Off-the-Grid Recovery of Piecewise Constant Images from Few Fourier Samples

Greg Ongie† and Mathews Jacob‡

October 5, 2015

Abstract

We introduce a method to recover a continuous domain representation of a piecewise constant two-dimensional image from few low-pass Fourier samples. Assuming the edge set of the image is localized to the zero set of a trigonometric polynomial, we show the Fourier coefficients of the partial derivatives of the image satisfy a linear annihilation relation. We present necessary and sufficient conditions for unique recovery of the image from finite low-pass Fourier samples using the annihilation relation. We also propose a practical two-stage recovery algorithm which is robust to model-mismatch and noise. In the first stage we estimate a continuous domain representation of the edge set of the image. In the second stage we perform an extrapolation in Fourier domain by a least squares two-dimensional linear prediction, which recovers the exact Fourier coefficients of the underlying image. We demonstrate our algorithm on the super-resolution recovery of MRI phantoms and real MRI data from low-pass Fourier samples, and show the superiority of the method over standard approaches for single-image super-resolution MRI.

1 Introduction

Many studies in image processing show that natural images can be represented efficiently as piecewise smooth functions [24, 38, 13]. For instance, the Mumford-Shah formulation poses the recovery of a piecewise smooth image as a variational optimization problem that jointly estimates a segmentation of the image and the smooth functions on the regions defined by the segmentation [27, 28]. This problem is typically solved using iterative algorithms that alternate between the evolution a level-set curve and the recovery of the full signal [8, 43, 3]. These constrained level-set methods have been demonstrated to yield considerably improved reconstructions in challenging inverse problems [17, 20, 40, 3, 4, 15, 34]. Unfortunately, these algorithms are often slow and vulnerable to local minima problems, which restrict their utility in applications.

The recovery of piecewise smooth one-dimensional signals using harmonic retrieval/linear prediction has been actively investigated in signal processing [23, 45, 25, 16, 2]. This includes the so-called finite-rate-of-innovation (FRI) framework [2] which extends Prony’s method [39, 11] to the recovery of a much wider class of signals, including non-uniform 1-D splines with unknown knot locations. While some extensions of 1-D FRI schemes to multidimensional polygonal signals are available [26, 36, 9], these specialized models do not provide good approximations of natural images. Recent work on continuous domain total variation norm minimization can also be applied to the recovery of a weighted linear combination of 1-D Diracs at

---

*This work is supported by grants NSF CCF-0844812, NSF CCF-1116067, NIH 1R21HL109710-01A1, ACS RSG-11-267-01-CCE, and onr-n000141310202.

†Department of Mathematics, University of Iowa, Iowa City, Iowa. (gregory-ongie@uiowa.edu).

‡Department of Electrical and Computer Engineering, University of Iowa, Iowa City, Iowa. (mathews-jacob@uiowa.edu)
arbitrary locations from few Fourier samples \cite{6,42,7}; these methods are off-the-grid continuous domain generalizations of compressed sensing theory \cite{14,1}. While the framework can account for piecewise constant 1-D signals \cite{6} and linear combination of Diracs in 2-D \cite{46} by analyzing the derivative of the signal, its extension to piecewise constant signals in two or higher dimensions is not straightforward. For instance, the partial derivatives piecewise constant images have singularities supported on curves that bound the constant regions. We observe that this violates two of the main assumptions required for continuous domain total variation super-resolution recovery: (1) that the number of singularities is finite, and (2) their minimum separation distance is bounded below.

Recently, Pan et al. \cite{32} proposed an extension of FRI methods to recover piecewise complex analytic images in continuous domain. Specifically, they assume that the complex derivative of the image is supported on the zero level-set of a band-limited periodic function, called the annihilating polynomial. The authors rely on the annihilation of the Fourier transform of the derivatives by convolution with the Fourier coefficients of the annihilating polynomial to recover the edge set, and eventually the image. The work in this paper is mainly motivated by \cite{32}, and aims to overcome its limitations. For instance, the recovery of an arbitrary piecewise complex analytic signal from few Fourier coefficients is ill-posed. This is because we may choose complex analytic functions such that a finite collection of its Fourier coefficients vanish. Even in recovering the edge set, no sampling conditions that guarantee perfect recovery were available in \cite{32}. Another problem is that the complex analytic signal model in \cite{32} is not realistic for natural images.

The main focus of this paper is to develop a novel theoretical framework and efficient algorithms to recover piecewise constant images, whose discontinuities are localized to zero level-sets of band-limited functions, from few Fourier samples. We introduce a two step algorithm, which has similarities to Prony’s method. The first step involves the determination of the discontinuities of the image as the zero set of a band-limited function, which translates to a set of linear annihilation conditions in the Fourier domain. Once the edge set is recovered, we recover the full image as the solution to an optimization problem. This work connects the classical Mumphord-Shah algorithm, which alternates between level-set estimation and signal estimation, with FRI theory, which provides recovery guarantees from finite number of Fourier samples.

Unlike the piecewise analytic model in \cite{32}, the number of degrees of freedom of our proposed piecewise constant representation is finite, which enables the successful recovery of the signal from finite number of Fourier coefficients. We introduce necessary and sufficient conditions for the recovery of the edge set and the edge set aware recovery of the signal. The one-to-one correspondence between the degree of the annihilating polynomial and the number of exponentials in the signal, which is used to establish guarantees in the traditional FRI settings, cannot be exploited in our setting since the number of exponentials is no longer finite. We introduce novel proof techniques based on the algebraic geometry of curves defined as zero sets of trigonometric polynomials to establish the uniqueness guarantees.

We also introduce several refinements to make the two-step algorithm robust to noise and model mismatch. Since the number of coefficients necessary to describe the edge set is not typically known in advance, we rely on singular value decomposition of the block Toeplitz matrix corresponding to the linear annihilation relation for model order determination. Once the null-space of the matrix is estimated, we pose the extrapolation of the Fourier coefficients of the signal from the known samples as an optimization problem. The algorithm determines the signal such that the energy of the convolution between its Fourier coefficients and the null-space filters is minimized. We show that this approach is approximately equivalent to estimating the edge set as a sum of squares polynomial built from a basis of null-space filters. This approach is similar to the projection method used in the multiple signal classification algorithm (MUSIC) algorithm \cite{33} used in direction-of-arrival estimation and other signal processing applications. We demonstrate the proposed scheme on the the super-resolution of magnetic resonance (MR) images from low-resolution Fourier samples. The experiments show that the framework is superior to other standard
discrete spatial domain approaches to single-image super-resolution schemes for MRI. In particular, our scheme shows advantages in being able to resolve fine image features, preserve strong edges, and suppress noise artifacts.

Preliminary versions of this work were presented in conference papers [30] and [29]. In particular, Theorems 1 and 2 were stated in [29] without proof. In addition to the proofs of these results, we provide sampling guarantees for edge aware signal recovery that were absent in [29]. While the basic version of the two step algorithm presented here was introduced in [30], it is considerably generalized and validated using MRI data in this work.

2 Overview

Similar to Prony’s method and FRI signal recovery, we introduce a two step algorithm for the recovery of piecewise constant images from few low-pass Fourier samples, assuming the edge set of the image is well-approximated by the zero level-set of a bandlimited periodic function, called the annihilating polynomial. In the first step, we estimate the edge set by solving for the annihilating polynomial. In the second step, we use the annihilating polynomial to perform an edge-aware recovery of the full signal. We start with a brief review of FRI signal recovery in 1-D to make the presentation self-contained and to motivate our proposed 2-D recovery scheme.

2.1 Review of 1-D FRI theory

The FRI framework introduced in [45] can be viewed as an extension of Prony’s method, also known as the annihilating filter method, which enables the recovery of a wide class of signals from sub-Nyquist sampling rates (we also draw attention to a similar work [23] that predates [45]). Here we focus on the FRI recovery of continuous domain 1-D piecewise constant signals \( f : [0, 1] \to \mathbb{C} \), given as

\[
 f(x) = \sum_{i=1}^{K-1} a_i 1_{[x_i, x_{i+1})}(x);
\]

from the low-pass Fourier coefficients

\[
 \hat{f}[k] = \int_0^1 f(x) e^{-j2\pi k x} dx; \quad k \in \mathbb{Z}, \ |k| \leq N.
\]

for some fixed \( N \). In \([1]\), \( 1_U \) denotes the characteristic function of the set \( U \), \( a_i \in \mathbb{C} \) are the signal amplitudes, and \( 0 < x_1 < x_2 < \cdots < x_N < 1 \) are the edge locations. See Figure 1a for an illustration of this signal model and the two-step FRI recovery described below.

First, observe that the derivative of \( f \) (in the sense of distributions) is the stream of Diracs:

\[
 \partial f(x) = \sum_{i=1}^{K} b_i \delta(x - x_i);
\]

where \( b_i = a_i - a_{i-1} \), for \( i = 1, \ldots, K \) with \( a_K = a_0 = 0 \); note the amplitudes \( a_i \) are recoverable from the \( b_i \) given the assumed boundary conditions. Therefore, the Fourier coefficients of \( \partial f \) are a linear combination of exponentials

\[
 \hat{\partial f}[k] = \sum_{i=1}^{K} b_i e^{-j2\pi k x_i},
\]
which we can compute from the known samples \( \hat{f}[k] \) by the Fourier domain relation

\[
\hat{\partial f}[k] = -j2\pi k \hat{f}[k].
\]

Recovering \( f \) now reduces to estimating the \( 2K \) unknown parameters \( \{b_i, x_i\}_{i=1}^K \) from the known \( 2N + 1 \) samples if (3). This can be achieved using Prony’s method, or one of its robust variants [11]. The solution relies on determining the so-called annihilating polynomial, i.e. the trigonometric polynomial defined as

\[
\mu(x) = \prod_{i=1}^K \left( e^{j2\pi x} - e^{j2\pi x_i} \right) = \sum_{k=0}^K c[k] e^{j2\pi k x}.
\]

which satisfies \( \mu(x_i) = 0 \) for each of the edge locations \( \{x_i\}_{i=1}^K \), and these are its only zeros. An easy computation shows that the Fourier coefficients \( \hat{\partial f}[k] \) are annihilated by convolution with the annihilating polynomial coefficients \( c[k] \), i.e.

\[
\sum_{k=0}^K c[k] \hat{f}[\ell - k] = 0, \quad \text{for all } \ell \in \mathbb{Z}
\]

This can also be seen from fact that, in spatial domain,

\[ \mu \partial f = 0, \]

in the sense of distributions. Therefore, provided we have at least \( 2K \) contiguous samples of \( \hat{\partial f} \) (i.e. \( N \geq K \)) we can form a linear Toeplitz system from (5), and solve for the \( K + 1 \) unknown coefficients \( c[k] \). Then we may solve for the coefficients \( b_i \), and hence the amplitudes \( a_i \), by substituting the \( \{x_i\}_{i=1}^N \) into the model (3) and solving an overdetermined linear system.

The main tool enabling FRI recovery is the existence of a trigonometric polynomial whose zero set coincides with the edge locations of the piecewise constant signal, or equivalently, the singularities of its derivative. Similar to [32], we propose extending this scheme to 2-D by considering the infinite one-dimensional zero sets of multi-variable trigonometric polynomials.

### 2.2 Image Model

We focus on a continuous domain representation of images, that is, as functions \( f : \mathbb{R}^2 \rightarrow \mathbb{C} \) with compact support. For simplicity we will assume the support is always contained within \([0,1]^2\). We are particularly interested in images that can be modeled as 2-D piecewise constant functions

\[
f(r) = \sum_{i=1}^n a_i 1_{\Omega_i}(r), \quad \text{for all } \ r = (x, y) \in [0,1]^2,
\]

where \( a_i \in \mathbb{C} \), and \( 1_{\Omega} \) is the characteristic function of the set \( \Omega \):

\[
1_{\Omega}(r) = \begin{cases} 
1 & \text{if } r \in \Omega \\
0 & \text{else}
\end{cases}
\]

Here we assume each region \( \Omega_i \subseteq [0,1]^2 \) has piecewise smooth boundary \( \partial \Omega_i \). We also assume the representation (6) is expressed with the fewest number of characteristic functions such that the regions \( \Omega_i \) are pairwise disjoint. In this case the set of discontinuous points of \( f \) coincides with \( E = \bigcup_{i=1}^n \partial \Omega_i \), which we call the edge set of the image.

Our aim is to recover the piecewise constant image \( f(r) \) from few of its Fourier coefficients

\[
\hat{f}[k] = \int_{[0,1]^2} f(r) e^{-j2\pi k \cdot r} \, dr, \quad k \in \Gamma \subseteq \mathbb{Z}^2
\]
Figure 1: Overview of proposed extension of FRI to 2-D piecewise constant images. In (a) we illustrate the FRI recovery of a 1-D continuous domain piecewise constant signal \( f(x) \), as described in §2.1. The distributional derivative \( \partial f(x) \) is a stream of Diracs supported on the edge set of the signal. Hence, multiplication by a trigonometric polynomial \( \mu(x) \) which is zero on the edge set will annihilate \( \partial f(x) \) in spatial domain. This descends to a linear convolution relation in Fourier domain. Therefore, given sufficiently many uniform low-pass Fourier samples \( \hat{f}[k] \) we may compute \( \hat{\partial f}[k] \), and form a linear system to solve for the finite Fourier coefficients \( c[k] \) of \( \mu(x) \). In (b) we illustrate the proposed extension of this scheme to 2-D presented in §2.3. Here \( f(r) \), \( r = (x, y) \), is now the characteristic function for a region \( \Omega \) in the plane. We show the distributional gradient of \( f \) is supported on the one-dimensional edge set \( \partial \Omega \). Multiplication by a trigonometric polynomial \( \mu(r) \) which vanishes on \( \partial \Omega \) will annihilate \( \nabla f \) in spatial domain, and this likewise translates to a linear convolution annihilation relation in Fourier domain, which we may use to solve for coefficients \( c[k] \) of \( \mu(r) \).
where $\Gamma$ is some finite sampling set.

