

# Solving inverse problems in imaging with Shearlab.jl

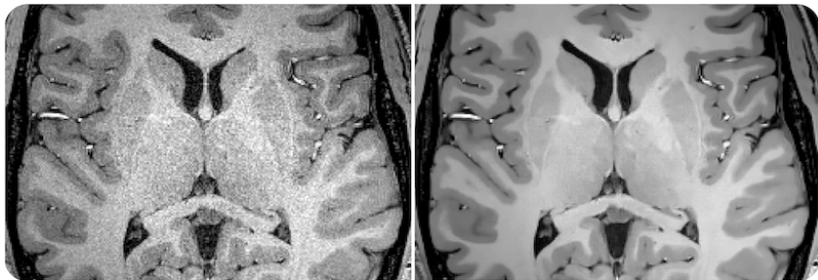
Héctor Andrade Loarca  
(Technische Universität Berlin)

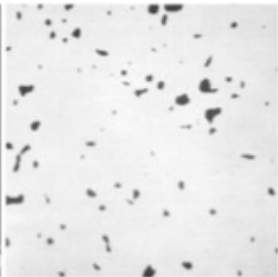
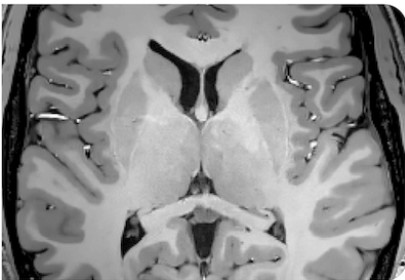
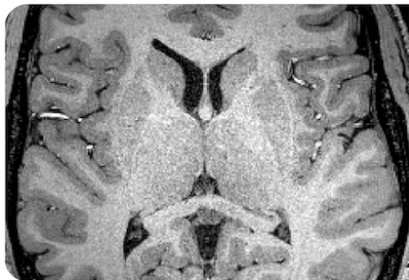
**SIAM IS 2018**

TU Berlin

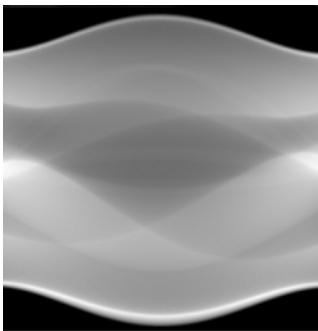
8<sup>th</sup> of June, 2018











# Inverse problems in Imaging

## Goal

Recover parameters characterizing a system under investigation from measurements (e.g. recover image from data).

# Inverse problems in Imaging

## Goal

Recover parameters characterizing a system under investigation from measurements (e.g. recover image from data).

## Mathematical formulation

Recover  $f_{\text{true}} \in X$  from data

$$g = \mathcal{T}(f_{\text{true}}) + \delta g$$

where  $g \in Y, \mathcal{T} : X \longrightarrow Y$  and  $\delta g \in Y$ .

# Inverse problems in Imaging

## Goal

Recover parameters characterizing a system under investigation from measurements (e.g. recover image from data).

## Mathematical formulation

Recover  $f_{\text{true}} \in X$  from data

$$g = \mathcal{T}(f_{\text{true}}) + \delta g$$

where  $g \in Y, \mathcal{T} : X \longrightarrow Y$  and  $\delta g \in Y$ .

- **Classical solution:** Minimization of the miss-fit against data:

$$\min_{f \in X} \mathcal{L}(\mathcal{T}(f), g)$$

$\mathcal{L} : Y \times Y \longrightarrow \mathbb{R}$  is a transformation of the negative data log-likelihood  $(-\log P(f|g))$ , e.g.  $\mathcal{L}(f) = \|\mathcal{T}(f) - g\|_2^2$ .



# Ill-posedness and regularization

## Hadamard well-posedness

Existence and uniqueness of solution for all data and continuous dependence of solution on the data.

Ill-posed problems tend to produce overfitting when minimizing the data miss-fit, but they are the most common in applications (CT, EEG, MRI,...).

