriately depending on R .

Now, we focus our attention on the higher-order terms. The naive idea is to only compute coefficients indexed by the finite subset $\mathcal{P}[M]^k \subseteq \mathcal{P}^k$, namely,

$$\{U[p]f * f_0 : p \in \mathcal{P}[M]^k, k = 0, 1, \dots, K\}.$$

This truncation tosses away the high frequency terms and might seem reasonable since Proposition 3.1.3 showed that the complex modulus pushes higher frequencies to lower frequencies. Indeed, for $f \in L^2(\mathbb{R}^d)$ and $p \in \mathcal{P}$, the proposition showed that $(|f * f_p|)^\wedge$ is non-zero in a neighborhood of the origin even though $(f * f_p)^\wedge$ is compactly supported away from the origin. However, the complex modulus can also push lower frequencies to higher frequencies. To see why, we note that the function $|f * f_p|$ is continuous, but in general, it is not $C^1(\mathbb{R}^d)$; even if we make the very mild assumption that $\nabla f_p \in L^1(\mathbb{R}^d)$, we can only conclude that $|f * f_p|$ has one distributional derivative belonging to $L^2(\mathbb{R}^d)$. Thus, the decay of $(|f * f_p|)^\wedge$ is quite slow, even though $(f * f_p)^\wedge$ is compactly supported! This observation shows that we must be careful when truncating $\mathcal{S}_{\mathcal{F}}$. However, we shall see that our choice of truncation in fact works, but only requires an alternative argument.

Theorem 3.2.2. *Let $\mathcal{F} = \{f_0\} \cup \{f_p\}_{p \in \mathcal{P}}$ be a uniform covering frame.*

(a) Upper bound: For all $f \in L^2(\mathbb{R}^d)$ and integers $M, K \geq 1$,

$$\|\mathcal{S}_{\mathcal{F}}[M, K](f)\|_{L^2\ell^2} \leq \|f\|_{L^2}.$$

(b) Lower bound: Let $\varepsilon \in [0, 1)$ and $R > 0$. There exist integers $K \geq 1$ and $M \geq C_{\text{tiling}}^{-1}R$ sufficiently large depending on ε , such that for all (ε, R) band-limited functions $f \in L^2(\mathbb{R}^d)$,

$$\|\mathcal{S}_{\mathcal{F}}[M, K](f)\|_{L^2\ell^2}^2 \geq (C_M^K(1 - \varepsilon^2) - C_{\text{decay}}^{K-1})\|f\|_{L^2}^2.$$

(c) Non-expansiveness: For all $f, g \in L^2(\mathbb{R}^d)$ and integers $M, K \geq 1$,

$$\|\mathcal{S}_{\mathcal{F}}[M, K](f) - \mathcal{S}_{\mathcal{F}}[M, K](g)\|_{L^2\ell^2} \leq \|f - g\|_{L^2}.$$

(d) Translation contraction: There exists a constant $C > 0$ depending only on \mathcal{F} such that for all $f \in L^2(\mathbb{R}^d)$, $y \in \mathbb{R}^d$, and integers $M, K \geq 1$, we have

$$\|\mathcal{S}_{\mathcal{F}}[M, K](T_y f) - \mathcal{S}_{\mathcal{F}}[M, K](f)\|_{L^2\ell^2} \leq C|y|\|\nabla f_0\|_{L^1}\|f\|_{L^2}.$$

(e) Additive diffeomorphism contraction: Let $\varepsilon \in [0, 1)$ and $R > 0$. There exists a universal constant $C > 0$, such that for all $f \in PW(\varepsilon, R)$, and all $\tau \in C^1(\mathbb{R}^d; \mathbb{R}^d)$ with $\|\nabla\tau\|_{L^\infty} \leq 1/(2d)$,

$$\|\mathcal{S}_{\mathcal{F}}[M, K](f \circ (I - \tau)) - \mathcal{S}_{\mathcal{F}}[M, K](f)\|_{L^2\ell^2} \leq C(R\|\tau\|_{L^\infty} + \varepsilon)\|f\|_{L^2}.$$

Proof.

(a) We apply Theorem 4.3.1 to obtain,

$$\|\mathcal{S}_{\mathcal{F}}[M, K](f)\|_{L^2\ell^2} \leq \|\mathcal{S}_{\mathcal{F}}(f)\|_{L^2\ell^2} = \|f\|_{L^2}.$$

(b) By Proposition 2.3.2, the assumption that f is almost band-limited, and that

$C_{\text{tiling}}M \geq R$, we have

$$\|f\|_{L^2}^2 = \|f * f_0\|_{L^2}^2 + \sum_{p \in \mathcal{P}[M]} \|f * f_p\|_{L^2}^2 + \varepsilon^2 \|f\|_{L^2}^2.$$

Applying Proposition 3.1.6 to the summation over $\mathcal{P}[M]$, we obtain,

$$(1 - \varepsilon^2) \|f\|_{L^2}^2 \leq \|f * f_0\|_{L^2}^2 + C_M^{-1} \sum_{p \in \mathcal{P}[M]} \|U[p]f * f_0\|_{L^2}^2 + C_M^{-1} \sum_{p \in \mathcal{P}[M]^2} \|U[p]f\|_{L^2}^2.$$

Continuing to apply Proposition 3.1.6, we see that

$$(1 - \varepsilon^2) \|f\|_{L^2}^2 \leq \sum_{k=0}^K C_M^{-k} \sum_{p \in \mathcal{P}[M]^k} \|U[p]f * f_0\|_{L^2}^2 + C_M^{-K} \sum_{p \in \mathcal{P}[M]^K} \|U[p]f\|_{L^2}^2.$$

Using that $C_M \in (0, 1)$ and Proposition 3.1.3, we have

$$(1 - \varepsilon^2) \|f\|_{L^2}^2 \leq C_M^{-K} \sum_{k=0}^K \sum_{p \in \mathcal{P}[M]^k} \|U[p]f * f_0\|_{L^2}^2 + C_M^{-K} C_{\text{decay}}^{K-1} \|f\|_{L^2}^2.$$

Rearranging, we obtain

$$\begin{aligned} \|\mathcal{S}_{\mathcal{F}}[M, K](f)\|_{L^2 \ell^2}^2 &= \sum_{k=0}^K \sum_{p \in \mathcal{P}[M]^k} \|U[p]f * f_0\|_{L^2}^2 \\ &\geq (C_M^K (1 - \varepsilon^2) - C_{\text{decay}}^{K-1}) \|f\|_{L^2}^2. \end{aligned}$$

Since $C_M \rightarrow 1$ as $M \rightarrow \infty$ and $C_{\text{decay}} \in (0, 1)$ independent of M , for fixed ε ,

we can pick K and M sufficiently large so that $C_M^K (1 - \varepsilon^2) - C_{\text{decay}}^{K-1} > 0$. Note

that this term represents the error due to approximating f by a band-limited

function, the width truncation, and the depth truncation.

(c)-(e) For any $f, g \in L^2(\mathbb{R}^d)$, we have

$$\|\mathcal{S}_{\mathcal{F}}[M, K](f) - \mathcal{S}_{\mathcal{F}}[M, K](g)\|_{L^2 \ell^2}^2 \leq \|\mathcal{S}_{\mathcal{F}}(f) - \mathcal{S}_{\mathcal{F}}(g)\|_{L^2 \ell^2}^2.$$

Applying Theorem 4.3.1 completes the proof.

□

3.3 Comparison and Discussion

In addition to the papers from Mallat’s group [106, 23, 131], we note that Wiatowski and Bölcskei [134] also constructed a scattering-like transform, which they called the *generalized feature extractor* Φ . Additionally, there is a rich history on the approximation properties of neural networks, but this viewpoint is significantly different from the scattering transform theory and there is little connection; nonetheless, see [26, 123] and the references therein for approximation theory results.

- (a) Generality and flexibility. We used Gabor frames as the model example, but our theory applies to any uniform covering frame. This is an important point because Gabor and wavelet frames are generated from a single function and are thus algebraically related; in contrast, the learned filters in CNNs are independent of each other and typically do not satisfy such rigid relationships [99].

So far, the theoretical papers of Mallat and collaborators have exclusively focused on wavelet scattering transforms.

Wiatowski and Bölcskei [134] studied a scattering framework that is more general than ours and Mallat’s. Instead of using the same semi-discrete wavelet frame for each layer of the network, they used (not necessarily tight) semi-discrete frames, and allowed each layer of the transformation to use a different frame. Their theory allowed for a variety of non-linearities at each layer, including the complex modulus, and allowed sub-sampling to be incorporated into

each layer.

- (b) Energy conservation. We showed in Theorem 3.2.1 that $\mathcal{S}_{\mathcal{F}}$ conserves the L^2 norm for any $f \in L^2(\mathbb{R}^d)$.

Mallat showed that $\mathcal{S}_{\mathcal{W}}$ is energy preserving but that result required a restrictive and technical *admissibility condition* on ψ , see [106, pages 1342-1343]. This is completely different from the usual admissibility condition related to the invertibility of the continuous wavelet transform. We cannot offer an intuitive explanation for what Mallat's complicated admissibility condition means. Numerical calculations have supported the assertion that an analytic cubic spline Battle-Lemarié wavelet is admissible for $d = 1$, [106, page 1345]. To our best knowledge, it is currently unknown if other Littlewood-Paley wavelets, such as curvelets [28] or shearlets [79, 80, 39, 40], are admissible.

Mallat's argument for energy conservation is qualitative and cannot be used to deduce a quantitative result because he approximated $f \in L^2(\mathbb{R}^d)$ with a function in the logarithmic Sobolev space. Motivated by this observation, Waldspurger [131, Theorem 3.1] gave mild assumptions on the generating wavelet ψ , see the reference for the explicit hypotheses. Under these assumptions, the following holds: there exists $r > 0$ and $a > 1$ such that for any integer $k \geq 2$ and real-valued $f \in L^2(\mathbb{R})$,

$$\sum_{\lambda \in \Lambda^k} \|U[\lambda]f\|_{L^2}^2 \leq \int_{\mathbb{R}} |\widehat{f}(\xi)|^2 \left(1 - \exp\left(-\frac{2\xi^2}{r^2 a^{2k}}\right)\right) d\xi. \quad (3.3)$$

This inequality quantifies the intuition that wavelet scattering coefficients become progressively concentrated in lower frequency regions. For a ψ satisfying

these assumptions, she showed that the resulting scattering transform conserves energy. However, her result only applies to one-dimensional real-valued functions, but it is possible that they can be adapted to more general situations. In particular, they do not apply to curvelets and shearlets.

In general, Φ is not energy preserving and possibly has trivial kernel. This is not surprising, because the lower bounds on $\|\mathcal{S}_{\mathcal{F}}(f)\|_{L^2\ell^2}$ and $\|\mathcal{S}_{\mathcal{W}}(f)\|_{L^2\ell^2}$ are related to the amount of energy the complex modulus pushes from high to low frequencies from one layer to the next. Thus, it would be surprising if any kind of frame and any type of nonlinearity has the same kind of effect.

- (c) Non-expansiveness. The non-expansiveness property holds for $\mathcal{S}_{\mathcal{F}}$, $\mathcal{S}_{\mathcal{W}}$, and Φ because this is a consequence of the frame property and network structure.
- (d) Translation contraction estimate. Wiatowski and Bölcskei did not provide a translation estimate for Φ .

Our translation estimate, Theorem 4.3.1, is similar to Mallat's translation estimate [106, Theorem 2.10]: There exists a constant $C > 0$ such that for all $f \in L^2(\mathbb{R}^d)$ and $y \in \mathbb{R}^d$,

$$\|\mathcal{S}_{\mathcal{W}}(T_y f) - \mathcal{S}_{\mathcal{W}}(f)\|_{L^2\ell^2} \leq C2^{-J}|y| \left(\sum_{k=0}^{\infty} \sum_{\lambda \in \Lambda^k} \|U[\lambda]f\|_{L^2}^2 \right).$$

To see why, our $\|\nabla f_0\|_{L^1}$ plays the same role as his $C2^{-J}$ because if $f_0(x) = 2^{-dJ}\phi(2^{-J}x)$ for some smooth $\phi \in L^1(\mathbb{R}^d)$, like in Mallat's case, then

$$\|\nabla f_0\|_{L^1} = \|\nabla \phi\|_{L^1} 2^{-J} = C2^{-J}.$$

The only difference is that our inequality is more transparent because it depends on $\|f\|_{L^2}$, whereas Mallat's estimate depends on the more complicated term $\sum_{k=0}^{\infty} \sum_{\lambda \in \Lambda^k} \|U[\lambda]f\|_{L^2}^2$. This term is finite if f belongs to a certain logarithmic Sobolev space and ψ satisfies the admissibility condition.

- (e) Diffeomorphism contraction estimate. Our diffeomorphism estimate, Theorem 4.3.1d, is essentially identical to the corresponding estimate for Φ [134, Theorem 1].

Mallat's diffeomorphism estimate [106, Theorem 2.12] is quite different. It says, for any $\tau \in C^2(\mathbb{R}^d; \mathbb{R}^d)$ with $\|\nabla\tau\|_{L^\infty}$ sufficiently small, there exists $C(J, \tau) > 0$ such that for all $f \in L^2(\mathbb{R}^d)$,

$$\|\mathcal{S}_{\mathcal{W}}(T_\tau f) - \mathcal{S}_{\mathcal{W}}(f)\|_{L^2\ell^2} \leq C(J, \tau) \sum_{k=0}^{\infty} \sum_{\lambda \in \Lambda^k} \|U[\lambda]f\|_{L^2}. \quad (3.4)$$

We caution that this estimate is only meaningful if $\sum_{k=0}^{\infty} \sum_{\lambda \in \Lambda^k} \|U[\lambda]f\|_{L^2} < \infty$; we expect that characterizing this class of functions is a difficult task. A sufficient (but perhaps not necessary) condition for this term being finite is that f belongs to a logarithmic Sobolev space and ψ satisfies the admissibility condition, and in which case, we have

$$\sum_{k=0}^{\infty} \sum_{\lambda \in \Lambda^k} \|U[\lambda]f\|_{L^2} \leq \sum_{k=0}^{\infty} \sum_{\lambda \in \Lambda^k} \|U[\lambda]f\|_{L^2}^2 < \infty.$$

For example, a function belongs to this logarithmic Sobolev space if its average modulus of continuity is bounded, and this condition is much weaker than band-limited. The point here is that, the currently known results for both $\mathcal{S}_{\mathcal{F}}$ and $\mathcal{S}_{\mathcal{W}}$ require additional regularity assumptions on f in order to establish

stability to diffeomorphisms, and removing all regularity assumptions appears to be challenging.

On the other hand, the inequality (3.4) has applications to finite depth wavelet scattering networks. Indeed, if one considers a transform that only includes K layers, then the summation on the right hand side terminates at $k = K$. Then we have

$$C(J, \tau) \sum_{k=0}^K \sum_{\lambda \in \Lambda^k} \|U[\lambda]f\|_{L^2} \leq C(J, \tau)(K + 1)\|f\|_{L^2},$$

and this inequality holds without additional regularity assumptions on f . As we have already mentioned, it is of interest if one could upper bound the left hand side independent of K .

- (f) Rotational invariance. By exploiting that \mathcal{W} is partially generated by a finite rotation group G , Mallat defined a variant of the windowed scattering operator $\tilde{\mathcal{S}}_{\mathcal{W}}$ that is G -invariant: $\tilde{\mathcal{S}}_{\mathcal{W}}(f \circ r) = \tilde{\mathcal{S}}_{\mathcal{W}}(f)$ for all $f \in L^2(\mathbb{R}^d)$ and $r \in G$, see [106, Section 5].

We shall construct a G -invariant rotational Fourier scattering transform $\mathcal{S}^{\mathcal{R}}$ in the next chapter.

In general, it is not possible to modify Φ in order to obtain a G -invariant $\tilde{\Phi}$, because the underlying frame elements need to be “compatible” with the action of G .

- (g) Finite scattering networks. The main advantage of the uniform covering frame approach is that we have a theory for appropriate truncations of $\mathcal{S}_{\mathcal{F}}$, and in par-

ticular, the lower bound in Theorem 3.2.2. Of course, the lower bound requires additional regularity assumptions, but in view of the discussion preceding the statement of the theorem, some kind of assumption is necessary.

In contrast, it is not known whether the analogous result holds for a finite width and depth truncation of \mathcal{S}_W because there is no known quantitative analogues of Propositions 3.1.3 and 3.1.6 for the wavelet case. It might be possible that Waldspurger’s quantitative estimate (3.3) can be used to obtain something similar to Proposition 3.1.3, but as we remarked earlier, her results only hold in one dimension. Wiatowski, Grohs, and Bölcskei [135] used ideas similar to that of Waldspurger’s to prove the exponential decay of energy property for the wavelet case, but their result only holds in dimension $d = 1$ and requires several restrictive assumptions including Sobolev regularity of f . While their results have applications to finite depth networks, they do not address finite width networks.

Finally, there is no analogous energy decay result for the generalized feature extractor Φ . It would be surprising if it were possible to prove that property without additional assumptions on the underlying frame.

Chapter 4: Rotational Fourier Scattering

The material in this chapter contains results from the paper [42], which was written by the author and Wojciech Czaja.

4.1 Rotational UCF

We already saw that certain semi-discrete Gabor frames are UCFs. We construct another important example of a UCF, and this frame is partially generated by rotations and modulation. We shall denote the family of functions by $\mathcal{R} = \mathcal{R}(A, B)$, and this implicitly depends on two fixed parameters: $A > 0$ and an integer $B \geq 1$. For each integer $m \geq 1$, let m^* denote the unique integer of the form 2^k such that

$$m \leq m^* < 2m.$$

Let R_m be the 2×2 counter-clockwise rotation matrix

$$R_m = \begin{pmatrix} \cos(B_m) & -\sin(B_m) \\ \sin(B_m) & \cos(B_m) \end{pmatrix}, \quad \text{where } B_m = \frac{2\pi}{m^*B}.$$

Let $G_m = G_m(B)$ be the finite rotation group generated by the following set of $d \times d$ matrices

$$\left\{ \begin{pmatrix} R_m & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}, \begin{pmatrix} 1 & & & & \\ & R_m & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}, \dots, \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 1 \\ & & & & & R_m \end{pmatrix} \right\}.$$

The identity element of G_m is denoted e . We shall see that G_1 is the most important group in this construction, so to simplify our notation, we write $G = G_1$. If H is a subgroup of G , then we write $H \leq G$. We have the nested subgroup property,

$$G = G_1 \leq G_2 \leq \dots \leq G_m \leq \dots.$$

It follows that for any $n \geq m$, each $r \in G_m$ is a bijection on G_n .

We need to introduce spherical coordinates. Let $\phi_1, \phi_2, \dots, \phi_{d-1}$ be the angular coordinates where $\phi_{d-1} \in [0, 2\pi)$ and $\phi_j \in [0, \pi]$ for $j = 1, 2, \dots, d-2$. If $(\xi_1, \xi_2, \dots, \xi_d)$ are the Euclidean coordinates of $\xi \in \mathbb{R}^d$, then its spherical coordinates are $(\rho, \phi_1, \dots, \phi_{d-1})$, where $\rho = |\xi|$,

$$\begin{aligned} \xi_d &= \rho \prod_{j=1}^{d-1} \sin(\phi_j), \\ \xi_k &= \rho \cos(\phi_k) \prod_{j=1}^{k-1} \sin(\phi_j), \end{aligned}$$

for $k = 1, 2, \dots, d-1$. The following proposition contains two well-known results about the existence of certain cutoff functions whose modulus squared forms a partition of unity.

Proposition 4.1.1 (Hernández and Weiss [83], Chapter 1.3).

(a) For any $A > 0$, there exists a non-negative and even $\eta_A \in C^\infty(\mathbb{R})$ supported in the closed interval $[-A, A]$ such that for all $x \in \mathbb{R}$,

$$\sum_{m \in \mathbb{Z}} |\eta_A(x - Am)|^2 = 1.$$

(b) For any integers $m \geq 1$ and $B \geq 1$, there exists a non-negative $\beta_{m,B} \in C^\infty(\mathbb{S}^{d-1})$, supported in the sector $\{\phi_1 : |\phi_1| \leq B_m\}$, such that for all $\omega \in \mathbb{S}^{d-1}$,

$$\sum_{r \in G_m} |\beta_{m,B}(r\omega)|^2 = 1.$$

We are ready to define the functions in the Fourier domain and in spherical coordinates. For any $\xi \in \mathbb{R}^d$, we can write $\xi = \rho\omega$ where $\rho \geq 0$ and $\omega \in \mathbb{S}^{d-1}$. Let f_0 be the smooth function such that

$$\widehat{f}_0(\xi) = \widehat{f}_0(\rho\omega) = \eta_A(\rho).$$

For each $m \geq 1$ and $r \in G_m$, let $f_{m,r}$ be the smooth function such that

$$\widehat{f}_{m,r}(\xi) = \widehat{f}_{m,r}(\rho\omega) = \eta_A(\rho - mA)\beta_{m,B}(r^{-1}\omega).$$

The next result shows that the sequence of functions,

$$\mathcal{R} = \{f_0\} \cup \{f_{m,r} : m \geq 1, r \in G_m\},$$

is a uniform covering frame, where the index set is

$$\mathcal{G} = \{(m, r) : m \geq 1, r \in G_m\}.$$

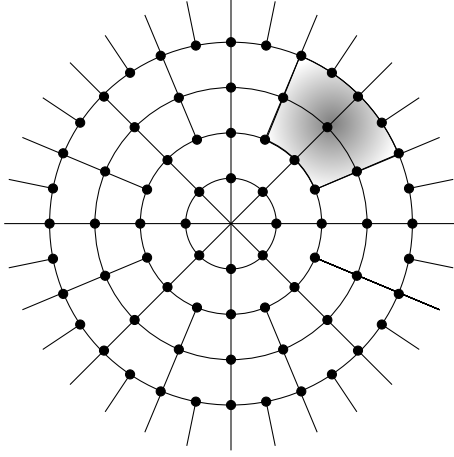


Figure 4.1: Let $d = 2$ and G be the group of rotations by angle $2\pi/8$. The black dots are elements of \mathcal{G} , when embedded in \mathbb{R}^2 , for the first four uniform Fourier scales, and the shaded gray region is the support of $(f_{3,2\pi/8})^\wedge$.

Theorem 4.1.2. *Let $\mathcal{R} = \{f_0\} \cup \{f_{m,r}\}_{(m,r) \in \mathcal{G}}$ be the sequence of functions defined above. Then, \mathcal{R} is a uniform covering frame for $L^2(\mathbb{R}^d)$.*

Proof. By construction, $\widehat{f_0}$ and $\widehat{f_p}$ are supported in a compact and connected sets. Note that $\widehat{f_0}(0) = \eta_A(0) = 1$, and since f_0 and f_p are Schwartz functions, they belong to $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d) \cap C^1(\mathbb{R}^d)$.

We check that the frame condition holds. By construction, we have

$$\begin{aligned} |\widehat{f_0}(\xi)|^2 + \sum_{m=1}^{\infty} \sum_{r \in G_m} |\widehat{f_{m,r}}(\xi)|^2 \\ = |\eta_A(\rho)|^2 + \sum_{m=1}^{\infty} |\eta_A(\rho - Am)|^2 \sum_{r \in G_m} |\beta_{m,B}(r^{-1}\omega)|^2 = 1, \end{aligned}$$

for all $\xi \in \mathbb{R}^d$. See Figure 4.1 for a visualization of \mathcal{G} and the tiling properties of \mathcal{R} .

It remains check that the uniform covering property holds. For each $m \geq 1$ and $r \in G_m$, since $\text{supp}(\widehat{f_{m,r}})$ is a rotation of $\text{supp}(\widehat{f_{m,e}})$, it suffices to check the uniform

covering property for the family of sets $\{\widehat{\text{supp}}(f_{m,e}): m \geq 1\}$. Further, it suffices to check the uniform covering property for the subset $\{\widehat{\text{supp}}(f_{m,e}): m^*B \geq 4\}$, since the complement of this set is $\{\widehat{\text{supp}}(f_{m,e}): m^*B < 4\}$, which has finite cardinality.

From here onwards, we assume $m^*B \geq 4$, or equivalently, $B_m \leq \pi/2$. Observe that $\widehat{f_{m,e}}$ is supported in the wedge,

$$W_m = \left\{ (\rho, \phi_1, \phi_2, \dots, \phi_{d-1}) : |\rho - Am| \leq A, |\phi_1| \leq B_m \right\}.$$

To show that $\{W_m: m^*B \geq 4\}$ satisfies the uniform covering property, it suffices to show that the maximum distance between any two points in W_m is bounded uniformly in m .

Let $\xi, \zeta \in W_m$ and let $(\rho, \phi_1, \phi_2, \dots, \phi_{d-1})$ and $(\gamma, \theta_1, \theta_2, \dots, \theta_{d-1})$ be their spherical coordinates, respectively. We have

$$|\xi_1 - \zeta_1| = |\rho \cos(\phi_1) - \gamma \cos(\theta_1)|.$$

Since $|\phi_1| \leq B_m \leq \pi/2$ and similarly for θ_1 , we see that

$$|\xi_1 - \zeta_1| \leq A(m+1) - A(m-1)\cos(B_m).$$

See Figure 5.1 for an illustration of this inequality. Using the standard trigonometric inequality, $|1 - \cos(t)| \leq t^2/2$ for all $t \in \mathbb{R}$, we see that

$$|\xi_1 - \zeta_1| \leq 2A + \frac{Am}{2} \left(\frac{2\pi}{m^*B} \right)^2 = 2A + \frac{2\pi^2 A}{B^2} \frac{m}{(m^*)^2}.$$

For $k = 2, 3, \dots, d-1$, we have

$$\begin{aligned} |\xi_k - \zeta_k| &= \left| \rho \cos(\phi_k) \prod_{j=1}^{k-1} \sin(\phi_j) - \gamma \cos(\theta_k) \prod_{j=1}^{k-1} \sin(\theta_j) \right| \\ &\leq \rho |\sin(\phi_1)| + \gamma |\sin(\theta_1)|. \end{aligned}$$

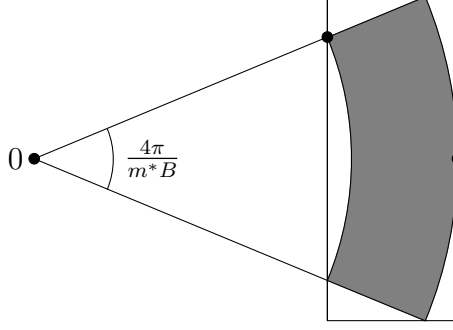


Figure 4.2: The shaded region represents all possible values of (ρ, ϕ_1) and (γ, θ_1) , where $A(m-1) \leq \rho \leq A(m+1)$ and $|\theta_1| \leq B_m$, and likewise for γ and ϕ_1 . The maximum horizontal distance between any points in this wedge is the width of the rectangle, and this distance is attained by the points $(A(m+1), 0)$ and $(A(m-1)\cos(2\pi/(m^*B)), A(m-1)\sin(2\pi/(m^*B)))$, located at the two black dots on the perimeter of the wedge.

Using that $\rho \leq A(m+1)$, the trigonometric inequality $|\sin(\phi_1)| \leq |\phi_1| \leq B_m$, and similarly for $|\sin(\theta_1)|$, we have

$$|\xi_k - \zeta_k| \leq 2A(m+1)B_m \leq \frac{4\pi A}{B} \frac{m+1}{m^*},$$

for $k = 2, 3, \dots, d-1$. The same argument shows that

$$|\xi_d - \zeta_d| \leq \rho|\sin(\phi_1)| + \gamma|\sin(\theta_1)| \leq \frac{4\pi A}{B} \frac{m+1}{m^*}.$$

The above inequalities imply the Euclidean distance between any two points in W_m is bounded uniformly in m , which completes the proof.

□

Definition 4.1.3. Fix $A > 0$ and an integer $B \geq 1$. Let

$$\mathcal{R} = \mathcal{R}(A, B) = \{f_0\} \cup \{f_{m,r}\}_{(m,r) \in \mathcal{G}}$$

be the sequence of functions constructed in this section. We call \mathcal{R} a *rotational uniform covering frame*.

Remark 4.1.4. Observe that $\widehat{f_{m,r}}(\xi) = \widehat{f_{m,e}}(r^{-1}\xi)$, so $f_{m,r}$ is a rotation of $f_{m,e}$. We also have $\widehat{f_{m+1,e}}(\xi) = \widehat{f_{1,e}}(\xi - mA)$, so each $f_{m,e}$ is a modulation of $f_{1,e}$. For this reason, we say rotational uniform covering frames are generated by rotations and modulations.

4.2 Rotational Fourier Scattering

If the uniform covering frame has additional group structure, it is possible to exploit that to construct a group invariant scattering transform. Let $\mathcal{R} = \{f_0\} \cup \{f_{m,r}\}_{(m,r) \in \mathcal{G}}$ be a rotational uniform covering frame and recall that its index set is $\mathcal{G} = \{(m,r): m \geq 1, r \in G_m\}$ with $G = G_1$. In order to derive a time-frequency scattering transform that is invariant under rotations G , we carefully study its group action on \mathcal{R} . This is carried out in the subsequent steps.

We define the following left and right group actions of G on \mathcal{G} . For each $r \in G$ and $(m,s) \in \mathcal{G}$, let $r(m,s) = (m,rs)$ and $(m,s)r = (m,sr)$. Both operations are well-defined in view of the nested subgroup property for $\{G_m\}$. For any $k \geq 1$, we define the left and right group actions of G on \mathcal{G}^k by extending the above definitions. Indeed, each $r \in G$ acts on $p \in \mathcal{G}^k$ on the left according to

$$r(p_1, p_2, \dots, p_k) = (rp_1, rp_2, \dots, rp_k),$$

and similarly for its right action.

The previous observations lead us to the following decomposition. We saw that G is a group action on \mathcal{G}^k . Then, there exists a set $\mathcal{Q}_k \subseteq \mathcal{G}^k$ such that we have the disjoint union

$$\mathcal{G}^k = \bigcup_{r \in G} r \mathcal{Q}_k. \quad (4.1)$$

Explicitly, we have $\mathcal{Q}_k = \mathcal{G}_0 \times \mathcal{G}^{k-1}$, where $\mathcal{G}_0 = \{(m, r): m \geq 1, r \in G_m/G\}$ and G_m/G is the quotient group, G_m modulo G .

