Recovery of Piecewise Smooth Images from Few Fourier Samples

Greg Ongie*, Mathews Jacob

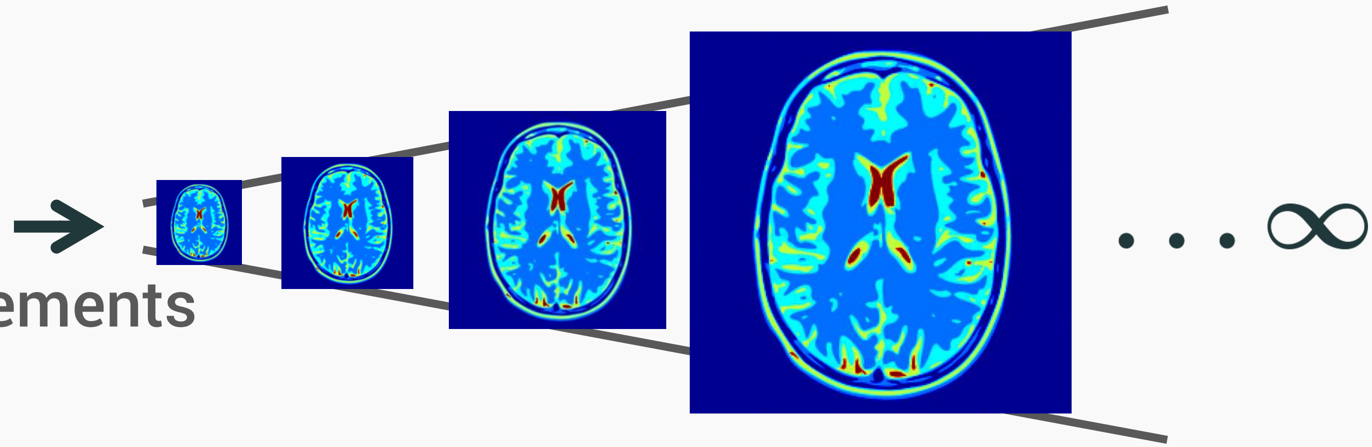
Computational Biomedical Imaging Group (CBIG)
University of Iowa

SampTA 2015
Washington, D.C.

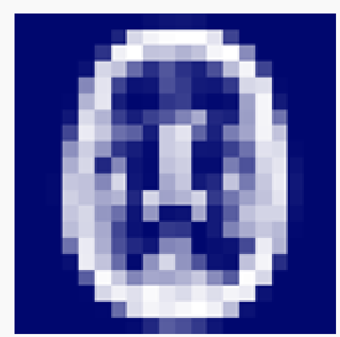
1. Introduction
2. Off-the-Grid Image Recovery: New Framework
3. Sampling Guarantees
4. Algorithms
5. Discussion & Conclusion

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

Few measurements

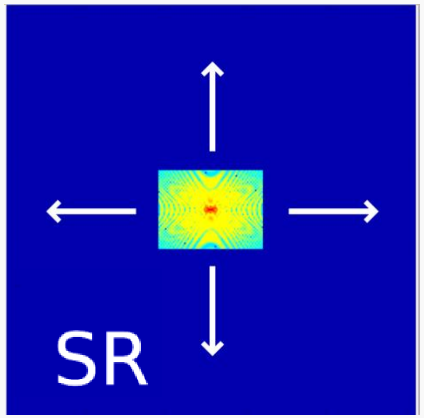


- Off-the-grid = Continuous domain representation
- Avoid discretization errors
- Continuous domain sparsity $\neq$ Discrete domain sparsity



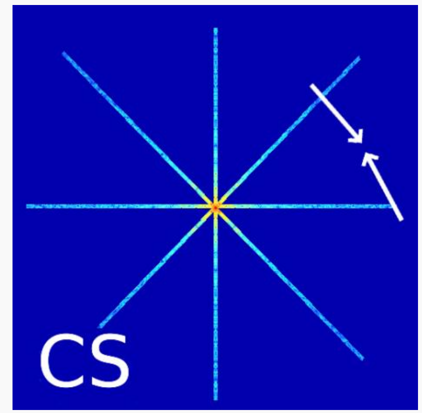
Wide range of applications

- Super-resolution MRI: Fourier undersampling approach



Fourier Extrapolation

vs.

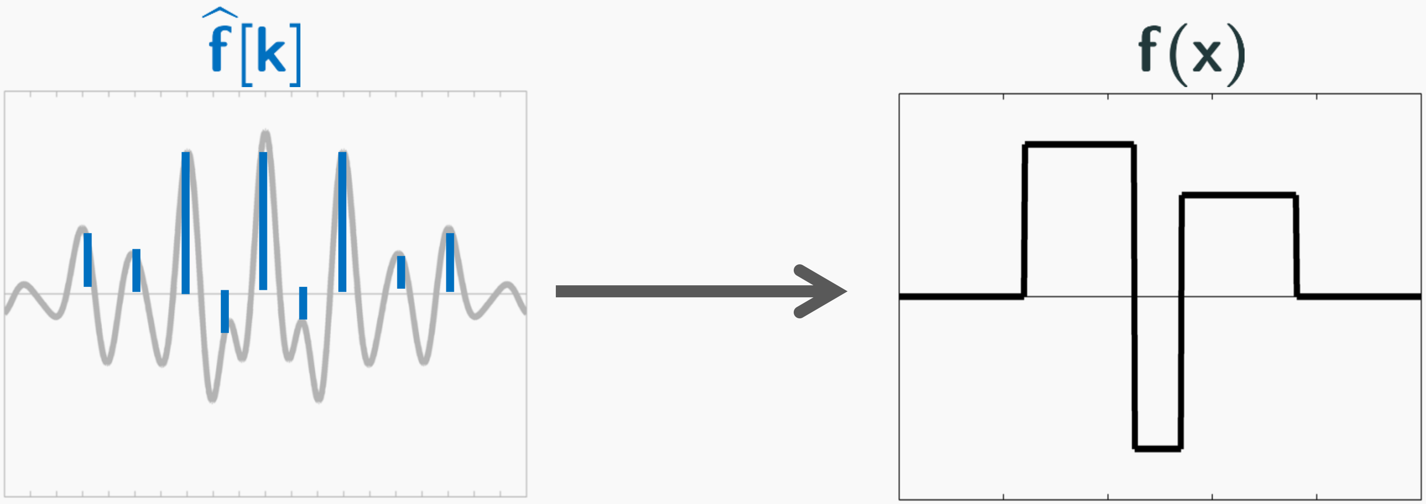


Fourier Interpolation

  – MRI Modalites: Multi-slice, Dynamic, MRSI

- Compressed Sensing MRI
- Outside MRI: Deconvolution Microscopy, Denoising, etc.