# Ill-posedness and regularization

## Hadamard well-posedness

Existence and uniqueness of solution for all data and continuous dependence of solution on the data.

Ill-posed problems tend to produce overfitting when minimizing the data miss-fit, but they are the most common in applications (CT, EEG, MRI,...).

- ▶ **Regularization:** Set of methods to avoid overfitting by slightly modify the original problem to increase its regularity.

# Ill-posedness and regularization

## Hadamard well-posedness

Existence and uniqueness of solution for all data and continuous dependence of solution on the data.

Ill-posed problems tend to produce overfitting when minimizing the data miss-fit, but they are the most common in applications (CT, EEG, MRI,...).

- ▶ **Regularization:** Set of methods to avoid overfitting by slightly modify the original problem to increase its regularity.
- ▶ **Variational regularization:** Uses a functional  $\mathcal{S} : X \longrightarrow \mathbb{R}$  (e.g.  $\|\cdot\|_1$ ) to encode a priori information about  $f_{\text{true}}$ , obtaining:

$$\min_{f \in X} [\mathcal{L}(\mathcal{T}(f), g) + \lambda \mathcal{S}(f)] \quad \text{for a fixed } \lambda \geq 0$$

# Image denoising

## Goal

Recover an image  $f \in X$  from noisy data:

$$g = f + \delta g$$

where  $\delta g$  is Gaussian white noise, with sd.  $\sigma$ .

# Image denoising

## Goal

Recover an image  $f \in X$  from noisy data:

$$g = f + \delta g$$

where  $\delta g$  is Gaussian white noise, with sd.  $\sigma$ .

- ▶ The risk of an estimator  $\tilde{f}$  is given by the MSE:

$$\mathbb{E} \|f - \tilde{f}\|_2^2$$

$\mathbb{E}$  is the expectation with respect of the probability distribution of  $\delta g$ .

## Goal

Recover an image  $f \in X$  from noisy data:

$$g = f + \delta g$$

where  $\delta g$  is Gaussian white noise, with sd.  $\sigma$ .

- ▶ The risk of an estimator  $\tilde{f}$  is given by the MSE:

$$\mathbb{E} \|f - \tilde{f}\|_2^2$$

$\mathbb{E}$  is the expectation with respect of the probability distribution of  $\delta g$ .

- ▶ The worst behaviour of the estimator is the supremum

$$\sup_{f \in X} \mathbb{E} \|f - \tilde{f}\|_2^2$$

the *Minimax* MSE will be

$$\inf_{\tilde{f}} \sup_{f \in X} \mathbb{E} \|f - \tilde{f}\|_2^2$$

## Frame

A frame for a Hilbert space  $X$  is a collection  $\Psi = \{\psi_i\}_{i \in \mathcal{I}} \subset X$  satisfying

$$A\|f\|_2 \leq \|\{\langle f, \psi_i \rangle\}_{i \in \mathcal{I}}\|_{\ell^2(\mathcal{I})} \leq B\|f\|_2 \quad \forall f \in X$$

for some  $0 < A \leq B < \infty$ .

## Frame

A frame for a Hilbert space  $X$  is a collection  $\Psi = \{\psi_i\}_{i \in \mathcal{I}} \subset X$  satisfying

$$A\|f\|_2 \leq \|\{\langle f, \psi_i \rangle\}_{i \in \mathcal{I}}\|_{\ell^2(\mathcal{I})} \leq B\|f\|_2 \quad \forall f \in X$$

for some  $0 < A \leq B < \infty$ .

## Theorem ((Labate et al., 2012))

*"If an image is sparse within a frame  $\{\psi_i\}_{i \in \mathcal{I}}$ , one can obtain a Minimax MSE estimator by thresholding the coefficients in the expansion of the noisy data:*

$$g = \sum_{i \in \mathcal{I}} \langle g, \psi_i \rangle \psi_i \quad "$$



## Goal

Recover an image  $f \in X$  from known data:

$$g = P_K(f)$$

where  $P_K$  is an orthogonal projection onto the known subspace  $X_K \triangleleft X$ .

