Computational Harmonic Analysis meets Imaging Sciences
Part II

Gitta Kutyniok
(Technische Universität Berlin)

BMS Summer School
Berlin, July 25 – August 5, 2016

Outline

1. Feature Extraction
   - Point- and Curvelike Structures
   - Application of Sparse Regularization
   - Asymptotic Result
   - Numerical Experiments

2. Magnetic Resonance Imaging
   - Sampling-Reconstruction Scheme
   - Compressed Sensing comes into Play
   - Optimality Result
   - Numerical Experiments

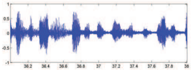
We start with Feature Extraction!

General Challenge in Data Analysis

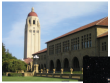
Modern Data in general is often composed of two or more morphologically distinct constituents, and we face the task of separating those components given the composed data.

Examples include...

- Audio data: Sinusoids and peaks.

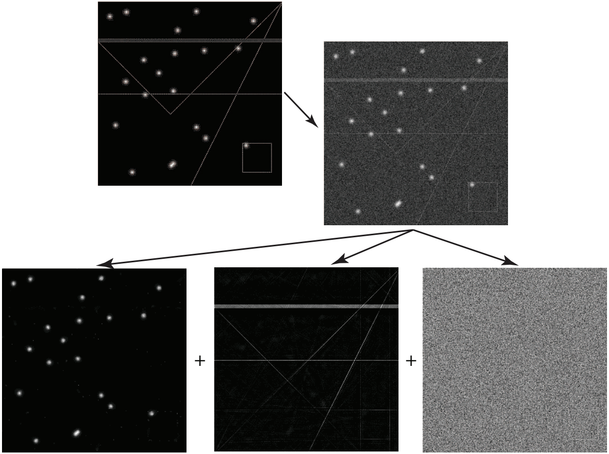


- Imaging data: Cartoon and texture.



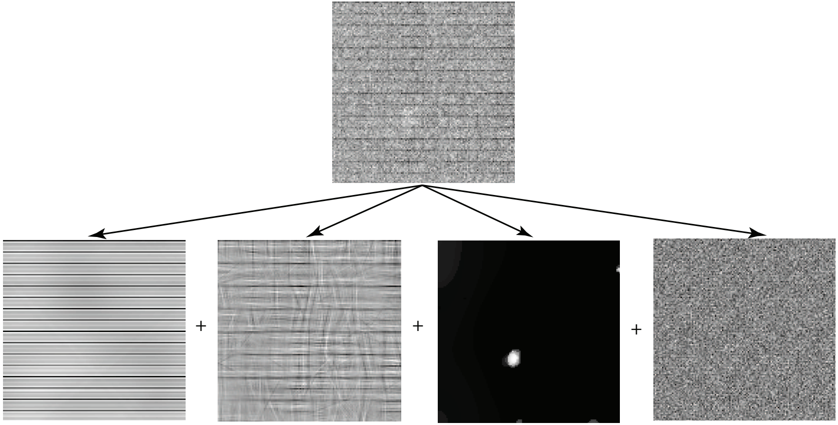
- High-dimensional data: Lower-dimensional structures of different dimensions.

Separating Artifacts in Images, I



(Source: J. L. Starck, M. Elad, D. L. Donoho; 2005 (Artificial Data))

Separating Artifacts in Images, II

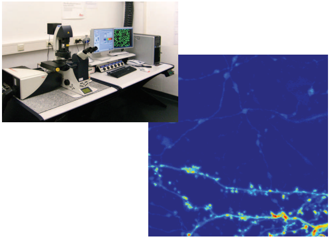


(Source: J. L. Starck, M. Elad, D. L. Donoho; 2005)

Problem from Neurobiology

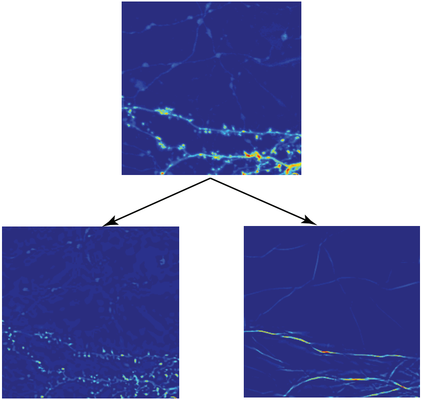
Alzheimer Research:

- Detection of characteristics of Alzheimer.
- Separation of spines and dendrites.



(Confocal-Laser Scanning-Microscopy)

Numerical Result



(Source: Brandt, K, Lim, Sündermann; 2010)

How does Sparse Regularization help with Component Separation?

‘Mathematical Model’

Model for 2 Components:

- Observe a signal x composed of two subsignals x_1 and x_2:

$$x = x_1 + x_2.$$

- Extract the two subsignals x_1 and x_2 from x, if only x is known.

'Mathematical Model'

Model for 2 Components:

- Observe a signal x composed of two subsignals x_1 and x_2:

$$x = x_1 + x_2.$$

- Extract the two subsignals x_1 and x_2 from x, if only x is known.

Isn't this impossible?

- There are two unknowns for every datum.

‘Mathematical Model’

Model for 2 Components:

- Observe a signal x composed of two subsignals x_1 and x_2:

$$x = x_1 + x_2.$$

- Extract the two subsignals x_1 and x_2 from x, if only x is known.

Isn’t this impossible?

- There are two unknowns for every datum.

But we have additional Information:

- The two components are geometrically different.

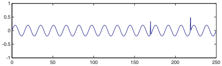
Birth of ℓ_1-Component Separation (2001)

Composition of Sinusoids and Spikes sampled at n points:

$$x = x_1^0 + x_2^0 = \Phi_1 c_1^0 + \Phi_2 c_2^0 = \left[\quad \Phi_1 \quad \Big| \quad \Phi_2 \quad \right] \begin{bmatrix} c_1^0 \\ c_2^0 \end{bmatrix},$$

where

- x, c_1^0, and c_2^0 are $n \times 1$.
- Φ_1 is the $n \times n$-Fourier matrix ($(\Phi_1)_{t,k} = e^{2\pi itk/n}$).
- Φ_2 is the $n \times n$-Identity matrix.



First Results of Compressed Sensing

Composition of Sinusoids and Spikes sampled at n points:

$$x = x_1^0 + x_2^0 = \Phi_1 c_1^0 + \Phi_2 c_2^0 = \left[\ \Phi_1 \ \middle|\ \Phi_2 \ \right] \left[\begin{array}{c} c_1^0 \\ c_2^0 \end{array} \right].$$

First Results of Compressed Sensing

Composition of Sinusoids and Spikes sampled at n points:

$$x = x_1^0 + x_2^0 = \Phi_1 c_1^0 + \Phi_2 c_2^0 = \left[\begin{array}{c|c} \Phi_1 & \Phi_2 \end{array} \right] \left[\begin{array}{c} c_1^0 \\ c_2^0 \end{array} \right].$$

Theorem (Bruckstein, Elad; 2002)(Donoho, Elad; 2003)

Let $A = (a_i)_{i=1}^N$ be an $n \times N$-matrix with normalized columns, $n << N$, and let c^0 satisfy

$$\|c^0\|_0 < \frac{1}{2}\left(1 + \mu(A)^{-1}\right),$$

with coherence $\mu(A) = \max_{i \neq j} |\langle a_i, a_j \rangle|$. Then

$$c^0 = \text{argmin} \|c\|_1 \text{ subject to } x = Ac.$$

First Results of Compressed Sensing

Composition of Sinusoids and Spikes sampled at n points:

$$x = x_1^0 + x_2^0 = \Phi_1 c_1^0 + \Phi_2 c_2^0 = \left[\quad \Phi_1 \quad \middle| \quad \Phi_2 \quad \right] \begin{bmatrix} c_1^0 \\ c_2^0 \end{bmatrix}.$$

Theorem (Bruckstein, Elad; 2002)(Donoho, Elad; 2003)

Let $A = (a_i)_{i=1}^N$ be an $n \times N$-matrix with normalized columns, $n << N$, and let c^0 satisfy

$$\|c^0\|_0 < \frac{1}{2}\left(1 + \mu(A)^{-1}\right),$$

with coherence $\mu(A) = \max_{i \neq j} |\langle a_i, a_j \rangle|$. Then

$$c^0 = \text{argmin}\|c\|_1 \text{ subject to } x = Ac.$$

Theorem (Donoho, Huo; 2001)

If $\#(\text{Sinusoids}) + \#(\text{Spikes}) = \|(c_1^0)\|_0 + \|(c_2^0)\|_0 < (1 + \sqrt{n})/2$, then

$$(c_1^0, c_2^0) = \text{argmin}(\|c_1\|_1 + \|c_2\|_1) \text{ subject to } x = \Phi_1 c_1 + \Phi_2 c_2.$$

Component Separation using Compressed Sensing

Let x be a signal composed of two subsignals x_1^0 and x_2^0:

$$x = x_1^0 + x_2^0.$$

Desiderata for two orthonormal bases Φ_1 and Φ_2:

- $x_i^0 = \Phi_i c_i^0$ with $\|c_i^0\|_0$ small, $i = 1, 2 \rightsquigarrow$ Sparsity!
- $\mu([\Phi_1|\Phi_2])$ small $\rightsquigarrow$ Morphological Difference!