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

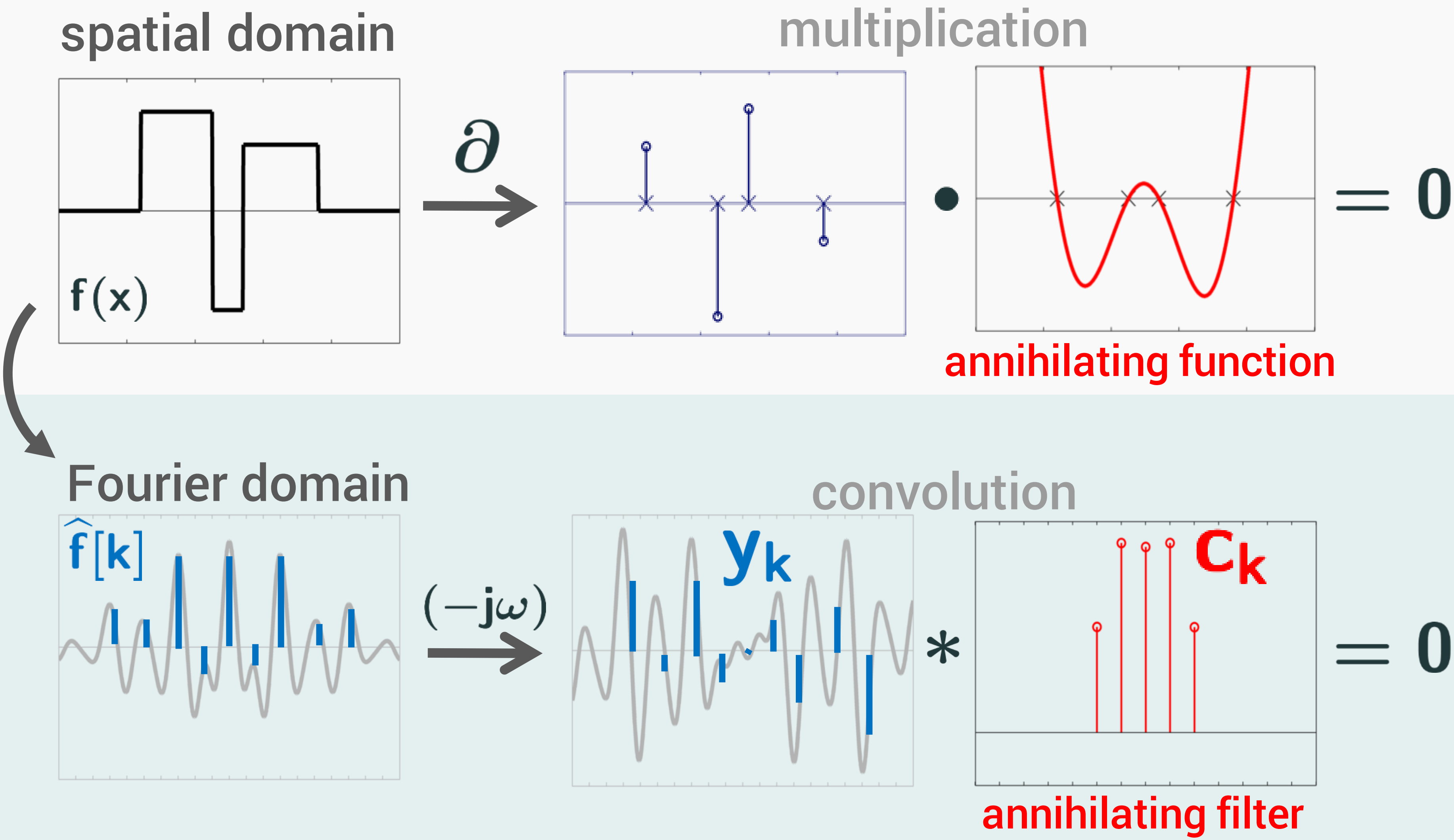


Uniform
Fourier samples

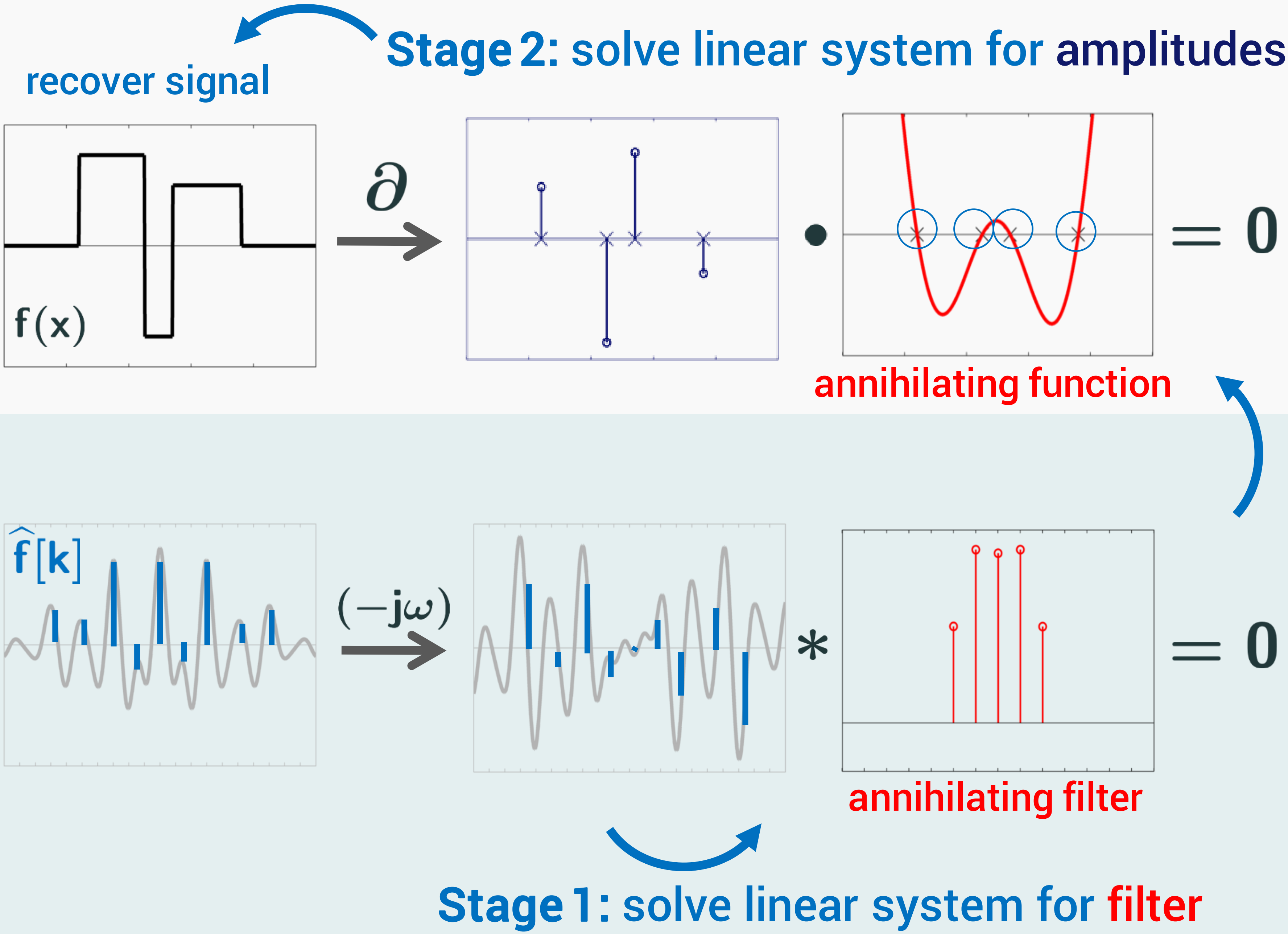
Off-the-grid
PWC signal

- Recent extension to 2-D images:

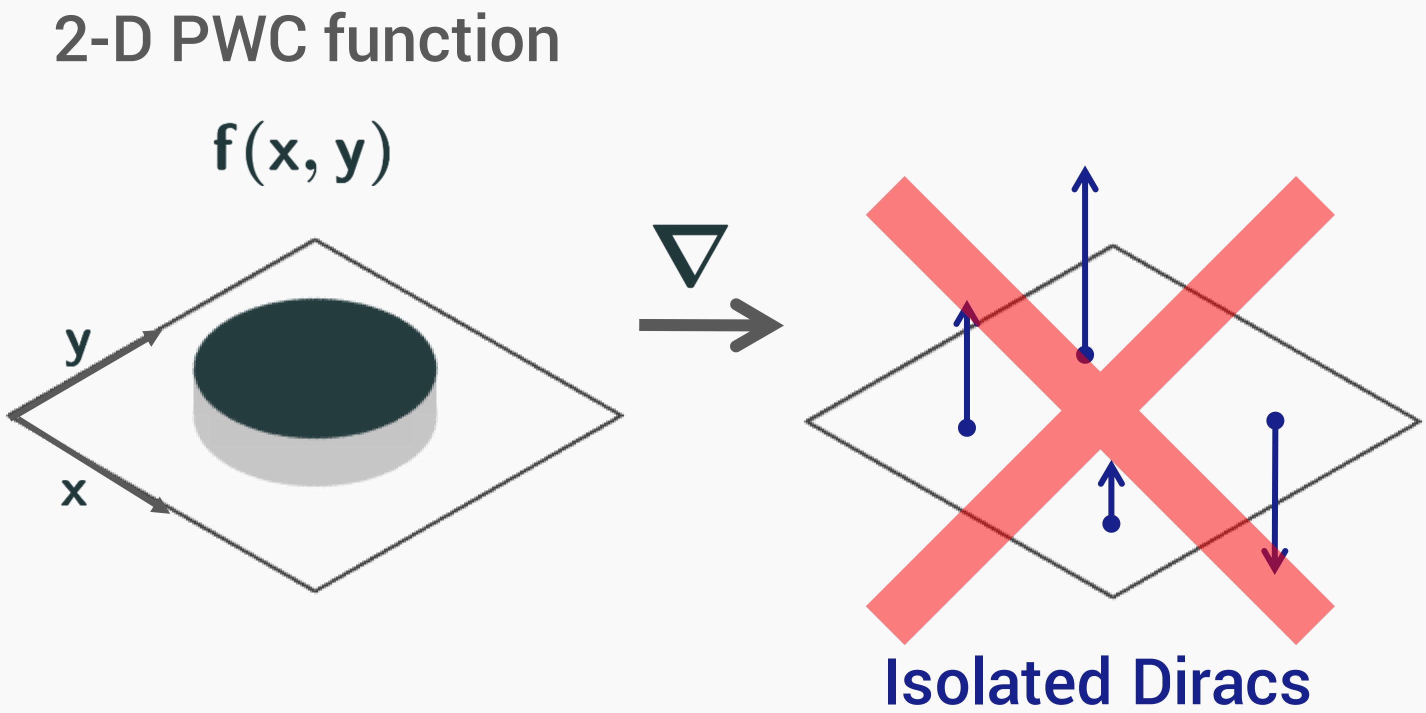
Pan, Blu, & Dragotti (2014), "Sampling Curves with FRI".



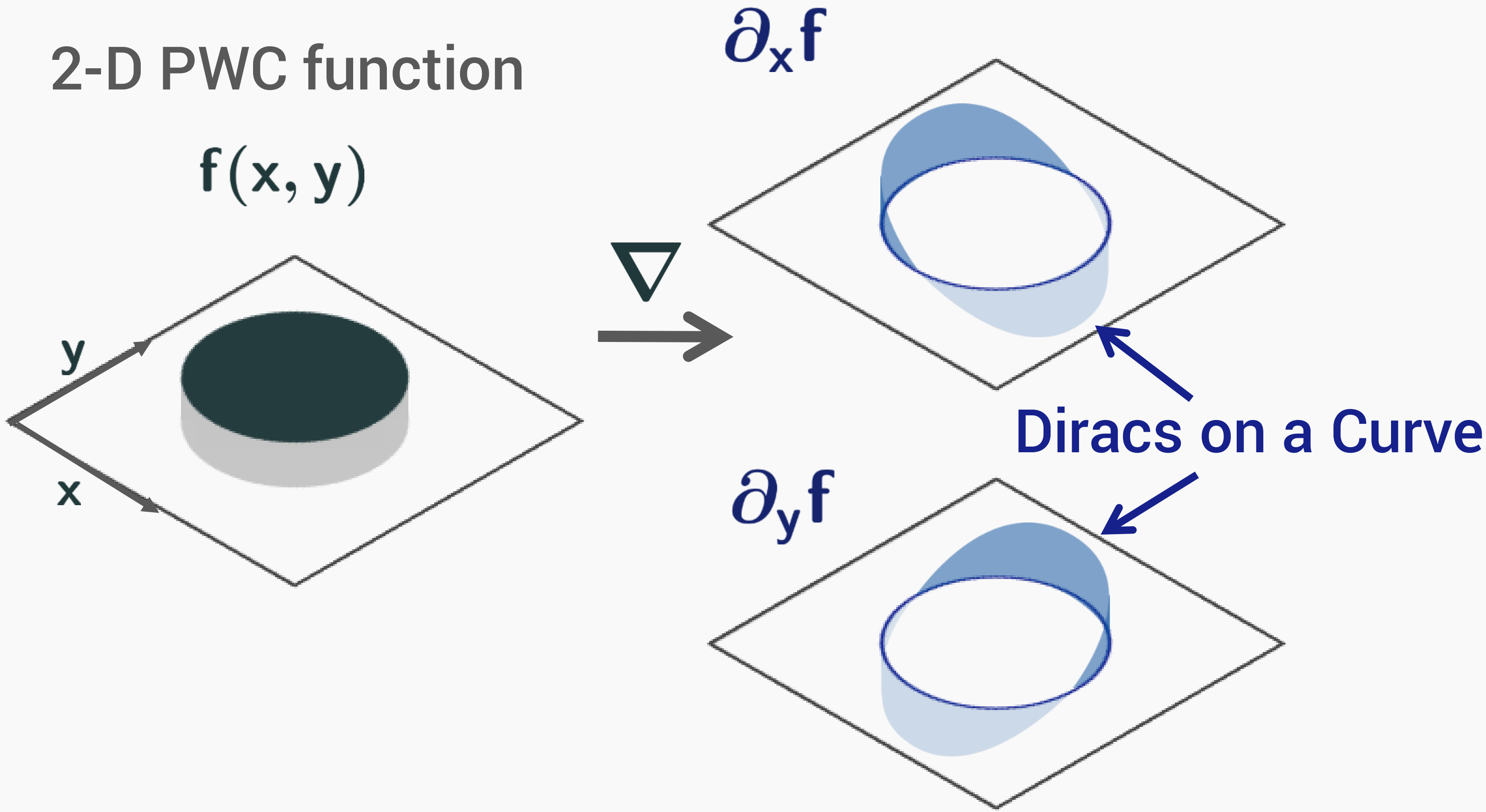
Annihilation Relation: $\sum_k y_{\ell-k} c_k = 0$


Stage 2: solve linear system for amplitudes
recover signal
f(x)
∂
= 0
annihilating function
$\hat{f}[k]$
$(-j\omega)$
*
= 0
annihilating filter
Stage 1: solve linear system for filter

