Sampling in inverse problems

Giovanni S. Alberti
MaLGa – Machine Learning Genoa Center
Department of Mathematics
University of Genoa

Harmonic Analysis E-Seminars

13 November 2024

Co-funded by the European Union (ERC, SAMPDE, 101041040). Views and opinions expressed are however those of the author only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the EU nor the granting authority can be held responsible for them.

Joint work with

Alessandro Felisi
(UniGe)

Matteo Santacesaria
(UniGe)

S. Ivan Trapasso
(PoliTo)

Compressed sensing for inverse problems and the sample complexity of the sparse Radon transform, J. Eur. Math. Soc., to appear

Outline

Sampling and inverse problems

Compressed sensing

Compressed sensing for inverse problems

Outline

Sampling and inverse problems

Compressed sensing

Compressed sensing for inverse problems

Sampling

- Function $f \in L^2(\mathcal{D})$
- Sampling points $t_l \in \mathcal{D}$, $l = 1, \dots, m$
- Sampling problem:

$$(f(t_l))_{l=1}^{m} \rightsquigarrow f$$

Sampling

- Function $f \in L^2(\mathcal{D})$
- Sampling points $t_l \in \mathcal{D}$, $l = 1, \dots, m$
- Sampling problem:

$$(f(t_l))_{l=1}^{m} \rightsquigarrow f$$

- Need assumptions on f
- Classically, f is ω-bandlimited (linear condition):

$$m \gtrsim \omega$$

Inverse problems

- $\mathcal{H}$ (e.g. $\mathcal{H} = L^2(\Omega)$): Hilbert space of inputs
- $F\colon \mathcal{H} \to L^2(\mathcal{D}; \mathcal{H}')$ linear **forward map** (compact)

[1]Natterer, The Mathematics of Computerized Tomography, 2001
Quinto, An Introduction to X-ray tomography and Radon Transforms, 2006

Inverse problems

- $\mathcal{H}$ (e.g. $\mathcal{H} = L^2(\Omega)$): Hilbert space of inputs
- $F \colon \mathcal{H} \to L^2(\mathcal{D}; \mathcal{H}')$ linear **forward map** (compact)
- Inverse problem:

$$F(u) \rightsquigarrow u$$

[1]Natterer, The Mathematics of Computerized Tomography, 2001
Quinto, An Introduction to X-ray tomography and Radon Transforms, 2006

Inverse problems

- $\mathcal{H}$ (e.g. $\mathcal{H} = L^2(\Omega)$): Hilbert space of inputs
- $F\colon \mathcal{H} \to L^2(\mathcal{D}; \mathcal{H}')$ linear **forward map** (compact)
- Inverse problem:

$$F(u) \rightsquigarrow u$$

Examples

1. **Deconvolution** (with Bessel operator):

$$F = (I - \Delta)^{-b/2}\colon L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2), \qquad F(u) = \kappa_b * u$$

where $b > 2$ and $\kappa_b \coloneqq \mathcal{F}^{-1}\left((1 + |\cdot|^2)^{-b/2}\right)$

[1]Natterer, The Mathematics of Computerized Tomography, 2001
Quinto, An Introduction to X-ray tomography and Radon Transforms, 2006

Inverse problems

- $\mathcal{H}$ (e.g. $\mathcal{H} = L^2(\Omega)$): Hilbert space of inputs
- $F\colon \mathcal{H} \to L^2(\mathcal{D}; \mathcal{H}')$ linear **forward map** (compact)
- Inverse problem:

$$F(u) \rightsquigarrow u$$

Examples

1. **Deconvolution** (with Bessel operator):

$$F = (I - \Delta)^{-b/2}\colon L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2), \qquad F(u) = \kappa_b * u$$

   where $b > 2$ and $\kappa_b \coloneqq \mathcal{F}^{-1}\left((1 + |\cdot|^2)^{-b/2}\right)$

2. **Radon transform**[1]:

$$\mathcal{R}\colon L^2(\mathcal{B}_1) \to L^2(\mathbb{S}^1; L^2(-1,1)), \qquad (\mathcal{R}u)(\theta) = \int_{\theta^\perp} u(y + \cdot\,\theta)dy \in L^2(-1,1)$$

[1] Natterer, The Mathematics of Computerized Tomography, 2001
Quinto, An Introduction to X-ray tomography and Radon Transforms, 2006

Sampling in inverse problems

- $F\colon \mathcal{H} \to L^2(\mathcal{D};\mathcal{H}')$ linear forward map
- Sampling points $t_l \in \mathcal{D}$, $l = 1, \dots, m$
- Inverse problems with sampling:

$$\left(F_{t_l} u := (Fu)(t_l)\right)_{l=1}^{m} \quad \rightsquigarrow \quad u$$

UniGe | MaLGa

Sampling in inverse problems

- $F \colon \mathcal{H} \to L^2(\mathcal{D}; \mathcal{H}')$ linear forward map
- Sampling points $t_l \in \mathcal{D}$, $l = 1, \ldots, m$
- Inverse problems with sampling:

$$\left(F_{t_l} u := (Fu)(t_l)\right)_{l=1}^{m} \quad \rightsquigarrow \quad u$$

Examples

1. Deconvolution: $F \colon L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$

$$\left((\kappa_b * u)(t_l)\right)_{l=1}^{m} \quad \rightsquigarrow \quad u$$

Sampling in inverse problems

- $F\colon \mathcal{H} \to L^2(\mathcal{D};\mathcal{H}')$ linear forward map
- Sampling points $t_l \in \mathcal{D}$, $l = 1,\dots,m$
- Inverse problems with sampling:

$$\left(F_{t_l}u := (Fu)(t_l)\right)_{l=1}^m \quad \rightsquigarrow \quad u$$

Examples

1. Deconvolution: $F\colon L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$

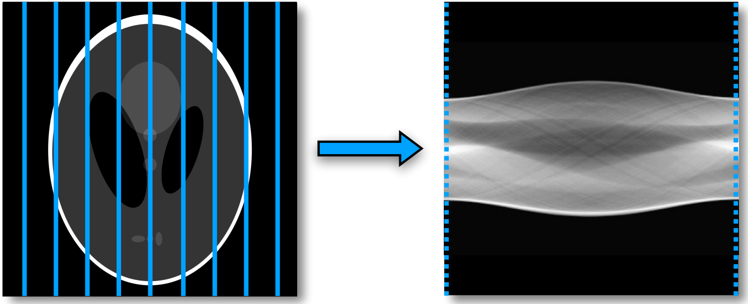
$$\left((\kappa_b * u)(t_l)\right)_{l=1}^m \quad \rightsquigarrow \quad u$$

2. Radon transform: $\mathcal{R}\colon L^2(\mathcal{B}_1) \to L^2(\mathbb{S}^1; L^2(-1,1))$

$$\left(\mathcal{R}_{\theta_l}u\right)_{l=1}^m \quad \rightsquigarrow \quad u$$

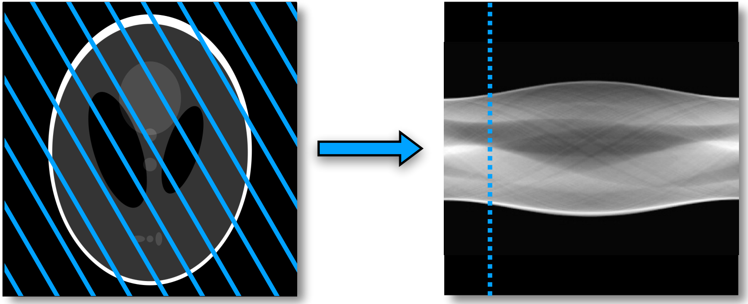
The sparse Radon transform

$$\mathcal{R}_\theta u(s) = \int_{\theta^\perp} u(y + s\theta)dy, \qquad \theta = \theta_1$$



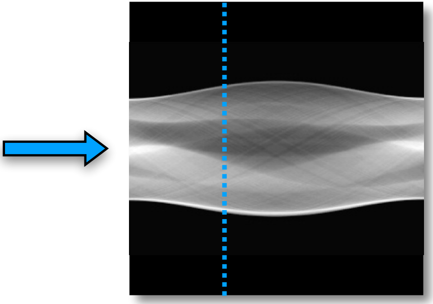
The sparse Radon transform

$$\mathcal{R}_\theta u(s) = \int_{\theta^\perp} u(y + s\theta)\,dy, \qquad \theta = \theta_2$$



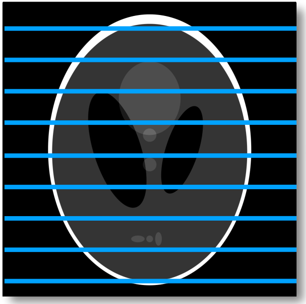
The sparse Radon transform

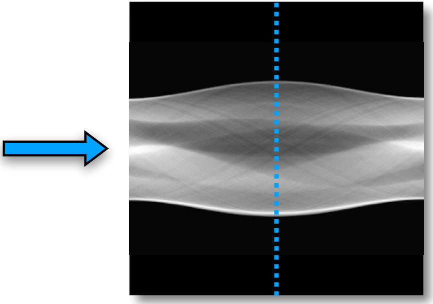
$$\mathcal{R}_\theta u(s) = \int_{\theta^\perp} u(y + s\theta)dy, \qquad \theta = \theta_3$$



The sparse Radon transform

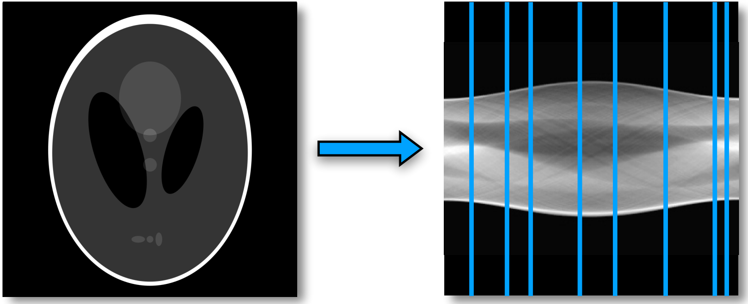
$$\mathcal{R}_\theta u(s) = \int_{\theta^\perp} u(y + s\theta)dy, \qquad \theta = \theta_4$$





The sparse Radon transform

$$(\mathcal{R}u(\theta_1,\cdot),\dots,\mathcal{R}u(\theta_m,\cdot)), \quad \theta_1,\dots,\theta_m \in \mathbb{S}^1$$



Sampling in inverse problems

- $F\colon \mathcal{H} \to L^2(\mathcal{D};\mathcal{H}')$ linear forward map
- Sampling points $t_l \in \mathcal{D}$, $l = 1, \dots, m$
- Inverse problems with sampling:

$$(F_{t_l} u)_{l=1}^m \quad \rightsquigarrow \quad u$$

UniGe | MaLGa

[2] R. A. DeVore, Nonlinear approximation, 1998

Sampling in inverse problems

- $F \colon \mathcal{H} \to L^2(\mathcal{D}; \mathcal{H}')$ linear forward map
- Sampling points $t_l \in \mathcal{D}$, $l = 1, \ldots, m$
- Inverse problems with sampling:

$$(F_{t_l} u)_{l=1}^m \quad \rightsquigarrow \quad u$$

Need **prior assumptions**:

- **Linear**: u bandlimited (or, more generally, smooth) is not realistic in most cases

UniGe | MaLGa

[2] R. A. DeVore, Nonlinear approximation, 1998

Sampling in inverse problems

- $F\colon \mathcal{H} \to L^2(\mathcal{D}; \mathcal{H}')$ linear forward map
- Sampling points $t_l \in \mathcal{D}$, $l = 1, \dots, m$
- Inverse problems with sampling:

$$(F_{t_l} u)_{l=1}^m \quad \rightsquigarrow \quad u$$

Need **prior assumptions**:

- **Linear**: u bandlimited (or, more generally, smooth) is not realistic in most cases
- **Nonlinear**[2]: u sparse...

[2] R. A. DeVore, Nonlinear approximation, 1998

Sparsity

- $\{\phi_n\}_n$ **orthonormal/Riesz basis** of $\mathcal{H}$

Sparsity

- $\{\phi_n\}_n$ **orthonormal/Riesz basis** of $\mathcal{H}$
- $u \in \mathcal{H}$ is s-**sparse** if

$$\#\{n : \langle u, \phi_n \rangle \neq 0\} \leqslant s$$

Sparsity

- $\{\phi_n\}_n$ **orthonormal/Riesz basis** of $\mathcal{H}$
- $u \in \mathcal{H}$ is s-**sparse** if

$$\#\{n : \langle u, \phi_n \rangle \neq 0\} \leqslant s$$

- compare with a **linear** condition:

$$\{n : \langle u, \phi_n \rangle \neq 0\} \subseteq \{1, \dots, s\}$$

Sparsity

- $\{\phi_n\}_n$ **orthonormal/Riesz basis** of $\mathcal{H}$
- $u \in \mathcal{H}$ is s-**sparse** if

$$\#\{n : \langle u, \phi_n \rangle \neq 0\} \leqslant s$$

- compare with a **linear** condition:

$$\{n : \langle u, \phi_n \rangle \neq 0\} \subseteq \{1, \ldots, s\}$$

Figure: **Important Genoese**

Figure: **Wavelet coefficients**

Main goal

Problem:

$$(F_{t_l} u)_{l=1}^{m} \quad \rightsquigarrow \quad u$$

Main goal

Problem:

$$(F_{t_l} u)_{l=1}^{m} \quad \rightsquigarrow \quad u$$

The main goal is to understand

- how to choose the samples $t_1, \ldots, t_m \in \mathcal{D}$

Main goal

Problem:

$$(F_{t_l} u)_{l=1}^{m} \quad \rightsquigarrow \quad u$$

The main goal is to understand

- how to choose the samples $t_1, \ldots, t_m \in \mathcal{D}$
- and how many are needed ($m = ?$)

Outline

Sampling and inverse problems

Compressed sensing

Compressed sensing for inverse problems

An example: Magnetic Resonance Imaging

Measurements: $\mathrm{F} =$ Fourier transform

Compressed sensing[3]

Setup:

[3] E. J. Candès, J. K. Romberg, T. Tao. Stable signal recovery from incomplete and inaccurate measurements. Comm. Pure Appl. Math. 59(8) (2006), 1207-1223
D. L. Donoho. Compressed sensing. IEEE Trans. Inf. Theory, 52(4) (2006), 1289–1306

Compressed sensing[3]

Setup:

- **Unknown**: $u^\dagger \in \mathbb{C}^M$ is s-sparse
- $\{\psi_t\}_t$ orthonormal basis (MRI: Fourier)
- Random subsampling: $t_l \in \{1, \dots, M\}$ chosen uniformly at random

[3] E. J. Candès, J. K. Romberg, T. Tao. Stable signal recovery from incomplete and inaccurate measurements. Comm. Pure Appl. Math. 59(8) (2006), 1207-1223
D. L. Donoho. Compressed sensing. IEEE Trans. Inf. Theory, 52(4) (2006), 1289–1306

Compressed sensing[3]

Setup:

- **Unknown**: $u^\dagger \in \mathbb{C}^M$ is s-sparse
- $\{\psi_t\}_t$ orthonormal basis (MRI: Fourier)
- Random subsampling: $t_l \in \{1, \ldots, M\}$ chosen uniformly at random

Problem:

$$(\langle u^\dagger, \psi_{t_l} \rangle)_{l=1}^m \quad \rightsquigarrow \quad u^\dagger$$

[3] E. J. Candès, J. K. Romberg, T. Tao. Stable signal recovery from incomplete and inaccurate measurements. Comm. Pure Appl. Math. 59(8) (2006), 1207-1223
D. L. Donoho. Compressed sensing. IEEE Trans. Inf. Theory, 52(4) (2006), 1289–1306

Compressed sensing[3]

Setup:

- **Unknown**: $u^\dagger \in \mathbb{C}^M$ is s-sparse
- $\{\psi_t\}_t$ orthonormal basis (MRI: Fourier)
- Random subsampling: $t_l \in \{1, \ldots, M\}$ chosen uniformly at random

