Super-resolution via Convex Programming

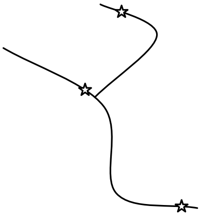
Carlos Fernandez-Granda
(Joint work with Emmanuel Candès)

Structure and Randomness in System Identification and Learning, IPAM

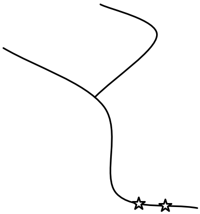
1/17/2013

Index

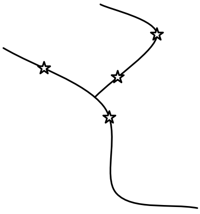
Single-molecule imaging



Frame 1

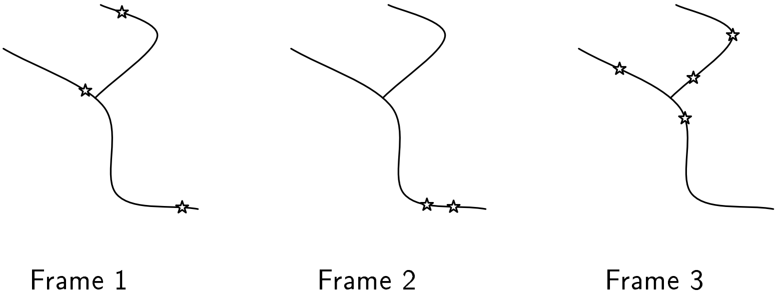


Frame 2



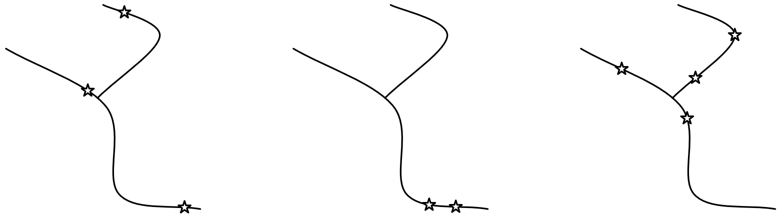
Frame 3

Single-molecule imaging



- Microscope receives light from fluorescent molecules

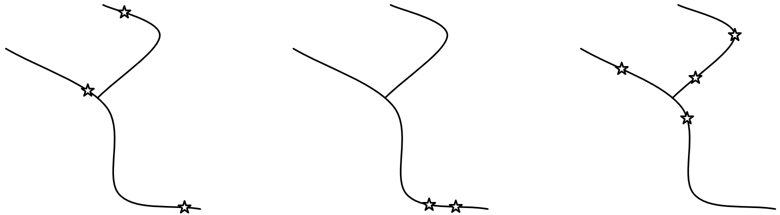
Single-molecule imaging



Frame 1 Frame 2 Frame 3

- Microscope receives light from fluorescent molecules
- **Few** molecules are active in each frame $\Rightarrow$ **sparsity**

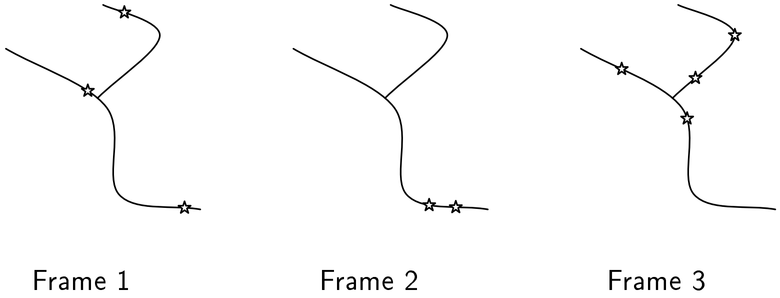
Single-molecule imaging



Frame 1 Frame 2 Frame 3

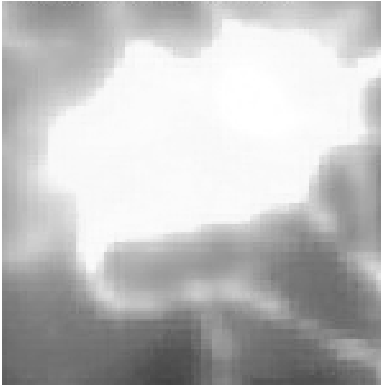
- Microscope receives light from fluorescent molecules
- **Few** molecules are active in each frame $\Rightarrow$ **sparsity**
- Multiple ($\sim$ 10000) frames are recorded and processed individually

Single-molecule imaging

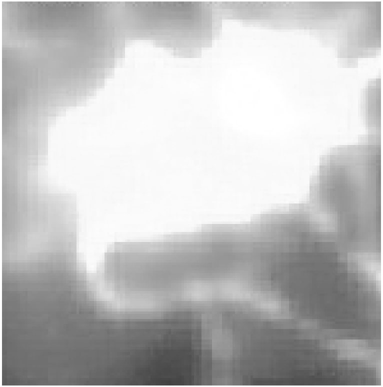


- Microscope receives light from fluorescent molecules
- **Few** molecules are active in each frame $\Rightarrow$ **sparsity**
- Multiple ($\sim$ 10000) frames are recorded and processed individually
- Results from all frames are **combined** to reveal the underlying signal

Single-molecule imaging

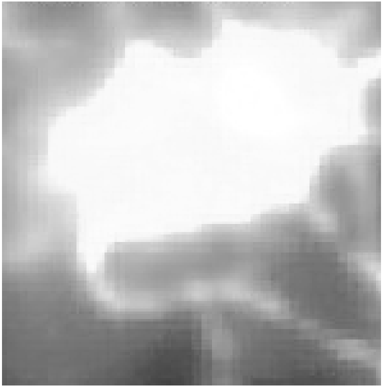


Single-molecule imaging



- **Bad news** : the **resolution** of our measurements is too low

Single-molecule imaging



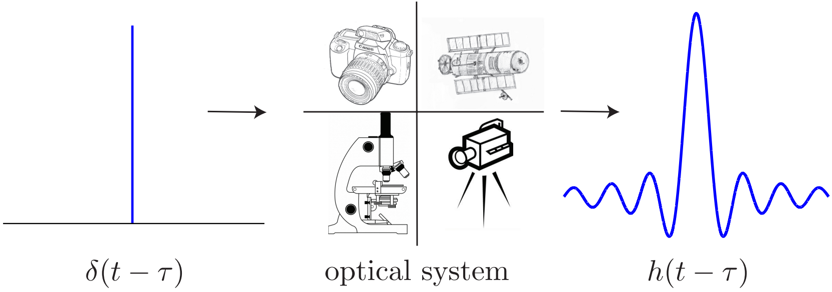
- **Bad news** : the **resolution** of our measurements is too low
- **Good news** : there is **structure** in the signal

Limits of resolution in imaging

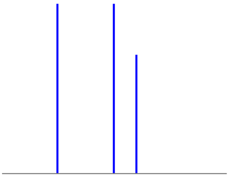
In any optical imaging system **diffraction** imposes a fundamental limit on resolution

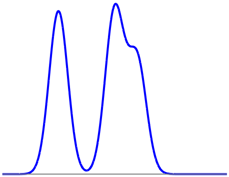
Limits of resolution in imaging

In any optical imaging system **diffraction** imposes a fundamental limit on resolution



What is super-resolution ?