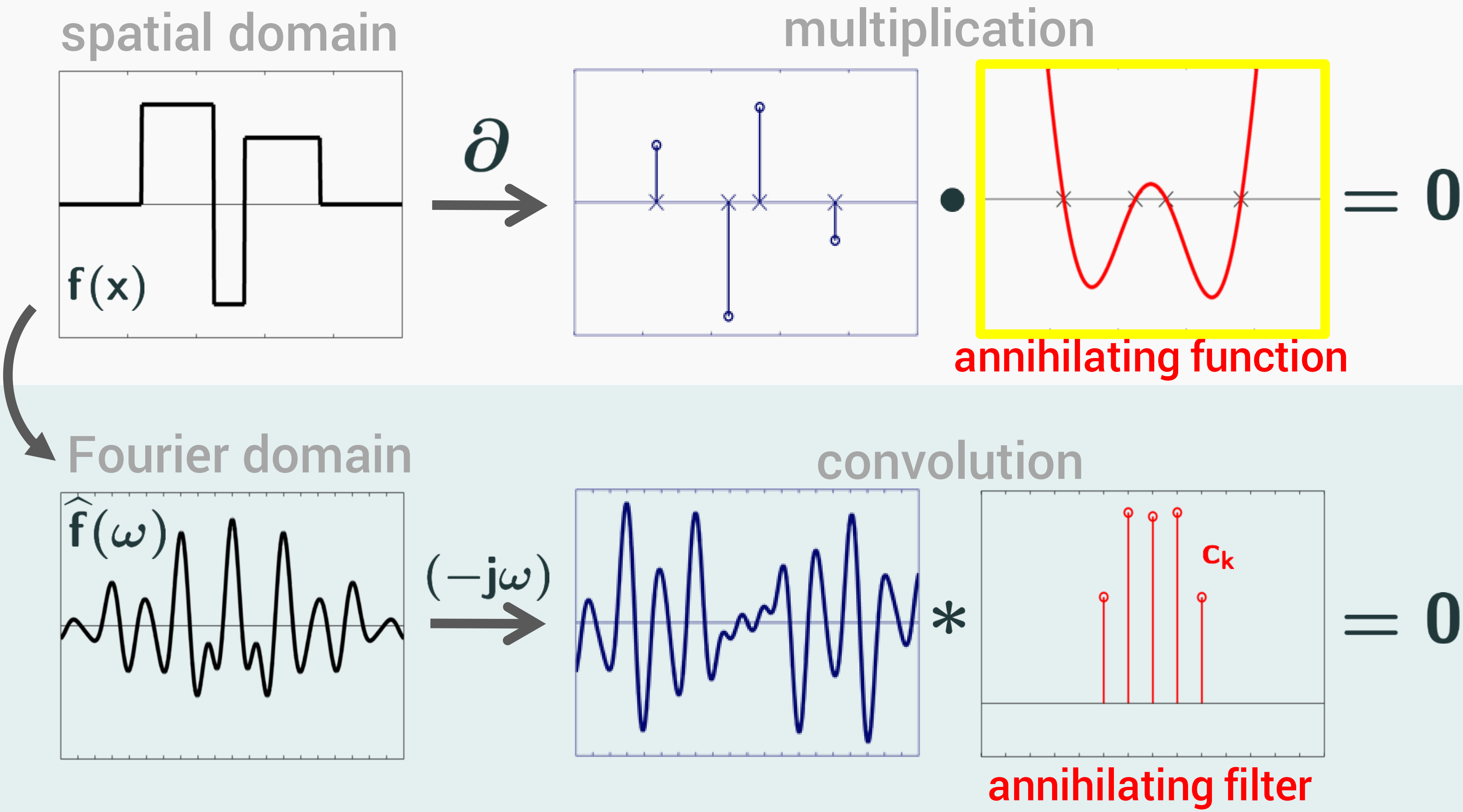
Challenges extending FRI to higher dimensions: Singularities not isolated



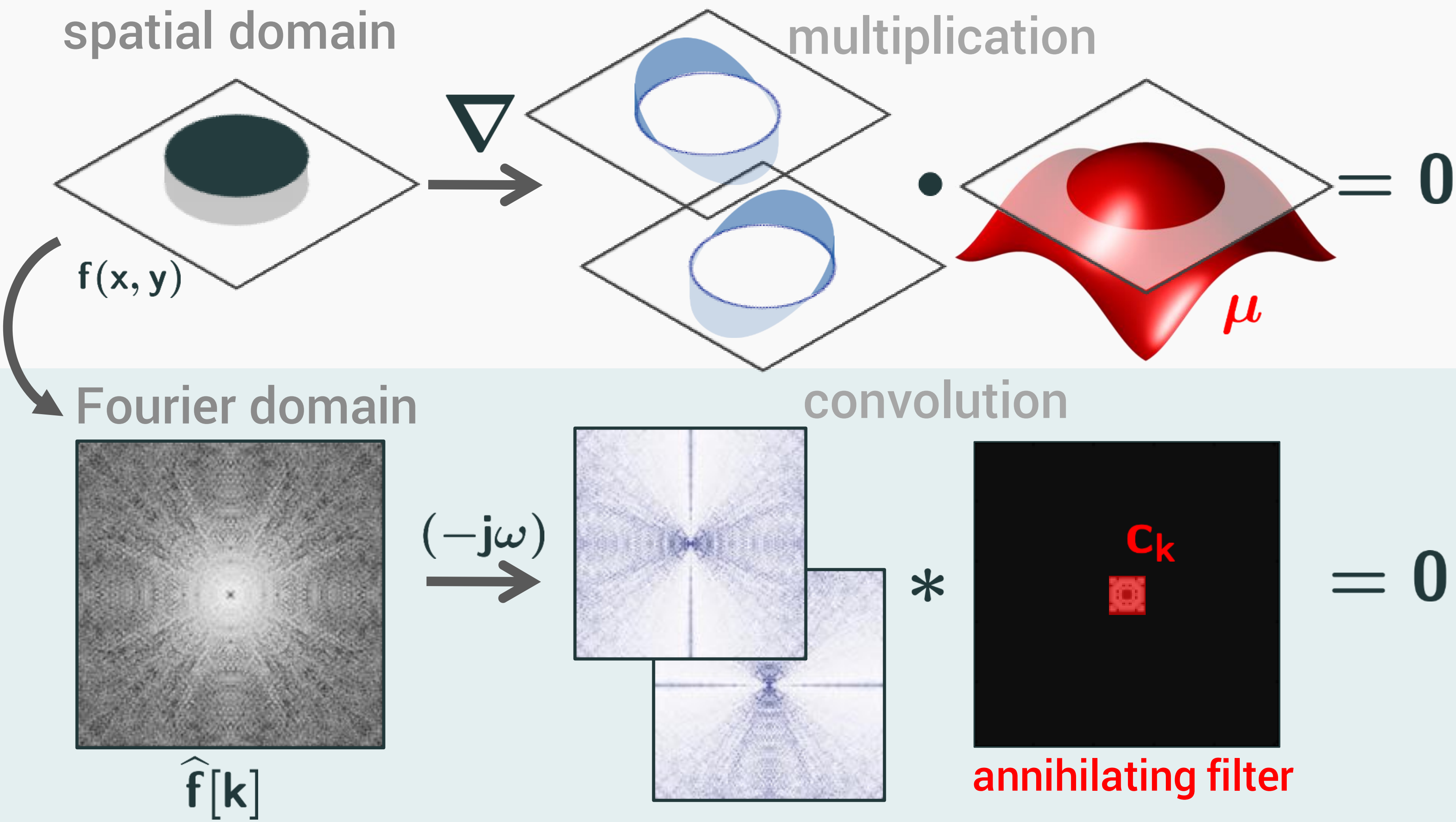
Challenges extending FRI to higher dimensions: Singularities not isolated



Recall 1-D Case...

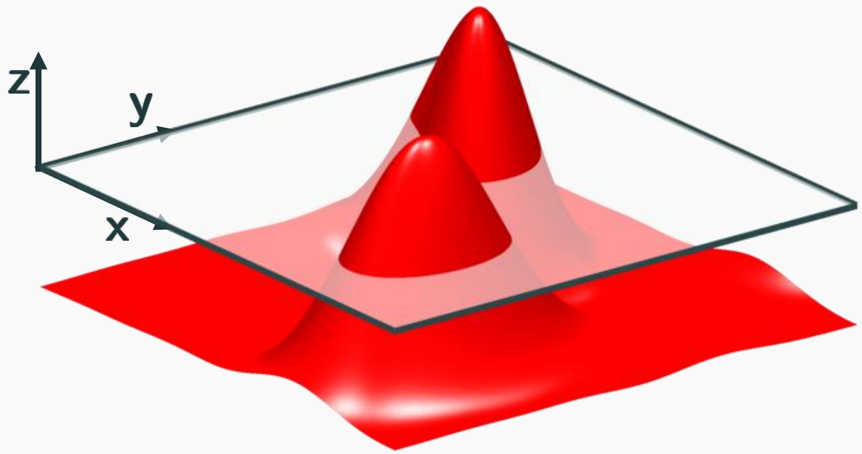


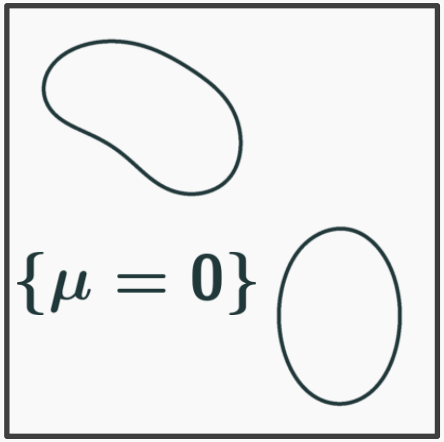
2-D PWC functions satisfy an annihilation relation



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]



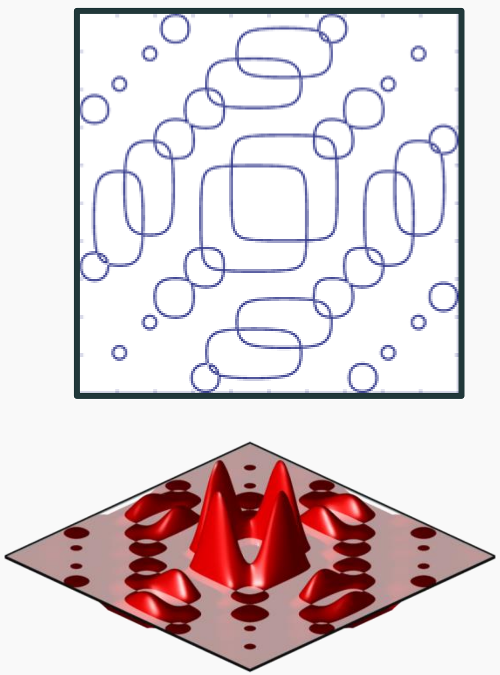


$$\mu(x, y) = \sum_{(k,l)\in\Lambda} c_{k,l}\, e^{j2\pi(kx+ly)}$$

"FRI Curve"

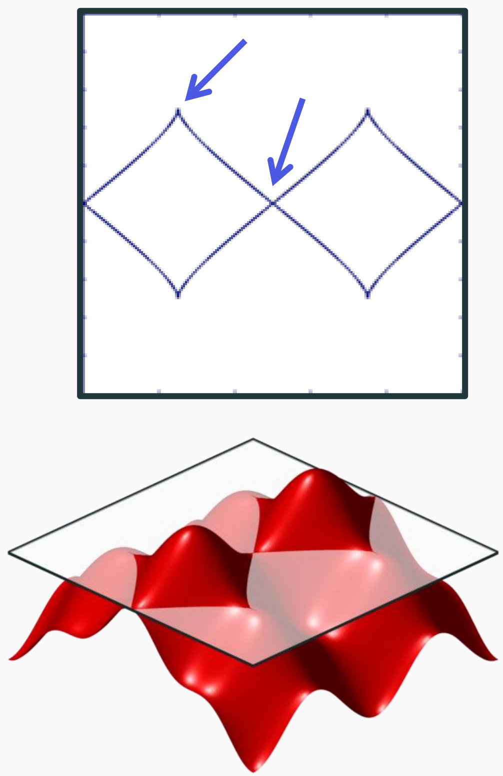
FRI curves can represent complicated edge geometries with few coefficients

