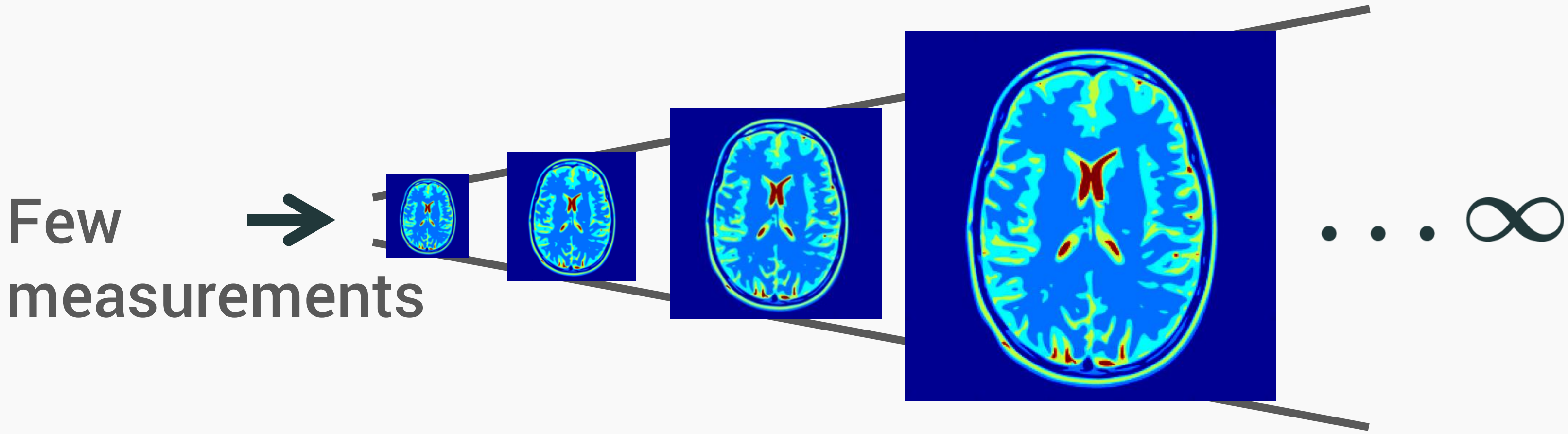
Off-the-Grid Compressive Imaging: Recovery of Piecewise Constant Images from Few Fourier Samples

Greg Ongie
PhD Candidate
Department of Applied Math and Computational Sciences
University of Iowa

April 25, 2016
U. Michigan, CSP Seminar

Our goal is to develop theory and algorithms for compressive off-the-grid imaging

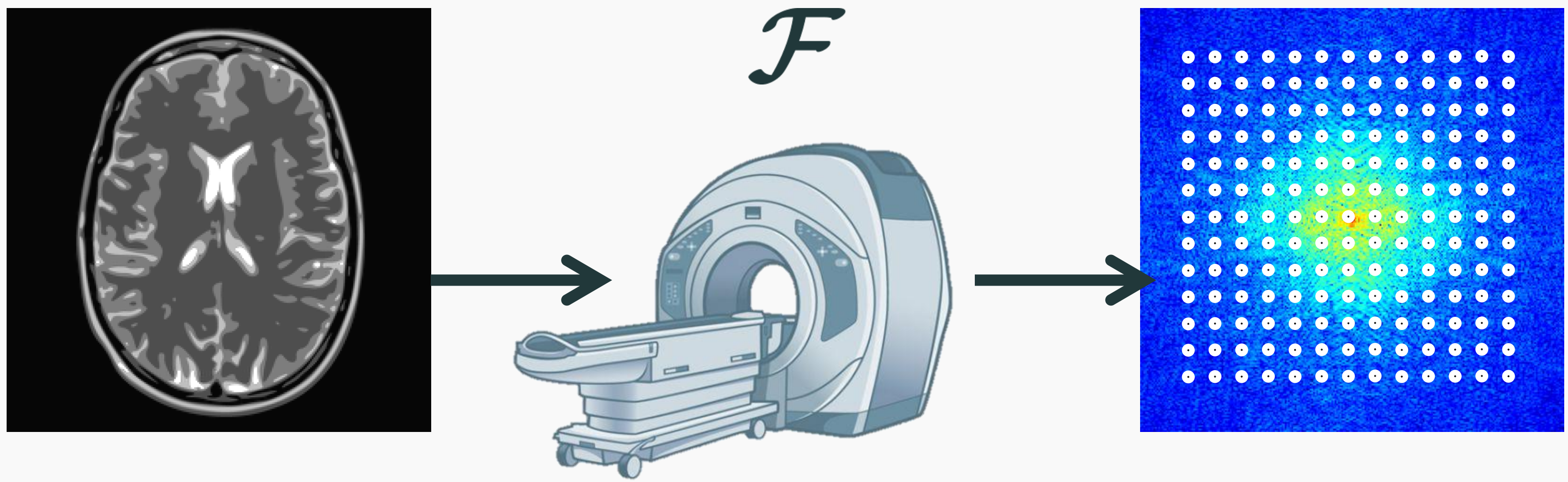


Off-the-grid = Continuous domain representation

Compressive off-the-grid *imaging*:

Exploit continuous domain modeling to improve image recovery from few measurements

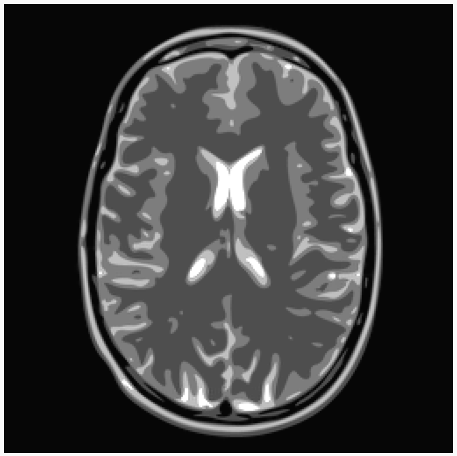
Motivation: MRI Reconstruction



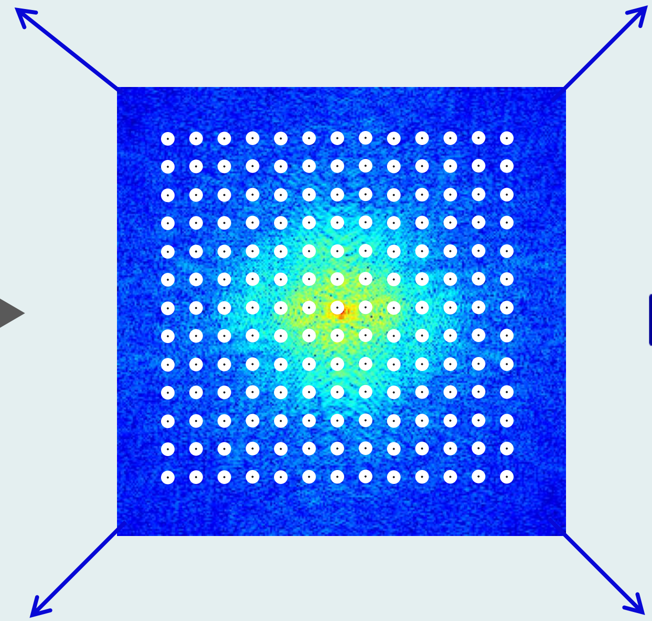
Main Problem:

Reconstruct image from Fourier domain samples

Related: Computed Tomography, Florescence Microscopy



$$\xrightarrow{\mathcal{F}}$$



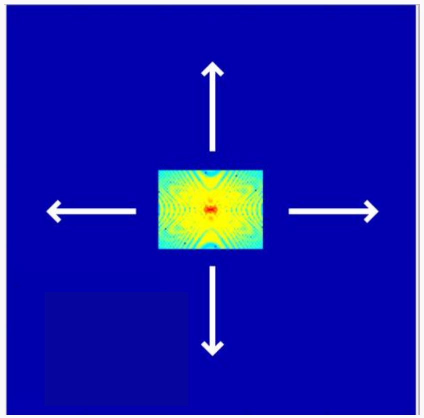
$$\mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in [0, 1]^{\mathbf{d}}$$

$$\widehat{\mathbf{f}}[\mathbf{k}] := \int_{[0,1]^{\mathbf{d}}} \mathbf{f}(\mathbf{x}) e^{-j2\pi \mathbf{k}\cdot\mathbf{x}} \mathbf{dx}$$

Uniform Fourier Samples =
Fourier Series Coefficients

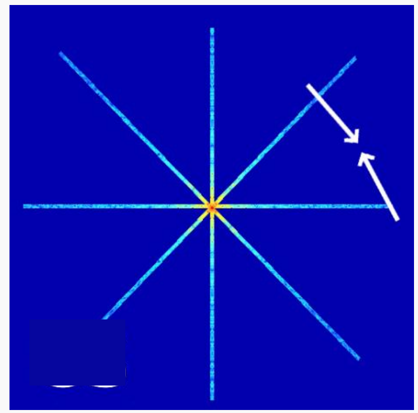
Types of “Compressive” Fourier Domain Sampling

low-pass

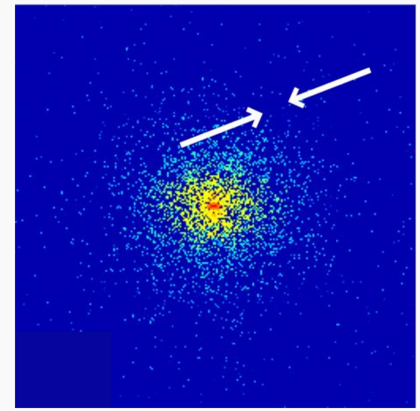


vs.

radial



random



Fourier Extrapolation

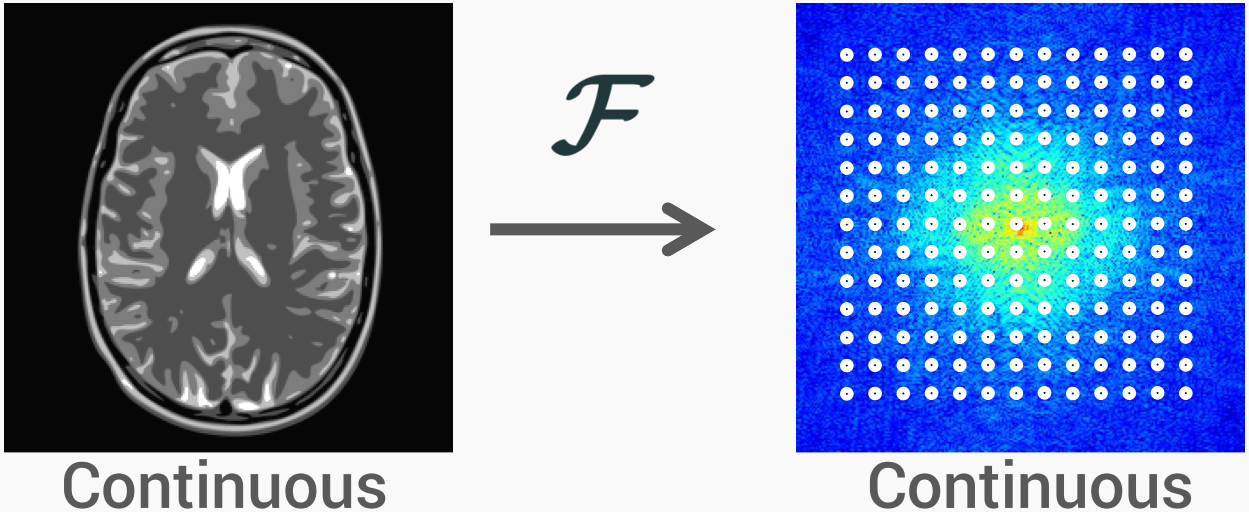
Fourier Interpolation

Super-resolution recovery

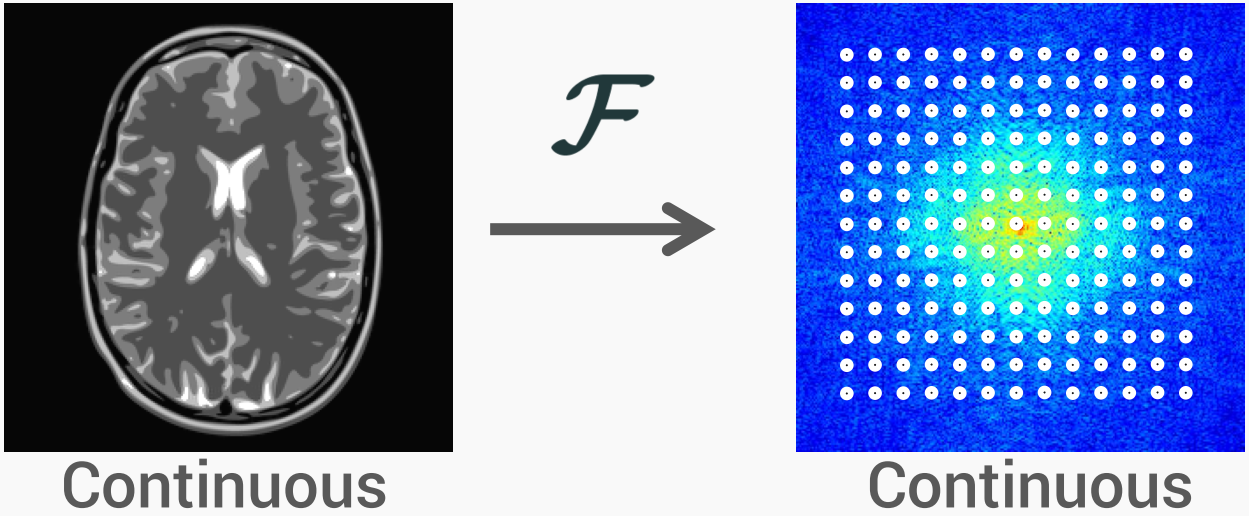
“Compressed Sensing” recovery

CURRENT DISCRETE PARADIGM

“True” measurement model:

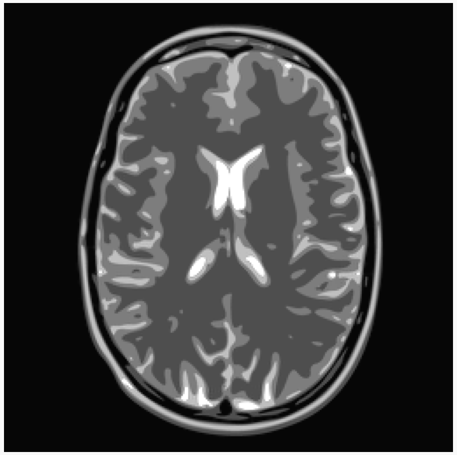


“True” measurement model:



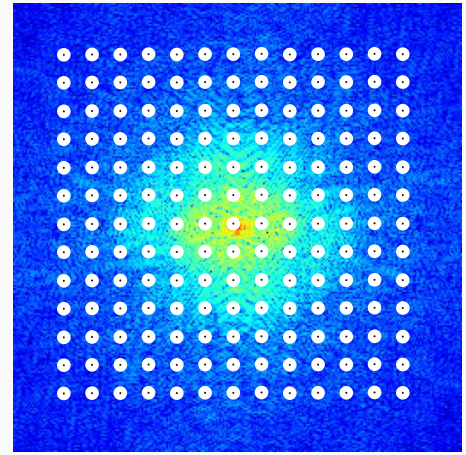
Approximated measurement model:

DFT Reconstruction

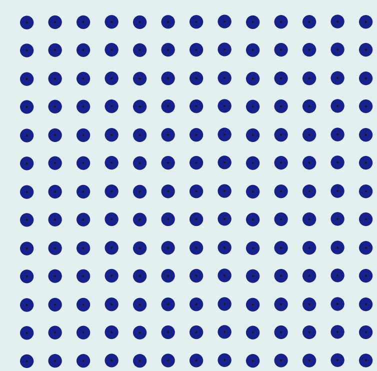


Continuous

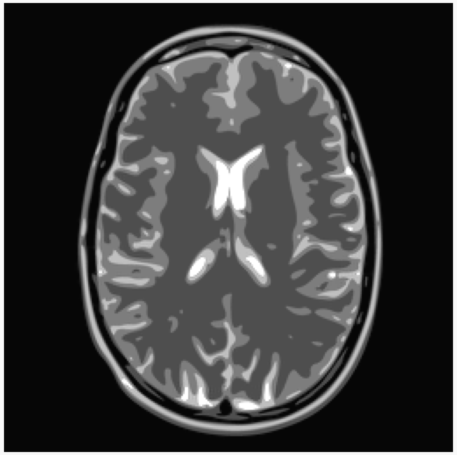
$\xrightarrow{\mathcal{F}}$



Continuous

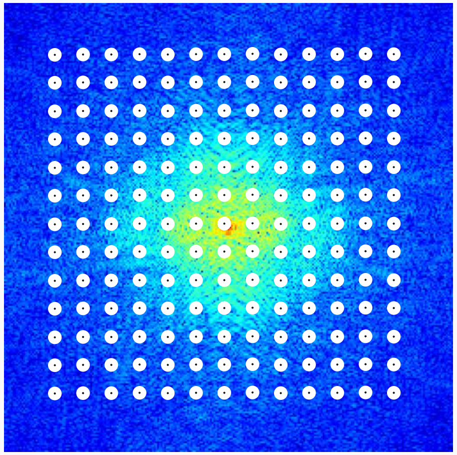


DFT Reconstruction

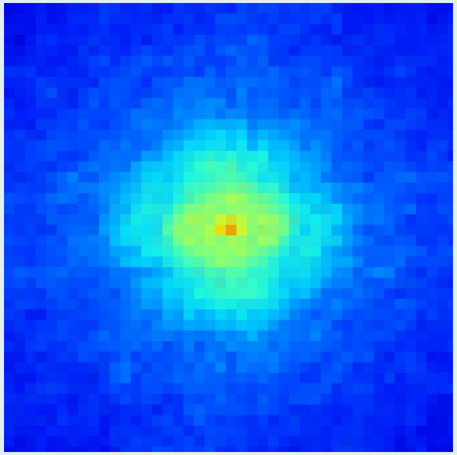


Continuous

$\xrightarrow{\mathcal{F}}$

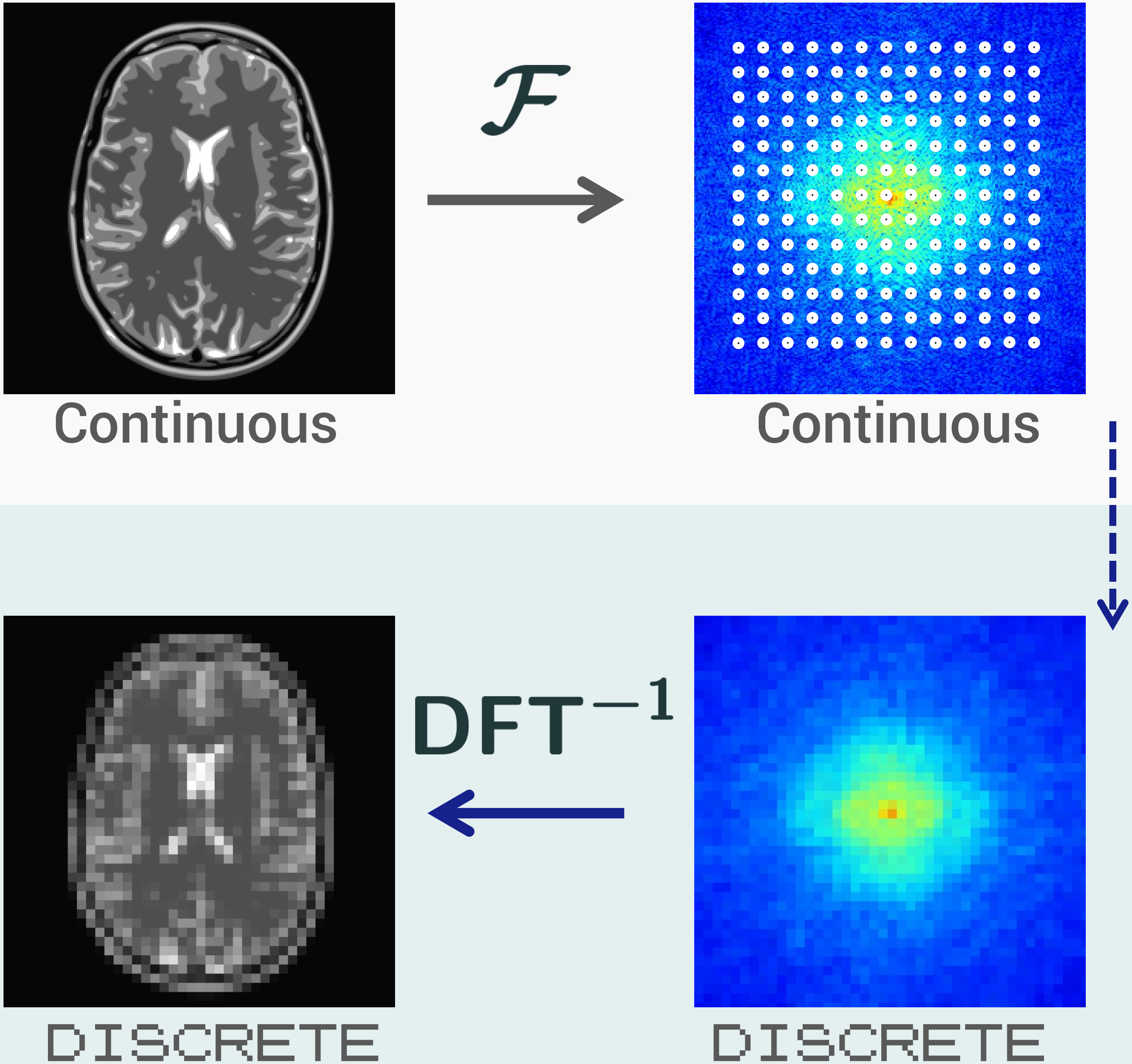


Continuous

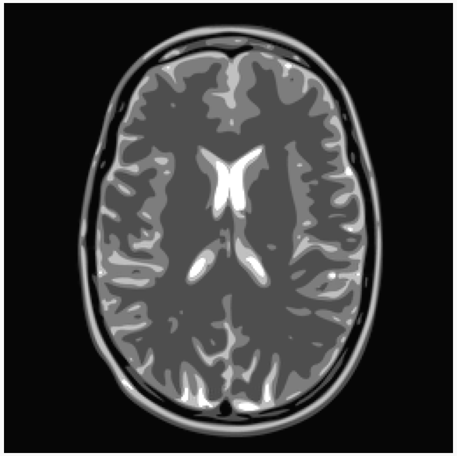


DISCRETE

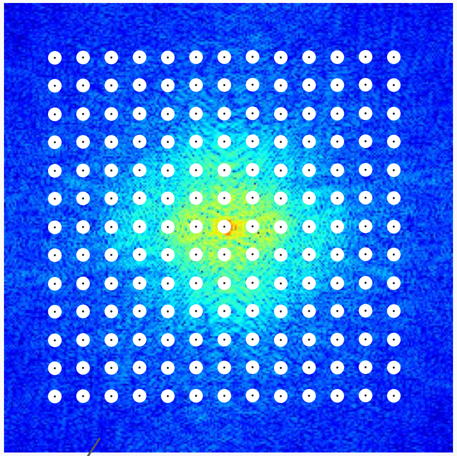
DFT Reconstruction



"Compressed Sensing" Recovery

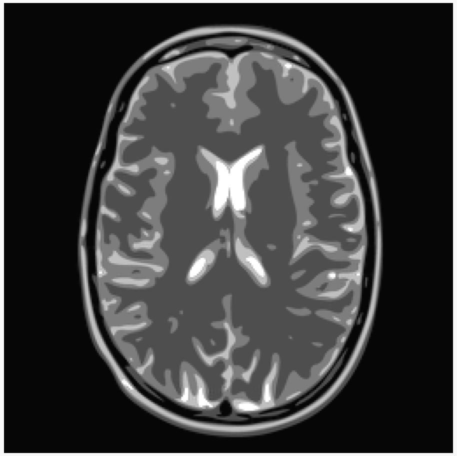


$\xrightarrow{\mathcal{F}}$

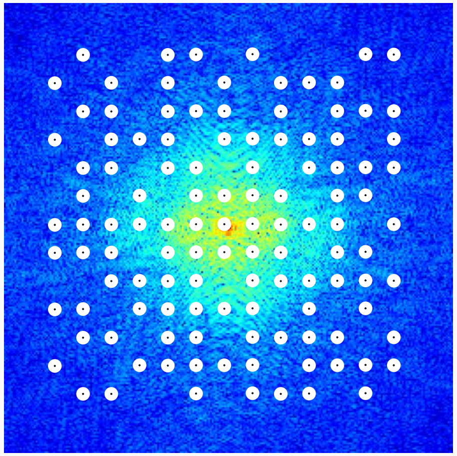


Full sampling is costly!
(or impossible—e.g. Dynamic MRI)