Retrieving fine-scale structure from low-resolution data

What is super-resolution ?

Retrieving fine-scale structure from low-resolution data

Equivalently, extrapolating the high end of the spectrum

Super-resolution of point sources

Super-resolution of structured signals from bandlimited data arises in many applications

Super-resolution of point sources

Super-resolution of structured signals from bandlimited data arises in many applications

- Reconstruction of sub-wavelength structure in conventional imaging
- Reconstruction of sub-pixel structure in electronic imaging
- Spectral analysis of multitone signals
- Similar problems in signal processing, radar, spectroscopy, medical imaging, astronomy, geophysics, etc.

Super-resolution of point sources

Super-resolution of structured signals from bandlimited data arises in many applications

- Reconstruction of sub-wavelength structure in conventional imaging
- Reconstruction of sub-pixel structure in electronic imaging
- Spectral analysis of multitone signals
- Similar problems in signal processing, radar, spectroscopy, medical imaging, astronomy, geophysics, etc.

The signal of interest is often modeled as a **superposition of point sources**

Super-resolution of point sources

Super-resolution of structured signals from bandlimited data arises in many applications

- Reconstruction of sub-wavelength structure in conventional imaging
- Reconstruction of sub-pixel structure in electronic imaging
- Spectral analysis of multitone signals
- Similar problems in signal processing, radar, spectroscopy, medical imaging, astronomy, geophysics, etc.

The signal of interest is often modeled as a **superposition of point sources**

- Celestial bodies in astronomy
- Line spectra in speech analysis
- Fluorescent molecules in single-molecule microscopy

Mathematical model

- Signal : superposition of delta measures

$$x = \sum_j a_j \delta_{t_j} \qquad a_j \in \mathbb{C},\ t_j \in T \subset [0,1]$$

Mathematical model

- Signal : superposition of delta measures

$$x = \sum_j a_j \delta_{t_j} \qquad a_j \in \mathbb{C},\ t_j \in T \subset [0,1]$$

- Measurements : low-pass Fourier coefficients

$$y = \mathcal{F}_n\, x$$

$$y(k) = \int_0^1 e^{-i2\pi kt}\, x\,(\mathrm{d}t) = \sum_j a_j e^{-i2\pi k t_j}, \quad k \in \mathbb{Z},\ |k| \le f_c$$

Number of measurements : $n = 2\, f_c + 1$

Mathematical model

- Signal : superposition of delta measures

$$x = \sum_j a_j \delta_{t_j} \qquad a_j \in \mathbb{C},\ t_j \in T \subset [0,1]$$

- Measurements : low-pass Fourier coefficients

$$y = \mathcal{F}_n\, x$$
$$y(k) = \int_0^1 e^{-i2\pi kt} x\,(\mathrm{d}t) = \sum_j a_j e^{-i2\pi k t_j}, \quad k \in \mathbb{Z},\ |k| \leq f_c$$

Number of measurements : $n = 2\, f_c + 1$

Under what conditions is it possible to recover x from y ?

Equivalent problem : line spectra estimation

Swapping time and frequency

Equivalent problem : line spectra estimation

Swapping time and frequency

- Signal : superposition of sinusoids

$$x(t) = \sum_j a_j e^{i2\pi\omega_j t} \qquad a_j \in \mathbb{C},\ \omega_j \in T \subset [0,1]$$

Equivalent problem : line spectra estimation

Swapping time and frequency

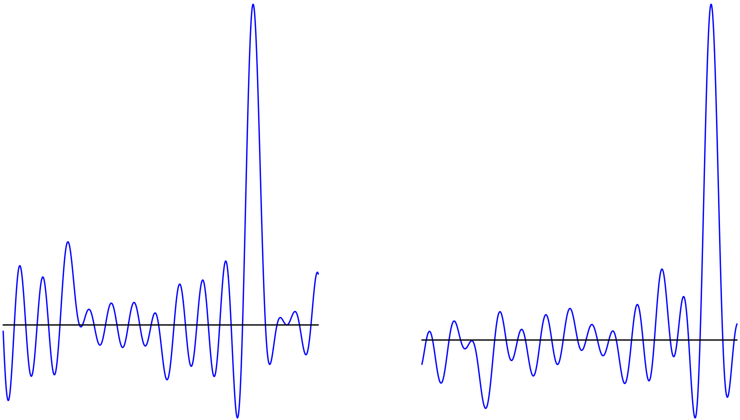
- Signal : superposition of sinusoids

$$x(t) = \sum_j a_j e^{i2\pi\omega_j t} \qquad a_j \in \mathbb{C},\ \omega_j \in T \subset [0,1]$$

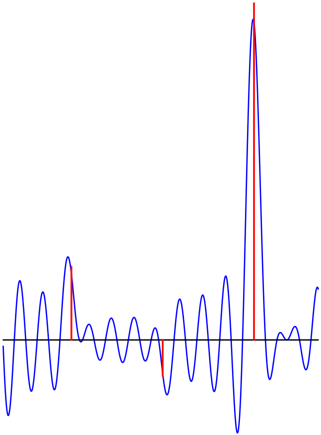
- Measurements : equispaced samples

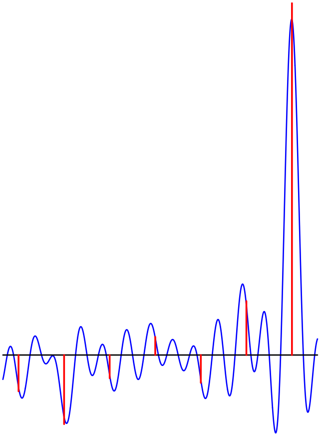
$$x(1), x(2), x(3), \ldots x(n)$$

Can you find the spikes ?



