Towards a Mathematical Theory of Super-resolution

Carlos Fernandez-Granda

www.stanford.edu/~cfgranda/

Information Theory Forum, Information Systems Laboratory, Stanford

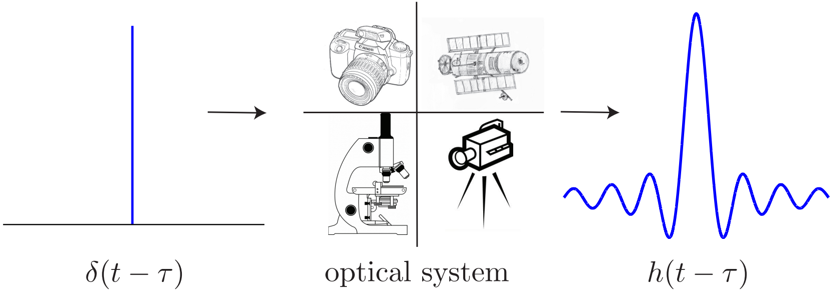
10/18/2013

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

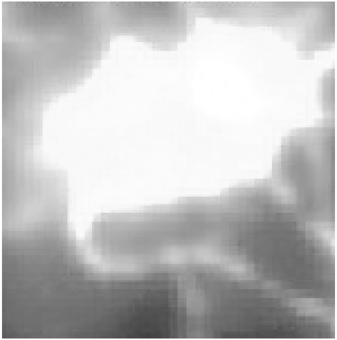
The resolving power of lenses, however perfect, is limited (**Lord Rayleigh**)

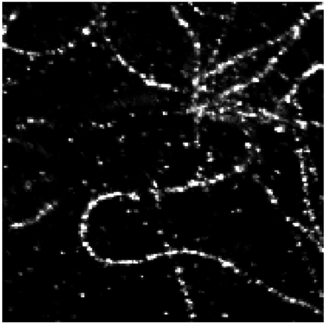


Diffraction imposes a fundamental limit on the resolution of optical systems

Motivation

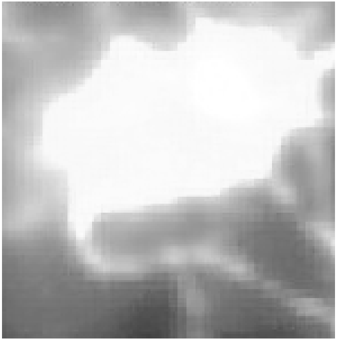
Similar problems arise in electronic imaging, signal processing, radar, spectroscopy, medical imaging, astronomy, geophysics, etc.

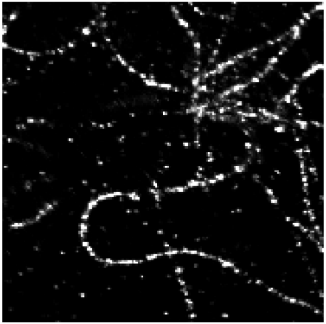




Motivation

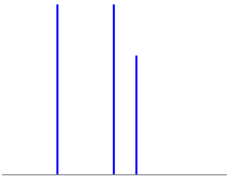
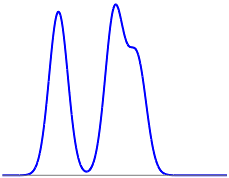
Similar problems arise in electronic imaging, signal processing, radar, spectroscopy, medical imaging, astronomy, geophysics, etc.





Signals of interest are often point sources : celestial bodies (astronomy), line spectra (signal processing), molecules (fluorescence microscopy), ...

Super-resolution

Aim : estimating fine-scale structure from low-resolution data

Super-resolution

Aim : estimating fine-scale structure from low-resolution data

Equivalently, extrapolating the high end of the spectrum

Mathematical model

- Signal : superposition of Dirac measures with support T

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

Mathematical model

- Signal : superposition of Dirac measures with support T

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

- Measurements : low-pass filtering with cut-off frequency f_c

$$y = \mathcal{F}_c\, x \quad \text{(vector of low-pass Fourier coefficients)}$$

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

Equivalent problem : line-spectra estimation

Swapping time and frequency

- Signal : superposition of sinusoids

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

Equivalent problem : line-spectra estimation

Swapping time and frequency

- Signal : superposition of sinusoids

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

- Measurements : equisposed samples

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

Equivalent problem : line-spectra estimation

Swapping time and frequency

- Signal : superposition of sinusoids

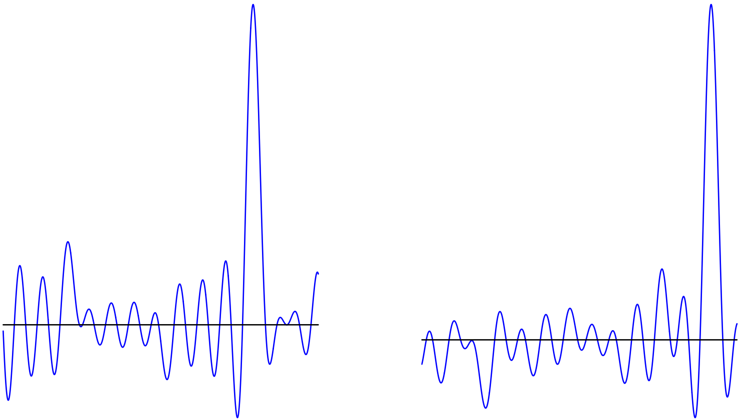
$$x(t) = \sum_j a_j e^{i2\pi f_j t} \qquad a_j \in \mathbb{C},\ f_j \in T \subset [0,1]$$

- Measurements : equispaced samples

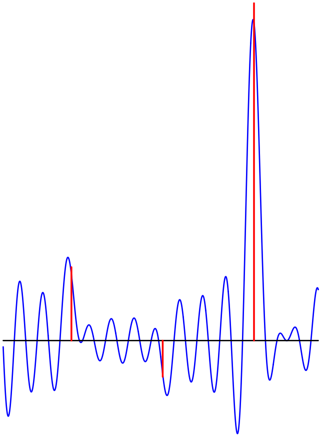
$$x(1), x(2), x(3), \ldots x(n)$$

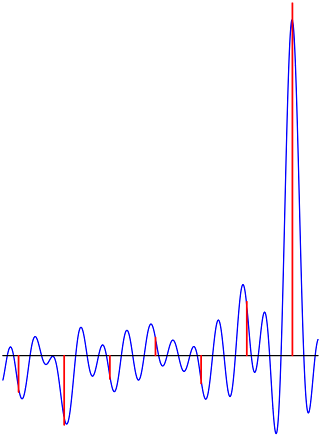
- Classical problem in signal processing

Can you find the spikes?