Solve

$$(c_1^*, c_2^*) = \text{argmin}(\|c_1\|_1 + \|c_2\|_1) \text{ subject to } x = \Phi_1 c_1 + \Phi_2 c_2$$

and derive the approximate components

$$x_i^0 \approx x_i^* = \Phi_i c_i^*, \quad i = 1, 2.$$

Two Paths

Avalanche of Recent Work

Problem: Solve $x = Ac^0$ with A an $n \times N$-matrix ($n < N$).

Deterministic World:

- Mutual coherence of $A = (a_k)_k$.
- Bound $\|c^0\|_0$ dependent on $\mu(A)$.
- Efficiently solve the problem $x = Ac^0$.
- Contributors: *Bruckstein, Cohen, Dahmen, DeVore, Donoho, Elad, Fuchs, Gribonval, Huo, K, Rauhut, Temlyakov, Tropp, ...*

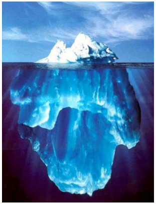
Random World:

- Restricted isometry constants of a random $A = (a_k)_k$.
- Bound $\|c^0\|_0$ by $n/(2\log(N/n))(1 + o(1))$.
- Efficiently solve the problem $x = Ac^0$ with high probability.
- Contributors: *Candès, Donoho, Fornasier, K, Krahmer, Rauhut, Romberg, Tanner, Tao, Tropp, Vershynin, Ward, ...*

Novel Direction for Sparsity

Geometric Clustering:

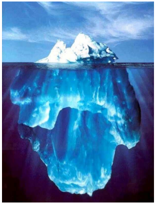
- $x = Ac^0$ with A an $n \times N$-matrix ($n < N$).
- Nonzeros of c^0 often
  - arise not in arbitrary patterns,
  - but are rather highly structured.
- Interactions between columns of A in ill-posed problems
  - is not arbitrary,
  - but rather geometrically driven.



Novel Direction for Sparsity

Geometric Clustering:

- $x = Ac^0$ with A an $n \times N$-matrix ($n < N$).
- Nonzeros of c^0 often
  - arise not in arbitrary patterns,
  - but are rather highly structured.
- Interactions between columns of A in ill-posed problems
  - is not arbitrary,
  - but rather geometrically driven.



Other results on "structured sparsity":

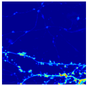
- Joint sparsity, fusion frame sparsity, block sparsity, ...
- Contributors: Boufounos, Ehler, Eldar, Gribonval, Fornasier, K, Rauhut, Schnass, Vandergheynst, Vershynin, Ward, ...

How can these Ideas be applied to

Separation of Points and Curves?

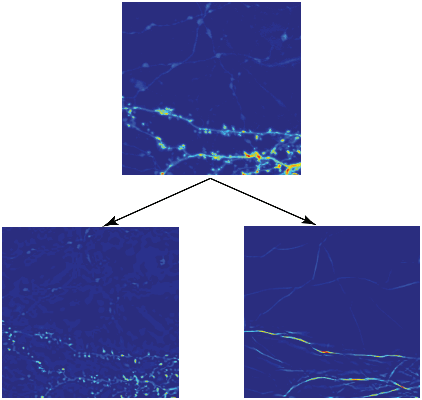
Back to Neurobiological Imaging

- Two morphologically distinct components:
  - Points
  - Curves



- Choose suitable representation systems which provide optimally sparse representations of
  - pointlike structures $\longrightarrow$ Wavelets
  - curvelike structures $\longrightarrow$ Shearlets
- Minimize the ℓ_1 norm of the coefficients.
- This forces
  - the pointlike objects into the wavelet part of the expansion
  - the curvelike objects into the shearlet part.

Empirical Separation of Spines and Dendrites

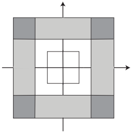


Wavelet Expansion　　Shearlet Expansion

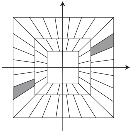
(Source: Brandt, K, Lim, Sündermann; 2010)

Chosen Pair

Optimal for Pointlike Structures:
Orthonormal Wavelets are a basis with perfectly isotropic generating elements at different scales.

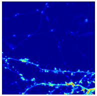


Optimal for Curvelike Structures:
Shearlets (K, Labate; 2006) are a highly directional frame with increasingly anisotropic elements at fine scales ($\longrightarrow$ www.ShearLab.org).



Microlocal Model

Neurobiological Geometric Mixture in 2D:



Point Singularity:

$$\mathcal{P}(x) = \sum_{i=1}^{P} |x - x_i|^{-3/2}$$

Curvilinear Singularity:

$$\mathcal{C} = \int \delta_{\tau(t)} dt, \quad \tau \text{ a closed } C^2\text{-curve.}$$

Observed Signal:

$$f = \mathcal{P} + \mathcal{C}$$

Scale-Dependent Decomposition

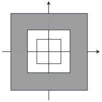
Observed Object:

$$f = \mathcal{P} + \mathcal{C}.$$

Subband Decomposition:

Wavelets and shearlets use the same scaling subbands!

$$f_j = \mathcal{P}_j + \mathcal{C}_j, \quad \mathcal{P}_j = \mathcal{P} \star F_j \text{ and } \mathcal{C}_j = \mathcal{C} \star F_j.$$



ℓ_1-Decomposition:

$$(W_j, S_j) = \text{argmin} \|(\langle W_j, \psi_\lambda \rangle)_\lambda\|_1 + \|(\langle S_j, \sigma_\eta \rangle)_\eta\|_1 \text{ s.t. } f_j = W_j + S_j$$

Asymptotic Separation

Theorem (Donoho, K; 2013)

$$\frac{\|W_j - \mathcal{P}_j\|_2 + \|S_j - \mathcal{C}_j\|_2}{\|\mathcal{P}_j\|_2 + \|\mathcal{C}_j\|_2} \to 0, \qquad j \to \infty.$$

At all sufficiently fine scales, nearly-perfect separation is achieved!

Analysis of Decomposition within one Scale

Signal Model:

$$x = x_1^0 + x_2^0 \in \mathcal{H}$$

Remarks:

- Given two Parseval frames Φ_1, Φ_2 ($\Phi_i(\Phi_i^T x) = x$ for all x).
- Too many decompositions $x = \Phi_1 c_1 + \Phi_2 c_2$.
- Use $x = \Phi_1(\Phi_1^T x_1) + \Phi_2(\Phi_2^T x_2)$, where $x = x_1 + x_2$.
- Norm is placed on analysis rather than synthesis side.

Decomposition Technique:

$$(x_1^\star, x_2^\star) = \text{argmin}_{x_1,x_2} \|\Phi_1^T x_1\|_1 + \|\Phi_2^T x_2\|_1 \text{ subject to } x = x_1 + x_2$$

Relative Sparsity and Cluster Coherence

Let $\Phi_1 = (\varphi_{1,i})_{i \in I_1}$ and $\Phi_2 = (\varphi_{2,i})_{i \in I_2}$.