Multiple curves & intersections



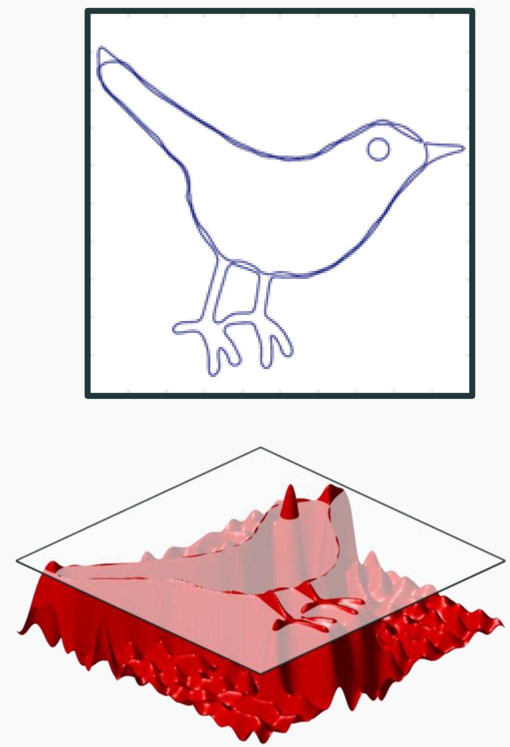
13x13 coefficients

Non-smooth points



7x9 coefficients

Approximate arbitrary curves



25x25 coefficients

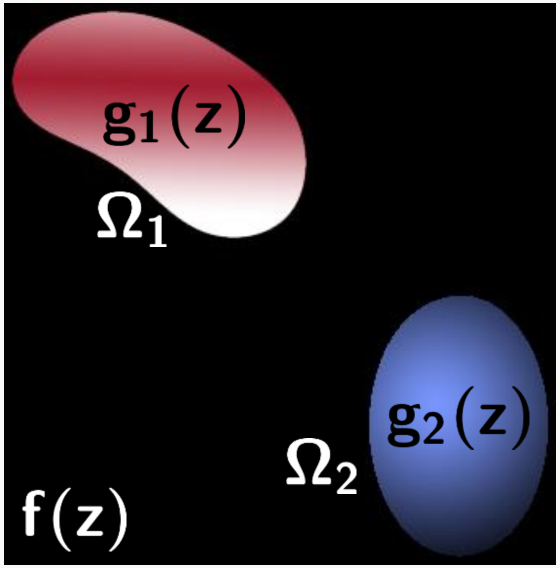
1. Introduction
2. **Off-the-Grid Image Recovery: New Framework**
3. Sampling Guarantees
4. Algorithms
5. Discussion & Conclusion

We give an improved theoretical framework for higher dimensional FRI recovery

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

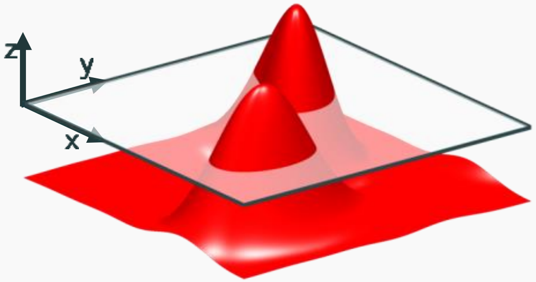
$$f(z) = \sum_{i=1}^{N} g_i(z) \cdot 1_{\Omega_i}(z)$$

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



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

  Infinite DoF

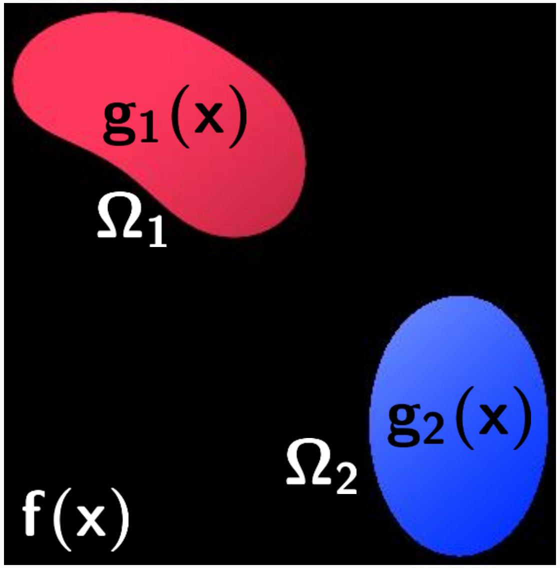


We give an improved theoretical framework for higher dimensional FRI recovery

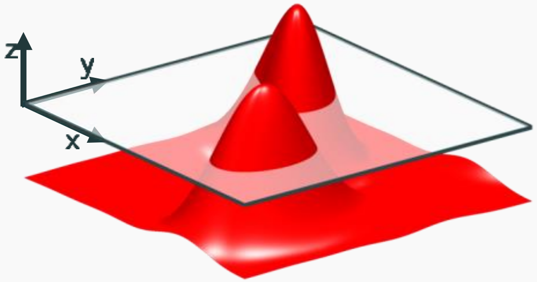
- Proposed model:
  piecewise smooth signals

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

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



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

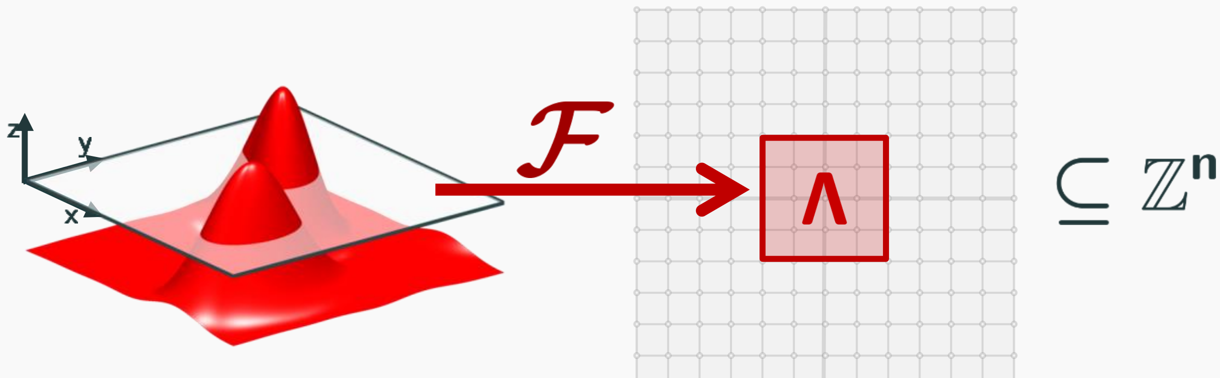


Annhilation relation for PWC signals

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



Annhilation relation for PWC signals

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

for μ bandlimited to Λ then

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

any 1st order partial derivative

Proof idea: Show $\mu \cdot \partial \mathrm{f} = 0$ "distributionally"

Use convolution theorem

Proof:

Write $\mathbf{f} = \sum_i \mathbf{a_i} \cdot \mathbf{1}_{\Omega_i} \implies \partial\mathbf{f} = \sum_i \mathbf{a_i} \cdot \boxed{\partial\mathbf{1}_{\Omega_i}}$

Proof:

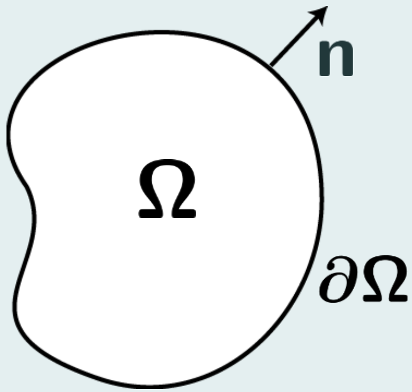
Write $\mathbf{f} = \sum_i a_i \cdot \mathbf{1}_{\Omega_i} \implies \partial \mathbf{f} = \sum_i a_i \cdot \partial \mathbf{1}_{\Omega_i}$

Distributional derivative of indicator function:

smooth test function

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

divergence theorem



Proof:

Write $\mathbf{f} = \sum_i a_i \cdot \mathbf{1}_{\Omega_i} \implies \partial \mathbf{f} = \sum_i a_i \cdot \partial \mathbf{1}_{\Omega_i}$

Distributional derivative of indicator function:

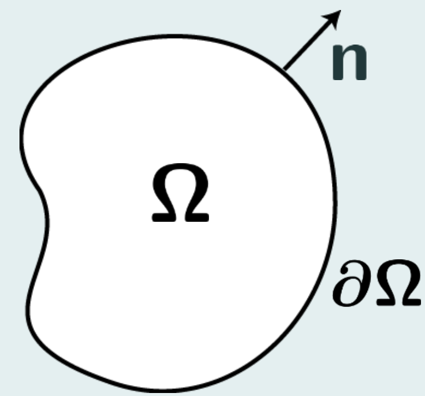
smooth test function

$$\langle \partial_j \mathbf{1}_\Omega, \varphi \rangle = -\langle \mathbf{1}_\Omega, \partial_j \varphi \rangle$$

$$= -\int_\Omega \partial_j \varphi \, dx$$

divergence theorem

$$= -\oint_{\partial\Omega} \varphi \, n_j \, d\sigma$$



$$\implies \mu \cdot \partial_j \mathbf{1}_\Omega = 0$$

$$\langle \mu \cdot \partial_j \mathbf{1}_\Omega, \varphi \rangle = \langle \partial_j \mathbf{1}_\Omega, \mu \, \varphi \rangle = -\oint_{\partial\Omega} \mu \, \varphi \, n_j \, d\sigma = 0$$

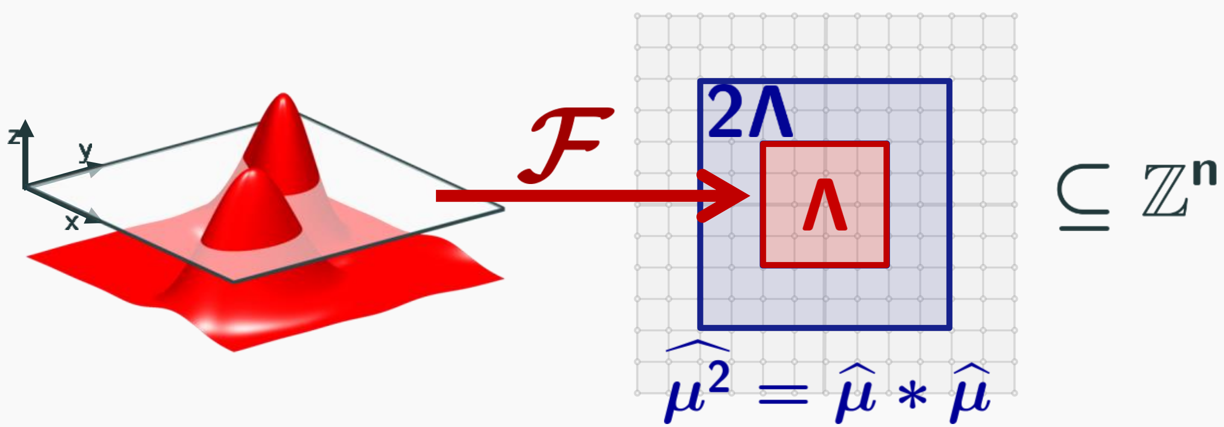
Since $\mu = 0$ on $\partial\Omega$

Annhilation relation for PW linear signals

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

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

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

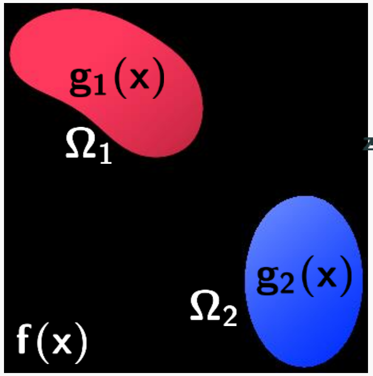
annihilated by μ^2

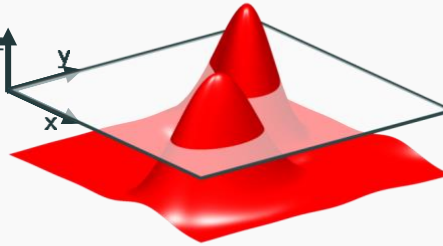
Can extend annihilation relation to a wide class of **piecewise smooth** images.