"Compressed Sensing" Recovery

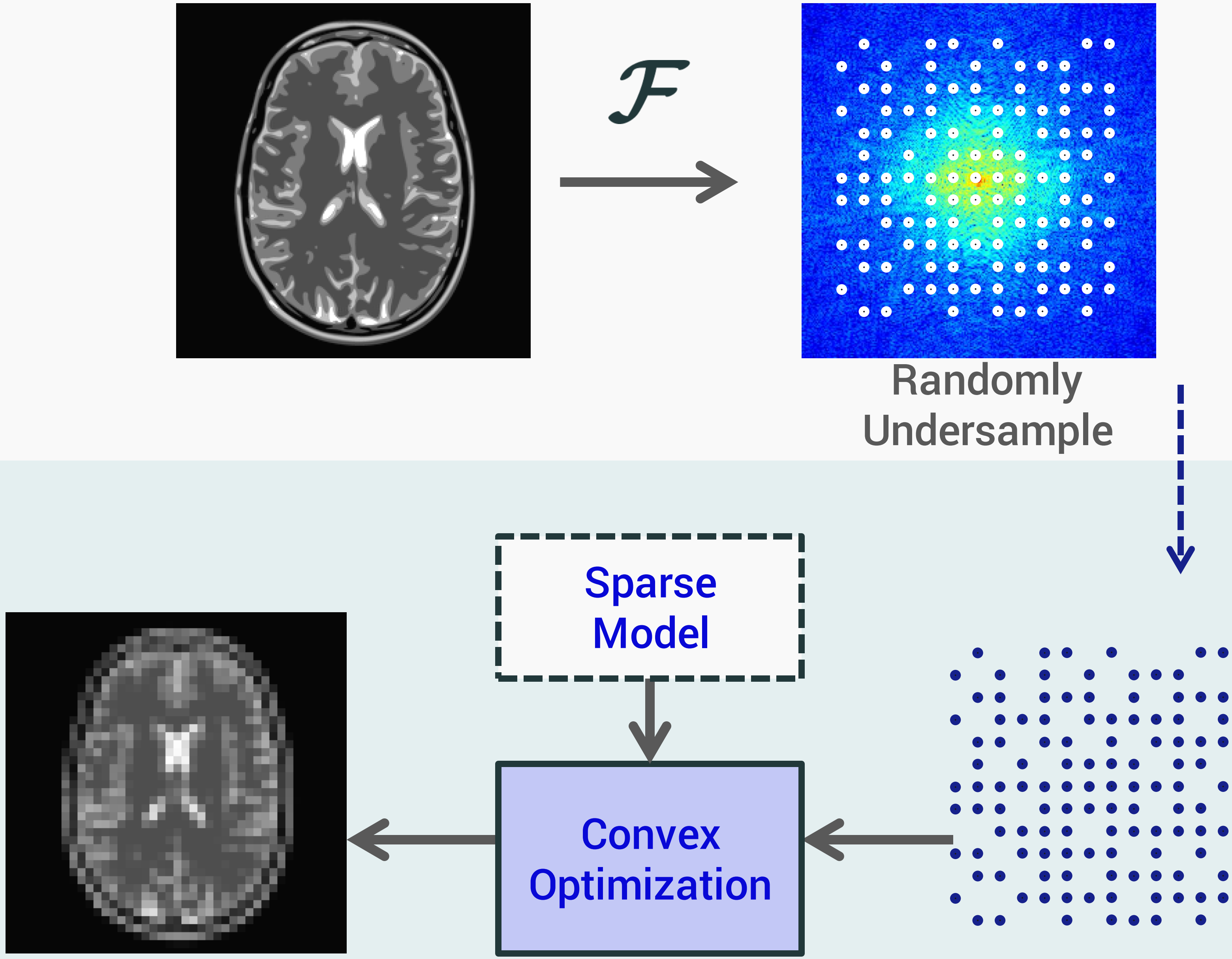


$\xrightarrow{\mathcal{F}}$



Randomly
Undersample

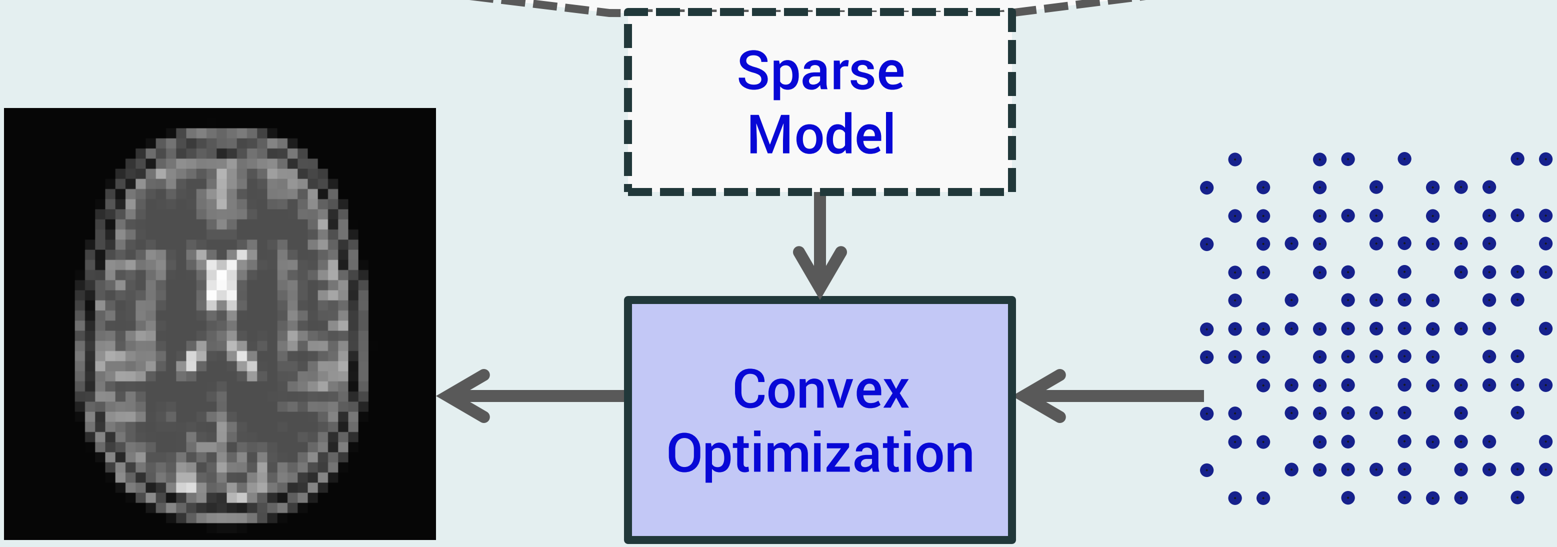
“Compressed Sensing” Recovery



Example:
Assume discrete gradient
of image is sparse

↕

Piecewise constant model



Recovery by Total Variation (TV) minimization

TV semi-norm: $\|\mathbf{g}\|_{\mathrm{TV}} = \sum_{\mathbf{i},\mathbf{j}} \sqrt{|\mathbf{g}_{\mathbf{i}+1,\mathbf{j}} - \mathbf{g}_{\mathbf{i},\mathbf{j}}|^2 + |\mathbf{g}_{\mathbf{i},\mathbf{j}+1} - \mathbf{g}_{\mathbf{i},\mathbf{j}}|^2}$

i.e., L1-norm of discrete gradient magnitude

$\sum_{\mathbf{i},\mathbf{j}}$ 

Recovery by Total Variation (TV) minimization

TV semi-norm: $\|\mathbf{g}\|_{TV} = \sum_{i,j} \sqrt{|\mathbf{g}_{i+1,j} - \mathbf{g}_{i,j}|^2 + |\mathbf{g}_{i,j+1} - \mathbf{g}_{i,j}|^2}$

i.e., L1-norm of discrete gradient magnitude

$$\sum_{i,j}$$

$$\min_{\mathbf{g}\in\mathbb{C}^{N\times N}} \|\mathbf{g}\|_{TV} \quad \text{subject to} \quad \mathbf{F}_{\Omega}\mathbf{g} = \mathbf{F}_{\Omega}\mathbf{f} \qquad (\mathbf{TV}\text{-min})$$

Recovery by Total Variation (TV) minimization

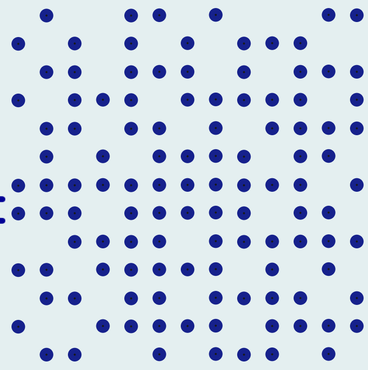
TV semi-norm: $\|\mathbf{g}\|_{\mathbf{TV}} = \sum_{\mathbf{i},\mathbf{j}} \sqrt{|\mathbf{g}_{\mathbf{i}+1,\mathbf{j}} - \mathbf{g}_{\mathbf{i},\mathbf{j}}|^2 + |\mathbf{g}_{\mathbf{i},\mathbf{j}+1} - \mathbf{g}_{\mathbf{i},\mathbf{j}}|^2}$

i.e., L1-norm of discrete gradient magnitude

$\sum_{\mathbf{i},\mathbf{j}}$ 

$$\min_{\mathbf{g}\in\mathbb{C}^{\mathbf{N}\times\mathbf{N}}} \|\mathbf{g}\|_{\mathbf{TV}} \quad \text{subject to} \quad \mathbf{F}_{\Omega}\mathbf{g} = \mathbf{F}_{\Omega}\mathbf{f} \qquad (\mathbf{TV}\text{-min})$$

Restricted DFT

$\Omega =$ 

Sample locations

Recovery by Total Variation (TV) minimization

TV semi-norm: $\|\mathbf{g}\|_{\mathsf{TV}} = \sum_{\mathbf{i},\mathbf{j}} \sqrt{|\mathbf{g}_{\mathbf{i}+1,\mathbf{j}} - \mathbf{g}_{\mathbf{i},\mathbf{j}}|^2 + |\mathbf{g}_{\mathbf{i},\mathbf{j}+1} - \mathbf{g}_{\mathbf{i},\mathbf{j}}|^2}$

i.e., L1-norm of discrete gradient magnitude

$$\sum_{\mathbf{i},\mathbf{j}}$$

$$\min_{\mathbf{g}\in\mathbb{C}^{\mathsf{N}\times\mathsf{N}}} \|\mathbf{g}\|_{\mathsf{TV}} \quad \text{subject to} \quad \mathbf{F}_\Omega\mathbf{g} = \mathbf{F}_\Omega\mathbf{f} \qquad (\mathbf{TV}\text{-min})$$

Restricted DFT

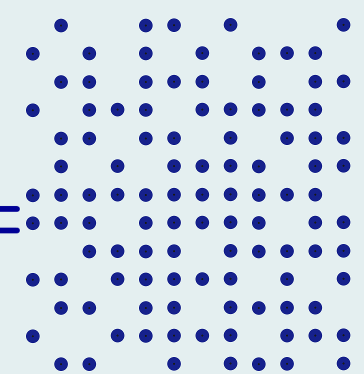
Convex optimization problem

Fast iterative algorithms:

ADMM/Split-Bregman,

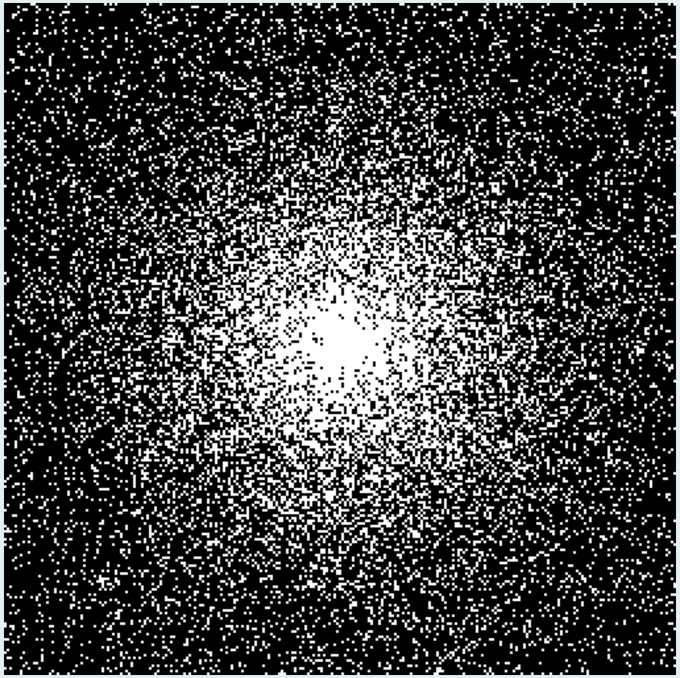
FISTA, Primal-Dual, etc.

$\Omega =$



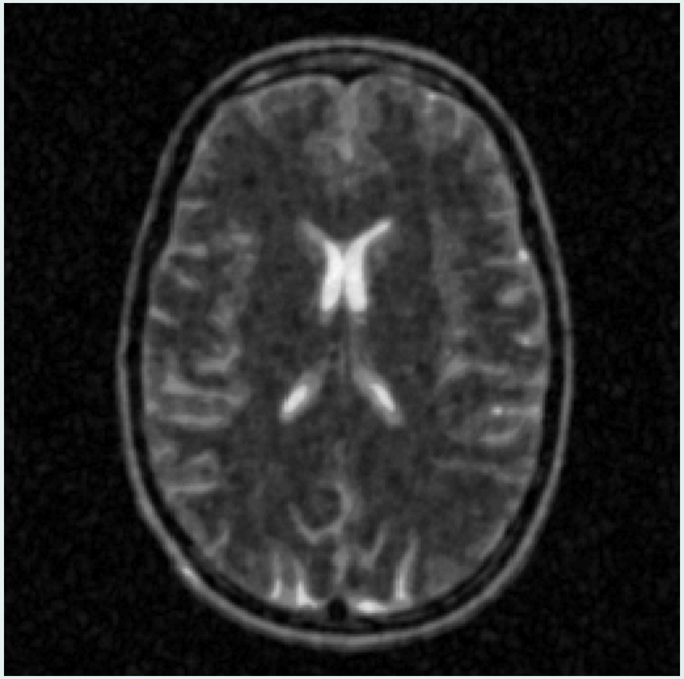
Sample locations

Example:



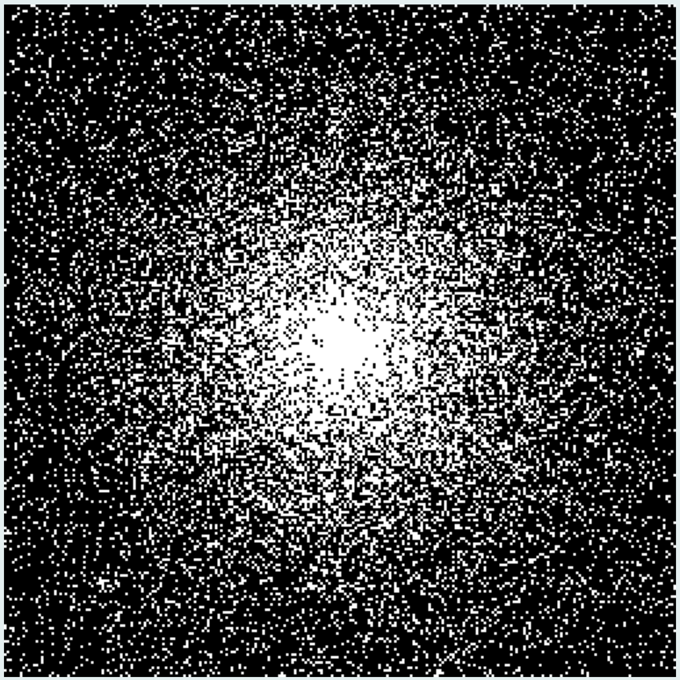
25% Random
Fourier samples
(variable density)

$\xrightarrow{\mathbf{DFT}^{-1}}$



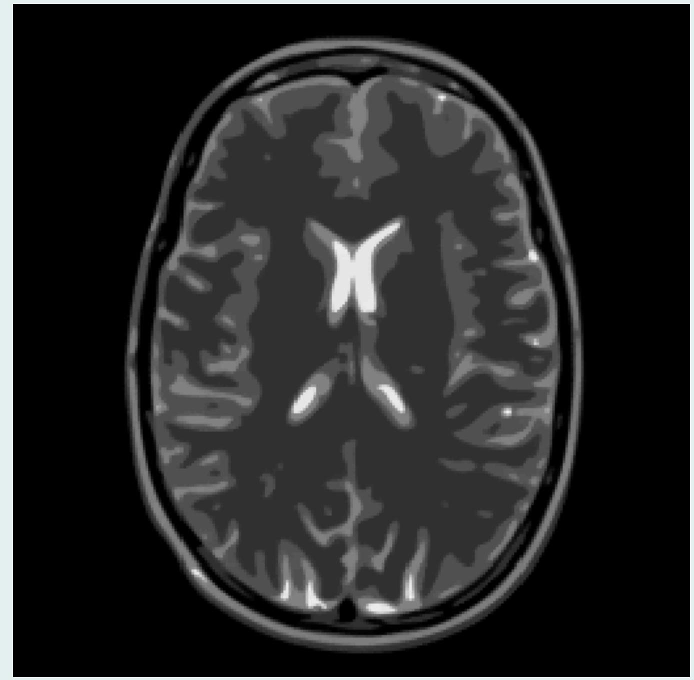
Rel. Error = 30%

Example:



25% Random
Fourier samples
(variable density)

TV-min



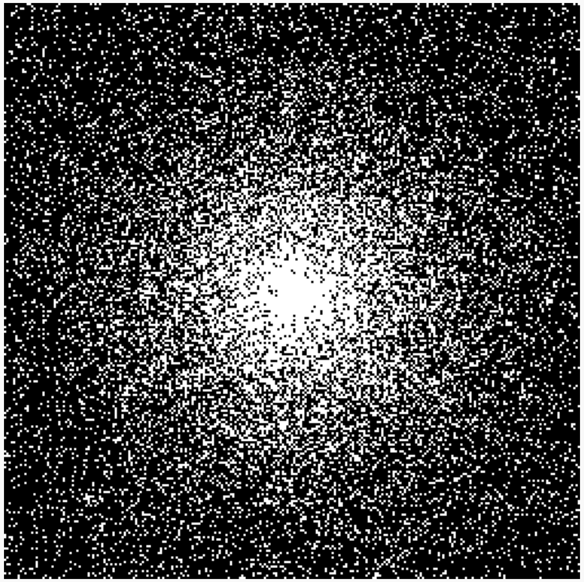
Rel. Error = 5%

Theorem [Krahmer & Ward, 2012]:

If $\mathbf{f} \in \mathbb{C}^{N \times N}$ has ***s*-sparse gradient**, then **f** is the unique solution to (TV-min) with high probability provided the number of random* Fourier samples **m** satisfies

$$\mathbf{m} \gtrsim \mathbf{s} \log^3(\mathbf{s}) \log^5(\mathbf{N})$$

* Variable density sampling



Summary of
DISCRETE PARADIGM

- Approximate $\mathcal{F} \rightarrow \mathbf{DFT}$

- **Fully sampled:**
  Fast reconstruction by $\mathbf{DFT}^{-1}$

- **Under-sampled (*Compressed sensing*):**
  Exploit sparse models & convex optimization
  – E.g. TV-minimization
  – Recovery guarantees

Summary of DISCRETE PARADIGM

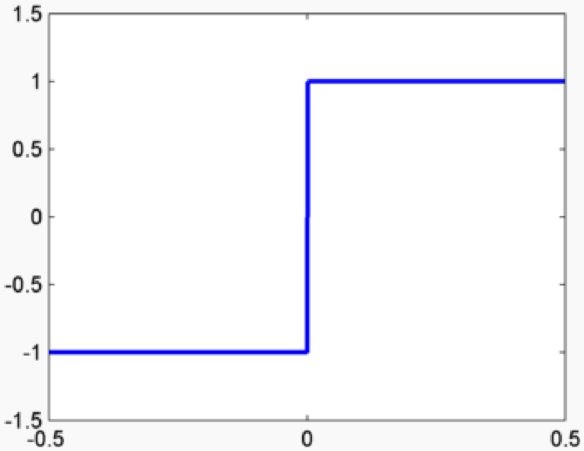
- Approximate $\mathcal{F} \longrightarrow \mathbf{DFT}$

- **Fully sampled:**
  Fast reconstruction by $\mathbf{DFT}^{-1}$

- **Under-sampled (*Compressed sensing*):**
  Exploit sparse models & convex optimization
  – E.g. TV-minimization
  – Recovery guarantees

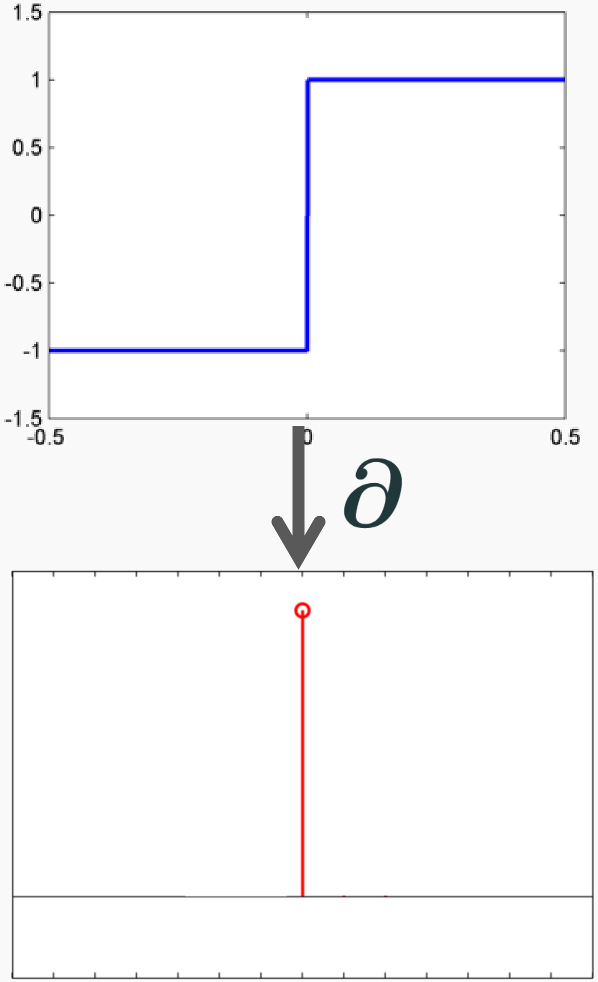
Problem: The DFT Destroys Sparsity!