Problem:

$$(\langle u^\dagger, \psi_{t_l} \rangle)_{l=1}^m \quad \rightsquigarrow \quad u^\dagger$$

with

$$m \gtrsim s$$

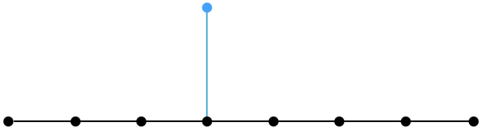
[3] E. J. Candès, J. K. Romberg, T. Tao. Stable signal recovery from incomplete and inaccurate measurements. Comm. Pure Appl. Math. 59(8) (2006), 1207-1223
D. L. Donoho. Compressed sensing. IEEE Trans. Inf. Theory, 52(4) (2006), 1289–1306

Coherence

In general, sparsity alone is **not enough**

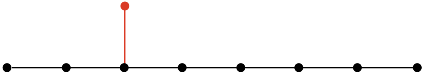
Coherence

In general, sparsity alone is **not enough**
Suppose we have a 1-sparse vector $u^\dagger \in \mathbb{R}^M$ $(M = 8)$ w.r.t. $\phi_n = \delta_n$



Coherence

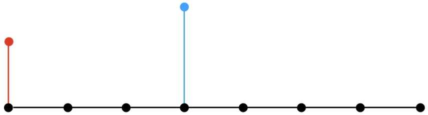
In general, sparsity alone is **not enough**
Suppose we have a 1-sparse vector $u^{\dagger} \in \mathbb{R}^{M}$ $(M = 8)$ w.r.t. $\phi_n = \delta_n$
Consider a sensing system $\psi_l = \delta_l$



UniGe | MaLGa

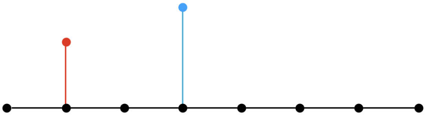
Coherence

In general, sparsity alone is **not enough**
Suppose we have a 1-sparse vector $u^\dagger \in \mathbb{R}^M$ $(M = 8)$ w.r.t. $\phi_n = \delta_n$
Consider a sensing system $\psi_l = \delta_l$; we count $m = 1$



Coherence

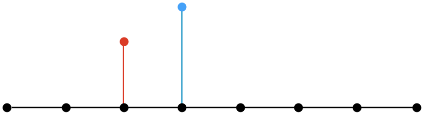
In general, sparsity alone is **not enough**
Suppose we have a 1-sparse vector $u^\dagger \in \mathbb{R}^M$ $(M = 8)$ w.r.t. $\phi_n = \delta_n$
Consider a sensing system $\psi_l = \delta_l$; we count $m = 2$



UniGe | MaLGa

Coherence

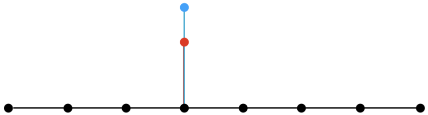
In general, sparsity alone is **not enough**
Suppose we have a 1-sparse vector $u^\dagger \in \mathbb{R}^M$ ($M = 8$) w.r.t. $\phi_n = \delta_n$
Consider a sensing system $\psi_l = \delta_l$; we count $m = 3$



UniGe | MaLGa

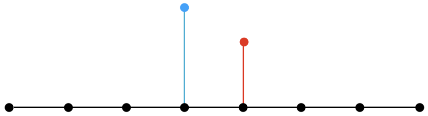
Coherence

In general, sparsity alone is **not enough**
Suppose we have a 1-sparse vector $u^\dagger \in \mathbb{R}^M$ ($M = 8$) w.r.t. $\phi_n = \delta_n$
Consider a sensing system $\psi_l = \delta_l$; we count $m = 4$



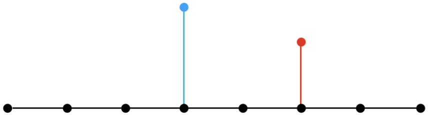
Coherence

In general, sparsity alone is **not enough**
Suppose we have a 1-sparse vector $u^\dagger \in \mathbb{R}^M$ ($M = 8$) w.r.t. $\phi_n = \delta_n$
Consider a sensing system $\psi_l = \delta_l$; we count $m = 5$



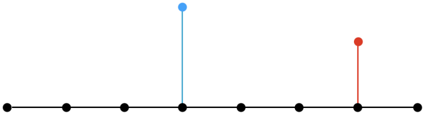
Coherence

In general, sparsity alone is **not enough**
Suppose we have a 1-sparse vector $u^\dagger \in \mathbb{R}^M$ $(M = 8)$ w.r.t. $\phi_n = \delta_n$
Consider a sensing system $\psi_l = \delta_l$; we count $m = 6$



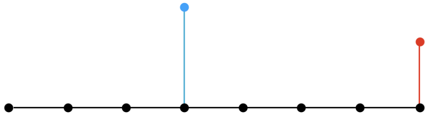
Coherence

In general, sparsity alone is **not enough**
Suppose we have a 1-sparse vector $u^\dagger \in \mathbb{R}^M$ ($M = 8$) w.r.t. $\phi_n = \delta_n$
Consider a sensing system $\psi_t = \delta_t$; we count $m = 7$



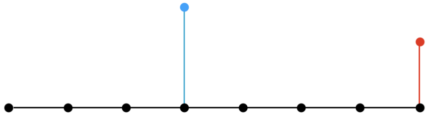
Coherence

In general, sparsity alone is **not enough**
Suppose we have a 1-sparse vector $u^\dagger \in \mathbb{R}^M$ $(M = 8)$ w.r.t. $\phi_n = \delta_n$
Consider a sensing system $\psi_t = \delta_t$; we count $m = 8$



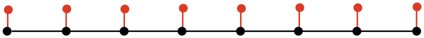
Coherence

In general, sparsity alone is **not enough**
Suppose we have a 1-sparse vector $u^\dagger \in \mathbb{R}^M$ $(M = 8)$ w.r.t. $\phi_n = \delta_n$
Consider a sensing system $\psi_t = \delta_t$; we count $m = M$