Can you find the spikes?





Prior art

Based on Prony's method : MUSIC, ESPRIT, matrix pencil, . . .

Prior art

Based on Prony's method : MUSIC, ESPRIT, matrix pencil, ...

- An estimate of the number of spikes is needed

Prior art

Based on Prony's method : MUSIC, ESPRIT, matrix pencil, ...

- An estimate of the number of spikes is needed
- Do not apply to multiple dimensions (not even 2D)

Prior art

Based on Prony's method : MUSIC, ESPRIT, matrix pencil, ...

- An estimate of the number of spikes is needed
- Do not apply to multiple dimensions (not even 2D)
- Noise must be Gaussian and white

Prior art

Based on Prony's method : MUSIC, ESPRIT, matrix pencil, . . .

- An estimate of the number of spikes is needed
- Do not apply to multiple dimensions (not even 2D)
- Noise must be Gaussian and white
- Incorporating additional structure is difficult (different point-spread functions, assumptions on the signal)

Prior art

Based on Prony's method : MUSIC, ESPRIT, matrix pencil, . . .

- An estimate of the number of spikes is needed
- Do not apply to multiple dimensions (not even 2D)
- Noise must be Gaussian and white
- Incorporating additional structure is difficult (different point-spread functions, assumptions on the signal)

This talk

- Non-parametric estimation

Prior art

Based on Prony's method : MUSIC, ESPRIT, matrix pencil, . . .

- An estimate of the number of spikes is needed
- Do not apply to multiple dimensions (not even 2D)
- Noise must be Gaussian and white
- Incorporating additional structure is difficult (different point-spread functions, assumptions on the signal)

This talk

- Non-parametric estimation
- Provably stable in the presence of noise

Prior art

Based on Prony's method : MUSIC, ESPRIT, matrix pencil, ...

- An estimate of the number of spikes is needed
- Do not apply to multiple dimensions (not even 2D)
- Noise must be Gaussian and white
- Incorporating additional structure is difficult (different point-spread functions, assumptions on the signal)

This talk

- Non-parametric estimation
- Provably stable in the presence of noise
- Flexible variational framework based on convex programming

Outline of the talk

Sparsity is not enough

Theory

Proof (sketch)

Implementation via semidefinite programming

Robustness to noise

Sparsity is not enough

Theory

Proof (sketch)

Implementation via semidefinite programming

Robustness to noise

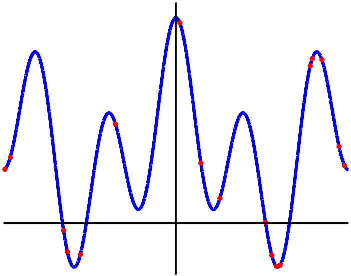
Compressed sensing vs super-resolution

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

Compressed sensing vs super-resolution

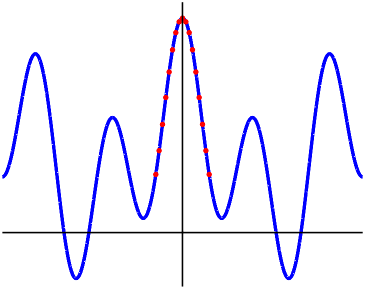
Estimation of sparse signals from undersampled measurements suggests connections to compressed sensing

Compressed sensing



spectrum interpolation

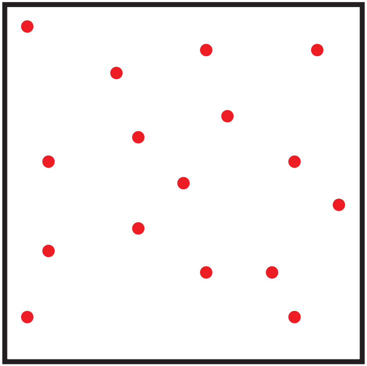
Super-resolution



spectrum extrapolation

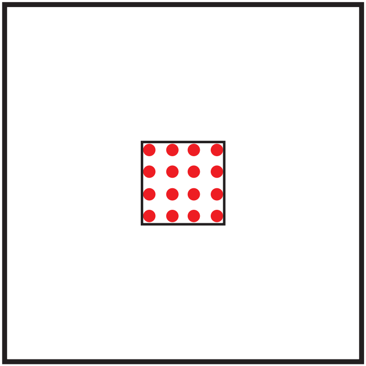
Compressed sensing vs super-resolution

Compressed sensing



spectrum interpolation

Super-resolution



spectrum extrapolation

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

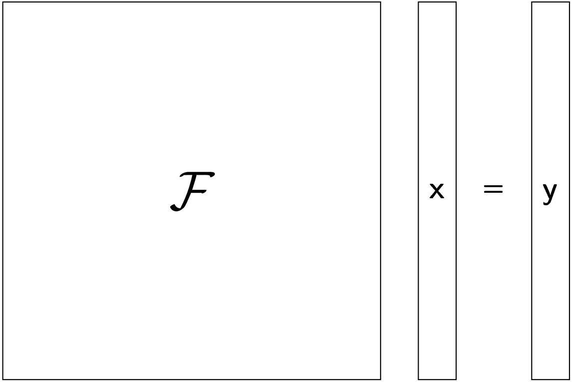
Compressed sensing

- 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
- Equivalently, the energy of all sparse signals is preserved by the randomized measurements (restricted isometry property)

