Towards a Mathematical Theory of Super-Resolution

Emmanuel Candès

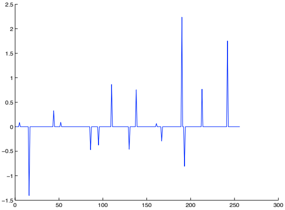
Optimization and Statisitical Learning, Les Houches, January 2013

Collaborator

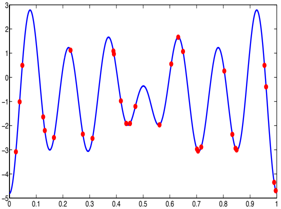
Carlos Fernandez-Granda (Stanford, EE)

Prelude: Compressed Sensing

Some origin

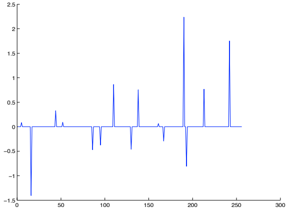


sparse signal

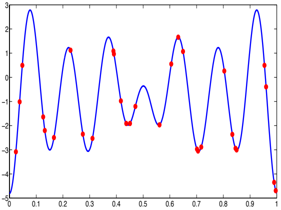


sample spectrum at random

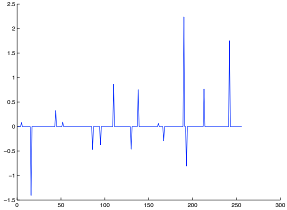
Some origin



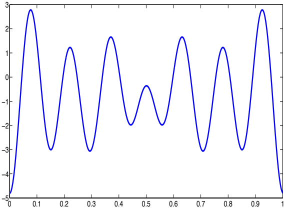
sparse signal



sample spectrum at random



min $\ell_1 \to$ exact



min $\ell_1 \to$ exact interpolation

An early result

- $x \in \mathbb{C}^N$
- Discrete Fourier transform

$$\hat{x}[\omega] = \sum_{t=0}^{N-1} x[t] e^{-i2\pi\omega t/N} \quad \omega = 0, 1, \ldots, N-1$$

An early result

- $x \in \mathbb{C}^N$
- Discrete Fourier transform

$$\hat{x}[\omega] = \sum_{t=0}^{N-1} x[t] e^{-i2\pi\omega t/N} \quad \omega = 0, 1, \ldots, N-1$$

Theorem (C., Romberg and Tao (04))

- *x: k-sparse*
- *n Fourier coefficients selected at random*

$$\ell_1 \text{ is exact if } n \gtrsim k \log N$$

An early result

- $x \in \mathbb{C}^N$
- Discrete Fourier transform

$$\hat{x}[\omega] = \sum_{t=0}^{N-1} x[t] e^{-i2\pi\omega t/N} \quad \omega = 0, 1, \ldots, N-1$$

Theorem (C., Romberg and Tao (04))

- *x: k-sparse*
- *n Fourier coefficients selected at random*

$$\ell_1 \text{ is exact if } n \gtrsim k \log N$$

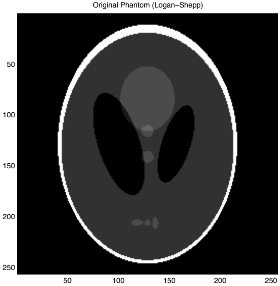
Extensions: C. and Plan (10)

- Can deal with noise (in essentially optimal way)
- Can deal with approximate sparsity

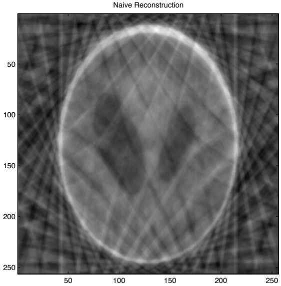
Other works: Donoho (04)

Extensions: reconstruction from undersampled freq. data

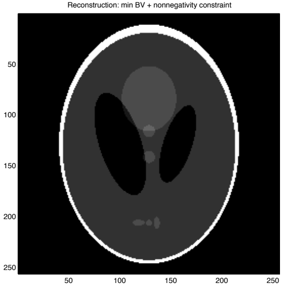
Minimize ℓ_1 norm of gradient subject to data constraints



original

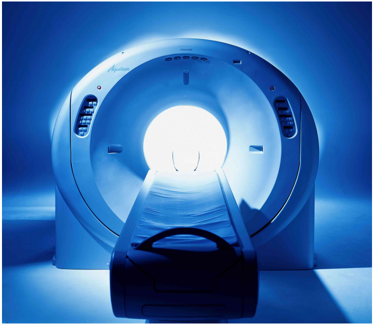


filtered backprojection



min $\ell_1 \rightarrow$ perfect

Magnetic resonance imaging



Acquire data by scanning in Fourier space

Impact on MR pediatrics

Lustig (UCB), Pauly, Vasanawala (Stanford)

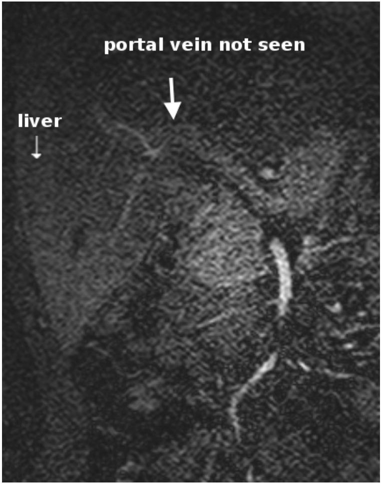
Parallel imaging (PI)

Compressed sensing + PI

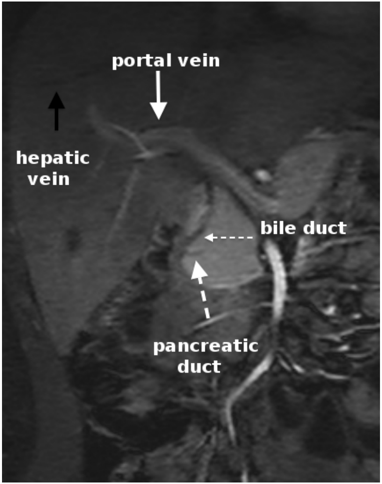
6 year old male abdomen: 8X acceleration

Impact on MR pediatrics

Lustig (UCB), Pauly, Vasanawala (Stanford)



Parallel imaging (PI)



Compressed sensing + PI

6 year old male abdomen: 8X acceleration

Agenda

Compressed sensing: Nyquist sampling is irrelevant

- Can sample at will/random
- Cvx opt. solves an interpolation problem exactly under sparsity constraints
- Robust to noise
- Essentially discrete and finite time theory: exceptions
  - Eldar et al.
  - Adcock, Hansen et al.

Agenda

Compressed sensing: Nyquist sampling is irrelevant

- Can sample at will/random
- Cvx opt. solves an interpolation problem exactly under sparsity constraints
- Robust to noise
- Essentially discrete and finite time theory: exceptions
  - Eldar et al.
  - Adcock, Hansen et al.