Continuous



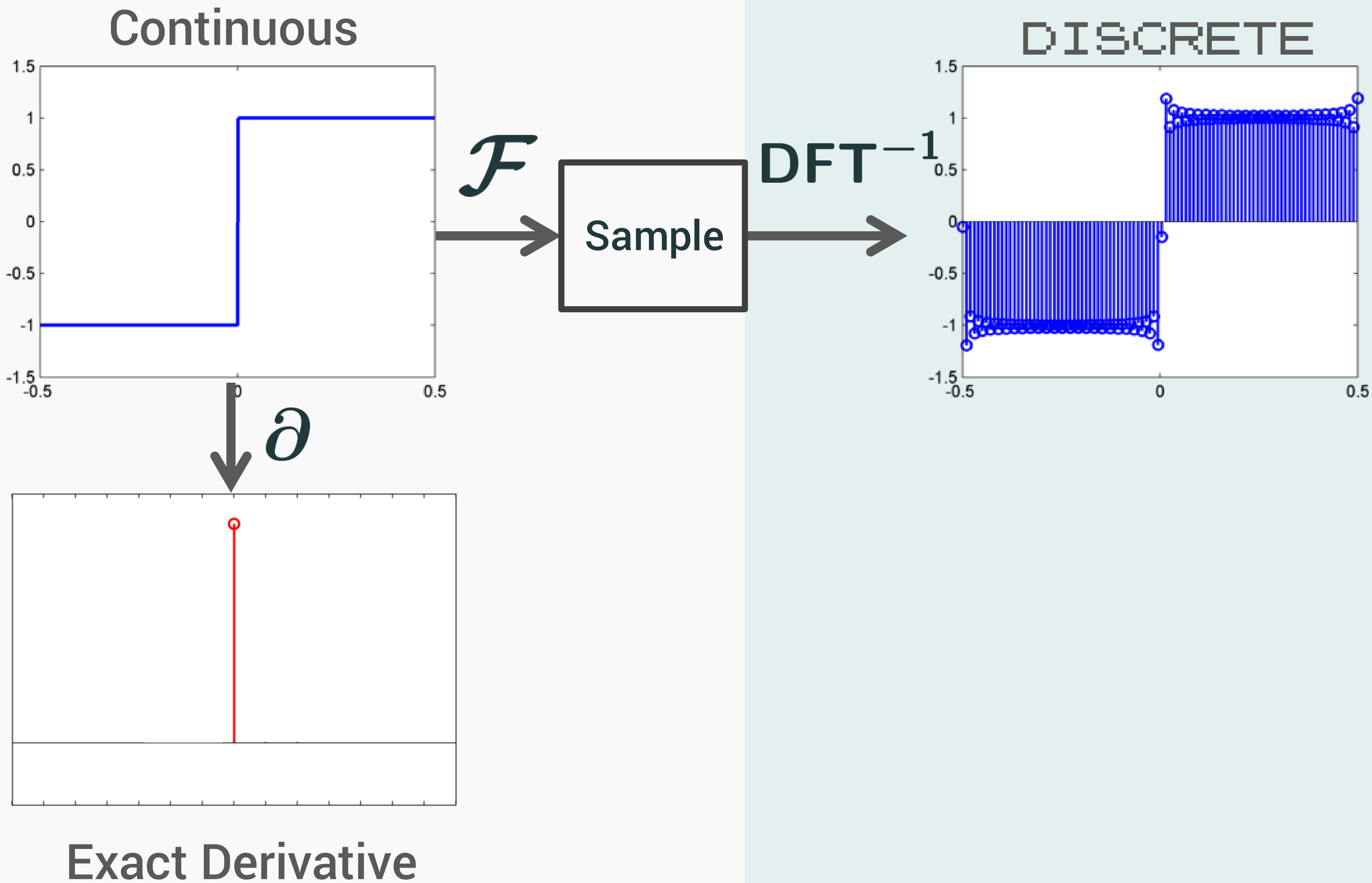
Problem: The DFT Destroys Sparsity!

Continuous

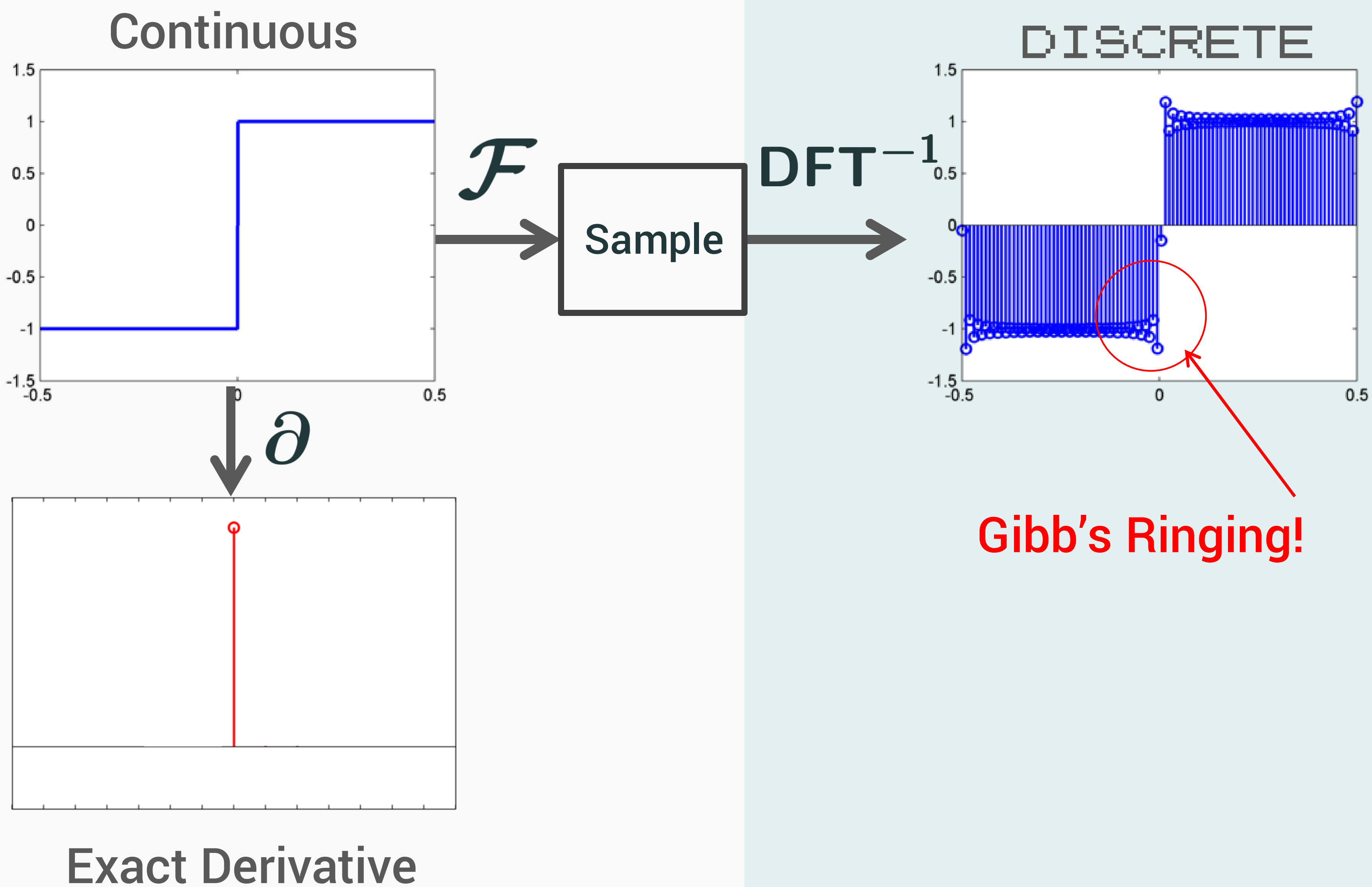


Exact Derivative

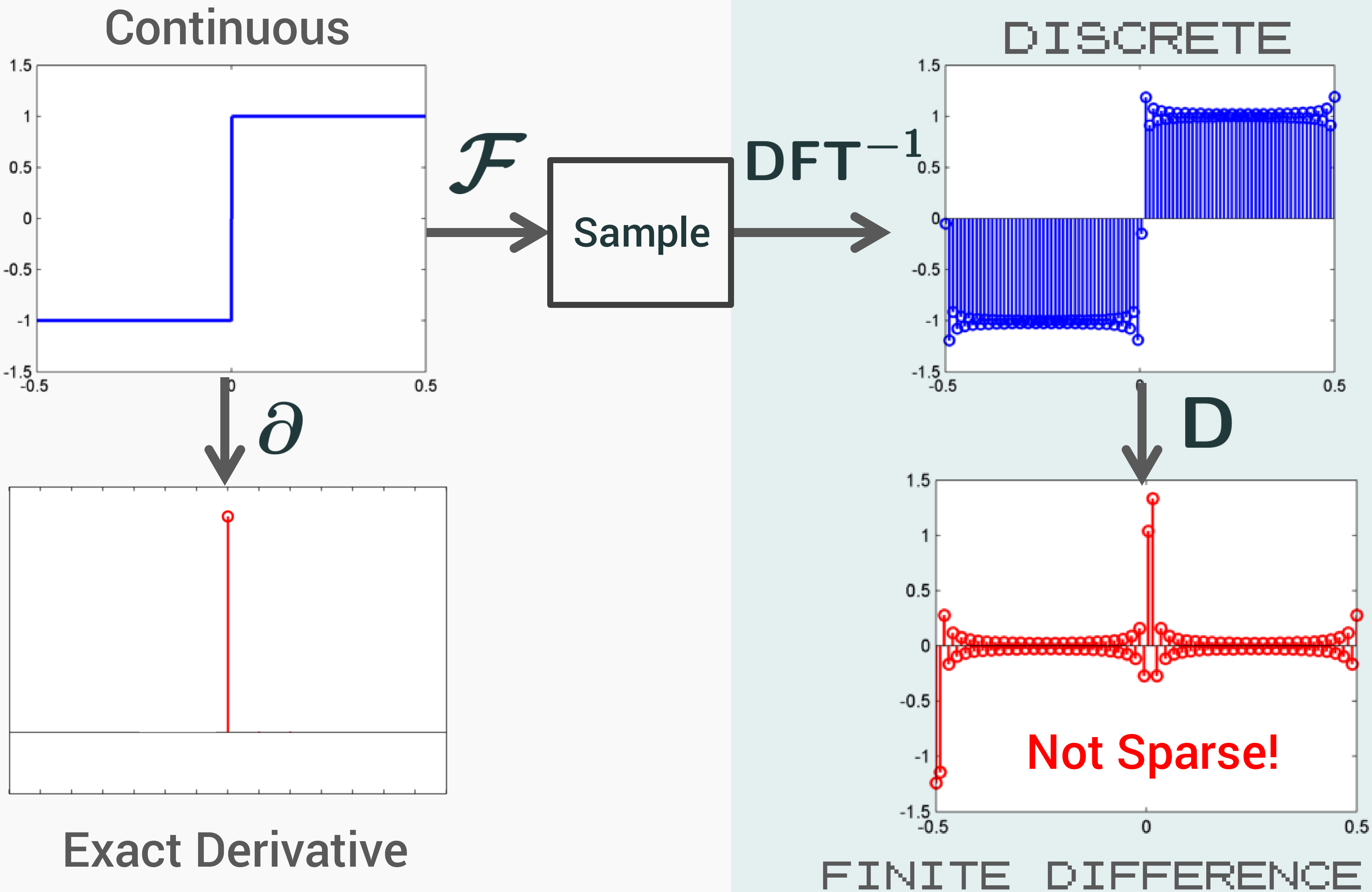
Problem: The DFT Destroys Sparsity!



Problem: The DFT Destroys Sparsity!



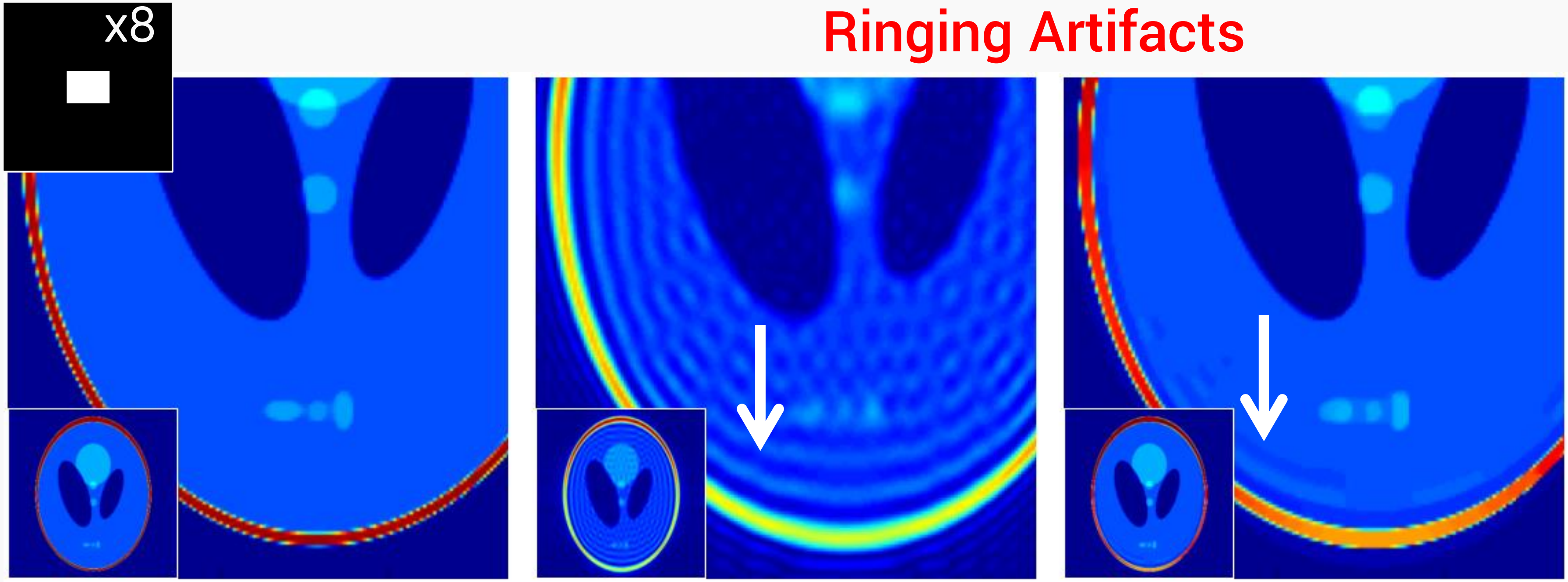
Problem: The DFT Destroys Sparsity!



Consequence: TV fails in super-resolution setting

Fourier

Ringing Artifacts

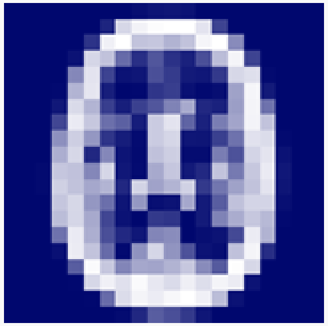


(a) Fully sampled (b) IFFT, SNR=10.8dB (c) TV, SNR=16.6dB

Can we move beyond the DISCRETE PARADIGM in Compressive Imaging?

Challenges:

- Continuous domain sparsity $\neq$ Discrete domain sparsity



- What are the continuous domain analogs of sparsity?
- Can we pose recovery as a convex optimization problem?
- Can we give recovery guarantees, *a la* TV-minimization?

New *Off-the-Grid* Imaging Framework: *Theory*

Classical Off-the-Grid Method: Prony (1795)

Uniform time samples

Off-the-grid frequencies

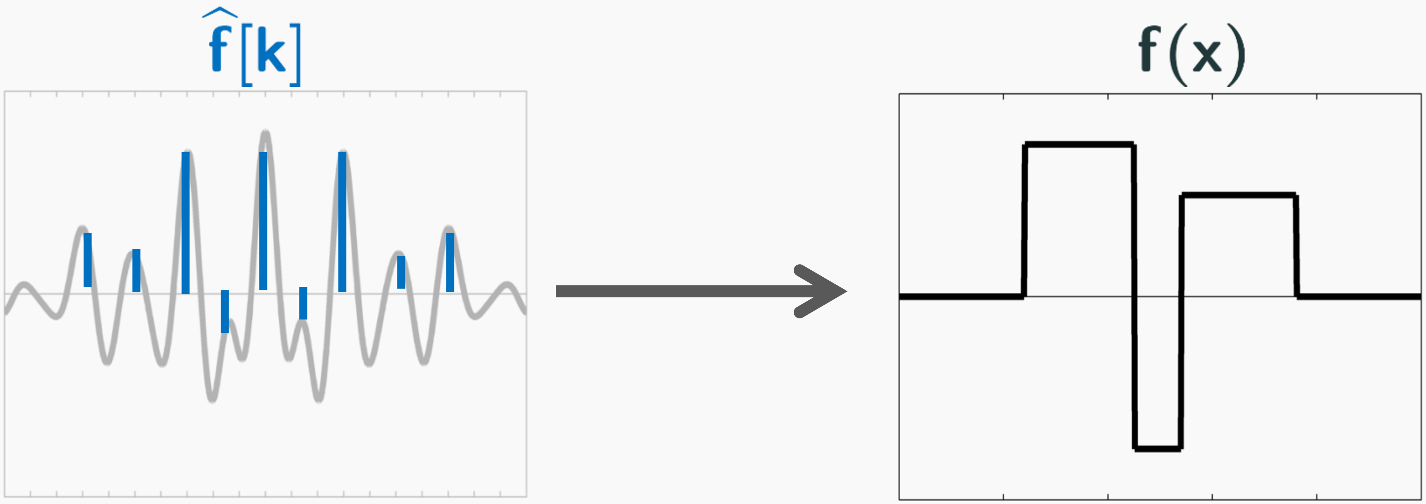
- Robust variants:

  Pisarenko (1973), MUSIC (1986), ESPRIT (1989),

  Matrix pencil (1990) . . . Atomic norm (2011)

Main inspiration: Finite-Rate-of-Innovation (FRI)

[Vetterli et al., 2002]

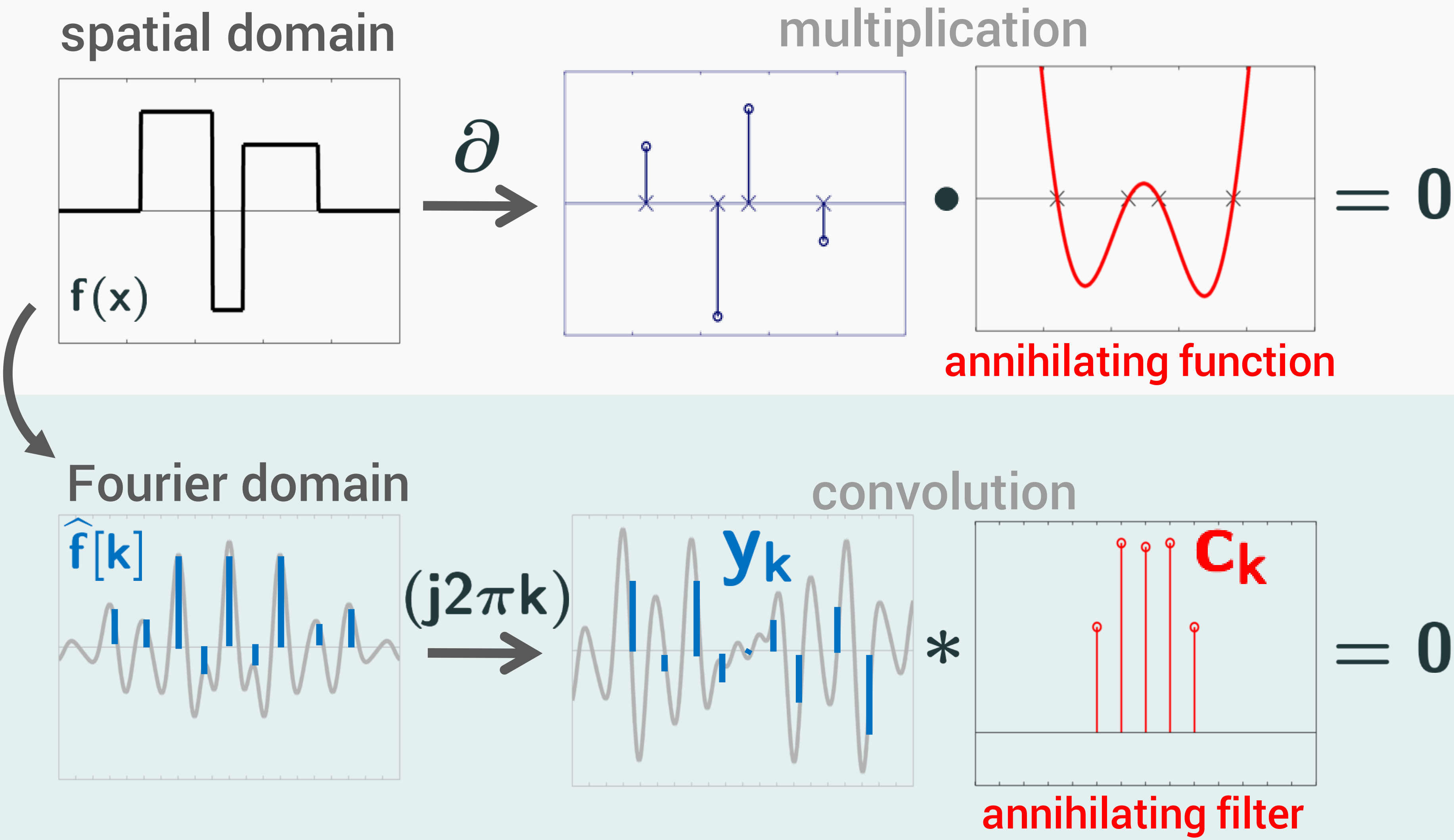


Uniform
Fourier samples

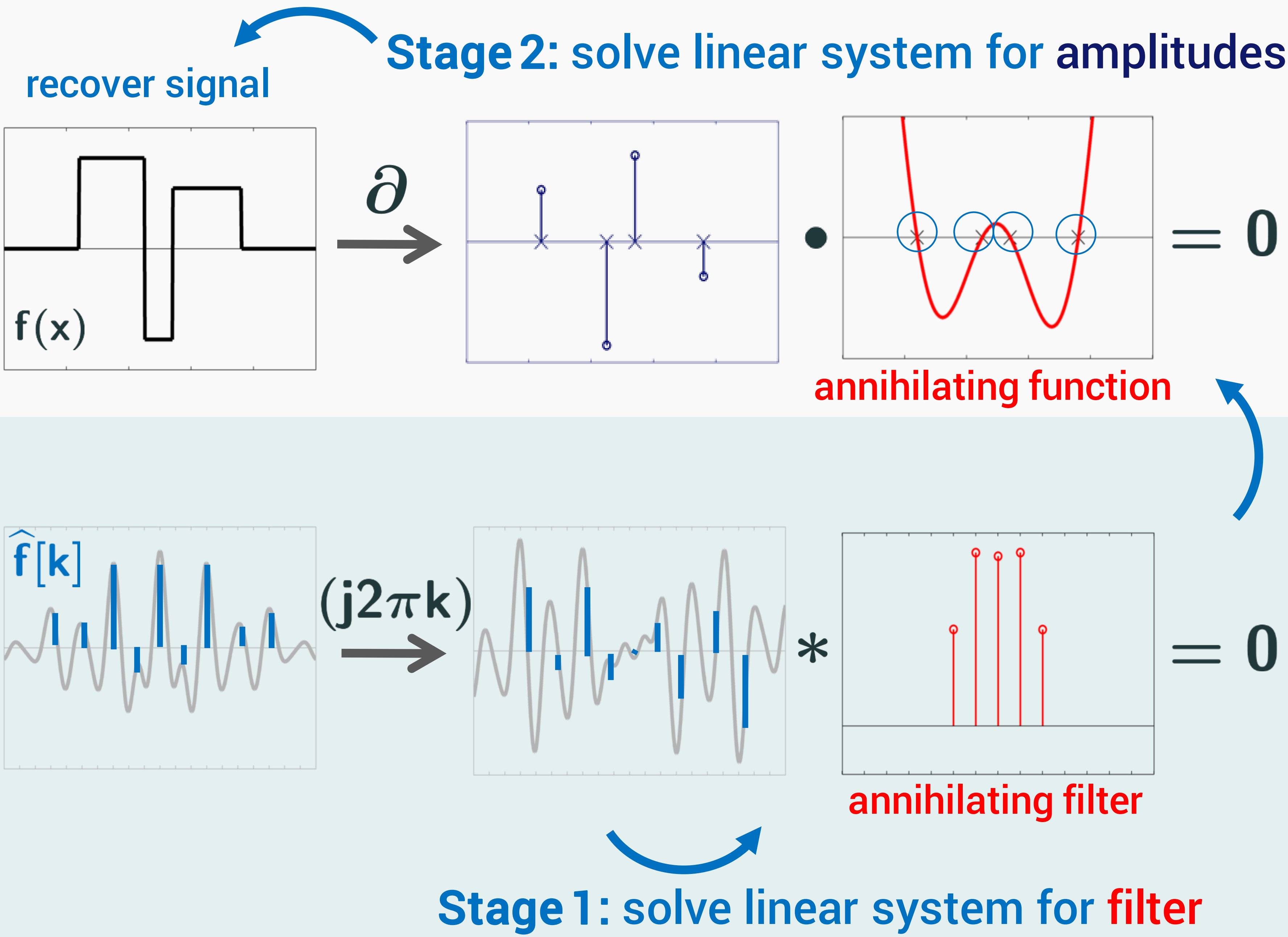
Off-the-grid
PWC signal

- Recent extension to 2-D images:

  Pan, Blu, & Dragotti (2014), "Sampling Curves with FRI".

