A Convex-Programming Framework for Super-Resolution

Carlos Fernandez-Granda

www.stanford.edu/~cfgranda/

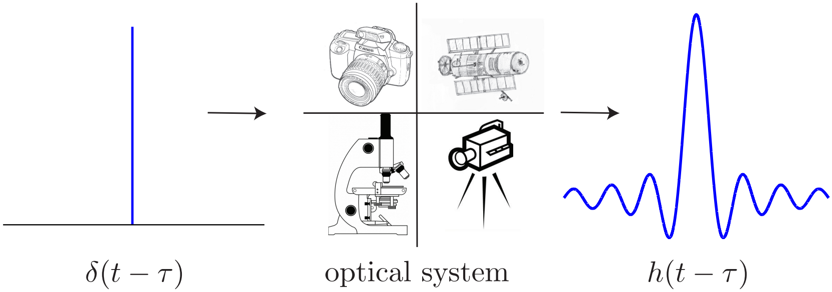
March/February 2014

Acknowledgements

- This work was supported by a Fundación La Caixa Fellowship and a Fundación Caja Madrid Fellowship
- **Collaborator** : Emmanuel Candès (Department of Mathematics and of Statistics, Stanford)

Motivation : Limits of resolution in imaging

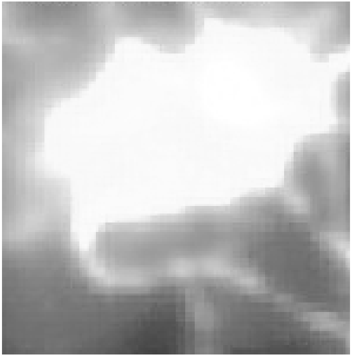
The resolving power of lenses, however perfect, is limited (**Lord Rayleigh**)



Diffraction imposes a fundamental limit on the resolution of optical systems

Aim

Estimation from data that have limited resolution

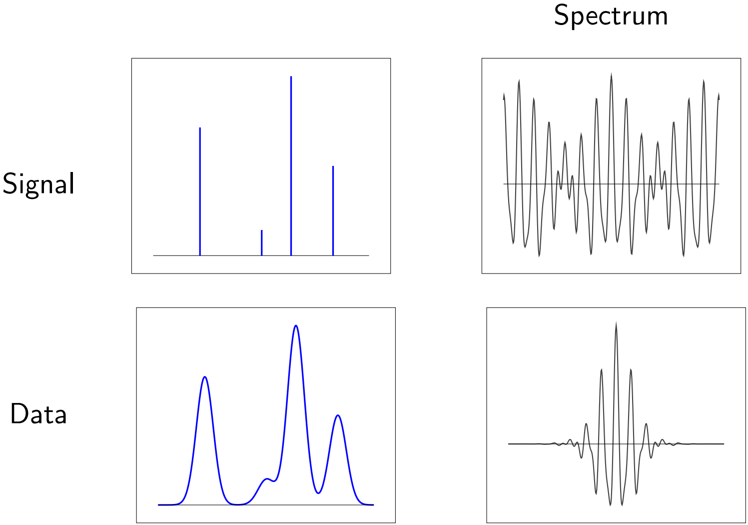


- Microscopy
- Astronomy
- Electronic imaging
- Medical imaging
- Signal processing
- Radar
- Spectroscopy
- Geophysics
- ...

Super-resolution

- **Optics** : Data-acquisition techniques to overcome the diffraction limit
- **Image processing** : Methods to upsample images onto a finer grid while preserving edges and hallucinating textures
- **This talk** : Signal estimation from low-pass measurements

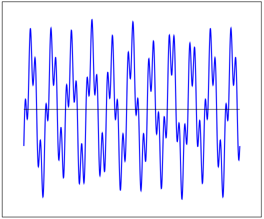
Spatial Super-resolution

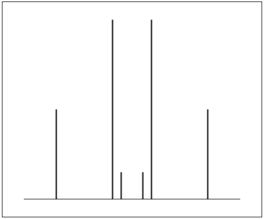


Spectral Super-resolution

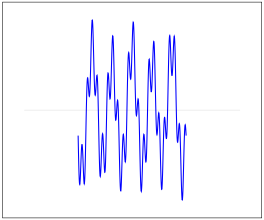
Spectrum

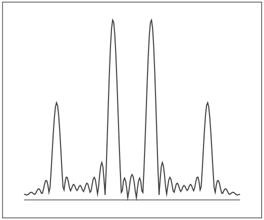
Signal





Data





Point sources

- In many applications signals of interest are point sources :
  - Celestial bodies (astronomy)
  - Fluorescent molecules (microscopy)
  - Line spectra (spectroscopy, signal processing)

Point sources

- In many applications signals of interest are point sources :
  - Celestial bodies (astronomy)
  - Fluorescent molecules (microscopy)
  - Line spectra (spectroscopy, signal processing)
- Traditional approaches
  1. Fitting point-spread function to each source (matched filtering)
     - Sensitive to noise and high dynamic ranges
  2. Algorithms based on Prony's method : MUSIC, ESPRIT, ...
     - Parametric (number of sources must be known)
     - Extension to 2D is very computationally intensive
     - Strong assumptions on noise (Gaussian, white), signal and measurement model

Statistical estimation via convex programming

- In the 70s and 80s, ℓ_1-norm minimization proposed for deconvolution in seismography [Claerbout, Muir '73],[Levy, Fullagar '81], [Santosa, Symes '86]
- Later, huge impact of convex-programming techniques in high-dimensional statistics
    1. Well-developed theory
    2. Robustness to noise, even in non-asymptotic regimes
    3. Flexibility
- Very little theory on estimation from low-resolution data (original problem tackled by geophysicists)

Super-resolution via convex programming

- Can we super-resolve using optimization ? Under what conditions ?
- Is the method stable to noise ?
- How do we adapt to different signal, noise and measurement models ?

Super-resolution via convex programming

- Can we super-resolve using optimization ? Under what conditions ?
- Is the method stable to noise ?
- How do we adapt to different signal, noise and measurement models ?

- This talk : Framework for estimation from low-resolution data
  1. Precise theoretical analysis
  2. Non-asymptotic stability guarantees
  3. Natural extensions handle
     - Piecewise-smooth functions
     - Clustered point sources
     - Demixing of sines and spikes
     - Super-resolution from multiple measurements

Outline of the talk

Basic model

Estimation from noisy data

A general framework

Basic model

Estimation from noisy data

A general framework

Mathematical model

- Signal : superposition of Dirac measures with support T

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

- Data : low-pass Fourier coefficients with cut-off frequency f_c

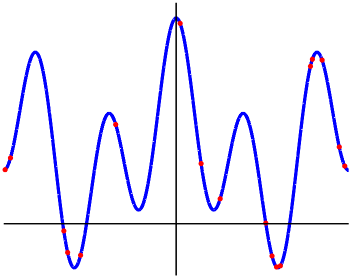
$$y = \mathcal{F}_c\, x$$

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

Compressed sensing vs super-resolution

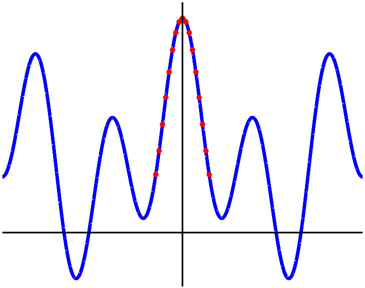
Estimation of sparse signals from undersampled measurements suggests connections to compressed sensing

Compressed sensing



spectrum interpolation

Super-resolution



spectrum extrapolation

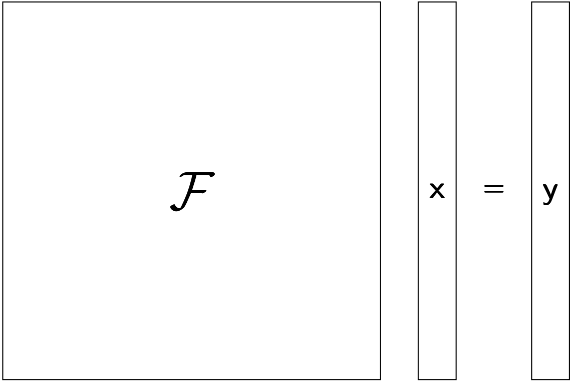
Compressed sensing

- Compressed sensing : stable estimation from random Fourier coefficients [Candès, Tao, Romberg '04]
- **Crucial insight** : measurement operator is well conditioned when acting upon sparse signals
- Equivalently, the energy of all sparse signals is preserved in the data (*restricted isometry property*)
- Most analyses of sparse-regression methods in high-dimensional statistics are based on similar conditions (*restricted-eigenvalue condition*, *restricted strong convexity*, *null-space property*)

Compressed sensing

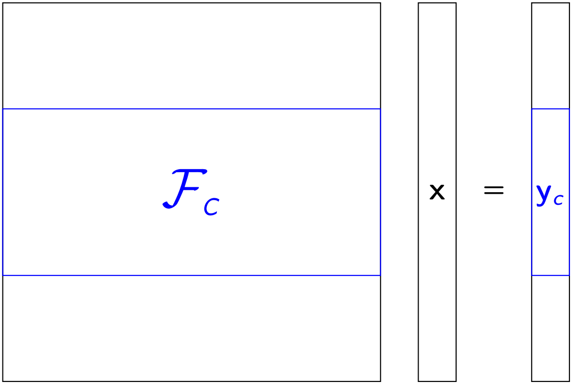
- Compressed sensing : stable estimation from random Fourier coefficients [Candès, Tao, Romberg '04]
- **Crucial insight** : measurement operator is well conditioned when acting upon sparse signals
- Equivalently, the energy of all sparse signals is preserved in the data (*restricted isometry property*)
- Most analyses of sparse-regression methods in high-dimensional statistics are based on similar conditions (*restricted-eigenvalue condition*, *restricted strong convexity*, *null-space property*)
- Do they hold in super-resolution ?