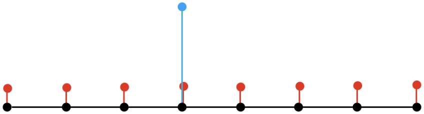
Coherence

In general, sparsity alone is **not enough**
Suppose we have a 1-sparse vector $u^\dagger \in \mathbb{R}^M$ $(M = 8)$ w.r.t. $\phi_n = \delta_n$
Consider another sensing system $(\psi_t)_t$



Coherence

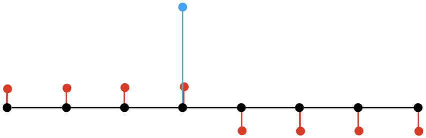
In general, sparsity alone is **not enough**
Suppose we have a 1-sparse vector $u^\dagger \in \mathbb{R}^M$ $(M = 8)$ w.r.t. $\phi_n = \delta_n$
Consider another sensing system $(\psi_t)_t$; we count $m = 1$



UniGe | MaLGa

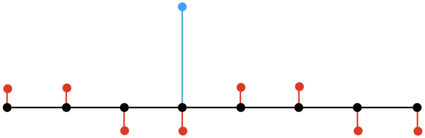
Coherence

In general, sparsity alone is **not enough**
Suppose we have a 1-sparse vector $u^\dagger \in \mathbb{R}^M$ ($M = 8$) w.r.t. $\phi_n = \delta_n$
Consider another sensing system $(\psi_t)_t$; we count $m = 2$



Coherence

In general, sparsity alone is **not enough**
Suppose we have a 1-sparse vector $u^\dagger \in \mathbb{R}^M$ ($M = 8$) w.r.t. $\phi_n = \delta_n$
Consider another sensing system $(\psi_t)_t$; we count $m = 3$

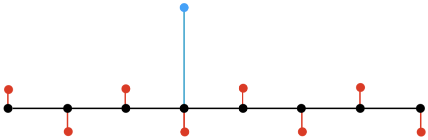


Coherence

In general, sparsity alone is **not enough**

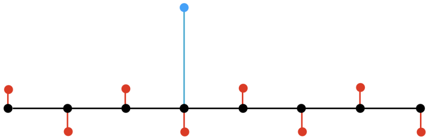
Suppose we have a 1-sparse vector $u^\dagger \in \mathbb{R}^M$ ($M = 8$) w.r.t. $\phi_n = \delta_n$

Consider another sensing system $(\psi_t)_t$; we count $m = 4$



Coherence

In general, sparsity alone is **not enough**
Suppose we have a 1-sparse vector $u^\dagger \in \mathbb{R}^M$ $(M = 8)$ w.r.t. $\phi_n = \delta_n$
Consider another sensing system $(\psi_t)_t$; we count $m = 1 + \log_2 M$



UniGe | MaLGa

Coherence

This feature is measured by the coherence between the **sparsifying dictionary** and the **sensing system**

Coherence

This feature is measured by the coherence between the **sparsifying dictionary** and the **sensing system**

$$B = \sqrt{M} \cdot \max_{n,t} |\langle \phi_n, \psi_t \rangle|$$

Coherence

$$B = \sqrt{M} \cdot \max |\langle \phi_n, \psi_t \rangle|$$

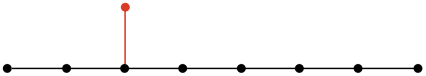


Figure: **First example**: $B = \sqrt{M}$

Coherence

$$B = \sqrt{M} \cdot \max |\langle \phi_n, \psi_t \rangle|$$

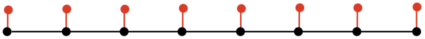


Figure: **Second example**: $B = 1$

Recovery estimate[4]

- $u^\dagger \in \mathbb{C}^M$ **unknown**
- $u^\dagger$ is s-**sparse** w.r.t. $\{\phi_n\}_{n=1}^M$
- **minimization problem**

$$\widehat{u} \in \arg\min_{u \in \mathbb{C}^M} \|(\langle u, \phi_n \rangle)_n\|_1 \quad : \quad \langle u, \psi_{t_l} \rangle = \langle u^\dagger, \psi_{t_l} \rangle, \; l = 1, \ldots, m$$

[4] S. Foucart, H. Rauhut. A mathematical introduction to compressive sensing. 2013

Recovery estimate[4]

- $u^\dagger \in \mathbb{C}^M$ **unknown**
- $u^\dagger$ is s-**sparse** w.r.t. $\{\phi_n\}_{n=1}^M$
- **minimization problem**

$$\widehat{u} \in \arg\min_{u \in \mathbb{C}^M} \|(\langle u, \phi_n \rangle)_n\|_1 \quad : \quad \langle u, \psi_{t_l} \rangle = \langle u^\dagger, \psi_{t_l} \rangle, \; l = 1, \ldots, m$$

Theorem

If

$$m \gtrsim B^2 s \cdot \text{log factors}$$

then

$$\widehat{u} = u^\dagger$$

with overwhelming probability.