Definition:

- For each $i = 1, 2$, x_i^0 is relatively sparse in Φ_i w.r.t. Λ_i, if

$$\|1_{\Lambda_1^c} \Phi_1^T x_1^0\|_1 + \|1_{\Lambda_2^c} \Phi_2^T x_2^0\|_1 \leq \delta.$$

  We call Λ_1 and Λ_2 sets of significant coefficients.

- We define cluster coherence for Λ_1 by

$$\mu_c(\Lambda_1) = \max_{j \in I_2} \sum_{i \in \Lambda_1} |\langle \varphi_{1,i}, \varphi_{2,j} \rangle|.$$

Central Estimate

Theorem (Donoho, K; 2013):
Suppose x_1^0 and x_2^0 are relatively sparse with Λ_1 and Λ_2 sets of significant coefficients. Then

$$\|x_1^\star - x_1^0\|_2 + \|x_2^\star - x_2^0\|_2 \leq \frac{2\delta}{1 - 2\mu_c},$$

where

$$\mu_c = \max(\mu_c(\Lambda_1), \mu_c(\Lambda_2)).$$

- δ: Relative sparsity measure.
- μ_c: Cluster coherence.

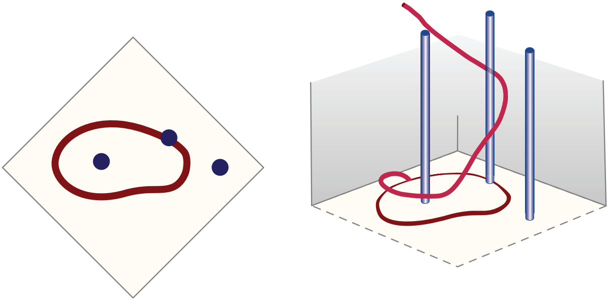
Application of Previous Result

- x: Filtered signal f_j $(= \mathcal{P}_j + \mathcal{C}_j)$.
- Φ_1: Wavelets filtered with F_j.
- Φ_2: Shearlets filtered with F_j.
- Λ_1: Significant wavelet coefficients of $\langle \psi_\lambda, \mathcal{P}_j \rangle$.
- Λ_2: Significant shearlet coefficients of $\langle \sigma_\eta, \mathcal{C}_j \rangle$.
- δ: Degree of approximation by significant coefficients.
- $\mu_c(\Lambda_1), \mu_c(\Lambda_2)$: Cluster coherence of wavelets-shearlets.
- Estimate of error: $\frac{2\delta}{1-2\mu_c}$.

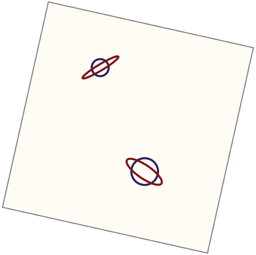
Application of Previous Result

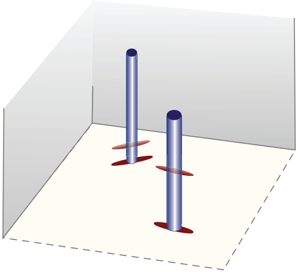
- x: Filtered signal f_j $(= \mathcal{P}_j + \mathcal{C}_j)$.
- Φ_1: Wavelets filtered with F_j.
- Φ_2: Shearlets filtered with F_j.
- Λ_1: Significant wavelet coefficients of $\langle \psi_\lambda, \mathcal{P}_j \rangle$?
- Λ_2: Significant shearlet coefficients of $\langle \sigma_\eta, \mathcal{C}_j \rangle$?
- δ: Degree of approximation by significant coefficients.
- $\mu_c(\Lambda_1), \mu_c(\Lambda_2)$: Cluster coherence of wavelets-shearlets.
- Estimate of error: $\frac{2\delta}{1-2\mu_c} = o(\|\mathcal{P}_j\|_2 + \|\mathcal{C}_j\|_2)$ as $j \to \infty$.

Singular Support and Wavefront Set of $\mathcal{P}$ and $\mathcal{C}$



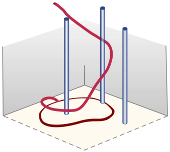
Phase Space Portrait of Wavelets and Shearlets

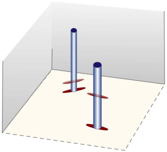




Cluster Coherence

- Wavelets in $\Lambda_1 \approx$ vertical tubes clustering around the point singularities of $\mathcal{P}$.
- Shearlets in $\Lambda_2 \approx$ tubes clustering around the curvilinear phase portrait of $\mathcal{C}$.
- Single wavelet is incoherent with ensemble of shearlets in Λ_2.
- Single shearlet is incoherent with ensemble of wavelets in Λ_1.





Key Idea from Microlocal Analysis

- Hart Smith's Phase Space Metric:

$$d((s,t);(s',t')) = \left|\left\langle e_s, t-t' \right\rangle\right| + \left|\left\langle e_{s'}, t-t' \right\rangle\right| + |t-t'|^2 + |s-s'|^2.$$

- 'Approximate' Sets of Significant Wavelet Coefficients:

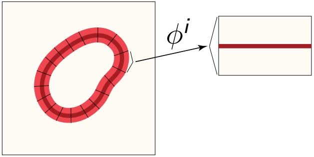
$$\Lambda_{1,j} = \{\text{wavelet lattice}\} \cap \{(s,t) : d((s,t); WF(\mathcal{P})) \leq \eta_j a_j\}.$$

- 'Approximate' Sets of Significant Shearlet Coefficients:

$$\Lambda_{2,j} = \{\text{shearlet lattice}\} \cap \{(s,t) : d((s,t); WF(\mathcal{C})) \leq \eta_j a_j\}.$$

Analysis of the Curvilinear Part

- The diffeomorphism ϕ^i



allows us to perform computations for distribution $\mathcal{L}_w$:

$$\langle \mathcal{L}_w, f \rangle = \int_{-\rho}^{\rho} w(t) f(t,0) dt.$$

- Use linear operator M_{ϕ^i} for transformation; use the 'model'

$$|M_{\phi^i}(\eta, \eta')| \leq c_N \cdot 2^{|j-j'|}(1 + \min(2^j, 2^{j'}) \cdot d((s,t), \chi_{\phi^i}(s',t')))^{-N}$$

Essential Estimates

Proposition:

- $(\Lambda_{1,j})$ and $(\Lambda_{2,j})$ have the following two properties:
  - asymptotically negligible cluster coherences:

$$\mu_c(\Lambda_{1,j}), \mu_c(\Lambda_{2,j}) \to 0, \qquad j \to \infty.$$

  - asymptotically negligible cluster approximation errors:

$$\delta_j = \delta_{1,j} + \delta_{2,j} = o(\|\mathcal{P}_j\|_2 + \|\mathcal{C}_j\|_2), \qquad j \to \infty.$$

Asymptotic Separation

Application of the abstract separation estimate then implies:

Theorem (Donoho, K; 2013)

$$\frac{\|W_j - \mathcal{P}_j\|_2 + \|S_j - \mathcal{C}_j\|_2}{\|\mathcal{P}_j\|_2 + \|\mathcal{C}_j\|_2} \to 0, \qquad j \to \infty.$$

At all sufficiently fine scales, nearly-perfect separation is achieved!

Recovery of Fourier Data

or: Fast Data Acquisition in MRI

Fourier Sampling

Important Situation:

Pointwise Samples of the Fourier transform!

Applications:

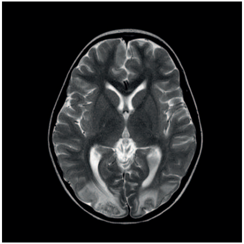
- Magnetic Resonance Imaging (MRI)
- Electron Microscopy
- Fourier Optics
- X-ray Computed Tomography
- Reflection Seismology
- ...