Without any further assumptions on the edge set, unique recovery of images in the form \([6]\) from finitely many Fourier samples is clearly ill-posed since the class of piecewise smooth curves has infinite degrees of freedom. Hence, we restrict the edge set to belong to a parametric model described by finitely many parameters. Motivated by the FRI curves framework [32], we will constrain the edge set to be the zero set of a 2-D band-limited periodic function:

$$C = \{ r \in [0, 1]^2 : \mu(r) = 0 \}, \quad \text{where} \quad \mu(r) = \sum_{k \in \Lambda} c[k] e^{2\pi i k \cdot r}, \quad \forall r \in [0, 1]^2, \tag{9}$$

where the coefficients $c[k] \in \mathbb{C}$, and $\Lambda$ is any finite subset of $\mathbb{Z}^2$.

We call any function $\mu$ described by \([9]\) a trigonometric polynomial, and the zero set $C = \{ \mu = 0 \}$ a trigonometric curve, provided $C$ is infinite and has no isolated points. Note a piecewise constant image cannot have isolated points in its edge set, so this restriction on the zero set is fully compatible with our image model. We elaborate more on the properties of trigonometric curves in §2.4; see also Fig. 2.

### 2.3 Annihilation property

We now show that the Fourier coefficients of the partial derivatives of piecewise constant images \([6]\) having edge set described by a trigonometric curve \([9]\) satisfy an annihilation relation similar to \([5]\) derived for 1-D piecewise constant signals.

First consider the case of a single characteristic function $1_\Omega$ where $\Omega$ is a simply connected region with piecewise smooth boundary $\partial \Omega$. We will show that the product of the gradient of $1_\Omega$ by any smooth periodic function $\psi$ vanishing on $\partial \Omega$ is identically zero

$$\psi \nabla 1_\Omega = 0. \tag{10}$$

However, as $1_\Omega$ is non-smooth, the above equality can only be understood in a weak sense. In particular, it is useful to interpret the partial derivatives of $1_\Omega$ as tempered distributions, i.e. as linear functionals acting on test functions of Schwartz class $\mathcal{S}$ (see, e.g., [41] or [17]). Since the domain $[0, 1]^2$ is compact, in this case $\mathcal{S}$ coincides with all $C^\infty$ smooth complex-valued periodic functions on $[0, 1]^2$. Letting $\varphi \in \mathcal{S}$ denote any such test function, we have:

$$\langle \partial_x 1_\Omega, \varphi \rangle = -\langle 1_\Omega, \partial_x \varphi \rangle = - \int_\Omega \partial_x \varphi \, dr = - \oint_{\partial \Omega} \varphi \, dy \tag{11}$$

where the last step follows by Green’s theorem. Likewise, one can show

$$\langle \partial_y 1_\Omega, \varphi \rangle = \oint_{\partial \Omega} \varphi \, dx. \tag{12}$$

Generalizing the above relation to the gradient $\nabla 1_\Omega$ by setting $\varphi = (\varphi_1, \varphi_2)$ for any $\varphi_1, \varphi_2 \in \mathcal{S}$, we obtain

$$\langle \nabla 1_\Omega, \varphi \rangle = -\langle 1_\Omega, \nabla \cdot \varphi \rangle = - \int_\Omega \nabla \cdot \varphi \, dr = - \oint_{\partial \Omega} \varphi \cdot \mathbf{n} \, ds \tag{13}$$

where $\mathbf{n}$ is the unit outward normal to the curve $\partial \Omega$.

If $\psi$ is any smooth function that vanishes on $\partial \Omega$, i.e. $\psi(r) = 0$ for all $r \in \partial \Omega$, then the product $\psi \nabla 1_\Omega$ is well-defined as a tempered distribution, and we have

$$\langle \psi \nabla 1_\Omega, \varphi \rangle = \langle \nabla 1_\Omega, \psi \varphi \rangle = - \oint_{\partial \Omega} \psi (\varphi \cdot \mathbf{n}) \, ds = 0, \tag{14}$$
and thus we have shown (10) holds in the distributional sense.

The above analysis easily generalizes to piecewise constant functions by linearity: if \( f \) is a piecewise constant function (6), and \( \phi \) is any smooth function that vanishes on the edge set \( E = \bigcup_{i=1}^n \partial \Omega_i \), then

\[
\phi \nabla f = \sum_{i=1}^n a_i \left( \phi \nabla 1_{\Omega_i} \right) = 0.
\]

(15)

Now assume that \( \phi = \mu \) is a trigonometric polynomial, and hence is described by a finite number of Fourier coefficients, which we denote by \( \{ c[k] \}_{k \in \Lambda} \), where \( \Lambda \subseteq \mathbb{Z}^2 \) is its finite Fourier support set. Taking Fourier transforms of (15), and applying the convolution theorem for tempered distributions yields the following annihilation relation:

**Proposition 1.** Let \( f = \sum_{i=1}^n a_i 1_{\Omega_i} \) be piecewise constant where the edge set \( E = \bigcup_{i=1}^n \partial \Omega_i \) is a subset of the trigonometric curve \( \{ \mu = 0 \} \). Then the Fourier transform of the partial derivatives of \( f \) are annihilated by convolution with the Fourier coefficients \( \{ c[k] \}_{k \in \Lambda} \) of \( \mu \):

\[
\sum_{k \in \Lambda} c[k] \hat{\nabla} f[\ell - k] = 0, \quad \forall \ell \in \mathbb{Z}^2.
\]

(16)

where \( \hat{\nabla} f[k] = (\hat{\partial_x} f[k], \hat{\partial_y} f[k]) \).

Due to the above property, we call \( \mu \) an **annihilating polynomial** for \( f \), and the coefficients \( \{ c[k] \}_{k \in \Lambda} \) an **annihilating filter**. We show in Section 3 that the above relations can be used to recover an annihilating polynomial, and hence the edge set, from finitely many low-pass Fourier coefficients of the image.

### 2.4 Properties of the edge set model: Trigonometric curves

Our image model, which assumes the edge set coincides with a trigonometric curve (9), may at first seem overly restrictive. However, trigonometric curves can represent a wide variety of curve topologies, even with relatively few coefficients. For instance, trigonometric curves can have multiple connected components, non-smooth cusps, and can self-intersect in complicated ways. Moreover, any compactly supported plane curve can be approximated to an arbitrary degree of accuracy by a trigonometric curve, provided the number of coefficients is large enough; see Fig. 2.

Trigonometric curves can be characterized in terms similar to algebraic plane curves, i.e. the zero sets of two-variable complex polynomials in the complex plane. In particular, one can associate with every algebraic plane curve a unique minimal degree polynomial whose zero set generates the curve. Similarly, we prove there is a unique minimal degree trigonometric polynomial associated with any trigonometric curve \( C \), which we call the **minimal polynomial** for \( C \). Here we define the **degree** of a trigonometric polynomial \( \mu \) to be the dimensions of the smallest rectangle that contains the frequency support set \( \Lambda \), which we denote as \( \deg(\mu) = (K,L) \). In Appendix A we prove the following:

**Proposition 2.** For every trigonometric curve \( C \) there is a unique (up to scaling) real-valued trigonometric polynomial \( \mu_0 \) with \( C = \{ \mu_0 = 0 \} \) such that for any other trigonometric polynomial \( \mu \) with \( C = \{ \mu = 0 \} \) we have \( \deg(\mu_0) \leq \deg(\mu) \) componentwise.

This proposition shows trigonometric curves can always be described as the zero set of a **real-valued** trigonometric polynomial, a property we will use implicitly from now on. Moreover, it allows us to uniquely define the **degree** of a trigonometric curve as the degree of its associated minimal polynomial. Intuitively, the degree gives some measure of inherent complexity of the curve. For example, the degree puts bounds on how many non-smooth points a curve may have and the number of connected curve components. One particularly important bound for this work is the following:
Proposition 3. If $C$ is a trigonometric curve of degree $(K, L)$, then the number of connected components of the complement set $[0, 1]^2 / C$ is at most $(K + L)^2$.

Consequently, for a piecewise constant image described by the edge set a trigonometric curve of degree $(K, L)$, the image cannot consist of more than $(K + L)^2$ distinct regions, and so the number of terms in the sum (6) is bounded by $(K + L)^2$.

Some further topological and algebraic properties of minimal polynomials can be found in Appendix B, which are used in proofs of the sampling theorems in §3 and §4.

Figure 2: Examples of trigonometric curves. The one-dimensional zero sets of two-variable trigonometric polynomials, i.e. trigonometric curves, are a topologically diverse class of plane curves, which can (a) have many distinct curve components, (b) non-smooth cusps and self-intersections (indicated by arrows), and (c) approximate arbitrary plane curves using few coefficients. Shown inset is a surface plot of the trigonometric polynomial whose zero set gives the corresponding trigonometric curve.

3 Recovery of edge set from finite Fourier samples

In this section, we focus on sampling guarantees for the first step of our recovery scheme. We give necessary and sufficient conditions on the number of samples required for unique recovery of the edge set of the signal from the annihilation relation. Similar sampling guarantees for edge-aware recovery of the full signal are provided in the next section.

Let $f$ be any piecewise constant image whose edge set is described by a trigonometric curve $\{\mu = 0\}$ where $\mu$ is a trigonometric polynomial with coefficients $\{c[k]\}_{k \in \Lambda}$. If we have access to all Fourier samples $\hat{f}[k]$ for all $k \in \Gamma$, then from the annihilation relation (16) we can form the finite linear system of equations:

$$\sum_{k \in \Lambda} c[k] \hat{\nabla} f[\ell - k] = 0, \quad \forall \, \ell \in \Gamma | \Lambda$$

(17)

where the index set $\Gamma | \Lambda \subseteq \mathbb{Z}^2$ is the set of all shifts $\ell \in \mathbb{Z}^2$ such that $\ell - k \in \Gamma$ for all $k \in \Lambda$. Here the expression $\hat{\nabla} f[k] = (\hat{\partial}_x f[k], \hat{\partial}_y f[k])$ is computed from the known samples using the Fourier domain relations $\hat{\partial}_x f[k] = -j 2\pi k_x \hat{f}[k]$, and $\hat{\partial}_y f[k] = -j 2\pi k_y \hat{f}[k]$, for any $k = [k_x, k_y] \in \mathbb{Z}^2$. Provided we have enough samples of $\hat{f}$ to construct a determined linear system of equations (17), we can potentially solve for the unknown coefficients $\{c[k]\}_{k \in \Lambda}$. We now derive conditions for the necessary and sufficient number of samples needed for unique recovery via the linear system (17).
3.1 Necessary condition for recovery of edge set

Suppose the Fourier sampling set \( \Gamma \) is rectangular of size \( K' \times L' \), and the minimal polynomial \( \mu \) for the edge set has degree \( (K, L) \), i.e. the smallest rectangular index set \( \Lambda \subseteq \mathbb{Z}^2 \) containing the Fourier support of \( \mu \) has a total of \( (K+1) \times (L+1) \) indices. By considering all possible integer shifts of \( \Lambda \) inside \( \Gamma \), and accounting for both partial derivatives, in this case we may form \( 2(K' - K)(L' - L) \) valid equations from (17). Because we must solve for \( (K+1)(L+1) \) unknown coefficients supported within \( \Lambda \), and since we can assume the coefficients are non-trivial, we have the following necessary condition for recovery of the coefficients of \( \mu \), or equivalently, the edge set:

**Proposition 4.** Let \( f \) be piecewise constant such that the edge set is a trigonometric curve of degree \( (K, L) \). A necessary condition to recover the edge set from (17), is to collect samples of \( \hat{f} \) on a \((K', L')\) rectangular grid such that

\[
2(K' - K)(L' - L) \geq (K+1)(L+1) - 1.
\]

To illustrate this bound, suppose \( \mu \) has degree \( (K, K) \), and we take Fourier samples from a square grid of size \( K' \times K' \). Then we need \( K' \geq 1.71K \), or roughly \( 3K^2 \) Fourier samples to recover the edge set. Our numerical experiments on simulated data (see Fig. 3) indicate the above necessary condition might also be sufficient for unique recovery; that is, we hypothesize the coefficients of the minimal polynomial \( \{c[k]\}_{k \in \Lambda} \) are the only non-trivial solution to the system of equations (17) when it is fully determined.

3.2 Sufficient conditions for recovery of edge set

We now focus on sufficient conditions for the recovery of the edge set of a piecewise constant image from (17). For simplicity we assume that all index sets in \( \mathbb{Z}^2 \) are rectangular and symmetric about the origin. However, the analysis in this section can be readily extended to arbitrary shaped sampling sets \( \Gamma \subseteq \mathbb{Z}^2 \) with sufficiently many contiguous uniform sampling locations.

