Demixing Sines and Spikes:

Spectral Super-resolution in the Presence of Outliers

Carlos Fernandez-Granda

www.cims.nyu.edu/~cfgranda

Analysis Seminar, Courant

12/1/2016

Acknowledgements

Joint work with Gongguo Tang, Xiaodong Wang and Le Zheng

Project funded by NSF award DMS-1616340

Compressed sensing

Spectral super-resolution

Spectral super-resolution in the presence of outliers

Demixing via semidefinite programming

Greedy demixing + local optimization

Compressed sensing

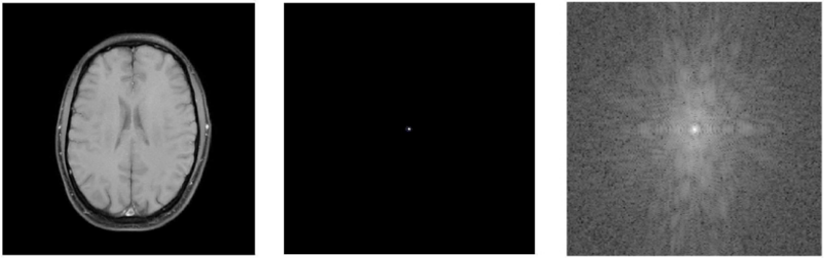
Spectral super-resolution

Spectral super-resolution in the presence of outliers

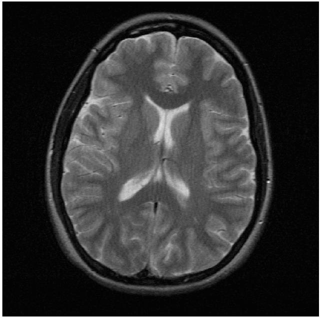
Demixing via semidefinite programming

Greedy demixing + local optimization

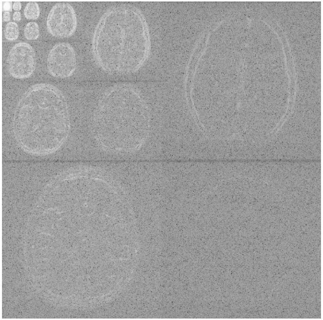
Magnetic resonance imaging



Images are sparse/compressible



Wavelet coefficients



Magnetic resonance imaging

Data: Samples from spectrum

Problem: Sampling is time consuming (annoying, patient might move)

Images are compressible ($\approx$ sparse)

Can we recover compressible signals from less data?

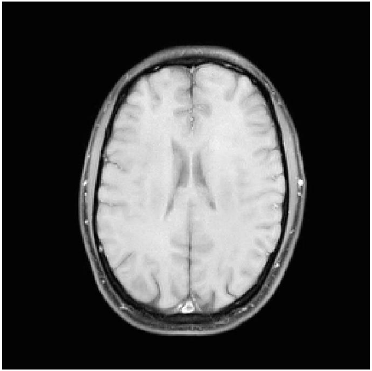
Compressed sensing

1. Undersample data randomly

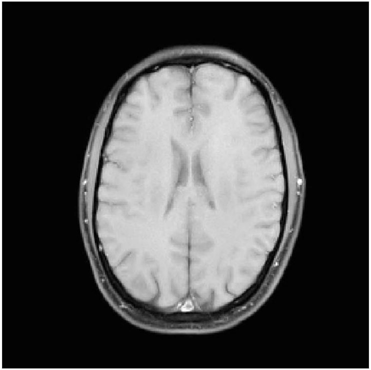
2. Solve the optimization problem

$$\begin{array}{ll} \textit{minimize} & ||\text{wavelet transform of estimate}||_1 \\ \textit{subject to} & \text{frequency samples of estimate} = \text{data} \end{array}$$

Compressed sensing in MRI



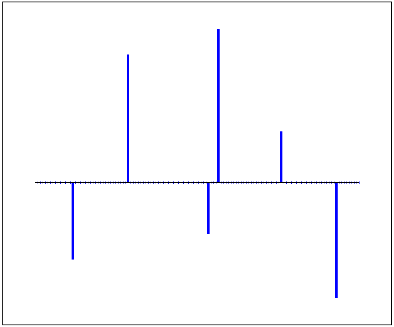
x2 Undersampling



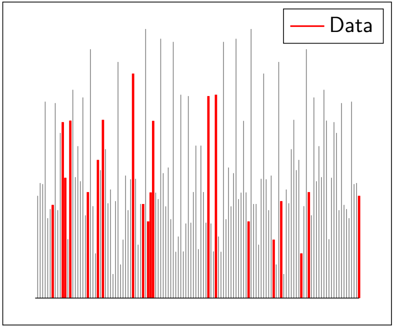
Compressed sensing (basic model)

1. Undersample the spectrum randomly

Signal



Spectrum



Compressed sensing (basic model)

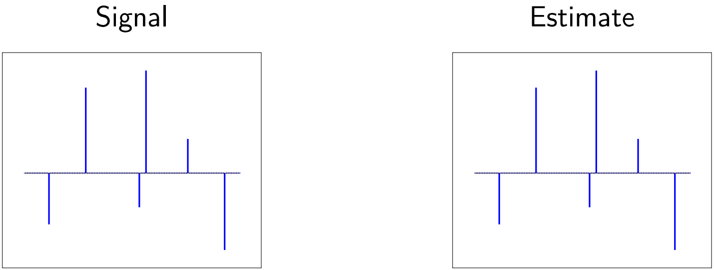
2. Solve the optimization problem

$$\begin{aligned} &\textit{minimize} \quad && ||\text{estimate}||_1 \\ &\textit{subject to} \quad && \text{frequency samples of estimate} = \text{data} \end{aligned}$$

Compressed sensing (basic model)

2. Solve the optimization problem

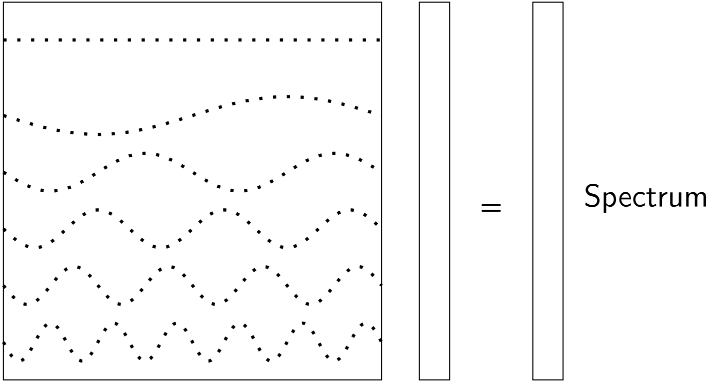
$$\begin{aligned} &\textit{minimize} && \|\text{estimate}\|_1 \\ &\textit{subject to} && \text{frequency samples of estimate} = \text{data} \end{aligned}$$



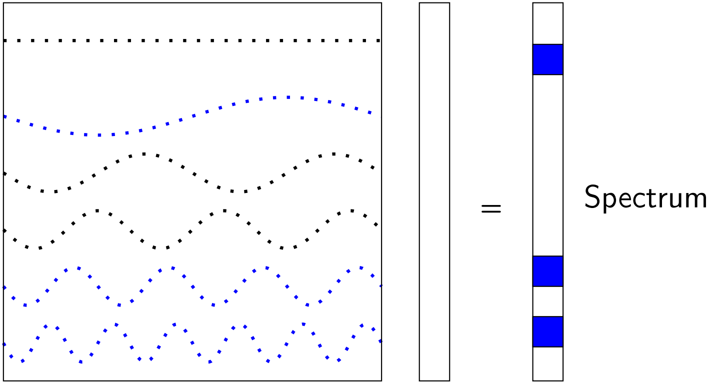
Theoretical questions

1. Is the problem well posed?
2. When can we guarantee that ℓ_1-norm minimization works?

Is the problem well posed?

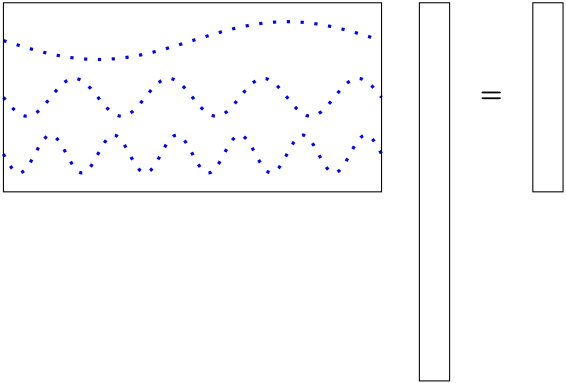


Is the problem well posed?



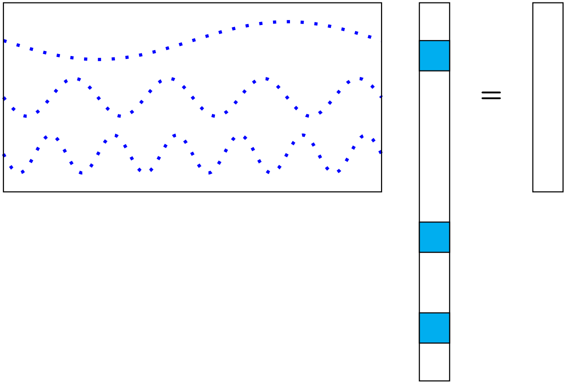
Measurements = random DFT coefficients

Is the problem well posed?



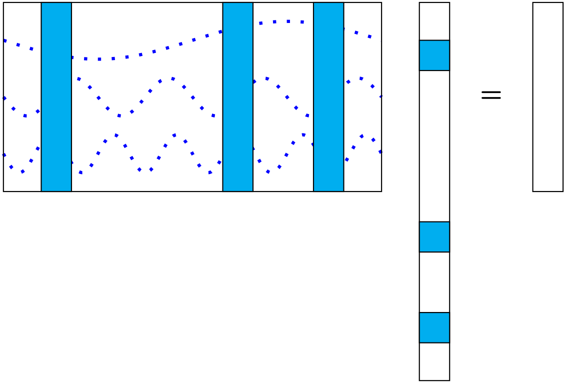
Measurements = random DFT coefficients

Is the problem well posed?



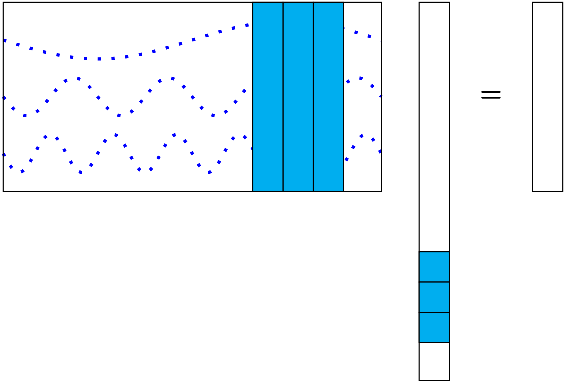
What is the effect of the measurement operator on sparse vectors?

Is the problem well posed?



Are sparse submatrices always well conditioned?

Is the problem well posed?



Are sparse submatrices always well conditioned?