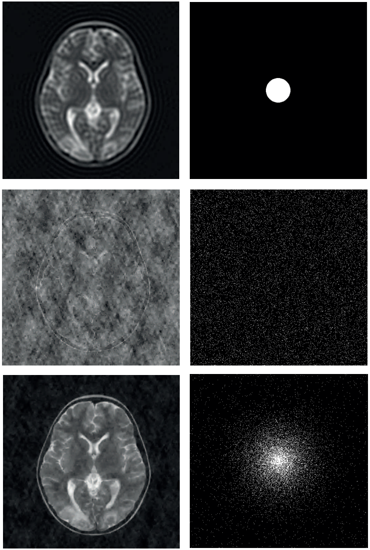
Common Model:

Let $f \in L^2(\mathbb{R}^2)$ with additional regularity assumptions, and $\Delta \subseteq \mathbb{Z}^2$.
Reconstruct f from

$$(\hat{f}(n))_{n \in \Delta} = (\langle f, e_n \rangle)_{n \in \Delta}, \quad e_n(x) := e^{2\pi i \langle x, n \rangle}.$$

Sampling of Fourier Data





(Source: Lim; 2014)

General Sampling Strategy

- Fourier measurements:

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}.$$

- Orthonormal basis:

$$\{\psi_\lambda\}_{\lambda \in \Lambda}.$$

- Sparse representation:

$$f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda.$$

- Reconstruction:

$$\left(\langle f, e_n \rangle = \sum_{\lambda \in \Lambda} \langle \psi_\lambda, e_n \rangle c_\lambda \right)_{n \in \Delta} \mapsto (c_\lambda)_{\lambda \in \Lambda}.$$

General Sampling Strategy

- Fourier measurements:

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}.$$

- Orthonormal basis:

$$\{\psi_\lambda\}_{\lambda \in \Lambda}.$$

- Sparse representation: $\longrightarrow$ Model for f?

$$f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda.$$

- Reconstruction:

$$\left(\langle f, e_n \rangle = \sum_{\lambda \in \Lambda} \langle \psi_\lambda, e_n \rangle c_\lambda \right)_{n \in \Delta} \mapsto (c_\lambda)_{\lambda \in \Lambda}.$$

General Sampling Strategy

- Fourier measurements:

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}.$$

- Orthonormal basis: $\longrightarrow$ Choice of $\{\psi_\lambda\}_{\lambda \in \Lambda}$?

$$\{\psi_\lambda\}_{\lambda \in \Lambda}.$$

- Sparse representation: $\longrightarrow$ Model for f?

$$f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda.$$

- Reconstruction:

$$\left(\langle f, e_n \rangle = \sum_{\lambda \in \Lambda} \langle \psi_\lambda, e_n \rangle c_\lambda \right)_{n \in \Delta} \mapsto (c_\lambda)_{\lambda \in \Lambda}.$$

General Sampling Strategy

- Fourier measurements: $\longrightarrow$ Sampling Scheme?

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}.$$

- Orthonormal basis: $\longrightarrow$ Choice of $\{\psi_\lambda\}_{\lambda \in \Lambda}$?

$$\{\psi_\lambda\}_{\lambda \in \Lambda}.$$

- Sparse representation: $\longrightarrow$ Model for f?

$$f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda.$$

- Reconstruction:

$$\left(\langle f, e_n \rangle = \sum_{\lambda \in \Lambda} \langle \psi_\lambda, e_n \rangle c_\lambda \right)_{n \in \Delta} \mapsto (c_\lambda)_{\lambda \in \Lambda}.$$

General Sampling Strategy

- Fourier measurements: $\longrightarrow$ Sampling Scheme?

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}.$$

- Orthonormal basis: $\longrightarrow$ Choice of $\{\psi_\lambda\}_{\lambda \in \Lambda}$?

$$\{\psi_\lambda\}_{\lambda \in \Lambda}.$$

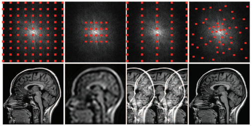
- Sparse representation: $\longrightarrow$ Model for f?

$$f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda.$$

- Reconstruction: $\longrightarrow$ Reconstruction Algorithm?

$$\left(\langle f, e_n \rangle = \sum_{\lambda \in \Lambda} \langle \psi_\lambda, e_n \rangle c_\lambda \right)_{n \in \Delta} \mapsto (c_\lambda)_{\lambda \in \Lambda}.$$

Compressed Sensing Type Approaches



- Lustig, Donoho, Pauly; 2007
  $\rightsquigarrow$ Sparse MRI: Spirals, $L^2(\mathbb{R}^2)$, Wavelets, ℓ_1.
  $$\min_g \|\Psi g\|_1 \quad \text{s.t.} \quad \|\hat{g}|_\Delta - \hat{f}|_\Delta\|_2 \leq \varepsilon.$$
- Krahmer, Ward; 2014
  $\rightsquigarrow$ Variable Density Sampling, $\mathbb{C}^{N\times N}$, Haar Wavelets, TV.
- Adcock, Hansen, K, Ma; 2014
  $\rightsquigarrow$ Block Sampling, $L^2(\mathbb{R}^2)$, Wavelets, Generalized Sampling.
- Adcock, Hansen, Poon, Roman; 2014
  $\rightsquigarrow$ Multilevel Sampling, $\mathcal{H}$, ONS, ℓ_1.
- Shi, Yin, Sankaranarayanan, Baraniuk; 2014
  $\rightsquigarrow$ Dynamic MRI: Variable Density Sampling, $\mathbb{R} \times \mathbb{R}^n$, Wavelets, ℓ_1.
- ...

Appropriate Notion of Optimality?

Ingredients:

- Continuum Model $\mathcal{C} \subseteq L^2(\mathbb{R}^2)$.
  - Acquiring data in a continuous world.
  - Optimal best N-term approximation rate:

$$\|f - f_N\|_2 \lesssim N^{-\alpha} \text{ as } N \to \infty \text{ for all } f \in \mathcal{C},$$

where $f_N = \sum_{\lambda \in \Lambda_N} c_\lambda \psi_\lambda$ for some frame $(\psi_\lambda)_{\lambda \in \Lambda} \subseteq L^2(\mathbb{R}^2)$.

Appropriate Notion of Optimality?

Ingredients:

- Continuum Model $\mathcal{C} \subseteq L^2(\mathbb{R}^2)$.
  - Acquiring data in a continuous world.
  - Optimal best N-term approximation rate:

$$\|f - f_N\|_2 \lesssim N^{-\alpha} \text{ as } N \to \infty \text{ for all } f \in \mathcal{C},$$

  where $f_N = \sum_{\lambda \in \Lambda_N} c_\lambda \psi_\lambda$ for some frame $(\psi_\lambda)_{\lambda \in \Lambda} \subseteq L^2(\mathbb{R}^2)$.
- Sampling Schemes $\Delta_M \subseteq \mathbb{Z}^2$, $\#\Delta_M = M$ and $M \to \infty$.

Appropriate Notion of Optimality?

Ingredients:

- Continuum Model $\mathcal{C} \subseteq L^2(\mathbb{R}^2)$.
  - Acquiring data in a continuous world.
  - Optimal best N-term approximation rate:

$$\|f - f_N\|_2 \lesssim N^{-\alpha} \text{ as } N \to \infty \text{ for all } f \in \mathcal{C},$$

  where $f_N = \sum_{\lambda \in \Lambda_N} c_\lambda \psi_\lambda$ for some frame $(\psi_\lambda)_{\lambda \in \Lambda} \subseteq L^2(\mathbb{R}^2)$.
- Sampling Schemes $\Delta_M \subseteq \mathbb{Z}^2$, $\#\Delta_M = M$ and $M \to \infty$.
- Reconstruction Procedure $\mathcal{R} : \mathcal{C} \times \Delta \to L^2(\mathbb{R}^2)$, $\Delta = \bigcup_M \{\Delta_M\}$.

Appropriate Notion of Optimality?

Ingredients:

- Continuum Model $\mathcal{C} \subseteq L^2(\mathbb{R}^2)$.
  - Acquiring data in a continuous world.
  - Optimal best N-term approximation rate:

$$\|f - f_N\|_2 \lesssim N^{-\alpha} \text{ as } N \to \infty \text{ for all } f \in \mathcal{C},$$

  where $f_N = \sum_{\lambda \in \Lambda_N} c_\lambda \psi_\lambda$ for some frame $(\psi_\lambda)_{\lambda \in \Lambda} \subseteq L^2(\mathbb{R}^2)$.
- Sampling Schemes $\Delta_M \subseteq \mathbb{Z}^2$, $\#\Delta_M = M$ and $M \to \infty$.
- Reconstruction Procedure $\mathcal{R} : \mathcal{C} \times \Delta \to L^2(\mathbb{R}^2)$, $\Delta = \bigcup_M \{\Delta_M\}$.