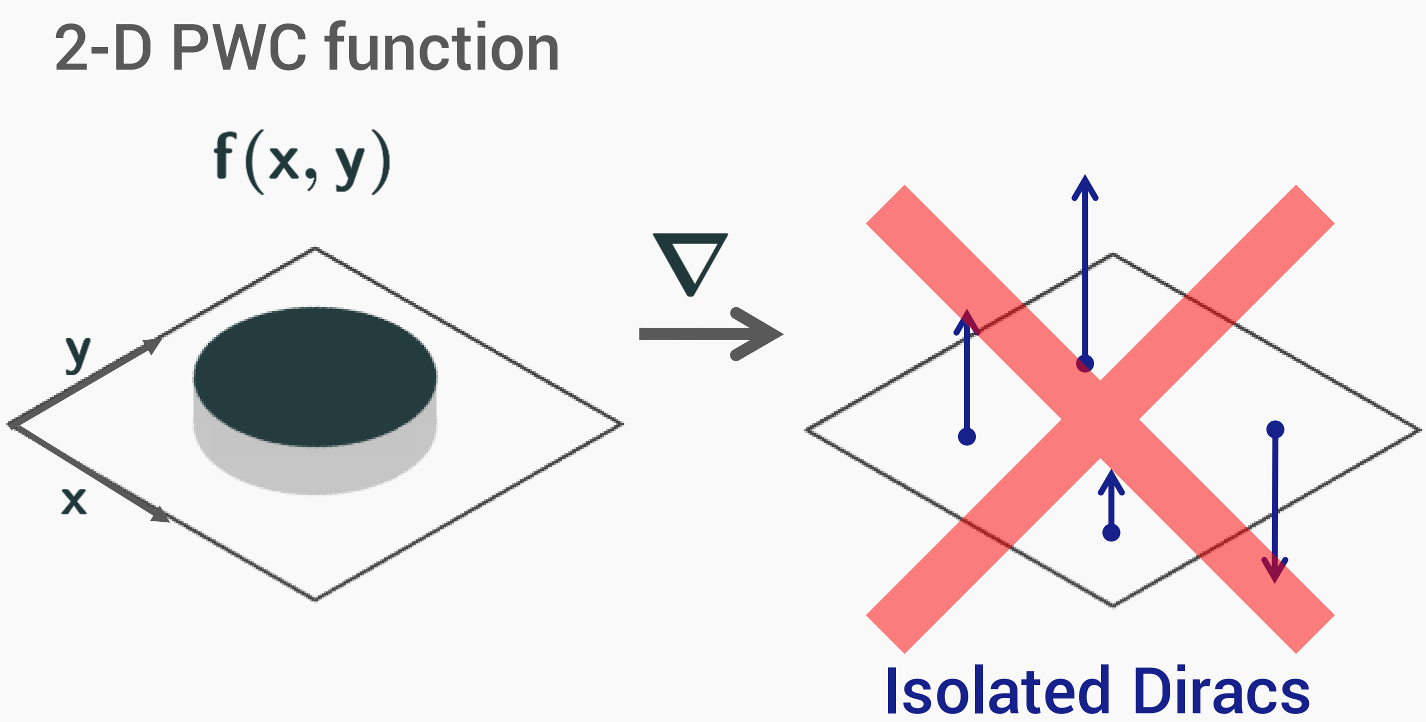


Annihilation Relation: $\sum_k y_{\ell-k} c_k = 0$

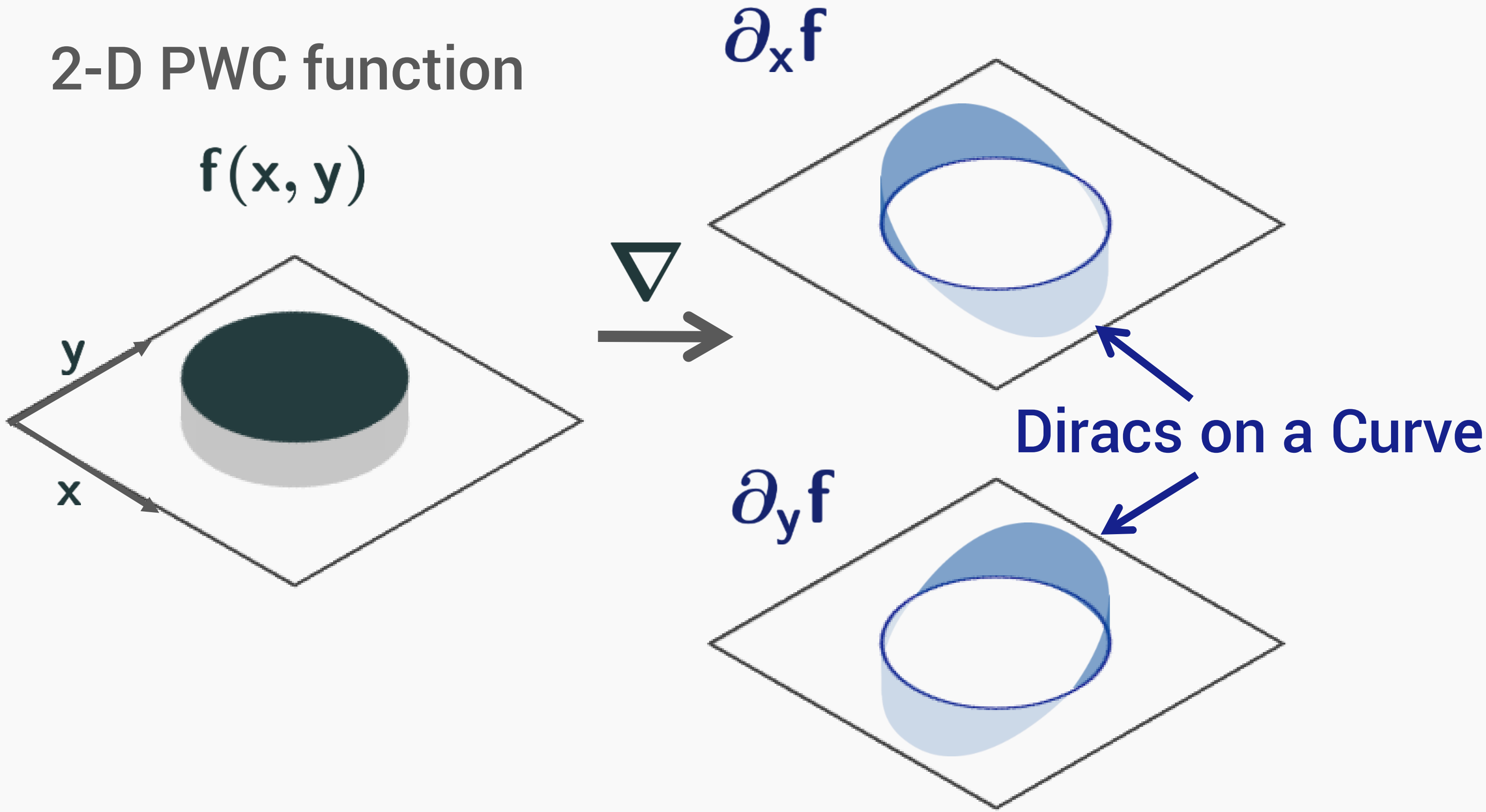


Stage 2: solve linear system for amplitudes
recover signal
f(x)
∂
annihilating function
= 0
f̂[k]
(j2πk)
*
annihilating filter
= 0
Stage 1: solve linear system for filter

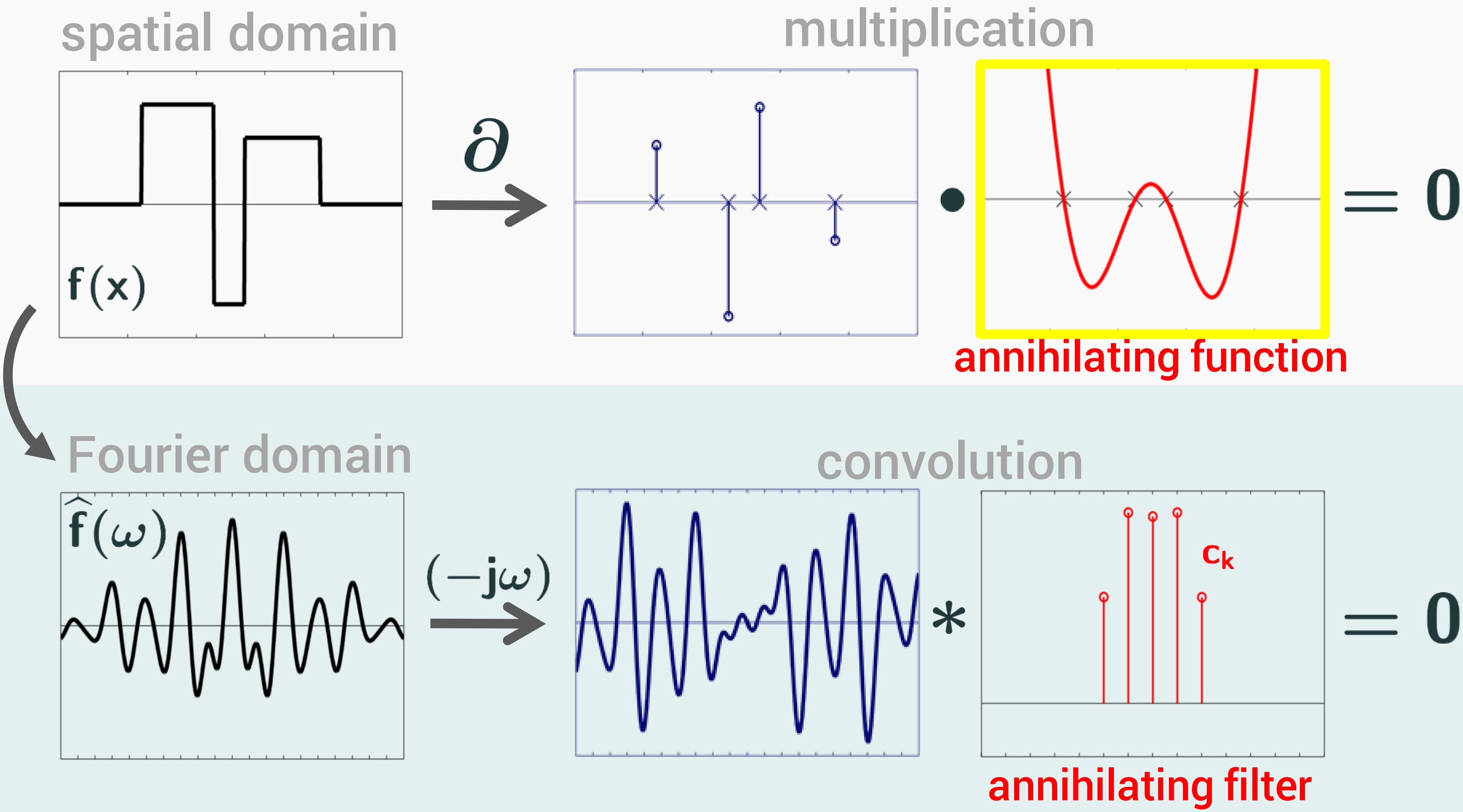
Challenges extending FRI to higher dimensions:
Singularities not isolated



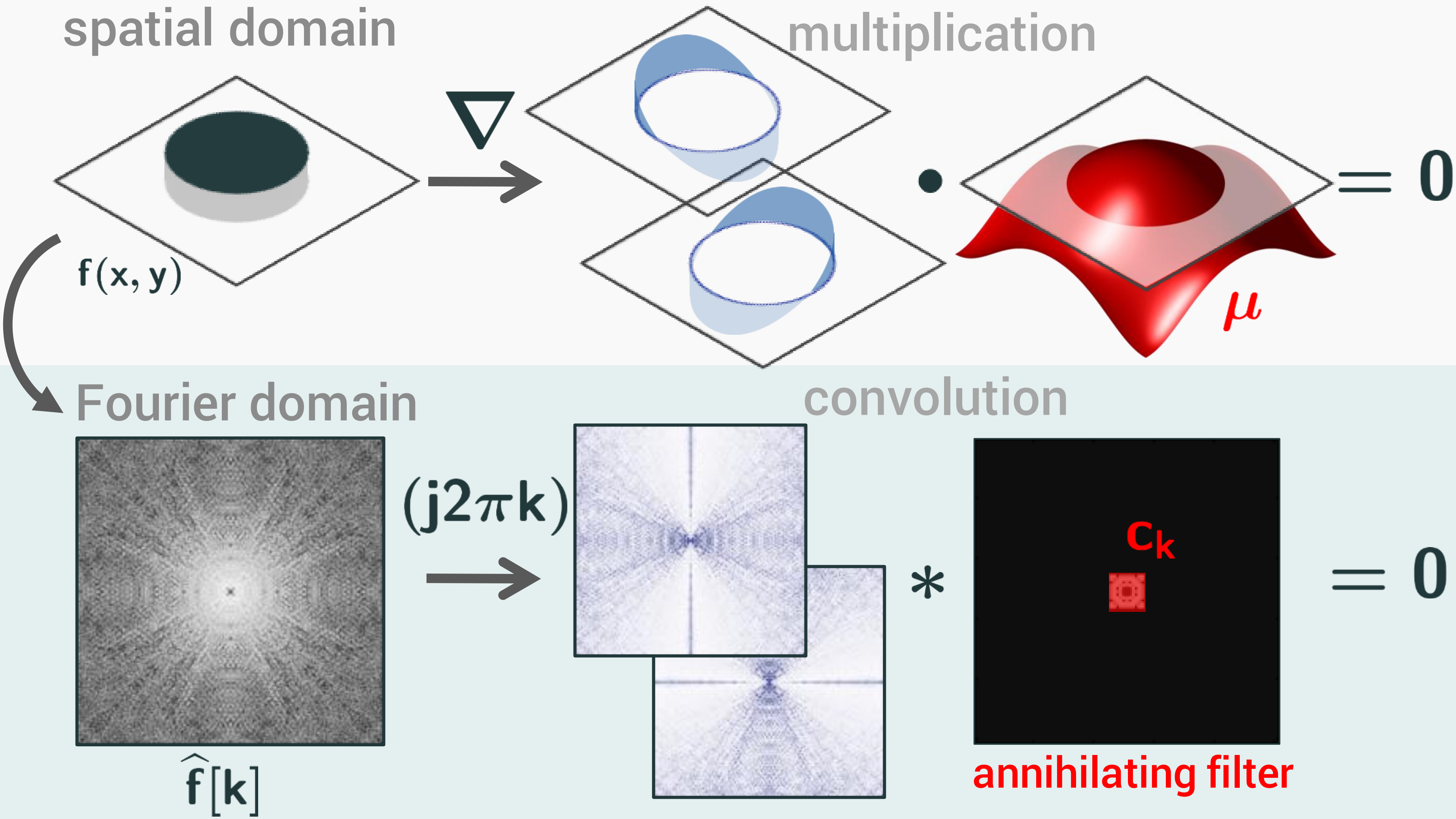
Challenges extending FRI to higher dimensions:
Singularities not isolated



Recall 1-D Case...

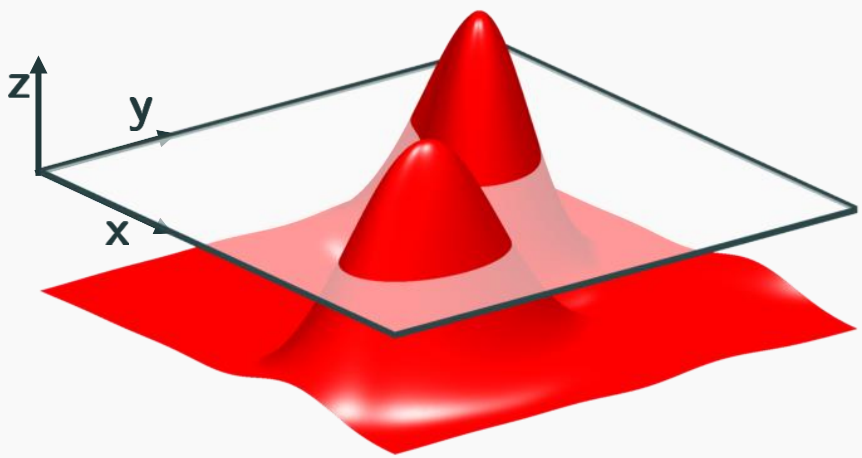


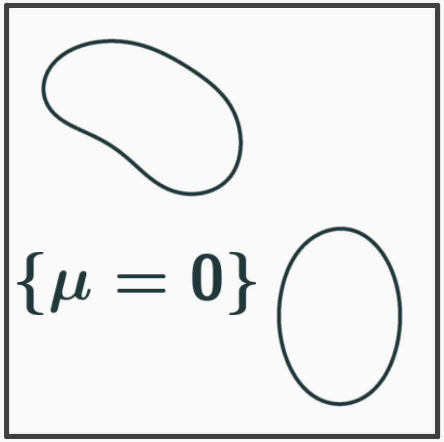
2-D PWC functions satisfy an annihilation relation



Annihilation relation: $\sum_k \widehat{\nabla f}[\ell - k] c_k = 0$

Can recover edge set when it is the
zero-set of a 2-D trigonometric polynomial
[Pan et al., 2014]



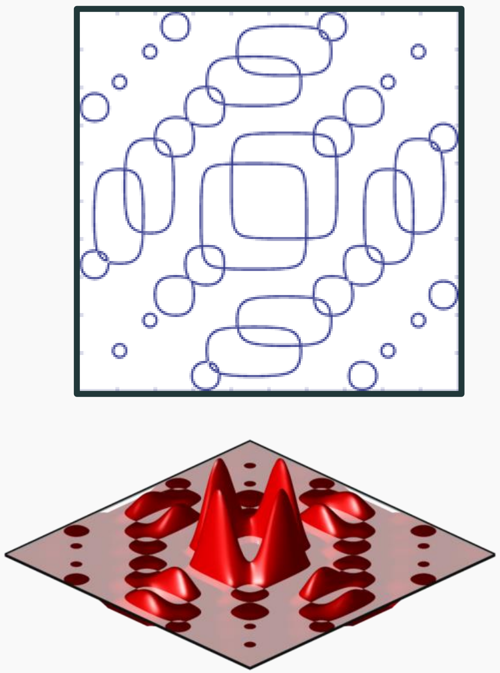


$$\mu(x, y) = \sum_{(k,l)\in\Lambda} c_{k,l}\, e^{j2\pi(kx+ly)}$$

"FRI Curve"

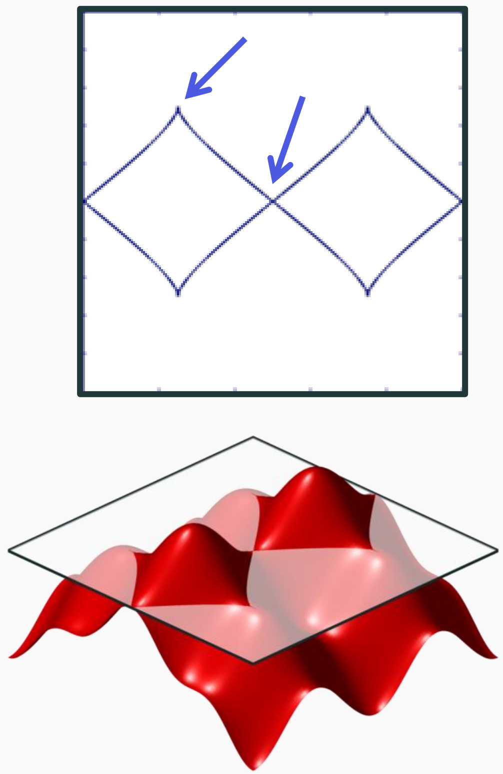
FRI curves can represent complicated edge geometries with few coefficients

Multiple curves & intersections



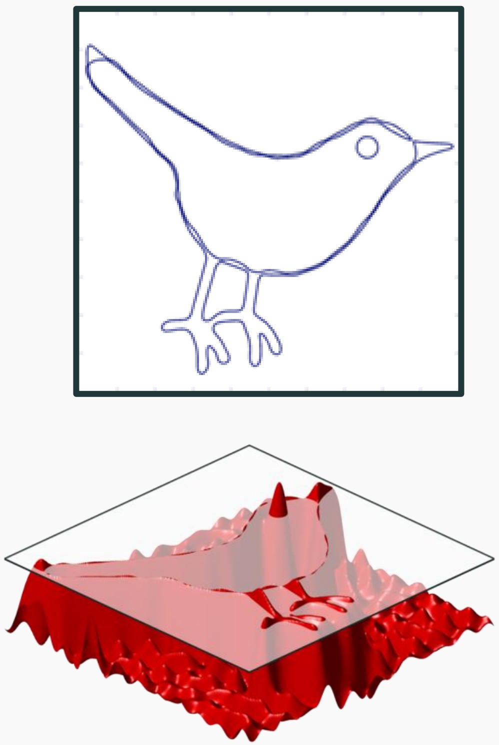
13x13 coefficients

Non-smooth points



7x9 coefficients

Approximate arbitrary curves



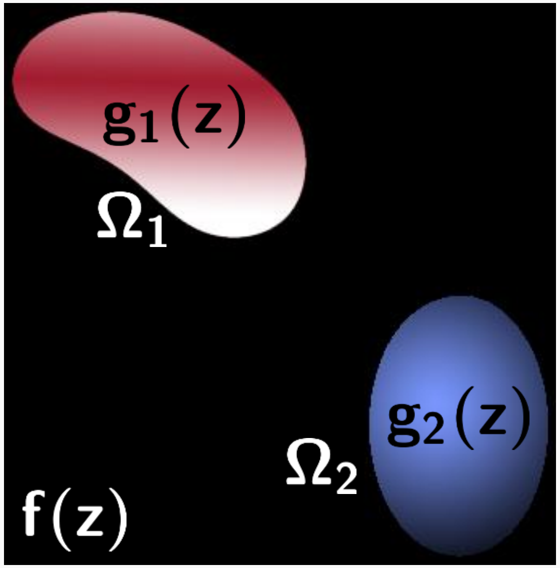
25x25 coefficients

We give an improved theoretical framework for higher dimensional FRI recovery

- *[Pan et al., 2014]* derived annihilation relation for piecewise complex analytic signal model

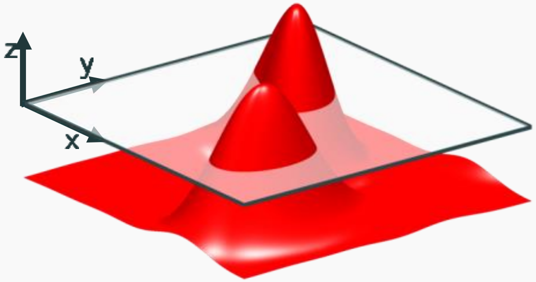
$$f(z) = \sum_{i=1}^{N} g_i(z) \cdot 1_{\Omega_i}(z)$$

$$\text{s.t.} \quad g_i \text{ analytic in } \Omega_i$$



- Not suitable for natural images
- 2-D only
- Recovery is ill-posed:

  Infinite DoF

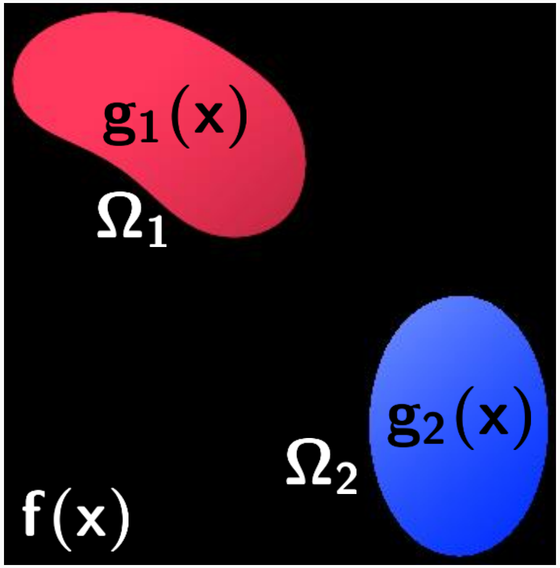


We give an improved theoretical framework for higher dimensional FRI recovery *[O. & Jacob, SampTA 2015]*

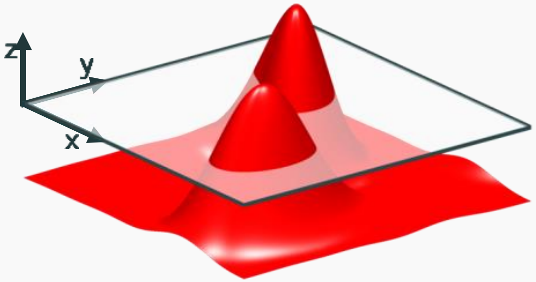
- Proposed model: piecewise smooth signals

$$f(x) = \sum_{i=1}^{N} g_i(x) \cdot 1_{\Omega_i}(x)$$

$$\text{s.t.} \quad g_i \text{ smooth in } \Omega_i$$



- Extends easily to n-D
- Provable sampling guarantees
- Fewer samples necessary for recovery

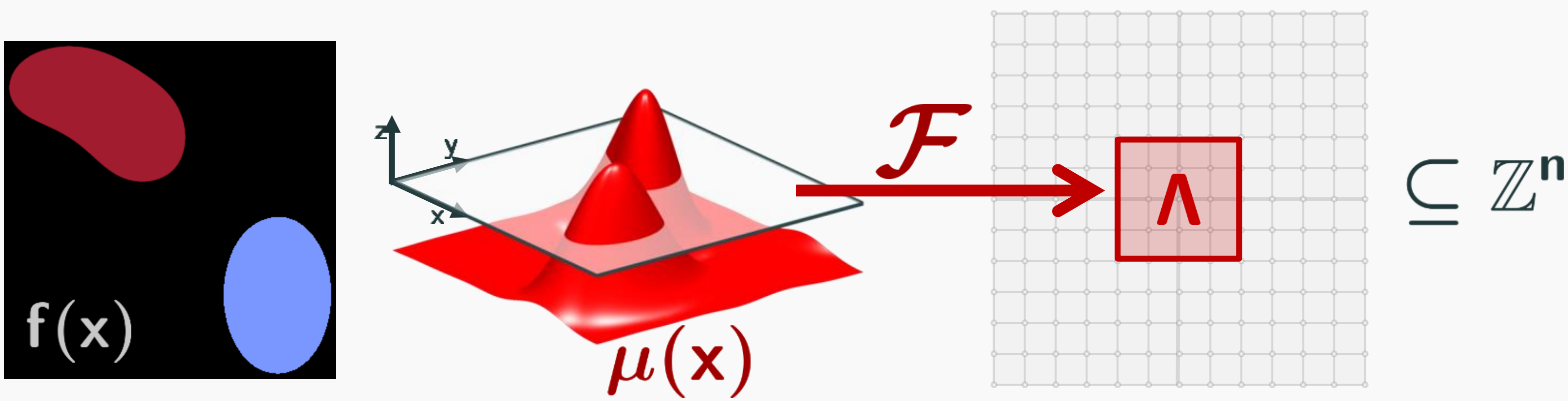


Annhilation relation for PWC signals

Prop: If *f* is PWC with edge set $\mathbf{E} \subseteq \{\mu = 0\}$ for μ bandlimited to $\mathbf{\Lambda}$ then

$$\sum_{\mathbf{k}\in\mathbf{\Lambda}} \widehat{\mu}[\mathbf{k}]\widehat{\partial \mathbf{f}}[\ell - \mathbf{k}] = \mathbf{0}, \quad \forall \ell \in \mathbb{Z}^{\mathbf{n}}$$

any 1st order partial derivative



Annhilation relation for PWC signals

Prop: If *f* is PWC with edge set $\mathbf{E} \subseteq \{\mu = \mathbf{0}\}$

for μ bandlimited to $\mathbf{\Lambda}$ then

$$\sum_{\mathbf{k}\in\mathbf{\Lambda}} \widehat{\mu}[\mathbf{k}]\widehat{\partial\mathbf{f}}[\ell - \mathbf{k}] = \mathbf{0}, \quad \forall \ell \in \mathbb{Z}^{\mathbf{n}}$$

any 1st order partial derivative

Proof idea:

Show $\mu \cdot \partial\mathbf{f} = \mathbf{0}$ as tempered distributions

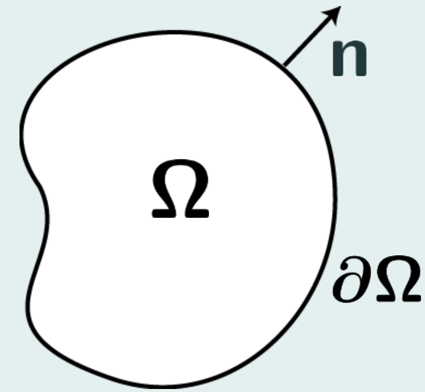
Use convolution theorem

Distributional derivative of indicator function:

smooth test function

$$
\begin{aligned}
\langle \partial_j \mathbf{1}_\Omega, \varphi \rangle &= -\langle \mathbf{1}_\Omega, \partial_j \varphi \rangle \\
&= -\int_\Omega \partial_j \varphi \, \mathbf{dx} \\
&= -\oint_{\partial\Omega} \varphi \, \mathbf{n_j} \, \mathbf{d}\sigma
\end{aligned}
$$

divergence theorem



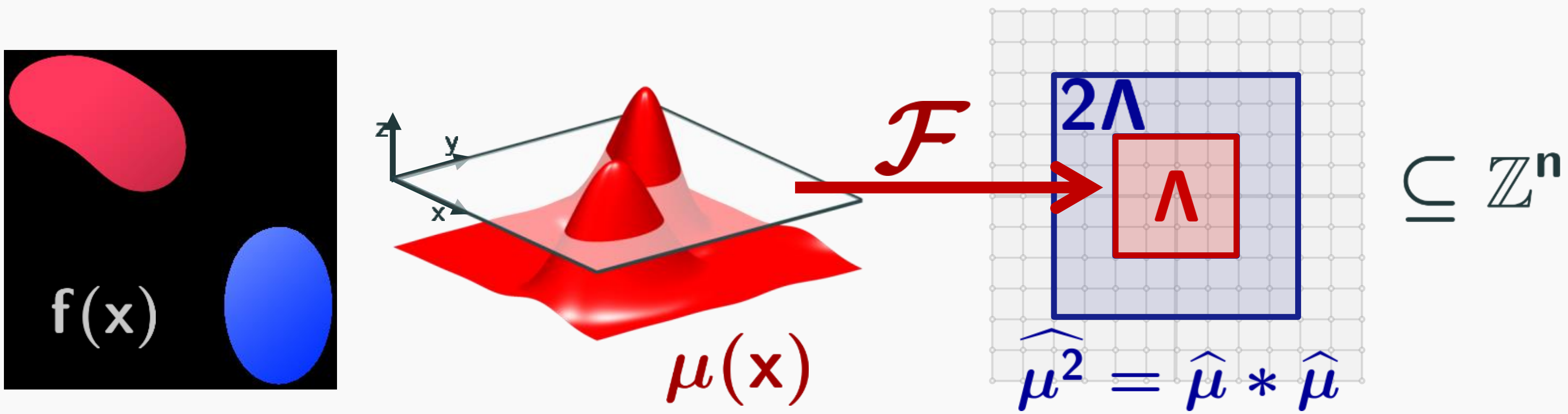
Weighted curve integral

Annhilation relation for PW linear signals

Prop: If **f** is PW linear, with edge set $\mathbf{E} \subseteq \{\mu = 0\}$ and μ bandlimited to **Λ** then

$$\sum_{\mathbf{k} \in 2\Lambda} \widehat{\mu^2}[\mathbf{k}]\,\widehat{\partial^2 \mathbf{f}}[\ell - \mathbf{k}] = 0, \quad \forall \ell \in \mathbb{Z}^{\mathbf{n}}$$

any 2nd order partial derivative



Annhilation relation for PW linear signals

Prop: If **f** is PW linear, with edge set $\mathbf{E} \subseteq \{\mu = 0\}$

and μ bandlimited to **Λ** then

$$\sum_{\mathbf{k}\in 2\boldsymbol{\Lambda}} \widehat{\mu^2}[\mathbf{k}]\,\widehat{\partial^2 \mathbf{f}}[\ell - \mathbf{k}] = \mathbf{0}, \quad \forall \ell \in \mathbb{Z}^{\mathbf{n}}$$

any 2nd order partial derivative

Proof idea: $\mathbf{f} = \mathbf{g} \cdot 1_\Omega, \quad \mathbf{g}$ **linear**

product rule x2 $\partial^2 \mathbf{f} = \cancel{\partial^2 \mathbf{g} \cdot 1_\Omega} + 2\partial \mathbf{g} \cdot \partial 1_\Omega + \mathbf{g} \cdot \partial^2 1_\Omega$

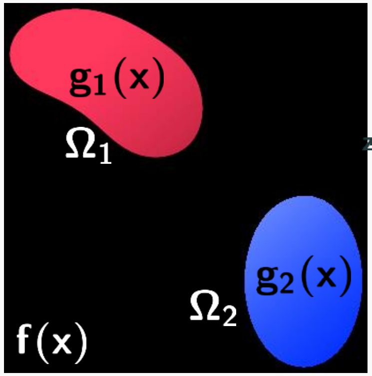
annihilated by μ^2

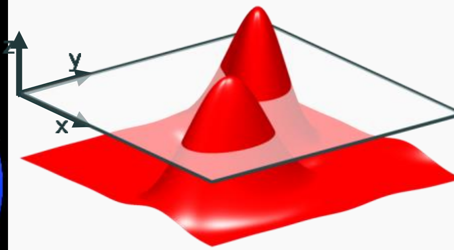
Can extend annihilation relation to a wide class of piecewise smooth images.

$$\mathbf{f}(\mathbf{x}) = \sum_{\mathbf{i}=1}^{\mathbf{N}} \mathbf{g}_{\mathbf{i}}(\mathbf{x}) \cdot 1_{\Omega_{\mathbf{i}}}(\mathbf{x})$$

$$\text{s.t.} \quad \mathbf{D}\mathbf{g}_{\mathbf{i}} = \mathbf{0} \text{ in } \Omega_{\mathbf{i}}$$

Any constant coeff. differential operator

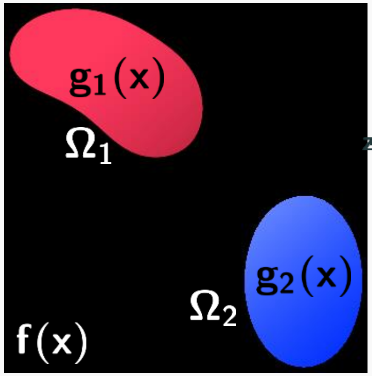


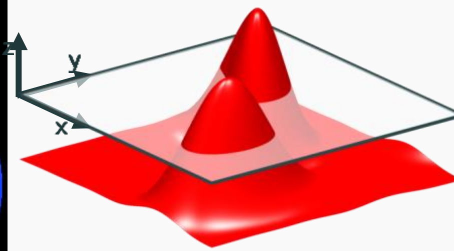


Can extend annihilation relation to a wide class of piecewise smooth images.

$$f(x) = \sum_{i=1}^{N} g_i(x) \cdot 1_{\Omega_i}(x)$$

$$\text{s.t.} \quad D g_i = 0 \text{ in } \Omega_i$$



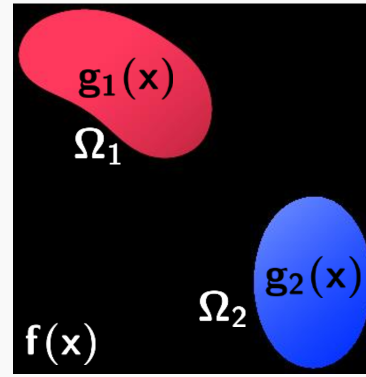


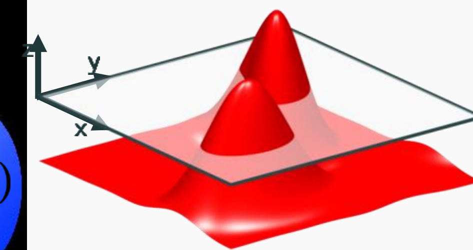
Signal Model:	Choice of Diff. Op.:	
PW Constant	$D = \nabla$	1^{st} order
PW Analytic*	$D = \partial_x + j\partial_y$	

Can extend annihilation relation to a wide class of **piecewise smooth** images.

$$f(x) = \sum_{i=1}^{N} g_i(x) \cdot 1_{\Omega_i}(x)$$

$$\text{s.t.} \quad D g_i = 0 \text{ in } \Omega_i$$



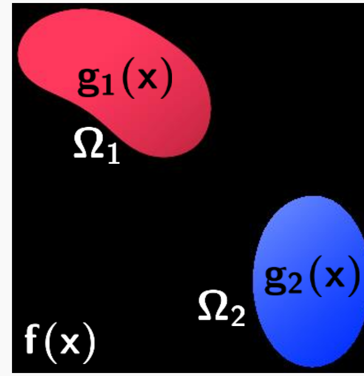


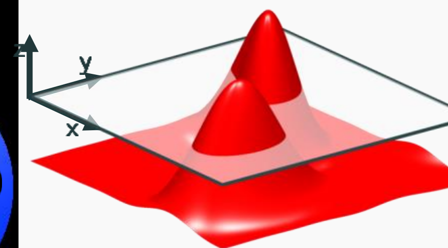
Signal Model:	Choice of Diff. Op.:	
PW Constant	$D = \nabla$	1st order
PW Analytic*	$D = \partial_x + j\partial_y$	1st order
PW Harmonic	$D = \Delta$	2nd order
PW Linear	$D = \{\partial_{xx}, \partial_{xy}, \partial_{yy}\}$	2nd order

Can extend annihilation relation to a wide class of **piecewise smooth** images.

$$f(x) = \sum_{i=1}^{N} g_i(x) \cdot 1_{\Omega_i}(x)$$

$$\text{s.t.} \quad D g_i = 0 \text{ in } \Omega_i$$





Signal Model:	Choice of Diff. Op.:	
PW Constant	$D = \nabla$	1st order
PW Analytic*	$D = \partial_x + j\partial_y$	
PW Harmonic	$D = \Delta$	2nd order
PW Linear	$D = \{\partial_{xx}, \partial_{xy}, \partial_{yy}\}$	
PW Polynomial	$D = \{\partial^\alpha\}_{\lvert\alpha\rvert=n}$	nth order

Sampling theorems:

Necessary and sufficient number of Fourier samples for

1. Unique recovery of edge set/annihilating polynomial
2. Unique recovery of full signal given edge set
   - Not possible for PW analytic, PW harmonic, etc.
   - Prefer PW polynomial models

➔ Focus on 2-D PW constant signals

Challenges to proving uniqueness

1-D FRI Sampling Theorem [Vetterli et al., 2002]:

A continuous-time PWC signal with K jumps can be uniquely recovered from 2K+1 uniform Fourier samples.

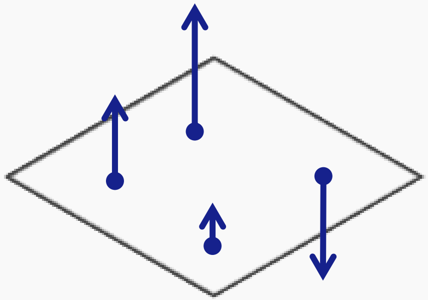
Proof (a la Prony's Method):

Form Toeplitz matrix $\mathbf{T}$ from samples, use uniqueness of Vandermonde decomposition: $\mathbf{T} = \mathbf{V}\mathbf{D}\mathbf{V}^H$

"Caratheodory Parametrization"

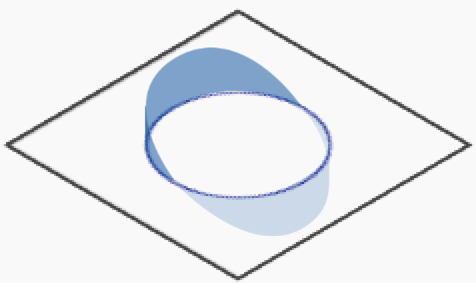
Challenges proving uniqueness, cont.

Extends to *n*-D if singularities isolated [Sidiropoulos, 2001]



$$\xrightarrow{\mathcal{F}} \quad \hat{\mathbf{f}}[\mathbf{k}] = \sum_{\mathbf{i}} \mathbf{a_i} \mathbf{e}^{-\mathbf{j}2\pi\mathbf{k}\cdot\mathbf{x_i}}$$

Not true in our case--singularities supported on curves:



$$\xrightarrow{\mathcal{F}} \quad \widehat{\nabla \mathbf{f}}[\mathbf{k}] = \oint_{\partial\Omega} \mathbf{e}^{-\mathbf{j}2\pi\mathbf{k}\cdot\mathbf{x}} \mathbf{n}\, \mathbf{ds}$$

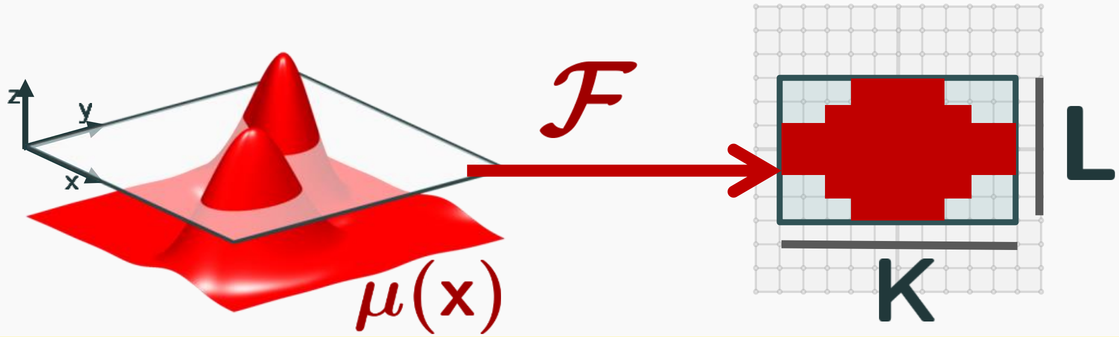
Requires new techniques:

– Spatial domain interpretation of annihilation relation

– Algebraic geometry of trigonometric polynomials

Minimal (Trigonometric) Polynomials

Define $\mathbf{deg}(\mu) = (\mathsf{K}, \mathsf{L})$ to be the dimensions of the smallest rectangle containing the Fourier support of μ



Prop: Every zero-set of a trig. polynomial **C** with no isolated points has a *unique* real-valued trig. polynomial μ_0 of minimal degree such that if $\mathsf{C} = \{\mu = 0\}$ then $\mathbf{deg}(\mu_0) \leq \mathbf{deg}(\mu)$ and $\mu = \gamma \cdot \mu_0$

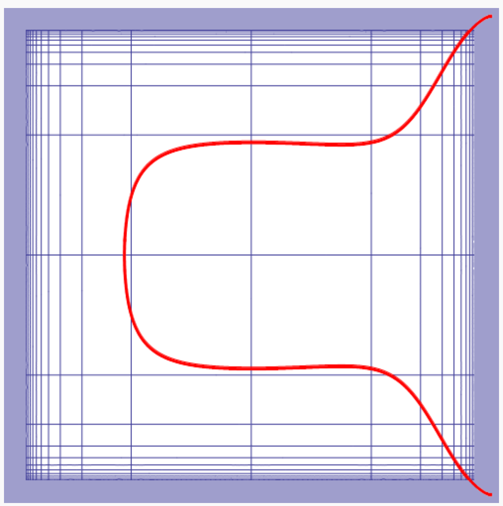
Degree of min. poly. = analog of sparsity/complexity of edge set

Proof idea: Pass to Real Algebraic Plane Curves

Zero-sets of trig polynomials of degree (K,L)

are in 1-to-1 correspondence with

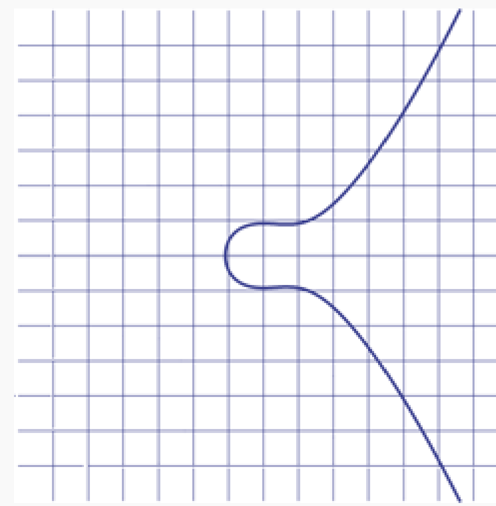
Real algebraic plane curves of degree (K,L)



$\mathbb{T}^2$

$$\mu(z, w) = 0;$$
$$|z| = |w| = 1$$

Conformal change of variables



$\mathbb{R}^2$

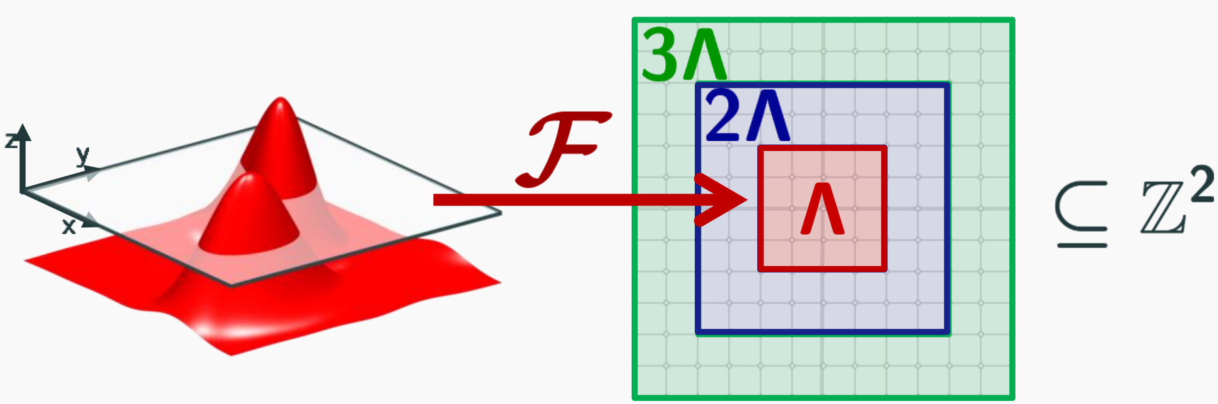
$$p(t, s) = 0;$$
$$t, s \in \mathbb{R}^2$$

Uniqueness of edge set recovery

Theorem: If $\mathbf{f}$ is PWC* with edge set $\mathbf{E} = \{\mu = 0\}$ with μ minimal and bandlimited to Λ then $\mathbf{c} = \widehat{\mu}$ is the unique solution to