Simple experiment



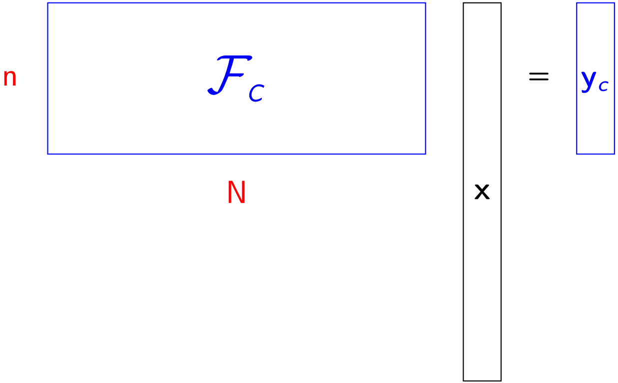
Discretize support to lie on a grid with $N = 4096$ points

Simple experiment



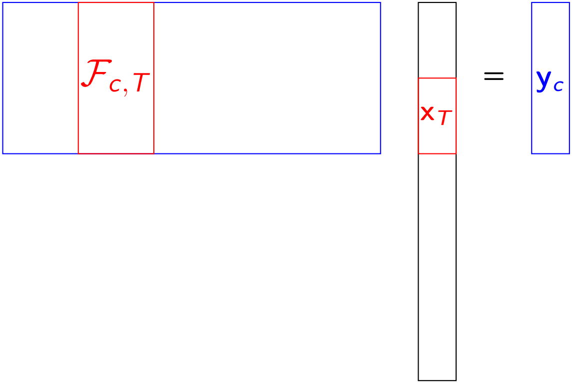
Measure n low-pass DFT coefficients, super-resolution factor (SRF) : N/n

Simple experiment



Measure n low-pass DFT coefficients, super-resolution factor (SRF) : N/n

Simple experiment



Restrict support of the signal to an interval of 48 contiguous points

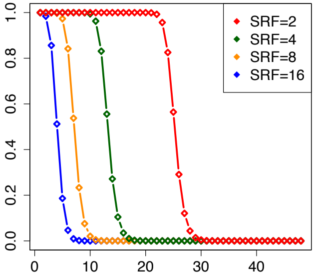
Simple experiment

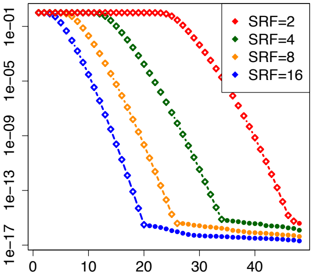
$$\mathcal{F}_{c,T} \; \mathbf{x}_T = \mathbf{y}_c$$

Compute singular values of resulting linear operator

Sparsity is not enough

Most clustered sparse signals are suppressed by low-pass filtering



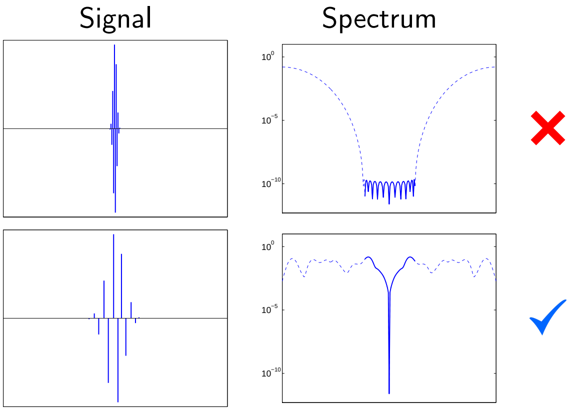


For SRF $= 4$, there is a subspace S of dimension 24 where for all unit-normed $x \in S$ $||\mathcal{F}_c\, x||_2 \leq 10^{-7}$

For such signals estimation is impossible by any method at signal-to-noise ratios below 145 dB

Theory : prolate spheroidal sequences [Slepian '78]

Sparsity is not enough

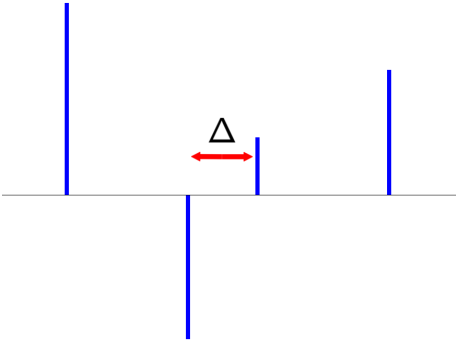


More refined conditions are necessary to restrict our signal model

Minimum separation

Definition : The minimum separation Δ of a discrete set T is

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



Total-variation norm

- Continuous counterpart of the ℓ_1 norm
- If $x = \sum_j a_j \delta_{t_j}$ then $||x||_{\text{TV}} = \sum_j |a_j|$
- Not the total variation of a piecewise-constant function

Total-variation norm

- Continuous counterpart of the ℓ_1 norm
- If $x = \sum_j a_j \delta_{t_j}$ then $||x||_{\text{TV}} = \sum_j |a_j|$
- Not the total variation of a piecewise-constant function
- **Formal definition :** For a complex measure ν

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

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

Estimation via convex programming

In a zero-noise limit, i.e. $y = \mathcal{F}_c \, x$, we solve

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

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

Estimation via convex programming

In a zero-noise limit, i.e. $y = \mathcal{F}_c\, x$, we solve

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

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

Theorem [Candès, F. '12]

If the minimum separation of the signal support T obeys

$$\Delta \geq 2\,/f_c := 2\,\lambda_c,$$

then recovery is exact

Nonparametric approach (no previous knowledge of the number of spikes)

Estimation via convex programming

In a zero-noise limit, i.e. $y = \mathcal{F}_c\, x$, we solve

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

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

Theorem [Candès, F. '14]

If the minimum separation of the signal support T obeys

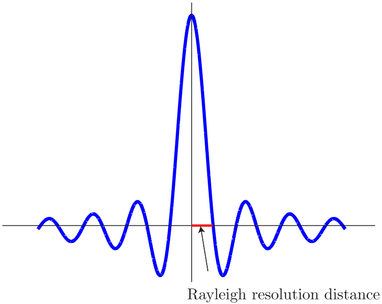
$$\Delta \geq 1.38\, / f_c := 1.38\, \lambda_c,$$

then recovery is exact

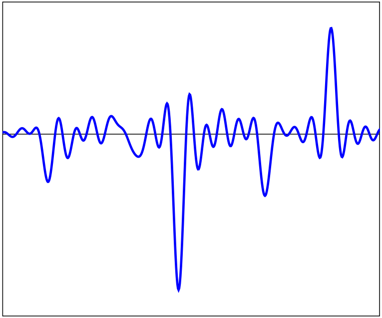
Nonparametric approach (no previous knowledge of the number of spikes)

Minimum separation

Point-spread function



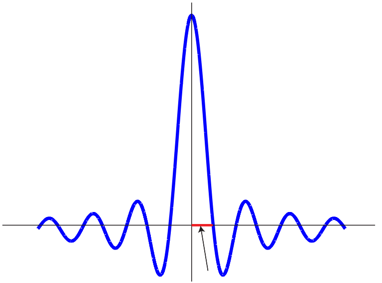
$\Delta = 1.38\,\lambda_c$



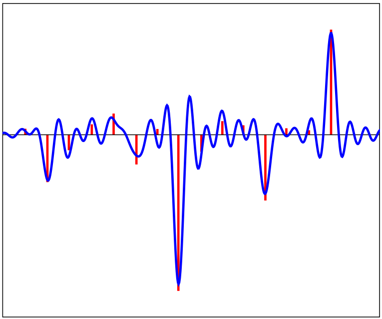
$\lambda_c/2$ is the Rayleigh resolution limit

