Optimization-based sparse recovery: Compressed sensing vs. super-resolution

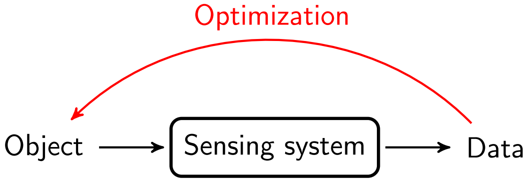
Carlos Fernandez-Granda, Google

Computational Photography and Intelligent Cameras, IPAM

2/5/2014

- This work was supported by a Fundación La Caixa Fellowship and a Fundación Caja Madrid Fellowship
- Joint work with Emmanuel Candès (Stanford)

Optimization-based recovery



Outline

Two inverse problems

When is the problem well posed?

When do optimization-based methods succeed?

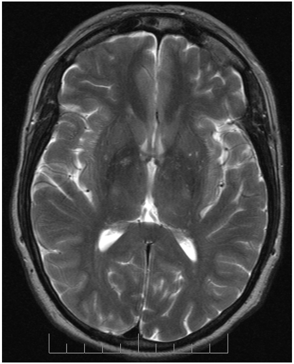
Two inverse problems

When is the problem well posed?

When do optimization-based methods succeed?

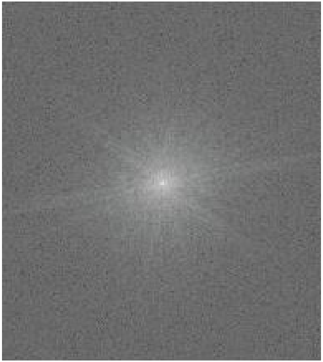
Compressed sensing

Object



Compressible

Data

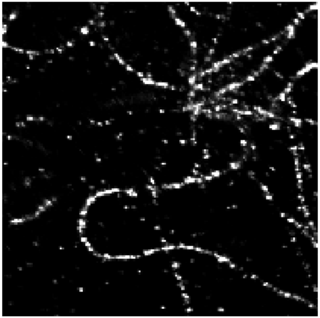


Randomized

Mathematical model: Random Fourier coefficients of sparse signal

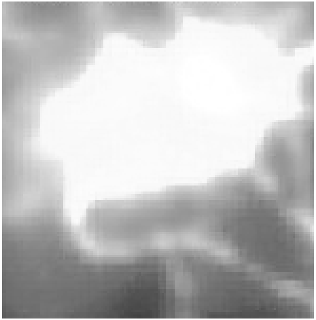
Super-resolution

Object



Point sources

Data



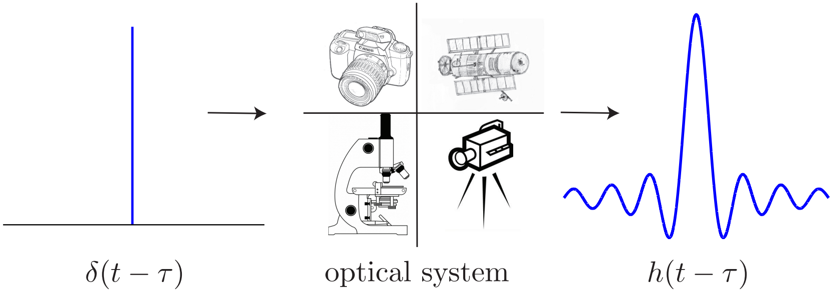
Low-pass blur

Mathematical model: Low-pass Fourier coefficients of sparse signal

(Figures courtesy of V. Morgenshtern)

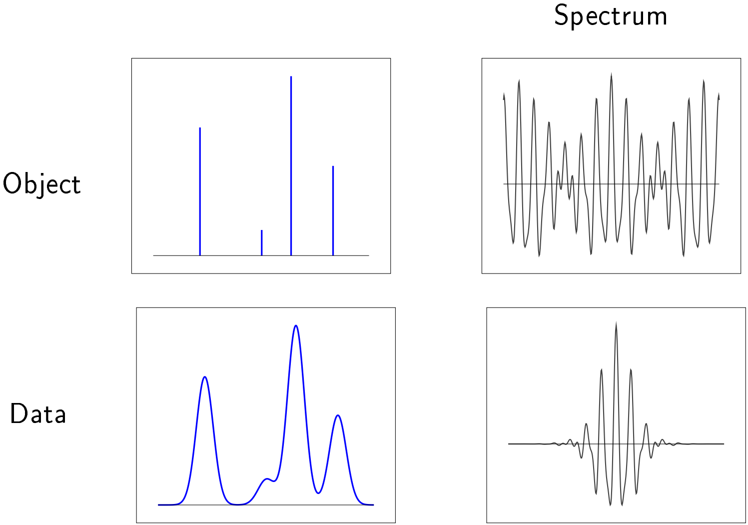
Motivation: Limits of resolution in imaging