Can you find the spikes?





Compressed sensing

- Compressed sensing theory establishes robust recovery of spikes from **random** Fourier measurements [Candès, Romberg & Tao 2004]

Compressed sensing

- Compressed sensing theory establishes robust recovery of spikes from **random** Fourier measurements [Candès, Romberg & Tao 2004]
- **Crucial insight** : measurement operator is well conditioned when acting upon sparse signals (restricted isometry property)

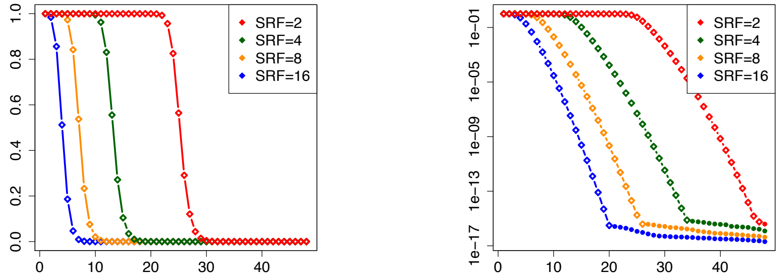
Compressed sensing

- Compressed sensing theory establishes robust recovery of spikes from **random** Fourier measurements [Candès, Romberg & Tao 2004]
- **Crucial insight** : measurement operator is well conditioned when acting upon sparse signals (restricted isometry property)
- Is this the case in the super-resolution setting ?

Compressed sensing

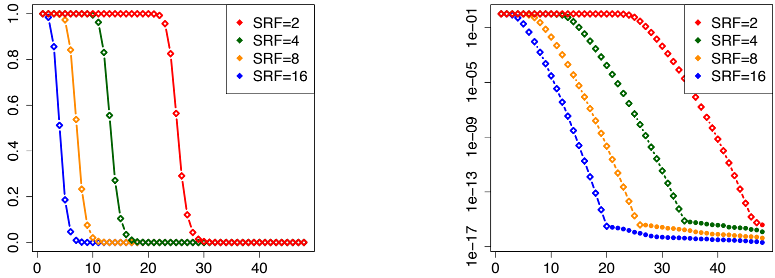
- Compressed sensing theory establishes robust recovery of spikes from **random** Fourier measurements [Candès, Romberg & Tao 2004]
- **Crucial insight** : measurement operator is well conditioned when acting upon sparse signals (restricted isometry property)
- Is this the case in the super-resolution setting ?
- Simple experiment :
  - discretize support to $N = 4096$ points
  - restrict signal to an interval of 48 contiguous points
  - measure n DFT coefficients $\Rightarrow$ super-resolution factor (SRF) $= \frac{N}{n}$
  - how well conditioned is the inverse problem **if we know the support** ?

Simple experiment



For SRF $= 4$ measuring any unit-normed vector in a subspace of dimension 24 results in a signal with norm less than 10^{-7}

Simple experiment



For SRF $= 4$ measuring any unit-normed vector in a subspace of dimension 24 results in a signal with norm less than 10^{-7}

At an SNR of 145 dB, recovery is impossible **by any method**

Prolate spheroidal sequences

There seems to be a sharp transition between signals that are preserved and signals that are completely suppressed.

Prolate spheroidal sequences

There seems to be a sharp transition between signals that are preserved and signals that are completely suppressed.

This phenomenon can be characterized asymptotically by using Slepian's **prolate spheroidal sequences**

Prolate spheroidal sequences

There seems to be a sharp transition between signals that are preserved and signals that are completely suppressed.

This phenomenon can be characterized asymptotically by using Slepian's **prolate spheroidal sequences**

These sequences are the eigenvectors of the concatenation of a time limiting $\mathcal{P}_T$ and a frequency limiting operator $\mathcal{P}_W$, $0 < W \leq 1/2$

Prolate spheroidal sequences

There seems to be a sharp transition between signals that are preserved and signals that are completely suppressed.

This phenomenon can be characterized asymptotically by using Slepian's **prolate spheroidal sequences**

These sequences are the eigenvectors of the concatenation of a time limiting $\mathcal{P}_T$ and a frequency limiting operator $\mathcal{P}_W$, $0 < W \leq 1/2$

Asymptotically WT eigenvalues cluster near one while the rest are almost zero [Slepian 1978]

Consequences of Slepian's work

- For any value of SRF above one, there exist signals almost in the null-space of the measurement operator

Consequences of Slepian's work

- For any value of SRF above one, there exist signals almost in the null-space of the measurement operator
- The norm of the measurements corresponding to k-sparse signals for a fixed SRF decreases *exponentially* in k

Consequences of Slepian's work

- For any value of SRF above one, there exist signals almost in the null-space of the measurement operator
- The norm of the measurements corresponding to k-sparse signals for a fixed SRF decreases *exponentially* in k
- For SRF $= 1.05$ there is a 256-sparse signal such that $||y||_2 \leq 10^{-7}$

Consequences of Slepian's work

- For any value of SRF above one, there exist signals almost in the null-space of the measurement operator
- The norm of the measurements corresponding to k-sparse signals for a fixed SRF decreases *exponentially* in k
- For SRF $= 1.05$ there is a 256-sparse signal such that $||y||_2 \leq 10^{-7}$
- For SRF $= 4$ there is a 48-sparse signal such that $||y||_2 \leq 10^{-33}$

Consequences of Slepian's work

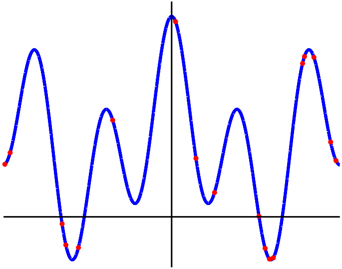
- For any value of SRF above one, there exist signals almost in the null-space of the measurement operator
- The norm of the measurements corresponding to k-sparse signals for a fixed SRF decreases *exponentially* in k
- For SRF $= 1.05$ there is a 256-sparse signal such that $||y||_2 \leq 10^{-7}$
- For SRF $= 4$ there is a 48-sparse signal such that $||y||_2 \leq 10^{-33}$
- For any interval T of length k, there is an irretrievable subspace of dimension $(1 - 1/\text{SRF})\, k$ supported on T

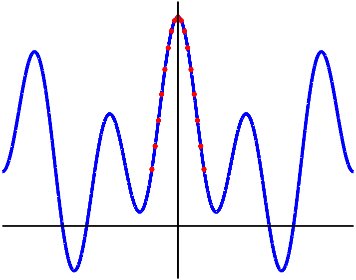
Consequences of Slepian's work

- For any value of SRF above one, there exist signals almost in the null-space of the measurement operator
- The norm of the measurements corresponding to k-sparse signals for a fixed SRF decreases *exponentially* in k
- For SRF $= 1.05$ there is a 256-sparse signal such that $||y||_2 \leq 10^{-7}$
- For SRF $= 4$ there is a 48-sparse signal such that $||y||_2 \leq 10^{-33}$
- For any interval T of length k, there is an irretrievable subspace of dimension $(1 - 1/\text{SRF})\, k$ supported on T
- When SRF > 2 many clustered sparse signals are **killed** by the measurement process

Compressed sensing vs super-resolution

What is the difference ?