Minimum separation

Point-spread function



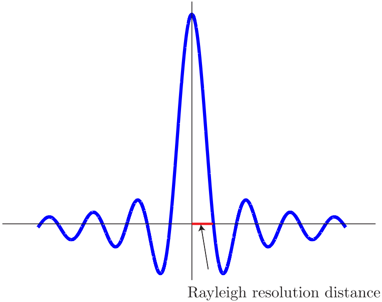
$\Delta = 1.38\,\lambda_c$



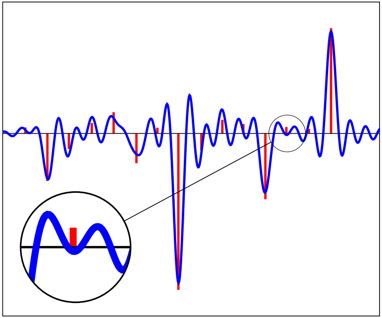
$\lambda_c/2$ is the Rayleigh resolution limit

Minimum separation

Point-spread function

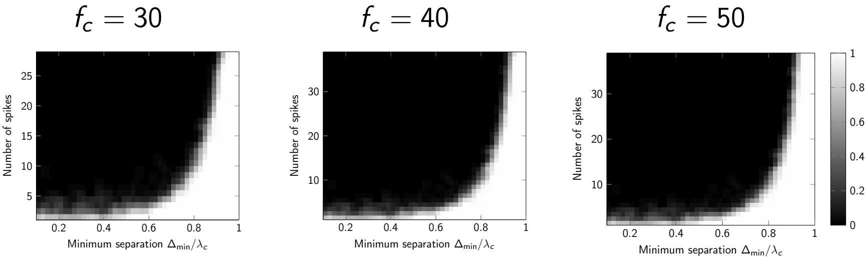


$\Delta = 1.38\,\lambda_c$



$\lambda_c/2$ is the Rayleigh resolution limit

Numerical evaluation of minimum separation



Conjecture : TV-norm minimization succeeds if $\Delta \geq \lambda_c$

Sparse estimation from correlated covariates

If we discretize the support

- Sparse recovery via ℓ_1-norm minimization in an overcomplete Fourier dictionary
- Theory based on dictionary incoherence [Donoho, Stark '89], [Tropp '06] is very weak, due to high column correlation
- If the ambient dimension is 20 000 and we have 1 000 measurements, how many spikes can we recover ?

Sparse estimation from correlated covariates

If we discretize the support

- Sparse recovery via ℓ_1-norm minimization in an overcomplete Fourier dictionary
- Theory based on dictionary incoherence [Donoho, Stark '89], [Tropp '06] is very weak, due to high column correlation
- If the ambient dimension is 20 000 and we have 1 000 measurements, how many spikes can we recover ?

  Previous theory [Dossal, Mallat '05] : 3 spikes

Sparse estimation from correlated covariates

If we discretize the support

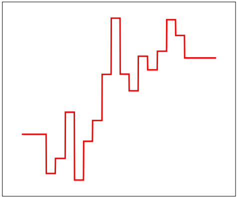
- Sparse recovery via ℓ_1-norm minimization in an overcomplete Fourier dictionary
- Theory based on dictionary incoherence [Donoho, Stark '89], [Tropp '06] is very weak, due to high column correlation
- If the ambient dimension is 20 000 and we have 1 000 measurements, how many spikes can we recover ?

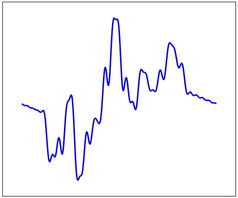
  Previous theory [Dossal, Mallat '05] : 3 spikes

  Our result : 362 spikes

Piecewise-constant functions

- Signal : piecewise-constant function
- Measurements : low-pass Fourier coefficients





Corollary

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

yields exact recovery if $\Delta \geq 1.38\,\lambda_c$

Similar result for cont. differentiable piecewise-smooth functions

Higher dimensions

- Signal : superposition of point sources (delta measures) in 2D
- Measurements : low-pass 2D Fourier coefficients

Theorem [Candès, F. 2012]

TV-norm minimization yields exact recovery if

$$\Delta \geq 2.38 \, \lambda_c$$

In dimension d, $\Delta \geq C_d \, \lambda_c$, where C_d only depends on d

Sketch of proof : Dual polynomial

A sufficient condition for

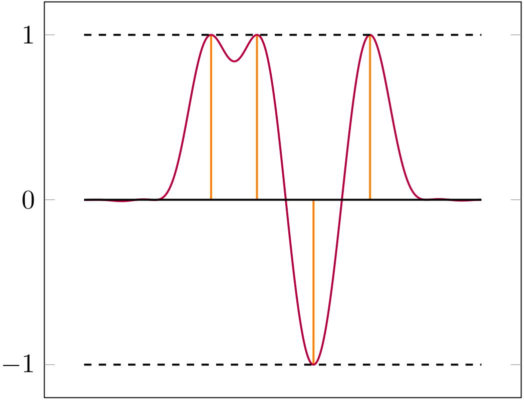
$$x = \sum_{j \in T} a_j \delta_{t_j} = \sum_{j \in T} |a_j| \, e^{i\phi_j} \delta_{t_j}$$

to be the unique solution is that there exists q such that

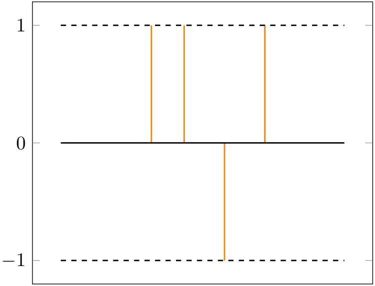
1. $q(t) = \sum_{k=-f_c}^{f_c} b_k e^{i2\pi kt}$ (low pass polynomial)
2. $q(t_j) = e^{i\phi_j}$, $t_j \in T$ (interpolates the sign of the signal on T)
3. $|q(t)| < 1$, $t \in T^c$

q is a subgradient of the TV norm at the signal x that is orthogonal to the null space of the measurement operator

Sketch of proof : Dual polynomial



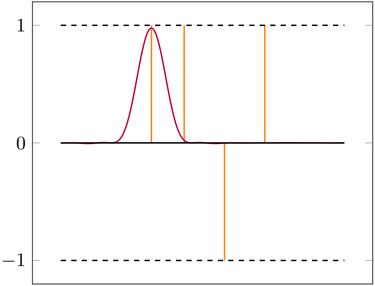
Sketch of proof : Construction by interpolation



1st idea : Interpolation with a low-frequency fast-decaying kernel K

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

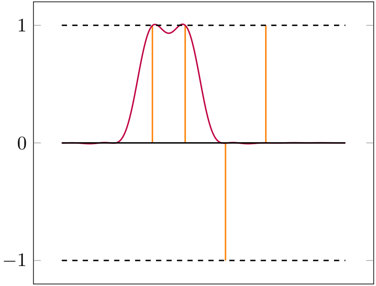
Sketch of proof : Construction by interpolation



1st idea : Interpolation with a low-frequency fast-decaying kernel K

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

Sketch of proof : Construction by interpolation



1st idea : Interpolation with a low-frequency fast-decaying kernel K

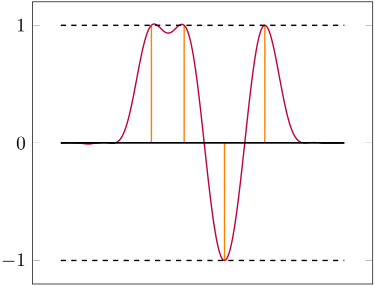
$$q(t) = \sum_{t_j \in T} \alpha_j \, K(t - t_j),$$

Sketch of proof : Construction by interpolation

1st idea : Interpolation with a low-frequency fast-decaying kernel K

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

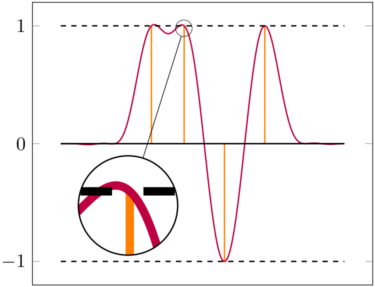
Sketch of proof : Construction by interpolation



1st idea : Interpolation with a low-frequency fast-decaying kernel K

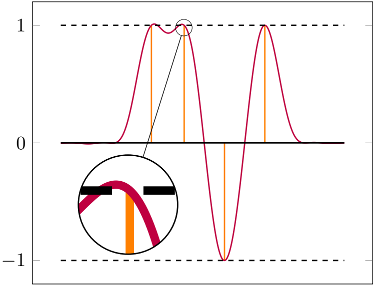
$$q(t) = \sum_{t_j \in T} \alpha_j \, K(t - t_j),$$