This lecture: super-resolution

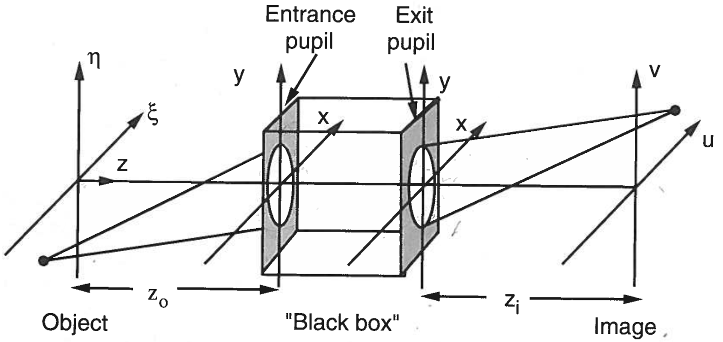
- Can only sample low frequencies
- Cvx opt solves an extrapolation problem exactly under sparsity constraints
- Some robustness (sometimes) to noise
- Continuous time theory

Motivation

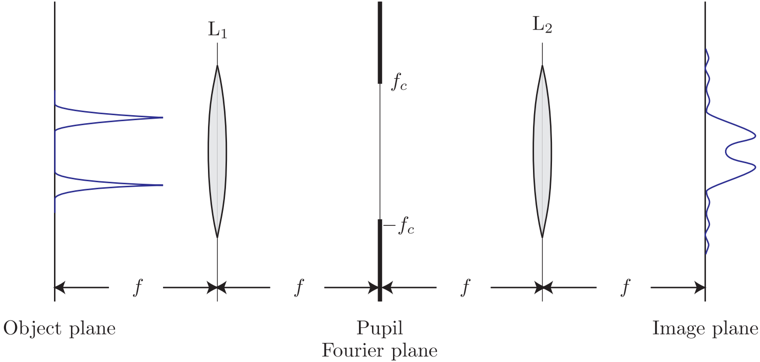
Diffraction limited systems

The physical phenomenon called diffraction *is of the utmost importance in the theory of optical imaging systems*

Joseph Goodman



Diffraction limited systems: canonical example



4f optical system

Mathematical model

$$f_{\mathsf{obs}}(t) = (h * f)(t)$$
$$\hat{f}_{\mathsf{obs}}(\omega) = \hat{h}(\omega)\hat{f}(\omega)$$

h : point spread function (PSF)

$\hat{h}$: transfer function (TF)

Bandlimited imaging systems

Bandlimited system

$$|\omega| > \Omega \quad \Rightarrow \quad |\hat{h}(\omega)| = 0$$

$\hat{f}_{\mathsf{obs}}(\omega) = \hat{h}(\omega)\,\hat{f}(\omega) \to$ suppresses *all* high-frequency components

Bandlimited imaging systems

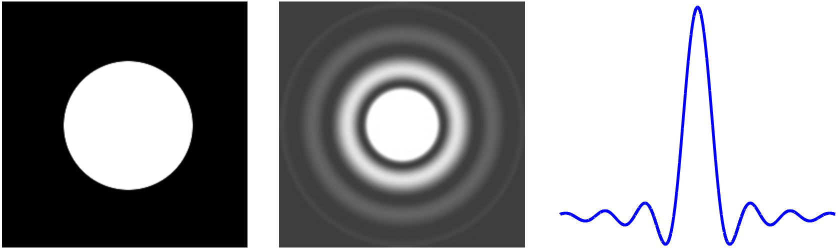
Bandlimited system

$$|\omega| > \Omega \quad \Rightarrow \quad |\hat{h}(\omega)| = 0$$

$\hat{f}_{\text{obs}}(\omega) = \hat{h}(\omega)\,\hat{f}(\omega) \rightarrow$ suppresses *all* high-frequency components

Example: coherent imaging

$$\hat{h}(\omega) = 1_P(\omega) \qquad \text{indicator of pupil element}$$

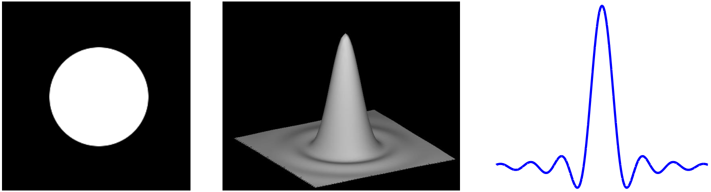


TF
Pupil

PSF
Airy disk

cross-section (PSF)

Examples



TF PSF cross-section (PSF)

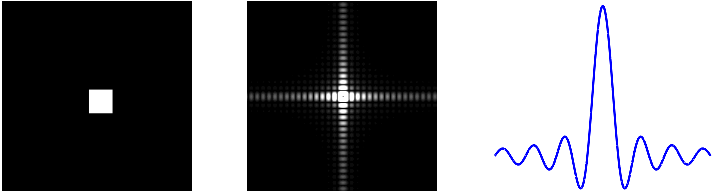
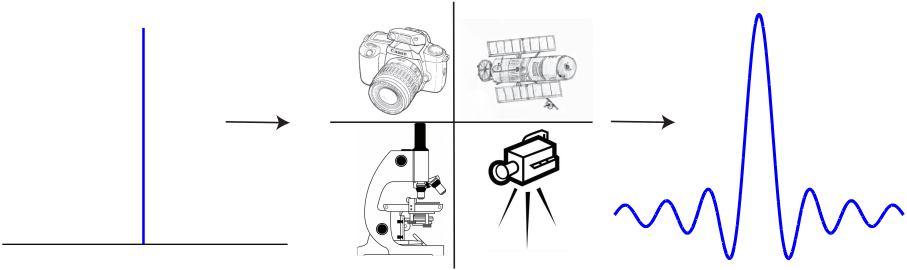
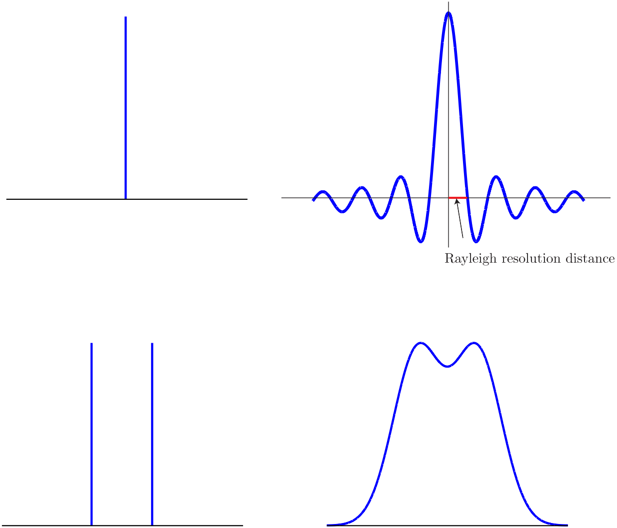


Image of point source



$\delta(t-\tau)$ optical system $h(t-\tau)$

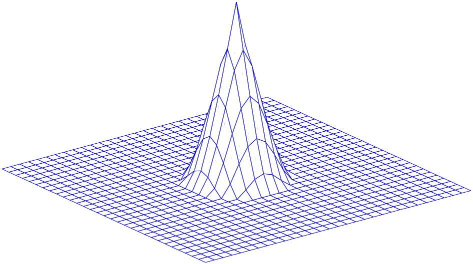
Rayleigh resolution limit



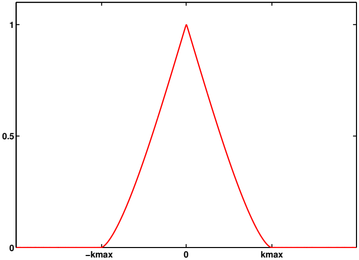
Lord Rayleigh

Incoherent imaging

$$I_{\mathsf{obs}} = I * h_{\mathsf{inc}} \qquad h_{\mathsf{inc}}(t) = |h_{\mathsf{coh}}(t)|^2$$