Compressed sensing

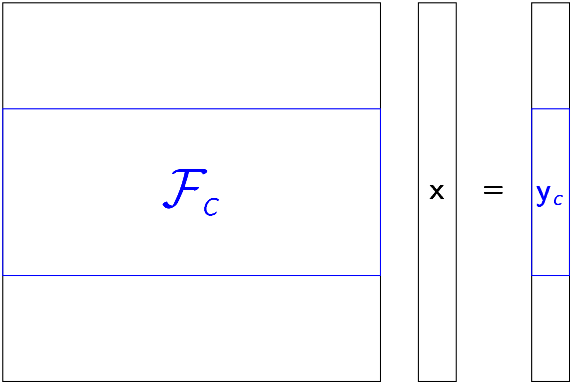
- Compressed sensing theory establishes robust recovery of spikes from random Fourier measurements [Candès, Romberg & Tao 2004]
- **Crucial insight** : measurement operator is well conditioned when acting upon sparse signals
- Equivalently, the energy of all sparse signals is preserved by the randomized measurements (restricted isometry property)
- This is a necessary condition for stable estimation, but is it the case in super-resolution ?

Simple experiment



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

Simple experiment



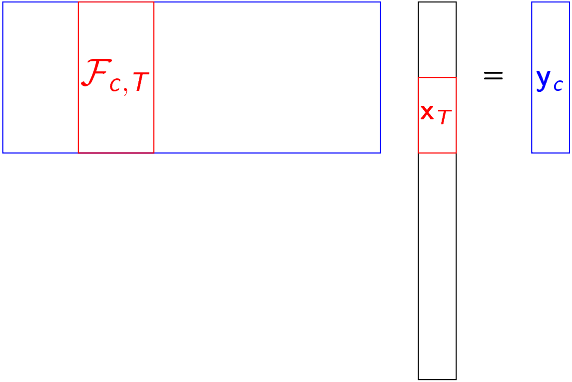
Measure n low-pass DFT coefficients, super-resolution factor (SRF) : N/n

Simple experiment

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

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

Simple experiment



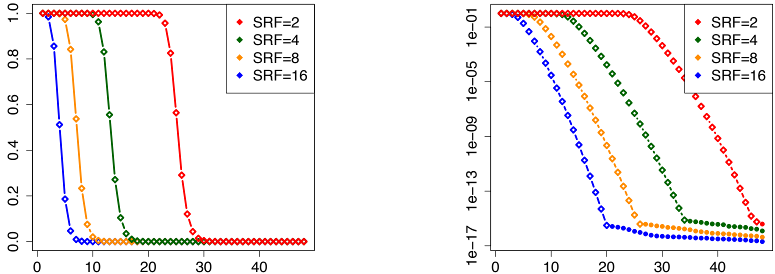
Restrict support of the signal to an interval of 48 contiguous points

Simple experiment

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

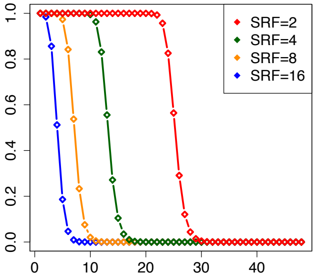
Compute singular values of resulting linear operator

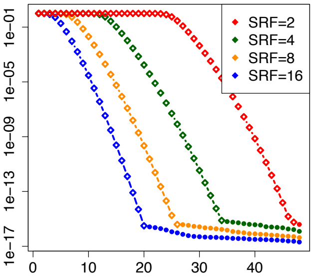
Simple experiment



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

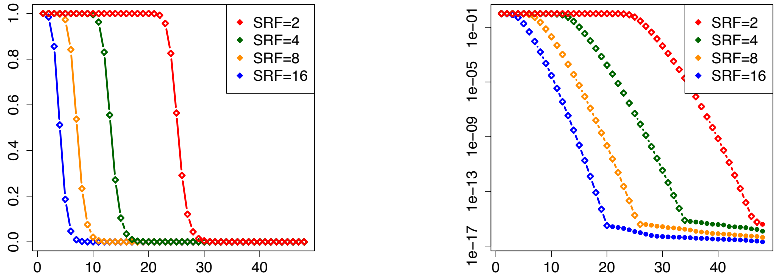




For SRF $= 4$ measuring any unit-normed vector in a subspace of dimension 24 results in a signal with norm less than 10^{-7}

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

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}

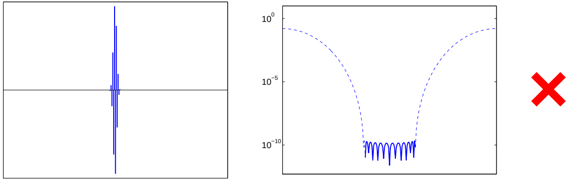
At an SNR of 145 dB, recovery is impossible by any method

This phenomenon is characterized asymptotically by Slepian's seminal work on prolate spheroidal sequences

Conclusion

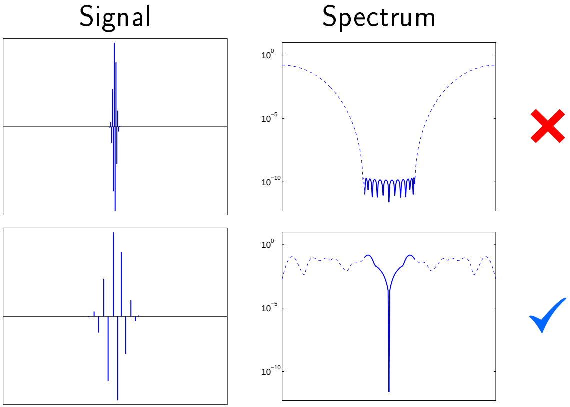
Sparsity is not enough

Signal　　　Spectrum



Conclusion

Sparsity is not enough



Additional conditions are necessary to restrict our signal model

Sparsity is not enough

Theory

Proof (sketch)