Sketch of proof : Construction by interpolation



Problem : Magnitude of polynomial locally exceeds 1

Sketch of proof : Construction by interpolation

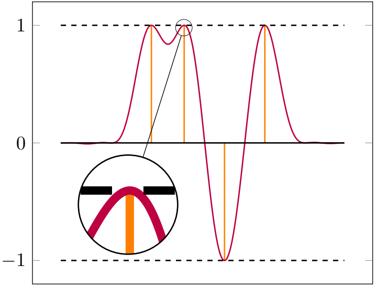


Problem : Magnitude of polynomial locally exceeds 1

Solution : Add correction term and force $q'(t_k) = 0$ for all $t_k \in T$

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

Sketch of proof : Construction by interpolation



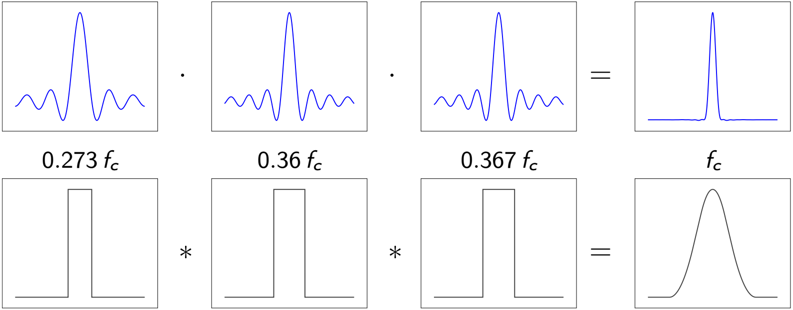
Problem : Magnitude of polynomial locally exceeds 1

Solution : Add correction term and force $q'(t_k) = 0$ for all $t_k \in T$

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

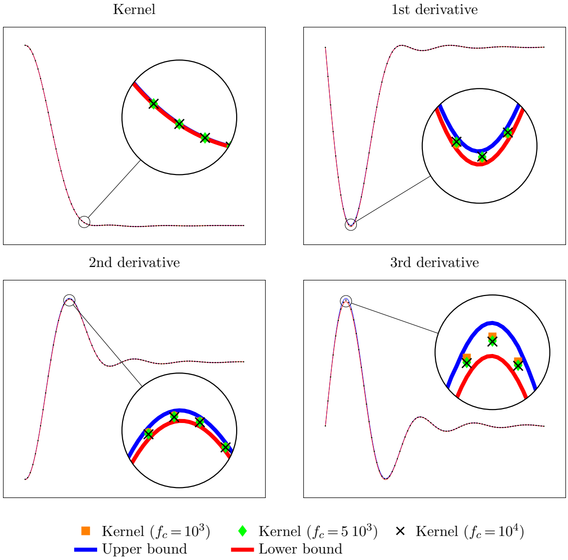
Sketch of proof : Interpolation kernel

Key step : Designing a good interpolation kernel



Trade-off between *spikiness* at the origin and asymptotic decay

Sketch of proof : Non-asymptotic bounds on kernel



Dual polynomial as theoretical tool

Subsequent work builds on our construction to analyze

- Stability of super-resolution [Candès, F. '13], [F. '13], [Azais, De Castro, Gamboa '13], [Duval, Peyré '13]
- Denoising of line spectra [Tang, Bhaskar, Recht '13]
- Compressed sensing off the grid [Tang, Bhaskar, Shah, Recht '13]
- Recovery of splines from their projection onto spaces of algebraic polynomials [Bendory, Dekel, Feuer '13], [De Castro, Mijoule '14]
- Recovery of point sources from spherical harmonics [Bendory, Dekel, Feuer '13]

Practical implementation

- **Primal problem** :

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

Infinite-dimensional variable $\tilde{x}$ (measure in $[0, 1]$)

First option : Discretizing + ℓ_1-norm minimization

Practical implementation

- **Primal problem** :

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

Infinite-dimensional variable $\tilde{x}$ (measure in $[0, 1]$)

First option : Discretizing + ℓ_1-norm minimization

- **Dual problem** :

$$\max_{\tilde{u} \in \mathbb{C}^n} \mathrm{Re}\left[y^* \tilde{u}\right] \quad \text{subject to} \quad ||\mathcal{F}_c^*\, \tilde{u}||_\infty \leq 1, \quad n := 2f_c + 1$$

Finite-dimensional variable $\tilde{u}$, but infinite-dimensional constraint

$$\mathcal{F}_c^*\, \tilde{u} = \sum_{k \leq |f_c|} \tilde{u}_k e^{i2\pi kt}$$

Second option : Solving the dual problem

Lemma : Semidefinite representation

The Fejér-Riesz Theorem and the semidefinite representation of polynomial sums of squares imply that

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

is equivalent to

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

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

Consequence : The dual problem is a tractable semidefinite program

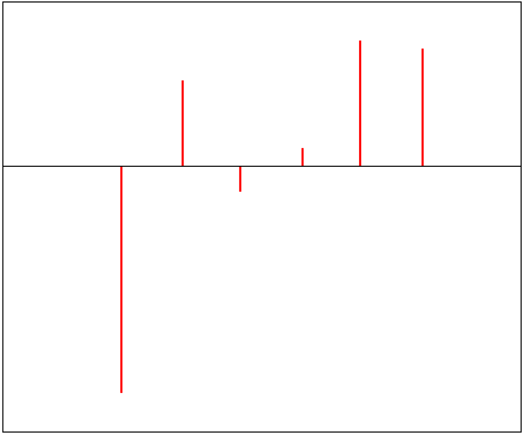
Support-locating polynomial

How do we obtain an estimator from the dual solution ?

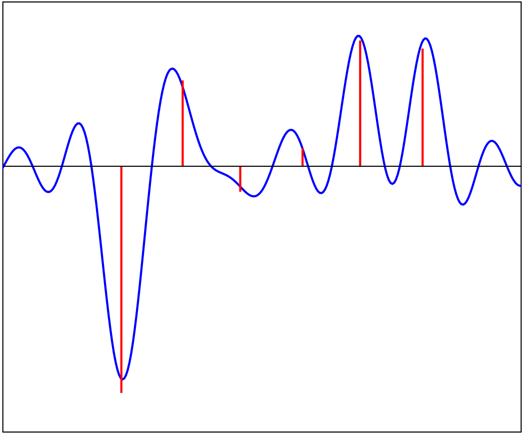
Dual solution vector : Fourier coefficients of low-pass polynomial that interpolates the sign of the primal solution (follows from strong duality)

Idea : Use the polynomial to locate the support of the signal

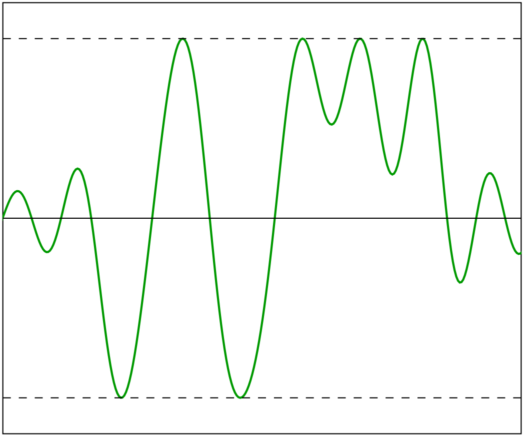
Super-resolution via semidefinite programming



Super-resolution via semidefinite programming

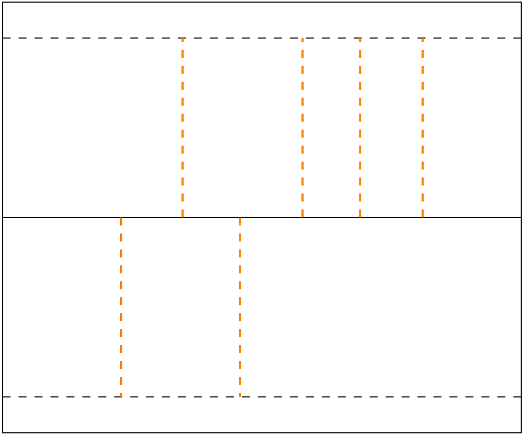


Super-resolution via semidefinite programming



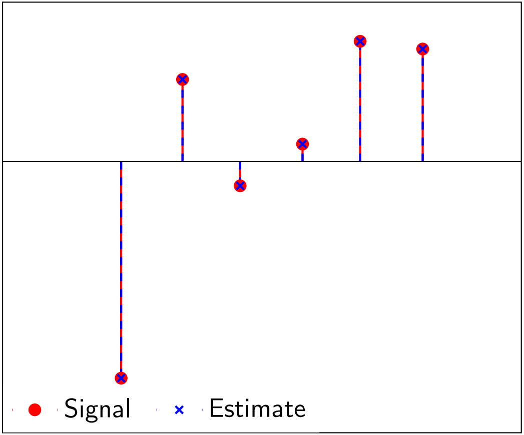
1. Solve semidefinite program to obtain dual solution

Super-resolution via semidefinite programming



2. Locate points at which corresponding polynomial has unit magnitude

Super-resolution via semidefinite programming



3. Estimate amplitudes via least squares

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

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

Estimation from noisy data

We assume additive noise with norm bounded by δ

$$y = \mathcal{F}_c x + z$$

Our estimator is the solution to

$$\min_{\tilde{x}} ||\tilde{x}||_{\mathrm{TV}} \quad \text{subject to} \quad ||\mathcal{F}_c \, \tilde{x} - y||_2 \leq \delta,$$