2D TF



cross section (TF)

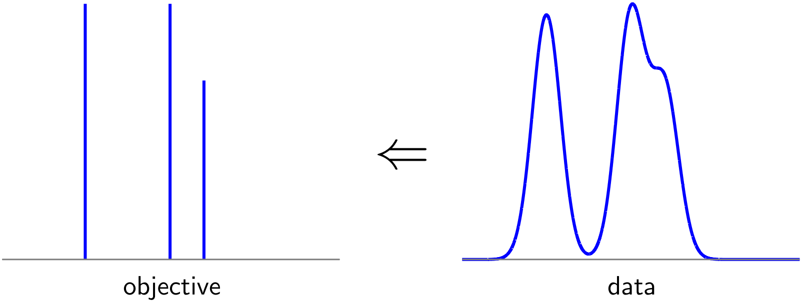
Other examples of low-pass data

$$f_{\text{obs}} = f * h \qquad h \text{ bandlimited}$$

- out-of-focus blur
- atmospheric turbulence blur
- motion blur
- near-field accoustic holography
- ...

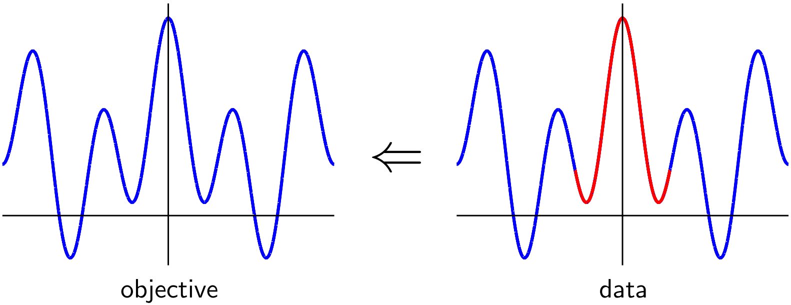
The Super-Resolution Problem

Super-resolution: spatial viewpoint



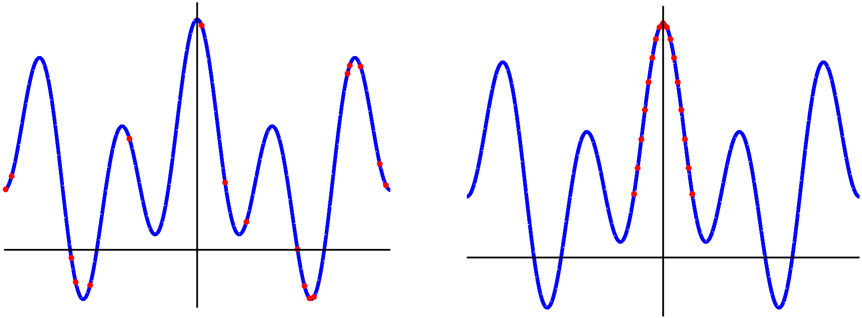
ill-posed deconvolution to break the diffraction limit

Super-resolution: frequency viewpoint



ill-posed extrapolation

Random vs. low-frequency sampling: 1D

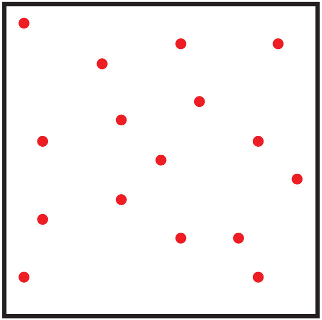


Random sampling (CS)

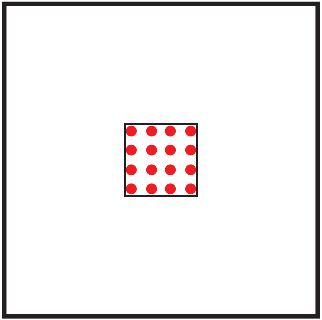
Low-frequency sampling (SR)

Very different from compressive sensing (CS)

Random vs. low-frequency sampling: 2D



Random sampling (CS)



Low-frequency sampling (SR)

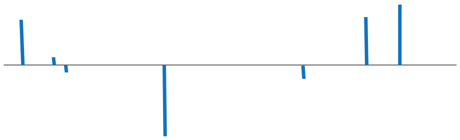
Very different from compressive sensing (CS)

A Mathematical Theory of Super-resolution

Mathematical model

- Signal:

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



- Data: $n = 2f_c + 1$ low-frequency coefficients (Nyquist sampling)

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

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

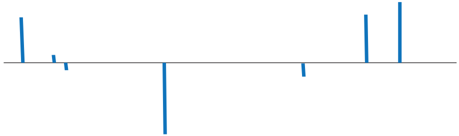
- Resolution limit: ($\lambda_c/2$ is Rayleigh distance)

$$1/f_c = \lambda_c$$

Mathematical model

- Signal:

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



- Data: $n = 2f_c + 1$ low-frequency coefficients (Nyquist sampling)

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

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

- Resolution limit: ($\lambda_c/2$ is Rayleigh distance)

$$1/f_c = \lambda_c$$

Question

Can we resolve the signal beyond this limit?

Equivalent problem: spectral estimation

Swap time and frequency

- Signal
$$x(t) = \sum_j a_j e^{i2\pi\omega_j t} \qquad a_j \in \mathbb{C},\ \omega_j \in [0,1]$$
- Observe samples $x(0), x(1), \ldots, x(n-1)$

Equivalent problem: spectral estimation

Swap time and frequency

- Signal

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

- Observe samples $x(0), x(1), \ldots, x(n-1)$

Question

Can we resolve the frequencies beyond the Heisenberg limit?

Recovery by minimum total-variation

Recover signal by solving

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

Total-variation norm: '$\|x\|_{\mathsf{TV}} = \int |x(\mathrm{d}t)|$'

- Continuous analog of ℓ_1 norm
- If $x = \sum_j a_j \delta_{\tau_j}$, $\|x\|_{\mathsf{TV}} = \sum_j |a_j|$
- If x absolutely continuous wrt Lebesgue, $\|x\|_{\mathsf{TV}} = \int |x(t)| \mathrm{d}t$

Noiseless recovery: main result

$$y(k) = \int_0^1 e^{-i2\pi kt} x(\mathsf{d}t) \qquad |k| \leq f_c$$

Min distance $\quad \Delta(T) = \inf_{(t,t') \in T \,:\, t \neq t'} |t - t'|_\infty \qquad T \subset [0,1]$

Noiseless recovery: main result

$$y(k) = \int_0^1 e^{-i2\pi kt} x(\mathsf{d}t) \qquad |k| \le f_c$$

Min distance $\qquad \Delta(T) = \inf_{(t,t') \in T \,:\, t \neq t'} |t - t'|_\infty \qquad T \subset [0,1]$

Theorem (C. and Fernandez Granda (2012))

If support T of x obeys

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

then min TV solution is exact! For real-valued x, a min dist. of $1.87\lambda_c$ suffices

- Infinite precision!