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



Diffraction imposes a fundamental limit on the resolution of optical systems

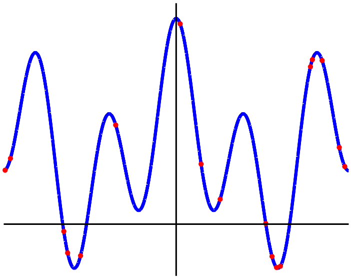
Super-resolution



Super-resolution

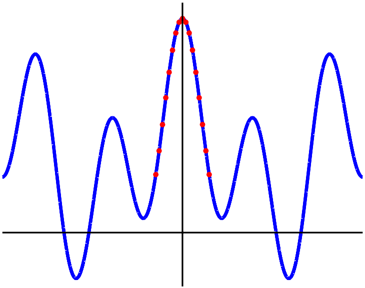
- **Optics**: Data-acquisition techniques to overcome the diffraction limit
- **Image processing**: Methods to upsample images onto a finer grid while preserving edges and hallucinating textures
- **This talk**: Estimation/deconvolution from low-pass measurements

Compressed sensing



Spectrum interpolation

Super-resolution



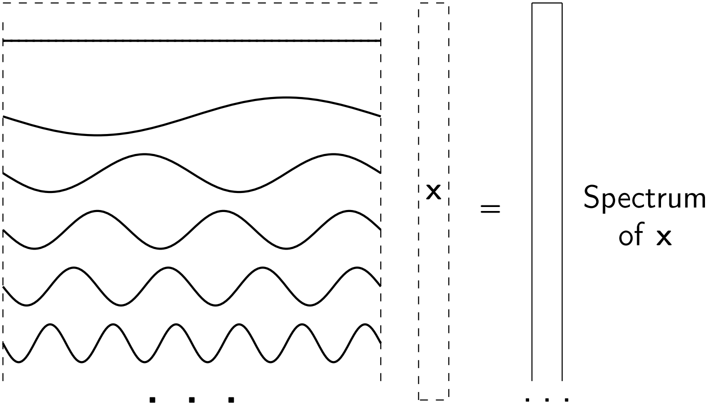
Spectrum extrapolation

Two inverse problems

When is the problem well posed?

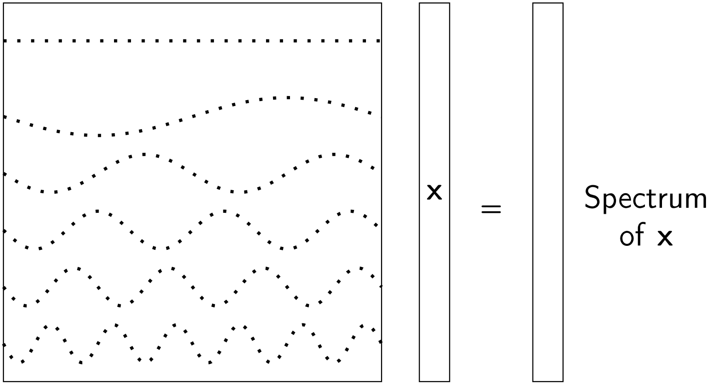
When do optimization-based methods succeed?