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

$$\text{s.t.} \quad \mathbf{Dg_i = 0} \text{ in } \mathbf{\Omega_i}$$

Any constant coeff. differential operator

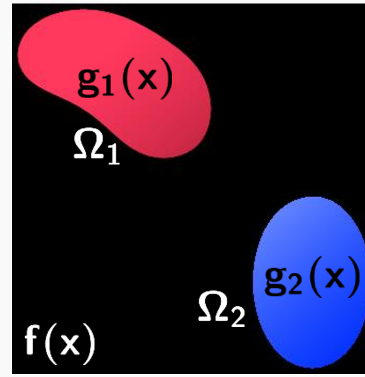


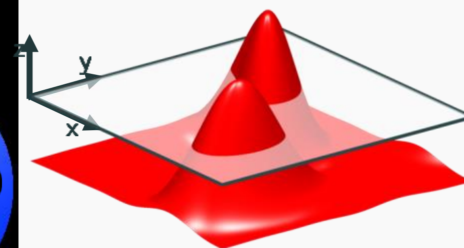


Can extend annihilation relation to a wide class of **piecewise smooth** images.

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

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





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

Annhilation relation for PW smooth images

Prop: If f is PW smooth, such that

1. the nulling operator D is nth order, and
2. the edge set $\mathbf{E} \subseteq \{\mu = 0\}$ with μ bandlimited to Λ then

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

1. Introduction
2. Off-the-Grid Image Recovery: New Framework
3. **Sampling Theorems**
4. Algorithms
5. Discussion & Conclusion

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 knots can be uniquely recovered from 2K+1 uniform Fourier samples.

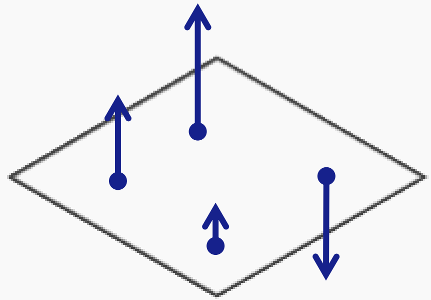
Proof (a la Prony):

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

"Caratheodory Parametrization"

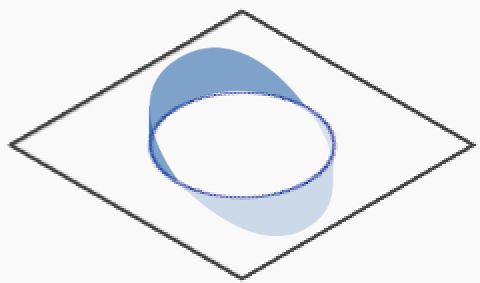
Challenges proving uniqueness, cont.

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



$$\xrightarrow{\mathcal{F}} \quad \mathbf{y}[\mathbf{k}] = \sum_{\mathbf{i}} \mathbf{a_i} e^{j2\pi \mathbf{k}\cdot \mathbf{x_i}}$$

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



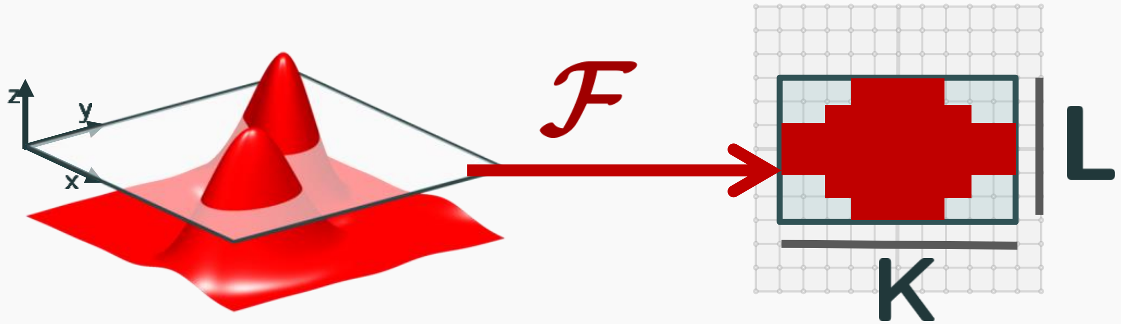
$$\xrightarrow{\mathcal{F}} \quad \mathbf{y}[\mathbf{k}] = \oint_{\partial\Omega} e^{j2\pi \mathbf{k}\cdot \mathbf{x}} \mathbf{n}\, ds$$

Requires new techniques:

- Spatial domain interpretation of annihilation relation
- Algebraic geometry of trig. polynomials

Minimal (Trig) 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 $\mathsf{C} = \{\mu = 0\}$

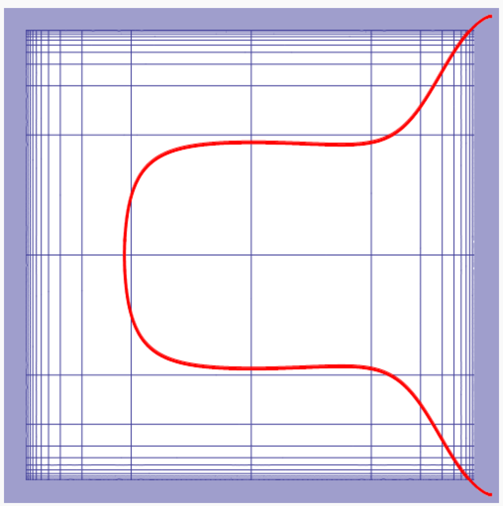
Then $\mathbf{deg}(\mu_0) \leq \mathbf{deg}(\mu)$ and $\mu = \gamma \cdot \mu_0$

Proof idea: Pass to Real Algebraic Plane Curves

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

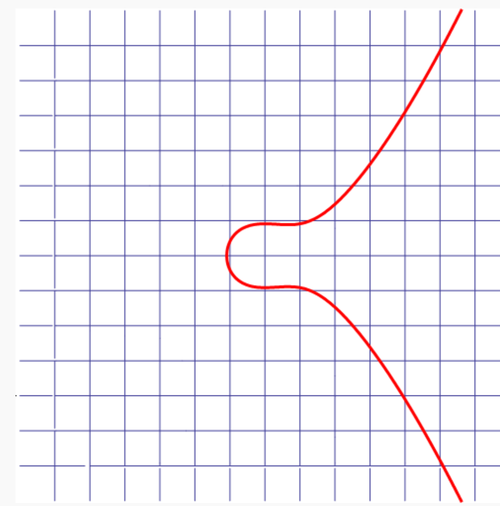
are in 1-to-1 correspondence with

real algebraic plane curves of degree (K,L)



$\mathbb{T}^2$

Conformal change of variables



$\mathbb{R}^2$

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

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

Uniqueness of edge set recovery

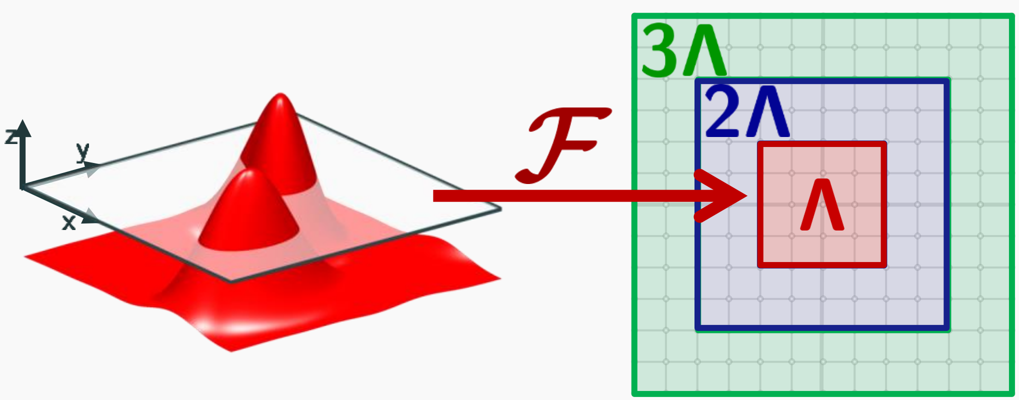
Theorem: If $\mathbf{f}$ is PWC* with edge set $\mathbf{E} = \{\mu = 0\}$

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 geometrical restrictions apply



$\subseteq \mathbb{Z}^2$

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

Proof Sketch:

- Let $\mathbf{d}[\mathbf{k}]$ be another solution:

$$\sum_{\mathbf{k}\in\boldsymbol{\Lambda}} \mathbf{d}[\mathbf{k}]\widehat{\boldsymbol{\nabla}\mathbf{f}}[\boldsymbol{\ell}-\mathbf{k}] = \mathbf{0} \quad \forall \boldsymbol{\ell} \in 2\boldsymbol{\Lambda}$$

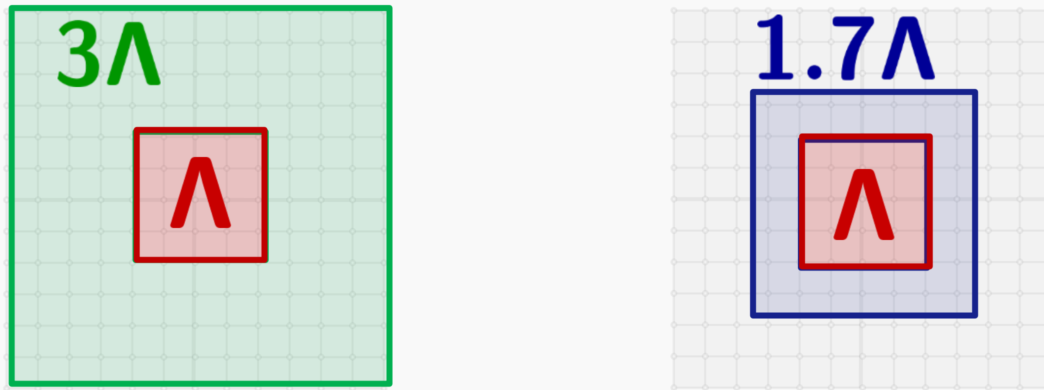
- Translate to spatial domain condition: $\mathbf{d}[\mathbf{k}] \leftrightarrow \boldsymbol{\eta}(\mathbf{x})$

$$\int_{\{\mu=0\}} \boldsymbol{\eta}\,(\boldsymbol{\varphi}\cdot\mathbf{n})\,\mathbf{ds} = \mathbf{0} \quad \forall \boldsymbol{\varphi} : \ \mathbf{supp}(\widehat{\boldsymbol{\varphi}}) \in 2\boldsymbol{\Lambda}$$

- Show this implies η must vanish on $\{\mu = 0\}$ and so $\eta = \mu$ since μ is minimal.