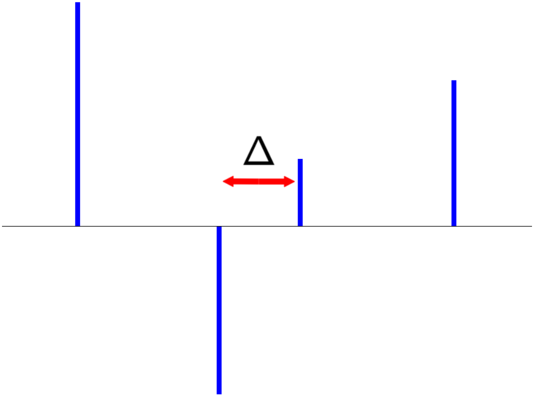
Implementation via semidefinite programming

Robustness to noise

Minimum separation

To exclude highly-clustered signals from our model, we control the minimum separation Δ of the support T

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



Total-variation norm

- Continuous counterpart of the ℓ_1 norm

Total-variation norm

- Continuous counterpart of the ℓ_1 norm
- If $x = \sum_j a_j \delta_{t_j}$ then $||x||_{\text{TV}} = \sum_j |a_j|$

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 $[0,1]$)

Recovery via convex programming

In the absence of noise, i.e. if $y = \mathcal{F}_c\, x$, 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 $[0, 1]$

Recovery via convex programming

In the absence of noise, i.e. if $y = \mathcal{F}_c\, x$, 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. 2012]

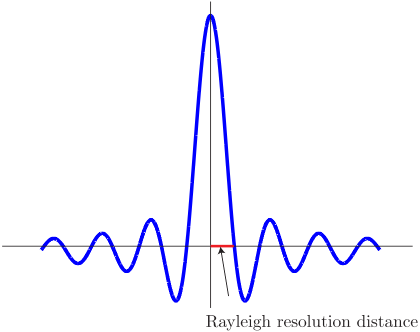
If the minimum separation of the signal support T obeys

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

then recovery is exact

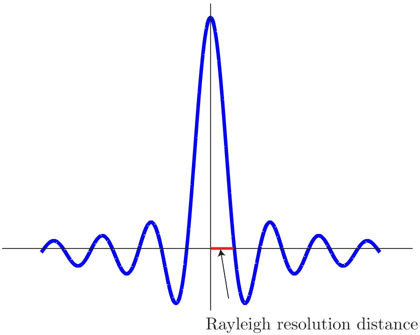
Minimum-distance condition

- $\lambda_c/2$ is the Rayleigh resolution limit (half-width of measurement filter)



Minimum-distance condition

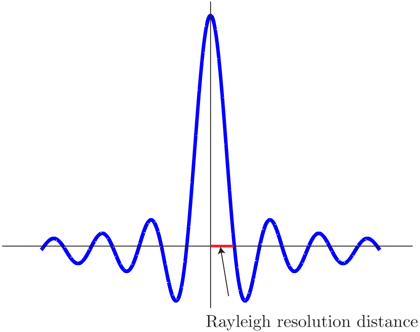
- $\lambda_c/2$ is the Rayleigh resolution limit (half-width of measurement filter)



- Numerical simulations show that TV-norm minimization fails if $\Delta < \lambda_c$

Minimum-distance condition

- $\lambda_c/2$ is the Rayleigh resolution limit (half-width of measurement filter)



- Numerical simulations show that TV-norm minimization fails if $\Delta < \lambda_c$
- If $\Delta < \lambda_c/2$ some signals are *almost* in the nullspace of the measurement operator (no method can achieve stable estimation)

Sparse recovery in overcomplete dictionaries

If we discretize the support

- Sparse recovery via ℓ_1-norm minimization in an overcomplete Fourier dictionary

Sparse recovery in overcomplete dictionaries

If we discretize the support

- Sparse recovery via ℓ_1-norm minimization in an overcomplete Fourier dictionary
- Previous theory based on dictionary incoherence is very weak, due to high column correlation

Sparse recovery in overcomplete dictionaries

If we discretize the support

- Sparse recovery via ℓ_1-norm minimization in an overcomplete Fourier dictionary
- Previous theory based on dictionary incoherence is very weak, due to high column correlation
- If $N = 20000$ and $n = 1000$, how many spikes can we recover ?

Sparse recovery in overcomplete dictionaries

If we discretize the support

- Sparse recovery via ℓ_1-norm minimization in an overcomplete Fourier dictionary
- Previous theory based on dictionary incoherence is very weak, due to high column correlation
- If $N = 20000$ and $n = 1000$, how many spikes can we recover ? Previous theory [Dossal 2005] : **3 spikes**

Sparse recovery in overcomplete dictionaries

If we discretize the support

- Sparse recovery via ℓ_1-norm minimization in an overcomplete Fourier dictionary
- Previous theory based on dictionary incoherence is very weak, due to high column correlation
- If $N = 20000$ and $n = 1000$, how many spikes can we recover ?
  Previous theory [Dossal 2005] : **3 spikes**
  Our result : $n/4 =$ **250 spikes**

Higher dimensions

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

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

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

Higher dimensions

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

Theorem [Candès, F. 2012]

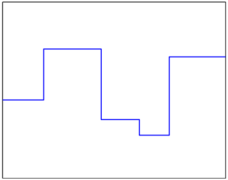
TV-norm minimization yields exact recovery if

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

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

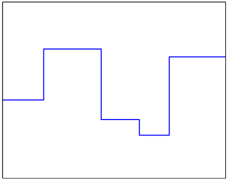
Extensions

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



Extensions

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



Corollary

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

yields exact recovery if $\Delta \geq 2\,\lambda_c$

Similar result for cont. differentiable piecewise-smooth functions

Sparsity is not enough

Theory

Proof (sketch)