Asymptotic Optimality: We call a sampling-reconstruction scheme $(\mathcal{C}, \Delta, \mathcal{R})$ asymptotically optimal, if, for all $f \in \mathcal{C}$,

$$\|f - \mathcal{R}(f, \Delta_M)\|_2 \lesssim M^{-\alpha} \text{ as } M \to \infty.$$

General Sampling Strategy

- Fourier measurements: $\longrightarrow$ Sampling Scheme?

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}.$$

- Orthonormal basis: $\longrightarrow$ Choice of $\{\psi_\lambda\}_{\lambda \in \Lambda}$?

$$\{\psi_\lambda\}_{\lambda \in \Lambda}.$$

- Sparse representation: $\longrightarrow$ Model for f?

$$f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda.$$

- Reconstruction: $\longrightarrow$ Reconstruction Algorithm?

$$\left(\langle f, e_n \rangle = \sum_{\lambda \in \Lambda} \langle \psi_\lambda, e_n \rangle c_\lambda \right)_{n \in \Delta} \mapsto (c_\lambda)_{\lambda \in \Lambda}.$$

General Sampling Strategy

- Fourier measurements: $\longrightarrow$ Sampling Scheme?

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}.$$

- Orthonormal basis: $\longrightarrow$ Choice of $\{\psi_\lambda\}_{\lambda \in \Lambda}$?

$$\{\psi_\lambda\}_{\lambda \in \Lambda}.$$

- Sparse representation:

$$f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda, \text{ where } f \text{ is a cartoon-like function.}$$

- Reconstruction: $\longrightarrow$ Reconstruction Algorithm?

$$\left(\langle f, e_n \rangle = \sum_{\lambda \in \Lambda} \langle \psi_\lambda, e_n \rangle c_\lambda \right)_{n \in \Delta} \mapsto (c_\lambda)_{\lambda \in \Lambda}.$$

General Sampling Strategy

- Fourier measurements: $\longrightarrow$ Sampling Scheme?

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}.$$

- Shearlet frame:

$$\{\psi_\lambda\}_{\lambda \in \Lambda}.$$

- Sparse representation:

$$f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda, \text{ where } f \text{ is a cartoon-like function.}$$

- Reconstruction: $\longrightarrow$ Reconstruction Algorithm?

$$\left(\langle f, e_n \rangle = \sum_{\lambda \in \Lambda} \langle \psi_\lambda, e_n \rangle c_\lambda \right)_{n \in \Delta} \mapsto (c_\lambda)_{\lambda \in \Lambda}.$$

Frame Theory

Problem: Let $\{\psi_\lambda\}_{\lambda\in\Lambda}$ be a frame for $\mathcal{H}$. In general, it is not true that

$$f = \sum_{\lambda\in\Lambda} \langle f, \psi_\lambda \rangle \, \psi_\lambda \quad \text{for all } f \in \mathcal{H}.$$

Theorem: We have

$$f = \sum_{\lambda\in\Lambda} \langle f, \psi_\lambda \rangle \, \tilde{\psi}_\lambda \quad \text{for all } f \in \mathcal{H},$$

where $\{\tilde{\psi}_\lambda := S^{-1}\psi_\lambda\}_{\lambda\in\Lambda}$ is the associated (canonical) dual frame and S the associated frame operator.

Problem with Frames

- Fourier measurements: $\longrightarrow$ Sampling Scheme?

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}.$$

- Shearlet frame:

$$\{\psi_\lambda\}_{\lambda \in \Lambda}.$$

- Sparse representation:

$$f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda, \text{ where } f \text{ is a cartoon-like function.}$$

- Reconstruction: $\longrightarrow$ Reconstruction Algorithm?

$$\left(\langle f, e_n \rangle = \sum_{\lambda \in \Lambda} \langle \psi_\lambda, e_n \rangle c_\lambda \right)_{n \in \Delta} \mapsto (c_\lambda)_{\lambda \in \Lambda}.$$

Problem with Frames

- Fourier measurements: $\longrightarrow$ Sampling Scheme?

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}.$$

- Shearlet frame:

$$\{\psi_\lambda\}_{\lambda \in \Lambda}.$$

- Sparse representation:

$$f = \sum_{\lambda \in \Lambda} c_\lambda \tilde{\psi}_\lambda, \text{ where } c_\lambda = \langle f, \psi_\lambda \rangle \text{ and } f \text{ is a cartoon-like function.}$$

- Reconstruction: $\longrightarrow$ Reconstruction Algorithm?

$$\left(\langle f, e_n \rangle = \sum_{\lambda \in \Lambda} \langle \psi_\lambda, e_n \rangle c_\lambda \right)_{n \in \Delta} \mapsto (c_\lambda)_{\lambda \in \Lambda}.$$

Problem with Frames

- Fourier measurements: $\longrightarrow$ Sampling Scheme?

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}.$$

- Shearlet frame:

$$\{\psi_\lambda\}_{\lambda \in \Lambda}.$$

- Sparse representation:

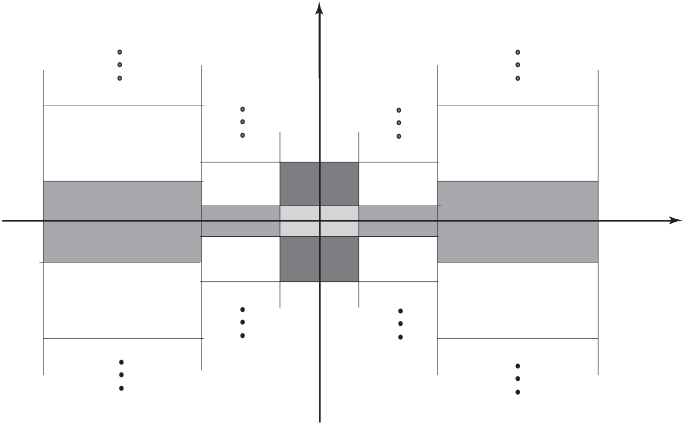
$$f = \sum_{\lambda \in \Lambda} c_\lambda \tilde{\psi}_\lambda, \text{ where } c_\lambda = \langle f, \psi_\lambda \rangle \text{ and } f \text{ is a cartoon-like function.}$$

- Reconstruction: $\longrightarrow$ Reconstruction Algorithm?

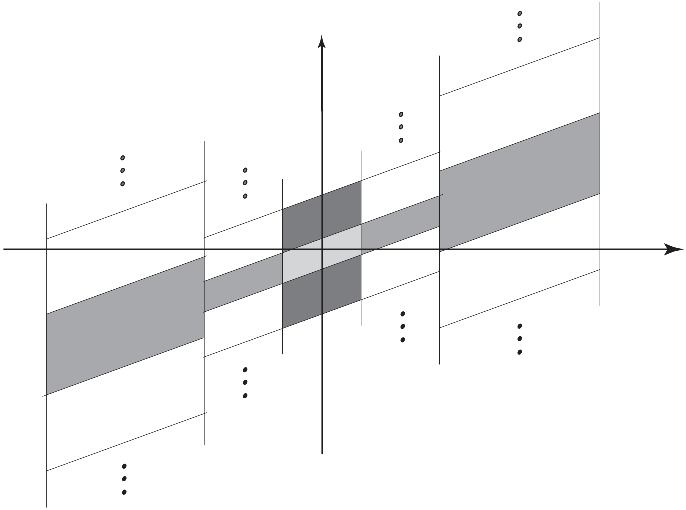
$$\left(\langle f, e_n \rangle = \sum_{\lambda \in \Lambda} \langle \tilde{\psi}_\lambda, e_n \rangle c_\lambda \right)_{n \in \Delta} \mapsto (c_\lambda)_{\lambda \in \Lambda}.$$

Dualizable Shearlets...