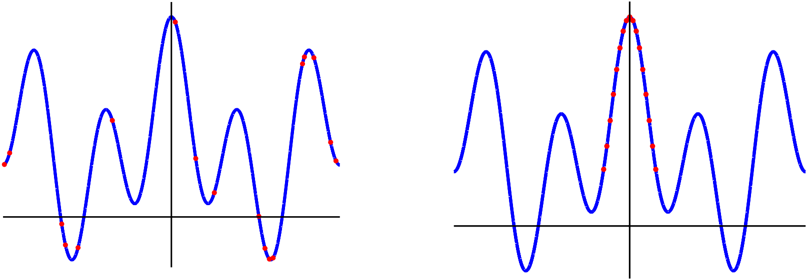




- Compressed sensing : spectrum **interpolation**

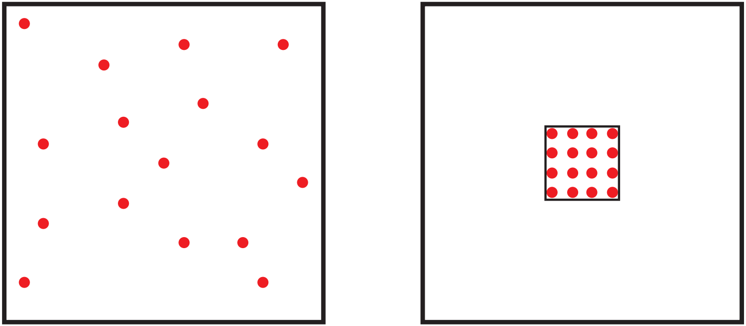
Compressed sensing vs super-resolution

What is the difference ?



- Compressed sensing : spectrum **interpolation**
- Super-resolution : spectrum **extrapolation**

Compressed sensing vs super-resolution



- Compressed sensing : spectrum **interpolation**
- Super-resolution : spectrum **extrapolation**

Conclusion

- Robust super-resolution of arbitrary sparse signals is impossible

Conclusion

- Robust super-resolution of arbitrary sparse signals is impossible
- We can only hope to recover signals that are not too clustered

Conclusion

- Robust super-resolution of arbitrary sparse signals is impossible
- We can only hope to recover signals that are not too clustered
- In this work, this structural assumption is captured by introducing the **minimum separation** of the support T of a signal

$$\Delta(T) = \inf_{(t,t') \in T \,:\, t \neq t'} |t - t'|$$

Noiseless case

Recovery by solving

$$\min_{\tilde{x}} ||\tilde{x}||_{\mathrm{TV}} \quad \text{subject to} \quad \mathcal{F}_n \, \tilde{x} = y,$$

over all finite complex measures $\tilde{x}$ supported on $[0, 1]$

Noiseless case

Recovery by solving

$$\min_{\tilde{x}} ||\tilde{x}||_{\mathrm{TV}} \quad \text{subject to} \quad \mathcal{F}_n \tilde{x} = y,$$

over all finite complex measures $\tilde{x}$ supported on $[0, 1]$

Total-variation norm of a complex measure ν :

$$||\nu||_{\mathrm{TV}} = \sup \sum_{j=1}^{\infty} |\nu(B_j)|,$$

(supremum over all finite partitions B_j of $[0, 1]$)

Noiseless case

Recovery by solving

$$\min_{\tilde{x}} ||\tilde{x}||_{\mathrm{TV}} \quad \text{subject to} \quad \mathcal{F}_n \, \tilde{x} = y,$$

over all finite complex measures $\tilde{x}$ supported on $[0,1]$

Total-variation norm of a complex measure ν :

$$||\nu||_{\mathrm{TV}} = \sup \sum_{j=1}^{\infty} |\nu(B_j)| \,,$$

(supremum over all finite partitions B_j of $[0,1]$)

- **not** the total variation of a piecewise constant function

Noiseless case

Recovery by solving

$$\min_{\tilde{x}} ||\tilde{x}||_{\mathrm{TV}} \quad \text{subject to} \quad \mathcal{F}_n\, \tilde{x} = y,$$

over all finite complex measures $\tilde{x}$ supported on $[0, 1]$

Total-variation norm of a complex measure ν :

$$||\nu||_{\mathrm{TV}} = \sup \sum_{j=1}^{\infty} |\nu(B_j)|,$$

(supremum over all finite partitions B_j of $[0, 1]$)

- **not** the total variation of a piecewise constant function
- continuous counterpart of the ℓ_1 norm

Noiseless case

Recovery by solving

$$\min_{\tilde{x}} ||\tilde{x}||_{\mathrm{TV}} \quad \text{subject to} \quad \mathcal{F}_n\, \tilde{x} = y,$$

over all finite complex measures $\tilde{x}$ supported on $[0,1]$

Total-variation norm of a complex measure ν :

$$||\nu||_{\mathrm{TV}} = \sup \sum_{j=1}^{\infty} |\nu(B_j)|\,,$$

(supremum over all finite partitions B_j of $[0,1]$)

- **not** the total variation of a piecewise constant function
- continuous counterpart of the ℓ_1 norm
- if $\sum_j a_j \delta_{t_j}$ then $||x||_{\mathrm{TV}} = \sum_j |a_j|$

Noiseless recovery

$$y(k) = \int_0^1 e^{-i2\pi kt} x(\mathrm{d}t) = \sum_j a_j e^{-i2\pi kt_j}, \quad k \in \mathbb{Z}, |k| \leq f_c$$

Theorem [Candès, F. 2012]

If the minimum separation of the signal obeys

$$\Delta(T) \geq 2/f_c := 2\lambda_c,$$

then x is recovered exactly by total-variation norm minimization

Noiseless recovery

$$y(k) = \int_0^1 e^{-i2\pi kt} x(\mathrm{d}t) = \sum_j a_j e^{-i2\pi kt_j}, \quad k \in \mathbb{Z}, |k| \le f_c$$