Estimation from noisy data

We assume additive noise with norm bounded by δ

$$y = \mathcal{F}_c x + z$$

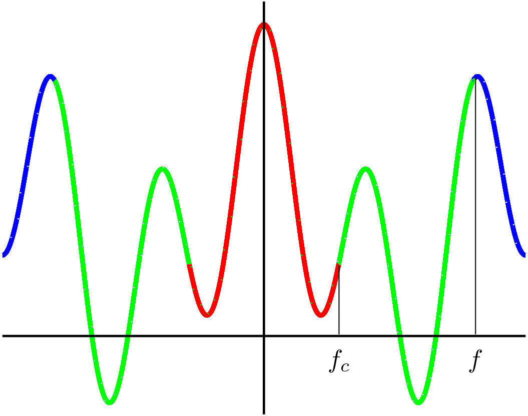
Our estimator is the solution to

$$\min_{\tilde{x}} ||\tilde{x}||_{\text{TV}} \quad \text{subject to} \quad ||\mathcal{F}_c\, \tilde{x} - y||_2 \leq \delta,$$

Metrics to quantify estimation accuracy :

1. Approximation error at a higher resolution
2. Support-detection error

Super-resolution factor : spectral viewpoint

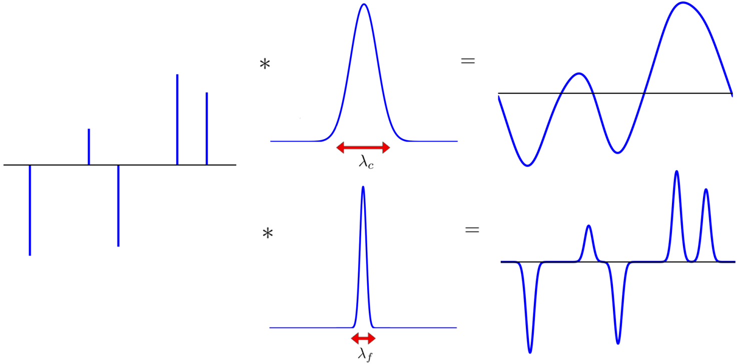


Super-resolution factor

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

Super-resolution factor : spatial viewpoint

Signal at a resolution λ : convolution with a kernel ϕ_λ of width λ



Super-resolution factor

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

Approximation at a higher resolution

At the resolution of the measurements

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

How does the estimate degrade at a higher resolution ?

Approximation at a higher resolution

At the resolution of the measurements

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

How does the estimate degrade at a higher resolution ?

Theorem [Candès, F. 2012]

If $\Delta \geq 1.38 / f_c$ then the solution $\hat{x}$ to

$$\min_{\tilde{x}} ||\tilde{x}||_{\text{TV}} \quad \text{subject to} \quad ||\mathcal{F}_c \, \tilde{x} - y||_2 \leq \delta,$$

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

Practical implementation at a noise level δ

- **Primal problem** :

$$\min_{\tilde{x}} ||\tilde{x}||_{\mathrm{TV}} \quad \text{subject to} \quad ||\mathcal{F}_c\, \tilde{x} - y||_2 \leq \delta$$

First option : Discretizing + ℓ_1-norm minimization

- **Dual problem** :

$$\max_{\tilde{u} \in \mathbb{C}^n} \mathrm{Re}\left[y^* \tilde{u}\right] - \delta\, ||\tilde{u}||_2 \quad \text{subject to} \quad ||\mathcal{F}_c^*\, \tilde{u}||_\infty \leq 1, \quad n := 2f_c + 1$$

Second option : Solving the dual problem

Practical implementation at a noise level δ

- **Primal problem** :

$$\min_{\tilde{x}} ||\tilde{x}||_{\mathrm{TV}} \quad \text{subject to} \quad ||\mathcal{F}_c\, \tilde{x} - y||_2 \leq \delta$$

  First option : Discretizing + ℓ_1-norm minimization

- **Dual problem** :

$$\max_{\tilde{u} \in \mathbb{C}^n} \mathrm{Re}\left[y^* \tilde{u}\right] - \delta\, ||\tilde{u}||_2 \quad \text{subject to} \quad ||\mathcal{F}_c^*\, \tilde{u}||_\infty \leq 1, \quad n := 2f_c + 1$$

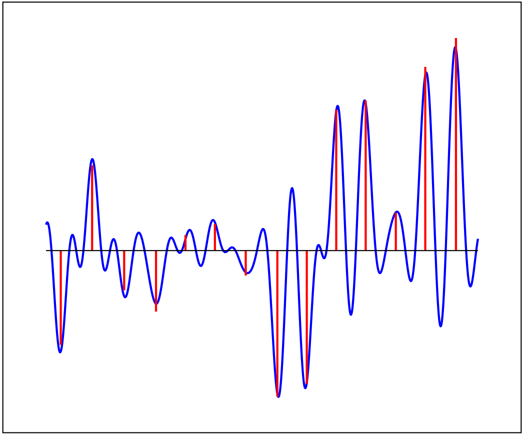
  Second option : Solving the dual problem

- **Dual solution :**

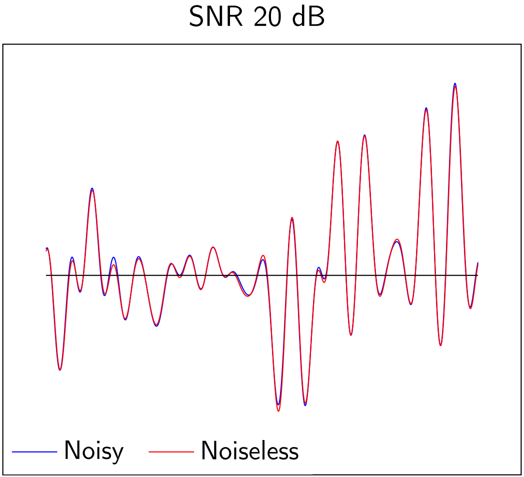
  Coefficients of polynomial that interpolates sign of primal solution