Noiseless recovery: main result

$$y(k) = \int_0^1 e^{-i2\pi kt} x(\mathrm{d}t) \qquad |k| \le f_c$$

Min distance $\quad \Delta(T) = \inf_{(t,t') \in T \,:\, t \neq t'} |t - t'|_\infty \qquad T \subset [0,1]$

Theorem (C. and Fernandez Granda (2012))

If support T of x obeys

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

then min TV solution is exact! For real-valued x, a min dist. of $1.87\lambda_c$ suffices

- Infinite precision!
- Whatever the amplitudes!

Noiseless recovery: main result

$$y(k) = \int_0^1 e^{-i2\pi kt} x(\mathsf{d}t) \qquad |k| \le f_c$$

Min distance $\qquad \Delta(T) = \inf_{(t,t') \in T \,:\, t \neq t'} |t - t'|_\infty \qquad T \subset [0,1]$

Theorem (C. and Fernandez Granda (2012))

If support T of x obeys

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

then min TV solution is exact! For real-valued x, a min dist. of $1.87\lambda_c$ suffices

- Infinite precision!
- Whatever the amplitudes!
- Can recover $(2\lambda_c)^{-1} = f_c/2 = n/4$ spikes from n low-freq. samples!

Noiseless recovery: main result

$$y(k) = \int_0^1 e^{-i2\pi kt} x(\mathrm{d}t) \qquad |k| \le f_c$$

Min distance $\quad \Delta(T) = \inf_{(t,t') \in T \,:\, t \neq t'} |t - t'|_\infty \qquad T \subset [0,1]$

Theorem (C. and Fernandez Granda (2012))

If support T of x obeys

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

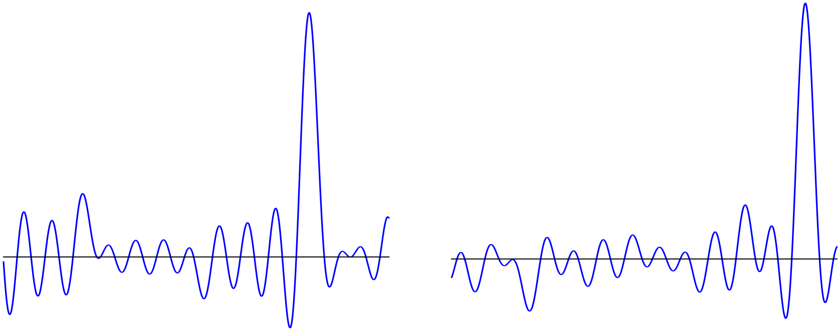
then min TV solution is exact! For real-valued x, a min dist. of $1.87\lambda_c$ suffices

- Infinite precision!
- Whatever the amplitudes!
- Can recover $(2\lambda_c)^{-1} = f_c/2 = n/4$ spikes from n low-freq. samples!

- Have a proof for $1.85\lambda_c$
- Can be improved (but not much)

Flooded spikes

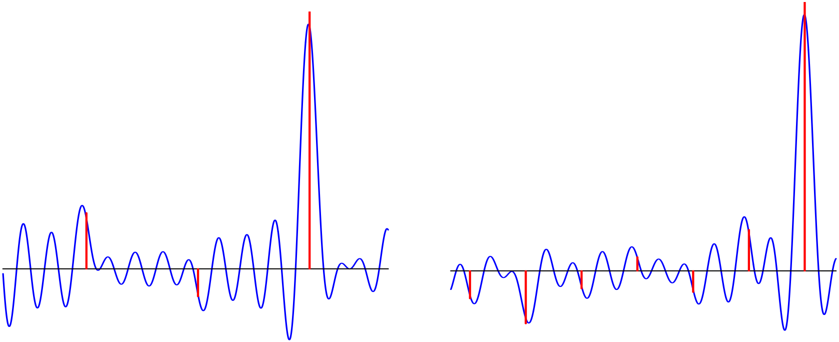
- Sparse spike train obeys min distance assumption
- Low-frequency data



Where are the spikes?

Flooded spikes

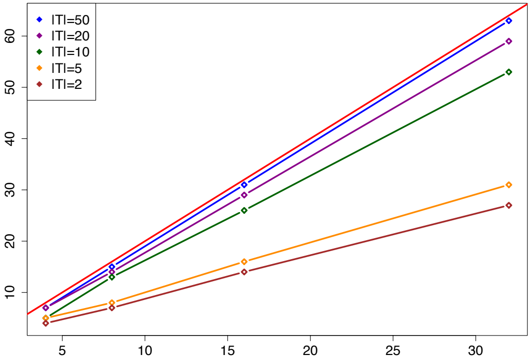
- Sparse spike train obeys min distance assumption
- Low-frequency data



Where are the spikes?

Lower bound

- Put $k = |T|$ spikes on an equispaced grid at fixed distance
- Search for amplitudes s. t. ℓ_1 fails



Min distances at which exact recovery by ℓ_1 min fails to occur against $\lambda_c/2$
At red curve, min distance would be exactly equal to λ_c
ℓ_1 fails if distance is below λ_c

Super-resolution in higher dimensions

- Signal

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

- Data: low-frequency coefficients (Nyquist sampling)

$$y(k) = \int_{[0,1]^2} e^{-i2\pi\langle k,t\rangle} x(\mathrm{d}t) = \sum_j a_j e^{-i2\pi\langle k,t_j\rangle} \qquad \begin{array}{l} k = (k_1,k_2) \in \mathbb{Z}^2 \\ |k_1|,|k_2| \le f_c \end{array}$$

- Resolution limit: $1/f_c = \lambda_c$

Super-resolution in higher dimensions

- Signal
$$x = \sum_j a_j \delta_{\tau_j} \qquad a_j \in \mathbb{C},\ \tau_j \in T \subset [0,1]^2$$
- Data: low-frequency coefficients (Nyquist sampling)
$$y(k) = \int_{[0,1]^2} e^{-i2\pi\langle k,t\rangle} x(\mathrm{d}t) = \sum_j a_j e^{-i2\pi\langle k,t_j\rangle} \qquad \begin{array}{l} k = (k_1,k_2) \in \mathbb{Z}^2 \\ |k_1|, |k_2| \le f_c \end{array}$$
- Resolution limit: $1/f_c = \lambda_c$

Theorem (C. and Fernandez Granda (2012))

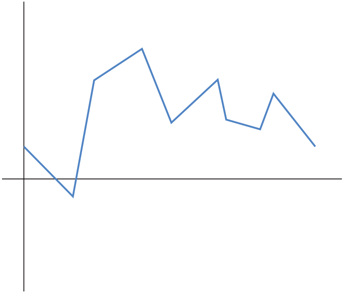
If support T of x obeys
$$\Delta(T) \ge 2.38\,\lambda_c$$
then min TV solution is exact!

