# Radon Transform Inversion using the Shearlet Representation

*Demetrio Labate*

*University of Houston*

*Joint work with F. Colonna, G. Easley, K. Guo*

# Introduction

The *Radon transform* of a function $f(x)$, $x \in \mathbb{R}^2$, is defined as

$$(Rf)(\theta, t) = \int_{L(\theta,t)} f(x)\, dx,$$

where $L(\theta, t)$ is the straight line $x \cdot \theta = t$, for $\theta \in S^1$, $t \in \mathbb{R}$.

The Radon transform arises in situations where one wants to determine structural properties of an object by methods that utilize projected information, as occurring, for example, in $X$-Ray Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Synthetic Aperture Radars (SAR) and Seismic Tomography.

# Introduction

The problem of interest consists in *recovering the object $f$ from the Radon data*

$$y(\theta, t) = R\, f(\theta, t)$$

or (in the presence of noise)

$$y(\theta, t) = R\, f(\theta, t) + z(\theta, t).$$

In principle, the mathematical problem was solved by Radon [1917], who obtained the inversion formula:

$$f(x) = \frac{1}{4\pi^2} \int_{S^1} \int_{\mathbb{R}} \frac{\frac{d}{dt} R(\theta, t)}{x \cdot \theta - t}\, dt\, d\theta.$$

# Introduction

Problem solved?

# Introduction

Problem solved? Not really.

# Introduction

Problem solved? Not really.

• How to convert the inversion formula of the Radon transform into an accurate computational algorithm is far from obvious.

• In particular, in most cases it is not possible to measure all the data $R(\theta, t)$. In this situation, serious issues of uniqueness and stability of solutions arise. This is further complicated if the data are corrupted by noise.

# Outline

1. Ill-posed Problems
2. Shearlet Representation
3. Shearlet-based Inversion of Radon Data

# Outline

1. $\boxed{\text{Ill-posed Problems}}$
2. Shearlet Representation
3. Shearlet-based Inversion of Radon Data

# Ill-posed Problems

For problems in mathematical physics, Hadamard postulated three properties which he deemed should be satisfied by all "reasonable" problems:

1. Existence of a solution.
2. Uniqueness of a solution.
3. Continuous dependence of a solution on the data.

A problem satisfying all these three requirements is *well-posed*.

Unfortunately, many important problems are not well-posed.

# Ill-posed Problems

Consider the problem of recovering $f \in \mathcal{H}$ from the linear equation:

$$A\, f = g.$$

This inverse problem might fail to be well-posed because of the following:

1. It is not solvable since $g$ is not in the range of $A$.
2. It is not uniquely solvable since $A$ is not injective.
3. The solution does not depend continuously on the data since $A^{-1}$ is not continuous.

# Ill-posed Problems

Issues 1. and 2. can be dealt with by using the *generalized inverse*:

$$A^+ = (A^*A)^{-1}A^*.$$

so that $f^+ = A^+ g$.

Using the *singular value decomposition*:

$$A = \sum_k \sigma_k \langle f, f_k \rangle g_k$$

where $(f_k)$, $(g_k)$ are appropriate orthonormal systems and $(\sigma_k)$ are the singular values of $A$, the generalized inverse of $A$ is

$$A^+ g = \sum_k \sigma_k^{-1} \langle g, g_k \rangle f_k.$$

However, $A^+$ is not continuous in general (singular values $\to 0$).

In this case a *regularization* must be introduced.

# Ill-posed Problems

Example of regularization: *Truncated SVD*

Consider the SVD of $A$

$$A = \sum_k \sigma_k \, \langle f, f_k \rangle g_k.$$

Recall that, formally,

$$A^+ g = \sum_k \sigma_k^{-1} \, \langle g, g_k \rangle f_k.$$

Then a regularized inverse is obtained as

$$T_\gamma \, g = \sum_{\sigma_k \geq \gamma} \sigma_k^{-1} \, \langle g, g_k \rangle f_k.$$

# Ill-posed Problems

Example of regularization: *Tikhonov regularization*

The regularized inverse has the form:

$$T_\gamma\, g = \sum_k w_\gamma(\sigma_k)\, \sigma_k^{-1}\, \langle g, g_k \rangle f_k,$$

where $w_\gamma(\sigma_k) = \frac{1}{1+\gamma/\sigma_k^2}$.

This is equivalent to considering the regularized inverse operators

$$T_\gamma = \left(A^* A + \gamma I\right)^{-1} A^*.$$

# Radon Transform

The singular value decomposition of the Radon transform $R$ is of the form

$$Rf = \sum_{m=0}^{\infty} \sigma_m \langle f, h_m \rangle \, g_m,$$

where

$$\sigma_m^2 = \frac{2\pi}{m+1}.$$

Since the singular value decay slowly to 0, the ill-posedness of the Radon problem is not very pronounced.

If only incomplete data $(Rf)(\theta, t)$ are known, the problem may become severely ill-posed.

# Radon Transform

Traditionally, the function $f$ is estimated by regularizing the Radon inverse as

$$\tilde{f} = \sum_{m=0}^{\infty} w_m \, \sigma_m^{-1} \, [Rf, f_m] \, h_m,$$

where $[\cdot, \cdot]$ is the inner product in the space $\mathcal{D}$ of $L^2$ functions on $S^1 \times \mathbb{R}$ with antipodal symmetry.

The window coefficients $w_m$ "regularize" the effect of small singular values $\sigma_m$.

Drawback: *As a result of the regularization, the high frequency information in $f$ is damped*, as evident in typical regularized reconstructions which are blurred versions of the original.

# Radon Transform

• The limitation of SVD is due to the fact that the basis functions employed in the decomposition *derive entirely from the operator and not from the object* to be recovered. As a result, these basis functions cannot approximate all kind of objects effectively.

# Radon Transform

• The limitation of SVD is due to the fact that the basis functions employed in the decomposition *derive entirely from the operator and not from the object* to be recovered. As a result, these basis functions cannot approximate all kind of objects effectively.