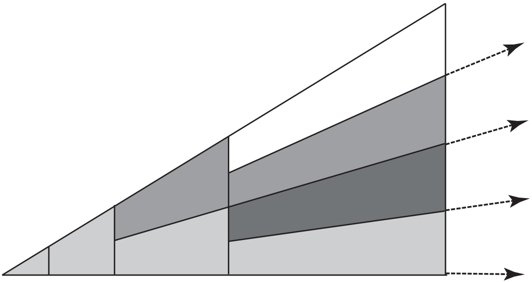
Intuition: Partition of Fourier Domain, shear= 0



Intuition: Partition of Fourier Domain, shear$\neq 0$



Intuition: Filters

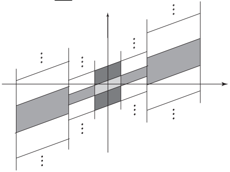


Shearlet Generators

Let $\gamma \in L^2(\mathbb{R}^2)$ be compactly supported such that, for $\rho > 0$ fixed,

$$|\partial^d \hat{\gamma}(\xi)| \lesssim \frac{\min\{1, |\xi_1|^\alpha\}}{(1+|\xi_1|)^\beta (1+|\xi_2|)^\beta} \quad \text{for all } d \le R$$

with $R \geq 1, \alpha \geq 1 + \frac{6}{\rho}$, and $\beta > \alpha + 1$.



Observation:
For each s,

$$\{\gamma^s_{j,m} = 2^{\frac{3}{4}j}\gamma(A_j S_s \cdot - m) : j, m\} \quad \text{and} \quad \{\tilde{\gamma^s}_{j,m} = 2^{\frac{3}{4}j}\tilde{\gamma}(\tilde{A}_j S_s^* \cdot - m) : j, m\}$$

form orthonormal bases for $L^2(\mathbb{R}^2)$.

Dualizable Shearlet Frame

For some regularity parameter $\rho > 0$, define

$$\psi_{j,k,m} = \Theta_s * \gamma_{j,m}^s \quad \text{and} \quad \tilde{\psi}_{j,k,m} = \tilde{\Theta}_s * \tilde{\gamma}_{j,m}^s \quad \text{with } s = 2^{-j/2}k.$$

Dualizable Shearlet Frame

For some regularity parameter $\rho > 0$, define

$$\psi_{j,k,m} = \Theta_s * \gamma_{j,m}^s \quad \text{and} \quad \tilde{\psi}_{j,k,m} = \tilde{\Theta}_s * \tilde{\gamma}_{j,m}^s \quad \text{with } s = 2^{-j/2}k.$$

Theorem (K, Lim; 2014):
The dualizable shearlet system

$$\mathcal{SH} := \{\psi_{j,k,m}, \tilde{\psi}_{j,k,m} : j \geq 0, |k| < 2^{j/2}, m \in \mathbb{Z}^2\}$$

forms a compactly supported frame and a dual frame is given by

$$\left\{\mathcal{F}^{-1}\left(\frac{\hat{\psi}_{j,k,m}}{\sum_s |\hat{\Theta}_s|^2}\right), \mathcal{F}^{-1}\left(\frac{\hat{\tilde{\psi}}_{j,k,m}}{\sum_s |\hat{\tilde{\Theta}}_s|^2}\right) : \psi_{j,k,m}, \tilde{\psi}_{j,k,m} \in \mathcal{SH}\right\}.$$

Optimal Sparse Approximation inherited!

Theorem (K, Lim; 2014):
Let f be a cartoon-like function and let $\mathcal{SH} = (\psi_\lambda)_{\lambda \in \Lambda}$ be as before. Then, for any $\rho > 0$, there exists a positive constant C_ρ such that

$$\|f - f_N\|_2^2 \lesssim N^{-2+15\rho} \cdot (\log(N))^2,$$

where f_N is the N term approximation (of the N largest $\langle f, \psi_\lambda \rangle$'s) with respect to the dual frame of $\mathcal{SH}$, i.e.

$$f_N = \sum_{\lambda \in \Lambda_N} \langle f, \psi_\lambda \rangle \tilde{\psi}_\lambda.$$

Recall:

- Optimal rate: N^{-2}.
- Regularity parameter: $\rho > 0$.

General Sampling Strategy

- Fourier measurements: $\longrightarrow$ Sampling Scheme?

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}.$$

- Dualizable Shearlet frame:

$$\{\psi_\lambda\}_{\lambda \in \Lambda}.$$

- Sparse representation:

$$f = \sum_{\lambda \in \Lambda} c_\lambda \tilde{\psi}_\lambda, \text{ where } c_\lambda = \langle f, \psi_\lambda \rangle \text{ and } f \text{ is a cartoon-like function.}$$

- Reconstruction: $\longrightarrow$ Reconstruction Algorithm?

$$\left(\langle f, e_n \rangle = \sum_{\lambda \in \Lambda} \langle \tilde{\psi}_\lambda, e_n \rangle c_\lambda \right)_{n \in \Delta} \mapsto (c_\lambda)_{\lambda \in \Lambda}.$$

Directional Sampling Strategy

Sampling Strategy: Dualizable Shearlet Systems

Recall: We have $(k \leftrightarrow s)$

$$\langle f, \psi_{j,k,m} \rangle = \langle f, \Theta_s * \gamma^s_{j,m} \rangle = \langle \overline{\Theta}_s * f, \gamma^s_{j,m} \rangle = c^s_{j,m}.$$

Sampling Strategy: Dualizable Shearlet Systems

Recall: We have ($k \leftrightarrow s$)

$$\langle f, \psi_{j,k,m} \rangle = \langle f, \Theta_s * \gamma^s_{j,m} \rangle = \langle \overline{\Theta}_s * f, \gamma^s_{j,m} \rangle = c^s_{j,m}.$$

Determining the measurement vector:

$$\overline{\Theta}_s * f = \sum_{(j,m) \in \Lambda_s} c^s_{j,m} \gamma^s_{j,m}$$

Sampling Strategy: Dualizable Shearlet Systems

Recall: We have ($k \leftrightarrow s$)

$$\langle f, \psi_{j,k,m} \rangle = \langle f, \Theta_s * \gamma^s_{j,m} \rangle = \langle \overline{\Theta}_s * f, \gamma^s_{j,m} \rangle = c^s_{j,m}.$$

Determining the measurement vector:

$$\overline{\Theta}_s * f = \sum_{(j,m)\in\Lambda_s} c^s_{j,m} \gamma^s_{j,m} \implies \langle \overline{\Theta}_s * f, e_n \rangle = \sum_{(j,m)\in\Lambda_s} \langle \gamma^s_{j,m}, e_n \rangle c^s_{j,m}$$

Sampling Strategy: Dualizable Shearlet Systems

Recall: We have ($k \leftrightarrow s$)

$$\langle f, \psi_{j,k,m} \rangle = \langle f, \Theta_s * \gamma^s_{j,m} \rangle = \langle \overline{\Theta}_s * f, \gamma^s_{j,m} \rangle = c^s_{j,m}.$$

Determining the measurement vector:

$$\overline{\Theta}_s * f = \sum_{(j,m)\in\Lambda_s} c^s_{j,m} \gamma^s_{j,m} \implies \langle \overline{\Theta}_s * f, e_n \rangle = \sum_{(j,m)\in\Lambda_s} \langle \gamma^s_{j,m}, e_n \rangle c^s_{j,m}$$

$$\implies \langle P^s_J(\overline{\Theta}_s * f), e_n \rangle = \sum_{(j,m)\in\Lambda_{J,s}} \langle \gamma^s_{j,m}, e_n \rangle c^s_{j,m}$$

Sampling Strategy: Dualizable Shearlet Systems

Recall: We have ($k \leftrightarrow s$)

$$\langle f, \psi_{j,k,m} \rangle = \langle f, \Theta_s * \gamma^s_{j,m} \rangle = \langle \overline{\Theta}_s * f, \gamma^s_{j,m} \rangle = c^s_{j,m}.$$

Determining the measurement vector:

$$\overline{\Theta}_s * f = \sum_{(j,m)\in\Lambda_s} c^s_{j,m}\gamma^s_{j,m} \implies \langle \overline{\Theta}_s * f, e_n \rangle = \sum_{(j,m)\in\Lambda_s} \langle \gamma^s_{j,m}, e_n \rangle c^s_{j,m}$$

$$\implies \langle P^s_J(\overline{\Theta}_s * f), e_n \rangle = \sum_{(j,m)\in\Lambda_{J,s}} \langle \gamma^s_{j,m}, e_n \rangle c^s_{j,m}$$