Compressed sensing



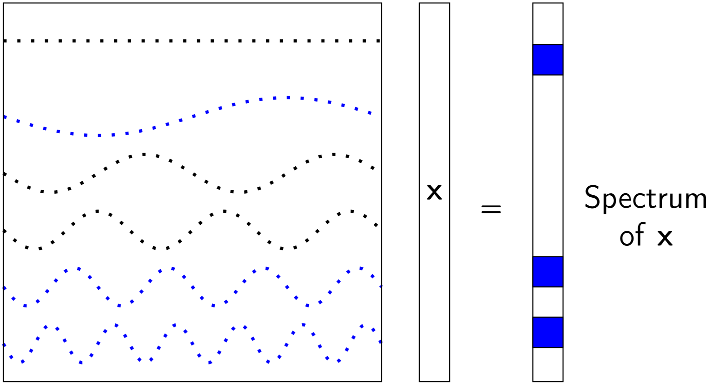
Fourier series of a measure/function $\mathbf{x}$ with domain $[0, 1]$

Compressed sensing



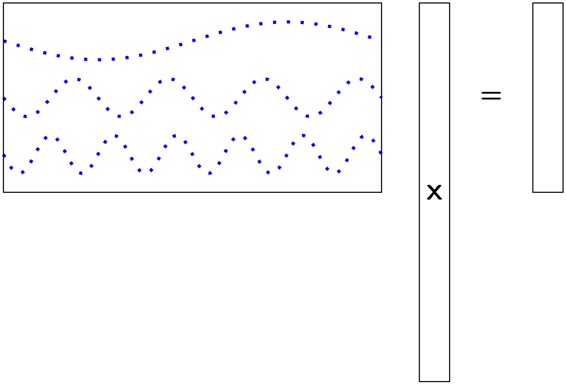
Discrete Fourier transform (DFT) of a vector $\mathbf{x} \in \mathbb{R}^N$

Compressed sensing



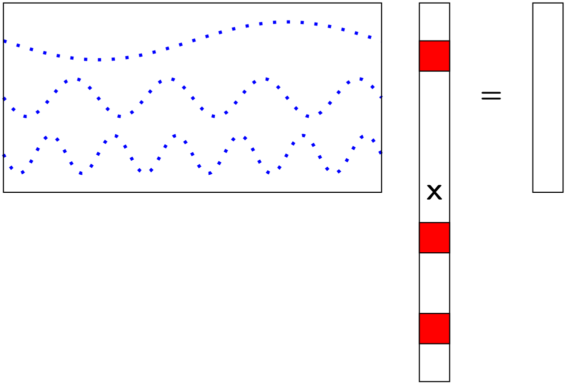
Data: Random DFT coefficients

Compressed sensing



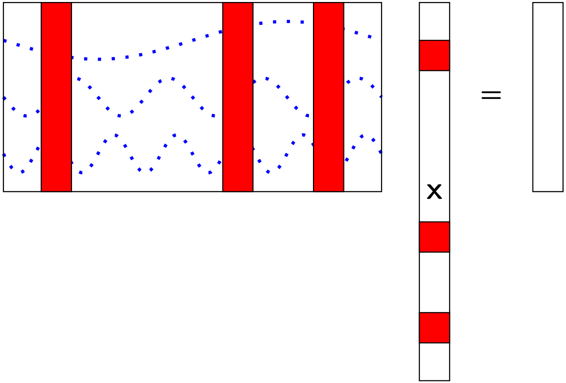
Data: Random DFT coefficients

Compressed sensing



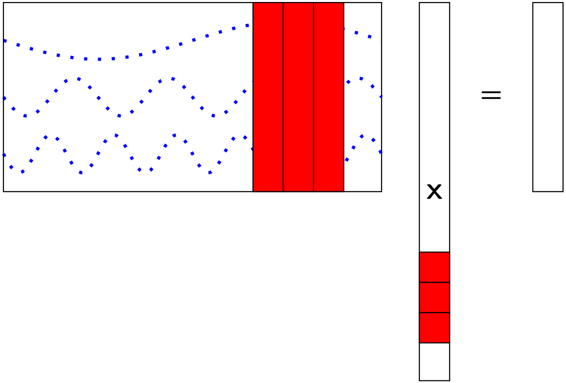
What is the effect of the measurement operator on sparse vectors?