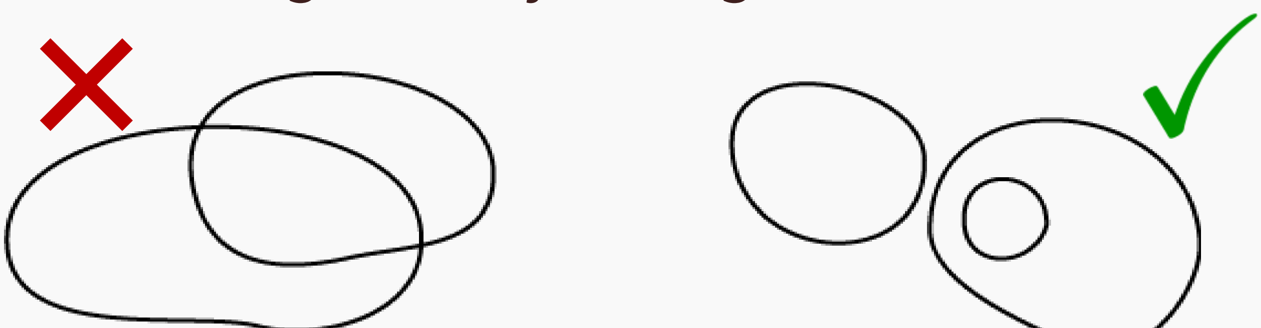
Current Limitations to Uniqueness Theorem

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

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

when the sampling set $\Gamma \supseteq 3\Lambda$

*Some geometrical 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 Λ then $\mathbf{g} = \mathbf{f}$ is the unique solution to

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

when the sampling set $\Gamma \supseteq 3\Lambda$

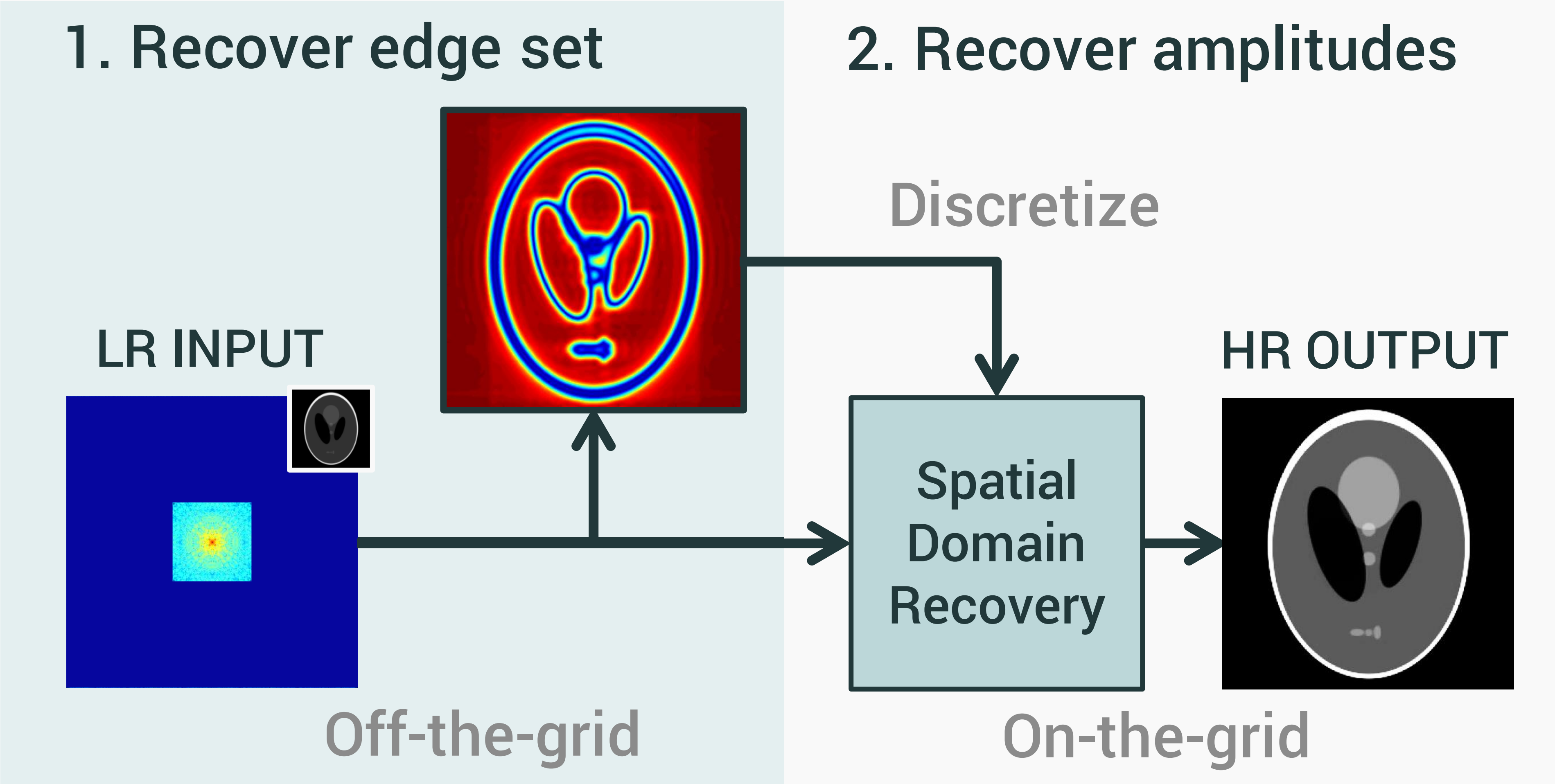
*Some geometrical restrictions apply

Equivalently,

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

1. Introduction
2. Off-the-Grid Image Recovery: New Framework
3. Sampling Theorems
4. **Algorithms**
5. Discussion & Conclusion

Previously: Two-stage Super-resolution MRI
Piecewise Constant Signal Model [O. & Jacob, 2015]

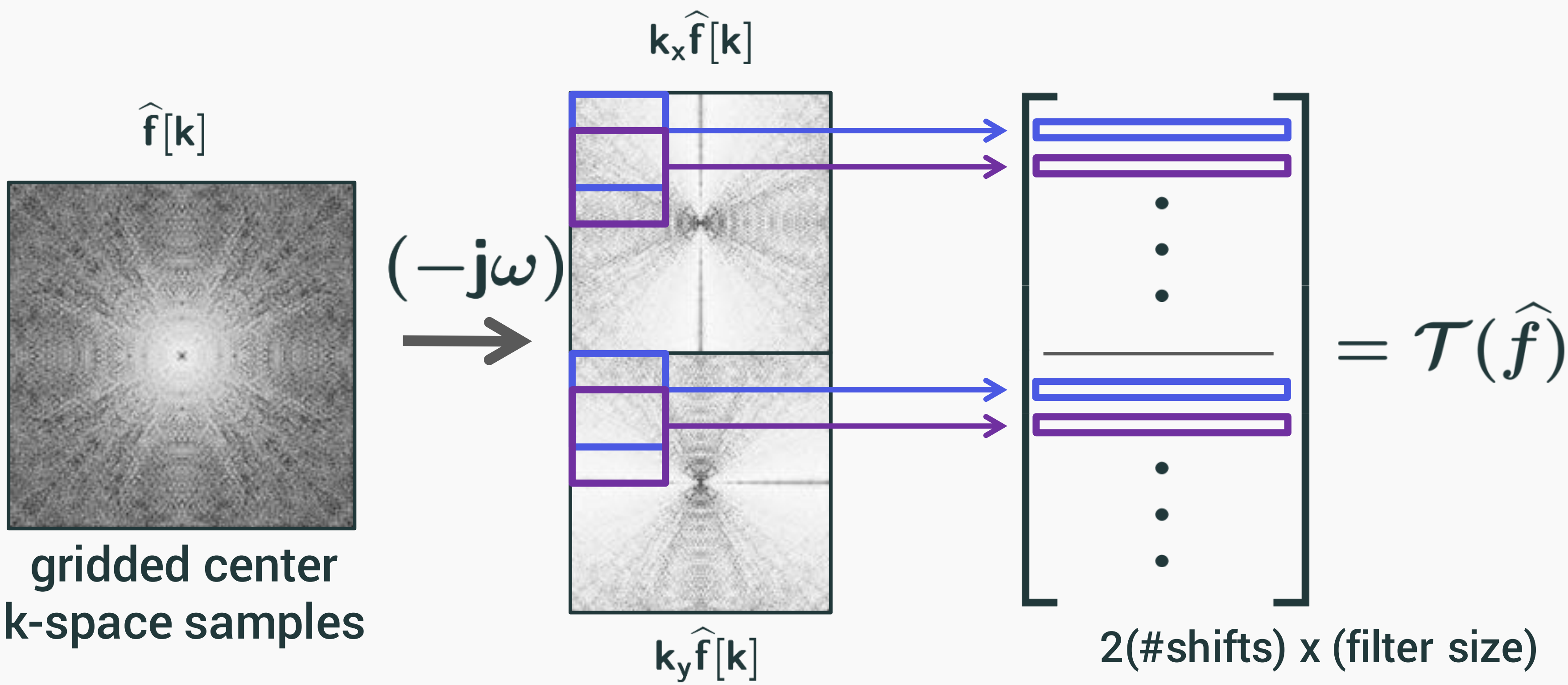


Matrix representation of annihilation

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

2-D convolution matrix
(block Toeplitz)

vector of filter coefficients

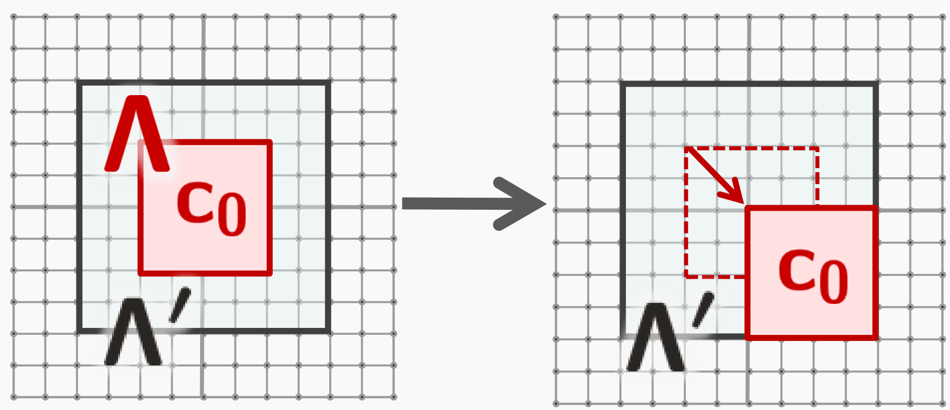


Basis of algorithms:
Annihilation matrix is low-rank

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

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

Fourier domain



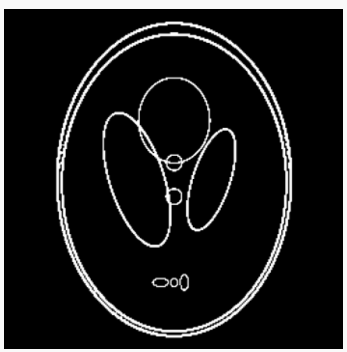
Spatial domain $\mu(x, y) \longrightarrow e^{j2\pi(kx+ly)}\mu(x, y)$

Basis of algorithms:
Annihilation matrix is low-rank

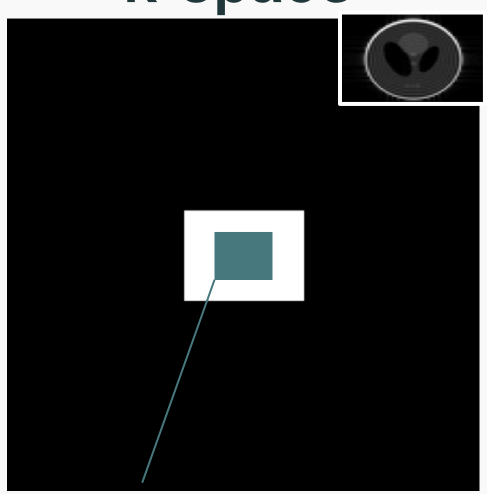
Prop: If the level-set function is bandlimited to Λ and the assumed filter support $\Lambda' \supset \Lambda$ then