Hence, we preliminarily set

$$y_n := \langle P^s_J(\overline{\Theta}_s * f), e_n \rangle.$$

Sampling Strategy: Dualizable Shearlet Systems

Recall: We have ($k \leftrightarrow s$)

$$\langle f, \psi_{j,k,m} \rangle = \langle f, \Theta_s * \gamma^s_{j,m} \rangle = \langle \overline{\Theta}_s * f, \gamma^s_{j,m} \rangle = c^s_{j,m}.$$

Determining the measurement vector:

$$\overline{\Theta}_s * f = \sum_{(j,m)\in\Lambda_s} c^s_{j,m} \gamma^s_{j,m} \implies \langle \overline{\Theta}_s * f, e_n \rangle = \sum_{(j,m)\in\Lambda_s} \langle \gamma^s_{j,m}, e_n \rangle c^s_{j,m}$$

$$\implies \langle P^s_J(\overline{\Theta}_s * f), e_n \rangle = \sum_{(j,m)\in\Lambda_{J,s}} \langle \gamma^s_{j,m}, e_n \rangle c^s_{j,m}$$

Hence, we preliminarily set

$$y_n := \langle P^s_J(\overline{\Theta}_s * f), e_n \rangle.$$

Remark: In practice, $P^s_J(\overline{\Theta}_s * f) \approx \overline{\Theta}_s * f$, hence $y_n = \widehat{\overline{\Theta}_s}(n) \cdot \hat{f}(n)$.

General Sampling Scheme

- Fourier measurements: $\longrightarrow$ Sampling Scheme?

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}.$$

- Dualizable Shearlet frame:

$$\{\psi_\lambda\}_{\lambda \in \Lambda}.$$

- Sparse representation:

$$f = \sum_{\lambda \in \Lambda} c_\lambda \tilde{\psi}_\lambda, \text{ where } c_\lambda = \langle f, \psi_\lambda \rangle \text{ and } f \text{ is a cartoon-like function.}$$

- Reconstruction: $\longrightarrow$ Reconstruction Algorithm?

$$\left(\langle f, e_n \rangle = \sum_{\lambda \in \Lambda} \langle \tilde{\psi}_\lambda, e_n \rangle c_\lambda \right)_{n \in \Delta} \mapsto (c_\lambda)_{\lambda \in \Lambda}.$$

General Sampling Scheme

- Fourier measurements: $\longrightarrow$ Sampling Scheme?

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}.$$

- Dualizable Shearlet frame:

$$\{\psi_\lambda\}_{\lambda \in \Lambda}.$$

- Sparse representation:

$$f = \sum_{\lambda \in \Lambda} c_\lambda \tilde{\psi}_\lambda, \text{ where } c_\lambda = \langle f, \psi_\lambda \rangle \text{ and } f \text{ is a cartoon-like function.}$$

- Reconstruction:

$$(c_\lambda)_{\lambda \in \Lambda} = \text{argmin}_{(\tilde{c}_\lambda)_{\lambda \in \Lambda}} \|(\tilde{c}_\lambda)_{\lambda \in \Lambda}\|_1 \text{ s.t. } \Big(\langle f, e_n \rangle = \sum_{\lambda \in \Lambda} \langle \tilde{\psi}_\lambda, e_n \rangle \tilde{c}_\lambda \Big)_{n \in \Delta}.$$

Shear-Adapted Density Sampling

Linear System of Equations:

$$\langle P_J^s(\overline{\Theta}_s * f), e_n \rangle = \sum_{(j,m)\in\Lambda_{J,s}} \langle \gamma_{j,m}^s, e_n \rangle c_{j,m}^s.$$

Shear-Adapted Density Sampling

Linear System of Equations:

$$\langle P_J^s(\overline{\Theta}_s * f), e_n \rangle = \sum_{(j,m)\in\Lambda_{J,s}} \langle \gamma_{j,m}^s, e_n \rangle c_{j,m}^s.$$

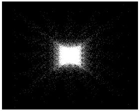
Introducing Randomness:

$$\frac{1}{\sqrt{p_s(n_{s,\ell})}} \langle P_J^s(\overline{\Theta}_s * f), e_{n_{s,\ell}} \rangle = \sum_{(j,m)\in\Lambda_{J,s}} \underbrace{\left[\frac{1}{\sqrt{p_s(n_{s,\ell})}} \langle \gamma_{j,m}^s, e_{n_{s,\ell}} \rangle \right]}_{\Phi_s :=} c_{j,m}^s,$$

where

- $s \in \mathbb{S}_{J/2} := \{0\} \cup \{\frac{q}{2^{j/2}} : |q| < 2^{j/2}, q \in 2\mathbb{Z}+1, j = 0, \ldots, J\}$,
- $\{n_{s,\ell} : \ell = 1, \ldots, L_s\} \subseteq \mathbb{Z}^2 \cap [-2^{J(1+\rho)}, 2^{J(1+\rho)}]^2$ is chosen according to a probability density function

$$p_s(n) = \frac{c_s}{J^2(1+|n_1|)(1+|n_2 - s n_1|)}.$$



Sparse Sampling Strategy

Theorem (K, Lim; 2015):

Let f be a cartoon-like function which is $C^{2,r}$, $r \in [\frac{1}{4}, 1)$ smooth apart from a C^2-discontinuity curve of non-vanishing curvature. Further, let

- $\rho > 0$ be fixed (regularity),
- $J > 0$ be 'sufficiently large' (limiting scale),
- $y_s := \left(\sqrt{p_s(n_{s,\ell})}^{-1} \langle P_J^s(\overline{\Theta}_s * f), e_{n_{s,\ell}} \rangle\right)_{\ell=1,\ldots,L_s}$, (measurements),
- $\Phi_s := \left(\sqrt{p_s(n_{s,\ell})}^{-1} \langle \gamma_{j,m}^s, e_{n_{s,\ell}} \rangle\right)_{(j,m)\in\Lambda_{J,s},\ell=1,\ldots,L_s}$ (sampling matrix).

Sparse Sampling Strategy

Theorem (K, Lim; 2015):

Let f be a cartoon-like function which is $C^{2,r}$, $r \in [\frac{1}{4}, 1)$ smooth apart from a C^2-discontinuity curve of non-vanishing curvature. Further, let

- $\rho > 0$ be fixed (regularity),
- $J > 0$ be 'sufficiently large' (limiting scale),
- $y_s := \left(\sqrt{p_s(n_{s,\ell})}^{-1} \langle P_J^s(\overline{\Theta}_s * f), e_{n_{s,\ell}} \rangle\right)_{\ell=1,\ldots,L_s}$, (measurements),
- $\Phi_s := \left(\sqrt{p_s(n_{s,\ell})}^{-1} \langle \gamma_{j,m}^s, e_{n_{s,\ell}} \rangle\right)_{(j,m)\in\Lambda_{J,s},\ell=1,\ldots,L_s}$ (sampling matrix).

For each $s \in \mathbb{S}_{J/2}$,

$$(\hat{c}_\lambda)_{\lambda\in\Lambda_{J,s}} = \operatorname{argmin}_c \|c\|_1 \text{ subject to } \Phi_s c = y_s,$$

Sparse Sampling Strategy

Theorem (K, Lim; 2015):

Let f be a cartoon-like function which is $C^{2,r}$, $r \in [\frac{1}{4}, 1)$ smooth apart from a C^2-discontinuity curve of non-vanishing curvature. Further, let

- $\rho > 0$ be fixed (regularity),
- $J > 0$ be 'sufficiently large' (limiting scale),
- $y_s := \left(\sqrt{p_s(n_{s,\ell})}^{-1} \langle P_J^s(\overline{\Theta}_s * f), e_{n_{s,\ell}} \rangle \right)_{\ell=1,\ldots,L_s}$, (measurements),
- $\Phi_s := \left(\sqrt{p_s(n_{s,\ell})}^{-1} \langle \gamma_{j,m}^s, e_{n_{s,\ell}} \rangle \right)_{(j,m)\in\Lambda_{J,s},\ell=1,\ldots,L_s}$ (sampling matrix).