Compressed sensing



Crucial insight: restricted operator is well conditioned when acting upon any sparse signal (*restricted isometry property*) [Candès, Tao 2006]

Compressed sensing



Crucial insight: restricted operator is well conditioned when acting upon any sparse signal (*restricted isometry property*) [Candès, Tao 2006]

Super-resolution

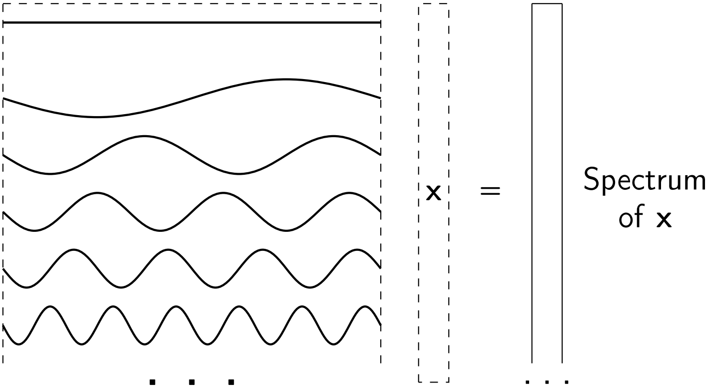
- **Signal**: superposition of spikes (Dirac measures) in the unit interval

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

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

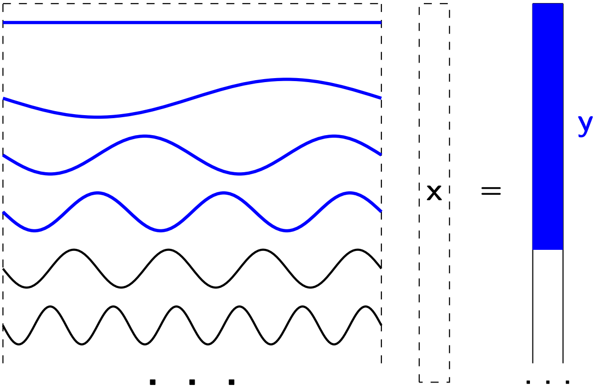
$$y(k) = \int_0^1 e^{-i2\pi kt} x(\mathrm{d}t) = \sum_j a_j e^{-i2\pi k t_j}, \quad k \in \mathbb{Z},\ |k| \leq f_c$$

Super-resolution



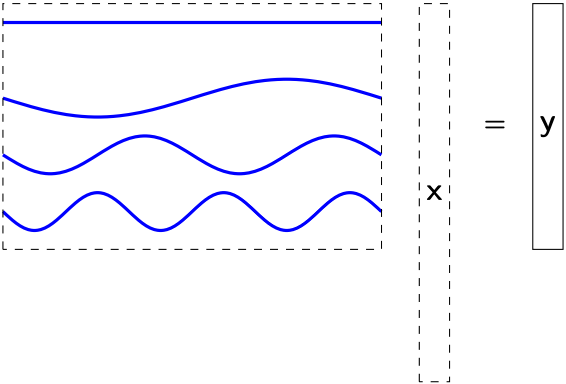
Fourier series of a measure $\mathbf{x}$ with domain $[0, 1]$

Super-resolution



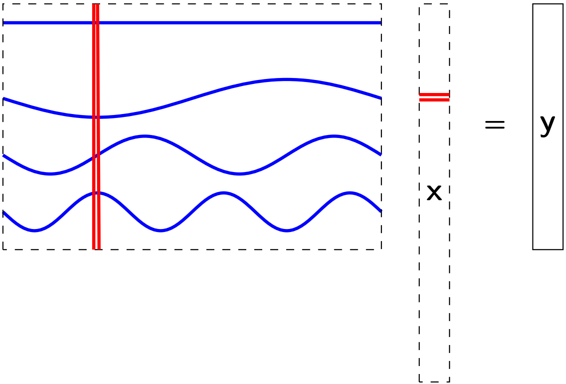
Data: Low-pass Fourier coefficients

Super-resolution



Data: Low-pass Fourier coefficients

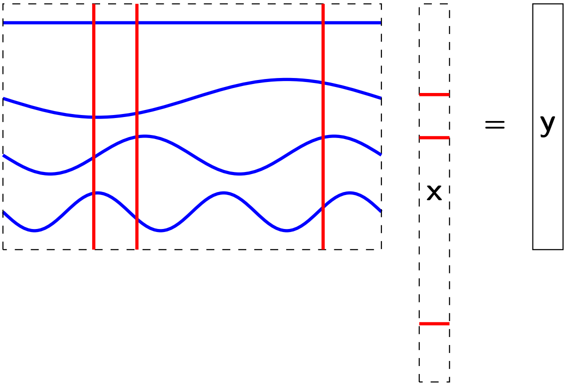
Super-resolution



Problem: If the support is clustered, the problem may be ill posed

In super-resolution sparsity is not enough!