Restricted isometry property (RIP)

An $m \times n$ matrix A satisfies the restricted isometry property if there is $0 < \delta < 1$ such that for any s-sparse vector $\mathbf{x}$

$$(1-\delta)\,||\boldsymbol{x}||_2 \leq ||A\boldsymbol{x}||_2 \leq (1+\delta)\,||\boldsymbol{x}||_2$$

Random Fourier matrices satisfy the RIP with high probability if s is $\mathcal{O}\,(\textit{measurements})$ up to log factors (Candès, Tao 2006)

$2s$-RIP implies that for any s-sparse signals $\boldsymbol{x_1}, \boldsymbol{x_2}$

$$||A\boldsymbol{x_2} - A\boldsymbol{x_1}||_2$$

Restricted isometry property (RIP)

An $m \times n$ matrix A satisfies the restricted isometry property if there is $0 < \delta < 1$ such that for any s-sparse vector $\mathbf{x}$

$$(1-\delta)\,||\boldsymbol{x}||_2 \leq ||A\boldsymbol{x}||_2 \leq (1+\delta)\,||\boldsymbol{x}||_2$$

Random Fourier matrices satisfy the RIP with high probability if s is $\mathcal{O}\,(\textit{measurements})$ up to log factors (Candès, Tao 2006)

2s-RIP implies that for any s-sparse signals $\boldsymbol{x_1}, \boldsymbol{x_2}$

$$||A\boldsymbol{x_2} - A\boldsymbol{x_1}||_2 = ||A\,(\boldsymbol{x_2} - \boldsymbol{x_1})||_2$$

Restricted isometry property (RIP)

An $m \times n$ matrix A satisfies the restricted isometry property if there is $0 < \delta < 1$ such that for any s-sparse vector $\mathbf{x}$

$$(1-\delta)\,||\boldsymbol{x}||_2 \leq ||A\boldsymbol{x}||_2 \leq (1+\delta)\,||\boldsymbol{x}||_2$$

Random Fourier matrices satisfy the RIP with high probability if s is $\mathcal{O}\,(\textit{measurements})$ up to log factors (Candès, Tao 2006)

$2s$-RIP implies that for any s-sparse signals $\boldsymbol{x_1}, \boldsymbol{x_2}$

$$\begin{aligned} ||A\boldsymbol{x_2} - A\boldsymbol{x_1}||_2 &= ||A\,(\boldsymbol{x_2} - \boldsymbol{x_1})||_2 \\ &\geq (1-\delta)\,||\boldsymbol{x_2} - \boldsymbol{x_1}||_2 \end{aligned}$$

Theoretical questions

1. Is the problem well posed?

2. When can we guarantee that ℓ_1-norm minimization works?

Characterizing the minimum ℓ_1-norm estimate

- Aim: Show that the original signal $\mathbf{x}$ is the solution of

$$\begin{aligned} \textit{minimize} \quad & \left|\left|\boldsymbol{x}'\right|\right|_1 \\ \textit{subject to} \quad & A\,\boldsymbol{x}' = \boldsymbol{y} \end{aligned}$$

- This is guaranteed by the existence of a dual certificate

Dual certificate

$\boldsymbol{q} \in \mathbb{C}^m$ is a dual certificate associated to $\boldsymbol{x}$ if

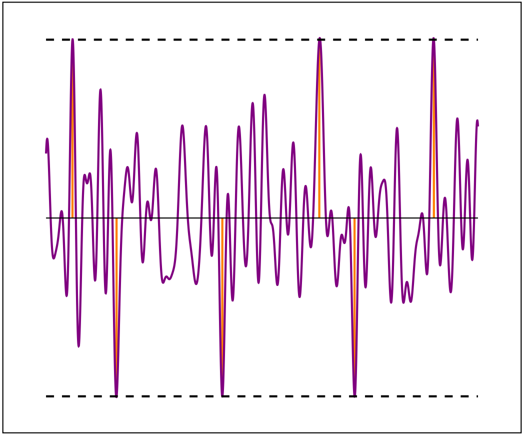
$$\boldsymbol{v} := A^* \boldsymbol{q}$$

satisfies

$$\boldsymbol{v}_i = \frac{\boldsymbol{x}_i}{|\boldsymbol{x}_i|} \qquad \text{if } \boldsymbol{x}_i \neq 0$$

$$|\boldsymbol{v}_i| < 1 \qquad \text{if } \boldsymbol{x}_i = 0$$

Example of $\boldsymbol{v}$



Linear combination of row vectors that interpolates the sign of the signal

Dual certificate

$\boldsymbol{v}$ is a subgradient of the ℓ_1 norm at $\boldsymbol{x}$

For any other feasible point $\boldsymbol{x} + \boldsymbol{h}$ such that $A\boldsymbol{h} = 0$

$$||\boldsymbol{x} + \boldsymbol{h}||_1$$

Dual certificate

$\boldsymbol{v}$ is a subgradient of the ℓ_1 norm at $\boldsymbol{x}$

For any other feasible point $\boldsymbol{x} + \boldsymbol{h}$ such that $A\boldsymbol{h} = 0$

$$||\boldsymbol{x} + \boldsymbol{h}||_1 \geq ||\boldsymbol{x}||_1 + \langle \boldsymbol{v}, \boldsymbol{h} \rangle$$

Dual certificate

$\boldsymbol{v}$ is a subgradient of the ℓ_1 norm at $\boldsymbol{x}$

For any other feasible point $\boldsymbol{x} + \boldsymbol{h}$ such that $A\boldsymbol{h} = 0$

$$\begin{aligned} ||\boldsymbol{x} + \boldsymbol{h}||_1 &\geq ||\boldsymbol{x}||_1 + \langle \boldsymbol{v}, \boldsymbol{h} \rangle \\ &= ||\boldsymbol{x}||_1 + \langle A^* \boldsymbol{q}, \boldsymbol{h} \rangle \end{aligned}$$

Dual certificate

$\boldsymbol{v}$ is a subgradient of the ℓ_1 norm at $\boldsymbol{x}$

For any other feasible point $\boldsymbol{x} + \boldsymbol{h}$ such that $A\boldsymbol{h} = 0$

$$\begin{aligned} ||\boldsymbol{x} + \boldsymbol{h}||_1 &\geq ||\boldsymbol{x}||_1 + \langle \boldsymbol{v}, \boldsymbol{h} \rangle \\ &= ||\boldsymbol{x}||_1 + \langle A^* \boldsymbol{q}, \boldsymbol{h} \rangle \\ &= ||\boldsymbol{x}||_1 + \langle \boldsymbol{q}, A\boldsymbol{h} \rangle \end{aligned}$$

Dual certificate

$\boldsymbol{v}$ is a subgradient of the ℓ_1 norm at $\boldsymbol{x}$

For any other feasible point $\boldsymbol{x}+\boldsymbol{h}$ such that $A\boldsymbol{h}=0$

$$\begin{aligned} ||\boldsymbol{x}+\boldsymbol{h}||_1 &\geq ||\boldsymbol{x}||_1 + \langle \boldsymbol{v}, \boldsymbol{h} \rangle \\ &= ||\boldsymbol{x}||_1 + \langle A^* \boldsymbol{q}, \boldsymbol{h} \rangle \\ &= ||\boldsymbol{x}||_1 + \langle \boldsymbol{q}, A\boldsymbol{h} \rangle \\ &= ||\boldsymbol{x}||_1 \end{aligned}$$

Dual certificate

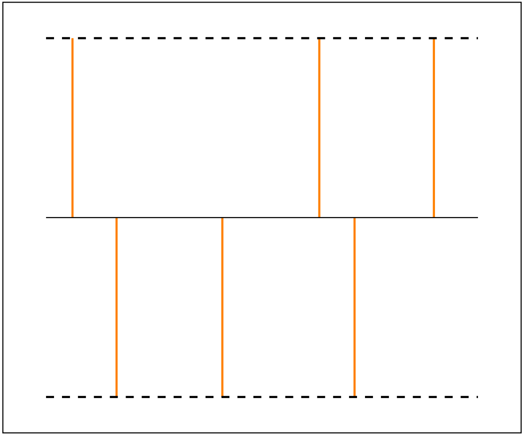
$\boldsymbol{v}$ is a subgradient of the ℓ_1 norm at $\boldsymbol{x}$

For any other feasible point $\boldsymbol{x} + \boldsymbol{h}$ such that $A\boldsymbol{h} = 0$

$$\begin{aligned} \|\boldsymbol{x} + \boldsymbol{h}\|_1 &\geq \|\boldsymbol{x}\|_1 + \langle \boldsymbol{v}, \boldsymbol{h} \rangle \\ &= \|\boldsymbol{x}\|_1 + \langle A^* \boldsymbol{q}, \boldsymbol{h} \rangle \\ &= \|\boldsymbol{x}\|_1 + \langle \boldsymbol{q}, A\boldsymbol{h} \rangle \\ &= \|\boldsymbol{x}\|_1 \end{aligned}$$

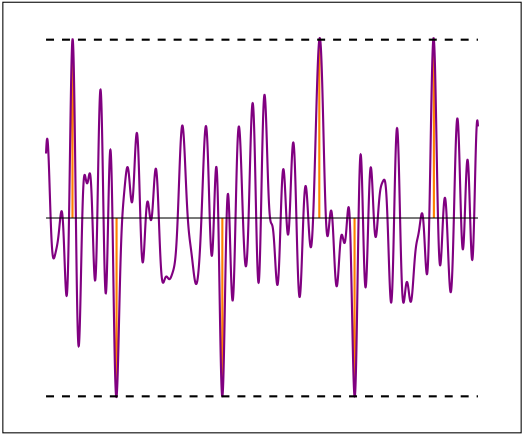
By a (slightly) more complicated argument $\boldsymbol{x}$ is the unique solution

Dual certificate for compressed sensing



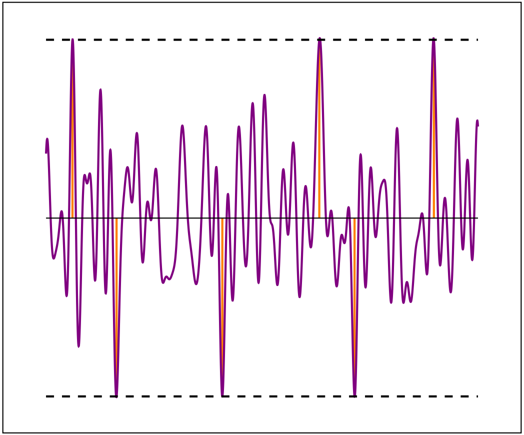
Aim: Show that a dual certificate exists for *any* sparse support and sign pattern

Certificate for compressed sensing



Idea: Minimum-energy interpolator has closed-form solution

Certificate for compressed sensing



Valid certificate if the sparsity is $\mathcal{O}$ (measurements) up to log factors

(Candès, Romberg, Tao 2006)

Compressed sensing

Spectral super-resolution

Spectral super-resolution in the presence of outliers

Demixing via semidefinite programming

Greedy demixing + local optimization

Spectral super-resolution

Goal: Estimate the spectrum of a multisinusoidal signal from a finite number of samples

Fundamental problem in signal processing

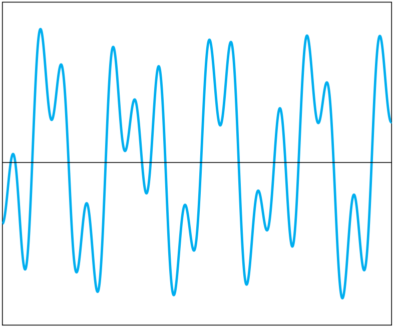
Classic techniques:

- Linear nonparametric methods: windowed periodogram
- Prony-based methods: MUSIC, matrix pencil, ESPRIT...