Example

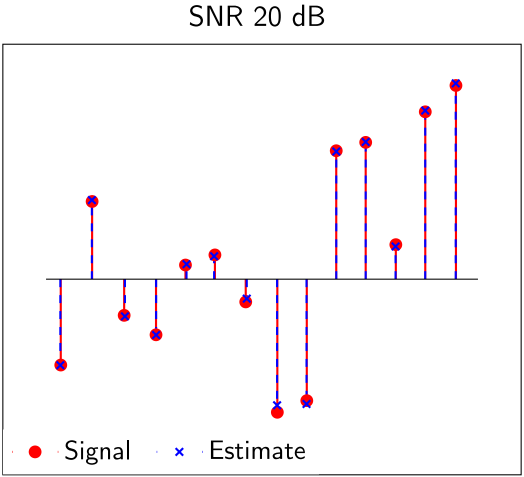
Minimum separation : $1.5\,\lambda_c$



Example

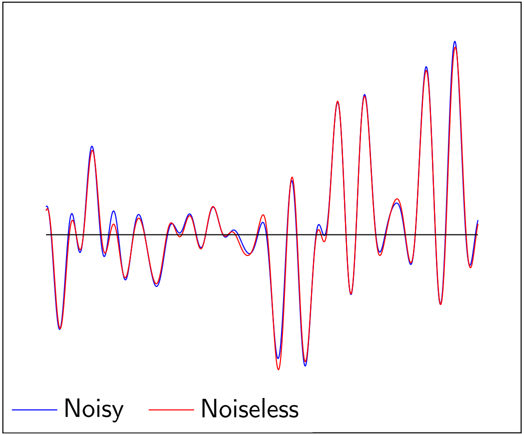


Example

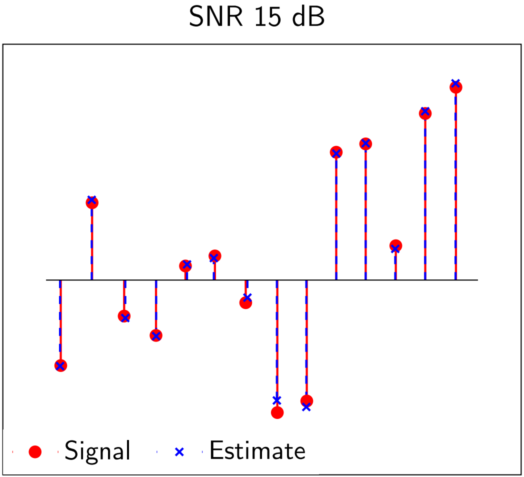


Example

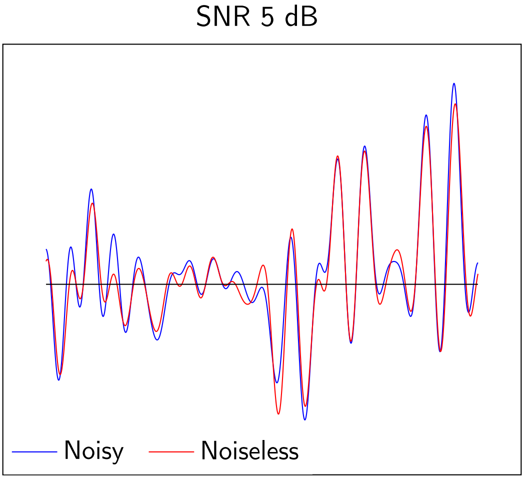
SNR 15 dB



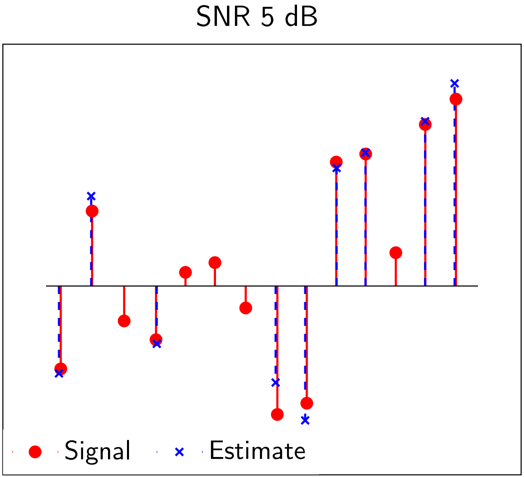
Example



Example



Example



Support-detection accuracy

- Original support : T
- Estimated support : $\widehat{T}$

Theorem [F. 2013]

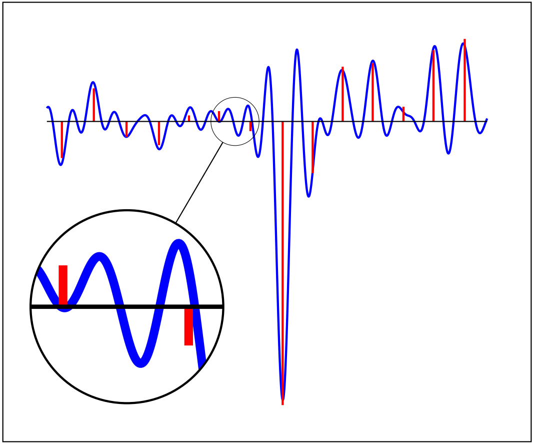
For any $t_i \in T$, if $|a_i| > C_1\delta$ there exists $\hat{t}_i \in \widehat{T}$ such that

$$\left|t_i - \hat{t}_i\right| \leq \frac{1}{f_c}\sqrt{\frac{C_2\delta}{|a_i| - C_1\delta}}$$

No dependence on the amplitude of the signal at other locations

Consequence

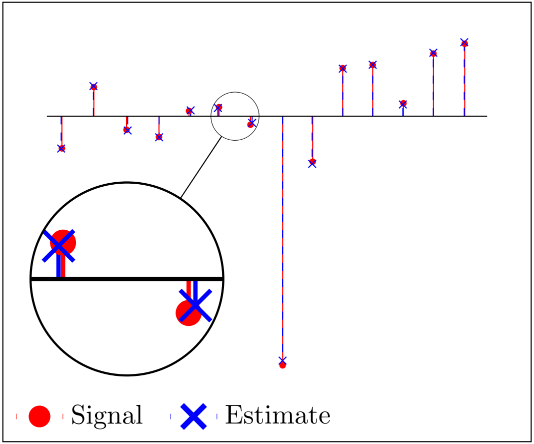
Robustness of the algorithm to high dynamic ranges



SNR 20 dB (15 dB without the large spike)

Consequence

Robustness of the algorithm to high dynamic ranges



SNR 20 dB (15 dB without the large spike)

Some comments

- Non-asymptotic results, whereas most theory for Prony-based methods is asymptotic (convergence of sample autocorrelation matrices)
- Usual proof techniques from high-dimensional statistics do not apply
  1. Conditions (restricted-isometry property, restricted-eigenvalue condition, etc.) do not hold
  2. Estimation takes place over a continuous domain
- Proofs combine insights from harmonic analysis and convex optimization (generalization of dual polynomials)

Basic model

Estimation from noisy data

A general framework

A general framework

Incorporating different assumptions on the signal, the noise and the sensing process is important in applications

We can do this by adapting the cost function and constraints of the optimization problem

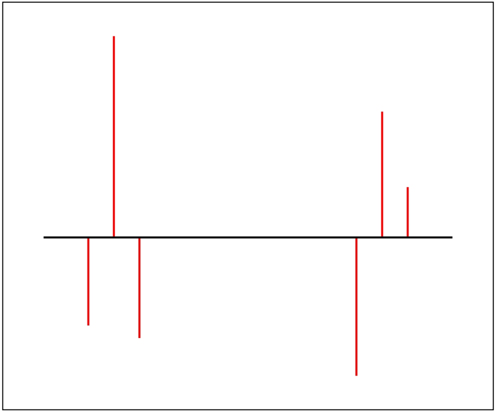
This section :

1. Super-resolution of clustered point sources
2. Demixing of sines and spikes
3. Super-resolution from multiple measurements

Super-resolution of clustered sources

Aim : Super-resolving signals structured in small clusters

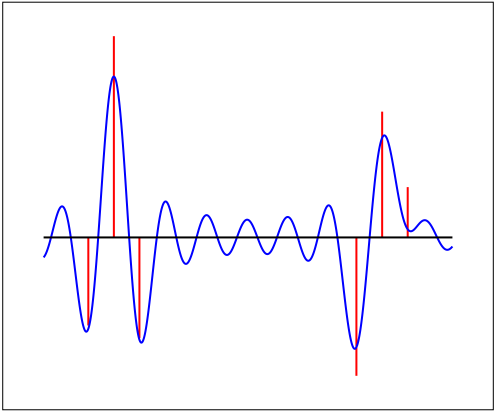
Super-resolution of clustered sources



Clustered point sources

Minimum separation $= 0.6\,\lambda_c$, SNR $= 25$ dB

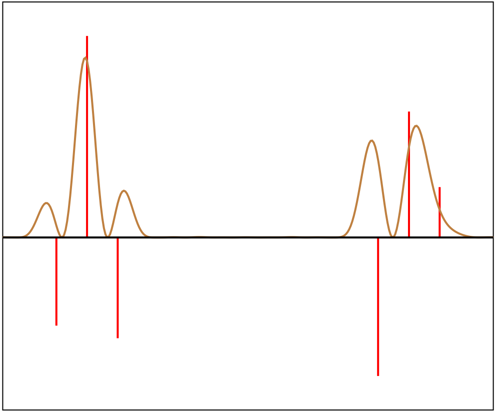
Super-resolution of clustered sources



Clustered point sources

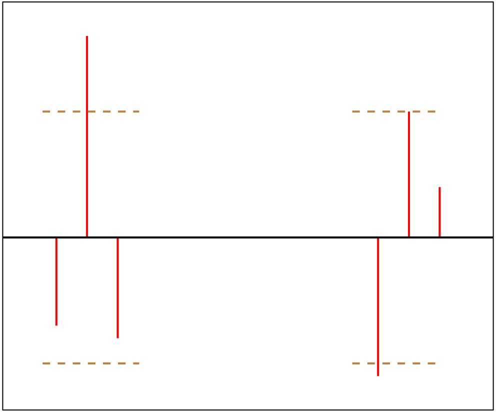
Minimum separation $= 0.6\,\lambda_c$, SNR $= 25$ dB

Super-resolution of clustered sources



Computing a coarse estimate S of the support is easy

Super-resolution of clustered sources



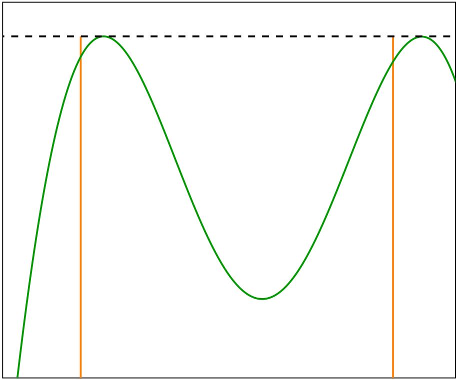
Computing a coarse estimate S of the support is easy

Super-resolution of clustered sources

Support-locating polynomial obtained from solving the dual of

$$\min_{\tilde{x}} ||\tilde{x}||_{\mathrm{TV}} \quad \text{subject to} \quad ||\mathcal{F}_c\, \tilde{x} - y||_2 \leq \delta$$

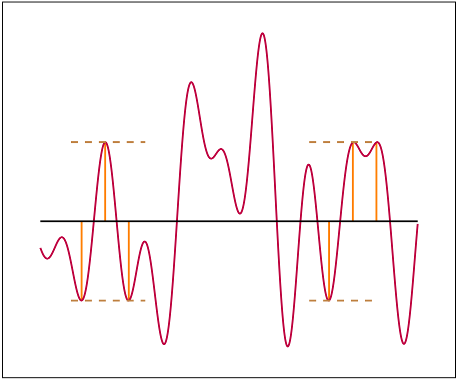
Super-resolution of clustered sources



Support-locating polynomial obtained from solving the dual of

$$\min_{\tilde{x}} ||\tilde{x}||_{\mathrm{TV}} \quad \text{subject to} \quad ||\mathcal{F}_c\, \tilde{x} - y||_2 \leq \delta$$