Super-resolution



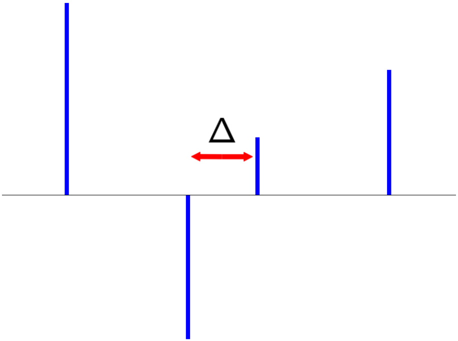
If the support is spread out, there is still hope

We need conditions beyond sparsity

Minimum separation

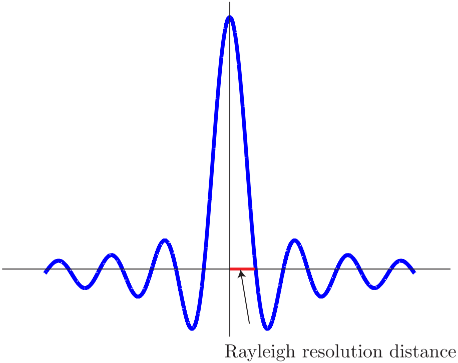
The minimum separation Δ of the support of x is

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



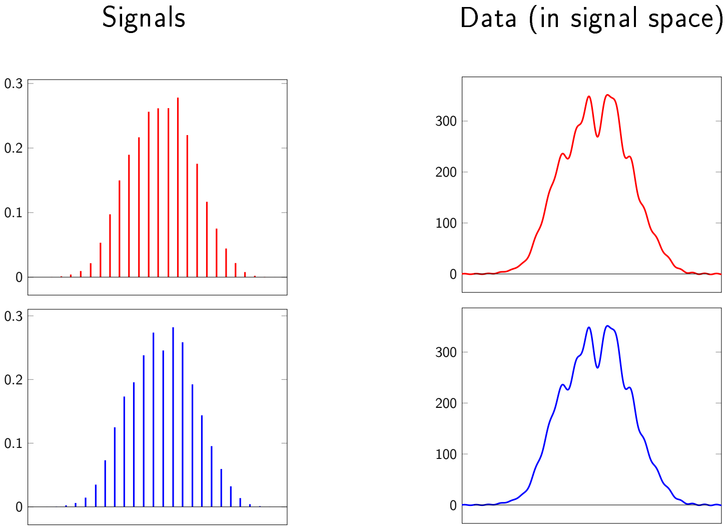
Conditioning of submatrix with respect to Δ

If $\Delta < \lambda_c := 1/f_c$ the problem is ill posed

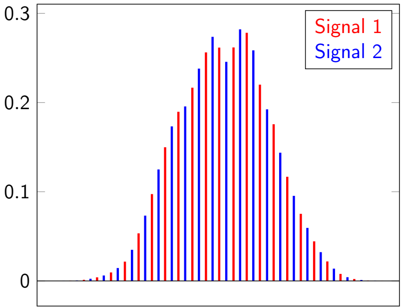


λ_c is the diameter of the main lobe of the point-spread function (twice the Rayleigh distance)

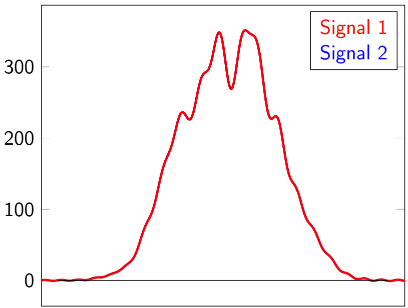
Example: 25 spikes, $f_c = 10^3$, $\Delta = 0.8/f_c$