We can consider the action of G on \mathcal{R} . Each $r \in G$ acts on $L^2(\mathbb{R}^d)$ by the rule

$$r: f(x) \mapsto f_r(x) = f(rx).$$

Using the definition of $f_{m,s}$, for any $r \in G$,

$$(f_{m,s})_r(x) = f_{m,s}(rx) = f_{m,r^{-1}s}(x).$$

This defines a group action of G on \mathcal{R} . Further, this calculation shows that the invariant subsets of \mathcal{R} under G are $\{f_0\}$ and $\{f_{m,s}: s \in G_m\}$, for each $m \geq 1$. For any $r \in G$ and $(m, s) \in \mathcal{G}$, by the previous identity and a change of variables,

$$(f_r * f_{m,s})(x) = (f * (f_{m,s})_{r^{-1}})(rx) = (f * f_{m,rs})_r(x).$$

By iterating this identity, for all integers $k \geq 1$ and $p \in \mathcal{G}^k$, we obtain

$$(U[p]f_r)(x) = (U[rp]f)_r(x).$$

By definition, f_0 is rotationally invariant, so we see that

$$(U[p]f_r * f_0)(x) = ((U[rp]f)_r * f_0)(x) = (U[rp]f * f_0)_r(x). \quad (4.2)$$

We now carry out the same steps but for specific finite subsets of \mathcal{G}^k and \mathcal{Q}_k .

Fix integers $k \geq 1$ and $M \geq 1$. We define the finite set

$$\mathcal{G}[M] = \{(m, r) : 1 \leq m \leq M, r \in G_m\}.$$

By construction of \mathcal{R} , we have

$$|\widehat{f}_0(\xi)|^2 + \sum_{p \in \mathcal{G}[M]} |\widehat{f}_p(\xi)|^2 = \begin{cases} 1 & \text{if } \xi \in \overline{B_{AM}(0)}, \\ 0 & \text{if } \xi \notin \overline{B_{AM}(0)}. \end{cases} \quad (4.3)$$

Let $\mathcal{G}[M]^k$ be the product of $\mathcal{G}[M]$ with itself k times. Using the same definition as before, we see that G is a group action on $\mathcal{G}[M]^k$. Similar to before, there exists a finite set $\mathcal{Q}[M, k] \subseteq \mathcal{G}[M]^k$ such that we have the disjoint union

$$\mathcal{G}[M]^k = \bigcup_{r \in G} r\mathcal{Q}[M, k]. \quad (4.4)$$

Explicitly, we have $\mathcal{Q}[M, k] = \mathcal{G}_0[M] \times \mathcal{G}[M]^k$ where

$$\mathcal{G}_0[M] = \{(m, r) : 1 \leq m \leq M, r \in G_m/G\}.$$

For all integers $k \geq 1$ and $p \in \mathcal{G}[M]^k$, the same argument as before shows that

$$(U[p]f_r * f_0)(x) = (U[rp]f * f_0)_r(x). \quad (4.5)$$

Definition 4.2.1. Given a rotational uniform covering frame \mathcal{R} , the associated *rotational Fourier scattering transform* $\mathcal{S}^{\mathcal{R}}$ is formally defined as

$$\begin{aligned} \mathcal{S}^{\mathcal{R}}(f) &= \left\{ |G|^{-1/2} \|(f * f_0)_r\|_{\ell_r^2(G)} \right\} \cup \left\{ \|(U[rq]f * f_0)_r\|_{\ell_r^2(G)} : q \in \mathcal{Q}_k, k \geq 1 \right\} \\ &= \left\{ |G|^{-1/2} \left(\sum_{r \in G} |(f * f_0)_r|^2 \right)^{1/2} \right\} \cup \left\{ \left(\sum_{r \in G} |(U[rq]f * f_0)_r|^2 \right)^{1/2} : q \in \mathcal{Q}_k, k \geq 1 \right\}. \end{aligned}$$

Definition 4.2.2. Given a rotational uniform covering frame \mathcal{R} and integers $M \geq 1$ and $K \geq 1$, the associated *truncated rotational Fourier scattering transform* $\mathcal{S}^{\mathcal{R}}[M, K]$ is formally defined as

$$\begin{aligned} & \mathcal{S}^{\mathcal{R}}[M, K](f) \\ &= \left\{ |G|^{-1/2} \|(f * f_0)_r\|_{\ell_r^2(G)} \right\} \cup \left\{ \|(U[rq]f * f_0)_r\|_{\ell_r^2(G)} : q \in \mathcal{Q}[M, k], 1 \leq k \leq K \right\}. \end{aligned}$$

Remark 4.2.3. We explain our choice of notation. Throughout this document, we used \mathcal{F} to denote a generic uniform covering frame whereas we used \mathcal{R} to mean the rotational variant. By assumption, $\mathcal{S}_{\mathcal{F}}$ denotes the Fourier scattering transform and the subscript emphasizes that it depends on a fixed \mathcal{F} . Similarly, the super-script in $\mathcal{S}^{\mathcal{R}}$ indicates that the rotational Fourier scattering transform depends on a fixed \mathcal{R} . By Theorem 4.1.2, \mathcal{R} is a perfectly valid uniform covering frame, so it can be used as the underlying frame in the Fourier scattering transform, and this operator is denoted $\mathcal{S}_{\mathcal{R}}$. However, we emphasize that while both $\mathcal{S}_{\mathcal{R}}$ and $\mathcal{S}^{\mathcal{R}}$ use the same frame \mathcal{R} , they are *distinct* operators, and likewise for their truncations. This can be easily seen from their definitions or by comparing their network structures. While the sub-script and super-script notation might seem odd, we use them to differentiate between these two operators. In the next two propositions, we derive their quantitative relationship.

Proposition 4.2.4. Let $\mathcal{R} = \{f_0\} \cup \{f_{m,r}\}_{(m,r) \in \mathcal{G}}$ be a rotational uniform covering

frame. For all $f, g \in L^2(\mathbb{R}^d)$, we have

$$\begin{aligned}\|\mathcal{S}^{\mathcal{R}}(f)\|_{L^2\ell^2} &= \|\mathcal{S}_{\mathcal{R}}(f)\|_{L^2\ell^2}, \\ \|\mathcal{S}^{\mathcal{R}}(f) - \mathcal{S}^{\mathcal{R}}(g)\|_{L^2\ell^2} &\leq \|\mathcal{S}_{\mathcal{R}}(f) - \mathcal{S}_{\mathcal{R}}(g)\|_{L^2\ell^2}.\end{aligned}$$

Proof. To prove the equality, we use the decomposition (4.1) to obtain

$$\begin{aligned}\|\mathcal{S}^{\mathcal{R}}(f)\|_{L^2\ell^2}^2 &= \frac{1}{|G|} \sum_{r \in G} \|(f * f_0)_r\|_{L^2}^2 + \sum_{k=1}^{\infty} \sum_{q \in \mathcal{Q}_k} \sum_{r \in G} \|(U[rq]f * f_0)_r\|_{L^2}^2 \\ &= \|f * f_0\|_{L^2}^2 + \sum_{k=1}^{\infty} \sum_{q \in \mathcal{Q}_k} \sum_{r \in G} \|U[rq]f * f_0\|_{L^2}^2 \\ &= \|f * f_0\|_{L^2}^2 + \sum_{k=1}^{\infty} \sum_{p \in \mathcal{G}^k} \|U[p]f * f_0\|_{L^2}^2 = \|\mathcal{S}_{\mathcal{R}}(f)\|_{L^2\ell^2}^2.\end{aligned}$$

To prove the inequality in the proposition, we apply the reverse triangle inequality for the $\ell^2(G)$ norm in the definition of $\mathcal{S}^{\mathcal{R}}$ and the decomposition (4.1),

$$\begin{aligned}\|\mathcal{S}^{\mathcal{R}}(f) - \mathcal{S}^{\mathcal{R}}(g)\|_{L^2\ell^2}^2 &\leq \frac{1}{|G|} \sum_{r \in G} \|(f * f_0)_r - (g * f_0)_r\|_{L^2}^2 \\ &\quad + \sum_{k=1}^{\infty} \sum_{q \in \mathcal{Q}_k} \sum_{r \in G} \|(U[rq]f * f_0)_r - (U[rq]g * f_0)_r\|_{L^2}^2 \\ &= \|f * f_0 - g * f_0\|_{L^2}^2 \\ &\quad + \sum_{k=1}^{\infty} \sum_{q \in \mathcal{Q}_k} \sum_{r \in G} \|U[rq]f * f_0 - U[rq]g * f_0\|_{L^2}^2 \\ &= \|f * f_0 - g * f_0\|_{L^2}^2 + \sum_{k=1}^{\infty} \sum_{p \in \mathcal{G}^k} \|U[p]f * f_0 - U[p]g * f_0\|_{L^2}^2 \\ &= \|\mathcal{S}_{\mathcal{R}}(f) - \mathcal{S}_{\mathcal{R}}(g)\|_{L^2\ell^2}^2.\end{aligned}$$

□

Proposition 4.2.5. *Let $\mathcal{R} = \{f_0\} \cup \{f_{m,r}\}_{(m,r) \in \mathcal{G}}$ be a rotational uniform covering frame. There exists $C = C_{\text{tiling}} > 0$ such that for any integers $M, K \geq 1$ and integer*

$N \leq CM$,

$$\|\mathcal{S}^{\mathcal{R}}[M, K](f)\|_{L^2\ell^2} \geq \|\mathcal{S}_{\mathcal{R}}[N, K](f)\|_{L^2\ell^2}.$$

Proof. By comparing the identities (2.2) and (4.3), we see that there exists a constant $C_{\text{tiling}} > 0$ such that whenever $N \leq C_{\text{tiling}}M$, we have $\mathcal{P}[N]^k \subseteq \mathcal{G}[M]^k$. This fact and the disjoint decomposition (4.4) imply

$$\begin{aligned} \|\mathcal{S}^{\mathcal{R}}[M, K](f)\|_{L^2\ell^2}^2 &= \frac{1}{|G|} \sum_{r \in G} \|(f * f_0)_r\|_{L^2}^2 + \sum_{k=1}^K \sum_{q \in \mathcal{Q}[M, k]} \sum_{r \in G} \|(U[rq]f * f_0)_r\|_{L^2}^2 \\ &= \|f * f_0\|_{L^2}^2 + \sum_{k=1}^K \sum_{p \in \mathcal{G}[M]^k} \|U[p]f * f_0\|_{L^2}^2 \\ &\geq \|f * f_0\|_{L^2}^2 + \sum_{k=1}^K \sum_{p \in \mathcal{P}[N]^k} \|U[p]f * f_0\|_{L^2}^2 \\ &= \|\mathcal{S}_{\mathcal{R}}[N, K](f)\|_{L^2\ell^2}^2. \end{aligned}$$

□

4.3 Rotational Invariance and Properties

We are ready to state and prove the main theorem, which shows that $\mathcal{S}^{\mathcal{R}}$ and $\mathcal{S}^{\mathcal{R}}[M, K]$ are effective feature extractors. We already did most of the work in the construction of $\mathcal{S}^{\mathcal{R}}$ and $\mathcal{S}^{\mathcal{R}}[M, K]$, and in proving Propositions 4.2.4 and 4.2.5. The basic strategy is to quantitatively relate $\mathcal{S}^{\mathcal{R}}$ to $\mathcal{S}_{\mathcal{R}}$ and then use known results about the latter.

Theorem 4.3.1. *Let $\mathcal{R} = \{f_0\} \cup \{f_{m,r}\}_{(m,r) \in \mathcal{G}}$ be a rotational uniform covering frame.*

(a) G-invariance: For all integers $M, K \geq 1$, $f \in L^2(\mathbb{R}^d)$, and $r \in G$,

$$\mathcal{S}^{\mathcal{R}}(f_r) = \mathcal{S}^{\mathcal{R}}(f) \quad \text{and} \quad \mathcal{S}^{\mathcal{R}}[M, K](f_r) = \mathcal{S}^{\mathcal{R}}[M, K](f).$$

(b) Upper bound: For all integers $M, K \geq 1$ and $f \in L^2(\mathbb{R}^d)$,

$$\|\mathcal{S}^{\mathcal{R}}[M, K](f)\|_{L^2\ell^2} \leq \|\mathcal{S}^{\mathcal{R}}(f)\|_{L^2\ell^2} = \|f\|_{L^2}.$$

(c) Lower bound: Let $\varepsilon \in [0, 1)$ and $R > 0$. For sufficiently large $M, K \geq 1$, there exists $C \in (0, 1)$ depending on $\mathcal{R}, \varepsilon, M, K$, such that for all (ε, R) band-limited functions $f \in L^2(\mathbb{R}^d)$,

$$\|\mathcal{S}^{\mathcal{R}}[M, K](f)\|_{L^2\ell^2} \geq C\|f\|_{L^2}.$$

(d) Non-expansiveness: For all integers $M, K \geq 1$ and $f, g \in L^2(\mathbb{R}^d)$,

$$\|\mathcal{S}^{\mathcal{R}}[M, K](f) - \mathcal{S}^{\mathcal{R}}[M, K](g)\|_{L^2\ell^2} \leq \|\mathcal{S}^{\mathcal{R}}(f) - \mathcal{S}^{\mathcal{R}}(g)\|_{L^2\ell^2} \leq \|f - g\|_{L^2}.$$

(e) Translation contraction: There exists $C > 0$ depending only on \mathcal{R} such that for all integers $M, K \geq 1$, $f \in L^2(\mathbb{R}^d)$, and $y \in \mathbb{R}^d$,

$$\begin{aligned} & \|\mathcal{S}^{\mathcal{R}}[M, K](T_y f) - \mathcal{S}^{\mathcal{R}}[M, K](f)\|_{L^2\ell^2} \\ & \leq \|\mathcal{S}^{\mathcal{R}}(T_y f) - \mathcal{S}^{\mathcal{R}}(f)\|_{L^2\ell^2} \\ & \leq C|y| \|\nabla f_0\|_{L^1} \|f\|_{L^2}. \end{aligned}$$

(f) Additive diffeomorphism contraction: There exists a constant $C > 0$ such that for all $\varepsilon \in [0, 1)$, $R > 0$, (ε, R) band-limited $f \in L^2(\mathbb{R}^2)$, and $\tau \in C^1(\mathbb{R}^2; \mathbb{R}^2)$

with $\|\nabla\tau\|_{L^\infty} \leq 1/4$,

$$\begin{aligned} & \|\mathcal{S}^{\mathcal{R}}[M, K](f \circ (I - \tau)) - \mathcal{S}^{\mathcal{R}}[M, K](f)\|_{L^2\ell^2} \\ & \leq \|\mathcal{S}^{\mathcal{R}}(f \circ (I - \tau)) - \mathcal{S}^{\mathcal{R}}(f)\|_{L^2\ell^2} \\ & \leq C(R\|\tau\|_{L^\infty} + \varepsilon)\|f\|_{L^2}. \end{aligned}$$

Proof.

(a) For the first term in $\mathcal{S}^{\mathcal{R}}(f_r)$ and $\mathcal{S}^{\mathcal{R}}[M, K](f_r)$, we use that f_0 is rotationally invariant to obtain

$$\sum_{s \in G} |(f_r * f_0)_s|^2 = \sum_{s \in G} |(f * f_0)_{rs}|^2 = \sum_{s \in G} |(f * f_0)_s|^2.$$

For the remaining terms in $\mathcal{S}^{\mathcal{R}}(f_r)$, fix an integer $k \geq 1$ and $q \in \mathcal{Q}_k$. By identity (4.2) and re-indexing the following sum, we have

$$\sum_{s \in G} |(U[sq]f_r * f_0)_s|^2 = \sum_{s \in G} |(U[rsq]f * f_0)_{rs}|^2 = \sum_{s \in G} |(U[sq]f * f_0)_s|^2.$$

This proves that $\mathcal{S}^{\mathcal{R}}(f_r) = \mathcal{S}^{\mathcal{R}}(f)$. For the remaining terms in $\mathcal{S}^{\mathcal{R}}[M, K](f_r)$, fix an integer $1 \leq k \leq K$ and $q \in \mathcal{Q}[M, k]$. By identity (4.5) and re-indexing the following sum, we have

$$\sum_{s \in G} |(U[sq]f_r * f_0)_s|^2 = \sum_{s \in G} |(U[rsq]f * f_0)_{rs}|^2 = \sum_{s \in G} |(U[sq]f * f_0)_s|^2.$$

This proves that $\mathcal{S}^{\mathcal{R}}[M, K](f_r) = \mathcal{S}^{\mathcal{R}}[M, K](f)$ and $\mathcal{S}^{\mathcal{R}}(f_r) = \mathcal{S}^{\mathcal{R}}(f)$.

(b) By definition and Proposition 4.2.4, we have

$$\|\mathcal{S}^{\mathcal{R}}[M, K](f)\|_{L^2\ell^2} \leq \|\mathcal{S}^{\mathcal{R}}(f)\|_{L^2\ell^2} = \|\mathcal{S}_{\mathcal{R}}(f)\|_{L^2\ell^2}.$$

We apply Theorem 3.2.1 to complete the proof.

(c) By Proposition 4.2.5, there exists $C_{\text{tiling}} > 0$ such that for all $N \leq C_{\text{tiling}}M$, we have

$$\|\mathcal{S}^{\mathcal{R}}[M, K](f)\|_{L^2\ell^2} \geq \|\mathcal{S}_{\mathcal{R}}[N, K](f)\|_{L^2\ell^2}.$$

Using the lower bound for $\mathcal{S}_{\mathcal{R}}[N, K]$ in Theorem 3.2.2 completes the proof.

(d) By definition and Proposition 4.2.4, we have

$$\|\mathcal{S}^{\mathcal{R}}[M, K](f) - \mathcal{S}^{\mathcal{R}}[M, K](g)\|_{L^2\ell^2} \leq \|\mathcal{S}_{\mathcal{R}}(f) - \mathcal{S}_{\mathcal{R}}(g)\|_{L^2\ell^2}.$$

Since $\mathcal{S}_{\mathcal{R}}$ is non-expansive, see Theorem 3.2.1, the result follows.

(e) By definition and Proposition 4.2.4,

$$\|\mathcal{S}^{\mathcal{R}}[M, K](T_y f) - \mathcal{S}^{\mathcal{R}}[M, K](f)\|_{L^2\ell^2} \leq \|\mathcal{S}_{\mathcal{R}}(T_y f) - \mathcal{S}_{\mathcal{R}}(f)\|_{L^2\ell^2}.$$

The rest follows by using the translation estimate for $\mathcal{S}_{\mathcal{R}}$, see Theorem 3.2.1.

(f) Again, we use the definition and Proposition 4.2.4 to deduce

$$\|\mathcal{S}^{\mathcal{R}}[M, K](f \circ (I - \tau)) - \mathcal{S}^{\mathcal{R}}[M, K](f)\|_{L^2\ell^2} \leq \|\mathcal{S}_{\mathcal{R}}(f \circ (I - \tau)) - \mathcal{S}_{\mathcal{R}}(f)\|_{L^2\ell^2}.$$

We use the diffeomorphism estimate, Theorem 3.2.1, to complete the proof.

□

Remark 4.3.2. We saw in the proof of Theorem 4.3.1 that if we replaced the $\ell^2(G)$ norm in the definition of $\mathcal{S}^{\mathcal{R}}$ and $\mathcal{S}^{\mathcal{R}}[M, K]$ with the $\ell^p(G)$ norm for any $1 \leq p < \infty$, then the resulting operators would still be G -invariant. Additionally, since G is a finite set, the $\ell^p(G)$ and $\ell^2(G)$ norms are equivalent up to constants depending on p and $|G|$, so our estimates carry generalize to the $\ell^p(G)$ case. We chose the $\ell^2(G)$

norm in the definitions of $\mathcal{S}^{\mathcal{R}}$ and $\mathcal{S}^{\mathcal{R}}[M, K]$ because it is more natural to work with a Hilbert space as opposed to a Banach space.

4.4 Relationship to Directional Representations

While our construction of rotational uniform covering frames is motivated by neural networks, since they are also time-frequency representations and are partially generated by rotations, it is natural to ask whether they are related to recent developments in directional Fourier analysis. In order to make precise comparisons, we first summarize several important works on directional Fourier and wavelet analysis. See the theses [111, 132] for additional discussion on directional representations and applications.

- (a) Candès introduced a wavelet system that decomposes a function according to its direction, scale, and location. For an appropriate non-zero function ψ , see [26, Definition 1], and $(a, b, u) \in \mathbb{R}^+ \times \mathbb{R} \times \mathbb{S}^{d-1}$, a *ridgelet* is

$$\psi_{a,b,u}(x) = a^{-1/2}\psi(a^{-1}(u \cdot x - b)). \quad (4.6)$$

To see why this function is directionally sensitive, observe that $\psi_{a,b,u}$ is constant on hyperplanes perpendicular to u and oscillates in the direction of u . Candès also introduced a discrete ridgelet system, and proved that any $L^2(\mathbb{R}^d)$ can be reconstructed from its continuous and discrete coefficients, see [26, Theorems 1 and 2] for precise statements.

- (b) In contrast to the above wavelet inspired approach, Grafakos and Sansing con-

structed directional time-frequency representations. For a non-zero $g \in \mathcal{S}(\mathbb{R})$ and $(m, t, u) \in \mathbb{R} \times \mathbb{R} \times \mathbb{S}^{d-1}$, a *weighted Gabor ridge function* is

$$g_{m,t,u}(x) = D^{(n-1)/2}(g_{m,t})(u \cdot x), \quad \text{where} \quad g_{m,t}(s) = e^{2\pi i m \cdot (s-t)} g(s-t),$$

and $D^{(n-1)/2}$ is the Fourier multiplier with symbol $|\xi|^{(n-1)/2}$. Similar to ridgelets, a weighted Gabor ridge function is constant along hyperplanes perpendicular to u and oscillates in the direction of u . They established a $L^2(\mathbb{R}^d)$ reconstruction formula for this continuous family, see [75, Theorem 3], but were unable to obtain a discrete Parseval frame from this continuous family. This discretization problem was partially resolved: by omitting the multiplier $D^{(n-1)/2}$, it is possible to obtain a discrete frame, consisting of (un-weighted) Gabor ridge functions, for certain subspaces of $L^2(\mathbb{R}^d)$, see [43, Theorem 5.6] for a precise statement.

- (c) While both of the previous representations relied on the ridge function $x \mapsto x \cdot u$, Candès and Donoho used an entirely different approach to construct a family of wavelets called *curvelets* [28]. They are constructed using a decomposition of the Fourier domain in the same spirit as the *second dyadic decomposition* [125, pages 377 and 403]. Each curvelet oscillates in a certain direction and its shape satisfies the anisotropic scaling relation, width \sim length². We omit their precise definitions since they are quite technical to state and such details are not relevant to our current discussion.
- (d) Finally, shearlets [96, 80, 39, 40, 20, 112] are also wavelets that extract directional information from functions. Unlike curvelets, they are generated using shearing operations as opposed to rotations.

With these examples of directional representations in mind, we return our attention to RUCFs. We first show that each frame element oscillates in a certain direction, which is not surprising since they are partially generated by rotations. We already observed that $f_p \in \mathcal{S}(\mathbb{R}^d)$, so by Fourier inversion and a change of variables, for all $x \in \mathbb{R}^d$,

$$f_{m,r}(x) = \int_0^\infty \int_{\mathbb{S}^{d-1}} e^{2\pi i \rho(x \cdot r\omega)} \eta_A(\rho - Am) \beta_{m,B}(\omega) \rho^{d-1} d\sigma(\omega) d\rho, \quad (4.7)$$

where σ is the surface measure of \mathbb{S}^{d-1} . Recall that $\beta_{m,B}$ is supported in the cap $\{\phi_1: |\phi_1| \leq B_m\}$, which implies the above integral is taken over a subset of the sphere where $\omega \sim e_1$. We consider two separate cases.

- (a) If $x \in \mathbb{R}^d$ is parallel to re_1 , then $x \cdot r\omega \sim 1$ and the phase in the integrand of (4.7) rapidly changes. Since η_A and $\beta_{m,B}$ are non-negative, we expect $f_{m,r}$ to oscillate in the direction of re_1 . Further, due to the compact support of η_A , we have $\rho \sim mA$, so $f_{m,r}$ oscillates at frequency approximately mA .
- (b) If $x \in \mathbb{R}^d$ is perpendicular to re_1 , then $x \cdot r\omega \sim 0$, and the phase in the integrand of (4.7) changes slowly. Since η_A and $\beta_{m,B}$ are essentially constant, we do not expect $f_{m,r}$ to oscillate in directions perpendicular to re_1 .

Figure 4.3 contains several numerically computed examples of $f_{m,r}$ and they agree with the qualitative description that we presented.

In view of the directional bias of the frame elements, one might suspect that the frame coefficient $f * f_{m,r}$ carries directional frequency information about f . Indeed,

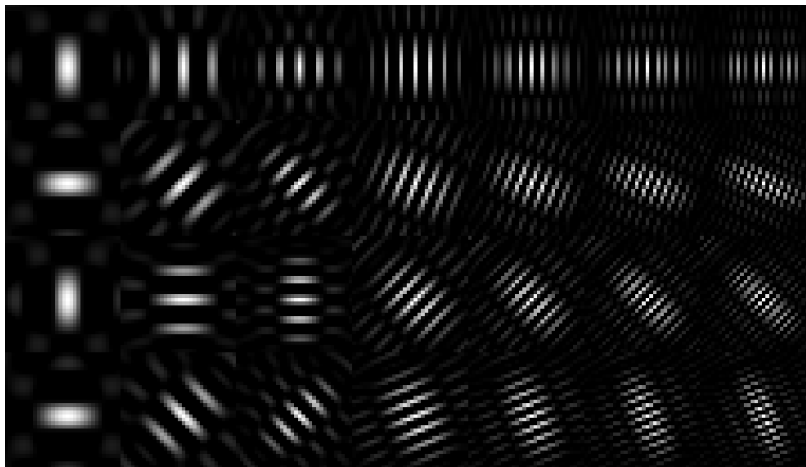


Figure 4.3: Let $d = 2$ and G be the group generated by rotations by $2\pi/4$. The figures are zoomed-in intensity plots of the rotational uniform covering frame elements. The (n, m) -th image corresponds to f_{m,r_n} where $r_n = 2\pi(n - 1)/(4m^*)$.

by Plancherel,

$$\|f * f_{m,r}\|_{L^2}^2 = \int_0^\infty \int_{\mathbb{S}^{d-1}} |\widehat{f}(\rho r \omega)|^2 |\eta_A(\rho - mA)|^2 |\beta_{m,B}(\omega)|^2 \rho^{d-1} d\sigma(\omega) d\rho.$$

By the definitions of η_A and $\beta_{m,B}$, the integral is taken over the region where $\rho \sim mA$ and $\omega \sim e_1$. This equation shows that $\|f * f_{m,r}\|_{L^2}$ is a weighted average of how much f oscillates at frequencies roughly mA and in approximately the re_1 direction.

We have shown that rotational uniform covering frame elements are directionally sensitive and their frame coefficients carry directional information. These properties are also shared by ridgelets, weighted Gabor ridge functions, and curvelets. However, since ridgelets and weighted Gabor ridge functions are constant on hyperplanes, they are not in $L^p(\mathbb{R}^d)$ for any $0 < p < \infty$. In contrast, curvelets and rotational uniform covering frame elements are smooth and well-localized in both space and frequency, and consequently, they belong to many popular function spaces.

Of course, curvelets are not uniform covering frames, but their constructions have significant similarities. In view of these observations, it is reasonable to say that Gabor ridge functions are time-frequency analogues of ridgelets, while rotational uniform covering frame elements are time-frequency analogues of curvelets.

Chapter 5: Applications of Fourier Scattering

Some of the material in this chapter also appeared in the papers [41, 42] by the author and Wojciech Czaja, while the remainder of the material has not yet been published.

5.1 The Discretization Problem

For data analysis and signal processing applications, all functions must be converted into arrays. We show how to construct discrete and finite uniform covering frames.

Let \mathcal{F} be a UCF with index set \mathcal{P} , and note that \mathcal{F} does not necessarily have to be a RUCF. Recall that a *discrete frame* for $L^2(\mathbb{R}^d)$ is a sequence of functions $\{\phi_j\}_{j \in \mathcal{J}} \subseteq L^2(\mathbb{R}^d)$ indexed by a possibly uncountable index set \mathcal{J} such that there exist $0 < A \leq B < \infty$ and for all $f \in L^2(\mathbb{R}^d)$,

$$A\|f\|_{L^2}^2 \leq \sum_{j \in \mathcal{J}} |\langle f, \phi_j \rangle|^2 \leq B\|f\|_{L^2}^2.$$

Similar to the semi-discrete case, there is an associated reproducing formula for recovering f , see [58, 47].

Semi-discrete Gabor frames arose in sampling the frequencies of the STFT, while discrete Gabor frames arose in sampling both the space and frequency domain of the STFT. Consider the discrete set $\{(x_j, \xi_j)\}_{j \in \mathcal{J}}$ and let \tilde{g} be the involution of

g. By a similar calculation, we have

$$V_g f(x_j, \xi_j) = e^{-2\pi i x_j \cdot \xi_j} \int_{\mathbb{R}^d} f(y) \overline{(T_{x_j} M_{\xi_j} \tilde{g})(y)} dy.$$

Again, the phase factor $e^{-2\pi i x_j \cdot \xi_j}$ is unimportant. The above can be interpreted as the inner product of f with the sequence of functions $\{T_{x_j} M_{\xi_j} \tilde{g}\}_{j \in \mathcal{J}}$, which establishes its relationship to frames. This collection of functions can be studied from a representation theory viewpoint. Indeed, the operator $T_x M_\xi$ is a unitary representation of the Heisenberg group on $L^2(\mathbb{R}^d)$, see [77]. We do not adapt this point of view, as we are interested in more general time-frequency representations that are not associated with this group structure.

To obtain discrete UCFs, we take advantage of the fact that f_p is band-limited for each $p \in \mathcal{P}$ and use ideas from classical sampling theory. By definition, the diameter of $\text{supp}(\widehat{f}_p)$ is bounded uniformly in p . Let $S_p = (S_{p,1}, S_{p,2}, \dots, S_{p,d})$ be the side lengths of the smallest rectangle that contains $\text{supp}(\widehat{f}_p)$. Similarly, let $S_0 = (S_{0,1}, S_{0,2}, \dots, S_{0,d})$ be the side lengths of the smallest rectangle that contains $\text{supp}(\widehat{f}_0)$.