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

Example:
Shepp-Logan



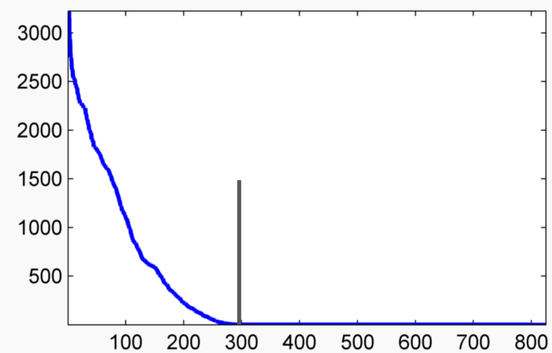
k-space



Assumed filter: 33x25

Samples: 65x49

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



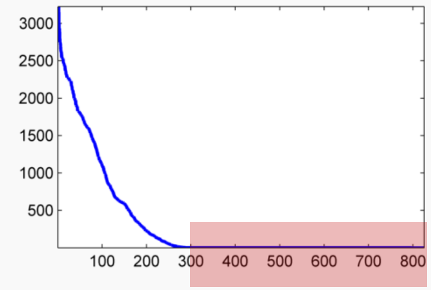
Rank $\approx$ 300

Stage 1: Robust annihilting filter estimation

1. Compute SVD

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

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

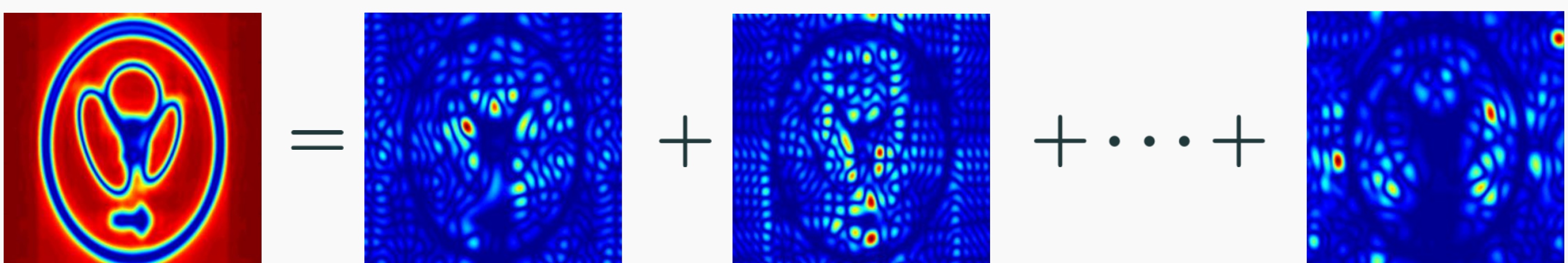


2. Identify null space

$$\mathbf{V} = [\mathbf{V_S}\ \mathbf{V_N}], \qquad \mathbf{V_N} = [\mathbf{c}_1, \ldots, \mathbf{c}_n]$$

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

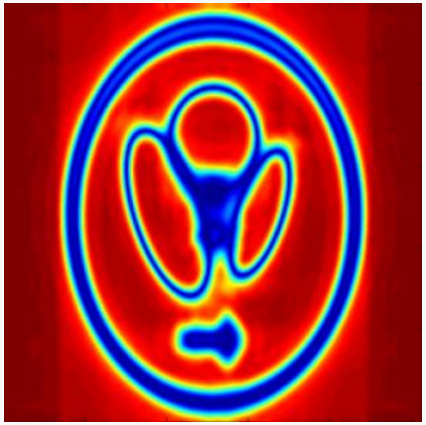
Stage 2: Weighted TV Recovery

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

discretize

relax

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



Edge weights

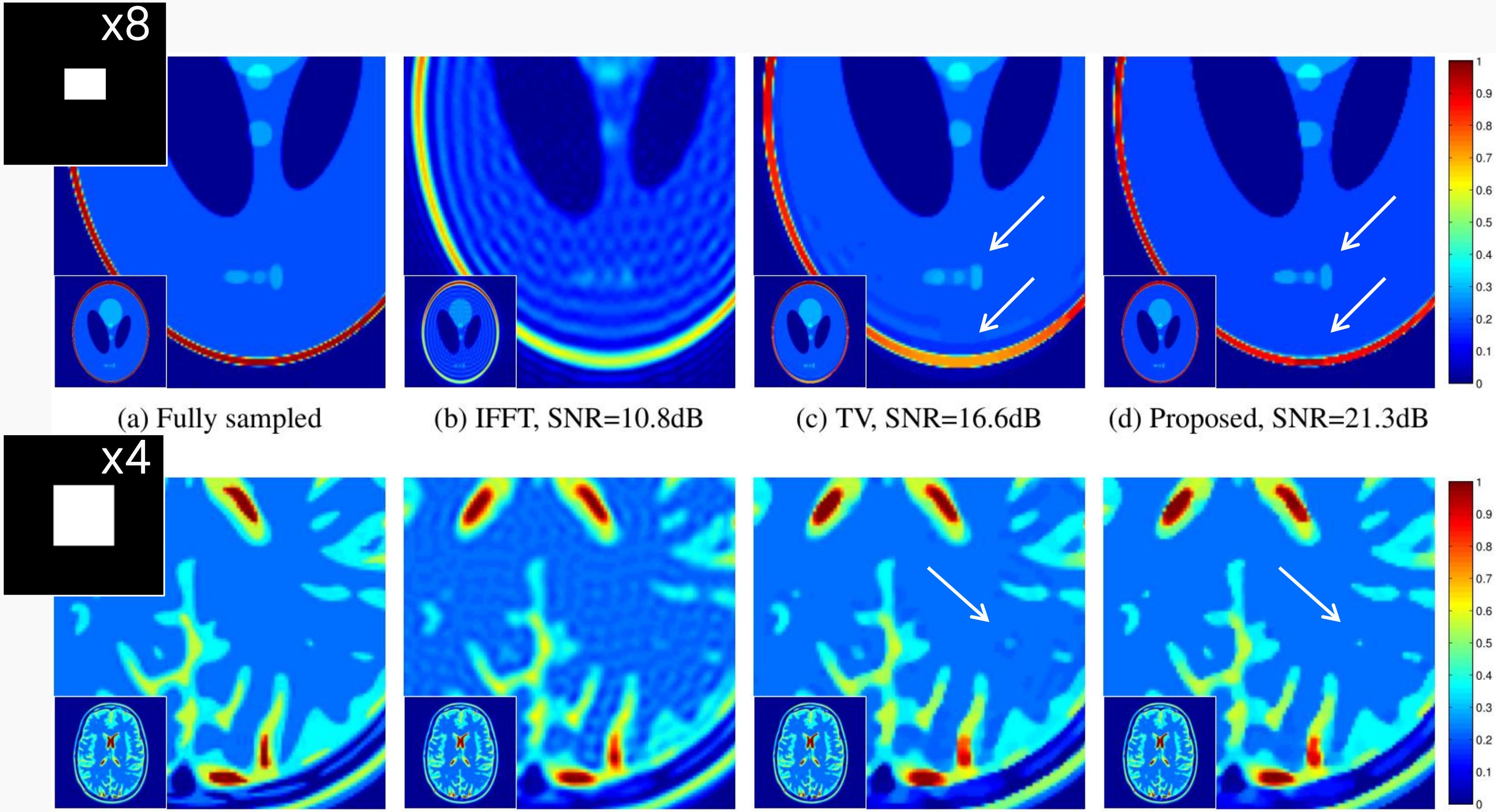
x = discrete spatial domain image

D = discrete gradient

A = Fourier undersampling operator

b = k-space samples

Recovery of MRI Medical Phantoms

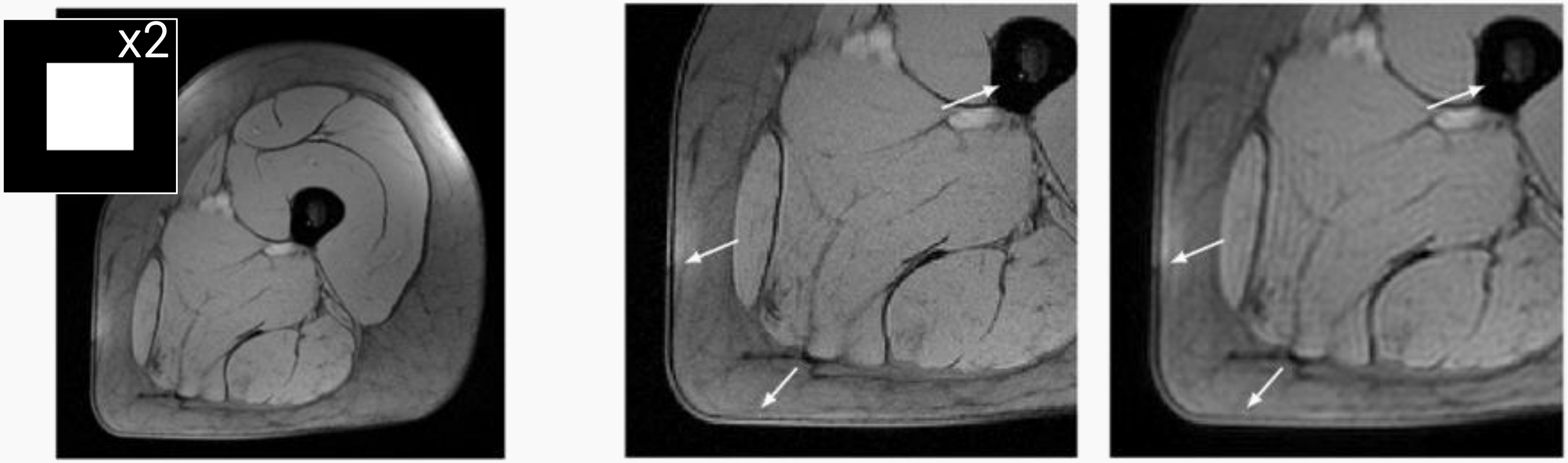


(a) Fully sampled (b) IFFT, SNR=10.8dB (c) TV, SNR=16.6dB (d) Proposed, SNR=21.3dB

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

Analytical phantoms from [Guerquin-Kern, 2012]

Recovery of Real MR Data



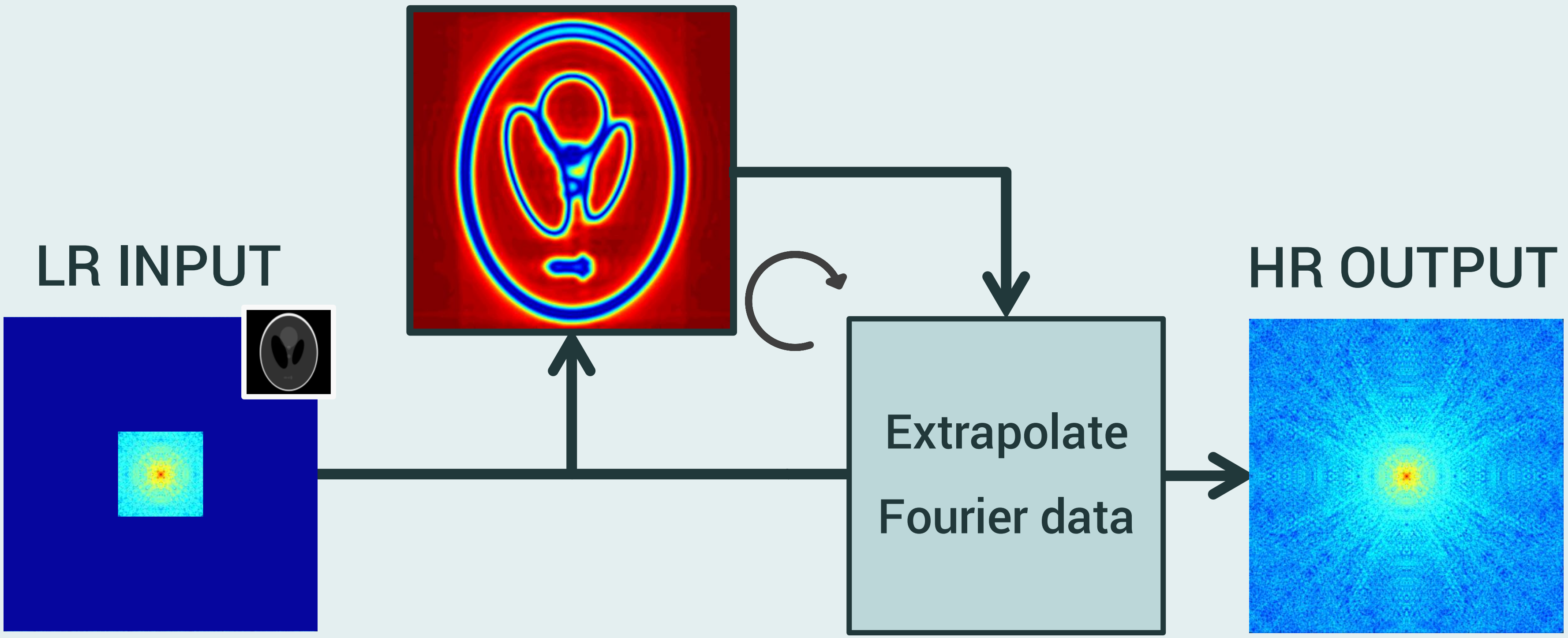
(a) Fully-sampled (b) Fully-sampled (zoom) (c) Zero-padded, SNR=20.1dB