Example: 25 spikes, $f_c = 10^3$, $\Delta = 0.8/f_c$



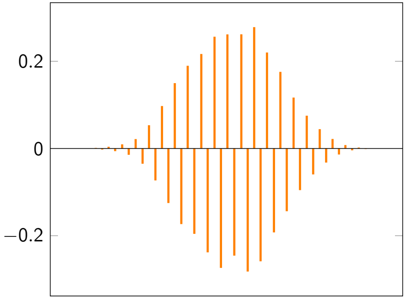
Signals



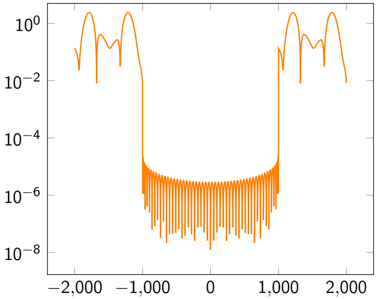
Data (in signal space)

Example: 25 spikes, $f_c = 10^3$, $\Delta = 0.8/f_c$

The difference is almost in the null space of the measurement operator



Difference



Spectrum

Two inverse problems

When is the problem well posed?

When do optimization-based methods succeed?

Estimation via convex programming

Measurement model: Underdetermined linear system (we need further assumptions)

$$Ax \approx y$$

Idea: Impose nonparametric assumptions on structure by minimizing a cost function

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

In our case: ℓ_1 norm to enforce sparsity

Subgradients

Definition: g_x is a subgradient of a convex function $\mathcal{C}$ at x if for any vector v

$$\mathcal{C}(x+v) \geq \mathcal{C}(x) + \langle g_x, v \rangle$$

Lemma

If there is a subgradient of $\mathcal{C}$ at x in the range of A^*, $g_x = A^* u$, and $Ax = y$ then x is a solution of

$$\min_{\tilde{x}} \quad \mathcal{C}(\tilde{x}) \qquad \text{subject to} \quad A\tilde{x} = y$$

Proof: For all x' such that $Ax' = y$, so that $A(x' - x) = 0$,

$$\begin{aligned} \mathcal{C}(x') &\geq \mathcal{C}(x) + \langle g_x, x' - x \rangle \\ &= \mathcal{C}(x) + \langle A^* u, x' - x \rangle \\ &= \mathcal{C}(x) + \langle u, A(x' - x) \rangle = \mathcal{C}(x) \end{aligned}$$

Dual certificate for the ℓ_1 norm

Lemma

x is a solution to

$$\min_{\tilde{x}} \quad ||\tilde{x}||_1 \qquad \text{subject to} \quad A\tilde{x} = y$$

if $Ax = y$ and there exists $g_x = A^* u$ such that

$$\begin{aligned} g_x(j) &= \text{sign}\{x(j)\} && \text{if } j \in \text{support}(x) \\ |g_x(j)| &\leq 1 && \text{if } j \notin \text{support}(x) \end{aligned}$$

Dual certificate for the ℓ_1 norm

Lemma

x is the unique solution to

$$\min_{\tilde{x}} \quad ||\tilde{x}||_1 \qquad \text{subject to} \quad A\tilde{x} = y$$

if $Ax = y$ and there exists $g_x = A^* u$ such that

$$\begin{aligned} g_x(j) &= \text{sign}\{x(j)\} && \text{if } j \in \text{support}(x) \\ |g_x(j)| &< 1 && \text{if } j \notin \text{support}(x) \end{aligned}$$

Dual certificate for the ℓ_1 norm

Lemma

x is solution to

$$\min_{\tilde{x}} \quad ||\tilde{x}||_1 \qquad \text{subject to} \quad A\tilde{x} = y$$

if $Ax = y$ and there exists $g_x = A^* u$ such that

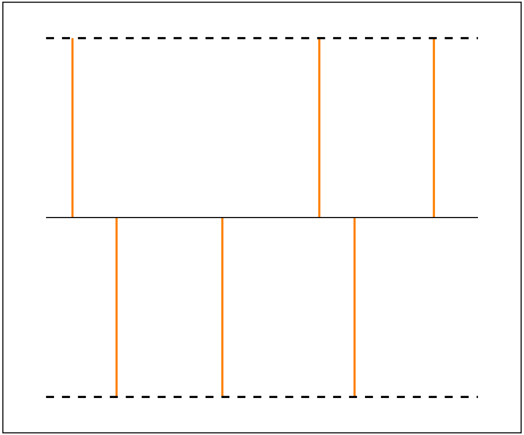
$$\begin{aligned} g_x(j) &= \text{sign}\{x(j)\} && \text{if } j \in \text{support}(x) \\ |g_x(j)| &< 1 && \text{if } j \notin \text{support}(x) \end{aligned}$$