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

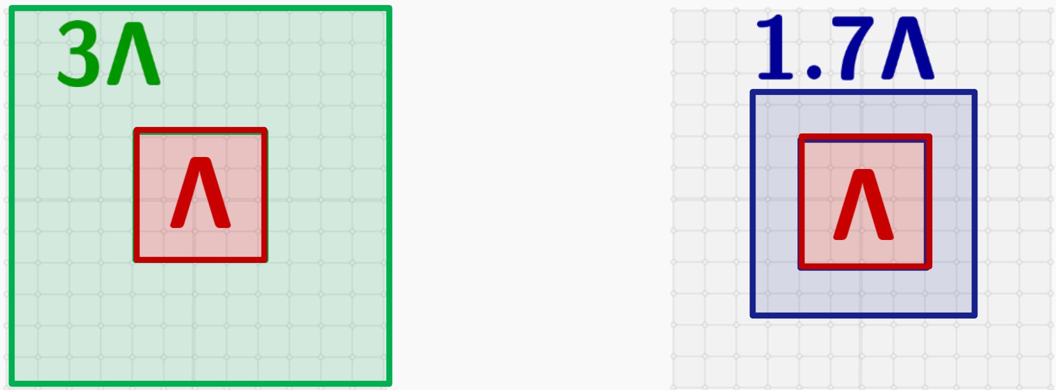
*Some geometric restrictions apply



Requires samples of $\widehat{\mathbf{f}}$ in 3Λ to build equations

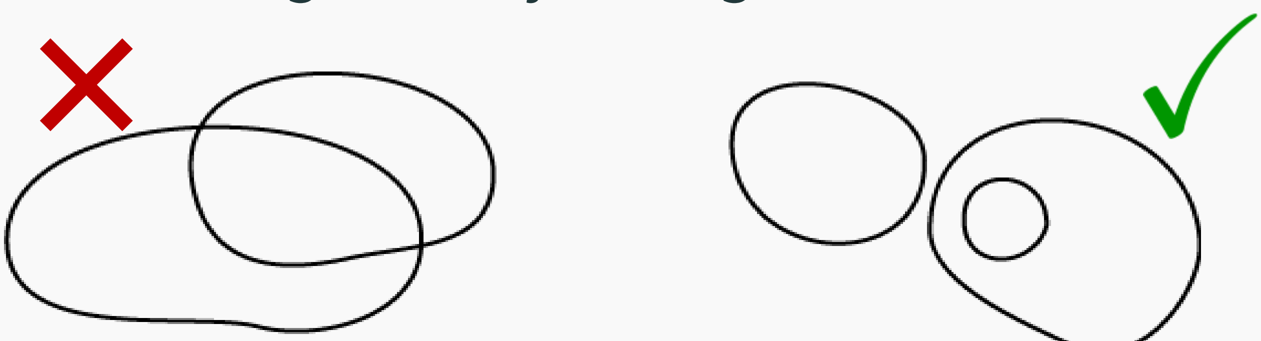
Current Limitations to Uniqueness Theorem

- Gap between necessary and sufficient # of samples:



Sufficient Necessary

- Restrictions on geometry of edge sets: *non-intersecting*



$\{\mu = 0\}$ $\{\mu = 0\}$

Uniqueness of signal (given edge set)

Theorem: If **f** is PWC* with edge set $\mathbf{E} = \{\mu = \mathbf{0}\}$ with μ minimal and bandlimited to $\mathbf{\Lambda}$ then $\mathbf{g} = \mathbf{f}$ is the unique solution to

$$\mu \cdot \nabla \mathbf{g} = \mathbf{0} \ \text{ s.t. } \ \widehat{\mathbf{f}}[\mathbf{k}] = \widehat{\mathbf{g}}[\mathbf{k}], \mathbf{k} \in \mathbf{\Gamma}$$

when the sampling set $\mathbf{\Gamma} \supseteq 3\mathbf{\Lambda}$

*Some geometric restrictions apply

Uniqueness of signal (given edge set)

Theorem: If **f** is PWC* with edge set $\mathbf{E} = \{\mu = 0\}$ with μ minimal and bandlimited to $\mathbf{\Lambda}$ then $\mathbf{g} = \mathbf{f}$ is the unique solution to

$$\mu \cdot \nabla \mathbf{g} = \mathbf{0} \quad \text{s.t.} \quad \widehat{\mathbf{f}}[\mathbf{k}] = \widehat{\mathbf{g}}[\mathbf{k}], \mathbf{k} \in \mathbf{\Gamma}$$

when the sampling set $\mathbf{\Gamma} \supseteq 3\mathbf{\Lambda}$

*Some geometric restrictions apply

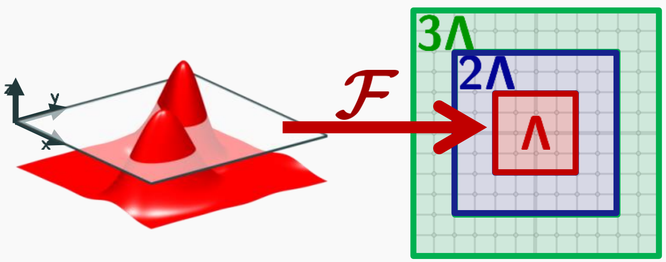
Equivalently,

$$\mathbf{f} = \arg\min_{\mathbf{g}} ||\mu \cdot \nabla \mathbf{g}|| \quad \text{s.t.} \quad \widehat{\mathbf{f}}[\mathbf{k}] = \widehat{\mathbf{g}}[\mathbf{k}], \mathbf{k} \in \mathbf{\Gamma}$$

Summary of Proposed *Off-the-Grid Framework*

- Extend Prony/FRI methods to recover multidimensional singularities (curves, surfaces)

- Unique recovery from *uniform* Fourier samples: # of samples proportional to degree of edge set polynomial

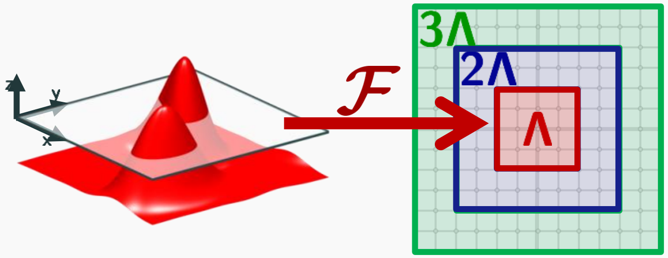


- Two-stage recovery
  1. Recover edge set by solving linear system
  2. Recover amplitudes

Summary of Proposed *Off-the-Grid Framework*

- Extend Prony/FRI methods to recover multidimensional singularities (curves, surfaces)

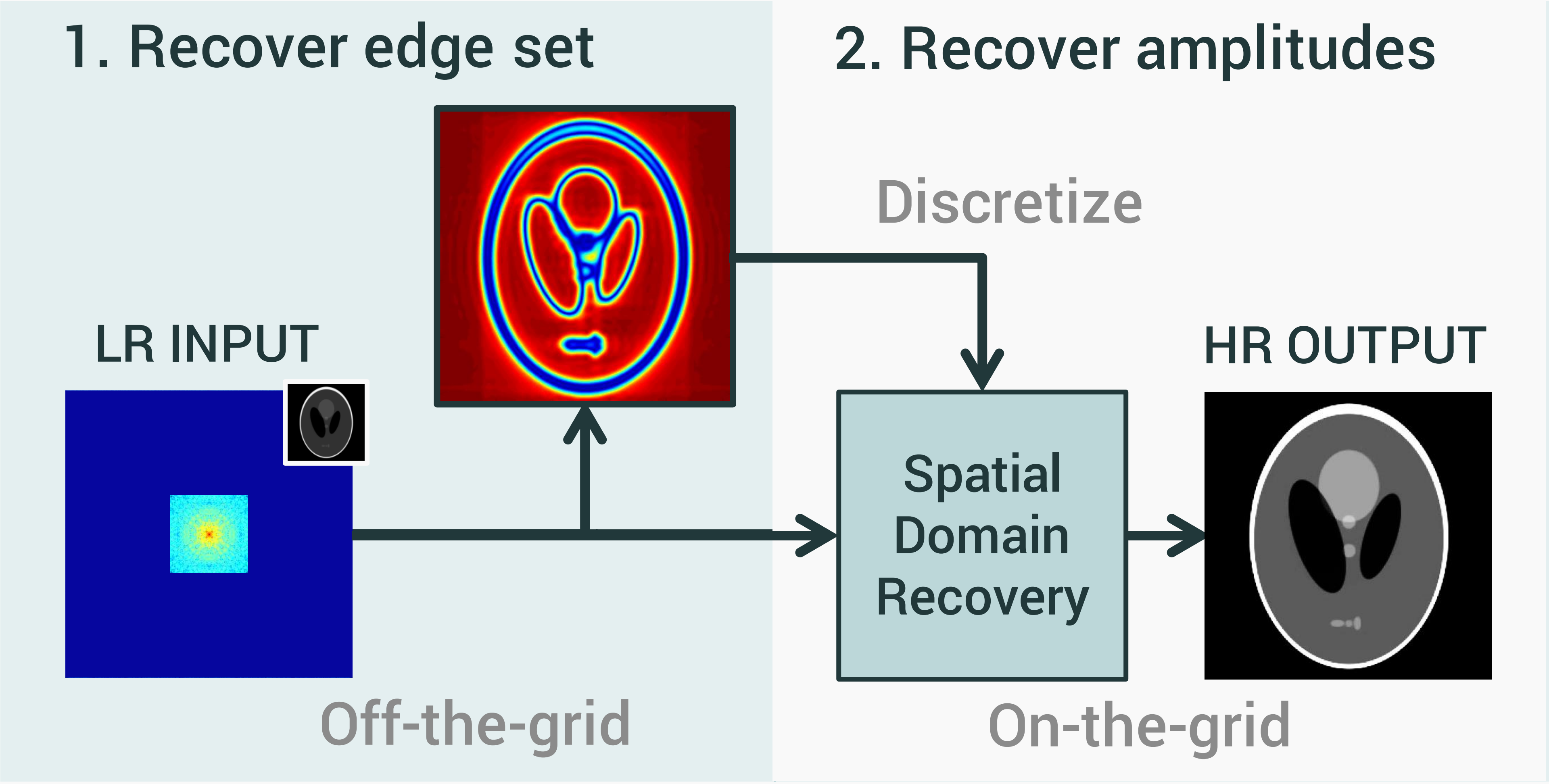
- Unique recovery from *uniform* Fourier samples: # of samples proportional to degree of edge set polynomial



- Two-stage recovery
  1. Recover edge set by solving linear system (Robust?)
  2. Recover amplitudes (How?)

New *Off-the-Grid* Imaging Framework: *Algorithms*

Two-stage Super-resolution MRI Using Off-the-Grid Piecewise Constant Signal Model [O. & Jacob, ISBI 2015]

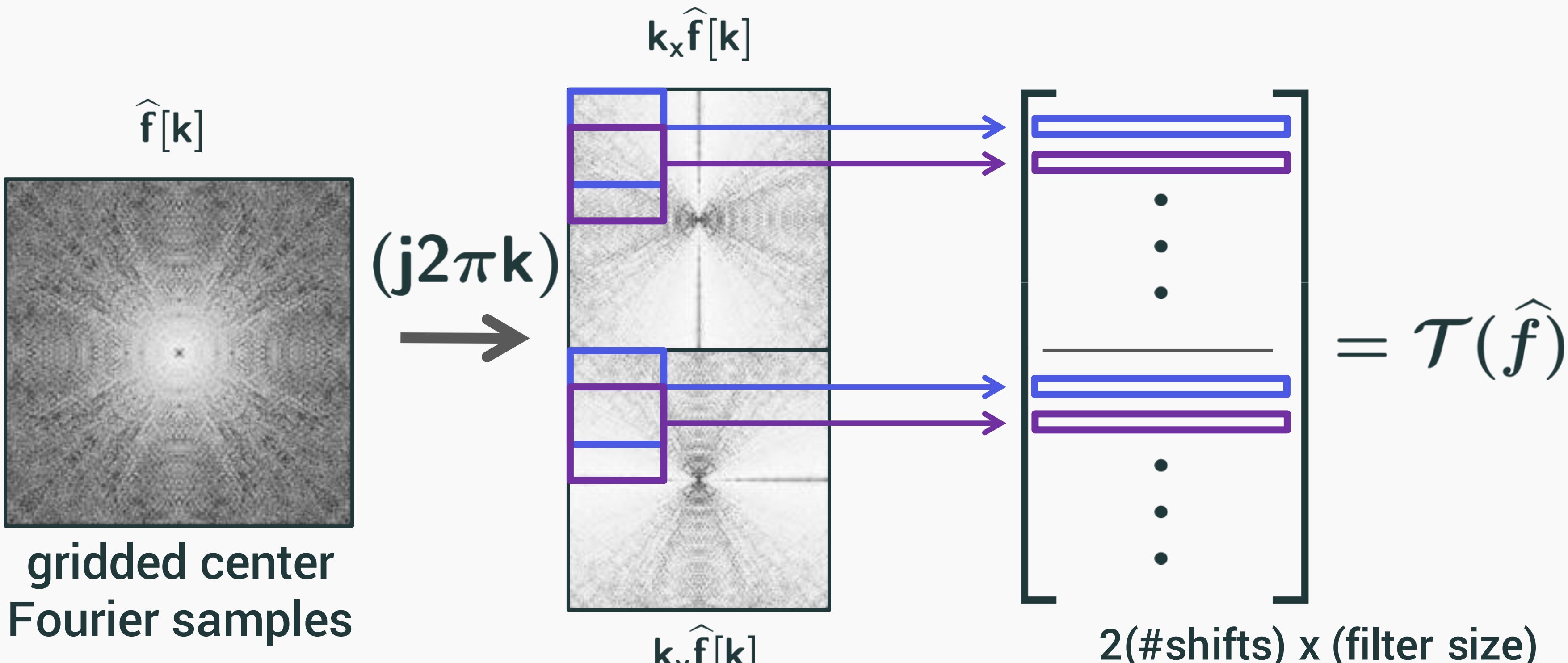


Matrix representation of annihilation

$$\mathcal{T}(\widehat{f})\,\boldsymbol{c} = 0$$

2-D convolution matrix
(block Toeplitz)

vector of filter coefficients

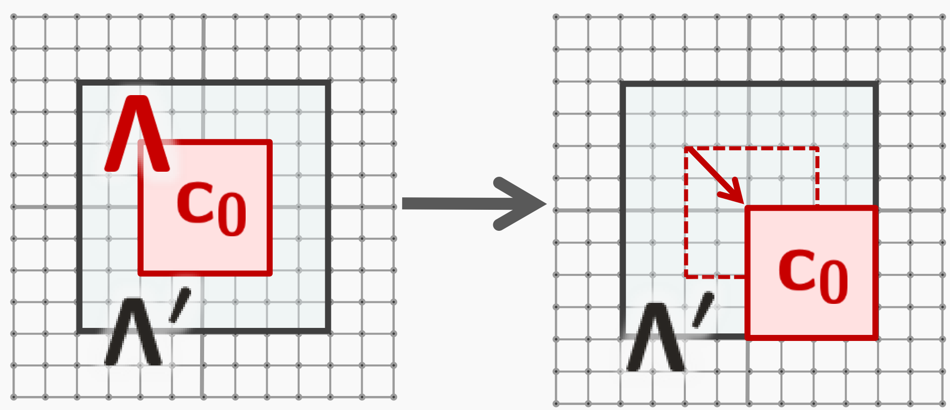


Basis of algorithms:
Annihilation matrix is low-rank

Prop: If the level-set function is bandlimited to Λ and the assumed filter support $\Lambda' \supset \Lambda$ then

$$\mathrm{rank}[\mathcal{T}(\widehat{\mathbf{f}})] \leq |\Lambda'| - (\#\text{shifts } \Lambda \text{ in } \Lambda')$$

Fourier domain



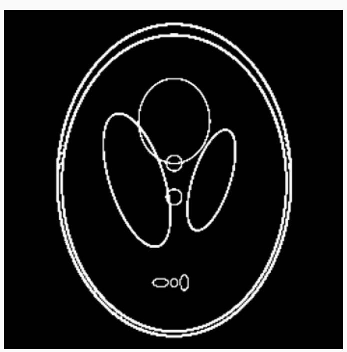
Spatial domain $\mu(x, y) \longrightarrow e^{j2\pi(kx+ly)}\mu(x, y)$

Basis of algorithms:
Annihilation matrix is low-rank

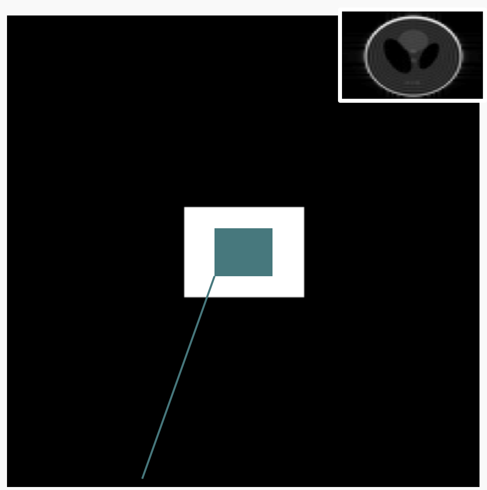
Prop: If the level-set function is bandlimited to Λ and the assumed filter support $\Lambda' \supset \Lambda$ then

$$\mathrm{rank}[\mathcal{T}(\widehat{\mathbf{f}})] \leq |\Lambda'| - (\#\mathrm{shifts}\ \Lambda\ \mathrm{in}\ \Lambda')$$

Example:
Shepp-Logan



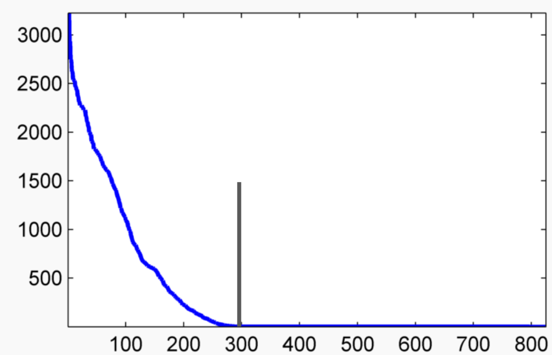
Fourier domain



Assumed filter: 33x25

Samples: 65x49

$$\sigma(\mathcal{T}(\widehat{f}))$$



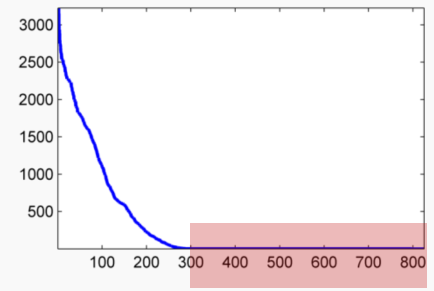
Rank $\approx$ 300

Stage 1: Robust annihilting filter estimation

1. Compute SVD

$$\mathcal{T}(\widehat{\mathbf{f}}) = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$$

$$\sigma(\mathcal{T}(\widehat{\mathbf{f}}))$$

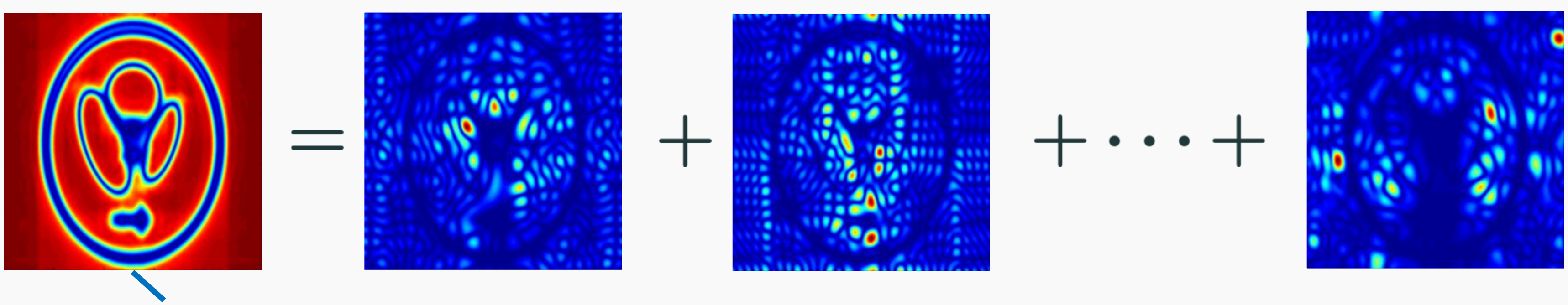


2. Identify **null space**

$$\mathbf{V} = [\mathbf{V}_S \ \mathbf{V}_N], \qquad \mathbf{V}_N = [\mathbf{c}_1, \ldots, \mathbf{c}_n]$$

3. Compute sum-of-squares average

$$\mu = |\mathcal{F}^{-1}\mathbf{c}_1|^2 + |\mathcal{F}^{-1}\mathbf{c}_2|^2 + \cdots + |\mathcal{F}^{-1}\mathbf{c}_n|^2$$



