Optimal Compressive Imaging for Fourier Data

Gitta Kutyniok

(Technische Universität Berlin)

joint with

Wang-Q Lim (Technische Universität Berlin)

UCL-Duke Workshop "Sensing & Analysis of High-Dimensional Data"
University College London, September 4 – 5, 2014

Sensing Data

In the age of Big Data, acquiring data is of tremendous importance, but a highly difficult task. Each new technology requires a sensibly adapted methodology.

Common desiderata:

- Few samples to acquire the data.
- Efficient reconstruction procedure.
- Good approximation rates.
- Robustness to noise, etc.

Sensing Data

In the age of Big Data, acquiring data is of tremendous importance, but a highly difficult task. Each new technology requires a sensibly adapted methodology.

Common desiderata:

- Few samples to acquire the data.
- Efficient reconstruction procedure.
- Good approximation rates.
- Robustness to noise, etc.

Novel Approach:

- Compressed Sensing.
- Frame Theory.

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:

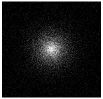
Let $f \in L^2(\mathbb{R}^2)$ with additional regularity assumptions, and $\Delta \subseteq \mathbb{Z}^2$.
Reconstruct f from

$$(\hat{f}(n))_{n \in \Delta}.$$

General Sampling Strategy using Sparsity

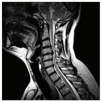
- Fourier measurements of $f \in L^2(\mathbb{R}^2)$:

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}, \quad \text{where } \Delta \subseteq \mathbb{Z}^2,\ e_n(x) := e^{2\pi i \langle x, n \rangle}.$$



- Sparse representation:

$$f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda \quad \text{with some frame } (\psi_\lambda)_{\lambda \in \Lambda} \subseteq L^2(\mathbb{R}^2).$$



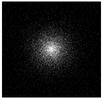
- Reconstruction:

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

General Sampling Strategy using Sparsity

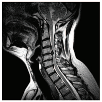
- Fourier measurements of $f \in L^2(\mathbb{R}^2)$: $\longrightarrow$ Choice of Δ?

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}, \quad \text{where } \Delta \subseteq \mathbb{Z}^2,\ e_n(x) := e^{2\pi i \langle x, n \rangle}.$$



- Sparse representation:

$$f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda \quad \text{with some frame } (\psi_\lambda)_{\lambda \in \Lambda} \subseteq L^2(\mathbb{R}^2).$$



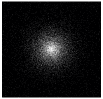
- Reconstruction:

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

General Sampling Strategy using Sparsity

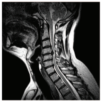
- Fourier measurements of $f \in L^2(\mathbb{R}^2)$: $\longrightarrow$ Choice of Δ?

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}, \quad \text{where } \Delta \subseteq \mathbb{Z}^2,\ e_n(x) := e^{2\pi i \langle x, n \rangle}.$$



- Sparse representation: $\longrightarrow$ Model for f and choice of $(\psi_\lambda)_{\lambda \in \Lambda}$?

$$f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda \quad \text{with some frame } (\psi_\lambda)_{\lambda \in \Lambda} \subseteq L^2(\mathbb{R}^2).$$



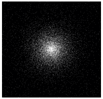
- Reconstruction:

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

General Sampling Strategy using Sparsity

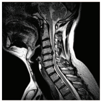
- Fourier measurements of $f \in L^2(\mathbb{R}^2)$: $\longrightarrow$ Choice of Δ?

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta}, \quad \text{where } \Delta \subseteq \mathbb{Z}^2,\ e_n(x) := e^{2\pi i \langle x, n \rangle}.$$



- Sparse representation: $\longrightarrow$ Model for f and choice of $(\psi_\lambda)_{\lambda \in \Lambda}$?

$$f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda \quad \text{with some frame } (\psi_\lambda)_{\lambda \in \Lambda} \subseteq L^2(\mathbb{R}^2).$$



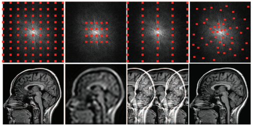
- Reconstruction: $\longrightarrow$ Reconstruction Algorithm?

