IEEE International Symposium on Biomedical Imaging
April 18-21, 2017, Melbourne, Australia

Tutorial 3:

Continuous domain sparse recovery of biomedical imaging data using structured low-rank approaches

JONG CHUL YE, KAIST
MATHEWS JACOB, UNIV. OF IOWA

Joint work by several authors

CBIG, Univ. Iowa

1. Dr. Gregory Ongie
2. Dr. Merry Mani
3. Mr. Arvind Balachandrasekaran
4. Ms. Ipshita Bhattacharya
5. Ms. Sampurna Biswas

KAIST

1. Dr. Kyong Hwan Jin
2. Mr. Juyoung Lee
3. Dongwook Lee
4. Junhong Min

Overview

Part 1: Theory & Algorithms

Part 2: Applications

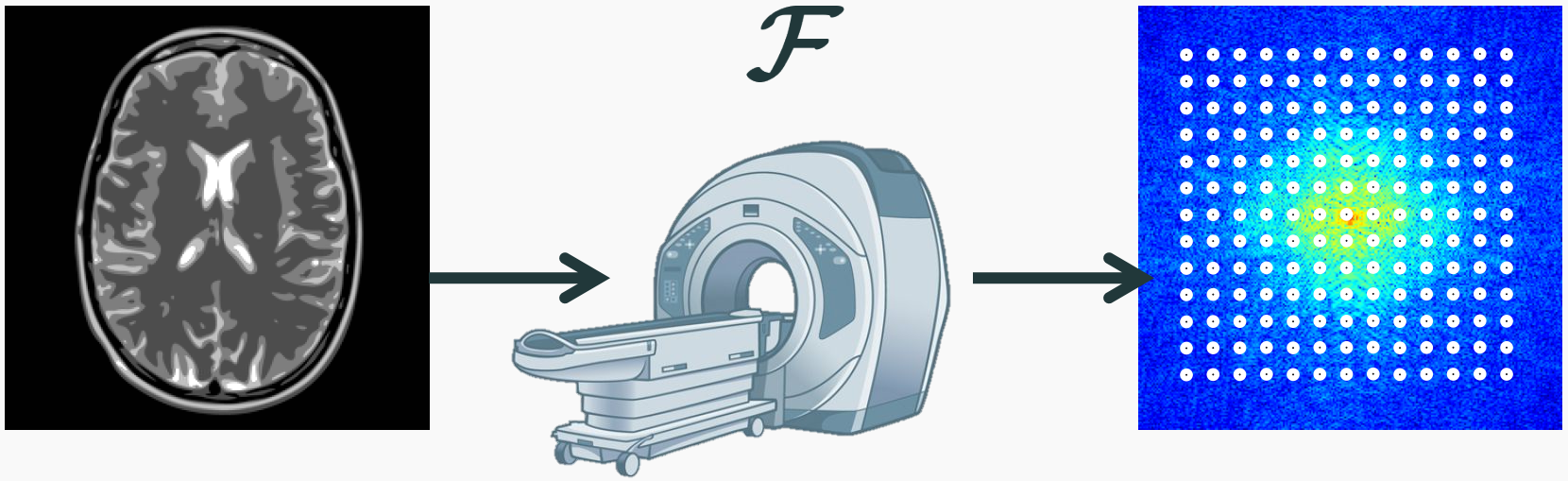
Overview

1. Introduction
2. Review of Compressive Sensing
3. FRI extrapolation from uniform samples
4. Structured low-rank interpolation for non-uniform samples
5. Fast implementations
6. Biomedical applications

Overview

1. Introduction
2. Review of Compressive Sensing
3. FRI extrapolation from uniform samples
4. Structured low-rank interpolation for non-uniform samples
5. Fast implementations
6. Biomedical applications

Motivation: MRI reconstruction

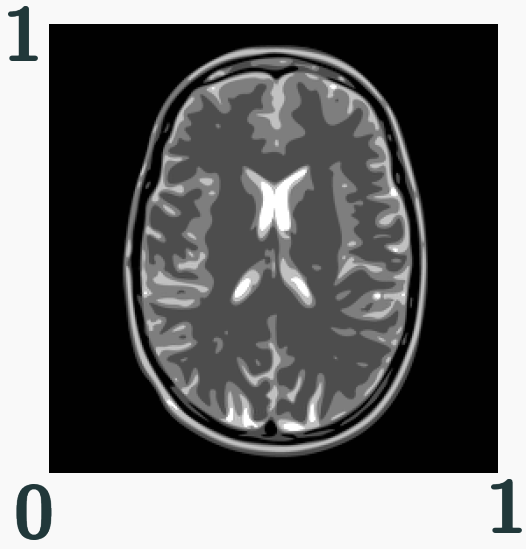


Main Problem:

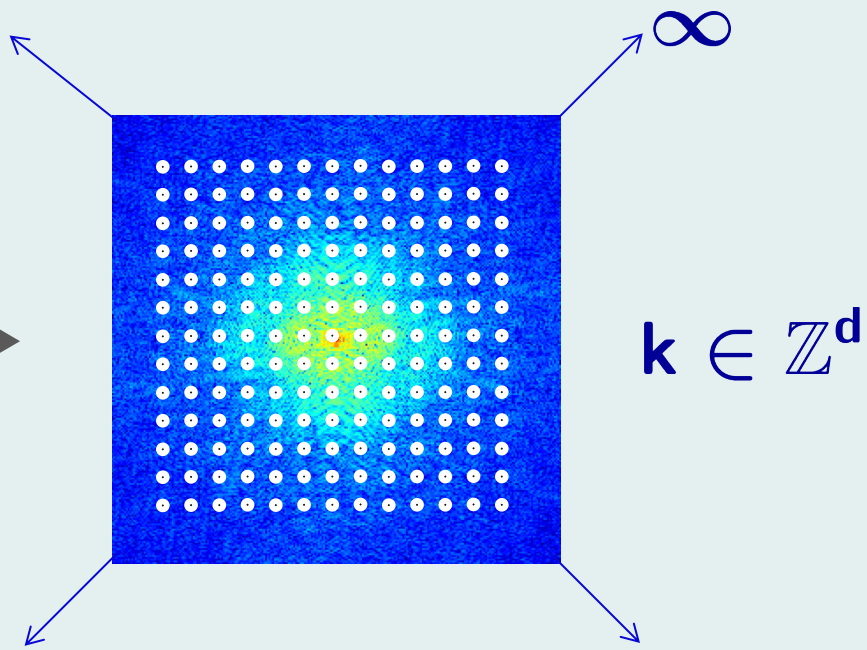
Reconstruct image from Fourier domain samples

Related: Computed Tomography, Florescence Microscopy

Motivation: MRI Reconstruction



$\mathcal{F}$



$\mathbf{k} \in \mathbb{Z}^{\mathbf{d}}$

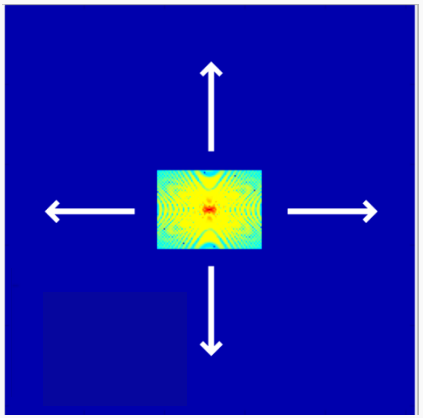
$\mathbf{f(x)}, \quad \mathbf{x} \in [\mathbf{0},\mathbf{1}]^{\mathbf{d}}$

$$\widehat{\mathbf{f}}[\mathbf{k}] := \int_{[\mathbf{0},\mathbf{1}]^{\mathbf{d}}} \mathbf{f(x)} \mathbf{e}^{-\mathbf{j}2\pi\mathbf{k}\cdot\mathbf{x}} \mathbf{dx}$$

Uniform Fourier Samples =
Fourier Series Coefficients

Types of "Compressive" Fourier Domain Sampling

low-pass

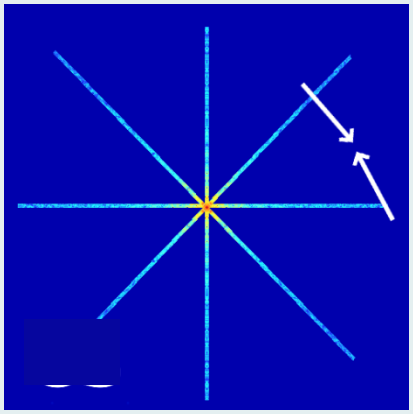


Fourier
Extrapolation

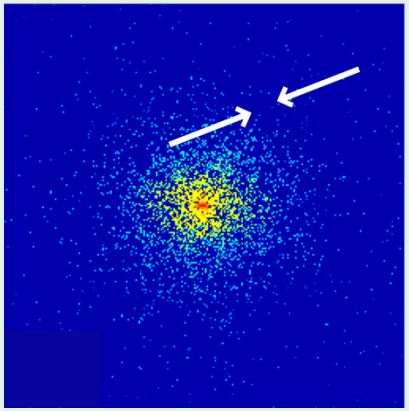
↓

Super-resolution
recovery

radial



random



Fourier
Interpolation

↓

"Compressed Sensing"
recovery

Extrapolation: super-resolution microscopy

The Nobel Prize in Chemistry 2014



Photo: Matt Staley/HHMI

Eric Betzig

Prize share: 1/3



© Bernd Schuller, Max-Planck-Institut

Stefan W. Hell

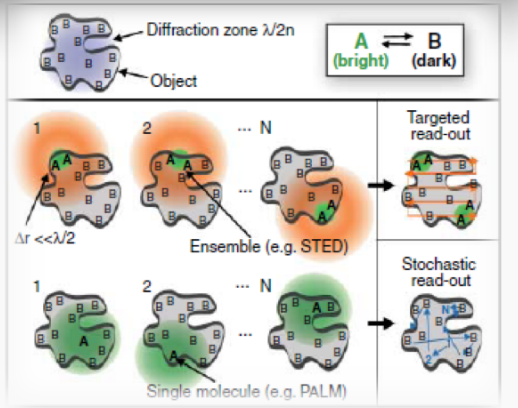
Prize share: 1/3

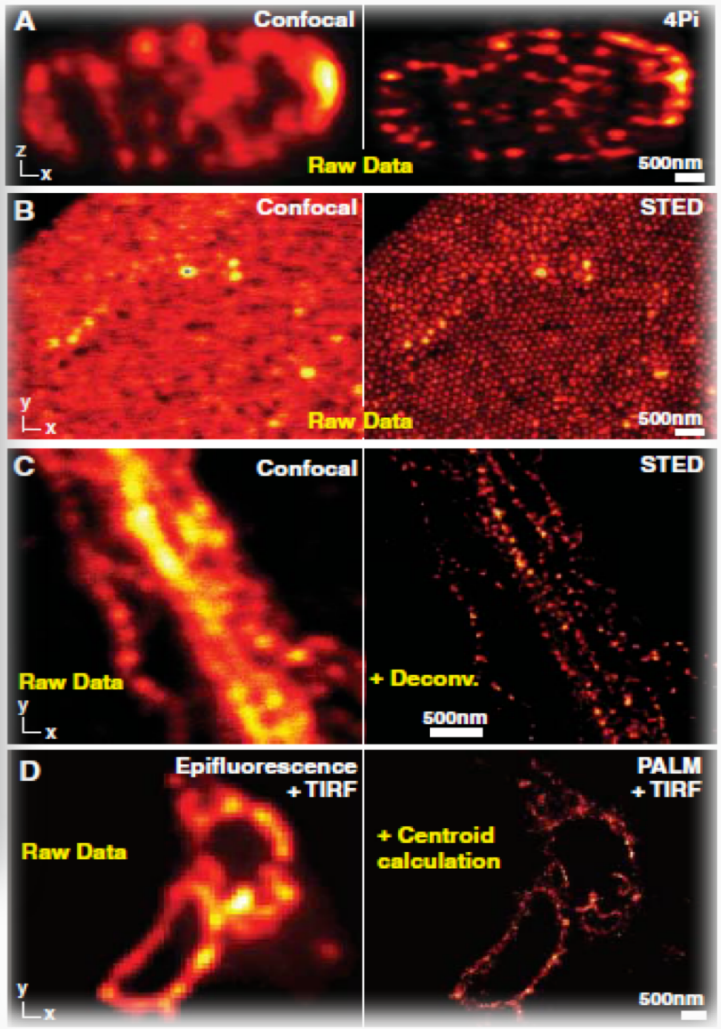


Photo: K. Lowder via Wikimedia Commons, CC-BY-SA-3.0

William E. Moerner

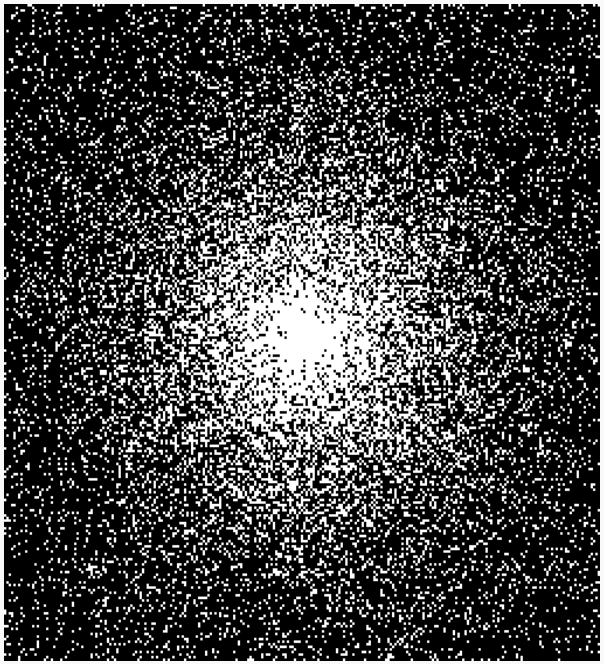
Prize share: 1/3





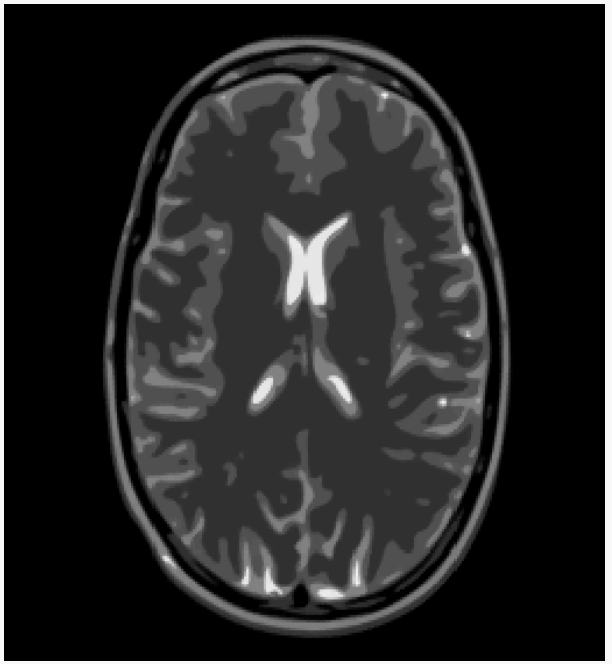
S. Hell et al, Science 2007.

Interpolation: accelerated MRI



25% Random
Fourier samples
(variable density)

TV-min



Rel. Error = 5%

Overview

1. Introduction
2. Review of Compressive Sensing
3. FRI extrapolation from uniform samples
4. Structured low-rank interpolation for non-uniform samples
5. Fast implementations
6. Biomedical applications

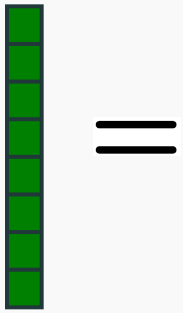
Compressed Sensing (CS)

- Incoherent projection
- Underdetermined system
- Sparse unknown vector

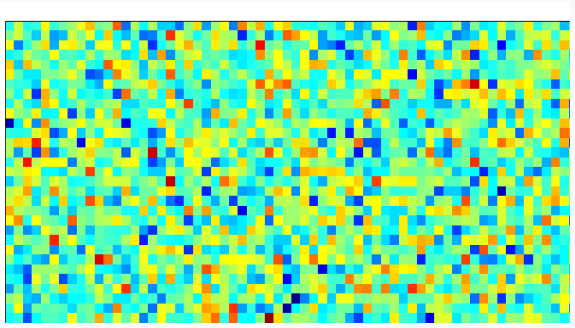
b **A** **x**

$M \times 1$

measurements



$=$





$N \times 1$

sparse signal

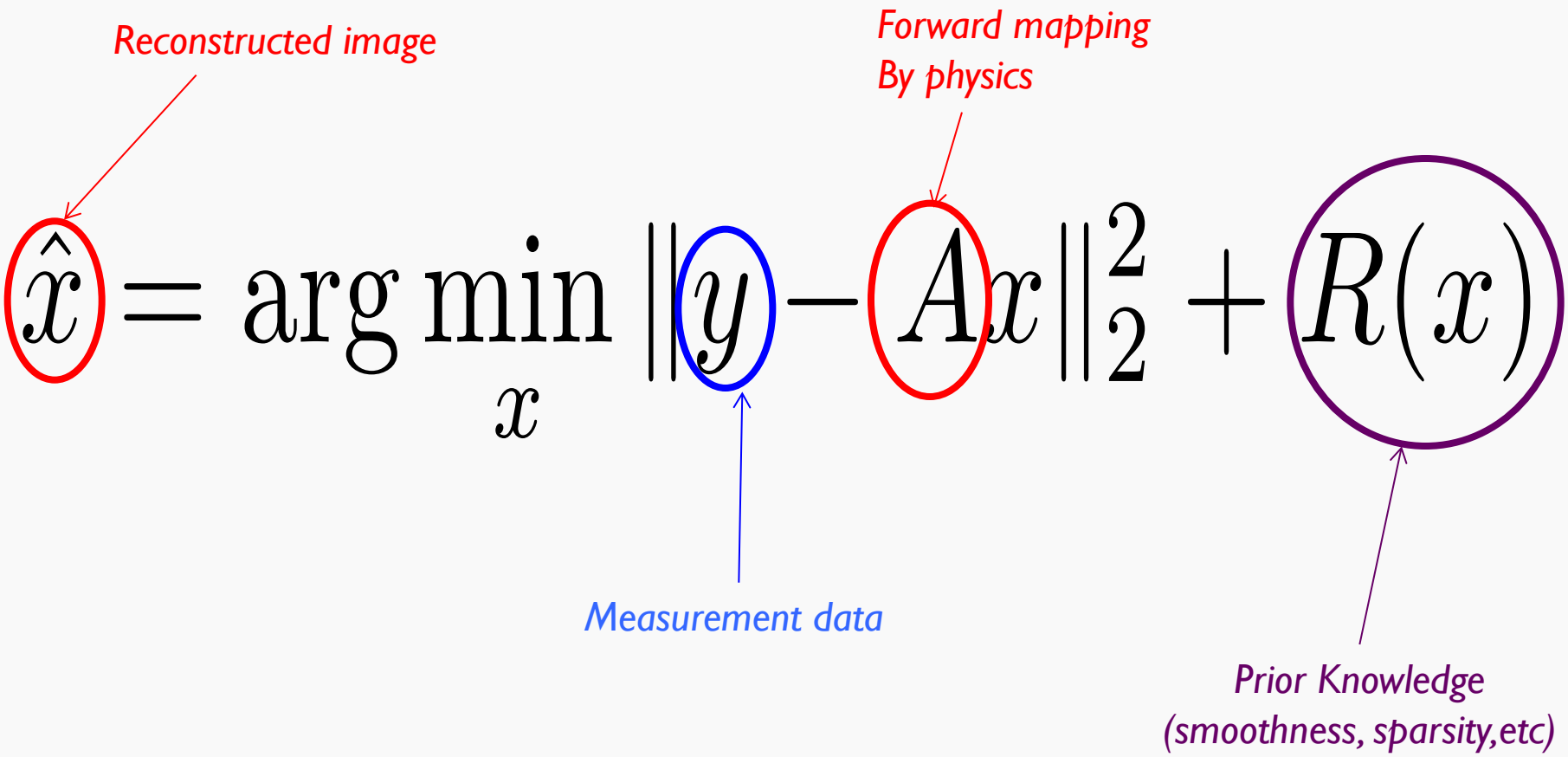
K

 # non-zeros

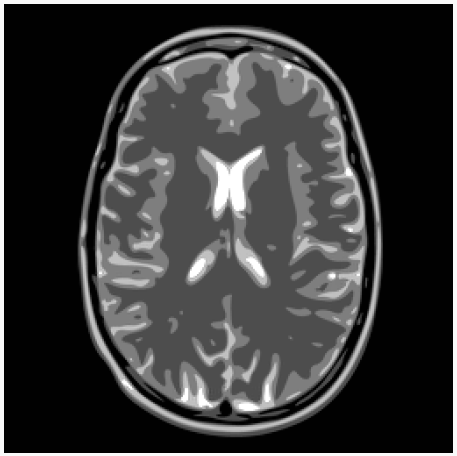
$$M \approx K \log(N) \ll N$$

Courtesy of Dr. Dror Baron

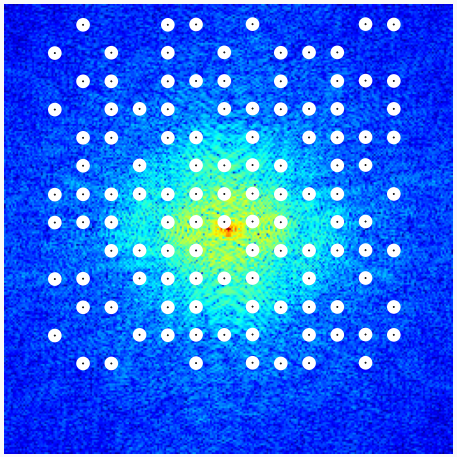
Sparse-Low Rank Recovery in Nutshell



Application to biomedical imaging



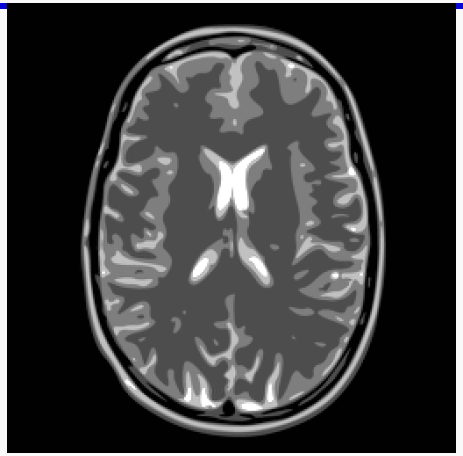
$\mathcal{F}$



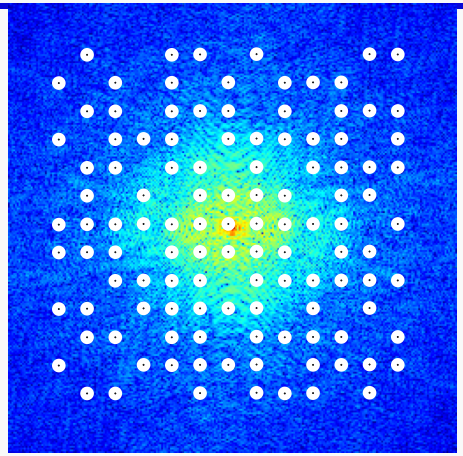
Full sampling is costly!

Randomly undersample

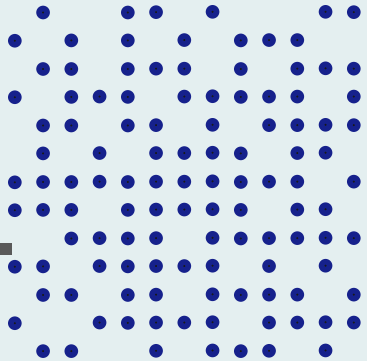
Application to biomedical imaging



$\mathcal{F}$

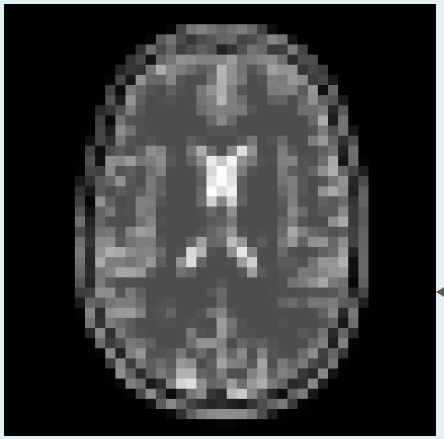


Randomly Undersample

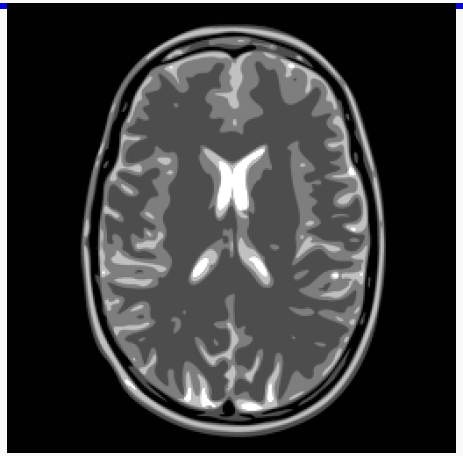


Sparse Model

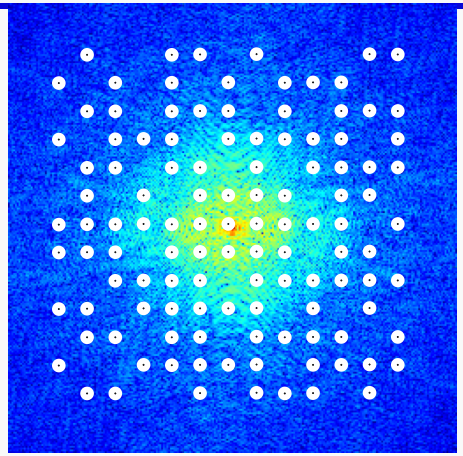
Convex Optimization



Analysis formulation of Compressed Sensing

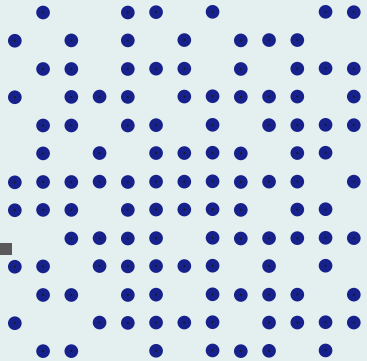


$\mathcal{F}$

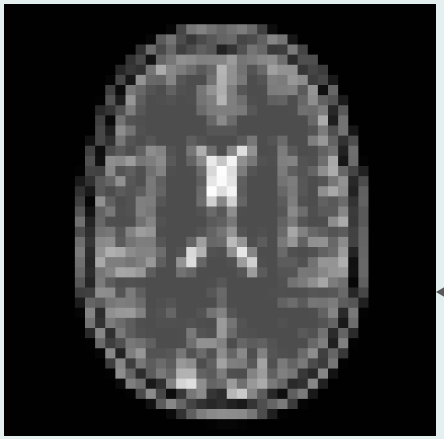


Randomly Undersample

Sparse Model



Convex Optimization

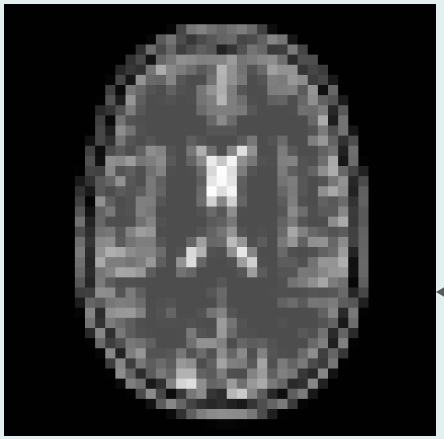


Example:

Assume discrete gradient of image is sparse

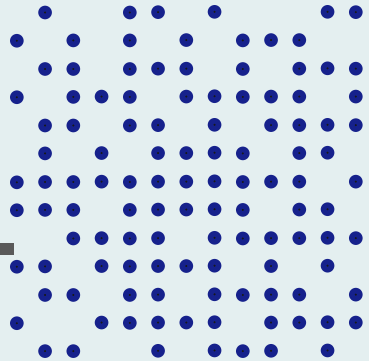


Piecewise constant model



Sparse Model

Convex Optimization



Recovery by Total Variation (TV) minimization

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

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

$$\sum_{i,j}$$



Recovery by Total Variation (TV) minimization

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

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

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



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

Recovery by Total Variation (TV) minimization

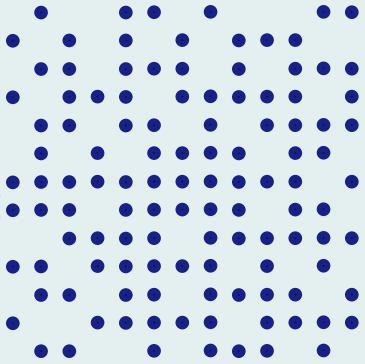
TV semi-norm: $\|\mathbf{g}\|_{\mathsf{TV}} = \sum_{\mathsf{i,j}} \sqrt{|\mathbf{g}_{\mathsf{i}+1,\mathsf{j}} - \mathbf{g}_{\mathsf{i,j}}|^2 + |\mathbf{g}_{\mathsf{i,j}+1} - \mathbf{g}_{\mathsf{i,j}}|^2}$

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

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

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

Restricted DFT

$\Omega =$ 

Sample locations

Recovery by Total Variation (TV) minimization

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

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

$$\sum_{i,j}$$



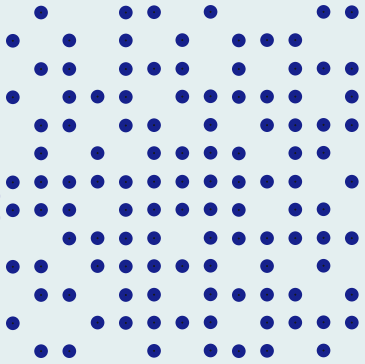
$$\min_{\mathbf{g} \in \mathbb{C}^{N \times N}} \|\mathbf{g}\|_{\mathsf{TV}} \quad \text{subject to} \quad \mathbf{F}_\Omega \mathbf{g} = \mathbf{F}_\Omega \mathbf{f} \qquad (\mathbf{TV}\text{-min})$$

Convex optimization problem

Fast iterative algorithms:
ADMM/Split-Bregman,
FISTA, Primal-Dual, etc.

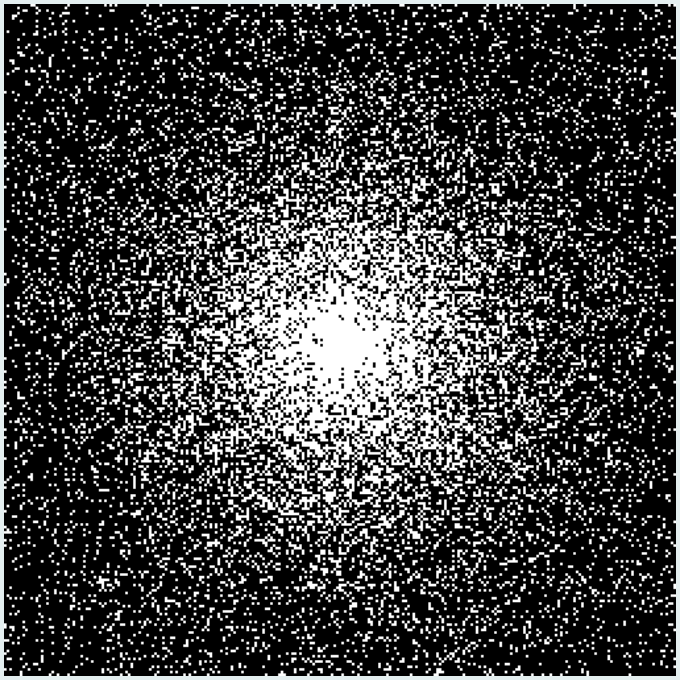
Restricted DFT

$\Omega =$



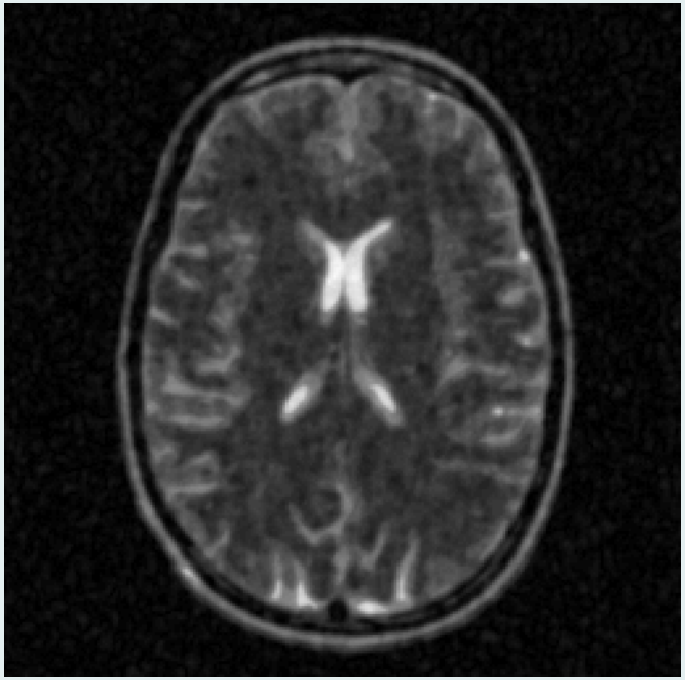
Sample locations

Recovery using zero filled IFFT



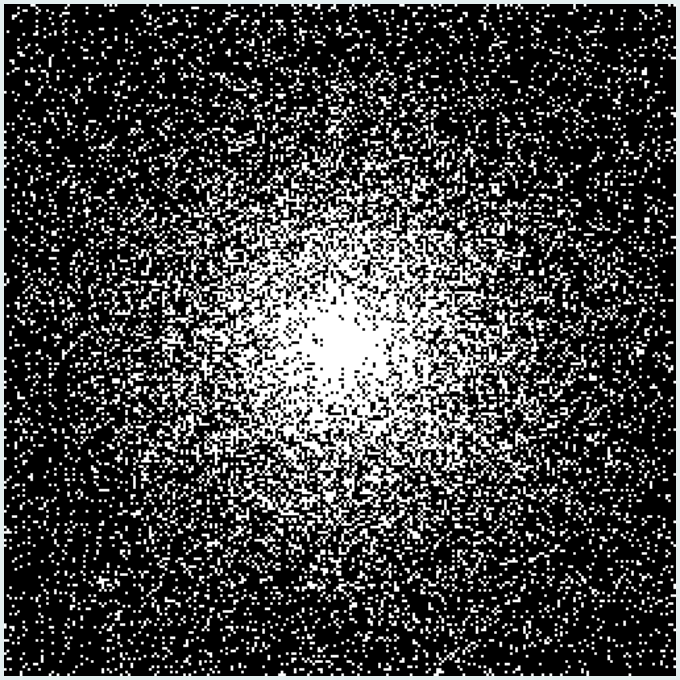
25% Random
Fourier samples
(variable density)

DFT^{-1}



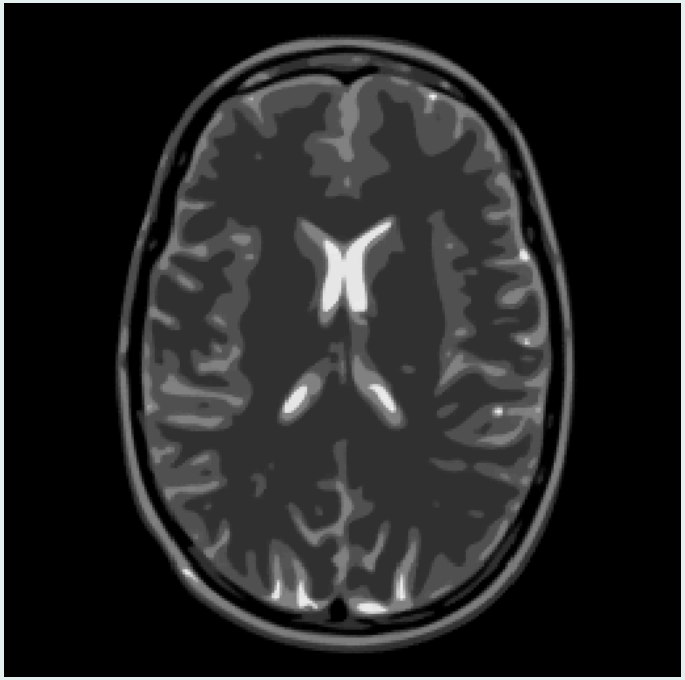
Rel. Error = 30%

Recovery using TV minimization



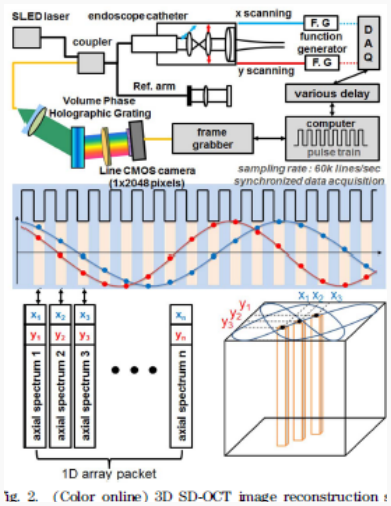
25% Random
Fourier samples
(variable density)

TV-min



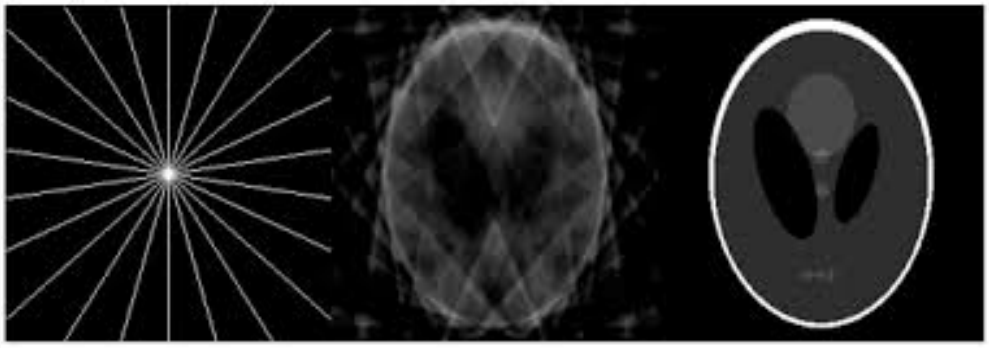
Rel. Error = 5%

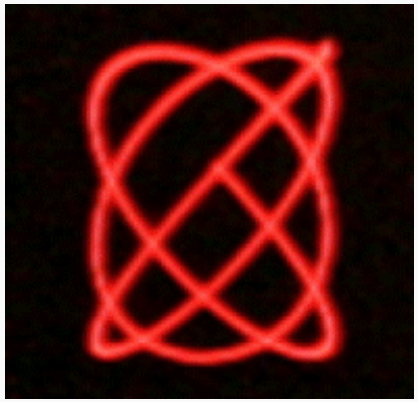
Limitations of CS



ig. 2. (Color online) 3D SD-OCT image reconstruction

- Discrete domain theory
- Explicit form of sensing matrix
- RIP issue → no direct interpolation





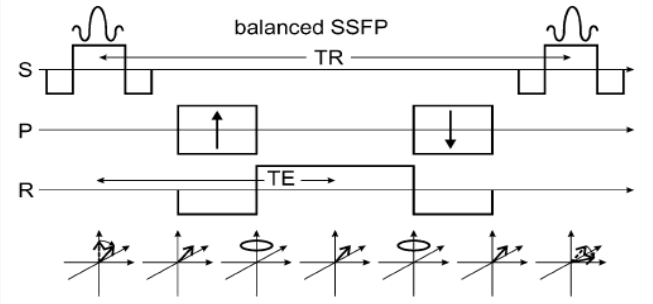


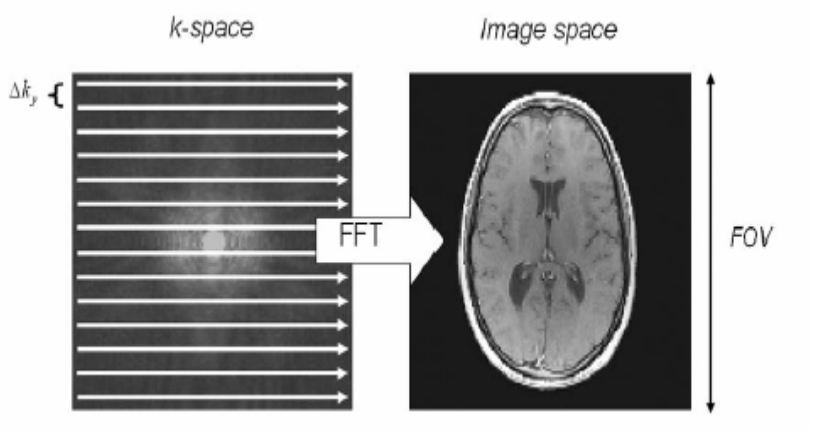


Analytic Reconstruction

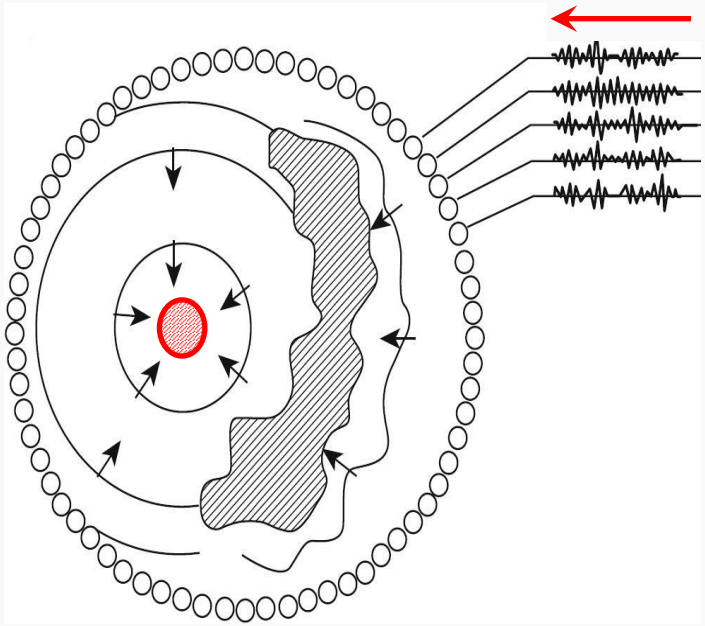
Beautiful analytic reconstruction results from fully sampled data

(a) MR Imaging





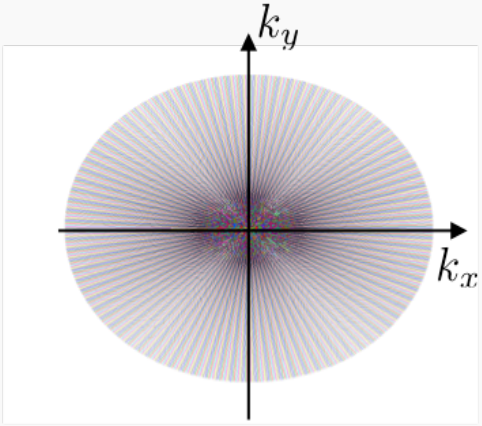
(b) Time-reversal of a scattered wave



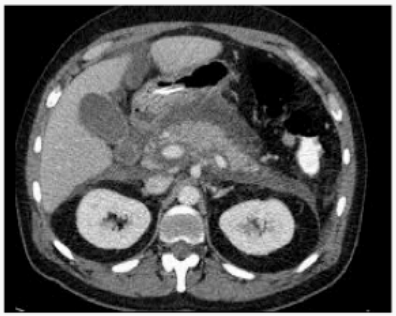
$$\mathcal{I}_1(x) = \int_\Omega f(z) \int_0^T \int_{\partial\Omega} \frac{\partial G(x,y,T,t)}{\partial \nu_y} \frac{\partial \Gamma}{\partial t}(z,y,0,T-t) d\sigma(y) dt\, dz.$$

Project Goal: Unification of CS and analytic reconstruction for biomedical imaging using a 2-layer approach

Dense samples



Analytic Reconstruction



Rank $H(x)$

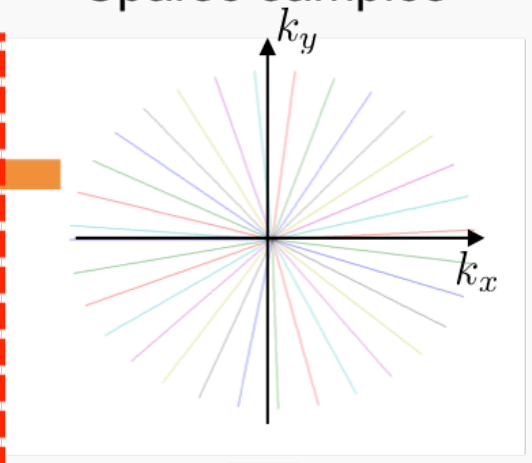
"The Missing Link"

?

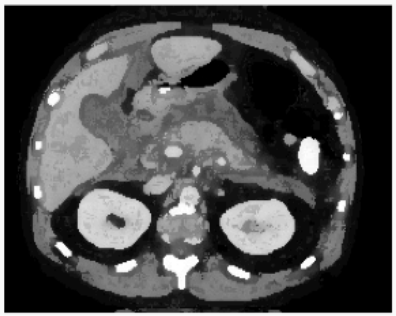
dual

$\|x\|_0$

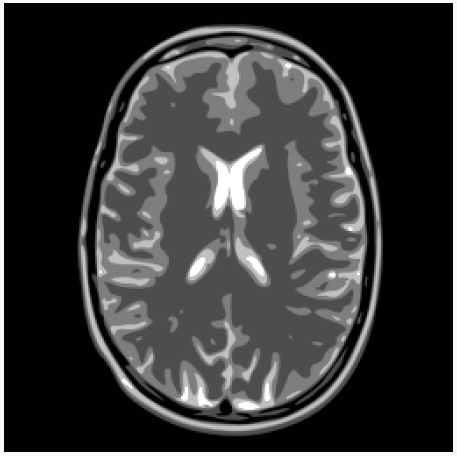
Sparse samples



Compressed Sensing

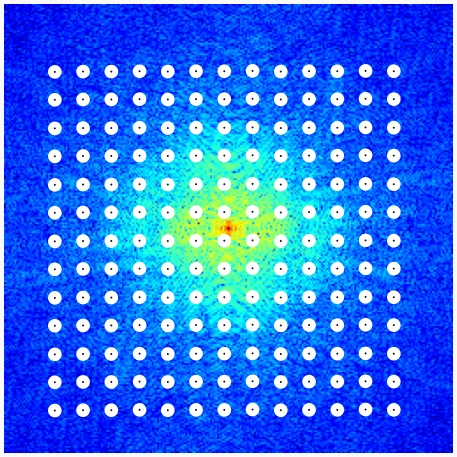


“True” measurement model:



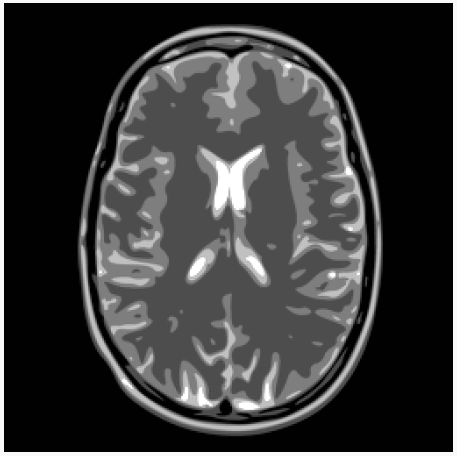
Continuous

$\mathcal{F}$

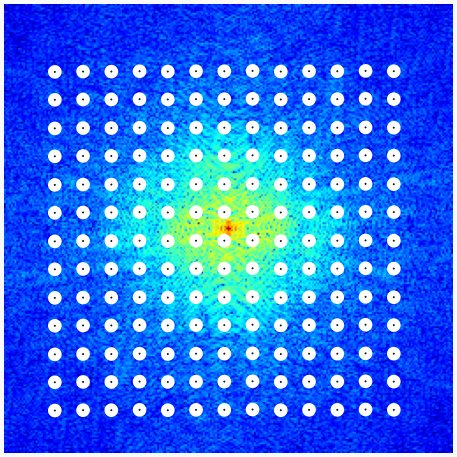


Continuous

“True” measurement model:



$\mathcal{F}$

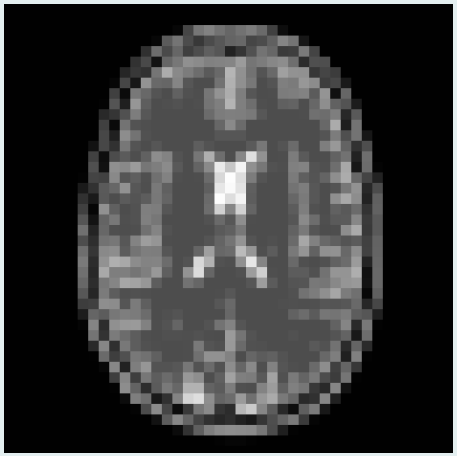


Continuous

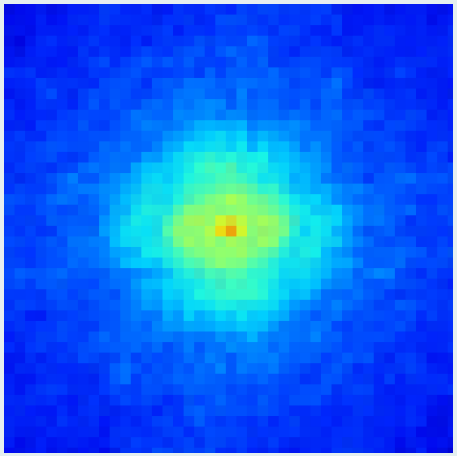
Continuous

Approximated measurement model:

$\approx$



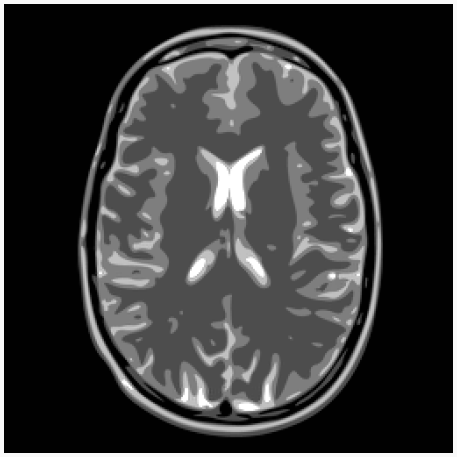
DFT



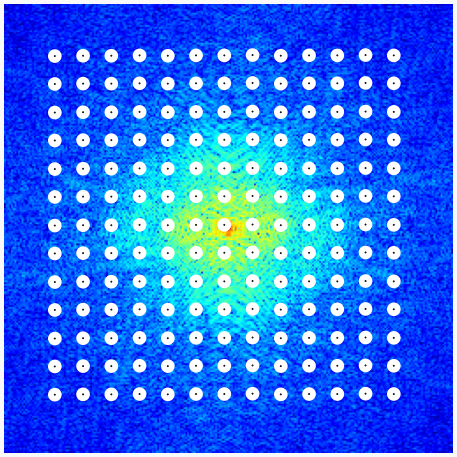
DISCRETE

DISCRETE

DFT Reconstruction

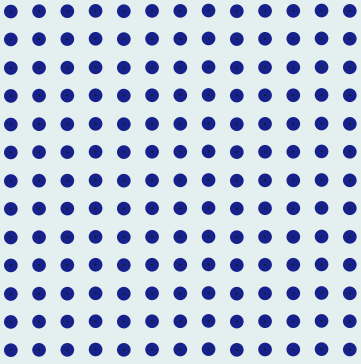


$\mathcal{F}$

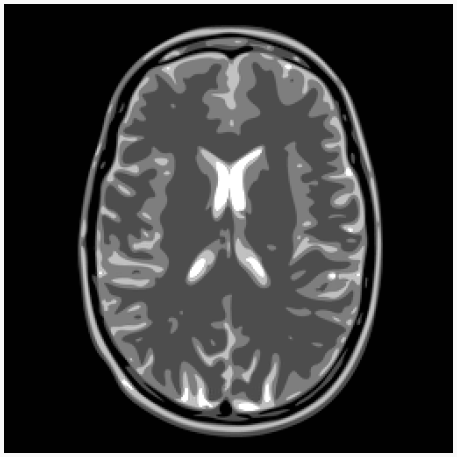


Continuous

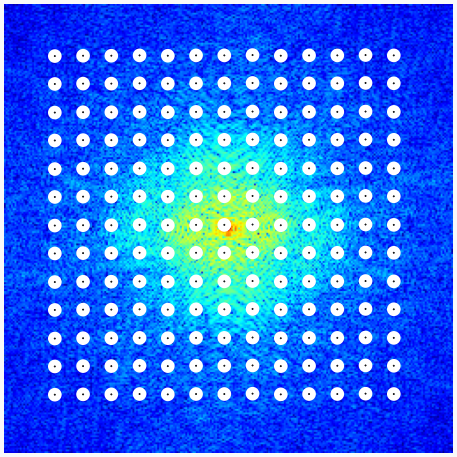
Continuous



DFT Reconstruction

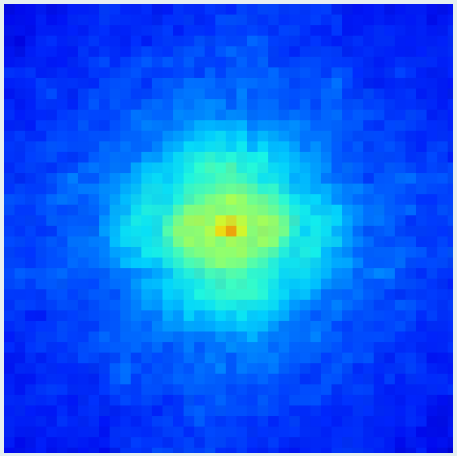


$\mathcal{F}$



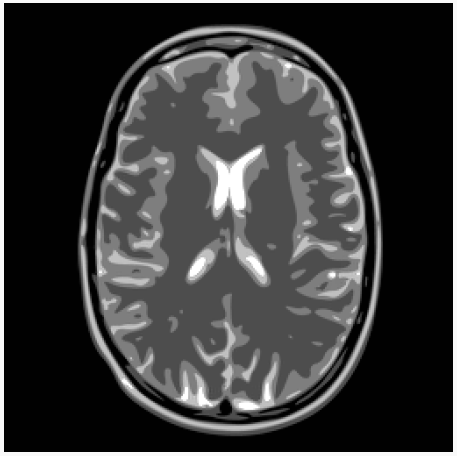
Continuous

Continuous



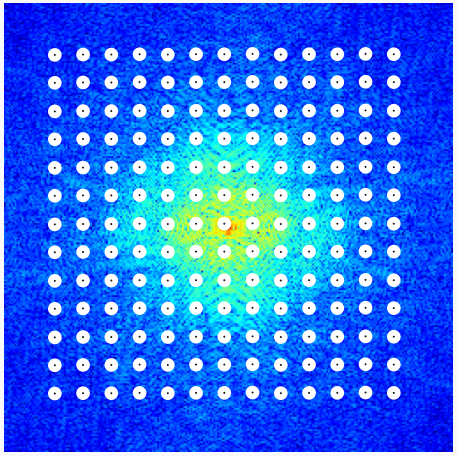
DISCRETE

DFT Reconstruction

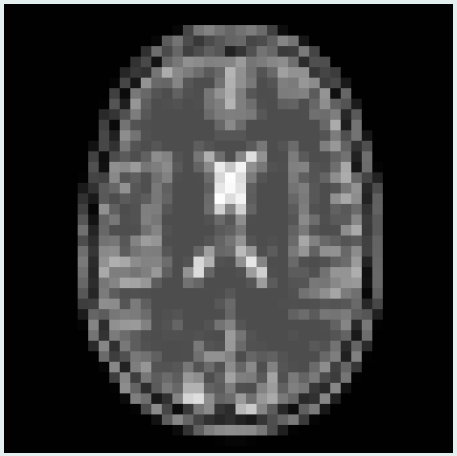


Continuous

$\mathcal{F}$

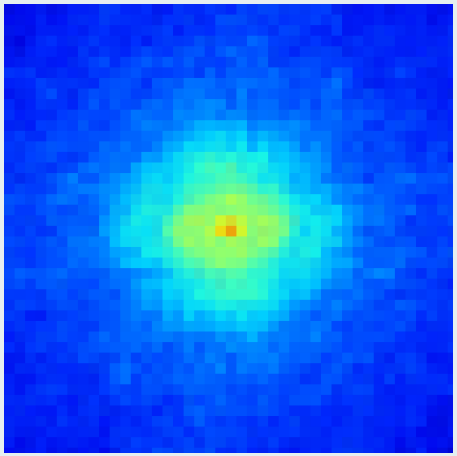


Continuous



DISCRETE

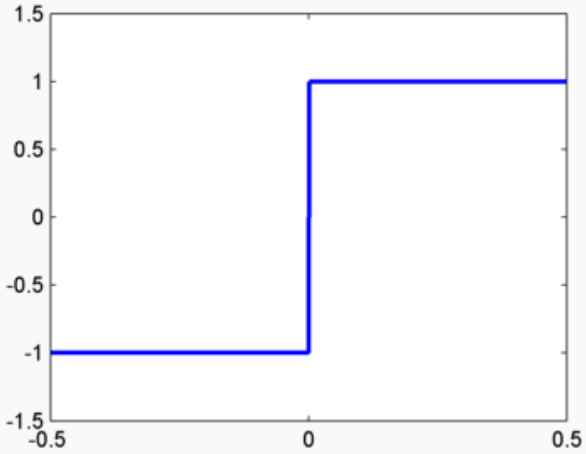
DFT^{-1}



DISCRETE

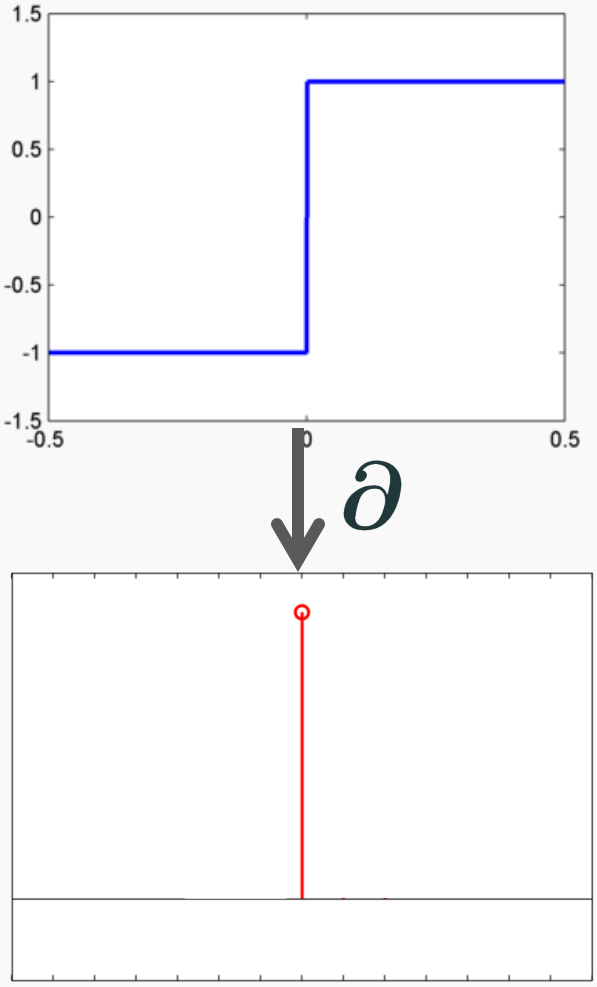
Challenge: Discrete approximation destroys sparsity!

Continuous



Challenge: Discrete approximation destroys sparsity!

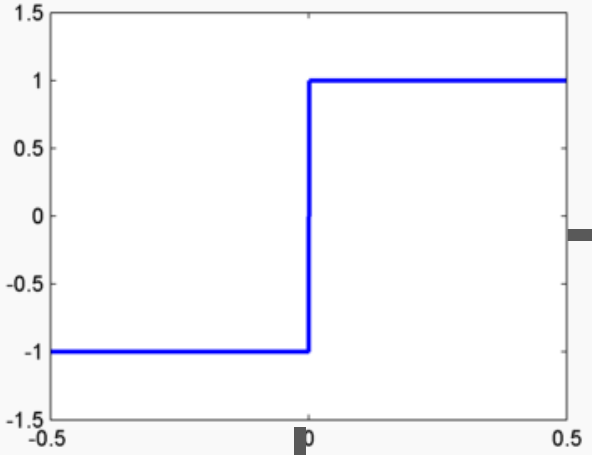
Continuous



Exact Derivative

Challenge: Discrete approximation destroys sparsity!

Continuous

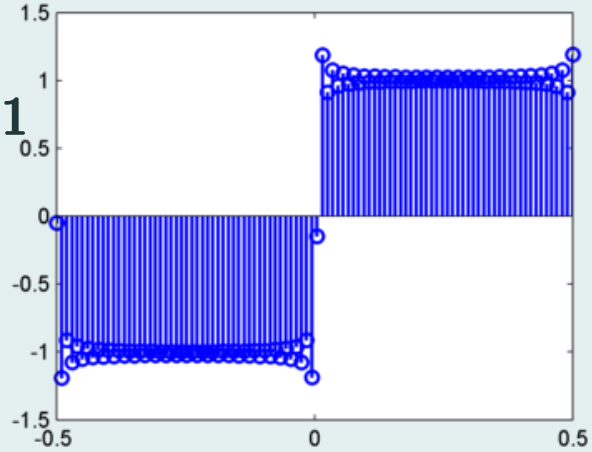


$\mathcal{F}$

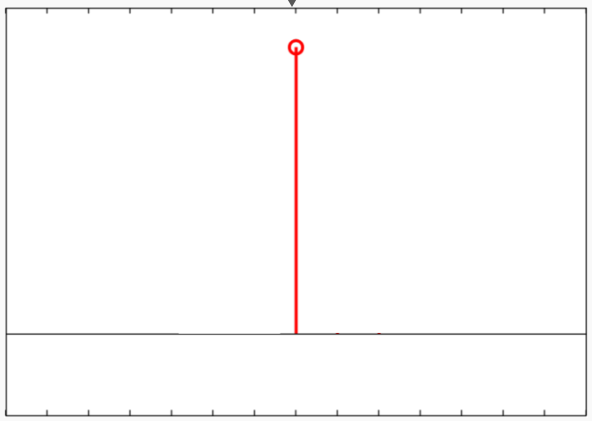
Sample

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

DISCRETE



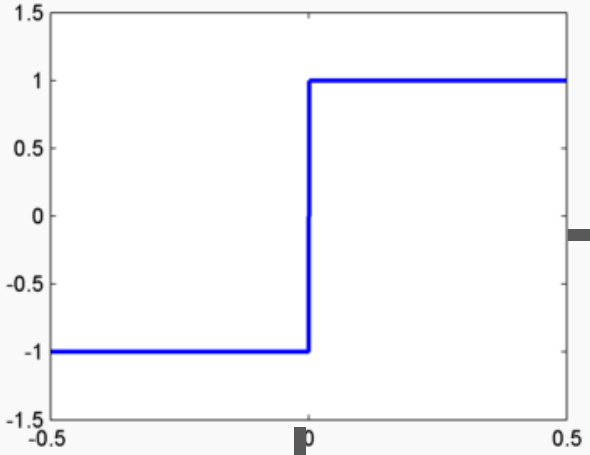
∂



Exact Derivative

Challenge: Discrete approximation destroys sparsity!

Continuous

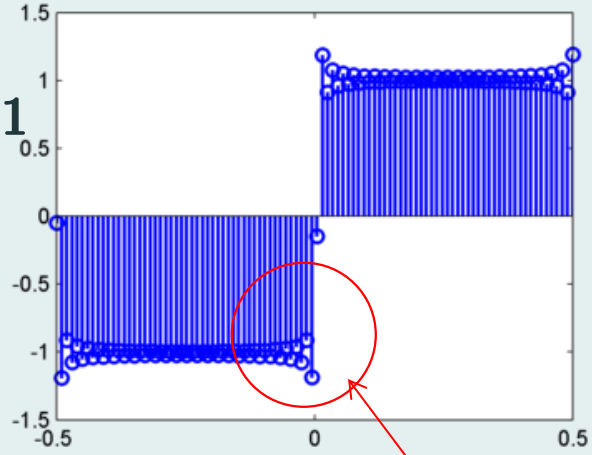


$\mathcal{F}$

Sample

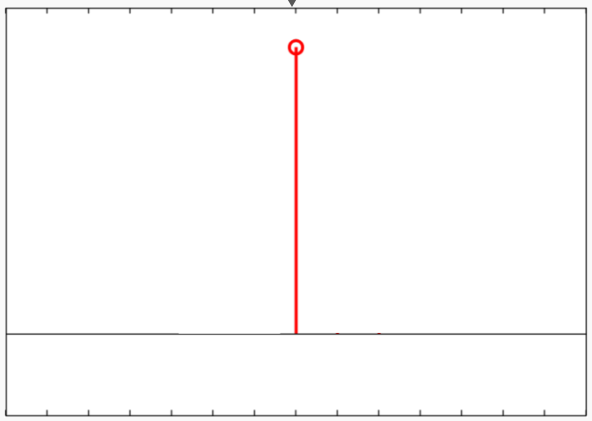
$\mathbf{DFT}^{-1}$

DISCRETE



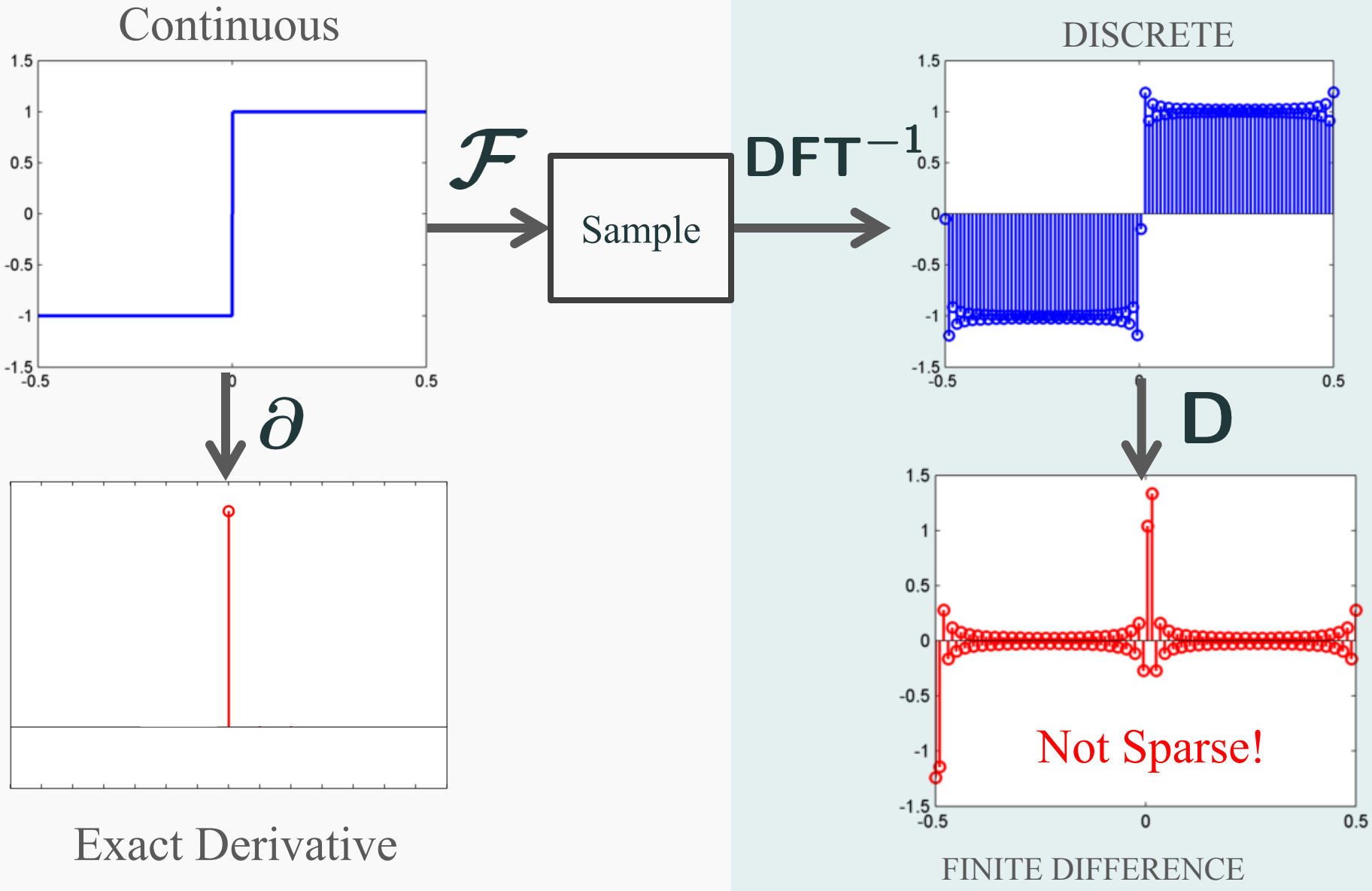
Gibb's Ringing!

∂



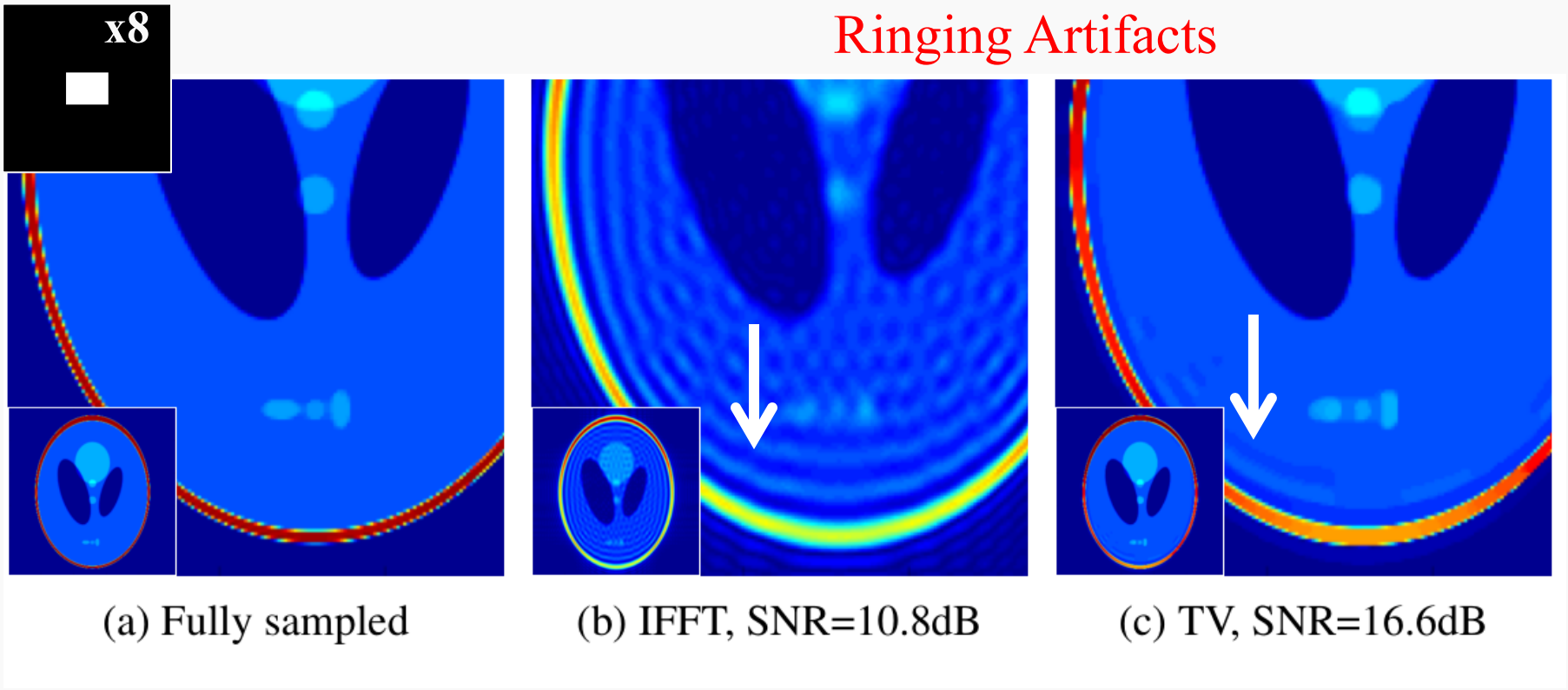
Exact Derivative

Challenge: Discrete approximation destroys sparsity!



Super-resolution setting: ringing artifacts !!