Theorem [Candès, F. 2012]

If the minimum separation of the signal obeys

$$\Delta(T) \ge 2/f_c := 2\lambda_c,$$

then x is recovered exactly by total-variation norm minimization

- Infinite precision

Noiseless recovery

$$y(k) = \int_0^1 e^{-i2\pi kt} x(\mathrm{d}t) = \sum_j a_j e^{-i2\pi kt_j}, \quad k \in \mathbb{Z}, |k| \leq f_c$$

Theorem [Candès, F. 2012]

If the minimum separation of the signal obeys

$$\Delta(T) \geq 2/f_c := 2\lambda_c,$$

then x is recovered exactly by total-variation norm minimization

- Infinite precision
- Recovers $(2\lambda_c)^{-1} = f_c/2 = n/4$ spikes from n low-frequency measurements

Noiseless recovery

$$y(k) = \int_0^1 e^{-i2\pi kt} x(\mathrm{d}t) = \sum_j a_j e^{-i2\pi k t_j}, \quad k \in \mathbb{Z}, |k| \leq f_c$$

Theorem [Candès, F. 2012]

If the minimum separation of the signal obeys

$$\Delta(T) \geq 2/f_c := 2\lambda_c,$$

then x is recovered exactly by total-variation norm minimization

- Infinite precision
- Recovers $(2\lambda_c)^{-1} = f_c/2 = n/4$ spikes from n low-frequency measurements
- If x is real, then

$$\Delta(T) \geq 2/f_c := 1.87\,\lambda_c,$$

Lower bound

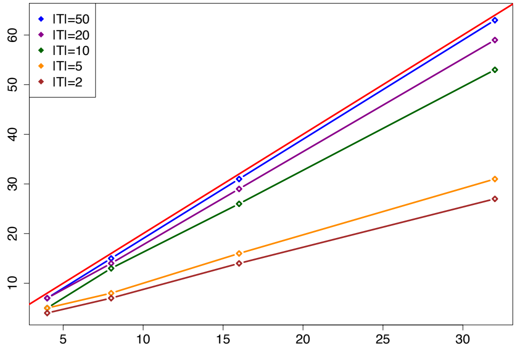
- Consider signals supported on a finite grid of N points

Lower bound

- Consider signals supported on a finite grid of N points
- We look for adversarial sign patterns

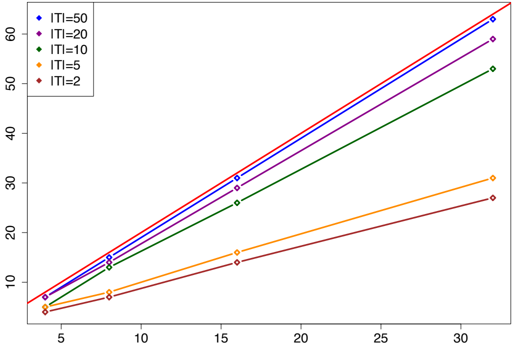
Lower bound

- Consider signals supported on a finite grid of N points
- We look for adversarial sign patterns
- We plot the ratio N/n (x-axis) against the largest minimum distance (y-axis) at which ℓ_1-norm minimization fails



Lower bound

- Consider signals supported on a finite grid of N points
- We look for adversarial sign patterns
- We plot the ratio N/n (x-axis) against the largest minimum distance (y-axis) at which ℓ_1-norm minimization fails



- Red line corresponds to $\Delta(T) = \lambda_c$

Higher dimensions

- Signal :

$$x = \sum_j a_j \delta_{t_j} \qquad a_j \in \mathbb{R},\ t_j \in T \subset [0,1]^2$$

- Measurements : low-pass 2D Fourier coefficients

$$y(k) = \int_{[0,1]^2} e^{-i2\pi \langle k,t \rangle} x(\mathrm{d}t) = \sum_j a_j e^{-i2\pi \langle k,t_j \rangle}, \quad k \in \mathbb{Z}^2,\ |k| \leq f_c$$

Higher dimensions

- Signal :

$$x = \sum_j a_j \delta_{t_j} \qquad a_j \in \mathbb{R},\ t_j \in T \subset [0,1]^2$$

- Measurements : low-pass 2D Fourier coefficients

$$y(k) = \int_{[0,1]^2} e^{-i2\pi\langle k,t\rangle} x\,(\mathrm{d}t) = \sum_j a_j e^{-i2\pi\left\langle k,t_j\right\rangle}, \quad k \in \mathbb{Z}^2,\ |k| \leq f_c$$

Theorem [Candès, F. 2012]

TV norm minimization yields exact recovery if

$$\Delta(T) \geq 2.38\,/f_c := 2.38\,\lambda_c,$$

Higher dimensions

- Signal :

$$x = \sum_j a_j \delta_{t_j} \qquad a_j \in \mathbb{R},\ t_j \in T \subset [0,1]^2$$

- Measurements : low-pass 2D Fourier coefficients

$$y(k) = \int_{[0,1]^2} e^{-i2\pi\langle k,t\rangle} x\,(\mathrm{d}t) = \sum_j a_j e^{-i2\pi\left\langle k,t_j\right\rangle}, \quad k \in \mathbb{Z}^2,\ |k| \le f_c$$

Theorem [Candès, F. 2012]

TV norm minimization yields exact recovery if

$$\Delta(T) \ge 2.38\,/f_c := 2.38\,\lambda_c,$$

In dimension d, $\Delta(T) \ge C_d\,\lambda_c$, where C_d only depends on d

Extensions

- Signal : $\ell - 1$ times continuously differentiable piecewise smooth function

$$x = \sum_{t_j \in T} \mathbf{1}_{(t_{j-1}, t_j)} p_j(t), \ t_j \in T \subset [0, 1]$$

where $p_j(t)$ is a polynomial of degree ℓ
- Measurements : low-pass Fourier coefficients

$$y(k) = \int_0^1 e^{-i2\pi kt} x\,(\mathrm{d}t)\,, \quad k \in \mathbb{Z},\ |k| \leq f_c$$

Extensions

- Signal : $\ell - 1$ times continuously differentiable piecewise smooth function

$$x = \sum_{t_j \in T} \mathbf{1}_{(t_{j-1}, t_j)} p_j(t), \; t_j \in T \subset [0,1]$$

where $p_j(t)$ is a polynomial of degree ℓ

- Measurements : low-pass Fourier coefficients

$$y(k) = \int_0^1 e^{-i2\pi kt} x\,(\mathrm{d}t)\,, \quad k \in \mathbb{Z}, \, |k| \leq f_c$$