This talk: optimization-based spectral super-resolution

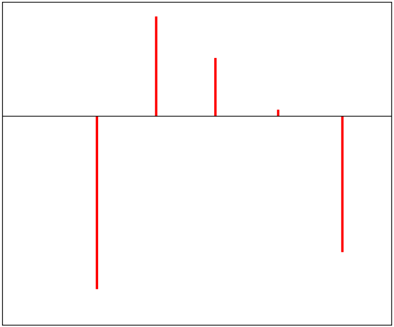
Spectral super-resolution

Signal



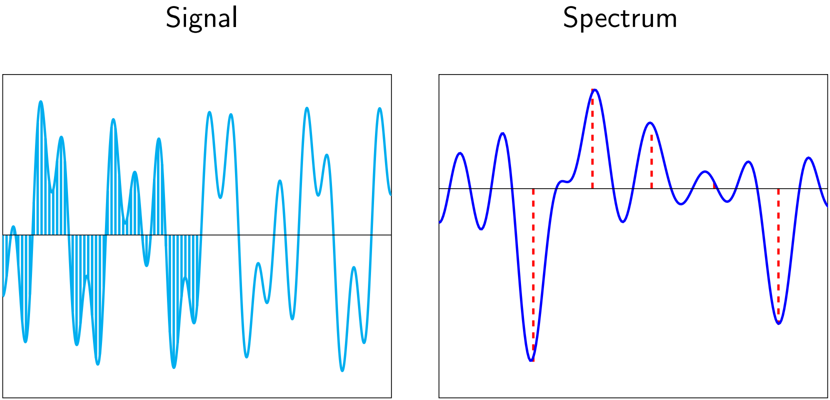
$g(t) := \sum_{j=1}^{k} \boldsymbol{x}_j \exp(i2\pi f_j t)$

Spectrum



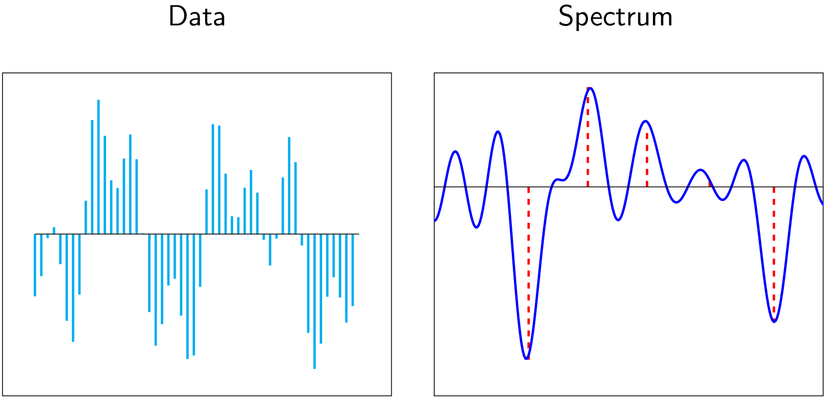
$\mu := \sum_{j=1}^{k} \boldsymbol{x}_j \, \delta(f - f_j)$

Spectral super-resolution



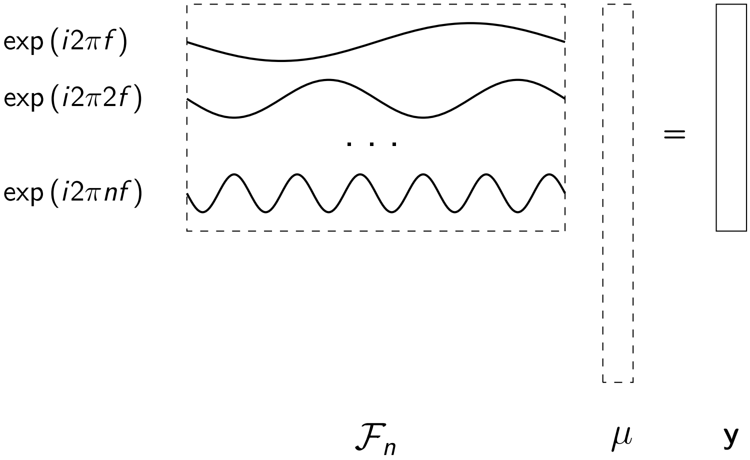
Data: $g(1), \quad g(2), \quad \cdots \quad g(n)$

Spectral super-resolution



Data: $g(l) = \int_0^1 \exp(i2\pi f l)\, \mathrm{d}\mu(f), \quad 1 \leq l \leq n$

Underdetermined linear system: $\boldsymbol{y} = \mathcal{F}_n \, \mu$

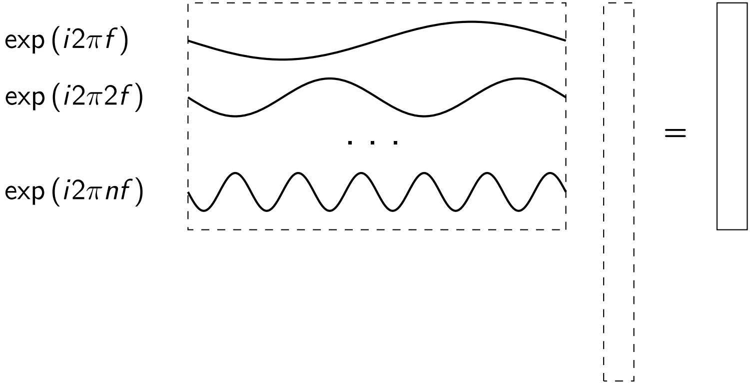


Theoretical questions

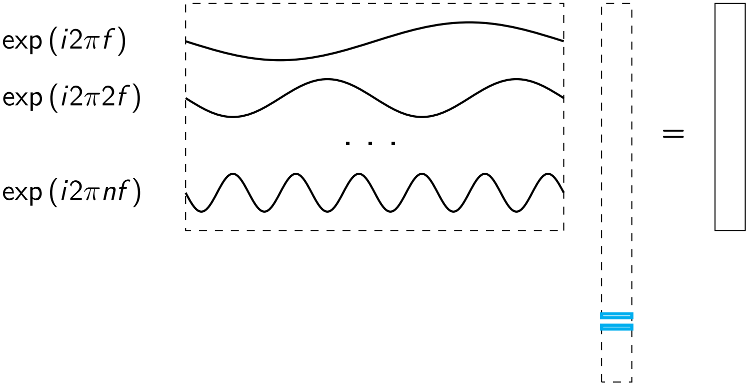
1. Is the problem well posed?

2. When can we guarantee that optimization-based approaches work?

Is the problem well posed?

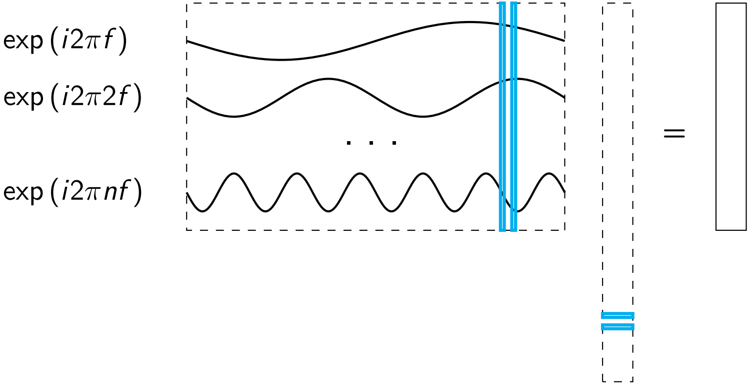


Is the problem well posed?



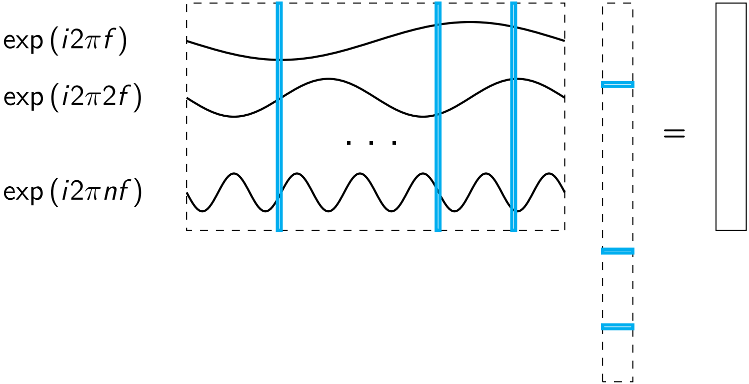
Effect of measurement operator on sparse vectors?

Is the problem well posed?



Submatrix can be very ill conditioned!

Is the problem well posed?

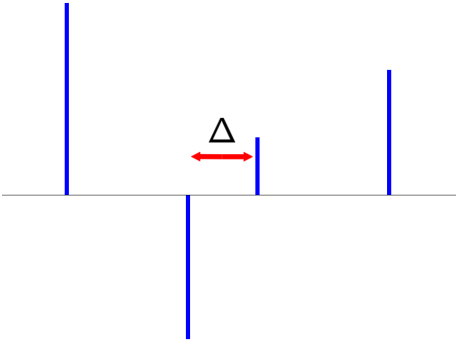


If the support is spread out there is hope

Minimum separation

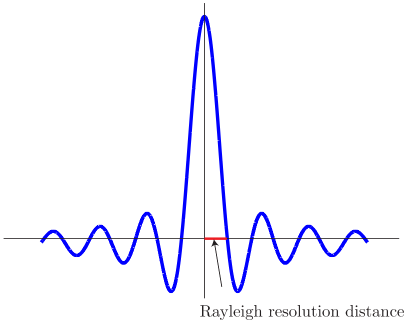
The minimum separation Δ of the support T of μ is

$$\Delta = \inf_{(f,f') \in \text{support}(\mu) : f \neq f'} |f - f'|$$



Conditioning of submatrix with respect to Δ

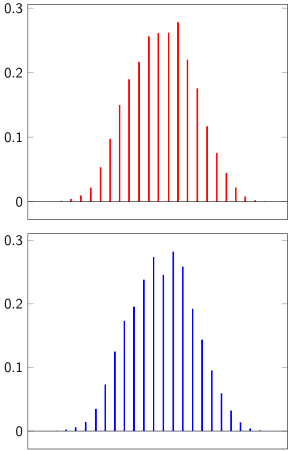
- If $\Delta < 2/(n-1)$ the problem is ill posed
- If $\Delta > 2/(n-1)$ the problem becomes well posed
- Proved asymptotically by Slepian and non-asymptotically by Moitra



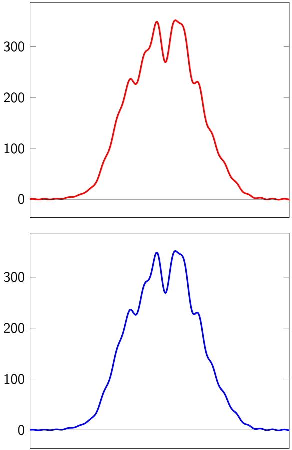
$2/(n-1)$ is the diameter of the main lobe of the impulse response of the measurement operator (twice the Rayleigh distance in optics)

Example: 25 spectral lines, $n = 2001$, $\Delta = 1.6/(n-1)$

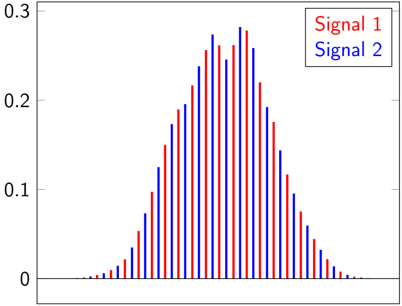
Spectrum of the signals



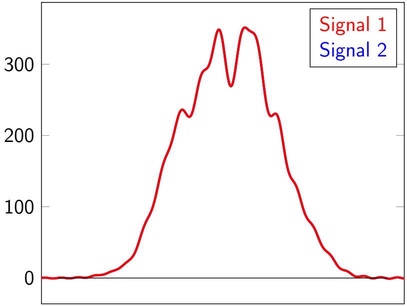
Spectrum of the data



Example: 25 spectral lines, $n = 2001$, $\Delta = 1.6/(n-1)$



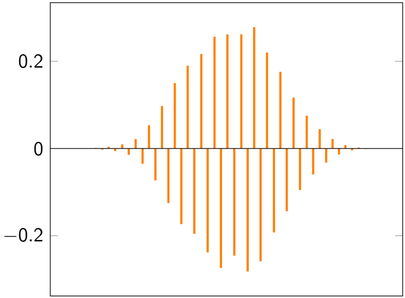
Spectrum of the signals



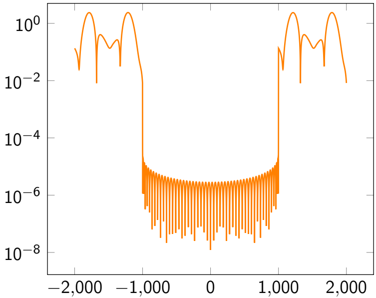
Spectrum of the data

Example: 25 spectral lines, $n = 2001$, $\Delta = 1.6/(n-1)$

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



Difference of signal spectra



Difference of signals

Theoretical questions

1. Is the problem well posed?

2. When can we guarantee that optimization-based approaches work?

Total-variation norm

- Continuous counterpart of the ℓ_1 norm
- If $\mu = \sum_j \boldsymbol{x}_j \delta_{f_j}$ then $||\mu||_{\mathsf{TV}} = ||\boldsymbol{x}||_1$
- Not the total variation of a piecewise-constant function
- Formal definition: For a complex measure ν