Implementation via semidefinite programming

Robustness to noise

Optimality conditions

Consider the problem

$$\min_{\tilde{x}} f(\tilde{x}) \quad \text{subject to} \quad A\,\tilde{x} = y,$$

where f is convex

Lemma

If there exists a subgradient $g(x)$ of f at a feasible point x such that $g(x) = A^* v$ for some v, then x is a solution

Optimality conditions

Consider the problem

$$\min_{\tilde{x}} f(\tilde{x}) \quad \text{subject to} \quad A\tilde{x} = y,$$

where f is convex

Lemma

If there exists a subgradient $g(x)$ of f at a feasible point x such that $g(x) = A^* v$ for some v, then x is a solution

For any h such that $Ah = 0$

$$\begin{aligned} f(x+h) &\geq f(x) + \langle g(x), h \rangle \quad \text{by definition of subgradient and convexity of } f \\ &= f(x) + \langle A^* v, h \rangle \\ &= f(x) + \langle v, Ah \rangle \\ &= f(x) \end{aligned}$$

Certificate of optimality

q is a subgradient of the total-variation norm at $x = \sum_{j \in T} |a_j| \, e^{i\phi_j} \delta_{t_j}$ if

$$\begin{cases} q(t_j) = e^{i\phi_j}, & t_j \in T \\ |q(t)| \leq 1, & t \in [0,1] \setminus T \end{cases} \tag{1}$$

Certificate of optimality

q is a subgradient of the total-variation norm at $x = \sum_{j \in T} |a_j| \, e^{i\phi_j} \delta_{t_j}$ if

$$\begin{cases} q(t_j) = e^{i\phi_j}, & t_j \in T \\ |q(t)| \leq 1, & t \in [0,1] \setminus T \end{cases} \tag{1}$$

To certify optimality we also need

$$q(t) = \mathcal{F}_c^* v = \sum_{k=-f_c}^{f_c} v_k e^{i2\pi kt}$$

Certificate of optimality

q is a subgradient of the total-variation norm at $x = \sum_{j \in T} |a_j| \, e^{i\phi_j} \delta_{t_j}$ if

$$\begin{cases} q(t_j) = e^{i\phi_j}, & t_j \in T \\ |q(t)| \leq 1, & t \in [0,1] \setminus T \end{cases} \tag{1}$$

To certify optimality we also need

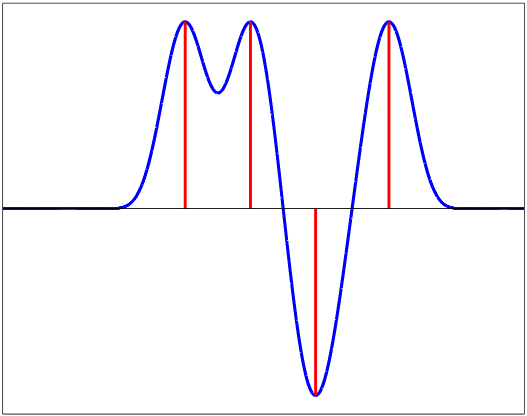
$$q(t) = \mathcal{F}_c^* v = \sum_{k=-f_c}^{f_c} v_k e^{i2\pi kt}$$

If (1) is strengthened to

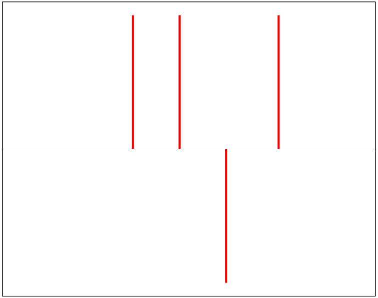
$$|q(t)| < 1, \quad t \in [0,1] \setminus T$$

then x is the unique solution

Certificate of optimality



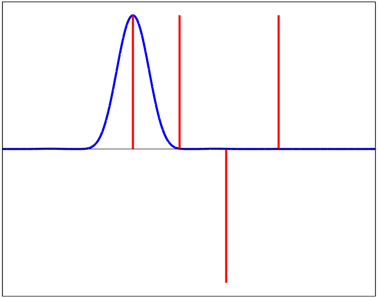
Construction of the certificate



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

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

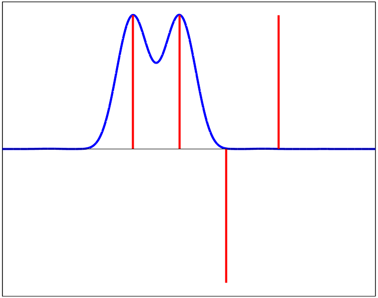
Construction of the certificate



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

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

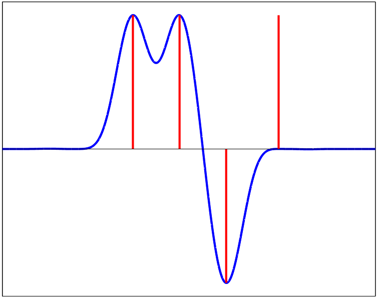
Construction of the certificate



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

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

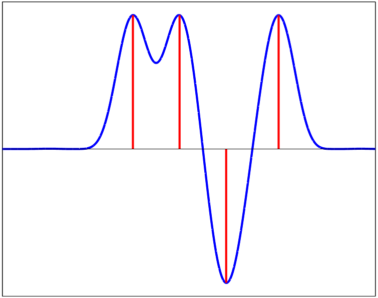
Construction of the certificate



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

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

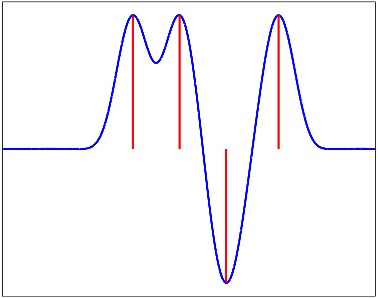
Construction of the certificate



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

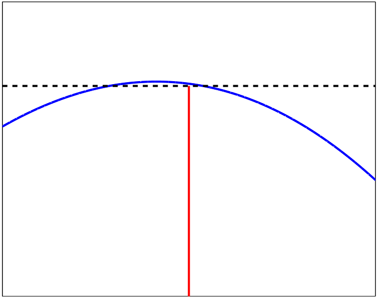
$$q(t) = \sum_{t_j \in T} \alpha_j K(t - t_j),$$