First, we introduce some additional notation to aid in the statement of our results: Let \( \Gamma, \Lambda \subseteq \mathbb{Z}^2 \) be rectangular index sets symmetric about the origin. We write \( \Gamma|\Lambda \) to denote the index set \( \{\ell \in \Gamma : (\ell - k) \in \Gamma \text{ for all } k \in \Lambda\} \), and write \( \Gamma*\Lambda \) to denote the set \( \{\ell - k : k \in \Lambda, \ell \in \Gamma\} \). With these definitions, one can check that if \( \gamma \) and \( \mu \) are trigonometric polynomials with coefficients supported within \( \Gamma \) and \( \Lambda \), respectively, then \( \eta = \gamma\mu \) has coefficients supported within \( \Delta = \Gamma*\Lambda = \Lambda*\Gamma \), and \( \gamma = \eta/\mu \) has coefficients supported in \( \Gamma = \Delta|\Lambda \). We also write \( 2\Lambda = \Lambda*\Lambda, 3\Lambda = 2\Lambda*\Lambda \), etc. Note that \( m\Lambda \) is a dilation of the set \( \Lambda \) by a factor of \( m \) in each dimension: if \( \Lambda = \{(k, \ell) \in \mathbb{Z}^2 : |k| \leq K, |\ell| \leq L\} \), then \( m\Lambda = \{(k, \ell) \in \mathbb{Z}^2 : |k| \leq mK, |\ell| \leq mL\} \).

For simplicity, we first state our result in the case of a single characteristic function. We give the proof of this result, and all the remaining proofs in [3] [4] in Appendix [4].

**Theorem 1.** Let \( f = 1_\Omega \) be the characteristic function of a simply connected region \( \Omega \) with boundary \( \partial \Omega = \{\mu = 0\} \), \( \mu \) the minimal polynomial with coefficients \( \{c[k]\}_{k \in \Lambda} \). Then the \( c[k] \) can be uniquely recovered (up to scaling) as the only non-trivial solution to the equations

\[
\sum_{k \in \Lambda} c[k] \nabla f[\ell - k] = 0, \text{ for all } \ell \in 2\Lambda.
\]

The proof of Theorem 1 entails showing any other trigonometric polynomial \( \eta \) having coefficients \( \{d[k]\}_{k \in \Lambda} \) satisfying (18) must vanish on \( \partial \Omega \). From this it follows that \( \eta \) is a scalar multiple of \( \mu \) by properties of minimal polynomials. We now extend the above result to piecewise constant signals, provided the boundaries of the regions do not intersect:
Figure 3: Exact recovery of edge set of piecewise constant images from the minimum necessary number of Fourier samples. The piecewise constant images shown in (a)&(d) were generated to have edge set given by the trigonometric curve \( \{ \mu = 0 \} \) of degree \((6, 6)\), i.e. the known minimal polynomial \( \mu \) is described by \(7 \times 7\) coefficients. We recover \( \mu \) by solving the linear system (17) built from \(11 \times 11\) Fourier samples of the original image, which were approximated numerically to high precision. This is the necessary minimum number of samples specified by Prop. 4, and is less than the sufficient number of samples specified by Thm. 2 (\(19 \times 19\) samples). The recovered minimal polynomial \( \mu \) is shown in (b)&(e) with \( \{ \mu = 0 \} \) plotted in black; we observe the recovered coefficients match the known minimal polynomial coefficients to within machine precision. The normalized singular values of the linear system (17) are plotted in (c)&(f). Note the last singular value is orders of magnitude smaller than the others, which indicates numerically that the recovered solution is unique.

**Theorem 2.** Let \( f = \sum_{i=1}^{n} a_i 1_{\Omega_i} \) be piecewise constant such that the boundary curves \( \partial \Omega_i \), \( i = 1, \ldots, n \), are pairwise disjoint and are given by trigonometric curves: \( \partial \Omega_i = \{ \mu_i = 0 \} \), \( \mu_i \) the minimal polynomial. Then, the coefficients \( \{ c[k] \}_{k \in \Lambda} \), of \( \mu = \mu_1 \cdots \mu_n \), and equivalently the edge set \( E = \bigcup_{i=1}^{n} \partial \Omega_i \), can be uniquely recovered (up to scaling) as the only non-trivial solution of

\[
\sum_{k \in \Lambda} c[k] \hat{\nabla} f[\ell - k] = 0, \quad \text{for all } \ell \in 2\Lambda.
\]

Note that to form the equations in (18) and (19) requires access to Fourier samples \( \hat{f}[k] \) for all \( k \in 3\Lambda \), which is greater than the necessary number of samples given in Theorem 4 (approximately \(1.7\Lambda\)). In other words, there is a gap in our results between the necessary and sufficient number of samples needed for unique recovery of the edge set. We conjecture that Theorems 1 and 2 can in fact be sharpened to the necessary number of samples, and extended to piecewise constant signals where the boundaries of the regions are allowed to intersect.
We also provide a generalization of Theorem 2 that will be useful in practical applications when the degree of the edge set is unknown. Specifically, assume that the system (17) is built with a larger coefficient support set \( \Lambda' \) than the minimal support \( \Lambda \). In this case the solution to (17) is not unique: any multiple of the minimal polynomial \( \eta = \mu \gamma \) such that \( \gamma \) has coefficients in \( \Lambda' \mid \Lambda \) is also a solution. This is because \( \eta \) has coefficients supported within \( \Lambda' \), and satisfies \( \{ \mu = 0 \} \subseteq \{ \eta = 0 \} \), hence is also an annihilating polynomial. We now show that every solution to (17) must be multiple of the minimal polynomial under some restrictions on the edge set.

**Theorem 3.** Let \( f \) be as in Thm. 2. Suppose \( \mu = \mu_1 \cdots \mu_n \) has coefficients in \( \Lambda \), and let \( \Lambda' \) be another rectangular index set with \( \Lambda \subseteq \Lambda' \). If the coefficients \( \{ d[k] \}_{k \in \Lambda'} \) of a trigonometric polynomial \( \eta \) are a non-trivial solution of

\[
\sum_{k \in \Lambda'} d[k] \hat{\nabla} f[\ell - k] = 0, \quad \text{for all } \ell \in \Lambda' * \Lambda,
\]

then \( \mu \) divides \( \eta \), that is \( \eta = \mu \gamma \) where \( \gamma \) is another trigonometric polynomial with coefficients supported in \( \Lambda' \mid \Lambda \).

### 4 Edge set aware recovery of the image

In this section we establish conditions for the unique recovery of a piecewise constant image from few of its low-pass Fourier samples with knowledge of the minimal polynomial describing its edge set. We will study the problem of identifiability of the signal in continuous domain: given ideal Fourier samples \( \hat{f}[k], k \in \Gamma \), and the minimal polynomial \( \mu \), when is \( f \) is the unique function satisfying the annihilation relation, \( \mu \nabla f = 0 \)? Here we will restrict solutions \( f \) to belong to \( L^1([0,1]^2) \), and interpret \( \nabla f \) in a a distributional sense.

First, we show any function \( g \in L^1([0,1]^2) \) satisfying the annihilation relation can be compactly parametrized as a piecewise constant function:

**Proposition 5.** Let \( \mu \) be any trigonometric polynomial. Suppose \( g \in L^1([0,1]^2) \) satisfies the annihilation relation

\[
\mu \nabla g = 0 \tag{21}
\]

where equality holds in the distributional sense. Then, \( g \) can be parameterized as the piecewise constant function

\[
g = \sum_{i=1}^{n} b_i 1_{\Omega_i} \tag{22}
\]

almost everywhere, where \( b_i \in \mathbb{C} \) and \( \Omega_i \) are the connected components of \( \{ \mu = 0 \}^C \).

The above proposition implies that the recovery of the image now simplifies to the estimation of the unknown amplitudes \( \{ b_i \}_{i=1}^{n} \), provided the regions \( \Omega_i \) are known. One approach of recovering the signal then would be to substitute the parametric model (22) back into the annihilation system (17), and solve for the unknown amplitudes, analogous to the amplitude recovery step in Prony’s method. However, this direct approach is not feasible in practice, since to identify the regions \( \Omega_i \) would require us to factor a high degree multivariate polynomial having arbitrary complex coefficients. For this reason, obtaining an exact off-the-grid parametric representation of the full signal in spatial domain is especially challenging in our setting.

We avoid this direct approach and instead pose recovery as the solution to the following optimization problem in spatial domain:

\[
\min_{g \in L^1} \| \mu \nabla g \| \quad \text{such that } \hat{g}[k] = \hat{f}[k], \quad \text{for all } k \in \Gamma, \tag{23}
\]
where $\| \cdot \|$ is any suitable norm. Alternatively, we can solve for the solution entirely in Fourier domain:

$$\min_{\hat{g}[k], k \in \mathbb{Z}} \| c \ast \nabla \hat{g} \| \text{ such that } \hat{g}[k] = \hat{f}[k], \text{ for all } k \in \Gamma.$$  \hfill (24)

where $\{c[k]\}_{k \in \Lambda}$ are the finitely supported coefficients of the minimal polynomial $\mu$, $\ast$ denotes discrete 2-D convolution, and again $\| \cdot \|$ is some suitable norm. In the formulation (24), the solution is an extrapolation in Fourier domain from the known samples in $\Gamma$ to the unknown coefficients, which can be viewed as linear prediction in 2-D. Note that (23) and (24) are still infinite-dimensional optimization problems, and not solvable in practice. However, (23) and (24) will motivate the practical algorithms we introduce in §5.

We now derive necessary and sufficient conditions for the unique recovery of the signal from (23) or equivalently (24).

4.1 Necessary conditions

A piecewise constant function $g$ in the form (22) with known regions $\{\Omega_i\}_{i=1}^n$ has $n$ remaining degrees of freedoms corresponding to the unknown region amplitudes $\{b_i\}_{i=1}^n$. Hence the necessary minimum number of samples for unique recovery of the full signal—using any algorithm—is $n$, the number of connected components of $\{\mu = 0\}^C$. However, this quantity will be much smaller than the number of available samples.

For instance, if $\mu$ is a minimal polynomial of degree $(K, L)$, by Prop. 3 we have the bound $n \leq (K + L)^2 \sim 4KL$ when $K \sim L$. However, this bound is not necessarily tight, and from empirical evidence we believe the tight bound is $n \leq KL$, which is only likely to occur in highly artificial cases; see the comments in §A.3. By Prop. 4 the necessary minimum number of samples required to uniquely reconstruct the edge set from the annihilation equations is roughly $3KL$, which is already close to the provable maximum bound on the number of regions, and well above the empirical bound. Therefore, once the edge set is obtained, no additional samples should be necessary to uniquely identify the region amplitudes.

4.2 Sufficient conditions

Under the same non-intersecting assumption on the edge set of the piecewise constant function $f$ as in the Theorems of §3, we may show the low-pass samples $\hat{f}[k]$ for all $k \in \Lambda$, where $\Lambda$ is the coefficient support set of the minimal polynomial, are sufficient to uniquely identify $f$ from the annihilation relation with knowledge of minimal polynomial:

**Theorem 4.** Let $f$ be as in Thm. 2 with known minimal polynomial $\mu$ having coefficients in $\Lambda$, and let the Fourier sampling set $\Gamma \supseteq \Lambda$. If $g \in L^1([0,1]^2)$ satisfies

$$\mu \nabla g = 0 \quad \text{and} \quad \hat{g}[k] = \hat{f}[k], k \in \Gamma$$  \hfill (25)

then $g = f$ almost everywhere.

We conjecture that this result also holds for piecewise constant functions having an edge set described by an arbitrary trigonometric curve $C$, not just those consisting of non-intersecting curves. However, for this conjecture to hold, it is necessary that the connected components of $[0,1]^2/C$ be bounded by $\#\Lambda = (K + 1)(L + 1)$ where $(K, L)$ is the degree of $C$, since otherwise the degrees of freedom of the signal would outnumber the available samples. This is consistent with our observations in §A.3.
4.3 Recovery from finite annihilation relation

The conditions derived in §4.1 and §4.2 settle the problem of unique recovery from using the infinite dimensional annihilation relation (25). A more subtle question, relevant to practical applications, is whether the Fourier extrapolation solution is uniquely determined from finitely many annihilation equations, i.e. for some large extrapolation set $\Delta \subseteq \mathbb{Z}^2$, and some small sampling set $\Gamma \subseteq \mathbb{Z}^2$ do the conditions

$$\sum_{k \in \Lambda} c[k] \nabla g[\ell - k] = 0 \text{ for all } \ell \in \Delta \mid \Lambda \text{ and } \hat{g}[k] = \hat{f}[k] \text{ for all } k \in \Gamma.$$  \hspace{1cm} (26)

imply $\hat{g}[k] = \hat{f}[k]$ for all $k \in \Delta$? In other words, can we uniquely extrapolate $\hat{f}$ to a larger grid $\Delta \subseteq \mathbb{Z}^2$ using the Fourier domain annihilation relation (26) given few of its samples? Note that (25) may be thought of as the limiting case of (26), where the extrapolation set $\Delta$ expands to all of $\mathbb{Z}^2$. Our numerical experiments (see Fig. 4) indicate that, given the minimal polynomial, unique recovery is in fact possible from (26) with the necessary minimum number of samples given in §4.1. We hypothesize that a sufficient condition analogous to Theorem 4 also holds in this case. However, establishing this condition is less straightforward than in the infinite dimensional setting (25) since in this context we lack the parameterization result (22). We plan to pursue this question in a future work.