Fourier



Overview

1. Introduction

2. Review of Compressive Sensing

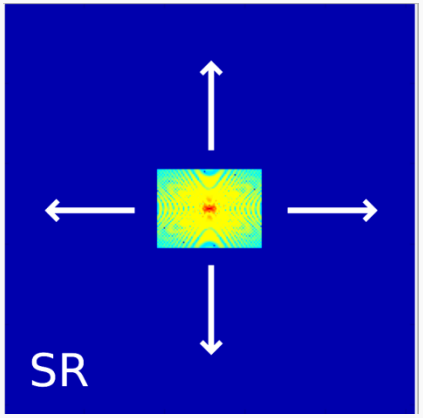
3. FRI extrapolation from uniform samples

    1-D Theory

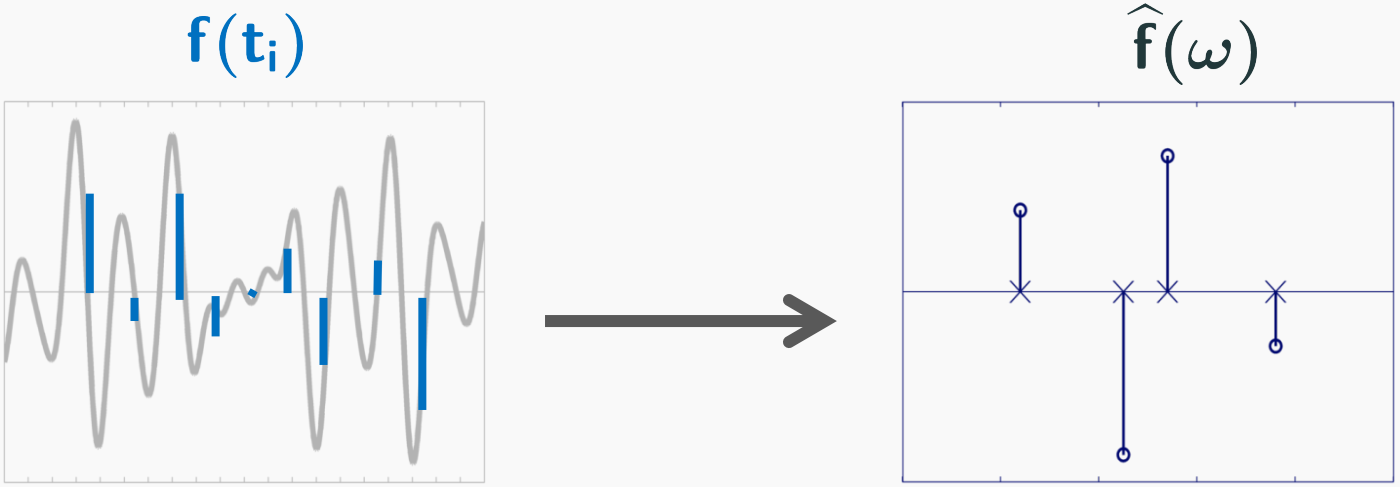
4. Structured low-rank interpolation for non-uniform samples

5. Fast implementations

6. Biomedical applications



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



Uniform time samples

Off-the-grid frequencies

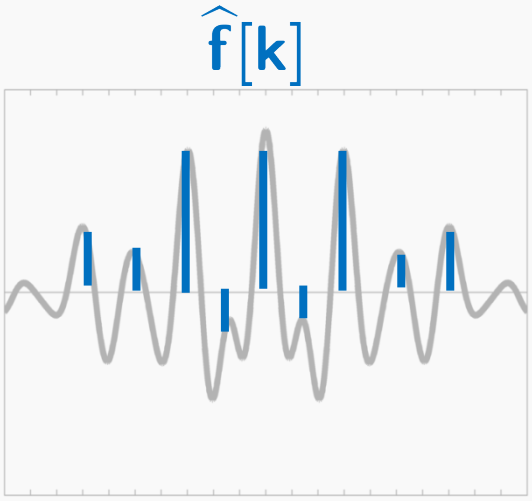
- Robust variants:

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

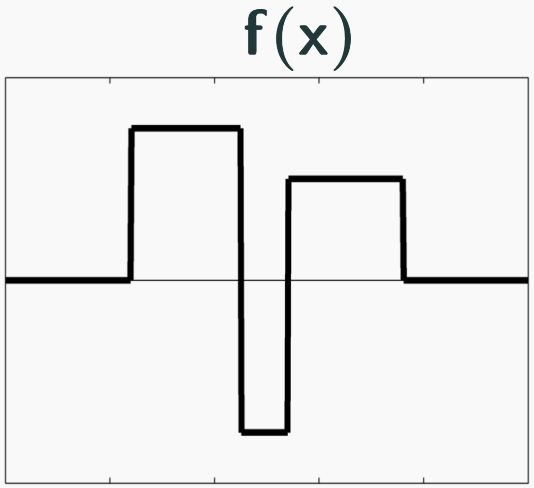
  Matrix pencil (1990) . . . Atomic norm (2011)

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

[Vetterli et al., 2002]



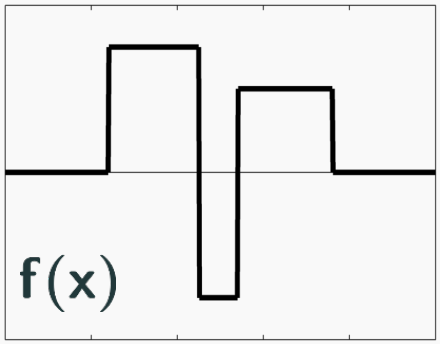
Uniform
Fourier samples



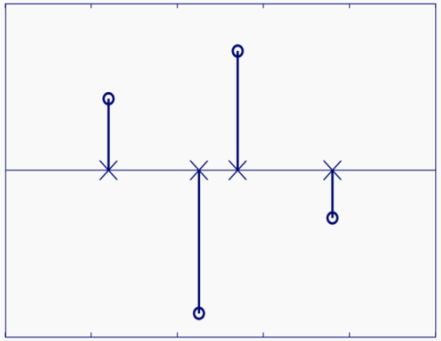
Off-the-grid
PWC signal

spatial domain

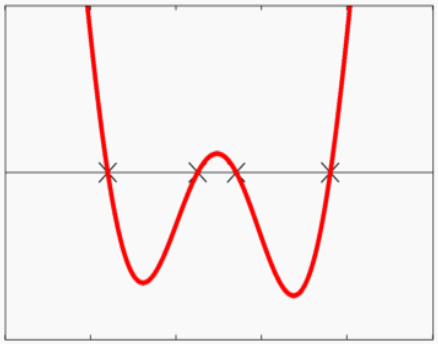
multiplication



∂



•

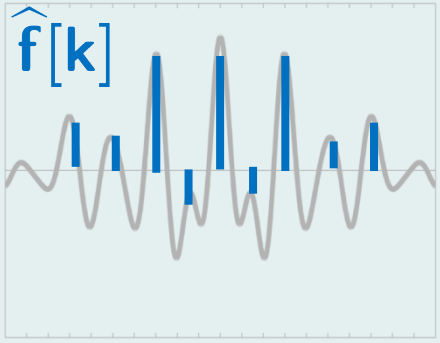


$= 0$

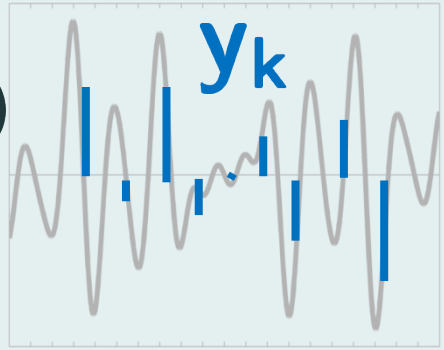
annihilating function

Fourier domain

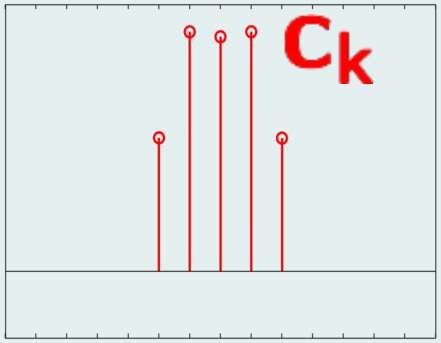
convolution



$(j2\pi k)$



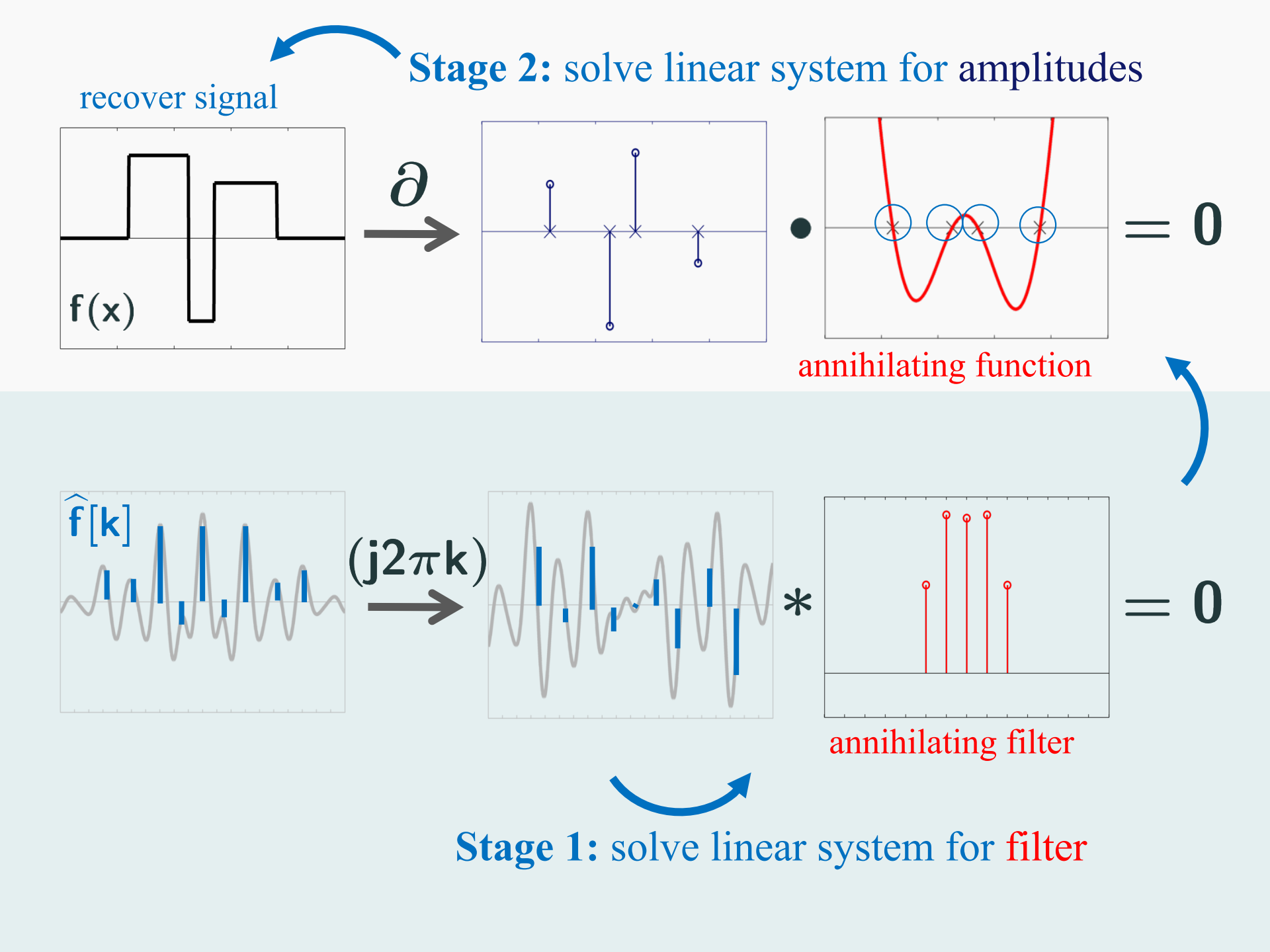
$*$



$= 0$

annihilating filter

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


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

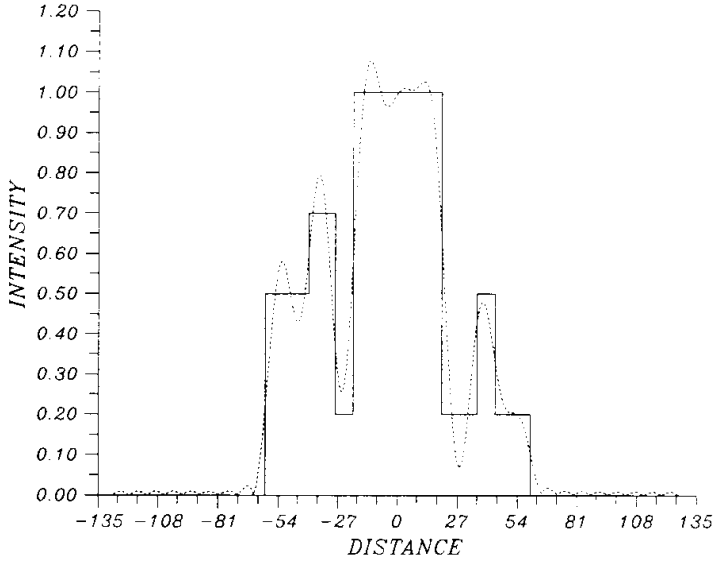
Similar 1-D FRI idea by [Liang & Hacke 1989]

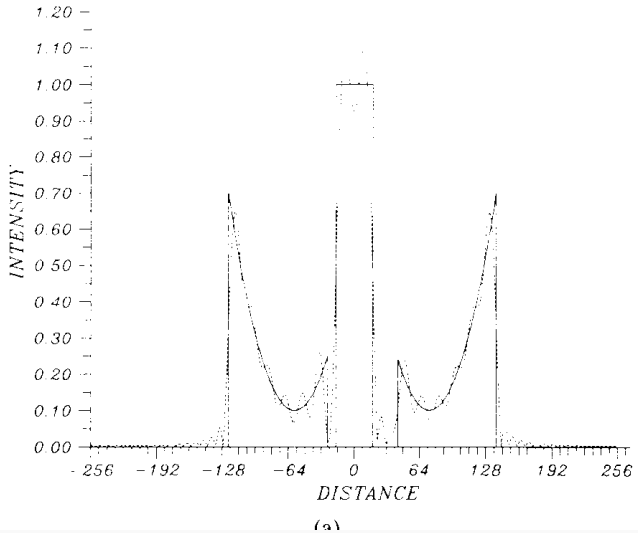
IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, VOL. 37, NO. 4, APRIL 1989

Superresolution Reconstruction Through Object Modeling and Parameter Estimation

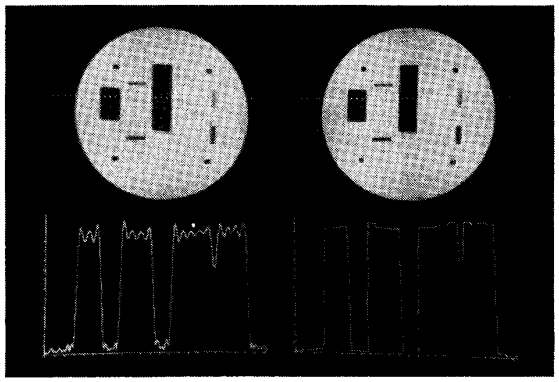
E. MARK HAACKE, ZHI-PEI LIANG, AND STEVEN H. IZEN

***Abstract*—Fourier transform reconstruction with limited data is often encountered in tomographic imaging problems. Conventional techniques, such as FFT-based methods, the spatial-support-limited extrapolation method, and the maximum entropy method, have not been optimal in terms of both Gibbs ringing reduction and resolution enhancement. In this correspondence, a new method based on object modeling and parameter estimation is proposed to achieve superresolution reconstruction.**





(a)



(a) (b)

Fig. 2. (a) Fourier reconstruction of a phantom from real magnetic resonance data using 256 data points in the vertical direction and 64 points in the horizontal direction. (b) Same as (a), but vertical direction is reconstructed using the proposed method. An example profile through the phantom shows the improvement in image behavior.

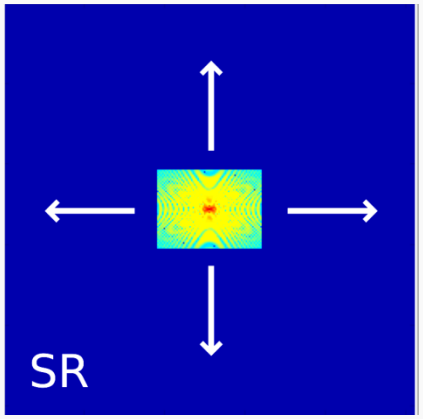
Overview

1. Introduction

2. Review of Compressive Sensing

3. FRI extrapolation from uniform samples

    2-D Theory



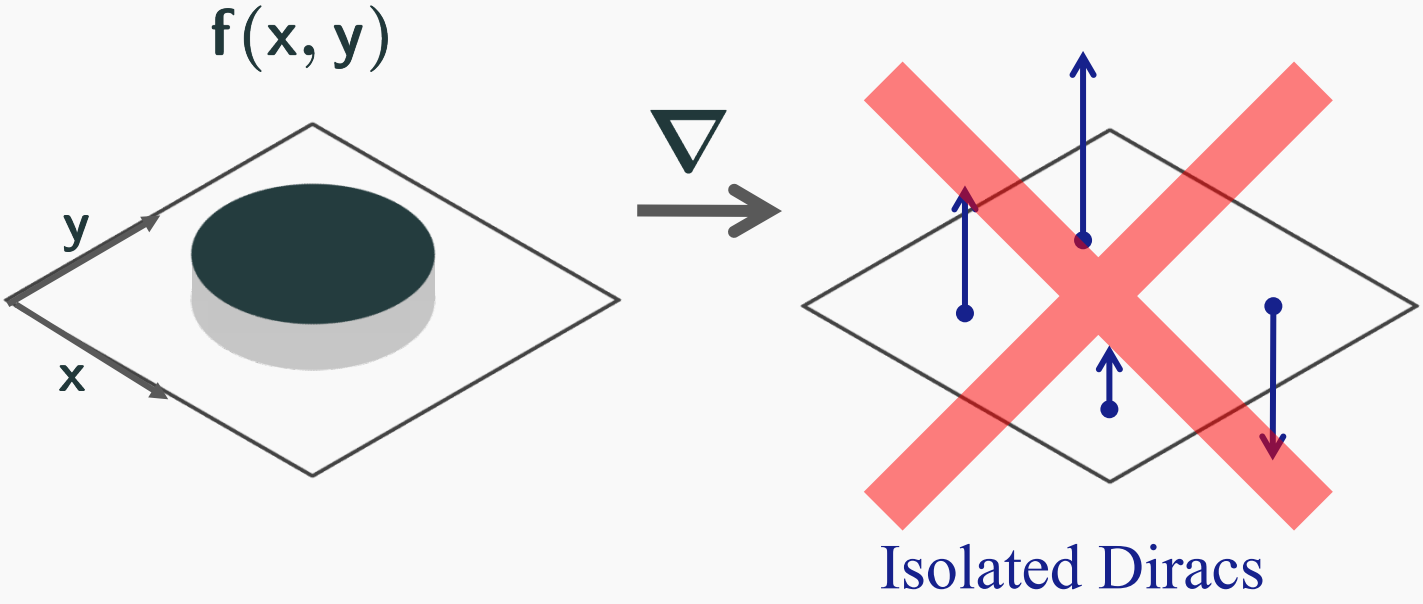
4. Structured low-rank interpolation for non-uniform samples

5. Fast implementations

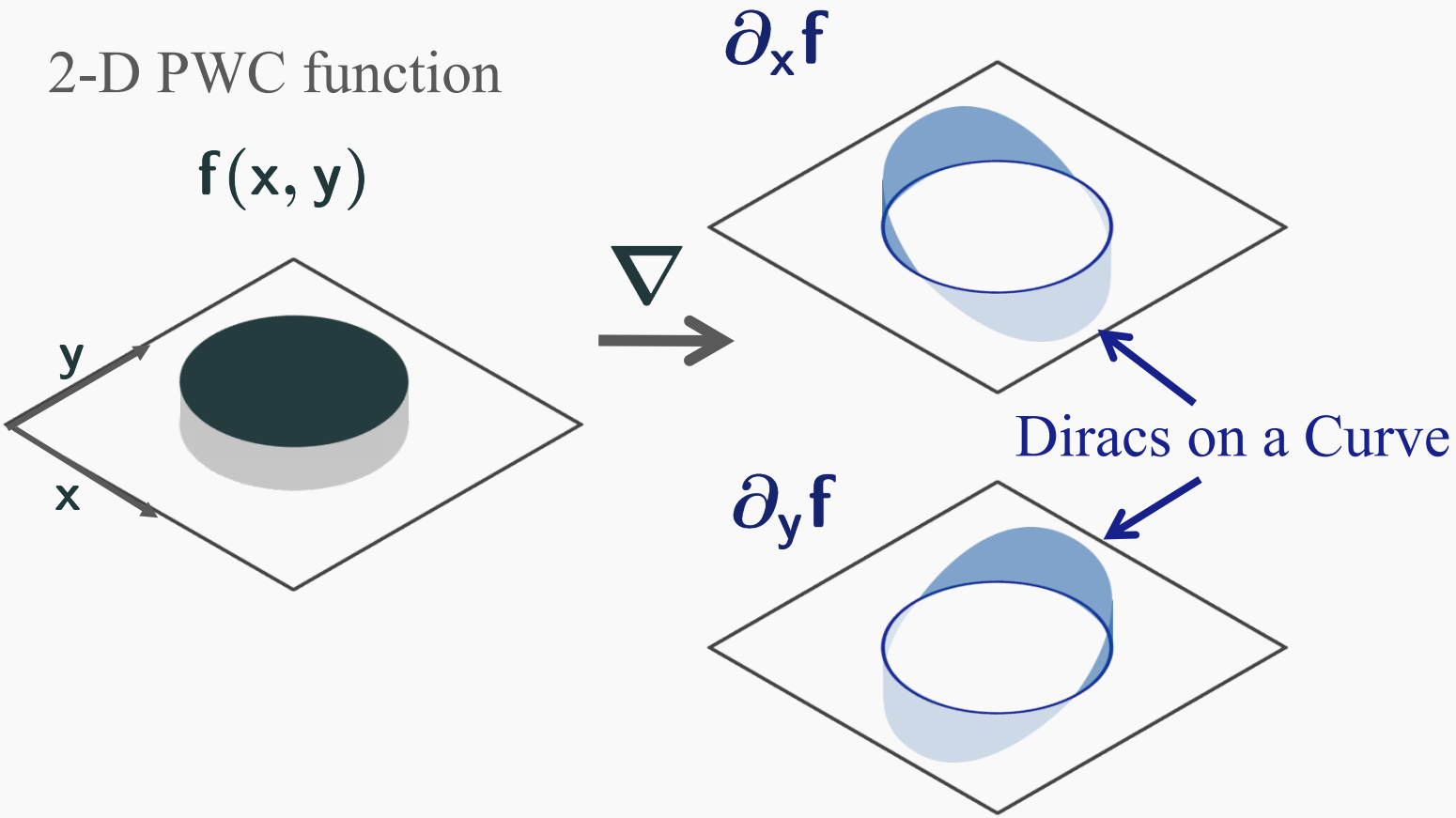
6. Biomedical applications

Extension to higher dims: Singularities not isolated

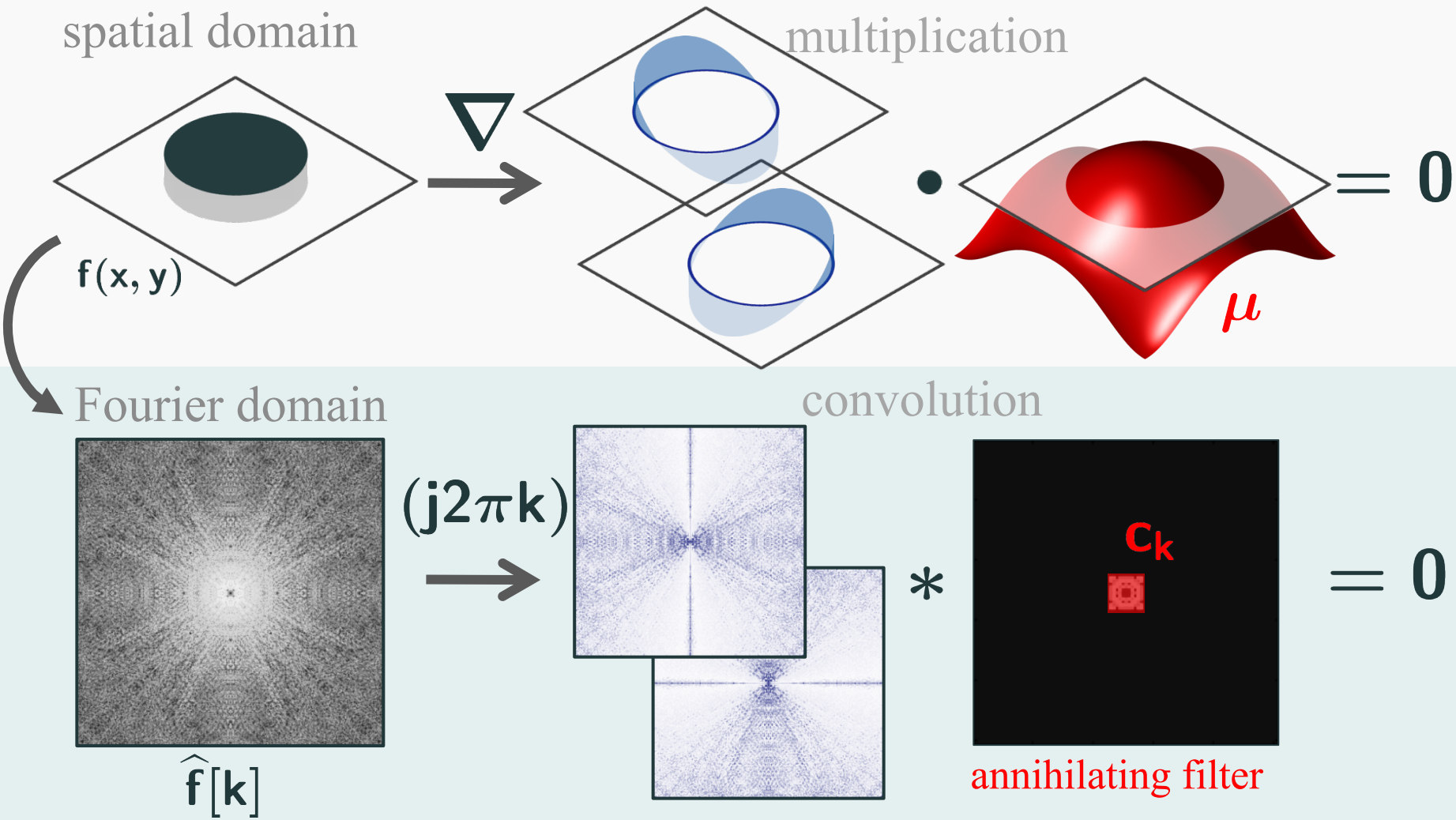
2-D PWC function



Extension to higher dims: Singularities not isolated



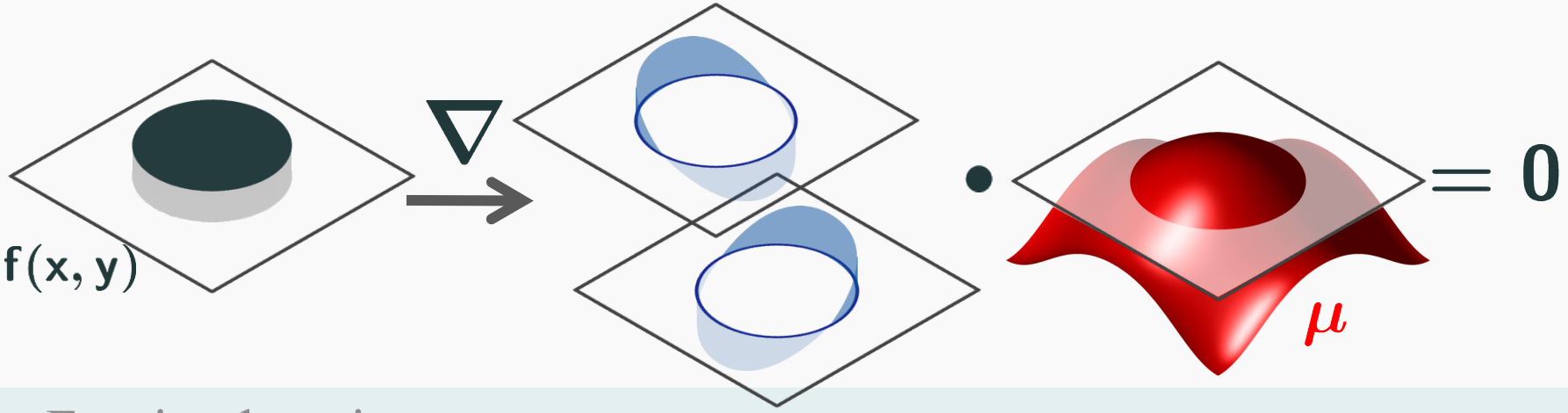
2-D PWC functions satisfy an annihilation relation



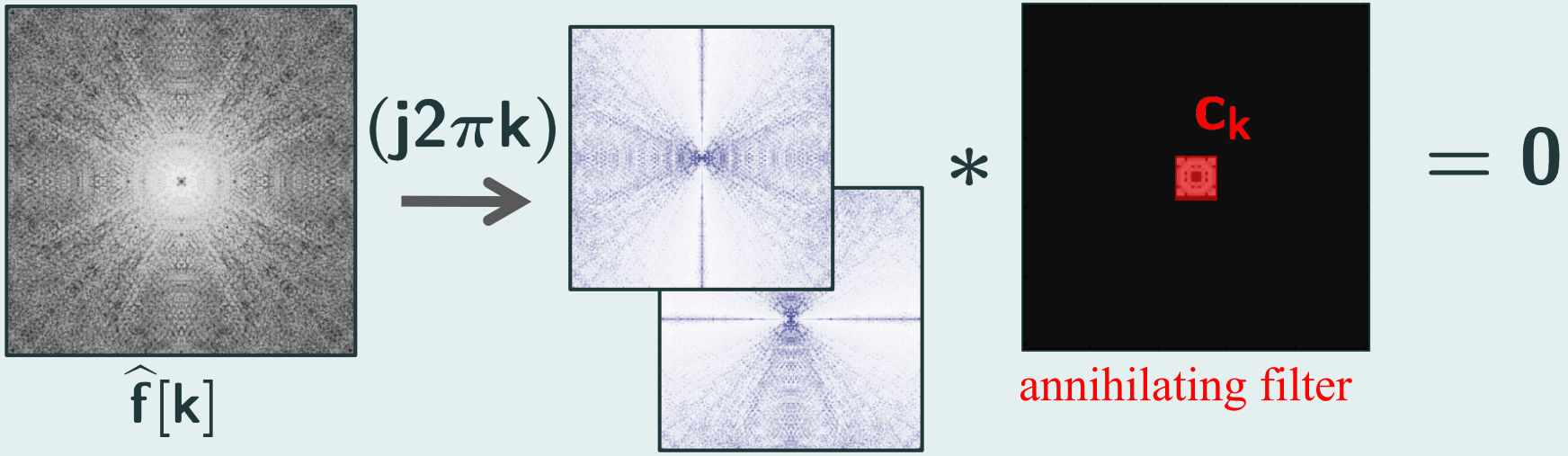
Annihilation relation: $\sum_k \widehat{\nabla f}[\ell - k]\, c_k = 0$

Image recovery

Stage 2: extrapolate given filter

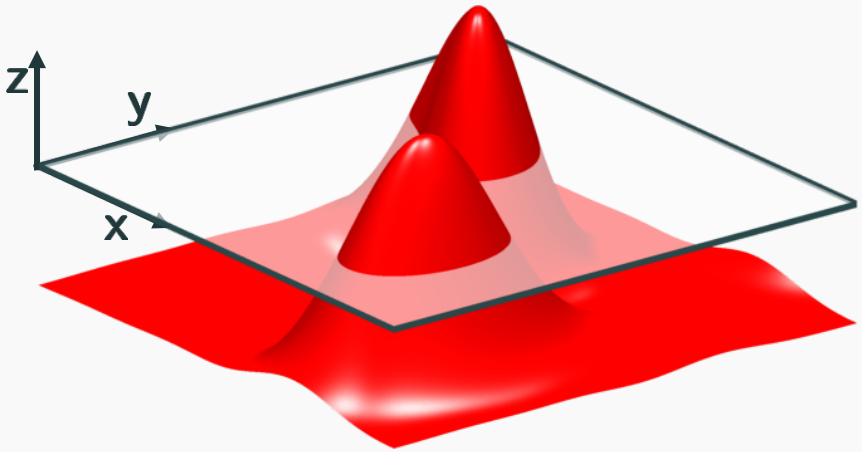


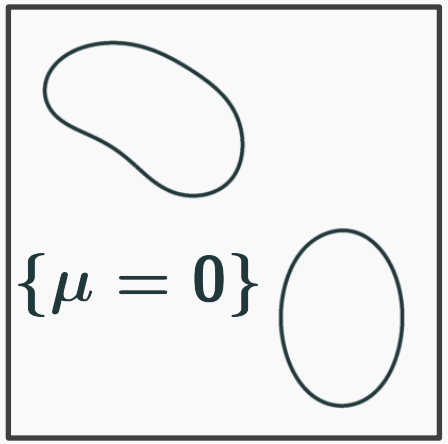
Fourier domain

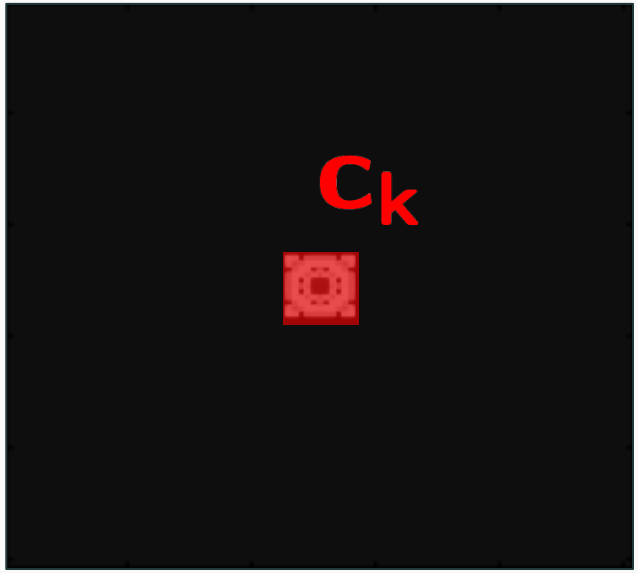


Stage 1: solve linear system for filter

Bandlimited curves





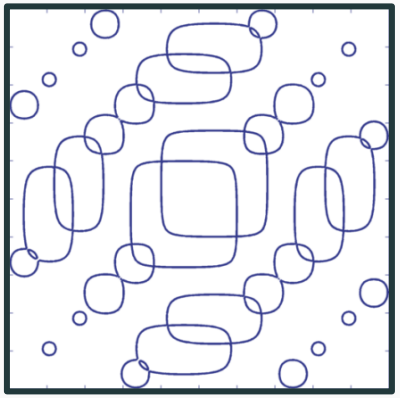


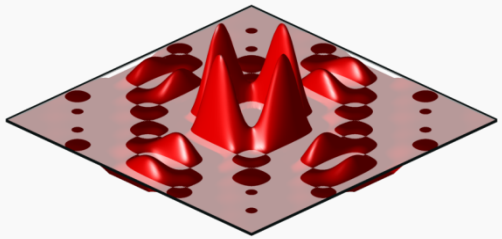
"FRI Curve"

Pan, Vetterli & Blu, IEEE Trans IP, 2015

Bandlimited curves can represent complex shapes

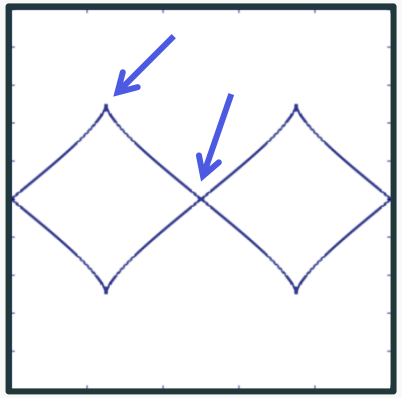
Multiple curves & intersections

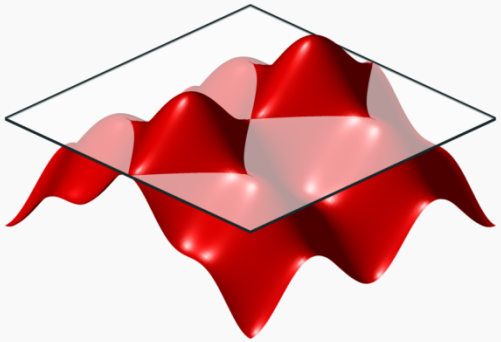




13x13 coefficients

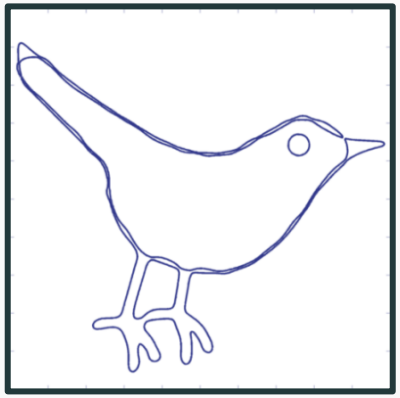
Non-smooth points

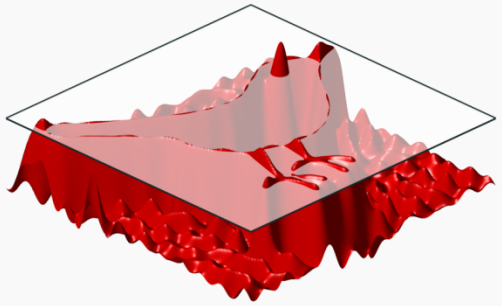




7x9 coefficients

Approximate arbitrary curves





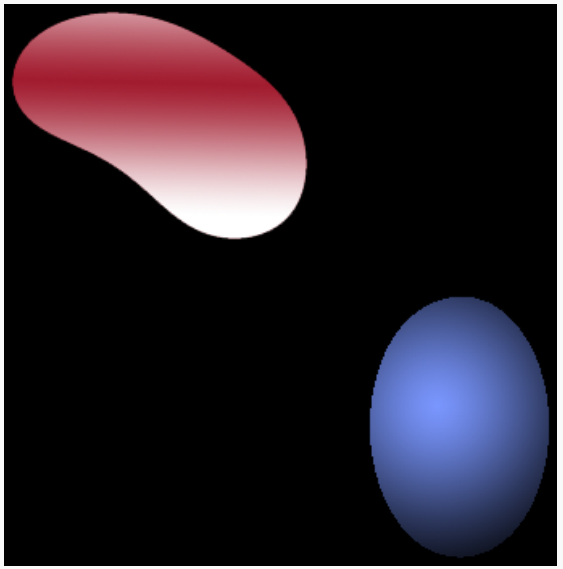
25x25 coefficients

Piecewise analytic model

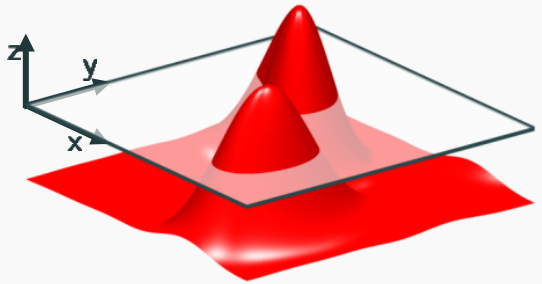
- Signal model: piecewise analytic signal

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

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



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



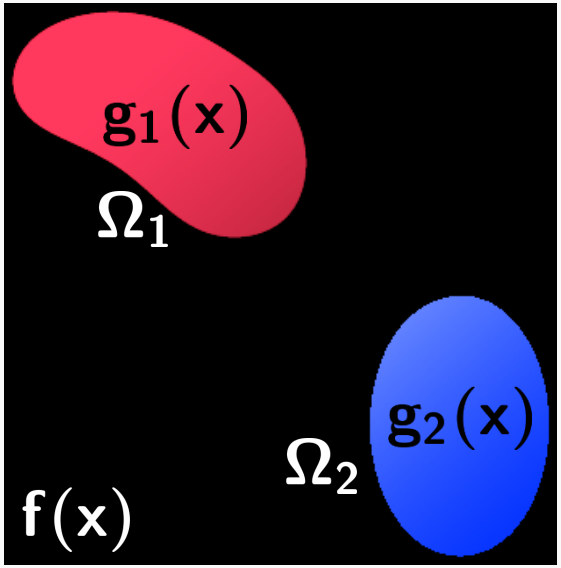
Pan, Vetterli & Blu, IEEE Trans IP, 2015

Piecewise smooth model

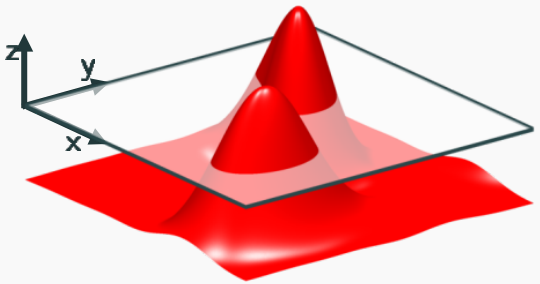
- New model: piecewise smooth signals

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

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



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



Ongie & Jacob, SIAM J Imag. Science, in press

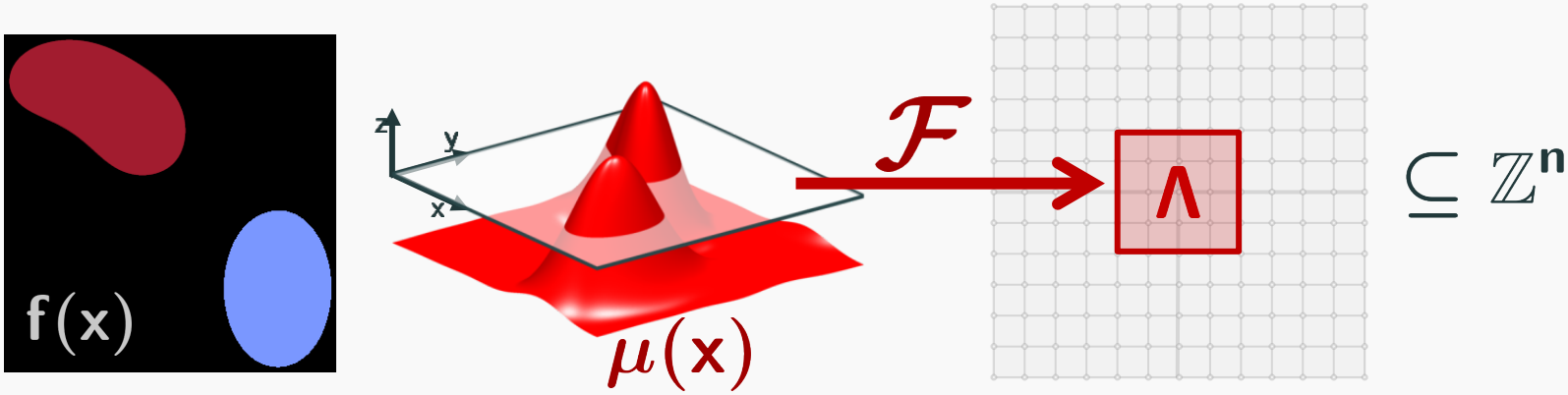
Eg. Piecewise constant functions: annihilation

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

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

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

any 1st order partial derivative

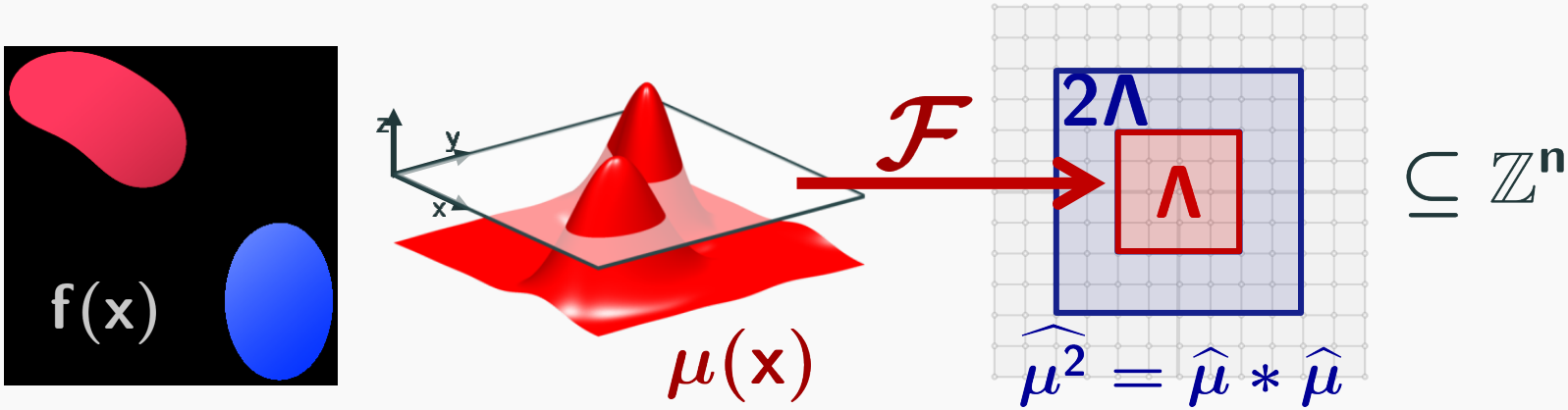


Eg. Piecewise linear functions: annihilation

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

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

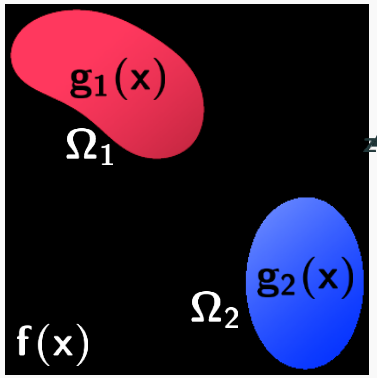
any 2nd order partial derivative

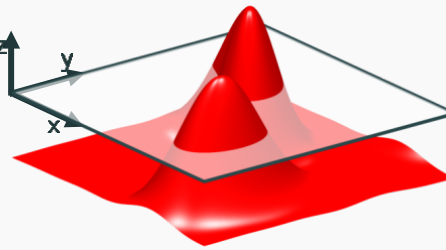


General signal models

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

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

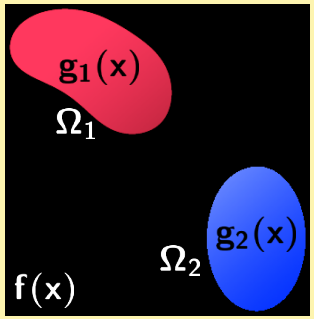


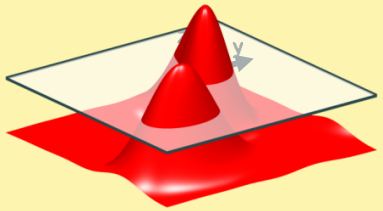


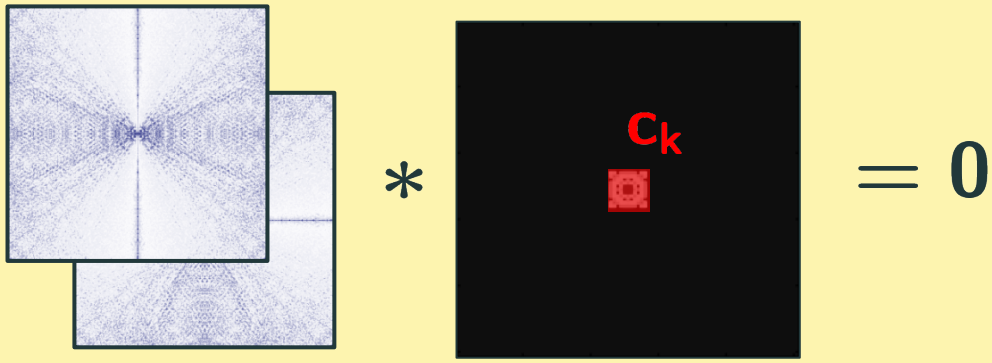
Signal Model:	Choice of Diff. Op.:	
PW Constant	$D = \nabla$	1st order
PW Analytic*	$D = \partial_x + j\partial_y$	
PW Harmonic	$D = \Delta$	2nd order
PW Linear	$D = \{\partial_{xx}, \partial_{xy}, \partial_{yy}\}$	
PW Polynomial	$D = \{\partial^{\alpha}\}_{\lvert\alpha\rvert=n}$	nth order

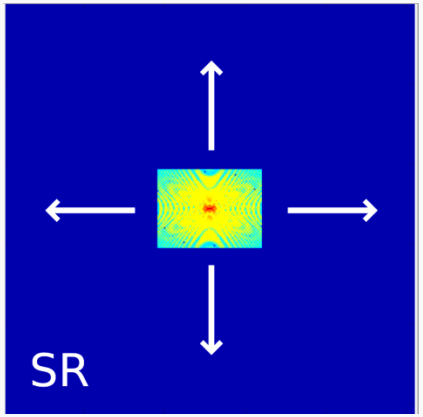
Can we extrapolate ?

Piecewise smooth signals satisfy annihilation conditions









How much should we sample to extrapolate?

Recovery guarantees: challenges

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.

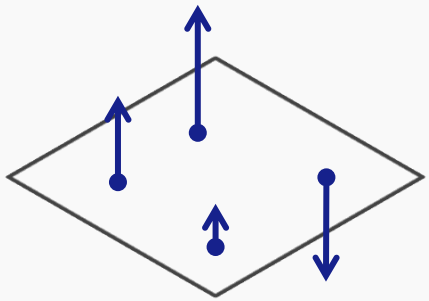
Proof (a la Prony's Method):

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

"Caratheodory Parametrization"

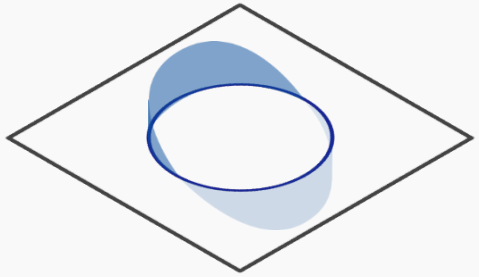
Recovery guarantees: challenges

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



$$\xrightarrow{\mathcal{F}} \quad \hat{f}[k] = \sum_i a_i e^{-j2\pi k \cdot x_i}$$

Not true when singularities supported on curves:



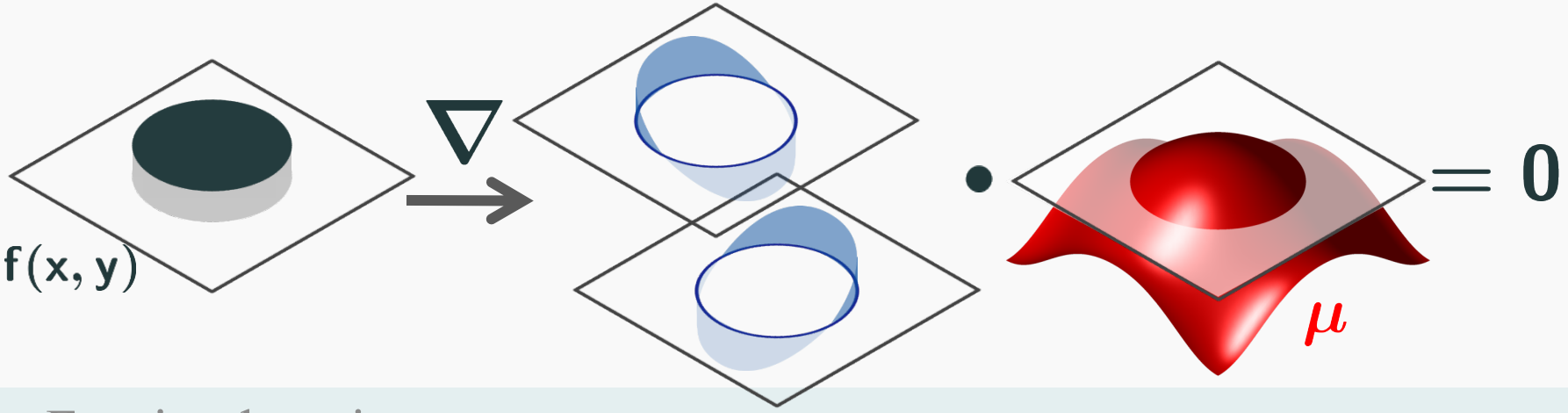
$$\xrightarrow{\mathcal{F}} \quad \widehat{\nabla f}[k] = \oint_{\partial\Omega} e^{-j2\pi k \cdot x} n \, ds$$

Requires new techniques:

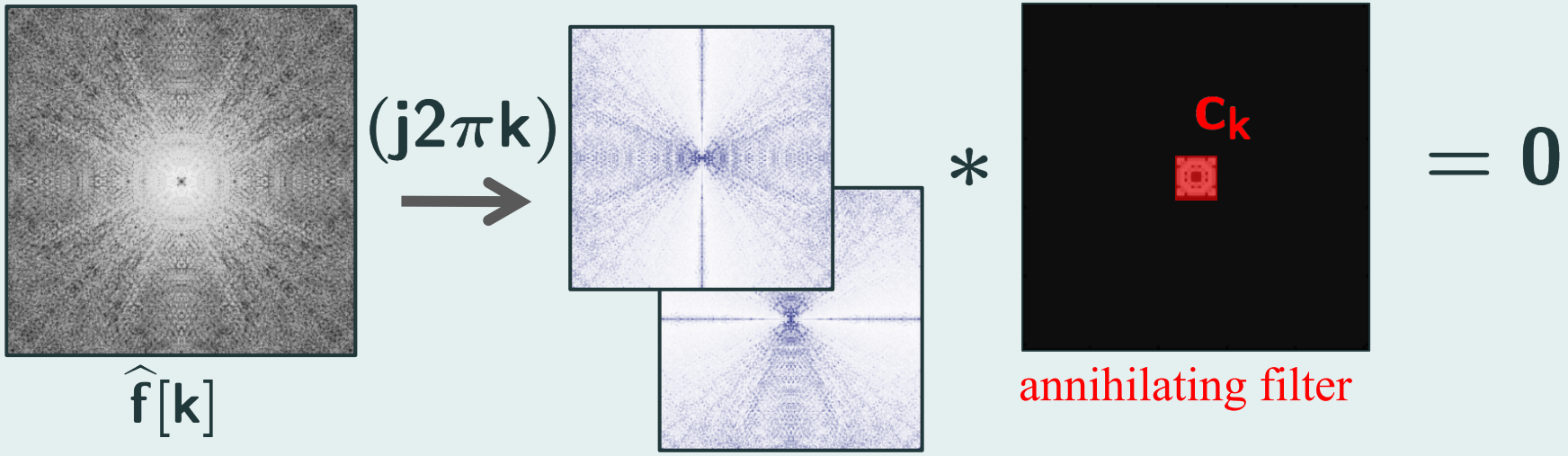
- Spatial domain interpretation of annihilation relation
- Algebraic geometry of trigonometric polynomials

Image recovery

Stage 2: extrapolate given filter



Fourier domain



Stage 1: solve linear system for filter

Step 1. When can you recover the filter ?

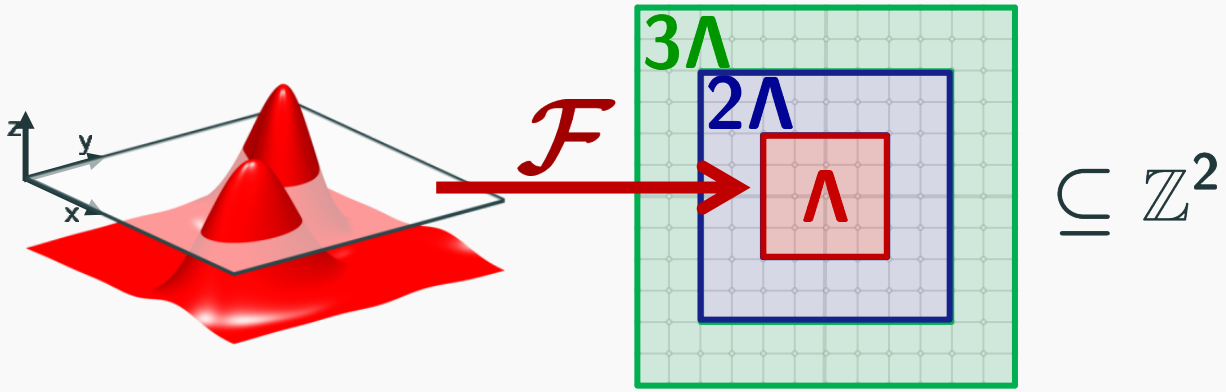
Theorem: If $\mathbf{f}$ is PWC* with edge set $\mathbf{E} = \{\mu = 0\}$

with μ minimal and bandlimited to $\boldsymbol{\Lambda}$ then

$\mathbf{c} = \widehat{\mu}$ is the unique solution to

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

*Some geometric restrictions apply



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

Ongie & Jacob, SIAM J Imag. Science, in press

Step 2. When can you recover the signal given the filter ?

Theorem: If $\mathbf{f}$ 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

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

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

Ongie & Jacob, SIAM J Imag. Science, in press

Step 2. When can you extrapolate given the filter ?

Theorem: If $\mathbf{f}$ 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

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

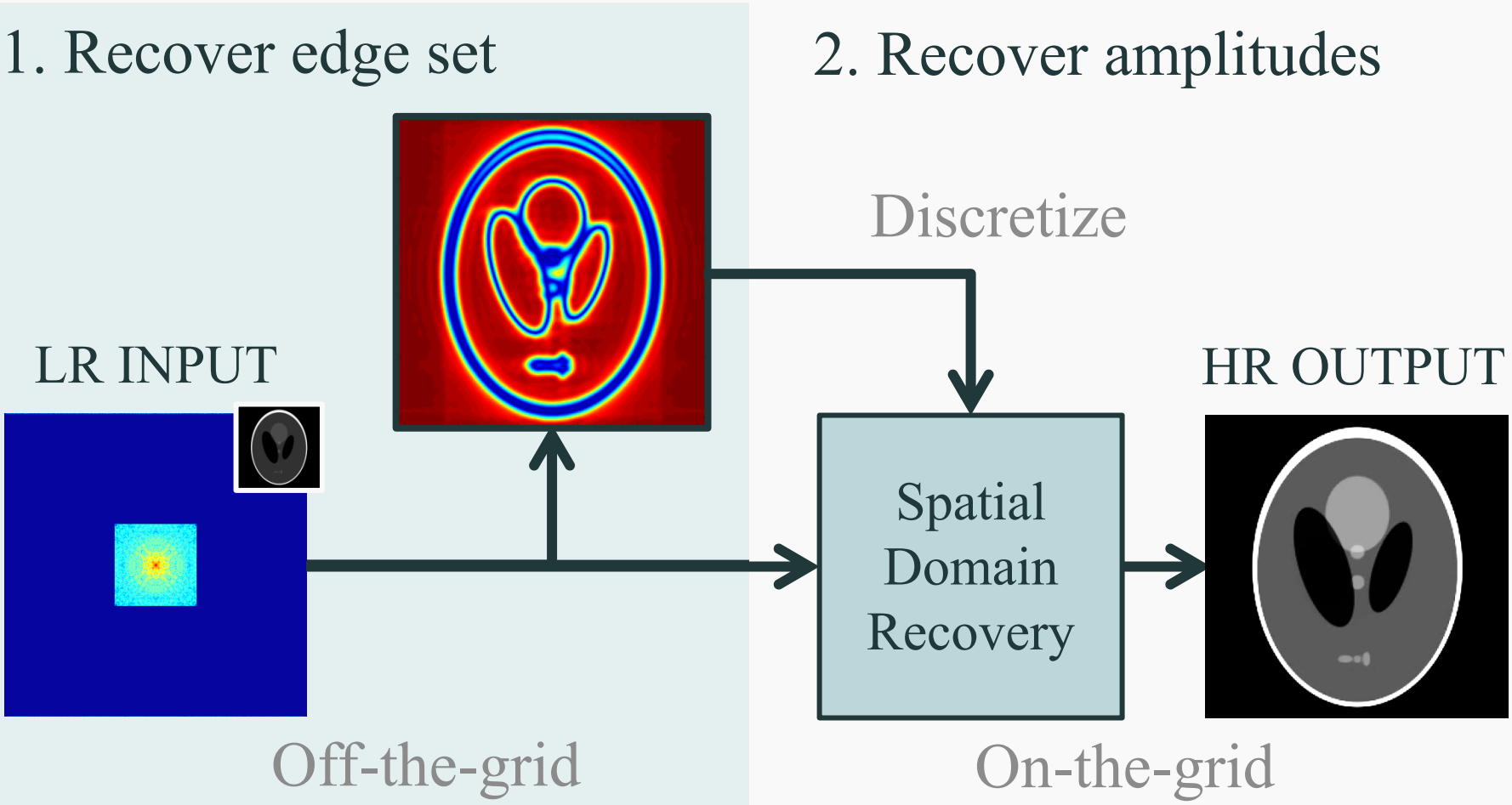
when the sampling set $\mathbf{\Gamma} \supseteq \mathbf{3\Lambda}$

Equivalently,

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

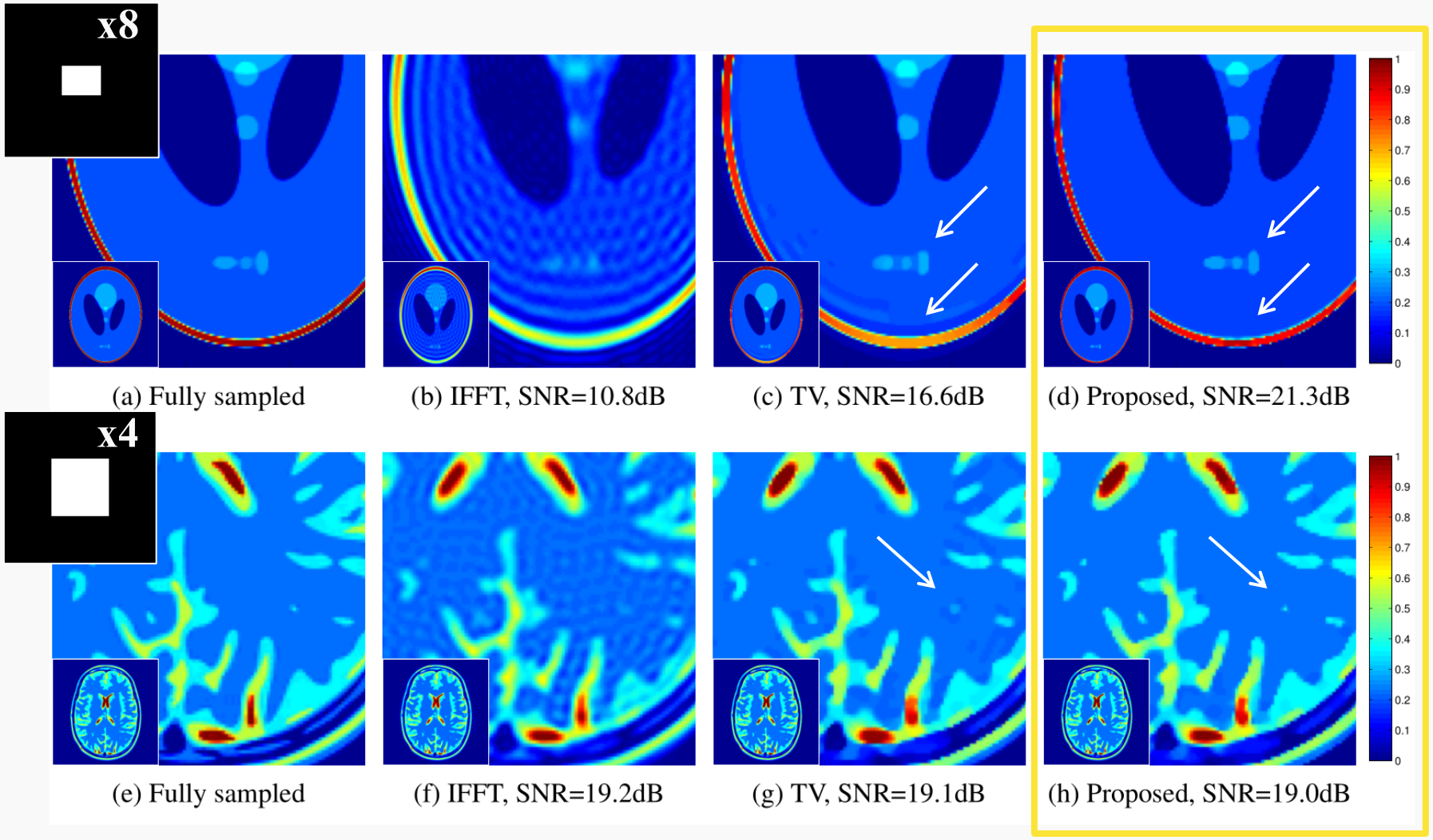
Ongie & Jacob, SIAM J Imag. Science, in press

Super-resolution image recovery



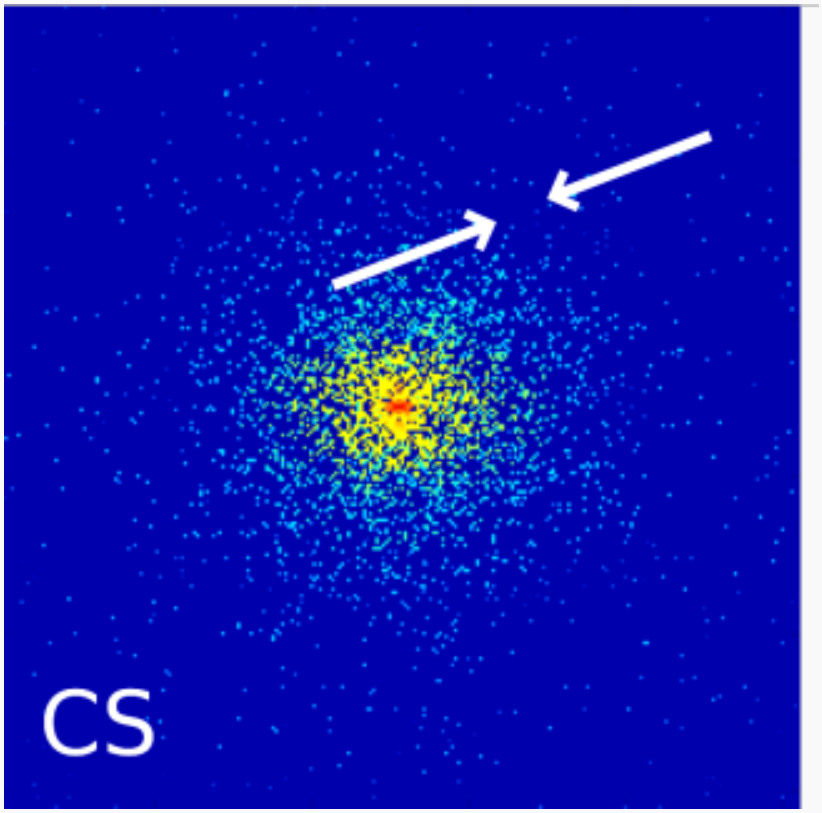
Ongie & Jacob, SIAM J Imag. Science, in press

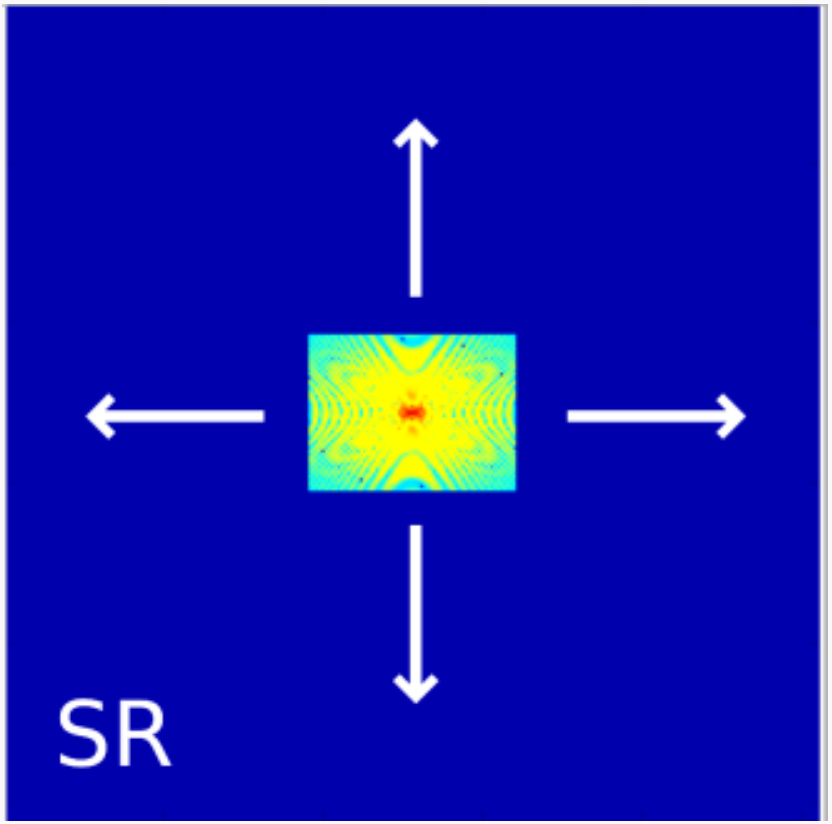
Super-resolution of MRI Medical Phantoms



Ongie & Jacob, SIAM J Imag. Science, in press

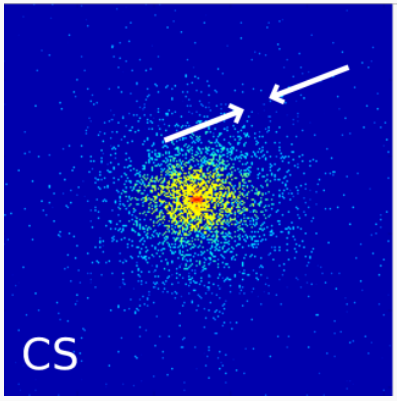
Can we generalize to non-uniform setting ??





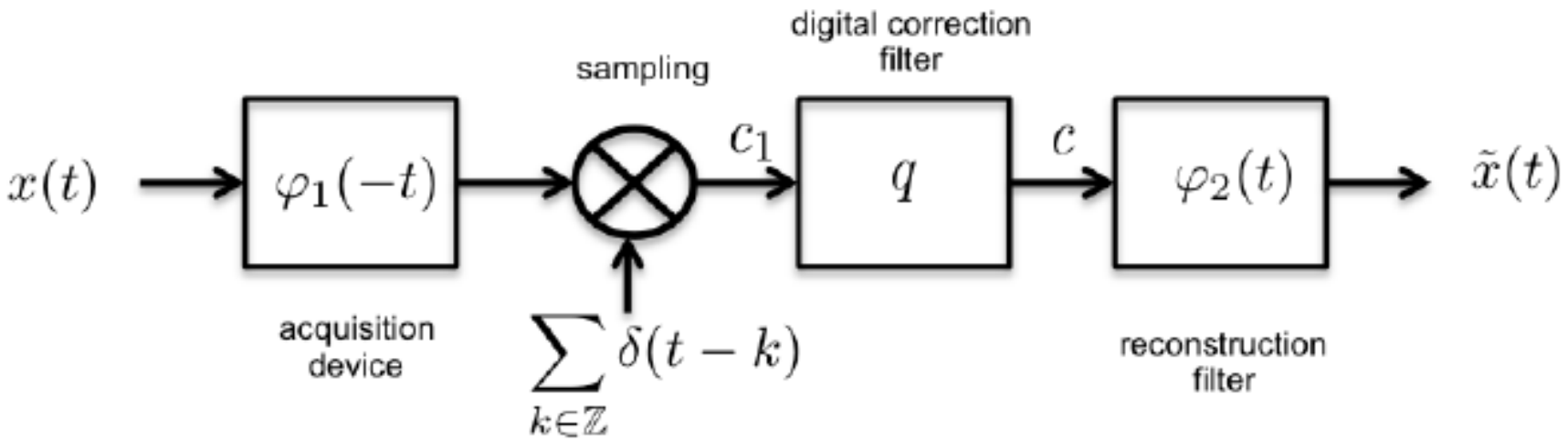
Improve recovery using non-uniform sampling

Overview

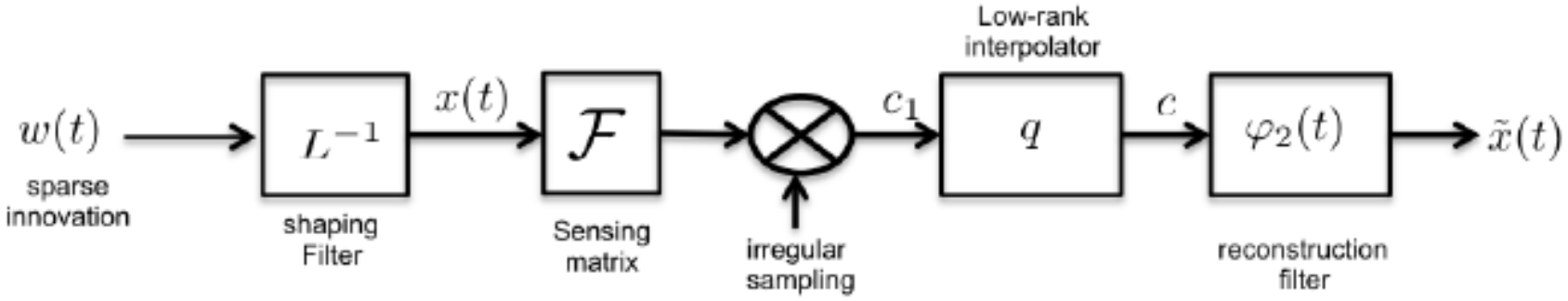


1. Introduction
2. Review of Compressive Sensing
3. FRI extrapolation from uniform samples
4. Structured low-rank interpolation for non-uniform samples
   - 1-D Theory
5. Fast implementations
6. Biomedical applications

Sampling vs low-rank interpolation



(a)

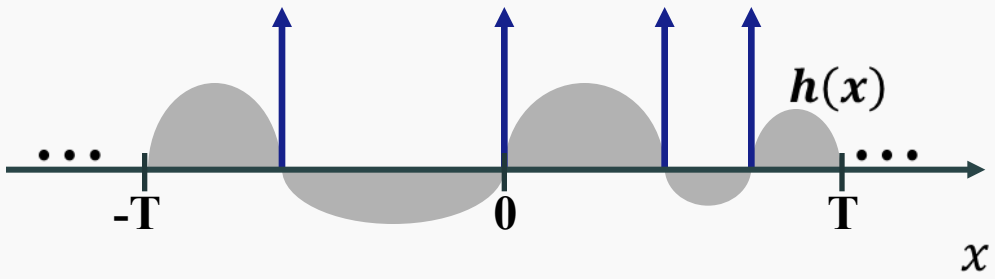


(b)

Key idea: annihilating filter

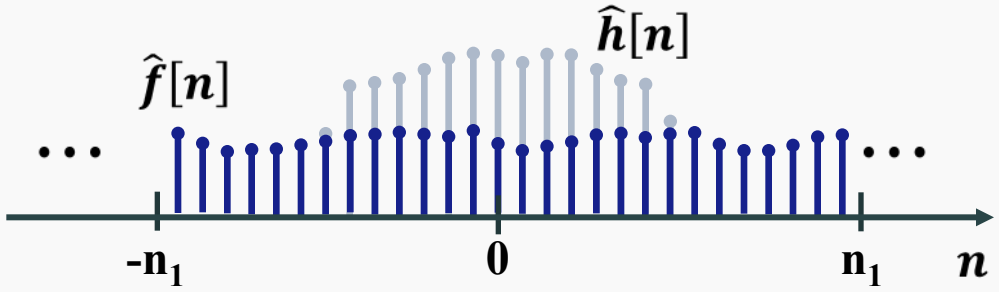
*** FRI Sampling theory**

$f(x)$: periodic stream of K Diracs



$h(x)$: annihilating function

$$f(x) \cdot h(x) = 0$$



$\hat{h}[n]$: annihilating filter

$$\hat{f}[n] * \hat{h}[n] = 0$$

Length of $\hat{h}[n] \geq$ k+1

Low rank Hankel matrix

** Jin KH et al. IEEE TCI (to appear)* ** Ye JC et al. IEEE TIT, 2016* ** Jin KH et al.,IEEE TIP, 2015*

*** ALOHA : Annihilating filter based LOw rank Hankel matrix Approach**

Finite length convolution

$$(\hat{h} * \hat{f})[n] = \sum_{I=0}^{\kappa-1} \hat{h}[I]\hat{f}[n-I] = 0$$

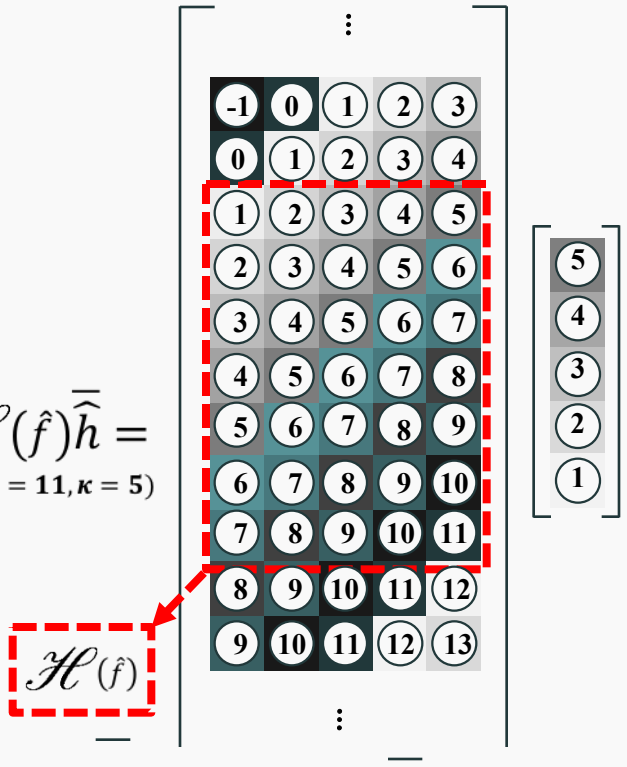
κ : *# of annihilating filter coef.*

Matrix Representation

$$\mathscr{C}(\hat{f})\bar{\hat{h}} = 0$$

$$\mathscr{C}(\hat{f})\bar{\hat{h}} = \mathscr{H}(\hat{f})$$

$(n_1 = 11, \kappa = 5)$



Sparsity in spatial domain ⇔ low rankness in k-space

Low-Rank Hankel matrix minimization

$$\text{Rank}\mathcal{H}(\hat{\boldsymbol{f}}) = k$$

* Jin KH et al IEEE TCI, 2016
* Jin KH et al.,IEEE TIP, 2015
* Ye JC et al., IEEE TIT, 2016

Missing elements can be found by low rank Hankel structured matrix completion

$$\min_{\mathbf{m}} \quad \|\mathcal{H}(\mathbf{m})\|_*$$

$$\text{subject to} \quad P_\Omega(\mathbf{m}) = P_\Omega(\hat{f})$$

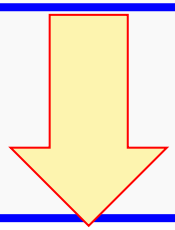
$\|\cdot\|_*$ Nuclear norm P_Ω Projection on sampling positions

General TV Signals

$$\mathrm{L}f(x) = \sum_{j=0}^{k-1} a_j \delta(x - x_j), \quad x_j \in [0, \tau].$$

$$\mathrm{L} := a_K \mathrm{D}^K + a_{K-1}\mathrm{D}^{K-1} + \ldots + a_1 \mathrm{D} + a_0$$

Piecewise smooth
Splines, polynomials



Weighted Fourier data

$$\mathcal{F}\{\mathrm{L}f(x)\} = \hat{l}(\omega)\hat{f}(\omega) = \sum_{j=0}^{k-1} a_j e^{-i\omega x_j}$$

$$l(\omega) = a_K (i\omega)^K + a_{K-1}(i\omega)^{K-1} + \ldots + a_1 (i\omega) + a_0$$

Existence of Annihilating Filter

Annihilating filter for weighted Fourier data

$$\hat{h}(\omega) * \left(\hat{l}(\omega) \hat{f}(\omega) \right) = 0$$

General Low-Rank Hankel Matrix Completion

$$(P) \quad \min_{\mathbf{m} \in \mathbb{C}^n} \quad \text{RANK} \mathscr{H}(\mathbf{m})$$

$$\text{subject to} \quad P_\Omega(\mathbf{m}) = P_\Omega(\hat{\mathbf{l}} \odot \hat{\mathbf{f}}) \ ,$$

Extension to general signal models

Stream of Diracs

$$x(t) = \sum_{l\in\mathbb{Z}} \sum_{i=0}^{r-1} c_i \delta(t - t_i - l\tau) \ ,$$

$$rank = r$$

$$\hat{h}(z) = \sum_{l=0}^{r} \hat{h}[l] z^{-l} = \prod_{j=0}^{r-1} (1 - e^{-i2\pi t_j/\tau} z^{-1})$$

Stream of differentiated Diracs

$$x(t) = \sum_{l\in\mathbb{Z}} \sum_{j=0}^{d_j} \ldots$$

$$rank = \sum_j d_j$$

With a proper weighting, the Hankel matrix of the weighted k-space data → low ranked.