# Image inpainting

## Goal

Recover an image  $f \in X$  from known data:

$$g = P_K(f)$$

where  $P_K$  is an orthogonal projection onto the known subspace  $X_K \triangleleft X$ .

## Sparse Regularization/CS approach (Genzel, Kutyniok):

" If a signal (image) is sparse within a frame  $\Psi$ , it can be recovered from highly underdetermined, non-adaptive linear measurements by  $\ell^1$ -regularization (Davenport et al., 2012), i.e.

$$\min_{\tilde{f} \in X} \|\{\langle \tilde{f}, \psi_i \rangle\}_{i \in \mathcal{I}}\|_{\ell^1(\mathcal{I})} \quad \text{s.t. } P_K(\tilde{f}) = g = P_K(f) \quad "$$

## Theorem (Genzel, Kutyniok; 2014)

Let  $\delta > 0$  and  $\Lambda \subset \mathcal{I}$  be a  $\delta$ -**cluster** for  $f$  with respect to a frame  $\Psi$  (i.e.  $\|\mathbb{1}_{\Lambda^c} T_\Psi f\|_{\ell^1} \leq \delta$ ). If  $\mu_c(\Lambda, P_M \Psi) < 1/2$  and  $f^*$  is the minimizer of the problem, then

$$\|\{\langle f^* - f, \psi_i \rangle\}_{i \in \mathcal{I}}\|_{\ell^1(\mathcal{I})} \leq \frac{2\delta}{1 - \mu_c(\Lambda, P_M \Psi)}$$

## Cluster coherence

$$\mu_c(\Lambda, P_M \Psi) := \max_{j \in \mathcal{I}} \sum_{i \in \Lambda} |\langle P_M \psi_i, P_M \psi_j \rangle|$$

## Theorem (Genzel, Kutyniok; 2014)

Let  $\delta > 0$  and  $\Lambda \subset \mathcal{I}$  be a  $\delta$ -**cluster** for  $f$  with respect to a frame  $\Psi$  (i.e.  $\|\mathbb{1}_{\Lambda^c} T_{\Psi} f\|_{\ell^1} \leq \delta$ ). If  $\mu_c(\Lambda, P_M \Psi) < 1/2$  and  $f^*$  is the minimizer of the problem, then

$$\|\{\langle f^* - f, \psi_i \rangle\}_{i \in \mathcal{I}}\|_{\ell^1(\mathcal{I})} \leq \frac{2\delta}{1 - \mu_c(\Lambda, P_M \Psi)}$$

## Cluster coherence

$$\mu_c(\Lambda, P_M \Psi) := \max_{j \in \mathcal{I}} \sum_{i \in \Lambda} |\langle P_M \psi_i, P_M \psi_j \rangle|$$

- ▶ **Conclusion:** One can use sparsifying frames on images to perform denoising and inpainting. The quality depends on the level of sparsity.
- ▶ **Problem:** Pick a good frame for the image space.

# Image space: Cartoon-like functions

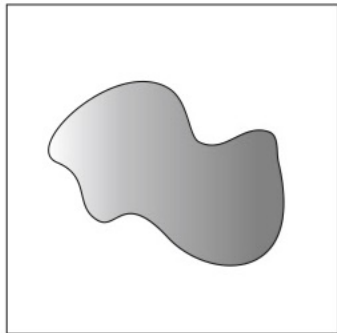
## Definition

Let  $f : \mathbb{R}^2 \rightarrow \mathbb{C}$ ,  $f \in \mathcal{E}^2(\mathbb{R}^2)$  if  $f = f_0 + \chi_B f_1$ , with  $B \subset [0, 1]^2$ ,  $\partial B \in C^2$  and with bounded curvature. Moreover,  $f_i \in C^2(\mathbb{R}^2)$  with  $\|f_i\|_{C^2} \leq 1$  and  $\text{supp} f_i \subset [0, 1]^2$  for  $i = 0, 1$ .