• The approach that we propose adapts the paradigm of the *Wavelet-Vaguelette Decomposition*, introduced by Donoho [1995].
In this approach, the singular-value system is replaced by a representation system which does not depend solely on the operator, but is chosen to best capture the features of the objects to be recovered.

# Outline

1. Ill-posed Problems
2. $\boxed{\text{Shearlet Representation}}$
3. Shearlet-based Inversion of Radon Data

# Directional Multiscale Representations

Traditional multiscale systems (i.e., wavelets) are optimal in dealing point-discontinuities only. They have a limited capability to deal singularities in higher dimensions such as edges and moving fronts.

The *shearlet representation* belongs to a class of representations which have been introduced to provide *sparse* representations for multidimensional functions with distributed discontinuities.

Another well known example of such representation is the *curvelet representation* of Candès and Donoho.

# Shearlet Representation [GLLWW04,KL05]

The shearlet representation provides a directional multiscale decomposition of functions within the framework of affine systems.

A **Shearlet System**, for $n = 2$, is an affine-like system of the form

$$\{\psi_{j,\ell,k} = |\det A|^{j/2}\,\psi(B^\ell\,A^j x - k) :\ j,\ell \in \mathbb{Z}, k \in \mathbb{Z}^2\},$$

where

$$B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} 2 & 0 \\ 0 & \sqrt{2} \end{pmatrix},$$

and $\psi$ is a well-localized function.

$A$ is the *anisotropic dilation matrix* and $B$ is the *shear matrix*.

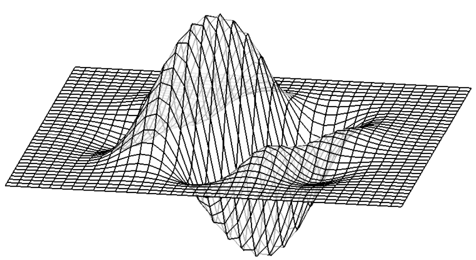
# Shearlet Representation

The generating function of the shearlet system $\{\psi_{j,\ell,k}\}$ is chosen to be a band-limited, well-localized function, of the form

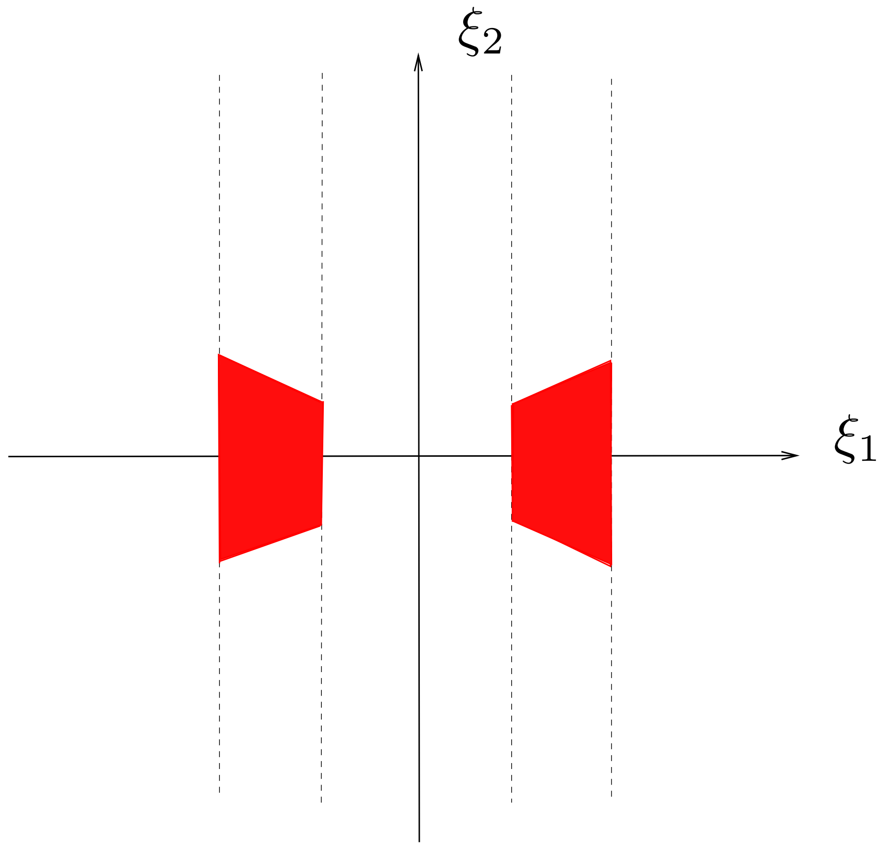
$$\hat{\psi}(\omega) = \hat{\psi}_1(\omega_1)\,\hat{\psi}_2\Big(\frac{\omega_2}{\omega_1}\Big).$$

with : $\operatorname{supp}\hat{\psi}_1 \subset [-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2]$ and $\operatorname{supp}\hat{\psi}_2 \subset [-1, 1]$.