Non-uniform spline

$$\mathrm{L}x = \sum_{j=0}^{r-1} c_j \delta(t - t_j)$$

$$\mathrm{L} := a_K \partial^K + a_{K-1} \partial^{K-1} + \ldots + a_1 \partial + a_0$$

$$rank = r$$

$$\hat{h}(z) = \sum_{l=0}^{r} \hat{h}[l] z^{-l} = \prod_{j=0}^{r-1} (1 - e^{-i2\pi t_j/\tau} z^{-1})$$

Piecewise smooth polynomial

$$x^{(q+1)}(t) = \sum_{l\in\mathbb{Z}} \sum_{j=0}^{q} c_{l_j} \delta^{(j)}(t - t_l)$$

$$rank = rq$$

$$\hat{h}(z) = \prod_{j=0}^{r-1} (1 - u_j z^{-1})^q$$

* Ye JC et al.,IEEE TIT 2016

Performance Guarantees

Exact Recovery

$$\min_{\mathbf{m}} \|\mathcal{H}(\mathbf{m})\|_*$$

$$\text{subject to} \quad P_\Omega(\mathbf{m}) = P_\Omega(\hat{\mathbf{f}})$$

$$m \geq c_1 \mu c_s k \log^\alpha n$$

$$\alpha = \begin{cases} 2, & \text{on grid} \\ 4, & \text{off grid} \end{cases}$$

Stable Recovery

$$\min_{\mathbf{m}} \|\mathcal{H}(\mathbf{m})\|_*$$

$$\text{subject to} \quad \|P_\Omega(\mathbf{m}) - P_\Omega(\hat{\mathbf{f}})\| \leq \delta$$

$$\|\mathcal{H}(\mathbf{m}) - \mathcal{H}(\hat{\mathbf{f}})\|_F \leq c_2 n^2 \delta$$

* Ye JC et al.,IEEE TIT 2016

Mutual Coherence for FRI

$$\mu \leq \max\left\{\frac{\zeta_{n-d+1}}{\sigma_{\min}\left(\mathcal{V}^*_{n-d+1}\mathcal{V}_{n-d+1}\right)}, \frac{\zeta_d}{\sigma_{\min}\left(\mathcal{V}^*_d\mathcal{V}_d\right)}\right\}$$

Confluent Vandermonde matrix

$$\mathscr{H}(\hat{\mathbf{x}}) = \mathcal{V}_{n-d+1}\,\mathcal{B}\,\mathcal{V}_d^T,$$

$$\zeta_N = N\left[\frac{(N-1)!}{(N-l_{\max})!}\right]^2$$

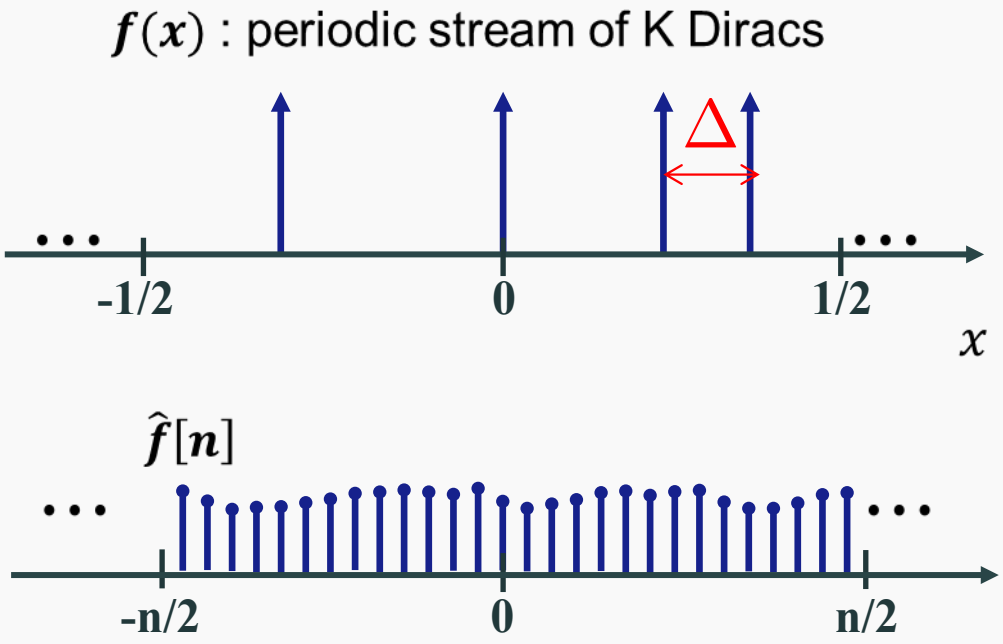
Multiplicity of roots

$$\mu \leq \frac{n/2}{n/2 - 1/\Delta - 1}$$

Using extreme function for bounding singular value See Moitra (2015)

** Ye JC et al.,IEEE TIT 2016*

Relation to Super-resolution: Minimum separation



$$\Delta > \frac{2}{n}$$

Same as
Candes et al (2013)
Tang et al (2015)

$$\mu \leq \frac{n/2}{n/2 - 1/\Delta - 1}$$

Using extreme function
for bounding singular value
See Moitra (2015)

** Ye JC et al.,IEEE TIT 2016*

Link to discrete domain CS

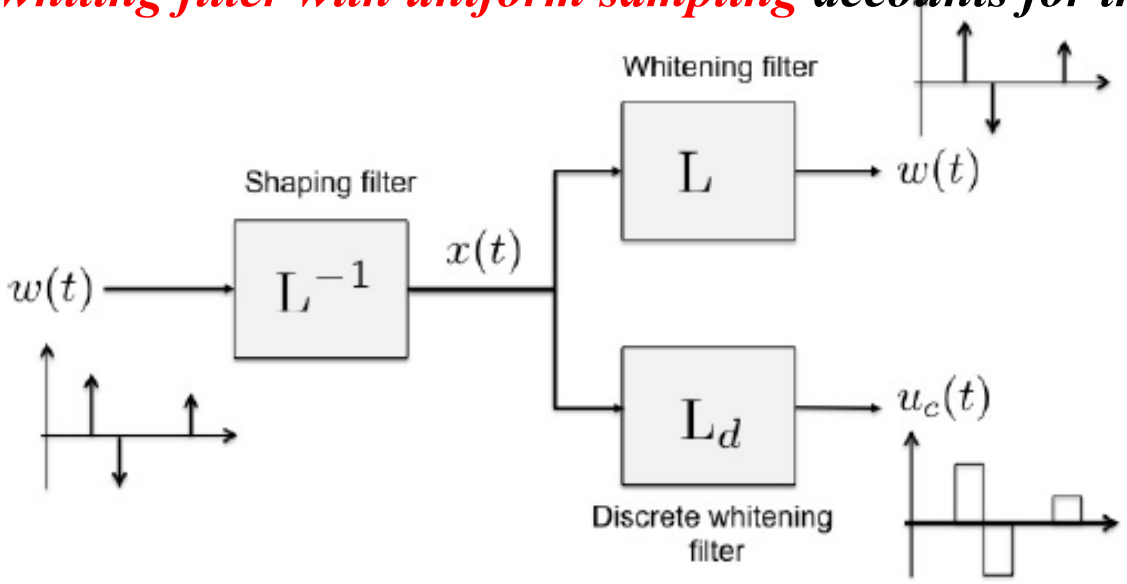
** Ye JC et al.,IEEE TIT 2016*

On grid model using cardinal setup

- ***Unknown singularities are located on integer grid***

$$\mathrm{L}x(t) = \sum_{l \in \mathbb{Z}} a[l]\delta(t-l)$$

- ***Discrete whiting filter with uniform sampling accounts for the sparsity***



Off-Grid vs On-Grid : Hankel

* Ye JC et al.,IEEE TIT 2016

Hankel Matrix: off-grid

$$\begin{bmatrix} \hat{y}[0] & \hat{y}[1] & \cdots & \hat{y}[d-1] \\ \hat{y}[1] & \hat{y}[2] & \cdots & \hat{y}[d] \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}[n-d] & \hat{y}[n-d+1] & \cdots & \hat{y}[n-1] \end{bmatrix}$$

Wrap-around Hankel Matrix: on-grid

$$\left[\begin{array}{cccc} \hat{w}_0[0] & \hat{w}_0[1] & \cdots & \hat{w}_0[d-1] \\ \hat{w}_0[1] & \hat{w}_0[2] & \cdots & \hat{w}_0[d] \\ \vdots & \vdots & \ddots & \vdots \\ \hat{w}_0[n-d] & \hat{w}_0[n-d+1] & \cdots & \hat{w}_0[n-1] \\ \hline \hat{w}_0[n-d+1] & \hat{w}_0[n-d+2] & \cdots & \hat{w}_0[0] \\ \vdots & \vdots & \ddots & \vdots \\ \hat{w}_0[n-1] & \hat{w}_0[0] & \cdots & \hat{w}_0[d-2] \end{array}\right]$$

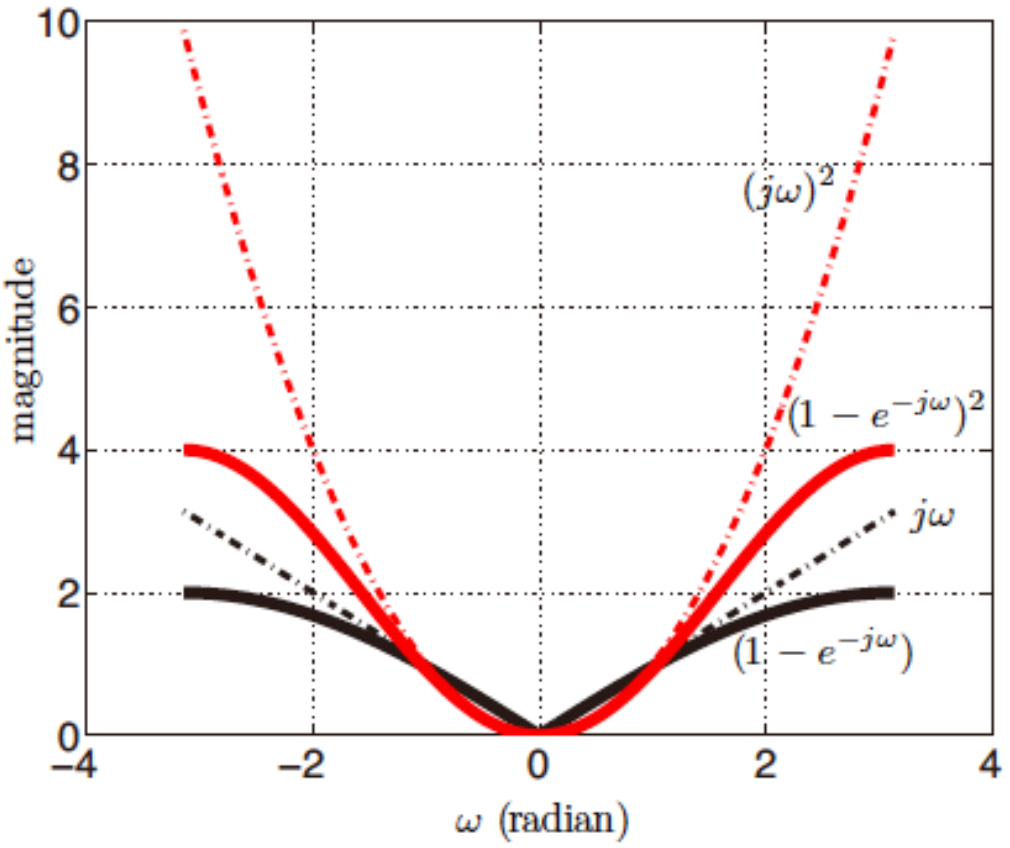
Periodic repetition

$$m \geq c_1 \mu c_s k \log^{\alpha} n$$

$$\alpha = \begin{cases} 2, & \text{on grid} \\ 4, & \text{off grid} \end{cases}$$

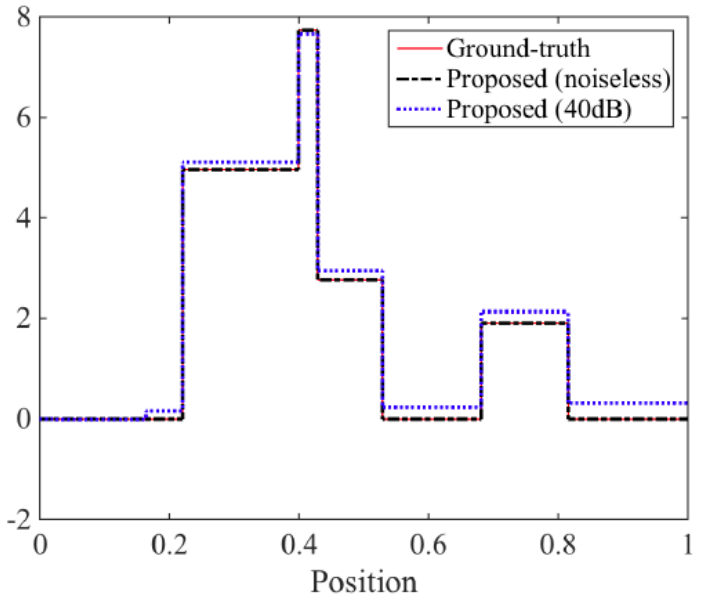
Off-grid vs On-grid: weighting

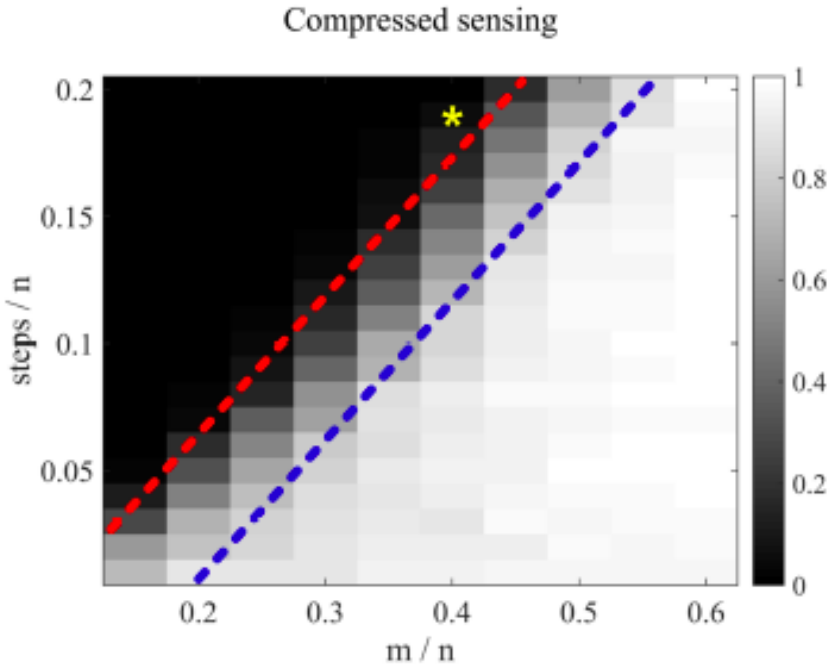
Regularized Weighting → more stable

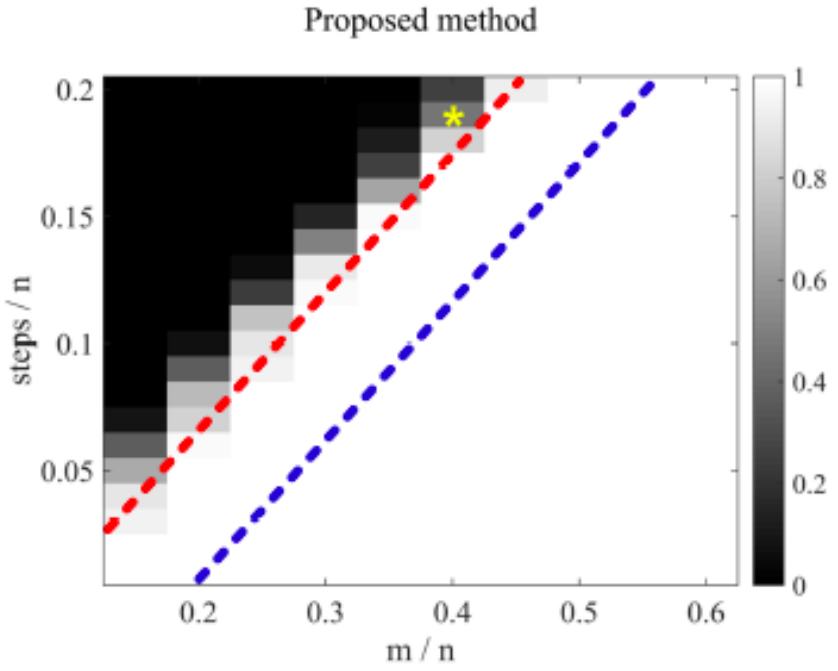


** Ye JC et al.,IEEE TIT 2016*

Phase transition:
piecewise constant signals





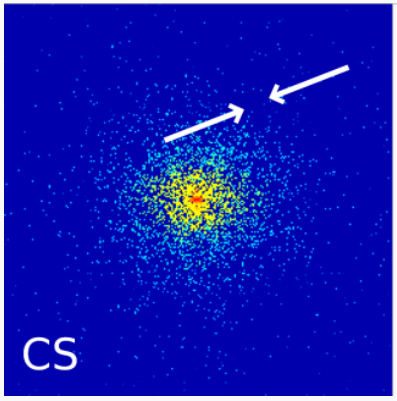


Overview

1. Introduction

2. Review of Compressive Sensing

3. FRI extrapolation from uniform samples



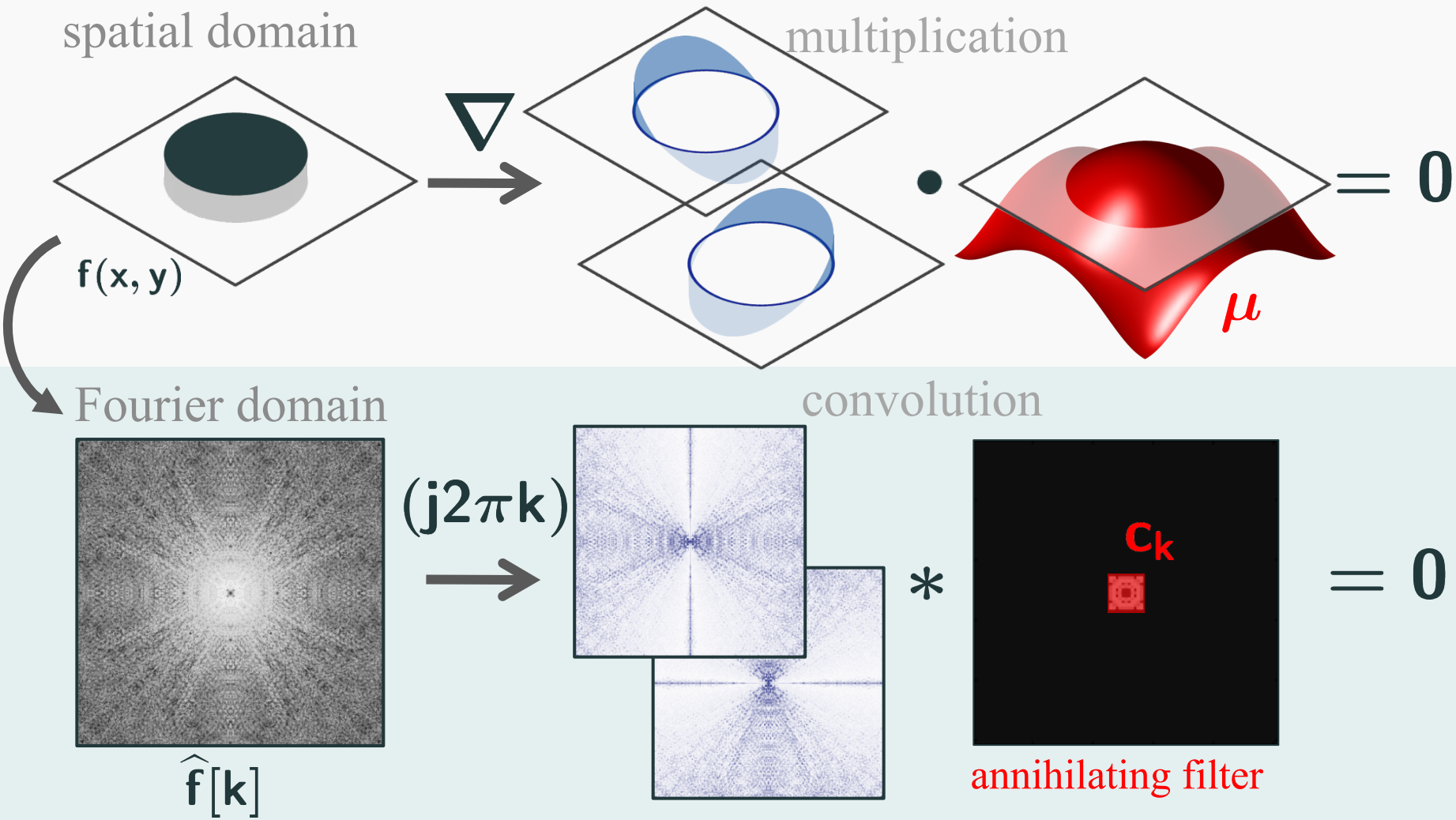
4. Structured low-rank interpolation for non-uniform samples

- 2-D Theory

5. Fast implementations

6. Biomedical applications

2-D PWC functions satisfy an annihilation relation



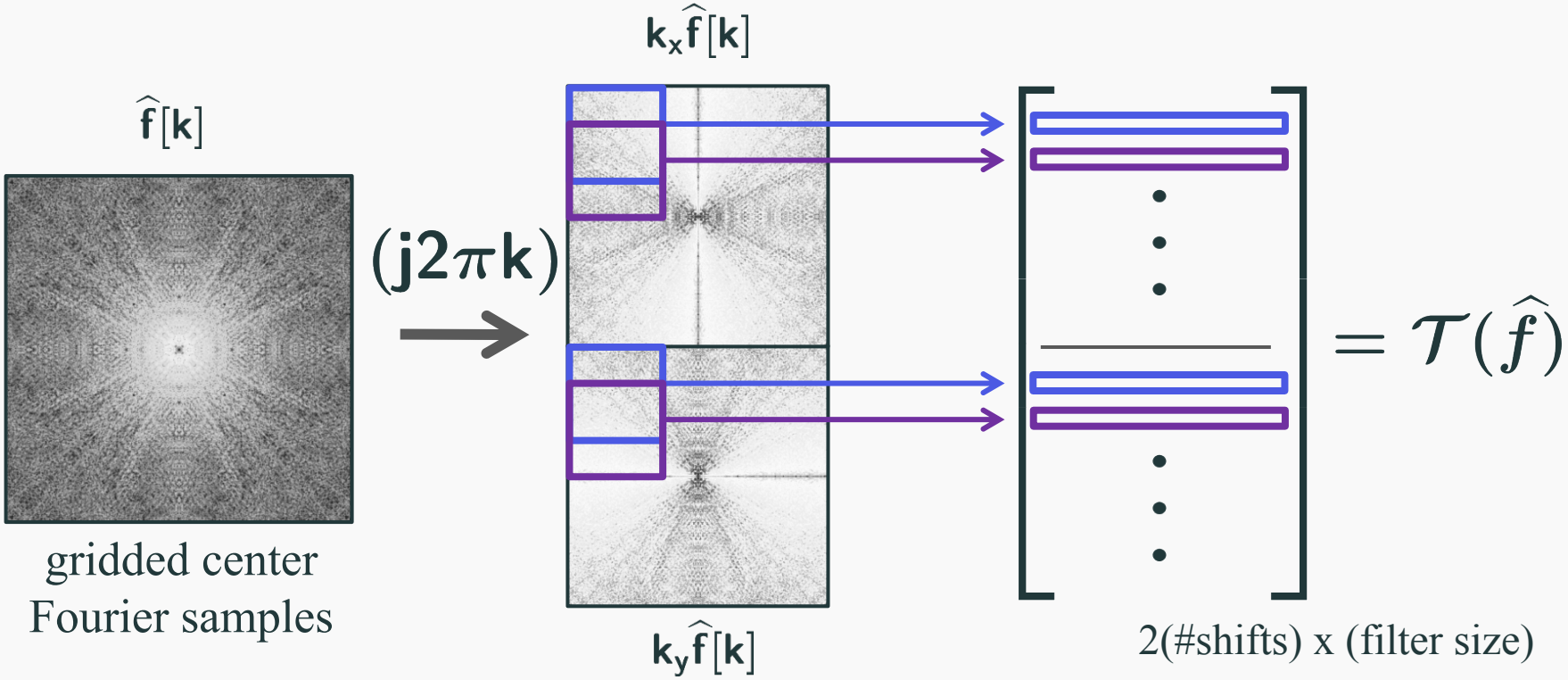
Annihilation relation: $\sum_k \widehat{\nabla f}[\ell - k]\, c_k = 0$

Matrix representation of annihilation

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

2-D convolution matrix
(block Toeplitz)

vector of filter coefficients



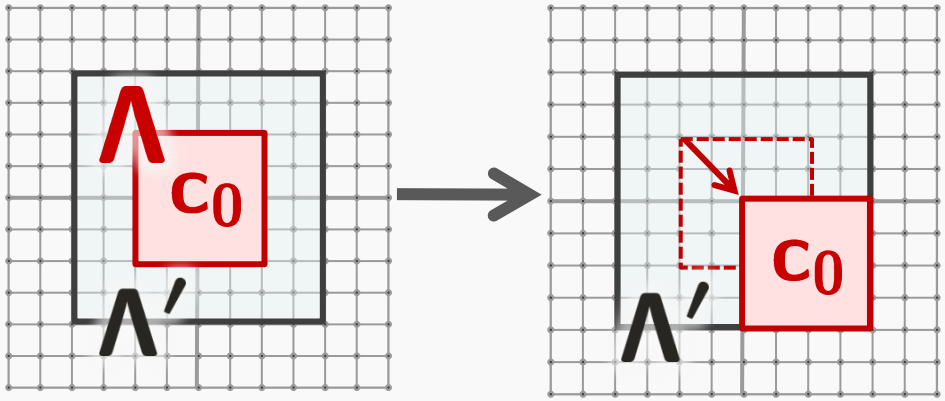
Basis of algorithms: Annihilation matrix is low-rank

Prop: If the level-set function is bandlimited to Λ

and the assumed filter support $\Lambda' \supset \Lambda$ then

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

Fourier domain



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

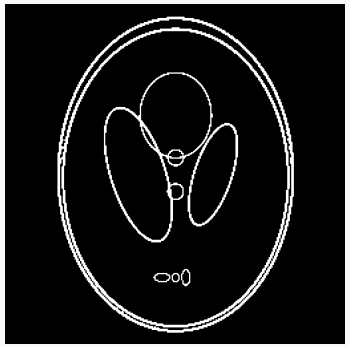
Basis of algorithms: Annihilation matrix is low-rank

Prop: If the level-set function is bandlimited to $\boldsymbol{\Lambda}$

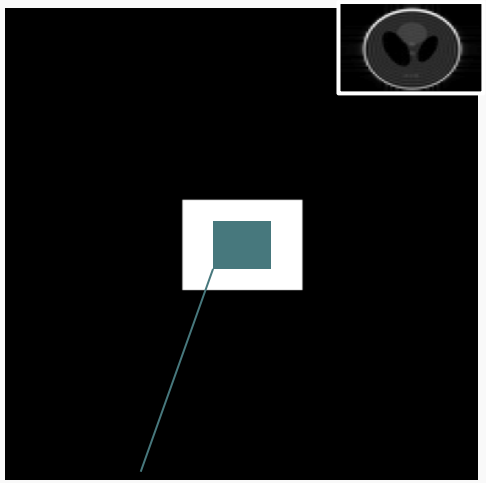
and the assumed filter support $\boldsymbol{\Lambda}' \supset \boldsymbol{\Lambda}$ then

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

Example:
Shepp-Logan



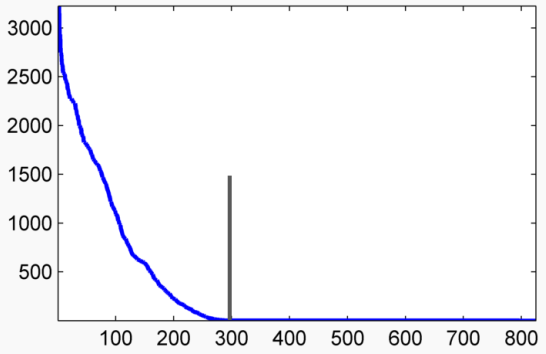
Fourier domain



Assumed filter: 33x25

Samples: 65x49

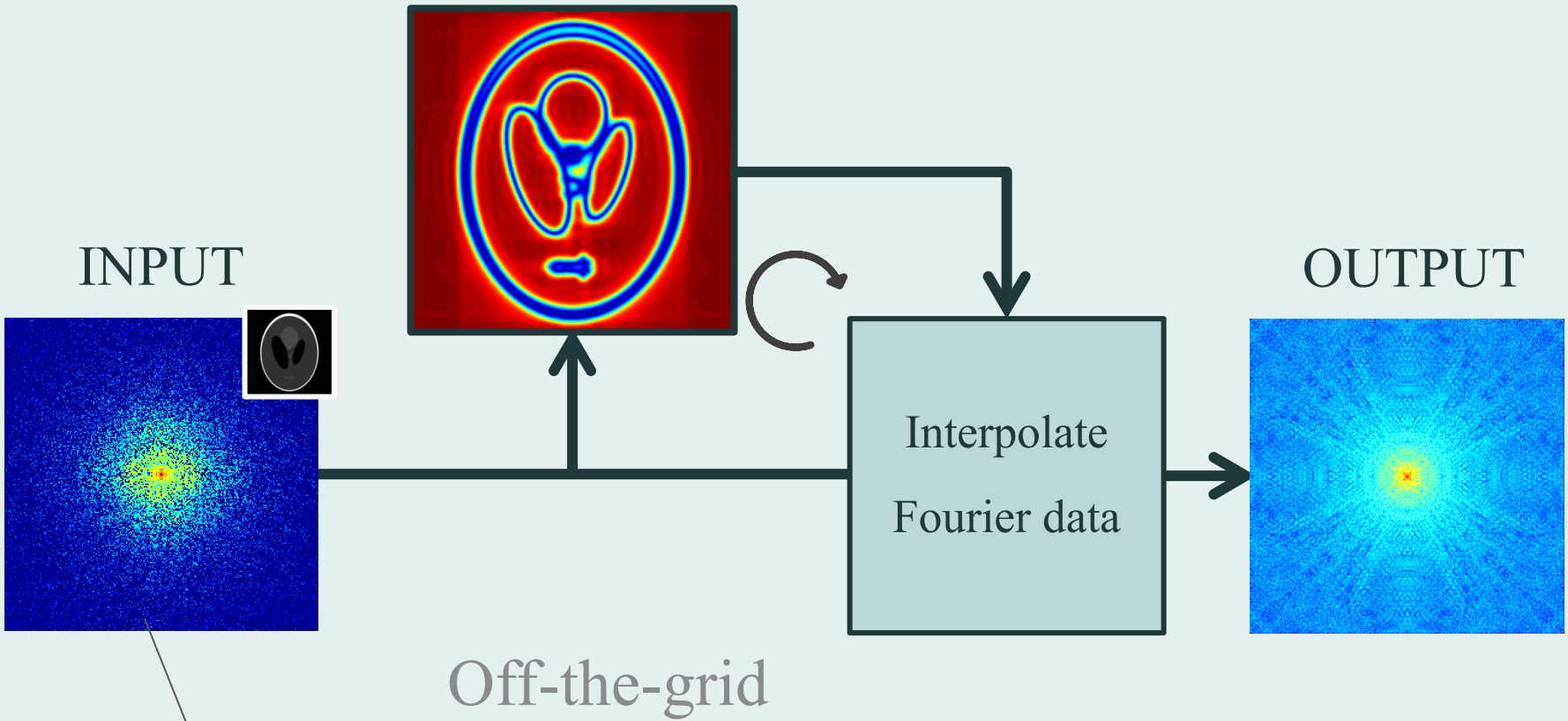
$\sigma(\mathcal{T}(\widehat{f}))$



Rank ≈ 300

One Step Algorithm

Jointly estimate edge set and amplitudes



Accommodate random samples

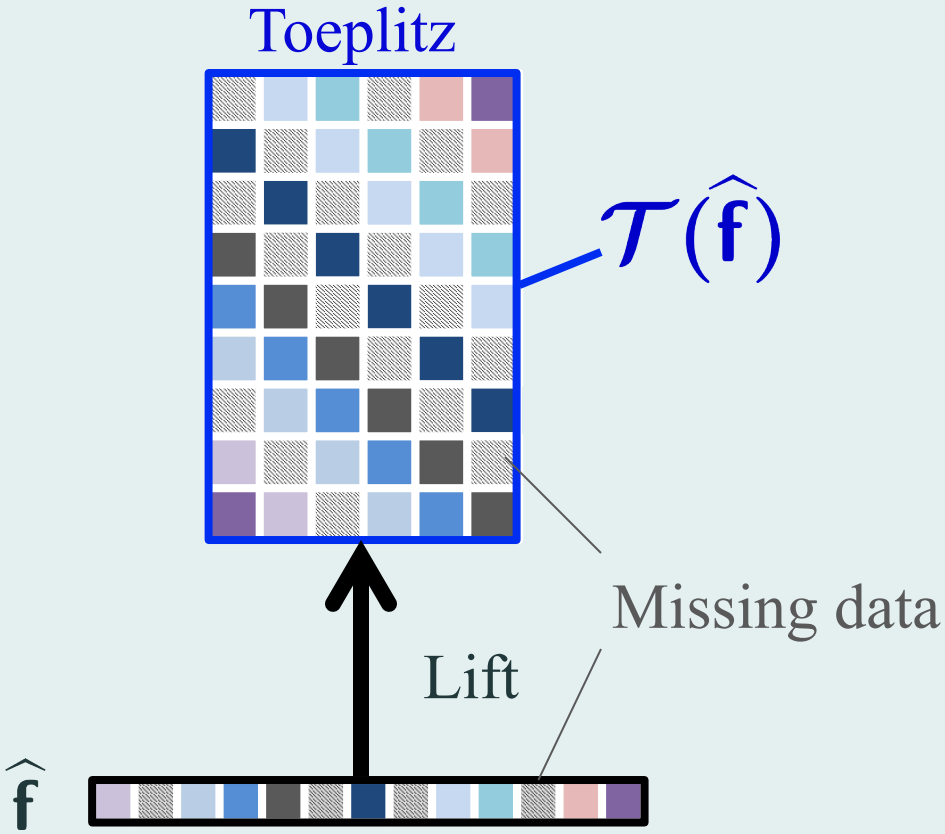
Recovery as a structured low-rank matrix completion

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

Recovery as a structured low-rank matrix completion

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

1-D Example:

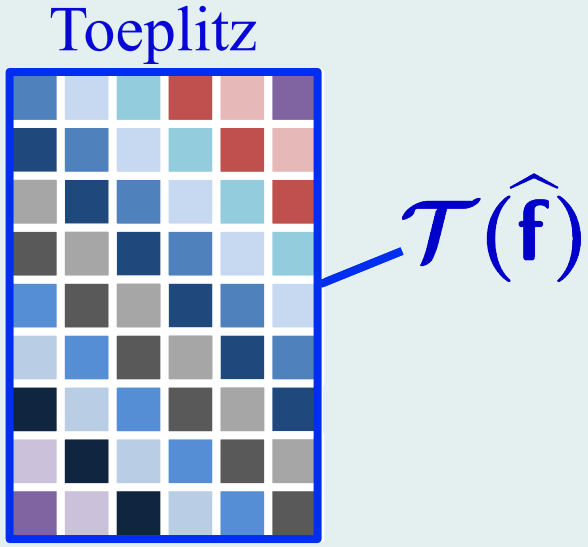


Recovery as a structured low-rank matrix completion

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

1-D Example:

Complete matrix

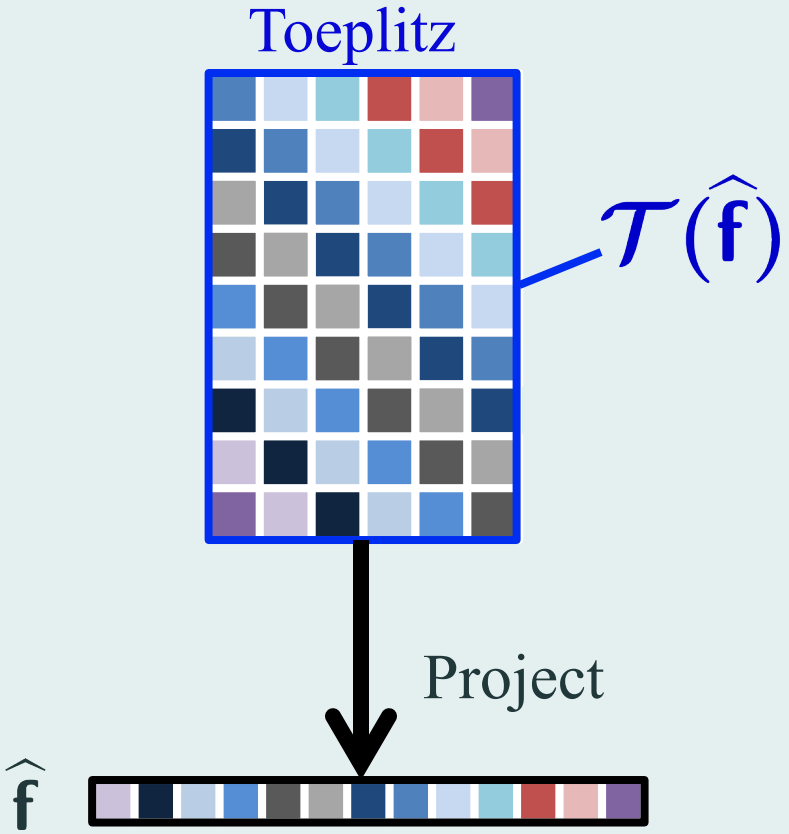


$\widehat{\mathbf{f}}$ 

Recovery as a structured low-rank matrix completion

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

1-D Example:



Recovery as a structured low-rank matrix completion

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

NP-Hard!

Recovery as a structured low-rank matrix completion

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

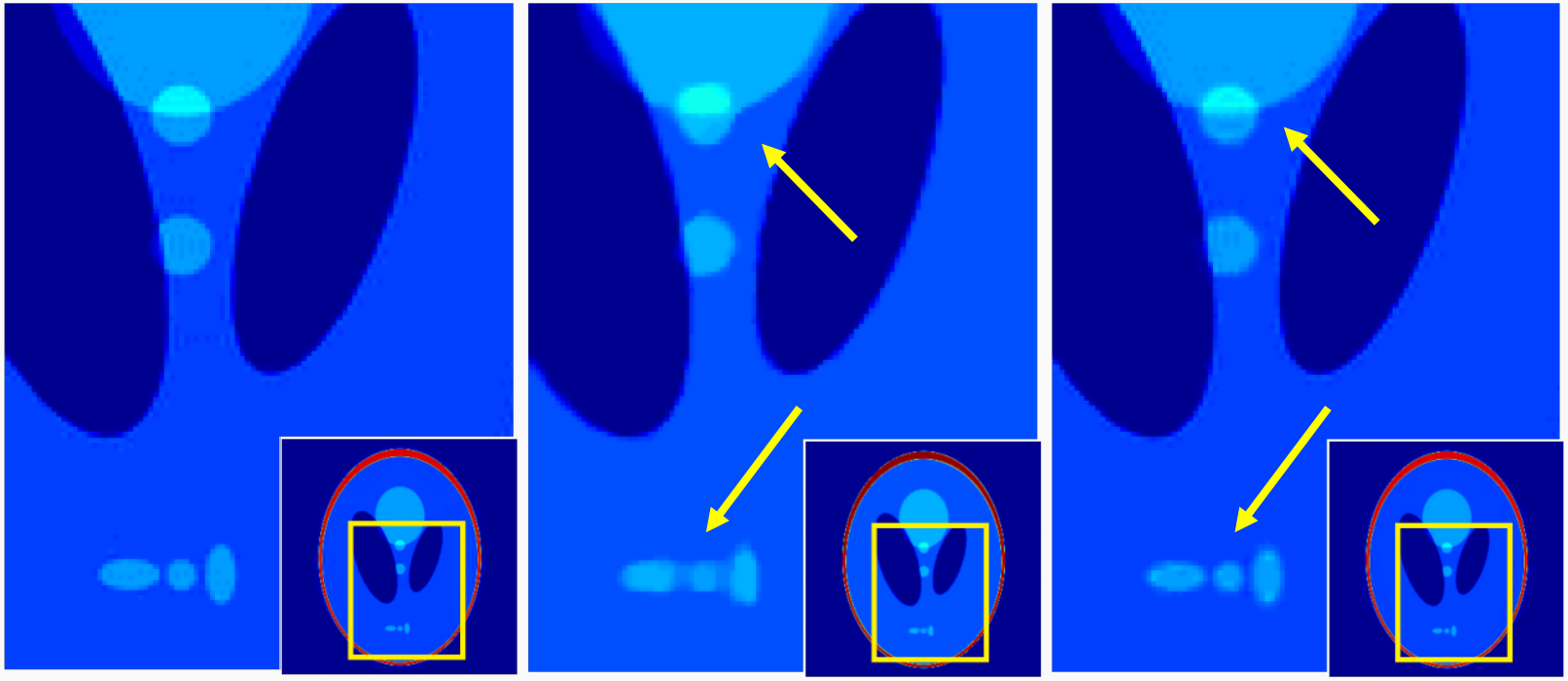


Convex Relaxation

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

Nuclear norm – sum of singular values

Recovery from 20-fold random undersampled data



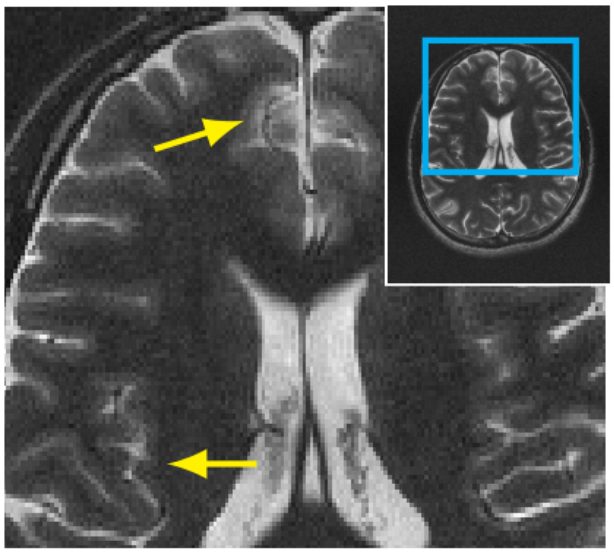
Fully sampled

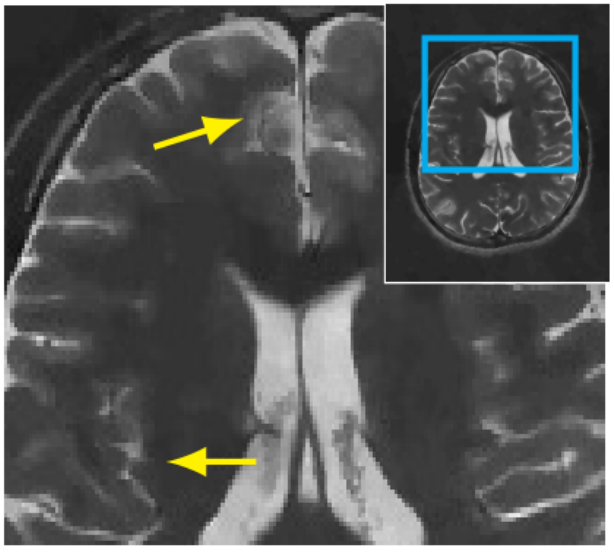
TV regularized recovery

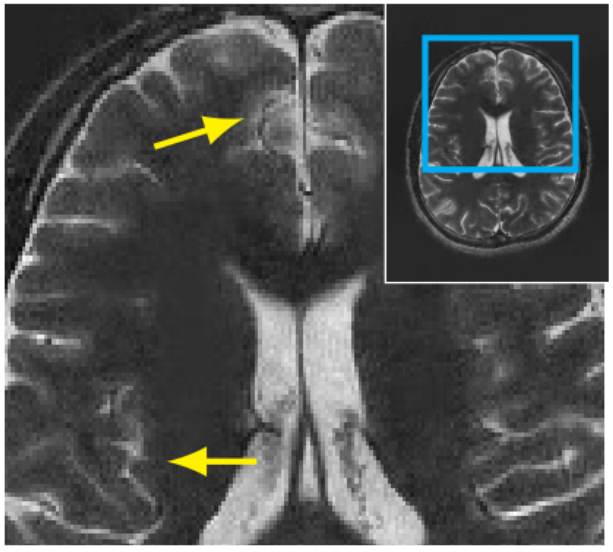
Structured low-rank recovery

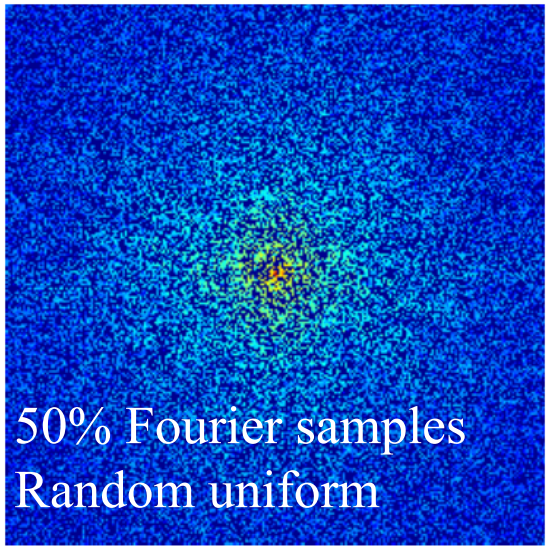
[Ongie & Jacob, SAMPTA 15](https://arxiv.org/abs/1609.07429)
https://arxiv.org/abs/1609.07429

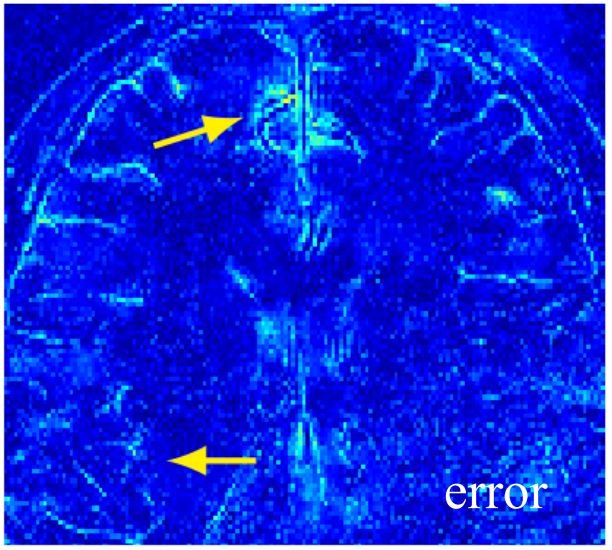
Fully sampled TV (SNR=17.8dB) GIRAF (SNR=19.0)

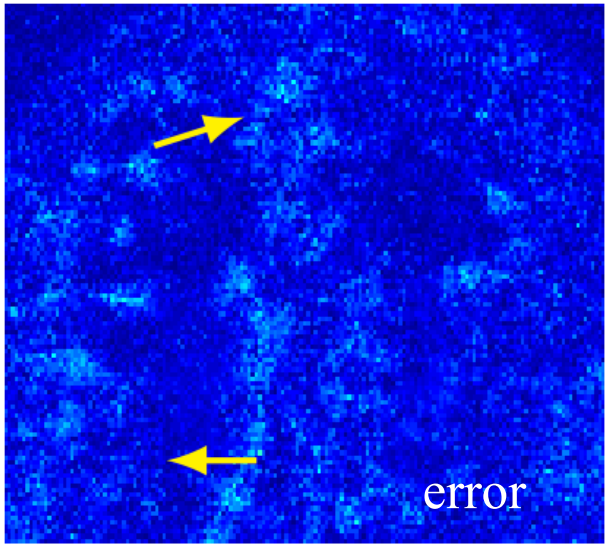












Ongie & Jacob, SAMPTA 15
https://arxiv.org/abs/1609.07429

Performance guarantee

Let f be a piecewise constant signal with edge set, which is the zero level set of a bandlimited function. Assume that f is sampled uniformly at m locations random on a Fourier domain grid Γ. Then, f can be recovered from the samples using a SLR approach if

$$m > \rho_1 \, c_s \, r \, \log^4 |\Gamma|$$

ρ_1 = incoherence measure of edge-set

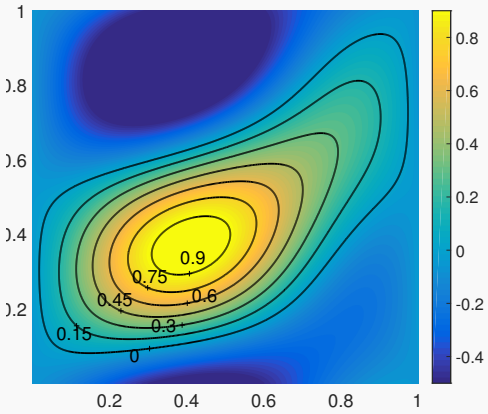
r = rank of $\mathcal{T}(\hat{f})$

c_s = ratio of grid size to filter size

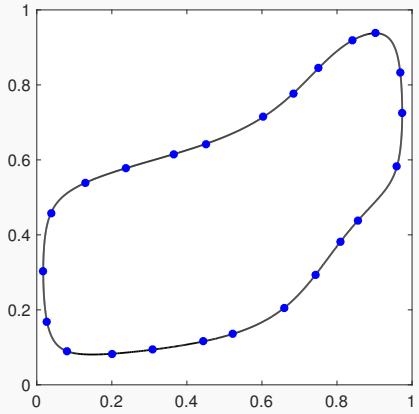
Ongie & Jacob, ICIP16, https://arxiv.org/abs/1703.01405

Incoherency measure ρ_1

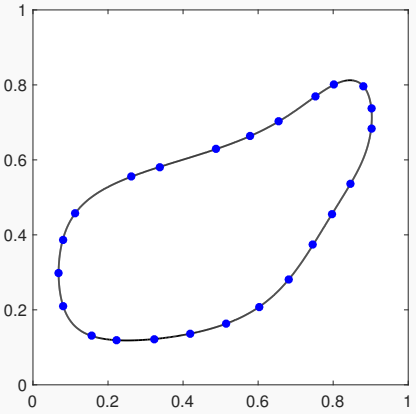
Intuition: minimum separation distance when packing r points on the edge-set curve, where $\mathbf{r} = \mathbf{rank}\,\mathcal{T}(\widehat{\mathbf{f}})$



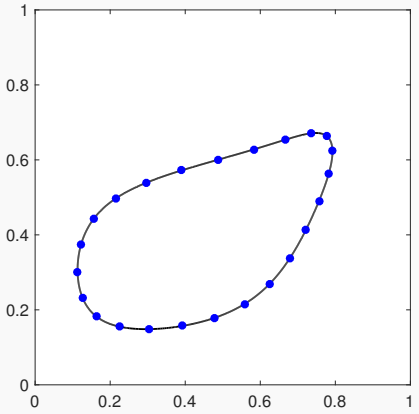
(a) Level-sets of μ_0



(b) $\rho \leq 8.0$



(c) $\rho \leq 264.9$



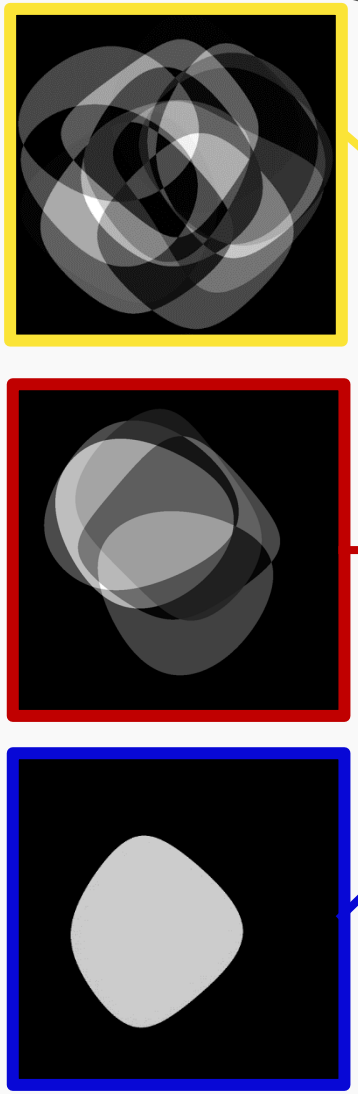
(d) $\rho \leq 5.0 \times 10^4$

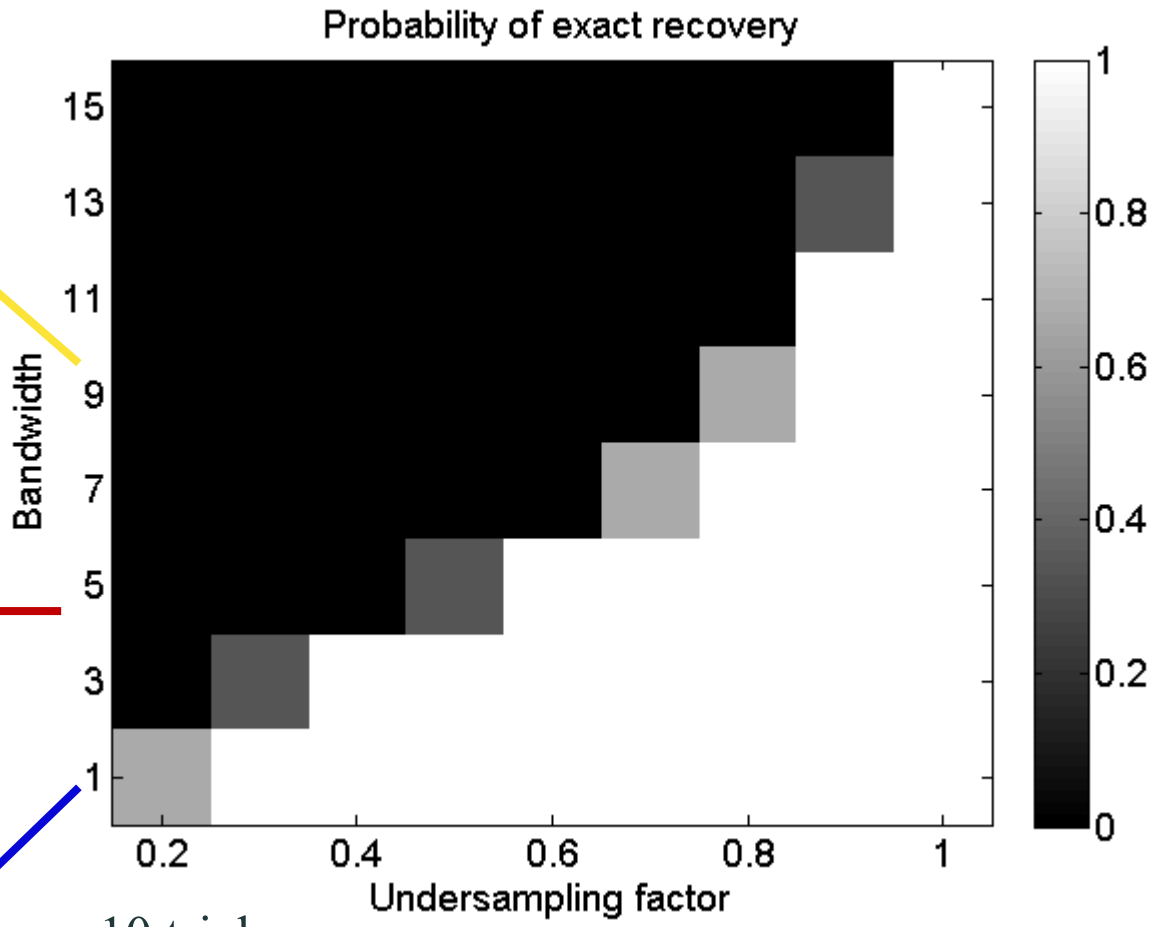
Small regions: high incoherence & more measurements

Complex boundaries: high rank/bandwidth

Phase transitions

Randomly generated synthetic PWC images





- 10 trials
- Uniform random Fourier samples
- 64x64 Fourier sampling window

[Ongie & Jacob, ICIP16, https://arxiv.org/abs/1703.01405]

Related structured low-rank methods in MRI

Magnetic Resonance in Medicine 64:457–471 (2010)

SPIRiT: Iterative Self-consistent Parallel Imaging Reconstruction From Arbitrary *k*-Space

Michael Lustig[1,2]* and John M. Pauly[2]

Discrete formulation exploiting multichannel acquisition

Low-Rank Modeling of Local k-Space Neighborhoods (LORAKS) for Constrained MRI

Justin P. Haldar, *Member, IEEE*

Discrete formulation exploiting sparsity, smoothly varying phase, and multichannel acquisition

Overview

1. Introduction
2. Review of Compressive Sensing
3. FRI extrapolation from uniform samples
4. Structured low-rank interpolation for non-uniform samples
5. Fast algorithms
6. Biomedical applications

Nuclear norm minimization

$$\min_{\hat{\mathbf{f}}} \|\mathbf{A}\hat{\mathbf{f}} - \mathbf{b}\|^2 + \lambda\|\mathbf{X}\|_* \quad \text{s.t.} \quad \mathbf{X} = \mathcal{T}(\hat{\mathbf{f}})$$

ADMM = Singular value thresholding (SVT)

1. Singular value thresholding step

   -compute *full SVD* of X!

2. Solve linear least squares problem

   -analytic solution or CG solve



Alternating projections ["SAKE," Shin 14], ["LORAKS," Haldar, 14]

$$\min_{\widehat{\mathbf{f}}} \|\mathbf{A}\widehat{\mathbf{f}} - \mathbf{b}\|^2 \quad \text{s.t.} \quad \boxed{\mathbf{X} = \mathcal{T}(\widehat{\mathbf{f}})}$$

$$\boxed{\text{rank}\ \mathbf{X} \leq \mathbf{r}}$$

Alternating projection algorithm (Cadzow)

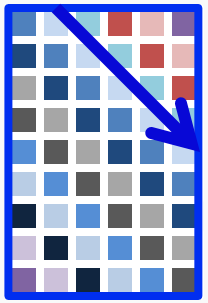
1. Project onto space of rank r matrices

   -Compute *truncated SVD:* $\mathbf{X}^\star = \mathbf{U}\boldsymbol{\Sigma}_r\mathbf{V}^H$

2. Project onto space of structured matrices

   -Average along "diagonals"



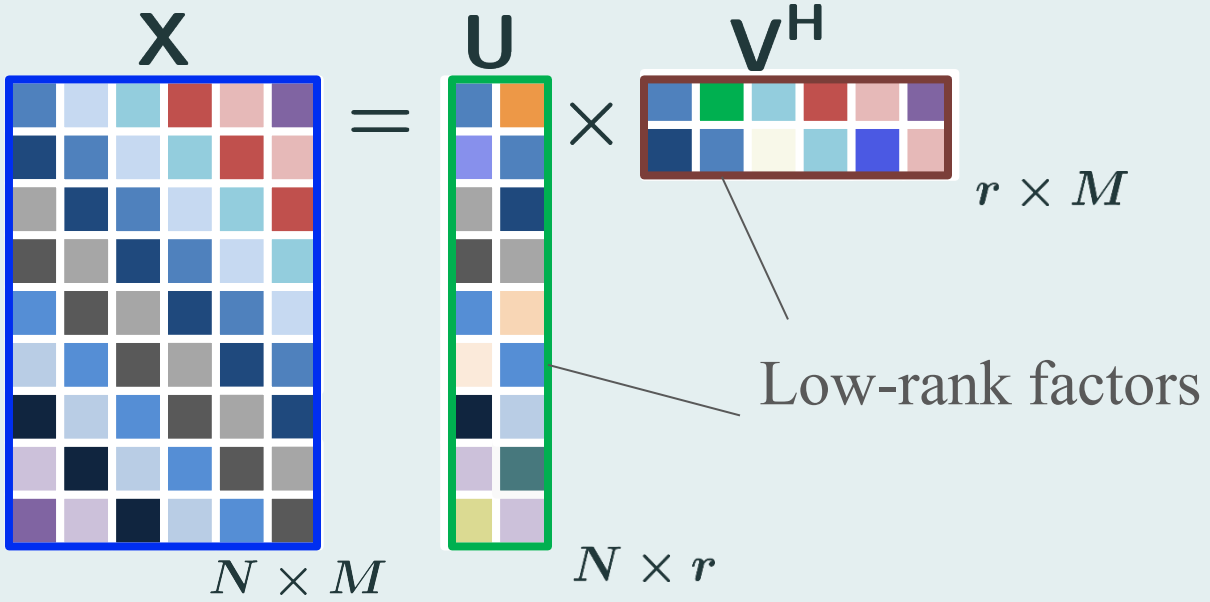


U,V factorization [O.& Jacob, SampTA 15, Jin et al., ISBI 15]

$$\min_{\widehat{\mathbf{f}}} \|\mathbf{A}\widehat{\mathbf{f}} - \mathbf{b}\|^2 + \lambda \|\mathbf{X}\|_* \quad \text{s.t.} \quad \mathbf{X} = \mathcal{T}(\widehat{\mathbf{f}})$$

"U,V factorization trick"

$$\|\mathbf{X}\|_* = \min_{\mathbf{X}=\mathbf{U}\mathbf{V}^H} \frac{1}{2} \left(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2 \right)$$



U,V factorization [O.& Jacob, SampTA 15, Jin et al., ISBI 15]

$$\min_{\hat{\mathbf{f}},\mathbf{U},\mathbf{V}} \|\mathbf{A}\hat{\mathbf{f}} - \mathbf{b}\|^2 + \frac{\lambda}{2}\left(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2\right)$$

$$\text{s.t.} \quad \mathbf{U}\mathbf{V}^H = \mathcal{T}(\hat{\mathbf{f}})$$

UV factorization approach

1. ~~Singular value thresholding step~~

   ~~-compute *full SVD* of X!~~

   SVD-free → fast matrix inversion steps

2. Solve linear least squares problem

   -analytic solution or CG solve



Main challenge : Computational complexity & memory

2-D

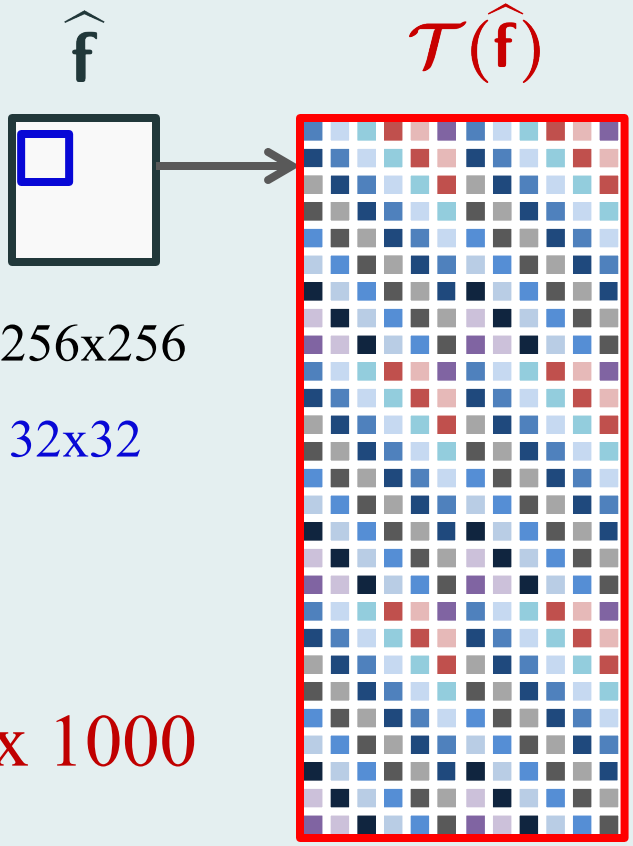
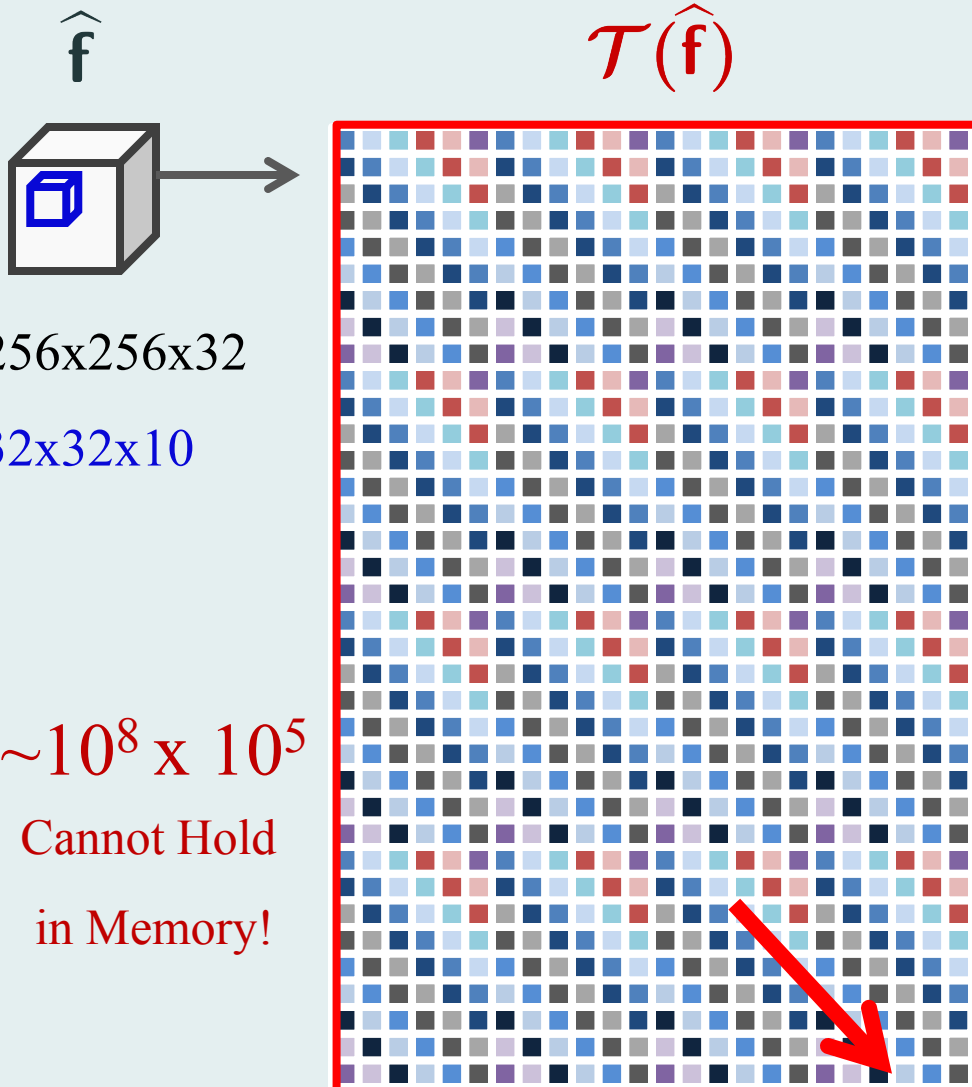


Image: 256x256

Filter: 32x32

$\sim 10^6 \times 1000$

3-D



256x256x32

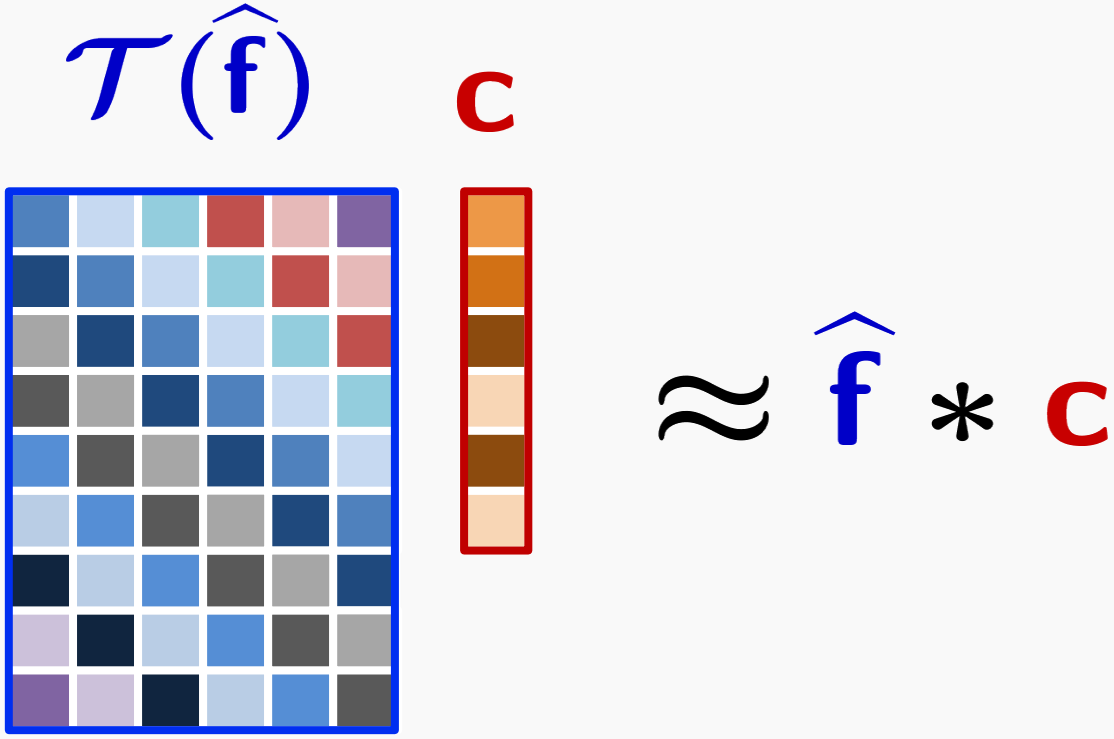
32x32x10

$\sim 10^8 \times 10^5$

Cannot Hold

in Memory!

Exploit convolutional structure of the matrix



Fast evaluation using FFT

Direct computation of small Gram matrix: avoid storage

IRLS algorithm along with structure exploitation

- Original IRLS: To recover low-rank matrix X, iterate

$$\mathbf{W} \leftarrow (\mathbf{X}^{\mathbf{H}}\mathbf{X} + \epsilon\mathbf{I})^{-\frac{1}{2}}$$

$$\mathbf{X} \leftarrow \arg\min_{\mathbf{X}} \|\mathbf{X}\mathbf{W}^{\frac{1}{2}}\|_{\mathbf{F}}^{2} + \lambda\|\mathbf{A}\mathbf{X} - \mathbf{B}\|_{\mathbf{F}}^{2}$$

IRLS

- Original IRLS: To recover low-rank matrix X, iterate

$$\mathbf{W} \leftarrow (\mathbf{X}^{\mathsf{H}}\mathbf{X} + \epsilon \mathbf{I})^{-\frac{1}{2}}$$

$$\mathbf{X} \leftarrow \arg\min_{\mathbf{X}} \|\mathbf{X}\mathbf{W}^{\frac{1}{2}}\|_{\mathsf{F}}^2 + \lambda \|\mathbf{A}\mathbf{X} - \mathbf{B}\|_{\mathsf{F}}^2$$

- We adapt to structured case: $\mathbf{X} = \mathcal{T}(\hat{f})$

$$\mathbf{W} \leftarrow (\mathcal{T}(\hat{\mathbf{f}})^{\mathsf{H}}\mathcal{T}(\hat{\mathbf{f}}) + \epsilon \mathbf{I})^{-\frac{1}{2}}$$

$$\hat{\mathbf{f}} \leftarrow \arg\min_{\hat{\mathbf{f}}} \|\mathcal{T}(\hat{\mathbf{f}})\mathbf{W}^{\frac{1}{2}}\|_{\mathsf{F}}^2 + \lambda \|\mathbf{A}\hat{\mathbf{f}} - \mathbf{b}\|^2$$

IRLS algorithm

- Original IRLS: To recover low-rank matrix X, iterate

$$\mathbf{W} \leftarrow (\mathbf{X}^{\mathsf{H}}\mathbf{X} + \epsilon \mathbf{I})^{-\frac{1}{2}}$$

$$\mathbf{X} \leftarrow \arg\min_{\mathbf{X}} \|\mathbf{X}\mathbf{W}^{\frac{1}{2}}\|_{\mathsf{F}}^2 + \lambda \|\mathbf{A}\mathbf{X} - \mathbf{B}\|_{\mathsf{F}}^2$$

- We adapt to structured case: $\mathbf{X} = \mathcal{T}(\hat{f})$

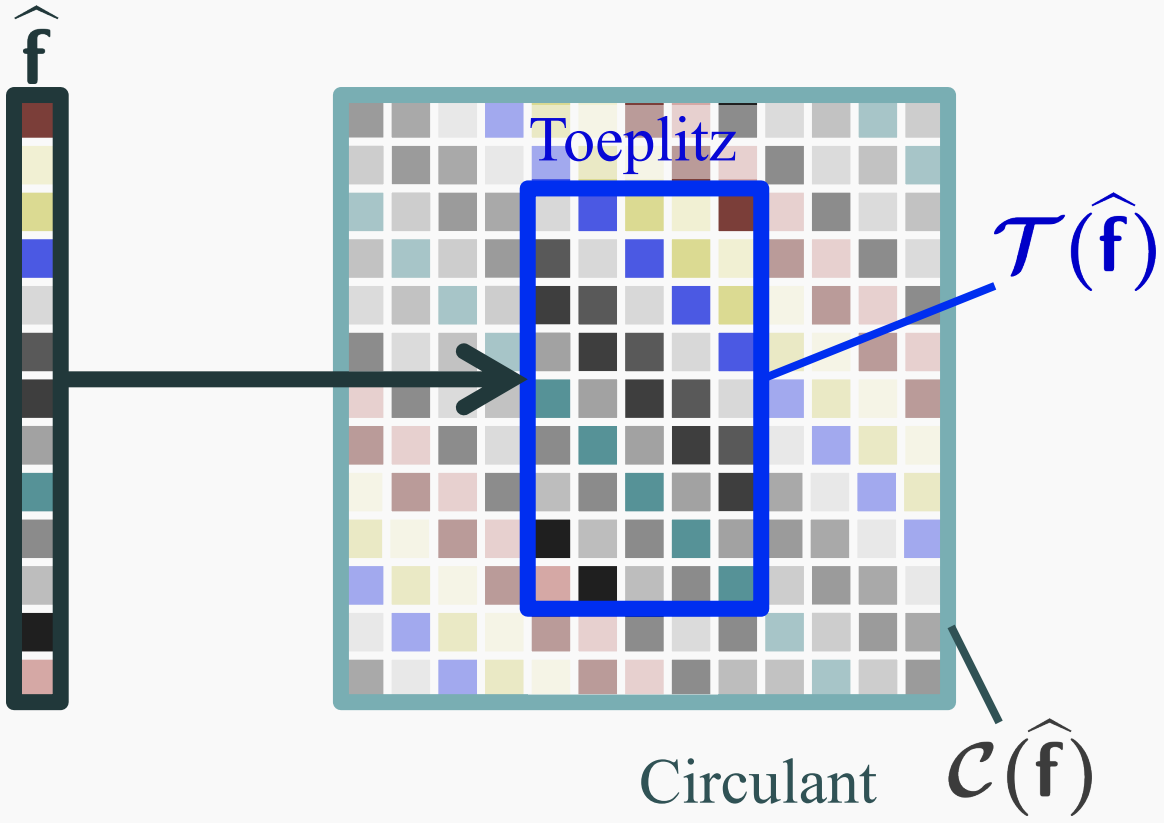
$$\mathbf{W} \leftarrow (\mathcal{T}(\hat{\mathbf{f}})^{\mathsf{H}}\mathcal{T}(\hat{\mathbf{f}}) + \epsilon \mathbf{I})^{-\frac{1}{2}}$$

$$\hat{\mathbf{f}} \leftarrow \arg\min_{\hat{\mathbf{f}}} \|\mathcal{T}(\hat{\mathbf{f}})\mathbf{W}^{\frac{1}{2}}\|_{\mathsf{F}}^2 + \lambda \|\mathbf{A}\hat{\mathbf{f}} - \mathbf{b}\|^2$$

Without modification, this approach is still slow!