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

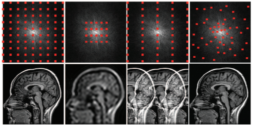
Compressed Sensing Type Approaches

- Lustig, Donoho, Pauly; 2007
  $\rightsquigarrow$ Sparse MRI: Spirals, $L^2(\mathbb{R}^2)$, Wavelets, ℓ_1.

$$\min_g \|\Psi g\|_1 \quad \text{s.t.} \quad \|\hat{g}|_\Delta - \hat{f}|_\Delta\|_2 \leq \varepsilon.$$

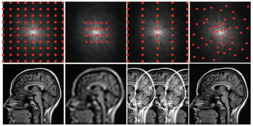


Compressed Sensing Type Approaches



- 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.

Compressed Sensing Type Approaches

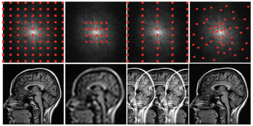


- Lustig, Donoho, Pauly; 2007
  $\rightsquigarrow$ Sparse MRI: Spirals, $L^2(\mathbb{R}^2)$, Wavelets, ℓ_1.
  $$\min_g \|\Psi g\|_1 \quad \text{s.t.} \quad \|\hat{g}|_\Delta - \hat{f}|_\Delta\|_2 \leq \varepsilon.$$
- Krahmer, Ward; 2014
  $\rightsquigarrow$ Variable Density Sampling, $\mathbb{C}^{N\times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

Compressed Sensing Type Approaches

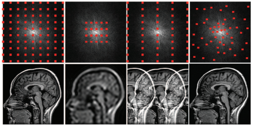
- Lustig, Donoho, Pauly; 2007
  $\rightsquigarrow$ Sparse MRI: Spirals, $L^2(\mathbb{R}^2)$, Wavelets, ℓ_1.

$$\min_g \|\Psi g\|_1 \quad \text{s.t.} \quad \|\hat{g}|_\Delta - \hat{f}|_\Delta\|_2 \leq \varepsilon.$$



- 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.

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 $f \in \mathcal{C}$,

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

Looking ahead...

Key Ideas of our Approach

- Continuum Model
  $\rightsquigarrow$ Functions governed by anisotropic structures.
- Frame
  $\rightsquigarrow$ Dualizable (compactly supported) shearlets.
- Sampling Scheme
  $\rightsquigarrow$ Directional variable density sampling.
- Reconstruction Procedure
  $\rightsquigarrow$ ℓ_1 minimization.

Main Results:

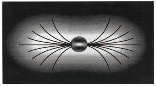
- *This leads to a provably asymptotically optimal scheme.*
- *This scheme outperforms in particular all wavelet-based schemes.*

Let's start with a suitable Model...

Anisotropic/Cartoon Structures

Images:

- Governing structure in images.
- Justified by neurophysiology.

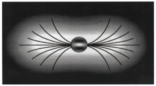


Field et al., 1993

Anisotropic/Cartoon Structures

Images:

- Governing structure in images.
- Justified by neurophysiology.



Field et al., 1993

Definition (Donoho; 2001):

The set of cartoon-like functions $\mathcal{E}^2(\mathbb{R}^2)$ is defined by

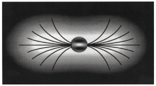
$$\mathcal{E}^2(\mathbb{R}^2) = \{f \in L^2(\mathbb{R}^2) : f = f_0 + f_1 \cdot \chi_B\},$$

where $B \subset [0,1]^2$ with ∂B a closed C^2-curve, $f_0, f_1 \in C_0^2([0,1]^2)$.

Anisotropic/Cartoon Structures

Images:

- Governing structure in images.
- Justified by neurophysiology.



Field et al., 1993

Definition (Donoho; 2001):

The set of cartoon-like functions $\mathcal{E}^2(\mathbb{R}^2)$ is defined by

$$\mathcal{E}^2(\mathbb{R}^2) = \{f \in L^2(\mathbb{R}^2) : f = f_0 + f_1 \cdot \chi_B\},$$

where $B \subset [0,1]^2$ with ∂B a closed C^2-curve, $f_0, f_1 \in C_0^2([0,1]^2)$.