- Recovery :

$$\min \|\tilde{x}^{(\ell+1)}\|_{\mathrm{TV}} \quad \text{subject to} \quad \mathcal{F}_n \tilde{x} = y$$

Extensions

- Signal : $\ell - 1$ times continuously differentiable piecewise smooth function
$$x = \sum_{t_j \in T} \mathbf{1}_{(t_{j-1}, t_j)} p_j(t), \ t_j \in T \subset [0,1]$$
where $p_j(t)$ is a polynomial of degree ℓ
- Measurements : low-pass Fourier coefficients
$$y(k) = \int_0^1 e^{-i2\pi k t} x\,(\mathrm{d}t)\,, \quad k \in \mathbb{Z}, \, |k| \le f_c$$
- Recovery :
$$\min \|\tilde{x}^{(\ell+1)}\|_{\mathrm{TV}} \quad \text{subject to} \quad \mathcal{F}_n \tilde{x} = y$$

Corollary

TV norm minimization yields exact recovery if $\Delta(T) \ge 2\,\lambda_c$

Sparse recovery

If we discretize the support :

- Sparse recovery problem in an overcomplete Fourier dictionary

Sparse recovery

If we discretize the support :

- Sparse recovery problem in an overcomplete Fourier dictionary
- Previous theory based on dictionary incoherence is very weak, due to high column correlation

Sparse recovery

If we discretize the support :

- Sparse recovery problem in an overcomplete Fourier dictionary
- Previous theory based on dictionary incoherence is very weak, due to high column correlation
- If $N = 20,000$ and $n = 1,000$, [Dossal 2005] : recovery of 3 spikes, our result : $n/4 = 250$ spikes

Sparse recovery

If we discretize the support :

- Sparse recovery problem in an overcomplete Fourier dictionary
- Previous theory based on dictionary incoherence is very weak, due to high column correlation
- If $N = 20,000$ and $n = 1,000$, [Dossal 2005] : recovery of 3 spikes, our result : $n/4 = 250$ spikes

Additional structural assumptions allow for a more precise theoretical analysis

Sketch of the proof

Sufficient condition for exact recovery of a signal supported on T :

For any $v \in \mathbb{C}^{|T|}$ with $|v_j| = 1$, there exists a low-frequency trigonometric polynomial

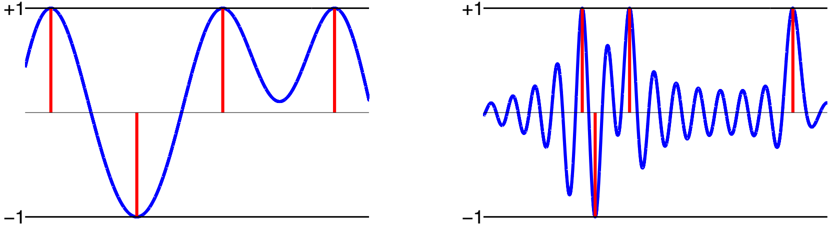
$$q(t) = \sum_{k=-f_c}^{f_c} c_k e^{i2\pi kt} \quad (1)$$

obeying

$$\begin{cases} q(t_j) = v_j, & t_j \in T, \\ |q(t)| < 1, & t \in [0,1] \setminus T. \end{cases} \quad (2)$$

Sketch of the proof

Interpolating the sign pattern with a low frequency polynomial becomes challenging if the minimum separation is small



Sketch of the proof

Interpolation with low-frequency kernel K

$$q(t) = \sum_{t_j \in T} \alpha_j K(t - t_j),$$

where α is a vector of coefficients, and q is constrained to satisfy

$$q(t_k) = \sum_{t_j \in T} \alpha_j K\left(t_k - t_j\right) = v_k, \quad \forall t_k \in T,$$

Sketch of the proof

Interpolation with low-frequency kernel K

$$q(t) = \sum_{t_j \in T} \alpha_j K(t - t_j),$$

where α is a vector of coefficients, and q is constrained to satisfy

$$q(t_k) = \sum_{t_j \in T} \alpha_j K(t_k - t_j) = v_k, \quad \forall t_k \in T,$$

If K is a sinc function this is equivalent to a least squares construction

Sketch of the proof

Interpolation with low-frequency kernel K

$$q(t) = \sum_{t_j \in T} \alpha_j K(t - t_j),$$

where α is a vector of coefficients, and q is constrained to satisfy

$$q(t_k) = \sum_{t_j \in T} \alpha_j K\left(t_k - t_j\right) = v_k, \quad \forall t_k \in T,$$

If K is a sinc function this is equivalent to a least squares construction

Kernels with faster decay, such as the Fejér kernel or the Gaussian kernel [Kahane 2011], yield better results

Sketch of the proof

Interpolation with low-frequency kernel K

$$q(t) = \sum_{t_j \in T} \alpha_j K(t - t_j),$$

where α is a vector of coefficients, and q is constrained to satisfy

$$q(t_k) = \sum_{t_j \in T} \alpha_j K\left(t_k - t_j\right) = v_k, \quad \forall t_k \in T,$$

If K is a sinc function this is equivalent to a least squares construction

Kernels with faster decay, such as the Fejér kernel or the Gaussian kernel [Kahane 2011], yield better results

However, this does **not** allow to construct a valid **continuous** dual polynomial

Sketch of the proof

Adding a correction term

$$q(t) = \sum_{t_j \in T} \alpha_j K(t - t_j) + \beta_j K'(t - t_j)$$

and an extra constraint

$$q(t_k) = v_k, \qquad q'(t_k) = 0, \quad \forall t_k \in T$$

forces $|q|$ to reach a local maximum at each element of T

Sketch of the proof

Adding a correction term

$$q(t) = \sum_{t_j \in T} \alpha_j K(t - t_j) + \beta_j K'(t - t_j)$$

and an extra constraint

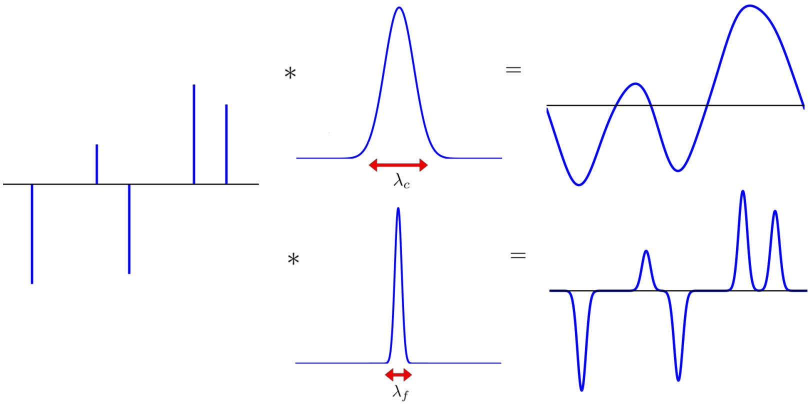
$$q(t_k) = v_k, \qquad q'(t_k) = 0, \quad \forall t_k \in T$$

forces $|q|$ to reach a local maximum at each element of T

Choosing K to be the square of a Fejér kernel allows to show that

- there exist α and β satisfying the constraints
- $|q|$ is strictly bounded by one on the off-support