Ongie & Jacob, ISBI16
https://arxiv.org/abs/1609.07429

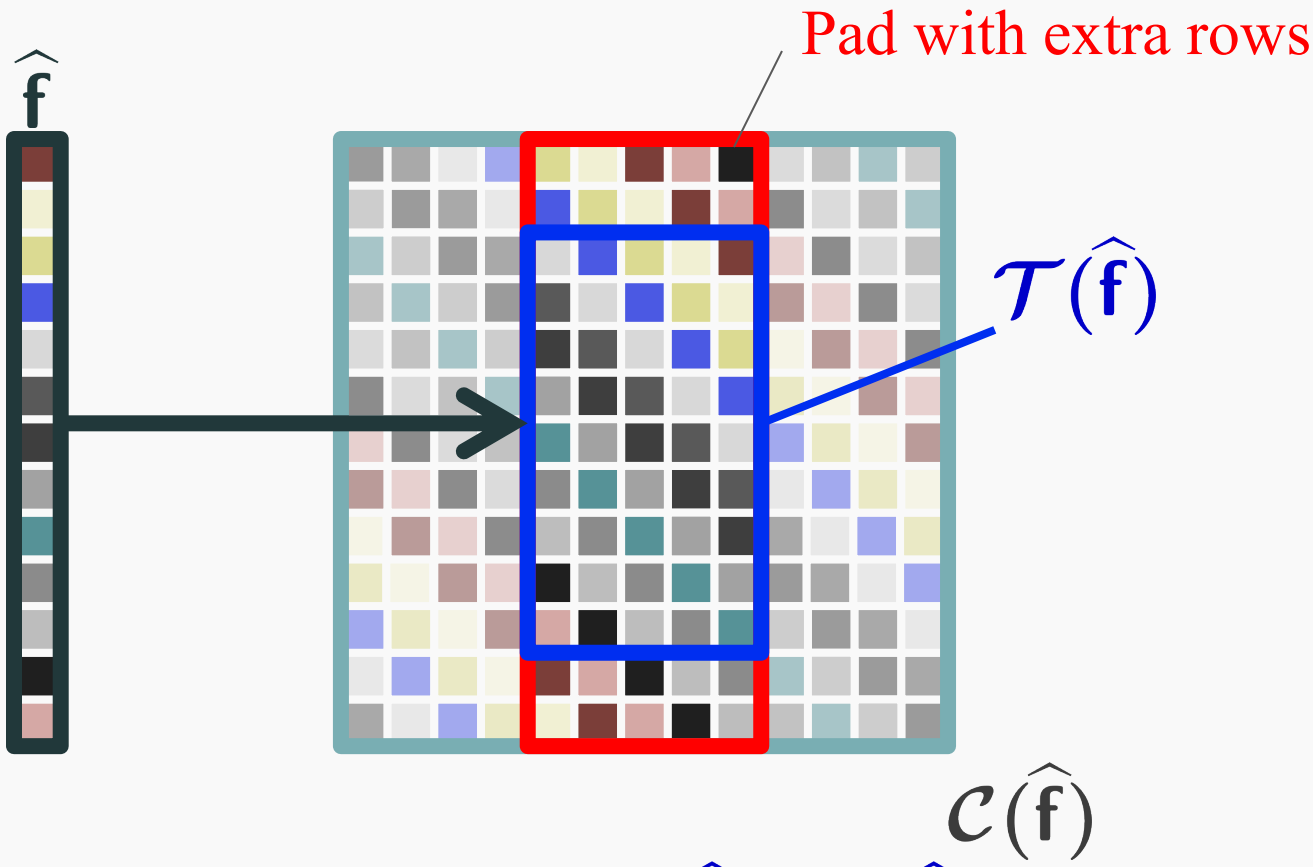
Idea 1: Embed Toeplitz lifting in circulant matrix



*Fast matrix-vector products with $\mathcal{T}(\widehat{\mathbf{f}})$ by FFTs

Ongie & Jacob, ISBI16
https://arxiv.org/abs/1609.07429

Idea 2: Approximate matrix lifting



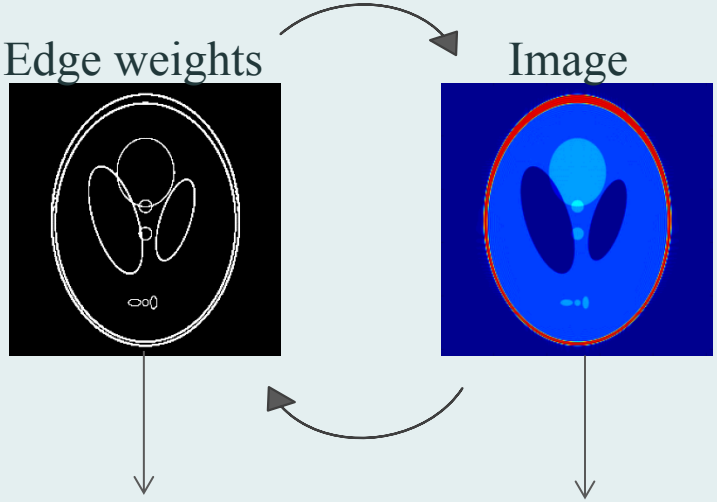
*Fast computation of $\mathcal{T}(\hat{\mathbf{f}})^{\mathsf{H}}\mathcal{T}(\hat{\mathbf{f}})$ by FFTs

Ongie & Jacob, ISBI16
https://arxiv.org/abs/1609.07429

GIRAF: fast [O. & Jacob, 2016 (arXiv)]

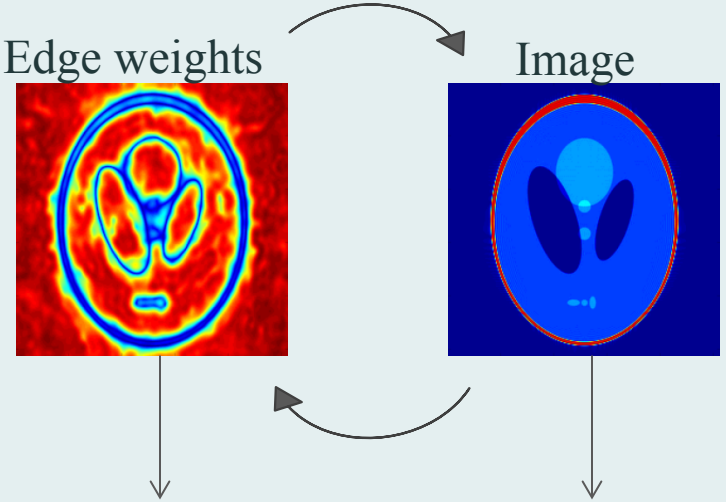
Complexity similar to IRLS for TV minimization

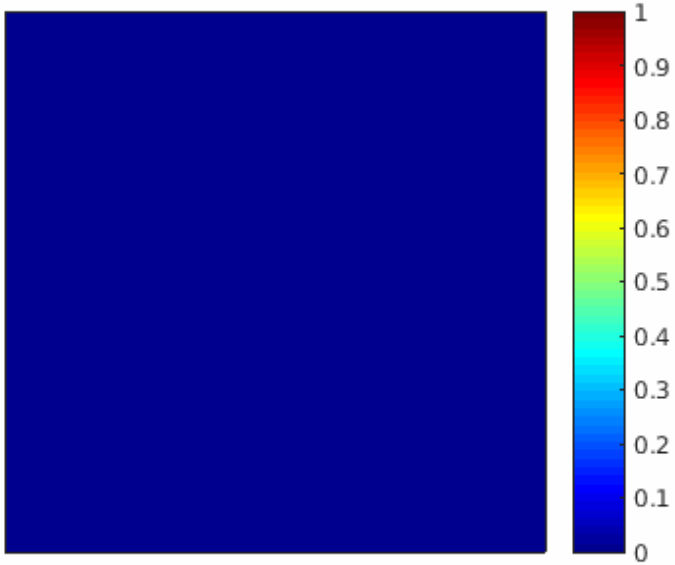
IRLS TV-minimization

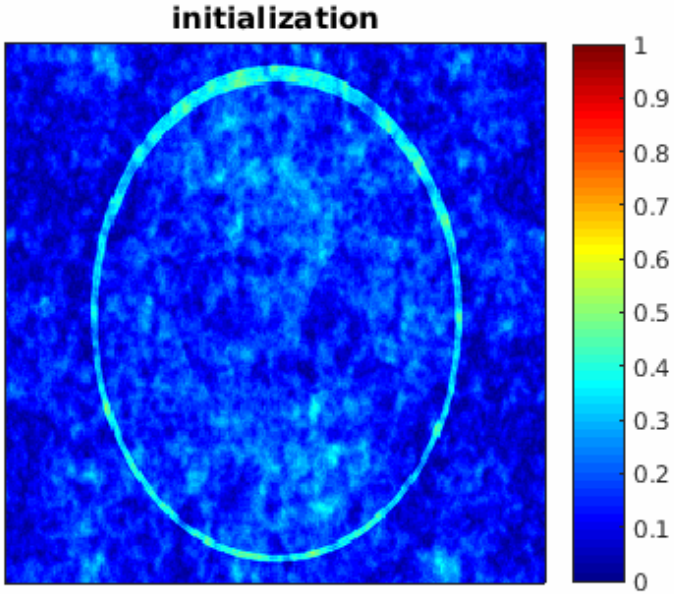


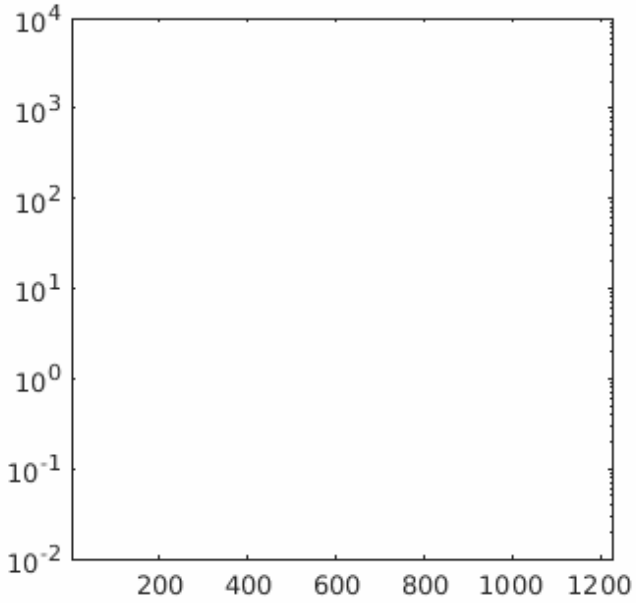
$$w_{i,j} = \frac{1}{|(\nabla f)_{i,j}| + \epsilon}$$

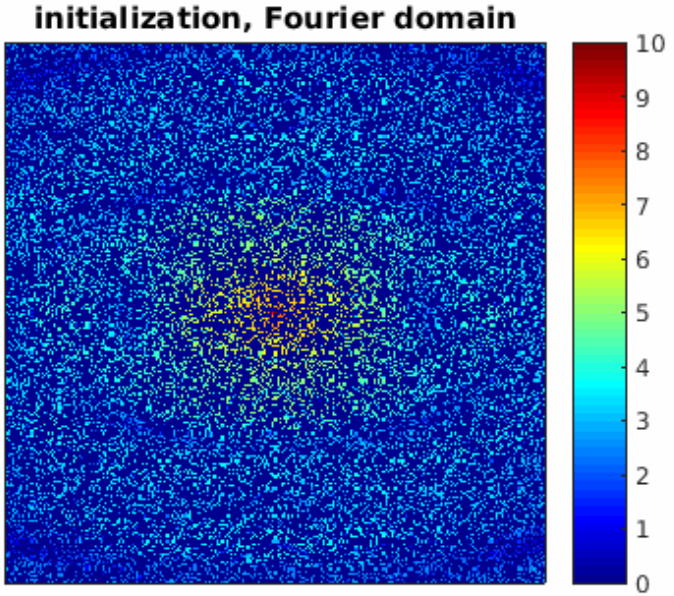
GIRAF algorithm



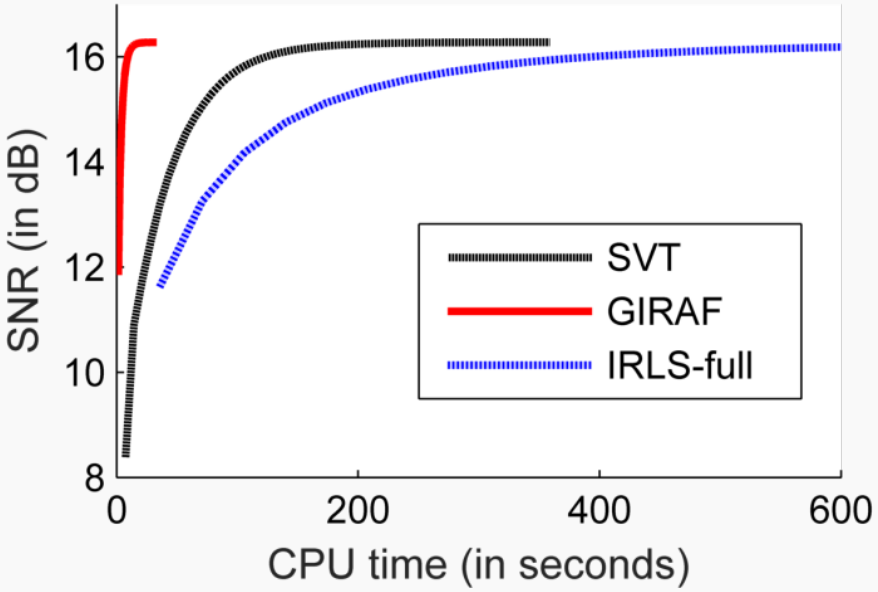

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0


initialization
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0


10^4
10^3
10^2
10^1
10^0
10^{-1}
10^{-2}
200
400
600
800
1000
1200


initialization, Fourier domain
10
9
8
7
6
5
4
3
2
1
0

Convergence speed of GIRAF

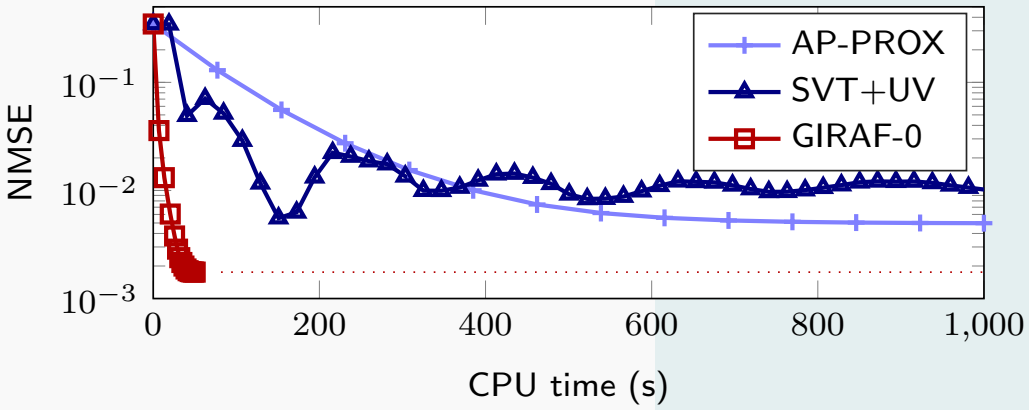


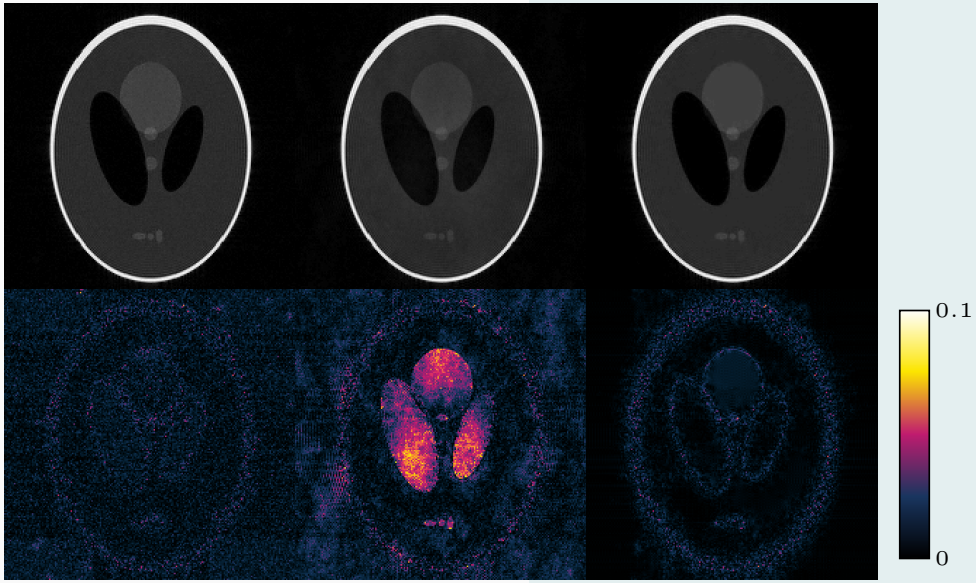
	15×15 filter		31×31 filter	
Algorithm	# iter	total	# iter	total
SVT	7	110s	11	790 s
GIRAF	6	20s	7	44 s

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

Ongie & Jacob, ISBI16
https://arxiv.org/abs/1609.07429
Software available at https://research.engineering.uiowa.edu/cbig/software

Convergence speed of GIRAF





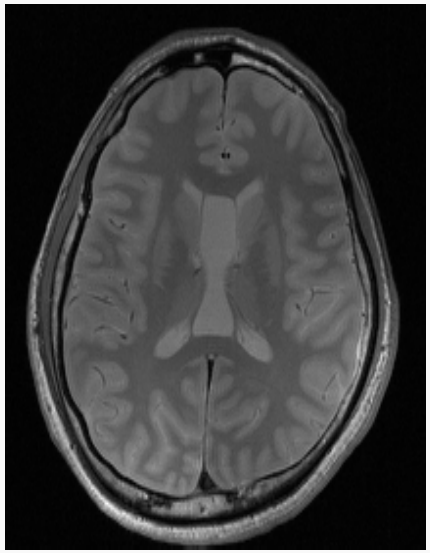
AP-PROX	SVT+UV	GIRAF-0
NMSE $= 4.9e{-}3$	NMSE $= 11.6e{-}3$	NMSE $= 1.8e{-}3$
Runtime: 1000 s	Runtime: 1090 s	Runtime: 49 s

Overview

1. Introduction
2. Review of Compressive Sensing
3. FRI extrapolation from uniform samples
4. Structured low-rank interpolation for non-uniform samples
5. Fast implementations
6. Biomedical applications
   a. Applications to MRI
   b. Other applications

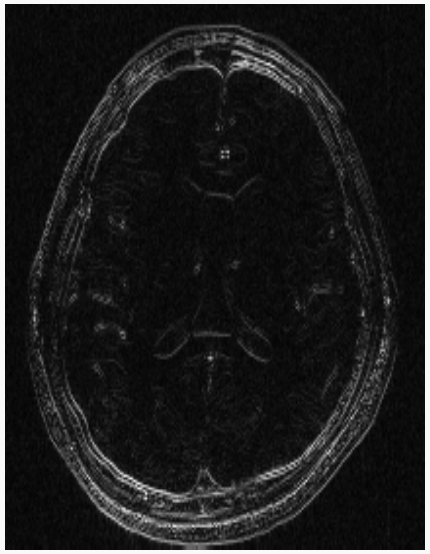
TV-domain sparse signal cases

y
(PE)



x (RO)

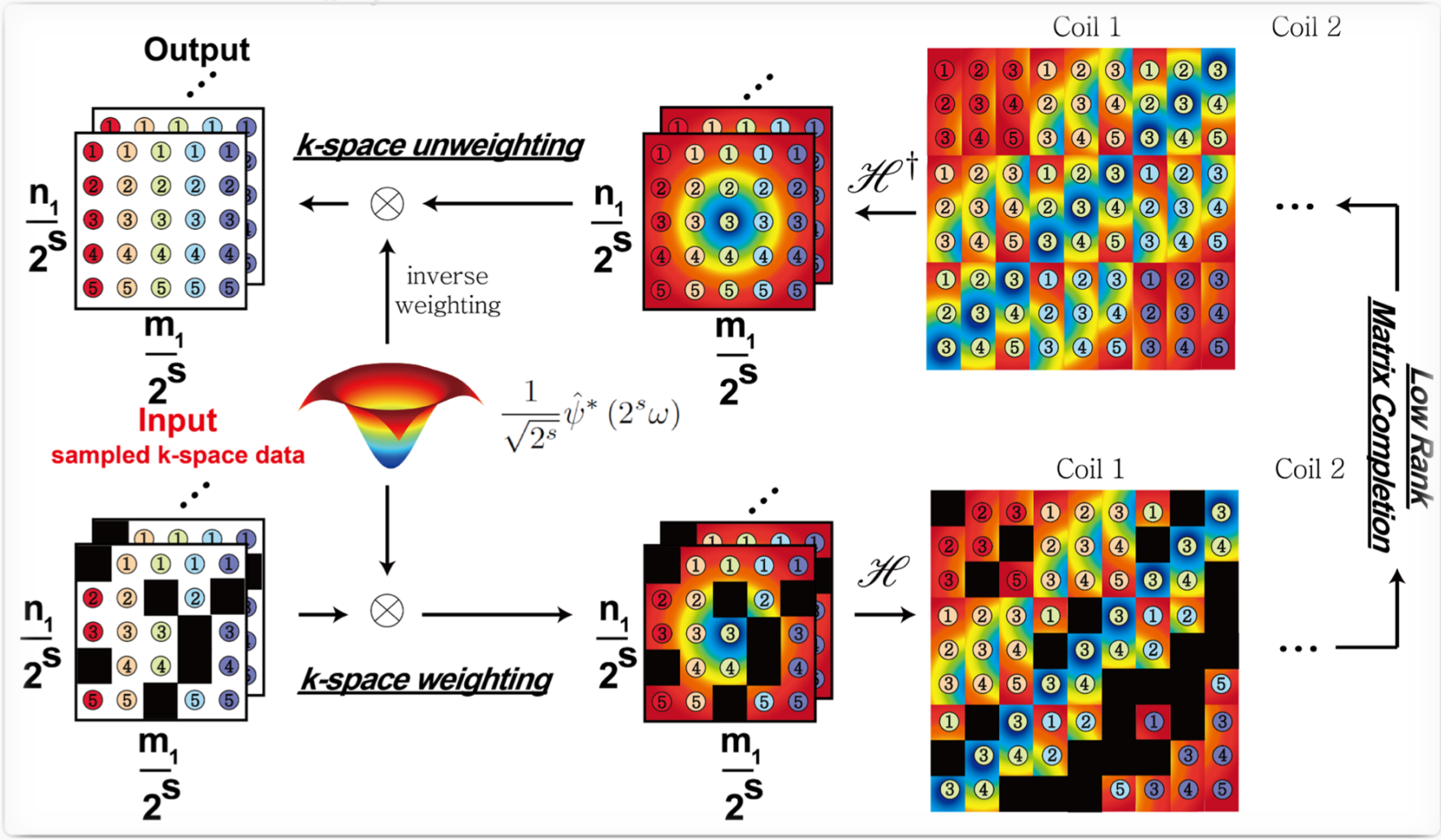
Sparse In
TV
domain



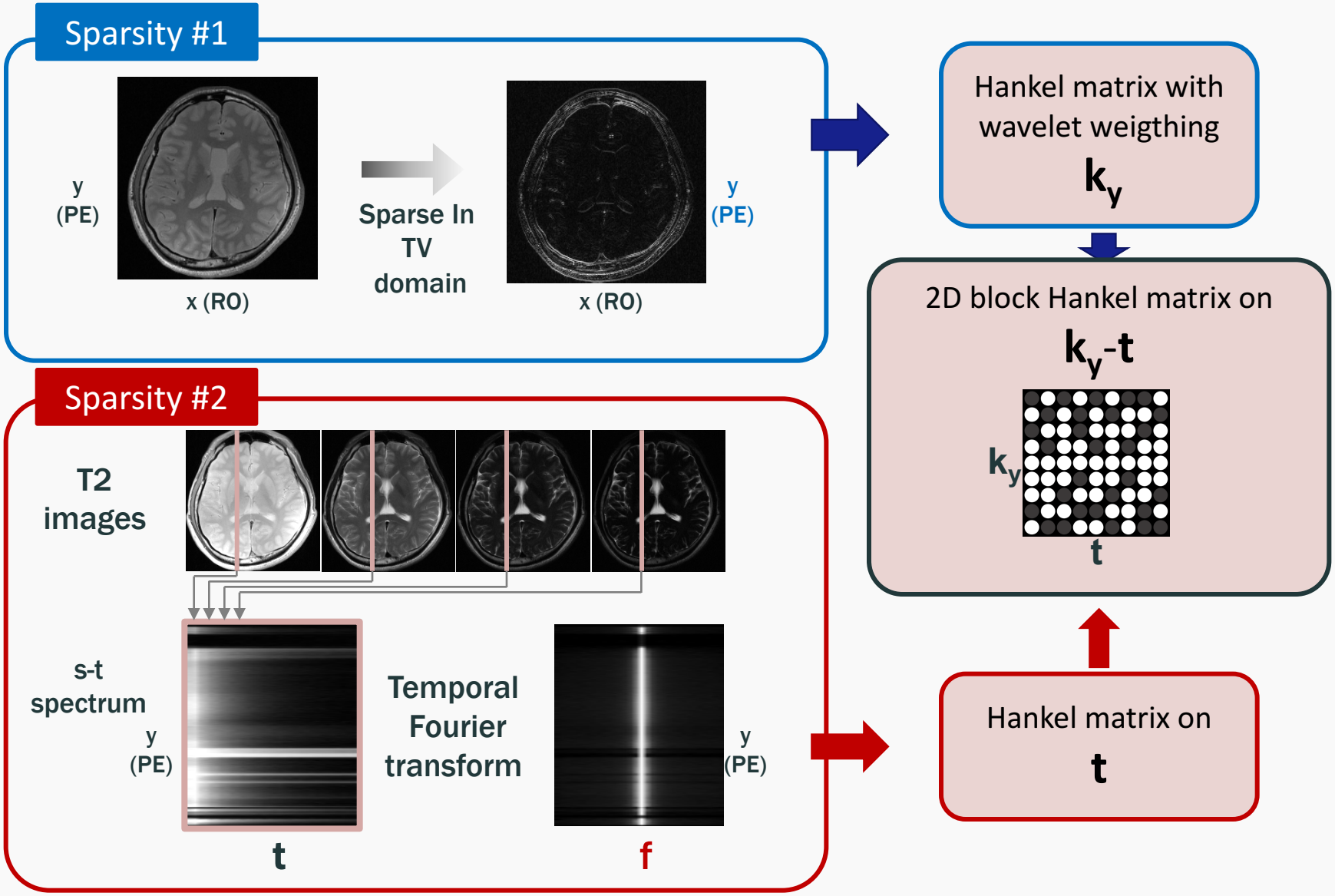
y
(PE)

x (RO)

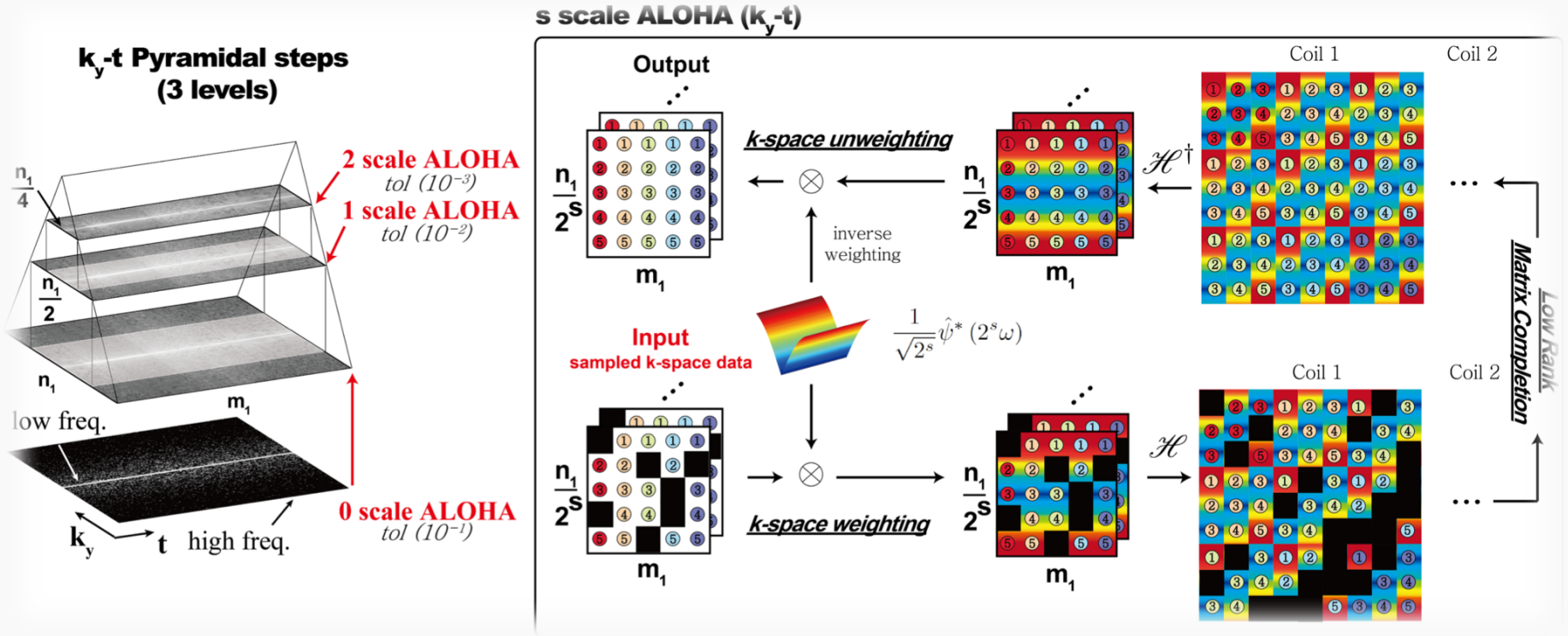
TV-domain sparse signal cases



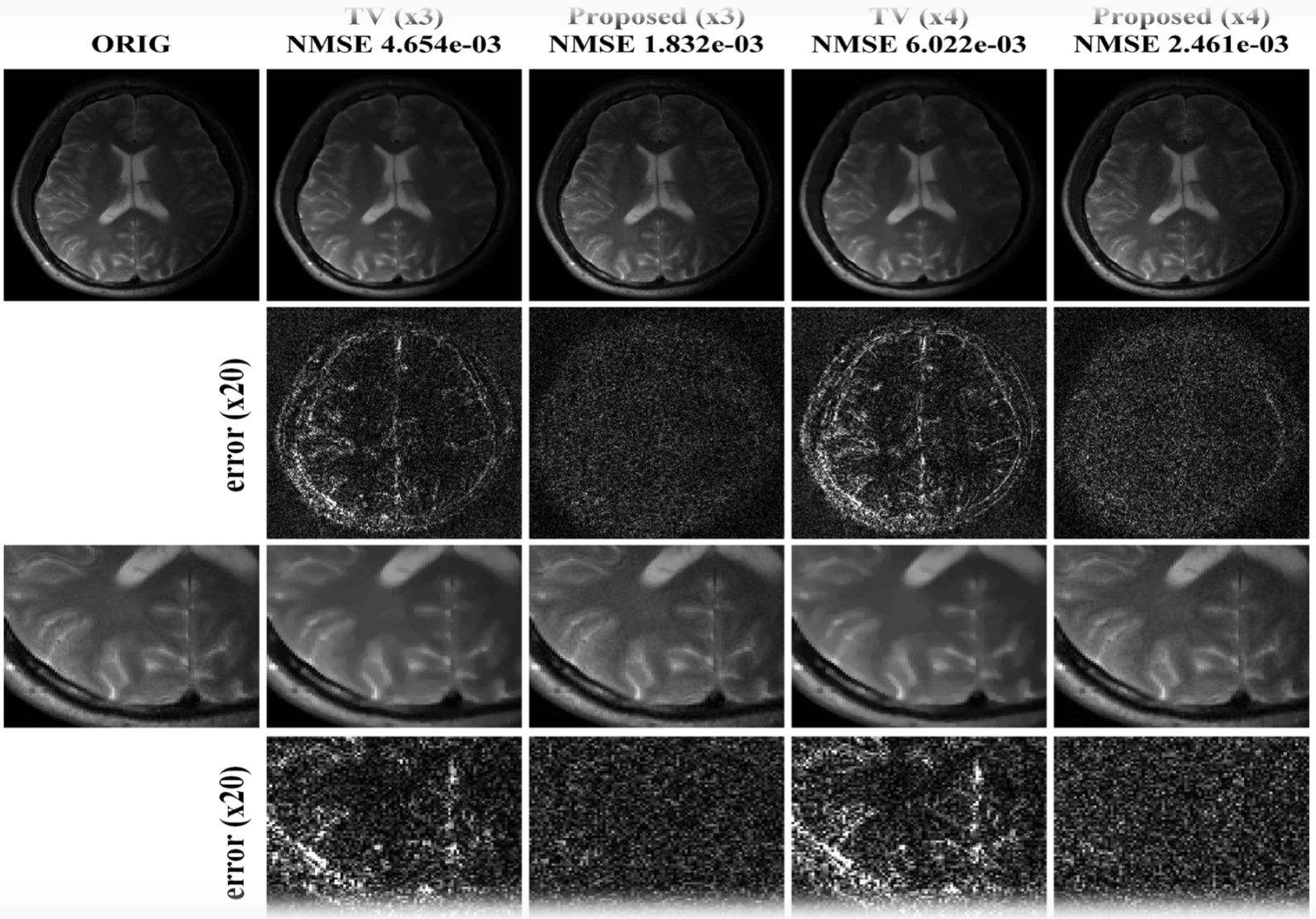
k-t dynamic sparse signal cases



k-t dynamic sparse signal cases



Single coil static MRI



Rank Bound for Parallel Imaging

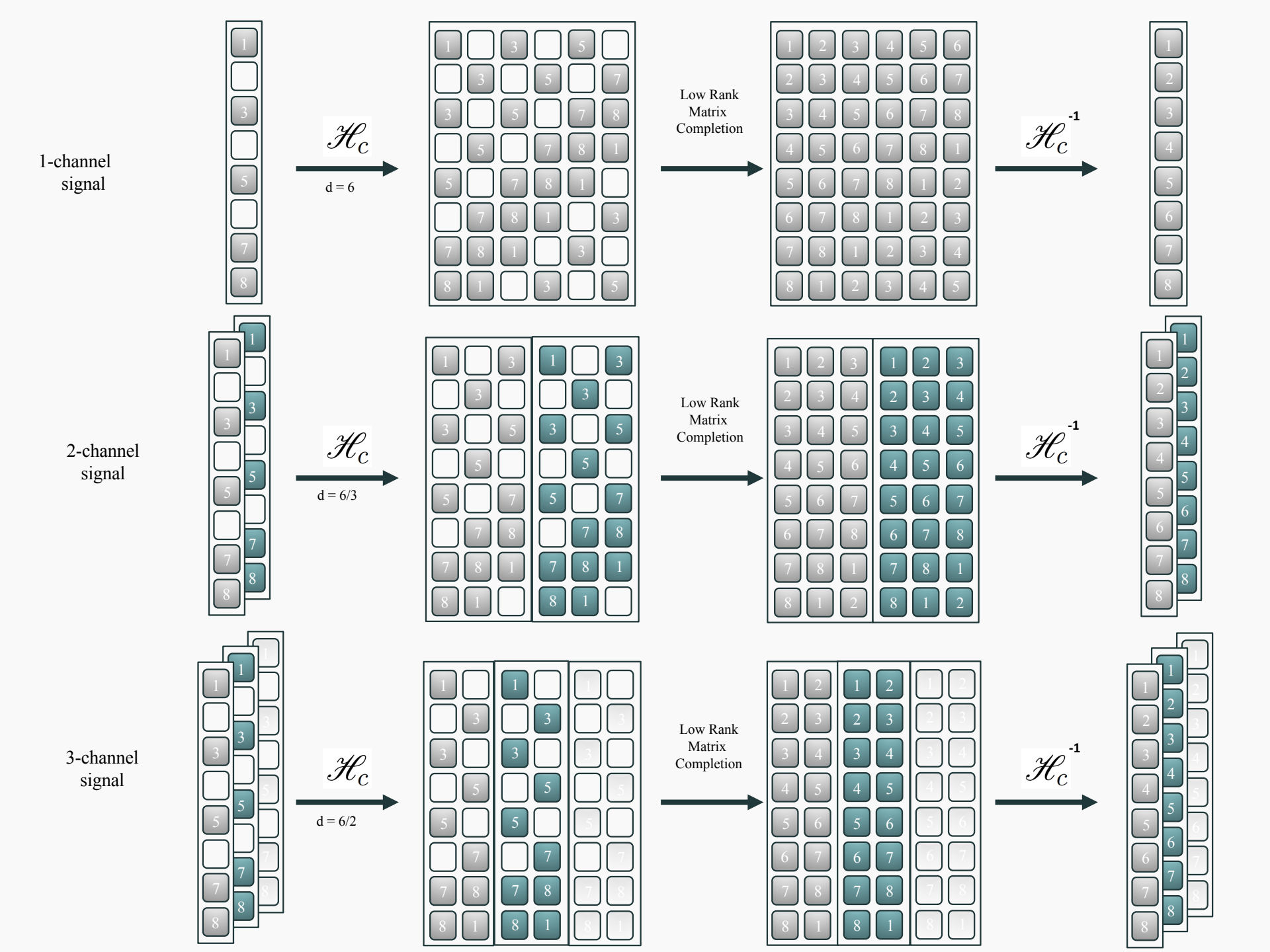
$$\mathcal{Y}_h = \left[\mathscr{H}_c(\hat{\mathbf{l}} \odot \hat{\mathbf{g}}_1) \quad \cdots \quad \mathscr{H}_c(\hat{\mathbf{l}} \odot \hat{\mathbf{g}}_{N_c}) \right]$$

$$\mathrm{RANK}\,\mathcal{Y}_h \leq \mathrm{RANK}\,\mathscr{H}_c(\hat{\mathbf{w}}) + \mathrm{RANK}\,\mathscr{H}_c(\hat{\mathbf{f}}_{tr})$$

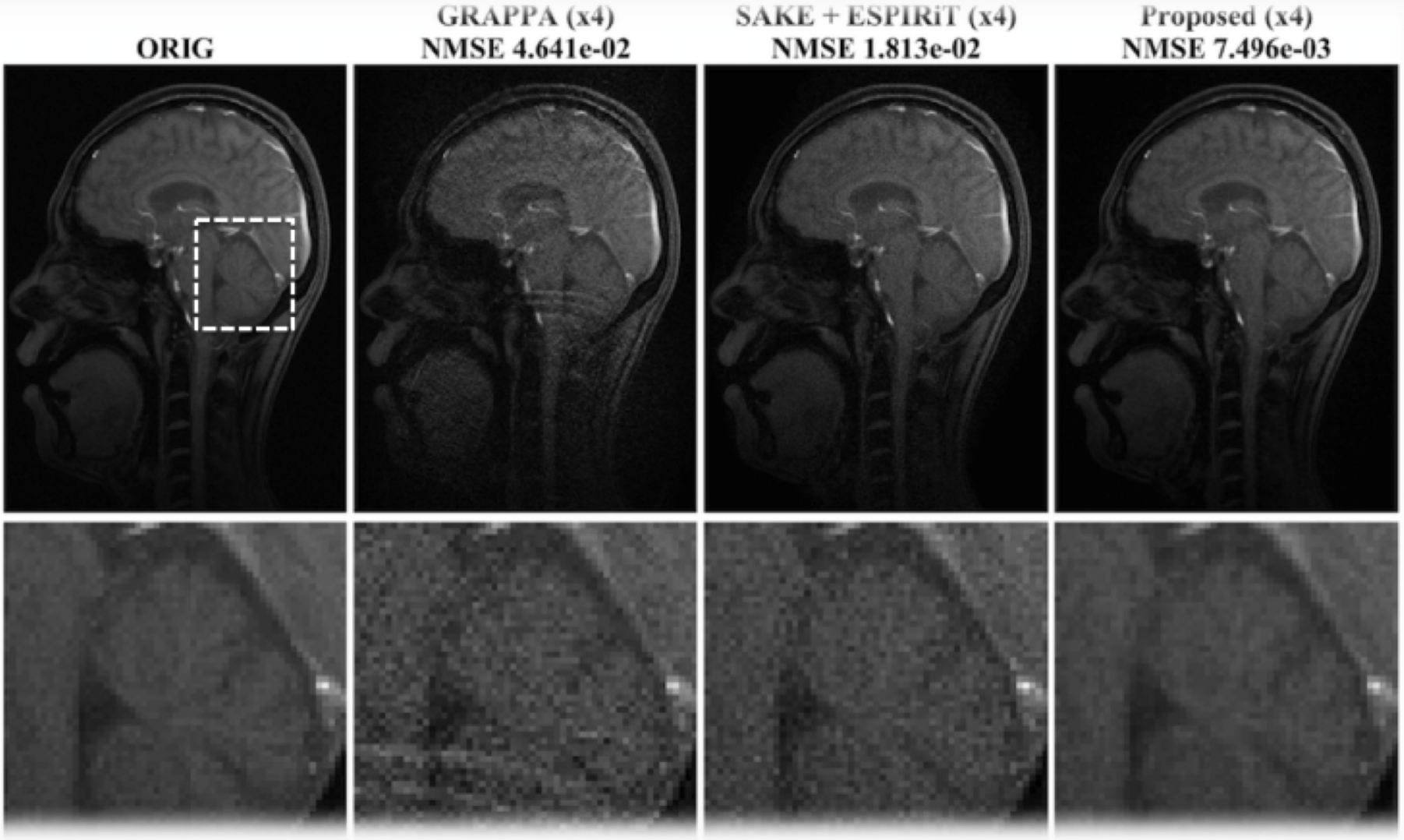
Sparsity of common image
In transform domain

Sparsity of sensitivity map
In Fourier domain

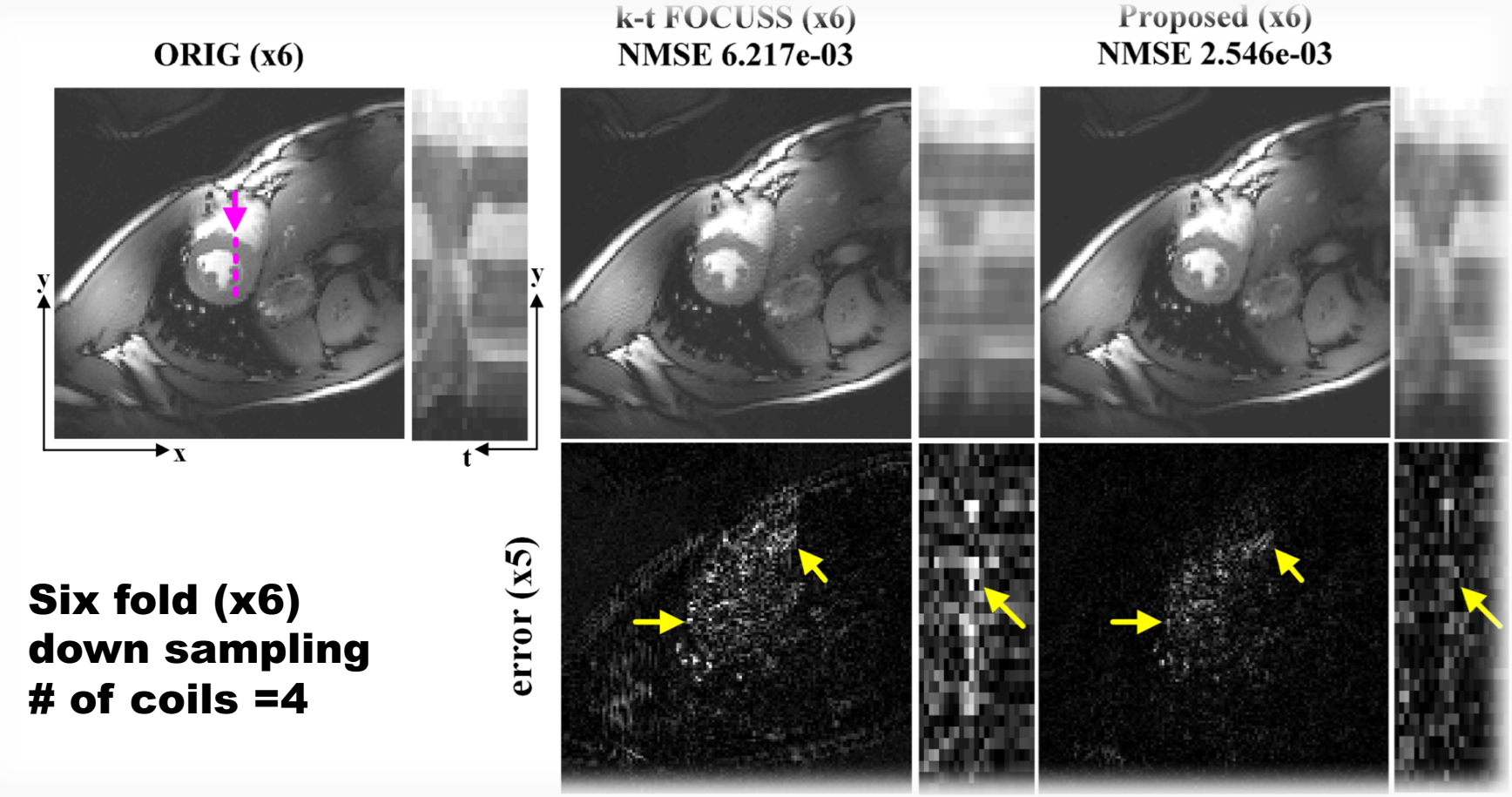
Jin et al, IEEE TCI, 2016


1-channel signal
$\mathscr{H}_C$
d = 6
Low Rank Matrix Completion
$\mathscr{H}_C^{-1}$
2-channel signal
$\mathscr{H}_C$
d = 6/3
Low Rank Matrix Completion
$\mathscr{H}_C^{-1}$
3-channel signal
$\mathscr{H}_C$
d = 6/2
Low Rank Matrix Completion
$\mathscr{H}_C^{-1}$

Parallel MRI



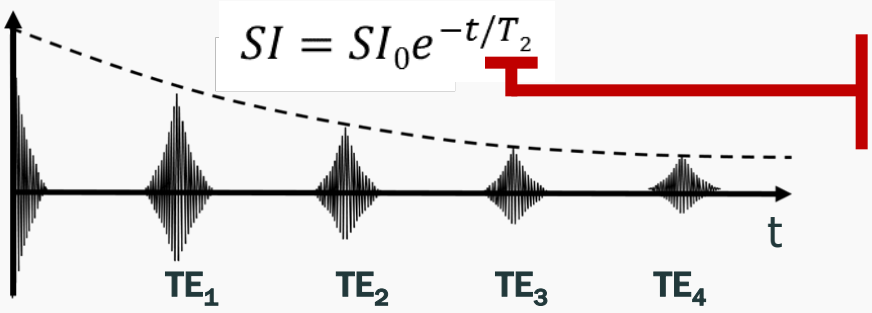
Dynamic MRI – multi coil



Six fold (x6) down sampling
# of coils =4

MR Parameter Mapping

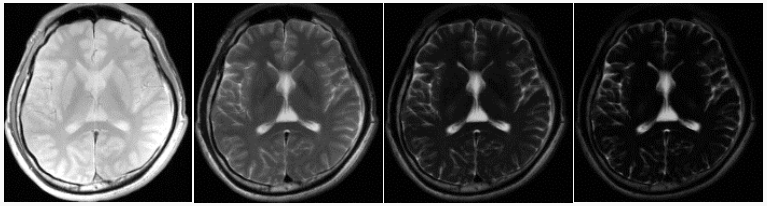
What is MR parameter mapping?



Finding the **quantitative** value of each tissue

e.g. Multi-Echo Spin-Echo (ME-SE, T2 mapping)

Pros | Cons



e.g. Multi-echo images for T2 mapping

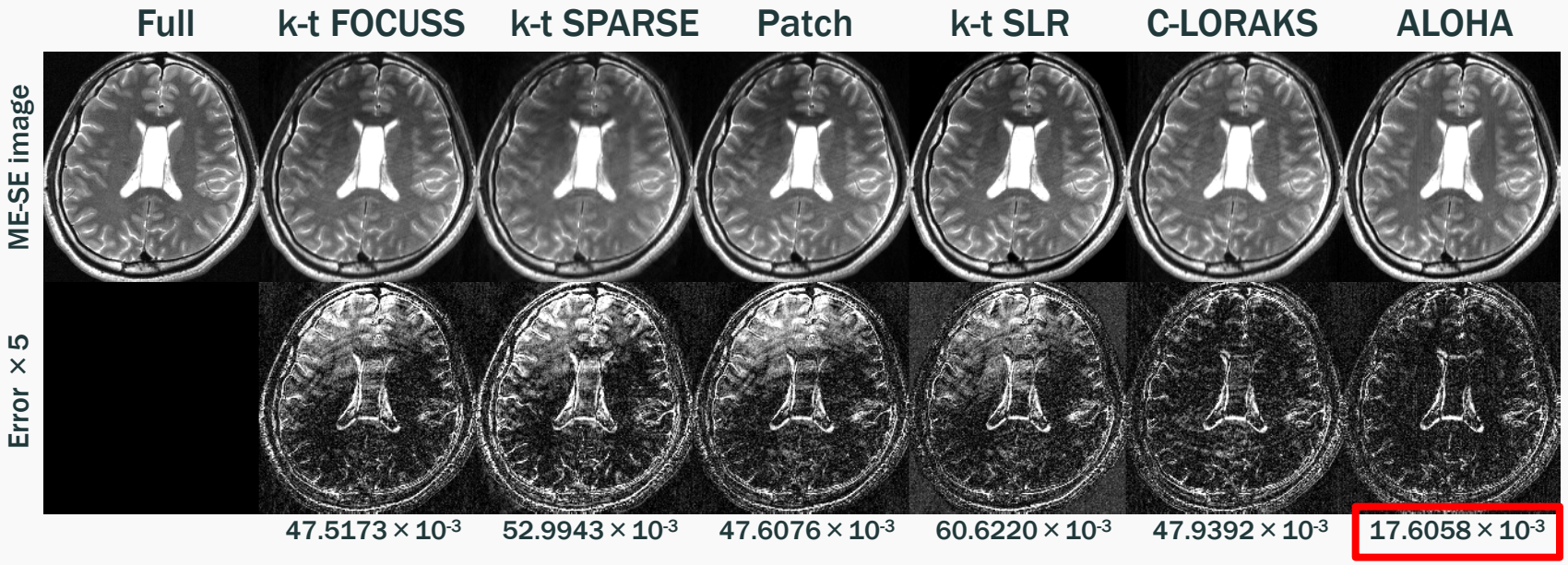
Long scan time

- Needs multiple scans
- Variation of TI, TE, FA, etc.

[1] Siemonsen, S et al., Radiology, 2008

[2] Rugg-Gunn, F. J., et al. *NEUROLOGY*, 2005

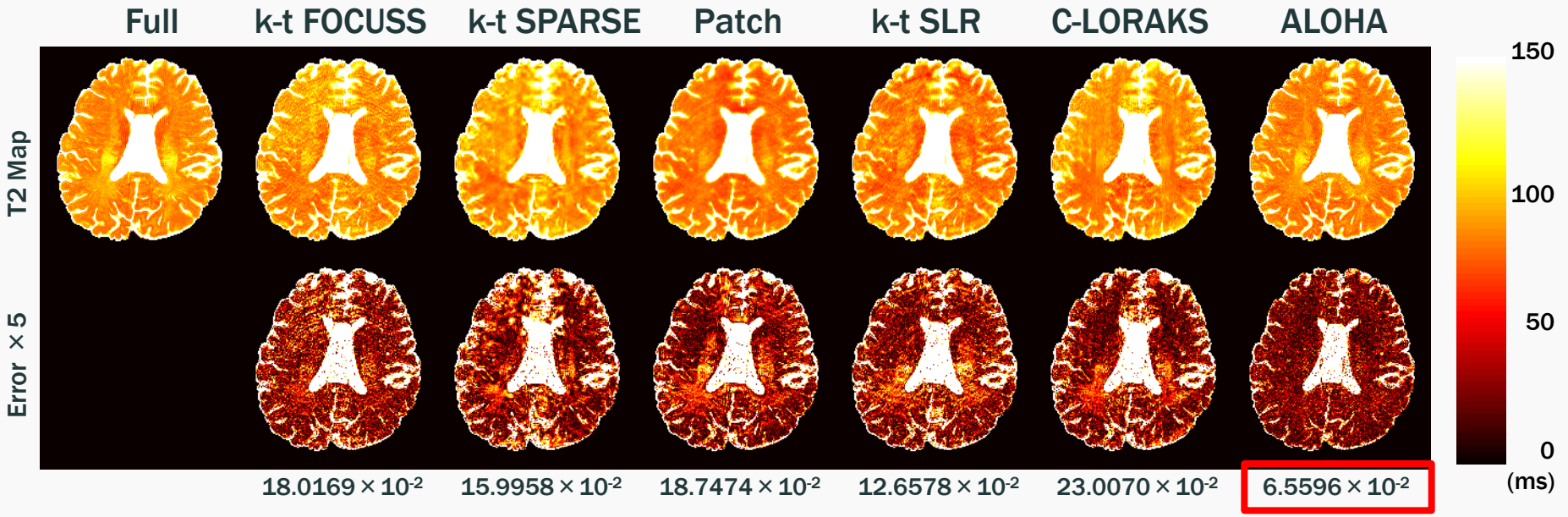
Result : in vivo acceleration study (ME-SE, T2)



Reconstruction of x12.8 accelerated scan – ME-SE (4th echo)

Lee et al, MRM, 2015

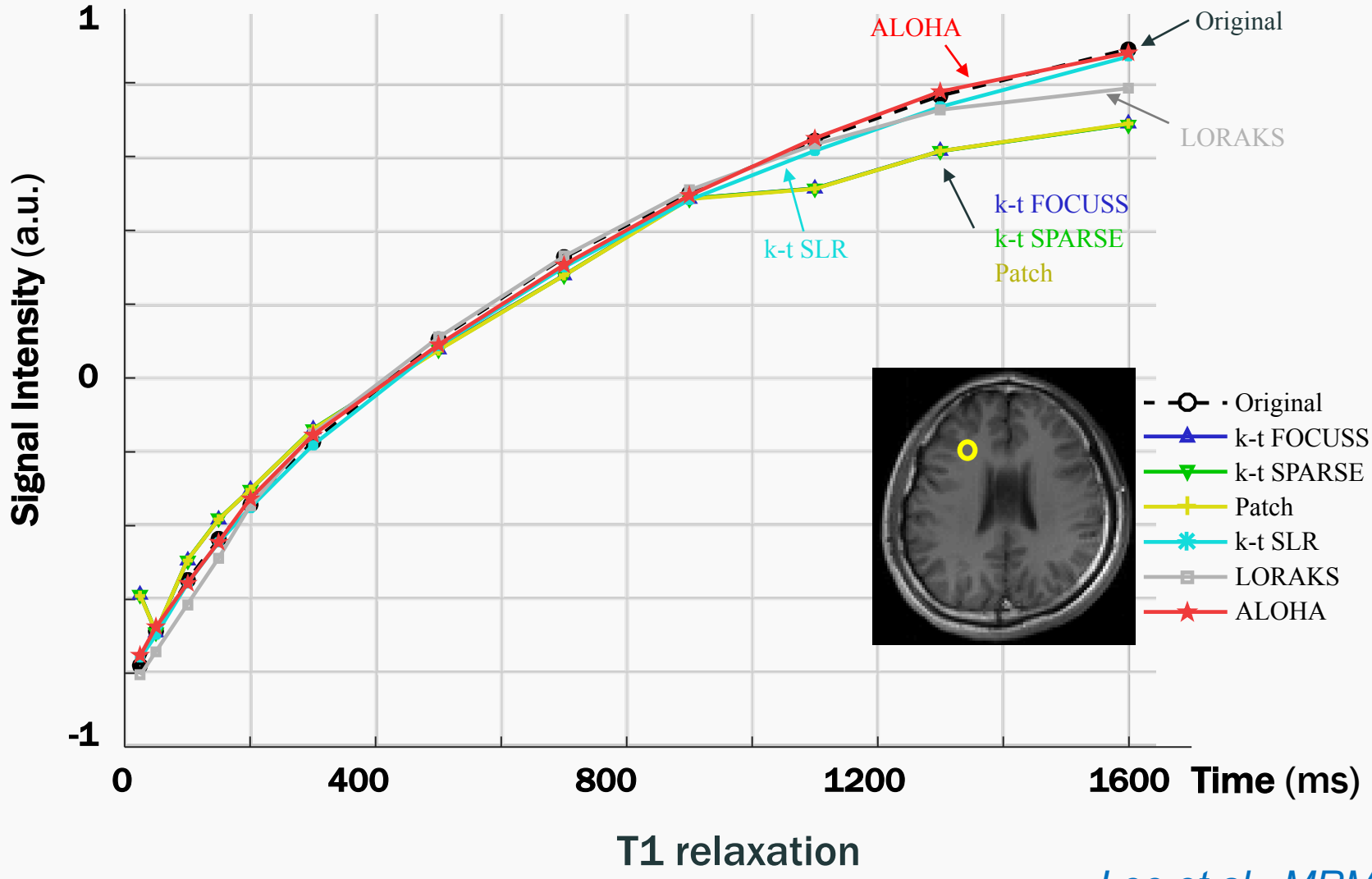
Result : in vivo acceleration study (ME-SE, T2)



Mapping from the reconstruction of x12.8 accelerated scan – T2 mapping

Lee et al, MRM, 2015

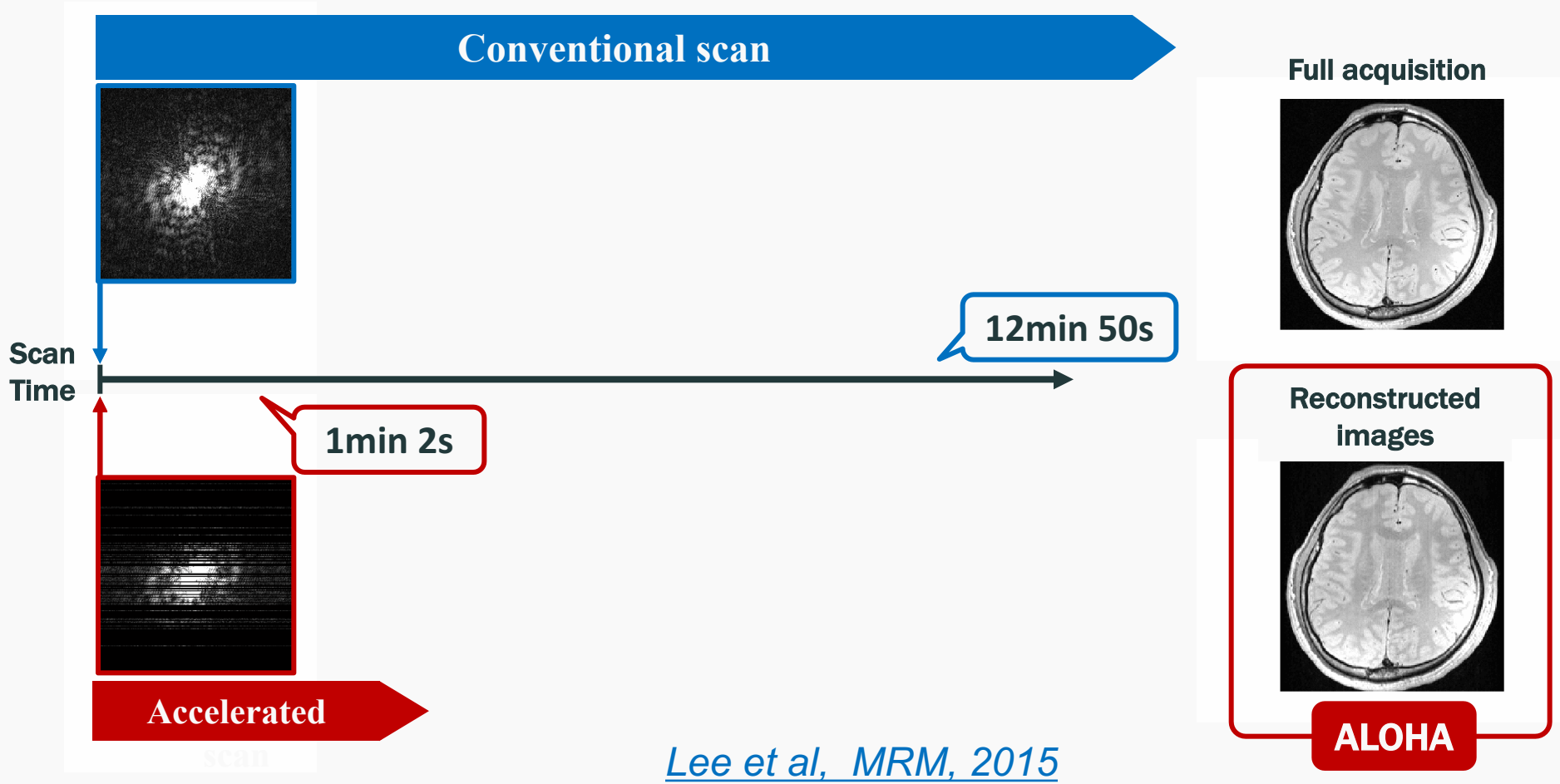
Result : Signal intensity curves (SE-IR, T1)



Lee et al, MRM, 2015

Summary of Results

- **Goal : Acceleration of MR Parameter mapping by undersampling and reconstruction**



Lee et al, MRM, 2015

Extension to 3-D applications using GIRAF

2-D

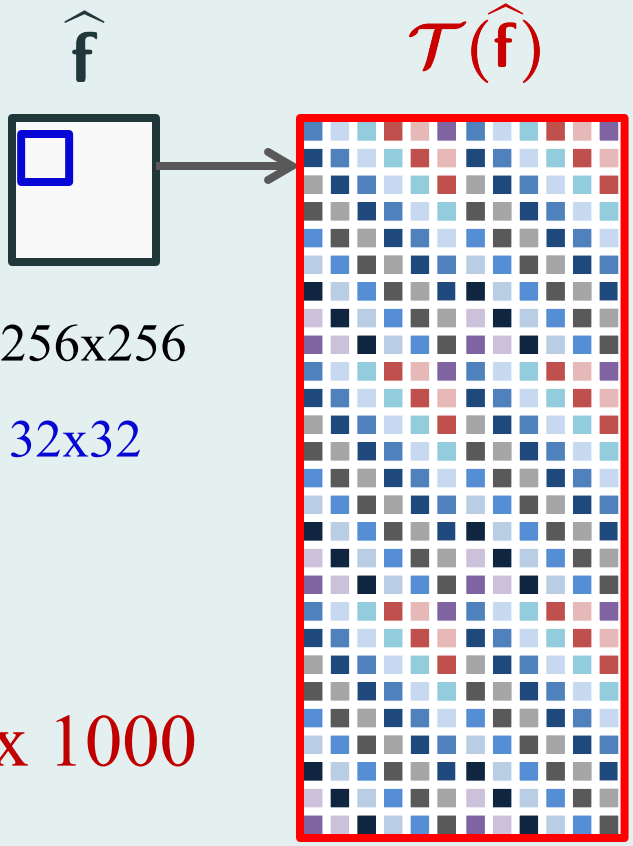
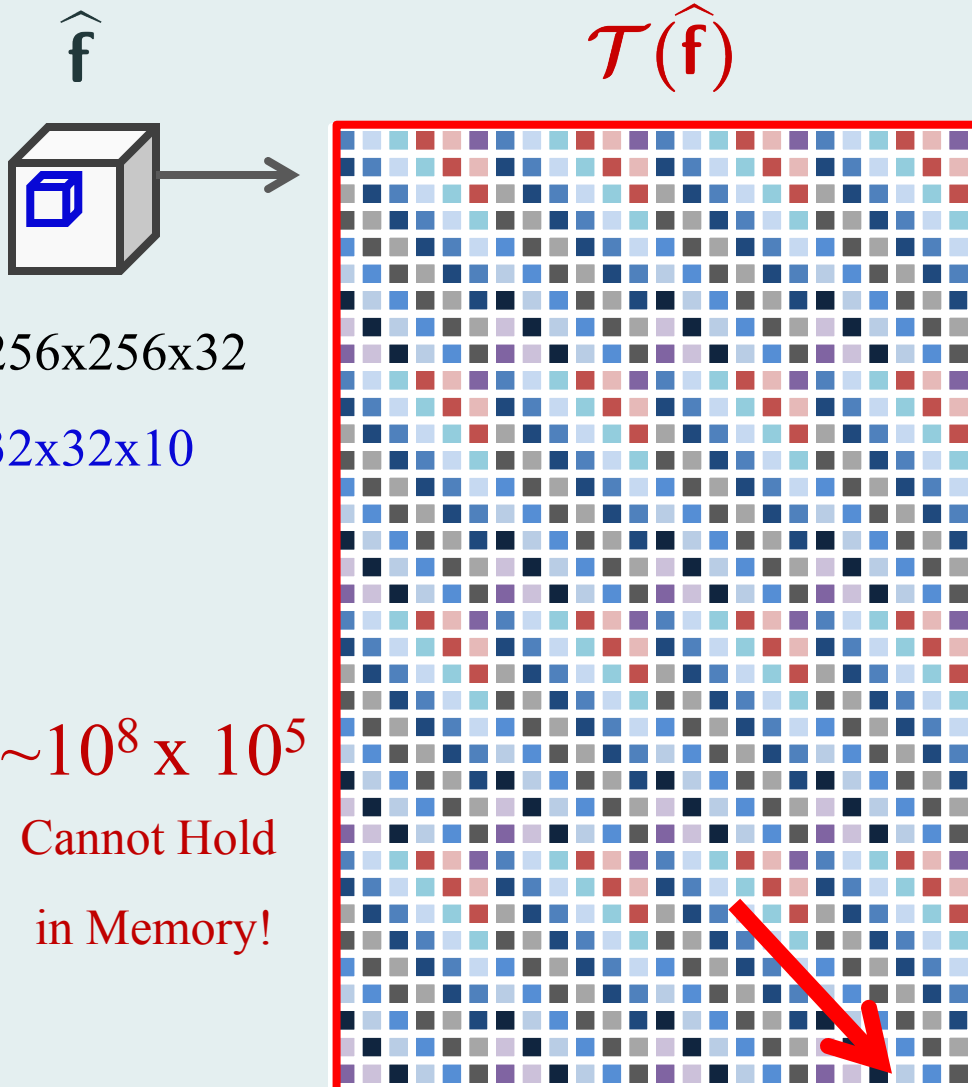


Image: 256x256

Filter: 32x32

$\sim 10^6 \times 1000$

3-D



256x256x32

32x32x10

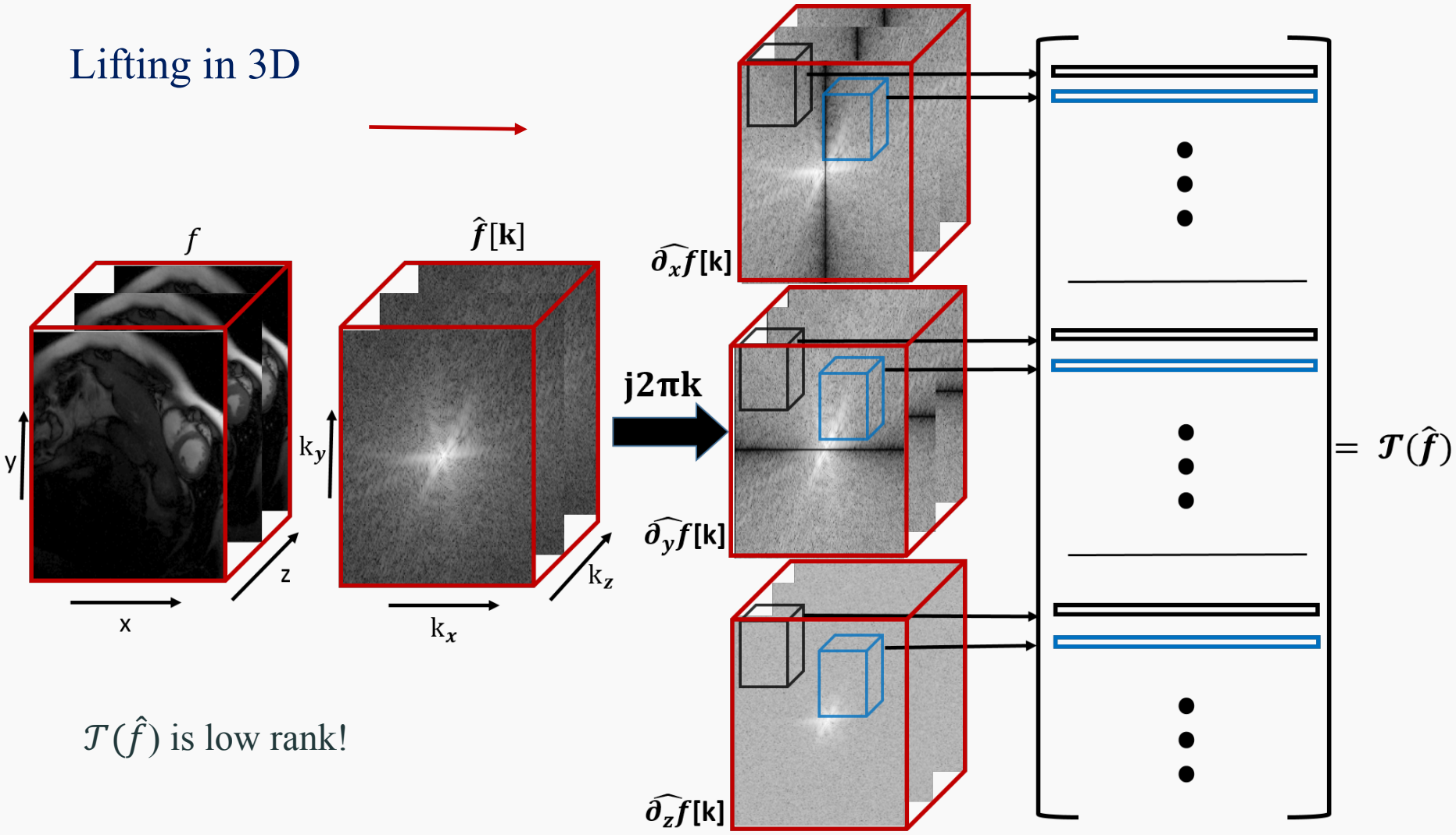
$\sim 10^8 \times 10^5$

Cannot Hold

in Memory!

3D applications: dynamic MRI

Lifting in 3D

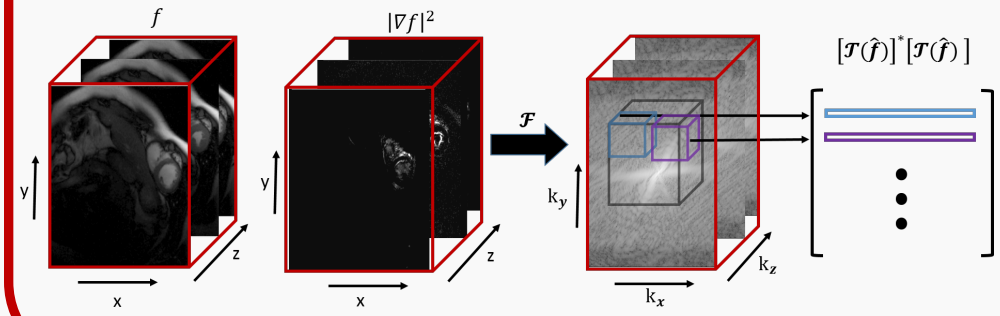


$\mathcal{T}(\hat{f})$ is low rank!

Fast 3-D implementation using GIRAF

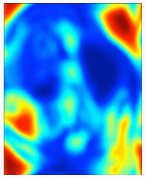
Weight Update:

Gram matrix: 2 FFTs and no matrix product.



Square root - one small eigen decomposition

Fourier data update:



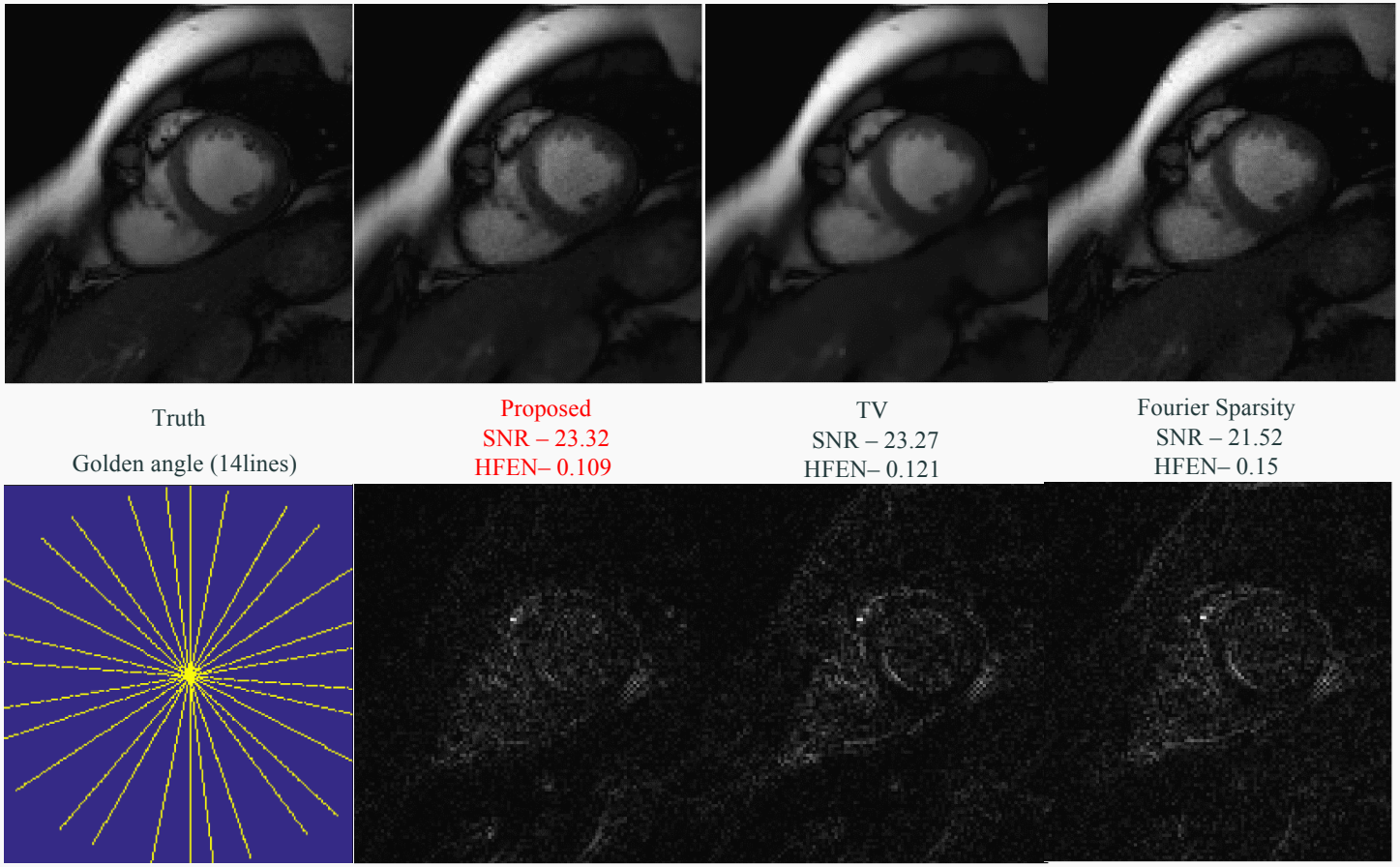
1 frame of μ_{sos}

$\|\mathcal{T}(\mathcal{Q}\mathcal{F}_t\hat{\mathbf{f}})\mathbf{W}^{\frac{1}{2}}\|_F^2 \rightarrow$ simplifies to $\|\mathcal{Q}(\mathcal{F}_t\hat{\mathbf{f}}) \star \hat{\mu}_{sos}\|_F^2$

Where, $\boldsymbol{\mu_{sos}} = \sqrt{\sum_{i=1}^{N} |\mathcal{F}^{-1}(\mathbf{w_i})|^2}$

Need few iterations of CG to solve.

Cardiac CINE MRI

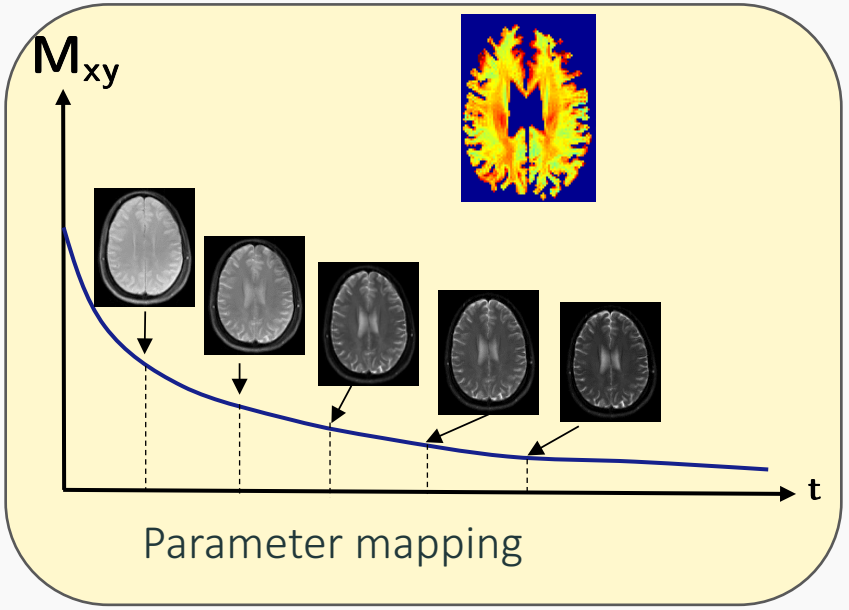


Balachandrasekaran & Jacob, ICIP 16

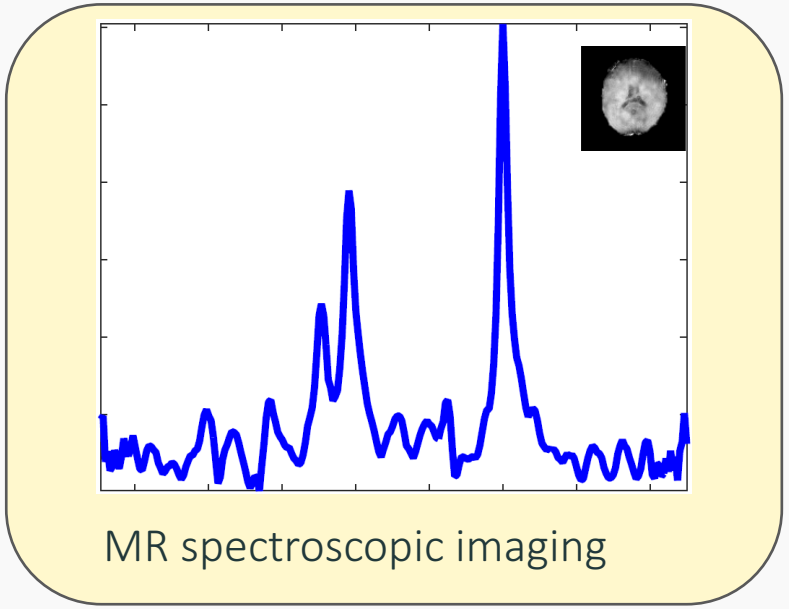
Exponential signal model

Linear combination of exponentials

$$\rho[\mathbf{r}, \mathbf{n}] = \sum_{\mathbf{i}=1}^{\mathbf{L}} \alpha_{\mathbf{i}}(\mathbf{r})\ \beta_{\mathbf{i}}(\mathbf{r})^{\mathbf{n}}$$



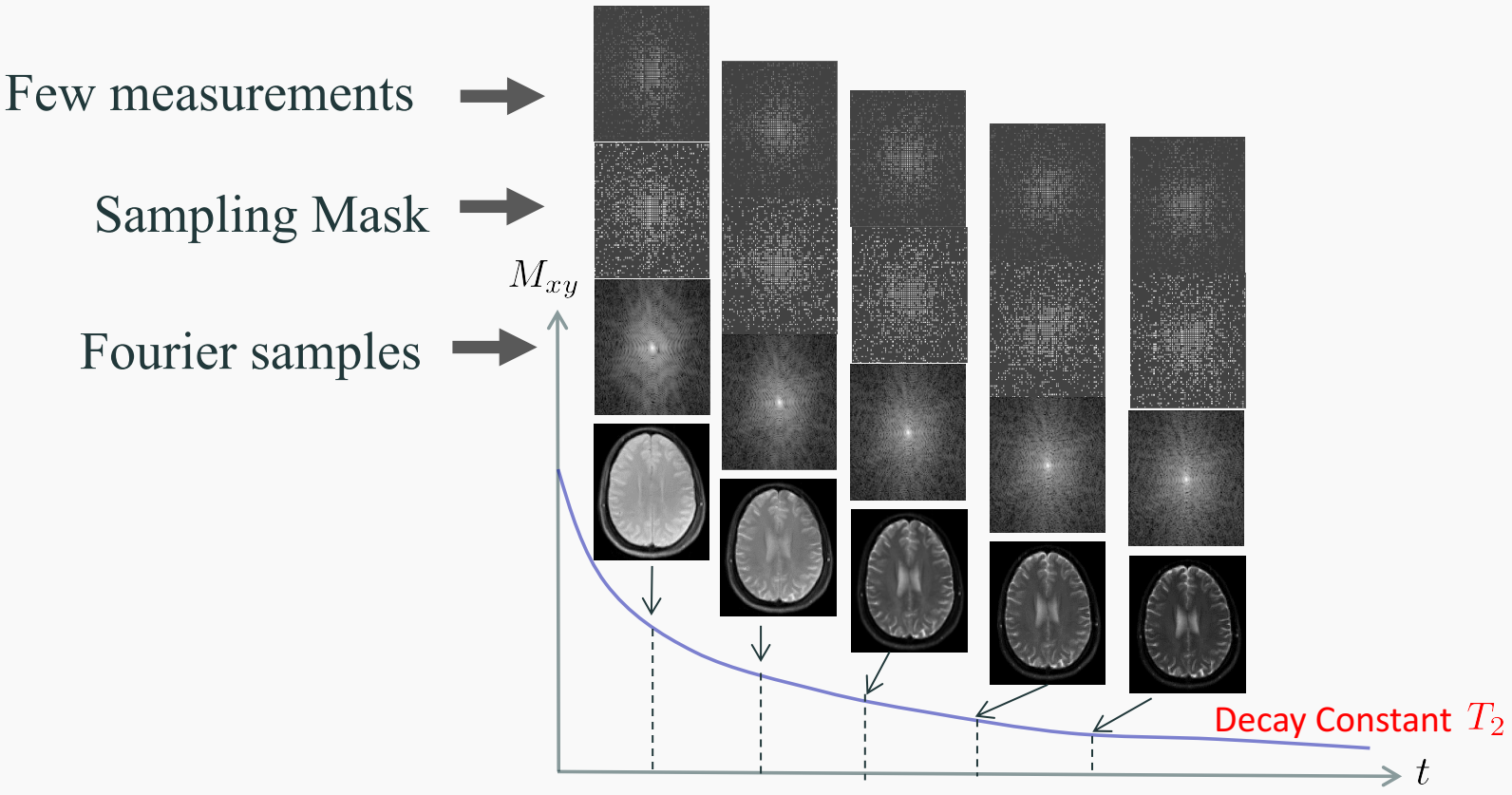
Parameter mapping



MR spectroscopic imaging

Acceleration the acquisition of exponentials

$$\rho[\mathbf{r}, \mathbf{n}] = \sum_{\mathbf{i}=1}^{\mathbf{L}} \alpha_{\mathbf{i}}(\mathbf{r})\ \beta_{\mathbf{i}}(\mathbf{r})^{\mathbf{n}}$$



MR parameter mapping

State of the art

Exploit correlations between voxel profiles

- Low rank methods [Doneva et al,..]
- Dictionary learning methods [BCS;Bhave et al]

Exploit correlations within an exponential voxel profile

- Pixel by pixel structured low-rank [MORASSA: Peng et al,2016]

Structured low-rank with wavelet models

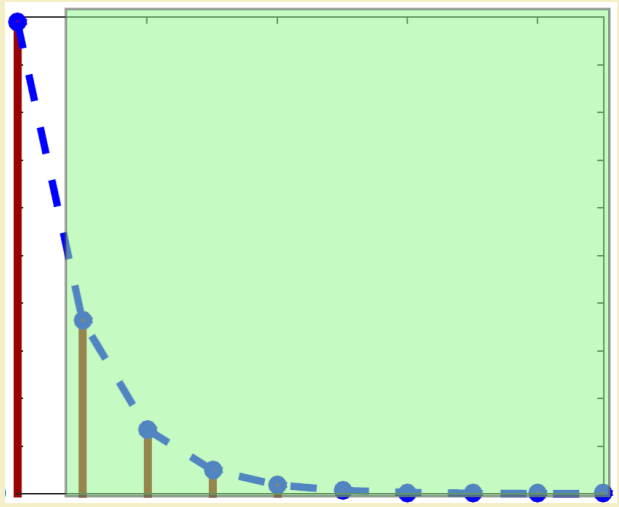
- Exploit sparsity of kx-t planes in wavelet domain [ALOHA]

Unify above strategies !!