Super-resolution of clustered sources

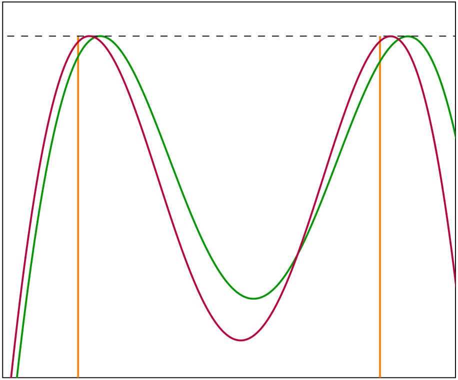


Support-locating polynomial obtained from solving the dual of

$$\min_{\tilde{x}} ||\tilde{x}||_{\text{TV}} \quad \text{subject to} \quad ||\mathcal{F}_c\, \tilde{x} - y||_2 \leq \delta \quad x_{S^c} = 0$$

The magnitude of the polynomial is only constrained on S

Super-resolution of clustered sources

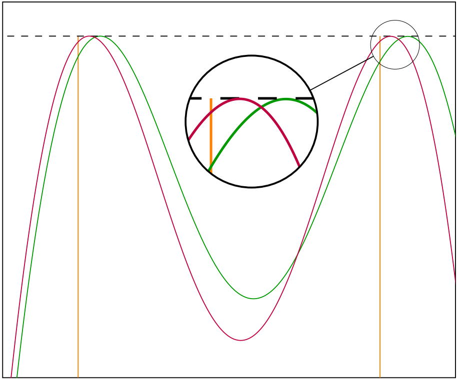


Support-locating polynomial obtained from solving the dual of

$$\min_{\tilde{x}} ||\tilde{x}||_{\text{TV}} \quad \text{subject to} \quad ||\mathcal{F}_c \, \tilde{x} - y||_2 \leq \delta \quad x_{S^c} = 0$$

The magnitude of the polynomial is only constrained on S

Super-resolution of clustered sources

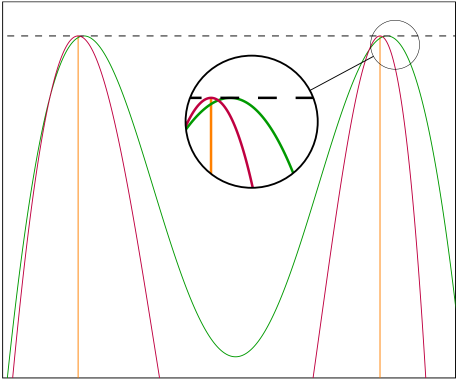


Support-locating polynomial obtained from solving the dual of

$$\min_{\tilde{x}} ||\tilde{x}||_{\text{TV}} \quad \text{subject to} \quad ||\mathcal{F}_c\, \tilde{x} - y||_2 \leq \delta \quad x_{S^c} = 0$$

The magnitude of the polynomial is only constrained on S

Super-resolution of clustered sources

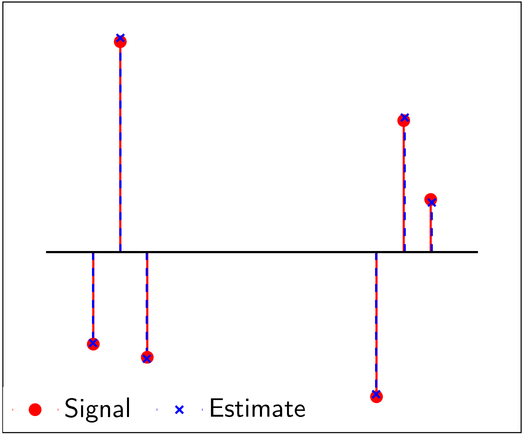


Support-locating polynomial obtained from solving the dual of

$$\min_{\tilde{x}} ||\tilde{x}||_{\mathrm{TV}} \quad \text{subject to} \quad ||\mathcal{F}_c\, \tilde{x} - y||_2 \leq \delta \quad x_{S^c} = 0$$

Without noise, we have exact recovery

Super-resolution of clustered sources



Joint work with Raf Mertens (Stanford)

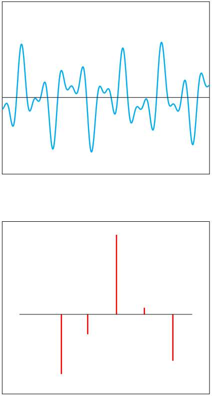
Demixing of sines and spikes

Aim : Super-resolving the spectrum of a multi-sinusoidal signal (sines) in the presence of impulsive events (spikes)

Demixing of sines and spikes

Sines

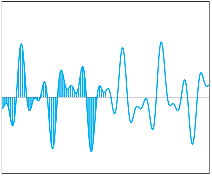
Spectrum



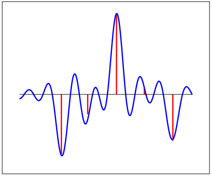
x

Demixing of sines and spikes

Sines



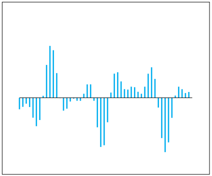
Spectrum



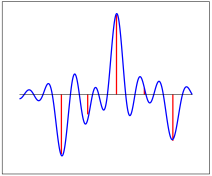
$\mathcal{F}_c \, x$

Demixing of sines and spikes

Sines

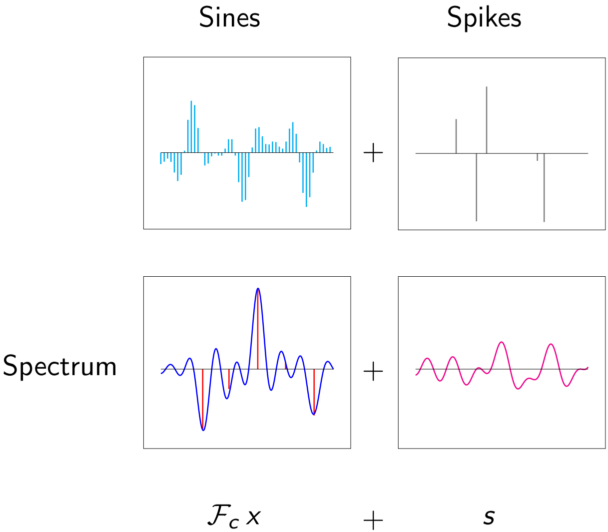


Spectrum

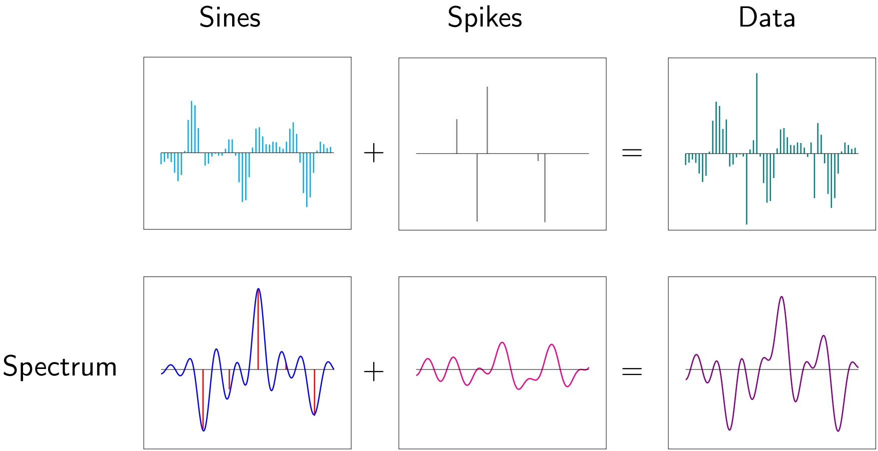


$\mathcal{F}_c \, x$

Demixing of sines and spikes



Demixing of sines and spikes



$$\mathcal{F}_c \, x \quad + \quad s \quad = \quad y$$

Demixing of sines and spikes

Estimator : Solution to

$$\min_{\tilde{x},\,\tilde{s}} ||\tilde{x}||_{\mathrm{TV}} + \gamma\,||\tilde{s}||_1 \quad \text{subject to} \quad \mathcal{F}_c\,\tilde{x} + \tilde{s} = y$$

Dual problem :

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

Demixing of sines and spikes

Estimator : Solution to

$$\min_{\tilde{x},\,\tilde{s}} ||\tilde{x}||_{\text{TV}} + \gamma\,||\tilde{s}||_1 \quad \text{subject to} \quad \mathcal{F}_c\,\tilde{x} + \tilde{s} = y$$

Dual problem :