Recover common zeros

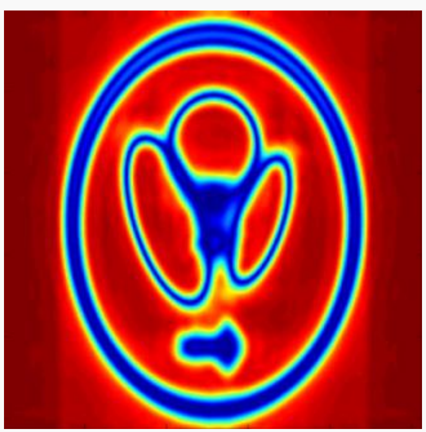
Stage 2: Weighted TV Recovery

$$\mathbf{f} = \arg\min_{\mathbf{g}} \|\mu \cdot \nabla \mathbf{g}\|_1 \quad \text{s.t.} \quad \widehat{\mathbf{f}}[\mathbf{k}] = \widehat{\mathbf{g}}[\mathbf{k}], \mathbf{k} \in \Gamma$$

discretize

relax

$$\min_{x} \sum_{i} w_i \cdot |(Dx)_i| + \lambda \|Ax - b\|^2$$



Edge weights

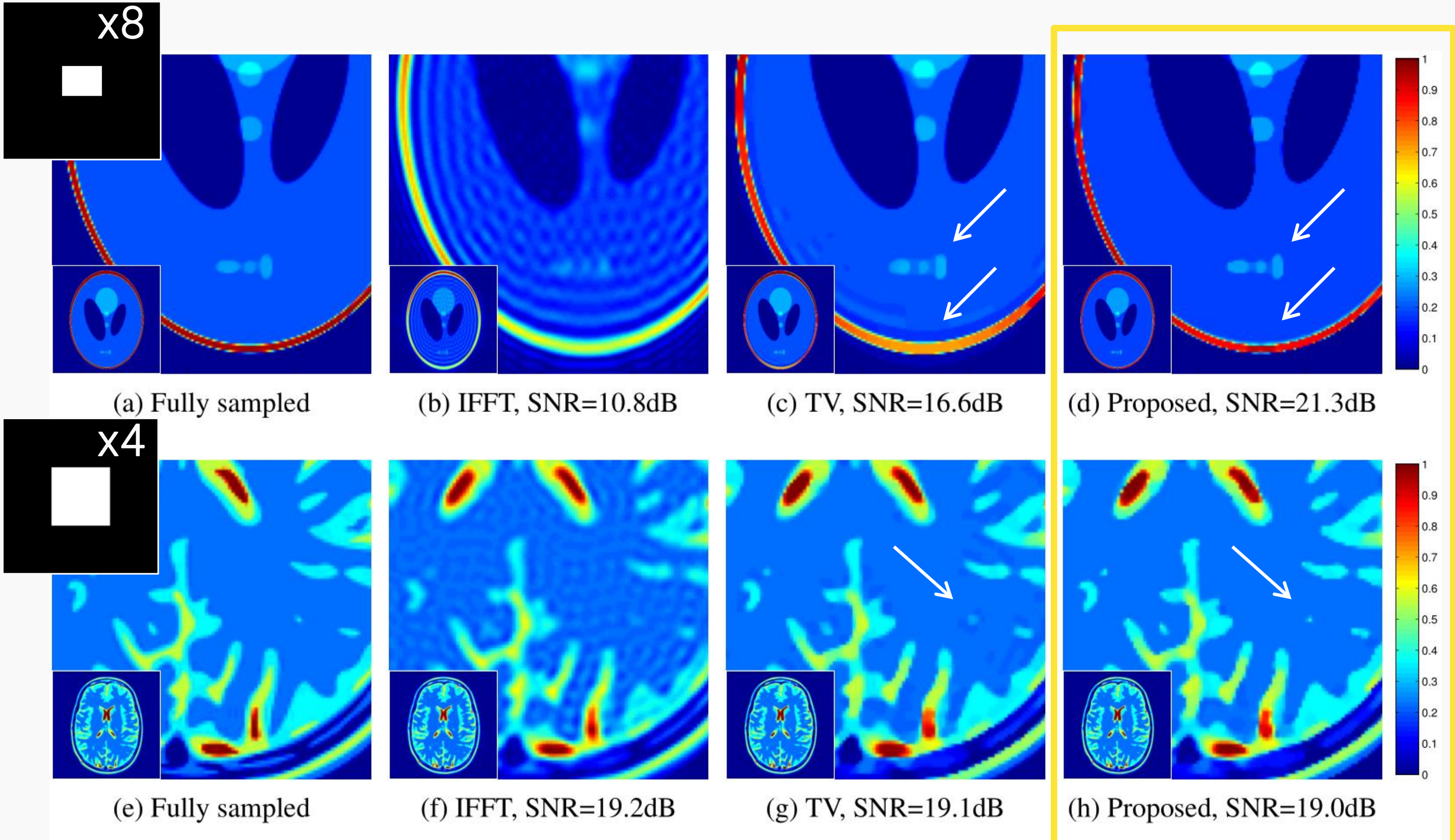
x = discrete spatial domain image

D = discrete gradient

A = Fourier undersampling operator

b = k-space samples

Super-resolution of MRI Medical Phantoms

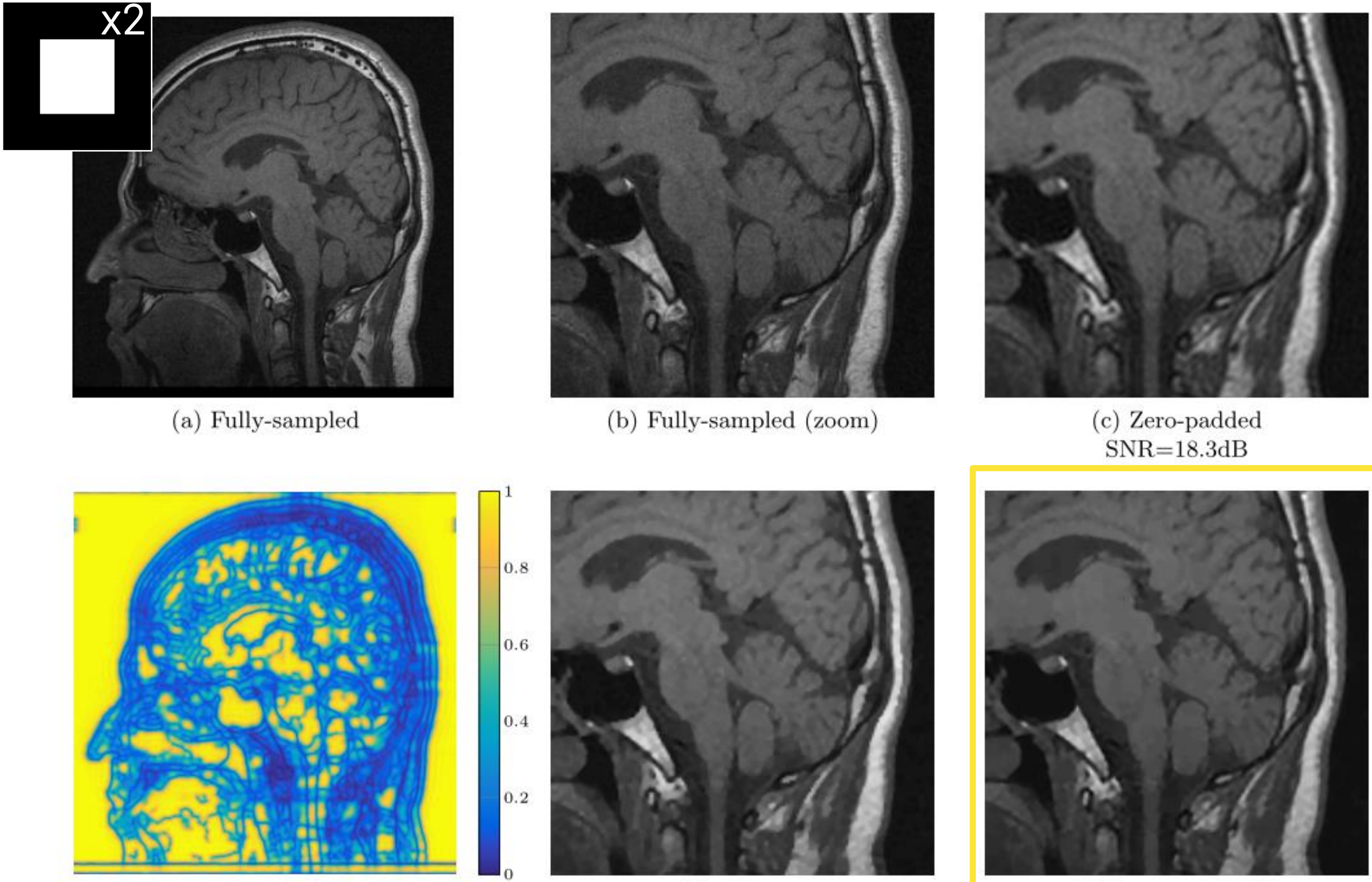


(a) Fully sampled (b) IFFT, SNR=10.8dB (c) TV, SNR=16.6dB (d) Proposed, SNR=21.3dB

(e) Fully sampled (f) IFFT, SNR=19.2dB (g) TV, SNR=19.1dB (h) Proposed, SNR=19.0dB

Analytical phantoms from *[Guerquin-Kern, 2012]*

Super-resolution of Real MRI Data



(a) Fully-sampled

(b) Fully-sampled (zoom)

(c) Zero-padded
SNR=18.3dB

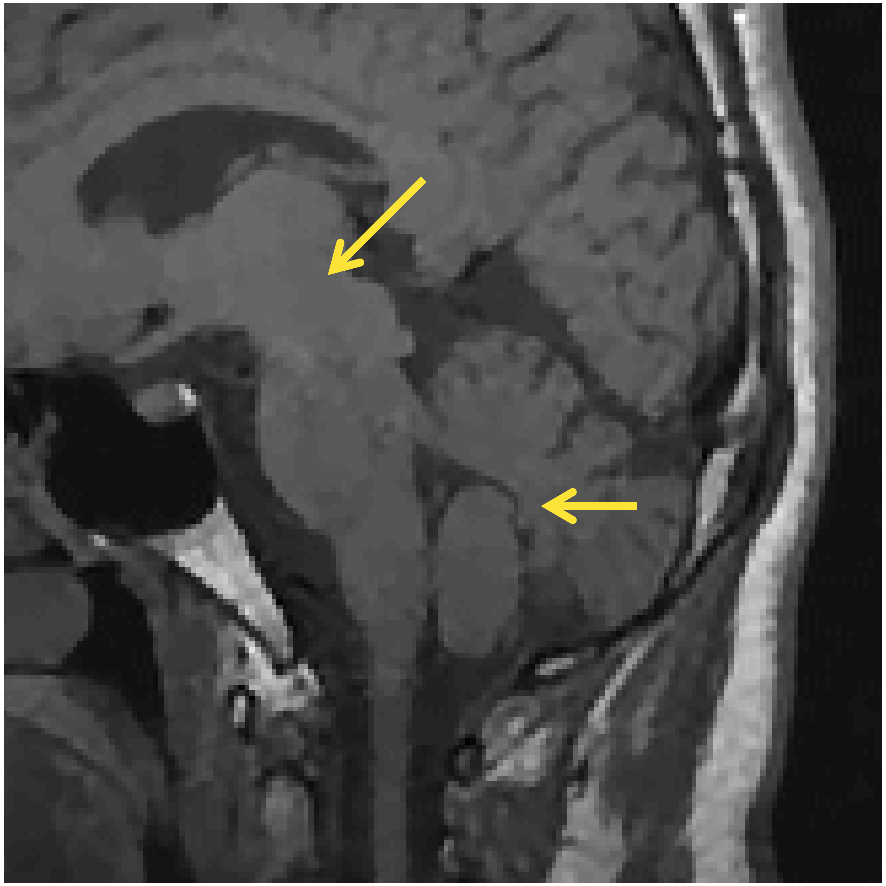
(d) Edge set estimate
(65×65 coefficients)

(e) TV reg.
SNR=18.5dB

(f) Proposed, LSLP
SNR=18.9dB

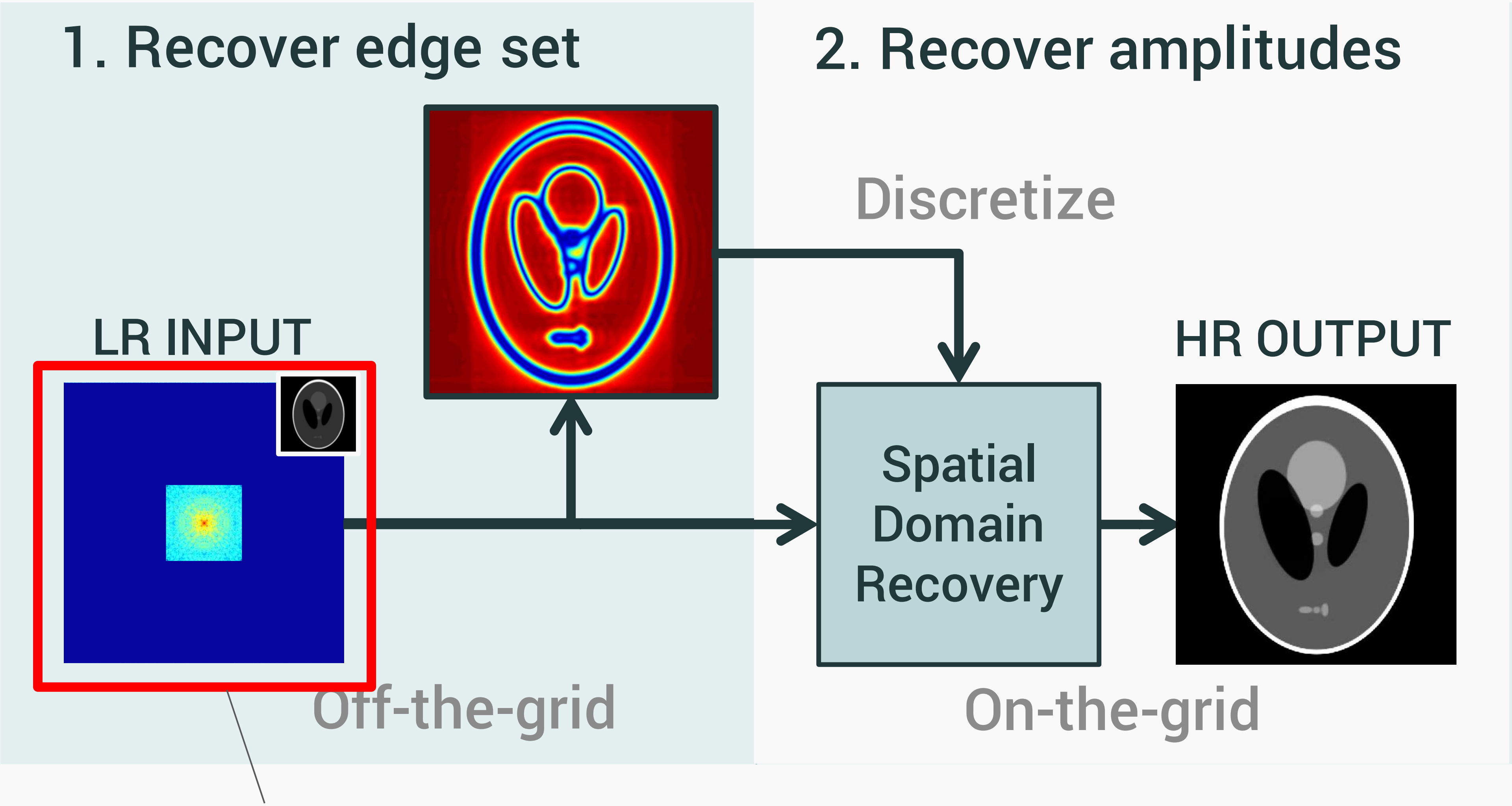
Super-resolution of Real MRI Data (Zoom)

(e) TV reg.
SNR=18.5dB



(f) Proposed, LSLP
SNR=18.9dB

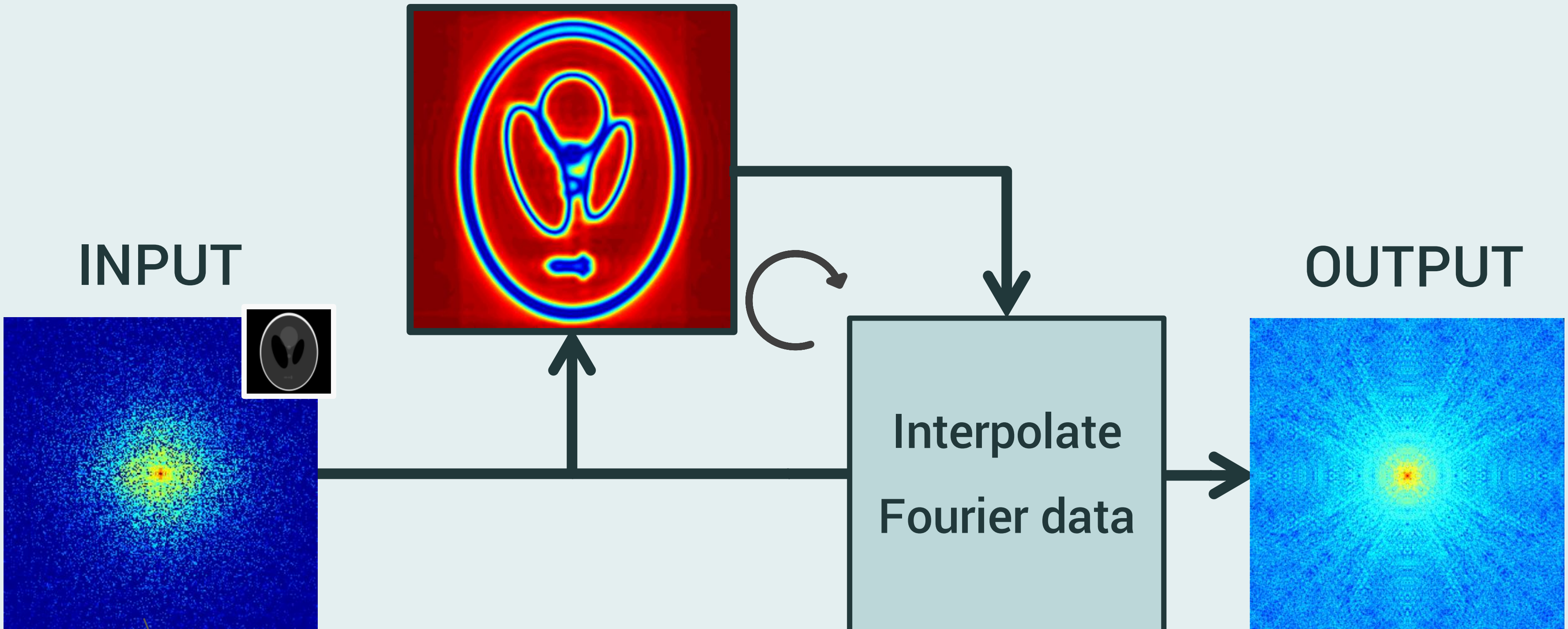
Two Stage Algorithm



Need uniformly sampled region!

One Stage Algorithm [O. & Jacob, SampTA 2015]

Jointly estimate edge set and amplitudes



Off-the-grid

Accommodate random samples

Pose recovery as a one-stage structured low-rank matrix completion problem

Recall: $\mathcal{T}(\widehat{\mathbf{f}})$ low rank $\leftrightarrow$ $\mathbf{f}$ piecewise constant

Toeplitz-like matrix built from Fourier data

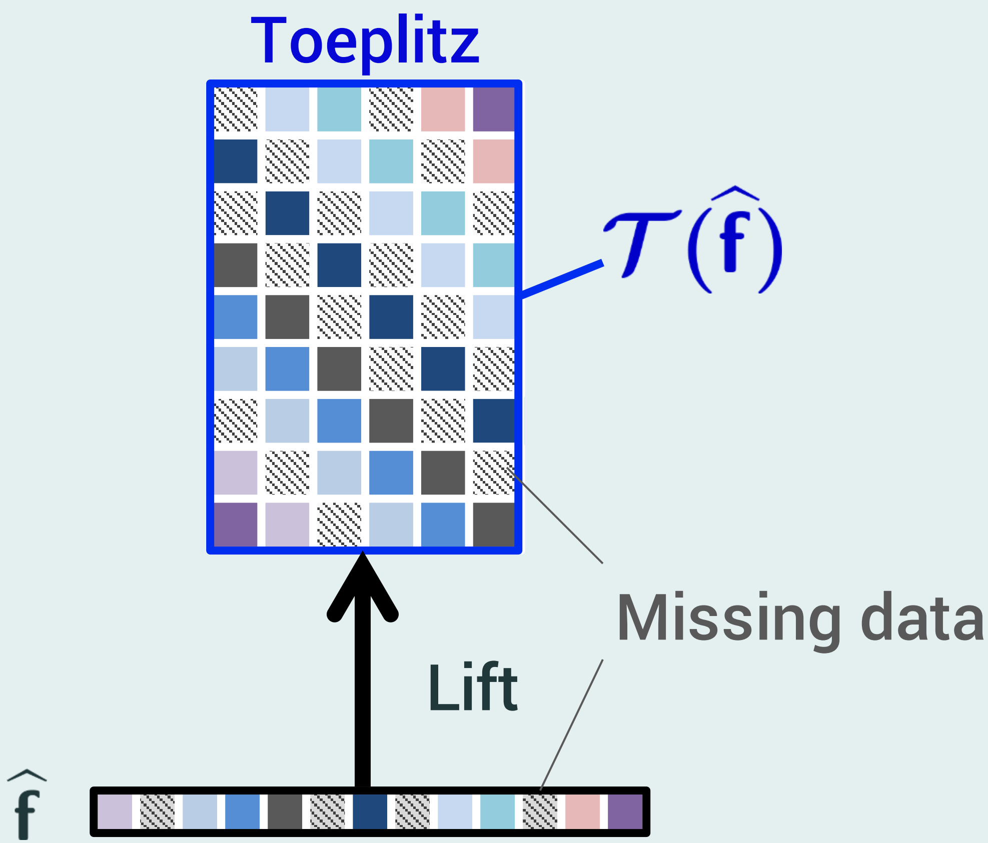
Pose recovery as a one-stage structured low-rank matrix completion problem

$$\min_{\widehat{\mathbf{f}}} \quad \mathrm{rank}[\mathcal{T}(\widehat{\mathbf{f}})] \quad \textbf{s.t.} \quad \widehat{\mathbf{f}}[\mathbf{k}] = \widehat{\mathbf{b}}[\mathbf{k}], \mathbf{k} \in \boldsymbol{\Gamma}$$

Pose recovery as a one-stage
structured low-rank matrix completion problem

$$\min_{\widehat{\mathbf{f}}} \quad \mathrm{rank}[\mathcal{T}(\widehat{\mathbf{f}})] \quad \textbf{s.t.} \quad \widehat{\mathbf{f}}[\mathbf{k}] = \widehat{\mathbf{b}}[\mathbf{k}], \mathbf{k} \in \mathbf{\Gamma}$$

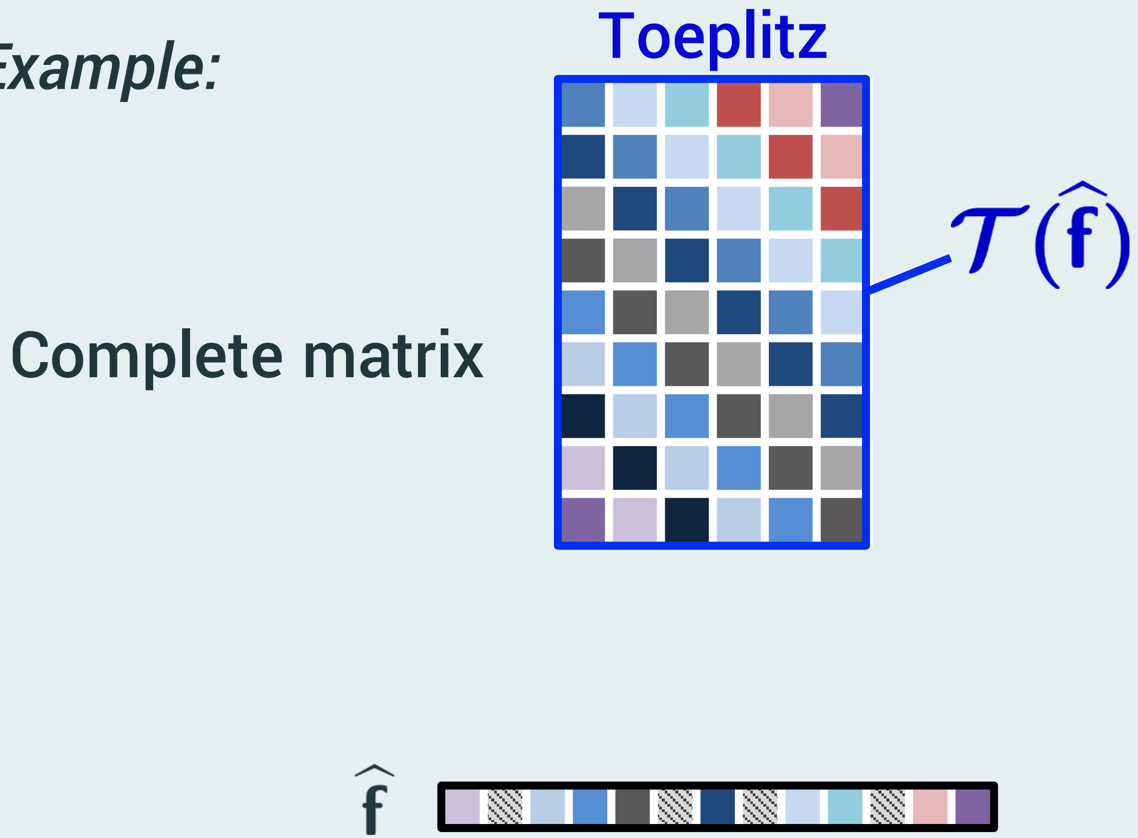
1-D Example:



Pose recovery as a one-stage structured low-rank matrix completion problem

$$\min_{\widehat{\mathbf{f}}} \quad \mathrm{rank}[\mathcal{T}(\widehat{\mathbf{f}})] \quad \textbf{s.t.} \quad \widehat{\mathbf{f}}[\mathbf{k}] = \widehat{\mathbf{b}}[\mathbf{k}], \mathbf{k} \in \mathbf{\Gamma}$$

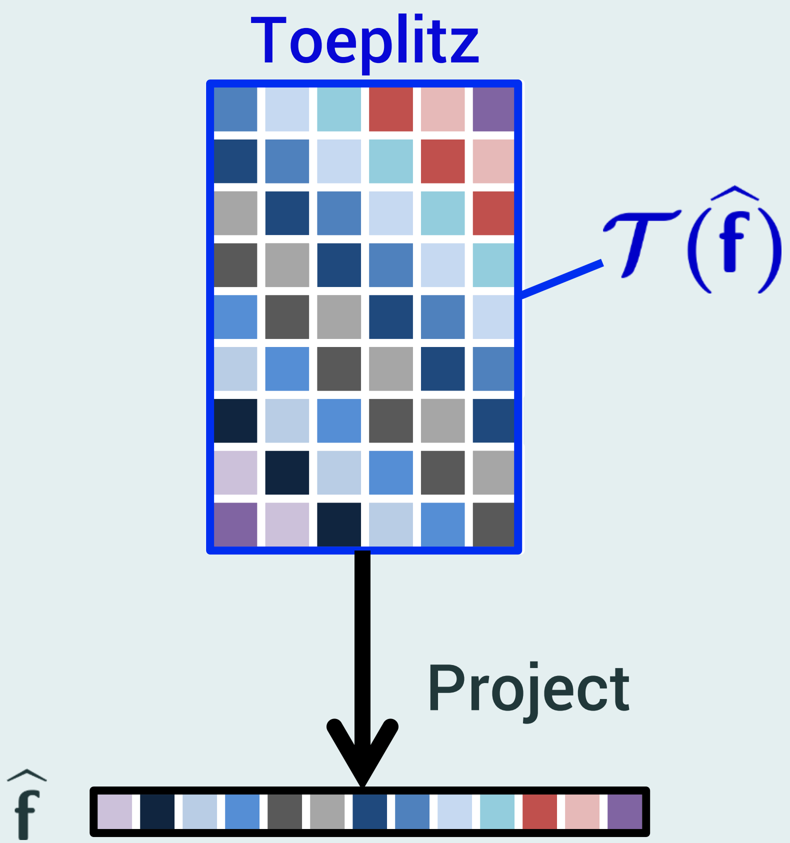
1-D Example:



Pose recovery as a one-stage **structured low-rank matrix completion** problem

$$\min_{\widehat{\mathbf{f}}} \quad \mathrm{rank}[\mathcal{T}(\widehat{\mathbf{f}})] \quad \textbf{s.t.} \quad \widehat{\mathbf{f}}[\mathbf{k}] = \widehat{\mathbf{b}}[\mathbf{k}], \mathbf{k} \in \boldsymbol{\Gamma}$$

1-D Example:



Pose recovery as a one-stage
structured low-rank matrix completion problem

$$\min_{\widehat{\mathbf{f}}} \quad \mathrm{rank}[\mathcal{T}(\widehat{\mathbf{f}})] \quad \textbf{s.t.} \quad \widehat{\mathbf{f}}[\mathbf{k}] = \widehat{\mathbf{b}}[\mathbf{k}], \mathbf{k} \in \boldsymbol{\Gamma}$$

NP-Hard!