Extensions

- Signal x is periodic and piecewise smooth

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

  - p_j polynomial of degree ℓ
  - x is $\ell - 1$ times continuously differentiable



- Data

$$y = \mathcal{F}_n x \quad y_k = \int_{[0,1]} x(t)\, e^{-i2\pi kt} \mathrm{d}t \quad |k| \leq f_c$$

- Recovery

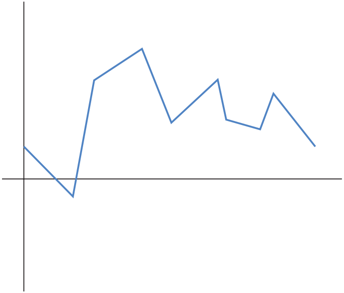
$$\min \|\tilde{x}^{(\ell+1)}\|_{\mathsf{TV}} \quad \text{subject to} \quad \mathcal{F}_n \tilde{x} = y$$

Extensions

- Signal x is periodic and piecewise smooth

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

  - p_j polynomial of degree ℓ
  - x is $\ell - 1$ times continuously differentiable



- Data

$$y = \mathcal{F}_n x \quad y_k = \int_{[0,1]} x(t)\, e^{-i2\pi kt} \mathrm{d}t \quad |k| \le f_c$$

- Recovery

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

Corollary

Under same assumptions, min TV solution is exact

Surprise: extreme coherence

$$\min \|\tilde{x}\|_{(\ell_1,\mathsf{TV})} \quad \text{subject to} \quad y = \mathcal{F}_n x$$

- $\mathcal{F}_n$ is $n \times \infty$ matrix with (normalized) column vectors indexed by time/space

$$f_t[k] = n^{-1/2} e^{i2\pi kt} \qquad |k| \le f_c$$

- Coherence is one! $\langle f_t, f_t' \rangle \to 1$ as $t' \to t$
- Yet perfect recovery!

Completely unexplained by current sparse recovery literature (which cannot deal with more than one spike)

Kahane's result

- $x \in \mathbb{C}^N$ with spacing $1/N$
- observe n low-frequency samples from DFT

Kahane (2011). Min ℓ_1 is exact if min separation obeys

$$\Delta(T) \geq 10\,\frac{1}{n}\sqrt{\log(N/n)}$$

Cannot pass to the continuum

Proof ideas

Recovery of x supported on $T \subset [0,1]$ exact if for any $v \in \mathbb{C}^{|T|}$ with $|v_j| = 1$ $\exists$

$$q(t) = \sum_{k=-f_c}^{f_c} c_k e^{i2\pi kt} \qquad \begin{cases} q(t_j) = v_j & t_j \in T \\ |q(t)| < 1, & t \in [0,1] \setminus T \end{cases}$$

low-freq. trig. polynomial

interpolating

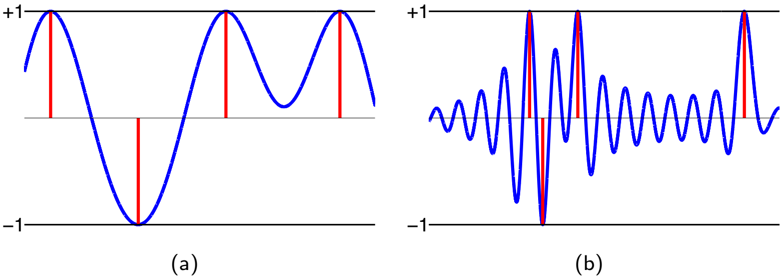


Figure: (a) separated spikes (b) clustered spikes

Construction of dual polynomial

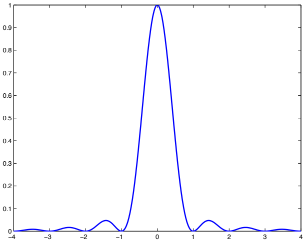
- Squared Fejér kernel

$$K(t) = \left[\frac{\sin\left(\left(\frac{f_c}{2}+1\right)\pi t\right)}{\left(\frac{f_c}{2}+1\right)\sin(\pi t)} \right]^4$$

Fourier coefficients of K supported on $\{-f_c, -f_c+1, \ldots, f_c\}$

- Dual polynomial

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



Fejér kernel

- Fit coefficients α, β so that for $t_j \in T$

$$\begin{cases} q(t_j) = v_j \\ q'(t_j) = 0 \end{cases}$$

- Proof: show this is well defined and $|q(t)| < 1$ on T^c

Other works and approaches to super-resolution

- Donoho ('89) [modulus of continuity under sparsity constraints]
- Eckhoff ('95) [algebraic approach to find singularities from first few freq. coeff.]
- Dragotti, Vetterli, Blu ('07) [algebraic approach, De Prony's method]
- Batenkov and Yomdin ('12) [algebraic approach]

Numerical Algorithms?

Formulation as a finite-dimensional problem

Primal problem

$$\min \ \|x\|_{\mathsf{TV}} \text{ s. t. } \mathcal{F}_n x = y$$

- Infinite-dimensional variable x
- Finitely many constraints

Dual problem

$$\max \ \mathrm{Re}\langle y, c\rangle \text{ s. t. } \|\mathcal{F}_n^* c\|_\infty \leq 1$$

- Finite-dimensional variable c
- Infinitely many constraints

$$(\mathcal{F}_n^* \, c)(t) = \sum_{|k| \leq f_c} c_k e^{i2\pi k t}$$

Formulation as a finite-dimensional problem

Primal problem

$$\min \ \|x\|_{\mathsf{TV}} \text{ s. t. } \mathcal{F}_n x = y$$

- Infinite-dimensional variable x
- Finitely many constraints

Dual problem

$$\max \ \mathrm{Re}\langle y, c\rangle \text{ s. t. } \|\mathcal{F}_n^* c\|_\infty \leq 1$$

- Finite-dimensional variable c
- Infinitely many constraints

$$(\mathcal{F}_n^* c)(t) = \sum_{|k| \leq f_c} c_k e^{i2\pi k t}$$

Semidefinite representability

$|(\mathcal{F}_n^* c)(t)| \leq 1$ for all $t \in [0,1]$ equivalent to

(1) there is Q Hermitian s. t.

$$\begin{bmatrix} Q & c \\ c^* & 1 \end{bmatrix} \succeq 0$$

(2) $\mathrm{trace}(Q) = 1$

(3) sums along superdiagonals vanish, $\sum_{i=1}^{n-j} Q_{i,i+j} = 0$ for $1 \leq j \leq n-1$

Semidefinite representability

$(\mathcal{F}_n^* \, c)(t) = \sum_{k=0}^{n-1} c_k e^{i2\pi kt}$

$$\|\mathcal{F}_n^* c\|_\infty \le 1 \quad \Longleftrightarrow \quad \begin{bmatrix} Q & c \\ c^* & 1 \end{bmatrix} \succeq 0, \quad \sum_{i=1}^{n-j} Q_{i,i+j} = \begin{cases} 1 & j = 0 \\ 0 & j = 1, 2, \ldots, n-1 \end{cases}$$

Semidefinite representability

$(\mathcal{F}_n^* c)(t) = \sum_{k=0}^{n-1} c_k e^{i2\pi kt}$

$$\|\mathcal{F}_n^* c\|_\infty \le 1 \quad \Longleftrightarrow \quad \begin{bmatrix} Q & c \\ c^* & 1 \end{bmatrix} \succeq 0, \quad \sum_{i=1}^{n-j} Q_{i,i+j} = \begin{cases} 1 & j = 0 \\ 0 & j = 1, 2, \ldots, n-1 \end{cases}$$