We introduce the following standard notation to simplify the formulas in this section. For multi-integers $n, m \in \mathbb{Z}^d$, we write $m \leq n$ to mean $m_j \leq n_j$ for all $1 \leq j \leq d$. If $n > 0$, let

$$\text{vol}(n) = \prod_{j=1}^d n_j \quad \text{and} \quad \frac{m}{n} = \left(\frac{m_1}{n_1}, \frac{m_2}{n_2}, \dots, \frac{m_d}{n_d} \right).$$

Then, for $n \in \mathbb{Z}^d$, let

$$f_{0,n}(x) = \text{vol}(S_0)^{-1/2} f_0\left(x + \frac{n}{S_0}\right),$$

$$f_{p,n}(x) = \text{vol}(S_p)^{-1/2} f_p\left(x + \frac{n}{S_p}\right).$$

Definition 5.1.1. We call the sequence of functions

$$\mathcal{F}_{\text{discrete}} = \{f_{0,n}\}_{n \in \mathbb{Z}^d} \cup \{f_{p,n}\}_{p \in \mathcal{P}, n \in \mathbb{Z}^d}$$

a *discrete rotational uniform covering frame*.

Proposition 5.1.2. $\mathcal{F}_{\text{discrete}}$ is a discrete Parseval frame for $L^2(\mathbb{R}^d)$.

Proof. We will repeatedly use the well-known fact that for any multi-integer $S \in \mathbb{Z}^d$ with $S > 0$, the sequence of functions,

$$\{\text{vol}(S)^{-1/2} e^{-2\pi i \xi \cdot n/S} : n \in \mathbb{Z}^d\},$$

is an orthonormal basis for $L^2(\mathbb{R}^d)$ functions that are supported in any rectangle of side length S .

By Parseval, we have

$$\langle f, f_{p,n} \rangle = \text{vol}(S_p)^{-1/2} \int_{\mathbb{R}^d} \widehat{f}(\xi) \overline{\widehat{f}_p(\xi)} e^{-2\pi i \xi \cdot n/S_p} d\xi.$$

Since \widehat{f}_p is supported in a translate of a rectangle with side length S , we see that

$$\sum_{n \in \mathbb{Z}^d} |\langle f, f_{p,n} \rangle|^2 = \int_{\mathbb{R}^d} |\widehat{f}(\xi)|^2 |\widehat{f}_p(\xi)|^2 d\xi.$$

We use the same argument to handle terms that involve the inner product of f with $f_{0,n}$. Doing so, we obtain

$$\sum_{n \in \mathbb{Z}^d} |\langle f, f_{0,n} \rangle|^2 = \int_{\mathbb{R}^d} |\widehat{f}(\xi)|^2 |\widehat{f}_0(\xi)|^2 d\xi.$$

Combining the previous equations and applying the frame identity 2.1, we conclude that

$$\begin{aligned}
& \sum_{n \in \mathbb{Z}^d} |\langle f, f_{0,n} \rangle|^2 + \sum_{p \in \mathcal{P}} \sum_{n \in \mathbb{Z}^d} |\langle f, f_{p,n} \rangle|^2 \\
&= \int_{\mathbb{R}^d} |\widehat{f}(\xi)|^2 |\widehat{f}_0(\xi)|^2 d\xi + \sum_{p \in \mathcal{P}} \int_{\mathbb{R}^d} |\widehat{f}(\xi)|^2 |\widehat{f_{m,r}}(\xi)|^2 d\xi \\
&= \int_{\mathbb{R}^d} |\widehat{f}(\xi)|^2 d\xi.
\end{aligned}$$

□

Recall that a *finite frame* for \mathbb{C}^d is a sequence of vectors $\Phi = \{\phi_j\}_{j \in \mathcal{J}} \subseteq \mathbb{C}^d$ indexed by a finite index set \mathcal{J} such that there exist $0 < A \leq B < \infty$ and for all $f \in \mathbb{C}^d$,

$$A\|f\|_2^2 \leq \sum_{j \in \mathcal{J}} |\langle f, \phi_j \rangle|^2 \leq B\|f\|_2^2.$$

Finite frames is its own subject area and there are numerous papers on this topic, see [13, 38, 11, 70, 95]. We highlight two main differences between finite and discrete frames.

- (a) In infinite dimensions and without additional assumptions on f besides belonging to $L^2(\mathbb{R}^d)$, a discrete frame must consist of infinitely many elements. For finite dimensional spaces, it is possible to use finitely many frame elements, and the smallest number one needs is exactly d .
- (b) Of course, any discrete frame must be converted to a finite frame in order to be used in practice. Finite frames have applications to digital signal processing. Compared to orthonormal bases, the extra redundancy is desirable for combating deletion of coefficients and for added stability to perturbations.

We construct finite uniform covering frames for d -order tensors F of size $N = (N_1, N_2, \dots, N_d)$. That is, we work with the finite dimensional space $\mathbb{C}^N = \mathbb{C}^{N_1} \times \mathbb{C}^{N_2} \times \dots \times \mathbb{C}^{N_d}$. If F and G are arrays of size $N > 0$, their Frobenius inner product is

$$\langle F, G \rangle = \sum_{1 \leq n \leq N} F(n) \overline{G(n)}$$

and the Frobenius norm is $\|F\|_2 = \sqrt{\langle F, F \rangle}$.

To define these finite frames, we use the d -dimensional discrete Fourier transform (DFT), and replace the continuous domain functions \widehat{f}_p with its samples on a lattice. Suppose F is an array of size N , and let $\widehat{F}(\xi)$ be the d -dimensional Fourier series of F evaluated at $\xi \in \mathbb{R}^d$,

$$\widehat{F}(\xi) = \sum_{1 \leq n \leq N} F(n) e^{-2\pi i \xi \cdot n / N}.$$

Its DFT is N -periodic and we view them as samples of its Fourier series evaluated at the points $m_j = -\overline{N}_j, -\overline{N}_j + 1, \dots, \overline{N}_j$, where $\overline{N} \in \mathbb{R}^d$ is defined as $\overline{N}_j = (N_j - 1)/2$. Call this set of points $\text{supp}(\widehat{F})$.

Let \mathcal{P}_N be the finite subset of \mathcal{P} such that $p \in \mathcal{P}_N$ if and only if $\text{supp}(\widehat{f}_p)$ has nontrivial intersection with the rectangle

$$[-\overline{N}_1, \overline{N}_1] \times [-\overline{N}_2, \overline{N}_2] \times \dots \times [-\overline{N}_d, \overline{N}_d].$$

Let $S_p - 1 \in \mathbb{Z}^d$ be the side length of the smallest rectangle containing $\text{supp}(\widehat{f}_p) \cap \text{supp}(\widehat{F})$, and let S_0 be defined analogously. For each $p \in \mathcal{P}_N$ and $0 \leq n \leq S_p - 1$,

let $F_{p,n}$ be an array of size N , which we define according to its DFT,

$$\begin{aligned}\widehat{F_{p,n}}(m) &= F_p(m)E_{p,n}(m), \\ F_p(m) &= \widehat{f_p}(m), \\ E_{p,n}(m) &= \text{vol}(S_p)^{-1/2}e^{-2\pi im \cdot n/S_p}.\end{aligned}$$

Similarly, we define the array $F_{0,n}$ by

$$\begin{aligned}\widehat{F_{0,n}}(m) &= F_0(m)E_{0,n}(m), \\ F_0(m) &= \widehat{f_0}(m), \\ E_{0,n}(m) &= \text{vol}(S_0)^{-1/2}e^{-2\pi im \cdot n/S_0}.\end{aligned}$$

Definition 5.1.3. Given $N = (N_1, N_2, \dots, N_d) > 0$, we call the set

$$\mathcal{F}_{\text{finite}} = \{F_{0,n}\}_{0 \leq n \leq S_0-1} \cup \{F_{p,n}\}_{p \in \mathcal{P}_N, 0 \leq n \leq S_p-1},$$

a *finite uniform covering frame*.

Proposition 5.1.4. For any $N = (N_1, N_2, \dots, N_d) > 0$, the set $\mathcal{F}_{\text{finite}}$ is a finite Parseval frame for \mathbb{C}^N .

Proof. By Parseval, we have

$$|\langle F, F_{p,n} \rangle|^2 = \frac{1}{\text{vol}(N)} \left| \sum_{-\bar{N} \leq m \leq \bar{N}} \widehat{F}(m) \overline{\widehat{F_p}(m) E_{p,n}(m)} \right|^2.$$

Since $\{E_{p,n}\}_{0 \leq n \leq S_p-1}$ is an orthonormal basis for the support of F_p with respect to the inner product $\langle \cdot, \cdot \rangle$, we have

$$\sum_{0 \leq n \leq S_p-1} |\langle F, F_{p,n} \rangle|^2 = \frac{1}{\text{vol}(N)} \sum_{-\bar{N} \leq m \leq \bar{N}} |\widehat{F}(m)|^2 |F_p(m)|^2.$$

Repeating the same argument shows that

$$\sum_{0 \leq n \leq S_0 - 1} |\langle F, F_{0,n} \rangle|^2 = \frac{1}{\text{vol}(N)} \sum_{-\bar{N} \leq m \leq \bar{N}} |\widehat{F}(m)|^2 |F_0(m)|^2.$$

Combining the previous equations, recalling that F_p are samples of \widehat{f}_p , and applying the frame identity (2.1), we have

$$\begin{aligned} & \sum_{0 \leq n \leq S_0 - 1} |\langle F, F_{0,n} \rangle|^2 + \sum_{p \in \mathcal{P}_N} \sum_{0 \leq n \leq S_p - 1} |\langle F, F_{p,n} \rangle|^2 \\ &= \frac{1}{\text{vol}(N)} \sum_{-\bar{N} \leq m \leq \bar{N}} |\widehat{F}(m)|^2 \left(|\widehat{F}_0(m)|^2 + \sum_{p \in \mathcal{P}_N} |\widehat{F}_p(m)|^2 \right) = \|F\|_2^2. \end{aligned}$$

□

5.2 Fast Fourier Scattering Transform

We propose the *fast Fourier scattering transform*, which is a simple algorithm that computes the Fourier scattering coefficients for an input vector. We use a finite uniform covering frame $\mathcal{F} = \mathcal{F}_{\text{discrete}}$, which can either be a traditional Gabor frame, a RUCF, or some other type of UCF. The algorithm also requires the user to input the depth of the network K , and the frequency covering parameter M . The choice of \mathcal{F} and the parameters usually depend on the type of application.

Algorithm 5.2.1. Fast Fourier scattering transform.

Input: Vector f , parameters $M, K \geq 1$, and $\mathcal{F} = \{f_0\} \cup \{f_p\}_{p \in \mathcal{P}[M]}$

for $k = 1, 2, \dots, K$

for each $p = (p', p_k) \in \mathcal{P}[M]^k$

 Compute $U[p]f = U[(p', p_k)]f = |U[p']f * f_{p_k}|$ and $U[p]f * f_0$

end

end

The reason we call this algorithm *fast* is because Theorem 3.2.2 quantifies the amount of information lost due to truncation, so we do not have to calculate an enormous number of scattering coefficients. Hence, we typically choose $K = 2$ and M large enough depending the regularity of the dataset. In fact, we can make the algorithm even faster and efficient:

- (a) In many applications, the input function f is real-valued. Suppose that for each $p \in \mathcal{P}[M]$, \widehat{f}_p is real-valued and there exists a unique $q \in \mathcal{P}[M]$ such that $\widehat{f}_p(\xi) = \widehat{f}_q(-\xi)$ for all $\xi \in \mathbb{R}^d$. Then, we have $|f * f_p| = |f * f_q|$. For any $k \geq 1$ and $p \in \mathcal{P}^k$, $U[p]f$ is also real-valued, so the same reasoning shows that for all $k \geq 1$ and $p \in \mathcal{P}[M]^k$, there exists a unique $q \in \mathcal{P}[M]^k$ such that $U[p]f = U[q]f$. Thus, we only need to compute half of the coefficients in the fast Fourier scattering transform. These assumptions hold, for example, if \mathcal{F} is a Gabor frame defined in Proposition 2.2.2 and the window $\widehat{f}_0 = \widehat{g}$ is real-valued and symmetric about the origin.
- (b) The proof of Proposition 3.1.6 shows that coefficients of the form $||f * f_p| * f_q|$ have small $L^2(\mathbb{R}^d)$ norm whenever $p \in \mathcal{P}[M]$, $q \in \mathcal{P}[N]$, and $M \ll N$. Since our algorithm still computes such coefficients, its runtime can be greatly reduced by not computing these coefficients.
- (c) Each scattering coefficient is band-limited, due to the convolution with f_0 . By the classical Shannon sampling theorem, we can down-sample the scattering

coefficients without any information loss, where the down-sampling rate is governed by the support of \widehat{f}_0 . This reduces the amount of memory required to store the coefficients.

5.3 Image Feature Comparison

Figure 5.1 compares the features generated by $\mathcal{S}_{\mathcal{F}}$ and $\mathcal{S}_{\mathcal{W}}$. The Matlab code that reproduces this figure, as well as Fourier scattering software, can be downloaded at the author’s personal website [100]. We use the publicly available ScatNet toolbox [108] to produce the analogous wavelet scattering features.

For the comparison, we use the standard 512×512 Lena image as our model example. Due to its varied content, Lena image allows to emphasize the differences between $\mathcal{S}_{\mathcal{F}}$ and $\mathcal{S}_{\mathcal{W}}$. We interpret the square image as samples of a function f , namely

$$\{f(m_1, m_2): m_1 = 1, 2, \dots, 512, m_2 = 1, 2, \dots, 512\}.$$

In order to make a fair comparison, we chose appropriate parameters. $\mathcal{S}_{\mathcal{F}}[M, K]$ requires specification of \mathcal{F} (for simplicity, we use a Gabor frame satisfying Proposition 2.2.2 with parameter a free to chose, $A = aI$, and real-valued even window $\widehat{f}_0 = \widehat{g}$), the integer M that controls the cardinality of $\mathcal{P}[M]$, and the depth of the network K . On the other hand, $\mathcal{S}_{\mathcal{W}}$ requires specification of the number of angles in the rotation group G , the range of dyadic scales (controlled by the coarsest scale J), and the depth of the network. It is appropriate to pick a and J such that \widehat{f}_0 and $\widehat{\varphi}_{2^J}$ are approximately supported in a ball of the same size. By this choice,

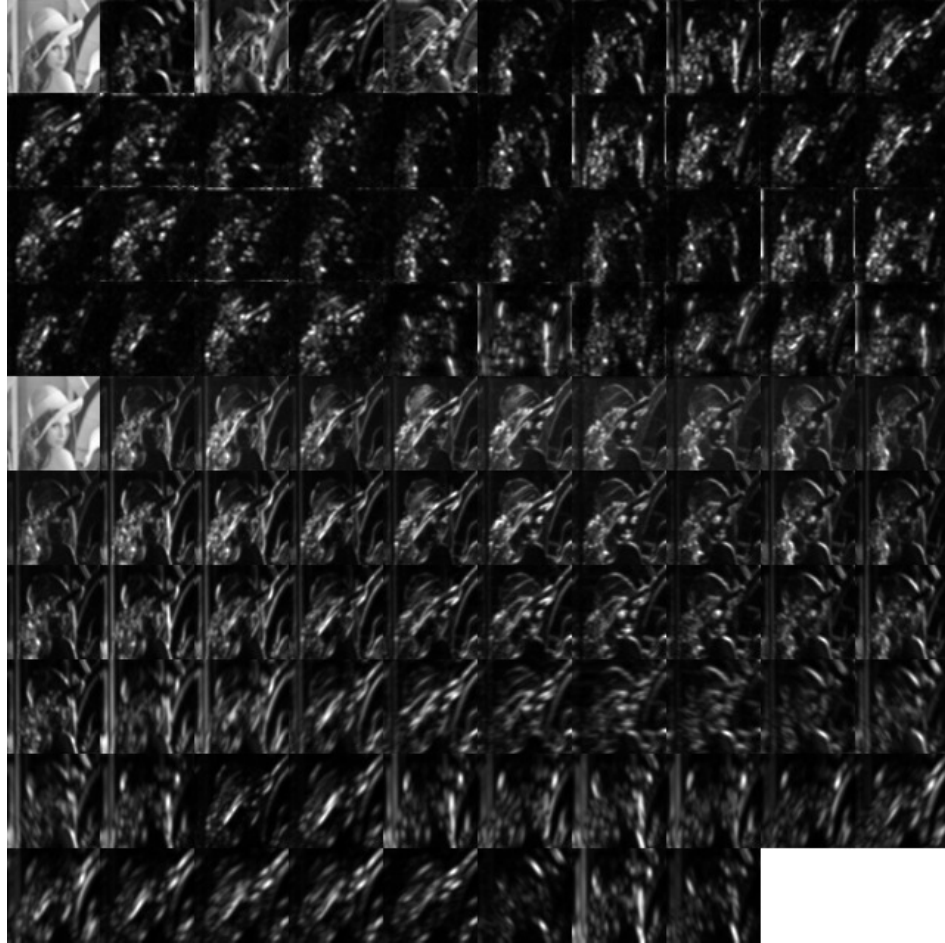


Figure 5.1: Fourier and windowed scattering transform coefficients whose norms exceed 0.5% and ordered by depth. The top four rows display (1, 33, 6) zero, first, and second order Fourier scattering transform coefficients. Bottom six rows display (1, 40, 17) zero, first, and second order windowed scattering coefficients.

the zero-th order coefficients of both transformations have approximately the same resolution. Next, we chose M and G so that both transforms have the same number of first-order coefficients. For both, we use a network with 2 layers. With these considerations, we computed $(1, 40, 560)$ and $(1, 40, 600)$ zero, first, and second order coefficients for $\mathcal{S}_{\mathcal{F}}$ and $\mathcal{S}_{\mathcal{W}}$, respectively.

While it is interesting to look at each coefficient, there are obviously too many to display. Hence, Figures 5.1 illustrates the Fourier and wavelet scattering coefficients with $L^2(\mathbb{R}^d)$ norm greater than $0.005 * \|f\|_2$. In order to clearly display the coefficients, each one is normalized to have a maximum of 1 (but only when plotted). The norm of the coefficients decrease as a function of the depth, so we chose 0.005 as the threshold parameter in order to display several of the second-order coefficients. While we only show the largest ones, it does not necessarily mean that they are the most important since it is plausible that smaller coefficients contain informative features.

We first concentrate on the Fourier scattering features. As seen in the Figure 5.1, the first-order coefficients extract distinct features of the image, and in particular, extract the most prominent edges in image. While the first-order coefficients capture individual features – various components of the hat, her hair, the feature, the background, and her facial features – the second-order coefficients appear to capture a combination of features. In general, it is difficult to substantiate what functions of the form $||f * f_p| * f_q|$ intuitively mean. This is partly because the Fourier transform is the standard tool for analyzing convolutions, but is not well suited for handling non-linear operators such as the complex modulus. For the wavelet case, Mallat

has heuristically argued that coefficients of the form $||f * \psi_{2^j, r}| * \psi_{2^k, s}|$ describe interactions between scales 2^{-j} and 2^{-k} [107]. By the same reasoning, coefficients of the form $||f * f_p| * f_q|$ describe the interactions between oscillations arising from the uniform Fourier scales $|p|_\infty$ and $|q|_\infty$.

Let us compare the features generated by $\mathcal{S}_{\mathcal{F}}$ and $\mathcal{S}_{\mathcal{W}}$. Their zero-order coefficients are almost identical due to our parameter choices; recall that we chose f_0 and φ_{2^j} to have approximately the same resolution. In some sense, the first-order coefficients illustrate the main differences between wavelet and Gabor transforms. Indeed, wavelets are multi-scale representations, so the first-order coefficients of $\mathcal{S}_{\mathcal{W}}$ have higher resolution than those of $\mathcal{S}_{\mathcal{F}}$, and the latter are of a fixed low resolution.

By inspection, the second-order coefficients of $\mathcal{S}_{\mathcal{F}}$ and $\mathcal{S}_{\mathcal{W}}$ appear to capture significantly different features. For example, the second-order Fourier scattering coefficients mainly capture the oscillatory pattern of the feather, whereas the second-order windowed scattering coefficients focused more on her hair and on the round shapes in the image. Again, it is hard to precisely describe what information the second-order terms capture. Finally, $\mathcal{S}_{\mathcal{F}}$ has a fewer number of second-order coefficients that exceed the threshold parameter, which is consistent with our theory that $\mathcal{S}_{\mathcal{F}}$ satisfies an exponential decay of energy property.

5.4 Anomaly Detection in Radiological Data

Anomaly detection is the process of detecting a small set of outliers within a dataset. One example is the detection of a few radioactive sources from the *gamma-ray spec-*

trum (number of gamma particles per energy level) of an entire scene. The SORDS dataset consists of three distinct radioactive sources, each hidden at a different location on the University of Maryland, College Park, campus. A vehicle mounted with a detector drove along the perimeter a total of 12 times, and for each loop, the detector collected the gamma-ray spectrum at various locations along the perimeter of the scene. See Figure 5.2 for more details.

The spectrum provides information about the compounds nearby. Strong sources, in comparison to the background radiation, can be easily detected by examining the *total count*, which is the ℓ^1 norm of the gamma-ray spectrum. Weaker sources do not have elevated total counts, and must be detected by examining the features apparent in the spectrum. These are detectable to the trained eye under ideal conditions, but real data is typically noisy and inaccurate. The rate at which a source emits gamma particles at a certain frequency can be modeled as a Poisson process, so important information might be missed if the collection time is insufficient. In the SORDS dataset, there is one strong source and two weak sources. See Figure 5.3 for representative spectra at different locations.

Upon examination of the dataset, we see that the ℓ^2 distance (or any other Euclidean metric) does not discriminate between the spectra between the background and sources. We transform the spectral information into a more discriminatory basis. The oscillations in the spectra are the key features, so it is natural to apply the Fourier scattering transform to each spectra. This transformation is well-suited for this data type, since the distance between the background and three source spectra are widened in the coefficient domain. See Figure 5.3 for the Fourier scattering

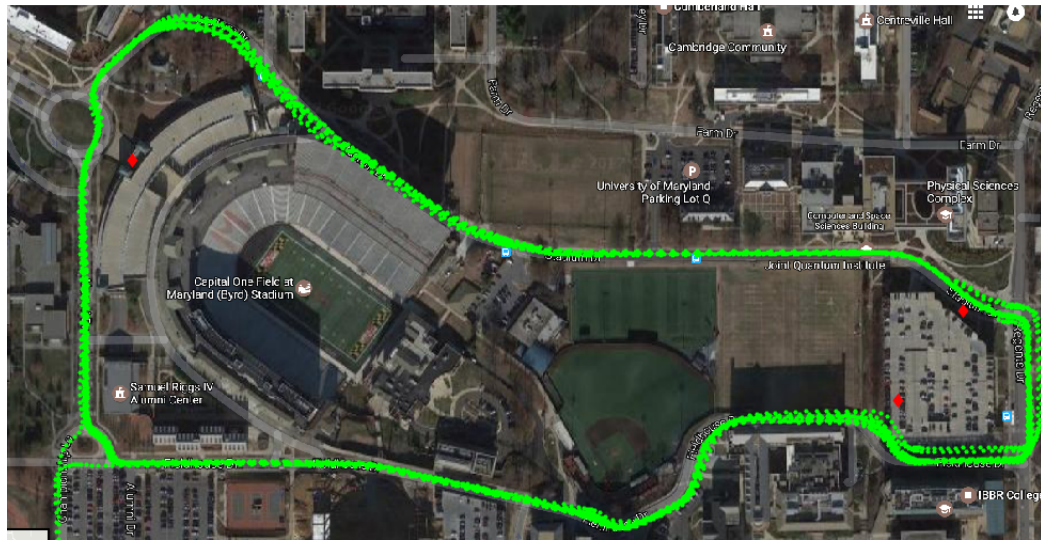
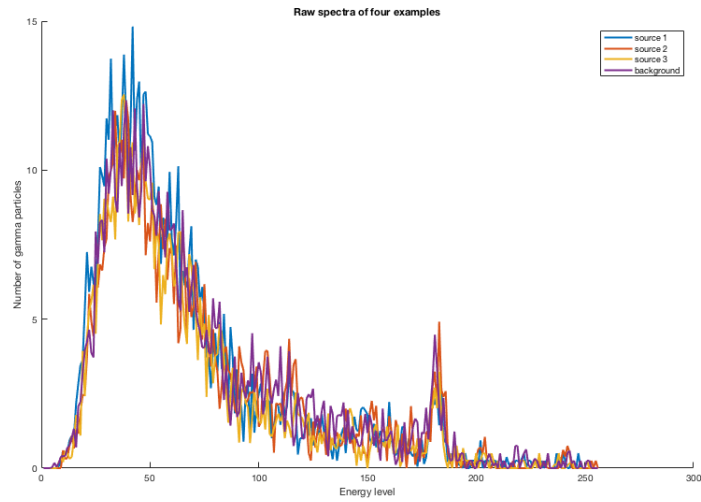


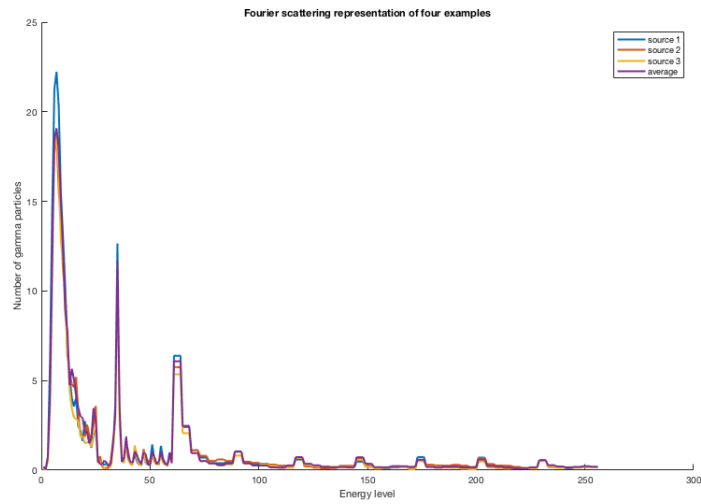
Figure 5.2: Three sources were hidden at locations on the University of Maryland, College Park campus, specified by the red diamonds. The gamma-ray spectrum was collected along the perimeter at the indicated green points. The detector passed around for a total of 12 times.

representation of these points.

We label the points nearby the sources as *radioactive* and the remaining locations as non-radioactive. We use the Fourier scattering representations of the spectrum corresponding to the first 6 loops, which approximately half of the total dataset, to train a support vector machine (SVM). The SVM is used to classify the Fourier scattering representation of the data from the remaining half of the dataset. The results are displayed in Figure 5.4. This experiment shows that the Fourier scattering transform de-noises the data, and transforms it into a suitable basis that standard machine learning algorithms can handle.



(a) Raw gamma-ray spectrum



(b) Fourier scattering representation spectrum

Figure 5.3: Examples of the gamma-ray spectrum at four different sample points: one near each source and one far away from the three sources. The blue is the spectra corresponding to the strong source. The red and the yellow spectra correspond to the weaker sources, and the red one is slightly easier to detect.

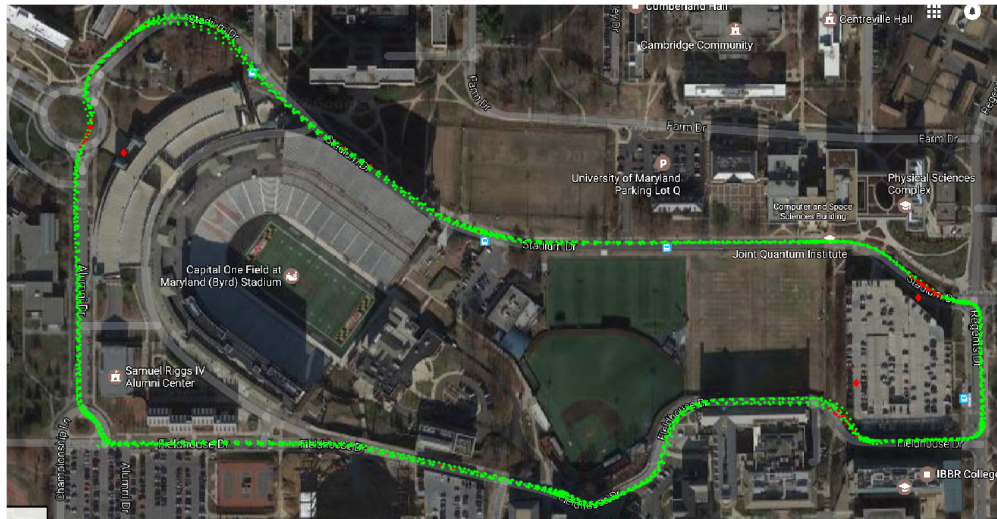


Figure 5.4: The red dots are classified as radioactive by the FST and SVM supervised learning algorithm. It correctly classifies nearly all of the points near the strong source, and most of the ones near the weaker sources. There are a total of 8 misclassified points that are nowhere near the sources.

Chapter 6: Background on Super-Resolution

6.1 Terminology

The term *super-resolution* varies depending on the field, and, consequently, there are various types of super-resolution problems. In some situations [116], super-resolution refers to the process of up-sampling an image onto a finer grid, which is a *spatial interpolation* procedure. In other situations [104], super-resolution refers to the process of recovering the object's high frequency information from its low frequency information, which is a *spectral extrapolation* procedure. In both situations, the super-resolution problem is ill-posed because the missing information can be arbitrary. However, it is possible to provide meaningful super-resolution algorithms by using *prior knowledge* of the data.

Although our point of view is mathematical, we do want to set the stage briefly by mentioning the following applications and scenarios, noting that these are just the tip of the iceberg, see [119, 91, 20, 44, 45].

- (a) In astronomy [118], each star can be modeled as a complex number times a Dirac δ -measure, and the Fourier transform of each star encodes important information about that star. However, an image of two stars that are close in

distance resembles an image of a single star. In this context, the super-resolution problem is to determine the number of stars and their locations, using the prior information that the actual object is a linear combination of Dirac δ -measures.

(b) In medical imaging [76], machines capture the structure of the patient's body tissues, in order to detect for anomalies in the patient. Their shapes and locations are the most important features, so each tissue can be modeled as the characteristic function of a closed set, or as a surface measure supported on the boundary of a set. The super-resolution problem is to capture the fine structures of the tissues, given that the actual object is a linear combination of singular measures.

(c) In certain applications, an image is obtained by convolving the object with the *point spread function* of an optical lens. Alternatively, the Fourier transform of the object is multiplied by a *modulation transfer function*. The resulting image's resolution is inherently limited by the Abbe diffraction limit, which depends on the illumination light's wavelength and on the diameter of the optical lens. Thus, the optical lens acts as a low-pass filter, see [104]. The purpose of super-resolution is to use prior knowledge about the object to obtain an accurate image whose resolution is higher than what can be measured by the optical lens.

In the super-resolution literature, the collection of object is modeled as a bounded measure μ on the torus \mathbb{T}^d . The measurement process provides us with a finite set $\Lambda \subseteq \mathbb{Z}^d$ and a function F defined on Λ . We say F is *spectral data on a set* $\Lambda \subseteq \mathbb{Z}^d$ if $F = \widehat{\mu} |_{\Lambda}$ for some $\mu \in M(\mathbb{T}^d)$. To relate this abstract formulation back

to our concrete motivation, we provide several examples of μ in various contexts.