Pose recovery as a one-stage
structured low-rank matrix completion problem

$$\min_{\widehat{\mathbf{f}}} \quad \mathrm{rank}[\mathcal{T}(\widehat{\mathbf{f}})] \ \ \textbf{s.t.} \ \ \widehat{\mathbf{f}}[\mathbf{k}] = \widehat{\mathbf{b}}[\mathbf{k}], \mathbf{k} \in \mathbf{\Gamma}$$

$$\Big\downarrow \ \textit{Convex Relaxation}$$

$$\min_{\widehat{\mathbf{f}}} \quad \|\mathcal{T}(\widehat{\mathbf{f}})\|_* \ \ \textbf{s.t.} \ \ \widehat{\mathbf{f}}[\mathbf{k}] = \widehat{\mathbf{b}}[\mathbf{k}], \mathbf{k} \in \mathbf{\Gamma}$$

Nuclear norm – sum of singular values

Computational challenges

- **Standard algorithms are slow:**

  Apply ADMM = Singular value thresholding (SVT)

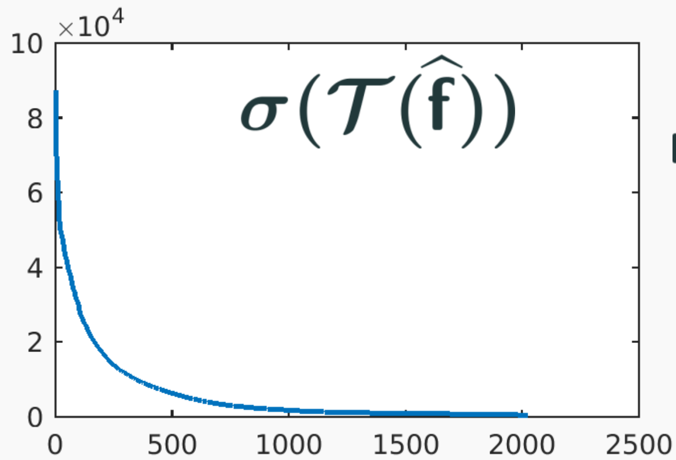
  Each iteration requires a large SVD:

  $$\mathbf{dim}(\mathcal{T}(\widehat{\mathbf{f}})) \approx \text{(\#pixels) x (filter size)}$$

  e.g. 10^6 x 2000

- **Real data can be "high-rank":**

e.g.
Singular values of
Real MR image



$$\mathbf{rank}(\mathcal{T}(\widehat{\mathbf{f}})) \approx 1000$$

Proposed Approach: Adapt IRLS algorithm

- IRLS: Iterative Reweighted Least Squares
- Proposed for low-rank matrix completion in [Fornasier, Rauhut, & Ward, 2011], [Mohan & Fazel, 2012]
- Adapt to structured matrix case:

$$\begin{cases} \mathbf{W} \leftarrow [\mathcal{T}(\widehat{\mathbf{f}})^*\mathcal{T}(\widehat{\mathbf{f}}) + \epsilon\mathbf{I}]^{-\frac{1}{2}} & \text{(weight matrix update)} \\ \widehat{\mathbf{f}} \leftarrow \arg\min_{\widehat{\mathbf{f}}} \|\mathcal{T}(\widehat{\mathbf{f}})\mathbf{W}^{\frac{1}{2}}\|_F^2 \ \text{ s.t. } \ \mathbf{P}\widehat{\mathbf{f}} = \mathbf{b} & \text{(LS problem)} \end{cases}$$

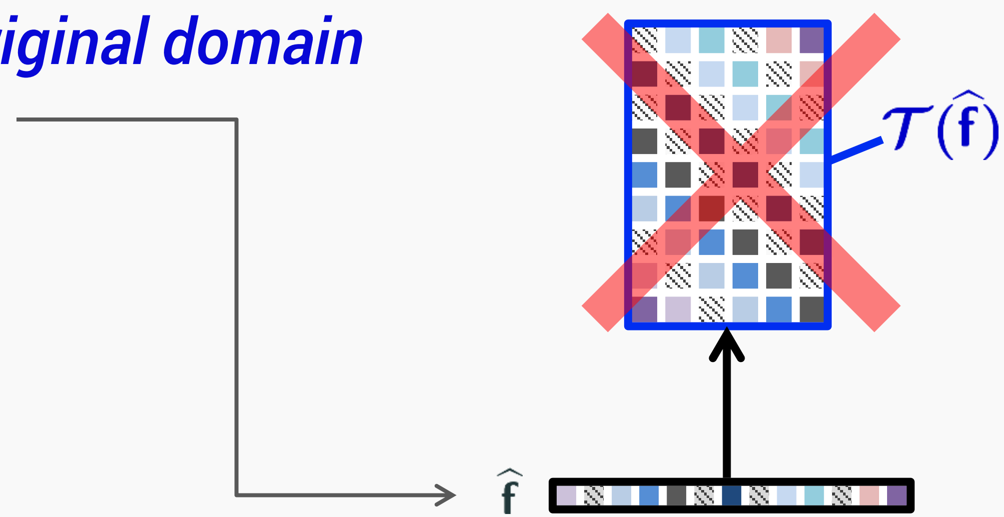
- Without modification, this approach is slow!

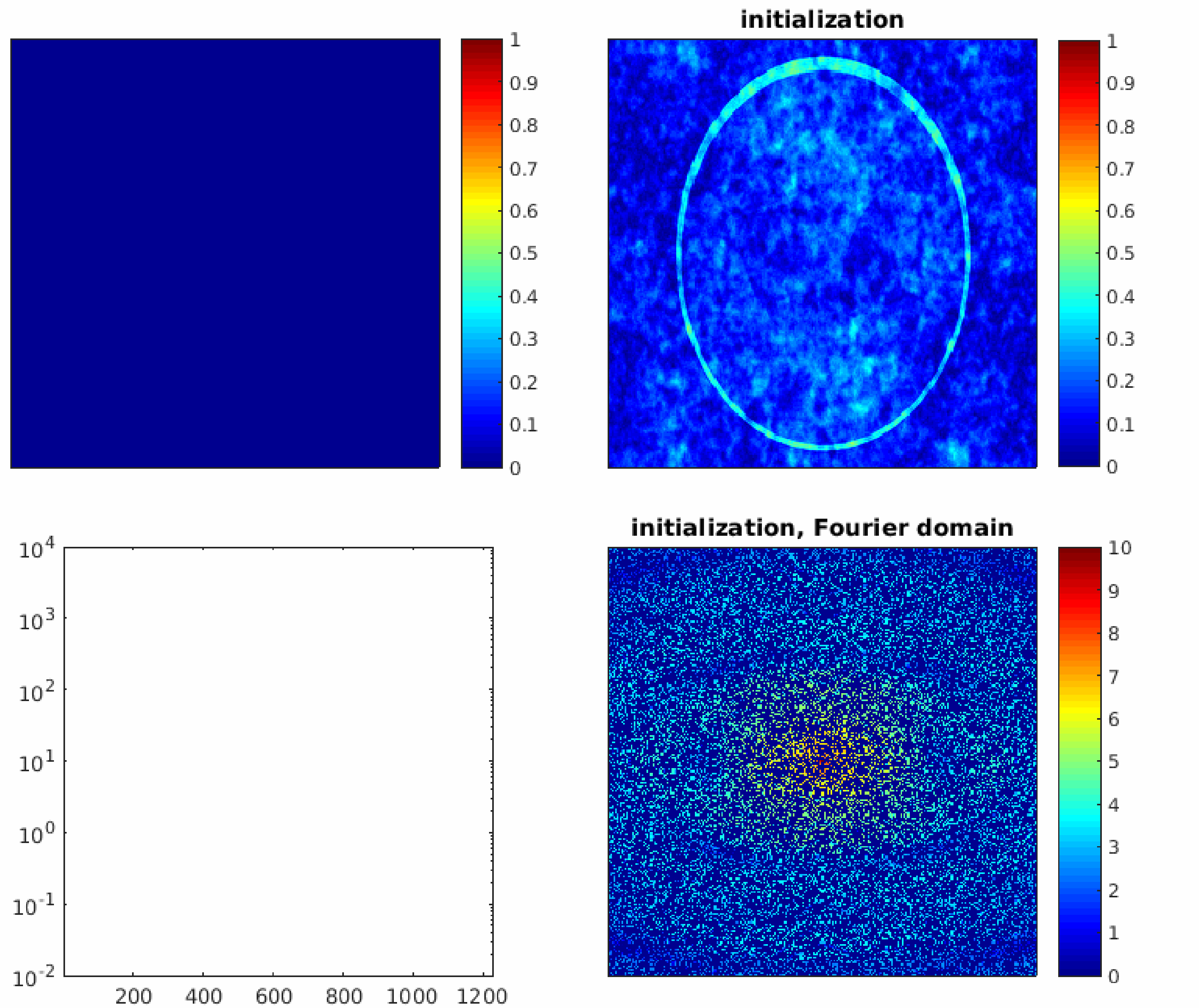
GIRAF algorithm [O. & Jacob, ISBI 2016]

- **GIRAF = Generic Iterative Reweighted Annihilating Filter**
- Exploit convolution structure to simplify IRLS algorithm:

$$\begin{cases} \mu \leftarrow \sum \lambda_i^{-\frac{1}{2}} \mu_i \quad \text{(annihilating filter update)} \\ \widehat{\mathbf{f}} \leftarrow \arg\min_{\widehat{\mathbf{f}}} \|\widehat{\mathbf{f}} * \widehat{\mu}\|_2^2 \quad \text{s.t.} \quad \mathbf{P}\widehat{\mathbf{f}} = \mathbf{b} \quad \text{(LS problem)} \end{cases}$$

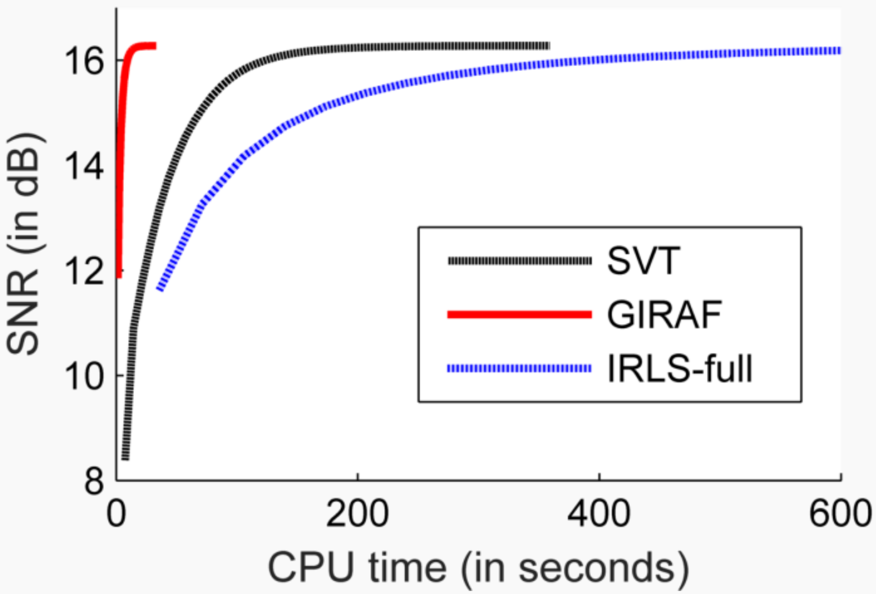
- Condenses weight matrix to *single* annihilating filter
- Solves problem in *original domain*





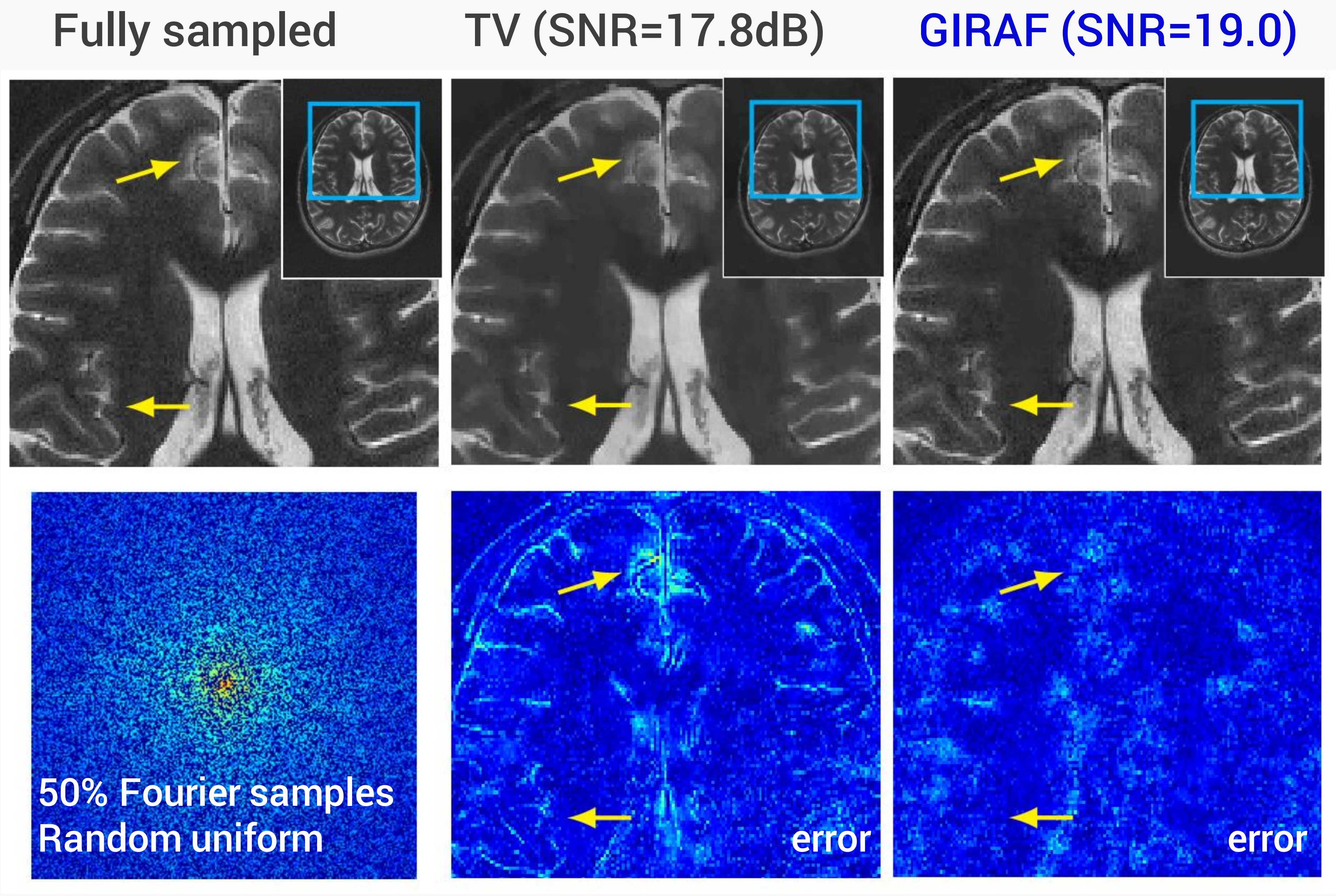
initialization
initialization, Fourier domain

Convergence speed of GIRAF



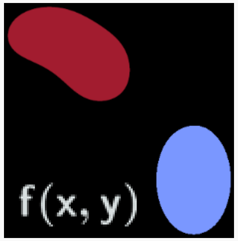
	15×15 filter		31×31 filter	
Algorithm	# iter	total	# iter	total
SVT	7	110s	11	790 s
GIRAF	6	20s	7	44 s

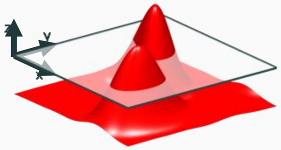
Table: iterations/CPU time to reach convergence tolerance of NMSE < 10^{-4}.


Fully sampled
TV (SNR=17.8dB)
GIRAF (SNR=19.0)
50% Fourier samples
Random uniform
error
error

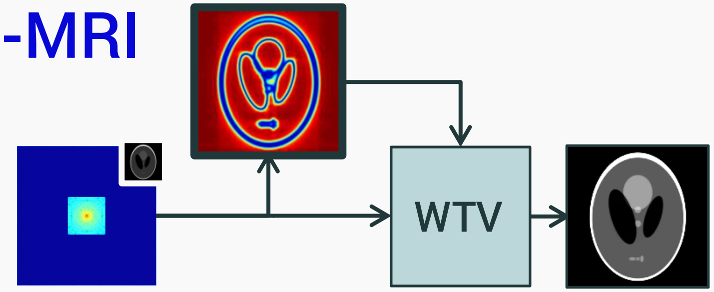
Summary

- New framework for off-the-grid image recovery
  - Extends FRI annihilating filter framework to piecewise polynomial images
  - Sampling guarantees





- Two stage recovery scheme for SR-MRI
  - Robust edge mask estimation
  - Fast weighted TV algorithm

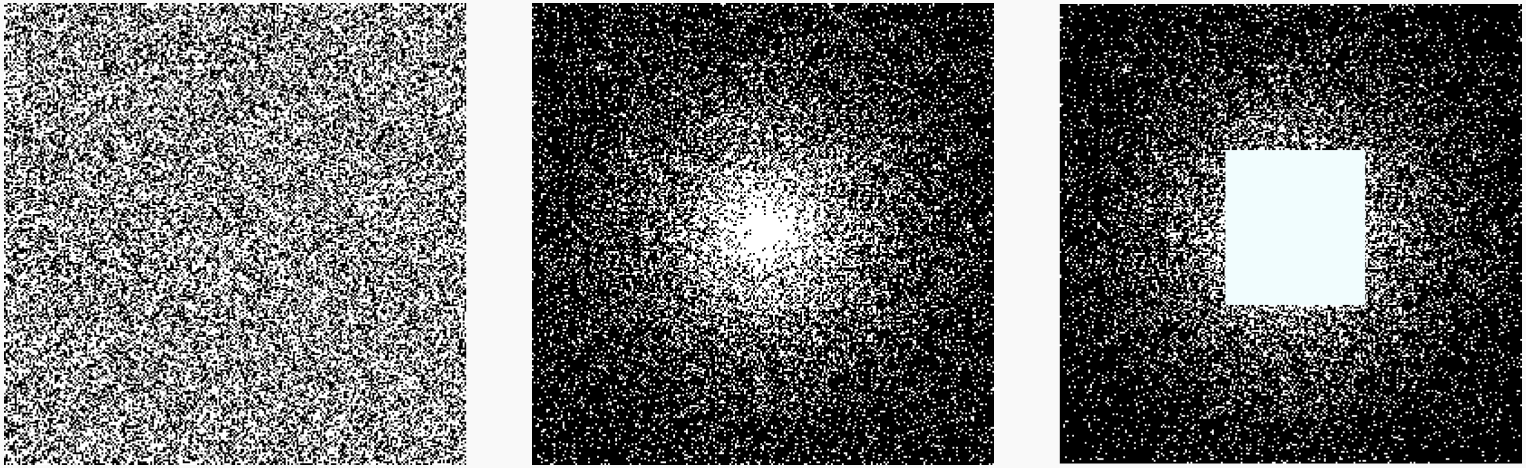


- One stage recovery scheme for CS-MRI
  - Structured low-rank matrix completion
  - Fast GIRAF algorithm

$$\min_{\hat{\mathbf{f}}} \|\mathcal{T}(\hat{\mathbf{f}})\|_*$$

Future Directions

- **Focus:** One stage recovery scheme for CS-MRI
  - Structured low-rank matrix completion $$\min_{\widehat{\mathbf{f}}} \|\mathcal{T}(\widehat{\mathbf{f}})\|_*$$
- Recovery guarantees for random sampling?
- What is the optimal random sampling scheme?



Thank You!

References

- Krahmer, F. & Ward, R. (2014). Stable and robust sampling strategies for compressive imaging. *Image Processing, IEEE Transactions on*, 23(2), 612-
- Pan, H., Blu, T., & Dragotti, P. L. (2014). Sampling curves with finite rate of innovation. *Signal Processing, IEEE Transactions on*, *62*(2), 458-471.
- Guerquin-Kern, M., Lejeune, L., Pruessmann, K. P., & Unser, M. (2012). Realistic analytical phantoms for parallel Magnetic Resonance Imaging.*Medical Imaging, IEEE Transactions on*, *31*(3), 626-636
- Vetterli, M., Marziliano, P., & Blu, T. (2002). Sampling signals with finite rate of innovation. *Signal Processing, IEEE Transactions on*, *50*(6), 1417-1428.
- Sidiropoulos, N. D. (2001). Generalizing Caratheodory's uniqueness of harmonic parameterization to N dimensions. *Information Theory, IEEE Transactions on*,*47*(4), 1687-1690.
- Ongie, G., & Jacob, M. (2015). Super-resolution MRI Using Finite Rate of Innovation Curves. *Proceedings of ISBI 2015, New York, NY*.
- Ongie, G. & Jacob, M. (2015). Recovery of Piecewise Smooth Images from Few Fourier Samples. *Proceedings of SampTA 2015, Washington D.C.*
- Ongie, G. & Jacob, M. (2015). Off-the-grid Recovery of Piecewise Constant Images from Few Fourier Samples. *Arxiv.org preprint.*
- Fornasier, M., Rauhut, H., & Ward, R. (2011). Low-rank matrix recovery via iteratively reweighted least squares minimization. *SIAM Journal on Optimization,* 21(4), 1614-1640.
- Mohan, K, and Maryam F. (2012). Iterative reweighted algorithms for matrix rank minimization." *The Journal of Machine Learning Research* 13.1 3441-3473.

Acknowledgements

- Supported by grants: NSF CCF-0844812, NSF CCF-1116067, NIH 1R21HL109710-01A1, ACS RSG-11-267-01-CCE, and ONR-N000141310202.