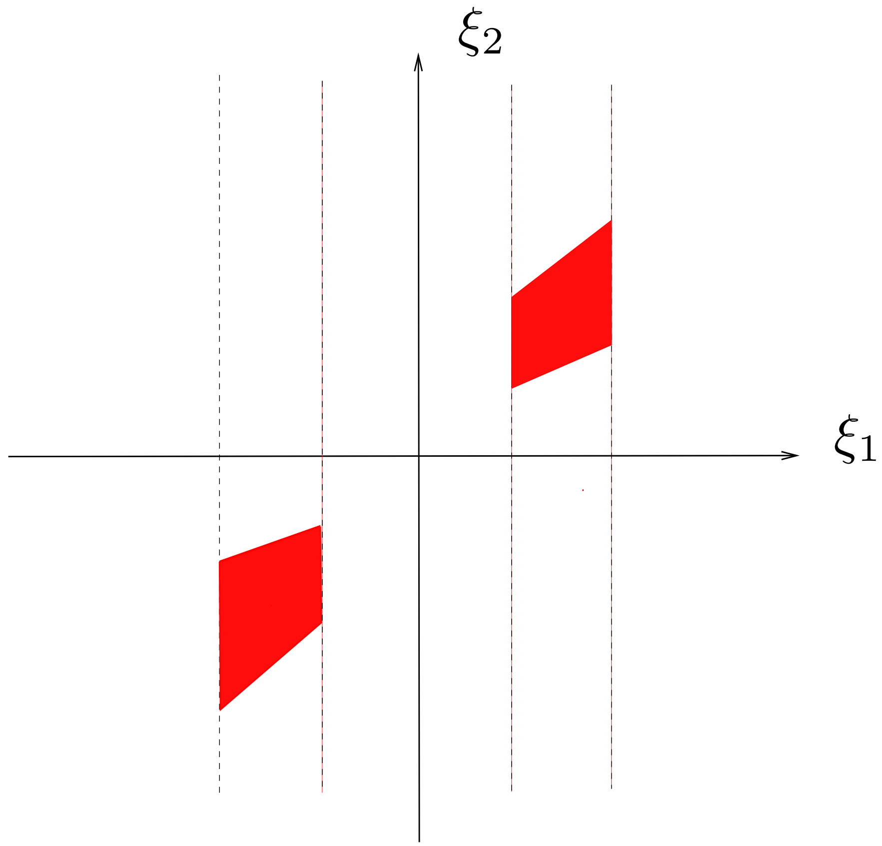
Each element $\psi_{j,\ell,k}$ is associated with the *scale* $2^{-j}$, the *location* $2^{-j}\,k$ and the *orientation* $2^{-j}\,\ell$.



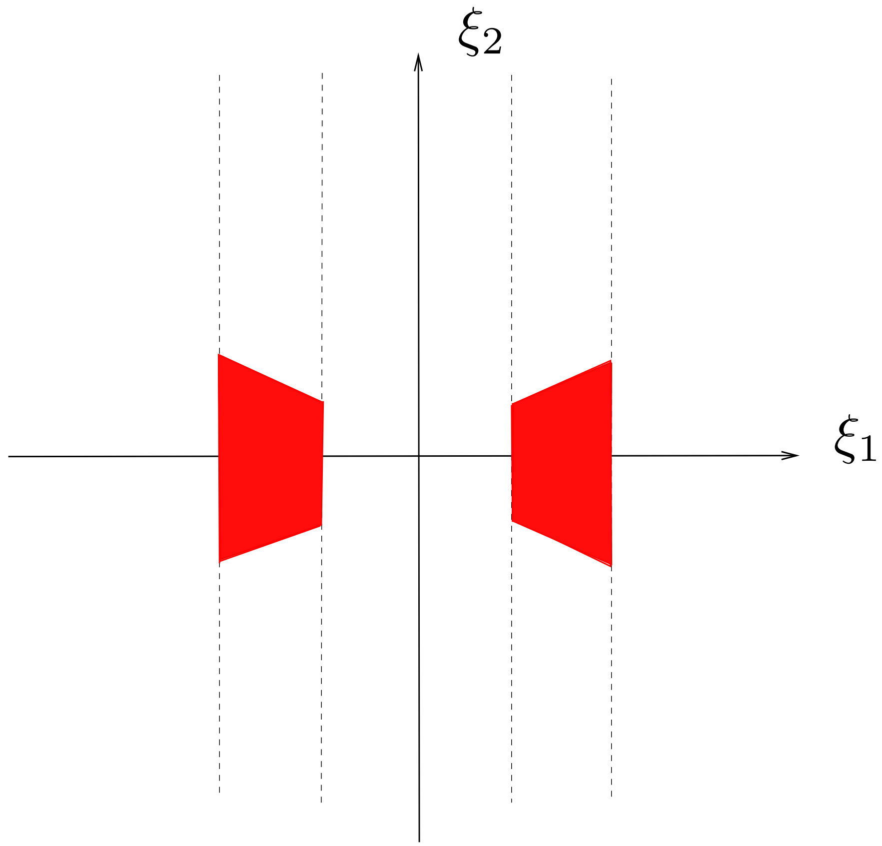
Frequency support of the mother shearlet $\psi$.



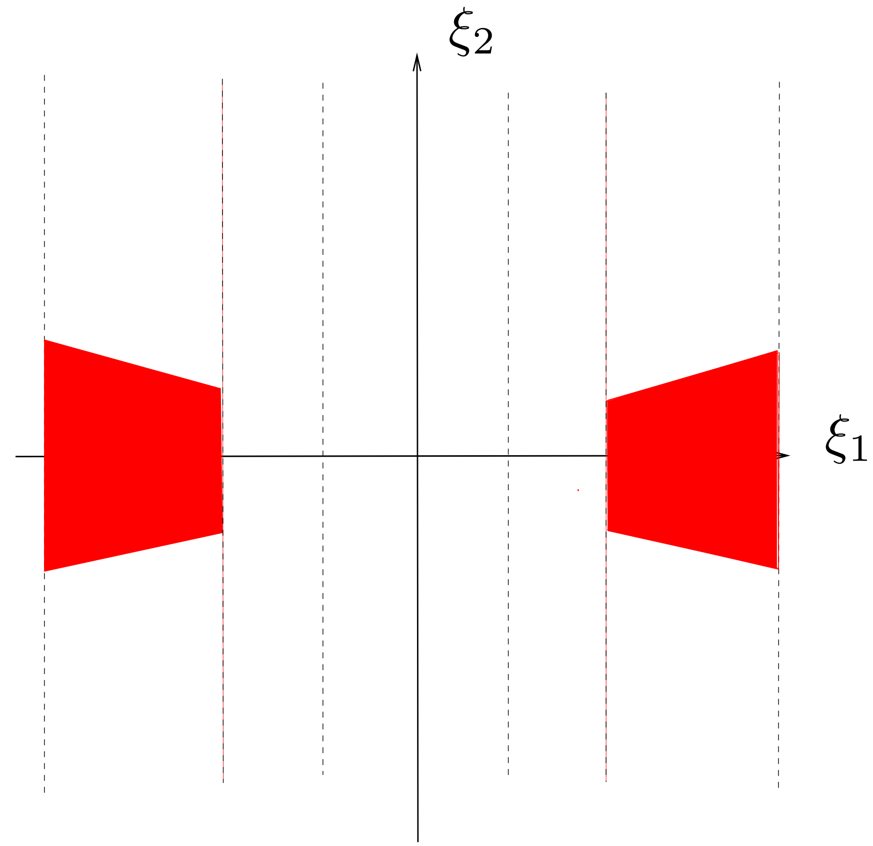
Action of the shear matrix $B$ on the shearlet $\psi$.