1-D signal satisfies an annihilation relation

$$y[n] = \sum_{m=0}^{L} \rho[n-m]h[m] = 0$$

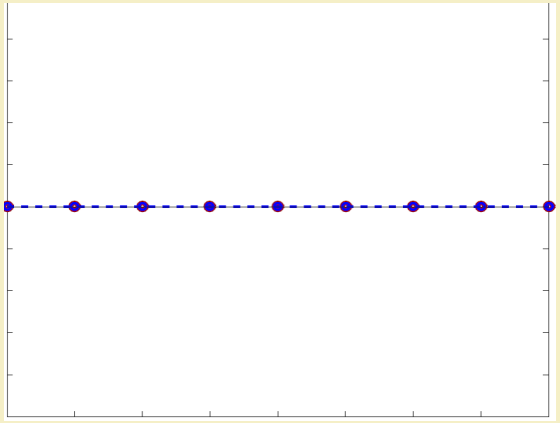
$\rho[n] = \alpha^n$



Possible shifts of filter

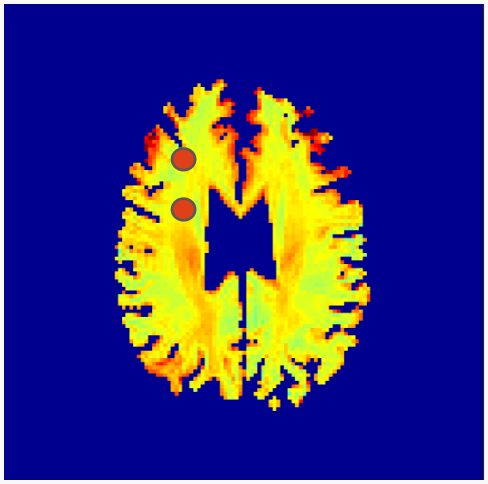
$* \; h[n] = [1, -\alpha^{-1}] \quad =$

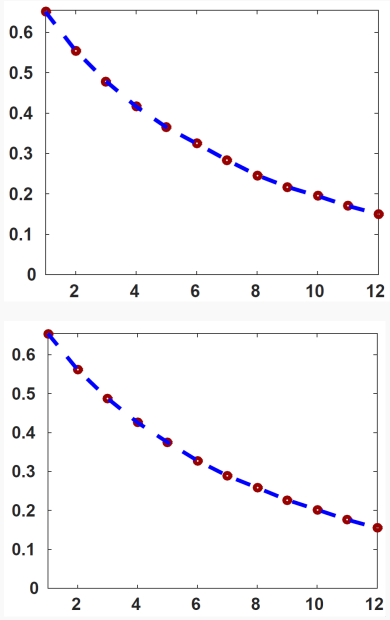
$y[n] = 0$

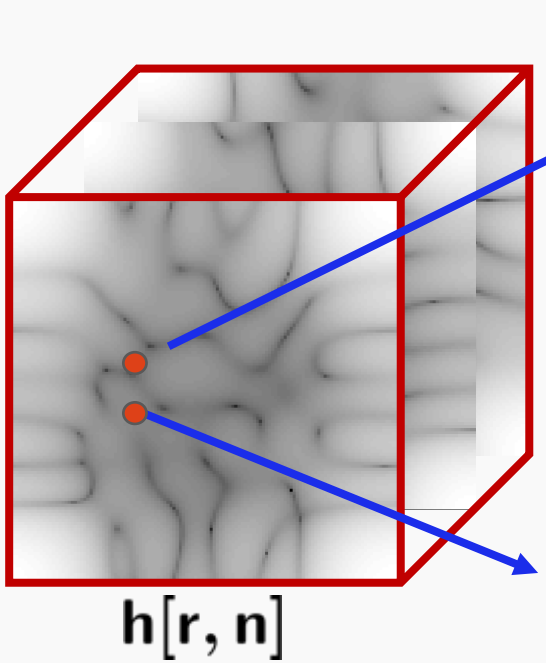


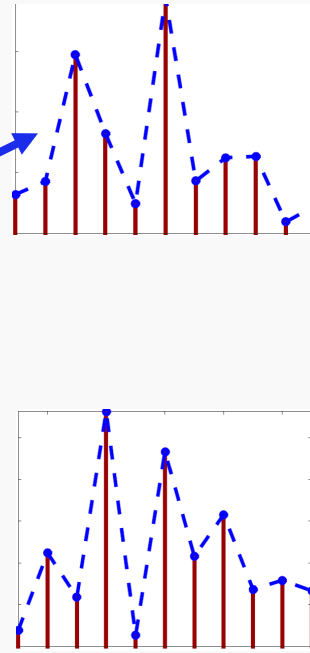
Exponential parameters are spatially smooth

Filter coefficients are spatially smooth

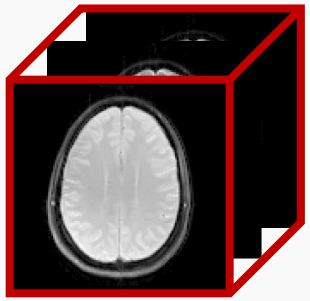








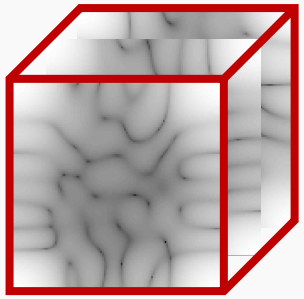
Spatially smooth filters: FIR in Fourier domain



$\rho[\mathbf{r}, \mathbf{n}]$

1-D

$*$

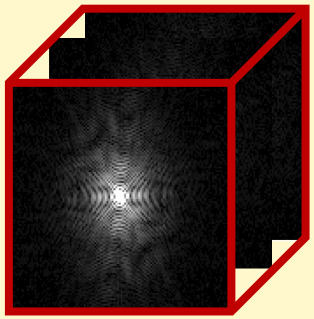


$\mathbf{h}[\mathbf{r}, \mathbf{n}]$

$= \mathbf{0}$

$\mathcal{F}_{\mathbf{2D}}$

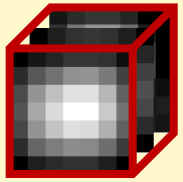
Fourier domain



$\widehat{\rho}[\mathbf{k}, \mathbf{n}]$

3-D

$\otimes$

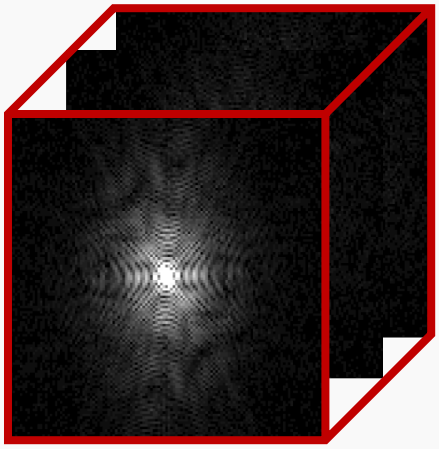


$\mathbf{d}[\mathbf{k}, \mathbf{n}]$

$= \mathbf{0}$

Annihilation relation: $\widehat{\rho}[\mathbf{k}, \mathbf{n}] \otimes \mathbf{d}[\mathbf{k}, \mathbf{n}] = \mathbf{0}$

Bandwidth of filter & spatial smoothness

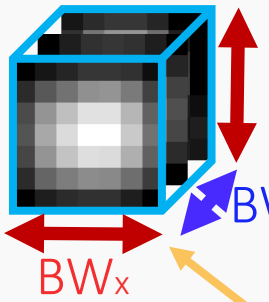


$\rho[\mathbf{k}, \mathbf{n}]$

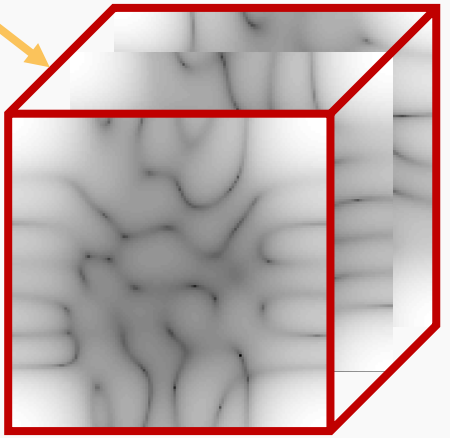
$\otimes$

3-D

$\mathbf{d}[\mathbf{k}, \mathbf{n}]$



$= \mathbf{0}$

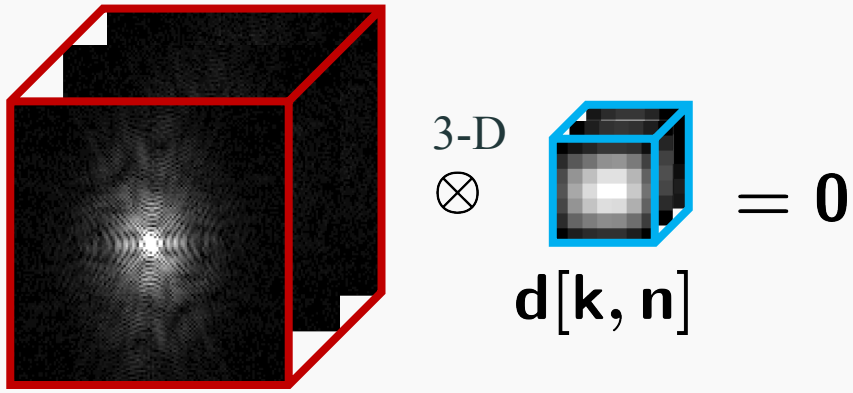


$\mathbf{h}[\mathbf{r}, \mathbf{n}]$

BW_x & BW_y: spatial smoothness of parameters.

BW_p: number of exponentials.

Annihilation as a matrix-vector product



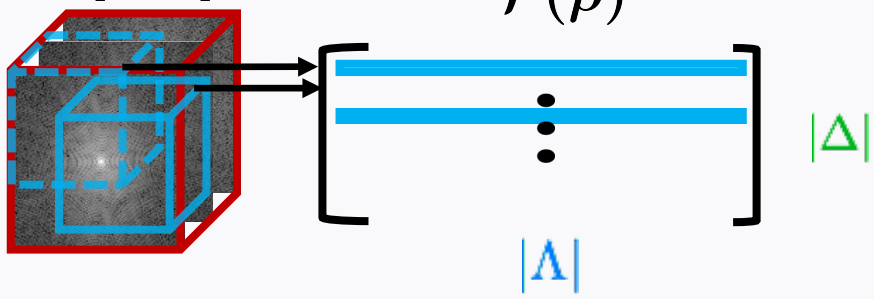
3-D ⊗

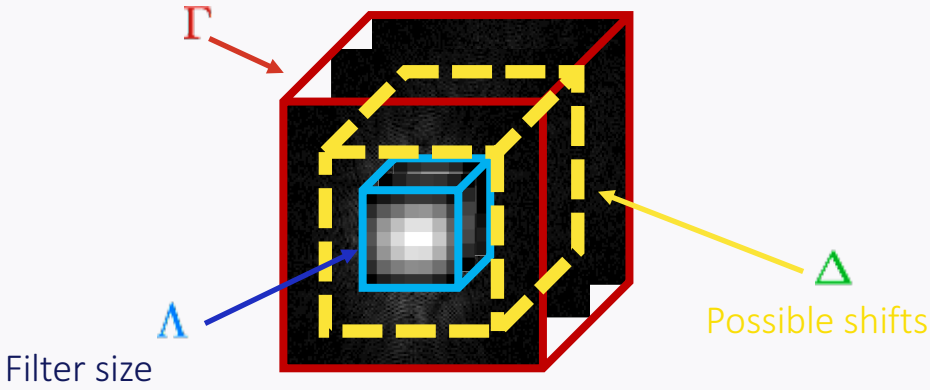
$= 0$

$\mathbf{d}[\mathbf{k}, \mathbf{n}]$

$\hat{\rho}[\mathbf{k}, \mathbf{n}]$

$\mathcal{T}(\hat{\rho})\ \mathbf{d} = \mathbf{0}$



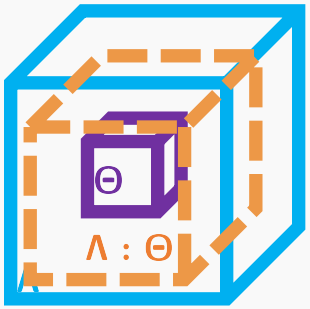


Minimal filter & assumed filter size

Minimal filter: exact size is unknown

Assumed filter: larger than minimal filter

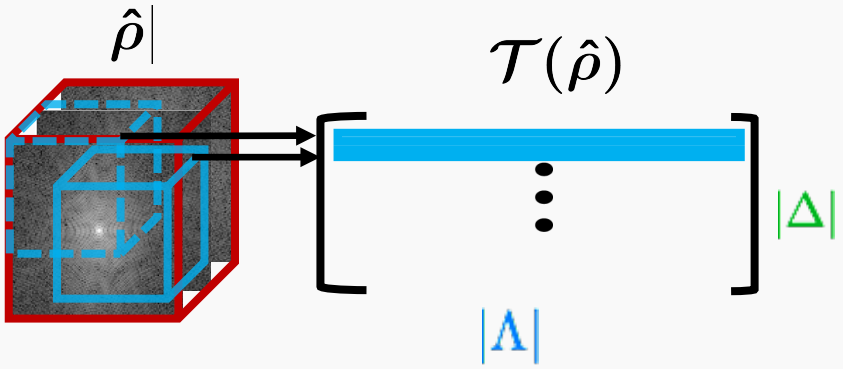
⇨ Several possible annihilating filters

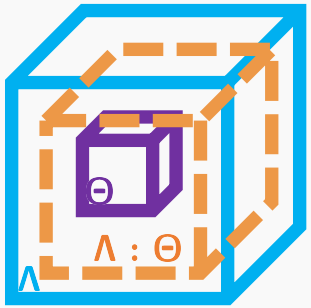


$$\underbrace{d[k,n]}_{\text{Assumed filter}} = \underbrace{c[k,n]}_{\text{Minimal filter}} \otimes \underbrace{e[k,n]}_{\text{FIR filter}}$$

Dimension of annihilating subspace $\geq |\Lambda : \Theta|$

Rank of the matrix





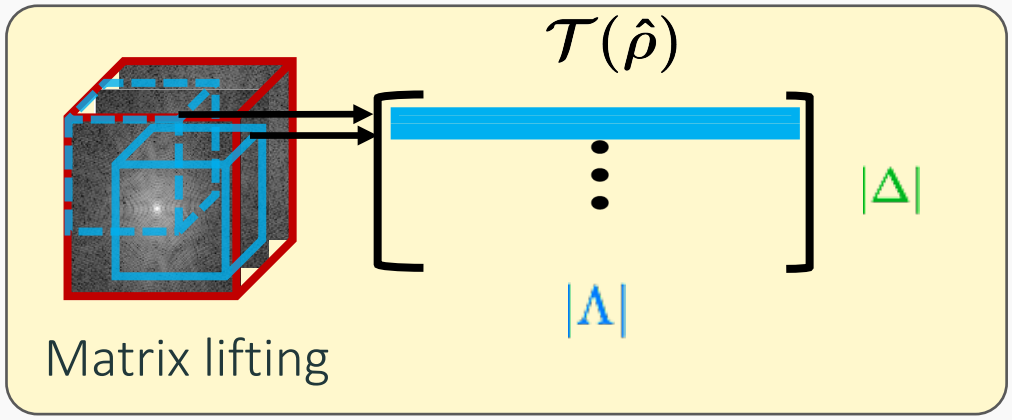
Dimension of annihilating subspace $\geq |\Lambda : \Theta|$

$$\mathbf{Rank}(\mathcal{T}(\hat{\rho})) \leq |\boldsymbol{\Lambda}| - |\boldsymbol{\Lambda} : \boldsymbol{\Theta}|$$

Structured low-rank optimization problem

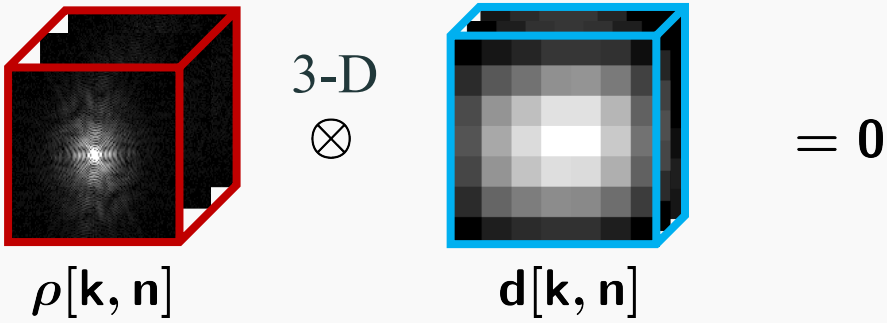
Find the signal that satisfies data consistency & minimizes rank

$$\hat{\rho}^{\star} = \arg\min_{\hat{\rho}} \|\mathcal{T}(\hat{\rho})\|_{\mathbf{p}} + \frac{\mu}{2}\|\mathcal{A}(\hat{\rho}) - \mathbf{b}\|_2^2$$



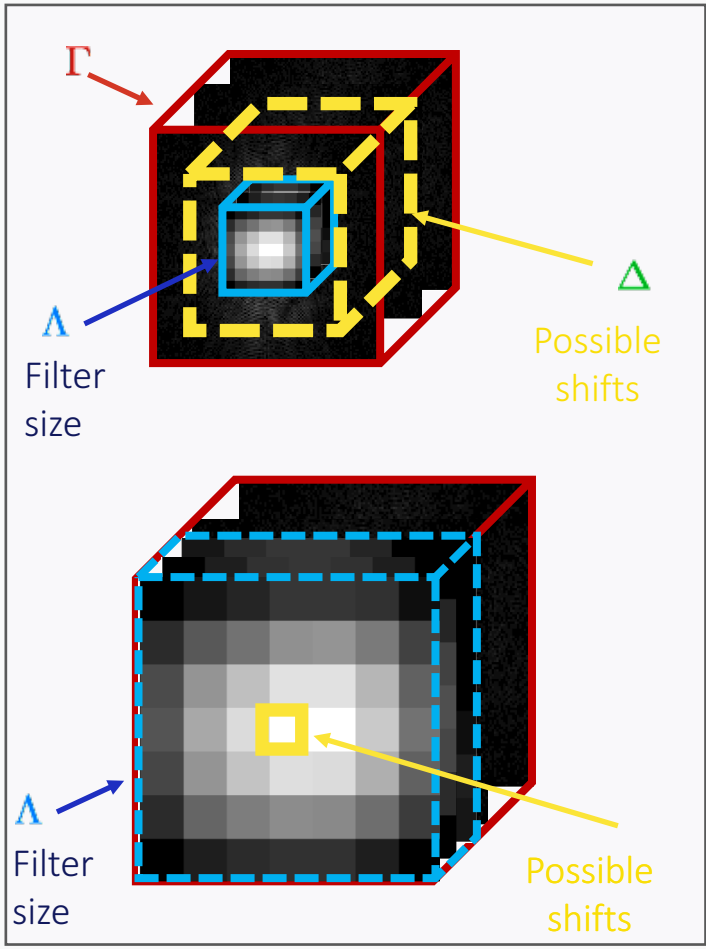
Special case: no spatial smoothness

Spatial dimension of filter is the same as Γ



Lifting: concatenation of pixel-by pixel Toeplitz matrices

$$\mathcal{T}(\rho) = [\mathcal{T}(\rho_1)|\ldots|\mathcal{T}(\rho_N)]$$



Relation to other structured low-rank priors

Pixel independent structured low-rank prior [MORASSA,Peng et al.,2016]

$$\{\rho_{\mathsf{m}}\} = \arg\min_{\{\rho_{\mathsf{m}}\}} \sum_{\mathsf{m}} \|\mathcal{T}(\rho_{\mathsf{m}})\|_* + \|\mathcal{A}(\boldsymbol{\rho}) - \mathsf{b}\|^2$$

Does not exploit spatial correlations

Proposed special case: nuclear norm of concatenated matrices

$$\mathcal{T}(\rho) = [\mathcal{T}(\rho_1)|\ldots|\mathcal{T}(\rho_N)]$$

Combination of low-rank & exponential priors: exploit spatial correlations

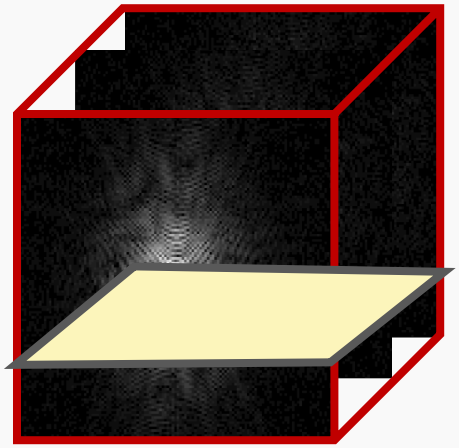
Only one small EVD per iteration: considerably faster

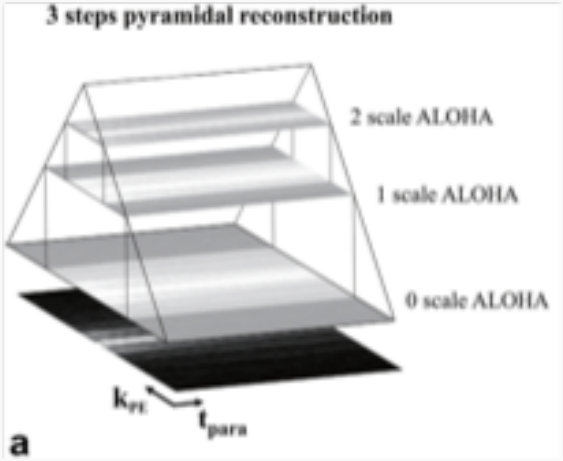
ALOHA based solution discussed earlier ..

2-D prior: fills in kx-t planes using structured low-rank interpolation

Signal sparse in wavelet domain

Does not exploit exponential decay of signal

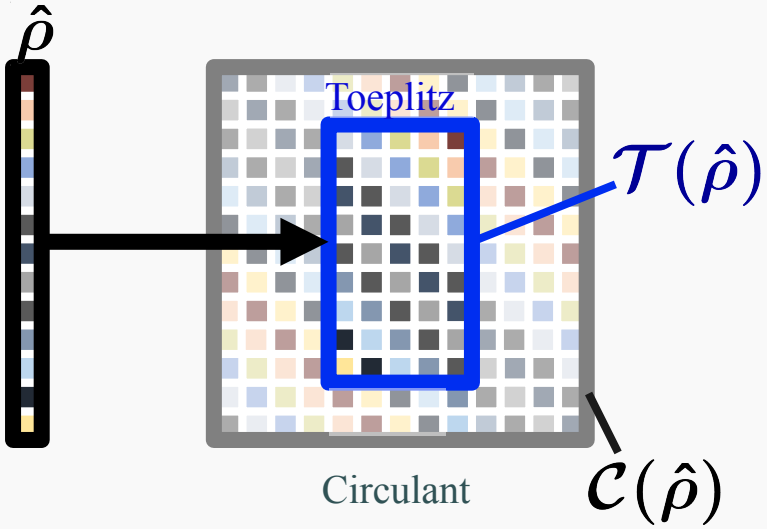


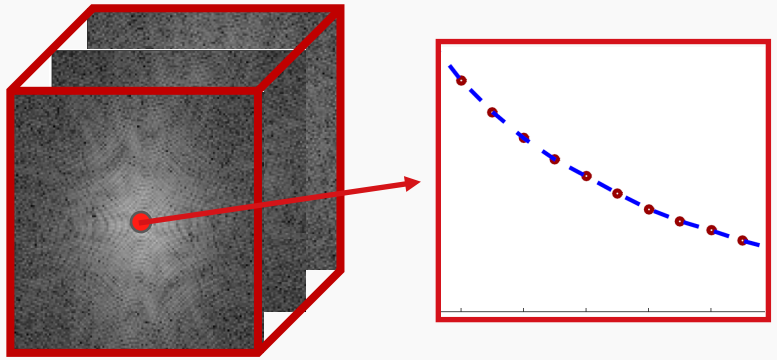


Use signal weighting: ALOHA-like prior

GIRAF: Circulant approximation is not accurate

Approximation is good if signal amplitude is negligible at boundaries





High signal amplitude: poor approximation

Solution: hybrid approach

Circular convolution in space

Linear convolution along parameter dimension

$$\mathbf{w}[\mathbf{k},\mathbf{n}] \otimes \hat{\boldsymbol{\rho}}[\mathbf{k},\mathbf{n}] = \sum_{\mathbf{m}} \underbrace{\sum_{p} \hat{\rho}[\mathbf{k}-\mathbf{p}, n-m] h[\mathbf{p}, m]}_{\approx\ \mathbf{g}_{\mathbf{n}-\mathbf{m},\mathbf{m}} = \hat{\boldsymbol{\rho}}_{\mathbf{n}-\mathbf{m}}\ *\ \mathbf{h}_{\mathbf{m}}}$$

Sum of spatial circular convolutions: computed using FFT

Weight evaluation

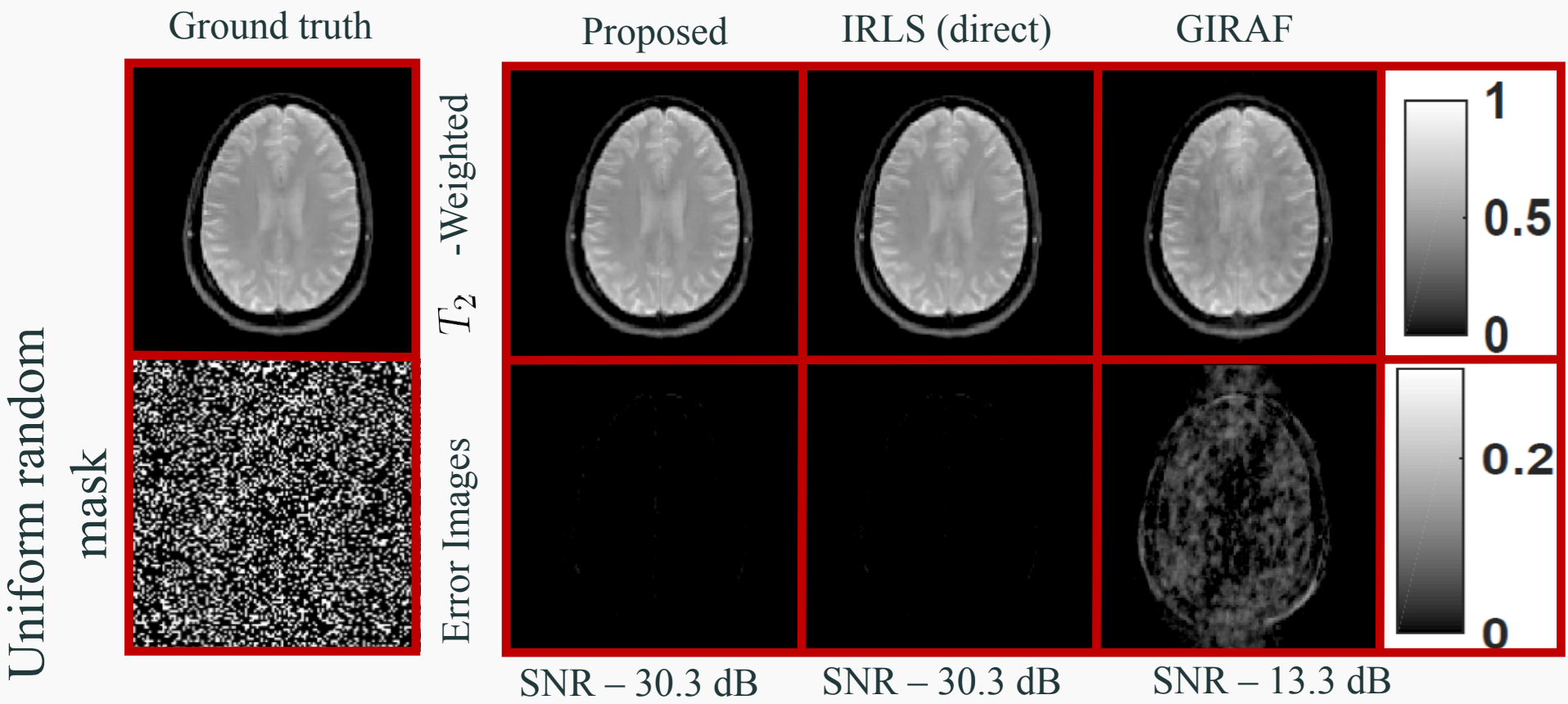
$$\mathbf{W} \leftarrow \left(\underbrace{\mathcal{T}(\hat{\rho})\mathcal{T}(\hat{\rho})^{H}}_{\mathbf{R}} + \epsilon \mathbf{I} \right)^{\frac{\mathbf{p}}{2}-1}$$

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_{1,1} & \mathbf{R}_{1,2} & \cdots & \mathbf{R}_{1,M} \\ \mathbf{R}_{2,1} & \mathbf{R}_{2,2} & \cdots & \mathbf{R}_{2,M} \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{R}_{M,1} & \mathbf{R}_{M,2} & \cdots & \mathbf{R}_{M,M} \end{pmatrix}$$

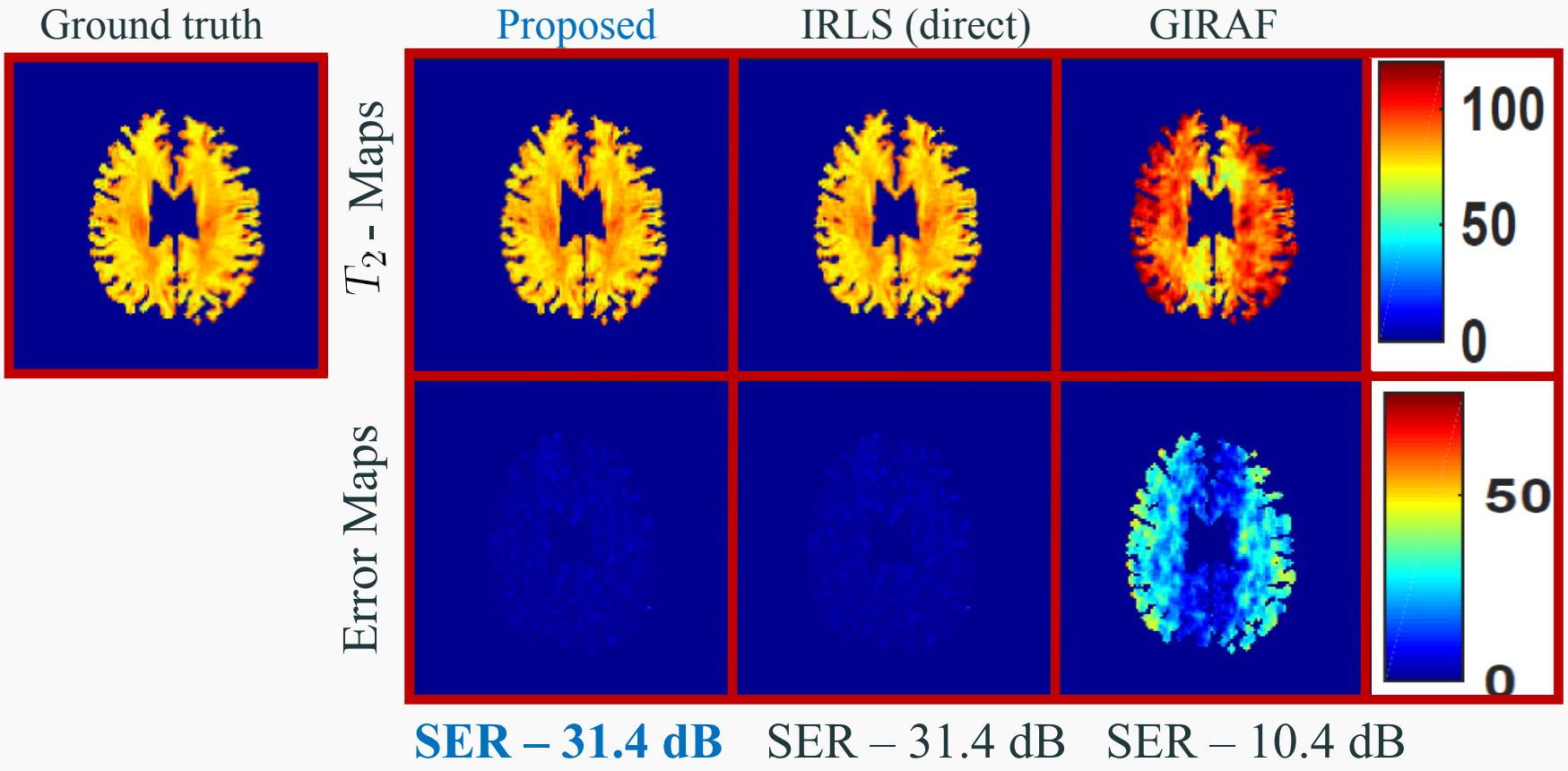
Sum of spatial circular convolutions: computed using FFT

$$\mathbf{R_{p,q}} = \sum_{\mathbf{i=1}}^{\mathbf{k}} \underbrace{\mathbf{T}(\hat{\rho}_{p+i-1})\mathbf{T}(\hat{\rho}_{q+i-1})^{*}}$$

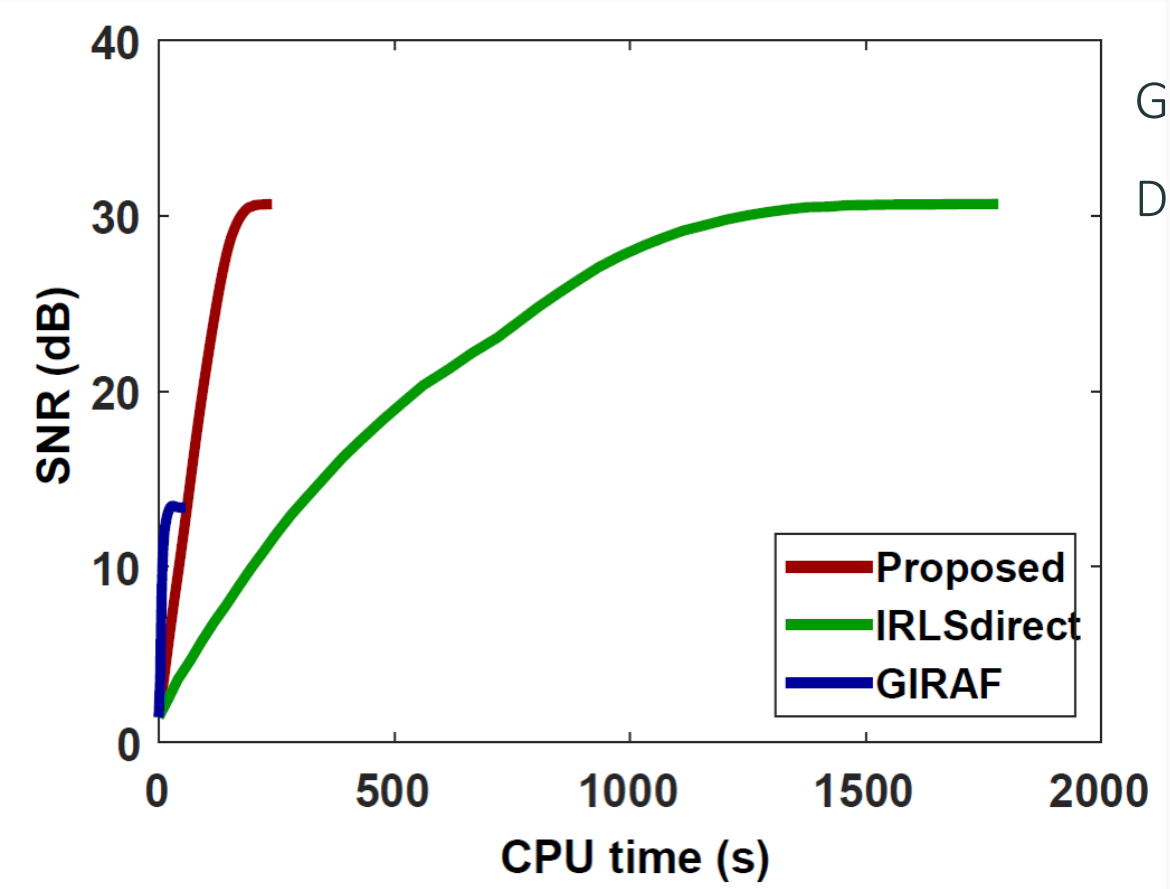
Validation using single channel data (sim)



Validation using single channel data (simulation)



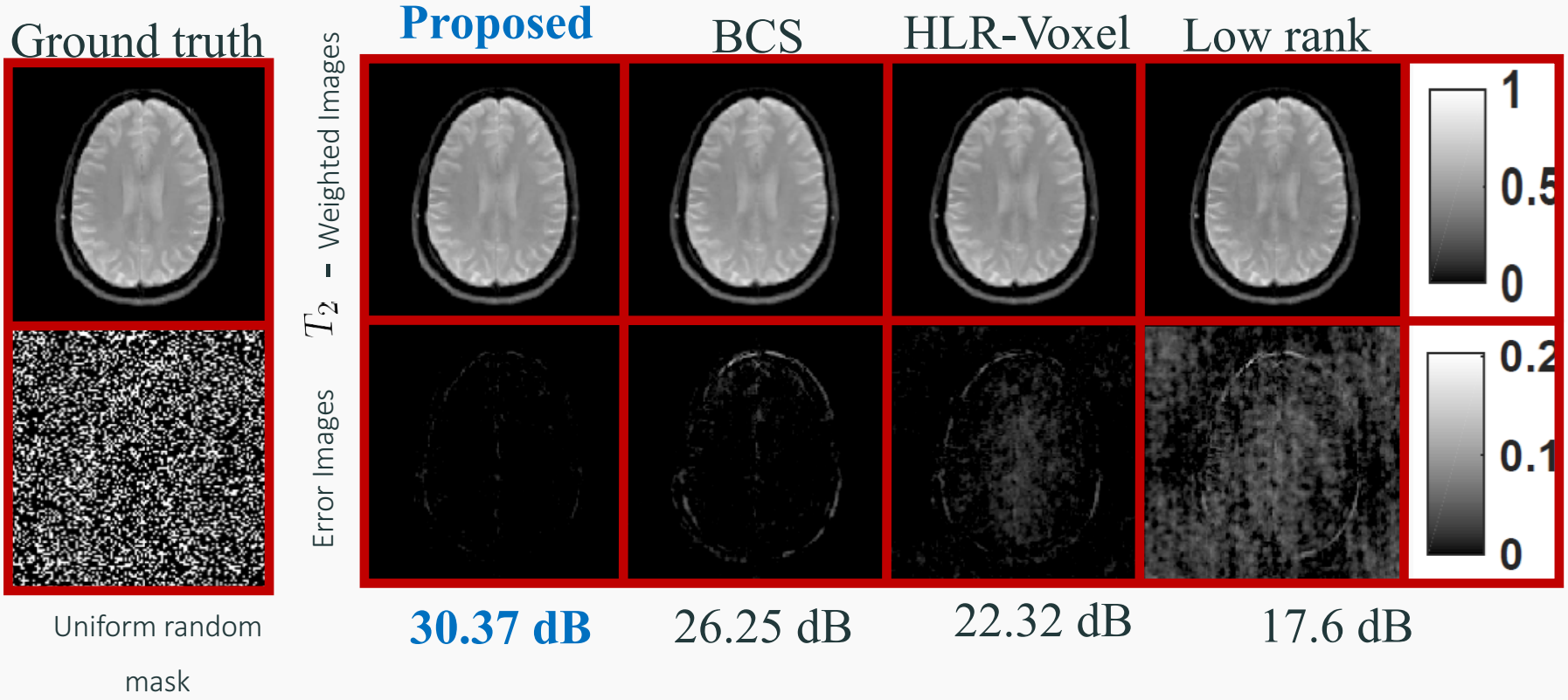
Fast convergence



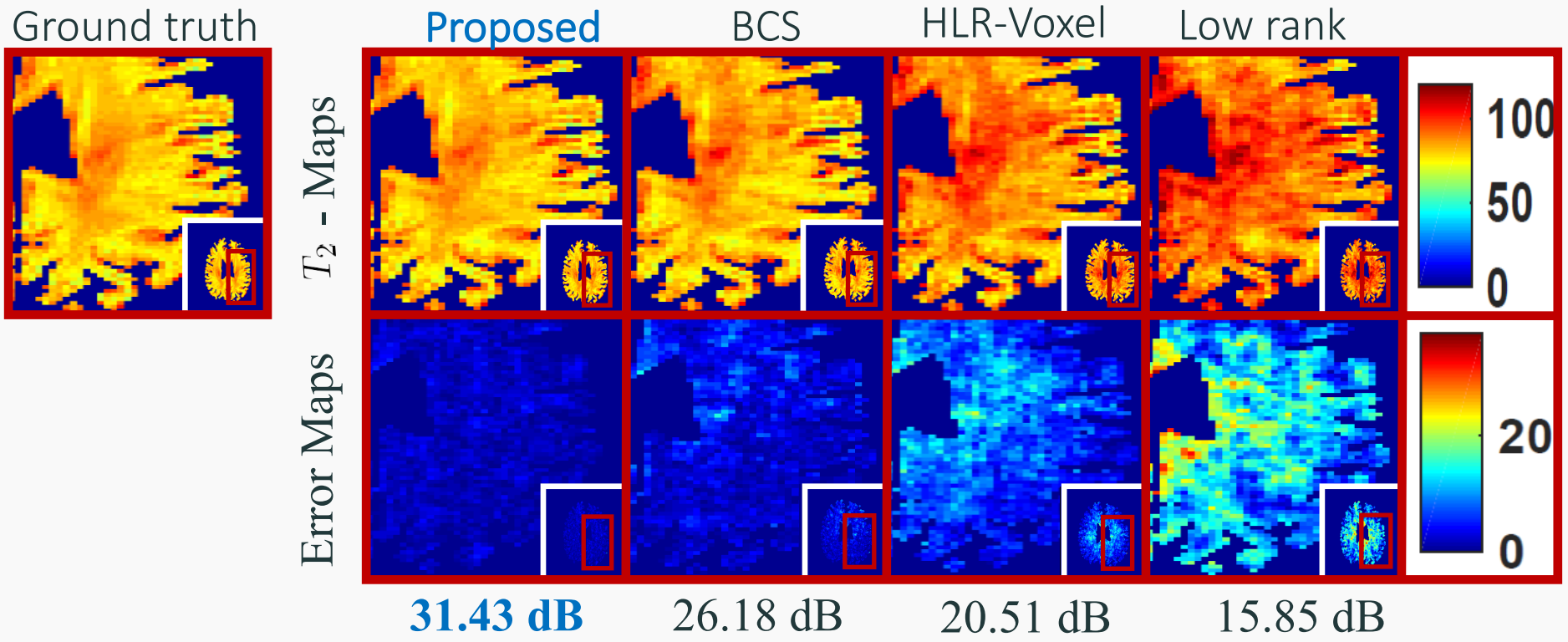
GIRAF: poor approximation

Direct IRLS: slow

Comparison with state of the art



Comparison with state of the art



Multichannel acquisitions

- TR=2500 ms,
- slice thickness = 5mm
- FOV = 22x22 cm^2
- Matrix size : 128x128,
- No of coils = 12,
- TE = 10 to 120 ms

Smaller filters provide better results

filter size	SNR
128x128x10	28.05
122x122x10	30.30
114x114x10	31.00
108x108x10	31.12
102x102x10	**31.21**
100x100x10	31.20

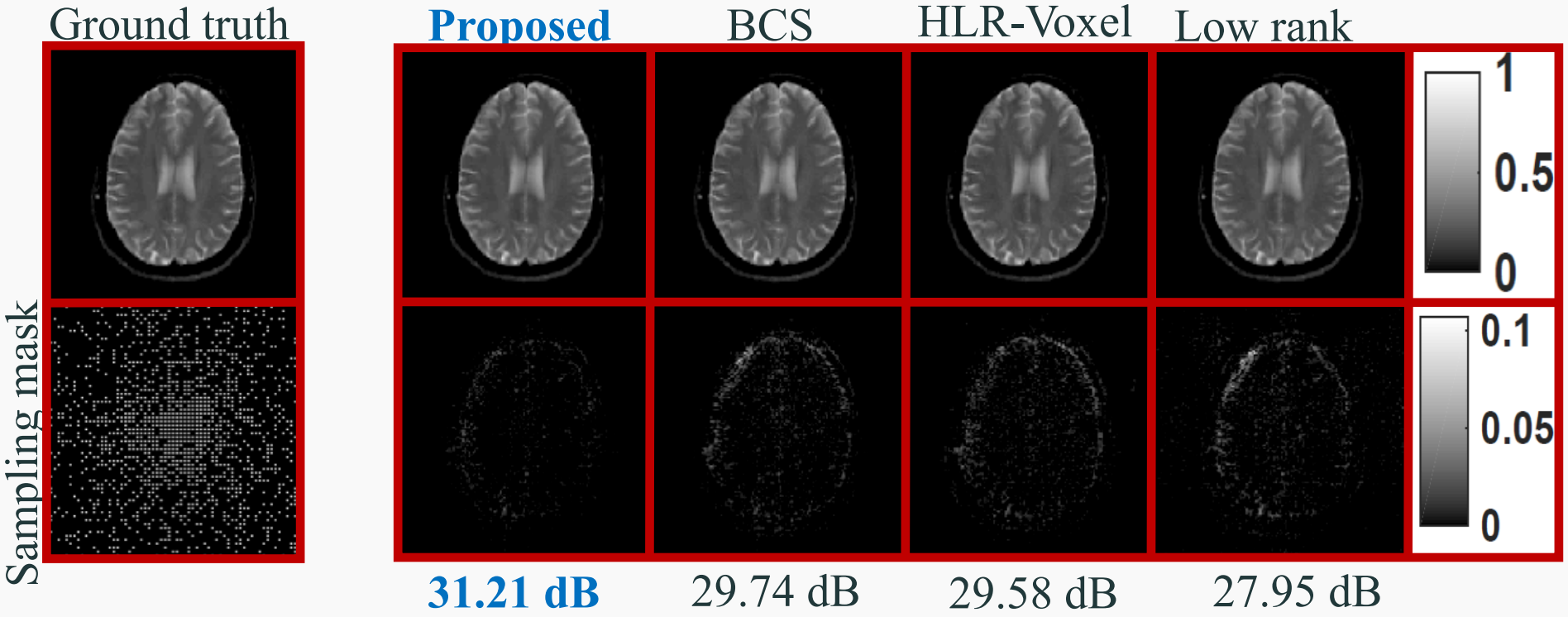
(a) Varying spatial dimension

filter size	SNR (dB)
102x102x11	30.80
102x102x10	**31.21**
102x102x7	31.13
102x102x4	30.96
102x102x2	30.78
102x102x1	29.88

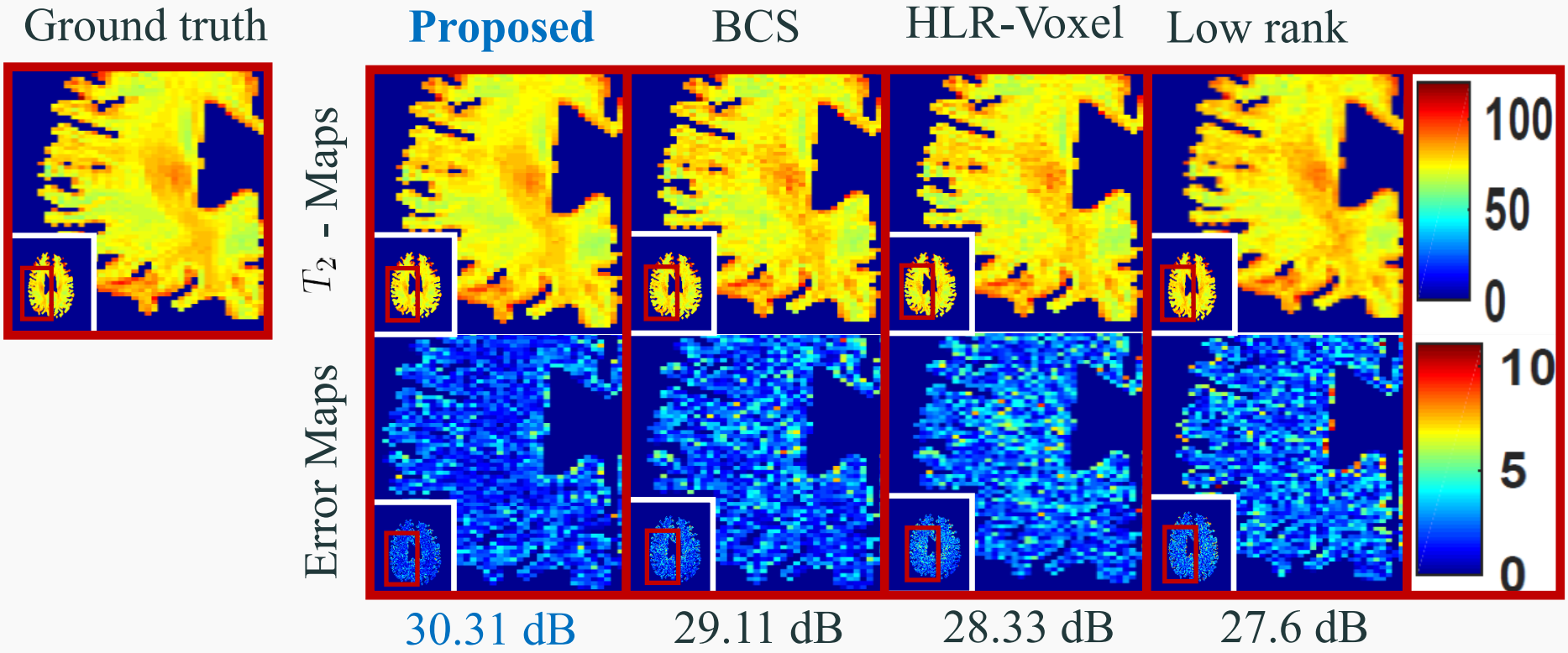
(b) Varying temporal dimensions

Smaller filters: spatial regularization

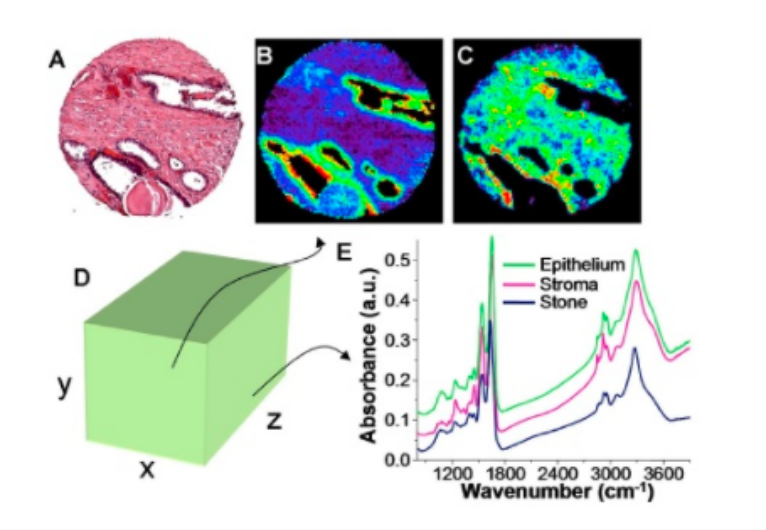
Multichannel acq: T2 weighted images



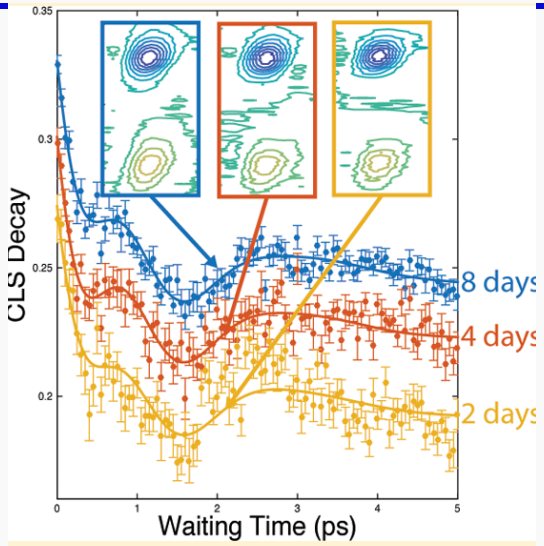
Multichannel acquisition: parameter maps



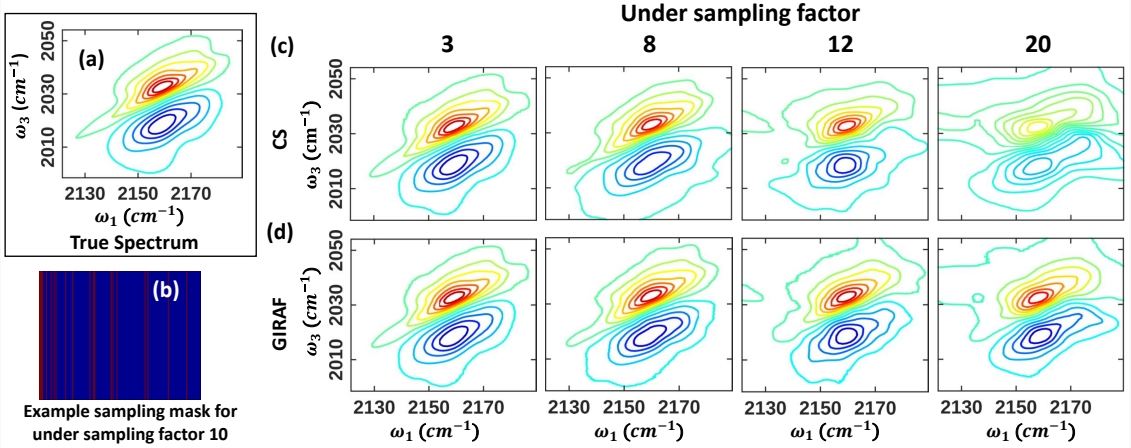
Infrared spectroscopy



1D IR spectroscopy



2D IR spectroscopy

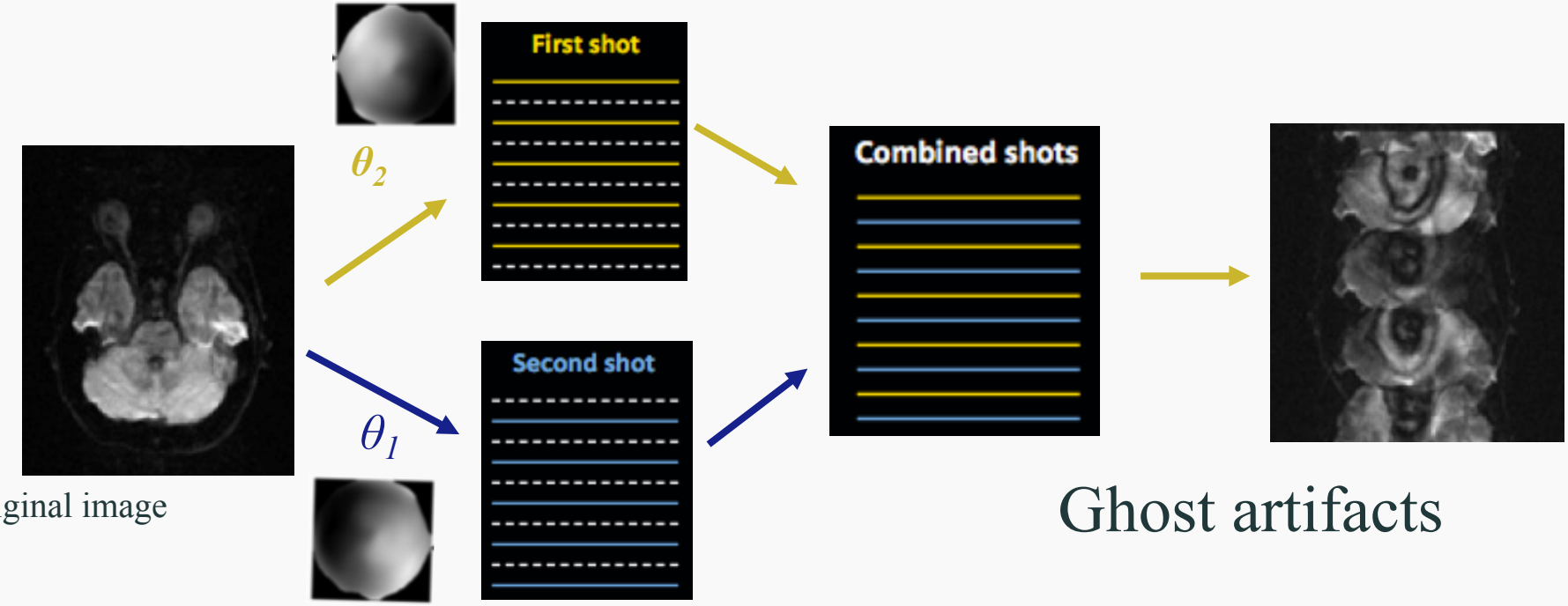


Accelerated imaging using GIRAF

Humston et al, Journal of Physical Chemistry
Bhattacharya et al, Optics Letters, submitted

Correction of Nyquist ghosts in multishot MRI [MUSSELS]

Motion-induced inter-shot phase errors



Self calibration methods: Image domain

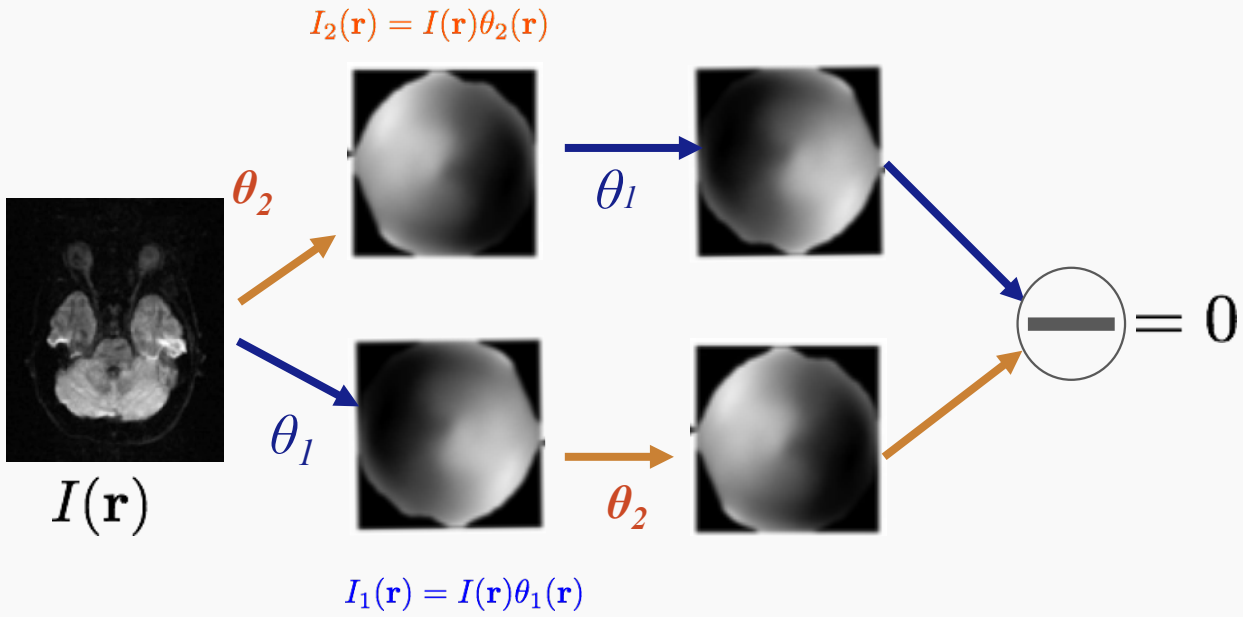


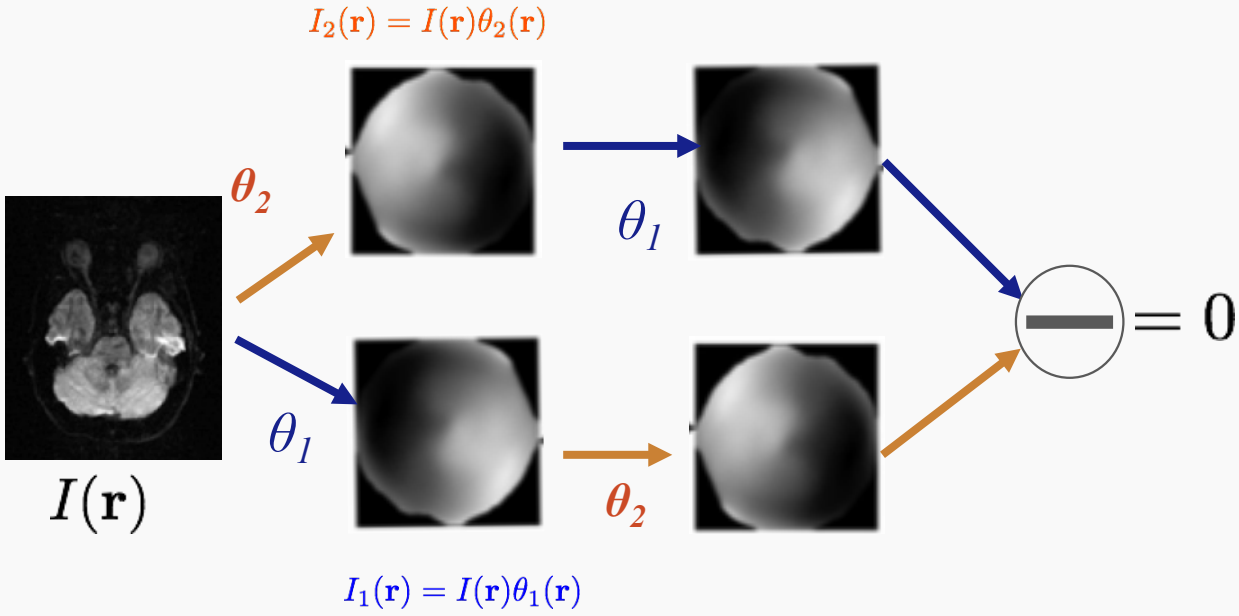
Image domain annihilation relation [Morrisson, Do & Jacob 2007]

$$I_2(\mathbf{r}) \cdot \theta_1(\mathbf{r}) - \hat{I}_1(\mathbf{r}) \cdot \hat{\theta}_2(\mathbf{r}) = 0$$

Model sensitivities as polynomials: EVD

Better than SOS estimates

Self calibration methods: Fourier domain



Fourier domain relation [Lustig 2012, Haldar 2014]

$$\hat{I}_2[\mathbf{k}] * \hat{\theta}_1[\mathbf{k}] - \hat{I}_1[\mathbf{k}] * \hat{\theta}_2[\mathbf{k}] = 0$$

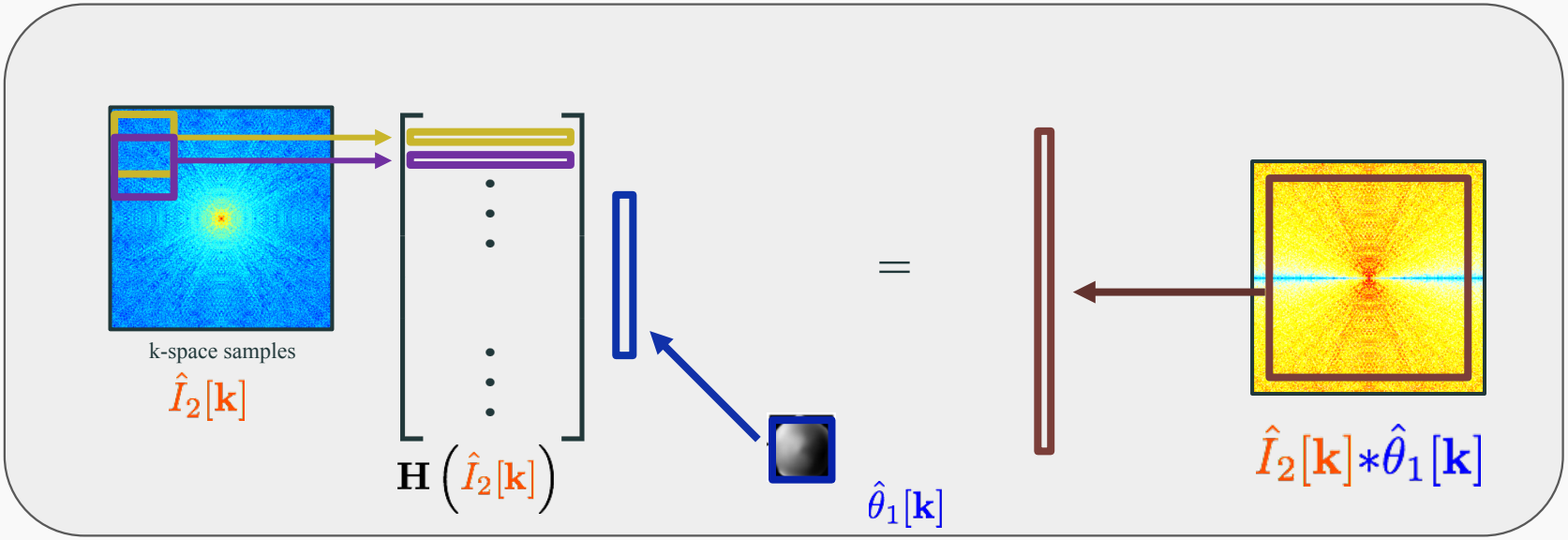
Phase: linear combination of exponentials → FIR filter

Self calibration methods: matrix form

Fourier domain relation

$$\hat{I}_2[\mathbf{k}] * \hat{\theta}_1[\mathbf{k}] - \hat{I}_1[\mathbf{k}] * \hat{\theta}_2[\mathbf{k}] = 0$$

Convolution: matrix multiplication



Self calibration methods: Fourier domain

Fourier domain relation

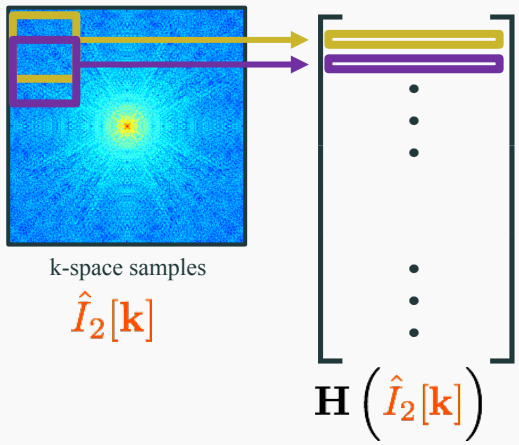
$$\hat{I}_2[\mathbf{k}] * \hat{\theta}_1[\mathbf{k}] - \hat{I}_1[\mathbf{k}] * \hat{\theta}_2[\mathbf{k}] = 0$$

Compact matrix representation

$$\underbrace{\left[\mathbf{H}\left(\hat{I}_2[\mathbf{k}]\right), \mathbf{H}\left(\hat{I}_1[\mathbf{k}]\right)\right]}_{\mathbf{Q}(I_1, I_2)} \begin{bmatrix} \hat{\theta}_1[\mathbf{k}] \\ -\hat{\theta}_2[\mathbf{k}] \end{bmatrix} = 0$$

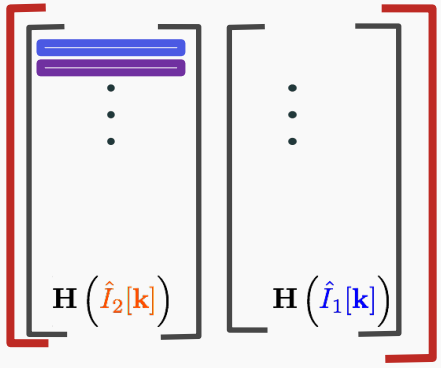
N shots: $\binom{N}{2}$ null space vectors

Q is low-rank & structured

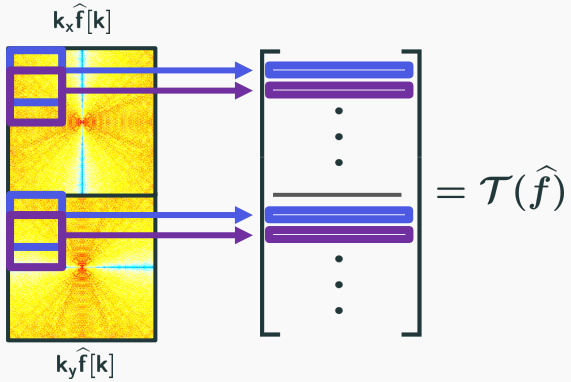


Smoothness regularized multishot MRI

Multi-shot recovery



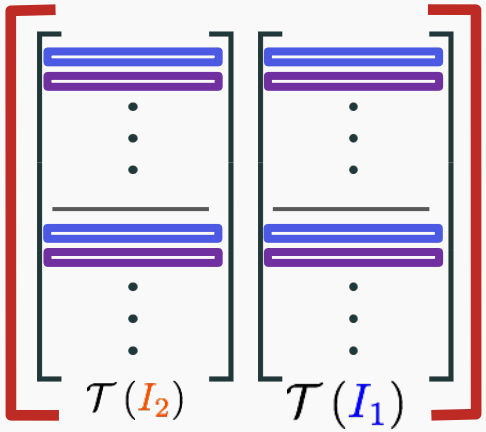
Smoothness regularization



Combine the matrix liftings

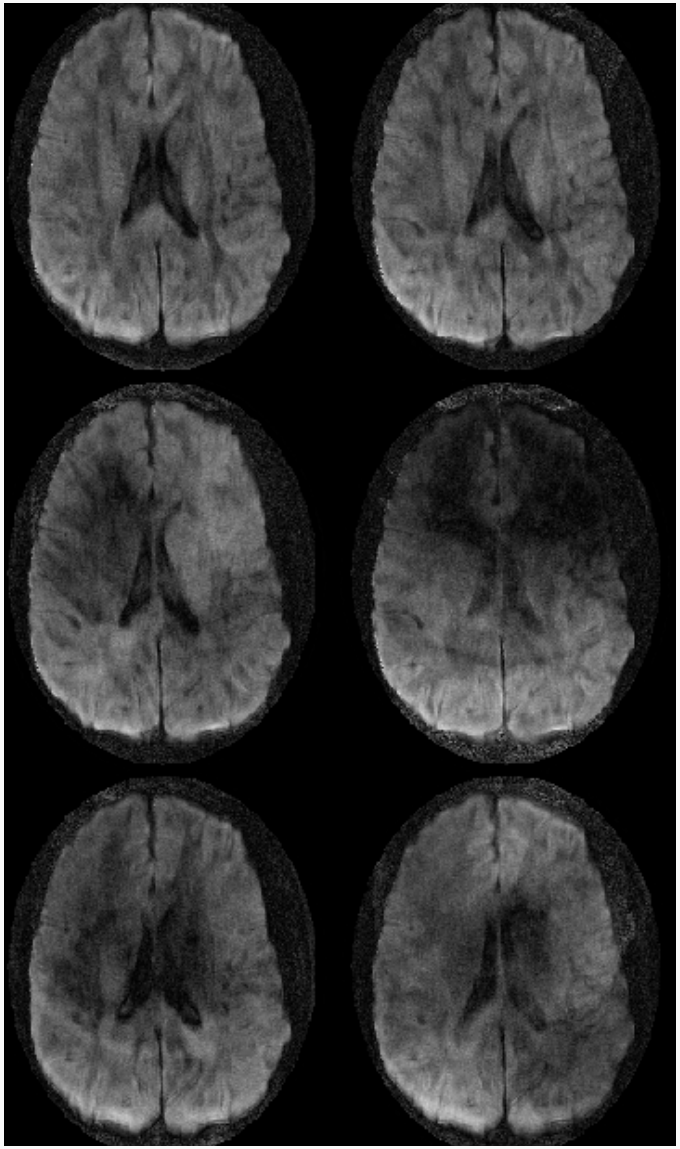
Structured low-rank recovery

$$\|\mathcal{A}(I_1, I_2) - \mathbf{b}\|^2 + \lambda \|\mathcal{G}(I_1, I_2)\|_*$$