(a) In astronomy [118], microscopy [104], line spectral estimation [126], source localization in remote sensing [66], and direction of arrival estimation [93], the unknown μ is a discrete measure. Its support represents the locations of the objects and its complex-valued coefficients encode their intensity and phase. The prior information is that the object of interest contains only a few number of point sources compared to the number of measurements. The critical scale is the *Rayleigh length*, which is the minimum separation between two point sources that the imaging system can resolve. In our case of Fourier measurements, as seen by the classical uncertainty principle, the Rayleigh length is the reciprocal of the number of measurements. When the Dirac measures in μ are separated by distances below the Rayleigh length, their recovery becomes challenging, but that is precisely the aim of super-resolution.

(b) In the context of image processing, we are given information that represents the low-resolution features of an unknown $\mu \in M(\mathbb{T}^d)$, and that this information is in the form,

$$f(x) = (\mu * \psi)(x) = \int_{\mathbb{T}^d} \psi(x - y) d\mu(y),$$

for some $\psi: \mathbb{T}^d \rightarrow \mathbb{C}$. For simplicity, we assume that $\widehat{\psi} = \mathbf{1}_\Lambda$, the characteristic function of some finite set $\Lambda \subseteq \mathbb{Z}^d$. Then, we have

$$f = (\mu * \psi) = (\widehat{\mu} \upharpoonright_\Lambda)^\vee.$$

Thus, we interpret $\widehat{\mu} \upharpoonright_\Lambda$ as the given low frequency information of μ , and the function $(\widehat{\mu} \upharpoonright_\Lambda)^\vee$ as the given low resolution image. A primary objective is

to determine if $\mu \in M(\mathbb{T}^d)$ can be recovered given $F = \hat{\mu} |_{\Lambda}$. The current literature has focused on discrete $\mu \in M(\mathbb{T}^d)$, see [30, 128, 69]. However, in the context of image processing, μ is the unknown high resolution image, and a typical image cannot be approximated realistically by a discrete measure, although such approximation is possible mathematically in the weak-* topology [7]. Hence, it is important to determine if non-discrete measures are solutions to the super-resolution problem.

- (c) A common approach for *discrete* image recovery in inverse problems in medical imaging is to minimize the ℓ^1 norm of the finite differences of the image, which promotes the recovery of piecewise constant images. A different but related approach is to study the recovery of *continuous* domain piecewise constant images. For example, [114] showed that the edge set of a piecewise constant image is uniquely identifiable from low-pass Fourier coefficients of the image when it coincides with the zero-set of a trigonometric polynomial, and [113] studied a practical approach for spectral extrapolation using low-rank Toeplitz matrix completion. Related works [115, 67] studied the recovery of multi-dimensional singular measures supported on curves defined by zero-sets of analytic functions. Our paper studies a similar formulation, but our recovery method, TV minimization, is quite different from the aforementioned papers.

6.2 Convex Approach

One method for producing an approximation of the unknown measure is to solve the *total variation minimization problem*: For a given finite subset $\Lambda \subseteq \mathbb{Z}^d$ and given spectral data F on Λ , find the solution(s) ν to

$$\inf_{\nu} \|\nu\|_{\text{TV}} \quad \text{such that} \quad \nu \in M(\mathbb{T}^d) \quad \text{and} \quad F = \widehat{\nu} \quad \text{on } \Lambda. \quad (\text{TV})$$

This is a *convex minimization problem* and we interpret a solution ν as a simple or least complicated high resolution extrapolation of F .

With regard to (TV), a fundamental question of uniqueness must be addressed. To see why this is so, if μ is not the unique solution, then the output of a numerical algorithm is not guaranteed to approximate μ . Thus, it is important to determine sufficient conditions such that μ is the unique solution. For this reason, we say that *reconstruction of μ from F on Λ is possible* if and only if μ is the unique solution to (TV). Of course, it could be theoretically possible to reconstruct μ by other means.

The mathematical theory for (TV) almost exclusively pertains to discrete measures satisfying a strong separation condition. If $\Lambda = \{-M, -M+1, \dots, M\}^d \subseteq \mathbb{Z}^d$, we say a discrete μ satisfies the *minimum separation condition with constant $C > 0$* if

$$\min_{x, y \in \text{supp}(\mu), x \neq y} \|x - y\|_{\ell^\infty(\mathbb{T}^d)} \geq \frac{C}{M}.$$

Theorem 6.2.1 (Candès and Fernandez-Granda, Theorems 1.2-1.3, [30], Fernandez-Granda, Theorem 2.2, [69]). *Let F be spectral data on $\Lambda = \{-M, -M+1, \dots, M\}^d \subseteq \mathbb{Z}^d$, where $F = \widehat{\mu} |_{\Lambda}$ for some discrete $\mu \in M(\mathbb{T}^d)$. In any of the*

following cases, μ is the unique solution to (TV).

(a) $d = 1$, $M \geq 128$, and μ satisfies the minimum separation condition with $C = 2$.

(b) $d = 1$, $M \geq 128$, and μ is real-valued and satisfies the minimum separation condition with $C = 1.87$.

(c) $d = 1$, $M \geq 10^3$, and μ satisfies the minimum separation condition with $C = 1.26$.

(d) $d = 2$, $M \geq 512$, and μ is real-valued and satisfies the minimum separation condition for $C = 2.38$.

To prove this theorem, the authors studied the following *dual problem*: Given spectral data F on a finite set $\Lambda \subseteq \mathbb{Z}^d$, find the solution(s) φ to

$$\max_{\varphi} \left| \sum_{m \in \Lambda} \widehat{\varphi}(m) F(m) \right| \quad \text{such that} \quad \|\varphi\|_{L^\infty} \leq 1 \quad \text{and} \quad \text{supp}(\widehat{\varphi}) \subseteq \Lambda. \quad (\text{TV}')$$

We call its solutions *dual polynomials*, and we let $\mathcal{D}(F, \Lambda)$ denote the set of dual polynomials.

While this viewpoint provides a satisfactory theoretical result, the computational aspect requires overcoming additional challenges. For $d = 1$, the authors proposed a *semi-definite program* (SDP) formulation of (TV'), see [30, Corollary 4.1], which relies on a spectral factorization theorem [59, Theorem 4.24].

Algorithm 6.2.2. Suppose μ and Λ satisfy the assumptions in Theorem 6.2.1a.

1. Compute a polynomial φ using the SDP reformulation of the dual problem.

2. If $|\varphi| \not\equiv 1$, then $\text{supp}(\mu) \subseteq S = \{x \in \mathbb{T}: |\varphi(x)| = 1\}$ consists of at most $2M$ elements.
3. Locate S using a root finding method.
4. Invert a linear system in order to solve for the coefficients of μ .

Candès and Fernandez-Granda pointed out that this method is incomplete due to the assumption in the second step of Algorithm 6.2.2. We emphasize that this can occur even if μ satisfies any of the first three conditions in Theorem 6.2.1.

Further, while the SDP reformation does not hold for $d \geq 2$, there are more involved iterative primal-dual algorithms that can be substituted for the first step of Algorithm 6.2.2, e.g., see [136, 52]. Alternatively, one could compute the solutions to (TV) directly without dealing with (TV'), and this would avoid some of the difficulties mentioned here, e.g., see [21], which uses a conditional gradient method.

We mention some related work on super-resolution, most of which adopt similar hypotheses as in Theorem 6.2.1. For example, [128] considered the recovery of discrete measures with random coefficients and a randomly selected low-frequency sampling set Λ . By considering different sampling schemes and assumptions on the discrete measure, it is possible to study other “super-resolution” problems [51, 4]. Finally, many these works were inspired by ℓ^1 minimization [35], compressed sensing [31, 56], and convex geometry [34].

We can extend the previous noiseless model to the situation where we only have access to imperfect measurements. Given a noise level $\delta > 0$, we say F_δ is *noisy spectral data* on a finite set $\Lambda \subseteq \mathbb{Z}^d$ if $F_\delta = \widehat{\mu}|_\Lambda + \eta$ for some $\mu \in M(\mathbb{T}^d)$ and

function η on Λ with $\|\eta\|_2 \leq \delta$.

The optimization method (TV) is unstable to small perturbations, so we must replace it with the following: For a given finite subset $\Lambda \subseteq \mathbb{Z}^d$ and given noisy spectral data F_δ on Λ , fix a parameter $\tau > 0$ and find the solution(s) ν to

$$\inf_{\nu} \frac{1}{2} \|F_\delta - \widehat{\nu}\|_{\ell^2(\Lambda)}^2 + \tau \|\nu\|_{\text{TV}}. \quad (\text{SR}_\tau)$$

This formulation is a generalization, to the measure case, of discrete compressed sensing algorithms [33].

The most important question in the study of regularization methods is to determine the relationship between the solutions of (TV) and (SR $_\tau$). If ν is a solution to (SR $_\tau$), then intuitively speaking, we expect that ν converges to μ if the parameter τ is chosen appropriately depending on the noise level and the noise level tends to zero. This intuition is somewhat correct, since it is possible to show convergence for a subsequence and in the weak-* sense.

Such convergence statements are qualitative, whereas we want a quantitative bound. This leads us to the question: What is a natural way of quantifying the error, $\nu - \mu$? Burger-Osher [24] argued that, since Tikhonov regularization is achieved in the weak-* topology, it would be surprising if it is possible to bound the error in the total variation norm. Since (SR $_\tau$) is a special case of Tikhonov regularization, it is reasonable that the same principle applies. Numerical results have shown that, in fact, the supports of μ and ν can be different [29, 60], which validates this heuristic. Thus, it is impossible to bound $\nu - \mu$ in the total variation norm.

Since super-resolution is concerned with the recovery of fine details from coarse

data, we would like to compare the difference $\nu - \mu$ at various scales. For this reason, for a fixed kernel K , we examine the smoothed out error $K * (\nu - \mu)$. Even with this simplification, quantifying the approximation error is difficult for general μ . For this reason, the available theory deals with μ satisfying any of the assumptions in Theorem 6.2.1, see [29]. Of course, there are metrics for quantifying the approximation rate, see [68, 5]. We also want to emphasize that we are not interested in only *de-noising* the measurements, which would only involve estimating $\hat{\mu}|_{\Lambda}$, which was carried out in [19, 127]; instead, we are interested in recovering the underlying measure μ , which is a more ambitious task.

6.3 Subspace Methods

One major drawback of the convex approach is that, Theorem 6.2.1 is essentially tight. Numerical experiments in [69] show that, for $d = 1$ and *even in the noiseless case*, the Diracs in μ need to be separated by at least $1/M$. Similar limitations are featured in greedy methods [65, 57]. While these algorithms are important and are used to circumvent discretization issues that arise in compressive sensing [31, 56], it is debatable whether they are actually super-resolution algorithms, precisely because they cannot be used to recover objects at scales smaller than the Rayleigh length.

In contrast, *subspace methods*, such as ESPRIT [120], Matrix Pencil Method (MPM) [88, 109], and MUSIC [103, 102], do not have this restriction on the separation of the Diracs. In the noiseless case, subspace methods can recover the discrete measure exactly. In the presence of noise, they are able to accurately recover the

measure provided that the noise is *sufficiently small*. These methods were partially inspired by the instability of the classical Prony’s method [117]. This is exciting because subspace methods have this stable super-resolution property and experimental results have verified that they are the best known algorithms for recovering discrete measures from noisy Fourier samples.

Unfortunately, the picture is incomplete because is not well-understood what “sufficiently small” means. One approach is to relate this problem to the conditioning of Vandermonde matrices. It is not surprising that there is a connection, since the condition number describes the stability of inverting an over-determined system. Let $T = \{t_n\}_{n=1}^S \subseteq \mathbb{T}$, and let V_T be the $M \times S$ Vandermonde matrix associated with T , which is defined as $(V_T)_{m,n} = e^{2\pi i m t_n}$ for $m = 0, 1, \dots, M - 1$. MUSIC and MPM are able to recover discrete measures supported in T provided that the noise is smaller, up to a constant times the smallest singular value of V_T , see [103, 109] for the explicit bounds.

However, it is not easy to accurately bound the smallest singular value. There is a certain phase transition for the smallest singular value. When the separation between all the points in T exceeds the Rayleigh length, then classical approximation theory results [130] give a sharp lower bound on the smallest singular value of V_T . While this is an important result, the main focus of super-resolution is to deal with the case where the separation is below the Rayleigh length. In this regime, for certain examples of T , the smallest singular value of V_T can decay exponentially as the separation decreases [109]. Thus, whenever any two points are closer than the Rayleigh length, the smallest singular is highly unstable and sensitive to the

configuration of the support set, which has made accurate quantitative analysis difficult. Understanding the geometric dependence is still an open problem in this subject area.

Chapter 7: Beurling Theory of Super-Resolution

The content of this chapter appeared in the paper [14], written by the author and John J. Benedetto.

7.1 Problem Statement

As we mentioned earlier, we are particularly interested in the super-resolution of discrete measures with separation below the Rayleigh length, and of singular measures. To do this, we study (TV) more abstractly. Several previous papers have relied on the following well-known relationship: $\mu \in M(\mathbb{T}^d)$ is the *unique* solution to (TV) given spectral data F on Λ if and only if $F = \widehat{\mu}|_{\Lambda}$ and there exists $\varphi \in \mathcal{D}(F, \Lambda)$ such that $\varphi = \text{sign}(\mu)$ ν -a.e. and $|\varphi(x)| < 1$ for all $x \notin \text{supp}(\nu)$, see [30, 60]. However, this characterization has not yielded useful results except in very special cases of μ and Λ , such as the one considered in Theorem 6.2.1. The technical reason is the following. Even if we reduce the space $\mathcal{D}(F, \Lambda)$ by identifying functions with the same global phase, in general, there is still no unique dual polynomial. This forces one to use generic tools, whereas not much is known about the sign of arbitrary trigonometric polynomials.

In order to obtain results that pertain to a richer class of measures, we do not

work with the aforementioned relationship. Instead, we develop a general theory that does not impose any conditions on μ and Λ . In fact, we study (TV) from an abstract, functional analysis perspective. Our methods are inspired by Beurling's work on minimal extrapolation [17, 18]. For recent results on non-uniform sampling, going back to [17] and balayage, see [3]. We later connect our results with the Candès and Fernandez-Granda theory on super-resolution [30].

We first introduce additional notation and adopt Beurling's language, since referring to the solutions of (TV) can be ambiguous when working with several different measures. We say $\nu \in M(\mathbb{T}^d)$ is an *extrapolation* of F on Λ if $F = \widehat{\nu}|_{\Lambda}$; and ν is a *minimal extrapolation* of F on Λ if it is a solution to (TV). The minimum value attained in (TV) is denoted by

$$\varepsilon = \varepsilon(F, \Lambda) = \inf\{\|\nu\|_{\text{TV}} : \nu \in M(\mathbb{T}^d) \text{ and } F = \widehat{\nu} \text{ on } \Lambda\}.$$

While ε is in general unknown, we shall see that, for many important applications, it can be deduced. To this end, let

$$\mathcal{E} = \mathcal{E}(F, \Lambda) = \{\nu \in M(\mathbb{T}^d) : \|\nu\|_{\text{TV}} = \varepsilon \text{ and } F = \widehat{\nu} \text{ on } \Lambda\}$$

denote the set of all minimal extrapolations. Thus, for any $\nu \in \mathcal{E}(F, \Lambda)$, we have $\varepsilon(F, \Lambda) = \|\nu\|_{\text{TV}}$. In order to understand the behavior of the minimal extrapolations, we study the set of dual polynomials, $\mathcal{D}(F, \Lambda)$. Both $\mathcal{E}(F, \Lambda)$ and $\mathcal{D}(F, \Lambda)$ are non-empty, see Proposition 7.2.1. It turns out that some of these dual polynomials can be characterized by the set,

$$\Gamma = \Gamma(F, \Lambda) = \{m \in \Lambda : |F(m)| = \varepsilon(F, \Lambda)\}.$$

This connection is made in Proposition 7.2.2. We exploit this characterization to prove Theorem 7.3.1, which provides results without assumptions on μ and Λ .

7.2 Some Functional Analysis Results

We prove basic facts about (TV) and (TV'). In contrast to previous approaches relying on convex analysis, e.g., see [60], our proofs only use basic functional analysis, which also illustrates its important role.

Proposition 7.2.1 contains the main functional analysis results. Let $C(\mathbb{T}^d)'$ be the dual space of continuous linear functionals on $C(\mathbb{T}^d)$. The celebrated Riesz representation theorem, see [7, Theorem 7.2.7, page 334], states that:

- (a) For each $\mu \in M(\mathbb{T}^d)$, there exists a bounded linear functional $\ell_\mu \in C(\mathbb{T}^d)'$ such that $\|\mu\|_{\text{TV}} = \|\ell_\mu\|$ and

$$\forall f \in C(\mathbb{T}^d), \quad \ell_\mu(f) = \langle f, \mu \rangle = \int_{\mathbb{T}^d} f(x) \overline{d\mu(x)}.$$

- (b) For each bounded linear functional $\ell \in C(\mathbb{T}^d)'$, there exists a unique $\mu \in M(\mathbb{T}^d)$ such that $\|\mu\|_{\text{TV}} = \|\ell\|$ and

$$\forall f \in C(\mathbb{T}^d), \quad \ell(f) = \langle f, \mu \rangle = \int_{\mathbb{T}^d} f(x) \overline{d\mu(x)}.$$

Proposition 7.2.1a shows that (TV) is well-posed, by using standard functional analysis arguments. However, this type of argument does not yield useful statements about the minimal extrapolations. Instead of working with $C(\mathbb{T}^d)$, we shall work with a subspace. Since the minimal extrapolations only depend on F on Λ , and not

the unknown measure μ that generates F , we consider the subspace

$$C(\mathbb{T}^d; \Lambda) = \left\{ f \in C(\mathbb{T}^d) : f(x) = \sum_{m \in \Lambda} \widehat{f}(m) e^{2\pi i m \cdot x} \right\}.$$

Proposition 7.2.1b shows that $C(\mathbb{T}^d; \Lambda)$ is a closed subspace of $C(\mathbb{T}^d)$, and that implies $C(\mathbb{T}^d; \Lambda)$ is a Banach space. Further, Proposition 7.2.1c demonstrates that

$$U = U(\mathbb{T}^d; \Lambda) = \{f \in C(\mathbb{T}^d; \Lambda) : \|f\|_{L^\infty} \leq 1\},$$

the closed unit ball of $C(\mathbb{T}^d; \Lambda)$, is compact.

The purpose of restricting to the subspace, $C(\mathbb{T}^d; \Lambda)$, is to identify the unknown $\mu \in M(\mathbb{T}^d)$ with the bounded linear functional, $L_\mu \in C(\mathbb{T}^d; \Lambda)'$, defined as

$$\forall f \in C(\mathbb{T}^d; \Lambda), \quad L_\mu(f) = \int_{\mathbb{T}^d} f(x) \overline{d\mu(x)}.$$

Although, by definition, $\|L_\mu\| = \sup_{f \in U} |L_\mu(f)|$, Proposition 7.2.1d,e show that we have the stronger statement,

$$\|L_\mu\| = \max_{f \in U} |L_\mu(f)| = \varepsilon.$$

The purpose of studying L_μ is to deduce information about the minimal extrapolations. As a consequence of the Radon-Nikodym theorem, for each $\mu \in M(\mathbb{T}^d)$, there exists a μ -measurable function $\text{sign}(\mu)$ such that $|\text{sign}(\mu)| = 1$ μ -a.e. and satisfies the identity

$$\forall f \in L^1_{|\mu|}(\mathbb{T}^d), \quad \int_{\mathbb{T}^d} f d|\mu| = \int_{\mathbb{T}^d} f \overline{\text{sign}(\mu)} d\mu.$$

See [7, Theorem 5.3.2, page 242, and Theorem 5.3.5, page 244] for further details.

Recall that the support of $\mu \in M(\mathbb{T}^d)$, $\text{supp}(\mu)$, is the complement of all open sets $A \subseteq \mathbb{T}^d$ such that $\mu(A) = 0$.

Proposition 7.2.1. *Let F be spectral data on a finite set $\Lambda \subseteq \mathbb{Z}^d$, where $F = \widehat{\mu} \upharpoonright_{\Lambda}$ for some $\mu \in M(\mathbb{T}^d)$.*

(a) $\mathcal{E}(F, \Lambda) \subseteq M(\mathbb{T}^d)$ is non-empty, weak-* compact, and convex.

(b) $C(\mathbb{T}^d; \Lambda)$ is a closed subspace of $C(\mathbb{T}^d)$.

(c) $U(\mathbb{T}^d; \Lambda)$ is a compact subset of $C(\mathbb{T}^d; \Lambda)$.

(d) $\varepsilon(F, \Lambda) = \|L_{\mu}\|$.

(e) $\varepsilon(F, \Lambda) = \max_{f \in U} |L_{\mu}(f)|$.

(f) $\mathcal{D}(F, \Lambda)$ is non-empty.

Proof.

(a) By definition of ε , there exists a sequence $\{\nu_j\}$ such that $\|\nu_j\|_{\text{TV}} \rightarrow \varepsilon$ and $\widehat{\nu}_j = F$ on Λ . Then, this sequence is bounded. By Banach-Alaoglu, after passing to a subsequence, we can assume there exists $\nu \in M(\mathbb{T}^d)$ such that $\nu_j \rightarrow \nu$ in the weak-* topology.

Let V be the closed unit ball of $C(\mathbb{T}^d)$. We have $\|\nu\|_{\text{TV}} \leq \varepsilon$ because

$$\|\nu\|_{\text{TV}} = \sup_{f \in V} |\langle f, \nu \rangle| = \sup_{f \in V} \lim_{j \rightarrow \infty} |\langle f, \nu_j \rangle| \leq \sup_{f \in V} \lim_{j \rightarrow \infty} \|f\|_{L^{\infty}} \|\nu_j\| \leq \varepsilon.$$

Moreover, for each $m \in \Lambda$, we have

$$F(m) = \lim_{j \rightarrow \infty} \widehat{\nu}_j(m) = \lim_{j \rightarrow \infty} \int_{\mathbb{T}^d} e^{-2\pi i m \cdot x} d\nu_j(x) = \int_{\mathbb{T}^d} e^{-2\pi i m \cdot x} d\nu(x) = \widehat{\nu}(m).$$

This shows that ν is an extrapolation, and thus, $\|\nu\|_{\text{TV}} \geq \varepsilon$. Therefore, $\nu \in \mathcal{E}$.

The proof that \mathcal{E} is weak-* compact is similar. Pick any sequence $\{\nu_j\} \subseteq \mathcal{E}$ and after passing to a subsequence, we can assume $\nu_j \rightarrow \nu$ in the weak-* topology for some $\nu \in M(\mathbb{T}^d)$. By the same argument, we see that $\nu \in \mathcal{E}$.

If \mathcal{E} contains exactly one measure, then \mathcal{E} is trivially convex. Otherwise, let $t \in [0, 1]$, $\nu_0, \nu_1 \in \mathcal{E}$, and $\nu_t = (1-t)\nu_0 + t\nu_1$. Then, ν_t is an extrapolation and thus, $\|\nu_t\| \geq \varepsilon$. By the triangle inequality, we have $\|\nu_t\| \leq (1-t)\|\nu_0\| + t\|\nu_1\| = \varepsilon$. Thus, $\nu_t \in \mathcal{E}$ for each $t \in [0, 1]$.

(b) Suppose $\{f_j\} \subseteq C(\mathbb{T}^d)$ and that there exists $f \in C(\mathbb{T}^d)$ such that $f_j \rightarrow f$ uniformly. Then, $\widehat{f_j}(m) \rightarrow \widehat{f}(m)$ for all $m \in \mathbb{Z}^d$. Since $\widehat{f_j}(m) = 0$ if $m \notin \Lambda$, we deduce that $\widehat{f}(m) = 0$ if $m \notin \Lambda$. This shows that $f \in C(\mathbb{T}^d; \Lambda)$ and thus, $C(\mathbb{T}^d; \Lambda)$ is a closed subspace of $C(\mathbb{T}^d)$.

(c) Let $\{f_j\} \subseteq U$. We first show that $\{f_j\}$ is a uniformly bounded equicontinuous family. By definition, $\|f_j\|_{L^\infty} \leq 1$, and there exist $a_{j,m} \in \mathbb{C}$ such that $f_j(x) = \sum_{m \in \Lambda} a_{j,m} e^{2\pi i m \cdot x}$. Note that $|a_{j,m}| \leq \|f_j\|_\infty \leq 1$. Then, for any $x, y \in \mathbb{T}^d$, we have

$$|f_j(x) - f_j(y)| = \left| \sum_{m \in \Lambda} a_{j,m} (e^{2\pi i m \cdot x} - e^{2\pi i m \cdot y}) \right| \leq \sum_{m \in \Lambda} |e^{2\pi i m \cdot x} - e^{2\pi i m \cdot y}|.$$

Let $\varepsilon > 0$ and $m \in \Lambda$. There exists $\delta_m > 0$ such that $|e^{2\pi i m \cdot x} - e^{2\pi i m \cdot y}| < \varepsilon$ whenever $|x - y| < \delta_m$. Let $\delta = \min_{m \in \Lambda} \delta_m$. Combining this with the previous inequalities, we have

$$|f_j(x) - f_j(y)| \leq \sum_{m \in \Lambda} |e^{2\pi i m \cdot x} - e^{2\pi i m \cdot y}| \leq \#\Lambda \varepsilon,$$

whenever $|x - y| < \delta$. This shows that $\{f_j\}$ is a uniformly bounded equicontinuous family.

By the Arzelà-Ascoli theorem, there exists $f \in C(\mathbb{T}^d)$, with $\|f\|_{L^\infty} \leq 1$, and a subsequence $\{f_{j_k}\}$ such that $f_{j_k} \rightarrow f$ uniformly. Thus, $\widehat{f_{j_k}}(m) \rightarrow \widehat{f}(m)$ for all $m \in \mathbb{Z}^d$, which shows that $f \in C(\mathbb{T}^d; \Lambda)$.

(d) Let $\nu \in \mathcal{E}$. Then,

$$\forall f \in U, \quad |L_\mu(f)| = |\langle f, \mu \rangle| = |\langle f, \nu \rangle| \leq \|f\|_{L^\infty} \|\nu\|_{\text{TV}} \leq \varepsilon,$$

which proves the upper bound, $\|L_\mu\| \leq \varepsilon$.

For the lower bound, we use the Hahn-Banach theorem to extend $L_\mu \in C(\mathbb{T}^d; \Lambda)'$ to $\ell \in C(\mathbb{T}^d)'$, where $\|L_\mu\| = \|\ell\|$. By the Riesz representation theorem, there exists a unique $\sigma \in M(\mathbb{T}^d)$ such that $\ell(f) = \langle f, \sigma \rangle$ for all $f \in C(\mathbb{T}^d)$ and $\|\sigma\| = \|\ell\|$.

In particular,

$$\forall f \in C(\mathbb{T}^d; \Lambda), \quad \langle f, \sigma \rangle = \ell(f) = L_\mu(f) = \langle f, \mu \rangle.$$

Set $f(x) = e^{-2\pi i m \cdot x}$, where $m \in \Lambda$, to deduce that $\widehat{\mu} = \widehat{\sigma}$ on Λ . This implies $\|\sigma\| \geq \varepsilon$. Combining these facts, we have

$$\varepsilon \leq \|\sigma\| = \|\ell\| = \|L_\mu\|.$$

This proves the lower bound.

(e) We know that $\|L_\mu\| = \varepsilon$. By definition, there exists $\{f_j\} \subseteq U$ such that $|L_\mu(f_j)| \geq \varepsilon - 1/j$. By compactness of U , there exists a subsequence $\{f_{j_k}\} \subseteq U$

and $f \in U$ such that $f_{j_k} \rightarrow f$ uniformly. We immediately have $|L_\mu(f)| \leq \|f\|_{L^\infty} \varepsilon \leq \varepsilon$. For the reverse inequality, as $k \rightarrow \infty$,

$$|L_\mu(f)| \geq |L_\mu(f_{j_k})| - |L_\mu(f - f_{j_k})| \geq \varepsilon - \frac{1}{j_k} - \|\mu\|_{\text{TV}} \|f - f_{j_k}\|_{L^\infty} \rightarrow \varepsilon.$$

This proves that $|L_\mu(f)| = \varepsilon$.

(f) By the previous part and a simple calculation,

$$\varepsilon = \max_{f \in U} |\langle f, \mu \rangle| = \max_{f \in U} \left| \sum_{m \in \Lambda} \widehat{f}(m) \overline{F(m)} \right| = \max_{f \in U} \left| \sum_{m \in \Lambda} \widehat{f}(m) F(m) \right|.$$

Hence, there exists a $\varphi \in U$ that attains this maximum, which by definition, is a dual polynomial.

□

Proposition 7.2.2a,b establish *uniform* statements: Even if there are several minimal extrapolations, they must all be supported in a common set and they must have similar sign patterns. Proposition 7.2.2c characterizes the case that $|\varphi| \equiv 1$ in terms of the set Γ . Intuitively, $\#\Gamma$ represents the number of “bad” dual polynomials. While it is desirable to have $\Gamma = \emptyset$, perhaps surprisingly, Proposition 7.2.2d shows that we can make strong statements even when $\#\Gamma$ is large.

Proposition 7.2.2. *Let F be spectral data on a finite set $\Lambda \subseteq \mathbb{Z}^d$, where $F = \widehat{\mu} \upharpoonright_\Lambda$ for some $\mu \in M(\mathbb{T}^d)$.*

(a) *For any $\varphi \in \mathcal{D}(F, \Lambda)$, we have*

$$\forall \nu \in \mathcal{E}(F, \Lambda), \quad \text{supp}(\nu) \subseteq \{x \in \mathbb{T}^d : |\varphi(x)| = 1\}.$$

(b) There exists $\varphi \in \mathcal{D}(F, \Lambda)$ such that

$$\forall \nu \in \mathcal{E}(F, \Lambda), \quad \varphi = \text{sign}(\nu) \quad \nu\text{-a.e.}$$

(c) There exists $\varphi \in \mathcal{D}(F, \Lambda)$ with $|\varphi| \equiv 1$ if and only if $\Gamma \neq \emptyset$.

(d) For each $m \in \mathbb{Z}^d$, define $\alpha_m \in \mathbb{R}/\mathbb{Z}$ by the formula $e^{-2\pi i \alpha_m} F(m) = |F(m)|$. If

$m \in \Gamma$, then

$$\forall \nu \in \mathcal{E}, \quad \text{sign}(\nu)(x) = e^{2\pi i \alpha_m} e^{2\pi i m \cdot x} \quad \nu\text{-a.e.}$$

Proof.