Construction of the certificate



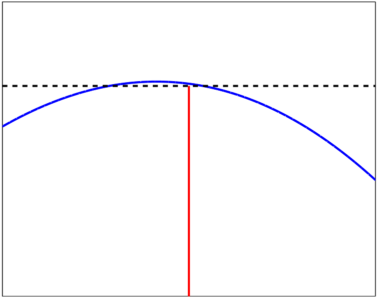
Problem : magnitude of polynomial locally exceeds 1

Construction of the certificate



Problem : magnitude of polynomial locally exceeds 1

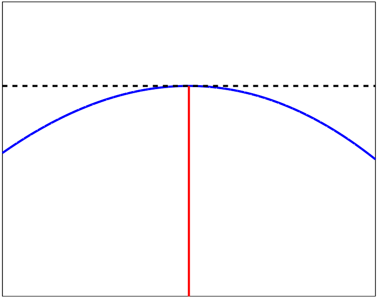
Construction of the certificate



Problem : magnitude of polynomial locally exceeds 1
Solution : add correction term and force $q'(t_k) = 0$ for all $t_k \in T$

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

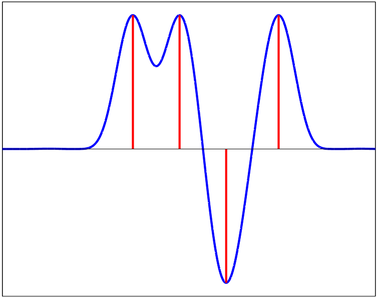
Construction of the certificate



Problem : magnitude of polynomial locally exceeds 1
Solution : add correction term and force $q'(t_k) = 0$ for all $t_k \in T$

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

Construction of the certificate



Problem : magnitude of polynomial locally exceeds 1

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

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

Sparsity is not enough

Theory

Proof (sketch)

Implementation via semidefinite programming

Robustness to noise

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 x (measure in $[0, 1]$)

First option : Discretize domain and apply 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 x (measure in $[0,1]$)

First option : Discretize domain and apply ℓ_1-norm minimization

- **Dual problem** :

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

Finite-dimensional variable u, but infinite-dimensional constraint

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

Second option : Recast dual problem as semidefinite program

Lemma : Semidefinite representation

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

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

is equivalent to

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

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

Using the dual solution

We can solve the dual problem, but how do we retrieve a primal solution ?

Using the dual solution

We can solve the dual problem, but how do we retrieve a primal solution ?

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

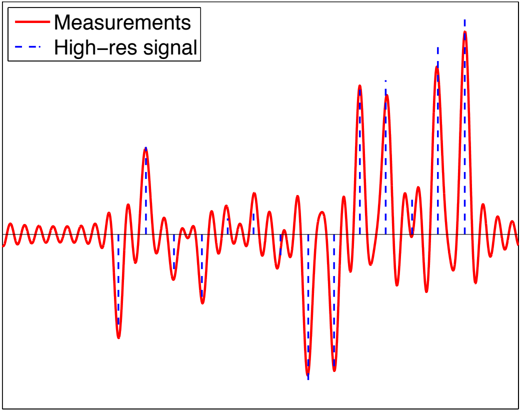
Using the dual solution

We can solve the dual problem, but how do we retrieve a primal solution ?

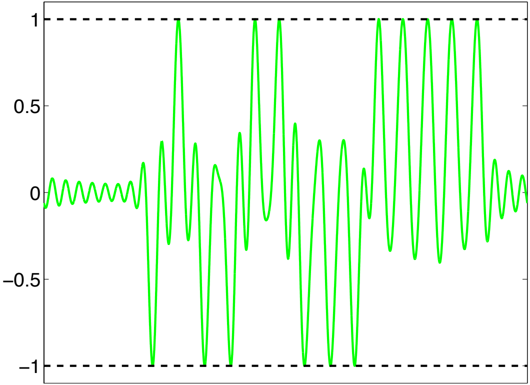
Dual solution vector : Fourier coefficients of low-pass polynomial that interpolates the sign of the primal solution (follows from strong duality)

Idea : Use the polynomial to locate the support of the signal and then estimate the amplitudes by least squares

Super-resolution via semidefinite programming

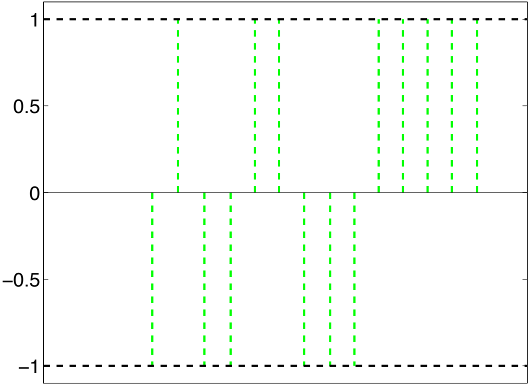


Super-resolution via semidefinite programming



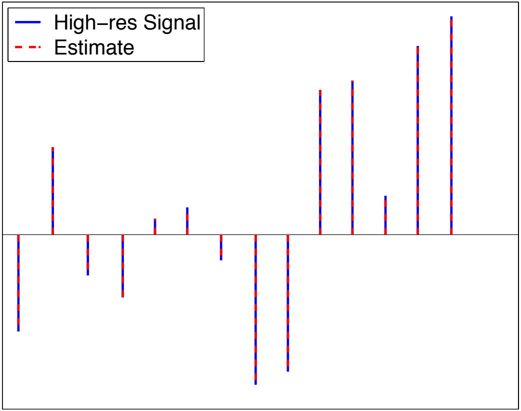
1. Solve semidefinite program to obtain dual solution

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

Sparsity is not enough

Theory

Proof (sketch)