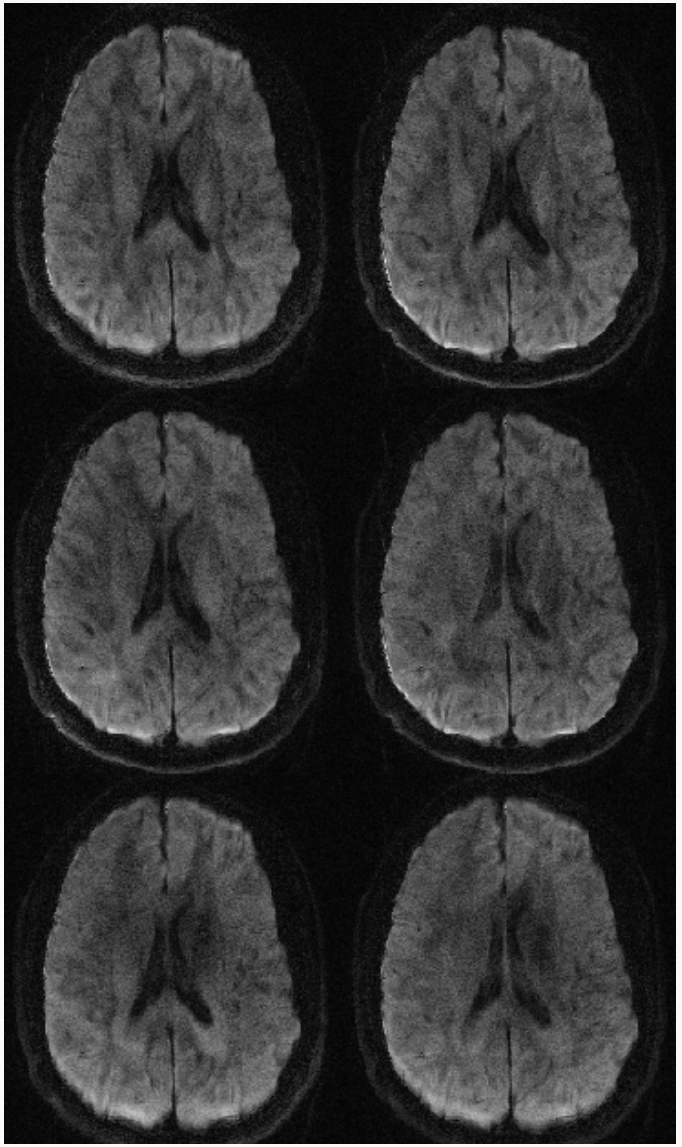


Mani & Jacob, Magnetic Resonance Medicine, in press

Structured low-rank recovery: MUSSELS

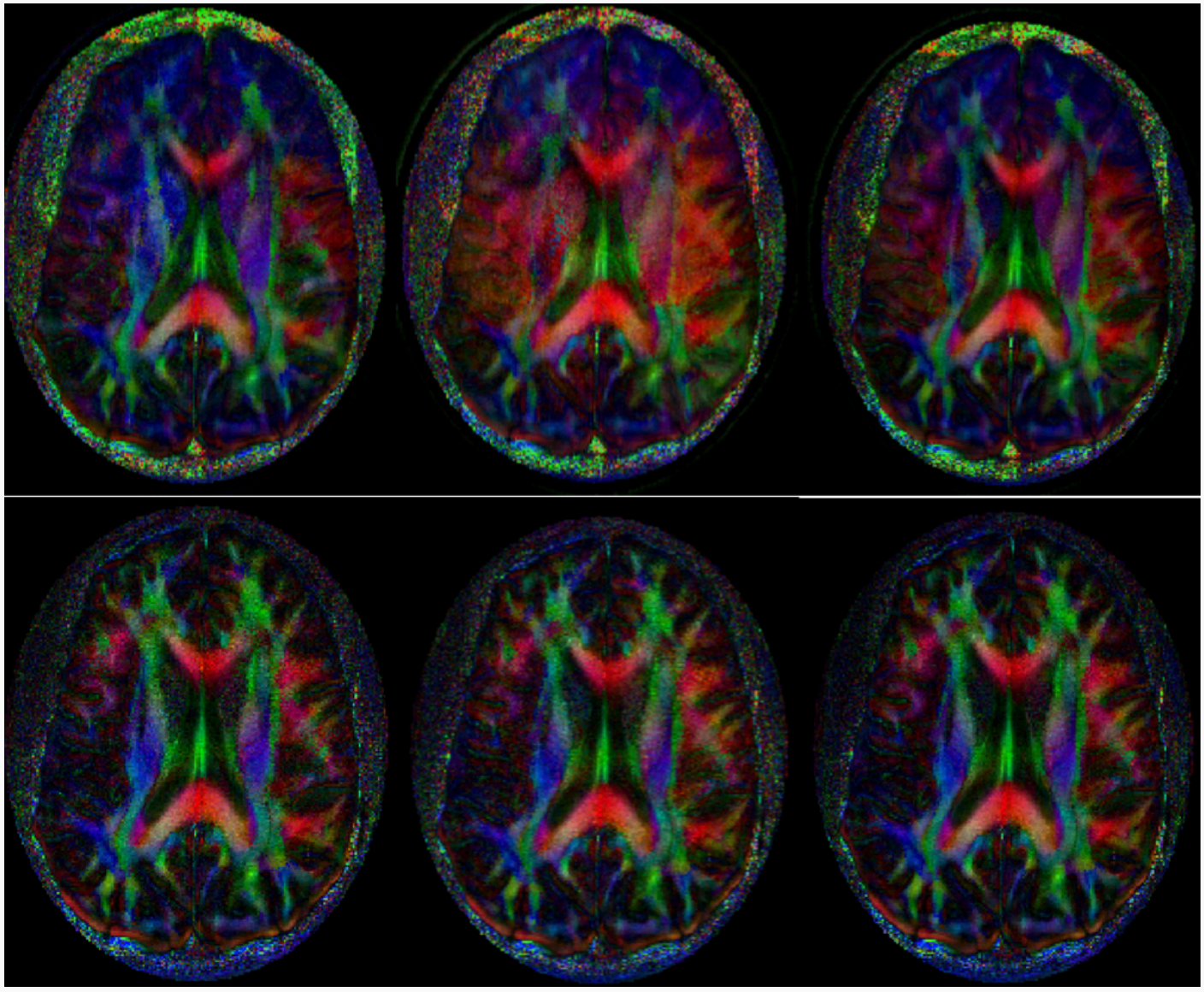


MUSE



MUSSELS

Comparison with MUSE (state of the art)

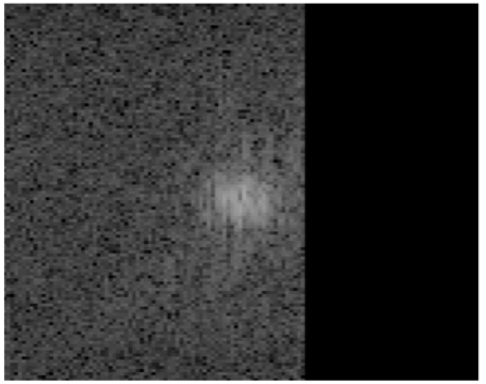


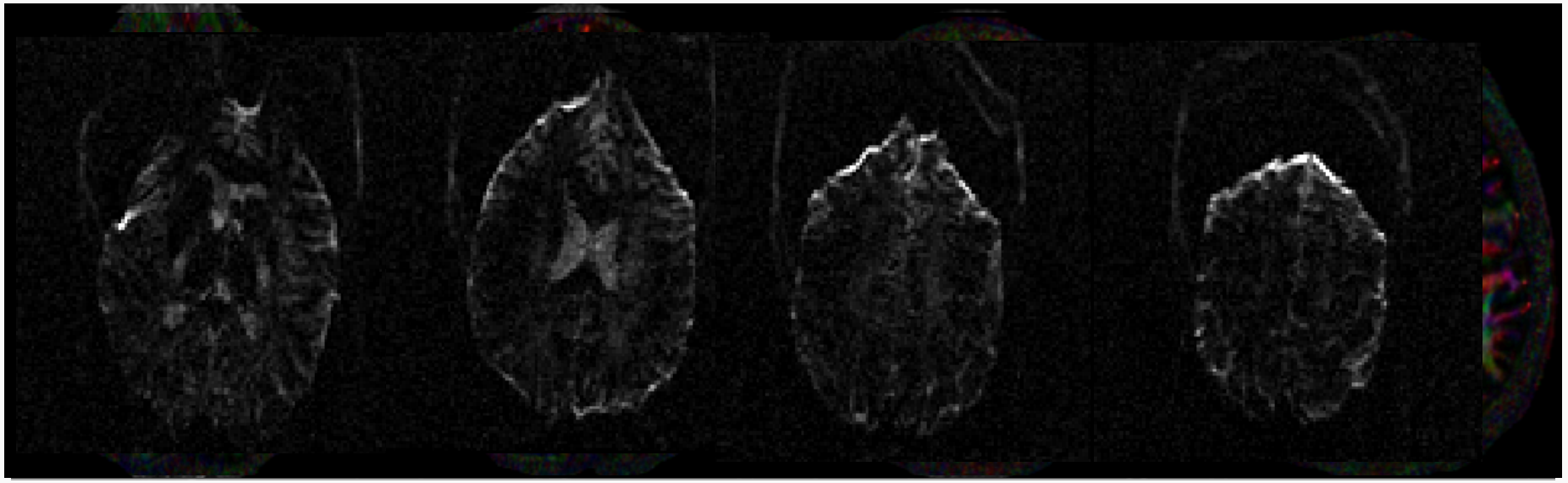
Mani & Jacob, Magnetic Resonance Medicine, in press

Odd even shifts & partial Fourier in EPI

Multishot & partial Fourier

Improved SNR & reduced distortion

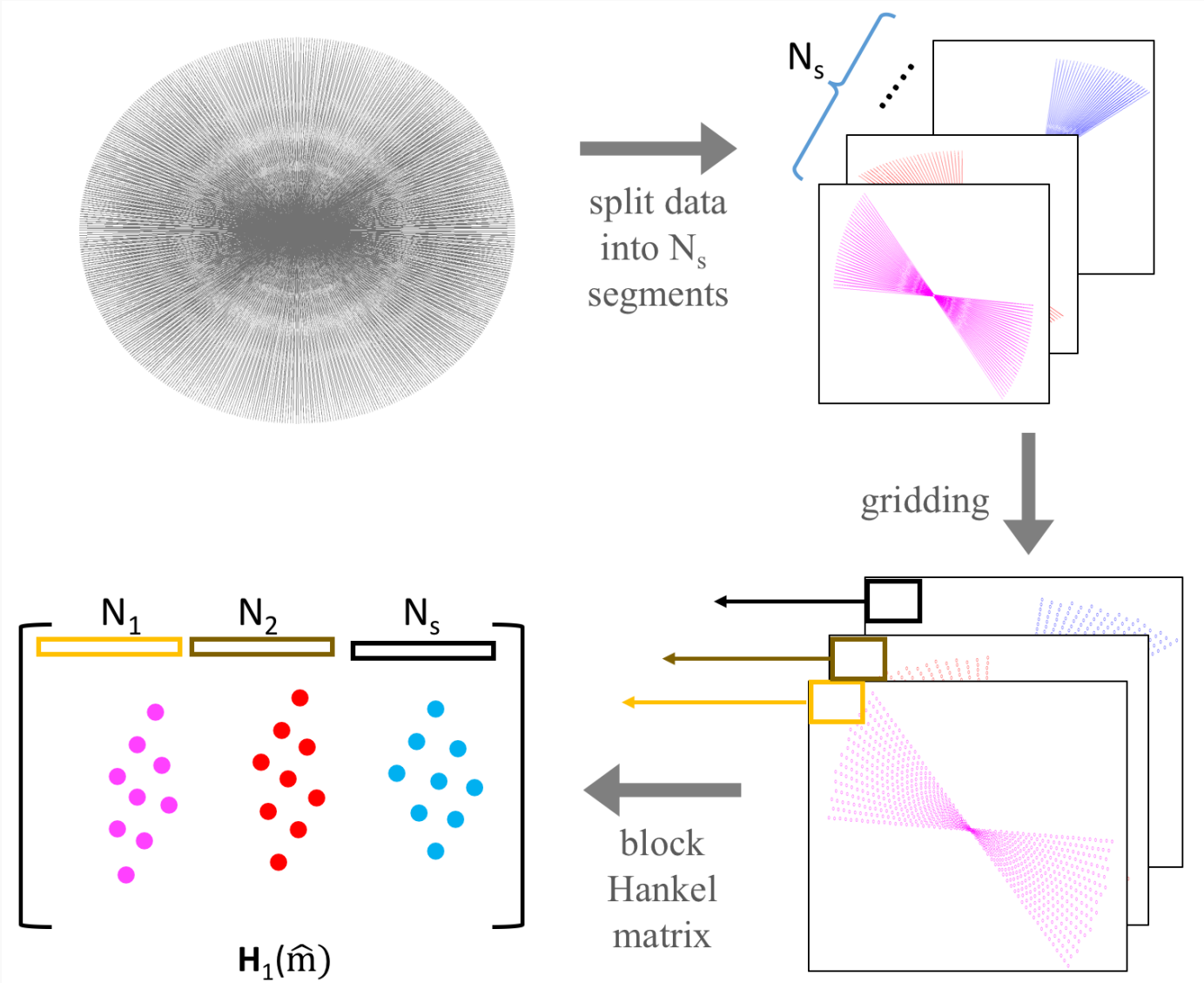




0.8 x 0.8 x 2mm; 3 avgs; 25 directions; b=700

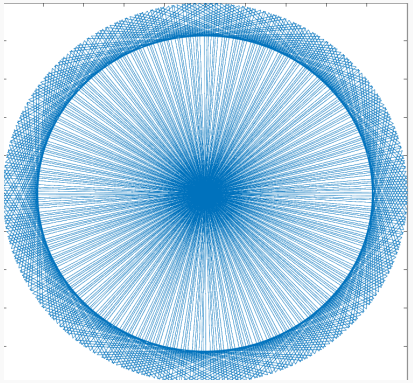
Mani & Jacob, Magnetic Resonance Medicine, in press, EMBC 2016

Radial trajectory correction

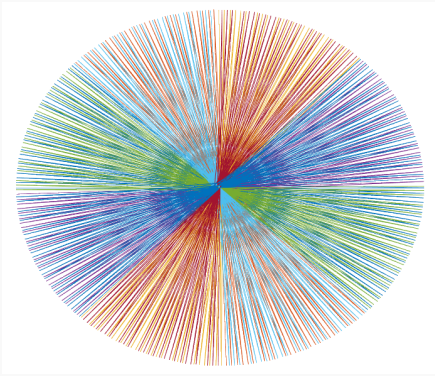


Results: MUSSELS based radial traj. correction

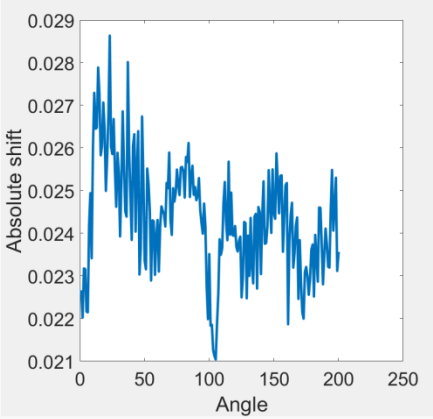
- Radial data:
  - 3T GE scanner
  - 256 spokes
  - 154 points per spoke
  - partial Fourier acq.
  - 32 channels



Ideal Radial trajectory

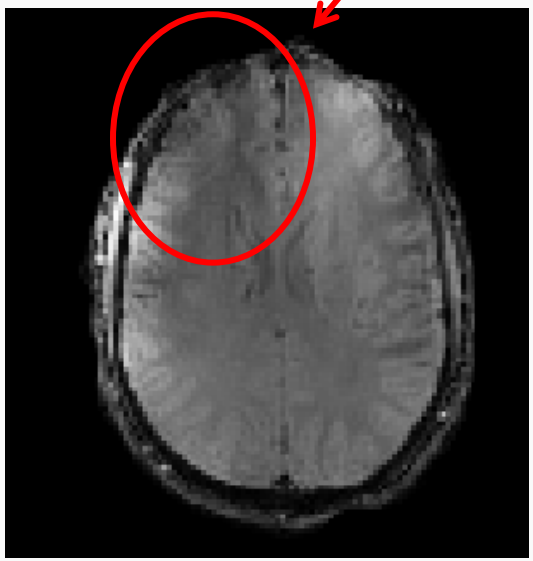


The 8 segments

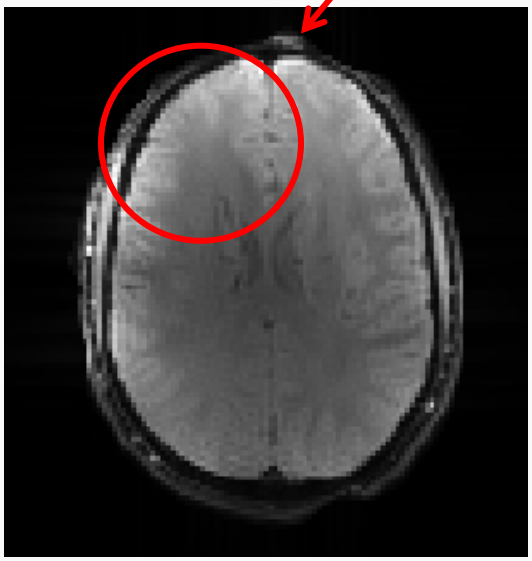


Plot of trajectory shift vs angle

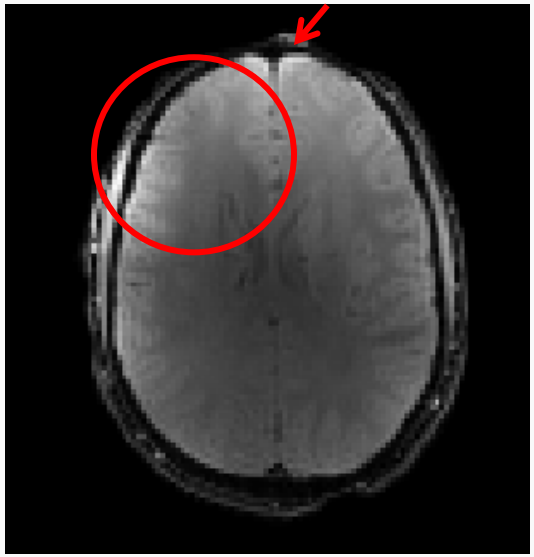
Recon using NUFFT



Recon using MUSSELS

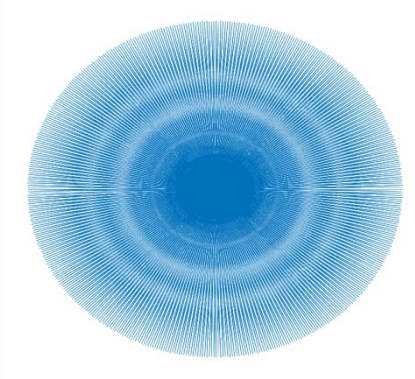


Recon using TrACR

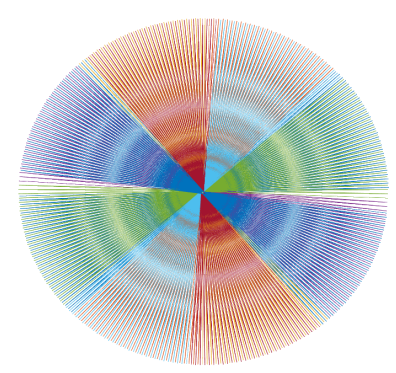


Results: MUSSELS based radial traj. correction

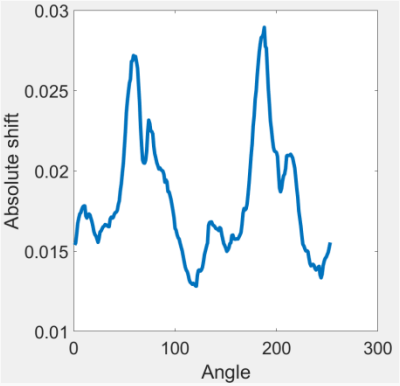
- Radial data:
  - 3T Siemens scanner
  - 512 spokes
  - 512 points per spoke
  - 5 channels



Ideal Radial trajectory

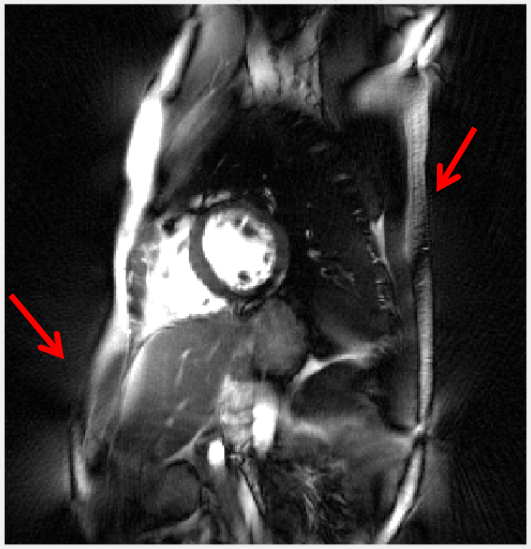


The 8 segments

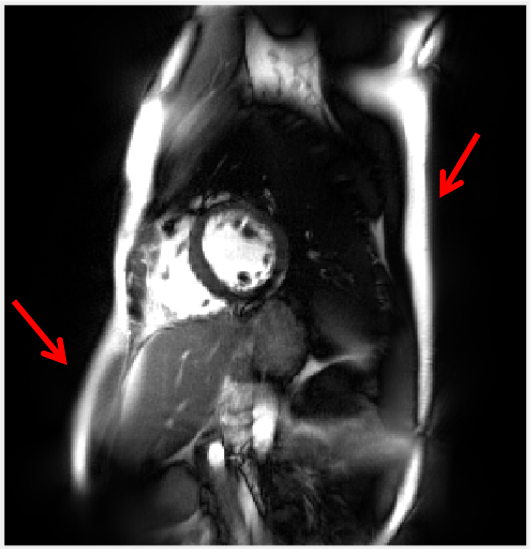


Plot of trajectory shift vs angle

Recon using NUFFT



Recon using MUSSELS

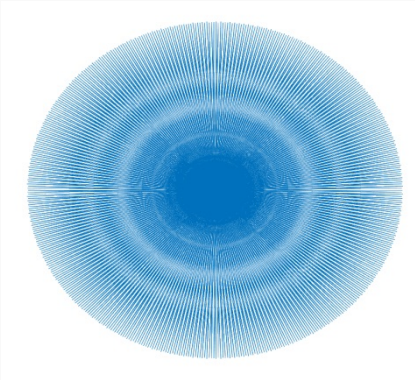


Recon using TrACR

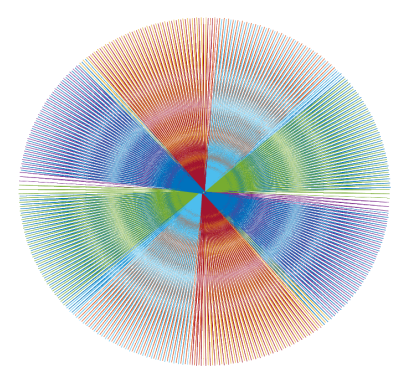


Results: MUSSELS based radial traj. correction

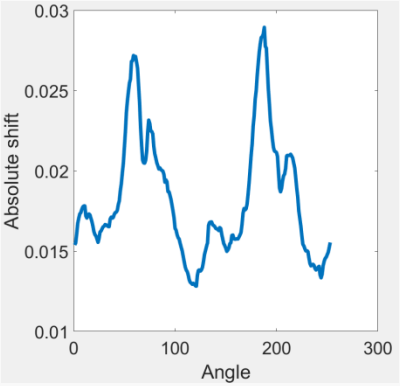
- Radial data:
  - 3T Siemens scanner
  - 512 spokes
  - 512 points per spoke
  - 5 channels



Ideal Radial trajectory

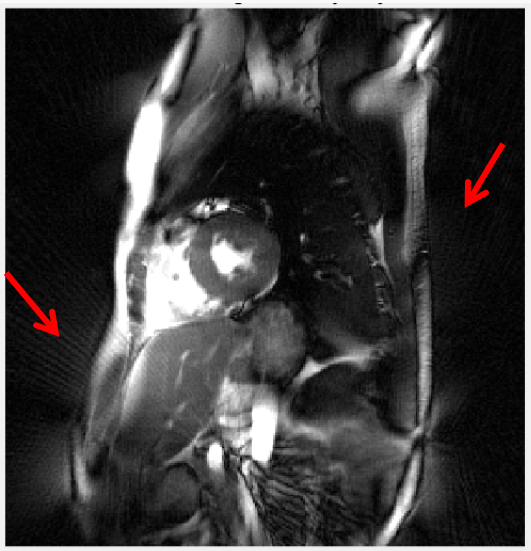


The 8 segments



Plot of trajectory shift vs angle

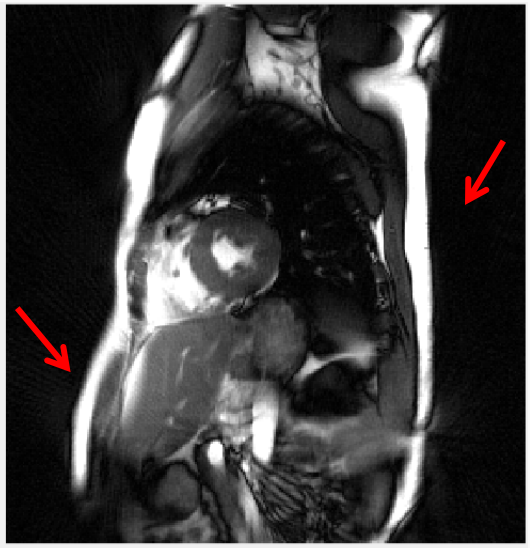
Recon using NUFFT



Recon using MUSSELS

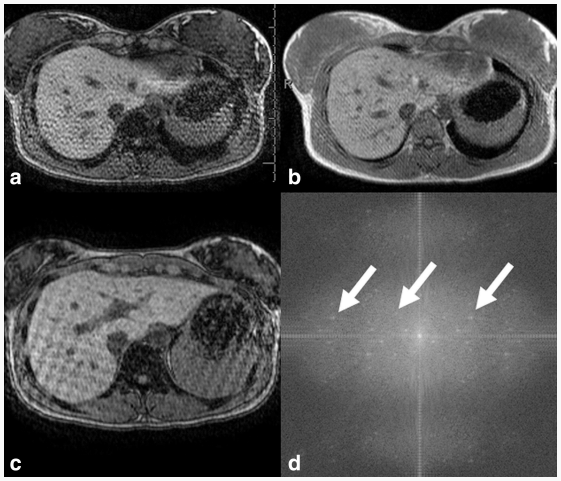


Recon using TrACR

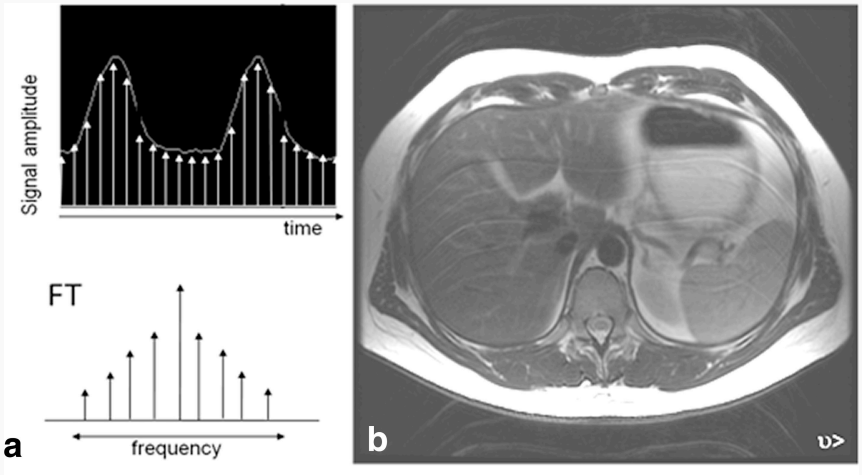


MR artifacts

- What is MR artifacts?
  During acquisition, external interruptions (ex. fluctuation power supply of gradient, motion of object, etc.) distort signals.



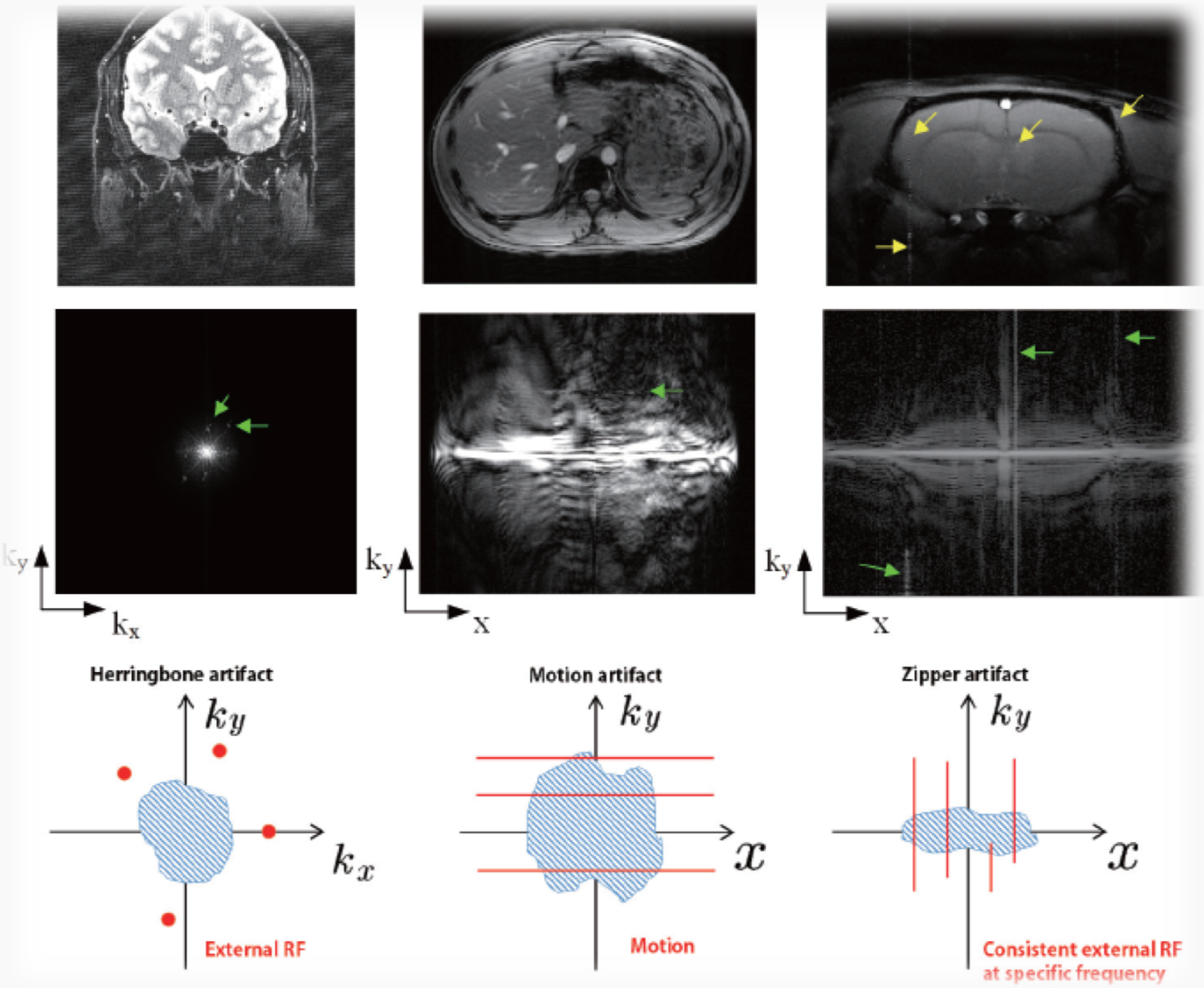
Spike noise



Respiratory motion

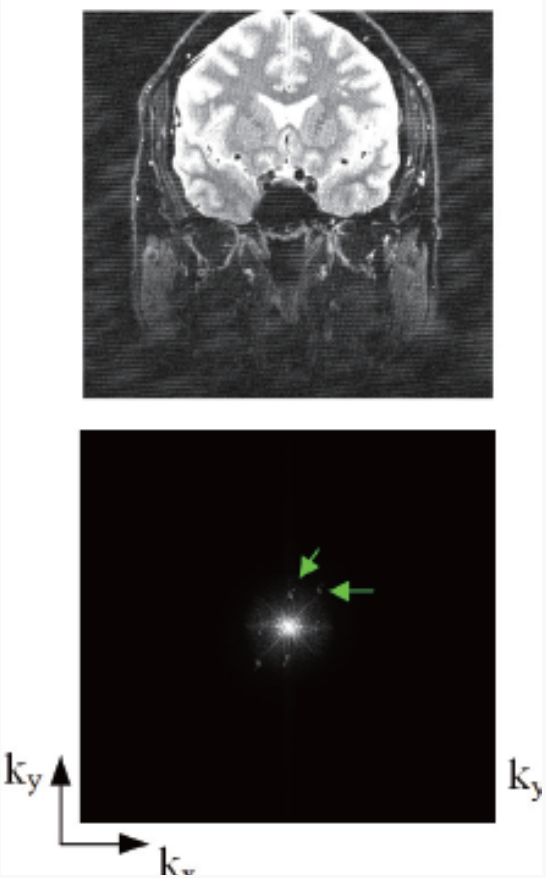
M. Graves, et. al., JMRI (2013)

Motivation

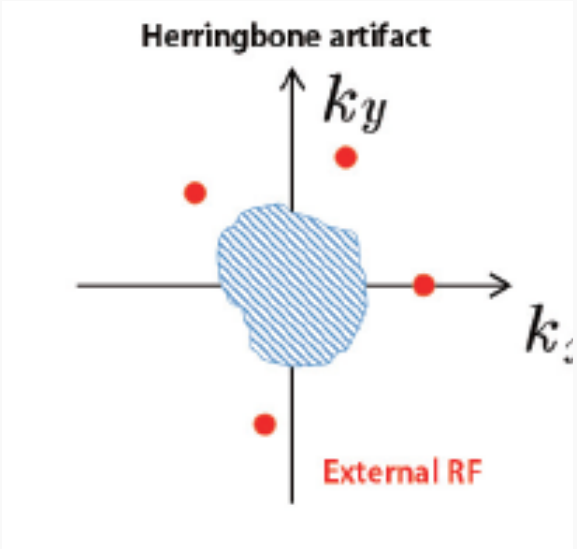

ky
kx
ky
x
ky
x
Herringbone artifact
ky
kx
External RF
Motion artifact
ky
x
Motion
Zipper artifact
ky
x
Consistent external RF
at specific frequency

Motivations: MR artifacts as sparse outliers

- ✓ Herringbone (spikes 2-D k-space)



$$\widetilde{M}(k_x, k_y) = \widehat{M}(k_x, k_y) + \underbrace{\sum_{j=1}^{S} \epsilon_j \delta[k_x - k_{x_j}, k_y - k_{y_j}]}_{\text{sparse outliers}},$$

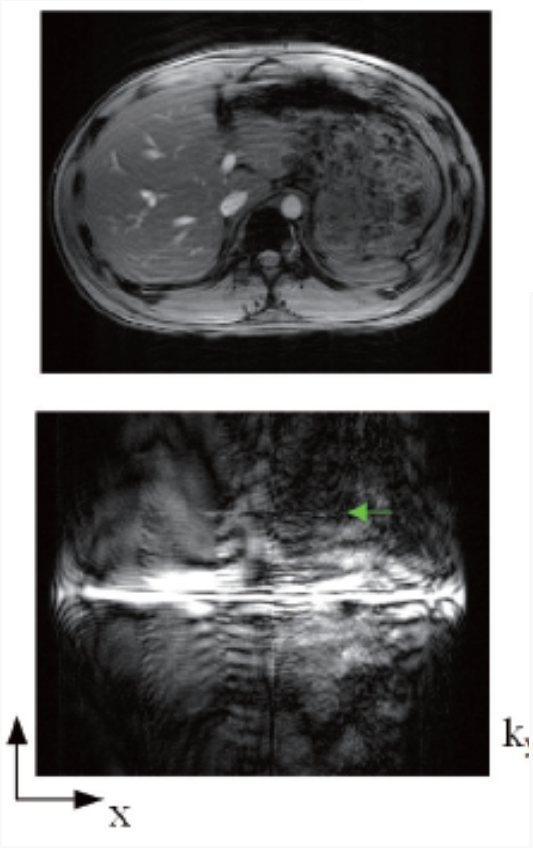


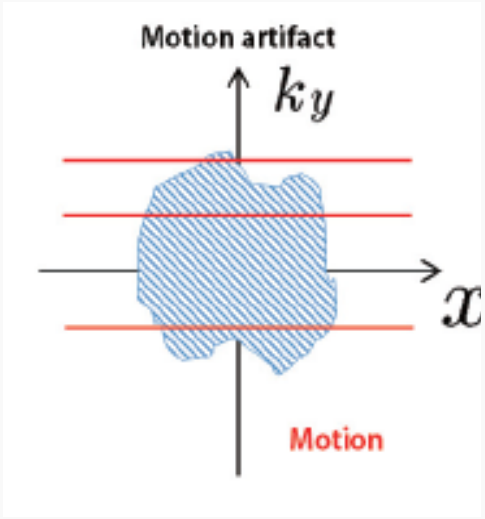
Motivations: MR artifacts as sparse outliers

- ✓ Motion artifact (spikes 1-D k-space parallel to readout)

$$\widetilde{M}(x,k_y) = \begin{cases} \widehat{M}(x,k_y)\exp(j2\pi k_y d(k_y)), & \text{when } k_y \in \{k_{y_1},\cdots,k_{y_S}\} \\ \widehat{M}(x,k_y), & \text{otherwise} \end{cases}$$

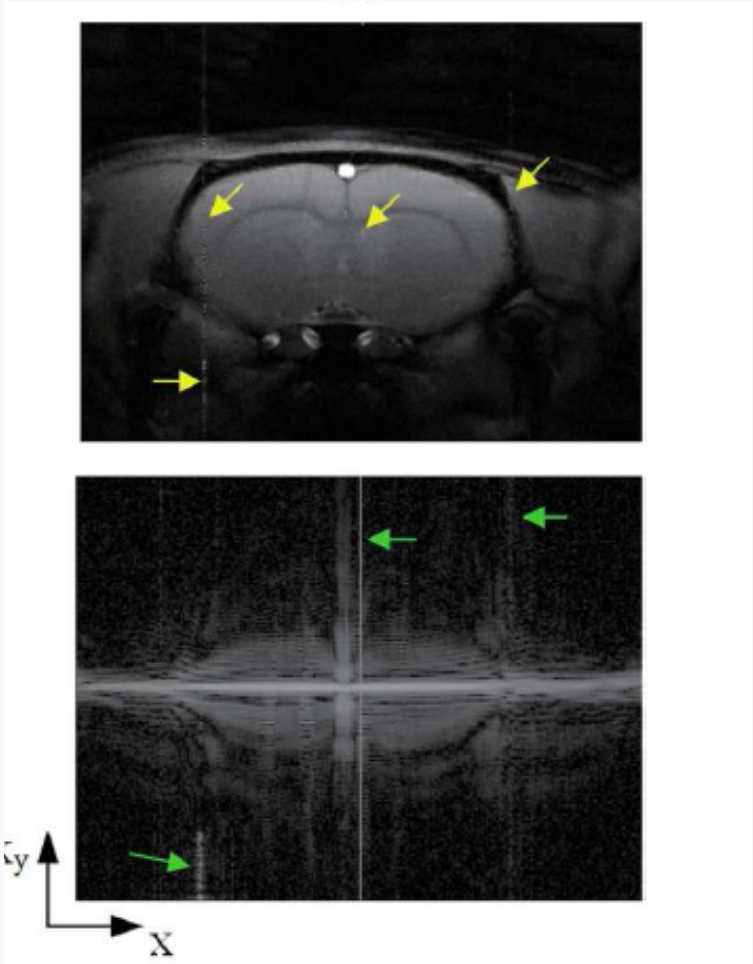
$$= \widehat{M}(x,k_y) + \underbrace{\sum_{j=1}^{S} \widehat{M}(x,k_y)(\exp(j2\pi k_y d(k_y)) - 1)\delta[k_y - k_{y_j}]}_{\text{sparse outliers}}$$



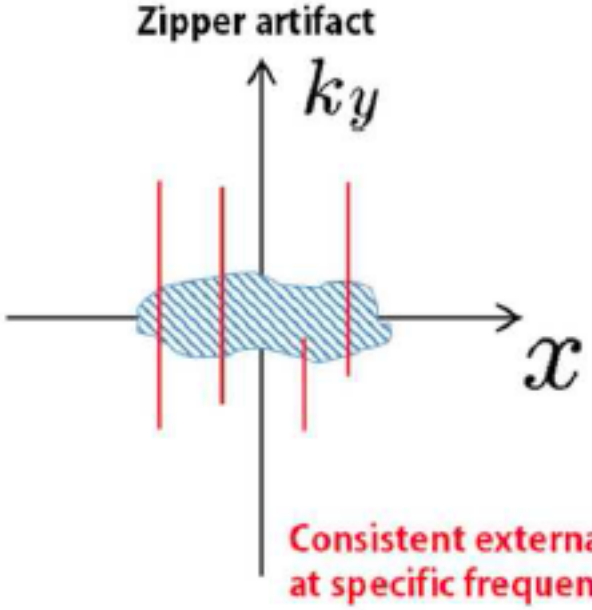


Motivations: MR artifacts as sparse outliers

Zipper artifact (spikes 1-D k-space perpendicular to readout)

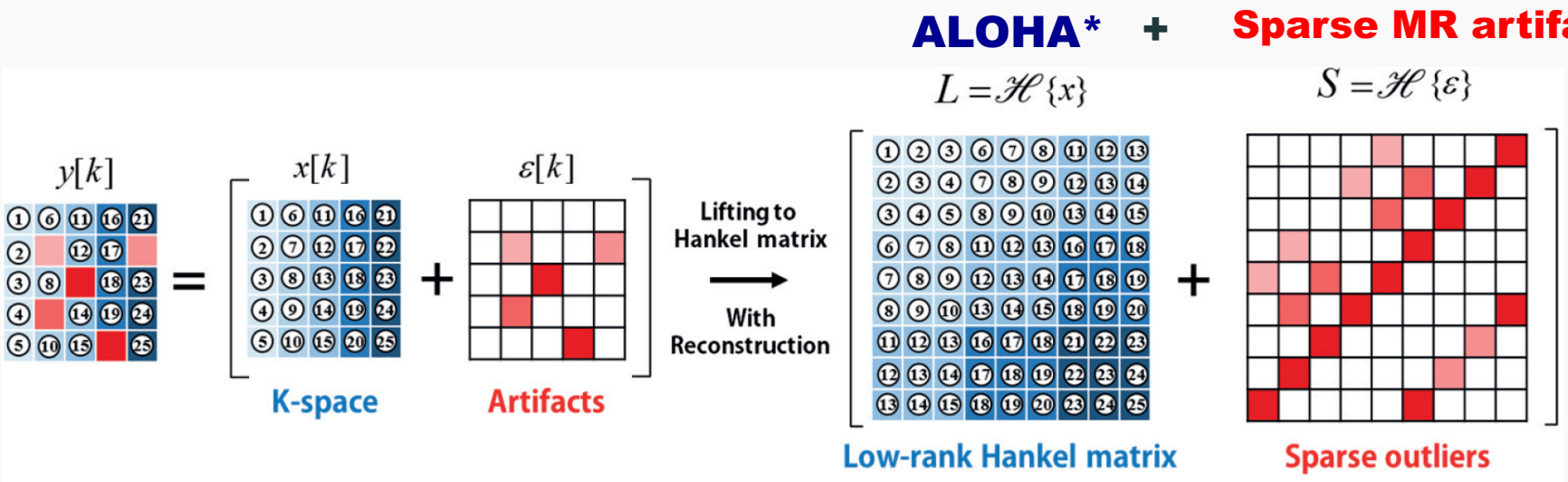


$$\widetilde{M}(x,k_y) = \widehat{M}(x,k_y) + \underbrace{\sum_{j=1}^{S} \epsilon_j \delta[x - x_j]}_{\text{sparse outliers}},$$



Key Observation : Sparse outliers

*** Sparse outlier is still sparse in weighted Hankel matrix**



- **ALOHA: Annihilating filter based LOw rank Hankel matrix Approach**

† E. Candes, et. al, JACM (2011), R. Otazo, et. al, MRM (2015)
*K.H. Jin, et. al, IEEE TIP (2015), K.H. Jin, et. al, arXiv (2015), J. C. Ye, et. al, arXiv (2015), J. Lee, et. al, MRM (2016), D. Lee, et. al, MRM (2016)

Robust ALOHA

RPCA for weighted Hankel matrix

$$\min_{\mathbf{M},\mathbf{E}} \; \|\mathscr{H}\{\mathbf{M}\}\|_* + \tau\|\mathbf{E}\|_1$$

signal — Sparse outlier

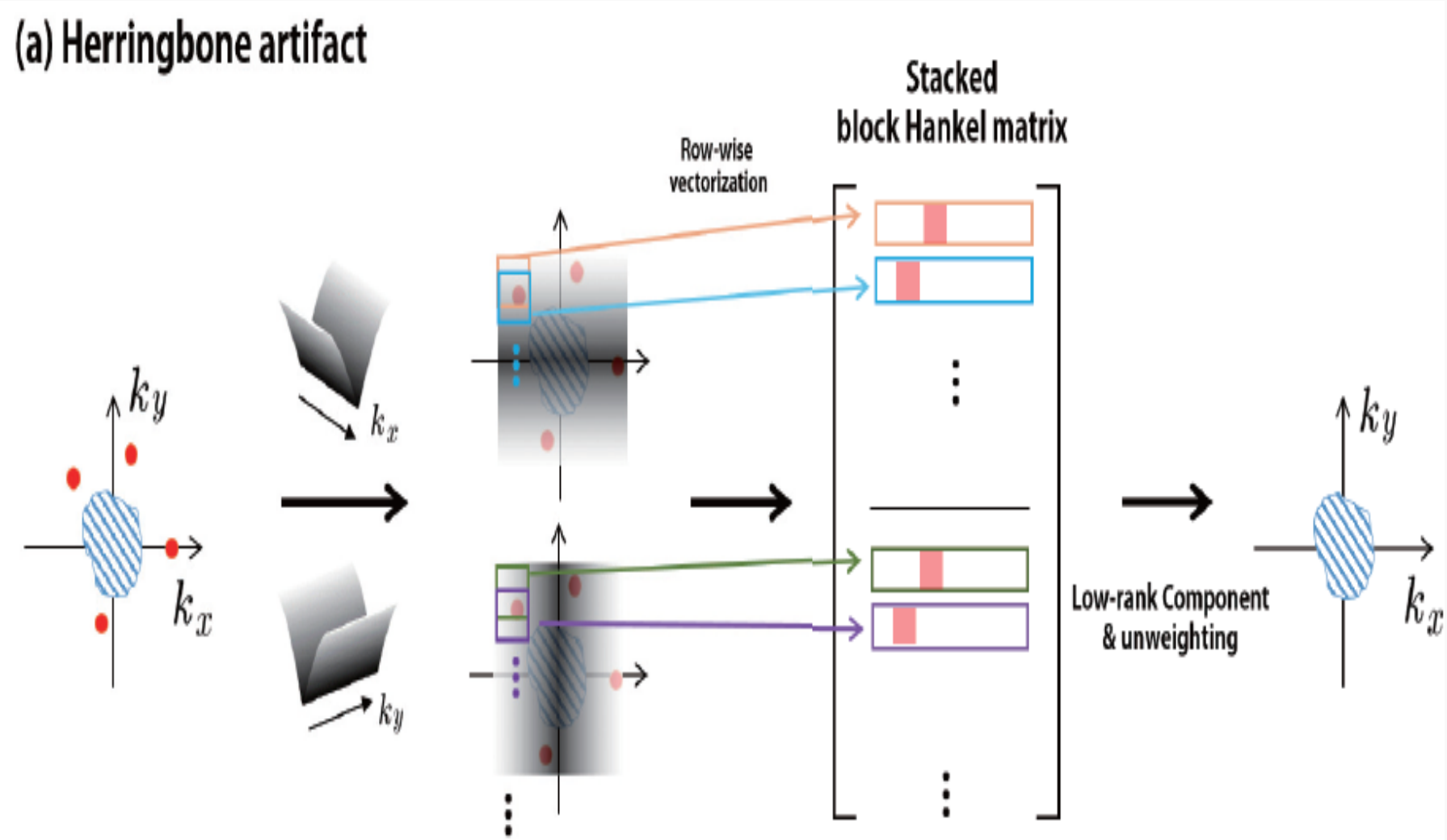
$$\text{subject to} \quad \mathbf{P}_\Omega\left(\widehat{\mathbf{W}} \odot \widehat{\mathbf{L}}\right) = \mathbf{P}_\Omega\left(\mathbf{M}+\mathbf{E}\right),$$

K-space weighting

- ✓ *Extension of ALOHA for decomposition of sparse outliers (E) out of mixed signal**
- ✓ *Can be addressed ADMM†*
- ✓ *K-space weighting*

* E. Candes, et. al, JACM (2011), R. Otazo, et. al, MRM (2015)
† S. Boyd, et. al., Foundations and Trends in Machine Learning (2011)
‡ Z. Wan, et. al., Mathematical Programming Computation (2012)

Algorithm Flowchart

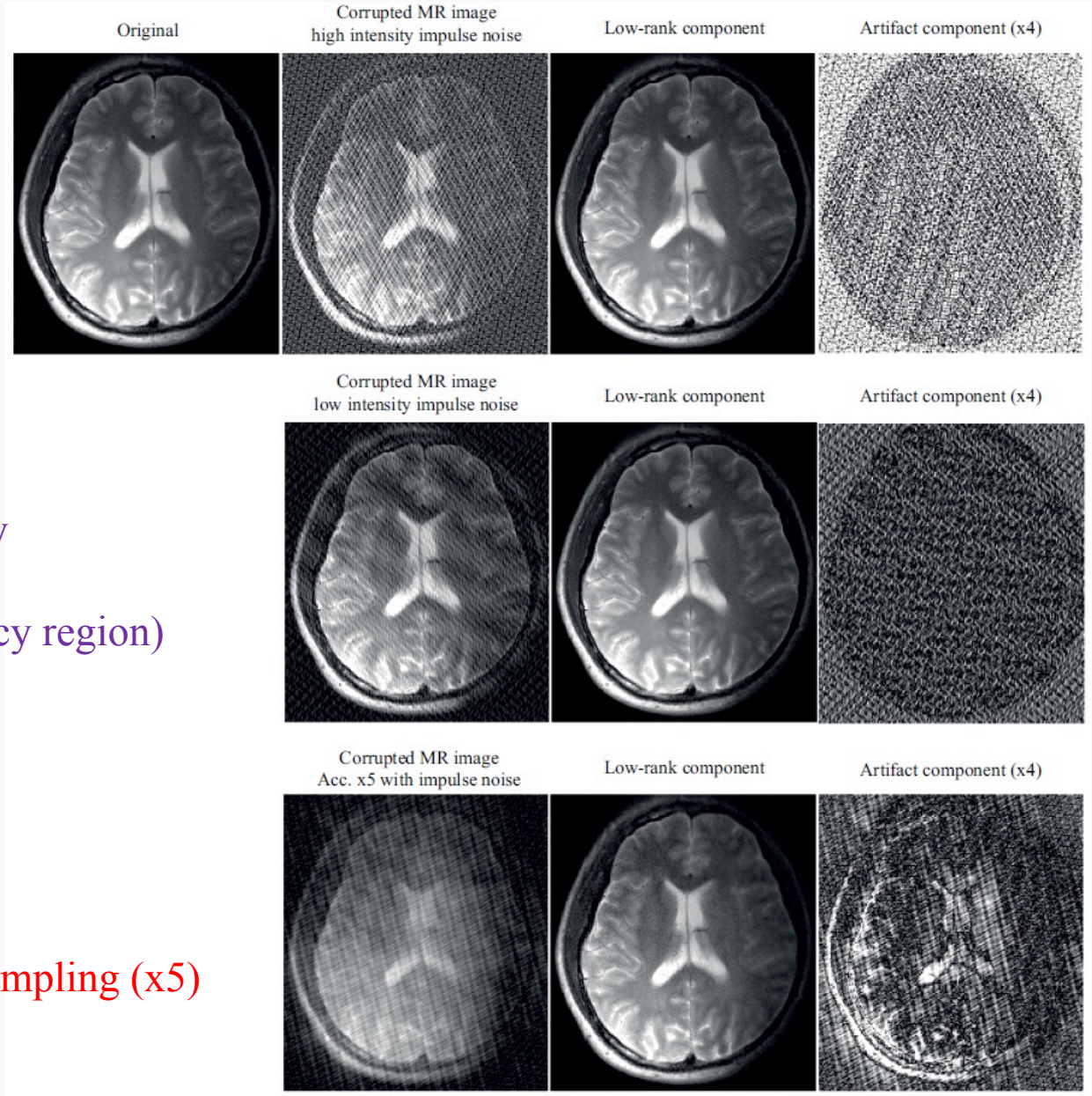


Retrospective results

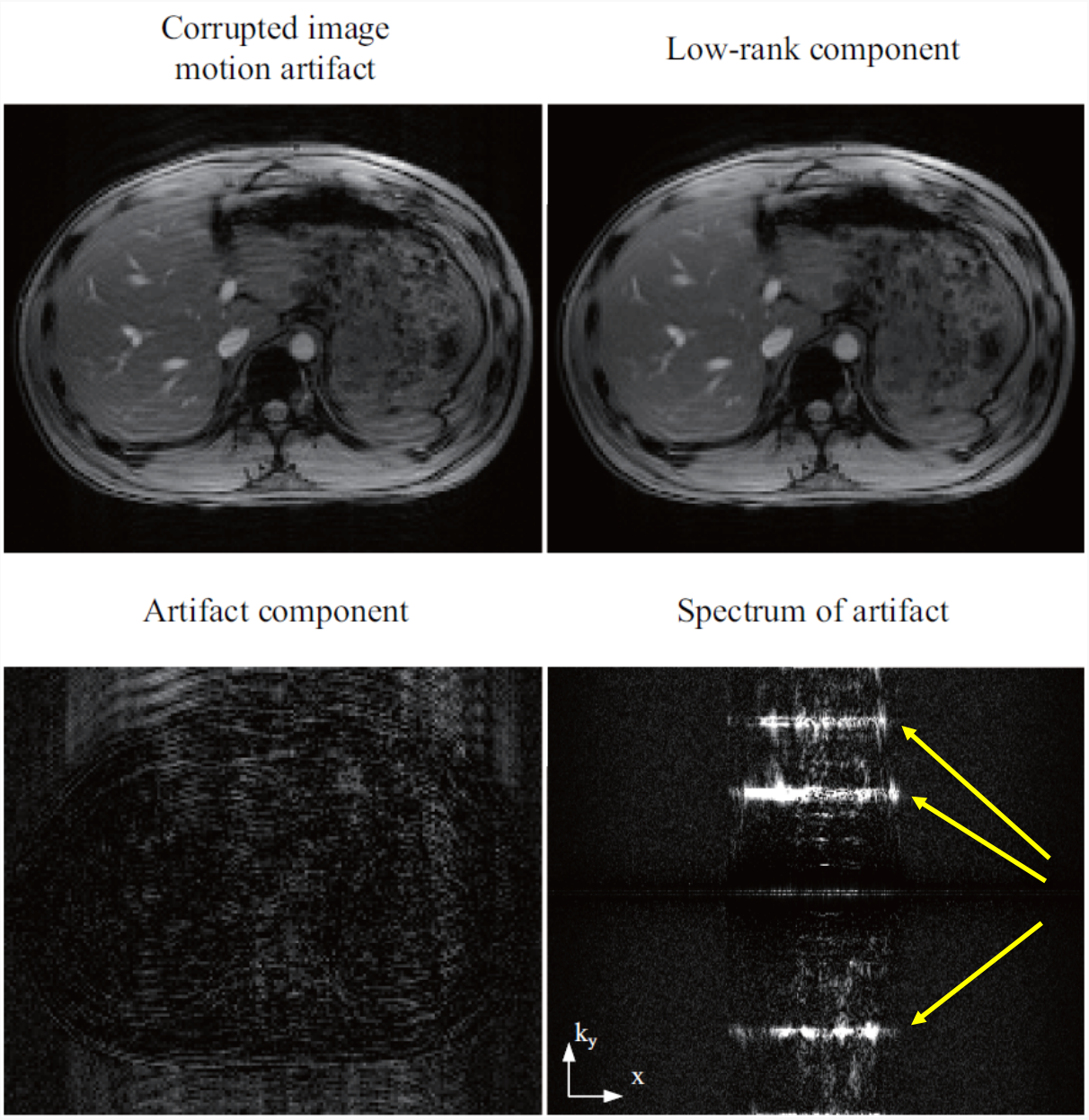
High intensity
Spike noise

Low intensity
Spike noise
(low frequency region)

Spike noise
with down sampling (x5)

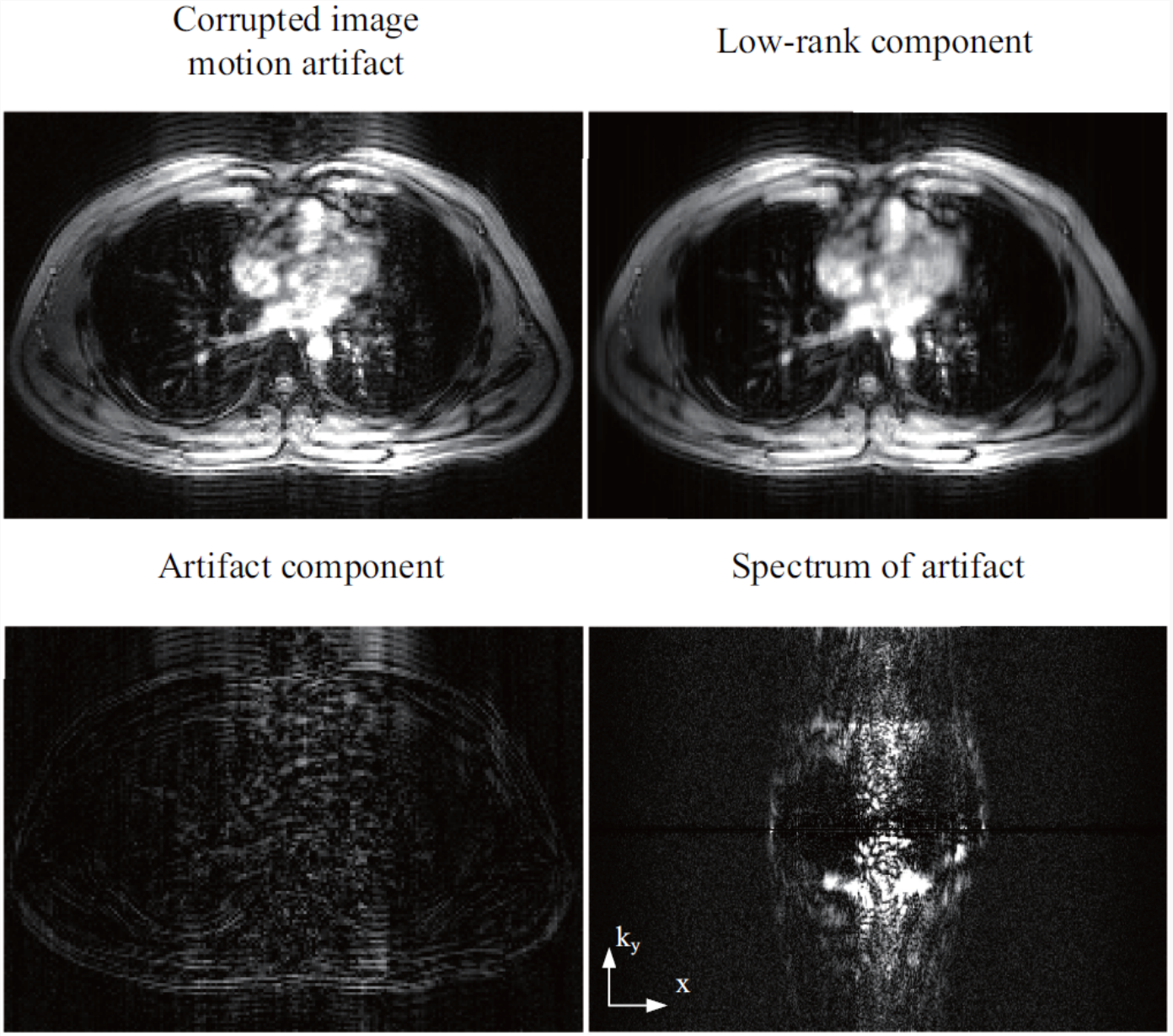


In Vivo Motion artifact

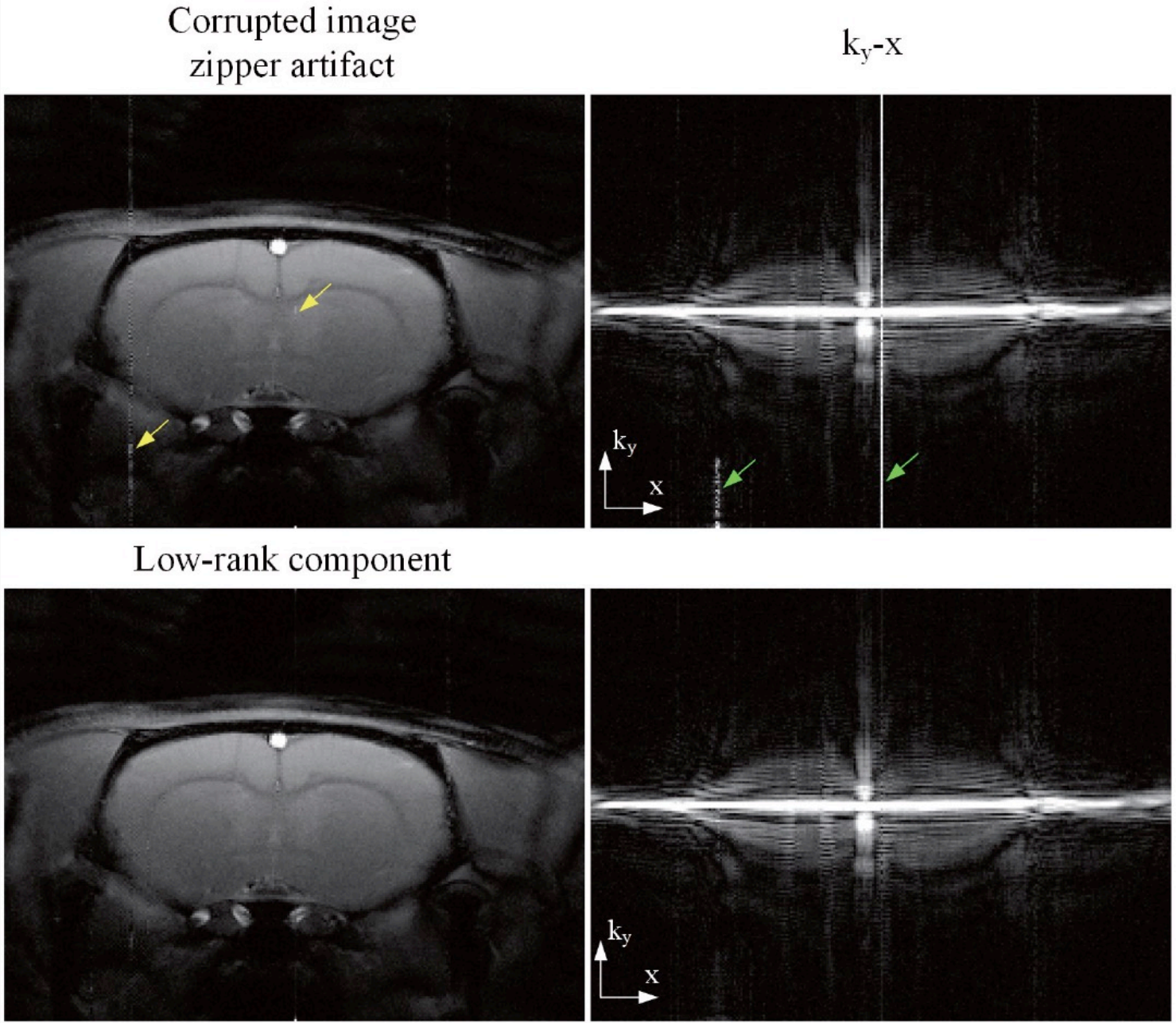


sudden motion
(3 times)

Cardiac Motion artifact

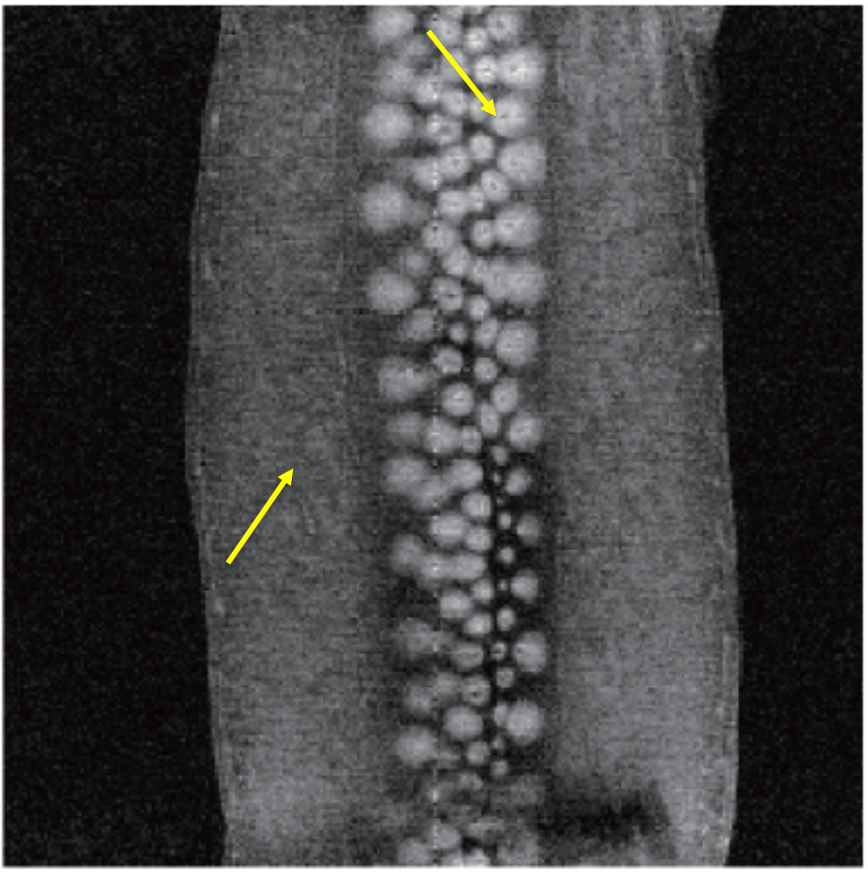


Zipper artifact

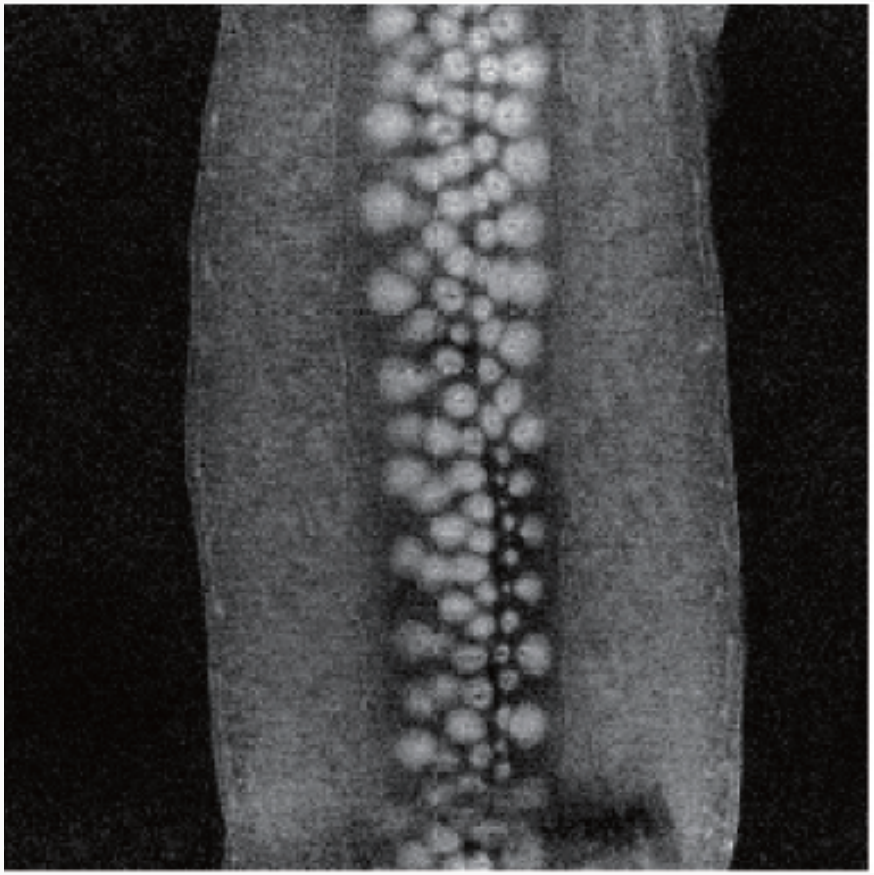


2-D herringbone (in-vivo)

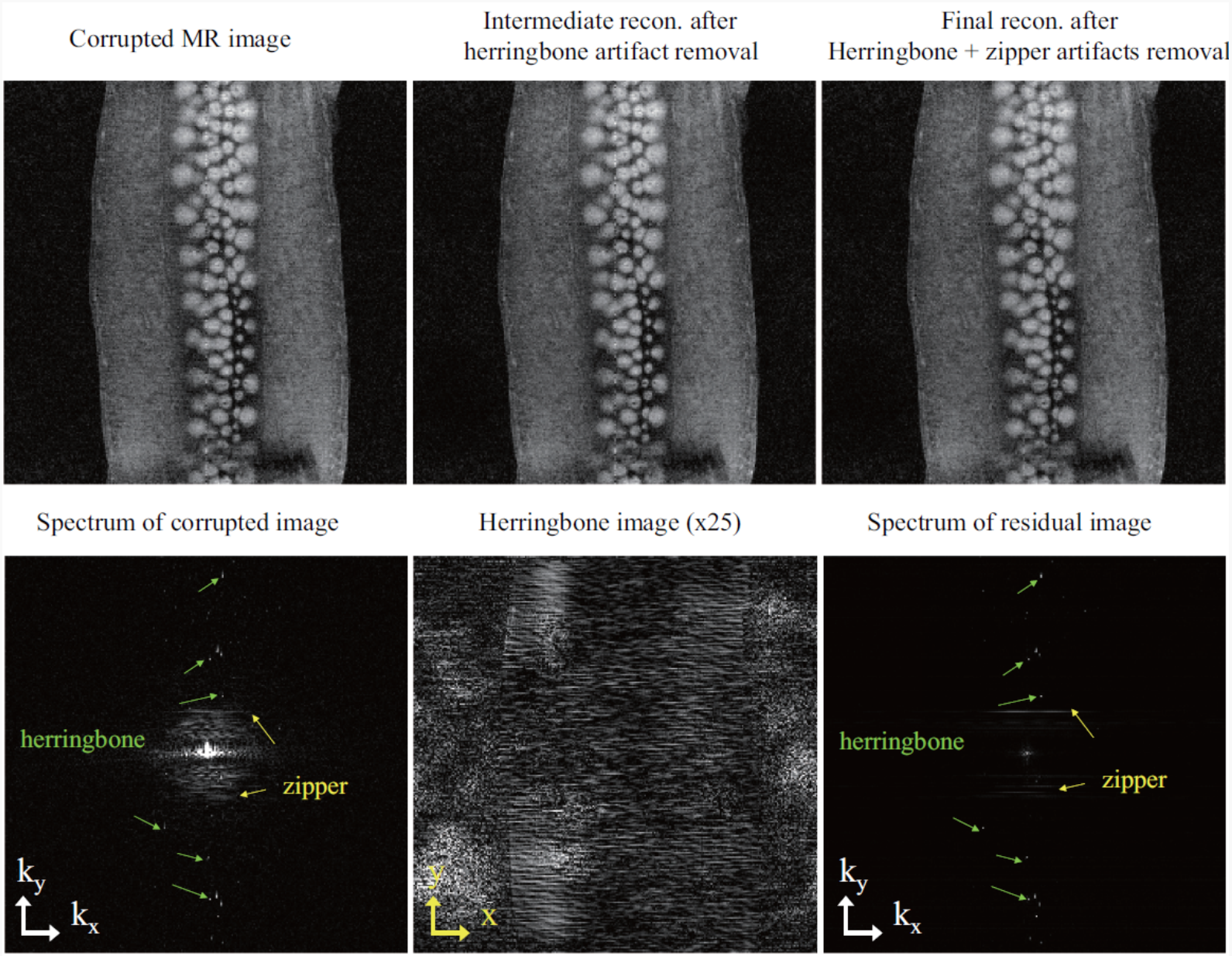
before



after

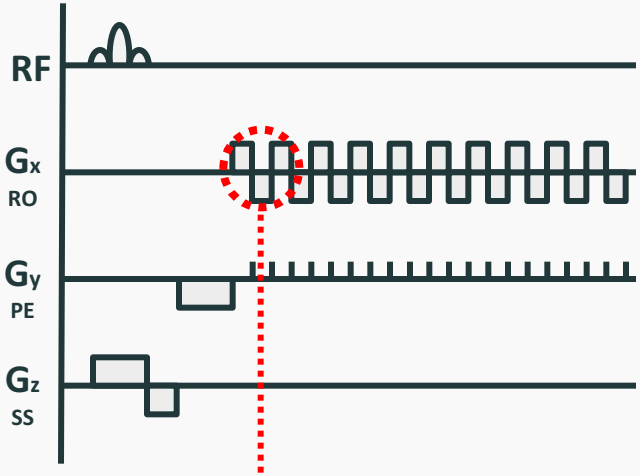


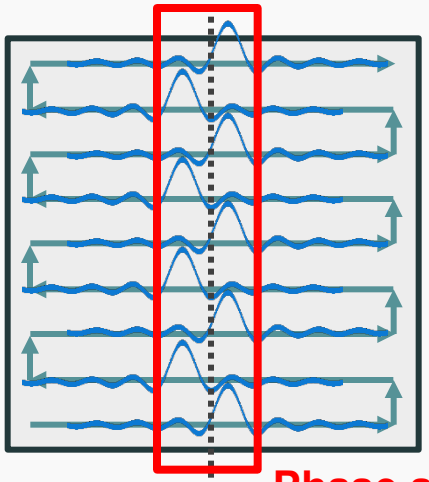
2-D herringbone (in-vivo)



EPI Ghost artifact

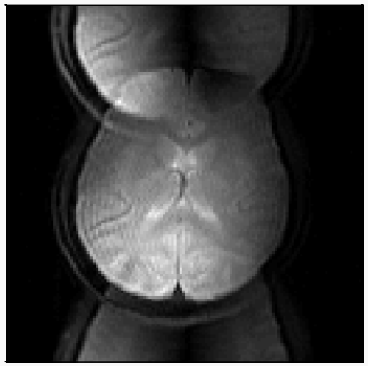
In EPI, Gradient is distorted by eddy currents and this causes phase shift



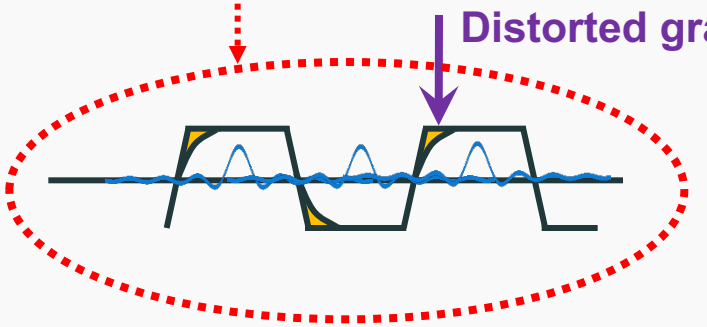


Phase shift

FT



Ghost artifact image

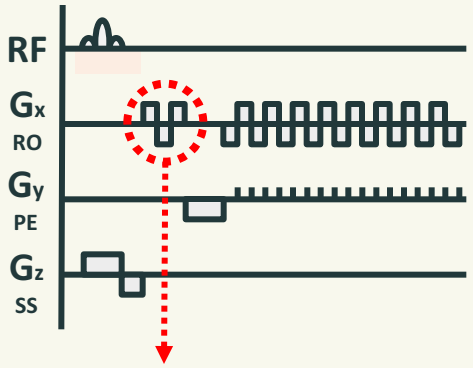


Even and odd echo mismatch causes ghost artifact!

Conventional correction

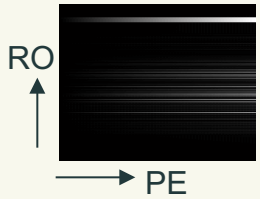
- **Navigator : pre-scan or reference scan**

Navigator-based



Without PE gradient

Calculate **difference** of phase between 1st -2nd line, 2nd -3rd line

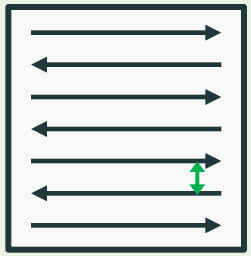


Make phase difference map

only possible to ***linear phase correction***

Navigator-free

- **Pulse sequence compensation**[1)]
- **Without any modification**[2)]
  - Using Parallel Imaging Information
  - others



Calculate Phase disparity from EPI data itself

lower performance compared to the reference-based approaches

1) Xiang QS et al., MRM, 2007
Poser BA et al., MRM, 2013

2) Zhang et al., MRM, 2004
Kim YC et al., JMRI, 2007

EPI model

EPI data can be expressed as

N : Total # of echoes
n : Index of each line
x : Read-out
y : Phase-encoding

Echo time

$$S_n(k_x,k_y)=\int\int m(x,y)e^{j2\pi\left[\Delta f(x,y)\left((TE+(n-N/2)ESP)+(-1)^n\left(\frac{k_x}{\gamma G_x}\right)\right)+k_x x+k_y y\right]}dxdy$$

Image intensity Frequency offset Echo spacing (time between each echo)

$$S_{n,+}(k_x,k_y)=\int\int m(x,y)e^{j2\pi\left[\Delta f(x,y)\left((TE+(n-N/2)ESP)+\left(\frac{k_x}{\gamma G_x}\right)\right)+k_x x+k_y y\right]}dxdy$$

Virtual k-space (**even** signals)

$$=\int\int A\cdot e^{\left(j2\pi\Delta f(x,y)\frac{k_x}{\gamma G_x}\right)}\cdot e^{j2\pi(k_x x+k_y y)}dxdy$$

$$S_{n,-}(k_x,k_y)=\int\int m(x,y)e^{j2\pi\left[\Delta f(x,y)\left((TE+(n-N/2)ESP)-\left(\frac{k_x}{\gamma G_x}\right)\right)+k_x x+k_y y\right]}dxdy$$

Different!

Virtual k-space (**odd** signals)

$$=\int\int A\cdot e^{\left(-j2\pi\Delta f(x,y)\frac{k_x}{\gamma G_x}\right)}\cdot e^{j2\pi(k_x x+k_y y)}dxdy$$

where $A=m(x,y)\exp\left(j2\pi\left[\Delta f(x,y)\left((TE+(n-N/2)ESP)\right)\right]\right)$

Sparsity of difference

The ghost generating phase term can be changed into a sine term

$$S_{n,\Delta}(k_x,k_y) = S_{n,+}(k_x,k_y) - S_{n,-}(k_x,k_y)$$

$$= \int\int A(x,y)\left(e^{j2\pi\Delta f(x,y)\frac{k_x}{\gamma G_x}} - e^{-j2\pi\Delta f(x,y)\frac{k_x}{\gamma G_x}}\right)e^{j2\pi\left(k_x x+k_y y\right)}\mathrm{d}x\mathrm{d}y$$

$$= \int\int A(x,y)2j\sin\left(2\pi\Delta f(x,y)\frac{k_x}{\gamma G_x}\right)e^{j2\pi\left(k_x x+k_y y\right)}\mathrm{d}x\mathrm{d}y$$

$$\sin\left(2\pi\Delta f(x,y)\frac{k_x}{\gamma G_x}\right) \simeq 2\pi\Delta f(x,y)\frac{k_x}{\gamma G_x}$$

$$S_{n,\Delta}(k_x,k_y) \simeq j2\pi k_x\int\int\frac{2}{\gamma G_x}A(x,y)\Delta f(x,y)e^{j2\pi\left(k_x x+k_y y\right)}\mathrm{d}x\mathrm{d}y$$

$$= \frac{2}{\gamma G_x}\mathcal{F}\left[\frac{\partial A(x,y)\Delta f(x,y)}{\partial x}\right]$$ **Sparse**

How can we use this sparsity?