(a) Using that $\varphi \in U(\mathbb{T}^d; \Lambda)$ and Proposition 7.2.1e, for all $\nu \in \mathcal{E}$, we have

$$|\langle \varphi, \nu \rangle| = |\langle \varphi, \mu \rangle| = |L_\mu(\varphi)| = \varepsilon = \|\nu\|_{\text{TV}}.$$

Since $\|\varphi\|_{L^\infty} \leq 1$ and $|\langle \varphi, \nu \rangle| = \|\nu\|_{\text{TV}}$, there exists $\theta \in \mathbb{R}$ such that $\varphi = e^{2\pi i \theta} \text{sign}(\nu)$ ν -a.e. Using that $|\varphi| = |\text{sign}(\nu)| = 1$ ν -a.e. and that ν is a Radon measure,

$$\text{supp}(\nu) \subseteq \overline{\{x \in \mathbb{T}^d: |\varphi(x)| = 1\}} = \{x \in \mathbb{T}^d: |\varphi(x)| = 1\}.$$

The last equality holds because the inverse image of the closed set $\{1\}$ under the continuous function $|\varphi|$ is closed.

(b) From the previous part, we know there exists a $f \in \mathcal{D}(F, \Lambda)$ such that $|\langle f, \mu \rangle| = \varepsilon$. Then, there exists $\theta \in \mathbb{R}$ such that $e^{2\pi i \theta} \langle f, \mu \rangle = \varepsilon$. Define the function $\varphi = e^{2\pi i \theta} f$, which is also a dual polynomial. Repeating the same argument from the previous part, we see that for all $\nu \in \mathcal{E}$, $\varphi = \text{sign}(\nu)$ ν -a.e.

(c) If $\varphi \in \mathcal{D}(F, \Lambda)$ and $|\varphi| \equiv 1$, then $\varphi = e^{2\pi i\theta} e^{2\pi im \cdot x}$ for some $\theta \in \mathbb{R}$ and $m \in \Lambda$.

Then, we have $\varepsilon = |\langle \varphi, \mu \rangle| = |F(m)|$ which shows that $m \in \Gamma$.

Conversely, let $m \in \Gamma$. Then we readily check that $\varphi(x) = e^{2\pi im \cdot x}$ is a dual polynomial and $|\varphi| \equiv 1$.

(d) Suppose $m \in \Gamma$ and $\nu \in \mathcal{E}$. Then

$$\begin{aligned} \int_{\mathbb{T}^d} e^{-2\pi i\alpha_m} e^{-2\pi im \cdot x} d\nu(x) &= e^{-2\pi i\alpha_m} \widehat{\nu}(m) \\ &= e^{-2\pi i\alpha_m} F(m) = |F(m)| = \varepsilon = \|\nu\|_{\text{TV}}. \end{aligned}$$

This shows $\text{sign}(\nu)(x) = e^{2\pi i\alpha_m} e^{2\pi im \cdot x}$ ν -a.e.

□

7.3 The Beurling Theory

We are ready to prove our main theorem.

Theorem 7.3.1. *Let F be spectral data on a finite set $\Lambda \subseteq \mathbb{Z}^d$.*

(a) *Suppose $\Gamma = \emptyset$. For any $\varphi \in \mathcal{D}(F, \Lambda)$, the closed set $S = \{x \in \mathbb{T}^d: |\varphi(x)| = 1\}$ has d -dimensional Lebesgue measure zero, and each minimal extrapolation of F on Λ is a singular measure supported in S .*

In particular, if $d = 1$, then S is a finite number of points and so each minimal extrapolation is a discrete measure supported in S .

(b) *Suppose $\#\Gamma \geq 2$. For each distinct pair $m, n \in \Gamma$, define $\alpha_{m,n} \in \mathbb{R}/\mathbb{Z}$ by*

$e^{2\pi i\alpha_{m,n}} = F(m)/F(n)$ and define the closed set,

$$S = \bigcap_{\substack{m,n \in \Gamma \\ m \neq n}} \{x \in \mathbb{T}^d : x \cdot (m - n) + \alpha_{m,n} \in \mathbb{Z}\}, \quad (7.1)$$

which is an intersection of $\binom{\#\Gamma}{2}$ periodic hyperplanes. Then, each minimal extrapolation of F on Λ is a singular measure supported in S .

In particular, if $d = 1$, then S is a finite number of points and so each minimal extrapolation is a discrete measure supported in S .

If $d \geq 2$ and there exist d linearly independent vectors, $p_1, p_2, \dots, p_d \in \mathbb{Z}^d$, such that

$$\{p_1, p_2, \dots, p_d\} \subseteq \{m - n : m, n \in \Gamma\},$$

then S is a lattice on \mathbb{T}^d .

Proof.

(a) Let $\varphi \in \mathcal{D}(F, \Lambda)$. By Proposition 7.2.2a,c, $|\varphi| \not\equiv 1$ because $\Gamma = \emptyset$ and each minimal extrapolation is supported in the closed set

$$S = \{x \in \mathbb{T}^d : |\varphi(x)| = 1\}.$$

Consider the function

$$\Phi(x) = 1 - |\varphi(x)|^2 = 1 - \sum_{m \in \Lambda} \sum_{n \in \Lambda} \widehat{\varphi}(m) \overline{\widehat{\varphi}(n)} e^{2\pi i(m-n) \cdot x}.$$

Then, the minimal extrapolations are supported in the closed set

$$S = \{x \in \mathbb{T}^d : \Phi(x) = 0\}.$$

Note that $\Phi \neq 0$ because $|\varphi| \neq 1$. Since Φ is a non-trivial real-analytic function, S is a set of d -dimensional Lebesgue measure zero. In particular, if $d = 1$, then S is a finite set of points.

- (b) Let $m, n \in \Gamma$. There exist $\alpha_m, \alpha_n \in \mathbb{R}/\mathbb{Z}$ defined in Proposition 7.2.2d, such that

$$\forall \nu \in \mathcal{E}, \quad \text{sign}(\nu)(x) = e^{2\pi i \alpha_m} e^{2\pi i m \cdot x} = e^{2\pi i \alpha_n} e^{2\pi i n \cdot x} \quad \nu\text{-a.e.}$$

Set $\alpha_{m,n} = \alpha_m - \alpha_n \in \mathbb{R}/\mathbb{Z}$. Then each minimal extrapolation is supported in

$$S_{m,n} = \{x \in \mathbb{T}^d : x \cdot (m - n) + \alpha_{m,n} \in \mathbb{Z}\}.$$

Thus, *each minimal extrapolation is supported in the set*

$$S = \bigcap_{\substack{m,n \in \Gamma \\ m \neq n}} S_{m,n} = \bigcap_{\substack{m,n \in \Gamma \\ m \neq n}} \{x \in \mathbb{T}^d : x \cdot (m - n) + \alpha_{m,n} \in \mathbb{Z}\}.$$

Note that $e^{2\pi i \alpha_{m,n}} = F(m)/F(n)$ because

$$e^{-2\pi i \alpha_m} F(m) = |F(m)| = \varepsilon = |F(n)| = e^{-2\pi i \alpha_n} F(n).$$

Suppose $\{p_1, \dots, p_d\}$ satisfies the hypothesis. By the support assertion that we just proved, there exists $\beta = (\beta_1, \beta_2, \dots, \beta_d) \in \mathbb{T}^d$ such that every minimal extrapolation is supported in

$$S = \bigcap_{j=1}^d \{x \in \mathbb{T}^d : x \cdot p_j + \beta_j \in \mathbb{Z}\}.$$

Let us explain the geometry of the situation before we proceed with the proof that S is a lattice. Note that $\{x \in \mathbb{T}^d : x \cdot p_j + \beta_j \in \mathbb{Z}\}$ is a family of parallel and periodically spaced hyperplanes. Since the vectors, p_1, p_2, \dots, p_d , are assumed

to be linearly independent, one family of hyperplanes is not parallel to any other family of hyperplanes. Hence, the intersection of d non-parallel and periodically spaced hyperplanes is a lattice, see Figure 7.1 for an illustration.

For a rigorous proof of this observation, first note that

$$S = \{x \in \mathbb{T}^d : Px + \beta \in \mathbb{Z}^d\},$$

where $P = (p_1, p_2, \dots, p_d)^t \in \mathbb{Z}^{d \times d}$ is invertible because its rows are linearly independent. Let $x_0 \in \mathbb{R}^d$ be the solution to $Px + \beta = 0$, and let $q_j \in \mathbb{Q}^d$ be the solution to $Px = e_j$, where e_j is the standard basis vector for \mathbb{R}^d . Then $p_j \cdot q_k = \delta_{j,k}$, and S is generated by the point x_0 and the lattice vectors, q_1, q_2, \dots, q_d .

□

Remark 7.3.2. Example 7.6.8 provides an example where $\#\Gamma = 3$ and there exists a singular continuous minimal extrapolation. This demonstrates that the conclusion of Theorem 7.3.1b is optimal.

We explain why Theorem 7.3.1 is an adaptation to the torus and a generalization to higher dimensions of Beurling's theorem. Let $M_b(\mathbb{R})$ be the space of complex Radon measures on \mathbb{R} with finite total variation norm. Since this is the only part that deals with measures on \mathbb{R} , we slightly abuse notation, and we denote the total variation norm on \mathbb{R} by $\|\cdot\|$ and the Fourier transform of $\mu \in M_b(\mathbb{R})$ by $\hat{\mu}: \mathbb{R} \rightarrow \mathbb{C}$.

Suppose we are given the set $\Lambda = [-\lambda, \lambda]$ and the spectral data $G = \hat{\mu}|_\Lambda$ for some $\mu \in M_b(\mathbb{R})$. Beurling studied the total variation problem on \mathbb{R} of finding

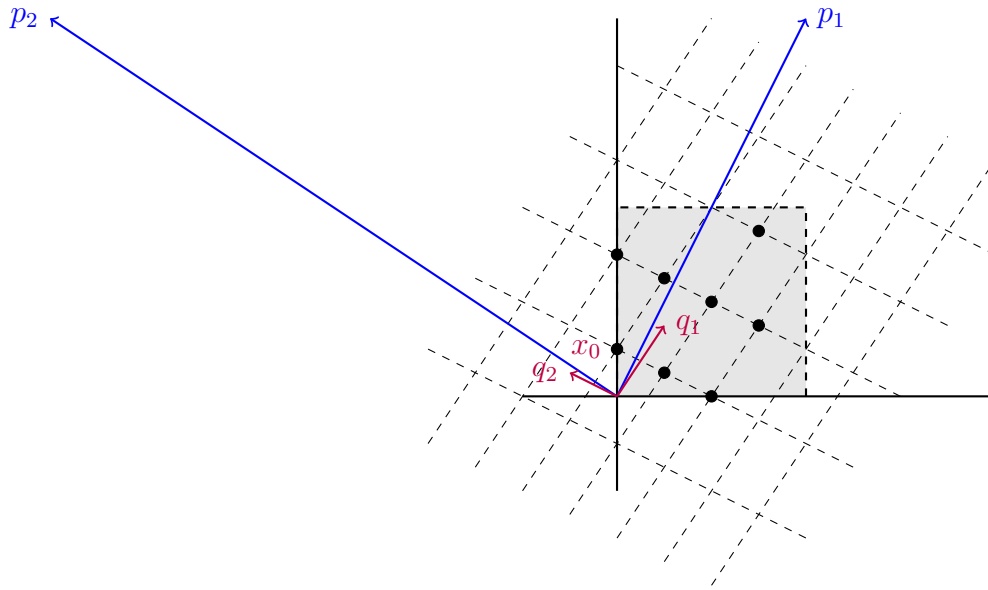


Figure 7.1: An illustration of Theorem 7.3.1b, where $d = 2$, $p_1 = (1, 2)$, $p_2 = (-3, 2)$, $\beta_1 = 1/2$, $\beta_2 = -1/2$, $q_1 = (1/4, 3/8)$, and $q_2 = (-1/4, 1/8)$. The family of hyperplanes are the dashed lines, the lattice S is the black dots, and the shaded region is $[0, 1)^2$.

solution(s) ν to

$$\inf_{\nu} \|\nu\|_{\text{TV}} \quad \text{such that} \quad \nu \in M_b(\mathbb{R}) \quad \text{and} \quad G = \widehat{\nu} \quad \text{on } \Lambda. \quad (7.2)$$

The solutions to this problem are called minimal extrapolations of G on Λ . He defined the quantities

$$m = \inf\{\|\nu\|_{\text{TV}} : \nu \in M_b(\mathbb{R}) \text{ and } G = \widehat{\nu} \text{ on } \Lambda\},$$

$$M = \{\nu \in M_b(\mathbb{R}) : G = \widehat{\nu} \text{ on } \Lambda \text{ and } \|\nu\|_{\text{TV}} = m\},$$

$$\Lambda_m = \{\gamma \in \Lambda : |G(\gamma)| = m\}.$$

Theorem 7.3.3 (Beurling, Theorem 2, page 362, [18]). *Let G be spectral data on $\Lambda = [-\lambda, \lambda]$, where $G = \widehat{\mu}|_{\Lambda}$ for some $\mu \in M_b(\mathbb{R})$.*

(a) *Suppose $\#\Lambda_m = 0$. There exist sequences, $\{a_k\} \subseteq \mathbb{C}$ and $\{x_k\} \subseteq \mathbb{R}$, for which*

$\#\{x_k : |x_k| < r\} = O(r)$ as $r \rightarrow \infty$, and such that

$$\nu = \sum_{k=1}^{\infty} a_k \delta_{x_k}$$

is the unique minimal extrapolation of G on Λ_m .

(b) *Suppose $\#\Lambda_m \geq 2$ and $\Lambda_m \neq \Lambda$. Then, Λ_m is a finite set, which allows us to*

define $\tau > 0$ as the smallest distance between any two points in Λ_m . Further,

there exist $\{a_k\} \subseteq \mathbb{C}$ and $x_0 \in \mathbb{R}$, such that

$$\nu = \sum_{k=-\infty}^{\infty} a_k \delta_{x_0 + \frac{k}{\tau}}$$

is the unique minimal extrapolation of G on Λ_m .

(c) *If $\Lambda_m = \Lambda$, then there exist $\alpha \in \mathbb{R}/\mathbb{Z}$ and $x \in \mathbb{R}$, such that $\nu = m e^{2\pi i \alpha} \delta_x$ is the*

unique minimal extrapolation of G on Λ_m .

Remark 7.3.4. It is difficult to deduce information about the minimal extrapolations when $\#\Lambda_m = 1$ because they might not be unique and there may exist positive absolutely continuous minimal extrapolations, e.g., see [18, 63, 64, 84] for specific examples and related work.

Remark 7.3.5. Beurling relied heavily on complex analysis and did not provide a higher dimensional version of Theorem 7.3.3. In contrast, we avoided the use of complex analysis in our proofs of Theorem 7.3.1 and the propositions leading to it. The extension to higher dimensions is not trivial, since as we saw, geometry plays a major role in $d \geq 2$. There are two important advantages of working with \mathbb{T}^d as opposed to \mathbb{R} from an application point of view. First, it is reasonable to assume that measures encountered in applications are compactly supported, and thus, their supports can be normalized to be the unit cube. Second, it is not clear how to solve (7.2) numerically, whereas there are algorithms for solving (TV).

7.4 Uniqueness and Non-uniqueness

While the mathematical theory we have developed connects the set Γ with the support of the minimal extrapolations, the difficulty of applying the theory is that in general, ε is unknown. However, in many important situations, it is possible to deduce the value of ε . We say $[A, B] \subseteq \mathbb{R}^+$ is an *admissibility range* for ε provided that $0 \leq A \leq \varepsilon \leq B$. The following proposition shows that we have $A = \|F\|_{\ell^\infty(\Lambda)}$ and $B = \|\mu\|_{\text{TV}}$ as an admissibility range, and B can be improved in certain situations. While the proof is elementary and is a direct consequence of

Hölder's inequality, we shall use this proposition throughout the remainder of this paper; and this also demonstrates its centrality and importance in our theory.

Proposition 7.4.1. *Let F be spectral data on a finite set $\Lambda \subseteq \mathbb{Z}^d$, where $F = \widehat{\mu} \upharpoonright_{\Lambda}$ for some $\mu \in M(\mathbb{T}^d)$. We have the lower and upper bounds,*

$$\|F\|_{\ell^\infty(\Lambda)} \leq \varepsilon \leq \|\mu\|_{TV}. \quad (7.3)$$

Further, if there exists an extrapolation $\nu \in M(\mathbb{T}^d)$ and $\|\nu\|_{TV} < \|\mu\|_{TV}$, then

$$\|F\|_{\ell^\infty(\Lambda)} \leq \varepsilon \leq \|\nu\|_{TV} < \|\mu\|_{TV}. \quad (7.4)$$

Proof. To see the lower bound for ε in (7.3) and (7.4), let σ be a minimal extrapolation of F on Λ . Then,

$$\sup_{m \in \Lambda} |F(m)| = \sup_{m \in \Lambda} |\widehat{\sigma}(m)| \leq \|\sigma\| = \varepsilon.$$

The upper bounds, $\varepsilon \leq \|\mu\|_{TV}$ and $\varepsilon \leq \|\nu\|_{TV}$ in (7.3) and (7.4), follow by definition of ε . □

Remark 7.4.2. By tightening the admissibility range for ε , we can deduce information about the minimal extrapolations. The simplest case is when $\varepsilon = \|F\|_{\ell^\infty(\Lambda)} = \|\mu\|_{TV}$, see Examples 7.6.2 and 7.6.3. The next simplest case is when there exists an extrapolation ν , such that $\varepsilon = \|F\|_{\ell^\infty(\Lambda)} = \|\nu\|_{TV}$, see Example 7.6.4. A more complicated case is when $\|F\|_{\ell^\infty(\Lambda)} < \varepsilon$, see Example 7.6.5.

Candès and Fernandez-Granda [30] focused entirely on the case that $\|\mu\|_{TV} = \varepsilon$ because this is a necessary condition to recover μ using (TV). In contrast to their

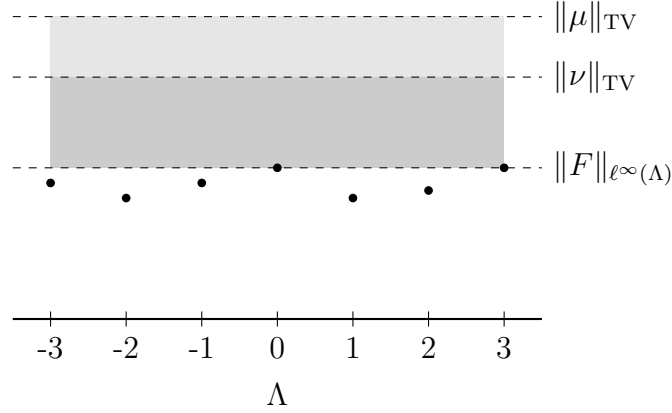


Figure 7.2: The black points represent the values of $|F|$ on Λ . The union of the light and dark gray regions is the admissibility range $[\|F\|_{\ell^\infty(\Lambda)}, \|\mu\|_{\text{TV}}]$, whereas the dark gray region is an improved admissibility range $[\|F\|_{\ell^\infty(\Lambda)}, \|\nu\|_{\text{TV}}]$.

result, our theory is better suited for situations when $\|F\|_{\ell^\infty(\Lambda)} = \varepsilon$. This is because Theorem 7.3.1b is strongest when $|F(m)| = \varepsilon$ for many points $m \in \Lambda$, i.e., when $\#\Gamma$ is large. Hence, our theory is better suited for deducing the impossibility of reconstruction using (TV).

Algorithm 6.2.2 fails precisely when it computes a dual polynomial φ such that $|\varphi| \equiv 1$. The following proposition shows that this occurs precisely when $\varepsilon = \|F\|_{\ell^\infty}$. Since we are given F on Λ , we then immediately get the exact value of ε .

Proposition 7.4.3. *Let F be spectral data on a finite set $\Lambda \subseteq \mathbb{Z}^d$, where $F = \widehat{\mu}|_\Lambda$ for some $\mu \in M(\mathbb{T}^d)$. Suppose there exists $\varphi \in \mathcal{D}(F, \Lambda)$ such that $|\varphi| \equiv 1$. Then, $\Gamma \neq \emptyset$ and $\varepsilon = \|F\|_{\ell^\infty(\Lambda)}$.*

Proof. By Proposition 7.2.2c, we must have $\Gamma \neq \emptyset$. Let $m \in \Gamma$ and by definition of Γ , we have $|F(m)| = \varepsilon$. This combined with the lower bound in Proposition 7.4.1 proves that $\varepsilon = \|F\|_{\ell^\infty(\Lambda)}$. □

Theorem 7.3.1 provides sufficient conditions for the minimal extrapolations to be supported in a discrete set, but it does not provide sufficient conditions for uniqueness. Since a family of discrete measures supported on a common set behaves essentially like a vector, we use basic linear algebra to address the question of uniqueness when the minimal extrapolations are supported in a discrete set. The following proposition is well-known, see, e.g., [51] for a similar result.

Proposition 7.4.4. *Let F be spectral data on a finite set $\Lambda = \{m_1, m_2, \dots, m_J\} \subseteq \mathbb{Z}^d$, where $F = \widehat{\mu} \upharpoonright_{\Lambda}$ for some $\mu \in M(\mathbb{T}^d)$. Suppose there exists a finite set, $\{x_k\}_{k=1}^K \subseteq \mathbb{T}^d$, such that each minimal extrapolation of F on Λ is supported in this set. Suppose that the matrix,*

$$E(m_1, \dots, m_J; x_1, \dots, x_K) = \begin{pmatrix} e^{-2\pi i m_1 \cdot x_1} & e^{-2\pi i m_1 \cdot x_2} & \dots & e^{-2\pi i m_1 \cdot x_K} \\ e^{-2\pi i m_2 \cdot x_1} & e^{-2\pi i m_2 \cdot x_2} & \dots & e^{-2\pi i m_2 \cdot x_K} \\ \vdots & \vdots & & \vdots \\ e^{-2\pi i m_J \cdot x_1} & e^{-2\pi i m_J \cdot x_2} & \dots & e^{-2\pi i m_J \cdot x_K} \end{pmatrix}, \quad (7.5)$$

has full column rank (this can only occur if $J \geq K$). Then, there is a unique minimal extrapolation of F on Λ .

Proof. Let ν be the difference of any two minimal extrapolations of F on Λ . Then ν is also supported in $\{x_k\}_{k=1}^K$ and it is of the form $\nu = \sum_{k=1}^K a_k \delta_{x_k}$. Since $\widehat{\nu} = 0$ on Λ , we have $0 = \widehat{\nu}(m_j) = \sum_{k=1}^K a_k e^{-2\pi i m_j \cdot x_k}$ for $j = 1, 2, \dots, J$, which is equivalent to the linear system $Ea = 0$, where $a = (a_1, \dots, a_K) \in \mathbb{C}^K$. By assumption, E has full column rank. This implies $a = 0$. \square

We address the situation when $\#\Gamma = 1$, i.e., the missing case of Theorem 7.3.1.

A measure $\mu \in M(\mathbb{T}^d)$ is *positive* if $\mu(A) \geq 0$ for all Borel sets $A \subseteq \mathbb{T}^d$. A sequence $\{a_m\}_{m \in \mathbb{Z}^d}$ is *positive-definite* if for all sequences $\{b_m\}_{m \in \mathbb{Z}^d}$ of finite support, we have

$$\sum_{m, n \in \mathbb{Z}^d} a_{m-n} b_m \overline{b_n} \geq 0.$$

Bochner's theorem, that is true for all locally compact abelian groups [121], asserts in our case of \mathbb{T}^d and \mathbb{Z}^d that a sequence is positive-definite if and only if it is the sequence of Fourier coefficients of a positive measure. In this special case of \mathbb{T}^d and \mathbb{Z}^d , Bochner's theorem is called Herglotz' theorem. Herglotz proved it in 1911 for $d = 1$, and natural modifications to the proof in [90, page 39] yield the statement in higher dimensions.

Proposition 7.4.5. *Let F be spectral data on a finite set $\Lambda \subseteq \mathbb{Z}^d$. Suppose there exists $n \in \Lambda$ such that $\{F(m+n)\}_{m \in \Lambda-n}$ extends to a positive-definite sequence on \mathbb{Z}^d . Then, $n \in \Gamma$, and each positive-definite extension of $\{F(m+n)\}_{m \in \Lambda-n}$ corresponds to a positive measure ν , such that $e^{2\pi i n \cdot x} \nu(x)$ is a minimal extrapolation of F on Λ .*

Proof. Extend $\{F(m+n)\}_{m \in \Lambda-n}$ to a positive-definite sequence on \mathbb{Z}^d . By Herglotz' theorem, there exists a positive measure $\nu \in M(\mathbb{T}^d)$ such that $\widehat{\nu}(m) = F(m+n)$ for all $m \in \Lambda - n$. Then, $\sigma(x) = e^{2\pi i n \cdot x} \nu(x)$ is an extrapolation of F on Λ , which implies

$$\|\nu\|_{\text{TV}} = \|\sigma\| \geq \varepsilon.$$

For the reverse inequality, since ν is a positive measure, we have $\|\nu\|_{\text{TV}} = \widehat{\nu}(0)$.

Then,

$$\|\sigma\| = \|\nu\|_{\text{TV}} = \widehat{\nu}(0) = |\widehat{\nu}(0)| = |F(n)| \leq \|F\|_{\ell^\infty(\Lambda)} \leq \varepsilon(F, \Lambda),$$

where the last inequality follows by Proposition 7.4.1. This shows that $|F(n)| = \|\nu\|_{\text{TV}} = \varepsilon$, which proves that $n \in \Gamma$ and $M_n\nu$ is a minimal extrapolation of F on Λ . \square

Remark 7.4.6. Beurling [18] and Esseen [63] essentially proved the analogue of Proposition 7.4.5 for the special case that $n = 0$, and for \mathbb{R} instead of \mathbb{T}^d . Proposition 7.4.5 generalizes their result to handle situations when $0 \notin \Lambda$. This is important because from the viewpoint of Proposition 7.5.1c, the super-resolution problem is invariant under simultaneous translations of F and Λ , which means that $0 \in \mathbb{Z}^d$ is no more special than any other point $n \in \mathbb{Z}^d$.

Remark 7.4.7. Proposition 7.4.5 suggests that the case $\#\Gamma = 1$ is special compared to the cases $\#\Gamma = 0$ or $\#\Gamma \geq 2$ because, when $\#\Gamma = 1$, there may exist absolutely continuous minimal extrapolations. In Example 7.6.2, $\#\Gamma = 1$, and there exist uncountably many discrete and positive absolutely continuous minimal extrapolations.

Remark 7.4.8. Suppose that F can be extended to a positive-definite sequence on \mathbb{Z} . In theory, there are an infinite number of such extensions, and one particular method of choosing such an extension is called the *Maximum Entropy Method* (MEM). According to this method, one extends F to the positive-definite sequence $\{a_m\}_{m \in \mathbb{Z}}$ whose corresponding density function $f \in L^1(\mathbb{T})$ is the unique maximizer of a specific logarithmic integral associated with the physical notion of entropy, e.g., see [6, Theorems 3.6.3 and 3.6.6]. MEM is related to spectral estimation methods [37], the maximum likelihood method [37], and moment problems [97].

Finally, we show that the surface measure of the zero set of a trigonometric polynomial is a minimal extrapolation. If $A, B \subseteq \mathbb{Z}^d$, let $A - B = \{a - b: a \in A, b \in B\}$.

Proposition 7.4.9. *Let $\Lambda \subseteq \mathbb{Z}^d$ be a finite set and $\varphi \in C(\mathbb{T}^d; \Lambda)$ such that $S = \{x \in \mathbb{T}^d: \varphi(x) = 0\}$ is non-trivial. Let F be spectral data on $\Lambda - \Lambda$, where $F = \widehat{\sigma}|_{\Lambda - \Lambda}$ and σ is the surface measure of S . Then, $\Gamma \neq \emptyset$, σ is a positive minimal extrapolation of F on $\Lambda - \Lambda$, and there exists $\Phi \in \mathcal{D}(F, \Lambda - \Lambda)$ such that $S = \{x \in \mathbb{T}^d: |\Phi(x)| = 1\}$.*

Proof. Of course, σ is an extrapolation of F on $\Lambda - \Lambda$, so we have $\|\sigma\| \geq \varepsilon$. To prove the reverse inequality, we observe that σ is a positive measure and by Proposition [7.4.1](#),

$$\|\sigma\| = |\widehat{\sigma}(0)| = |F(0)| \leq \|F\|_{\ell^\infty(\Lambda - \Lambda)} \leq \varepsilon(F, \Lambda - \Lambda).$$

This also shows that $0 \in \Gamma$.

Since S is invariant if we multiply φ by a constant, without loss of generality, assume that $\|\varphi\|_{L^\infty} \leq 1$. Define the function $\Phi = 1 - |\varphi|^2$, and note that $\Phi \in U(\mathbb{T}^d; \Lambda - \Lambda)$ and $S = \{x \in \mathbb{T}^d: \Phi(x) = 1\}$. Using these facts, we have $|\langle \Phi, \sigma \rangle| = \|\sigma\| = \varepsilon$. This shows that Φ is a dual polynomial. \square

7.5 Additional Results

Our next goal is to examine the symmetries of the minimal extrapolations. We are interested in the vector space operations, namely, addition of measures and the multiplication of measures by complex numbers. We are also interested in the operations that are well-behaved under the Fourier transform on \mathbb{T}^d , namely, translation,

modulation, convolution, and product of measures. For any $y \in \mathbb{R}^d$, let M_y be the modulation operator defined by $M_y f(x) = e^{2\pi i y \cdot x} f(x)$ and let T_y be the translation operator $T_y f(x) = f(x - y)$.