UniGe | MaLGa

[4] S. Foucart, H. Rauhut. A mathematical introduction to compressive sensing. 2013

Outline

Sampling and inverse problems

Compressed sensing

Compressed sensing for inverse problems

Setting[5]

- $\mathcal{H} = L^2(\Omega)$ with $\Omega \subset \mathbb{R}^2$ bounded
- $(\phi_{j,n})_{j,n}$: sufficiently nice **wavelet basis**
- $u^\dagger$ is **s-sparse** w.r.t. the wavelet basis

[5] B. Adcock, A. C. Hansen, C. Poon, B. Roman, Breaking the coherence barrier: A new theory for compressed sensing, Forum Math. Sigma, 2017.
E. Herrholz, G. Teschke, Compressive sensing principles and iterative sparse recovery for inverse and ill-posed problems, Inverse Probl., 2010

Setting[5]

- $\mathcal{H} = L^2(\Omega)$ with $\Omega \subset \mathbb{R}^2$ bounded
- $(\phi_{j,n})_{j,n}$: sufficiently nice **wavelet basis**
- $u^\dagger$ is **s-sparse** w.r.t. the wavelet basis

Forward map

$$u \longmapsto (F_{t_l} u)_{l=1}^m$$

not a subsampled isometry, but a **subsampled compact** operator

[5] B. Adcock, A. C. Hansen, C. Poon, B. Roman, Breaking the coherence barrier: A new theory for compressed sensing, Forum Math. Sigma, 2017.
E. Herrholz, G. Teschke, Compressive sensing principles and iterative sparse recovery for inverse and ill-posed problems, Inverse Probl., 2010

Difficulties with the Radon transform

Difficulties with the Radon transform

- From Jørgensen, Coban, Lionheart, McDonald and Withers, 2017:

  Compressive sensing connects the critical number of projections to the image sparsity, but does not cover CT. Empirical results suggest a similar connection.

- From Hansen, 2017:

  We used simulations studies to provide a foundation for the use of sparsity in CT where, unlike compressed sensing, it is not possible to give rigorous proofs.

Difficulties with the Radon transform

- From Jørgensen, Coban, Lionheart, McDonald and Withers, 2017:

  Compressive sensing connects the critical number of projections to the image sparsity, but does not cover CT. Empirical results suggest a similar connection.

- From Hansen, 2017:

  We used simulations studies to provide a foundation for the use of sparsity in CT where, unlike compressed sensing, it is not possible to give rigorous proofs.

Key tools:

1. Quasi-diagonalization of F
2. Relative coherence

Quasi-diagonalization

- $u^\dagger$ is **s-sparse**

[6] S. Mallat. A Wavelet Tour of Signal Processing. The Sparse Way, 2009

Quasi-diagonalization

- $u^\dagger$ is **s-sparse** – what about $Fu^\dagger$?

[6]S. Mallat. A Wavelet Tour of Signal Processing. The Sparse Way, 2009

Quasi-diagonalization

- $u^\dagger$ is **s-sparse** – what about $Fu^\dagger$?
- **smoothing operators** F:

$$\|Fu\|_{L^2} \asymp \|u\|_{H^{-b}}$$

[6] S. Mallat. A Wavelet Tour of Signal Processing. The Sparse Way, 2009

Quasi-diagonalization

- $u^\dagger$ is **s-sparse** – what about $Fu^\dagger$?
- **smoothing operators** F:

$$\|Fu\|_{L^2} \asymp \|u\|_{H^{-b}}$$

- **Littlewood-Paley properties** of wavelets[6]:

$$\|u\|_{H^{-b}}^2 \asymp \sum_{j,n} 2^{-2bj} |\langle u, \phi_{j,n} \rangle|^2$$

[6] S. Mallat. A Wavelet Tour of Signal Processing. The Sparse Way, 2009

Quasi-diagonalization

- $u^\dagger$ is **s-sparse** – what about $Fu^\dagger$?
- **smoothing operators** F:

$$\|Fu\|_{L^2} \asymp \|u\|_{H^{-b}}$$

- **Littlewood-Paley properties** of wavelets[6]:

$$\|u\|^2_{H^{-b}} \asymp \sum_{j,n} 2^{-2bj} |\langle u, \phi_{j,n}\rangle|^2$$

- **quasi-diagonalization** property:

$$\|Fu\|^2_{L^2} \asymp \sum_{j,n} 2^{-2bj} |\langle u, \phi_{j,n}\rangle|^2$$

[6] S. Mallat. A Wavelet Tour of Signal Processing. The Sparse Way, 2009

Quasi-diagonalization

- $u^\dagger$ is **s-sparse** – what about $Fu^\dagger$?
- **smoothing operators** F:

$$\|Fu\|_{L^2} \asymp \|u\|_{H^{-b}}$$

- **Littlewood-Paley properties** of wavelets[6]:

$$\|u\|^2_{H^{-b}} \asymp \sum_{j,n} 2^{-2bj} |\langle u, \phi_{j,n} \rangle|^2$$

- **quasi-diagonalization** property:

$$\|Fu\|^2_{L^2} \asymp \sum_{j,n} 2^{-2bj} |\langle u, \phi_{j,n} \rangle|^2$$

- **pseudo-sparsity property** on $Fu^\dagger$: use compressed sensing methods

[6] S. Mallat. A Wavelet Tour of Signal Processing. The Sparse Way, 2009

Relative coherence

What is the analogous concept of coherence in this case?

Relative coherence

What is the analogous concept of coherence in this case?