Figure 4: Edge set aware recovery of synthetic data. Piecewise constant images with $n = 3$ and $n = 7$ distinct connected regions were generated with edge set given by a known minimal annihilating polynomial having $7 \times 7$ coefficients (as shown in Fig. 3). Choosing $N$ samples uniformly at random from the $7 \times 7$ center Fourier coefficients of the image, we extrapolate the image in Fourier domain to $65 \times 65$ coefficients by solving (26) in a least-squares fashion. We plot the normalized root mean square error (NRMSE) of the recovery averaged over ten trials. Pictured above the plots is the inverse DFT of a reconstruction obtained from the indicated number of samples $N$. Note that in both cases we are able to recover the image nearly exactly from the necessary minimum number of samples, $N = 4$ and $N = 8$ respectively, corresponding to the number of regions (including the background region). Below this number, we still recover the Fourier coefficients of a piecewise constant image, however one with the incorrect region amplitudes.
5 Algorithms

We rely on a two step algorithm similar to Prony’s method, which involves (1) the estimation of the edge set of an image (2) the edge set aware recovery of the image. Theorems 2–3 together guarantee the unique recovery of the edge set from the linear annihilation relation (17), while Theorem 4 guarantees the unique recovery of the full piecewise constant image given its edge set. However, these guarantees only hold under ideal conditions, making direct estimation of the signal from (17) infeasible in practice. Here we introduce robust refinements of this procedure to enable recovery of the image in the presence of noise or model-mismatch.

5.1 Step 1: Estimation of an annihilating subspace

Directly estimating the edge set via the linear system (17) in practical applications is challenging due to the following problems:

1. **Model-order selection**: Theorems 2 and 3 assume the knowledge of the degree of the minimal polynomial describing the edge set of the image. However, this is often unknown in practical applications.

2. **Noise and model-mismatch**: The available Fourier samples are often corrupted by noise in practical settings. In addition, the assumption that the underlying image is piecewise constant with edge set described by a trigonometric curve will only be approximately true in practice.

We now introduce efficient approaches to overcome the above problems. The key idea is that rather than estimating a single annihilating filter, we estimate an entire annihilating subspace—a collection of linearly independent annihilating filters which jointly encode the edge set. This can be thought of as an extension of subspace methods in spectral estimation to the 2-D setting. Specifically, we draw parallels between our algorithm and the well-known MUSIC algorithm [33]. This gives us an efficient means to address the model-order selection problem. We also introduce an efficient algorithm for the robust estimation of the annihilating subspace in §5.1.2, which address the problem of model-mismatch and noisy measurements. This scheme is related to Cadzow denoising methods [5, 12] used for robust spectral estimation. Finally, once the annihilating subspace is estimated, in §5.2 we propose recovering the piecewise constant signal in Fourier domain by constraining its extrapolated Fourier coefficients to be orthogonal to the annihilating subspace.

For convenience, we rewrite the annihilation system (17) in matrix notation. First, we fix rectangular index sets $\Gamma, \Lambda \subseteq \mathbb{Z}^2$, which represent the sampling set, and annihilating filter support, respectively, and set $N = \#(\Gamma|\Lambda)$, $M = \#\Lambda$. Define $\mathcal{T}_x$ and $\mathcal{T}_y$ to be the linear operators mapping samples $\{\hat{f}[k]\}_{k \in \Gamma}$ to matrices $\mathcal{T}_x(\hat{f}) \in \mathbb{C}^{N \times M}$ and $\mathcal{T}_y(\hat{f}) \in \mathbb{C}^{N \times M}$ which represent the discrete convolution of $\partial_x f$ and $\partial_y f$ with a filter $\{c[k]\}_{k \in \Lambda}$ restricted to the index set $\Gamma|\Lambda$. In particular, $\mathcal{T}_x(\hat{f})$ and $\mathcal{T}_y(\hat{f})$ are block Toeplitz matrices with Toeplitz blocks, where the rows of $\mathcal{T}_x(\hat{f})$ indexed by $\ell \in \Gamma|\Lambda$ are given by $\{\partial_x f[\ell - k]\}_{k \in \Lambda}$ in vectorized form, and similarly for $\mathcal{T}_y(\hat{f})$; see Fig. 5. Finally, we define the linear operator $\mathcal{T}$ as the concatenation of $\mathcal{T}_x$ and $\mathcal{T}_y$, i.e.

$$\mathcal{T}(\hat{f}) = \begin{bmatrix} \mathcal{T}_x(\hat{f}) \\ \mathcal{T}_y(\hat{f}) \end{bmatrix} \in \mathbb{C}^{2N \times M}.$$  

Using $c \in \mathbb{C}^M$ to denote a vectorized version of $\{c[k]\}_{k \in \Lambda}$, we may reproduce the annihilation system (17) by writing

$$\mathcal{T}(\hat{f}) c = 0$$

Therefore, recovering an annihilating polynomial for $f$ is equivalent to finding a vector in the nullspace of the structured matrix $\mathcal{T}(\hat{f})$, which we call the annihilation matrix.
Figure 5: Construction of the annihilation matrix. Shown in (a) are the index sets in \( \mathbb{Z}^2 \) used to construct the annihilation matrix \( \mathcal{T}(\hat{f}) \) defined in (27). Here \( \Gamma \) in is the set of Fourier sampling locations, \( \Lambda \) (in red) is the assumed annihilating filter support, and \( \Gamma|\Lambda \) (in blue) is set of all integer shifts \( \ell \) such that \( \Lambda \) shifted by \( \ell \) is contained in \( \Gamma \). The construction of \( \mathcal{T}(\hat{f}) \) is shown schematically in (b). From known samples \( \hat{f}[k] \) for \( k \in \Gamma \) we compute \( \partial_x \hat{f}[k] \) and \( \partial_y \hat{f}[k] \) by a Fourier domain multiplication. Then each row of \( \mathcal{T}(\hat{f}) \) is obtained by vectorizing a rectangular Fourier domain patch of size \( \Lambda \) from the computed samples of \( \partial_x \hat{f} \) and \( \partial_y \hat{f} \).

5.1.1 Annihilating subspace estimation

When the degree of the minimal polynomial associated with the edge set is unknown in advance—which will typically be the case—we cannot build the annihilation matrix \( \mathcal{T}(\hat{f}) \) with the correct filter support to ensure its null space is one-dimensional. For instance, if the assumed filter support \( \Lambda' \) is strictly larger than support \( \Lambda \) of the minimal filter \( \{c[k]\}_{k \in \Lambda} \), the annihilation matrix will have a larger null space. This is because any filter of the form \( (c_0 \ast d)[k] \) having Fourier support \( k \in \Lambda' \) is also an annihilating filter by the associativity of convolution. In fact, Theorem 3 says that under certain conditions on \( f \), every null space vector of \( \mathcal{T}(\hat{f}) \) is of this form. Hence, under these conditions, \( \mathcal{T}(\hat{f}) \) is rank deficient with the following bound:

**Proposition 6.** Let \( f \) be a piecewise constant image with edge set \( C = \{\mu = 0\} \), where \( \mu \) is the minimal polynomial with coefficients in \( \Lambda \). If the assumed annihilating filter support \( \Lambda' \) strictly contains \( \Lambda \), then

\[
\text{rank}[\mathcal{T}(\hat{f})] \leq \#\Lambda' - \#(\Lambda'|\Lambda). \tag{28}
\]

Furthermore, equality holds if \( f \) satisfies the conditions of Thm. 3 and if the sampling set \( \Gamma \subseteq 2\Lambda + \Lambda' \).

We call the null space of \( \mathcal{T}(\hat{f}) \) the annihilating subspace. Since identifying the minimal annihilating filter within the annihilating subspace when it has large dimension is challenging, we propose to leverage the entire annihilating subspace for both edge set estimation and image reconstruction.

We observe that the edge set is still uniquely recoverable from the annihilating subspace under ideal conditions. To see this, let \( \mathbf{D} = [d_1, \ldots, d_R] \) be any orthonormal basis for the annihilating subspace. Define \( \mu_i \) to be the trigonometric polynomial having coefficients \( \mathbf{d}_i = \text{vec}(\{d_i[k]\}_{k \in \Lambda}) \). By the preceding discussion \( \mu_i \) must always have the form \( \mu_i = \gamma_i \mu \) for \( \mu \) the minimal polynomial and \( \gamma_i \) another trigonometric polynomial. We define a function \( \overline{\mu} \) as the following sum-of-squares average

\[
\overline{\mu}(\mathbf{r}) = \sqrt{\sum_{i=1}^{R} |\mu_i(\mathbf{r})|^2} = \sqrt{\sum_{i=1}^{R} |\gamma_i(\mathbf{r})\mu(\mathbf{r})|^2}. \tag{29}
\]
Note we always have \( \overline{\mu}(r) = 0 \) when \( \mu(r) = 0 \). We can also prove the following:

**Proposition 7.** The function \( \overline{\mu} \) has at most finitely many zeros on \( [0, 1]^2/\{\mu = 0\} \).

Therefore, we can identify the edge set by finding the zeros of \( \overline{\mu}(r) \), and removing any isolated points. This method of estimating the edge set can also be interpreted as a higher dimensional analogue of the technique used in spectral MUSIC to estimate the locations of isolated spectral peaks. Recall that with spectral MUSIC one performs this estimate by computing the distance of pure exponential signals from a particular eigenspace of the signal covariance matrix. The following result shows we may also interpret \( \overline{\mu}(r) \) in an analogous way:

**Proposition 8.** Let \( r \in [0, 1]^2 \), \( \mathbf{D} \in \mathbb{C}^{M \times R} \) be an orthonormal basis of the nullspace of \( \mathcal{T}(\hat{f}) \), and define \( \mathbf{e}(r) \in \mathbb{C}^M \) by \( \mathbf{e}(r)[k] = e^{j2\pi(k,r)} \) for all \( k \in \Lambda \). Then we have

\[
\overline{\mu}(r) = \| \mathbf{D}^H \mathbf{e}(r) \|_2
\]

that is, \( \overline{\mu}(r) \) is the \( \ell^2 \)-norm of the projection of \( \mathbf{e}(r) \) onto the annihilating subspace.

While we do not directly estimate the edge set in this fashion, the function \( \overline{\mu} \) play an important role in the second step of our proposed algorithm in §5.2.

### 5.1.2 Annihilation subspace estimation in the presence of noise

When the measurements are corrupted by noise or there is model mismatch, we propose performing a denoising step to better identify the annihilating subspace. This denoising step is similar to Cadzow methods used in FRI algorithms [2, 32], as well as recently proposed algorithms for structured low-rank matrix completion problems [19, 10, 31, 21]. Given noisy Fourier samples \( b[k] = f[k] + n[k] \), \( k \in \Gamma \), where we assume \( n[k] \) is a vector of i.i.d. mean-zero Gaussian noise, we attempt find the \( f^* \) closest in Euclidean norm that satisfies the rank bound in Proposition 6:

\[
\hat{f}^* = \arg\min_{\tilde{g} \in \mathbb{C}^M, \|\tilde{g}\|_2^2 \leq \lambda} \sum_{k \in \Gamma} |\tilde{g}[k] - b[k]|^2 \quad \text{such that} \quad \text{rank}(\mathcal{T}(\tilde{g})) \leq r.
\]  

Here the parameter \( r \) is estimated according to (28). This can be viewed as a maximum likelihood estimation, where the rank constraint on \( \mathcal{T}(\tilde{g}) \) enforces our modeling assumption; see [12] for a similar approach in 1-D. Note the rank constraint in (30) is non-convex, which means in general we are only able to compute an local minimizer of (30). One common heuristic for solving problems of the type (30) is to perform a Cadzow denoising on a lifted matrix. However, as observed in [32], this procedure is slow to converge in this context. Instead, we attempt to solve (30) more directly with a variable splitting approach adopted from [19]:

\[
\min_{\tilde{g} \in \mathbb{C}^M, \|\tilde{g}\|_2^2 \leq \lambda} \sum_{k \in \Gamma} |\tilde{g}[k] - b[k]|^2 + \lambda \|\mathcal{T}(\tilde{g}) - \mathbf{T}\|_F^2
\]

where \( \lambda \) is a regularization parameter balancing data fidelity and the low-rank penalty term. This suggests the following alternating minimization scheme: set \( \hat{f}^{(0)} = b \) and for \( n \geq 1 \), solve

\[
\mathbf{T}^{(n+1)} = \arg\min_{\mathbf{T} \in \mathbb{C}^{M \times R}, \text{rank}(\mathbf{T}) \leq r} \|\mathcal{T}(\hat{f}^{(n)}) - \mathbf{T}\|_F^2
\]

\[
\hat{f}^{(n+1)} = \arg\min_{\tilde{g} \in \mathbb{C}^M, \|\tilde{g}\|_2^2 \leq \lambda} \sum_{k \in \Lambda} |\tilde{g}[k] - b[k]|^2 + \lambda \|\mathcal{T}(\tilde{g}) - \mathbf{T}^{(n+1)}\|_F^2
\]
The exact solution to subproblem (32) is easily computed by taking the singular value decomposition of $\mathcal{T}(\hat{f}(n)) = U\Sigma V^*$, and setting $T_{n+1} = U\Sigma_i V^*$, where $\Sigma_i$ is the diagonal matrix of singular values $\sigma_i$ with $\sigma_i = 0$ for all $i > r$. Also, subproblem (33) is quadratic in $\hat{g}$, and has an exact solution given by $\hat{g} = (b + T^*(T)) / (1 + \lambda \text{diag}(T^*T))$, where the operations above are understood element-wise. Here $T^*$ denotes the adjoint of $T$, which computes a weighted sum of entries of $T \in \mathbb{C}^{2N \times M}$. In particular, one can show $T^*T$ is a diagonal matrix. Due to the alternating scheme, the cost in (31) is guaranteed to monotonically decrease with updates (33). A detailed convergence analysis of the algorithm is beyond the scope of this work. We observe the algorithm converges well in practice. In all our experiments conducted in [6] a suitable solution was reached after $10 \sim 20$ iterations.

5.2 Step 2: Edge set aware image recovery

We now focus on the edge-aware recovery of the image using (24), which amounts to the Fourier domain extrapolation of the measured samples. This recovery can still be considered as “off-the-grid”, in the sense that we reconstruct the exact Fourier series coefficients of the underlying continuous domain image. We will also briefly consider the case where we solve for the image “on-the-grid” in spatial domain similar to (32).