Sparsity of difference (Cont.)

$$S_{n,\Delta}(k_x, k_y) = \mathcal{F}(\textit{Sparse signal})$$

⇨ **Hankel structure matrix constructed by $S_{n,\Delta}(k_x, k_y)$ is low-ranked**

$$\mathscr{H}(S_{n,\Delta})\mathfrak{h} = (\mathscr{H}(S_{n,+}) - \mathscr{H}(S_{n,-}))\mathfrak{h} = \mathbf{0}$$

$$\Downarrow$$

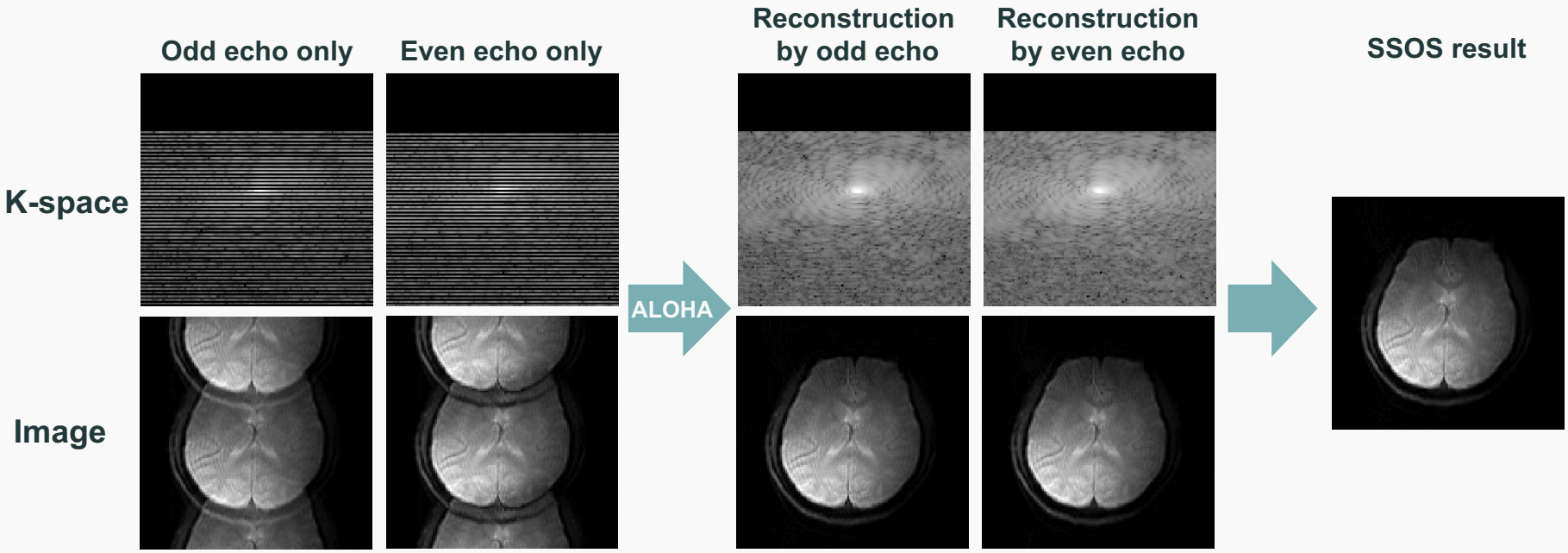
$$\left[\mathscr{H}(S_{n,+}) \quad \mathscr{H}(S_{n,-})\right] \begin{bmatrix} \mathfrak{h} \\ -\mathfrak{h} \end{bmatrix} = \mathbf{0}$$

Low rank structured matrix completion algorithm

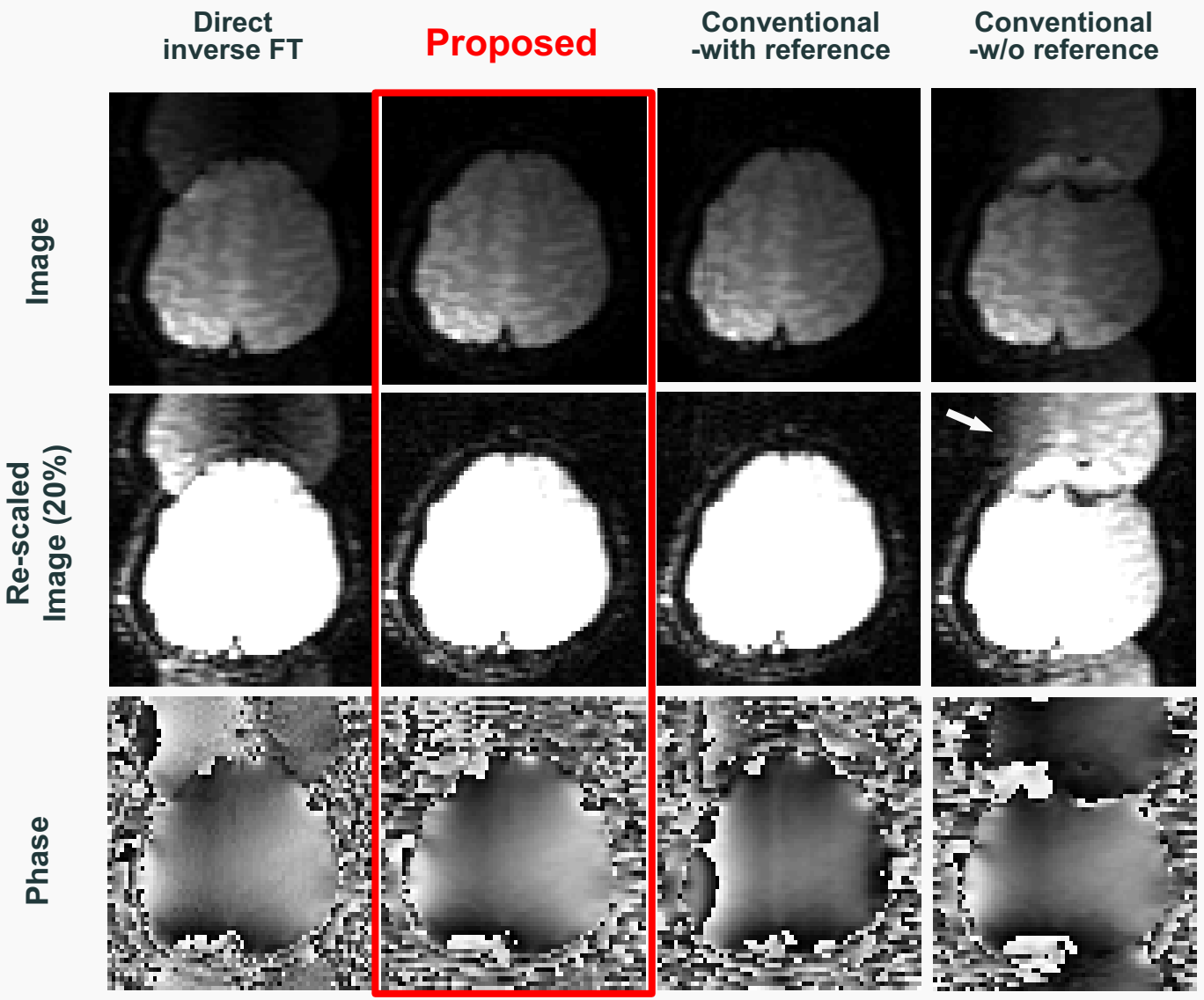
EPI ghost correction Problem ⇨ **k-space interpolation Problem using low rank structure**

Reconstruction flow

- **SE-EPI in-vivo data, 128x128 matrix size, 6/8 partial Fourier**



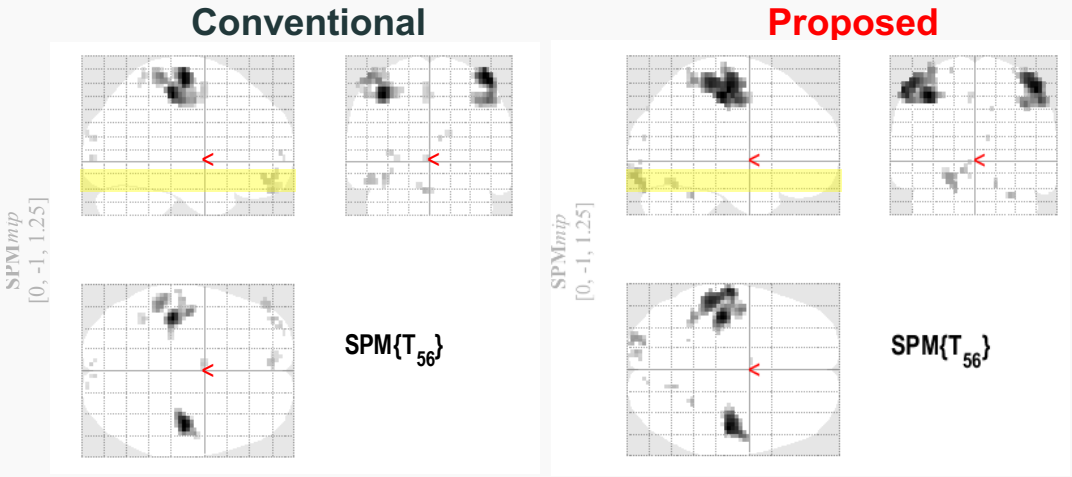
Result : GRE-EPI in-vivo

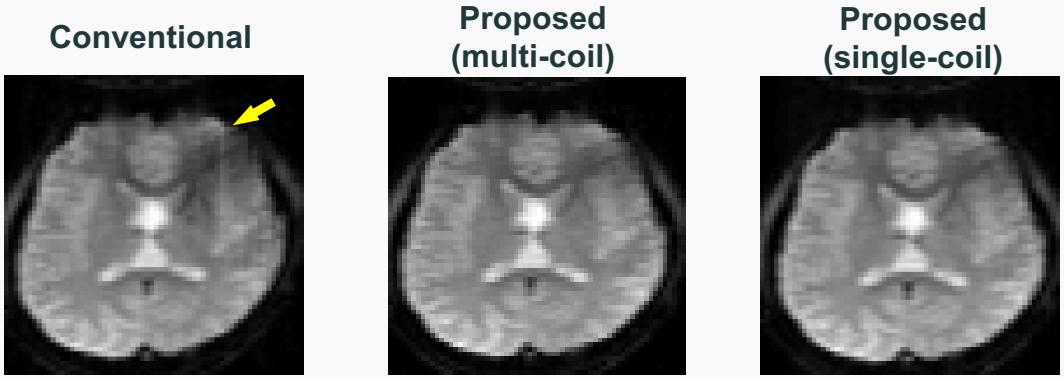


Result : fMRI analysis

fMRI analysis of GRE-EPI using SPM

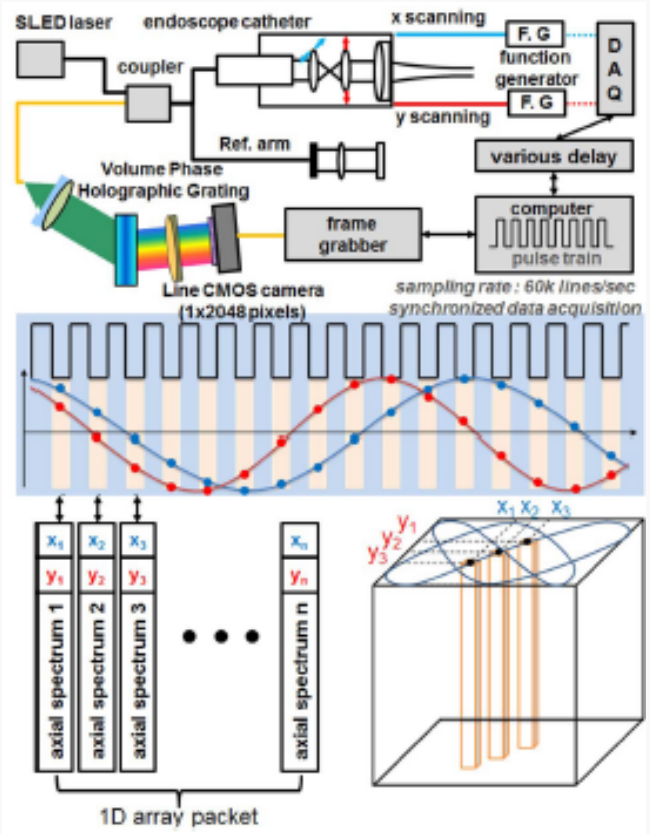
- Pair hand squeezing stimulation
  ⇒ Motor cortex activation
- Familywise error, $p \leq 0.05$



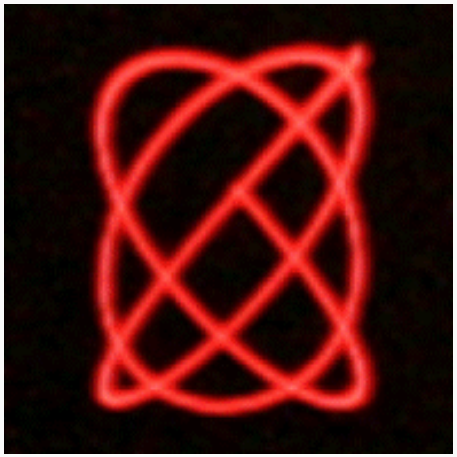


Applications to Image Processing

Inpainting & Impulse noise removal



ig. 2. (Color online) 3D SD-OCT image reconstruction



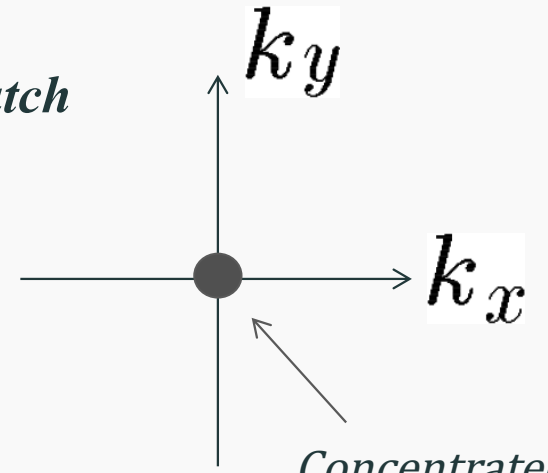




Spectral Domain Sparsity

Smooth patch



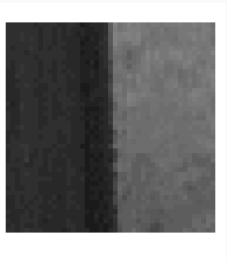


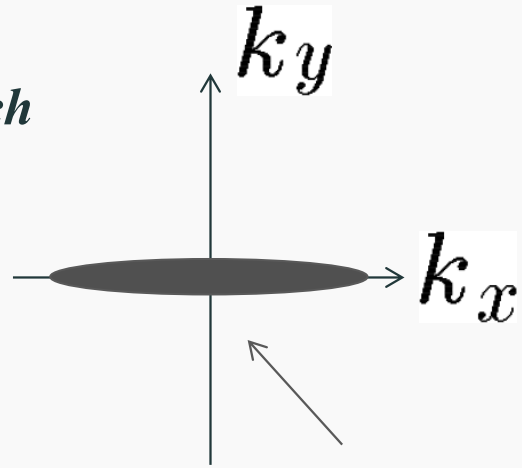
Concentrated at DC

Smoothness, texture, pattern

Sparse spectrum

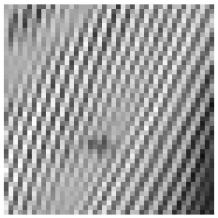
Edge patch

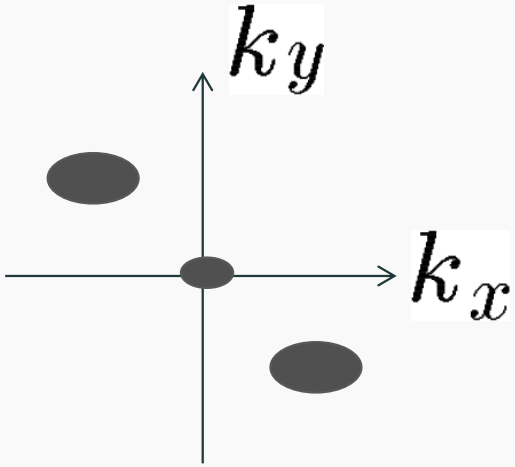




Elongated along Perpendicular to edge

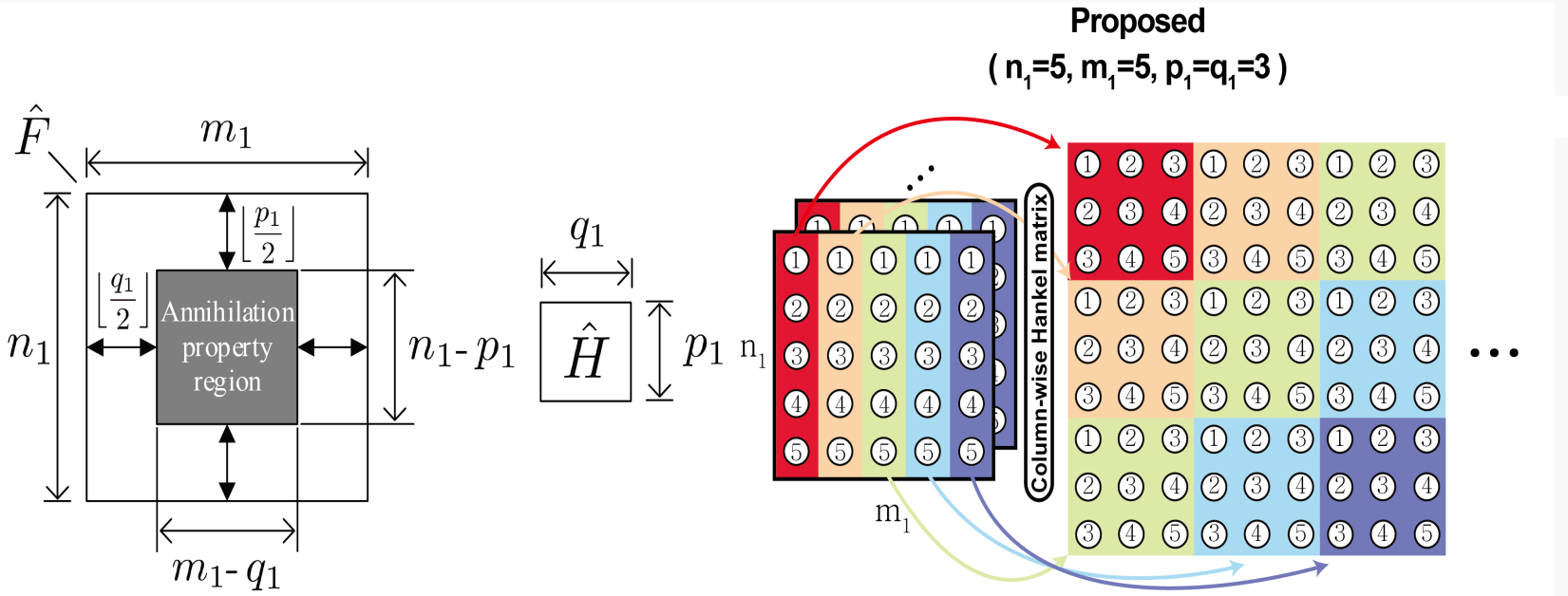
Texture patch





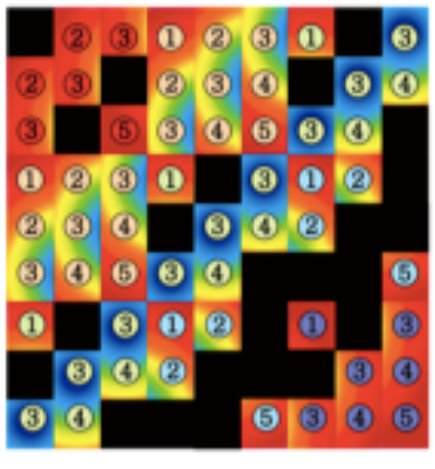
Distributed orthogonally w.r.t. texture

2-D Hankel matrix



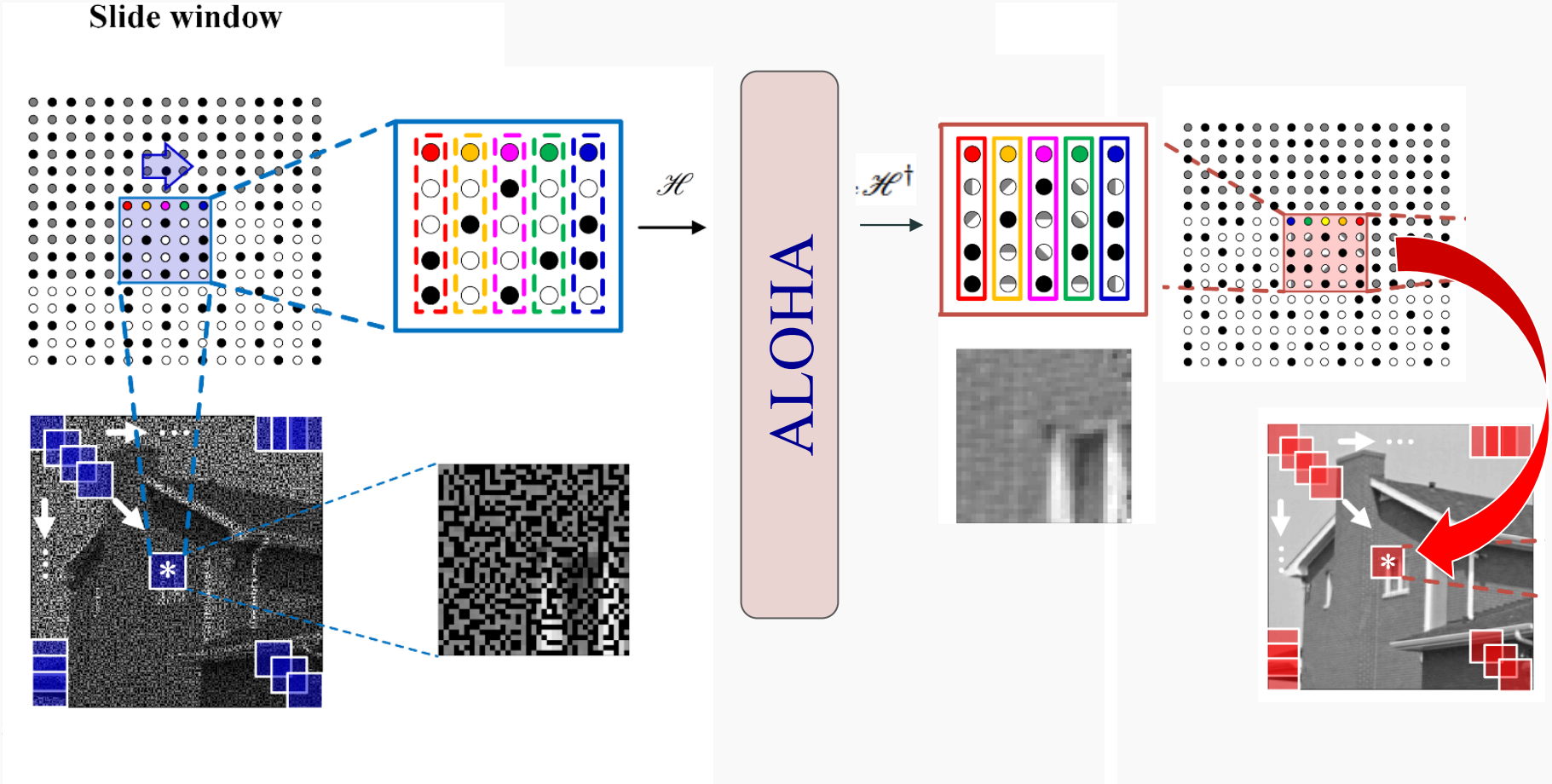


Hankel Matrix construction

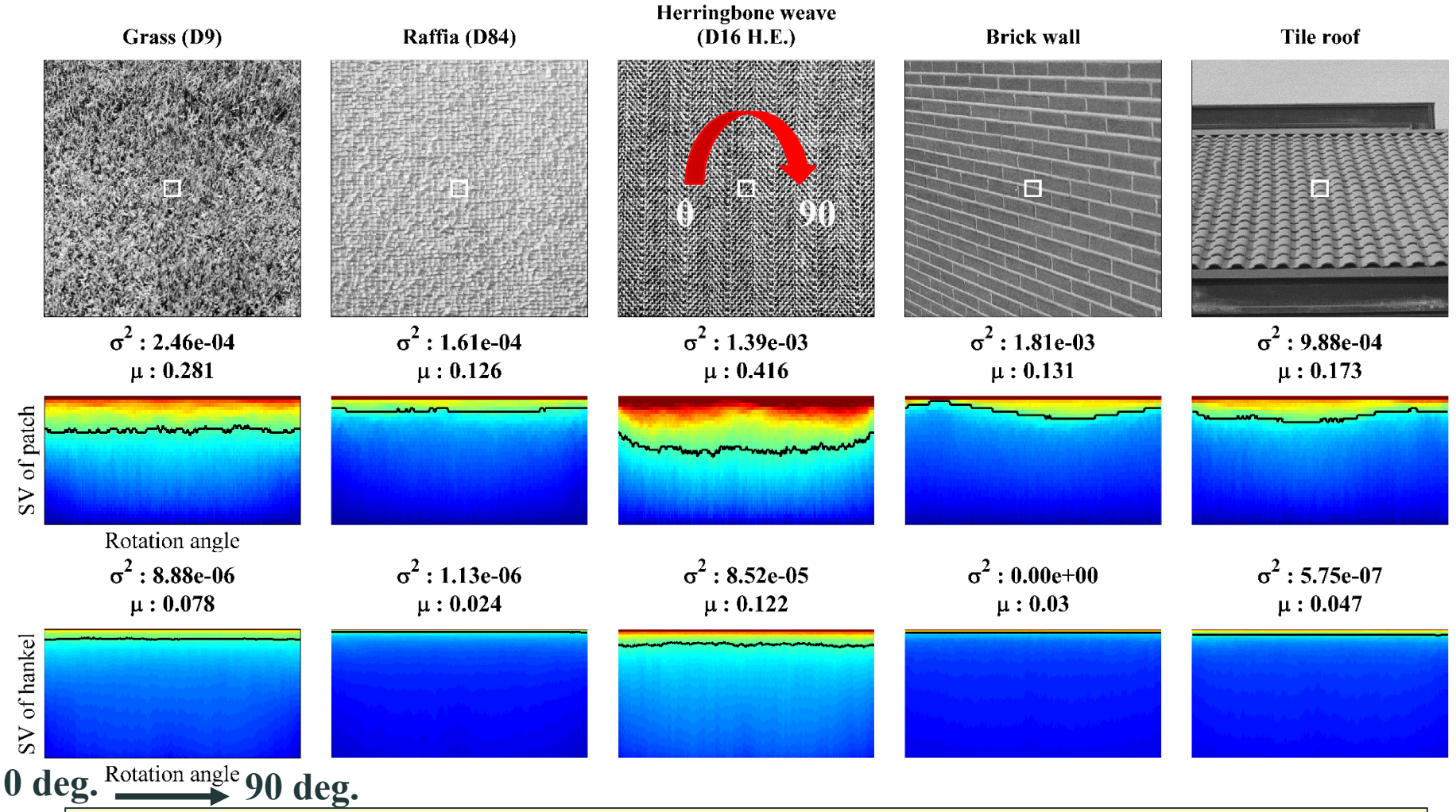


Why patch processing ?

- *Spectrum changes for each patch*
- *Need to adapt the local Image statistics*



Rotation invariant sparsity



Hankel structured matrix is intrinsic low rank !!

Experimental results (x5)

Barbara



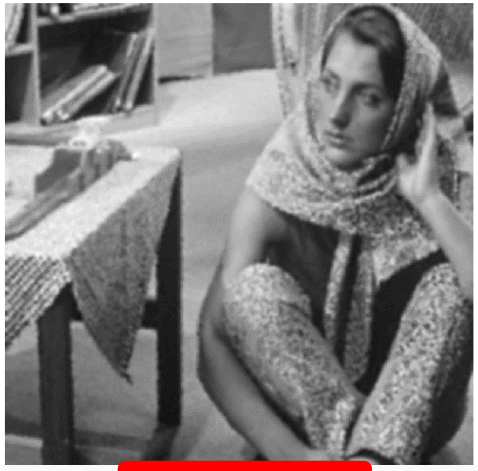
Missing (80%)
PSNR 6.54
SSIM 0.06374



Mesh
PSNR 23.09
SSIM 0.8042



Kernel
PSNR 22.91
SSIM 0.7551



Kernel (Steering)
PSNR 23.08
SSIM 0.7491



K-SVD
PSNR 24.27
SSIM 0.8075



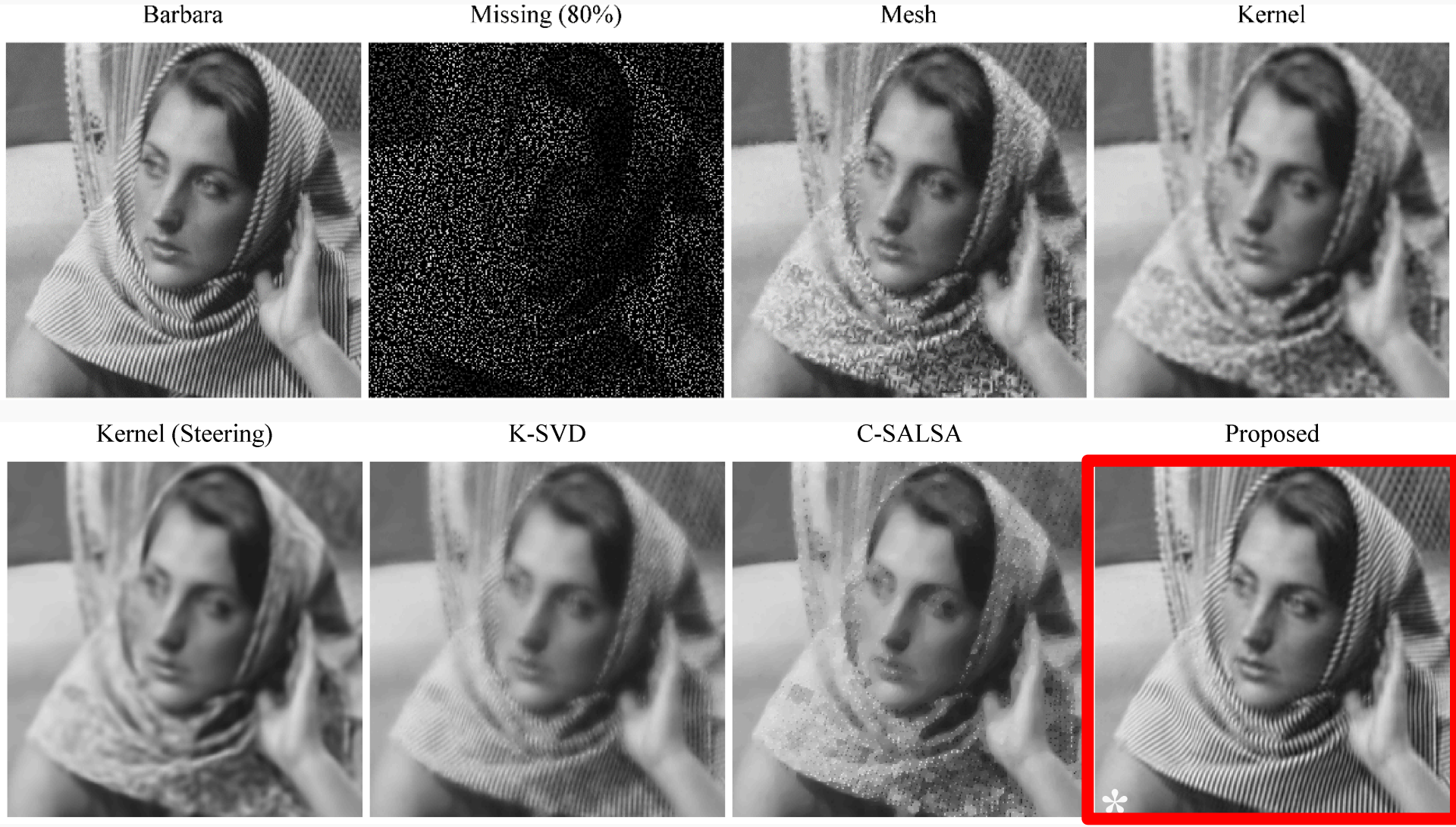
C-SALSA
PSNR 23.38
SSIM 0.7918



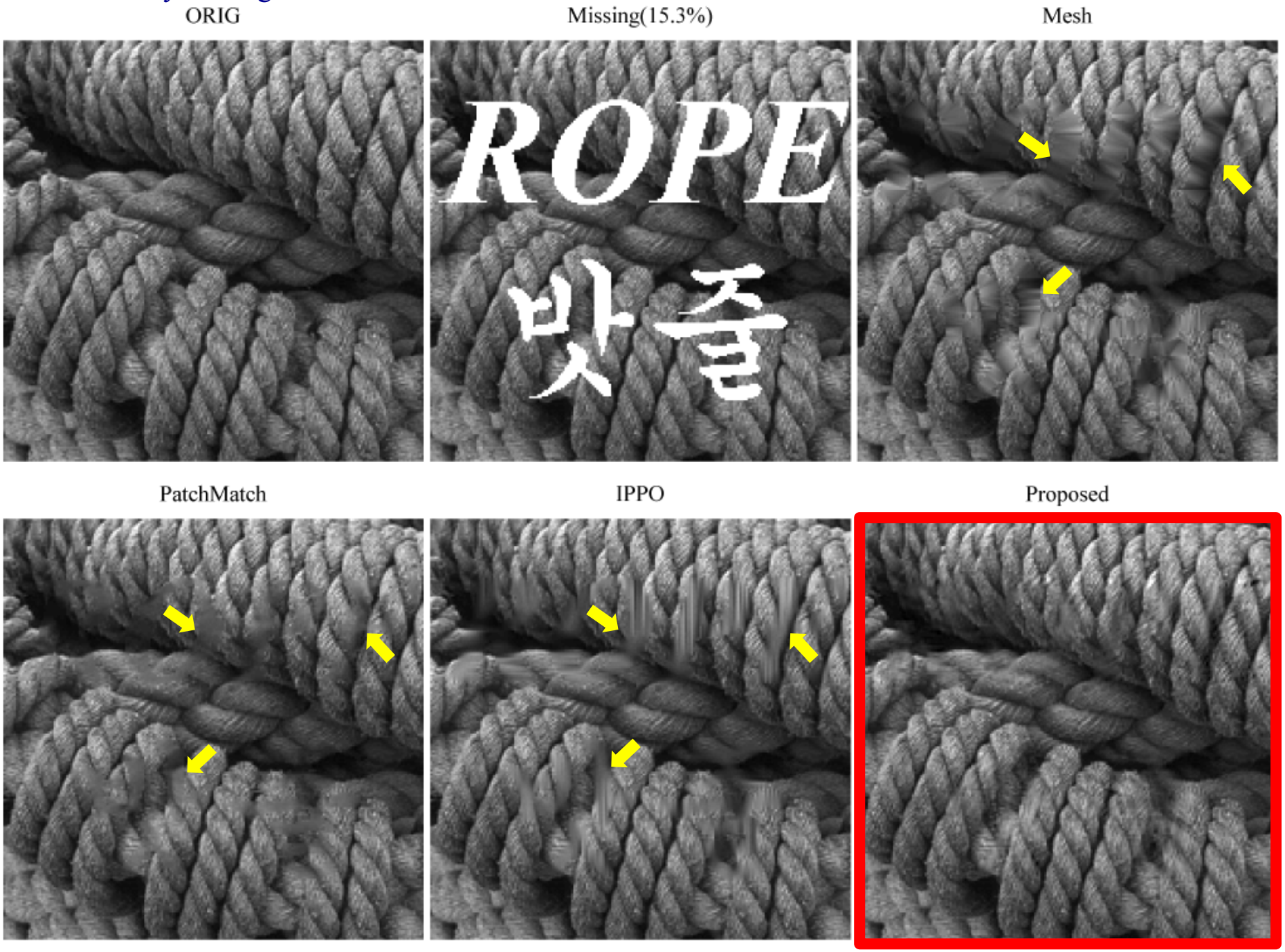
Proposed
PSNR 31.34
SSIM 0.9547



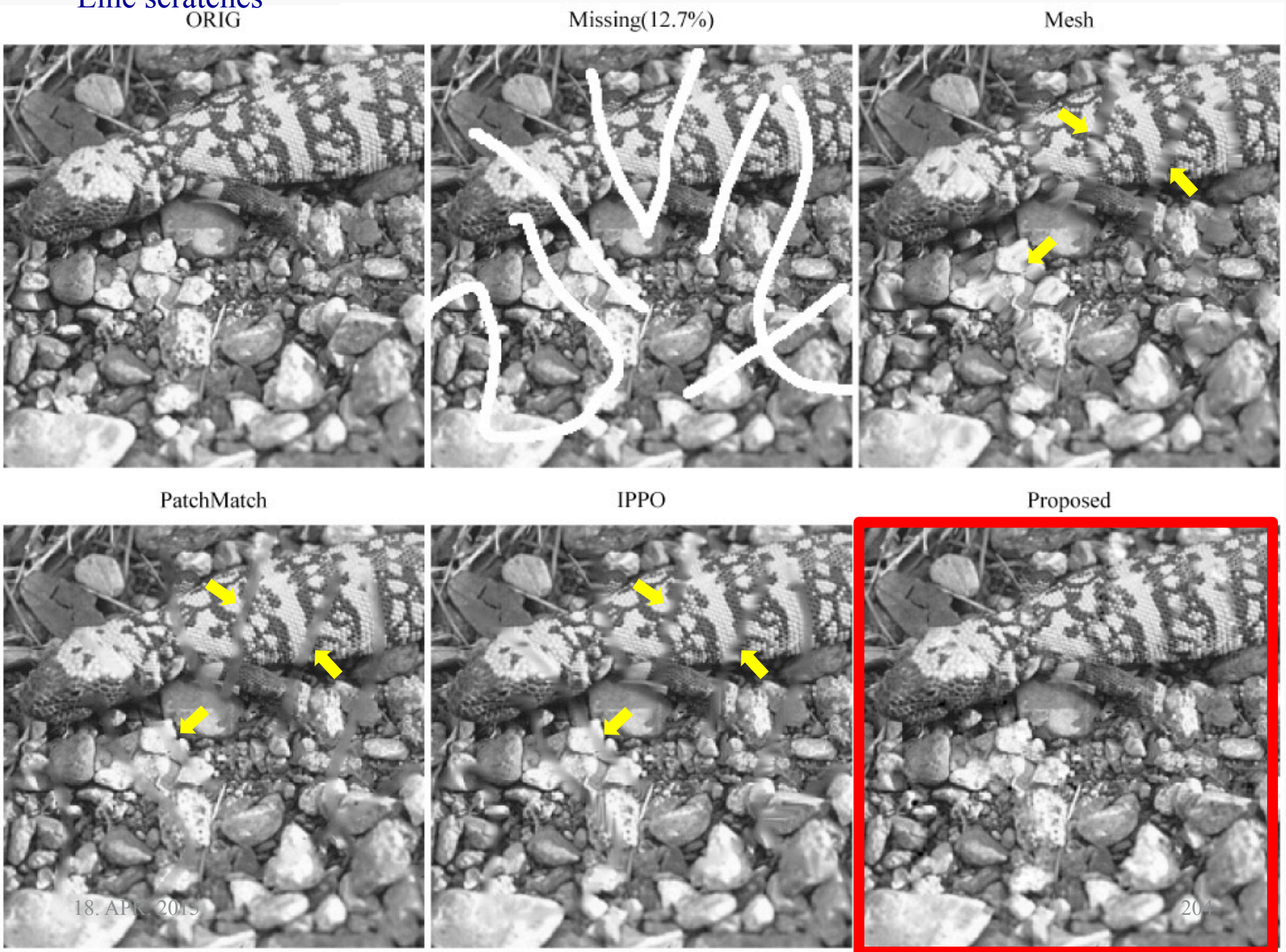
Experimental results (x5)



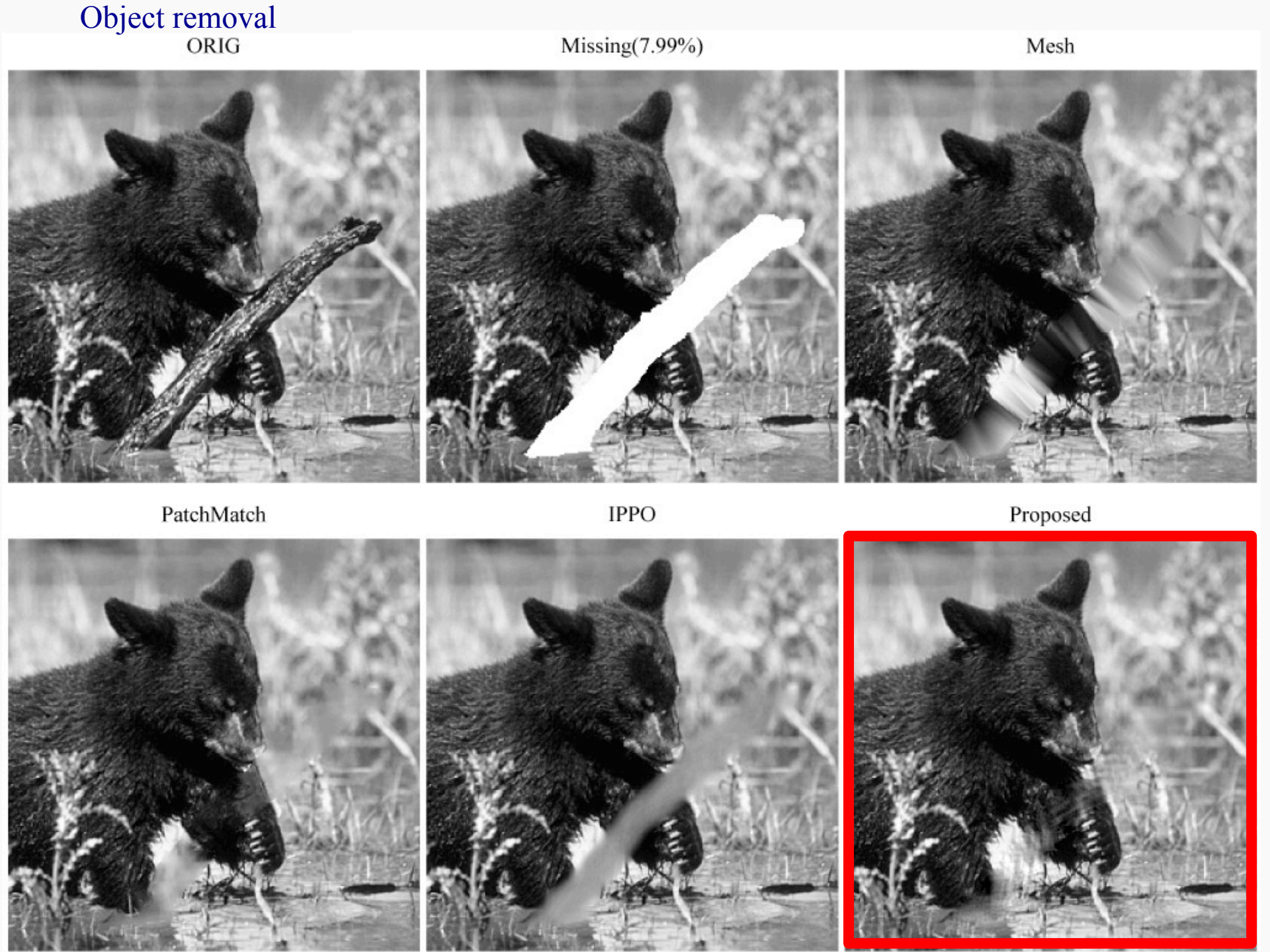
Text inlayed image reconstruction



Line scratches



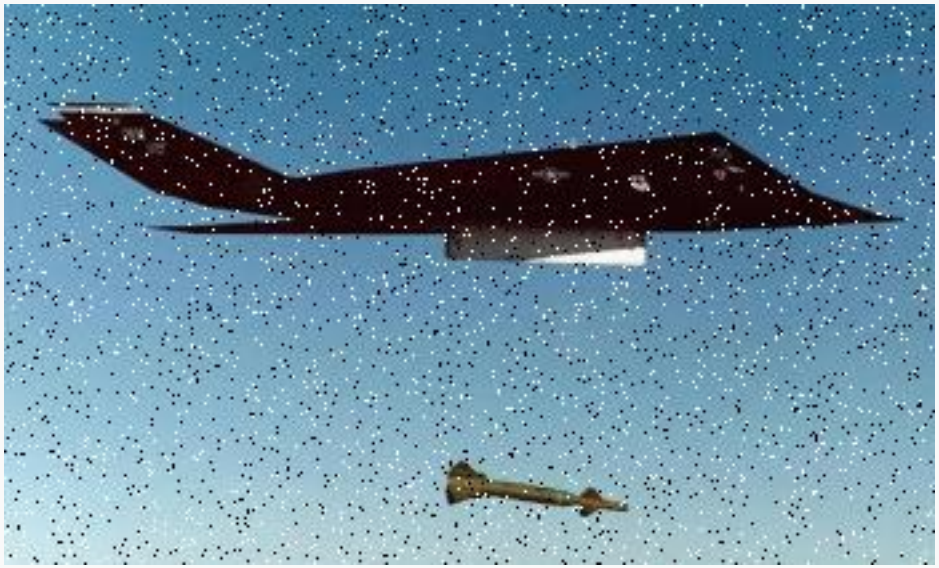
Object removal



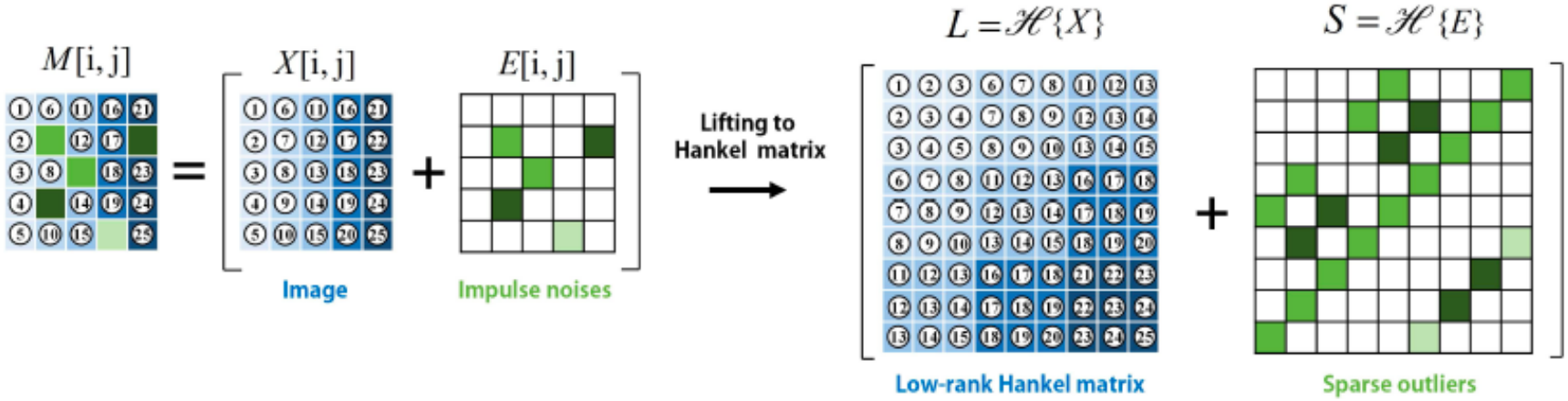


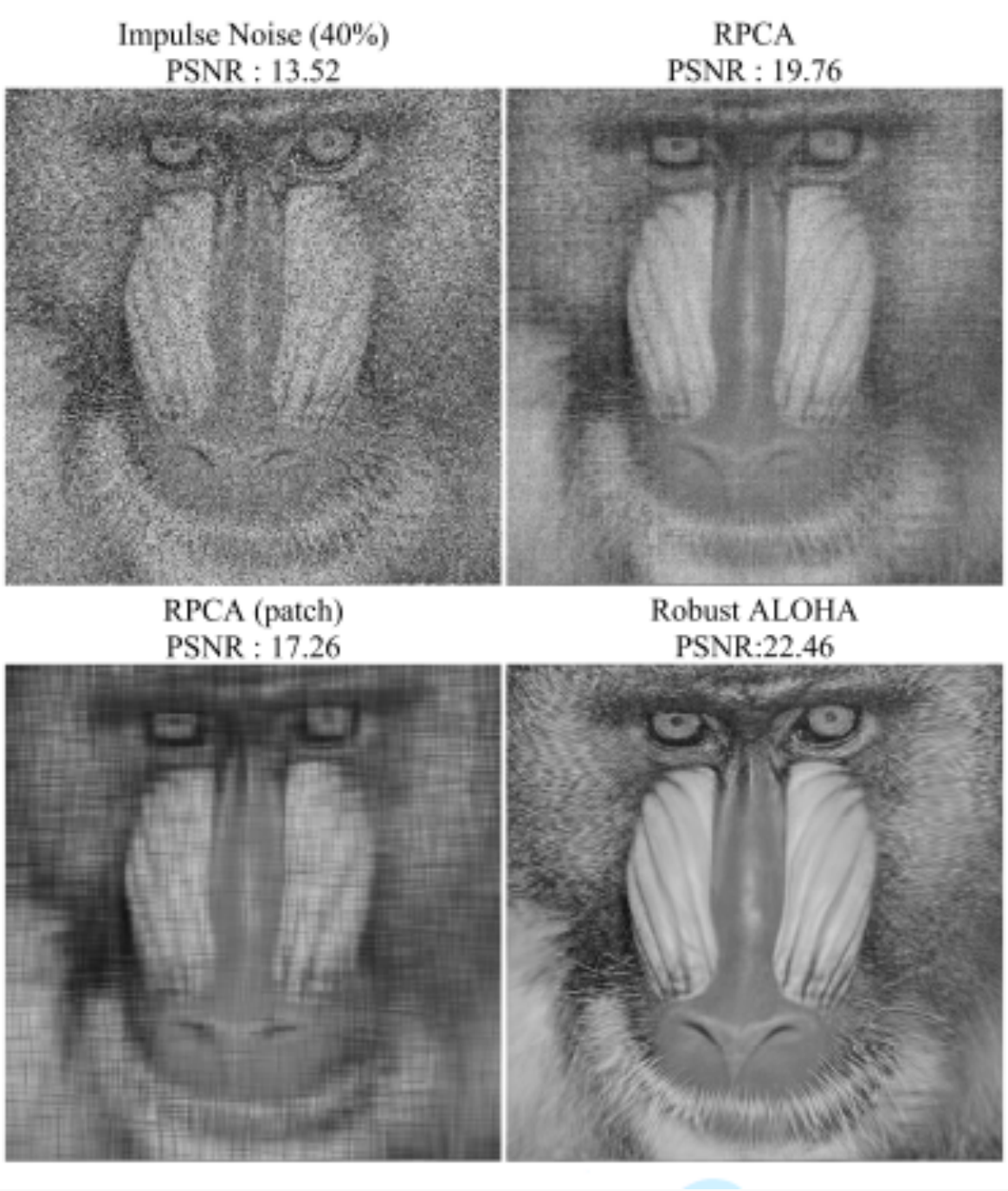
*K.H. Jin, et. al, IEEE TIP (2015)

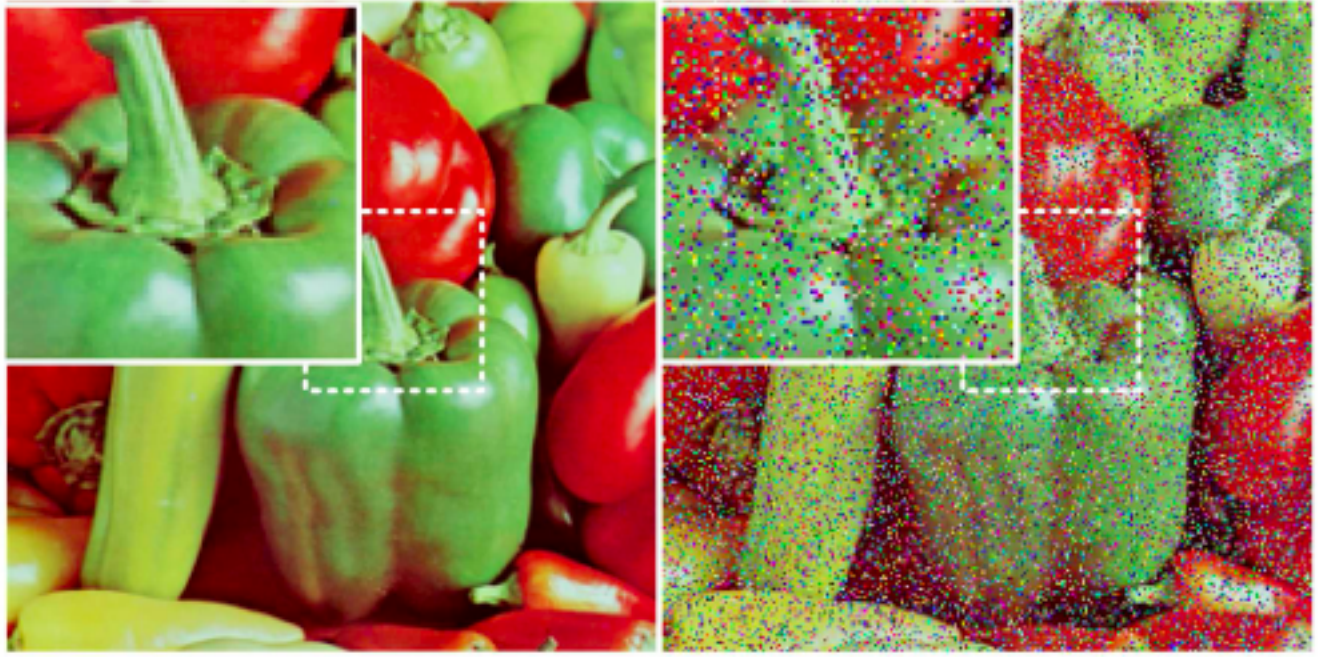
Impulse Noise Removal








Impulse Noise (40%)
PSNR : 13.52
RPCA
PSNR : 19.76
RPCA (patch)
PSNR : 17.26
Robust ALOHA
PSNR:22.46

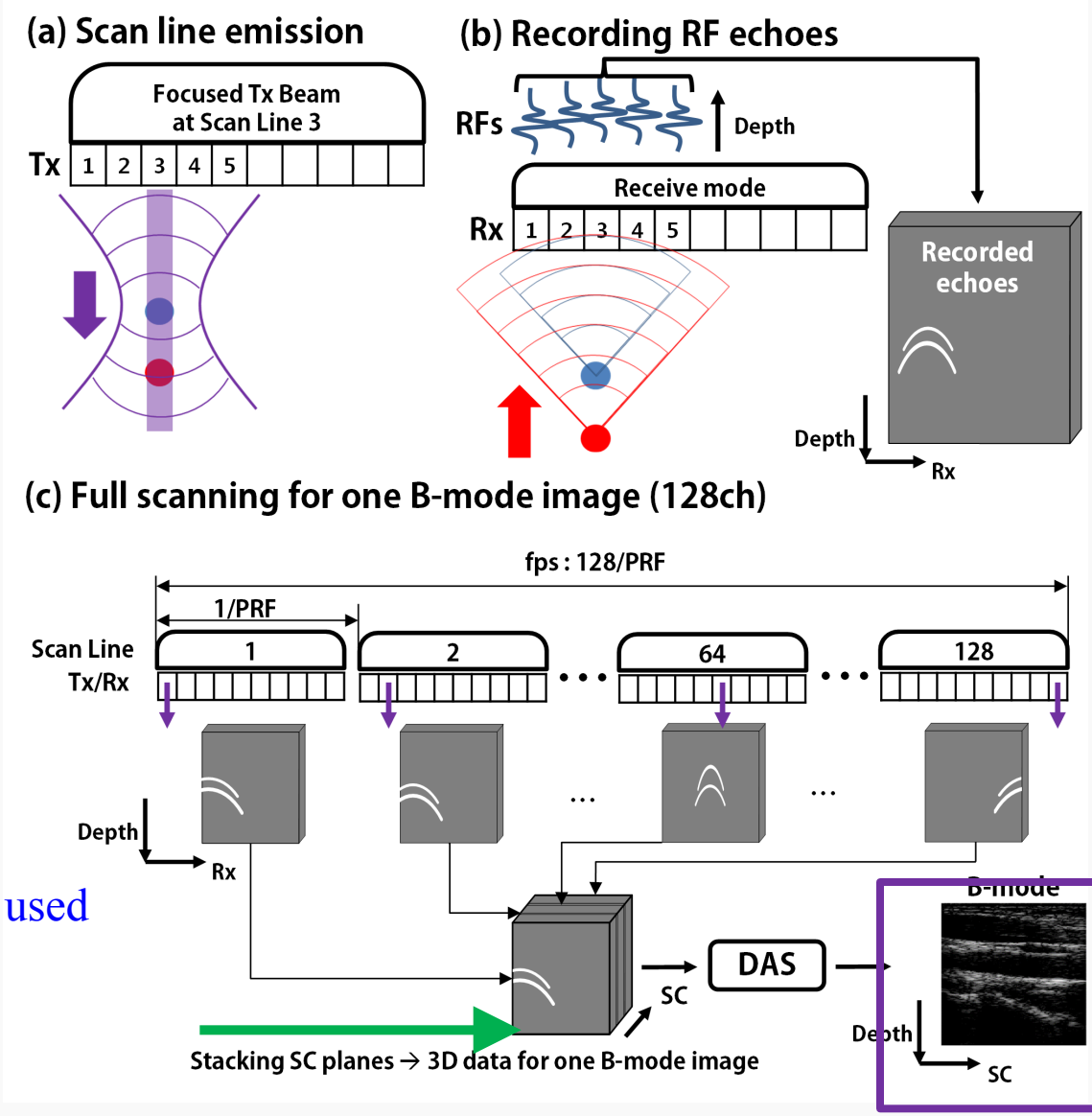


TVL1
PSNR : 28.68

Robust ALOHA
PSNR : 29.68



Application to B-mode US Imaging

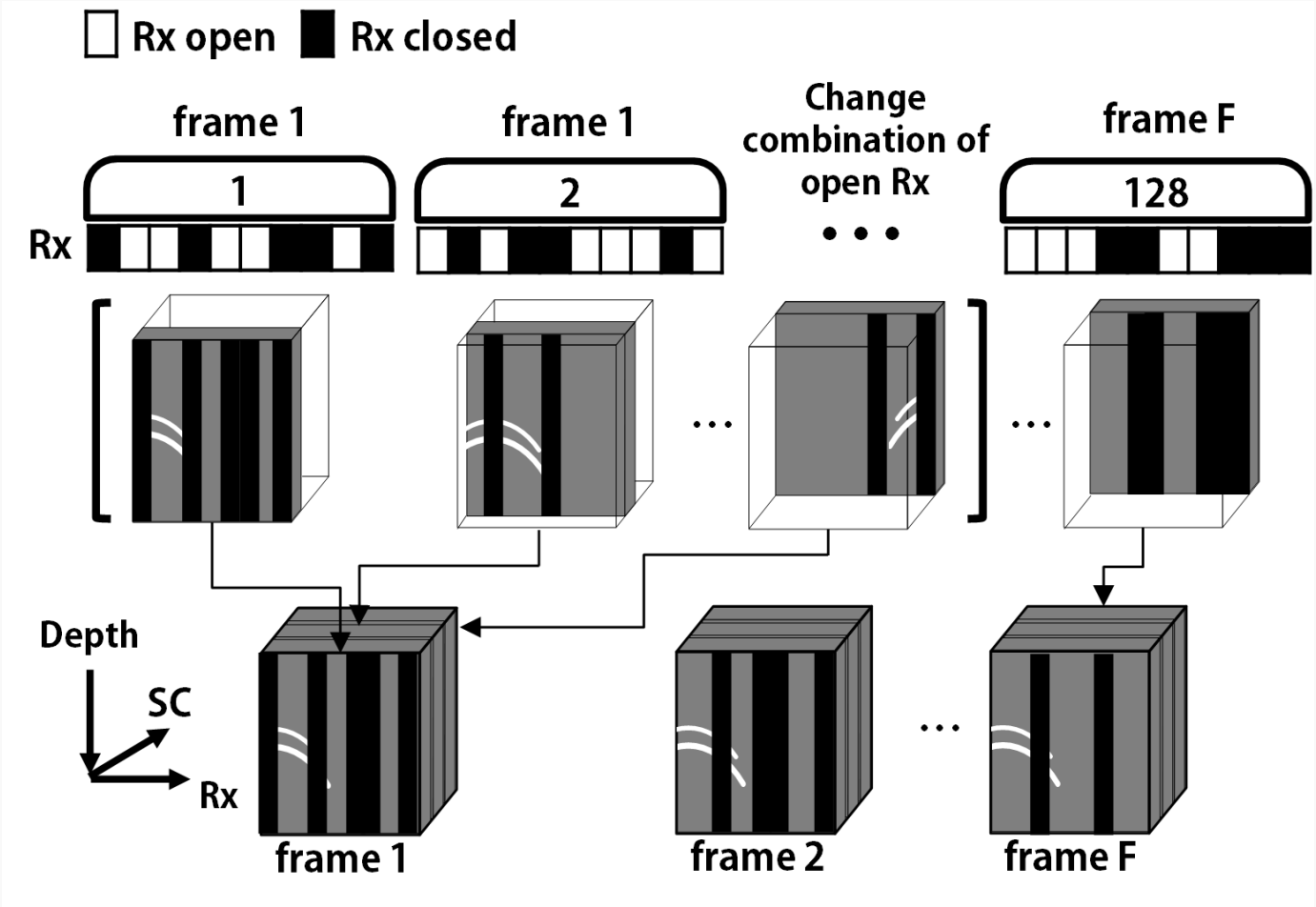


All Rx should be used
→ high power consumption, High data rate

1. Probes deliver beamformed B-mode image, only.

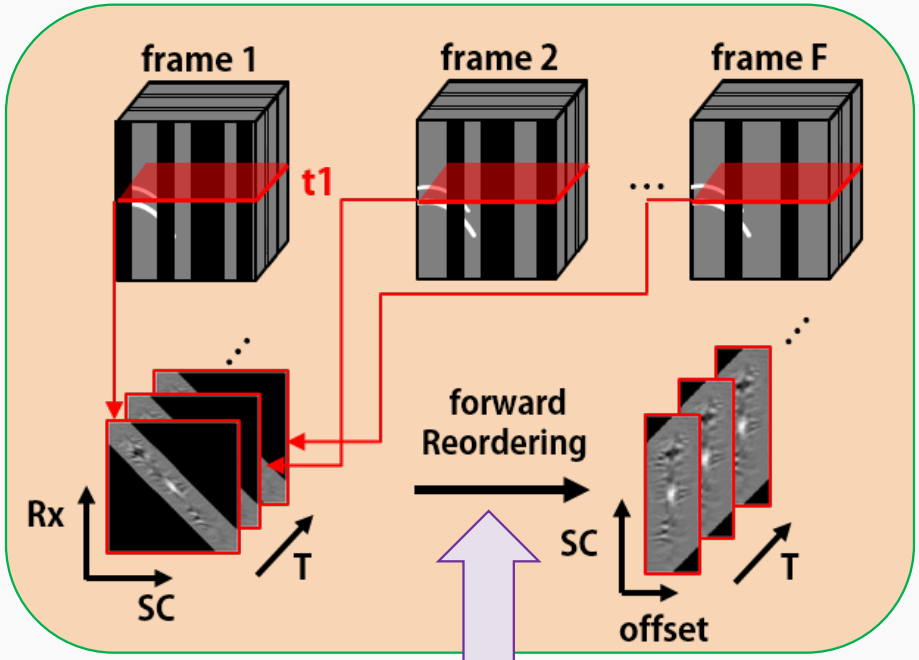
2. After DAS, raw measurements discarded.

Sub-sampled Dynamic Aperture B-mode Imaging



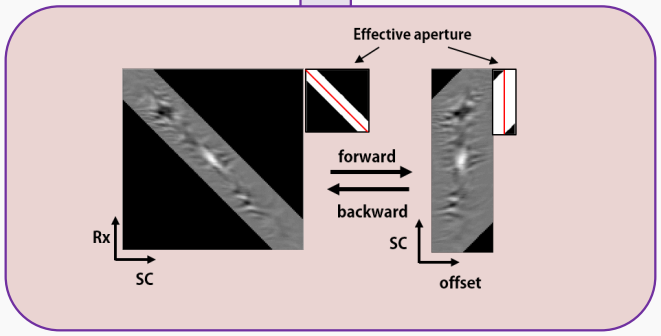
Low-Rankness of B-mode US Data

Temporal slices of
Pre-beamformed RF data



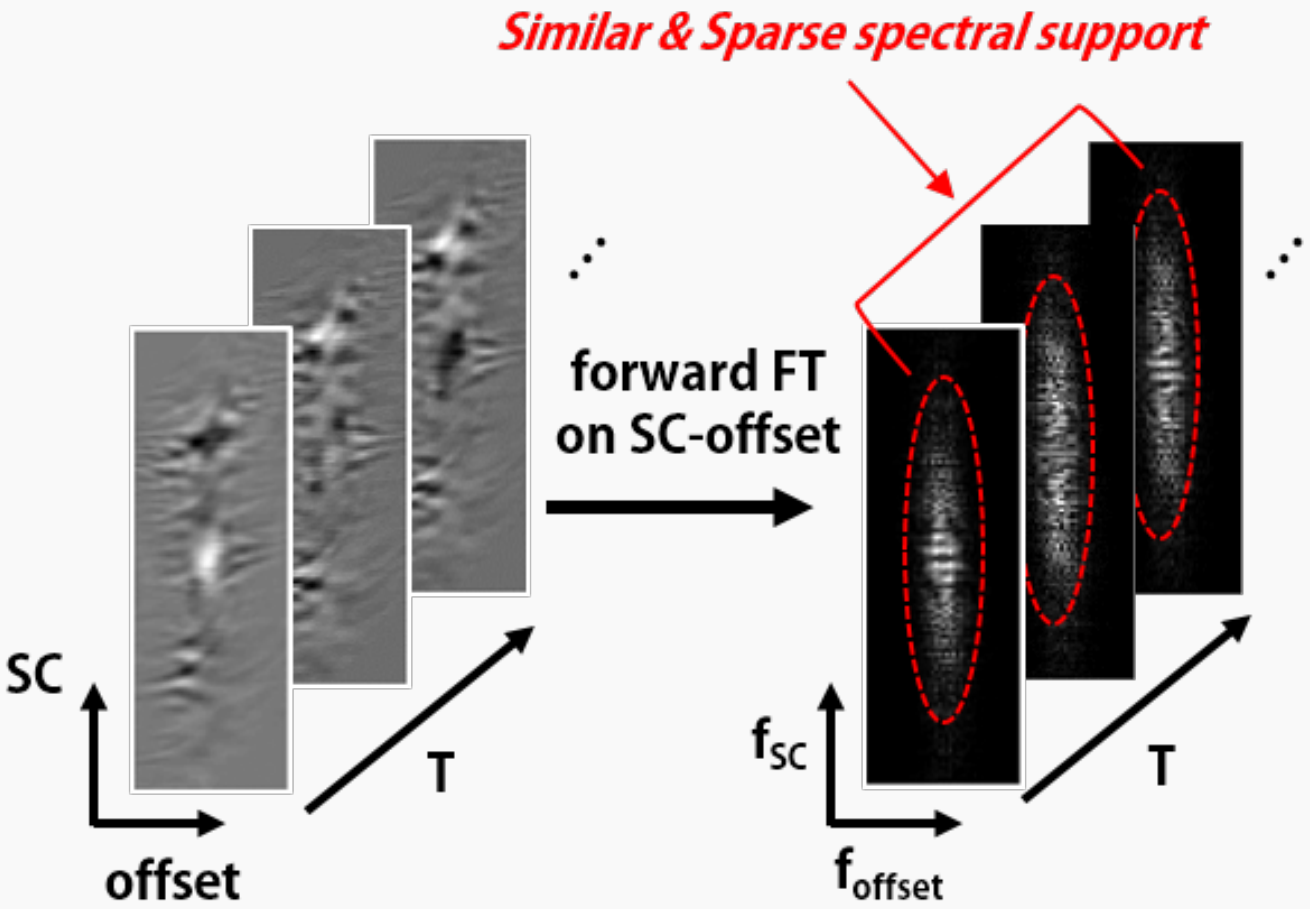
Series of
reordered data
→
Stacked
Hankel matrix

Multi-channel
ALOHA
Interpolation



Reordering of pre-beamformed
into scanline-offset domain
for low-rank property

Sparsity of the Spectrum



Low-Rankness of B-mode US Data

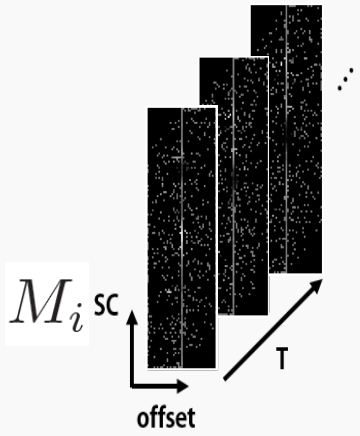
Exploiting Temporal Redundancy

→ inter-temporal annihilating filter

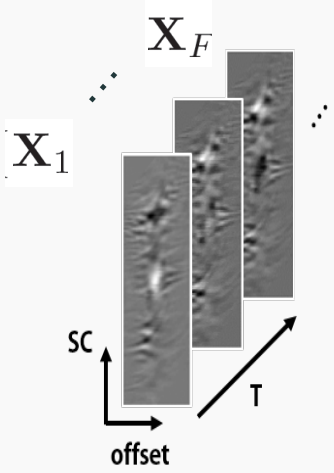
$$\min_{\mathcal{X}} \quad \|\mathcal{Z}\{\mathcal{X}\}\|_*$$

$$\text{subject to} \quad \mathcal{Z}\{\mathcal{X}\} = [\mathscr{H}\{\mathbf{X}_1\} \cdots \mathscr{H}\{\mathbf{X}_F\}].$$

$$X_i(j,k) = M_i(j,k),$$



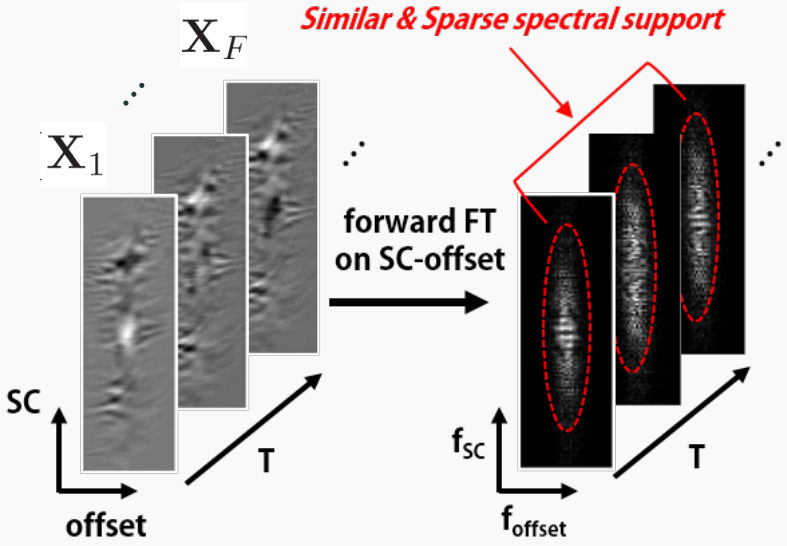
Reordered *subsampled pre-beamformed* data



Reconstructed slices

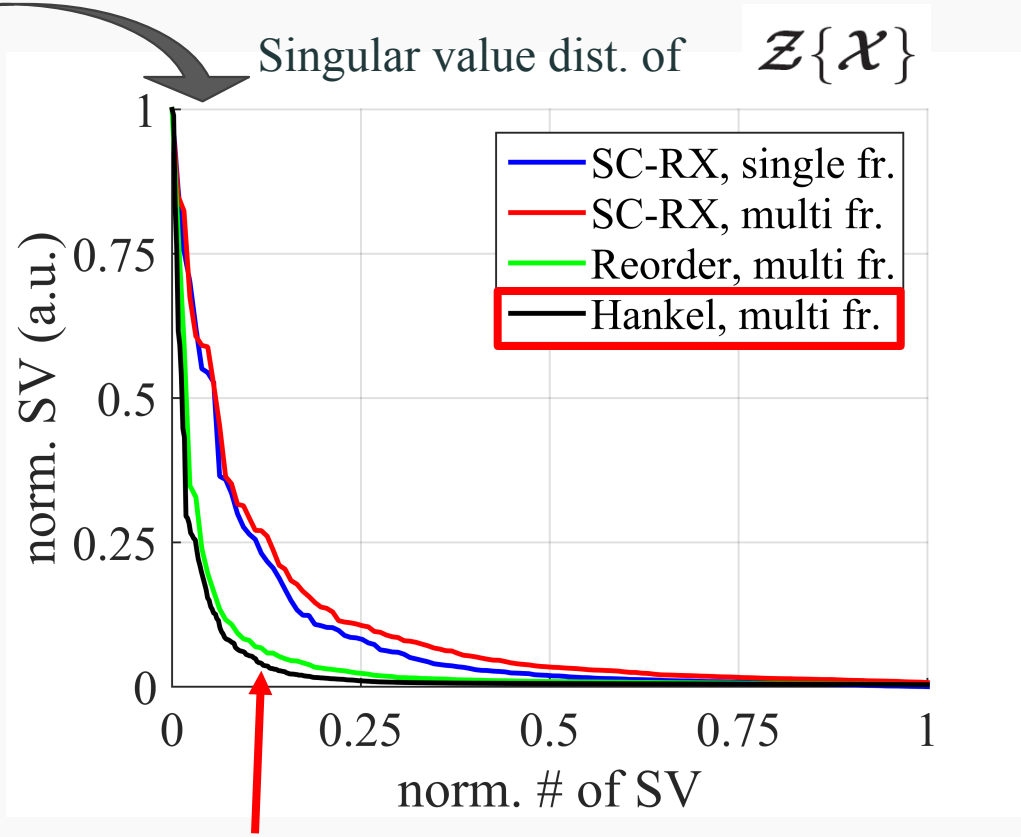
Low-Rankness of B-mode US Data

Matrix →Hankel matrix



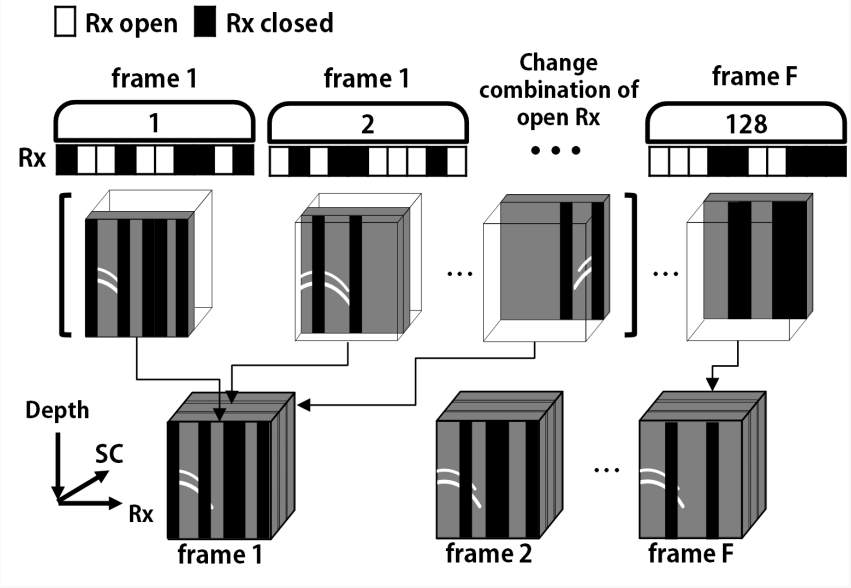
Spatio-temporal redundancy

→ achieved by multichannel Hankel matrix



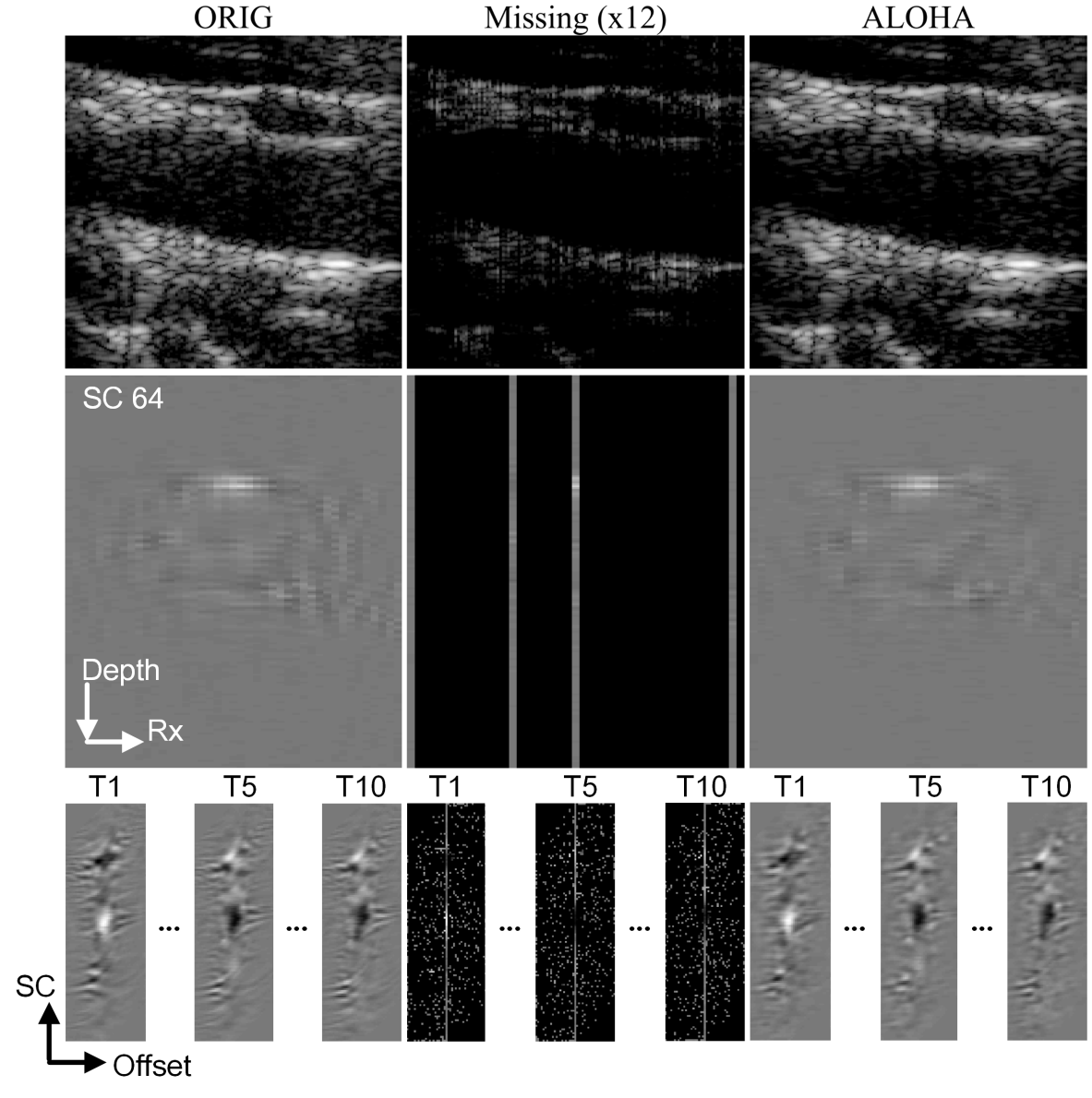
Lowest rank over several settings!