Proposition 7.5.1. *Let F be spectral data on a finite set $\Lambda \subseteq \mathbb{Z}^d$. Let $a \in \mathbb{C}$ be non-zero, $n \in \mathbb{Z}^d$, and $y \in \mathbb{R}^d$.*

(a) *Multiplication by constants is bijective: $a\varepsilon(F, \Lambda) = \varepsilon(aF, \Lambda)$, and $\nu \in \mathcal{E}(F, \Lambda)$ if and only if $a\nu \in \mathcal{E}(aF, \Lambda)$.*

(b) *Translation is bijective: $\varepsilon(\mu, \Lambda) = \varepsilon(M_y \mu, \Lambda)$, and $\nu \in \mathcal{E}(F, \Lambda)$ if and only if $T_y \nu \in \mathcal{E}(M_y F, \Lambda)$.*

(c) *Minimal extrapolation is invariant under simultaneous shifts of F and Λ : $\varepsilon(F, \Lambda) = \varepsilon(T_n F, \Lambda + n)$, and $\mathcal{E}(\mu, \Lambda) = \mathcal{E}(T_n \mu, \Lambda + n)$.*

(d) *The product of minimal extrapolations is a minimal extrapolation for the product: For $j = 1, 2$, let F_j be spectral data on a finite set $\Lambda_j \subseteq \mathbb{Z}^{d_j}$ and let $\nu_j \in \mathcal{E}(F_j, \Lambda_j)$. Then $\varepsilon(F_1, \Lambda_1)\varepsilon(F_2, \Lambda_2) = \varepsilon(F_1 \otimes F_2, \Lambda_1 \times \Lambda_2)$, and $\nu_1 \times \nu_2 \in \mathcal{E}(F_1 \otimes F_2, \Lambda_1 \times \Lambda_2)$.*

Proof.

(a) If $\nu \in \mathcal{E}(F, \Lambda)$, then $a\nu$ is an extrapolation of aF on Λ . Suppose $a\nu \notin \mathcal{E}(aF, \Lambda)$.

Then there exists σ such that $\widehat{\sigma} = aF$ on Λ and $\|\sigma\| < \|a\nu\|_{\text{TV}}$. Thus, $\widehat{\sigma}/a = F$ on Λ and $\|\sigma/a\| < \|\nu\|_{\text{TV}}$, which contradicts the assumption that $\nu \in \mathcal{E}(F, \Lambda)$.

The converse follows by a similar argument.

(b) If $\nu \in \mathcal{E}(F, \Lambda)$, then $T_y\nu$ is an extrapolation of M_yF on Λ . Suppose $T_y\nu \notin \mathcal{E}(M_yF, \Lambda)$. Then there exists σ such that $\widehat{\sigma} = M_yF$ on Λ and $\|\sigma\| < \|T_y\nu\|_{\text{TV}} = \|\nu\|_{\text{TV}}$. Then $(T_{-y}\sigma)^\wedge = F$ on Λ and $\|T_{-y}\sigma\| = \|\sigma\| < \|\nu\|_{\text{TV}}$, which contradicts the assumption that $\nu \in \mathcal{E}(F, \Lambda)$. The converse follows by a similar argument.

(c) If $f \in U(\mathbb{T}^d; \Lambda)$, then $M_n f \in U(\mathbb{T}^d; \Lambda + n)$. We have,

$$|\langle f, \mu \rangle| = \left| \sum_{m \in \Lambda} \widehat{f}(m) \overline{\widehat{\mu}(m)} \right| = \left| \sum_{m \in \Lambda + n} (M_n f)^\wedge(m) \overline{(M_n \mu)^\wedge(m)} \right| = |\langle M_n f, M_n \mu \rangle|.$$

Using Proposition 7.2.1e, we see that $\varepsilon(F, \Lambda) = \varepsilon(T_n F, \Lambda + n)$.

If $\nu \in \mathcal{E}(F, \Lambda)$, then $M_n \nu$ is an extrapolation of $T_n F$ on $\Lambda + n$, and $\|M_n \nu\|_{\text{TV}} = \|\nu\|_{\text{TV}} = \varepsilon(F, \Lambda) = \varepsilon(T_n F, \Lambda + n)$. Thus, $M_n \nu \in \mathcal{E}(T_n F, \Lambda + n)$. The converse follows similarly.

(d) For convenience, let $\mu = \mu_1 \times \mu_2$, $\nu = \nu_1 \times \nu_2$, $\Lambda = \Lambda_1 \times \Lambda_2$, $F = F_1 \otimes F_2$, $\varepsilon_j = \varepsilon(F_j, \Lambda_j)$, and $\varepsilon = \varepsilon(F, \Lambda)$. Since ν is an extrapolation of F on Λ , by Proposition 7.4.1, we have

$$\varepsilon \leq \|\nu\|_{\text{TV}} = \|\nu_1\| \|\nu_2\| = \varepsilon_1 \varepsilon_2.$$

To see the reverse inequality, by Proposition 7.2.1f, there exist $\varphi_j \in \mathcal{D}(F_j, \Lambda_j)$, for $j = 1, 2$. Let $\varphi = \varphi_1 \otimes \varphi_2$ and observe that $\varphi \in U(\mathbb{T}^d; \Lambda)$. By Proposition 7.2.1e, we have

$$\varepsilon = \max_{f \in U(\mathbb{T}^d; \Lambda)} |\langle f, \mu \rangle| \geq |\langle \varphi, \mu \rangle| \geq |\langle \varphi_1, \mu_1 \rangle \langle \varphi_2, \mu_2 \rangle| = \varepsilon_1 \varepsilon_2.$$

This shows that $\varepsilon = \varepsilon_1 \varepsilon_2$, and, since $\|\nu\|_{\text{TV}} = \varepsilon_1 \varepsilon_2$, we conclude that $\nu \in \mathcal{E}(F, \Lambda)$.

□

While minimal extrapolation is well-behaved under translation, it is not well behaved under modulation. This is because the Fourier transform of modulation is translation, and so $\widehat{\mu}|_{\Lambda}$ and $(M_n\mu)^\wedge|_{\Lambda}$ are, in general, not equal. In contrast, the Fourier transform of translation is modulation, and so $\widehat{\mu}|_{\Lambda}$ and $(T_y\mu)^\wedge|_{\Lambda}$ only differ by a phase factor. We shall prove these statements in Proposition 7.5.2.

Proposition 7.5.2.

- (a) For $j = 1, 2$, there exist spectral data F_j on a finite subset $\Lambda \subseteq \mathbb{Z}$ and $\nu_j \in \mathcal{E}(F_j, \Lambda)$, such that $\nu_1 + \nu_2 \notin \mathcal{E}(F_1 + F_2, \Lambda)$.
- (b) There exist spectral data F on a finite subset $\Lambda \subseteq \mathbb{Z}$, $\nu \in \mathcal{E}(F, \Lambda)$, and $n \in \mathbb{Z}$, such that $M_n\nu \notin \mathcal{E}(T_nF, \Lambda)$.
- (c) For $j = 1, 2$, there exist spectral data F_j on a finite subset $\Lambda \subseteq \mathbb{Z}$, and $\nu_j \in \mathcal{E}(F_j, \Lambda)$, such that $\nu_1 * \nu_2 \notin \mathcal{E}(F_1F_2, \Lambda)$.

Proof.

- (a) Let $\mu_1 = \delta_0 + \delta_{1/2}$, $\mu_2 = -\delta_0 - \delta_{1/2}$, $\Lambda = \{-1, 0, 1\}$, and $F_j = \widehat{\mu}_j|_{\Lambda}$. By Example 7.6.2, we have $\nu_1 = \delta_0 + \delta_{1/2} \in \mathcal{E}(F_1, \Lambda)$, and $\nu_2 = -\delta_{1/4} - \delta_{3/4} \in \mathcal{E}(F_2, \Lambda)$. Then, $\mu_1 + \mu_2 = 0$, and so $\varepsilon(F_1 + F_2, \Lambda) = 0$. However, $\nu_1 + \nu_2 \notin \mathcal{E}(F_1 + F_2, \Lambda)$ because $\|\nu_1 + \nu_2\| = \|\delta_0 - \delta_{1/4} + \delta_{1/2} - \delta_{3/4}\| > 0$.
- (b) Let $\mu = \delta_0 + \delta_{1/2}$, $\Lambda = \{-1, 0, 1\}$, $n = -1$, and $F = \widehat{\mu}|_{\Lambda}$. By Example 7.6.2, we have $\nu = \delta_{1/4} + \delta_{3/2} \in \mathcal{E}(F, \Lambda)$. However, $M_{-1}\mu = \delta_0 - \delta_{1/2}$, and by Example 7.6.3, $\mathcal{E}(T_1F, \Lambda) = \{\delta_0 - \delta_{1/2}\}$. Thus, $M_{-1}\nu \notin \mathcal{E}(T_1F, \Lambda)$.

(c) Let $\mu_1 = \delta_0 + \delta_{1/2}$, $\mu_2 = \delta_0 - \delta_{1/2}$, $\Lambda = \{-1, 0, 1\}$, and $F_j = \widehat{\mu}_j|_\Lambda$. Then $F_1 F_2 = 0$ on Λ , which implies $\varepsilon(F_1 F_2, \Lambda) = 0$. By Examples 7.6.2 and 7.6.3, we have $\nu_1 = \mu_1 \in \mathcal{E}(F_1, \Lambda)$ and $\nu_2 = \mu_2 \in \mathcal{E}(F_2, \Lambda)$. However, $\nu_1 * \nu_2 \notin \mathcal{E}(F_1 F_2, \Lambda)$ because $\|\nu_1 * \nu_2\| = \|\delta_0 - \delta_1\| > 0$.

□

7.6 Examples

There are several reasons why we are interested in computing the minimal extrapolations of discrete measures. They are the simplest types of measures, and so, their minimal extrapolations can be computed rather easily. By Theorem 7.3.1, the minimal extrapolations of a non-discrete measure are sometimes discrete measures, so they appear naturally in our analysis.

Remark 7.6.1. In view of Proposition 7.5.1 a,b, and without loss of generality, we can assume any discrete measure $\mu = \sum_{k=1}^{\infty} a_k \delta_{x_k} \in M(\mathbb{T}^d)$, where $\sum_{k=1}^{\infty} |a_k| < \infty$, can be written as $\mu = \delta_0 + \sum_{k=2}^{\infty} a'_k \delta_{x'_k} \in M(\mathbb{T}^d)$, where $\sum_{k=2}^{\infty} |a'_k| < \infty$.

Example 7.6.2. Let $\Lambda = \{-1, 0, 1\}$ and $F = \widehat{\mu}|_\Lambda$, where $\mu = \delta_0 + \delta_{1/2}$. We have $F(0) = 2$, and $F(\pm 1) = 0$. Clearly $\|F\|_{\ell^\infty(\Lambda)} = \|\mu\|_{\text{TV}} = 2$. By Proposition 7.4.1, $\varepsilon = \|F\|_{\ell^\infty(\Lambda)} = \|\mu\|_{\text{TV}} = 2$, which implies μ is a minimum extrapolation of F on Λ .

Further, there is an uncountable number of discrete minimal extrapolations.

To see this, for each $y \in \mathbb{T}$ and any integer $K \geq 2$, define

$$\nu_{y,K} = \frac{2}{K} \sum_{k=0}^{K-1} \delta_{y + \frac{k}{K}}.$$

A straightforward calculation shows that $\nu_{y,K}$ is an extrapolation and that $\|\nu_{y,K}\| = \varepsilon$.

Also, we can construct positive absolutely continuous minimal extrapolations. One example is the constant function $f \equiv 2$ on \mathbb{T} . For other examples, let $N \geq 2$ and let $F_N \in C^\infty(\mathbb{T})$ be the Fejér kernel,

$$F_N(x) = \sum_{n=-N}^N \left(1 - \frac{|n|}{N+1}\right) e^{2\pi i n x}.$$

For any $c > 0$ such that $c \leq (2N+2)/(3N+1)$, extend F on Λ to the sequence $\{(a_{N,c})_m\}_{m \in \mathbb{Z}}$, where

$$(a_{N,c})_m = \begin{cases} 2 & m = 0, \\ c \left(1 - \frac{|m|}{N+1}\right) & 2 \leq |m| \leq N, \\ 0 & \text{otherwise.} \end{cases}$$

Consider the real-valued function

$$\begin{aligned} f_{N,c}(x) &= 2 + \sum_{n=-N}^{-2} (a_{N,c})_n e^{2\pi i n x} + \sum_{n=2}^N (a_{N,c})_n e^{2\pi i n x} \\ &= 2 + c \sum_{n=-N}^{-2} \left(1 - \frac{|n|}{N+1}\right) e^{2\pi i n x} + c \sum_{n=2}^N \left(1 - \frac{|n|}{N+1}\right) e^{2\pi i n x}. \end{aligned}$$

We check that $\widehat{f_{N,c}}(m) = (a_{N,c})_m$ for all $m \in \mathbb{Z}$, which implies $f_{N,c}$ is an extrapolation of μ . Using the upper bound on c , we have, for all $x \in \mathbb{T}$, that

$$2 \geq c + 2c \left(1 - \frac{1}{N+1}\right) \cos(2\pi x) = c + c \left(1 - \frac{1}{N+1}\right) e^{2\pi i x} + c \left(1 - \frac{1}{N+1}\right) e^{-2\pi i x}.$$

Using this inequality, and the definitions of F_N and $f_{N,c}$, we have

$$f_{N,c}(x) \geq c F_N \geq 0.$$

Since $f_{N,c} \geq 0$, we also have

$$\|f_{N,c}\|_1 = \int_{\mathbb{T}} f_{N,c}(x) dx = 2 + \int_{\mathbb{T}} \left(\sum_{n=-N}^{-2} (a_{N,c})_n e^{2\pi i n x} + \sum_{n=2}^N (a_{N,c})_n e^{2\pi i n x} \right) dx = 2 = \varepsilon.$$

Thus, for any $N \geq 2$ and $c \leq (2N + 2)/(3N + 1)$, $f_{N,c}$ is a *positive absolutely continuous* minimal extrapolation. Hence, we have constructed an uncountable number of positive absolutely continuous minimal extrapolations.

Example 7.6.3. Let $\Lambda = \{-1, 0, 1\}$ and $F = \hat{\mu}|_{\Lambda}$, where $\mu = \delta_0 - \delta_{1/2}$. We have $F(0) = 0$, and $F(\pm 1) = 2$. Further, we have $\|F\|_{\ell^\infty(\Lambda)} = \|\mu\|_{\text{TV}} = 2$, so that by Proposition 7.4.1, we have $\varepsilon = \|F\|_{\ell^\infty(\Lambda)} = \|\mu\|_{\text{TV}} = 2$. Thus, μ is a minimal extrapolation of F on Λ .

Consequently, $\Gamma = \{-1, 1\}$. By Theorem 7.3.1b, there exists $\alpha_{-1,1} \in \mathbb{R}/\mathbb{Z}$ satisfying

$$e^{2\pi i \alpha_{-1,1}} = \frac{F(-1)}{F(1)} = 1,$$

and the minimal extrapolations are supported in the set $\{x \in \mathbb{T} : 2x \in \mathbb{Z}\} = \{0, 1/2\}$.

This implies each $\nu \in \mathcal{E}$ is discrete and can be written as $\nu = a_1 \delta_0 + a_2 \delta_{1/2}$. In theory, a_1, a_2 depend on ν , so we cannot conclude uniqueness yet.

The matrix E from (7.5) is

$$E(-1, 0, 1; 0, 1/2) = \begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Clearly, E has full column rank. So, by Proposition 7.4.4, μ is the unique minimal extrapolation of F on Λ . Thus, reconstruction of μ from F on Λ is possible.

Example 7.6.4. Let $\Lambda = \{-1, 0, 1\}$ and $F = \widehat{\mu} |_{\Lambda}$, where $\mu = \delta_0 - \delta_{1/4}$. We have $F(\pm 1) = 1 \pm i = \sqrt{2}e^{\pm\pi i/4}$, and $F(0) = 0$. Note that $\|F\|_{\ell^\infty(\Lambda)} = \sqrt{2} < 2 = \|\mu\|_{\text{TV}}$, which shows that $\sqrt{2} \leq \varepsilon \leq 2$. We claim that $\varepsilon = \sqrt{2}$. To see this, consider $\nu = (-\delta_{3/8} + \delta_{7/8})/\sqrt{2}$. We verify that $\|\nu\|_{\text{TV}} = \sqrt{2}$ and that ν is an extrapolation of F on Λ . By Proposition 7.4.1, $\varepsilon = \sqrt{2}$, which implies ν is a minimal extrapolation of F on Λ . This also implies $\mu \notin \mathcal{E}(F, \Lambda)$, and so reconstruction of μ from F on Λ using (TV) is impossible.

We claim that ν is the unique minimal extrapolation. The matrix E from (7.5) is

$$E(-1, 0, 1; 3/8, 7/8) = \begin{pmatrix} e^{2\pi i 3/8} & e^{2\pi i 7/8} \\ 1 & 1 \\ e^{-2\pi i 3/8} & e^{-2\pi i 7/8} \end{pmatrix},$$

which we observe to have full column rank. By Proposition 7.4.4, we conclude that ν is the unique minimal extrapolation of F on Λ . Thus, reconstruction of ν from F on Λ is possible.

We explain the derivation of ν . We guess that $\varepsilon = \sqrt{2}$ and see what Theorem 7.3.1b implies. Under this assumption that $\varepsilon = \sqrt{2}$, we have $\Gamma = \{-1, 1\}$. By Theorem 7.3.1b, there exists $\alpha_{-1,1} \in \mathbb{R}/\mathbb{Z}$ satisfying

$$e^{2\pi i \alpha_{-1,1}} = \frac{F(1)}{F(-1)} = e^{\pi i/2},$$

and the minimal extrapolations are supported in $\{x \in \mathbb{T} : 2x + 1/4 \in \mathbb{Z}\} = \{3/8, 7/8\}$. Hence, if $\varepsilon = \sqrt{2}$, then every $\sigma \in \mathcal{E}(F, \Lambda)$ is of the form $\sigma = a_1\delta_{3/8} + a_2\delta_{7/8}$. Thus, by definition of a minimal extrapolation, $\|\sigma\| = \sqrt{2}$ and $F = \widehat{\sigma}$ on Λ . Using this information, we solve for the coefficients a_1, a_2 , and compute

that $|a_1| + |a_2| = \sqrt{2}$, $a_1 = -a_2$, and $a_1 = -\sqrt{2}/2$. Thus, we obtain that $\sigma = (-\delta_{3/8} + \delta_{7/8})/\sqrt{2}$. From here, we simply check that $\nu = \sigma$ is, in fact, a minimal extrapolation of F on Λ .

Example 7.6.5. Let $\Lambda = \{-1, 0, 1\}$ and $F = \hat{\mu} |_{\Lambda}$, where $\mu = \delta_0 + e^{\pi i/3} \delta_{1/3}$. We have $F(-1) = 0$, $F(0) = 1 + e^{\pi i/3} = \sqrt{3} e^{\pi i/6}$, and $F(1) = 1 + e^{-\pi i/3} = \sqrt{3} e^{-\pi i/6}$.

Suppose, for the purpose of obtaining a contradiction, that $\varepsilon = \|F\|_{\ell^\infty(\Lambda)} = \sqrt{3}$. Then $\Gamma = \{0, 1\}$. By Theorem 7.3.1b, there is $\alpha_{0,1} \in \mathbb{R}/\mathbb{Z}$ such that

$$e^{2\pi i \alpha_{0,1}} = \frac{F(0)}{F(1)} = e^{\pi i/3},$$

and each $\nu \in \mathcal{E}(F, \Lambda)$ is of the form $\nu = a\delta_{1/6}$ for some $a \in \mathbb{C}$. Then, $|\hat{\nu}| = |a|$ on \mathbb{Z} and, in particular, $F \neq \hat{\nu}$ on Λ , which is a contradiction.

Thus, $\varepsilon > \sqrt{3}$, i.e., $\Gamma = \emptyset$. Therefore, Theorem 7.3.1a applies, and so there is a finite set S such that $\text{supp}(\nu) \subseteq S$ for each $\nu \in \mathcal{E}(F, \Lambda)$. In particular, each $\nu \in \mathcal{E}(F, \Lambda)$ is discrete. Hence, we have to solve the optimization problem given in Proposition 7.2.1e, which is

$$\varepsilon = \max \left\{ \left| a\sqrt{3} e^{\pi i/6} + b\sqrt{3} e^{-\pi i/6} \right| : \forall x \in \mathbb{T}, |ae^{2\pi i x} + b + ce^{-2\pi i x}| \leq 1, a, b, c \in \mathbb{C} \right\}.$$

This optimization problem can be written as a semi-definite program, see [30, Corollary 4.1, page 936]. After obtaining numerical approximations to the optimizers of this problem, we guess that the the exact optimizers are

$$a = \frac{2}{3\sqrt{3}} e^{-\pi i/6}, \quad b = \frac{4}{3\sqrt{3}} e^{\pi i/6}, \quad c = -\frac{i}{3\sqrt{3}}.$$

These values of a, b, c are, in fact, the optimizers because $a\sqrt{3} e^{\pi i/6} + b\sqrt{3} e^{-\pi i/6} = 2$ and $|ae^{2\pi i x} + b + ce^{-2\pi i x}| \leq 1$ for all $x \in \mathbb{T}$. Thus, $\varepsilon = 2$ and μ is a minimal

extrapolation of F on Λ . Since $|ae^{2\pi ix} + b + ce^{-2\pi ix}| = 1$ precisely on $S = \{0, 1/3\}$, by Theorem 7.3.1a, the minimal extrapolations are supported in S .

The matrix E from (7.5) is

$$E(-1, 0, 1; 0, 1/3) = \begin{pmatrix} 1 & e^{-2\pi i/3} \\ 1 & 1 \\ 1 & e^{2\pi i/3} \end{pmatrix}.$$

Since E has full rank, then by Proposition 7.4.4, μ is the unique minimal extrapolation of F on Λ . Thus, reconstruction of μ from F on Λ is possible.

The following example illustrates that if μ is a sum of two Dirac measures, then their supports have to be sufficiently spaced apart in order for super-resolution of μ to be possible. This shows that, without any assumptions on the coefficients of the discrete measure, a minimum separation condition is necessary to super-resolve a sum of two Dirac measures, such as in [30].

Example 7.6.6. Let $\Lambda \subseteq \mathbb{Z}^d$ be any finite subset and let $F_y = \widehat{\mu}_y|_{\Lambda}$, where $\mu_y = \delta_0 - \delta_y$ for some non-zero $y \in \mathbb{T}^d$. We claim that if y is sufficiently small depending on Λ , then μ_y is not a minimal extrapolation of F_y on Λ . Note that $\|\mu_y\| = 2$ for any $y \in \mathbb{T}^d$. Let η denote the normalized Lebesgue measure on \mathbb{T}^d , and define the measures ν_y by the formula

$$\nu_y(x) = \sum_{m \in \Lambda} F_y(m) e^{2\pi i m \cdot x} \eta(x).$$

By construction, ν_y is an extrapolation of F_y on Λ because, for each $n \in \Lambda$,

$$\widehat{\nu}_y(n) = \int_{\mathbb{T}^d} e^{-2\pi i n \cdot x} d\nu_y(x) = \sum_{m \in \Lambda} F_y(m) \int_{\mathbb{T}^d} e^{-2\pi i(n-m) \cdot x} dx = F_y(n).$$

We let D_Λ be the generalized Dirichlet kernel, defined by the formula

$$D_\Lambda(x) = \sum_{m \in \Lambda} e^{2\pi i m \cdot x}.$$

Then we have,

$$\|\nu_y\| = \int_{\mathbb{T}^d} \left| \sum_{m \in \Lambda} F_y(m) e^{2\pi i m \cdot x} \right| dx = \int_{\mathbb{T}^d} |D_\Lambda(x) - D_\Lambda(x - y)| dx.$$

By a fundamental theorem of calculus argument and Bernstein's inequality for trigonometric polynomials, we have the upper bound

$$\begin{aligned} |D_\Lambda(x) - D_\Lambda(x - y)| &\leq |y| \|\nabla D_\Lambda\|_\infty \\ &\leq 2\pi\sqrt{d}|y| \max_{m \in \Lambda} |m| \|D_\Lambda\|_\infty \\ &= 2\pi\sqrt{d}|y| \max_{m \in \Lambda} |m| (\#\Lambda), \end{aligned}$$

where $|\cdot|$ denotes the Euclidean norm and ∇ is the gradient. Consequently, if $|y|$ is small enough so that

$$|y| < \frac{1}{\pi\sqrt{d} \max_{m \in \Lambda} |m| (\#\Lambda)},$$

then $\|\nu_y\| < 2 = \|\mu_y\|$. In this case, μ_y is not a minimal extrapolation of F_y on Λ . Note that this argument does not contradict Proposition 7.4.1 because $\|F_y\|_{\ell^\infty(\Lambda)} \rightarrow 0$ as $y \rightarrow 0$. Thus, for y sufficiently small, reconstruction of μ_y from F on Λ using (TV) is impossible.

Remark 7.6.7. The authors of [60, Corollary 1] showed that when $\Lambda = \{-M, -M+1, \dots, M\}$, there exists a real measure μ with minimum separation $1/(2M)$ that cannot be recovered using (TV), given its Fourier coefficients on Λ . For this particular case, their result is sharper than the one in Example 7.6.6 because they used a

special fact about trigonometric polynomials due to Turán [129]. In contrast, our result does not require any assumptions on Λ or whether μ is real or complex, and it holds in all dimensions.

Example 7.6.8. Let $\Lambda = \{-1, 0, 1\}^2 \setminus \{(1, -1), (-1, 1)\}$ and $F = \widehat{\mu} |_{\Lambda}$, where $\mu = \delta_{(0,0)} + \delta_{(1/2,1/2)}$. Then, $F(m) = 1 + e^{-\pi i(m_1+m_2)}$, and, in particular, $F(1, 1) = F(-1, -1) = F(0, 0) = 2$ and $F(\pm 1, 0) = F(0, \pm 1) = 0$. We deduce that $\varepsilon = \|\mu\|_{\text{TV}} = \|F\|_{\ell^\infty(\Lambda)} = 2$ from Proposition 7.4.1, and so μ is a minimal extrapolation of F on Λ .

Further, $\Gamma = \{(0, 0), (1, 1), (-1, -1)\}$, and so $\#\Gamma = 3$. According to the definition of $\alpha_{m,n}$ in Theorem 7.3.1b, set $\alpha_{(-1,-1),(0,0)} = \alpha_{(0,0),(1,1)} = \alpha_{(-1,-1),(1,1)} = 0$. By the conclusion of Theorem 7.3.1b, the minimal extrapolations are supported in the set $S = S_{(-1,-1),(0,0)} \cap S_{(0,0),(1,1)} \cap S_{(-1,-1),(1,1)}$, where

$$\begin{aligned} S_{(-1,-1),(0,0)} &= \{x \in \mathbb{T}^2 : x \cdot (-1, -1) \in \mathbb{Z}\} = \{x \in \mathbb{T}^2 : x_1 + x_2 \in \mathbb{Z}\}, \\ S_{(0,0),(1,1)} &= \{x \in \mathbb{T}^2 : x \cdot (-1, -1) \in \mathbb{Z}\} = \{x \in \mathbb{T}^2 : x_1 + x_2 \in \mathbb{Z}\}, \\ S_{(-1,-1),(1,1)} &= \{x \in \mathbb{T}^2 : x \cdot (-2, -2) \in \mathbb{Z}\} = \{x \in \mathbb{T}^2 : 2x_1 + 2x_2 \in \mathbb{Z}\}. \end{aligned}$$

It follows that the minimal extrapolations are supported in

$$S = S_{(-1,-1),(0,0)} \cap S_{(0,0),(1,1)} \cap S_{(-1,-1),(1,1)} = \{x \in \mathbb{T}^2 : x_1 + x_2 = 1\}.$$

We can construct other discrete minimal extrapolations besides μ . For each $y \in \mathbb{T}$ and for each integer $K \geq 2$, define the measure

$$\nu_{y,K} = \frac{2}{K} \sum_{k=0}^{K-1} \delta_{\left(y + \frac{k}{K}, 1 - y - \frac{k}{K}\right)}.$$

We claim $\nu_{y,K}$ is a minimal extrapolation. We have $\|\nu_{y,K}\| = \varepsilon$, and

$$\widehat{\nu_{y,K}}(m) = e^{-2\pi i(m_1 y - m_2 y)} e^{-2\pi i m_2} \frac{2}{K} \sum_{k=0}^{K-1} e^{-2\pi i(m_1 - m_2)k/K}.$$

We see that $\widehat{\nu_{y,K}} = F$ on Λ , which proves the claim.

We can also construct continuous singular minimal extrapolations. Let $\sigma = \sqrt{2}\sigma_S$, where σ_S is the *surface measure* of the Borel set S . We readily verify that $\|\sigma\| = \varepsilon$ and

$$\widehat{\sigma}(m) = \sqrt{2} \int_{\mathbb{T}^2} e^{-2\pi i m \cdot x} d\sigma_S = 2e^{-2\pi i m_2} \int_0^1 e^{-2\pi i(m_1 - m_2)t} dt = 2\delta_{m_1, m_2},$$

which proves that σ is a minimal extrapolation of F on Λ . In particular, S is the smallest set that contains the support of all the minimal extrapolations.

Since μ is not the unique minimal extrapolation, reconstruction of μ from F on Λ using (TV) is impossible.

Example 7.6.9. For an integer $q \geq 3$, let C_q be the *middle $1/q$ -Cantor set*, which is defined by $C_q = \bigcap_{k=0}^{\infty} C_{q,k}$, where $C_{q,0} = [0, 1]$ and

$$C_{q,k+1} = \frac{C_{q,k}}{q} \cup \left((1 - q) + \frac{C_{q,k}}{q} \right).$$

Let $F_q: [0, 1] \rightarrow [0, 1]$ be the *Cantor-Lebesgue function* on C_q , which is defined by the point-wise limit of the sequence $\{F_{q,k}\}$, where $F_{q,0}(x) = x$ and

$$F_{q,k+1}(x) = \begin{cases} \frac{1}{2}F_{q,k}(qx) & 0 \leq x \leq \frac{1}{q}, \\ \frac{1}{2} & \frac{1}{q} \leq x \leq \frac{q-1}{q}, \\ \frac{1}{2}F_{q,k}(qx - (q-1)) + \frac{1}{2} & \frac{q-1}{q} \leq x \leq 1. \end{cases}$$

By construction, $F_q(0) = 0$, $F_q(1) = 1$, and F_q is non-decreasing and uniformly continuous on $[0, 1]$. Thus, F_q can be uniquely identified with the measure $\sigma_q \in M(\mathbb{T})$, and $\|\sigma_q\| = 1$. Since $F'_q = 0$ a.e. and F_q does not have any jump discontinuities, σ_q is a continuous singular measure, with zero discrete part. The Fourier coefficients of σ_q are

$$\widehat{\sigma}_q(m) = (-1)^m \prod_{k=1}^{\infty} \cos(\pi m q^{-k}(1-q)),$$

see [137, pages 195-196]. In particular, for any integer $n \geq 1$, we have

$$\widehat{\sigma}_q(q^n) = (-1)^{q^n} \prod_{k=1}^{\infty} \cos(\pi q^{-k}(1-q)),$$

which is convergent and independent of n .