5.2.1 Fourier domain recovery: Extrapolation by linear prediction

Suppose that we have an estimate of a basis for the annihilating subspace $D = [d_1, \ldots, d_R]$. Motivated by the Fourier domain recovery formulations (24) and (26), we pose the recovery of the Fourier domain of the signal within an arbitrary rectangular domain $\Delta \subseteq \mathbb{Z}^2$ as

$$\min_{\hat{g}[k], k \in \Delta} \sum_{i=1}^R \| \nabla \hat{g} \ast d_i \|_2^2 \quad \text{subj. to} \quad \hat{g}[k] = \hat{f}[k], k \in \Gamma,$$

(34)

i.e. we extrapolate the signal by enforcing the Fourier derivatives to be annihilated by filters in the annihilating subspace. Here we restrict the convolution to the valid region $\Delta \cap \Lambda \subseteq \mathbb{Z}^2$. When the measurements are corrupted by noise (i.e. $b[k] = \hat{f}[k] + n[k], k \in \Gamma$), we relax the equality constraint in (34) and solve instead:

$$\min_{\hat{g}[k], k \in \Delta} \sum_{i=1}^R \| \nabla \hat{g} \ast d_i \|_2^2 + \lambda \| P_{\Gamma} \hat{g} - b \|_2^2.$$

(35)

Here, $\lambda$ is a tunable parameter balancing annihilation and data fidelity, and $P_{\Gamma}$ is projection onto the sampling set $\Gamma$. Note that (35) is a linear least-squares problem, and can be solved efficiently with a conjugate gradient iterative solver, provided we have a fast method of computing matrix-vector products with the gradient of the objective in (35). Towards this end, we rewrite (35) in terms of a linear operator $Q$ acting on $\hat{g}$:

$$\min_{\hat{g}[k], k \in \Delta} \| Q \hat{g} \|_F^2 + \lambda \| P_{\Gamma} \hat{g} - b \|_2^2.$$

(36)

where $\| \cdot \|_F$ denotes the Frobenius norm. We can express $Q$ as

$$Q = [Q_{x,1}, Q_{y,1}, \ldots, Q_{x,R}, Q_{y,R}]; \quad Q_{x,i} = PC_i M_x, \quad Q_{y,i} = PC_i M_y \quad \text{for all } i = 1, \ldots, R,$$

(37)

with $M_x$ and $M_y$ representing elementwise multiplication by $-j2\pi k_x$ and $-j2\pi k_y$, respectively, $C_i$ representing discrete convolution with the annihilating filter $d_i$, and $P$ represents projection onto the index set $\Delta \cap \Lambda$. Note that since $\hat{g}$ is restricted to $\Delta$, we may take $C_i$ to be a 2-D circular convolution, provided we zero-pad the filter $d_i$ and the input $\hat{g}$ sufficiently. Thus, we can write $C_i = FS_i F^H$ where $F$ is the 2-D DFT on any rectangular grid containing the set $\Delta \cap \Lambda$, and $S_i$ is elementwise multiplication by the inverse
DFT of \(d_i\), which is given by a spatial gridding of the annihilating polynomial \(\mu_i(r)\) having coefficients \(d_i\).

Taking the gradient of the objective in (36) and setting it to zero yields

\[
(Q^*Q + \lambda P_T^*P_T)\hat{g} = P_T^*b.
\]

From (37), the operator \(Q^*Q\) can be expanded as

\[
Q^*Q = \sum_{i=1}^{R} (M_x^*C_i^*P^*P_C_iM_x + M_y^*C_i^*P^*P_C_iM_y)
\]

Unfortunately, when the number of annihilating filters is large (i.e. \(R\) is large), applying this operator directly will be computationally costly, requiring \(8R\) two-dimensional FFT operations per application. However, if the resolution of the reconstruction grid \(\Delta\) is large enough, the projection operator \(P^*P\) will be close to the identity matrix \(I\). Making the approximation \(P^*P \approx I\) in (38), we obtain

\[
Q^*Q \approx M_x^*F^*RFM_x + M_y^*F^*RFM_y.
\]

where \(R = \sum_{i=1}^{R} S_i^*S_i\) with represents pointwise multiplication by spatially gridded samples of the sum-of-squares polynomial \(\mu_2^2 = \sum_{i=1}^{R} |\mu_i|^2\), as defined in (29). Note this only requires four 2-D FFTs per iteration to evaluate, as opposed to \(8R\). Working backwards, we may show approximating \(Q^*Q\) in this manner is equivalent to minimizing

\[
\min_{{\hat{g} | k, k \in \Delta}} \|\nabla B(g)\|_{L^2}^2 + \lambda\|P_T\hat{g} - b\|_2^2
\]

where here \(B(g)\) is the function obtained by bandlimiting \(g\) to the Fourier support set \(\Delta\), and \(\| \cdot \|_{L^2}^2\) is the spatial domain \(L^2\)-norm of vector-valued functions on \([0, 1]^2\). This shows that linear prediction using the entire annihilating subspace is roughly equivalent to (24); the main difference is that the minimal polynomial \(\mu\) is replaced by the sum-of-squares average polynomial \(\mu\) defined in (29). We propose solving (39) as a surrogate for (35) when the dimension of the annihilating subspace is large. We call (39) the least-squares linear prediction (LSLP) approach.

### 5.2.2 Spatial domain recovery: Weighted total variation

We also consider solving for the discrete samples of a function \(g : [0, 1]^2 \rightarrow \mathbb{C}^2\) on a rectangular grid \(\{r_{i,j}\}\) in spatial domain:

\[
\min_{{g}} \sum_{i,j} \|\nabla g(r_{i,j}) \mu(r_{i,j})\| + \lambda\|P_T\hat{g} - b\|_2^2,
\]

This approach of solving for the image on discrete spatial grid is similar to [32], and can be thought of as a discrete approximation of (23). We also use finite differences to approximate the gradient in (40) to obtain a fast algorithm. In this case, (40) can be cast as a weighted total variation minimization, and be solved efficiently with a primal-dual algorithm, such as [48]. We call (40) the weighted total variation approach (WTV).

### 6 Experiments

In this section we demonstrate the utility of our proposed algorithm for the super-resolution of MR images. A single-coil MR acquisition restricted to a single plane with sampling locations on a uniform Cartesian grid can be modeled as the Fourier coefficients of the underlying 2-D image according to (8).
Figure 6: Super-resolution of piecewise constant phantoms from ideal Fourier samples. Recovery of Shepp-Logan phantom from $65 \times 49$ ideal low-pass Fourier samples, and a brain phantom reconstructed onto $256 \times 256$ spatial grid reconstructed from $97 \times 97$ samples. In (a)-(j), spatial domain is shown above, and Fourier domain below (in log scale). Shown in (a)&(f) are the fully sampled images, (b)&(g) is the IFFT of the zero-padded samples, (c)&(h) is a standard TV regularized recovery, (d)&(i) is the recovery using the proposed least squares linear prediction (LSLP) method, and (e)&(j) the recovery using the proposed weighted-TV (WTV) formulation. Note the the LSLP more accurately extrapolates the true Fourier coefficients of the image. The sum-of-squares polynomial used as the edge set estimate for the proposed schemes in shown (k) and (l), plotted in log scale.
First, to test the validity of the proposed scheme, we experiment on simulated data generated from piecewise constant medical phantoms. To ensure resolution-independence, we compute the samples from analytical Fourier domain expressions of the MR phantoms, as done in [18]. We collect the center Fourier samples located within a rectangular window of size \((4K+1) \times (4L+1)\), and solve for an annihilating filter of size \((2K+1) \times (2L+1)\), i.e. the maximum possible filter size allowable, using the procedure outlined in Section 5. We then recover the image using the proposed algorithms in §5, and compare against standard total variation (TV) regularized recovery. In all cases we tune the regularization parameter \(\lambda\) to optimize the signal-to-noise ratio (SNR), defined as \(\text{SNR} = 20 \log_{10}(\|x_0\|_2/\|x - x_0\|_2)\) where \(x\) is the reconstructed image, and \(x_0\) is the ground truth image.

![Image](image.png)

**Figure 7:** Super-resolution of a single-coil MR brain image from retrospectively undersampled data. We reconstruct the image from 129 \(\times\) 129 center Fourier samples taken from a 256 \(\times\) 256 acquisition reconstructed onto a 256 \(\times\) 256 spatial grid. The original fully sampled data is shown in (a), and zoomed in (b). The zero-padded inverse DFT of the undersampled data is shown in (c). A TV regularized recovery is shown in (e), and the proposed weighted \(f^2\) recovery scheme is shown in (f) which uses the sum-of-squares polynomial shown in (d) as an edge set estimate. The proposed scheme shows an improvement in SNR over the other methods, and is able to resolve more of the fine features of the image.

The results of this first experiment are shown in Fig. 6. We note that our proposed scheme is able to recover the Shepp-Logan medical phantom with high accuracy from as few as 65 \(\times\) 49 center k-space samples, which corresponds to a 20-fold undersampling. A TV-regularized recovery from the same samples
Figure 8: Super-resolution of a single-coil MR brain image from retrospectively undersampled data. We reconstruct the image from 129 × 129 center k-space samples taken from a 256 × 256 acquisition reconstructed onto a 256 × 256 spatial grid. The original fully sampled data is shown in (a), and zoomed in (b). The zero-padded inverse DFT of the undersampled data is shown in (c). A TV regularized recovery is shown in (e), and the proposed LSLP recovery scheme is shown in (f) which uses the sum-of-squares polynomial shown in (d) as an edge set estimate. The proposed scheme shows an improvement in SNR over the other methods, and is able to resolve more of the fine features of the image.

shows significant edge blurring, and broadening of image features. We observe a similar benefit in the case of a brain phantom having a more complicated edge set geometry. In this case, the WTV approach visually appears to show some advantages over the weighted least squares approach, in that it is better able to recovery some of the finer image features, and suppress ringing artifacts. However, the LSLP is better able to recover the true Fourier coefficients, as measured by the SNR.

We also perform another set of experiments validating the method on real MR data; see Figures 7 and 8. Here we retrospectively undersampled the data to 129×129 center Fourier samples taken from a 256×256 acquisition and extrapolate to 256×256 Fourier samples. Note we expect there to be significant model-mismatch in this case, since the underlying image is not exactly piecewise constant, plus there is noise in the acquisition. Both of these complications make estimating the true model order challenging. Hence, as a rule of thumb, we set the model order \( r \) in (30) to half the number of the coefficients in the assumed size of the annihilating filter, and apply the denoising step in [§5.1.2](#) before estimating the...
annihilating subspace. Empirically we find this procedure is still able to accurately estimate the edge set, as shown in Figs. 7(d) and 8(d). Additionally, for real data, we find our proposed LSLP approach to be superior to the WTV approach, hence we restrict our comparisons to this case. One of the main benefits of the proposed scheme appears to be its ability to suppress artifacts due to noise relative to a TV-regularized recovery. Again, we notice less distortion in the fine features of the image relative to TV, as well.

7 Discussion

In this work, we have developed sampling guarantees for the unique recovery of a continuous domain piecewise constant image from few of its low-pass Fourier samples, provided we assume the edge set of the image is described by the zero set of a multivariable trigonometric polynomial. This is achieved by proving the distributional derivatives of such an image satisfy a linear annihilation relation in Fourier domain. One of our main contributions is to prove that, under ideal conditions, unique recovery of the image is possible from the annihilation relation.

We also proposed a two step algorithm for the recovery of images, which extrapolates the known Fourier samples of the image in Fourier domain. We argue such a recovery scheme deserves to be called “off-the-grid”, since we recover the exact Fourier coefficients of the underlying continuous domain image, and at no point do we discretize in spatial domain.

In the first step of the algorithm, we estimate a collection of annihilating filters from the low-pass samples, which jointly encode the edge set of the image. In the second step, we perform a least squares linear prediction using these filters to extrapolate the signal Fourier domain. We show the proposed scheme has advantages over traditional single-image super-resolution schemes in MRI by suppressing noise artifacts and preserving strong image edges.

While we focused on apply the proposed scheme to super-resolution MRI, we note that this work has potentially wide-ranging applications beyond those presented here. For instance, the first stage our algorithm estimates a continuous domain representation of the edge set of an image, which can be useful in image segmentation problems. Additionally, the super-resolution problem we study is not limited to MRI, and will work in any setting where one has a good estimate of the low-pass Fourier samples of image, such as in the image zoom problem studied in [32].

In future work, we plan to resolve the conjectures put forth in §2.4, which would extend our sampling theorems to piecewise constant images with edge set given by an arbitrary trigonometric curve. We also hope to give further insight into the robustness of this scheme with respect to noise and model mismatch, especially with regard to the rank properties of the annihilation matrix.

Finally, we observe that most of the theory presented here extends directly to higher dimensions. For example, in three dimensions, we could consider piecewise constant signals whose gradient is supported on a two-dimensional surface given as the zero set of a trigonometric polynomial in three variables. However, the computational burden in this case will be much more significant. For instance, the denoising procedure in §5.1.2 would require several singular value decompositions of a very large matrix. Therefore a goal of future research is to come up with efficient algorithms to further enable the recovery of higher dimensional signals.

A Algebraic Properties of Trigonometric Polynomials and Curves

Here we develop the algebraic properties of trigonometric polynomials and trigonometric curves in some detail. We will slightly expand our definition of a trigonometric polynomial (9) to include those with frequencies $k$ on a shifted lattice $\mathbb{Z}^2 + \sigma$ for some fixed $\sigma = (\sigma_1, \sigma_2)$ with $\sigma_i \in \{0, \frac{1}{2}\}$. This is to ensure
the results below extend to trigonometric polynomials having an even number of coefficients in one or both dimensions.