# Image space: Cartoon-like functions

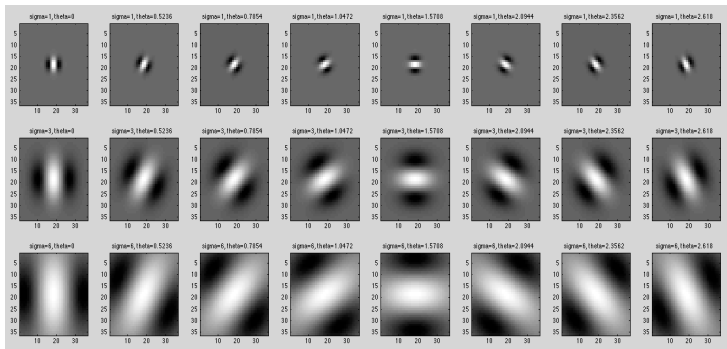
## Definition

Let  $f : \mathbb{R}^2 \rightarrow \mathbb{C}$ ,  $f \in \mathcal{E}^2(\mathbb{R}^2)$  if  $f = f_0 + \chi_B f_1$ , with  $B \subset [0, 1]^2$ ,  $\partial B \in C^2$  and with bounded curvature. Moreover,  $f_i \in C^2(\mathbb{R}^2)$  with  $\|f_i\|_{C^2} \leq 1$  and  $\text{supp} f_i \subset [0, 1]^2$  for  $i = 0, 1$ .



# Examples of frames for images

- ▶ Gabor frames (Gabor, 1946).
- ▶ Wavelet frames (Morlet et al., 1984).
- ▶ Curvelet frames (Candès et al., 1999).
- ▶ Shearlet frames (Kutyniok et al., 2005).



# Optimal approximation error for images

## Best N-term approx. error (Donoho, 2001)

Let  $\{\psi_\lambda\}_{\lambda \in \Lambda} \subset L^2(\mathbb{R}^2)$  a frame. The optimal best N-Term approximation error for any  $f \in \mathcal{E}^2(\mathbb{R}^2)$  is

$$\sigma_N(f, \{\psi_\lambda\}_{\lambda \in \Lambda}) = O(N^{-1})$$



# Optimal approximation error for images

## Best N-term approx. error (Donoho, 2001)

Let  $\{\psi_\lambda\}_{\lambda \in \Lambda} \subset L^2(\mathbb{R}^2)$  a frame. The optimal best N-Term approximation error for any  $f \in \mathcal{E}^2(\mathbb{R}^2)$  is

$$\sigma_N(f, \{\psi_\lambda\}_{\lambda \in \Lambda}) = O(N^{-1})$$

## Error of 2D-wavelets

$$\sigma_N(f, \{\psi_\lambda\}_\Lambda) \sim N^{-1/2}$$

# Optimal approximation error for images

## Best N-term approx. error (Donoho, 2001)

Let  $\{\psi_\lambda\}_{\lambda \in \Lambda} \subset L^2(\mathbb{R}^2)$  a frame. The optimal best N-Term approximation error for any  $f \in \mathcal{E}^2(\mathbb{R}^2)$  is

$$\sigma_N(f, \{\psi_\lambda\}_{\lambda \in \Lambda}) = O(N^{-1})$$

## Error of 2D-wavelets

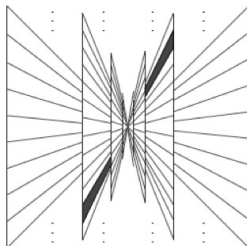
$$\sigma_N(f, \{\psi_\lambda\}_\Lambda) \sim N^{-1/2}$$

## Classical Shearlet Transform

$$\langle f, \psi_{j,k,m} \rangle = \int_{\mathbb{R}^2} f(x) \overline{\psi_{j,k,m}(x)} dx$$