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

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

Estimation via convex programming

For data of the form $\boldsymbol{y} = \mathcal{F}_n \, \mu$, we solve

$$\min_{\tilde{\mu}} ||\tilde{\mu}||_{\mathsf{TV}} \quad \text{subject to} \quad \mathcal{F}_n \, \tilde{\mu} = \boldsymbol{y},$$

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

Dual certificate

A dual certificate $\boldsymbol{q} \in \mathbb{C}^n$ of the TV norm at

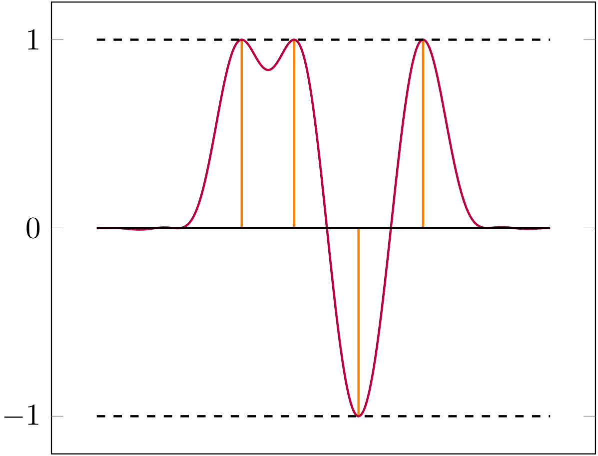
$$\mu := \sum_{j=1}^{k} \boldsymbol{x}_j \delta_{f_j} \qquad \boldsymbol{x} \in \mathbb{C}^k, \; f_j \in T$$

satisfies

$$\begin{aligned} Q(f) &:= \mathcal{F}_n^* \, \boldsymbol{q}(f) = \sum_{l=1}^{n} \boldsymbol{q}_l e^{-i2\pi l f} \\ Q(f_j) &= \frac{\boldsymbol{x}_j}{|\boldsymbol{x}_j|} && \text{if } f_j \in T \\ |Q(f)| &< 1 && \text{if } f \notin T \end{aligned}$$

We call Q a dual polynomial

Dual polynomial



Linear combination of low-pass sinusoids interpolating the sign

Dual certificate

Q is a subgradient of the TV norm at μ, in the sense that

$$||\mu + \nu||_{\text{TV}} \geq ||\mu||_{\text{TV}} + \langle Q, \nu \rangle \,, \quad \langle Q, \nu \rangle := \text{Re}\left[\int_{[0,1]} \overline{Q\,(f)}\; \text{d}\nu\,(f) \right]$$

For any $\mu + \nu$ such that $\mathcal{F}_n \nu = 0$

$$||\mu + \nu||_{\text{TV}}$$

Dual certificate

Q is a subgradient of the TV norm at μ, in the sense that

$$||\mu + \nu||_{\text{TV}} \geq ||\mu||_{\text{TV}} + \langle Q, \nu \rangle, \quad \langle Q, \nu \rangle := \text{Re}\left[\int_{[0,1]} \overline{Q(f)} \, \text{d}\nu(f) \right]$$

For any $\mu + \nu$ such that $\mathcal{F}_n \nu = 0$

$$||\mu + \nu||_{\text{TV}} \geq ||\mu||_{\text{TV}} + \langle Q, \nu \rangle$$

Dual certificate

Q is a subgradient of the TV norm at μ, in the sense that

$$||\mu + \nu||_{\mathsf{TV}} \geq ||\mu||_{\mathsf{TV}} + \langle Q, \nu \rangle, \quad \langle Q, \nu \rangle := \mathsf{Re}\left[\int_{[0,1]} \overline{Q(f)} \, \mathsf{d}\nu(f) \right]$$

For any $\mu + \nu$ such that $\mathcal{F}_n \nu = 0$

$$\begin{aligned} ||\mu + \nu||_{\mathsf{TV}} &\geq ||\mu||_{\mathsf{TV}} + \langle Q, \nu \rangle \\ &= ||\mu||_{\mathsf{TV}} + \langle \mathcal{F}_n^* \, \boldsymbol{q}, \nu \rangle \end{aligned}$$

Dual certificate

Q is a subgradient of the TV norm at μ, in the sense that

$$||\mu + \nu||_{\mathrm{TV}} \geq ||\mu||_{\mathrm{TV}} + \langle Q, \nu \rangle\,, \quad \langle Q, \nu \rangle := \mathrm{Re}\left[\int_{[0,1]} \overline{Q\,(f)}\, \mathrm{d}\nu\,(f)\right]$$

For any $\mu + \nu$ such that $\mathcal{F}_n \nu = 0$

$$\begin{aligned} ||\mu + \nu||_{\mathrm{TV}} &\geq ||\mu||_{\mathrm{TV}} + \langle Q, \nu \rangle \\ &= ||\mu||_{\mathrm{TV}} + \langle \mathcal{F}_n^* \boldsymbol{q}, \nu \rangle \\ &= ||\mu||_{\mathrm{TV}} + \langle \boldsymbol{q}, \mathcal{F}_n \nu \rangle \end{aligned}$$

Dual certificate

Q is a subgradient of the TV norm at μ, in the sense that

$$||\mu + \nu||_{\mathsf{TV}} \geq ||\mu||_{\mathsf{TV}} + \langle Q, \nu \rangle \,, \quad \langle Q, \nu \rangle := \operatorname{Re}\left[\int_{[0,1]} \overline{Q\,(f)}\; \mathrm{d}\nu\,(f) \right]$$

For any $\mu + \nu$ such that $\mathcal{F}_n \nu = 0$

$$\begin{aligned} ||\mu + \nu||_{\mathsf{TV}} &\geq ||\mu||_{\mathsf{TV}} + \langle Q, \nu \rangle \\ &= ||\mu||_{\mathsf{TV}} + \langle \mathcal{F}_n^* \, \boldsymbol{q}, \nu \rangle \\ &= ||\mu||_{\mathsf{TV}} + \langle \boldsymbol{q}, \mathcal{F}_n \nu \rangle \\ &= ||\mu||_{\mathsf{TV}} \end{aligned}$$

Dual certificate

Q is a subgradient of the TV norm at μ, in the sense that

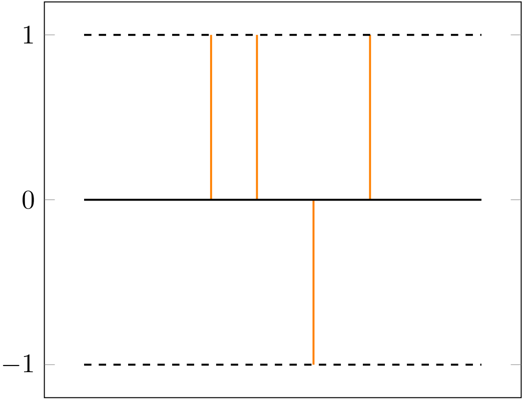
$$||\mu+\nu||_{\text{TV}} \geq ||\mu||_{\text{TV}} + \langle Q,\nu\rangle, \quad \langle Q,\nu\rangle := \text{Re}\left[\int_{[0,1]} \overline{Q(f)}\,\text{d}\nu(f)\right]$$

For any $\mu+\nu$ such that $\mathcal{F}_n\nu = 0$

$$\begin{aligned} ||\mu+\nu||_{\text{TV}} &\geq ||\mu||_{\text{TV}} + \langle Q,\nu\rangle \\ &= ||\mu||_{\text{TV}} + \langle \mathcal{F}_n^*\,\boldsymbol{q},\nu\rangle \\ &= ||\mu||_{\text{TV}} + \langle \boldsymbol{q},\mathcal{F}_n\nu\rangle \\ &= ||\mu||_{\text{TV}} \end{aligned}$$

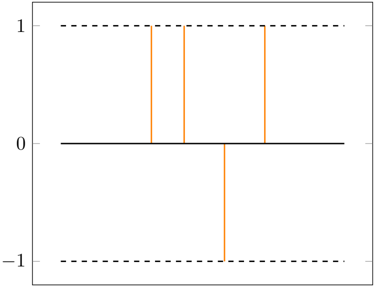
Existence of Q actually implies that μ is the unique solution

Certificate for super-resolution



Aim: Show that Q exists for any μ under a min. separation condition

Certificate for super-resolution



1st idea: Interpolation with a low-frequency fast-decaying kernel $\bar{K}$

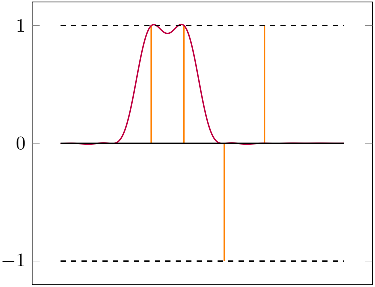
$$Q(f) = \sum_{j=1}^{k} \alpha_j \bar{K}(f - f_j)$$

Certificate for super-resolution

1st idea: Interpolation with a low-frequency fast-decaying kernel $\bar{K}$

$$Q(f) = \sum_{j=1}^{k} \boldsymbol{\alpha}_j \, \bar{K}\,(f - f_j)$$

Certificate for super-resolution



1st idea: Interpolation with a low-frequency fast-decaying kernel $\bar{K}$

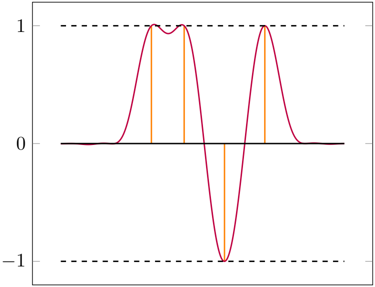
$$Q(f) = \sum_{j=1}^{k} \boldsymbol{\alpha}_j \, \bar{K}\,(f - f_j)$$

Certificate for super-resolution

1st idea: Interpolation with a low-frequency fast-decaying kernel $\bar{K}$

$$Q(f) = \sum_{j=1}^{k} \alpha_j \, \bar{K}\left(f - f_j\right)$$

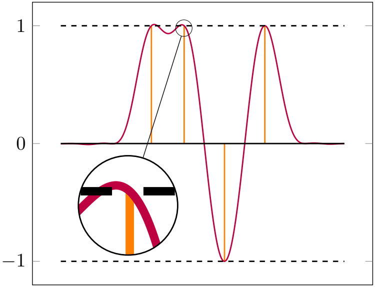
Certificate for super-resolution



1st idea: Interpolation with a low-frequency fast-decaying kernel $\bar{K}$

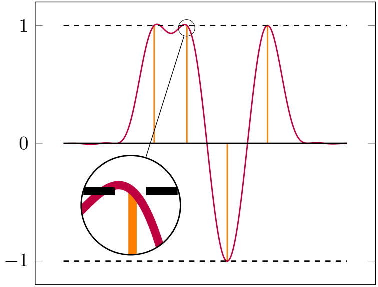
$$Q(f) = \sum_{j=1}^{k} \alpha_j \, \bar{K}\left(f - f_j\right)$$