We define the degree of \( \mu \), denoted as \( \text{deg}(\mu) = (K, L) \), to be the linear dimensions of the smallest rectangle \( R \) whose closure contains the frequency support set, and we say \( \mu \) is centered if the frequency support rectangle \( R \) is symmetric about origin \((k, l) = (0, 0)\). For example, 

\[
\mu_1(x, y) = 4\sin(2\pi x)\cos(4\pi y) = je^{j2\pi(-x-2y)} + je^{j2\pi(x-2y)} - je^{j2\pi(-x+2y)} - je^{j2\pi(x+2y)}
\]

has frequency support contained in the rectangle \([-1, 1] \times [-2, 2]\), hence is centered with degree \((2, 4)\). As another example, the trigonometric polynomial

\[
\mu_2(x, y) = 2e^{-2\pi x} + 1 + e^{2\pi x} - e^{2\pi(x+y)}
\]

has frequency support contained in \([-1, 1] \times [0, 1]\), hence degree \((2, 1)\), but is not centered. However, the trigonometric polynomial \(e^{-\pi y}\mu_2(x, y)\) is centered since it has frequency support in \([-1, 1] \times [-\frac{1}{2}, \frac{1}{2}]\). Note \(\mu_2\) and \(e^{-\pi y}\mu_2(x, y)\) have the same degree and the same zero set, since these properties are invariant under multiplication by a phase factor.

For any trigonometric polynomial \( \mu \) with \( \text{deg}(\mu) = (K, L) \) we can associate a unique complex polynomial \( \mathcal{P}[\mu] \) in \( \mathbb{C}[z, w] \) by making the substitutions \( e^{j2\pi x} \mapsto z, e^{j2\pi y} \mapsto w \), and multiplying by powers of \( z \) and \( w \) such that \( \mathcal{P}[\mu] \) has degree \( K \) as a polynomial in \( z \), and degree \( L \) as a polynomial in \( w \). Similarly, we may define an inverse mapping acting on polynomials \( q(z, w) \in \mathbb{C}[z, w] \) as \( \mathcal{P}^{-1}[q](x, y) := e^{-j2\pi Sx}e^{-j2\pi Ty}q(e^{j2\pi x}, e^{j2\pi y}) \), where \( S, T \) are chosen so that the resulting trigonometric polynomial is centered. For example, using the same \( \mu_1 \) and \( \mu_2 \) as above, we have

\[
\mathcal{P}[\mu_1](z, w) = zw^2(jz^{-1}w^{-2} + jzw^{-2} - jz^{-1}w^2 - zw^2) = j(1 + z^2 - w^4 - z^2w^4),
\]

\[
\mathcal{P}[\mu_2](z, w) = z(2z^{-1} + 1 + z - zw) = 2 + z + z^2 - z^2w
\]

and \( \mathcal{P}^{-1}[\mathcal{P}[\mu_1]](x, y) = \mu_1(x, y) \) whereas \( \mathcal{P}^{-1}[\mathcal{P}[\mu_2]](x, y) = e^{-j\pi x}\mu_2(x, y) \).

We also define \( E(x, y) = (e^{j2\pi x}, e^{j2\pi y}) \) for all \((x, y) \in [0, 1]^2\), which is a continuous bijective mapping of \([0, 1]^2\) onto the complex unit torus \( \mathbb{C}^2 = \{(z, w) \in \mathbb{C}^2 : |z| = |w| = 1\} \). Observe that \( \mu(x, y) = 0 \) if and only if \( \mathcal{P}[\mu](z, w) = 0 \) for \((z, w) = E(x, y) \), so that \( \mu = 0 \) on \( C \subseteq [0, 1]^2 \) if and only if \( \mathcal{P}[\mu] = 0 \) on \( E(C) \). This shows we may study the algebraic properties of trigonometric polynomials and their zero sets by studying their corresponding complex polynomials. Accordingly, for trigonometric polynomials \( \mu \) and \( \nu \) we will say \( \mu \) divides \( \nu \), denoted \( \mu \mid \nu \), when \( \mathcal{P}[\mu] \) divides \( \mathcal{P}[\nu] \) as polynomials in \( \mathbb{C}[z, w] \), or equivalently, if \( \nu = \mu \gamma \) for some trigonometric polynomial \( \gamma \); \( \mu \) and \( \nu \) are said to have a common factor if there is a \( \gamma \) with positive degree such that \( \gamma \mid \mu \) and \( \gamma \mid \nu \); and we say \( \mu \) is irreducible when \( \mathcal{P}[\mu] \) is irreducible as a polynomial in \( \mathbb{C}[z, w] \).

We will need the following well-known result regarding the maximum number of intersections of algebraic curves (see e.g. [35]):

**Theorem 5** (Bézout’s Theorem). Let \( p \) and \( q \) be non-constant polynomials in \( \mathbb{C}[z, w] \) with \( \text{deg}(f) = d_1 \) and \( \text{deg}(g) = d_2 \). If \( p \) and \( q \) have no common factor then the system of equations \( p(z, w) = q(z, w) = 0 \) has at most \( d_1 d_2 \) solutions.

Note that as one consequence of Bézout’s theorem, if \( f(z, w) \) is an irreducible polynomial and \( g(z, w) \) is any other polynomial such that \( f(z, w) = g(z, w) = 0 \) has infinitely many solutions, then \( f \) must divide \( g \).

We may readily translate these results to trigonometric polynomials \( \mu \) and \( \nu \) by mapping to the polynomials \( p = \mathcal{P}[\mu] \) and \( q = \mathcal{P}[\nu] \), and mapping back again. If \( \text{deg}(\mu) = (K, L) \), then \( f = \mathcal{P}[\mu] \) has total degree at most \( K + L \) as a polynomial in \( \mathbb{C}[z, w] \), and so we have the following result that we will also call “Bézout’s theorem”:
Theorem 6. Let $\mu$ and $\nu$ be trigonometric polynomials with $\deg(\mu) = (K, L)$ and $\deg(\nu) = (K', L')$. If $\mu$ and $\nu$ have no common factor then the system of equations $\mu(x, y) = \nu(x, y) = 0$ has at most $(K + L)(K' + L')$ solutions in $[0, 1]^2$.

Likewise, if the zero sets of trigonometric polynomials $\mu$ and $\nu$ have infinite intersection, and $\mu$ is irreducible, this implies $\mu \mid \nu$.

A.1 Existence of a minimal polynomial (Proof of Proposition 2)

Here we prove a slightly stronger version of Proposition 2, which shows the minimal polynomial satisfies an additional unique divisibility property. Recall that we say the zero set of a trigonometric polynomial $C = \{\mu = 0\}$ is a trigonometric curve if $C$ infinite and has no isolated points, i.e. if every connected component of $C$ is one-dimensional.

Proposition 9. For every trigonometric curve $C$ there is unique (up to scaling) real-valued trigonometric polynomial $\mu_0$ such that $C = \{\mu_0 = 0\}$ and for any other trigonometric polynomial $\mu$ with $\mu = 0$ on $C$ we have $\deg(\mu_0) \leq \deg(\mu)$ componentwise, and $\mu_0 \mid \mu$.

Proof. Let $\mu$ be any trigonometric polynomial with $C = \{\mu = 0\}$. Set $p = \mathcal{P}[\mu] \in \mathbb{C}[z, w]$. Then $p$ has a unique factorization (up to scaling) into irreducible factors: $p = p_1^{r_1} \cdots p_r^{r_r}$. Create a polynomial $p_0 = p_1^{r_1} \cdots p_r^{r_r}$, $k \leq r$, omitting any factors $p_i$ whose zero sets do not intersect $\mathbb{C}^2$, or intersect at only finitely many points. Define $\mu_0 = \mathcal{P}^{-1}[p_0]$, which by construction satisfies $\mu_0 = 0$ on $C/A$ where $A$ is possibly some finite set. However, by continuity, $\mu_0 = 0$ on all of $C$ because $C$ contains no isolated points by assumption. We will show $\mu_0$ is the desired minimal polynomial.

Note the algebraic curve $X = \{p_0 = 0\}$ is the union of irreducible curves $X_j = \{p_{ij} = 0\} \subseteq \mathbb{C}^2$. Define $C_j = E^{-1}(X_j \cap \mathbb{C}^2)$, then likewise we have a decomposition of $\mu_0$ as the union of the curves $C_j$. If $\mu$ is any other trigonometric polynomial with $C = \{\nu = 0\}$, then $\nu = 0$ on each $C_j$, and for $q = \mathcal{P}[\nu]$ we have $q = 0$ on each infinite set $E(C_j) = X_j \cap \mathbb{C}^2$. Therefore by Bézout’s inequality, $q = \mathcal{P}[\nu]$ must factor as $q = p_1^{m_1} \cdots p_r^{m_r}$ for some $r \in \mathbb{C}[z, w]$, hence $\deg(p_0) \leq \deg(q)$ and $p_0 \mid q$, which implies $\deg(\mu_0) \leq \deg(\nu)$ and $\mu_0 \mid \nu$.

We now show $\mu_0$ is real-valued under an appropriate choice of scaling. Without loss of generality assume $Re[\mu_0] \neq 0$. Since $Re[\mu_0] = \frac{1}{2}(\mu_0 + \mu_0^*)$ and $\mu_0$ is centered, we have $\deg(Re[\mu_0]) \leq \deg(\mu_0)$. Also, since $\{\mu_0 = 0\} \subseteq \{Re[\mu_0] = 0\}$ we have $\mu_0|Re[\mu_0]$ and so $Re[\mu_0] = c \cdot \mu_0$ for some $c \in \mathbb{C}$ by the uniqueness of $\mu_0$. Therefore $\frac{1}{2}\mu_0$ is real-valued.

Finally, to show uniqueness suppose $\nu_0$ is another centered trigonometric polynomial satisfying $C = \{\nu_0 = 0\}$ and $\deg(\nu_0) = \deg(\mu_0)$. Then $\mu_0 \mid \nu_0$, and so by degree constraints, $\nu_0 = c \cdot \mu_0$ for some $c \in \mathbb{C}$. 

A.2 Additional properties of minimal polynomials

Our uniqueness theorems also make use of the following result, which shows the zero set of the minimal polynomial and the zero set of its gradient have at most a finite intersection:

Proposition 10. Let $C$ be a trigonometric curve with minimal polynomial $\mu$. Then $\nabla \mu = 0$ for at most finitely many points on $C$.

Proof. Note the partial derivatives $\partial_x \mu$ and $\partial_y \mu$ are trigonometric polynomials with $\deg(\partial_x \mu) \leq \deg(\mu)$ and $\deg(\partial_y \mu) \leq \deg(\mu)$. First consider the case where $\mu$ is irreducible. If $\{\nabla \mu = 0\}$ and $C$ have an infinite intersection, then by Bézout’s theorem, $\mu$ must divide $\partial_x \mu$ and $\partial_y \mu$, and so by degree considerations, $\partial_x \mu = A \mu$ and $\partial_y \mu = B \mu$ for some non-zero constants $A, B$. By comparing coefficients, we see this is only possible if $\mu$ is the zero polynomial.
The general case follows by induction on the number of irreducible factors in \( \mu \): Suppose \( \mu = \mu_1 \cdots \mu_n \mu_0 \) where each \( \mu_i \) is a distinct irreducible factor (that \( \mu \) must have this form follows from the proof of Prop. [2]). Let \( \nu = \mu_1 \cdots \mu_{n-1} \), then we may write \( C = C_1 \cup C_2 \) where \( C_1 = \{ \nu = 0 \} \) and \( C_2 = \{ \mu_n = 0 \} \). By the product rule \( \nabla \mu = \mu_n \nabla \nu + \nu \nabla \mu_n \). Therefore, \( \nabla \mu = 0 \) on \( C_1 \) if only if \( \mu_n = 0 \) or \( \nabla \nu = 0 \) on \( C_1 \). However, \( \{ \mu_n = 0 \} \cap C_1 \) is finite by Bézout’s theorem, and \( \{ \nabla \nu = 0 \} \cap C_1 \) is finite by inductive hypothesis, hence \( \{ \nabla \mu = 0 \} \cap C_1 \) is finite. A similar argument shows \( \{ \nabla \mu = 0 \} \cap C_2 \) is finite, which proves the induction.

The following proposition shows that minimal polynomials must always change sign across an infinite segment of its zero set.

**Proposition 11.** Let \( C \) be a trigonometric curve with minimal polynomial \( \mu \). If \( \Omega_1 \) and \( \Omega_2 \) are distinct connected components of \( [0, 1]^2/C \) whose boundaries share an infinite segment of \( C \), then the sign of \( \mu \) must be different on \( \Omega_1 \) and \( \Omega_2 \).

**Proof.** Let \( \Omega_1 \) and \( \Omega_2 \) be any two connected components of \( [0, 1]^2/C \) whose boundaries share an infinite segment \( S \subseteq C \), and suppose, without loss of generality, \( \mu \) is positive on both \( \Omega_1 \) and \( \Omega_2 \). Then for all points in \( S \), \( \mu \) has a local minimum, and since \( \mu \) is smooth this implies \( \nabla \mu = 0 \) on all of \( S \), which violates Proposition 10. Therefore \( \mu \) must have different signs on \( \Omega_1 \) and \( \Omega_2 \).