Theorem (Donoho; 2001):

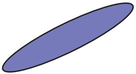
Let $(\psi_\lambda)_\lambda \subseteq L^2(\mathbb{R}^2)$ be a frame. Then the optimal asymptotic approximation error of $f \in \mathcal{E}^2(\mathbb{R}^2)$ is

$$\|f - f_N\|_2^2 \asymp N^{-2}, \quad N \to \infty, \quad \text{where } f_N = \sum_{\lambda \in \Lambda_N} c_\lambda \psi_\lambda.$$

Sparsifying Representation System

Parabolic scaling and shearing:

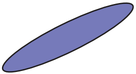
$$A_j = \begin{pmatrix} 2^j & 0 \\ 0 & 2^{j/2} \end{pmatrix} \quad \text{and} \quad S_k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}, \quad j,k \in \mathbb{Z}.$$



Sparsifying Representation System

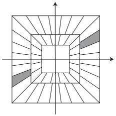
Parabolic scaling and shearing:

$$A_j = \begin{pmatrix} 2^j & 0 \\ 0 & 2^{j/2} \end{pmatrix} \quad \text{and} \quad S_k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}, \quad j, k \in \mathbb{Z}.$$



Definition (K, Labate; 2006):
The (cone-adapted) discrete shearlet system $\mathcal{SH}(\phi, \psi, \tilde{\psi})$ generated by $\phi \in L^2(\mathbb{R}^2)$ and $\psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ is the set

$$\{\phi(\cdot - m) : m \in \mathbb{Z}^2\},$$
$$\cup\{2^{3j/4}\psi(S_k A_j \cdot - m) : j \geq 0, |k| \leq \lceil 2^{j/2} \rceil, m \in \mathbb{Z}^2\},$$
$$\cup\{2^{3j/4}\tilde{\psi}(\tilde{S}_k \tilde{A}_j \cdot - m) : j \geq 0, |k| \leq \lceil 2^{j/2} \rceil, m \in \mathbb{Z}^2\}.$$

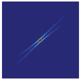


Compactly Supported Shearlets

Theorem (Kittipoom, K, Lim; 2012):

Let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be compactly supported, and let $\hat{\psi}$, $\hat{\tilde{\psi}}$ satisfy certain decay condition. Then $\mathcal{SH}(\phi, \psi, \tilde{\psi})$ forms a shearlet frame, i.e.,

$$A\|f\|_2^2 \leq \sum_{\sigma \in \mathcal{SH}(\phi,\psi,\tilde{\psi})} |\langle f, \sigma\rangle|^2 \leq B\|f\|_2^2 \quad \text{for all } f \in L^2(\mathbb{R}^2)$$



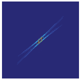
with controllable frame bounds.

$\rightsquigarrow$ Exemplary class with $B/A \approx 4$.

Compactly Supported Shearlets

Theorem (Kittipoom, K, Lim; 2012):
Let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be compactly supported, and let $\hat{\psi}$, $\hat{\tilde{\psi}}$ satisfy certain decay condition. Then $\mathcal{SH}(\phi, \psi, \tilde{\psi})$ forms a shearlet frame, i.e.,

$$A\|f\|_2^2 \le \sum_{\sigma \in \mathcal{SH}(\phi,\psi,\tilde{\psi})} |\langle f, \sigma \rangle|^2 \le B\|f\|_2^2 \quad \text{for all } f \in L^2(\mathbb{R}^2)$$

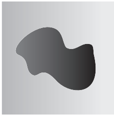


with controllable frame bounds.

$\rightsquigarrow$ Exemplary class with $B/A \approx 4$.