Certificate for super-resolution



Problem: Magnitude of certificate locally exceeds 1

Certificate for super-resolution

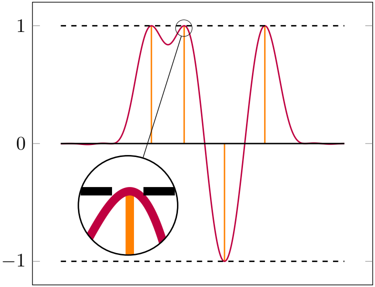


Problem: Magnitude of certificate locally exceeds 1

Solution: Add correction term and force the derivative of the certificate to equal zero on the support

$$Q(f) = \sum_{j=1}^{k} \alpha_j \, \bar{K}\,(f - f_j) + \beta_j \, \bar{K}'\,(f - f_j)$$

Certificate for super-resolution



Problem: Magnitude of certificate locally exceeds 1

Solution: Add correction term and force the derivative of the certificate to equal zero on the support

$$Q(f) = \sum_{j=1}^{k} \alpha_j \bar{K}(f - f_j) + \beta_j \bar{K}'(f - f_j)$$

Guarantees for spectral super-resolution

Theorem [Candès, F. 2012]

If the minimum separation of the spectrum support obeys

$$\Delta \geq \frac{4}{n-1}$$

then recovery via convex programming is exact

Theorem [Candès, F. 2012]

In 2D convex programming super-resolves spectral lines with a minimum separation of

$$\Delta \geq \frac{5.76}{n-1}$$

from samples of the form $g(1,1)$, $g(1,2)$, $\ldots$, $g(n,n)$

Guarantees for spectral super-resolution

Theorem [F. 2016]

If the minimum separation of the spectrum support obeys

$$\Delta \geq \frac{2.52}{n-1},$$

then recovery via convex programming is exact

Theorem [Candès, F. 2012]

In 2D convex programming super-resolves spectral lines with a minimum separation of

$$\Delta \geq \frac{5.76}{n-1}$$

from samples of the form $g(1,1)$, $g(1,2)$, $\ldots$, $g(n,n)$

Spectral super-resolution with missing data

Assume we observe a random subset of entries $\mathcal{S}$

New measurement operator $\mathcal{F}_{\mathcal{S}}$, for any measure ν

$$\mathcal{F}_{\mathcal{S}}\,\nu := (\mathcal{F}_n\,\nu)_{\mathcal{S}}$$

Signal: $\mu := \sum_{j=1}^{k} \boldsymbol{x}_j\,\delta\,(f - f_j)$

Data: $\boldsymbol{y}_{\mathcal{S}} := \mathcal{F}_{\mathcal{S}}\,\mu$

Can we still recover the signal?

Compressed sensing off the grid (Tang *et al* 2013)

Solving

$$\min_{\tilde{\mu}} ||\tilde{\mu}||_{\mathrm{TV}} \quad \text{subject to} \quad \mathcal{F}_{\mathcal{S}}\,\tilde{\mu} := \boldsymbol{y}_{\mathcal{S}}$$

achieves exact recovery with high prob. for $k = \mathcal{O}\left(|\mathcal{S}|\right)$ up to log factors if

$$\frac{\boldsymbol{x}_1}{|\boldsymbol{x}_1|}, \frac{\boldsymbol{x}_2}{|\boldsymbol{x}_2|}, \ldots, \frac{\boldsymbol{x}_k}{|\boldsymbol{x}_{\boldsymbol{k}}|}$$

are independent and uniformly distributed on the unit circle and

$$\Delta \geq \frac{4}{n-1}$$

Dual polynomial for compressed sensing off the grid

The only modification is the adjoint of the measurement operator

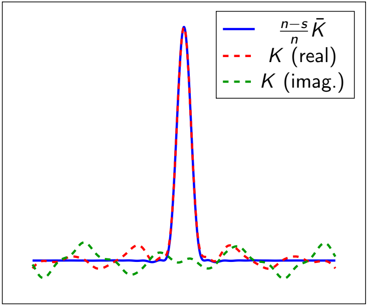
$$Q(f) := \mathcal{F}_{\mathcal{S}}^{*}\, \boldsymbol{q}(f) = \sum_{l \in \mathcal{S}} \boldsymbol{q}_l e^{-i2\pi l f}$$

$$Q(f_j) = \operatorname{sign}(\boldsymbol{x}_j) \qquad \text{if } f_j \in T$$

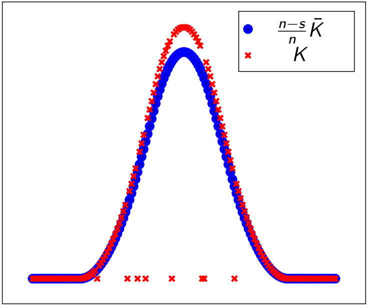
$$|Q(f)| < 1 \qquad \text{if } f \notin T$$

Idea: Interpolate with undersampled kernel

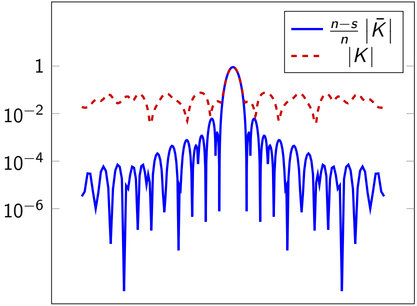
Random interpolation kernel



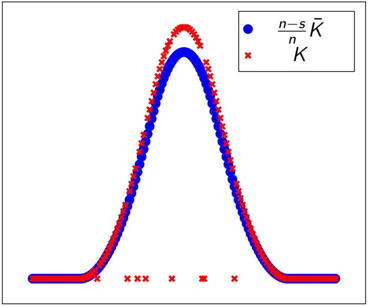
Spectrum (magnitude)



Random interpolation kernel

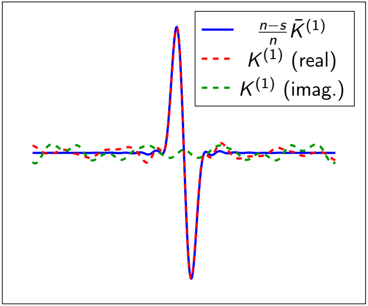


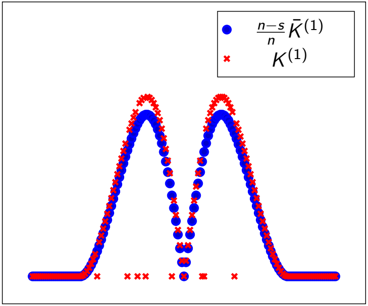
Spectrum (magnitude)



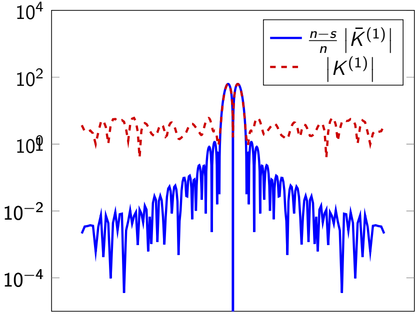
Random interpolation kernel (derivative)

Spectrum (magnitude)

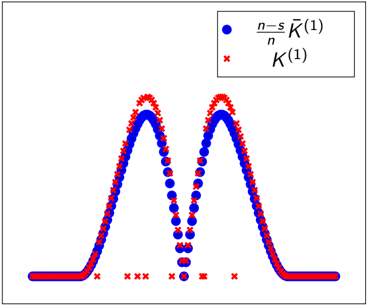




Random interpolation kernel (derivative)



Spectrum (magnitude)



Dual polynomial for compressed sensing off the grid

Construct dual polynomial via interpolation

$$Q(f) = \sum_{j=1}^{k} \boldsymbol{\alpha}_j \, K\left(f - f_j\right) \; + \; \boldsymbol{\beta}_j \, K'\left(f - f_j\right)$$

Valid dual polynomial with high probability as long as

$$\frac{\boldsymbol{x}_1}{|\boldsymbol{x}_1|}, \frac{\boldsymbol{x}_2}{|\boldsymbol{x}_2|}, \ldots, \frac{\boldsymbol{x}_k}{|\boldsymbol{x}_{\boldsymbol{k}}|}$$

are independent and uniformly distributed on the unit circle

Compressed sensing

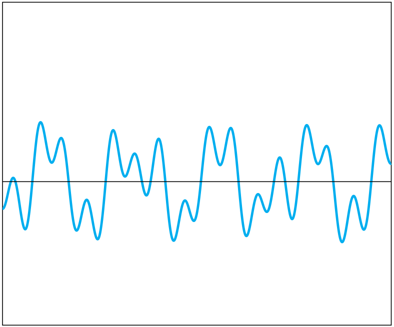
Spectral super-resolution

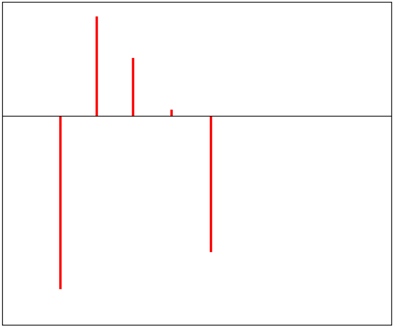
Spectral super-resolution in the presence of outliers

Demixing via semidefinite programming

Greedy demixing + local optimization

Spectral super-resolution in the presence of outliers

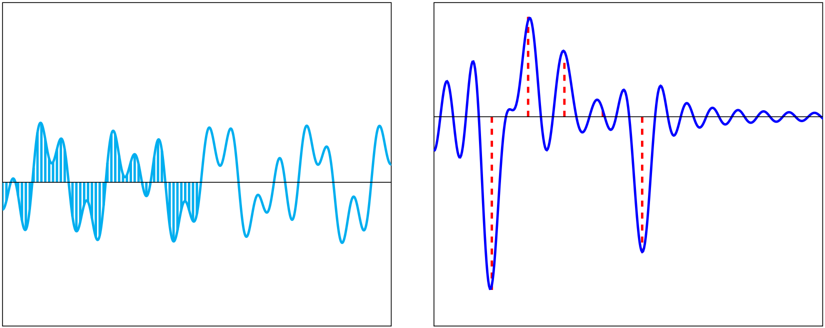




$g(t) := \sum_{j=1}^{k} \boldsymbol{x}_j \exp(i2\pi f_j t)$

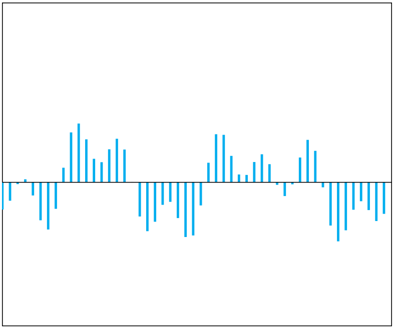
$\mu := \sum_{j=1}^{k} \boldsymbol{x}_j \, \delta(f - f_j)$

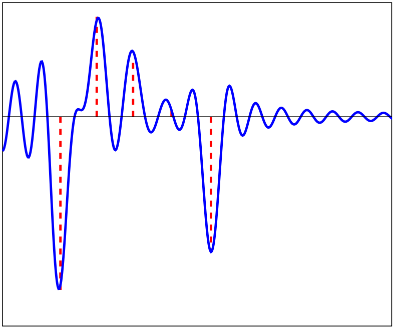
Spectral super-resolution in the presence of outliers



$$\mathcal{F}_n\,\mu = \begin{bmatrix} g\,(1) & g\,(2) & \cdots & g\,(n) \end{bmatrix}^T$$

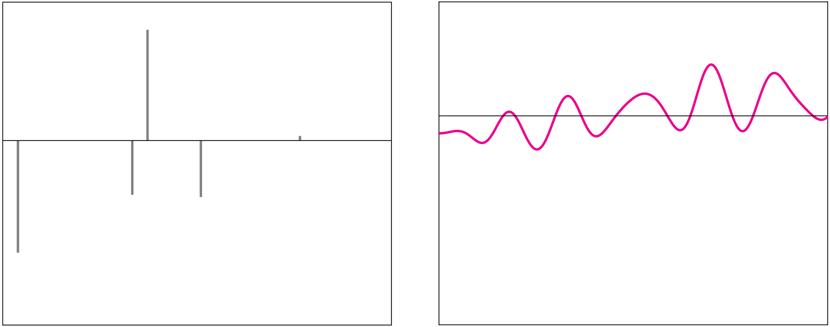
Spectral super-resolution in the presence of outliers