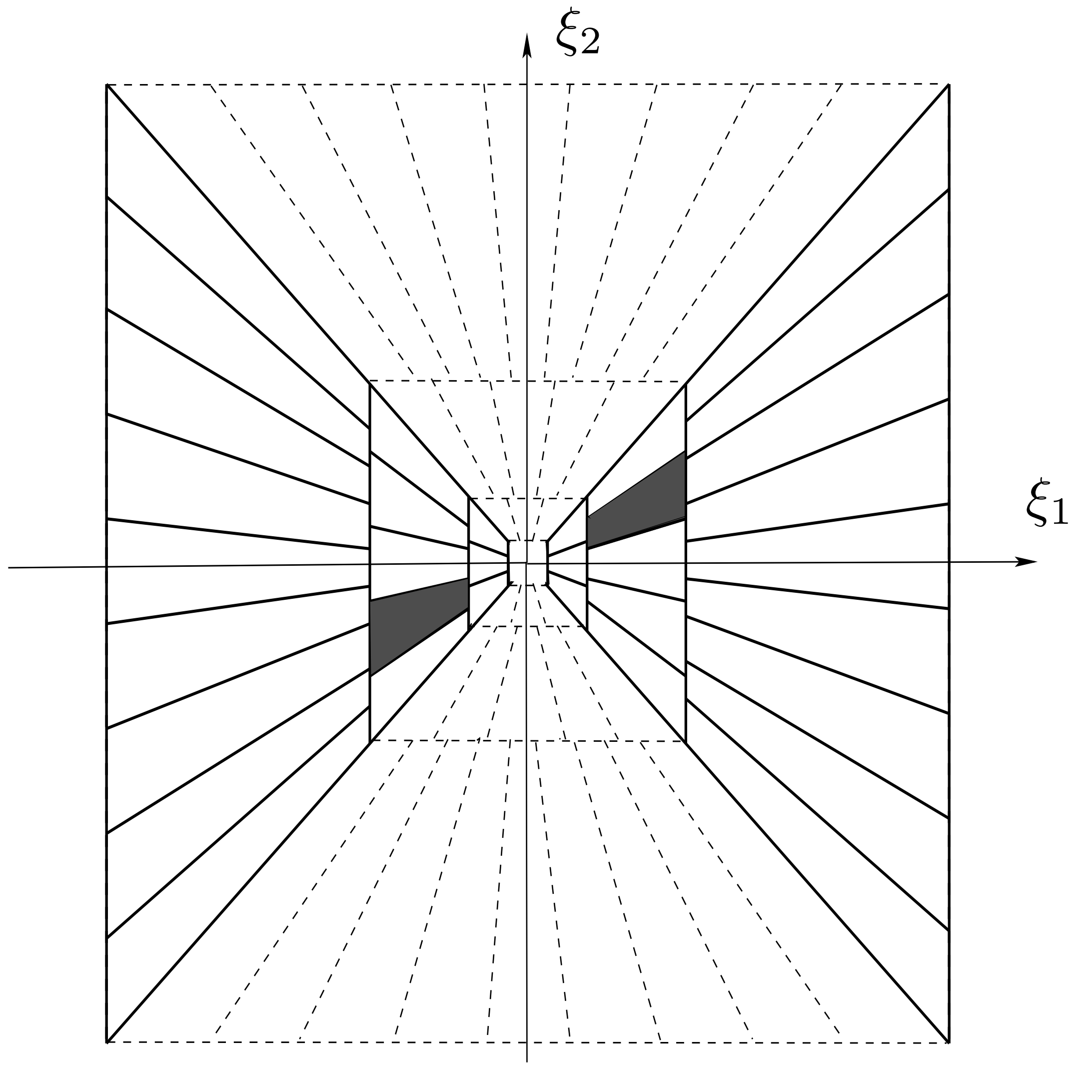
Frequency support of the mother shearlet $\psi$.



Action of the dilation matrix $A$ on the shearlet $\psi$.



Tiling of the frequency plane induced by the shearlets.



# Shearlet Representation

The shearlets tile the frequency plane and form a **Parseval frame** for $L^2(\mathbb{R}^2)$, that is