Theorem (K, Lim; 2011):
Let f be a cartoon-like function, let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be compactly supported, and let $\hat{\psi}$, $\hat{\tilde{\psi}}$ satisfy certain decay condition. Then $\mathcal{SH}(\phi, \psi, \tilde{\psi})$ provides an optimally sparse approximation of f, i.e.,

$$\|f - f_N\|_2^2 \lesssim N^{-2} \cdot (\log(N))^3, \quad N \to \infty.$$



Problem with Frames

False Statement for Shearlet Frame: We have

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta} \quad \text{and} \quad f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda,$$

and want to solve

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

Problem with Frames

False Statement for Shearlet Frame: We have

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta} \quad \text{and} \quad f = \sum_{\lambda \in \Lambda} \langle f, \psi_\lambda \rangle \, \psi_\lambda,$$

and want to solve

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

Problem with Frames

False Statement for Shearlet Frame: We have

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta} \quad \text{and} \quad f = \sum_{\lambda \in \Lambda} \langle f, \psi_\lambda \rangle \, \psi_\lambda,$$

and want to solve

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

Reason: For a frame $(\psi_\lambda)_{\lambda \in \Lambda}$, we have

$$f = \sum_{\lambda \in \Lambda} \langle f, \psi_\lambda \rangle \, \tilde{\psi}_\lambda \quad \text{for all } f \in L^2(\mathbb{R}^2),$$

where $\tilde{\psi}_\lambda = S^{-1}\psi_\lambda$ with $Sg = \sum_{\lambda \in \Lambda} \langle g, \psi_\lambda \rangle \, \psi_\lambda$.

Problem with Frames

False Statement for Shearlet Frame: We have

$$f \mapsto (\langle f, e_n \rangle)_{n\in\Delta} \quad \text{and} \quad f = \sum_{\lambda\in\Lambda} \langle f, \psi_\lambda \rangle \, \tilde{\psi}_\lambda,$$

and want to solve

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

Reason: For a frame $(\psi_\lambda)_{\lambda\in\Lambda}$, we have

$$f = \sum_{\lambda\in\Lambda} \langle f, \psi_\lambda \rangle \, \tilde{\psi}_\lambda \quad \text{for all } f \in L^2(\mathbb{R}^2),$$

where $\tilde{\psi}_\lambda = S^{-1}\psi_\lambda$ with $Sg = \sum_{\lambda\in\Lambda} \langle g, \psi_\lambda \rangle \, \psi_\lambda$.

Problem with Frames

False Statement for Shearlet Frame: We have

$$f \mapsto (\langle f, e_n \rangle)_{n \in \Delta} \quad \text{and} \quad f = \sum_{\lambda \in \Lambda} \langle f, \psi_\lambda \rangle \, \tilde{\psi}_\lambda,$$

and want to solve

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

Reason: For a frame $(\psi_\lambda)_{\lambda \in \Lambda}$, we have

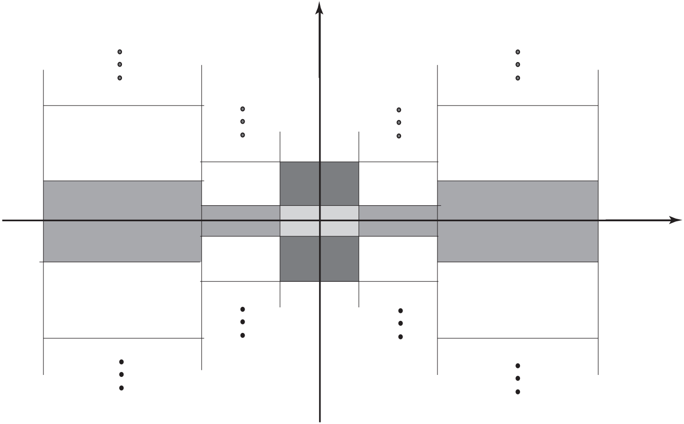
$$f = \sum_{\lambda \in \Lambda} \langle f, \psi_\lambda \rangle \, \tilde{\psi}_\lambda \quad \text{for all } f \in L^2(\mathbb{R}^2),$$

where $\tilde{\psi}_\lambda = S^{-1}\psi_\lambda$ with $Sg = \sum_{\lambda \in \Lambda} \langle g, \psi_\lambda \rangle \, \psi_\lambda$.