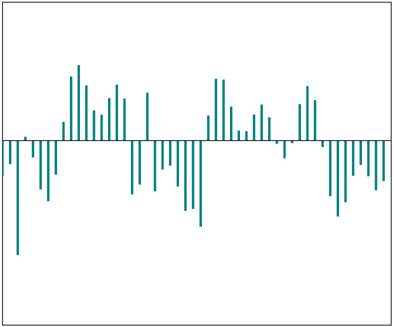
$$\mathcal{F}_n\,\mu = \begin{bmatrix} g(1) & g(2) & \cdots & g(n) \end{bmatrix}^T$$

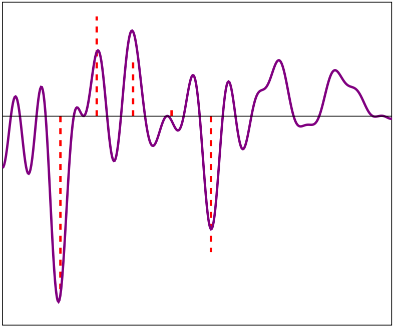
Spectral super-resolution in the presence of outliers



Some samples are completely corrupted by an s-sparse vector $\boldsymbol{z} \in \mathbb{C}^n$

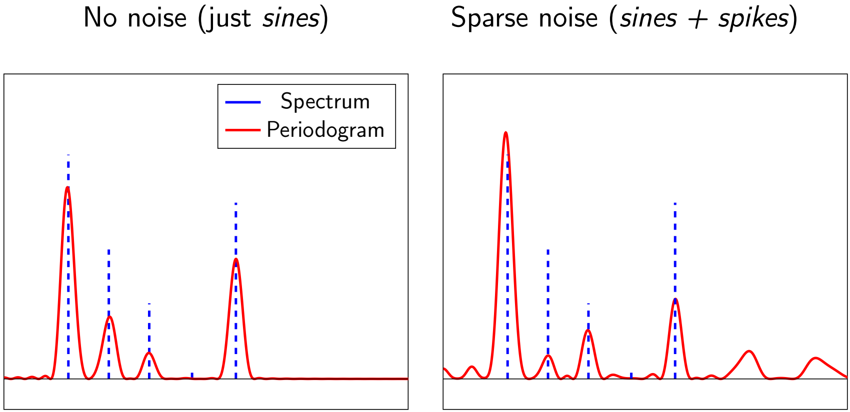
Spectral super-resolution in the presence of outliers





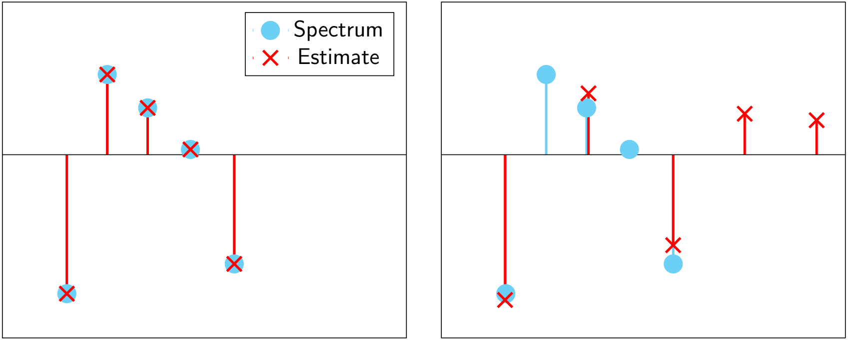
Data: $\quad \boldsymbol{y} := \mathcal{F}_n \, \mu + \boldsymbol{z}$

Linear nonparametric method: Gaussian periodogram



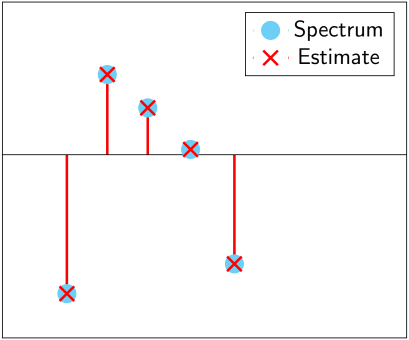
Prony-based method: MUSIC

No noise (just *sines*) | Sparse noise (*sines* + *spikes*)

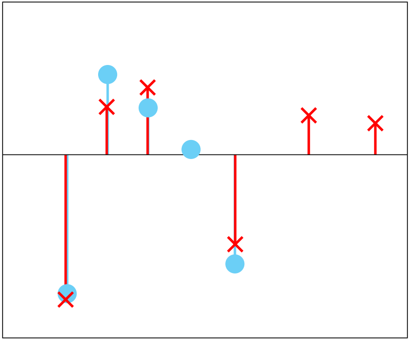


Optimization-based method (dense-noise model)

No noise (just *sines*)



Sparse noise (*sines* + *spikes*)



Optimization-based demixing

We incorporate a variable to model the sparse component

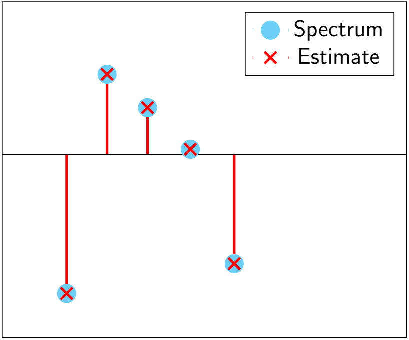
We promote sparsity of this component by penalizing its ℓ_1 norm

$$\min_{\tilde{\mu}, \tilde{\boldsymbol{z}}} ||\tilde{\mu}||_{\text{TV}} + \lambda \, ||\tilde{\boldsymbol{z}}||_1 \quad \text{subject to} \quad \mathcal{F}_n \, \tilde{\mu} + \tilde{\boldsymbol{z}} = \boldsymbol{y}$$

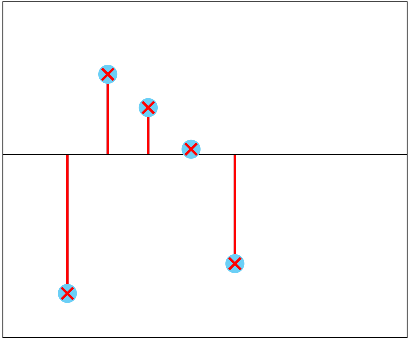
$\lambda > 0$ is a regularization parameter

Optimization-based method (dense + sparse noise model)

No noise (just *sines*)



Sparse noise (*sines + spikes*)



Guarantees for demixing

Theorem [F., Tang, Wang, Zheng 2016]

Solving the optimization for $\lambda = 1/\sqrt{n}$ recovers μ and $\boldsymbol{z}$ exactly with probability $1 - \epsilon$ as long as

$$k \leq C_k \left(\log \frac{n}{\epsilon} \right)^{-2} n,$$

$$s \leq C_s \left(\log \frac{n}{\epsilon} \right)^{-2} n,$$

for fixed numerical constants C_k, C_s

Number of sines and spikes are both $\mathcal{O}(n)$ up to logarithmic factors

Assumptions

- The minimum separation of the spectrum support obeys

$$\Delta \geq \frac{2.52}{n-1}$$

- Each entry of $\boldsymbol{z}$ is nonzero with probability s/n (independently)
- The phases of $\boldsymbol{x}$

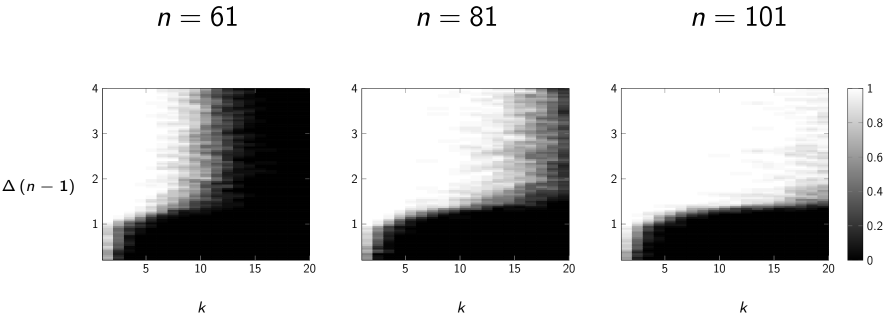
$$\frac{\boldsymbol{x}_1}{|\boldsymbol{x}_1|}, \frac{\boldsymbol{x}_2}{|\boldsymbol{x}_2|}, \dots, \frac{\boldsymbol{x}_k}{|\boldsymbol{x}_{\boldsymbol{k}}|}$$

and of the nonzero entries $\{i_1, \dots, i_s\}$ of $\boldsymbol{z}$

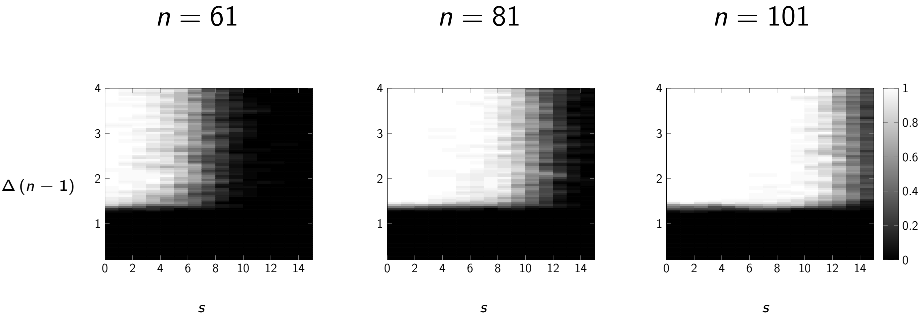
$$\frac{\boldsymbol{z}_{i_1}}{|\boldsymbol{z}_{i_1}|}, \frac{\boldsymbol{z}_{i_2}}{|\boldsymbol{z}_{i_2}|}, \dots, \frac{\boldsymbol{z}_{i_s}}{|\boldsymbol{z}_{i_s}|}$$

are independent and uniformly distributed on the unit circle

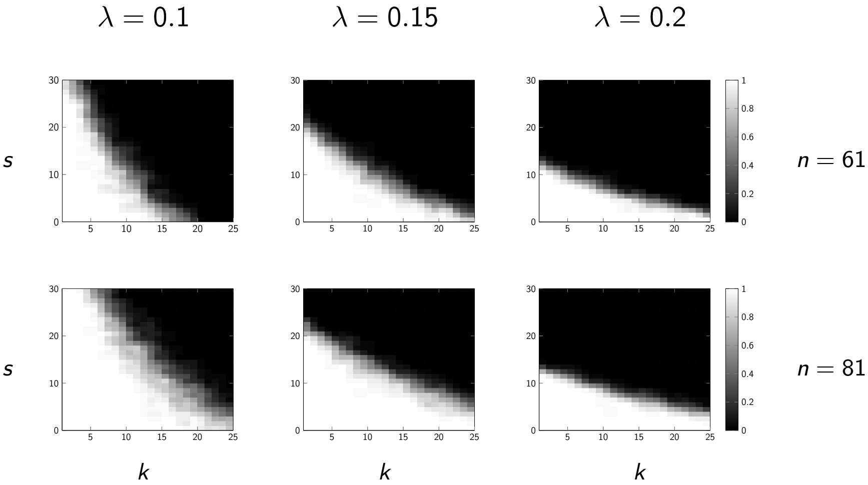
Experiments: $s := 10$



Experiments: $k := 15$



Experiments: $\Delta := 2/(n-1)$



Dual certificate for demixing

Dual certificate $\boldsymbol{q} \in \mathbb{C}^n$ and corresponding dual polynomial Q

$$Q(f) = \mathcal{F}_n^* \, \boldsymbol{q} = \sum_{j=1}^{n} \boldsymbol{q}_j \, e^{-i2\pi j f}$$