Implementation via semidefinite programming

Robustness to noise

Estimation from noisy data

- Without noise, we achieve perfect precision, i.e. infinite resolution

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

Estimation from noisy data

- Without noise, we achieve perfect precision, i.e. infinite resolution

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

- In practice, there is always noise

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

Understanding the performance in a noisy setting is crucial for applications

Estimation from noisy data

- Without noise, we achieve perfect precision, i.e. infinite resolution

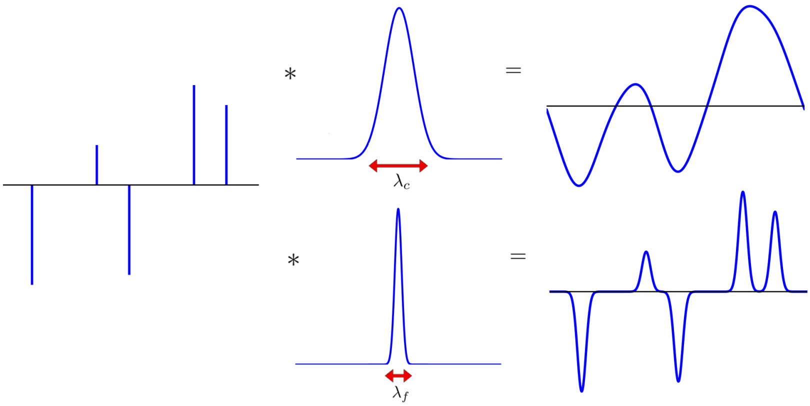
$$y = \mathcal{F}_c x$$

- In practice, there is always noise

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

  Understanding the performance in a noisy setting is crucial for applications
- Metrics :
  1. Approximation error at a higher resolution
  2. Support-detection error

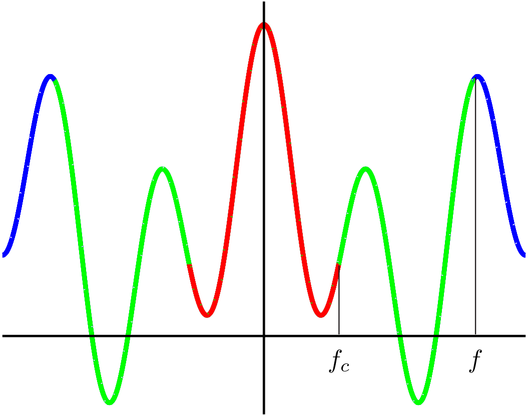
Super-resolution factor : spatial viewpoint



Super-resolution factor

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

Super-resolution factor : spectral viewpoint



Super-resolution factor

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

Approximation at a higher resolution

Resolution at scale λ is quantified by convolution with kernel ϕ_λ of width λ

At the resolution of the measurements

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

Approximation at a higher resolution

Resolution at scale λ is quantified by convolution with kernel ϕ_λ of width λ

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

Resolution at scale λ is quantified by convolution with kernel ϕ_λ of width λ

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]

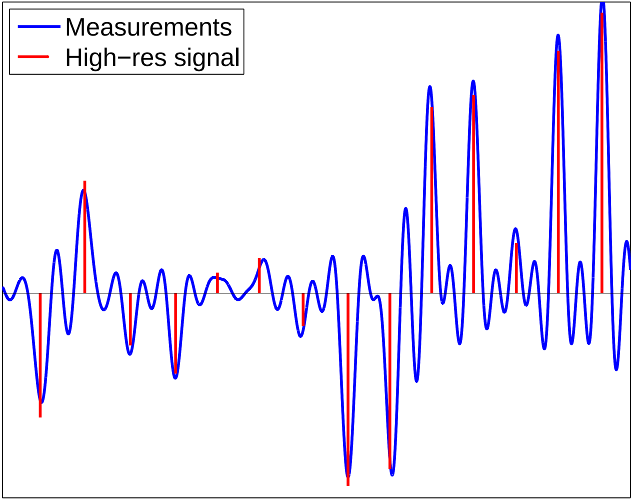
If $\Delta \geq 2/f_c$ then the solution $\hat{x}$ to

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

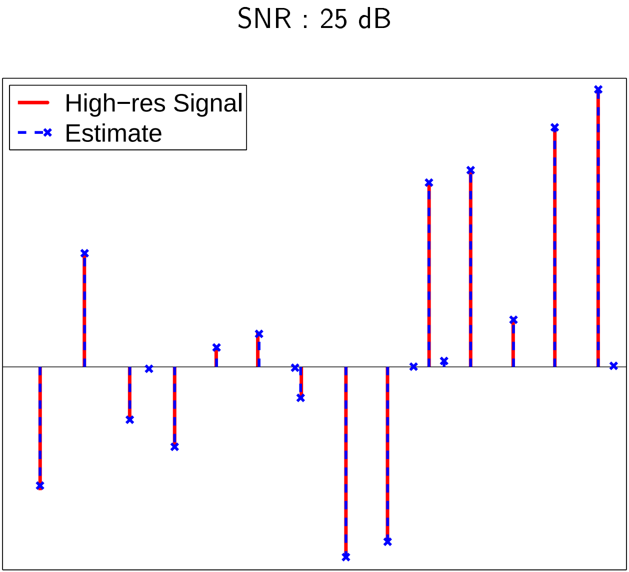
$$\text{satisfies} \quad \left|\left|\phi_{\lambda_f} * (\hat{x} - x)\right|\right|_{L_1} \lesssim \text{SRF}^2\, \delta$$