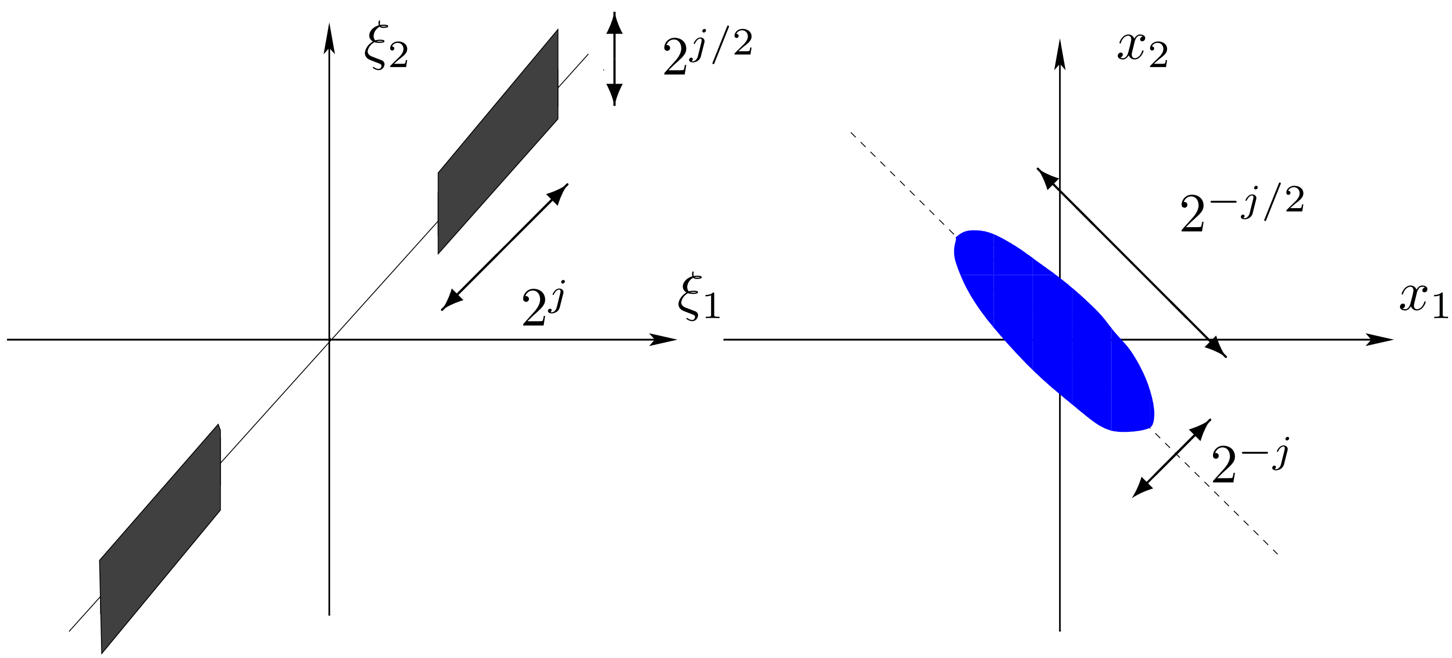
$$f = \sum_{j,\ell,k} \langle f, \psi_{j,\ell,k} \rangle \, \psi_{j,\ell,k}, \quad \forall f \in L^2(\mathbb{R}^2).$$

Notice that a Parseval frame need not be an orthonormal basis (it is not in this case). However the Plancherel formula holds, so that we still have a representation formula for any $f \in L^2(\mathbb{R}^2)$.

# Shearlet Representation

The shearlets $\{\psi_{j,k,\ell}\}$ are a Parseval frame of *well-localized wave-forms* at various *scales, locations and orientations*

The support is highly anisotropic and become increasingly elongated at finer scales.



# Shearlet Representation

Thanks to their localization and directional properties, shearlets are very efficient in dealing with functions containing discontinuities along curves.

**Theorem** Let $f$ be a two-dimensional $C^2$ function apart from discontinuities along $C^2$ curves. Let $\tilde{f}_N$ be the approximation of $f$ obtained by taking the $N$ largest coefficients in the shearlets expansion. Then the asymptotic approximation error obeys:

$$\|f - \tilde{f}_N\|^2 \leq C\,(\log N)^3\,N^{-2}, \quad N \to \infty.$$

This is essentially **optimal**.

Same approximation is obtained using the curvelets [Candès, Donoho], which have a very different mathematical construction.

# Outline

1. Ill-posed Problems
2. Shearlet Representation
3. Shearlet-based Inversion of Radon Data

# Shearlet-based Decomposition of $Rf$

The shearlet representation offers a combination of 2 useful features:

1. It nearly diagonalizes the operator $R^* R$.
2. It provides optimally sparse representation for $2D$-functions which are smooth away from discontinuities along curves.

Similar properties hold for more general operators and higher dimensions.

# Shearlet-based Decomposition of $Rf$

$R^* R$ is an elliptic pseudodifferential operator and

$$R^* R f(x) = \int_{\mathbb{R}^2} \frac{1}{|\xi|} e^{2\pi i x \xi} \hat{f}(\xi) \, d\xi$$

Both wavelets and shearlets have been proved to provide nearly diagonal representations of these operators.

# Shearlet-based Decomposition of $Rf$

For brevity of notation, we set $\mu = (j, \ell, k)$ and denote the element of the shearlet system by

$$\{\psi_\mu : \mu \in M\}.$$

We define the *companion shearlet system* as:

$$\{\psi_\mu^+ = 2^{-j}(-\Delta)^{1/4}\psi_\mu : \mu \in M\},$$

where $\Delta^\alpha$ is the fractional Laplacian, given by

$$((-\Delta)^\alpha f)\hat{}(\xi) = |\xi|^{2\alpha}\hat{f}(\xi).$$

It turns out that $\{\psi_\mu^+ : \mu \in M\}$ defines a frame for $L^2(\mathbb{R}^2)$.

# Shearlet-based Decomposition of $Rf$

Let $\widetilde{R} = (I \otimes D^1) \circ R$, where $D^\alpha$ is the 1-D fractional differentiation

$$(D^\alpha f)(t) = \int_{\mathbb{R}} |\omega|^{\frac{\alpha}{2}} \hat{f}(\omega)\, e^{2\pi i \omega t}\, d\omega.$$

Let $U_\mu = \widetilde{R}\, \psi_\mu^+$.