Let $\Lambda \subseteq \mathbb{Z}$ be any set containing 0, and let $F = \widehat{\sigma}_q|_{\Lambda}$. Since $F(0) = \|\sigma_q\| = 1$, we immediately see that $\varepsilon = 1$ and σ_q is a minimal extrapolation of F on Λ . Again, we cannot determine whether σ_q is the unique minimal extrapolation because Theorem 7.3.1 cannot handle the case $\#\Gamma = 1$, see Remark 7.4.7. For related examples that can be analyzed in terms of minimal extrapolations, see [85, 122, 89].

Example 7.6.10. Let $\sigma_A, \sigma_B \in M(\mathbb{T}^d)$ be the surface measures of the Borel sets $A = \{x \in \mathbb{T}^2: x_2 = 0\}$ and $B = \{x \in \mathbb{T}^2: x_2 = 1/2\}$, respectively. Let $\Lambda = \{-2, -1, \dots, 1\}^2$ and $F = \widehat{\mu}|_{\Lambda}$, where $\mu = \sigma_A + \sigma_B$. Then,

$$F(m) = \int_0^1 e^{2\pi i m_1 t} dt + \int_0^1 e^{2\pi i(m_1 t + m_2/2)} dt = \delta_{m_1, 0} + (-1)^{m_2} \delta_{m_1, 0}.$$

We immediately see that $\varepsilon = \|F\|_{\ell^\infty(\Lambda)} = \|\mu\|_{\text{TV}} = 2$, which implies μ is a minimal extrapolation of F on Λ . Then, $\Gamma = \{(0, 0), (0, 2), (0, -2)\}$, and, by Theorem 7.3.1b, the minimal extrapolations are supported in $\{x \in \mathbb{T}^2: x_2 = 0\} \cup \{x \in \mathbb{T}^2: x_2 = 1/2\}$.

Determining whether μ is the unique minimal extrapolation is beyond the theory we have developed herein, and we shall examine this uniqueness problem in [15].

7.7 Conclusion

The following topics encapsulate the contributions of this paper.

- (a) Qualitative behavior of minimal extrapolations. It is interesting to note that Γ provides significant information about the minimal extrapolations. Theorem 7.3.1 shows that, regardless of the dimension, when $\#\Gamma = 0$ or $\#\Gamma \geq 2$, the minimal extrapolations are always singular measures. The $\#\Gamma = 1$ case is pathological. Proposition 7.4.5 shows that this scenario is connected with the existence of positive absolutely continuous minimal extrapolations. Example 7.6.2 shows that there can exist both *uncountably* many discrete, as well as positive absolutely continuous, minimal extrapolations. Hence, this behavior is a fundamental feature of (TV) and is not an artifact of our analysis.
- (b) Computational consequences. One important aspect of Theorem 7.3.1 is its relationship with Algorithm 6.2.2 and similar variations. Proposition 7.4.3 shows that the algorithm fails precisely when $\varepsilon = \|F\|_{\ell^\infty(\Lambda)}$ and $\Gamma \neq \emptyset$. Since we are given the values of F on Λ , this immediately tells us what ε and Γ are. If $\#\Gamma \geq 2$, then the theorem is applicable and the minimal extrapolations are singular measures supported in the set defined in (7.1), which can be explicitly computed. We show how to apply our theorem to compute pertinent analytical examples. Hence, our main result is applicable even when Algorithm 6.2.2 and

similar variations fail.

- (c) Super-resolution of singular measures. When $d \geq 2$, Theorem 7.3.1 suggests that certain singular continuous measures are solutions to (TV). Somewhat surprisingly, this is indeed the case because Example 7.6.8 provides a singular continuous minimal extrapolation when $\#\Gamma \geq 2$. This also demonstrates that the conclusion of Theorem 7.3.1b is optimal. Proposition 7.4.9 shows that surface measures corresponding to the zero set of trigonometric polynomials are also minimal extrapolations.
- (d) Impossibility of super-resolution. Theorem 7.3.1 does not require structural assumptions on the finite subset $\Lambda \subseteq \mathbb{Z}^d$ or on the measure $\mu \in M(\mathbb{T}^d)$ that generates the spectral data F on Λ . Since the theorem also describes the support set of the minimal extrapolations, it is useful for determining whether it is possible for a given μ to be a minimal extrapolation. To illustrate this point, in Example 7.6.6, we provide a simple proof that shows, in general, a minimal separation condition is necessary to super-resolve a discrete measure.
- (e) Discrete-continuous correspondence. The total variation minimization problem can be viewed as a continuous analogue of the basis pursuit algorithm (BP). Indeed, let $\Phi \in \mathbb{C}^{N \times N}$ be the restriction of the $N \times N$ DFT matrix to the set $\Lambda \subseteq \{0, 1, \dots, N-1\}$, and let $y = \Phi x$. In this case, the basis pursuit algorithm is

$$\min_{\tilde{x} \in \mathbb{C}^N} \|\tilde{x}\|_{\ell^1} \quad \text{such that} \quad y = \Phi \tilde{x}.$$

By considering the case that $\mu = \sum_{n=1}^N x_n \delta_{n/N}$, it is straightforward to see that

(TV) is a generalization of (BP). From this point of view, Theorem 7.3.1 is a discrete-continuous correspondence result. Indeed, if $\#\Gamma \geq 2$, and either $d = 1$ or the vectors $\{p_j\}$ exist, then the minimal extrapolations are necessarily discrete measures whose support lie on a lattice; as we just discussed, such measures can be identified with vectors that are solutions to the discrete problem.

(f) Mathematical connections. Our results are closely related to Beurling's work on minimal extrapolation, see Theorem 7.3.3 Proposition 7.4.1 connects our results to the Candès and Fernandez-Granda theory [30] by introducing a concept called an admissibility range for ε . This connection is exploited to prove Proposition 7.4.3, which in turn, shows that our theorem is applicable to situations where Algorithm 6.2.2 fails.

Chapter 8: Stable Super-Resolution Limit

The material in this chapter originated from the paper [101], which was written by the author and Wenjing Liao.

8.1 Problem Statement

Let N be a positive integer and consider the discrete measure,

$$\mu = \sum_{n=0}^{N-1} x_n \delta_{n/N},$$

which is assumed to be supported on the grid of width $1/N$. We assume that x belongs to \mathbb{C}_S^N , the set of S -sparse vectors of length N . Suppose we are given the measurement vector $y \in \mathbb{C}^M$ with

$$y_m = \widehat{\mu}(m) + \eta_m, \quad \text{for } m = 0, 1, \dots, M-1,$$

where $\eta \in \mathbb{C}^M$ is some unknown additive noise that satisfies $\|\eta\|_2 \leq \delta$. We can write the measurement vector y as the linear system,

$$y = \Phi x + \eta,$$

where $\Phi \in \mathbb{C}^{M \times N}$ is the first M rows of the un-normalized $N \times N$ discrete Fourier transform (DFT) matrix:

$$\Phi_{m,n} = e^{-2\pi i m n / N}, \quad m = 0, 1, \dots, M-1, \quad n = 0, 1, \dots, N-1.$$

We want to recover μ , or equivalently x , from the measurements y in the super-resolution regime where $N \gg M$. In many applications, there are only a few point sources for which we need to recover at a high precision. For this reason, we can assume that the coefficient vector x is S -sparse, where S is much smaller than M . The noisy measurements are taken up to frequency M , so the Rayleigh length is $1/M$. The *super-resolution factor* is N/M , which is the number of grid points in one Rayleigh length. The challenge of recovering x increases as the super-resolution factor grows.

Subspace methods can recover μ if the noise is sufficiently small. Rather than analyze the performance of these algorithms, we take a more abstract approach, and study the theoretical super-resolution limit of all possible algorithms. The min-max error quantifies this approximation quality and can be used to evaluate the effectiveness of any algorithm, as the optimal method should be comparable to the min-max error.

Definition 8.1.1. Fix positive integers M, N, S such that $S \leq M \leq N$ and let $\delta > 0$. The S -min-max error is

$$E(M, N, S, \delta) = \inf_{\substack{\tilde{x}=A(y,M,N,S,\delta) \in \mathbb{C}^N \\ y=\Phi x+\eta}} \sup_{x \in \mathbb{C}_S^N} \sup_{\eta \in \mathbb{C}^M: \|\eta\|_2 \leq \delta} \|\tilde{x} - x\|_2.$$

Here, the infimum is taken over all functions A that takes the known information M, N, S, δ and the given data $y = \Phi x + \eta$ as inputs. In particular, the function A cannot depend on the unknown x and η .

The min-max error is a strong way of quantifying the error because the supremum is taken over all possible S -sparse vectors and noise bounded by δ . On the

other hand, this quantity does not assume that A is a particular algorithm, since the infimum is taken over all possible methods. Hence, no algorithm can achieve a min-max error smaller than $E(M, N, S, \delta)$.

To study the min-max error, we exploit its connection with the smallest singular value of certain restricted Fourier matrices.

Definition 8.1.2. Fix positive integers M, N, S such that $S \leq M \leq N$. The lower restricted isometry constant of order S is the quantity

$$\Theta(M, N, S) = \min_{|T|=S} \sigma_{\min}(\Phi_T),$$

where Φ_T is the restriction of Φ to columns indexed by the set $T \subseteq \{0, 1, \dots, N-1\}$ and $\sigma_{\min}(\Phi_T)$ is the smallest singular value of Φ_T .

The lower restricted isometry constant is related to the restricted isometry constant from compressive sensing [27], but with a major difference. The recovery guarantees from standard compressive sensing theory require Φ_T to be well-conditioned for all support sets T , whereas in our situation, the sub-matrices Φ_T can be highly ill-conditioned. There is an almost characterization of the min-max error in terms of the lower restricted isometry constant.

The main purpose of this chapter is to derive a lower bound on the lower restricted isometry constant, and we show this is sharp, up to constants depending on S , by considering an explicit example. This immediately provides a sharp upper bound on the min-max error, and has consequences to the stability of subspace methods.

8.2 Characterizations

We detail the relationship between the min-max error and the lower restricted isometry constant.

Proposition 8.2.1. *Fix positive integers M, N, S such that $2S \leq M \leq N$, and let $\delta > 0$. Then,*

$$\frac{\delta}{2\Theta(M, N, 2S)} \leq E(M, N, S, \delta) \leq \frac{2\delta}{\Theta(M, N, 2S)}.$$

Proof. We prove the upper bound first. Fix $x \in \mathbb{C}_S^N$, $\eta \in \mathbb{C}^M$ with $\|\eta\|_2 \leq \delta$, and suppose we are given $y = \Phi x + \eta$. Let $\tilde{x} \in \mathbb{C}^N$ be the sparsest vector such that $\|\Phi \tilde{x} - y\|_2 \leq \delta$, and this choice of \tilde{x} is independent of x and η . The vector \tilde{x} exists because x also satisfies this inequality constraint. By definition, we have

$$E(M, N, S, \delta) \leq \|\tilde{x} - x\|_2.$$

Since x is chosen as sparse as possible, we have $\|\tilde{x}\|_0 \leq \|x\|_0 \leq S$, which implies $\|\tilde{x} - x\|_0 \leq 2S$. By definition of the smallest singular value,

$$\frac{1}{\Theta(M, N, 2S)} = \sup_{|T|=2S} \frac{1}{\sigma_{\min}(\Phi_T)} \geq \frac{\|\tilde{x} - x\|_2}{\|\Phi(\tilde{x} - x)\|_2}.$$

By construction, we have

$$\|\Phi(\tilde{x} - x)\|_2 \leq \|\Phi \tilde{x} - y\|_2 + \|\Phi x - y\|_2 \leq 2\delta.$$

Combining previous inequalities completes the proof of the upper bound.

We focus our attention on the lower bound. By definition of the smallest singular value, there exists $u \in \mathbb{C}_{2S}^N$ of unit norm such that

$$\Theta(M, N, 2S) = \min_{|T|=2S} \sigma_{\min}(\Phi_T) = \|\Phi u\|_2.$$

Pick any vectors $v_1, v_2 \in \mathbb{C}_S^N$ such that

$$\frac{\delta}{\Theta(M, N, 2S)} u = v_1 - v_2.$$

Suppose we are given the data

$$y = \Phi v_1 = \Phi v_2 + \Phi(v_1 - v_2).$$

The previous three equations imply

$$\|\Phi v_2 - y\|_2 = \|\Phi(v_1 - v_2)\|_2 = \frac{\delta}{\Theta(M, N, 2S)} \|\Phi u\|_2 \leq \delta.$$

This proves that y is both the noiseless partial Fourier samples of v_1 and the noisy partial Fourier samples of v_2 with noise $\Phi(v_1 - v_2)$ of noise level at most δ . Recalling that u is of unit norm, for all functions A , let $\tilde{x} = A(y, M, N, S, \delta)$ and

$$\frac{\delta}{\Theta(M, N, 2S)} = \|v_1 - v_2\|_2 \leq 2 \max_{k=1,2} \|\tilde{x} - v_k\|_2 \leq 2E(M, N, S, \delta).$$

This completes the proof of the lower bound. □

Proposition 8.2.1 shows that it suffices to obtain tight bounds on $\Theta(M, N, S, \delta)$.

To do this, we prove a powerful duality relationship between the smallest singular value of Φ_T and an interpolation of measures supported on T by trigonometric polynomials.

Definition 8.2.2. Fix positive integers M, N, S such that $S \leq M \leq N$. For any $v \in \mathbb{C}_S^N$, let $\mathcal{P}(M, v)$ denote the set of trigonometric polynomials such that $f \in \mathcal{P}(M, v)$ if and only if $\text{supp}(\widehat{f}) \subseteq \{0, 1, \dots, M-1\}$ and $f(n/N) = v_n$ for each $n \in \text{supp}(v)$.

Proposition 8.2.3. Fix positive integers M, N, S such that $S \leq M \leq N$. For any support set $T \subseteq \{0, 1, \dots, N-1\}$ with cardinality S , we have

$$\frac{1}{\sigma_{\min}(\Phi_T)} = \sup_{\substack{\|v\|_2=1 \\ \text{supp}(v) \subseteq T}} \inf_{f \in \mathcal{P}(M, v)} \|f\|_{L^2}.$$

Proof. Fix a unit norm vector $v \in \mathbb{C}^N$ supported in T . The set of all trigonometric polynomials f with Fourier transform supported in $\{0, 1, \dots, M-1\}$ can be written in the form

$$f(t) = \sum_{m=0}^{M-1} \widehat{f}(m) e^{2\pi i m t}.$$

Let $u \in \mathbb{C}^M$ be the vector with $u_m = \widehat{f}(m)$ and note that $\|u\|_2 = \|f\|_{L^2}$. This establishes a bijection between each such f and a vector $u \in \mathbb{C}^M$. Thus, there exists a bijection between the set $\mathcal{P}(M, v)$ and the set of vectors $u \in \mathbb{C}^M$ such that

$$v_n = \sum_{m=0}^{M-1} u_m e^{2\pi i m n / N} \quad \text{for all } n \in T.$$

This condition is equivalent to the existence of a solution $u \in \mathbb{C}^M$ to the under-determined system of equations $(\Phi_T)^* u = v_T$, where $v_T \in \mathbb{C}^S$ denotes the restriction of v to T . Thus, we have

$$\inf_{f \in \mathcal{P}(M, v)} \|f\|_{L^2} = \inf_{(\Phi_T)^* u = v_T} \|u\|_2.$$

The minimum Euclidean norm solution to an under-determined system is given by the Moore-Penrose pseudoinverse, so

$$\inf_{(\Phi_T)^* u = v_T} \|u\|_2 = \|((\Phi_T)^*)^\dagger v_T\|_2.$$

To complete the proof, we note that

$$\sup_{\substack{\|v\|_2=1 \\ \text{supp}(v) \subseteq T}} \|((\Phi_T)^*)^\dagger v_T\|_2 = \frac{1}{\sigma_{\min}(\Phi_T)}.$$

□

8.3 Sparse Lagrange Polynomials

In view of Proposition 8.2.3, in order to obtain a lower bound on Θ , it suffices to construct, for each unit norm $v \in \mathbb{C}_S^N$, a particular interpolating polynomial $f(v) \in \mathcal{P}(M, v)$ and then upper bound $\|f(v)\|_{L^2}$ uniformly in v .

A natural first attempt is to use the classical Lagrange interpolating polynomial to carry out the interpolation, but we shall establish in the following proposition, that it yields a suboptimal bound. This might not be surprising since the Lagrange polynomials are known to exhibit poor behavior when the nodes consist of closely spaced points.

Definition 8.3.1. Fix positive integers S, N such that $S \leq N$. For any support set $T \subseteq \{0, 1, \dots, N-1\}$ with cardinality S , the Lagrange polynomials adapted to T is the set $\{L_n\}_{n \in T}$, where

$$L_n(t) = \prod_{k \in T \setminus \{n\}} \frac{e^{2\pi i t} - e^{2\pi i k/N}}{e^{2\pi i n/N} - e^{2\pi i k/N}}.$$

For any vector $v \in \mathbb{C}^N$ supported in T , we call

$$L(v) = \sum_{n \in T} v_n L_n$$

the Lagrange interpolating polynomial of v .

Proposition 8.3.2. Fix positive integers N, S such that $S \leq N$. For any unit norm $v \in \mathbb{C}_S^N$, we have

$$\|L(v)\|_{L^2} \leq \sqrt{S} \left(\frac{N}{2}\right)^{S-1}$$

Proof. Let $T = \text{supp}(v)$ and $\{L_n\}$ be the Lagrange polynomials adapted to T . By triangle inequality and Cauchy-Schwarz, we have

$$\|L(v)\|_{L^2} \leq \sum_{n \in T} |v_n| \|L_n\|_{L^2} \leq \left(\sum_{n \in T} \|L_n\|_{L^2}^2 \right)^{1/2}. \quad (8.1)$$

For distinct $n, k \in T$, we have

$$|e^{2\pi i n/N} - e^{2\pi i k/N}| \geq 4 \left| \frac{n}{N} - \frac{k}{N} \right|_{\mathbb{T}} \geq \frac{4}{N}.$$

Then, for all $t \in \mathbb{T}$ and $n \in T$,

$$|L_n(t)| \leq \prod_{k \in T \setminus \{n\}} \frac{2}{|e^{2\pi i n/N} - e^{2\pi i k/N}|} \leq \left(\frac{N}{2} \right)^{S-1}.$$

In particular, this implies

$$\|L_n\|_{L^2} \leq \left(\frac{N}{2} \right)^{S-1}. \quad (8.2)$$

Combining this with inequality (8.1) completes the proof. \square

Observe that in the proof of this proposition, since v is arbitrary, we cannot expect to do significantly better than inequality (8.1), and consequently, $\|L_n\|_{L^2}$ is the crucial quantity to estimate. The next proposition shows that, without additional assumptions on the support set T , our estimate (8.2) is optimal up to constants depending only on S . This suggests that it is impossible to obtain the optimal lower bound for $\sigma_{\min}(\Phi_T)$ using the Lagrange polynomials.

Proposition 8.3.3. *Let S be a positive integer and $T = \{0, 1, \dots, S-1\}$. For any positive function h such that $\lim_{t \rightarrow \infty} h(t) = 0$ and any $C(S) > 0$, there exists sufficiently large positive integers M, N with $M \leq N$ such that if $\{L_n\}_{n \in T}$ is the Lagrange polynomials adapted to T , then*

$$\|L_n\|_{L^2} > C(S)h(M)N^{S-1}.$$

Proof. Assume that $N \geq 8(S - 1)$, so that $n/N \subseteq [0, 1/8]$ for all $n \in T$. For all distinct $n, k \in T$, we have

$$|e^{2\pi in/N} - e^{2\pi ik/N}| \leq \frac{2\pi S}{N}.$$

This implies, for all $t \in \mathbb{T}$ and $n \in T$,

$$|L_n(t)| \geq \left(\frac{N}{2\pi S}\right)^{S-1} \prod_{k \in T \setminus \{n\}} |e^{2\pi it} - e^{2\pi ik/N}|.$$

If $t \in \mathbb{T}$ is far away from T , say $t \in [4/8, 5/8]$, then

$$|e^{2\pi it} - e^{2\pi ik/N}| \geq \frac{3}{2}.$$

This implies,

$$\begin{aligned} \|L_n\|_{L^2} &\geq \left(\frac{N}{2\pi S}\right)^{S-1} \left(\int_{4/8}^{5/8} \prod_{k \in T \setminus \{n\}} |e^{2\pi it} - e^{2\pi ik/N}|^2 dt\right)^{1/2} \\ &\geq \frac{1}{\sqrt{8}} \left(\frac{3N}{4\pi S}\right)^{S-1}. \end{aligned}$$

Taking M and N sufficiently large depending on h and $C(S)$ completes the proof. \square

We shall construct interpolating polynomials that have smaller norms. As seen in Proposition 8.3.3, the challenging case is when the support set T contains a substantial number of points that are close to each other. The main observation that we make in the following construction is that the Lagrange polynomials have Fourier transform supported in $\{0, 1, \dots, S-1\}$, whereas we have available the much larger spectral set $\{0, 1, \dots, M-1\}$. This extra flexibility allows us to construct better behaved interpolating polynomials.

Theorem 8.3.4. *Fix positive integers M, N, S such that $S \leq M$ and $MS \leq N$.*

For any support set $T \subseteq \{0, 1, \dots, N-1\}$ with cardinality S , there exists a set

of trigonometric polynomials $\{H_n\}_{n \in T}$ such that $H_n(k/N) = \delta_{n,k}$ for all $n, k \in T$, $\text{supp}(\widehat{H}_n) \subseteq \{0, 1, \dots, M-1\}$, and

$$\|H_n\|_{L^2} \leq CS^S \frac{1}{\sqrt{M}} \left(\frac{N}{M}\right)^{S-1}.$$

Proof. We define the auxiliary integer

$$P = \left\lfloor \frac{M}{S} \right\rfloor, \quad \text{and} \quad R = \left\lfloor \frac{P}{2} \right\rfloor.$$

Fix a support set T with cardinality S . For each $n \in T$, define the subsets

$$U_n = \left\{ k \in T : 0 < \left| \frac{n}{N} - \frac{k}{N} \right|_{\mathbb{T}} < \frac{M}{NS} \right\},$$

$$V_n = \left\{ k \in T : \left| \frac{n}{N} - \frac{k}{N} \right|_{\mathbb{T}} \geq \frac{M}{NS} \right\},$$

and the trigonometric polynomial

$$G_n(t) = \prod_{k \in U_n} \frac{e^{2\pi i P t} - e^{2\pi i P k / N}}{e^{2\pi i P n / N} - e^{2\pi i P k / N}} \prod_{k \in V_n} \frac{e^{2\pi i t} - e^{2\pi i k / N}}{e^{2\pi i n / N} - e^{2\pi i k / N}}.$$

We need to justify why G_n is well-defined. Clearly the terms over the product of $k \in V_n$ are non-zero. If $k \in U_n$, then

$$P \left| \frac{n}{N} - \frac{k}{N} \right|_{\mathbb{T}} < \frac{M}{S} \frac{1}{N} \leq 1,$$

which implies

$$\left| \frac{Pn}{N} - \frac{Pk}{N} \right|_{\mathbb{T}} = P \left| \frac{n}{N} - \frac{k}{N} \right|_{\mathbb{T}} \geq \frac{P}{N}.$$

By construction, we have $G_n(k/N) = \delta_{k,n}$ for all $n, k \in T$ and \widehat{G}_n is supported in

$\{0, 1, \dots, (S-1)P\}$. We proceed to estimate $\|G_n\|_{L^\infty}$. We have

$$\begin{aligned} \|G_n\|_{L^\infty} &\leq \prod_{k \in U_n} \frac{2}{|e^{2\pi i Pn/N} - e^{2\pi i Pk/N}|} \prod_{k \in V_n} \frac{2}{|e^{2\pi in/N} - e^{2\pi ik/N}|} \\ &\leq \prod_{k \in U_n} \frac{1}{2} \left| \frac{Pn}{N} - \frac{Pk}{N} \right|_{\mathbb{T}}^{-1} \prod_{k \in V_n} \frac{1}{2} \left| \frac{n}{N} - \frac{k}{N} \right|_{\mathbb{T}}^{-1} \\ &\leq \left(\frac{N}{2} \right)^{S-1} \left(\frac{1}{P} \right)^{\#U_n} \left(\frac{S}{M} \right)^{\#V_n}. \end{aligned}$$

While G_n satisfies the properties for interpolation, its norm is still too large as a factor of $1/\sqrt{M}$ is missing. We can get interpolating polynomials with smaller norms by multiplying by an appropriate Fejér kernel. Consider the following normalized Fejér kernel of order R , defined by the formula,

$$F_R(t) = \frac{1}{R} \sum_{m=-R}^R \left(1 - \frac{|m|}{R+1} \right) e^{2\pi imt}.$$

We note that $\widehat{F_R}$ is supported in the set $\{-R, -R+1, \dots, R\}$ and is normalized so that $F_R(0) = 1$. Its L^2 norm is approximately $C/R^{1/2}$, since by Parseval, we have

$$\begin{aligned} \|F_R\|_{L^2} &= \frac{1}{R(R+1)} \left((R+1)^2 + 2 \sum_{m=1}^R m^2 \right)^{1/2} \\ &= \frac{1}{R(R+1)} \left(\frac{2}{3}R^3 + 2R^2 + \frac{7}{3}R + 1 \right)^{1/2}. \end{aligned}$$

Finally, we are ready to define the desired polynomials. For each $n \in T$, let

$$H_n(t) = e^{2\pi i R(t-n/N)} F_R\left(t - \frac{n}{N}\right) G_n(t).$$

By construction of the G_n and F_R , we have the interpolation property $H_n(k/N) = \delta_{k,n}$, for all $n, k \in T$. Additionally, we see that $\widehat{H_n}$ is supported in

$$\{0, 1, \dots, 2R\} + \{0, 1, \dots, (S-1)P\} \subseteq \{0, 1, \dots, M-1\}.$$

Finally, to compute the L^2 norm of H_n , we use Hölder’s inequality,

$$\|H_n\|_{L^2} \leq \|F_R\|_{L^2} \|G_n\|_{L^\infty} \leq CS^S \frac{1}{\sqrt{M}} \left(\frac{N}{M}\right)^{S-1}.$$

□

Definition 8.3.5. Let M, N, S, T satisfy the assumptions in Theorem 8.3.4. We call $\{H_n\}_{n \in T}$ the sparse Lagrange polynomials adapted to T . For any $v \in \mathbb{C}^N$ supported in T , we call

$$H(v) = \sum_{n \in T} v_n H_n$$

the sparse Lagrange interpolating polynomial of v .

We call the function H_n “sparse” because for any M points on the torus and M values, it is possible to construct a trigonometric polynomial of degree $M - 1$ that interpolates those values at those points; in contrast, the sparse Lagrange polynomials have degree $M - 1$ and interpolate values at S points, where S can be much smaller than M . Of course, the main advantage of the sparse Lagrange polynomials is that they have significantly smaller norms compared to Lagrange polynomials, as seen in Theorem 8.3.4 and Figure 8.1.

8.4 Lower Bound using Polynomial Interpolation

The lower bound on the lower restricted isometry constant follows immediately from our machinery that we developed earlier.

Theorem 8.4.1. *Fix a positive integers S, M, N such that $S \leq M$ and $MS \leq N$.*

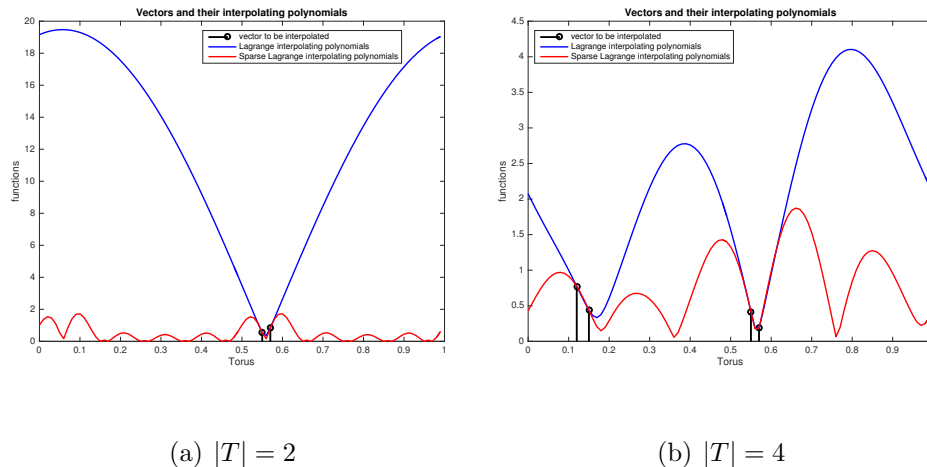


Figure 8.1: We set $M = 20$ and $N = 100$. These figures show two vectors supported on T where $|T| = 2$ in (a) and $|T| = 4$ in (b), the Lagrange interpolating polynomial L , and the sparse Lagrange interpolating polynomial H . The vectors and polynomials are complex-valued, so we display their magnitudes only. By construction, \widehat{L} is supported in $\{0, 1, \dots, |T| - 1\}$, while the \widehat{H} is supported in $\{0, 1, \dots, M - 1\}$. It is clear from these two examples that H has a smaller L^2 norm by incorporating the higher frequency components.

There exists a constant $A(S) > 0$ depending only on S such that

$$\Theta(M, N, S) \geq A(S) \sqrt{M} \left(\frac{M}{N}\right)^{S-1}.$$

Consequently, for all $\delta > 0$ and $2S \leq M$ and $2MS \leq N$, then

$$E(M, N, S, \delta) \leq \frac{2\delta}{A(2S)M} \left(\frac{N}{M}\right)^{2S-1}.$$

Proof. Fix a support set T with cardinality S and let $\{H_n\}_{n \in T}$ be the sparse Lagrange polynomials adapted to T . For any unit vector $v \in \mathbb{C}^N$ supported in T , we see that

$$H(t) = \sum_{n \in T} v_n H_n(t) \in \mathcal{P}(M, v).$$

By the duality principle, Proposition 8.2.3, we have

$$\frac{1}{\sigma_{\min}(\Phi_T)} \leq \sup_{\substack{\|v\|_2=1 \\ \text{supp}(v) \subseteq T}} \|H\|_{L^2} \leq \left(\sum_{n \in T} \|H_n\|_{L^2}^2 \right)^{1/2}.$$

By the upper bound on the sparse Lagrange polynomials, Theorem 8.3.4, we have

$$\frac{1}{\sigma_{\min}(\Phi_T)} \leq C S^{S+1/2} \frac{1}{\sqrt{M}} \left(\frac{N}{M}\right)^{S-1}.$$

This inequality holds for all support sets T with cardinality S , which completes the proof of the lower bound with

$$A(S) = \frac{C}{S^{S+1/2}}.$$

The second statement of the theorem follows immediately by applying Proposition 8.2.1. □

Remark 8.4.2. We have several comments about the numerology of the bounds.