Why (one way)?

$$\begin{bmatrix} Q & c \\ c^* & 1 \end{bmatrix} \succeq 0 \quad \Longleftrightarrow \quad Q - cc^* \succeq 0$$

$z = (z_0, \ldots, z_{n-1}),\ z_k = e^{i2\pi kt}$

$$z^* Q z = 1 \qquad\qquad z^* cc^* z = |c^* z|^2 = |(\mathcal{F}_n^* c)(t)|^2$$

SDP formulation

Dual as an SDP

$$\text{maximize} \quad \mathrm{Re}\langle y, c\rangle \qquad \text{subject to} \quad \begin{bmatrix} Q & c \\ c^* & 1 \end{bmatrix} \succeq 0$$
$$\sum_{i=1}^{n-j} Q_{i,i+j} = \delta_j \quad 0 \leq j \leq n-1$$

Algorithm

(1) Solve dual

(2) Check when $\sum_{|k| \leq f_c} c_k e^{i2\pi kt}$ has magnitude 1 $\rightarrow$ gives support T

SDP formulation

Dual as an SDP

$$\text{maximize} \quad \mathrm{Re}\langle y, c\rangle \qquad \text{subject to} \quad \begin{bmatrix} Q & c \\ c^* & 1 \end{bmatrix} \succeq 0$$
$$\sum_{i=1}^{n-j} Q_{i,i+j} = \delta_j \quad 0 \le j \le n-1$$

Algorithm

(1) Solve dual

(2) Check when $\sum_{|k|\le f_c} c_k e^{i2\pi kt}$ has magnitude 1 $\rightarrow$ gives support T

Find roots (on unit circle) of polynomial of degree $2n-2$

$$p_{2n-2}(e^{i2\pi t}) = 1 - |(\mathcal{F}_n^* c)(t)|^2 = 1 - \sum_{k=-2f_c}^{2f_c} u_k e^{i2\pi kt}, \qquad u_k = \sum_j c_j \bar{c}_{j-k}$$

At most $n-1$ roots! $\rightarrow$ Can solve for amplitudes

There is a solution with support size $n-1$. Not true in finite dimension!

Dual polynomial

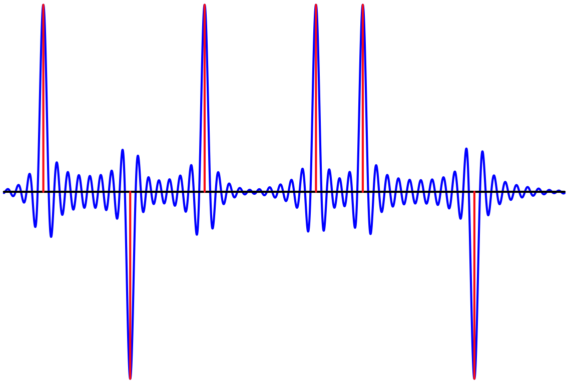


Figure: Sign of a real atomic measure x (red) and dual trigonometric polynomial $\mathcal{F}_n^* c$. Here, $f_c = 50$ so that we have $n = 101$ low-frequency coefficients.

Accuracy

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	$1.83\,10^{-7}$	$8.14\,10^{-8}$	$2.55\,10^{-8}$	$2.31\,10^{-8}$

Table: Numerical recovery of the signal support. There are approximately $f_c/4$ random locations in the unit interval.

Recovery example

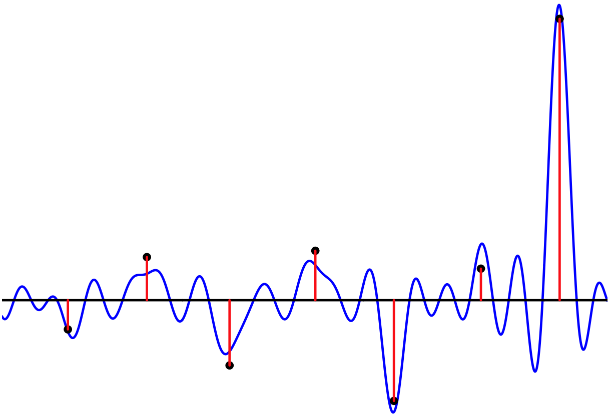


Figure: There are 21 spikes situated at arbitrary locations separated by at least $2\lambda_c$ and we observe 101 low-frequency coefficients ($f_c = 50$). In the plot, seven of the original spikes (black dots) are shown along with the corresponding low resolution data (blue line) and the estimated signal (red line).

Dual polynomial with random data

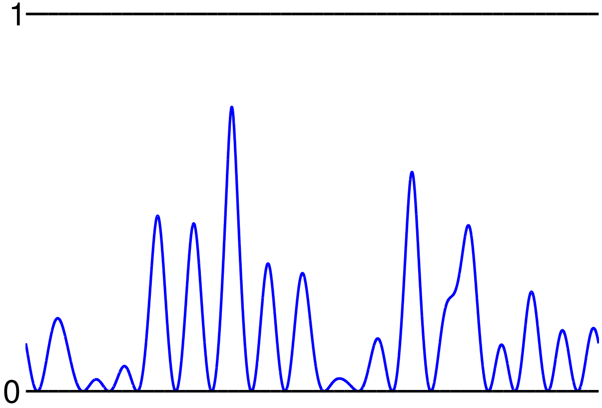
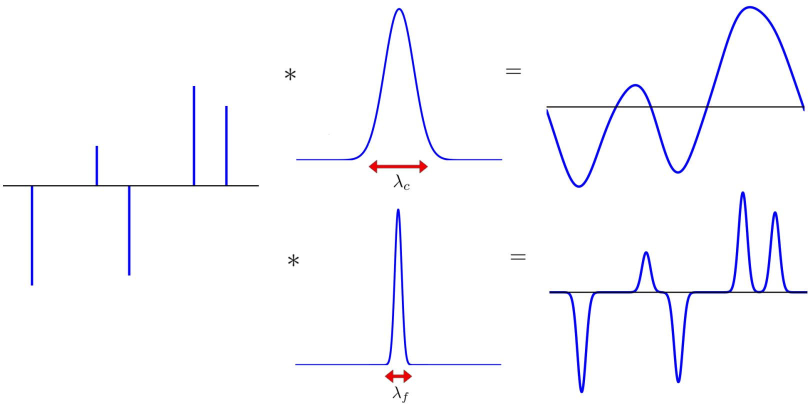


Figure: Trigonometric polynomial $1 - |(\mathcal{F}_n^* c)(t)|^2$ with random data $y \in \mathbb{C}^{21}$ ($n = 21$ and $f_c = 10$) with i.i.d. complex Gaussian entries. The polynomial has 16 roots.