(d) Edge mask (65×65 coefficients) (e) TV regularization, SNR=21.0dB (f) Proposed, SNR=21.1dB

4 Coil SENSE reconstruction w/phase

New Proposed One Stage Algorithm

Jointly estimate edge set and amplitudes



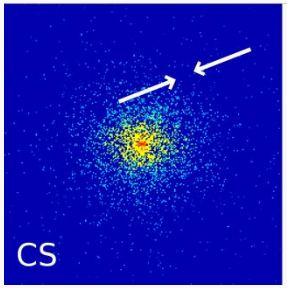
Off-the-grid

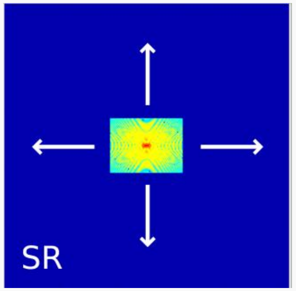
Pose recovery as a one-stage
structured low-rank matrix completion problem

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

$$\min_{\widehat{\mathbf{f}}} \quad \underbrace{\|\mathbf{P}\widehat{\mathbf{f}} - \mathbf{b}\|^2}_{\text{Data Consistency}} + \lambda \underbrace{\|\mathcal{T}(\widehat{\mathbf{f}})\|_*}_{\text{Regularization penalty}}$$

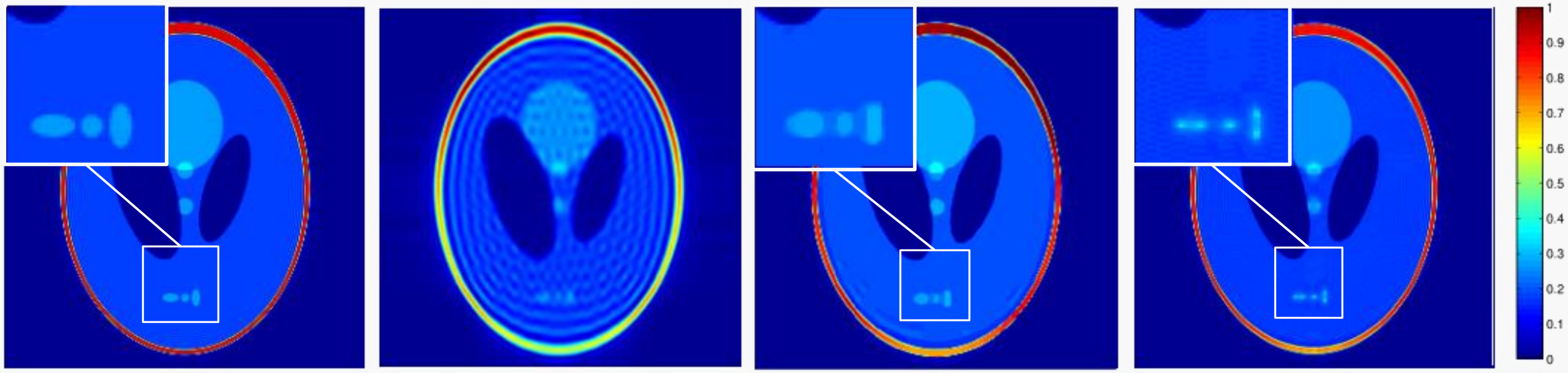
- Entirely off the grid
- Extends to CS paradigm

$\mathbf{P} =$ 

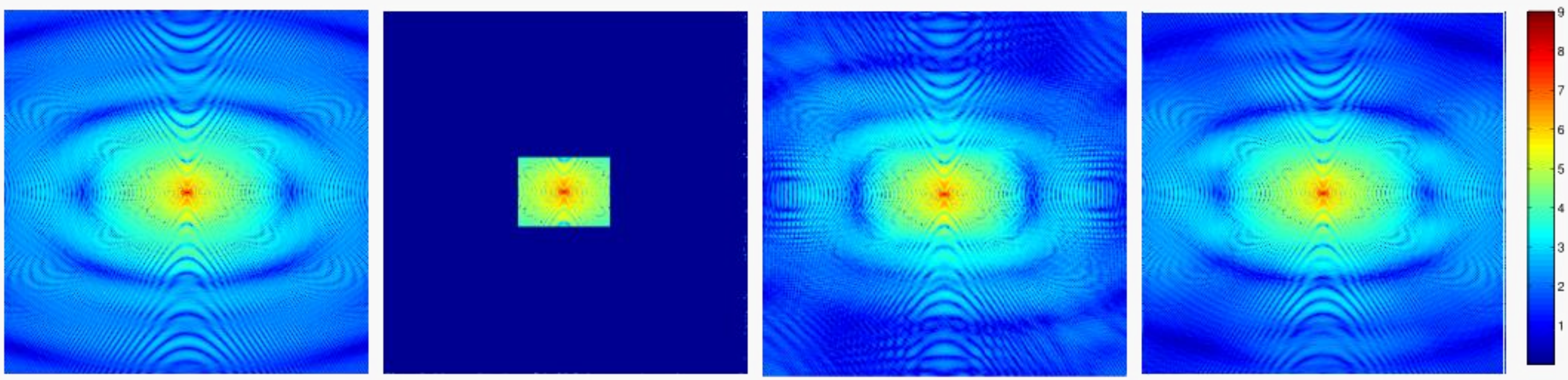
or 

- Use regularization penalty for other inverse problems

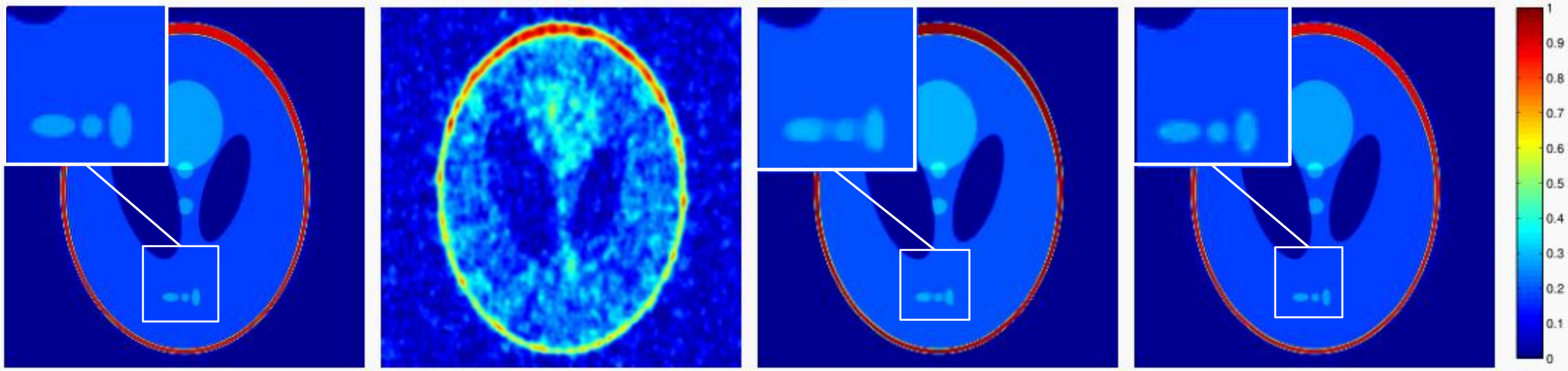
→off-the-grid alternative to TV, HDTV, etc



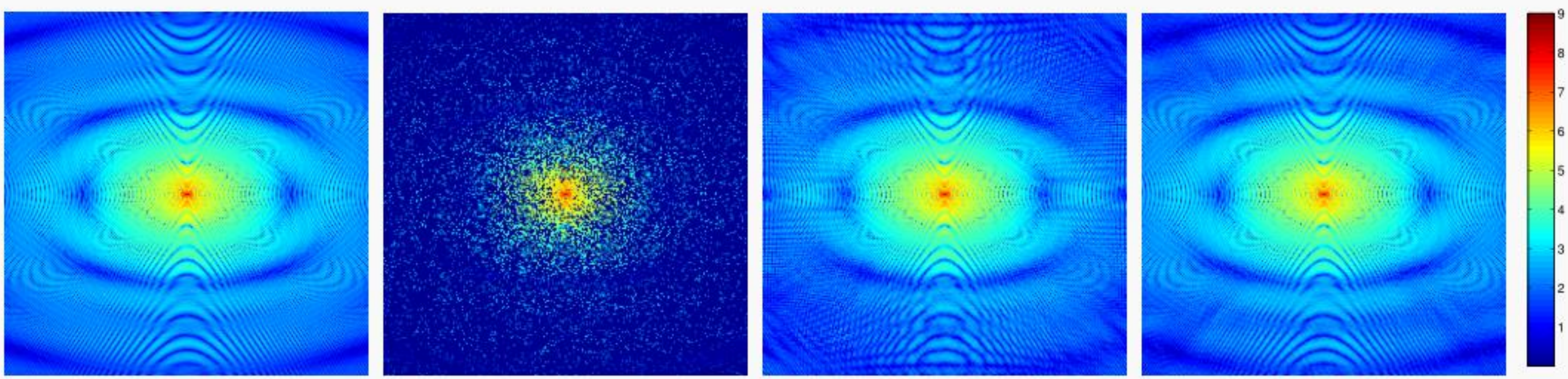
Fully sampled | Zeropadded IFFT | TV | rank min.



Fully sampled | Undersampled 20-fold | TV k-space | rank min k-space



Fully sampled   Zeropadded IFFT   TV   rank min.



Fully sampled   Undersampled 20-fold   TV k-space   rank min k-space

Computational challenges

- Naïve alg. is slow: ADMM + Singular value thresholding

$$\mathbf{dim}(\mathcal{T}(\widehat{\mathbf{f}})) \approx \text{ (2*window size) x (filter size)}$$

- Use matrix factorization trick:

$$\text{if } \mathbf{rank}(\mathbf{X}) \leq \mathbf{r},$$

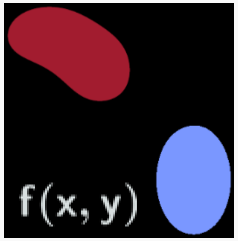
$$\|\mathbf{X}\|_* = \min_{\mathbf{X}=\mathbf{U}\mathbf{V}^H} \|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2$$

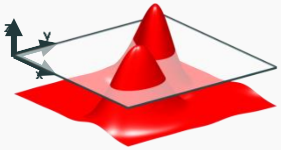
$$\mathbf{n} \times \mathbf{m} \qquad \mathbf{m} \times \mathbf{r} \qquad \mathbf{r} \times \mathbf{n}$$

- Future work: Exploit convolutional structure.

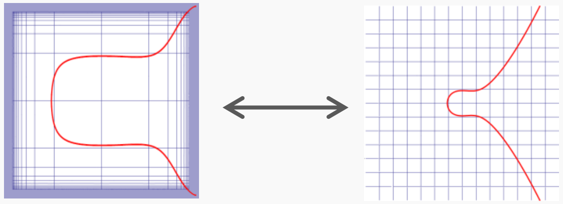
Summary

- New framework for higher dimensional FRI recovery
  - Extend annihilation relation to Piecewise smooth signal model





- Provide sampling guarantees for unique signal recovery
  - 2-D PWC Constant Signals
  - New Proof Techniques



- Novel Fourier domain structured low-rank penalty
  - Convex, Off-the-Grid, & widely applicable

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