**Theorem**

For each $f \in L^2(\mathbb{R}^2)$ we have the decomposition:

$$f = \sum_{\mu \in M} [Rf, U_\mu]\, 2^j \psi_\mu.$$

(A similar result is proved by Candès and Donoho using curvelets)

# Shearlet-based Decomposition of $Rf$

**Theorem** (*Shearlet Decomposition of $Rf$*)

*Sketch of proof*

We will use the *Radon Isometry:*

$$[\widetilde{R}f, \widetilde{R}g]_{\mathcal{D}} = \langle f, g \rangle_{L^2}, \quad f, g \in L^2.$$

Hence

$$\begin{aligned}
\langle f, \psi_\mu \rangle &= [\widetilde{R}f, \widetilde{R}\,\psi_\mu] \\
&= [(I \otimes D^1) \circ Rf, (I \otimes D^1) \circ R\,\psi_\mu] \\
&= [Rf, (I \otimes D^1) \circ (I \otimes D^1) \circ R\,\psi_\mu] \\
&= [Rf, (I \otimes D^1) \circ R \circ (-\Delta)^{\frac{1}{4}}\,\psi_\mu] \\
&= [Rf, \widetilde{R}(2^j \psi_\mu^+)] \\
&= 2^j\,[Rf, U_\mu].
\end{aligned}$$

# Shearlet-based Inversion of $Rf$

Consider the situation where $f \in \mathcal{E}^2$, the space of *functions with compact support on* $\mathbb{R}^2$ *which are* $C^2$ *away from* $C^2$ *edges.*

Suppose that we are given the Radon data

$$Y = Rf + \epsilon W,$$

where $\epsilon W$ is white Gaussian noise and $\epsilon$ is measuring the noise level.

Our goal is to recover $f$ from $Y$ and assess how the noise level will affect the estimation.

# Shearlet-based Inversion of $Rf$

After projecting $Y$ onto the frame $\{U_\mu : \mu \in M\}$, using the Shearlet Decomposition, we obtain the discrete data:

$$\begin{aligned} y_\mu = [Y, U_\mu] &= [Rf, U_\mu] + \epsilon\,[W, U\mu] \\ &= 2^{-j}\,\langle f, \psi_\mu \rangle + \epsilon\, n_\mu. \end{aligned}$$

Hence, the estimator $\tilde{f}$ of $f$ is obtained by applying an appropriate *thresholding rule* $\tilde{c}_\mu = T(y_\mu)$:

$$\tilde{f} = \sum_{\mu \in N(\epsilon)} \tilde{c}_\mu\, \psi_\mu.$$

where $N(\epsilon)$ is the *set of significant indices* (the indices which are nonzero after the thresholding) and depends on the noise level.

# Shearlet-based Inversion of $Rf$

**Theorem**

Let $f \in \mathcal{E}^2$ be the solution of the problem $Y = Rf + \epsilon W$ and $\tilde{f}$ be the estimator $\tilde{f} = \sum_{\mu \in N(\epsilon)} \tilde{c}_\mu \, \psi_\mu$. It satisfies

$$\sup_{\mathcal{E}^2} E\|f - \tilde{f}\|^2 \leq C \, \epsilon^{4/5} \, \big(\log(\epsilon^{-1})\big) \, , \; \epsilon \to 0.$$

# Shearlet-based Inversion of $Rf$

**Theorem**

Let $f \in \mathcal{E}^2$ be the solution of the problem $Y = Rf + \epsilon W$ and $\tilde{f}$ be the estimator $\tilde{f} = \sum_{\mu \in N(\epsilon)} \tilde{c}_\mu \, \psi_\mu$. It satisfies

$$\sup_{\mathcal{E}^2} E\|f - \tilde{f}\|^2 \leq C\, \epsilon^{4/5} \, \left(\log(\epsilon^{-1})\right), \ \epsilon \to 0.$$

This approach provides an essentially optimal convergence rate for the estimator (similar to [CD01]).

# Shearlet-based Inversion of $Rf$

**Theorem**

Let $f \in \mathcal{E}^2$ be the solution of the problem $Y = Rf + \epsilon W$ and $\tilde{f}$ be the estimator $\tilde{f} = \sum_{\mu \in N(\epsilon)} \tilde{c}_\mu \, \psi_\mu$. It satisfies

$$\sup_{\mathcal{E}^2} E\|f - \tilde{f}\|^2 \leq C \, \epsilon^{4/5} \, \left(\log(\epsilon^{-1})\right), \; \epsilon \to 0.$$

This approach provides an essentially optimal convergence rate for the estimator (similar to [CD01]).

This ouperforms the similar Wavelet-Vagulelette Decomposition, which yields convergence rate $O(\epsilon^{2/3})$, as $\epsilon \to 0$.

Traditional linear methods behave poorly in the presence of edges and only yield convergence rate $O(\epsilon^{1/2})$, as $\epsilon \to 0$.

# Shearlet-based Inversion of $Rf$

*Ideas from the proof of Theorem (1).*

The proof of the theorem uses the following facts which are a consequence of the optimal approximation properties of the shearlet system for functions in $\mathcal{E}^2$:

1. The neglected coefficients $\{\langle f, \psi_\mu \rangle : \mu \notin N(\epsilon)\}$ satisfy:

$$\sup_{f \in \mathcal{E}^2} \sum_{\mu \notin N(\epsilon)} |\langle f, \psi_\mu \rangle|^2 \leq C' \epsilon^{4/5}.$$

2. The risk proxy satisfies:

$$\sup_{f \in \mathcal{E}^2(A)} \sum_{\mu \in N(\epsilon)} \min(|\langle f, \psi_\mu \rangle|^2, 2^{2j} \epsilon^2) \leq C'' \epsilon^{4/5}.$$

# Shearlet-based Inversion of $Rf$

*Ideas from the proof of Theorem (2).*

3. The cardinality of $N(\epsilon)$ obeys:

$$\#N(\epsilon) \leq C''' \epsilon^{-2}.$$

Using these observations and the fact that the shearlet system is a Parseval frame, one can show that:

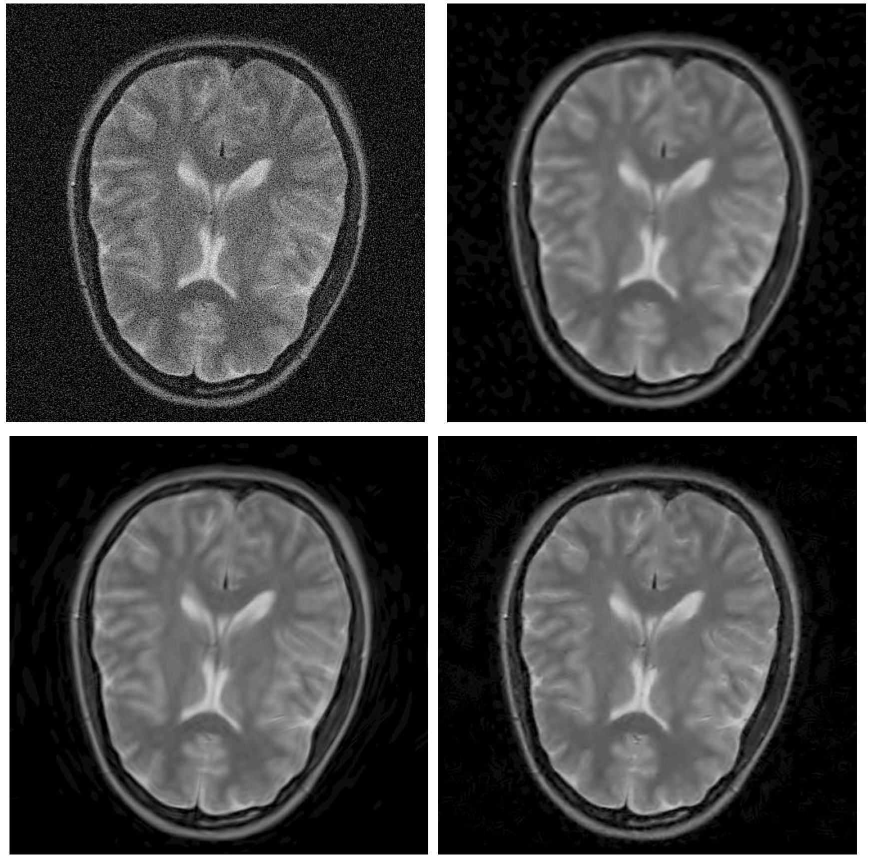
$$\begin{aligned} E\|\tilde{f} - f\|_2^2 &\leq E\|\tilde{c}_\mu - c_\mu\|_{\ell^2}^2 \\ &\leq E\left(\sum_{\mathcal{N}(\epsilon)} (\tilde{c}_\mu - c_\mu)^2\right) + E\left(\sum_{\mathcal{N}(\epsilon)^c} c_\mu^2\right) \\ &\leq C \log(\epsilon^{-1})\, \epsilon^{4/5}. \end{aligned}$$

## Numerical Demonstrations

Results prevously published in the literature [Lee, Lucier, 2001] have shown that a numerical algorithm implementing the wavelet-vaguelette decomposition outperforms traditional methods for the inversion of the Radon transform such as singular value decomposition and back-projection.

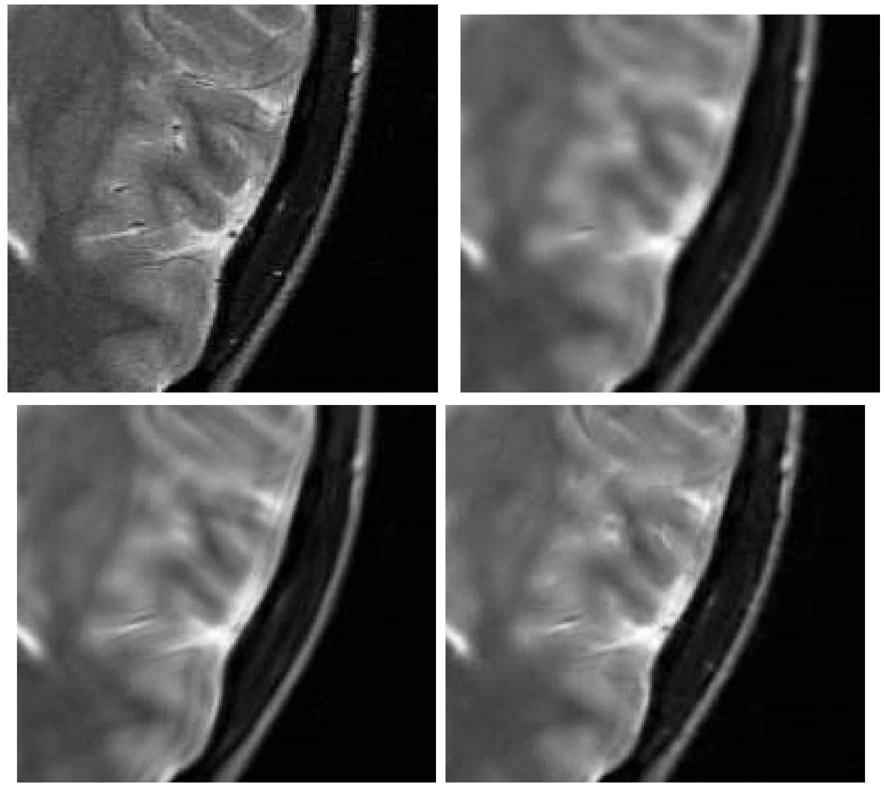
We have developed a numerical algorithm for the inversion of Radon data, based on the principles described above. This has been validated against similar routines based on wavelets and curvelets.