Note that this result puts certain topological constraints on the geometry of trigonometric curves. For instance, if \( C \) is any trigonometric curve, and \( r_0 \in C \), then in any sufficiently small neighborhood of \( r_0 \), the set \( C \setminus \{ r_0 \} \) must consist of an even number of branches, since otherwise the minimal polynomial would violate the parity rule in Prop. 11. This implies, for example, that within the region \([0, 1]^2\) trigonometric curves must always be closed. It also rules out such features such as triple junctions, where locally three curve segments terminate at a common intersection point.

### A.3 Bound on Connected Components (Proof of Proposition 3)

Let \( C \) be a trigonometric curve of degree \((K, L)\), i.e. the minimal polynomial \( \mu \) for \( C \) has degree \((K, L)\), and let \( \Omega_1, \ldots, \Omega_n \) be the connected components of \([0, 1]^2/C\). For each \( i = 1, \ldots, N \), \( \mu \) is continuous with \( \mu \neq 0 \) on \( \Omega_i \), \( \mu = 0 \) on \( \partial \Omega \), and \( \Omega = \Omega \cup \partial \Omega \) is compact, therefore \( \mu \) must attain its maximum or minimum at some point \( r_i \) in the interior \( \Omega \), where \( \nabla \mu (r_i) = 0 \). Therefore the number of components \( n \) is bounded by the cardinality of \( \{ \nabla \mu = 0 \} = \{ \partial_x \mu = 0 \} \cap \{ \partial_y \mu = 0 \} \). When \( \{ \nabla \mu = 0 \} \) is finite, the result follows by Bézout’s theorem since \( \deg(\partial_x \mu), \deg(\partial_y \mu) \leq (K, L) \).

If \( \{ \nabla \mu = 0 \} \) is not finite, i.e. \( \{ \partial_x \mu = 0 \} \) and \( \{ \partial_y \mu = 0 \} \) share a common component, we may obtain the same bound by a standard perturbation argument, which we summarize here (see, e.g., the appendix for further details).

Let \( v \in \mathbb{R}^2 \) be a sufficiently small vector. By continuity, the function \( \tilde{\mu}(r) = \mu(r) - \langle r, v \rangle \) also attains a maximum or minimum in each \( \Omega_i \), and \( \nabla \tilde{\mu} = \nabla \mu - v \) vanishes at these points, which shows \( n \) is also bounded by the cardinality of \( \{ \nabla \mu = v \} \). Finally, note that since \( \nabla \mu \) is a smooth map whose domain and range have the same dimension, for a generic vector \( v \) the set \( \{ \nabla \mu = v \} \) is finite by the regular level set theorem (Cor. 8.10, [22]), and the result again follows by Bézout’s theorem.

We note that the above bound appears to be an overestimate. Empirically, we find that the maximum number of connected components of the complement to a trigonometric curve of degree \((K, L)\) never exceeds \(KL\) (in the topology of the torus \( T^2 = (\mathbb{R}/\mathbb{Z})^2 \)). Moreover, this bound is only obtained in artificial cases, such as when the minimal polynomial is \( \mu(x, y) = \sin(\pi K x) \sin(\pi L y) \), whose zero set is the union of \( K \) vertical lines and \( L \) horizontal lines.
B Proofs of Theorems

B.1 Proof of Theorem 1

Suppose \( f = 1_{\Omega} \), with \( \Omega \) simply connected and \( \partial \Omega = \{ \mu = 0 \} \). Let \( d[k], k \in \Lambda \) be any non-trivial set of coefficients satisfying

\[
\sum_{k \in \Lambda} d[k] \nabla f[\ell - k] = 0, \quad \text{for all } \ell \in 2\Lambda.
\]  

We will show the coefficients \( d[k] \) must be a scalar multiple of \( c[k] \).

Let \( \eta \) denote the trigonometric polynomial having coefficients \( d[k] \), observe that the sums in (41) represent the Fourier coefficients of \( \eta \nabla f \) evaluated at \( \ell \in 2\Lambda \). Therefore, (41) can be expressed in spatial domain as

\[
P_{2\Lambda}(\eta \nabla f) = 0
\]

where \( P_{2\Lambda} \) is Fourier projection onto index set \( 2\Lambda \). If we let \( \varphi \) be any smooth test field, then by duality

\[
\langle P_{2\Lambda}(\eta \nabla f), \varphi \rangle = \langle \eta \nabla f, P_{2\Lambda} \varphi \rangle = 0,
\]

or, equivalently,

\[
\langle \eta \nabla f, \psi \rangle = 0, \quad \text{for all } \psi \in B_{2\Lambda}^2,
\]

where \( B_{2\Lambda}^2 \) denotes the space all smooth vector fields \( \varphi = (\varphi_1, \varphi_2) \) whose components \( \varphi_1, \varphi_2 \) are bandlimited to \( 2\Lambda \). Since \( f = 1_{\Omega} \), using (14) we have

\[
\oint_{\partial \Omega} \eta(n \cdot \mathbf{n}) ds = 0, \quad \text{for all } \psi \in B_{2\Lambda}^2.
\]  

Because \( \partial \Omega \) is a level set of \( \mu \), the outward normal is given pointwise by \( \mathbf{n}(r) = \pm \nabla \mu(r)/|\nabla \mu(r)| \) for all \( r \in \partial \Omega \), except when \( \nabla \mu(r) = 0 \), but this can only happen for at most finitely many \( r \in \partial \Omega \) by Prop. 10. Without loss of generality, assume \( \mu < 0 \) on \( \Omega \). Since \( \mu \) is a minimal polynomial, by Prop. 11 it must change sign across the any infinite segment of \( \partial \Omega \), hence \( \mu > 0 \) on \( \Omega^C \), so that \( \mathbf{n}(r) = \nabla \mu(r)/|\nabla \mu(r)| \) for all \( r \in \partial \Omega \). Setting \( \psi = \eta^* \nabla \mu \in B_{2\Lambda}^2 \) we have

\[
\oint_{\partial \Omega} |\eta|^2 |\nabla \mu| ds = 0.
\]

which is possible if and only if \( \eta = 0 \) on \( \partial \Omega \) since \( |\nabla \mu(r)| > 0 \) for at most finitely many \( r \in \partial \Omega \). Therefore \( \eta \) must be a scalar multiple of \( \mu \) by the uniqueness of minimal polynomials.

B.2 Proof of Theorems 2 and 3

Assume the \( \{d[k]\}_{k \in \Lambda} \) is another solution of the system, and \( \eta \) denote the associated trigonometric polynomial. Following the same steps as in the proof of Theorem 1 we may show

\[
\langle \eta \nabla f, \psi \rangle = 0, \quad \text{for all } \psi \in B_{2\Lambda}^2,
\]

Since \( f = \sum_{i=1}^n a_i 1_{\Omega_i} \), from (14) this gives

\[
\sum_{i=1}^n a_i \oint_{\partial \Omega_i} \eta(n \cdot \mathbf{n}) ds = 0, \quad \text{for all } \psi \in B_{2\Lambda}^2.
\]  

(43)
Recall we assume \( \partial \Omega_i = \{ \mu_i = 0 \} \) for all \( i = 1, \ldots, N \), with \( \mu_i \) minimal, and \( \mu = \mu_1 \cdots \mu_N \). Note that \( \cup_i \partial \Omega_i = \{ \mu = 0 \} \), and since we assumed the boundaries to be pairwise disjoint, \( \mu \) is also the minimal polynomial. Define \( \tilde{\mu}_i = \mu / \mu_i \). Then by the product rule

\[
\nabla \mu = \sum_{i=1}^{N} \tilde{\mu}_i \nabla \mu_i.
\]

Define the vector field

\[
\psi_j = \eta^* \tilde{\mu}_j \nabla \mu_j \quad \text{for all} \quad j = 1, \ldots, n.
\]

Note the components of \( \psi_j \) are trigonometric polynomials in \( B_{2\Lambda} \), hence \( \psi_j \in B_{2\Lambda}^2 \) for all \( j = 1, \ldots, n \). Since \( \tilde{\mu}_j = 0 \) on \( \partial \Omega_i \) for all \( i \neq j \), substituting \( \psi_j \) into (43) gives

\[
0 = \sum_{i=1}^{n} a_i \int_{\partial \Omega_i} |\eta|^2 \tilde{\mu}_j (\nabla \mu_j \cdot n) \, ds = a_j \int_{\partial \Omega_j} |\eta|^2 \tilde{\mu}_j (\nabla \mu_j \cdot n) \, ds = a_j \int_{\partial \Omega_j} |\eta|^2 (\nabla \mu \cdot n) \, ds.
\]

where in the last step we used \( \nabla \mu = \tilde{\mu}_j \nabla \mu_j \) on \( \partial \Omega_j \) by (44). Since we also assume each \( \Omega_i \) is connected, and the curves \( \partial \Omega_i \) do not intersect, we know \( \mu \) has fixed sign on \( \Omega_i \), and so the outward normal on each \( \partial \Omega_j \) is given by \( n = \nabla \mu / |\nabla \mu| \), up to a fixed choice of sign. Hence we have

\[
a_j \int_{\partial \Omega_j} |\eta|^2 |\nabla \mu| = 0, \quad \text{for all} \quad j = 1, \ldots, n,
\]

which implies \( \eta \) vanishes on each \( \partial \Omega_j \). Since \( \mu \) is the minimal polynomial, this is only possible if \( \eta \) is a scalar multiple of \( \mu \), which proves the claim in Theorem 2.

The proof of Theorem 3 is identical to the above, except we exchange the index set \( 2\Lambda \) with \( \Lambda \ast \Lambda' \). This allows us to make the same choice of \( \psi_j \) as in (45), which again implies \( \eta \) vanishes on each \( \partial \Omega_j \). Hence by Prop. 2 we have \( \eta = \mu \gamma \), where \( \gamma \in B_{\Lambda \ast \Lambda} \).

B.3 Proof of Proposition 5

Suppose \( g \in L^1([0,1]^2) \) satisfies \( \mu \nabla g = 0 \) and let \( \Omega \) be any connected component of \( \{ \mu = 0 \}^C \). Let \( \varphi \) be any arbitrary smooth test field with support \( K \) strictly contained within \( \Omega \). Since \( \mu = 0 \) only on \( \partial \Omega \), and \( \mu \) is smooth, the field \( \psi = \varphi / \mu \) is smooth, and we have

\[
\langle \mu \nabla g, \psi \rangle = \langle \nabla g, \varphi \rangle = 0.
\]

which shows the distribution \( \nabla g \) restricted to \( \Omega \) is identically zero. We now show this implies \( g \) restricted to \( \Omega \) is a constant function almost everywhere.

Let \( \phi_\epsilon \) be any smooth approximation of the identity, i.e. for all \( \epsilon > 0 \) define \( \phi_\epsilon(r) := \epsilon^2 \phi(\epsilon r) \) where \( \phi \) is any smooth function satisfying \( \phi \geq 0 \), \( \int \phi(r) \, dr = 1 \), \( \text{supp}(\phi) \subseteq \{ |r| \leq 1 \} \). Let \( \Omega' \) be any closed connected set inside \( \Omega \). Then there exists an \( \epsilon' > 0 \) such that \( \text{dist}(\Omega', \Omega) < \epsilon' \). Fix \( \epsilon < \epsilon' \). Since \( \nabla g \ast \phi_\epsilon \) is a smooth function, \( \nabla g|_\Omega = 0 \), and \( \text{supp}(\phi_\epsilon) \subseteq \Omega \) we have

\[
\nabla (g \ast \phi_\epsilon)(r) = (\nabla g \ast \phi_\epsilon)(r) = 0 \quad \text{for all} \quad r \in \Omega'
\]

which implies \( g \ast \phi_\epsilon \) is constant on \( \Omega' \). By properties of approximations of identity, \( g \ast \phi_\epsilon \to g \) pointwise almost everywhere as \( \epsilon \to 0 \), which shows \( g \) is the limit of constant functions on \( \Omega' \), and therefore itself constant on \( \Omega' \). Since \( \Omega' \subseteq \Omega \) was arbitrary, we may conclude \( g \) restricted to \( \Omega \) is a constant function almost everywhere, proving the claim.
B.4 Proof of Theorem 4

Suppose \( g \) is another solution, and set \( h = g - f \). We have
\[
\mu \nabla h = 0, \quad \text{and} \quad \hat{h}[k] = 0 \quad \text{for all} \; k \in \Gamma \supseteq \Lambda.
\]  
(46)

The proof amounts to showing that the only feasible solution to (46) is \( h = 0 \). By the first constraint and Prop. 5 we see that \( h \) is almost everywhere a piecewise constant function:
\[
h = \sum_{i=1}^{N} b_i 1_{\Omega_i},
\]  
(47)

where the constants \( b_i \) are unknown, and \( \Omega_i \) are the connected components of \( \{ \mu = 0 \}^C \). From \( \hat{h}[k] = 0, k \in \Gamma, \) we also have \( \hat{\nabla} h[k] = 0, k \in \Gamma, \) and so
\[
\hat{\nabla} h[k] = \sum_{i} b_i \nabla 1_{\Omega_i}[k] = \sum_{i} b_i \int_{\partial \Omega_i} e^{-j2\pi k \cdot r} n \ ds = 0, \quad \text{for all} \; k \in \Gamma.
\]

By linearity, this implies
\[
\sum_{i=1}^{N} b_i \int_{\partial \Omega_i} \varphi \cdot n \ ds = 0,
\]
for all vector fields \( \varphi \in B^2_\Gamma \). The proof now follows from the same argument in the proof to Theorem 2 (with \( \eta = 1 \)): Using the same notation as in (44), we may choose \( \varphi = \tilde{\mu}_j \nabla \mu_j \in B^2_\Lambda \subseteq B^2_\Gamma \), which implies
\[
b_j \int_{\partial \Omega_j} |\nabla \mu| = 0, \quad \text{for all} \; j = 1, \ldots, n,
\]
and so \( b_i = 0 \) for all \( i = 1, \ldots, N \), and hence \( h = 0 \) proving the claim.