Classically defined as

$$B = \sqrt{M} \sup_{n,t} |F_t(\phi_n)|$$

when we sampled with respect to the **uniform probability** on $[M]$

Relative coherence

$$B = \sqrt{M} \sup_{n,t} |F_t(\phi_n)|$$

$g(t) \equiv 1/M$ is the **probability density** w.r.t. the counting measure on $\mathcal{D} = \{1, \dots, M\}$

UniGe | MaLGa

Relative coherence

$$B = \sup_{n,t} \sqrt{\frac{1}{g(t)}} \cdot |F_t(\phi_n)|$$

$g(t) \equiv 1/M$ is the **probability density** w.r.t. the counting measure on $\mathcal{D} = \{1, \ldots, M\}$

UniGe | MaLGa

Relative coherence

$$B = \sup_{n,t} \sqrt{\frac{1}{g(t)}} \cdot |F_t(\phi_n)|$$

$g(t) \equiv 1/M$ is the **probability density** w.r.t. the counting measure on $\mathcal{D} = \{1, \ldots, M\}$

Sampling rule in general $\mathcal{D}$ (e.g. $\mathcal{D} \subseteq \mathbb{R}^d$):

$$g(t)\, dt$$

UniGe | MaLGa

Relative coherence

$$B = \sup_{n,t} \sqrt{\frac{1}{g(t)}} \cdot |F_t(\phi_n)|$$

Relative coherence

$$B = \sup_{n,t} \sqrt{\frac{1}{g(t)}} \cdot |F_t(\phi_n)|$$

Compact operator: need to normalize

$$B_{rel} = \sup_{n,t} \sqrt{\frac{1}{g(t)}} \cdot \frac{|F_t \phi_n|}{\|F\phi_n\|_{L^2}}$$

Relative coherence

$$B = \sup_{n,t} \sqrt{\frac{1}{g(t)}} \cdot |F_t(\phi_n)|$$

Compact operator: need to normalize

$$B_{rel} = \sup_{n,t} \sqrt{\frac{1}{g(t)}} \cdot \frac{|F_t \phi_n|}{\|F\phi_n\|_{L^2}}$$

We call this quantity relative coherence

Optimizing the probability density $g(t)$

The **optimal choice** for $g(t)$ (minimizing B_{rel}) depends on the decay in t of $|F_t \phi_{j,n}|$

UniGe | MaLGa

Optimizing the probability density $g(t)$

The **optimal choice** for $g(t)$ (minimizing B_{rel}) depends on the decay in t of $|F_t \phi_{j,n}|$

Examples

UniGe | MaLGa

Optimizing the probability density $g(t)$

The **optimal choice** for $g(t)$ (minimizing B_{rel}) depends on the decay in t of $|F_t\phi_{j,n}|$

Examples

1. **Deconvolution** for $b > 2$:

$$|\big((I-\Delta)^{-b/2}\phi_{j,n}\big)(t)| \lesssim \frac{e^{-C_b|d(t,\Omega)|}}{2^{(b-1)j}} \quad \Rightarrow \quad g(t) \propto e^{-C_b|d(t,\Omega)|}$$

UniGe | MaLGa

Optimizing the probability density $g(t)$

The **optimal choice** for $g(t)$ (minimizing B_{rel}) depends on the decay in t of $|F_t \phi_{j,n}|$

Examples

1. **Deconvolution** for $b > 2$:

$$|((I-\Delta)^{-b/2}\phi_{j,n})(t)| \lesssim \frac{e^{-C_b|d(t,\Omega)|}}{2^{(b-1)j}} \quad \Rightarrow \quad g(t) \propto e^{-C_b|d(t,\Omega)|}$$

2. **The Radon transform**:

$$\|\mathcal{R}_\theta \phi_{j,n}\|_{L^2([-1,1])} \lesssim 1 \quad \Rightarrow \quad g(\theta) = \frac{1}{2\pi}$$

Main abstract result[7]

Theorem

[7] G.S.A., A. Felisi, M. Santacesaria, S.I. Trapasso, JEMS, to appear

Main abstract result[7]

Theorem

- **Unknown**: $u^\dagger \in L^2(\Omega)$

[7] G.S.A., A. Felisi, M. Santacesaria, S.I. Trapasso, JEMS, to appear

Main abstract result[7]

Theorem

- **Unknown**: $u^\dagger \in L^2(\Omega)$
- **Sparsity**: $u^\dagger$ is s-sparse w.r.t. the wavelet basis $(\phi_{j,n})_{j,n}$

[7] G.S.A., A. Felisi, M. Santacesaria, S.I. Trapasso, JEMS, to appear

Main abstract result[7]

Theorem

- **Unknown**: $u^\dagger \in L^2(\Omega)$
- **Sparsity**: $u^\dagger$ is s-sparse w.r.t. the wavelet basis $(\phi_{j,n})_{j,n}$
- **Measurements**: $t_1, \ldots, t_m \in \mathcal{D}$ chosen i.i.d. w.r.t. $g(t)\,dt$ with

$$m \gtrsim B_{rel}^2 s \cdot \text{log factors}$$

[7] G.S.A., A. Felisi, M. Santacesaria, S.I. Trapasso, JEMS, to appear

Main abstract result[7]

Theorem

- **Unknown**: $u^\dagger \in L^2(\Omega)$
- **Sparsity**: $u^\dagger$ is s-sparse w.r.t. the wavelet basis $(\phi_{j,n})_{j,n}$
- **Measurements**: $t_1, \dots, t_m \in \mathcal{D}$ chosen i.i.d. w.r.t. $g(t)\,dt$ with

$$m \gtrsim B_{rel}^2 s \cdot \text{log factors}$$

- **Minimization problem**:

$$\hat{u} \in \arg\min_{u \in \mathcal{H}} \|(\langle u, \phi_{j,n} \rangle)_{j,n}\|_{1,w} \quad : \quad F_{t_l} u = F_{t_l} u^\dagger,\ l = 1, \dots, m$$

[7] G.S.A., A. Felisi, M. Santacesaria, S.I. Trapasso, JEMS, to appear

Main abstract result[7]

Theorem

- **Unknown**: $u^\dagger \in L^2(\Omega)$
- **Sparsity**: $u^\dagger$ is s-sparse w.r.t. the wavelet basis $(\phi_{j,n})_{j,n}$
- **Measurements**: $t_1, \ldots, t_m \in \mathcal{D}$ chosen i.i.d. w.r.t. $g(t)\,dt$ with

$$m \gtrsim B_{rel}^2 s \cdot \text{log factors}$$

- **Minimization problem**:

$$\hat{u} \in \operatorname*{arg\,min}_{u \in \mathcal{H}} \|(\langle u, \phi_{j,n} \rangle)_{j,n}\|_{1,w} \quad : \quad F_{t_l} u = F_{t_l} u^\dagger,\ l = 1, \ldots, m$$

Then

$$\hat{u} = u^\dagger$$

with overwhelming probability.