The range of A^* corresponds to

Compressed sensing: Random sinusoids

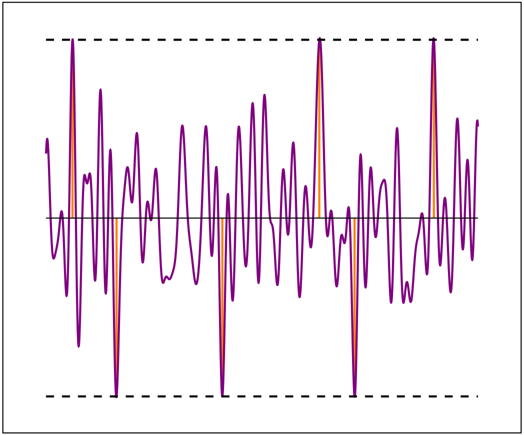
Super-resolution: Low-pass sinusoids

Dual certificate for compressed sensing



Least-squares interpolator

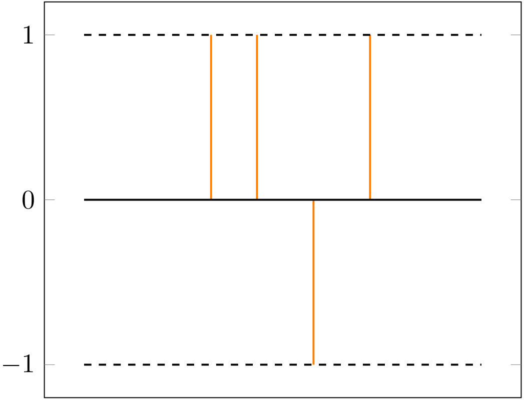
Dual certificate for compressed sensing



Works out for linear levels of sparsity (up to logarithmic factors)

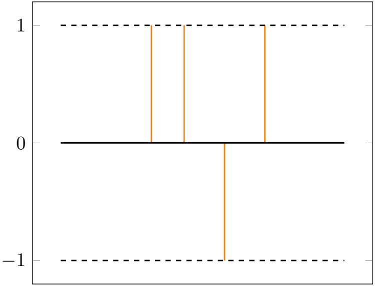
[Candès, Romberg, Tao 2006]