## Example: Numerical Inversion of $Rf$



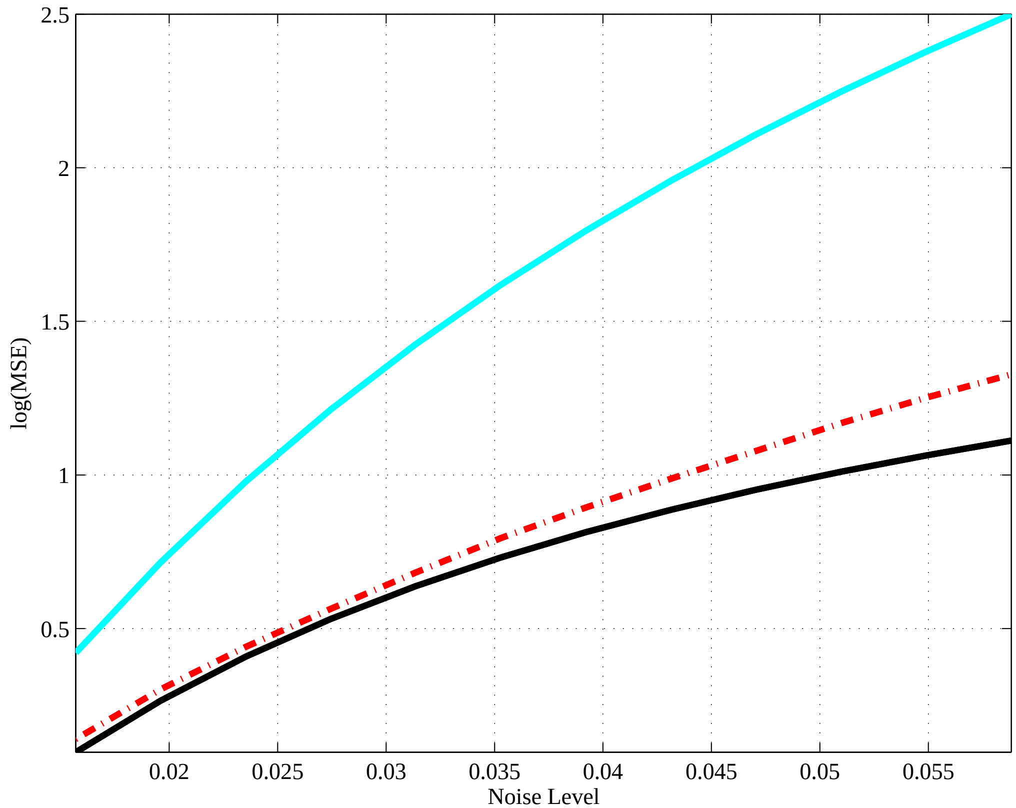
From the top, clockwise: unfiltered reconstruction (SNR=9.39 dB); wavelet-based estimate (SNR=20.14 dB); shearlet-based estimate (SNR=21.86 dB); curvelet-based estimate (SNR=20.85 dB)

# Example: Numerical Inversion of $Rf$



From the top, clockwise: wavelet-based estimate (SNR=20.14 dB); shearlet-based estimate (SNR=21.86 dB); curvelet-based estimate (SNR=20.85 dB)

# Example: Numerical Inversion of $Rf$




MSE performances of shearlet-based estimate (black line), curvelet-based estimate (dash-red line), and wavelet-based estimate (blue line) as a function of noise level.

# References available at:

www.math.uh.edu\~dlabate\publications.html
www.shearlets.org

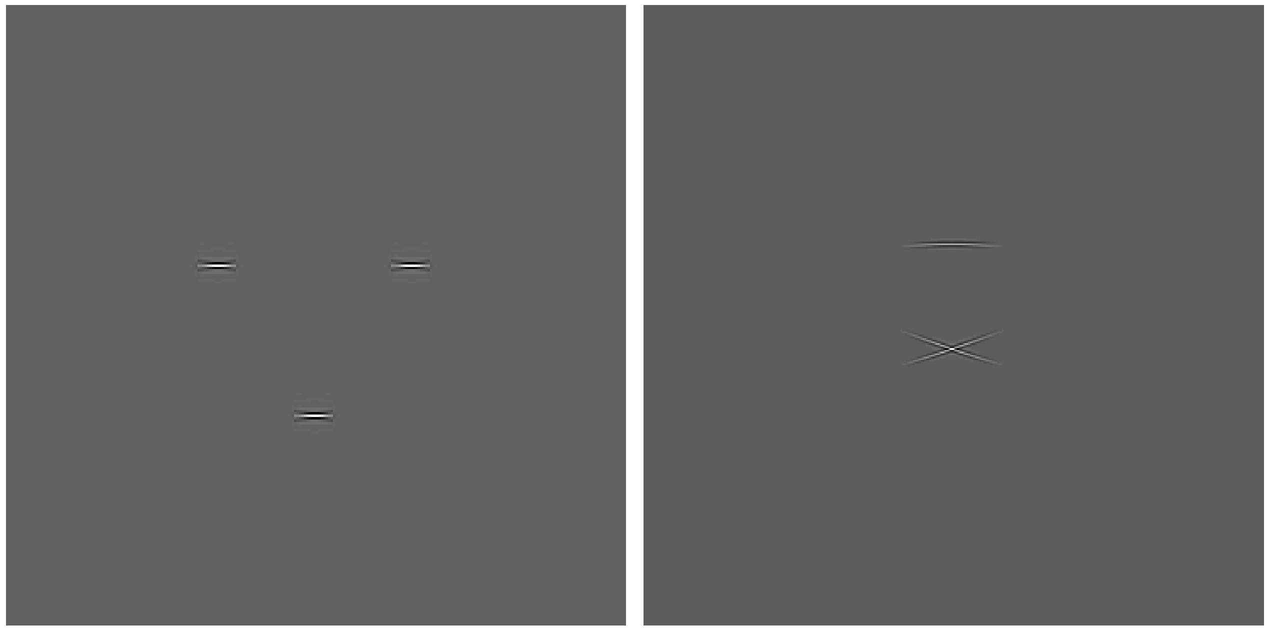


Figure 1: The image on the left illustrates some shearlet coefficients at a given direction and scale. The image on the right shows the corresponding companion shearlet coefficients in a Radon domain.