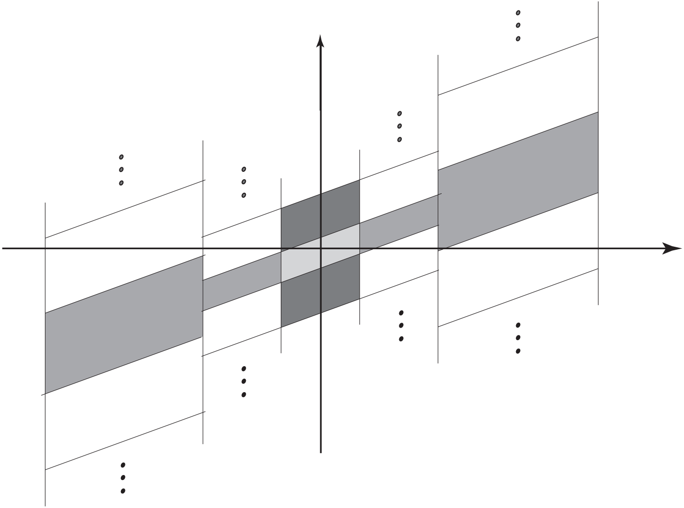
But: We do not know the dual shearlet frame $(\tilde{\psi}_\lambda)_{\lambda \in \Lambda}$ analytically!

Dualizable Shearlets...

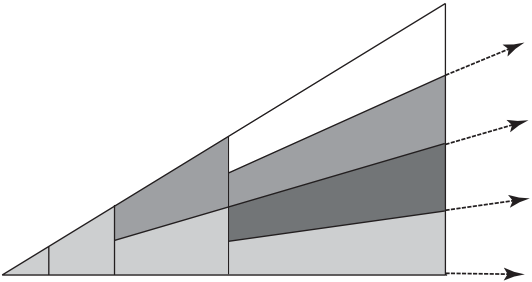
Intuition: Partition of Fourier Domain, shear$= 0$



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



Intuition: Filters

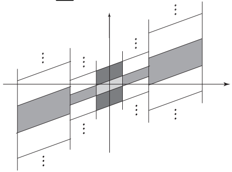


Shearlet Generators

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

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

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

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

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: $\rho > 0$.

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 ($k \leftrightarrow s$)

$$\langle f, \psi_{j,k,m} \rangle = \langle f, \Theta_s * \gamma_{j,m}^s \rangle = \langle \overline{\Theta}_s * f, \gamma_{j,m}^s \rangle = c_{j,m}^s.$$

Determining the measurement vector:

$$\overline{\Theta}_s * f = \sum_{(j,m)\in\Lambda_s} c_{j,m}^s \gamma_{j,m}^s \quad \Longrightarrow \quad \langle \overline{\Theta}_s * f, e_n \rangle = \sum_{(j,m)\in\Lambda_s} \langle \gamma_{j,m}^s, e_n \rangle c_{j,m}^s$$

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:

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

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

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 $y_n = \widehat{\overline{\Theta}_s}(n) \cdot \hat{f}(n)$.

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:

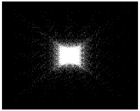
$$\langle P_J^s(\overline{\Theta}_s * f), e_n \rangle = \sum_{(j,m)\in\Lambda_{J,s}} \langle \gamma_{j,m}^s, e_n \rangle c_{j,m}^s.$$

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



$$p_s(n) = \frac{c_s}{J^2(1+|n_1|)(1+|n_2 - sn_1|)}.$$

Sparse Sampling Strategy

Theorem (K, Lim; 2014):

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).

Sparse Sampling Strategy

Theorem (K, Lim; 2014):

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

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

Sparse Sampling Strategy

Theorem (K, Lim; 2014):
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 $s \in \mathbb{S}_{J/2}$,

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

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

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

Sparse Sampling Strategy

Theorem (K, Lim; 2014):

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

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

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

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

Sparse Sampling Strategy

Theorem (K, Lim; 2014):