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

Dual solution : $\hat{u}$

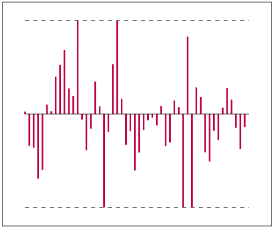
- $\hat{u}$ interpolates the sign of the primal solution $\hat{s}$
- $\mathcal{F}_c^*\,\hat{u}$ interpolates the sign of the primal solution $\hat{x}$

Demixing of sines and spikes

$\hat{u}$

$\mathcal{F}_c^* \, \hat{u}$

Dual solution

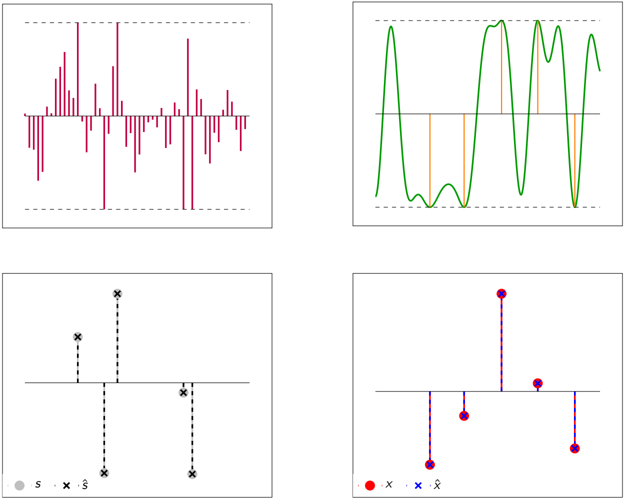


Demixing of sines and spikes

$\hat{u}$ | $\mathcal{F}_c^* \, \hat{u}$

Dual solution

Estimate



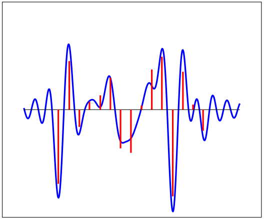
Spikes | Sines (spectrum)

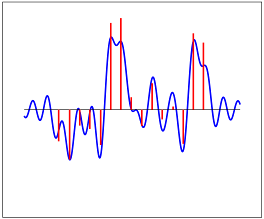
Super-resolution from multiple measurements

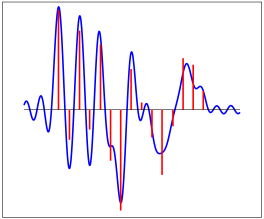
Aim : Super-resolving K signals with the same support

Motivation : Fluorescence microscopy (PALM, STORM), astronomy and communications

Super-resolution from multiple measurements



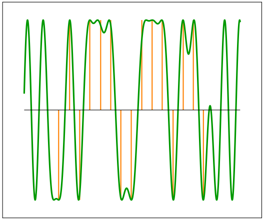




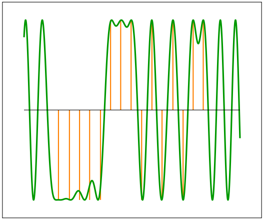
Minimum separation $= 0.7\,\lambda_c$

Super-resolution from multiple measurements

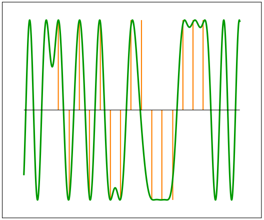
Dual sol. (signal 1)



Dual sol. (signal 2)



Dual sol. (signal 3)



Dual solutions obtained by solving separate problems

Group total variation

Estimator : Solution to minimizing group total-variation norm

- Continuous analog of $\ell_1 - \ell_2$ norm
- Promotes group sparsity
- If $X = \{x_1, x_2, x_3\}$, $a(t_j) \in \mathbb{C}^3$ for each $t_j \in T$ and

$$x_k = \sum_{t_j \in T} a(t_j)_k \, \delta_{t_j} \qquad \text{then} \qquad ||X||_{\text{GTV}} = \sum_{t_j \in T} ||a(t_j)||_2$$

Group total variation

Estimator : Solution to minimizing group total-variation norm

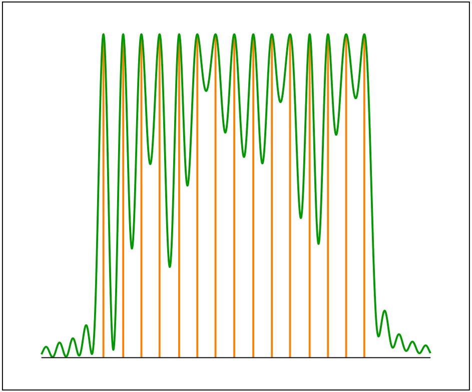
- Continuous analog of $\ell_1 - \ell_2$ norm
- Promotes group sparsity
- If $X = \{x_1, x_2, x_3\}$, $a(t_j) \in \mathbb{C}^3$ for each $t_j \in T$ and

$$x_k = \sum_{t_j \in T} a(t_j)_k \, \delta_{t_j} \qquad \text{then} \qquad ||X||_{\text{GTV}} = \sum_{t_j \in T} ||a(t_j)||_2$$

Dual solution :

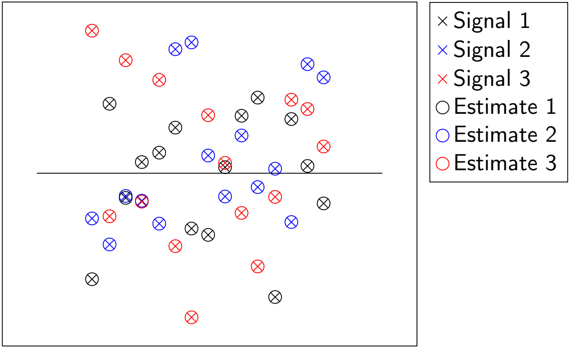
K-dimensional low-pass polynomial with unit magnitude on the estimate of the common support

Super-resolution from multiple measurements



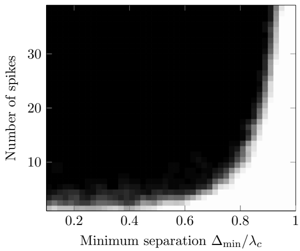
The estimator locates the support exactly

Super-resolution from multiple measurements

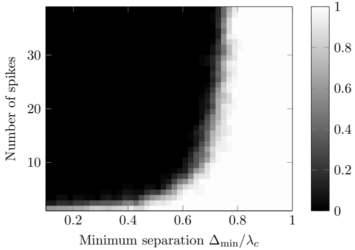


Minimum separation : As K grows, $\Delta_{\text{min}} \to \lambda_c/2$

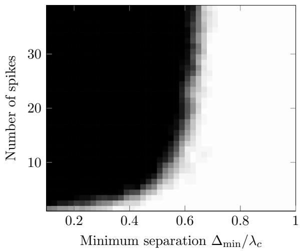
$K = 1$ (real amplitudes)



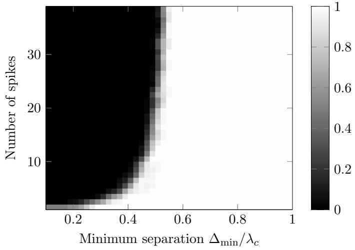
$K = 1$ (complex amplitudes)



$K = 2$ (complex amplitudes)



$K = 10$ (complex amplitudes)



Conclusion

Convex programming is a powerful tool for estimation from low-res data :

- Precise theoretical analysis
- Non-asymptotic stability guarantees
- Flexible framework

Conclusion

Convex programming is a powerful tool for estimation from low-res data :

- Precise theoretical analysis
- Non-asymptotic stability guarantees
- Flexible framework

Lots of work to do :

- Developing fast sdp solvers exploiting the structure in the dual problem
- Deconvolution from irregular samples
- Super-resolution of 2D curves
- *Blind deconvolution* : joint estimation of signal + point-spread function

Research directions

Generic goal in modern data processing :

Finding **low-dimensional structure** in **high-dimensional data**

This talk : Understanding the interaction between the data acquisition mechanism and the low-dimensional structure pays off !

Future directions :

- Sparse regression with highly-correlated design matrices
  e.g. dictionary of decaying exponentials
- Statistical processing of projected data
  e.g. dimensionality reduction in big-data
- Data-driven regularization :
  e.g. transform-invariant regularizers in computer vision

Image upsampling via transform-invariant regularization

Input

Upsampled image

Aim : Achieving large upsampling factors through data-driven regularizers that are approximately invariant to the projection onto the imaging plane

For more details

- **Towards a mathematical theory of super-resolution**. E. J. Candès and C. Fernandez-Granda. *Communications on Pure and Applied Math.*
- **Super-resolution from noisy data**. E. J. Candès and C. Fernandez-Granda. *Journal of Fourier Analysis and Applications* **19** (6), 1229-1254.
- **Support detection in super-resolution**. C. Fernandez-Granda. *Proceedings of SampTA 2013*, 145-148.
- **Super-resolution of point sources via convex programming**. C. Fernandez-Granda. Preprint.

Thank you