Stability

The super-resolution factor (SRF): spatial viewpoint

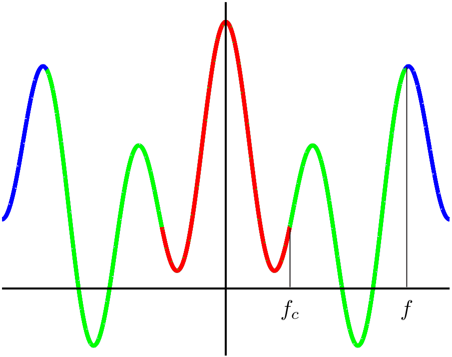


- Have data at resolution λ_c
- Wish resolution λ_f

Super-resolution factor

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

The super-resolution factor (SRF): frequency viewpoint



- Observe spectrum up to f_c
- Wish to extrapolate up to f

Super-resolution factor

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

Stability

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

Noisy data

$$y = \mathcal{F}_n x + w \quad \iff \quad \begin{aligned} \mathcal{F}_n^* y &= \mathcal{F}_n^* \mathcal{F}_n x + \mathcal{F}_n^* w \\ s &= \mathcal{P}_n x + z \end{aligned}$$

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

Bounded noise $\|z\|_{\mathsf{TV}} = \|z\|_{L_1} \leq \delta$

Stability

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

Noisy data

$$y = \mathcal{F}_n x + w \quad \Longleftrightarrow \quad \begin{aligned} \mathcal{F}_n^* y &= \mathcal{F}_n^* \mathcal{F}_n x + \mathcal{F}_n^* w \\ s &= \mathcal{P}_n x + z \end{aligned}$$

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

Bounded noise $\|z\|_{\mathsf{TV}} = \|z\|_{L_1} \leq \delta$

Recover signal by solving

$$\text{min } \|\tilde{x}\|_{\mathsf{TV}} \quad \text{subject to} \quad \|s - \mathcal{P}_n \tilde{x}\|_{\mathsf{TV}} \leq \delta$$

Stability

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

Noisy data

$$y = \mathcal{F}_n x + w \quad \Longleftrightarrow \quad \begin{array}{c} \mathcal{F}_n^* y = \mathcal{F}_n^* \mathcal{F}_n x + \mathcal{F}_n^* w \\ s = \mathcal{P}_n x + z \end{array}$$

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

Bounded noise $\|z\|_{\mathsf{TV}} = \|z\|_{L_1} \le \delta$

Recover signal by solving

$$\min \ \|\tilde{x}\|_{\mathsf{TV}} \quad \text{subject to} \quad \|s - \mathcal{P}_n \tilde{x}\|_{\mathsf{TV}} \le \delta$$

Theorem (C. and Fernandez Granda (2012))

If min dist. is at least $2\lambda_c$

$$\|(\hat{x} - x) * \varphi_{\lambda_c}\|_{\mathsf{TV}} \lesssim \delta$$

Stability

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

Noisy data

$$y = \mathcal{F}_n x + w \quad \Longleftrightarrow \quad \begin{aligned} \mathcal{F}_n^* y &= \mathcal{F}_n^* \mathcal{F}_n x + \mathcal{F}_n^* w \\ s &= \mathcal{P}_n x + z \end{aligned}$$

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

Bounded noise $\|z\|_{\mathsf{TV}} = \|z\|_{L_1} \le \delta$

Recover signal by solving

$$\min \ \|\tilde{x}\|_{\mathsf{TV}} \quad \text{subject to} \quad \|s - \mathcal{P}_n \tilde{x}\|_{\mathsf{TV}} \le \delta$$

Theorem (C. and Fernandez Granda (2012))

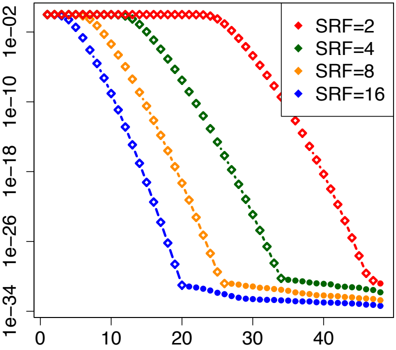
If min dist. is at least $2\lambda_c$

$$\|(\hat{x} - x) * \varphi_{\lambda_f}\|_{\mathsf{TV}} \lesssim \mathsf{SRF}^2 \cdot \delta$$

Limits of Super-resolution: Sparsity and Stability

Sparsity and stability

- Fixed grid of size $k = 48$ with spacing Rayleigh distance/SRF
- Compute eigenvalues of $\mathcal{P}_n$ with input on this grid



Analysis via Slepian's discrete prolate sequences

David Slepian

Analysis via Slepian's discrete prolate sequences (sketch)

$$s = \mathcal{P}_n(x + z)$$

1. Distance is Rayleigh/4 $\rightarrow$ there are eigenvalues/eigenvectors

$$\begin{aligned} \mathcal{P}_n x \approx \lambda\, x \quad & \lambda \approx 5.22\,\sqrt{k+1}\,e^{-3.23(k+1)} \\ k = 48 \quad & \lambda \leq 7 \times 10^{-68} \end{aligned}$$

Analysis via Slepian's discrete prolate sequences (sketch)

$$s = \mathcal{P}_n(x + z)$$

1. Distance is Rayleigh/4 $\rightarrow$ there are eigenvalues/eigenvectors

$$\begin{aligned} \mathcal{P}_n x \approx \lambda\, x \quad & \lambda \approx 5.22\,\sqrt{k+1}\, e^{-3.23(k+1)} \\ k = 48 \quad & \lambda \leq 7 \times 10^{-68} \end{aligned}$$

2. Distance is Rayleigh/1.05 (only seek to extend the spectrum by 5%)

$$\begin{aligned} \mathcal{P}_n x = \lambda\, x \quad & \lambda \approx 3.87\,\sqrt{k+1}\, e^{-0.15(k+1)} \\ k = 256 \quad & \lambda \leq 1.2 \times 10^{-15} \end{aligned}$$

Analysis via Slepian's discrete prolate sequences (sketch)

$$s = \mathcal{P}_n(x + z)$$

1. Distance is Rayleigh/4 $\rightarrow$ there are eigenvalues/eigenvectors

$$\begin{aligned} \mathcal{P}_n x \approx \lambda\, x \quad & \lambda \approx 5.22\,\sqrt{k+1}\,e^{-3.23(k+1)} \\ k = 48 \quad & \lambda \leq 7 \times 10^{-68} \end{aligned}$$

2. Distance is Rayleigh/1.05 (only seek to extend the spectrum by 5%)

$$\begin{aligned} \mathcal{P}_n x = \lambda\, x \quad & \lambda \approx 3.87\,\sqrt{k+1}\,e^{-0.15(k+1)} \\ k = 256 \quad & \lambda \leq 1.2 \times 10^{-15} \end{aligned}$$

3. (1) and (2) worse when spacing $\rightarrow 0$

Analysis via Slepian's discrete prolate sequences (sketch)

$$s = \mathcal{P}_n(x + z)$$

1. Distance is Rayleigh/4 $\rightarrow$ there are eigenvalues/eigenvectors

$$\begin{array}{cc} \mathcal{P}_n x \approx \lambda\, x & \lambda \approx 5.22\,\sqrt{k+1}\,e^{-3.23(k+1)} \\ k = 48 & \lambda \leq 7 \times 10^{-68} \end{array}$$