Example

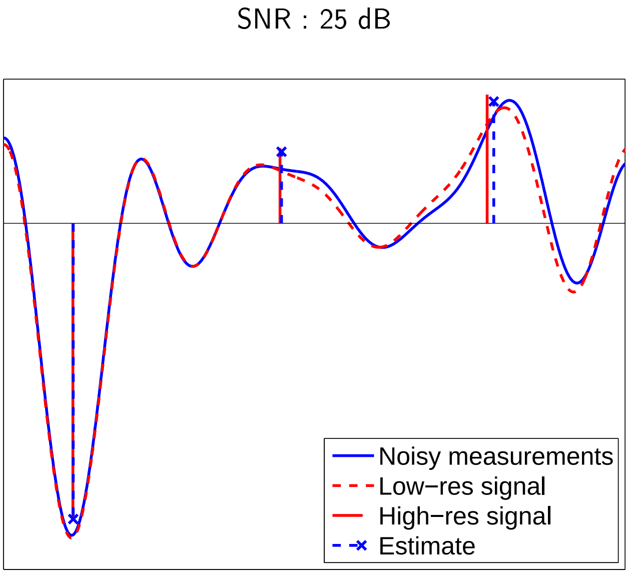
SNR : 25 dB



Example



Example


SNR : 25 dB
Noisy measurements
Low−res signal
High−res signal
Estimate

Support detection

- Original signal, support T

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

Support detection

- Original signal, support T

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

- Estimated signal, support $\widehat{T}$

$$\hat{x} = \sum_j \hat{a}_j \delta_{\hat{t}_j} \qquad \hat{a}_j \in \mathbb{C},\ \hat{t}_j \in \widehat{T} \subset [0,1]$$

Support detection

- Original signal, support T

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

- Estimated signal, support $\widehat{T}$

$$\hat{x} = \sum_j \hat{a}_j \delta_{\hat{t}_j} \qquad \hat{a}_j \in \mathbb{C},\ \hat{t}_j \in \widehat{T} \subset [0,1]$$

How accurately can we detect the support at a certain noise level δ?

Support-detection accuracy

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

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

Support-detection accuracy

Theorem [F. 2013]

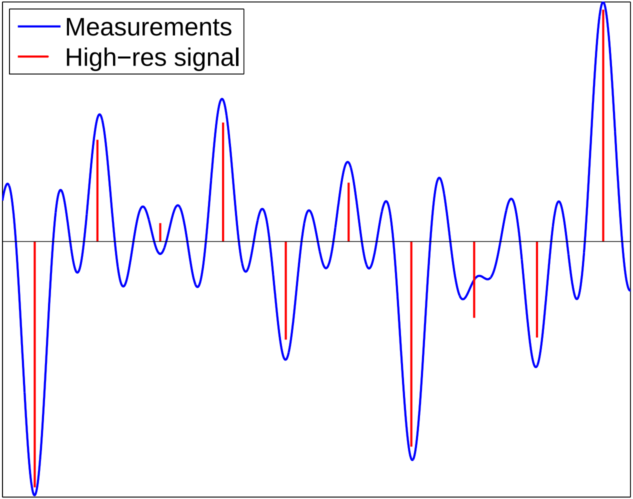
For any $t_i \in T$, if $|a_i| > C_1 \delta$ there exists $\hat{t}_i \in \widehat{T}$ such that

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

The support-detection accuracy is not affected by aliasing
(no dependence on the amplitude of the signal at other locations)

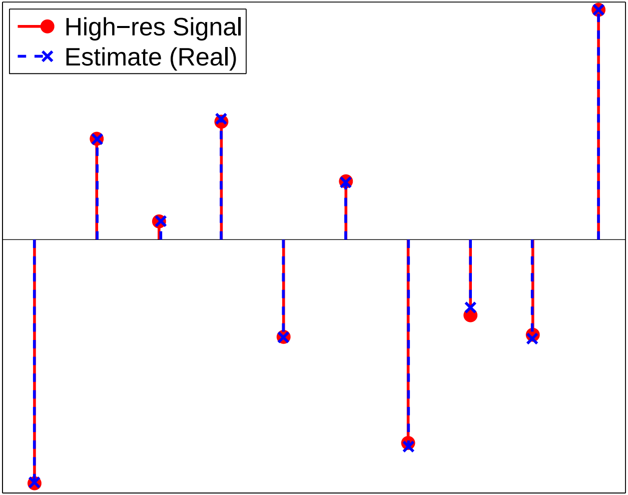
Consequence

Robustness of the algorithm to high dynamic ranges



Consequence

Robustness of the algorithm to high dynamic ranges



Conclusion and future work

- To obtain theoretical guarantees for super-resolution in realistic settings, we need conditions that avoid clustered supports

Conclusion and future work

- To obtain theoretical guarantees for super-resolution in realistic settings, we need conditions that avoid clustered supports
- Under a minimum-separation condition, convex programming achieves exact recovery

Conclusion and future work

- To obtain theoretical guarantees for super-resolution in realistic settings, we need conditions that avoid clustered supports
- Under a minimum-separation condition, convex programming achieves exact recovery
- The optimization problem can be recast as a tractable semidefinite program

Conclusion and future work

- To obtain theoretical guarantees for super-resolution in realistic settings, we need conditions that avoid clustered supports
- Under a minimum-separation condition, convex programming achieves exact recovery
- The optimization problem can be recast as a tractable semidefinite program
- The method is provably robust to noise

Conclusion and future work

- To obtain theoretical guarantees for super-resolution in realistic settings, we need conditions that avoid clustered supports
- Under a minimum-separation condition, convex programming achieves exact recovery
- The optimization problem can be recast as a tractable semidefinite program
- The method is provably robust to noise
- Research directions :
  - Super-resolution of images with sharp edges
  - Developing fast solvers to solve sdp formulation
  - Extending results to other overcomplete dictionaries

For more details

- **Towards a mathematical theory of super-resolution**. E. J. Candès and C. Fernandez-Granda. *Comm. on Pure and Applied Math.*
- **Super-resolution from noisy data**. E. J. Candès and C. Fernandez-Granda. *Journal of Fourier Analysis and Applications*
- **Support detection in super-resolution**. C. Fernandez-Granda. *Proceedings of SampTA 2013*
- **Prolate spheroidal wave functions, Fourier analysis, and uncertainty V - The discrete case**. D. Slepian. *Bell System Technical Journal*, 57 :1371-1430, 1978