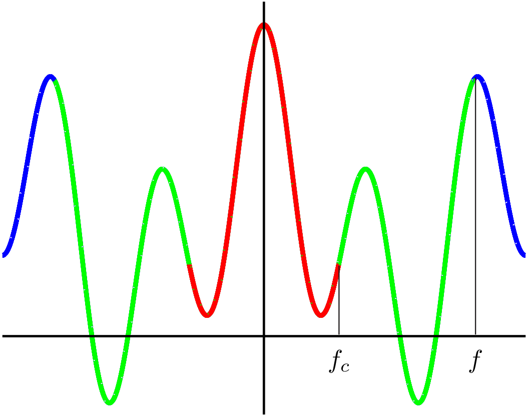
Super-resolution factor : spatial viewpoint



Super-resolution factor

$$\mathrm{SRF} = \frac{\lambda_c}{\lambda_f}$$

Super-resolution factor : spectral viewpoint



Super-resolution factor

$$\text{SRF} = \frac{f}{f_c}$$

Noise model

$$\mathcal{F}_n\, x(k) = \int_0^1 e^{-i2\pi kt} x\,(\mathrm{d}t)\,, \quad |k| \leq f_c$$

Noise model

$$\mathcal{F}_n\, x(k) = \int_0^1 e^{-i2\pi k t} x\,(\mathrm{d}t)\,, \quad |k| \leq f_c$$

Noise can be modeled in the spectral domain

$$y = \mathcal{F}_n\, x + w$$

Noise model

$$\mathcal{F}_n\, x(k) = \int_0^1 e^{-i2\pi kt} x\,(\mathrm{d}t)\,, \quad |k| \leq f_c$$

Noise can be modeled in the spectral domain

$$y = \mathcal{F}_n\, x + w$$

or spatial domain

$$s = \mathcal{F}_n^* y = \mathcal{P}_n\, x + z,$$

$\mathcal{P}_n$ projection onto the first n Fourier modes

Noise model

$$\mathcal{F}_n\, x(k) = \int_0^1 e^{-i2\pi kt} x\,(\mathrm{d}t)\,, \quad |k| \leq f_c$$

Noise can be modeled in the spectral domain

$$y = \mathcal{F}_n\, x + w$$

or spatial domain

$$s = \mathcal{F}_n^* y = \mathcal{P}_n\, x + z,$$

$\mathcal{P}_n$ projection onto the first n Fourier modes

Assumption :

$$||z||_{L_1} \leq \delta$$

Noise model

$$\mathcal{F}_n\, x(k) = \int_0^1 e^{-i2\pi kt} x\,(\mathrm{d}t)\,, \quad |k| \leq f_c$$

Noise can be modeled in the spectral domain

$$y = \mathcal{F}_n\, x + w$$

or spatial domain

$$s = \mathcal{F}_n^* y = \mathcal{P}_n\, x + z,$$

$\mathcal{P}_n$ projection onto the first n Fourier modes

Assumption :

$$||z||_{L_1} \leq \delta$$

Recovery algorithm

$$\min_{\tilde{x}} \; ||\tilde{x}||_{\mathrm{TV}} \quad \text{subject to} \quad ||\mathcal{P}_n \tilde{x} - s||_{L_1} \leq \delta$$

Robust recovery

Let ϕ_{λ_c} be a kernel with width λ_c and cut-off frequency f_c

If $||z||_{L_1} \leq \delta$, then

$$||\phi_{\lambda_c} * (x_{\text{est}} - x)||_{L_1} \approx ||z||_{L_1} \leq \delta,$$

Robust recovery

Let ϕ_{λ_c} be a kernel with width λ_c and cut-off frequency f_c

If $||z||_{L_1} \leq \delta$, then

$$||\phi_{\lambda_c} * (x_{\text{est}} - x)||_{L_1} \approx ||z||_{L_1} \leq \delta,$$

What can we expect for ϕ_{λ_f} ?

Robust recovery

Let ϕ_{λ_c} be a kernel with width λ_c and cut-off frequency f_c

If $||z||_{L_1} \leq \delta$, then

$$||\phi_{\lambda_c} * (x_{\text{est}} - x)||_{L_1} \approx ||z||_{L_1} \leq \delta,$$

What can we expect for ϕ_{λ_f} ?

Theorem [Candès, F. 2012]

If $\Delta(T) \geq 2/f_c$ then the solution x_{est} to the TV-norm minimization problem satisfies

$$\left|\left|\phi_{\lambda_f} * (x_{\text{est}} - x)\right|\right|_{L_1} \lesssim \text{SRF}^2\,\delta,$$

Practical implementation

- TV norm minimization is an infinite-dimensional optimization problem

Practical implementation

- TV norm minimization is an infinite-dimensional optimization problem
- Primal :

$$\min_{\tilde{x}} ||\tilde{x}||_{\mathrm{TV}} \quad \text{subject to} \quad \mathcal{F}_n \, \tilde{x} = y,$$

First option : Discretize $x \Rightarrow \ell_1$ norm minimization

Practical implementation

- TV norm minimization is an infinite-dimensional optimization problem
- Primal :

$$\min_{\tilde{x}} ||\tilde{x}||_{\mathrm{TV}} \quad \text{subject to} \quad \mathcal{F}_n \tilde{x} = y,$$

First option : Discretize $x \Rightarrow \ell_1$ norm minimization

- Dual :

$$\max_{u \in \mathbb{C}^n} \mathrm{Re}\left[y^* u\right] \quad \text{subject to} \quad ||\mathcal{F}_n^* u||_\infty \leq 1,$$