2. Distance is Rayleigh/1.05 (only seek to extend the spectrum by 5%)

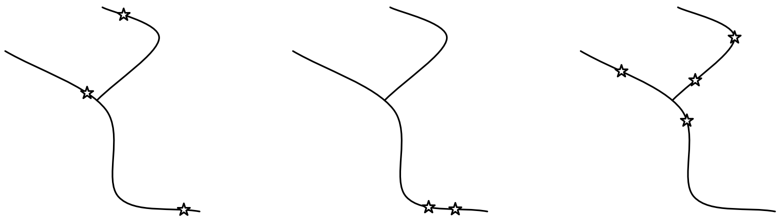
$$\begin{array}{cc} \mathcal{P}_n x = \lambda\, x & \lambda \approx 3.87\,\sqrt{k+1}\,e^{-0.15(k+1)} \\ k = 256 & \lambda \leq 1.2 \times 10^{-15} \end{array}$$

3. (1) and (2) worse when spacing $\rightarrow 0$
4. (1) approx holds for subspace of dimension $3k/4$

Application: Single Molecule Imaging in 3D Microscopy

Joint with Moerner Lab and Veniamin Morgenshtern (Stanford)

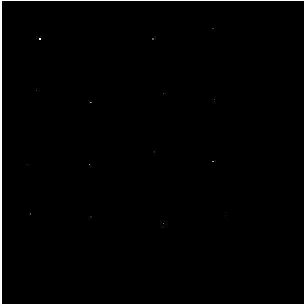
Structure of interest contains molecules that are “blinking”



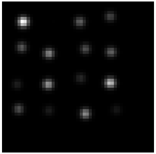
Frame 1 Frame 2 Frame 3

- **Few** molecules are active in each frame $\Rightarrow$ **sparsity!**
- **Multiple** (~ 10000) frames are recorded and processed individually
- **Results** from all frames are **combined** to reveal the underlying structure

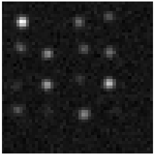
Optics acts as low-pass filter, detector adds noise



Original



Low-pass, subsampled

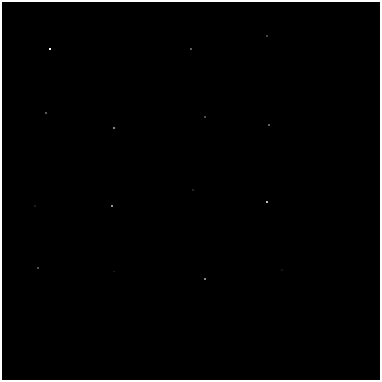


Noisy

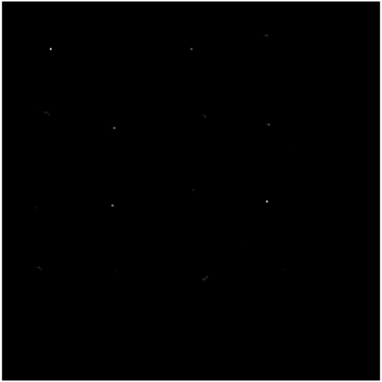
$$y = Lx + z$$

- x: signal
- y: output at the detector
- z: normal zero-mean noise
- L: models optics + subsampling (low-pass)

Noisy recovery



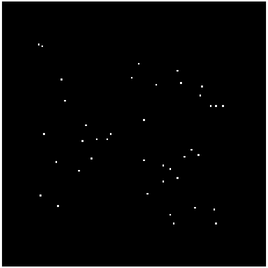
Original



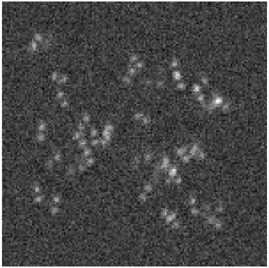
Estimate

Recovery of 3D signals

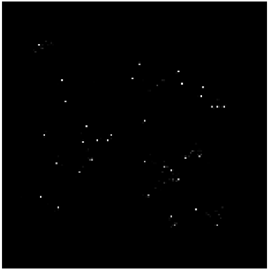
- Double-helix (DH) point spread function has **two** lobes
- The **angle** defined by these lobes encodes **z-position** of the molecule
- Appropriately modifying L, we can use the same algorithm to **reconstruct 3D signals from 2D data**



Original 3D signal, projected onto XY plane

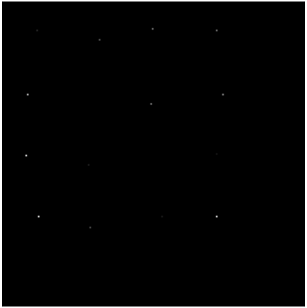


2D DH data

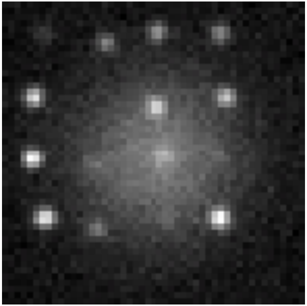


Estimated 3D signal, projected onto XY plane

Smooth background separation



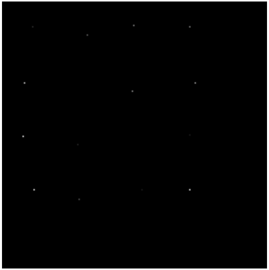
Original



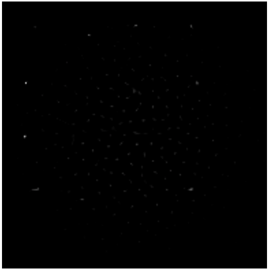
Data

minimize	$\frac{1}{2}\|y - L(x+p)\|_2^2 + \lambda\sigma\|x\|_{\mathsf{TV}}$
subject to	$x \geq 0$
	p low freq. trig. polynomial (background)

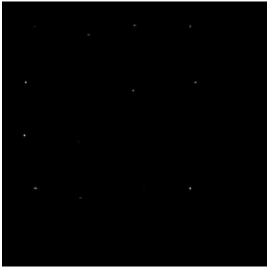
Smooth background separation (Cont'd)



Original



LASSO estimate (speckles)



Polynomial separation estimate (clean)

Summary

Distance between events	< Rayleigh	> Rayleigh
Noiseless TV recovery	✗	✓
Stability	✗ no method is stable	✓ min TV is stable

- Can super-resolve signals by convex programming
- Need structural assumptions for stable recovery
- Ongoing applications in 3D microscopy

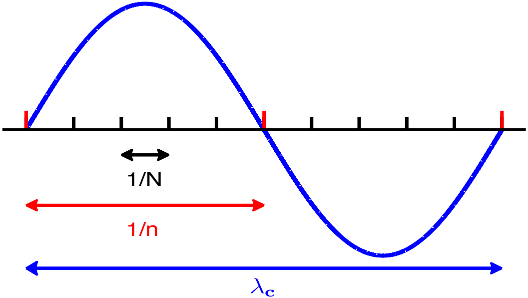
E. J. Candès, and C. Fernandez-Granda (2012). *Towards a mathematical theory of super-resolution*. To appear in Comm. Pure Appl. Math

E. J. Candès, and C. Fernandez-Granda (2012). *Super-resolution from noisy data*. `http://arxiv.org/abs/XXXX.YYYY`

The super-resolution factor (SRF)

$$\text{SRF} := \frac{\text{fine resolution}}{\text{coarse resolution}} := \frac{N}{n} \text{ (for discrete data)}$$

Wish to extend spectrum up until SRF $\times f_c$



Pictorial representation of SRF