B.5 Proof of Proposition 6

Let \( \mu \) be the minimal polynomial with coefficients \( \{c[k]\}_{k \in \Lambda} \), and for every \( \ell \in \Lambda' | \Lambda \) define the filter \( c_\ell \) supported in \( \Lambda' \) by
\[
c_\ell[k] = \begin{cases} 
c[k - \ell] & \text{if} \; k - \ell \in \Lambda \\
0 & \text{elsewhere}, \end{cases}
\]
that is, \( c_\ell \) is a shift of the minimal polynomial coefficients within the larger index set \( \Lambda \). By the Fourier shift theorem, \( c_\ell \) are the Fourier coefficients of \( e^{j2\pi \langle \ell, r \rangle} \mu(r) \). Therefore, \( c_\ell \) is also an annihilating filter since the zero sets of \( \mu(r) \) and \( e^{j2\pi \langle \ell, r \rangle} \mu(r) \) coincide. This implies \( T(\hat{f})c_\ell = 0 \) for all \( \ell \in \Lambda'|\Lambda \). Also, note that the vectors \( c_\ell \) are linearly independent because the location of their non-zero entries are distinct. Hence the nullspace of \( T(\hat{f}) \) has dimension at least \( \#(\Lambda'|\Lambda) \), which gives the desired rank bound.

Finally, supposing \( f \) satisfies the conditions of Theorem 2 and \( \Gamma \supseteq 2\Lambda' * \Lambda \) so that \( \Gamma|\Lambda \supseteq \Lambda * \Lambda_0 \), Theorem 3 shows that if \( d \) is any null space vector, then \( d \) are the Fourier coefficients of \( \mu \gamma \), for some trigonometric polynomial \( \gamma \) having coefficients in \( \Lambda'|\Lambda \), which shows that \( d \) belongs to the span of the \( c_\ell \), proving the claim.
Let \( r_0 \in \{ \mu_0 = 0 \}^C \). Then \( \overline{\mu}(r_0) = 0 \) if and only if \( \gamma_i(r_0) = 0 \) for all \( i = 1, \ldots, R \), where \( R = \#(\Lambda|\Lambda_0) \).
We will show the set \( P = \cap_{i=1}^R \{ \gamma_i = 0 \} \) is at most finite. Note that if \( P \) is infinite, by Bézout’s theorem, each \( \gamma_i \) must have a common non-constant factor \( \phi \), and so \( \gamma_i = \rho_i \phi \), where each \( \rho_i \) is a trigonometric polynomial with degree strictly smaller than \( \gamma_i \). This means each \( \rho_i \in B_{\Lambda} \) has coefficients supported within some index set \( \Lambda' \subseteq \Lambda|\Lambda_0 \) with \( \Lambda' \neq \Lambda|\Lambda_0 \). Since the coefficient vectors \( \rho_i \) are orthonormal, so are the \( \mu_i \) as functions in \( L^2([0,1]^2) \):
\[
\langle \mu_i, \mu_j \rangle = \int_{[0,1]^2} \rho_i(r) \cdot \rho_j^*(r) |\phi(r)|^2 \mu(r) \, dr = \delta_{i,j}.
\]
where \( \delta_{i,j} = 0 \) if \( i \neq j \) and \( \delta_{i,i} = 1 \) if \( i = j \). However, the above weighted integral also defines an inner product in which the \( \{ \rho_i \}_{i=1}^R \) are mutually orthogonal, and hence linearly independent. But this is impossible, since the \( \rho_i \) are contained in a space of dimension less than \( R \). Therefore the \( \gamma_i \) have no common factors, which by Bézout’s theorem implies the intersection of their zero sets is at most finite.

**B.7 Proof of Proposition 8**

**Proof.** Note that \( e(r) = P_{\Lambda} F\{ \delta_r \} \), where \( \delta_r \) is a Dirac delta centered at \( r \in [0,1]^2 \), and \( F \) is the (distributional) Fourier transform, and \( P_{\Lambda} \) is projection onto the index set \( \Lambda \). Since each column \( d_i \) in \( D \) is a filter supported in \( \Lambda \), we trivially have \( P_{\Lambda}(d_i) = d_i \). Therefore, by Parseval’s theorem,
\[
\| D^H e(r) \|^2 = \sum_{i=1}^n |\langle d_i, e(r) \rangle|^2 = \sum_{i=1}^n |\langle d_i, P_{\Lambda} F\{ \delta_r \} \rangle|^2
\]
\[
= \sum_{i=1}^n |\langle d_i, F\{ \delta_r \} \rangle|^2 = \sum_{i=1}^n |\langle \mu_i, \delta_r \rangle|^2 = \sum_{i=1}^n |\mu_i(r)|^2.
\]

**References**

[1] Richard G Baraniuk, *Compressive sensing*, IEEE signal processing magazine, 24 (2007).

[2] Thierry Blu, Pier-Luigi Dragotti, Martin Vetterli, Pina Marziliano, and Lionel Coulot, *Sparse sampling of signal innovations*, Signal Processing Magazine, IEEE, 25 (2008), pp. 31–40.

[3] Martin Burger, *A level set method for inverse problems*, Inverse problems, 17 (2001), p. 1327.

[4] Martin Burger and Stanley J Osher, *A survey on level set methods for inverse problems and optimal design*, European Journal of Applied Mathematics, 16 (2005), pp. 263–301.

[5] James A Cadzow, *Signal enhancement-a composite property mapping algorithm*, Acoustics, Speech and Signal Processing, IEEE Transactions on, 36 (1988), pp. 49–62.

[6] Emmanuel J Candès and Carlos Fernandez-Granda, *Super-resolution from noisy data*, Journal of Fourier Analysis and Applications, 19 (2013), pp. 1229–1254.

[7] , *Towards a mathematical theory of super-resolution*, Communications on Pure and Applied Mathematics, 67 (2014), pp. 906–956.
[8] Tony F Chan and Luminita A Vese, A level set algorithm for minimizing the mumford-shah functional in image processing, in Variational and Level Set Methods in Computer Vision, 2001. Proceedings. IEEE Workshop on, IEEE, 2001, pp. 161–168.

[9] Changsheng Chen, Pina Marziliano, and Alex C Kot, 2d finite rate of innovation reconstruction method for step edge and polygon signals in the presence of noise, Signal Processing, IEEE Transactions on, 60 (2012), pp. 2851–2859.

[10] Yuxin Chen and Yuejie Chi, Robust spectral compressed sensing via structured matrix completion, Information Theory, IEEE Transactions on, 60 (2014), pp. 6576–6601.

[11] Qi Cheng and H Yingbo, A review of parametric high-resolution methods, High-resolution and robust signal processing (H. Yingbo, A. Gershman, and Q. Cheng, eds.), Marcel Dekker, (2003).

[12] Laurent Condat and Akira Hirabayashi, Cadzow denoising upgraded: A new projection method for the recovery of dirac pulses from noisy linear measurements, Sampling Theory in Signal and Image Processing, 14 (2015), pp. p–17.

[13] Minh N Do and Martin Vetterli, The contourlet transform: an efficient directional multiresolution image representation, Image Processing, IEEE Transactions on, 14 (2005), pp. 1289–1306.

[14] David L Donoho, Compressed sensing, Information Theory, IEEE Transactions on, 52 (2006), pp. 1289–1306.

[15] Oliver Dorn and Dominique Lesselier, Level set methods for inverse scattering, Inverse Problems, 22 (2006), p. R67.

[16] Pier Luigi Dragotti, Martin Vetterli, and Thierry Blu, Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets strang–fix, Signal Processing, IEEE Transactions on, 55 (2007), pp. 1741–1757.

[17] Robert E Edwards, Fourier series: a modern introduction, vol. 2, Springer Science & Business Media, 2012.

[18] Matthieu Guerquin-Kern, L Lejeune, Klaas P Pruessmann, and Michael Unser, Realistic analytical phantoms for parallel magnetic resonance imaging, Medical Imaging, IEEE Transactions on, 31 (2012), pp. 626–636.

[19] JP Haldar, Low-rank modeling of local k-space neighborhoods (loraks) for constrained mri., IEEE transactions on medical imaging, 33 (2014), pp. 668–681.

[20] Mathews Jacob, Yoram Bresler, Vlad Toronov, X Zhang, and Andrew Webb, Level set algorithm for the reconstruction of functional activation in near-infrared spectroscopic imaging, Journal of Biomedical Optics, (2006).

[21] Kyong Hwan Jin, Dongwook Lee, and Jong Chul Ye, A general framework for compressed sensing and parallel mri using annihilating filter based low-rank hankel matrix, arXiv preprint arXiv:1504.00532, (2015).

[22] John M Lee, Smooth manifolds, Springer, 2003.

[23] Zhi-Pei Liang, EM Haacke, and CW Thomas, High-resolution inversion of finite fourier transform data through a localised polynomial approximation, Inverse Problems, 5 (1989), p. 831.
[24] Stephane G Mallat, A theory for multiresolution signal decomposition: the wavelet representation, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 11 (1989), pp. 674–693.

[25] Irena Maravic and Martin Vetterli, Sampling and reconstruction of signals with finite rate of innovation in the presence of noise, Signal Processing, IEEE Transactions on, 53 (2005), pp. 2788–2805.

[26] Peyman Milanfar, George C Verghese, W Clem Karl, and Alan S Willsky, Reconstructing polygons from moments with connections to array processing, Signal Processing, IEEE Transactions on, 43 (1995), pp. 432–443.

[27] D Mumford and J Shah, Boundary detection by minimizing functionals, Image understanding, (1988), pp. 19–43.

[28] David Mumford and Jayant Shah, Optimal approximations by piecewise smooth functions and associated variational problems, Communications on pure and applied mathematics, 42 (1989), pp. 577–685.

[29] Greg Ongie and Mathews Jacob, Recovery of piecewise smooth images from few fourier samples, Sampling Theory and Applications (SampTA), (2015).

[30] ———, Super-resolution MRI using finite rate of innovation, IEEE International Symposium on Biomedical Imaging: ISBI 2015., (2015).

[31] Parama Pal and PP Vaidyanathan, A grid-less approach to underdetermined direction of arrival estimation via low rank matrix denoising, Signal Processing Letters, IEEE, 21 (2014), pp. 737–741.

[32] Hanjie Pan, Thierry Blu, and Pier Luigi Dragotti, Sampling curves with finite rate of innovation, Signal Processing, IEEE Transactions on, 62 (2014).

[33] Ralph O Schmidt, Multiple emitter location and signal parameter estimation, Antennas and Propagation, IEEE Transactions on, 34 (1986), pp. 276–280.

[34] M Schweiger, O Dorn, A Zacharopoulos, I Nissila, and SR Arridge, 3d level set reconstruction of model and experimental data in diffuse optical tomography, Optics express, 18 (2010), pp. 150–164.

[35] Igor R Shafarevich, Basic algebraic geometry. 1. Varieties in projective space., Springer-Verlag, Berlin, 1994.

[36] Pancham Shukla and Pier Luigi Dragotti, Sampling schemes for multidimensional signals with finite rate of innovation, Signal Processing, IEEE Transactions on, 55 (2007), pp. 3670–3686.

[37] József Solymosi and Terence Tao, An incidence theorem in higher dimensions, Discrete & Computational Geometry, 48 (2012), pp. 255–280.

[38] J-L Starck, Emmanuel J Candès, and David L Donoho, The curvelet transform for image denoising, Image Processing, IEEE Transactions on, 11 (2002), pp. 670–684.

[39] Petre Stoica and Randolph L Moses, Introduction to spectral analysis, vol. 1, Prentice hall Upper Saddle River, NJ, 1997.

[40] Martin Storath, Andreas Weinmann, and Laurent Demaret, Jump-sparse and sparse recovery using potts functionals, (2014).
[41] Robert S Strichartz, *A guide to distribution theory and Fourier transforms*, World Scientific, 2003.

[42] Gongguo Tang, Badri Narayan Bhaskar, and Benjamin Recht, *Near minimax line spectral estimation*, in Information Sciences and Systems (CISS), 2013 47th Annual Conference on, IEEE, 2013, pp. 1–6.

[43] Andy Tsai, Anthony Yezzi Jr, and Alan S Willsky, *Curve evolution implementation of the mumford-shah functional for image segmentation, denoising, interpolation, and magnification*, Image Processing, IEEE Transactions on, 10 (2001), pp. 1169–1186.

[44] Luminita A Vese and Tony F Chan, *A multiphase level set framework for image segmentation using the mumford and shah model*, International journal of computer vision, 50 (2002), pp. 271–293.

[45] Martin Vetterli, Pina Marziliano, and Thierry Blu, *Sampling signals with finite rate of innovation*, Signal Processing, IEEE Transactions on, 50 (2002), pp. 1417–1428.

[46] Weiyu Xu, Jian-Feng Cai, Kumar Vlaj Mishra, Myung Cho, and Anton Kruger, *Precise semidefinite programming formulation of atomic norm minimization for recovering d-dimensional (d^2) off-the-grid frequencies*, in Information Theory and Applications Workshop (ITA), 2014, IEEE, 2014, pp. 1–4.

[47] Jong Chul Ye, Yoram Bresler, and Pierre Moulin, *A self-referencing level-set method for image reconstruction from sparse fourier samples*, International Journal of Computer Vision, 50 (2002), pp. 253–270.

[48] Mingqiang Zhu and Tony Chan, *An efficient primal-dual hybrid gradient algorithm for total variation image restoration*, UCLA CAM Report, (2008), pp. 08–34.