Dual certificate for super-resolution



Least-squares interpolator does not work

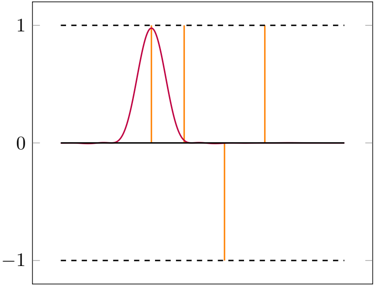
Dual certificate for super-resolution



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

$$g_x(t) = \sum_{t_j \in \text{support}(x)} \alpha_j \, K(t - t_j),$$

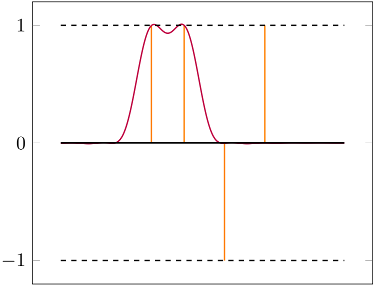
Dual certificate for super-resolution



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

$$g_x(t) = \sum_{t_j \in \text{support}(x)} \alpha_j \, K(t - t_j),$$

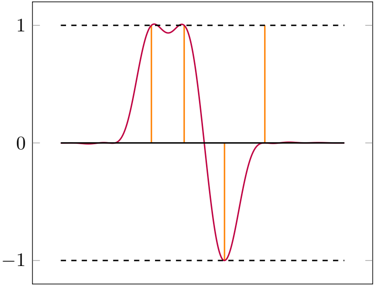
Dual certificate for super-resolution



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

$$g_x(t) = \sum_{t_j \in \text{support}(x)} \alpha_j \, K(t - t_j),$$

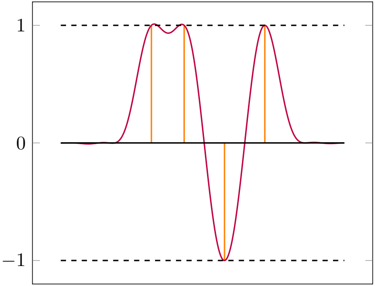
Dual certificate for super-resolution



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

$$g_x(t) = \sum_{t_j \in \text{support}(x)} \alpha_j K(t - t_j),$$

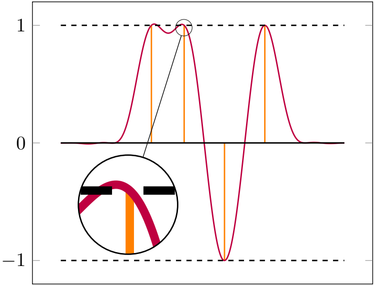
Dual certificate for super-resolution



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

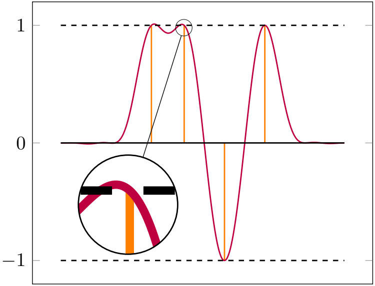
$$g_x(t) = \sum_{t_j \in \text{support}(x)} \alpha_j \, K(t - t_j),$$

Dual certificate for super-resolution



Problem: Magnitude of polynomial locally exceeds 1

Dual certificate for super-resolution

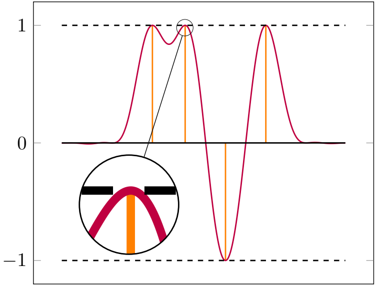


Problem: Magnitude of polynomial locally exceeds 1

Solution: Add correction term and force $g_x'(t_k) = 0$ for all $t_k \in \text{support}(x)$

$$g_x(t) = \sum_{t_j \in \text{support}(x)} \alpha_j K(t - t_j) + \beta_j K'(t - t_j)$$

Dual certificate for super-resolution



Problem: Magnitude of polynomial locally exceeds 1

Solution: Add correction term and force $g_x'(t_k) = 0$ for all $t_k \in \text{support}(x)$

$$g_x(t) = \sum_{t_j \in \text{support}(x)} \alpha_j \, K(t - t_j) + \beta_j \, K'(t - t_j)$$

Guarantees for super-resolution

Theorem [Candès, F. 2012]

If the minimum separation of the signal support obeys

$$\Delta \geq 2\,/f_c$$

then recovery via ℓ_1 norm minimization is exact

Guarantees for super-resolution

Theorem [Candès, F. 2014]

If the minimum separation of the signal support obeys

$$\Delta \geq 1.28\,/f_c,$$

then recovery via ℓ_1 norm minimization is exact

Guarantees for super-resolution

Theorem [Candès, F. 2014]

If the minimum separation of the signal support obeys

$$\Delta \geq 1.28 \,/ f_c,$$

then recovery via ℓ_1 norm minimization is exact

Theorem [Candès, F. 2012]

In 2D ℓ_1-norm minimization super-resolves point sources with a minimum separation of

$$\Delta \geq 2.38 \,/ f_c$$

where f_c is the cut-off frequency of the low-pass kernel

Guarantees for super-resolution

- Results hold for continuous version of the ℓ_1 norm (no discretization)
- Numerical simulations show that the method works for $\Delta \geq 1/f_c$
- Generalizations of dual certificate allow to prove robustness to noise [Candès, F. 2013], [F. 2013]
- If the signal is sparse, we can randomly undersample low-pass measurements [Tang, Bhaskar, Shah, Recht 2013]

Conclusion

Characterizing the interaction between the measurement operator and the structure of the object of interest is crucial to understand

- When the problem is well posed (conditioning of restricted operator)
- When optimization-based methods succeed (dual certificates)

For more details

- **Towards a mathematical theory of super-resolution**. E. J. Candès and C. Fernandez-Granda. *Communications on Pure and Applied Math* **67**(6), 906-956.
- **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.