for a measure μ with support T and sparse noise $\boldsymbol{z}$ with support Ω

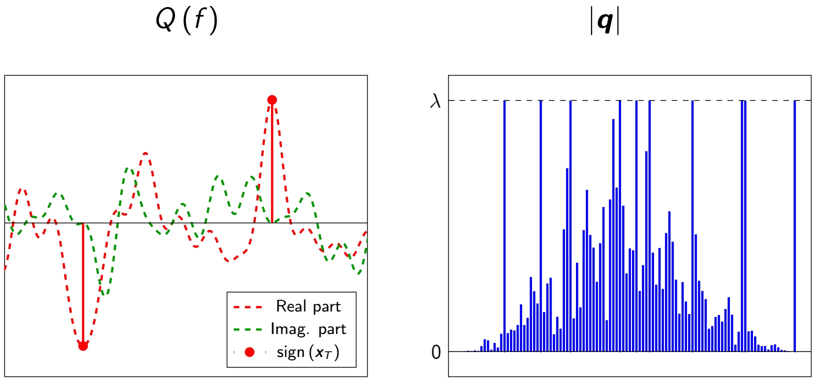
$$Q(f_j) = \frac{\boldsymbol{x}_j}{|\boldsymbol{x}_j|}, \qquad \forall f_j \in T$$

$$|Q(f)| < 1, \qquad \forall f \in T^c$$

$$\boldsymbol{q}_j = \lambda \frac{\boldsymbol{z}_l}{|\boldsymbol{z}_l|}, \qquad \forall j \in \Omega,$$

$$|\boldsymbol{q}_j| < \lambda, \qquad \forall l \in \Omega^c$$

Dual certificate for demixing



Dual certificate for demixing

Q is a "subgradient" of the TV norm at μ

$\frac{1}{\lambda}\boldsymbol{q}$ is a subgradient of the ℓ_1 norm at $\boldsymbol{z}$

For any other feasible pair $(\mu', \boldsymbol{z}')$ such that $\boldsymbol{y} = \mathcal{F}_n\,\mu' + \boldsymbol{z}' = \mathcal{F}_n\,\mu + \boldsymbol{z}$

$$\left|\left|\mu'\right|\right|_{\text{TV}} + \lambda\left|\left|\boldsymbol{z}'\right|\right|_1$$

Dual certificate for demixing

Q is a "subgradient" of the TV norm at μ

$\frac{1}{\lambda}\boldsymbol{q}$ is a subgradient of the ℓ_1 norm at $\boldsymbol{z}$

For any other feasible pair $(\mu', \boldsymbol{z}')$ such that $\boldsymbol{y} = \mathcal{F}_n\,\mu' + \boldsymbol{z}' = \mathcal{F}_n\,\mu + \boldsymbol{z}$

$$\left|\left|\mu'\right|\right|_{\text{TV}} + \lambda \left|\left|\boldsymbol{z}'\right|\right|_1 \geq \left|\left|\mu\right|\right|_{\text{TV}} + \left\langle Q, \mu' - \mu \right\rangle + \lambda \left|\left|\boldsymbol{z}\right|\right|_1 + \lambda \left\langle \frac{1}{\lambda}\boldsymbol{q}, \boldsymbol{z}' - \boldsymbol{z} \right\rangle$$

Dual certificate for demixing

Q is a "subgradient" of the TV norm at μ

$\frac{1}{\lambda}\boldsymbol{q}$ is a subgradient of the ℓ_1 norm at $\boldsymbol{z}$

For any other feasible pair $(\mu', \boldsymbol{z}')$ such that $\boldsymbol{y} = \mathcal{F}_n\,\mu' + \boldsymbol{z}' = \mathcal{F}_n\,\mu + \boldsymbol{z}$

$$\begin{aligned} \left|\left|\mu'\right|\right|_{\mathrm{TV}} + \lambda \left|\left|\boldsymbol{z}'\right|\right|_1 &\geq \left|\left|\mu\right|\right|_{\mathrm{TV}} + \left\langle Q, \mu' - \mu \right\rangle + \lambda \left|\left|\boldsymbol{z}\right|\right|_1 + \lambda \left\langle \frac{1}{\lambda}\boldsymbol{q}, \boldsymbol{z}' - \boldsymbol{z} \right\rangle \\ &\geq \left|\left|\mu\right|\right|_{\mathrm{TV}} + \left\langle \mathcal{F}_n^*\,\boldsymbol{q}, \mu' - \mu \right\rangle + \lambda \left|\left|\boldsymbol{z}\right|\right|_1 + \left\langle \boldsymbol{q}, \boldsymbol{z}' - \boldsymbol{z} \right\rangle \end{aligned}$$

Dual certificate for demixing

Q is a "subgradient" of the TV norm at μ

$\frac{1}{\lambda}\boldsymbol{q}$ is a subgradient of the ℓ_1 norm at $\boldsymbol{z}$

For any other feasible pair $(\mu', \boldsymbol{z}')$ such that $\boldsymbol{y} = \mathcal{F}_n \, \mu' + \boldsymbol{z}' = \mathcal{F}_n \, \mu + \boldsymbol{z}$

$$\begin{aligned} \left|\left|\mu'\right|\right|_{\text{TV}} + \lambda \left|\left|\boldsymbol{z}'\right|\right|_1 &\geq \left|\left|\mu\right|\right|_{\text{TV}} + \left\langle Q, \mu' - \mu \right\rangle + \lambda \left|\left|\boldsymbol{z}\right|\right|_1 + \lambda \left\langle \frac{1}{\lambda}\boldsymbol{q}, \boldsymbol{z}' - \boldsymbol{z} \right\rangle \\ &\geq \left|\left|\mu\right|\right|_{\text{TV}} + \left\langle \mathcal{F}_n^* \, \boldsymbol{q}, \mu' - \mu \right\rangle + \lambda \left|\left|\boldsymbol{z}\right|\right|_1 + \left\langle \boldsymbol{q}, \boldsymbol{z}' - \boldsymbol{z} \right\rangle \\ &= \left|\left|\mu\right|\right|_{\text{TV}} + \lambda \left|\left|\boldsymbol{z}\right|\right|_1 + \left\langle \boldsymbol{q}, \mathcal{F}_n \, \mu' + \boldsymbol{z}' - \mathcal{F}_n \, \mu - \boldsymbol{z} \right\rangle \end{aligned}$$

Dual certificate for demixing

Q is a "subgradient" of the TV norm at μ

$\frac{1}{\lambda}\boldsymbol{q}$ is a subgradient of the ℓ_1 norm at $\boldsymbol{z}$

For any other feasible pair $(\mu', \boldsymbol{z}')$ such that $\boldsymbol{y} = \mathcal{F}_n\, \mu' + \boldsymbol{z}' = \mathcal{F}_n\, \mu + \boldsymbol{z}$

$$
\begin{aligned}
\left|\left|\mu'\right|\right|_{\mathrm{TV}} + \lambda \left|\left|\boldsymbol{z}'\right|\right|_1 &\geq \left|\left|\mu\right|\right|_{\mathrm{TV}} + \left\langle Q, \mu' - \mu \right\rangle + \lambda \left|\left|\boldsymbol{z}\right|\right|_1 + \lambda \left\langle \frac{1}{\lambda}\boldsymbol{q}, \boldsymbol{z}' - \boldsymbol{z} \right\rangle \\
&\geq \left|\left|\mu\right|\right|_{\mathrm{TV}} + \left\langle \mathcal{F}_n^*\, \boldsymbol{q}, \mu' - \mu \right\rangle + \lambda \left|\left|\boldsymbol{z}\right|\right|_1 + \left\langle \boldsymbol{q}, \boldsymbol{z}' - \boldsymbol{z} \right\rangle \\
&= \left|\left|\mu\right|\right|_{\mathrm{TV}} + \lambda \left|\left|\boldsymbol{z}\right|\right|_1 + \left\langle \boldsymbol{q}, \mathcal{F}_n\, \mu' + \boldsymbol{z}' - \mathcal{F}_n\, \mu - \boldsymbol{z} \right\rangle \\
&= \left|\left|\mu\right|\right|_{\mathrm{TV}} + \lambda \left|\left|\boldsymbol{z}\right|\right|_1
\end{aligned}
$$

Dual certificate for demixing

Q is a "subgradient" of the TV norm at μ

$\frac{1}{\lambda}\boldsymbol{q}$ is a subgradient of the ℓ_1 norm at $\boldsymbol{z}$

For any other feasible pair $(\mu', \boldsymbol{z}')$ such that $\boldsymbol{y} = \mathcal{F}_n\, \mu' + \boldsymbol{z}' = \mathcal{F}_n\, \mu + \boldsymbol{z}$

$$\begin{aligned}
\left|\left|\mu'\right|\right|_{\mathrm{TV}} + \lambda \left|\left|\boldsymbol{z}'\right|\right|_1 &\geq \left|\left|\mu\right|\right|_{\mathrm{TV}} + \left\langle Q, \mu' - \mu \right\rangle + \lambda \left|\left|\boldsymbol{z}\right|\right|_1 + \lambda \left\langle \frac{1}{\lambda}\boldsymbol{q}, \boldsymbol{z}' - \boldsymbol{z} \right\rangle \\
&\geq \left|\left|\mu\right|\right|_{\mathrm{TV}} + \left\langle \mathcal{F}_n^*\, \boldsymbol{q}, \mu' - \mu \right\rangle + \lambda \left|\left|\boldsymbol{z}\right|\right|_1 + \left\langle \boldsymbol{q}, \boldsymbol{z}' - \boldsymbol{z} \right\rangle \\
&= \left|\left|\mu\right|\right|_{\mathrm{TV}} + \lambda \left|\left|\boldsymbol{z}\right|\right|_1 + \left\langle \boldsymbol{q}, \mathcal{F}_n\, \mu' + \boldsymbol{z}' - \mathcal{F}_n\, \mu - \boldsymbol{z} \right\rangle \\
&= \left|\left|\mu\right|\right|_{\mathrm{TV}} + \lambda \left|\left|\boldsymbol{z}\right|\right|_1
\end{aligned}$$

Existence of Q actually implies that $(\mu, \boldsymbol{z})$ is the unique solution

Dual certificate for demixing

$$Q(f) := Q_{\text{aux}}(f) + R(f)$$

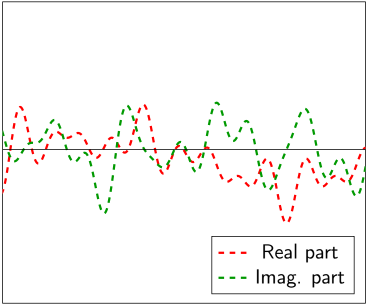
$$Q_{\text{aux}}(f) := \sum_{l \in \Omega^c} \boldsymbol{q}_l \, e^{-i2\pi l f}$$

$$R(f) := \lambda \sum_{l \in \Omega} \frac{\boldsymbol{z}_l}{|\boldsymbol{z}_l|} \, e^{-i2\pi l f}$$

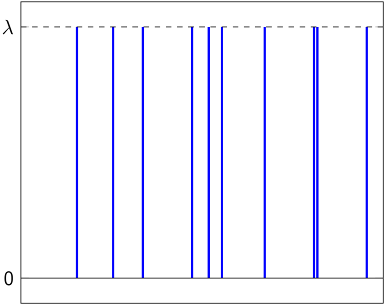
Satisfies condition

$$\boldsymbol{q}_j = \lambda \frac{\boldsymbol{z}_l}{|\boldsymbol{z}_l|} \qquad \forall j \in \Omega$$

$R(f)$



Spectrum (magnitude)



Dual certificate for demixing

We construct Q_{aux} via interpolation with a random kernel K (coeffs in Ω^c)

$$Q_{\text{aux}}(f) = \sum_{j=1}^{k} \boldsymbol{\alpha}_j \, K\left(f - f_j\right) \; + \; \boldsymbol{\beta}_j \, K'\left(f - f_j\right)$$