For each $s \in \mathbb{S}_{J/2}$,

$$(\hat{c}_\lambda)_{\lambda\in\Lambda_{J,s}} = \text{argmin}_c \|c\|_1 \text{ subject to } \Phi_s c = y_s,$$

Then with probability at least $1 - 2^{-J}$,

$$\left\| f - \sum_{s\in\mathbb{S}_{J/2}} \sum_{\lambda\in\Lambda_{J,s}} \hat{c}_\lambda \tilde{\psi}_\lambda \right\|_2^2 \lesssim 2^{-J(1-13\rho/2)} \quad \text{as } J \to \infty.$$

Sparse Sampling Strategy

Theorem (K, Lim; 2015):

Let f be a cartoon-like function which is $C^{2,r}$, $r \in [\frac{1}{4}, 1)$ smooth apart from a C^2-discontinuity curve of non-vanishing curvature. Further, let

- $\rho > 0$ be fixed (regularity),
- $J > 0$ be 'sufficiently large' (limiting scale),
- $y_s := \left(\sqrt{p_s(n_{s,\ell})}^{-1} \langle P_J^s(\overline{\Theta}_s * f), e_{n_{s,\ell}} \rangle\right)_{\ell=1,\dots,L_s}$, (measurements),
- $\Phi_s := \left(\sqrt{p_s(n_{s,\ell})}^{-1} \langle \gamma_{j,m}^s, e_{n_{s,\ell}} \rangle\right)_{(j,m)\in\Lambda_{J,s},\ell=1,\dots,L_s}$ (sampling matrix).

For each $s \in \mathbb{S}_{J/2}$, $(\sum_{s\in\mathbb{S}_{J/2}} L_s \lesssim J 2^{J/2(1+2\rho)} =: N)$

$$(\hat{c}_\lambda)_{\lambda\in\Lambda_{J,s}} = \text{argmin}_c \|c\|_1 \text{ subject to } \Phi_s c = y_s,$$

Then with probability at least $1 - 2^{-J}$,

$$\left\| f - \sum_{s\in\mathbb{S}_{J/2}} \sum_{\lambda\in\Lambda_{J,s}} \hat{c}_\lambda \tilde{\psi}_\lambda \right\|_2^2 \lesssim 2^{-J(1-13\rho/2)} \quad \text{as } J \to \infty.$$

Sparse Sampling Strategy

Theorem (K, Lim; 2015):

Let f be a cartoon-like function which is $C^{2,r}$, $r \in [\frac{1}{4}, 1)$ smooth apart from a C^2-discontinuity curve of non-vanishing curvature. Further, let

- $\rho > 0$ be fixed (regularity),
- $J > 0$ be 'sufficiently large' (limiting scale),
- $y_s := \left(\sqrt{p_s(n_{s,\ell})}^{-1} \langle P_J^s(\overline{\Theta}_s * f), e_{n_{s,\ell}} \rangle \right)_{\ell=1,\dots,L_s}$, (measurements),
- $\Phi_s := \left(\sqrt{p_s(n_{s,\ell})}^{-1} \langle \gamma_{j,m}^s, e_{n_{s,\ell}} \rangle \right)_{(j,m)\in\Lambda_{J,s},\ell=1,\dots,L_s}$ (sampling matrix).

For each $s \in \mathbb{S}_{J/2}$, $\left(\sum_{s\in\mathbb{S}_{J/2}} L_s \lesssim J2^{J/2(1+2\rho)} =: N\right)$

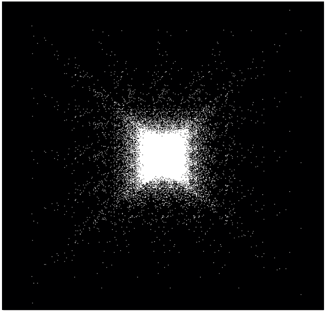
$$(\hat{c}_\lambda)_{\lambda\in\Lambda_{J,s}} = \operatorname{argmin}_c \|c\|_1 \text{ subject to } \Phi_s c = y_s,$$

Then with probability at least $1 - 2^{-J}$, $\rightsquigarrow$ Asymptotic Optimality!

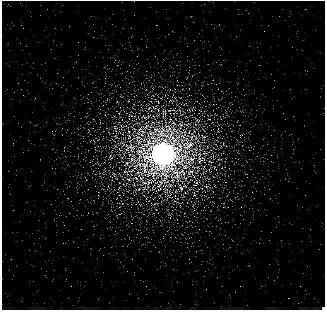
$$\left\| f - \sum_{s\in\mathbb{S}_{J/2}} \sum_{\lambda\in\Lambda_{J,s}} \hat{c}_\lambda \tilde{\psi}_\lambda \right\|_2^2 \lesssim 2^{-J(1-13\rho/2)} (= O(N^{-2+C\rho})) \quad \text{as } J \to \infty.$$

Numerical Experiments

Sampling Schemes

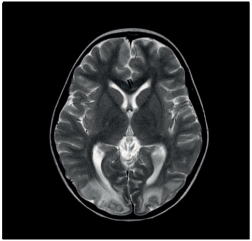


Directional Sampling Scheme

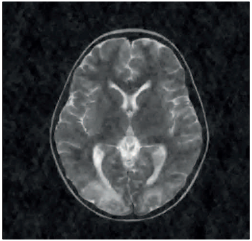


Variable Density Sampling Scheme

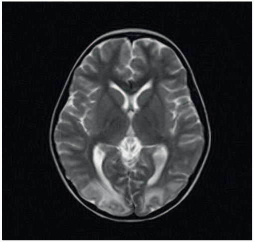
Numerical Results for 512x512 MRI Image



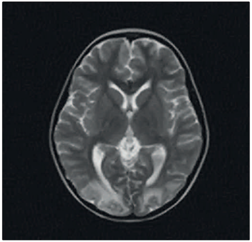
Original



Wavelets + Variable Density Sampling
(5% sampling rate, 24.9969dB)

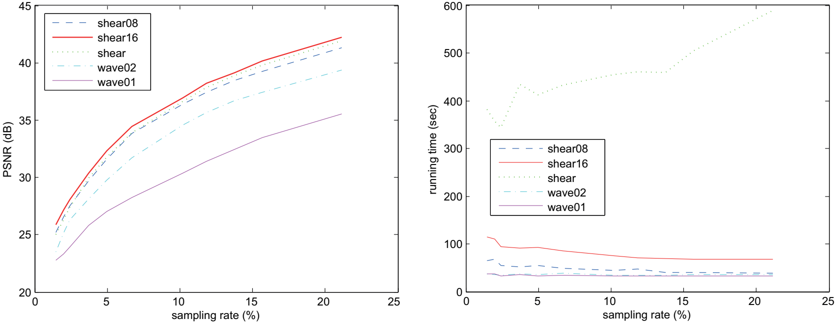


Shearlet Scheme
(5% sampling rate, 32.2845dB)



Wavelets + Directional Sampling
(5% sampling rate, 29.8138dB)

Approximation Curves for 512x512 MRI Image



- **shear08**: Directional sampling scheme with 8 directional filters.
- **shear16**: Directional sampling scheme with 16 directional filters.
- **shear**: Directional sampling scheme with (normal) shearlets.
- **wave02**: Directional sampling scheme with wavelets.
- **wave01**: Variable density sampling scheme with wavelets.

Let's conclude...

What to take Home...?

- Computational harmonic analysis and sparse approximation are a powerful combination to solve ill-posed inverse problems in imaging.
- Such a sparse regularization approach allows also precise theoretical results.
- We discussed the following inverse problems:
    - Feature Extraction
    - Magnetic Resonance Imaging
- Further applications include:
    - Inpainting
    - Edge Detection
    - ...

THANK YOU!

References available at:

www.math.tu-berlin.de/∼kutyniok

Code available at:

www.ShearLab.org

Related Books:

- Y. Eldar and G. Kutyniok
  Compressed Sensing: Theory and Applications
  Cambridge University Press, 2012.
- G. Kutyniok and D. Labate
  Shearlets: Multiscale Analysis for Multivariate Data
  Birkhäuser-Springer, 2012.