UniGe | MaLGa

[7] G.S.A., A. Felisi, M. Santacesaria, S.I. Trapasso, JEMS, to appear

A corollary for the Radon transform[8]

$$\big(\mathcal{R}u^\dagger(\theta_1,\cdot),\dots,\mathcal{R}u^\dagger(\theta_m,\cdot)\big)\in L^2(-1,1)^m \qquad \longrightarrow \qquad u^\dagger\in L^2(\mathcal{B}_1)$$

UniGe | MaLGa

[8]G.S.A., A. Felisi, M. Santacesaria, S.I. Trapasso, JEMS, to appear

A corollary for the Radon transform[8]

$$\big(\mathcal{R}u^\dagger(\theta_1,\cdot),\ldots,\mathcal{R}u^\dagger(\theta_m,\cdot)\big)\in L^2(-1,1)^m \qquad \longrightarrow \qquad u^\dagger\in L^2(\mathcal{B}_1)$$

Theorem

- Sparsity: unknown $u^\dagger \in L^2(\mathcal{B}_1)$ is s-sparse wrt an ONB of wavelets $(\phi_{j,n})_{j,n}$

UniGe | MaLGa

[8] G.S.A., A. Felisi, M. Santacesaria, S.I. Trapasso, JEMS, to appear

A corollary for the Radon transform[8]

$$\left(\mathcal{R}u^\dagger(\theta_1,\cdot),\ldots,\mathcal{R}u^\dagger(\theta_m,\cdot)\right)\in L^2(-1,1)^m \qquad \longrightarrow \qquad u^\dagger \in L^2(\mathcal{B}_1)$$

Theorem

- Sparsity: unknown $u^\dagger \in L^2(\mathcal{B}_1)$ is s-sparse wrt an ONB of wavelets $(\phi_{j,n})_{j,n}$
- Measurements: $\theta_1,\ldots,\theta_m \in [0,\pi]$ chosen uniformly at random with

$$m \gtrsim s \cdot \text{log factors}$$

[8]G.S.A., A. Felisi, M. Santacesaria, S.I. Trapasso, JEMS, to appear

A corollary for the Radon transform[8]

$$\left(\mathcal{R}u^\dagger(\theta_1,\cdot),\dots,\mathcal{R}u^\dagger(\theta_m,\cdot)\right)\in L^2(-1,1)^m \qquad \longrightarrow \qquad u^\dagger\in L^2(\mathcal{B}_1)$$

Theorem

- Sparsity: unknown $u^\dagger \in L^2(\mathcal{B}_1)$ is s-sparse wrt an ONB of wavelets $(\phi_{j,n})_{j,n}$
- Measurements: $\theta_1,\dots,\theta_m \in [0,\pi]$ chosen uniformly at random with

$$m \gtrsim s \cdot \text{log factors}$$

- Minimization problem:

$$\hat{u} \in \arg\min_u \|\Phi u\|_{1,w} \quad : \quad \mathcal{R}_{\theta_l} u = \mathcal{R}_{\theta_l} u^\dagger,\ \ l = 1,\dots,m$$

[8] G.S.A., A. Felisi, M. Santacesaria, S.I. Trapasso, JEMS, to appear

A corollary for the Radon transform[8]

$$\left(\mathcal{R}u^\dagger(\theta_1,\cdot),\dots,\mathcal{R}u^\dagger(\theta_m,\cdot)\right)\in L^2(-1,1)^m \qquad \longrightarrow \qquad u^\dagger \in L^2(\mathcal{B}_1)$$

Theorem

- Sparsity: unknown $u^\dagger \in L^2(\mathcal{B}_1)$ is s-sparse wrt an ONB of wavelets $(\phi_{j,n})_{j,n}$
- Measurements: $\theta_1,\dots,\theta_m \in [0,\pi]$ chosen uniformly at random with

$$m \gtrsim s \cdot \text{log factors}$$

- Minimization problem:

$$\hat{u} \in \arg\min_u \|\Phi u\|_{1,w} \quad : \quad \mathcal{R}_{\theta_l} u = \mathcal{R}_{\theta_l} u^\dagger, \ l = 1,\dots,m$$

Then, with high probability,

$$\hat{u} = u^\dagger$$

[8] G.S.A., A. Felisi, M. Santacesaria, S.I. Trapasso, JEMS, to appear

Conclusions

Past

- Theory of CS for random matrices and subsampled isometries (e.g. MRI)
- Empirical evidence for compressed sensing Radon transform

Present

- Abstract theory of sample complexity for inverse problems
- Rigorous theory of compressed sensing for the sparse Radon transform

Future

- Wavelets $\rightarrow$ shearlets, curvelets, etc.
- Generalisation to other ill-posed problems, possibly nonlinear

Paper

Slides

UniGe

MaLGa