1. The \sqrt{M} factor appears because the columns of Φ have norm \sqrt{M} . Had we had consider the normalized Fourier matrix Φ/\sqrt{M} instead, this factor in both the lower and upper bounds would disappear.
2. In view of the connection between $\Theta(M, N, S)$ and super-resolution in imaging, it makes sense from a physical point of view that this quantity only depends on the super-resolution factor N/M , and not on M or N individually.
3. The singular values of Φ_T and $\Phi_{\tilde{T}}$ are identical whenever $\tilde{T} = T + a \pmod{N}$ and $a \in \mathbb{Z}$, so without loss of generality, we can assume that $0 \in T$. Hence, even though the minimization in $\Theta(M, N, S)$ is taken over all subsets T with cardinality S , one of the elements in T is already fixed.

8.5 Upper Bound and Uncertainty Principles

An upper bound on the lower restricted isometry constant can be interpreted as an extremal case of the discrete uncertainty principle.

Theorem 8.5.1. *Fix a positive integer S . There exist constants $B(S), C(S) > 0$ such that if $M \geq S$ and $N \geq C(S)M^{3/2}$, then*

$$\Theta(M, N, S) \leq 2B(S)\sqrt{M}\left(\frac{M}{N}\right)^{S-1}.$$

Further, if $M \geq 2S$ and $N \geq C(2S)M^{3/2}$, then

$$E(M, N, S, \delta) \geq \frac{\delta}{4B(2S)}\frac{1}{\sqrt{M}}\left(\frac{N}{M}\right)^{2S-1}.$$

Proof. We first observe that

$$\Theta(M, N, S) = \min_{|T|=S} \sigma_{\min}(\Phi_T) = \min_{|T|=S} \inf_{\substack{u \neq 0 \\ \text{supp}(u) \subseteq T}} \frac{\|\Phi u\|_2}{\|u\|_2}.$$

The question then becomes, how small, in the ℓ^2 sense, can M consecutive discrete Fourier coefficients of a S -sparse N dimensional complex vector be? It suffices to consider a specific vector u supported in a set T of cardinality S . Let $T = \{0, 1, \dots, S-1\}$ and $u \in \mathbb{C}^N$ where

$$u_n = \begin{cases} (-1)^n \binom{S-1}{n} & \text{if } n \in T, \\ 0 & \text{otherwise.} \end{cases}$$

Then we have

$$\Theta(M, N, S) \leq \frac{\|\Phi u\|_2}{\|u\|_2} = \binom{2S-2}{S-1}^{-1/2} \|\Phi u\|_2.$$

To accurately calculate $\|\Phi u\|_2$, we identify u with the discrete measure

$$\mu_N = \sum_{n=0}^{N-1} u_n \delta_{n/N},$$

and recall the Dirichlet kernel of order M , defined as

$$D_M(t) = \sum_{m=0}^{M-1} e^{2\pi i m t}.$$

We have

$$\|\Phi u\|_2 = \left(\sum_{m=0}^{M-1} |\hat{\mu}_N(m)|^2 \right)^{1/2} = \|\mu_N * D_M\|_{L^2},$$

For all $t \in \mathbb{T}$, we have

$$(\mu_N * D_M)(t) = \sum_{n=0}^{S-1} (-1)^n \binom{S-1}{n} D_M\left(t - \frac{n}{N}\right).$$

The right hand side is the $(S - 1)$ -th order backwards finite difference of D_M . It is well-known that for each $t \in \mathbb{T}$, we have

$$N^{S-1}(\mu_N * D_M)(t) = D_M^{(S-1)}(t) + R_{S-1}(D_M)(t),$$

where the remainder term $|R_{S-1}(D_M)|$ is point-wise $O(1/N)$ as $N \rightarrow \infty$. In order to precisely determine how large we require M and N to be, we calculate the remainder term explicitly.

We first control the $S - 1$ derivative term. By the Bernstein inequality for trigonometric polynomials, we have

$$\|D_M^{(S-1)}\|_{L^2} \leq (\pi M)^{S-1} \|D_M\|_{L^2} \leq \pi^{S-1} M^{S-1/2}.$$

For the remainder term, by a Taylor expansion of D_M , for each $t \in \mathbb{T}$ and $0 \leq n \leq S - 1$, there exists $t_n \in (t - n/N, t)$ such that

$$D_M\left(t - \frac{n}{N}\right) = \sum_{k=0}^{S-1} \frac{D_M^{(k)}(t)}{k!} \left(\frac{n}{N}\right)^k (-1)^k + \frac{D_M^{(S)}(t_n)}{S!} \left(\frac{n}{N}\right)^S (-1)^S.$$

Hence, we see that

$$R_{S-1}(D_M)(t) = N^{S-1} \sum_{n=0}^{S-1} (-1)^{n+S} \binom{S-1}{n} \left(\frac{n}{N}\right)^S \frac{D_M^{(S)}(t_n)}{S!}.$$

We bound this term in the L^2 norm. By Bernstein's inequality,

$$\begin{aligned} \|R_{S-1}(D_M)\|_{L^2} &\leq \frac{1}{N} \frac{\|D_M^{(S)}\|_{L^\infty(\mathbb{T})}}{S!} \sum_{n=0}^{S-1} \binom{S-1}{n} n^S \\ &\leq \frac{1}{N} \frac{\pi^S M^{S+1}}{S!} \sum_{n=0}^{S-1} \binom{S-1}{n} n^S. \end{aligned}$$

Combining the previous inequalities, we deduce

$$\begin{aligned}
\Theta(M, N, S) &\leq \binom{2S-2}{S-1}^{-1/2} \|\mu_N * D_M\|_{L^2} \\
&\leq \binom{2S-2}{S-1}^{-1/2} \frac{1}{N^{S-1}} \left(\|D_M^{(S-1)}\|_{L^2} + \|R_{S-1}(D_M)\|_{L^2} \right) \\
&\leq B(S) \left(1 + \frac{M^{3/2}}{N} C(S) \right) \sqrt{M} \left(\frac{M}{N} \right)^{S-1},
\end{aligned}$$

where

$$B(S) = \binom{2S-2}{S-1}^{-1/2} \pi^{S-1}, \quad C(S) = \frac{\pi}{(S+1)!} \sum_{n=0}^{S-1} \binom{S-1}{n} n^S.$$

This completes the first statement of the theorem, and the second one follows immediately from Proposition [8.2.1](#). \square

8.6 Related Work

Among existing works, the papers by Donoho [\[55\]](#) and Demanet and Nguyen [\[53\]](#) are most closely related to this material. Both papers studied the min-max error for similar recovery problems, but with on a lattice subset of the *real line* \mathbb{R} and with *continuous* Fourier transform measurements. Consequently, their results do not apply to many popular algorithms such as MUSIC, Matrix pencil method, and TV-minimization, precisely because these methods operate under the assumption that the discrete measure lies on a *finite domain* (such as \mathbb{T}) and the known information consists of *discrete* Fourier transform coefficients. Moreover, their results hold asymptotically as the ratio of the number of measurements and lattice width shrinks to zero, but they did not show how small this ratio needs to be.

It is known that, *even in the noiseless setting*, without additional assumptions on the discrete measure's coefficients, total variation minimization fails if the separation is less than the Rayleigh length. For this reason, optimization approaches require additional assumptions on the measure in order to say something when the separation is below the Rayleigh length. The authors of [110, 54] require that all the amplitudes are *positive*. By connecting the TV-min problem with Beurling's theory of minimal extrapolation [17, 18], the authors [14] showed that if the sign of the measure satisfies certain algebraic properties, then it is a solution to the minimization problem.

A key quantity in this chapter, is the minimum singular value of rectangular Vandermonde matrices whose nodes are on the unit circle. Gautschi [73] gave an exact form of the inverse of a square Vandermonde matrix, but that does not yield bounds on the smallest singular value of rectangular Vandermonde matrices. When the separation is above the Rayleigh length, the author of [109] provided a sharp bound on the smallest singular value of rectangular Vandermonde matrices by using classical approximation theory results [130]. However, accurate estimation of the same quantity for when the nodes is below the Rayleigh length is much more difficult, because it is highly sensitive to the configuration of the support set. We conjecture that the lower bound presented in Theorem 8.4.1 is close to optimal when the support set T consists of closely spaced points, but it is definitely not optimal for other configurations.

Bibliography

- [1] Joakim Andén and Stéphane Mallat. Deep scattering spectrum. *IEEE Transactions on Signal Processing*, 62(16):4114–4128, 2014.
- [2] Sundaram Arivazhagan, L. Ganesan, and S. Priyal. Texture classification using Gabor wavelets based rotation invariant features. *Pattern Recognition Letters*, 27(16):1976–1982, 2006.
- [3] Enrico Au-Yeung and John J. Benedetto. Generalized Fourier frames in terms of balayage. *Journal of Fourier Analysis and Applications*, 21(3):472–508, 2015.
- [4] Céline Aubel, David Stotz, and Helmut Bölcskei. A theory of super-resolution from short-time fourier transform measurements. *Journal of Fourier Analysis and Applications*, pages 1–63, 2017.
- [5] Jean-Marc Azaïs, Yohann De Castro, and Fabrice Gamboa. Spike detection from inaccurate samplings. *Applied and Computational Harmonic Analysis*, 38(2):177–195, 2015.
- [6] John J. Benedetto. *Harmonic Analysis and Applications*. CRC Press Inc., 1996.
- [7] John J. Benedetto and Wojciech Czaja. *Integration and Modern Analysis*. Birkhäuser Boston, Inc., Boston, 2009.
- [8] John J. Benedetto, Wojciech Czaja, Przemysław Gadziński, and Alexander M. Powell. The Balian-Low theorem and regularity of Gabor systems. *The Journal of Geometric Analysis*, 13(2):239, 2003.
- [9] John J. Benedetto, Wojciech Czaja, and Andrei Ya Maltsev. The Balian–Low theorem for the symplectic form on \mathbb{R}^{2d} . *Journal of Mathematical Physics*, 44(4):1735–1750, 2003.

- [10] John J. Benedetto, Wojciech Czaja, Alexander M. Powell, and Jacob Stenbenz. An endpoint (1,) Balian-Low Theorem. *Mathematical Research Letters*, 13(2/3):467, 2006.
- [11] John J. Benedetto and Matthew Fickus. Finite normalized tight frames. *Advances in Computational Mathematics*, 18(2-4):357–385, 2003.
- [12] John J. Benedetto, Christopher Heil, and David F. Walnut. Gabor systems and the Balian-Low theorem. In *Gabor analysis and algorithms*, pages 85–122. Springer, 1998.
- [13] John J. Benedetto and Shidong Li. The theory of multiresolution analysis frames and applications to filter banks. *Applied and Computational Harmonic Analysis*, 5(4):389–427, 1998.
- [14] John J. Benedetto and Weilin Li. Super-resolution by means of Beurling minimal extrapolation. *arXiv preprint arXiv:1601.05761*, 2016.
- [15] John J. Benedetto and Weilin Li. Super-resolution of singular continuous measures on \mathbb{T}^d . *In Preparation*, 2018.
- [16] John J. Benedetto and David F. Walnut. Gabor frames for L_2 and related spaces. *Wavelets: Mathematics and Applications*, pages 97–162, 1993.
- [17] Arne Beurling. Balayage of Fourier-Stieltjes transforms. *The Collected Works of Arne Beurling*, 2:341–350, 1989.
- [18] Arne Beurling. Interpolation for an interval in \mathbb{R}^1 . *The Collected Works of Arne Beurling*, 2:351–365, 1989.
- [19] Badri Narayan Bhaskar, Gongguo Tang, and Benjamin Recht. Atomic norm denoising with applications to line spectral estimation. *IEEE Transactions on Signal Processing*, 61(23):5987–5999, 2013.
- [20] Edward H. Bosch, Wojciech Czaja, James M. Murphy, and Daniel Weinberg. Anisotropic representations for superresolution of hyperspectral data. *SPIE Defense+ Security*, pages 947213–947213, 2015.
- [21] Nicholas Boyd, Geoffrey Schiebinger, and Benjamin Recht. The alternating descent conditional gradient method for sparse inverse problems. *SIAM Journal on Optimization*, 27(2):616–639, 2017.
- [22] Richard P. Brent, Judy-Anne H. Osborn, and Warren D. Smith. Note on best possible bounds for determinants of matrices close to the identity matrix. *Linear Algebra and its Applications*, 466:21–26, 2015.
- [23] Joan Bruna and Stéphane Mallat. Invariant scattering convolution networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1872–1886, 2013.

- [24] Martin Burger and Stanley Osher. Convergence rates of convex variational regularization. *Inverse problems*, 20(5):1411, 2004.
- [25] Emmanuel Candes and Justin Romberg. Sparsity and incoherence in compressive sampling. *Inverse problems*, 23(3):969, 2007.
- [26] Emmanuel J. Candès. Harmonic analysis of neural networks. *Applied and Computational Harmonic Analysis*, 6(2):197–218, 1999.
- [27] Emmanuel J. Candès. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathématique*, 346(9-10):589–592, 2008.
- [28] Emmanuel J. Candès and David L. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Communications on Pure and Applied Mathematics*, 57(2):219–266, 2004.
- [29] Emmanuel J. Candès and Carlos Fernandez-Granda. Super-resolution from noisy data. *Journal of Fourier Analysis and Applications*, 19(6):1229–1254, 2013.
- [30] Emmanuel J. Candès and Carlos Fernandez-Granda. Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics*, 67(6):906–956, 2014.
- [31] Emmanuel J. Candès, Justin K. Romberg, and Terence Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, 2006.
- [32] Emmanuel J. Candes, Justin K. Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8):1207–1223, 2006.
- [33] Emmanuel J. Candès and Michael B. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25(2):21–30, 2008.
- [34] Venkat Chandrasekaran, Benjamin Recht, Pablo A. Parrilo, and Alan S. Willsky. The convex geometry of linear inverse problems. *Foundations of Computational Mathematics*, 12(6):805–849, 2012.
- [35] Scott S. Chen, David L. Donoho, and Michael A. Saunders. Atomic decomposition by basis pursuit. *SIAM Review*, 43(1):129–159, 2001.
- [36] Xiuyuan Cheng, Xu Chen, and Stéphane Mallat. Deep haar scattering networks. *Information and Inference: A Journal of the IMA*, 5(2):105–133, 2016.
- [37] Donald G. Childers. *Modern Spectrum Analysis*. IEEE Press, Piscataway, New Jersey, 1978.

- [38] Zoran Cvetkovic and Martin Vetterli. Oversampled filter banks. *IEEE Transactions on Signal Processing*, 46(5):1245–1255, 1998.
- [39] Wojciech Czaja and Emily J. King. Isotropic shearlet analogs for $L^2(\mathbb{R}^k)$ and localization operators. *Numerical Functional Analysis and Optimization*, 33(7-9):872–905, 2012.
- [40] Wojciech Czaja and Emily J. King. Anisotropic shearlet transforms for $L^2(\mathbb{R}^k)$. *Mathematische Nachrichten*, 287(8-9):903–916, 2014.
- [41] Wojciech Czaja and Weilin Li. Analysis of time-frequency scattering transforms. *Applied and Computational Harmonic Analysis*, 2017.
- [42] Wojciech Czaja and Weilin Li. Rotationally invariant time-frequency scattering transforms. *arXiv preprint arXiv:1710.06889*, 2017.
- [43] Wojciech Czaja, Benjamin Manning, James M. Murphy, and Kevin Stubbs. Discrete directional Gabor frames. *Applied and Computational Harmonic Analysis*, 2016.
- [44] Wojciech Czaja, James M Murphy, and Daniel Weinberg. Superresolution of remotely sensed images with anisotropic features. In *Sampling Theory and Applications (SampTA), 2015 International Conference on*, pages 317–321. IEEE, 2015.
- [45] Wojciech Czaja, James M. Murphy, and Daniel Weinberg. Single-image superresolution through directional representations. *arXiv preprint arXiv:1602.08575*, 2016.
- [46] Ingrid Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*, 36(5):961–1005, 1990.
- [47] Ingrid Daubechies. *Ten Lectures on Wavelets*, volume 61. SIAM, 1992.
- [48] Ingrid Daubechies, Alex Grossmann, and Yves Meyer. Painless nonorthogonal expansions. *Journal of Mathematical Physics*, 27(5):1271–1283, 1986.
- [49] John G. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America*, 2(7):1160–1169, 1985.
- [50] John G. Daugman. Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(7):1169–1179, 1988.
- [51] Yohann De Castro and Fabrice Gamboa. Exact reconstruction using Beurling minimal extrapolation. *Journal of Mathematical Analysis and Applications*, 395(1):336–354, 2012.

- [52] Yohann De Castro, Fabrice Gamboa, Didier Henrion, and J-B Lasserre. Exact solutions to super resolution on semi-algebraic domains in higher dimensions. *IEEE Transactions on Information Theory*, 63(1):621–630, 2017.
- [53] Laurent Demanet and Nam Nguyen. The recoverability limit for superresolution via sparsity. *arXiv preprint arXiv:1502.01385*, 2015.
- [54] Quentin Denoyelle, Vincent Duval, and Gabriel Peyré. Support recovery for sparse super-resolution of positive measures. *Journal of Fourier Analysis and Applications*, 23(5):1153–1194, 2017.
- [55] David L. Donoho. Superresolution via sparsity constraints. *SIAM Journal on Mathematical Analysis*, 23(5):1309–1331, 1992.
- [56] David L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.
- [57] Marco F. Duarte and Richard G. Baraniuk. Spectral compressive sensing. *Applied and Computational Harmonic Analysis*, 35(1):111–129, 2013.
- [58] Richard J. Duffin and Albert C. Schaeffer. A class of nonharmonic fourier series. *Transactions of the American Mathematical Society*, 72(2):341–366, 1952.
- [59] Bogdan Dumitrescu. *Positive Trigonometric Polynomials and Signal Processing Applications*. Springer, Dordrecht, 2007.
- [60] Vincent Duval and Gabriel Peyré. Exact support recovery for sparse spikes deconvolution. *Foundations of Computational Mathematics*, 15(5):1315–1355, 2015.
- [61] Michael Eickenberg, Georgios Exarchakis, Matthew Hirn, and Stephane Mallat. Solid harmonic wavelet scattering: Predicting quantum molecular energy from invariant descriptors of 3d electronic densities. In *Advances in Neural Information Processing Systems*, pages 6543–6552, 2017.
- [62] Michael Elad. Sparse and redundant representation modeling What next? *IEEE Signal Processing Letters*, 19(12):922–928, 2012.
- [63] Carl-Gustav Esseen. Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law. *Acta Mathematica*, 77(1):1–125, 1945.
- [64] Carl-Gustav Esseen. A note on Fourier-Stieltjes transforms and absolutely continuous functions. *Mathematics Scandinavica*, 2:153–157, 1954.
- [65] Albert Fannjiang and Wenjing Liao. Coherence-pattern guided compressive sensing with unresolved grids. *SIAM Journal on Imaging Sciences*, 5(1):179–202, 2012.

- [66] Albert Fannjiang, Thomas Strohmer, and Pengchong Yan. Compressed remote sensing of sparse objects. *SIAM Journal on Imaging Sciences*, 3(3):595–618, 2010.
- [67] Mitra Fatemi, Arash Amini, Loic Baboulaz, and Martin Vetterli. Shapes from pixels. *IEEE Transactions on Image Processing*, 25(3):1193–1206, 2016.
- [68] Carlos Fernandez-Granda. Support detection in super-resolution. In *Proceedings of the 10th International Conference on Sampling Theory and Applications*, pages 145–148, 2013.
- [69] Carlos Fernandez-Granda. Super-resolution of point sources via convex programming. *Information and Inference*, pages 1–53, 2016.
- [70] Matthew Fickus, Brody D. Johnson, Keri Kornelson, and Kasso A. Okoudjou. Convolutional frames and the frame potential. *Applied and Computational Harmonic Analysis*, 19(1):77–91, 2005.
- [71] Gerald B. Folland. *Harmonic Analysis in Phase Space*. Princeton University Press, 1986.
- [72] Dennis Gabor. Theory of communication. *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, 93(26):429–441, 1946.
- [73] Walter Gautschi. On inverses of Vandermonde and confluent Vandermonde matrices. *Numerische Mathematik*, 4(1):117–123, 1962.
- [74] Loukas Grafakos. *Classical fourier analysis*, volume 2. Springer, 2008.
- [75] Loukas Grafakos and Christopher Sansing. Gabor frames and directional time–frequency analysis. *Applied and Computational Harmonic Analysis*, 25(1):47–67, 2008.
- [76] Hayit Greenspan. Super-resolution in medical imaging. *The Computer Journal*, 52(1):43–63, 2009.
- [77] Karlheinz Gröchenig. *Foundations of time-frequency analysis*. Springer Science & Business Media, 2013.
- [78] Philipp Grohs, Sandra Keiper, Gitta Kutyniok, and Martin Schäfer. α -molecules. *Applied and Computational Harmonic Analysis*, 2015.
- [79] Kanghui Guo, Gitta Kutyniok, and Demetrio Labate. Sparse multidimensional representations using anisotropic dilation and shear operators. *Wavelets and Splines (Athens, GA, 2005)*, pages 189–201, 2006.
- [80] Kanghui Guo and Demetrio Labate. Optimally sparse multidimensional representation using shearlets. *SIAM Journal on Mathematical Analysis*, 39(1):298–318, 2007.

- [81] Kanghui Guo, Demetrio Labate, Wang-Q Lim, Guido Weiss, and Edward Wilson. Wavelets with composite dilations and their MRA properties. *Applied and Computational Harmonic Analysis*, 20(2):202–236, 2006.
- [82] Yoshihiko Hamamoto, Shunji Uchimura, Masanori Watanabe, Tetsuya Yasuda, Yoshihiro Mitani, and Shingo Tomita. A Gabor filter-based method for recognizing handwritten numerals. *Pattern Recognition*, 31(4):395–400, 1998.
- [83] Eugenio Hernández and Guido Weiss. *A First Course on Wavelets*. CRC Press, 1996.
- [84] Edwin Hewitt and Herman Rubin. The maximum value of a Fourier-Stieltjes transform. *Mathematica Scandinavica*, 3:97–102, 1955.
- [85] Einar Hille and J.D. Tamarkin. Remarks on a known example of a monotone continuous function. *American Mathematical Monthly*, 36(5):255–264, 1929.
- [86] Geoffrey Hinton, Li Deng, Dong Yu, George E. Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N. Sainath, et al. Deep neural networks for acoustic modeling in speech recognition. *IEEE Signal Processing Magazine*, 29(6):82–97, 2012.
- [87] Geoffrey E. Hinton, Simon Osindero, and Yee-Whye Teh. A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7):1527–1554, 2006.
- [88] Yingbo Hua and Tapan K. Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(5):814–824, 1990.
- [89] Jean-Pierre Kahane and Raphaël Salem. *Ensembles Parfaits et Séries Trigonométriques*. Hermann Paris, 1963.
- [90] Yitzhak Katznelson. *An Introduction to Harmonic Analysis*. Cambridge University Press, 2004.
- [91] Valery Khaidukov, Evgeny Landa, and Tijmen Jan Moser. Diffraction imaging by focusing-defocusing: An outlook on seismic superresolution. *Geophysics*, 69(6):1478–1490, 2004.
- [92] Wai Kin Kong, David Zhang, and Wenxin Li. Palmprint feature extraction using 2-D Gabor filters. *Pattern Recognition*, 36(10):2339–2347, 2003.
- [93] Hamid Krim and Viberg Mats. Two decades of array signal processing research: the parametric approach. *IEEE Signal Processing Magazine*, 13(4):67–94, 1996.
- [94] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, pages 1097–1105, 2012.

- [95] Gitta Kutyniok, Kasso A. Okoudjou, and Friedrich Philipp. Scalable frames and convex geometry. *Contemp. Math*, 626:19–32, 2014.
- [96] Demetrio Labate, Wang-Q Lim, Gitta Kutyniok, and Guido Weiss. Sparse multidimensional representation using shearlets. In *Optics & Photonics*, pages 59140U–59140U, 2005.
- [97] Henry J. Landau. Maximum entropy and the moment problem. *Bulletin of the American Mathematical Society*, 16(1):47–77, 1987.
- [98] Yann LeCun, B. Boser, John S. Denker, D. Henderson, Richard E. Howard, W. Hubbard, and Lawrence D. Jackel. Handwritten digit recognition with a back-propagation network. In *Advances in Neural Information Processing Systems*. Citeseer, 1990.
- [99] Honglak Lee, Chaitanya Ekanadham, and Andrew Y. Ng. Sparse deep belief net model for visual area V2. In *Advances in Neural Information Processing Systems*, pages 873–880, 2008.
- [100] Weilin Li. Personal website. <http://www.math.umd.edu/~wl298>, 2017.
- [101] Weilin Li and Wenjing Liao. Stable super-resolution limit and smallest singular value of restricted fourier matrices. *arXiv preprint arXiv:1709.03146*, 2017.
- [102] Wenjing Liao. Music for multidimensional spectral estimation: stability and super-resolution. *IEEE Transactions on Signal Processing*, 63(23):6395–6406, 2015.
- [103] Wenjing Liao and Albert Fannjiang. MUSIC for single-snapshot spectral estimation: Stability and super-resolution. *Applied and Computational Harmonic Analysis*, 40(1):33–67, 2016.
- [104] Jari Lindberg. Mathematical concepts of optical superresolution. *Journal of Optics*, 14(8):083001, 2012.
- [105] Stéphane Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1999.
- [106] Stéphane Mallat. Group invariant scattering. *Communications on Pure and Applied Mathematics*, 65(10):1331–1398, 2012.
- [107] Stéphane Mallat. Understanding deep convolutional networks. *Philosophical Transactions of the Royal Society A*, 374(2065), 2016.
- [108] Stéphane Mallat and et al. Scattering. <http://www.di.ens.fr/data/scattering>, 2017.
- [109] Ankur Moitra. Super-resolution, extremal functions and the condition number of vandermonde matrices. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 821–830. ACM, 2015.

- [110] Veniamin I. Morgenshtern and Emmanuel J. Candes. Super-resolution of positive sources: The discrete setup. *SIAM Journal on Imaging Sciences*, 9(1):412–444, 2016.
- [111] James M. Murphy. *Anisotropic Harmonic Analysis and Integration of Remotely Sensed Data*. PhD thesis, University of Maryland, College Park, 2015.
- [112] James M. Murphy, Jacqueline Le Moigne, and David J. Harding. Automatic image registration of multimodal remotely sensed data with global shearlet features. *IEEE Transactions on Geoscience and Remote Sensing*, 54(3):1685–1704, 2016.
- [113] Greg Ongie, Sampurna Biswas, and Mathews Jacob. Convex recovery of continuous domain piecewise constant images from non-uniform fourier samples. *IEEE Transactions on Signal Processing*, 2017.
- [114] Greg Ongie and Mathews Jacob. Off-the-grid recovery of piecewise constant images from few Fourier samples. *SIAM Journal on Imaging Sciences*, 9(3):1004–1041, 2016.
- [115] Hanjie Pan, Thierry Blu, and Pier Luigi Dragotti. Sampling curves with finite rate of innovation. *IEEE Transactions on Signal Processing*, 62(2):458–471, 2014.
- [116] Sung C. Park, Min K. Park, and Moon G. Kang. Super-resolution image reconstruction: a technical overview. *Signal Processing Magazine, IEEE*, 20(3):21–36, 2003.
- [117] Par R. Prony. Essai expérimentale et analytique. *J. de L’Ecole Polytechnique*, 1(22):24–76, 1795.
- [118] K.G. Puschmann and F. Kneer. On super-resolution in astronomical imaging. *Astronomy & Astrophysics*, 436(1):373–378, 2005.
- [119] Fred Rieke. *Spikes: exploring the neural code*. MIT press, 1999.
- [120] Richard Roy and Thomas Kailath. ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(7):984–995, 1989.
- [121] Walter Rudin. *Fourier Analysis on Groups*. John Wiley & Sons, Inc., New York, 1962.
- [122] Raphaël Salem. On singular monotonic functions of the Cantor type. *Journal of Mathematical Physics*, 21:69–82, 1942.
- [123] Uri Shaham, Alexander Cloninger, and Ronald R. Coifman. Provable approximation properties for deep neural networks. *Applied and Computational Harmonic Analysis*, 2016.

- [124] Laurent Sifre and Stéphane Mallat. Rotation, scaling and deformation invariant scattering for texture discrimination. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1233–1240, 2013.
- [125] Elias M. Stein. *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, volume 3. Princeton University Press, 1993.
- [126] Petre Stoica and Randolph L. Moses. *Introduction to Spectral Analysis*, volume 1. Prentice hall Upper Saddle River, 1997.
- [127] Gongguo Tang, Badri N. Bhaskar, and Benjamin Recht. Near minimax line spectral estimation. *IEEE Transactions on Information Theory*, 61(1):499–512, 2015.
- [128] Gongguo Tang, Badri N. Bhaskar, Parikshit Shah, and Benjamin Recht. Compressed sensing off the grid. *IEEE Transactions on Information Theory*, 59(11):7465–7490, 2013.
- [129] Paul Turán. On rational polynomials. *Acta Universitatis Szegediensis, Sect. Sci.*, pages 106–113, 1946.
- [130] Jeffrey D. Vaaler. Some extremal functions in fourier analysis. *Bulletin of the American Mathematical Society*, 12(2):183–216, 1985.
- [131] Irène Waldspurger. Exponential decay of scattering coefficients. In *Sampling Theory and Applications (SampTA), 2017 International Conference on*, pages 143–146. IEEE, 2017.
- [132] Daniel Weinberg. *Multiscale and Directional Representations of High-Dimensional Information Content in Remotely Sensed Data*. PhD thesis, University of Maryland, College Park, 2015.
- [133] Holger Wendland. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in Computational Mathematics*, 4(1):389–396, 1995.
- [134] Thomas Wiatowski and Helmut Bölcskei. A mathematical theory of deep convolutional neural networks for feature extraction. *IEEE Transactions on Information Theory*, 2017.
- [135] Thomas Wiatowski, Philipp Grohs, and Helmut Bölcskei. Energy propagation in deep convolutional neural networks. *IEEE Transactions on Information Theory*, 2017.
- [136] Weiyu Xu, Jian-Feng Cai, Kumar V. Mishra, Myung Cho, and Anton Kruger. Precise semidefinite programming formulation of atomic norm minimization for recovering d -dimensional ($D \geq 2$) off-the-grid frequencies. In *Information Theory and Applications Workshop, 2014*, pages 1–4. IEEE, 2014.

- [137] Antoni Zygmund. *Trigonometric Series*, volume 1. Cambridge University Press, 2002 (1959).