Second option : Recast as semidefinite program

Semidefinite representation

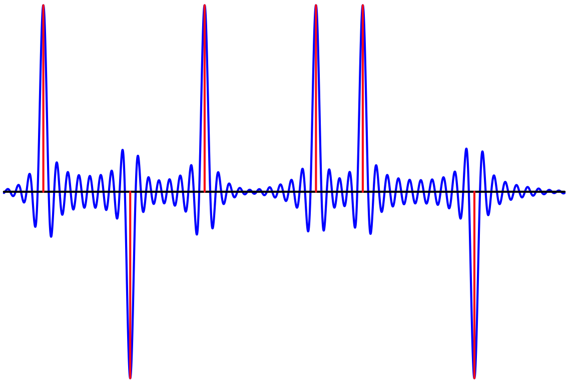
$$||\mathcal{F}_n^* \, u||_\infty \leq 1$$

is equivalent to

There exists a Hermitian matrix $Q \in \mathbb{C}^{n \times n}$ such that

$$\begin{bmatrix} Q & u \\ u^* & 1 \end{bmatrix} \succeq 0, \qquad \sum_{i=1}^{n-j} Q_{i,i+j} = \begin{cases} 1, & j = 0, \\ 0, & j = 1, 2, \ldots, n-1. \end{cases}$$

Support detection



By strong duality, $\mathcal{F}_n^* \, \hat{u}$ interpolates the sign of $\hat{x}$

Experiment

Support recovery by solving the SDP

f_c	25	50	75	100
Average error	$6.66\,10^{-9}$	$1.70\,10^{-9}$	$5.58\,10^{-10}$	$2.96\,10^{-10}$
Maximum error	$1.83\,10^{-7}$	$8.14\,10^{-8}$	$2.55\,10^{-8}$	$2.31\,10^{-8}$

For each f_c, 100 random signals with $|T| = f_c/4$ and $\Delta(T) \geq 2/f_c$

Related work

- Super-resolution of spike trains by ℓ_1 norm minimization in seismic prospecting [Claerbout, Muir 1973 ; Levy, Fullagar 1981 ; Santosa, Symes 1986]

Related work

- Super-resolution of spike trains by ℓ_1 norm minimization in seismic prospecting [Claerbout, Muir 1973 ; Levy, Fullagar 1981 ; Santosa, Symes 1986]
- Guarantees on recovery of positive k-sparse signals from $2k + 1$ Fourier coefficients by ℓ_1 norm minimization [Donoho, Tanner ; Fuchs 2005]

Related work

- Super-resolution of spike trains by ℓ_1 norm minimization in seismic prospecting [Claerbout, Muir 1973 ; Levy, Fullagar 1981 ; Santosa, Symes 1986]
- Guarantees on recovery of positive k-sparse signals from $2k+1$ Fourier coefficients by ℓ_1 norm minimization [Donoho, Tanner ; Fuchs 2005]
- Recovery of lacunary Fourier series coefficients [Kahane 2011]

Related work

- Super-resolution of spike trains by ℓ_1 norm minimization in seismic prospecting [Claerbout, Muir 1973 ; Levy, Fullagar 1981 ; Santosa, Symes 1986]
- Guarantees on recovery of positive k-sparse signals from $2k+1$ Fourier coefficients by ℓ_1 norm minimization [Donoho, Tanner ; Fuchs 2005]
- Recovery of lacunary Fourier series coefficients [Kahane 2011]
- Finite rate of innovation [Dragotti, Vetterli, Blu 2007]

Related work

- Super-resolution of spike trains by ℓ_1 norm minimization in seismic prospecting [Claerbout, Muir 1973 ; Levy, Fullagar 1981 ; Santosa, Symes 1986]
- Guarantees on recovery of positive k-sparse signals from $2k+1$ Fourier coefficients by ℓ_1 norm minimization [Donoho, Tanner ; Fuchs 2005]
- Recovery of lacunary Fourier series coefficients [Kahane 2011]
- Finite rate of innovation [Dragotti, Vetterli, Blu 2007]
- Bounds on the modulus of continuity of super-resolution [Donoho 1992]

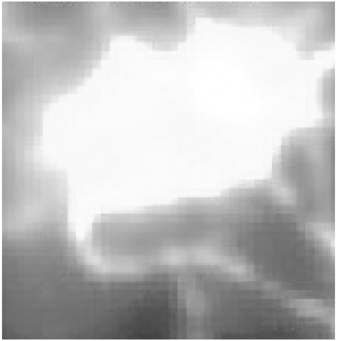
Related work

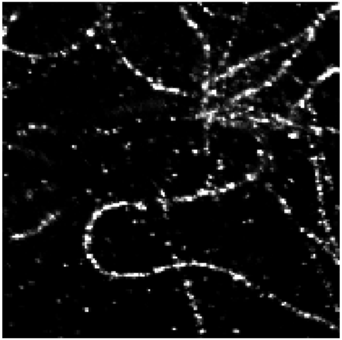
- Super-resolution of spike trains by ℓ_1 norm minimization in seismic prospecting [Claerbout, Muir 1973 ; Levy, Fullagar 1981 ; Santosa, Symes 1986]
- Guarantees on recovery of positive k-sparse signals from $2k + 1$ Fourier coefficients by ℓ_1 norm minimization [Donoho, Tanner ; Fuchs 2005]
- Recovery of lacunary Fourier series coefficients [Kahane 2011]
- Finite rate of innovation [Dragotti, Vetterli, Blu 2007]
- Bounds on the modulus of continuity of super-resolution [Donoho 1992]
- Super-resolution by greedy methods [Fannjiang, Liao 2012]

Related work

- Super-resolution of spike trains by ℓ_1 norm minimization in seismic prospecting [Claerbout, Muir 1973 ; Levy, Fullagar 1981 ; Santosa, Symes 1986]
- Guarantees on recovery of positive k-sparse signals from $2k+1$ Fourier coefficients by ℓ_1 norm minimization [Donoho, Tanner ; Fuchs 2005]
- Recovery of lacunary Fourier series coefficients [Kahane 2011]
- Finite rate of innovation [Dragotti, Vetterli, Blu 2007]
- Bounds on the modulus of continuity of super-resolution [Donoho 1992]
- Super-resolution by greedy methods [Fannjiang, Liao 2012]
- Random undersampling of low-frequency coefficients [Tang et al 2012]

Conclusion





Stable super-resolution is possible via tractable non-parametric methods based on convex optimization

Note on single-molecule imaging : Joint work with E. Candès, V. Morgenshtern and the Moerner lab at Stanford

Thank you