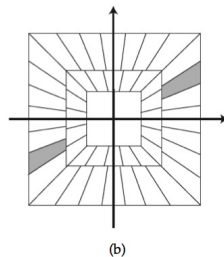
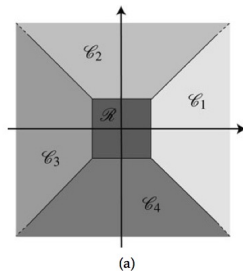
where

$$\mathcal{SH}(\psi) = \{ \psi_{j,k,m}(x) = 2^{3j/4} \psi(S_k A_j x - m) : (j, k) \in \mathbb{Z}^2, m \in \mathbb{Z}^2 \}$$



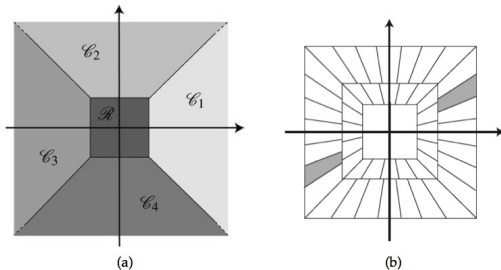
# Cone-adapted shearlet transform and optimal sparsity

$$\mathcal{SH}(\phi, \psi, \tilde{\psi}, c) := \mathcal{P}_{\mathcal{R}}\Phi(\phi, c1) \cup \mathcal{P}_{\mathcal{C}_1}\Psi(\psi, c) \cup \mathcal{P}_{\mathcal{C}_2}\tilde{\Psi}(\tilde{\psi}, c)$$



# Cone-adapted shearlet transform and optimal sparsity

$$\mathcal{SH}(\phi, \psi, \tilde{\psi}, c) := \mathcal{P}_{\mathcal{R}}\Phi(\phi, c1) \cup \mathcal{P}_{\mathcal{C}_1}\Psi(\psi, c) \cup \mathcal{P}_{\mathcal{C}_2}\tilde{\Psi}(\tilde{\psi}, c)$$



**Cone shearlets sparsity** (Band limited case: Lim, Labate; 2006),  
(Compactly supported case: Kutyniok, Lim, 2011)

Best  $N$ -term approximation error

$$\sigma_N(f, \{\psi_{j,k,m}\}_{j,k,m}) \sim N^{-1}(\log(N))^{3/2}$$

## ▶ Matlab

- ▶ FFST- Fast Finite Shearlet Transform (Häuser, Steidl, TU Kaiserslautern)  
<http://www.mathematik.uni-kl.de/imagepro/software/ffst/>
- ▶ 2D/3D Shearlet Toolbox (D. Labate, University of Houston)  
<https://www.math.uh.edu/~dlabate/software.html>
- ▶ **Shearlab3D** (G. Kutyniok, W.-Q.Lim, R. Reisenhofer, TU Berlin)  
<http://www.shearlab.org/>

## ▶ Python

- ▶ pyShearLab (Stefan Loock, U Göttingen)  
<http://na.math.uni-goettingen.de/pyshearlab/>

## ▶ Julia

- ▶ **Shearlab.jl** (H. Andrade, TU Berlin)  
<https://github.com/arsenal9971/Shearlab.jl>

# Current software

## ▶ Matlab

- ▶ FFST- Fast Finite Shearlet Transform (Häuser, Steidl, TU Kaiserslautern)  
<http://www.mathematik.uni-kl.de/imagepro/software/ffst/>
- ▶ 2D/3D Shearlet Toolbox (D. Labate, University of Houston)  
<https://www.math.uh.edu/~dlabate/software.html>
- ▶ **Shearlab3D** (G. Kutyniok, W.-Q.Lim, R. Reisenhofer, TU Berlin)  
<http://www.shearlab.org/>

## ▶ Python

- ▶ pyShearLab (Stefan Loock, U Göttingen)  
<http://na.math.uni-goettingen.de/pyshearlab/>

## ▶ Julia

- ▶ **Shearlab.jl** (H. Andrade, TU Berlin)  
<https://github.com/arsenal9971/Shearlab.jl>

## ▶ Lets code!

# Thanks!

## Questions?