In-vivo Acquisition



- Verasonics system with a Linear type probe (L7-4)
- Center freq:5MHz
- Sampling:20 MHz.
- 128 scanlines (SC) x 128 RX channels
- RX element
  - Width:133um
  - space between RX elements : 158um

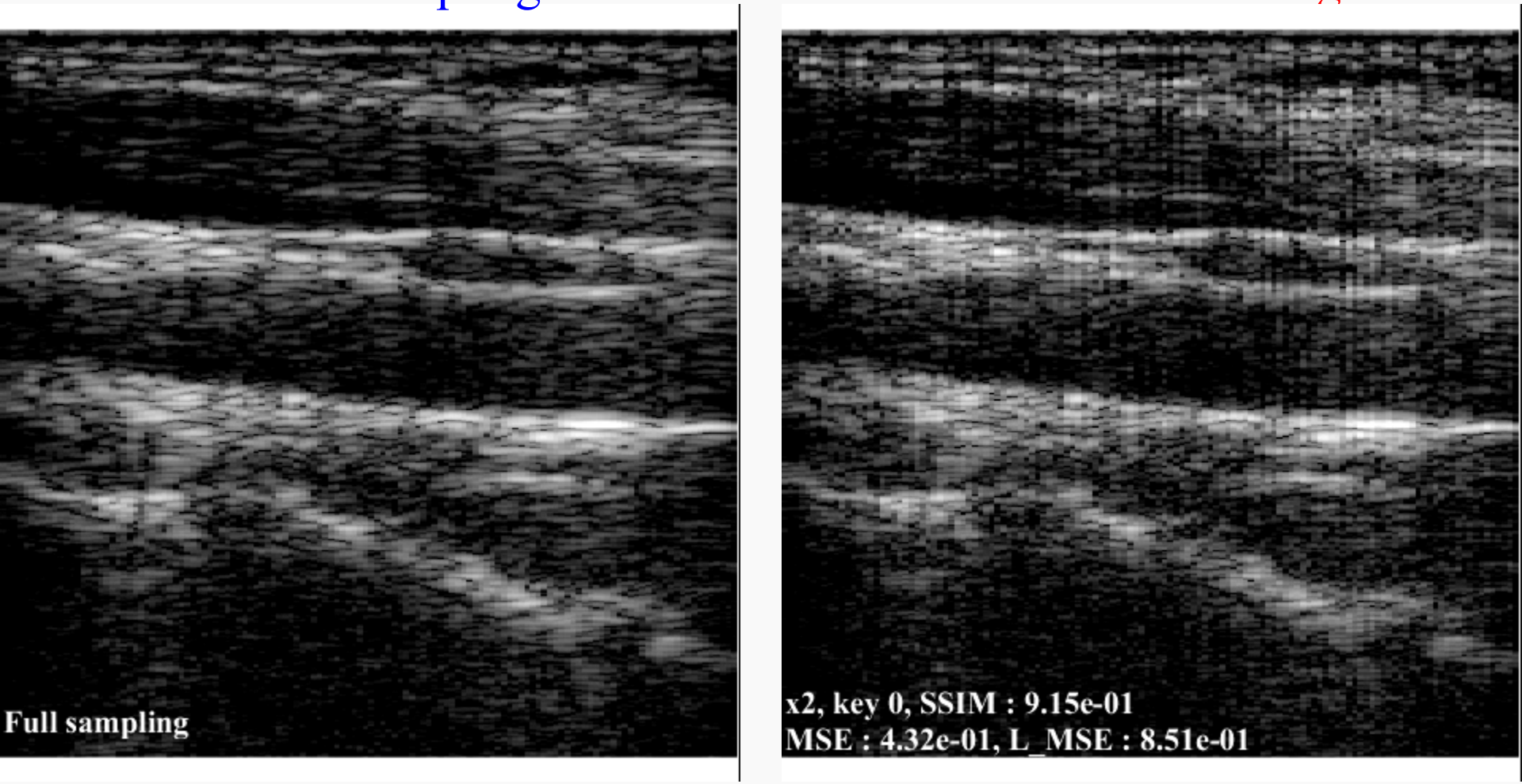
Snapshot image from dynamic scan



Dynamic reconstruction (x2)

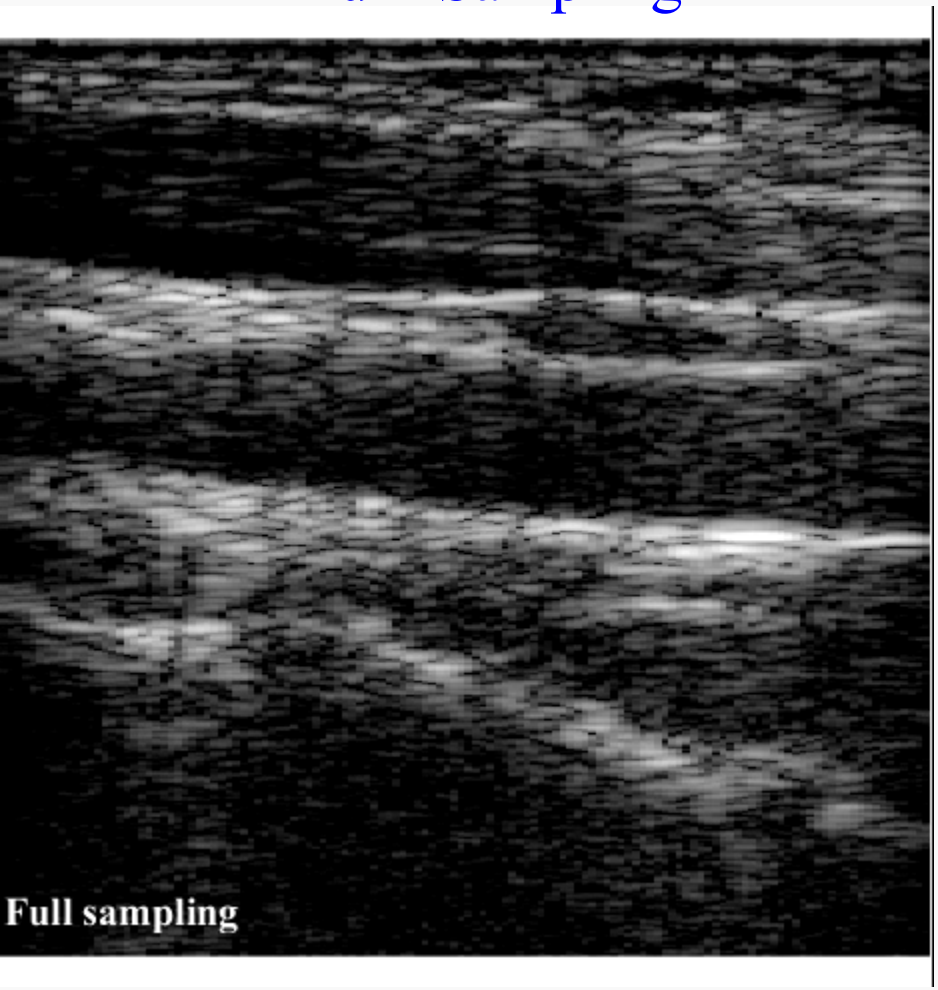
Full Sampling

Beam forming

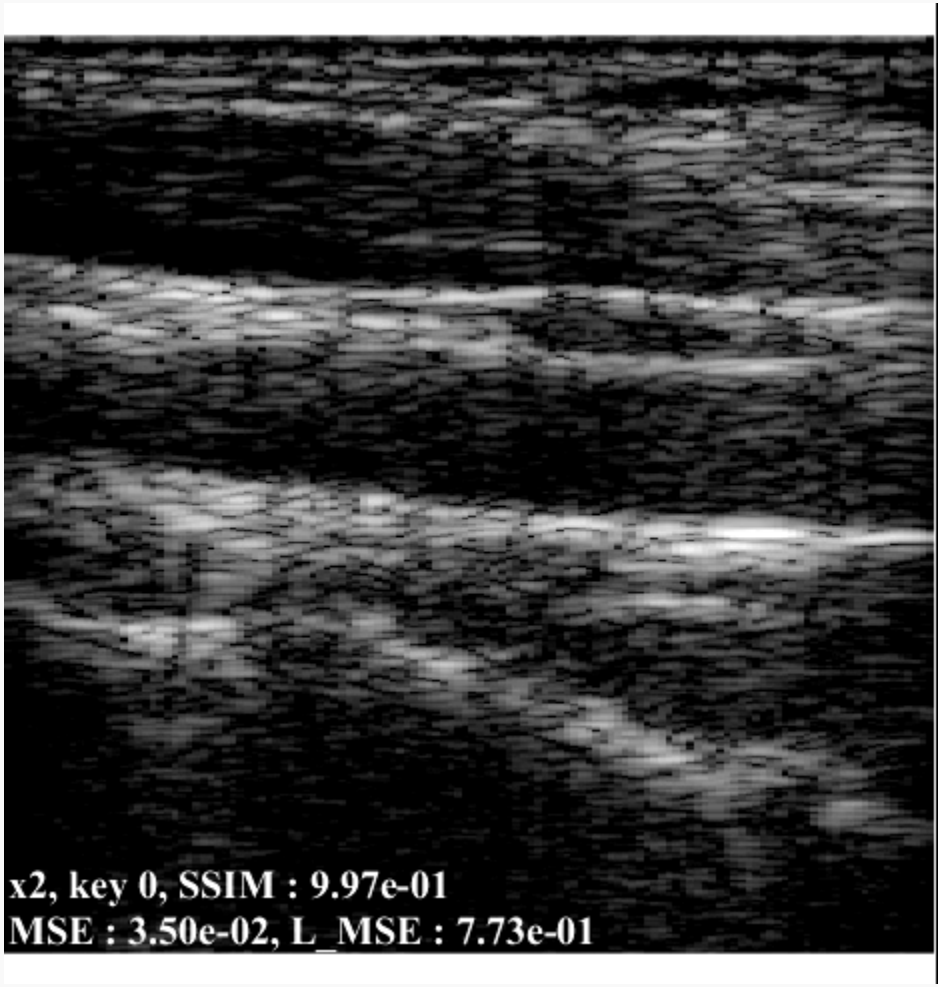


Dynamic reconstruction (x2)

Full Sampling



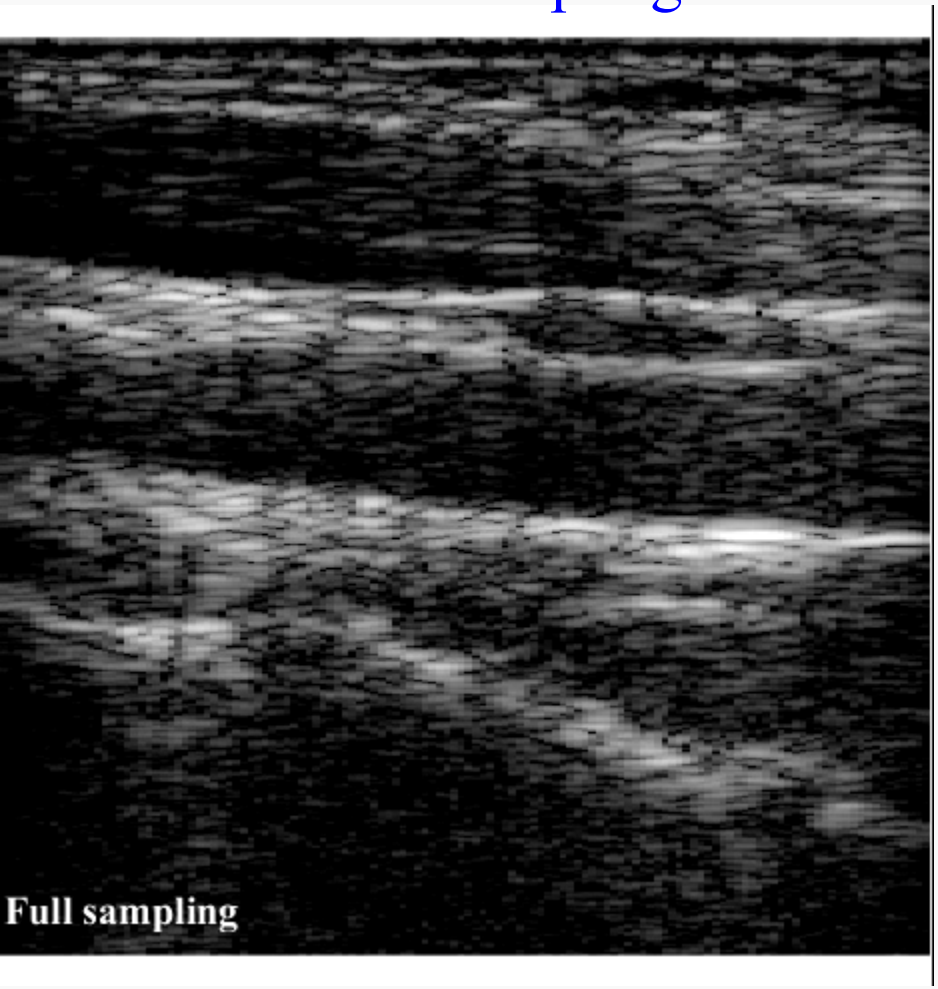
ALOHA

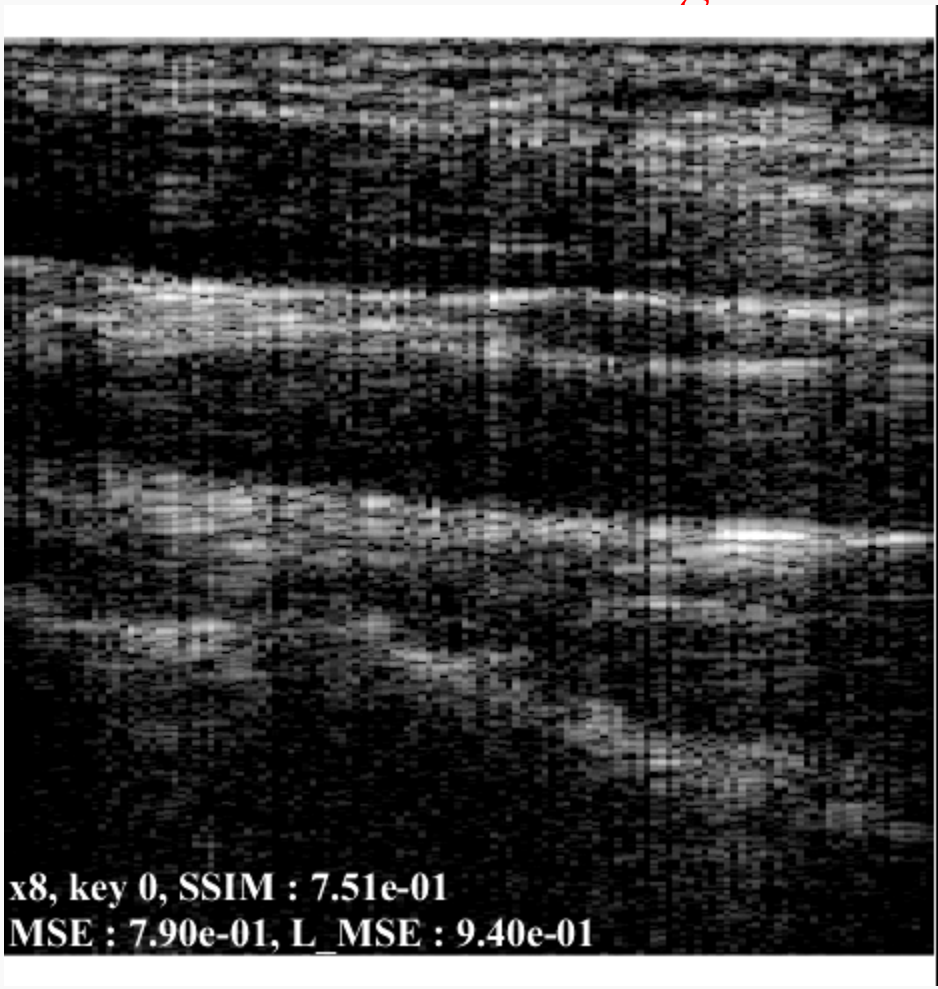


Dynamic reconstruction (x8)

Full Sampling

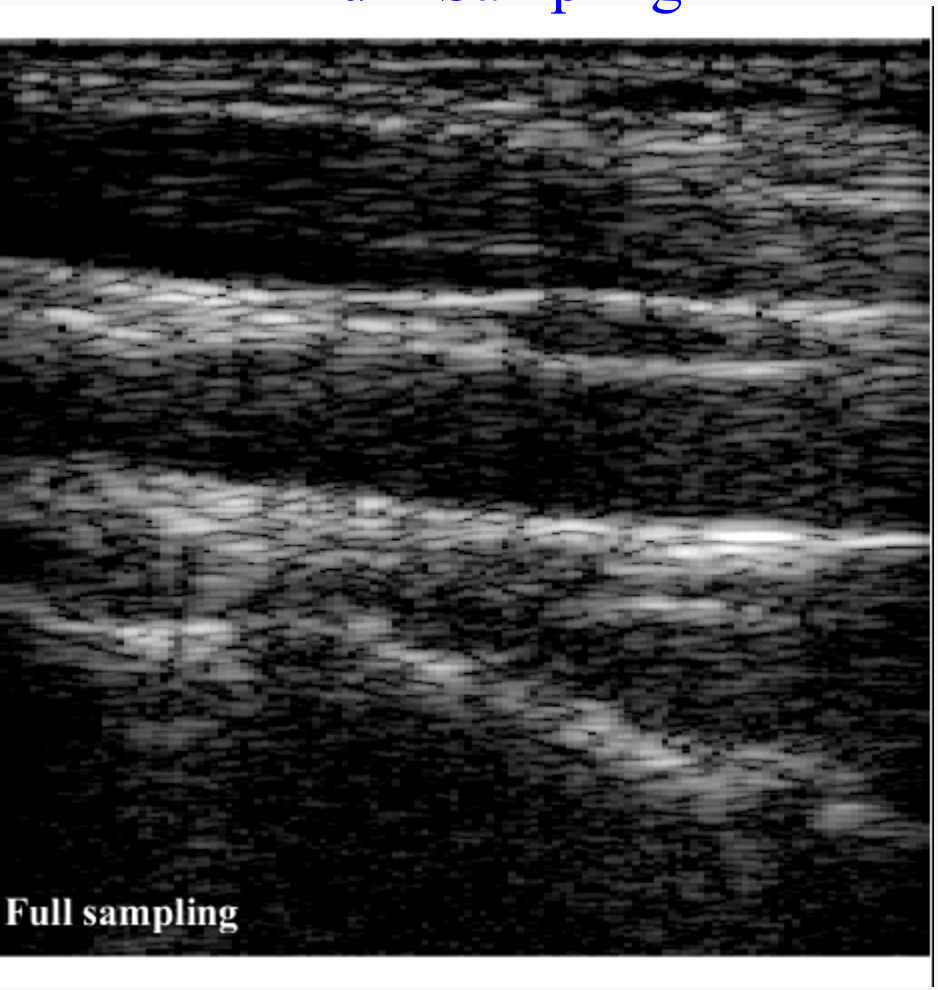
Beam forming



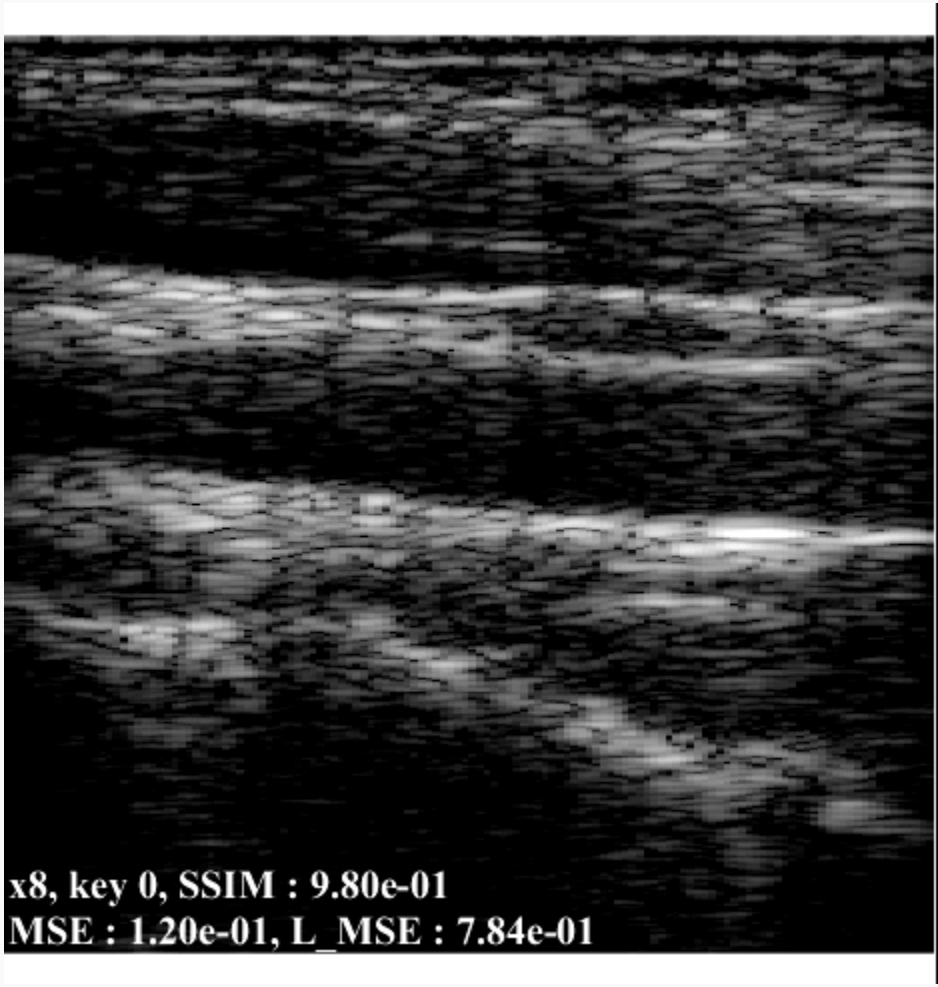


Dynamic reconstruction (x8)

Full Sampling



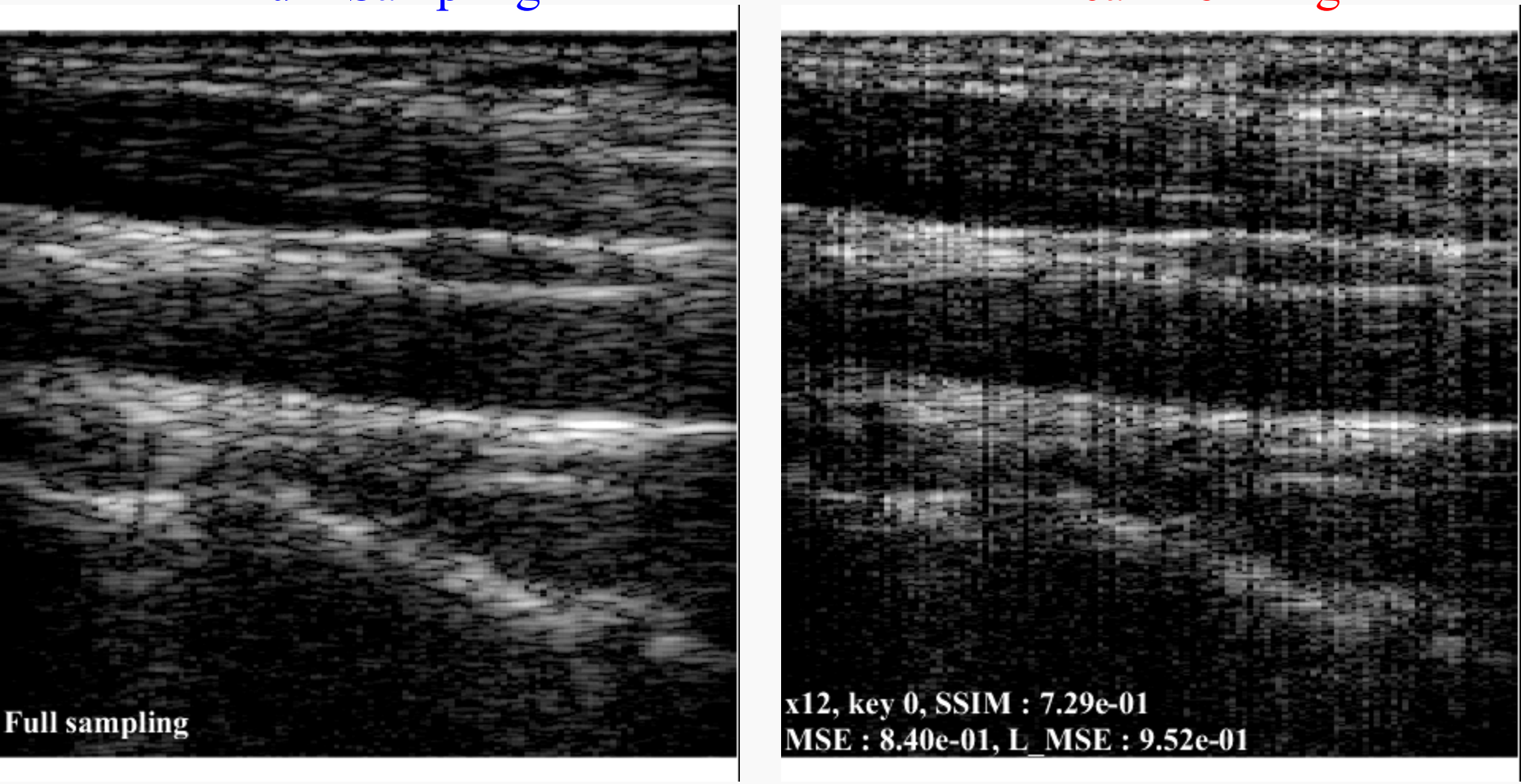
ALOHA



Dynamic reconstruction (x12)

Full Sampling

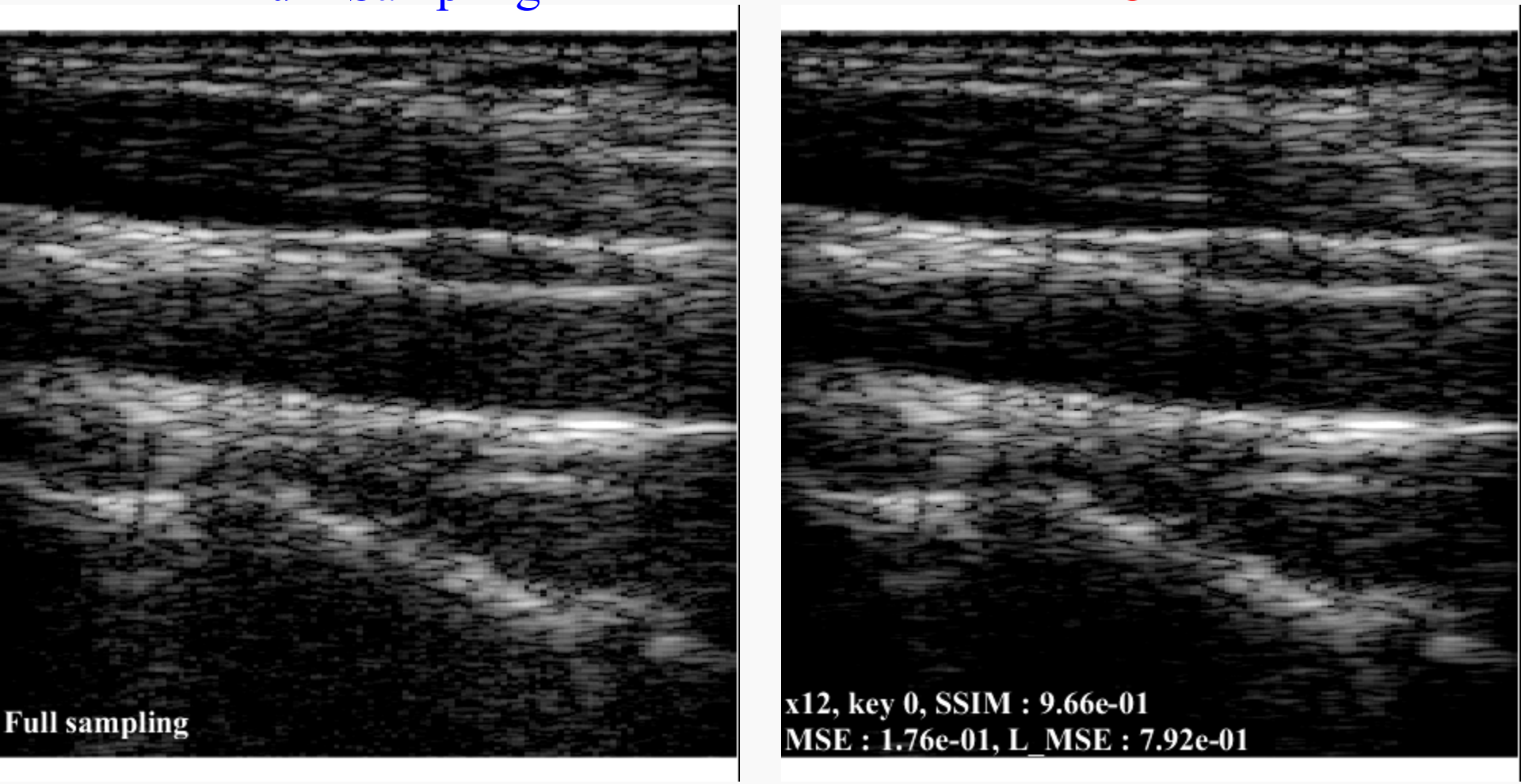
Beam forming



Dynamic reconstruction (x12)

Full Sampling

ALOHA

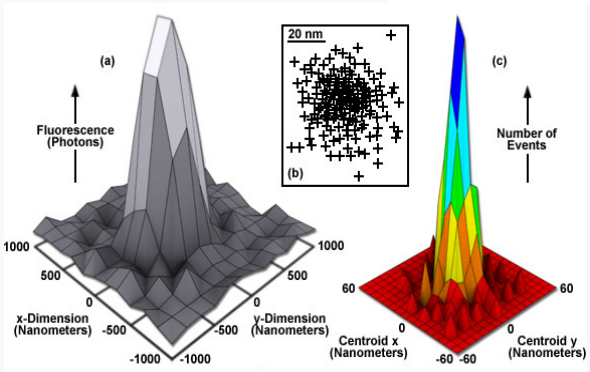


Localization microscopy

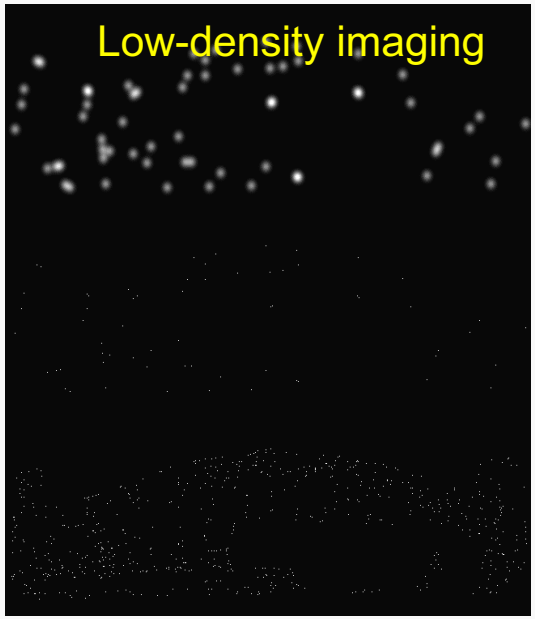
- Nanoscopy based on localization
  - Localization precision is not diffraction limited
  - Sparsely activated probes + localization => super-resolution image

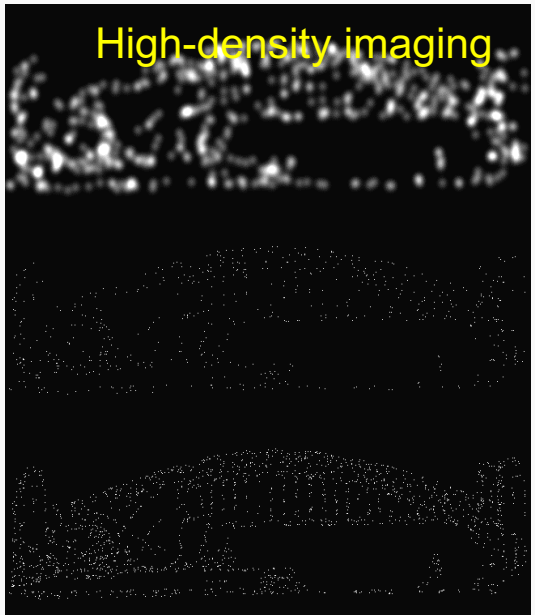
$$\sigma_{\mu_i} = \sqrt{\left(\frac{s_i^2}{N} + \frac{a^2/12}{N} + \frac{8\pi s_i^4 b^2}{a^2 N^2}\right)}$$

Thompson et al. BPJ 2002



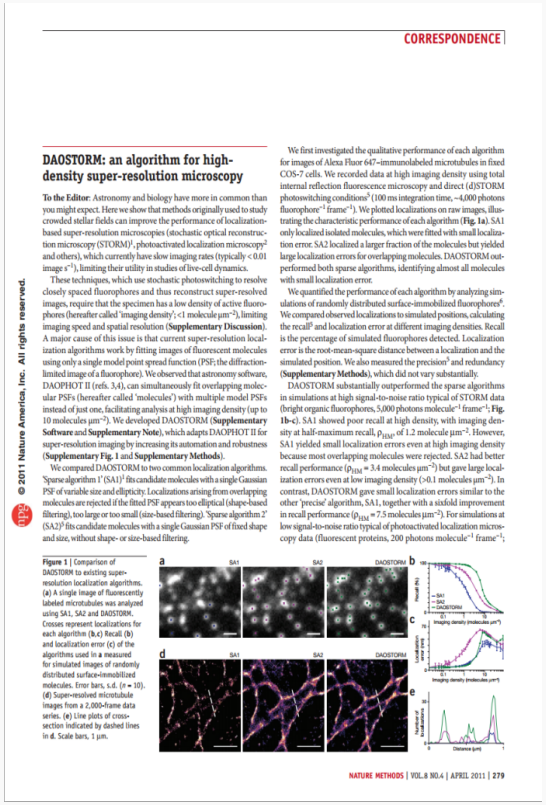
- However, sparse activation scheme has too slow temporal resolution for live imaging
  - Tens of seconds or several minutes
- High-density imaging for fast live imaging
  - Require a robust localization algorithm and system





Existing high density algorithm

Greedy approach



CORRESPONDENCE

DAOSTORM: an algorithm for high-density super-resolution microscopy

To the Editor: Astronomy and biology have more in common than you might expect. Here we show that methods originally used to study crowded stellar fields can improve the performance of localization-based super-resolution microscopies (stochastic optical reconstruction microscopy (STORM)[1], photoactivated localization microscopy[2] and others), which currently have slow imaging rates (typically < 0.01 image s^{-1}), limiting their utility in studies of live-cell dynamics.

These techniques, which use stochastic photoswitching to resolve closely spaced fluorophores and thus reconstruct super-resolved images, require that the specimen has a low density of active fluorophores (hereafter called 'imaging density'; <1 molecule μm^{-2}), limiting imaging speed and spatial resolution (**Supplementary Discussion**). A major cause of this issue is that current super-resolution localization algorithms work by fitting images of fluorescent molecules using only a single model point spread function (PSF; the diffraction-limited image of a fluorophore). We observed that astronomy software, DAOPHOT II (refs. 3,4), can simultaneously fit overlapping molecular PSFs (hereafter called 'molecules') with multiple model PSFs instead of just one, facilitating analysis at high imaging density (up to 10 molecules μm^{-2}). We developed DAOSTORM (**Supplementary Software** and **Supplementary Note**), which adapts DAOPHOT II for super-resolution imaging by increasing its automation and robustness (**Supplementary Fig. 1** and **Supplementary Methods**).

We compared DAOSTORM to two common localization algorithms. 'Sparse algorithm 1' (SA1)[5] fits candidate molecules with a single Gaussian PSF of variable size and ellipticity. Localizations arising from overlapping molecules are rejected if the fitted PSF appears too elliptical (shape-based filtering), too large or too small (size-based filtering). 'Sparse algorithm 2' (SA2)[6] fits candidate molecules with a single Gaussian PSF of fixed shape and size, without shape- or size-based filtering.

We first investigated the qualitative performance of each algorithm for images of Alexa Fluor 647–immunolabeled microtubules in fixed COS-7 cells. We recorded data at high imaging density using total internal reflection fluorescence microscopy and direct (d)STORM photoswitching conditions[5] (100 ms integration time, ~4,000 photons fluorophore^{-1} frame^{-1}). We plotted localizations on raw images, illustrating the characteristic performance of each algorithm (**Fig. 1a**). SA1 only localized isolated molecules, which were fitted with small localization error. SA2 localized a larger fraction of the molecules but yielded large localization errors for overlapping molecules. DAOSTORM outperformed both sparse algorithms, identifying almost all molecules with small localization error.

We quantified the performance of each algorithm by analyzing simulations of randomly distributed surface-immobilized fluorophores[6]. We compared observed localizations to simulated positions, calculating the recall[7] and localization error at different imaging densities. Recall is the percentage of simulated fluorophores detected. Localization error is the root-mean-square distance between a localization and the simulated position. We also measured the precision[5] and redundancy (**Supplementary Methods**), which did not vary substantially.

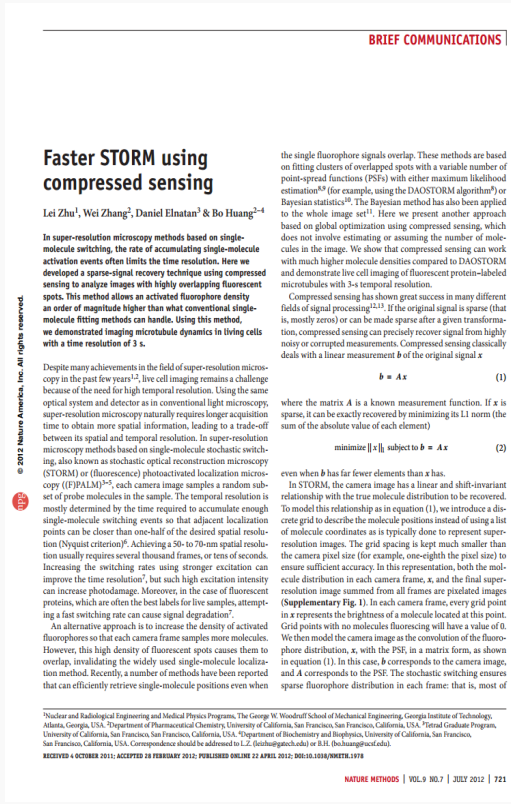
DAOSTORM substantially outperformed the sparse algorithms in simulations at high signal-to-noise ratio typical of STORM data (bright organic fluorophores, 5,000 photons molecule^{-1} frame^{-1}; **Fig. 1b–c**). SA1 showed poor recall at high density, with imaging density at half-maximum recall, ρ_{HM}, of 1.2 molecule μm^{-2}. However, SA1 yielded small localization errors even at high imaging density because most overlapping molecules were rejected. SA2 had better recall performance (ρ_{HM} = 3.4 molecules μm^{-2}) but gave large localization errors even at low imaging density (>0.1 molecules μm^{-2}). In contrast, DAOSTORM gave small localization errors similar to the other 'precise' algorithm, SA1, together with a sixfold improvement in recall performance (ρ_{HM} = 7.5 molecules μm^{-2}). For simulations at low signal-to-noise ratio typical of photoactivated localization microscopy data (fluorescent proteins, 200 photons molecule^{-1} frame^{-1};

Figure 1 | Comparison of DAOSTORM to existing super-resolution localization algorithms. (**a**) A single image of fluorescently labeled microtubules was analyzed using SA1, SA2 and DAOSTORM. Crosses represent localizations for each algorithm (**b,c**) Recall (**b**) and localization error (**c**) of the algorithms used in a measured for simulated images of randomly distributed surface-immobilized molecules. Error bars, s.d. ($n = 10$). (**d**) Super-resolved microtubule images from a 2,000-frame data series. (**e**) Line plots of cross-section indicated by dashed lines in d. Scale bars, 1 μm.

© 2011 Nature America, Inc. All rights reserved.

Holden, S.et al, Nat Methods, 2011

Sparsity based approach



BRIEF COMMUNICATIONS

Faster STORM using compressed sensing

Lei Zhu[1], Wei Zhang[2], Daniel Elnatan[3] & Bo Huang[2–4]

In super-resolution microscopy methods based on single-molecule switching, the rate of accumulating single-molecule activation events often limits the time resolution. Here we developed a sparse-signal recovery technique using compressed sensing to analyze images with highly overlapping fluorescent spots. This method allows an activated fluorophore density an order of magnitude higher than what conventional single-molecule fitting methods can handle. Using this method, we demonstrated imaging microtubule dynamics in living cells with a time resolution of 3 s.

Despite many achievements in the field of super-resolution microscopy in the past few years[1,2], live cell imaging remains a challenge because of the need for high temporal resolution. Using the same optical system and detector as in conventional light microscopy, super-resolution microscopy naturally requires longer acquisition time to obtain more spatial information, leading to a trade-off between its spatial and temporal resolution. In super-resolution microscopy methods based on single-molecule stochastic switching, also known as stochastic optical reconstruction microscopy (STORM) or (fluorescence) photoactivated localization microscopy ((F)PALM)[3–5], each camera image samples a random subset of probe molecules in the sample. The temporal resolution is mostly determined by the time required to accumulate enough single-molecule switching events so that adjacent localization points can be closer than one-half of the desired spatial resolution (Nyquist criterion)[6]. Achieving a 50- to 70-nm spatial resolution usually requires several thousand frames, or tens of seconds. Increasing the switching rates using stronger excitation can improve the time resolution[7], but such high excitation intensity can increase photodamage. Moreover, in the case of fluorescent proteins, which are often the best labels for live samples, attempting a fast switching rate can cause signal degradation[7].

An alternative approach is to increase the density of activated fluorophores so that each camera frame samples more molecules. However, this high density of fluorescent spots causes them to overlap, invalidating the widely used single-molecule localization method. Recently, a number of methods have been reported that can efficiently retrieve single-molecule positions even when the single fluorophore signals overlap. These methods are based on fitting clusters of overlapped spots with a variable number of point-spread functions (PSFs) with either maximum likelihood estimation[8,9] (for example, using the DAOSTORM algorithm[8]) or Bayesian statistics[10]. The Bayesian method has also been applied to the whole image set[11]. Here we present another approach based on global optimization using compressed sensing, which does not involve estimating or assuming the number of molecules in the image. We show that compressed sensing can work with much higher molecule densities compared to DAOSTORM and demonstrate live cell imaging of fluorescent protein–labeled microtubules with 3-s temporal resolution.

Compressed sensing has shown great success in many different fields of signal processing[12,13]. If the original signal is sparse (that is, mostly zeros) or can be made sparse after a given transformation, compressed sensing can precisely recover signal from highly noisy or corrupted measurements. Compressed sensing classically deals with a linear measurement $\boldsymbol{b}$ of the original signal $\boldsymbol{x}$

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} \quad (1)$$

where the matrix $\boldsymbol{A}$ is a known measurement function. If $\boldsymbol{x}$ is sparse, it can be exactly recovered by minimizing its L1 norm (the sum of the absolute value of each element)

$$\text{minimize } \|\boldsymbol{x}\|_1 \text{ subject to } \boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} \quad (2)$$

even when $\boldsymbol{b}$ has far fewer elements than $\boldsymbol{x}$ has.

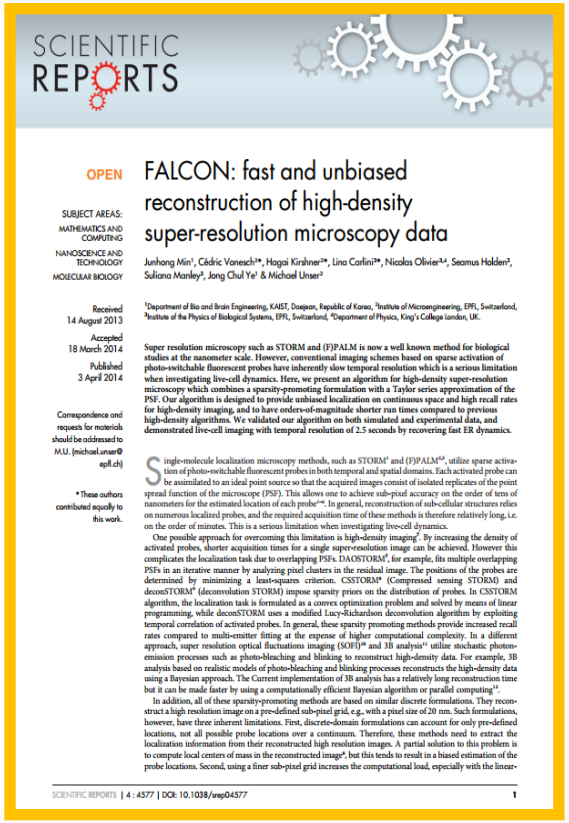
In STORM, the camera image has a linear and shift-invariant relationship with the true molecule distribution to be recovered. To model this relationship as in equation (1), we introduce a discrete grid to describe the molecule positions instead of using a list of molecule coordinates as is typically done to represent super-resolution images. The grid spacing is kept much smaller than the camera pixel size (for example, one-eighth the pixel size) to ensure sufficient accuracy. In this representation, both the molecule distribution in each camera frame, $\boldsymbol{x}$, and the final super-resolution image summed from all frames are pixelated images (**Supplementary Fig. 1**). In each camera frame, every grid point in $\boldsymbol{x}$ represents the brightness of a molecule located at this point. Grid points with no molecules fluorescing will have a value of 0. We then model the camera image as the convolution of the fluorophore distribution, $\boldsymbol{x}$, with the PSF, in a matrix form, as shown in equation (1). In this case, $\boldsymbol{b}$ corresponds to the camera image, and $\boldsymbol{A}$ corresponds to the PSF. The stochastic switching ensures sparse fluorophore distribution in each frame: that is, most of

[1]Nuclear and Radiological Engineering and Medical Physics Programs, The George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA. [2]Department of Pharmaceutical Chemistry, University of California, San Francisco, California, USA. [3]Tetrad Graduate Program, University of California, San Francisco, California, USA. [4]Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, California, USA. Correspondence should be addressed to L.Z. (leizhu@gatech.edu) or B.H. (bo.huang@ucsf.edu).

RECEIVED 4 OCTOBER 2011; ACCEPTED 28 FEBRUARY 2012; PUBLISHED ONLINE 22 APRIL 2012; DOI:10.1038/NMETH.1978

© 2012 Nature America, Inc. All rights reserved.

Zhu, L.et al, Nat Methods, 2012



SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:
MATHEMATICS AND COMPUTING
NANOSCIENCE AND TECHNOLOGY
MOLECULAR BIOLOGY

Received
14 August 2013
Accepted
18 March 2014
Published
3 April 2014

Correspondence and requests for materials should be addressed to M.U. (michael.unser@epfl.ch)

*These authors contributed equally to this work.

FALCON: fast and unbiased reconstruction of high-density super-resolution microscopy data

Junhong Min[1], Cédric Vonesch[2]*, Hagai Kirshner[2]*, Lina Carlini[2]*, Nicolas Olivier[2,4], Seamus Holden[2], Suliana Manley[3], Jong Chul Ye[1] & Michael Unser[2]

[1]Department of Bio and Brain Engineering, KAIST, Daejeon, Republic of Korea, [2]Institute of Microengineering, EPFL, Switzerland, [3]Institute of the Physics of Biological Systems, EPFL, Switzerland, [4]Department of Physics, King's College London, UK.

Super resolution microscopy such as STORM and (F)PALM is now a well known method for biological studies at the nanometer scale. However, conventional imaging schemes based on sparse activation of photo-switchable fluorescent probes have inherently slow temporal resolution which is a serious limitation when investigating live-cell dynamics. Here, we present an algorithm for high-density super-resolution microscopy which combines a sparsity-promoting formulation with a Taylor series approximation of the PSF. Our algorithm is designed to provide unbiased localization on continuous space and high recall rates for high-density imaging, and to have orders-of-magnitude shorter run times compared to previous high-density algorithms. We validated our algorithm on both simulated and experimental data, and demonstrated live-cell imaging with temporal resolution of 2.5 seconds by recovering fast ER dynamics.

Single-molecule localization microscopy methods, such as STORM[1] and (F)PALM[2,3], utilize sparse activation of photo-switchable fluorescent probes in both temporal and spatial domains. Each activated probe can be assimilated to an ideal point source so that the acquired images consist of isolated replicates of the point spread function of the microscope (PSF). This allows one to achieve sub-pixel accuracy on the order of tens of nanometers for the estimated location of each probe[1–4]. In general, reconstruction of sub-cellular structures relies on numerous localized probes, and the required acquisition time of these methods is therefore relatively long, i.e. on the order of minutes. This is a serious limitation when investigating live-cell dynamics.

One possible approach for overcoming this limitation is high-density imaging[2]. By increasing the density of activated probes, shorter acquisition times for a single super-resolution image can be achieved. However this complicates the localization task due to overlapping PSFs. DAOSTORM[4], for example, fits multiple overlapping PSFs in an iterative manner by analyzing pixel clusters in the residual image. The positions of the probes are determined by minimizing a least-squares criterion. CSSTORM[4] (Compressed sensing STORM) and deconSTORM[4] (deconvolution STORM) impose sparsity priors on the distribution of probes. In CSSTORM algorithm, the localization task is formulated as a convex optimization problem and solved by means of linear programming, while deconSTORM uses a modified Lucy-Richardson deconvolution algorithm by exploiting temporal correlation of activated probes. In general, these sparsity promoting methods provide increased recall rates compared to multi-emitter fitting at the expense of higher computational complexity. In a different approach, super resolution optical fluctuation imaging (SOFI)[4] and 3B analysis[14] utilize stochastic photon-emission processes such as photo-bleaching and blinking to reconstruct high-density data. For example, 3B analysis based on realistic models of photo-bleaching and blinking processes reconstructs the high-density data using a Bayesian approach. The Current implementation of 3B analysis has a relatively long reconstruction time but it can be made faster by using a computationally efficient Bayesian algorithm or parallel computing[14].

In addition, all of these sparsity-promoting methods are based on similar discrete formulations. They reconstruct a high resolution image on a pre-defined sub-pixel grid, e.g., with a pixel size of 20 nm. Such formulations, however, have three inherent limitations. First, discrete-domain formulations can account for only pre-defined locations, not all possible probe locations over a continuum. Therefore, these methods need to extract the localization information from their reconstructed high resolution images. A partial solution to this problem is to compute local centers of mass in the reconstructed image[4], but this tends to result in a biased estimation of the probe locations. Second, using a finer sub-pixel grid increases the computational load, especially with the linear-

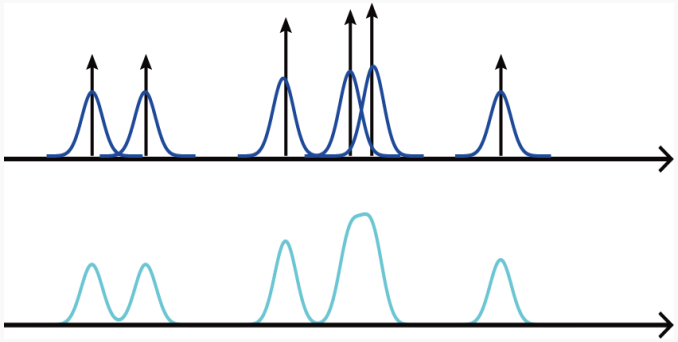
Min, J.et al, Sci. Rep, 2014

Better Localization Performance

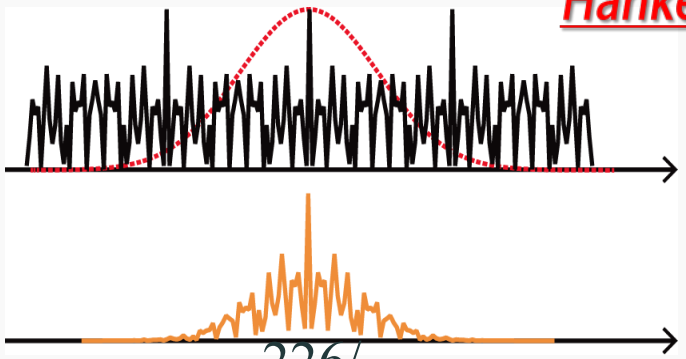
ALOHA for localization microscopy

ALOHA principle

Image: $g = h * f$ f: Sparse



Fourier: $\hat{g} = \hat{h} \odot \hat{f}$ $\hat{f}$: Low-rank Hankel matrix

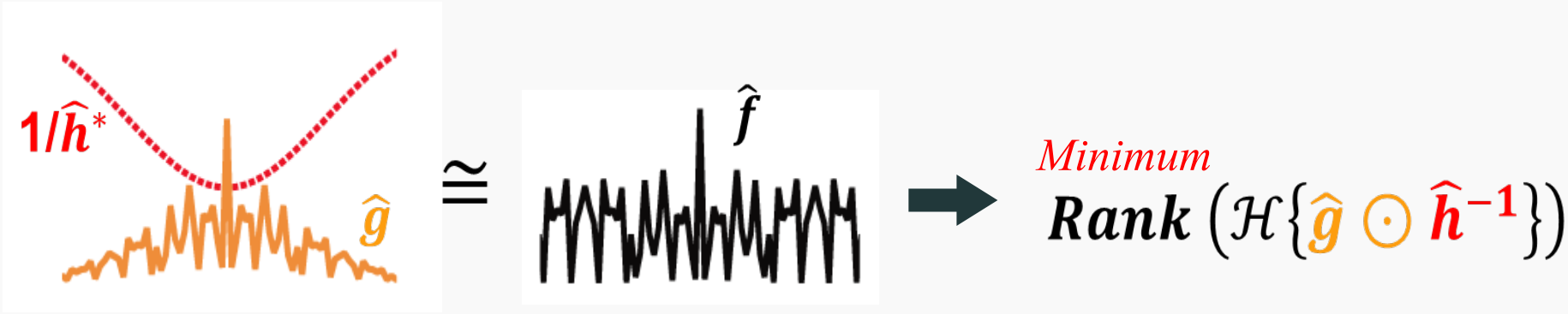


- ✓ *PSF estimation*
- ✓ *Deconvolution*
- ✓ *Grid-free localization*

PSF estimation

- HD localization algorithms usually assume that PSF is *known and fixed*
  - Requiring additional training low-density data set
  - *In live experiment, PSF is varying in time and space both.*

Key idea: *Optimal PSF h^* => minimum rank of Hankel matrix*



- Under symmetric Gaussian PSF model, its width (σ) is estimated by minimizing Schatten norm

$$\sigma^* = \min_{\sigma} \left\| \left(\mathcal{H}\{\hat{g} \odot \hat{h}_{\sigma}^{-1}\} \right) \right\|_{P\ (p<1)}$$

Grid-free localization

- Now, we have entire Fourier spectrum $\hat{f}$
- *Localization is nothing but spectral estimation problem!*

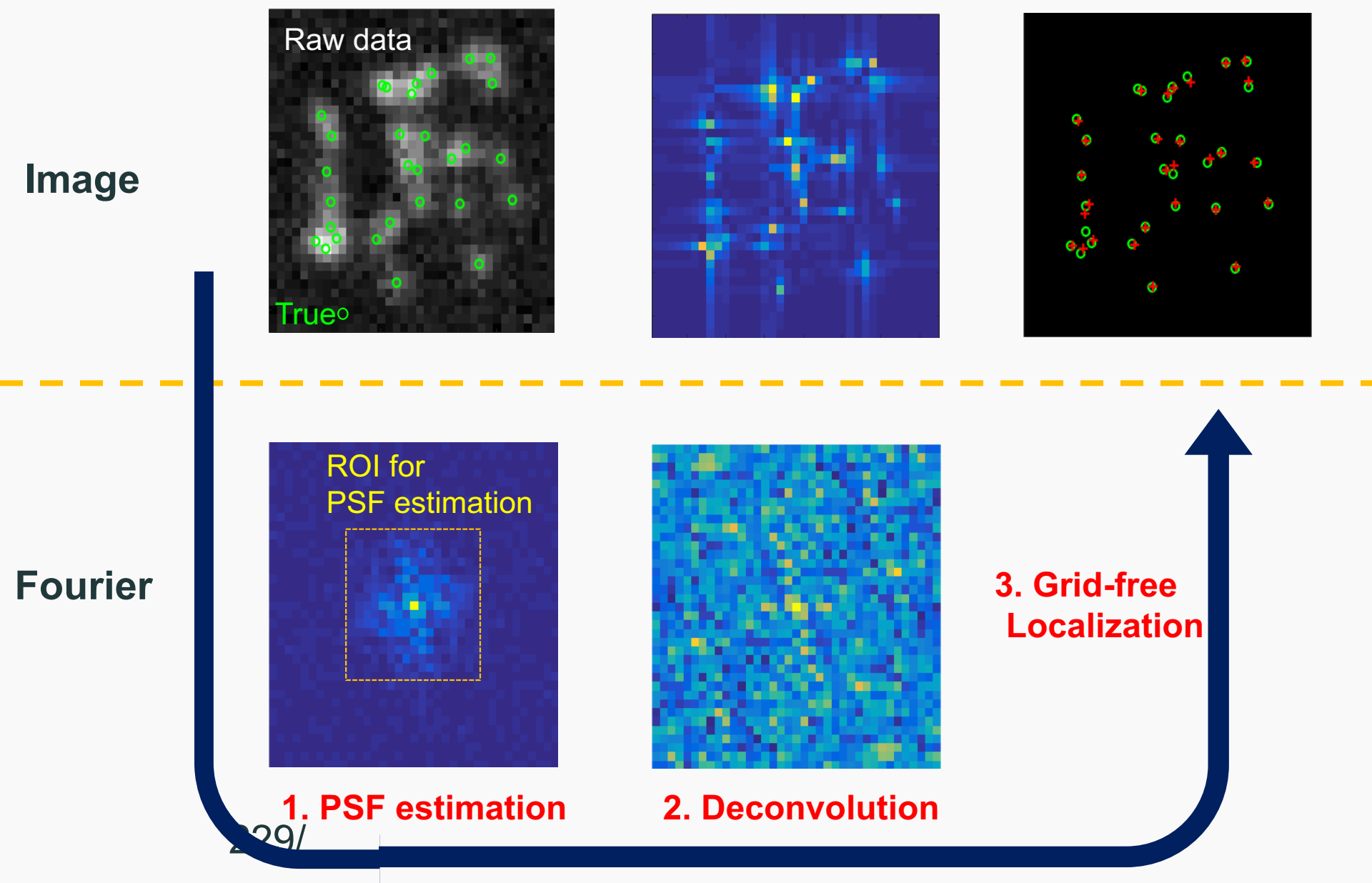
$$\hat{f}(m,n) = \sum_i c_i e^{-j2\pi(\frac{mx_i}{M}+\frac{ny_i}{N})} = \sum_i c_i p_i^m q_i^n$$

- We used ACMP (algebraically coupled matrix pencils) algorithm (Vanpoucke et al, 1994)
- Data matrix $Z^{M\times N}$ of rank k, having no shared harmonics of p_i, q_i

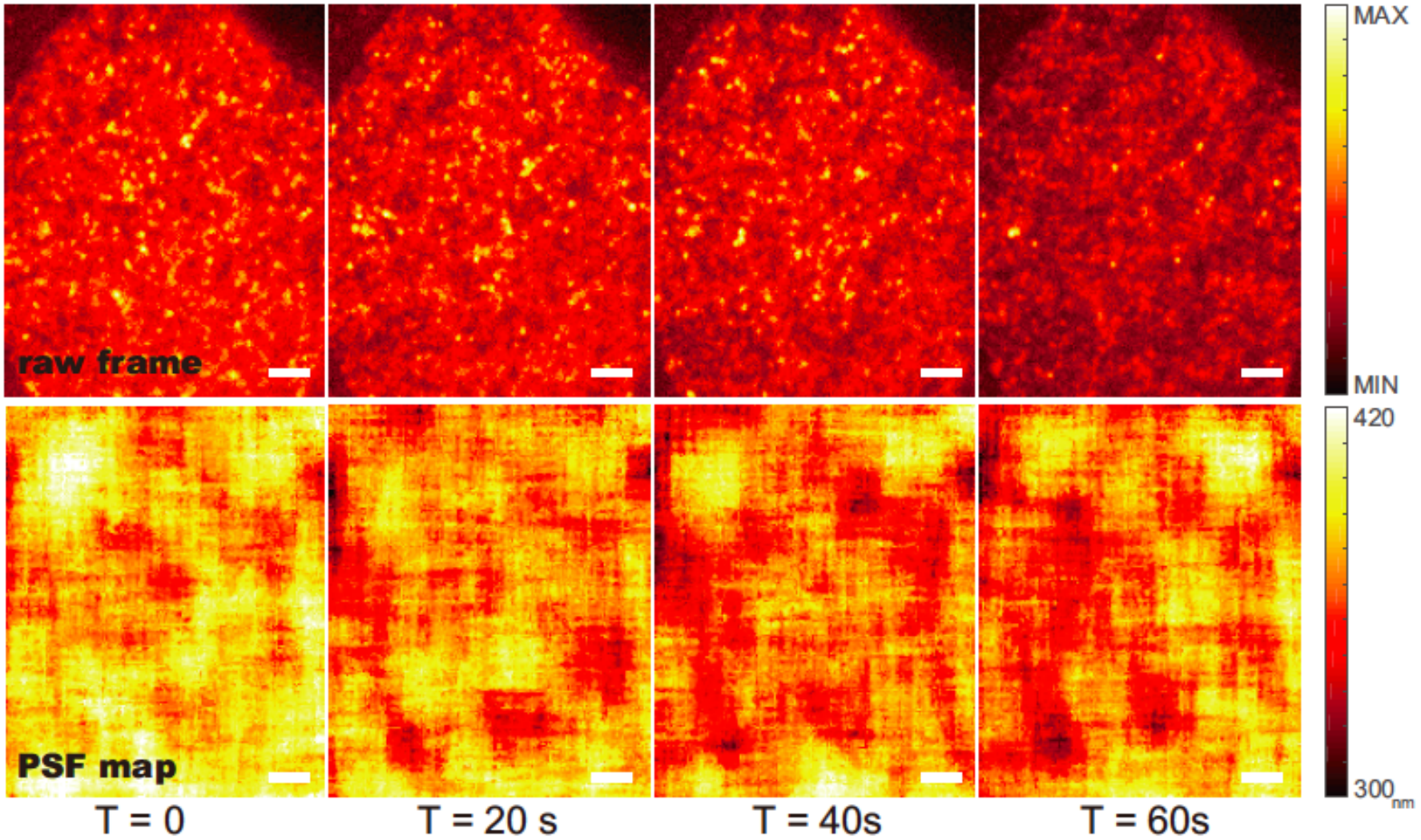
$$\mathrm{Z}^{M\times N} = P^{M\times k}\, C^{k\times k}\, Q'^{N\times k}$$

- In Matrix form: $\mathrm{Z}^{M\times N} = P^{M\times k}\, C^{k\times k}\, Q'^{N\times k}$
- *P,Q are Vandermonde matrix*, C is diagonal

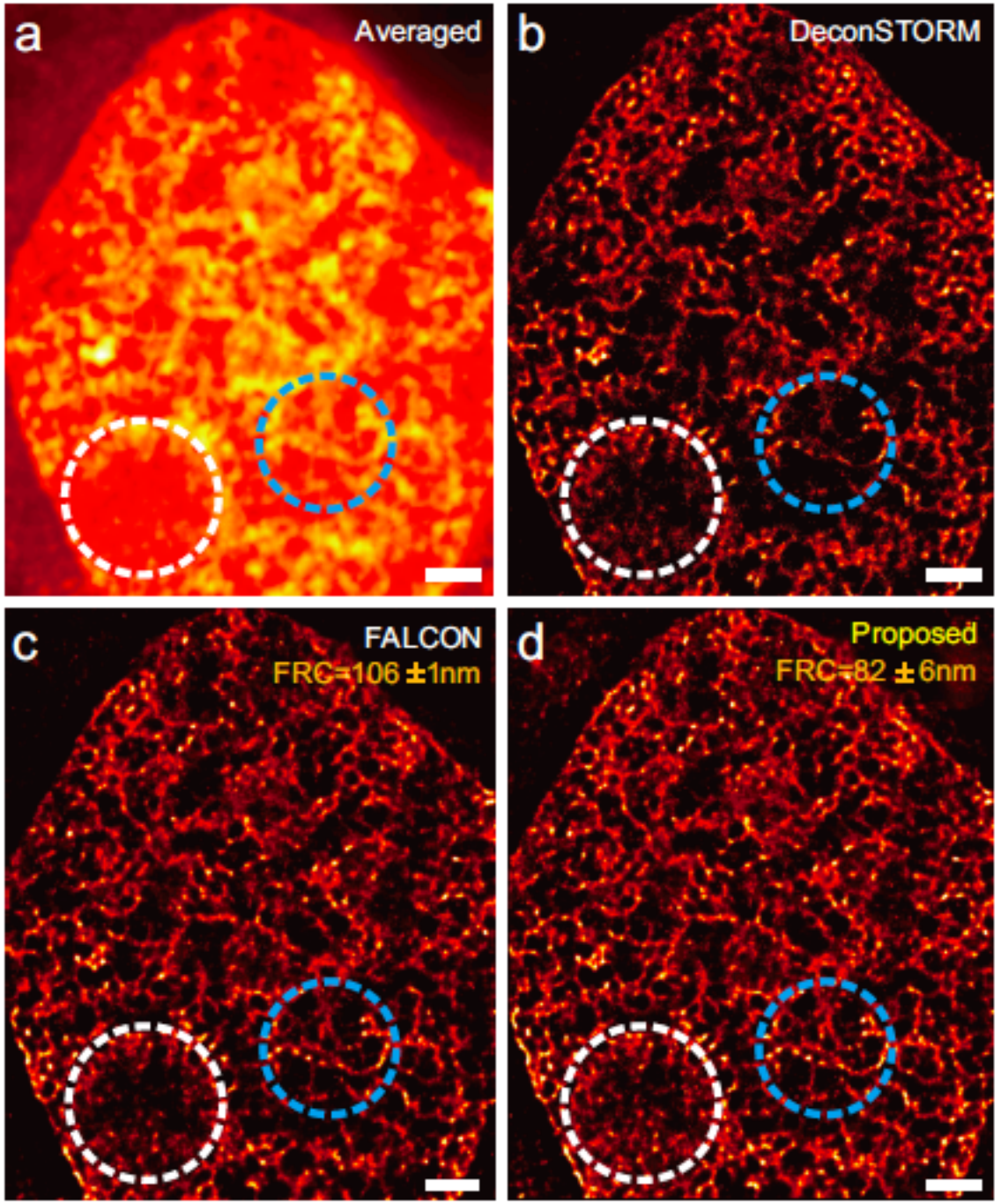
Algorithm procedure



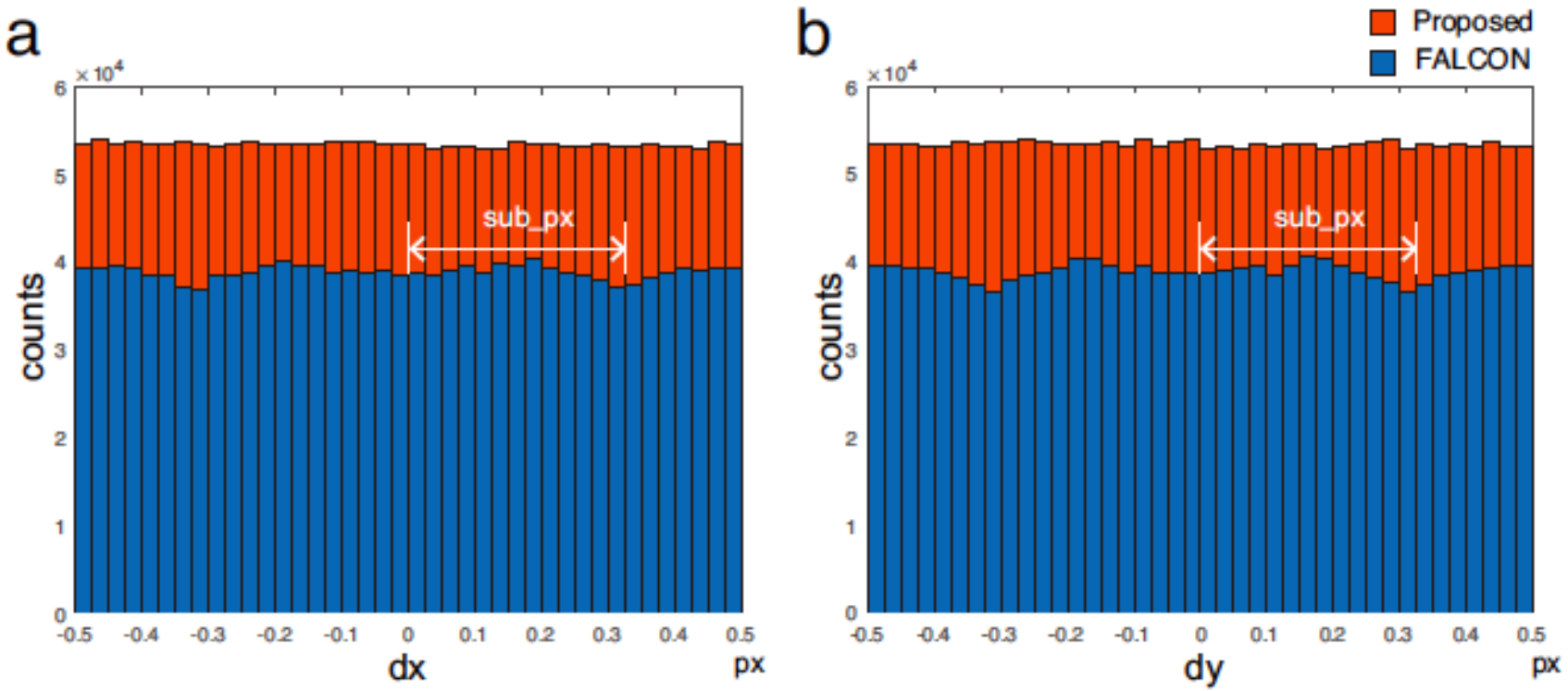
PSF variation along time



Reconstruction



Localization bias



Free MATLAB software available

https://research.engineering.uiowa.edu/cbig/software

http://bispl.weebly.com/aloha-for-mri.html

Conclusions

- Off-the-grid = Continuous domain representation
- *Compressive* off-the-grid *imaging*:

  Exploit continuous domain modeling to improve image recovery from few measurements
- Two realizations: extrapolation, interpolation
  - Extrapolation: FRI theory
  - Interpolation: Structured low-rank matrix completion
- Performance guarantee for structured low-rank approach
  - 1D, 2D theory → near optimal performance guarantee

Conclusions (cont.)

- Extensive applications
  - MRI
    - Compressed sensing MRI, parallel MRI
    - Super-resolution MRI
    - MR artifact removal
  - Image processing: inpainting, impulse noise denoising
  - Other imaging applications
    - US imaging
    - Optics
- A missing link between analytic recon and CS ?

Joint work by several authors

CBIG, Univ. Iowa

1. Dr. Gregory Ongie
2. Dr. Merry Mani
3. Mr. Arvind Balachandrasekaran
4. Ms. Ipshita Bhattacharya
5. Ms. Sampurna Biswas

KAIST

1. Dr. Kyong Hwan Jin
2. Mr. Juyoung Lee
3. Dongwook Lee
4. Junhong Min

References

- Jong Chul Ye, Jong Min Kim, Kyong Hwan Jin and Kiryung Lee, "Compressive sampling using annihilating filter-based low-rank interpolation", IEEE Trans. on Information Theory, vol. 63, no. 2, pp.777-801, Feb. 2017.

- 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," IEEE Trans. on Computational Imaging, vol 2, no. 4, pp. 480 - 495, Dec. 2016.

- Kyong Hwan Jin, Ji-Yong Um, Dongwook Lee, Juyoung Lee, Sung-Hong Park and Jong Chul Ye, " MRI artifact correction using sparse + low-rank decomposition of annihilating lter-based Hankel matrix", Magnetic Resonance in Medicine (in press), 2016

- Juyoung Lee, Kyong Hwan Jin, and Jong Chul Ye, "Reference-free single-pass EPI Nyquist ghost correction using annihilating filter-based low rank Hankel matrix (ALOHA)", Magnetic Resonance in Medicine, 10.1002/mrm.26077, Feb. 17, 2016.

References

- Dongwook Lee,, Kyong Hwan Jin, Eung Yeop Kim, Sung-Hong Park and Jong Chul Ye, "Acceleration of MR parameter mapping using annihilating filter-based low rank Hankel matrix (ALOHA)", Magnetic Resonance in Medicine, 10.1002/mrm.26081, Jan. 1, 2016.

- Kyong Hwan Jin and Jong Chul Ye, "Annihilating filter based low rank Hankel matrix approach for image inpainting", IEEE Trans. Image Processing, 2015 Nov;24(11):3498-511.

- KH Jin, JC Ye, Sparse+ low rank decomposition of annihilating filter-based Hankel matrix for impulse noise removal, arXiv preprint arXiv:1510.05559

- Min, J., Carlini, L., Unser, M., Manley, S., & Ye, J. C. (2015, September). Fast live cell imaging at nanometer scale using annihilating filter-based low-rank Hankel matrix approach. In SPIE Optical Engineering+ Applications (pp. 95970V-95970V). International Society for Optics and Photonics.

- Jin, Kyong Hwan, Yo Seob Han, and Jong Chul Ye. "Compressive dynamic aperture B-mode ultrasound imaging using annihilating filter-based low-rank interpolation." Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on. IEEE, 2016.

References

- G. Ongie, M. Jacob, Off-the-Grid Recovery of Piecewise Constant Images from Few Fourier Samples, SIAM Journal on Imaging Sciences, in press.
- G. Ongie, S. Biswas, M. Jacob, Convex recovery of continuous domain piecewise constant images from non-uniform Fourier samples, https://arxiv.org/abs/1703.01405
- Ongie, M. Jacob, GIRAF: A Fast Algorithm for Structured Low-Rank Matrix Recovery, https://arxiv.org/abs/1609.07429
- M. Mani, M. Jacob, D. Kelley, V. Magnotta, Multishot sensitivity encoded diffusion data recovery using structured low rank matrix completion (MUSSELS), Magnetic Resonance in Medicine, in press.
- G. Ongie, S. Biswas, M. Jacob, Structured matrix recovery of piecewise constant signals with performance guarantees, International Conference on Image Processing, 2016.
- A. Balachandrasekaran, G. Ongie, M. Jacob, Accelerated dynamic MRI using structured matrix completion, International Conference on Image Processing, 2016.

References

- G. Ongie, M. Jacob, A fast algorithm for stuctured low rank matrix recovery with applications to undersampled MRI reconstruction, ISBI, Prague, Czech Republic, 2016.

- G. Ongie, M. Jacob, Recovery of piecewise smooth images from few Fourier samples, Sampling Theory and Applications (SampTA), Washington D.C., 2015.

- G. Ongie, M. Jacob, Super-resolution MRI using finite rate of innovation curves, IEEE ISBI, New York City, USA, 2015.

- M. Mani, V. Magnotta, D. Kelley, M.Jacob, Comprehensive reconstruction of multishot multichannel diffusion MRI data using MUSSELS, Engineering in Biology and Medicine Conference, 2016.