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

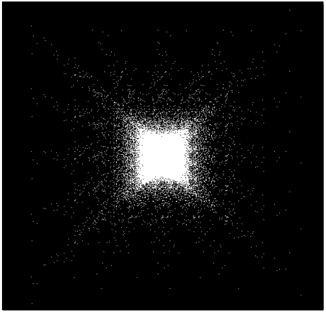
$$(\hat{c}_\lambda)_{\lambda\in\Lambda_{J,s}} = \text{argmin}_c \|c\|_1 \text{ subject to } \Phi_s c = y_s,$$

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

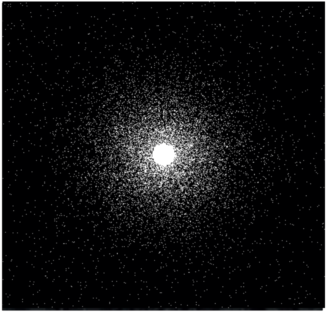
$$\left\| f - \sum_{s\in\mathbb{S}_{J/2}} \sum_{\lambda\in\Lambda_{J,s}} \hat{c}_\lambda \tilde{\psi}_\lambda \right\|_2^2 \lesssim 2^{-J(1-13\rho/2)} (= O(N^{-2+C\rho})) \quad \text{as } J \to \infty.$$

Numerical Experiments

Sampling Schemes

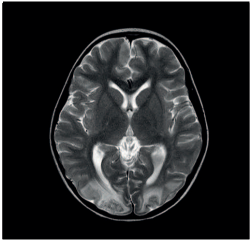


Directional Sampling Scheme

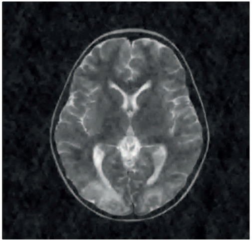


Variable Density Sampling Scheme

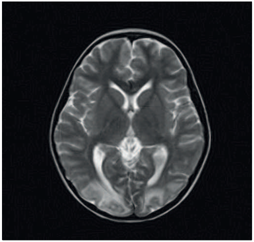
Numerical Results for 512x512 MRI Image



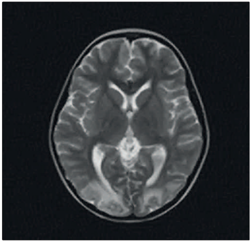
Original



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

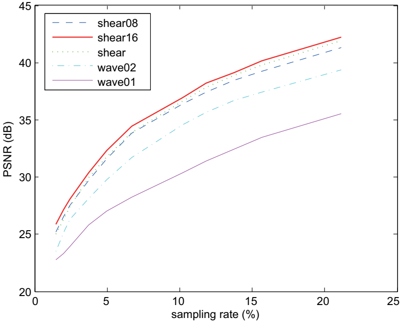


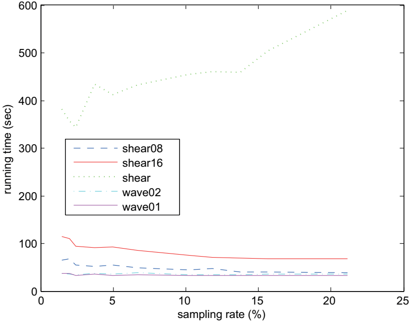
Shearlet Scheme
(5% sampling rate, 32.2845dB)



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

Approximation Curves for 512x512 MRI Image



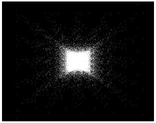


- **shear08**: Directional sampling scheme with 8 directional filters.
- **shear16**: Directional sampling scheme with 16 directional filters.
- **shear**: Directional sampling scheme with (normal) shearlets.
- **wave02**: Directional sampling scheme with wavelets.
- **wave01**: Variable density sampling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

Let's conclude...

What to take Home...?

- Sampling and reconstruction strategies are key for acquiring data in a continuous world.
- Applications such as MRI only allow Fourier samples.
- Sampling and Reconstruction Scheme for Fourier Data:
  - Cartoon-like functions as continuum model.
  - (Dualizable) Shearlets as sparsifying system.
  - Directional sampling scheme.
- Asymptotically optimal recovery could be proven.
- Numerical evidence of superiority of the scheme.



THANK YOU!

References available at:

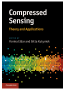
`www.math.tu-berlin.de/~kutyniok`

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.