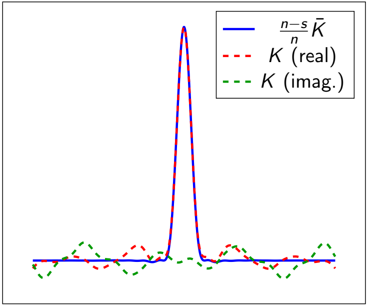
To ensure

$$Q\left(f_j\right) = \frac{\boldsymbol{x}_j}{|\boldsymbol{x}_j|}, \quad f_j \in T,$$
$$Q'\left(f_j\right) = 0, \qquad f_j \in T$$

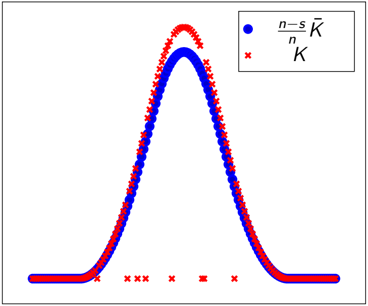
we enforce

$$Q_{\text{aux}}\left(f_j\right) = \frac{\boldsymbol{x}_j}{|\boldsymbol{x}_j|} - R\left(f_j\right), \quad f_j \in T$$
$$Q'_{\text{aux}}\left(f_j\right) = -R'\left(f_j\right), \qquad f_j \in T$$

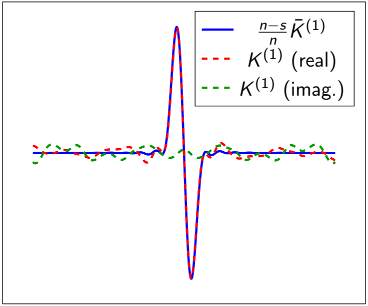
Random interpolation kernel



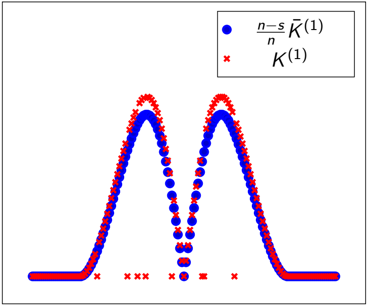
Spectrum (magnitude)



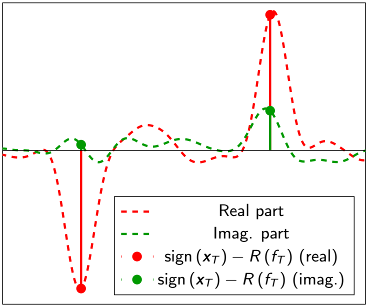
Random interpolation kernel (derivative)



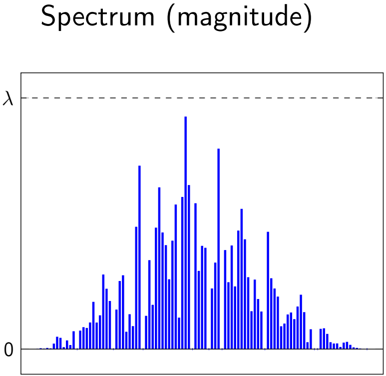
Spectrum (magnitude)



$Q_{\text{aux}}(f)$

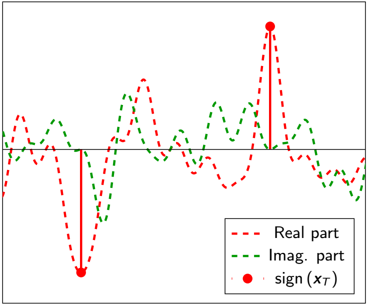


Real part
Imag. part
sign $(x_T) - R(f_T)$ (real)
sign $(x_T) - R(f_T)$ (imag.)

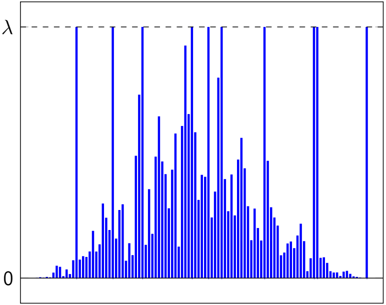


Spectrum (magnitude)
λ
0

$Q(f)$



Spectrum (magnitude)



Compressed sensing

Spectral super-resolution

Spectral super-resolution in the presence of outliers

Demixing via semidefinite programming

Greedy demixing + local optimization

Practical implementation

- **Primal problem**:

$$\min_{\tilde{\mu},\tilde{\boldsymbol{z}}} ||\tilde{\mu}||_{\mathrm{TV}} + \lambda\,||\tilde{\boldsymbol{z}}||_1 \quad \text{subject to} \quad \mathcal{F}_n\,\tilde{\mu} + \tilde{\boldsymbol{z}} = \boldsymbol{y}$$

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

First option: Discretizing + ℓ_1-norm minimization

Practical implementation

- **Primal problem**:

$$\min_{\tilde{\mu}, \tilde{\boldsymbol{z}}} ||\tilde{\mu}||_{\text{TV}} + \lambda ||\tilde{\boldsymbol{z}}||_1 \quad \text{subject to} \quad \mathcal{F}_n \tilde{\mu} + \tilde{\boldsymbol{z}} = \boldsymbol{y}$$

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

First option: Discretizing + ℓ_1-norm minimization

- **Dual problem**:

$$\max_{\boldsymbol{\eta} \in \mathbb{C}^n} \langle \boldsymbol{y}, \boldsymbol{\eta} \rangle \quad \text{subject to} \quad ||\mathcal{F}_n^* \boldsymbol{\eta}||_\infty \leq 1$$

$$||\boldsymbol{\eta}||_\infty \leq \lambda$$

Finite-dimensional variable $\boldsymbol{\eta}$, but infinite-dimensional constraint

$$\mathcal{F}_n^* \boldsymbol{\eta}(f) = \sum_{l=-n}^{n} \boldsymbol{\eta}_l e^{-i2\pi l f}$$

Second option: Solving the dual problem

Lemma: Semidefinite representation

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

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

is equivalent to

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

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

Consequence: The dual problem is a tractable semidefinite program

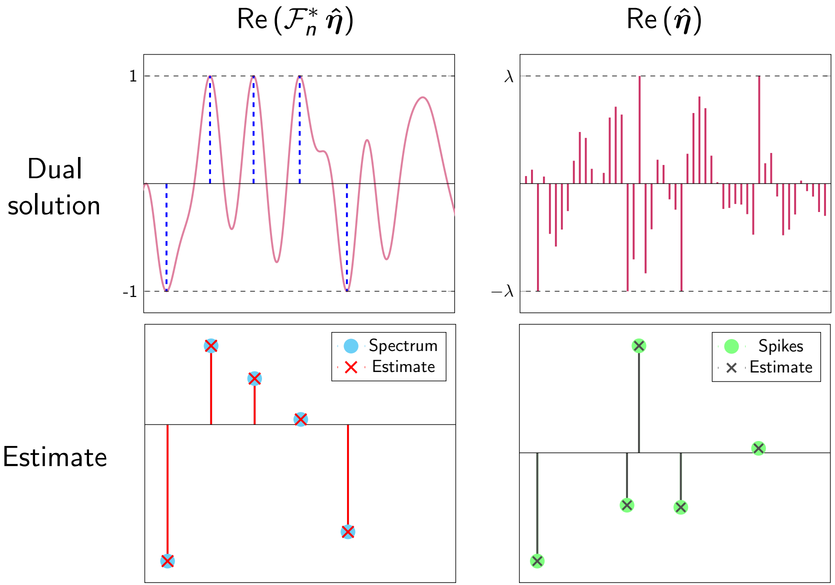
Support-locating polynomial

How do we obtain an estimator from the dual solution?

Dual solution vector: From strong duality

- $\hat{\boldsymbol{\eta}}$ interpolates the sign of the primal solution $\hat{\boldsymbol{z}}$
- $\mathcal{F}_n^* \hat{\boldsymbol{\eta}}$ interpolates the sign of the primal solution $\hat{\mu}$

Demixing via semidefinite programming



Compressed sensing

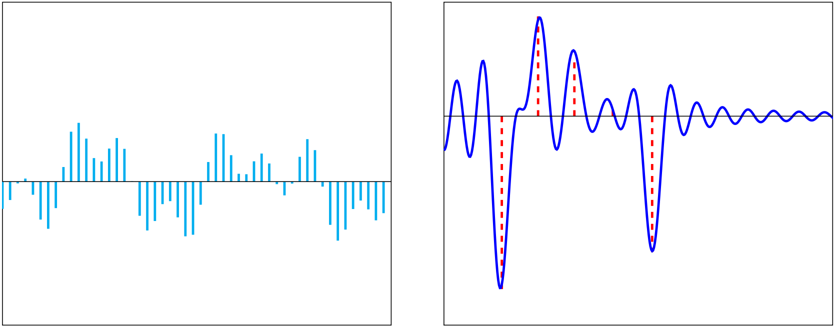
Spectral super-resolution

Spectral super-resolution in the presence of outliers

Demixing via semidefinite programming

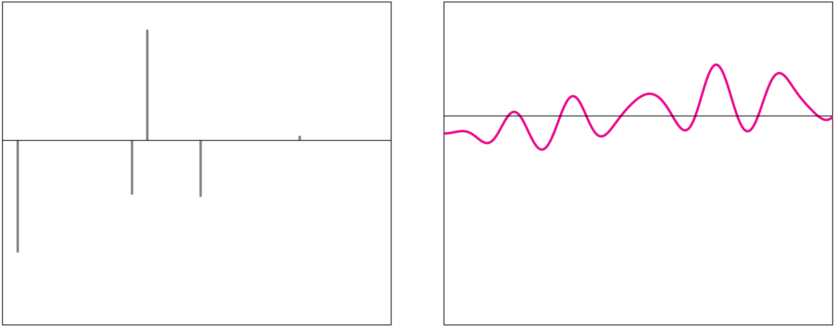
Greedy demixing + local optimization

Spectral super-resolution in the presence of outliers



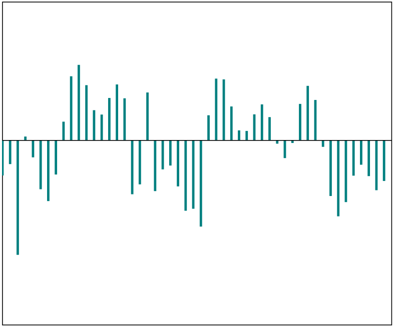
$$\begin{bmatrix} g(1) \\ g(2) \\ \cdots \\ g(n) \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{k} \boldsymbol{x}_j \exp(i2\pi f_j 1) \\ \sum_{j=1}^{k} \boldsymbol{x}_j \exp(i2\pi f_j 2) \\ \cdots \\ \sum_{j=1}^{k} \boldsymbol{x}_j \exp(i2\pi f_j n) \end{bmatrix} = \sum_{j=1}^{k} \boldsymbol{x}_j \begin{bmatrix} \exp(i2\pi f_j) \\ \exp(i2\pi 2 f_j) \\ \cdots \\ \exp(i2\pi n f_j) \end{bmatrix}$$

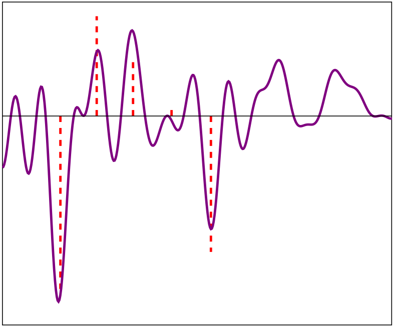
Spectral super-resolution in the presence of outliers



$$z = \sum_{l \in \Omega} z_l \begin{bmatrix} 0 \\ \cdots \\ 1 \\ \cdots \\ 0 \end{bmatrix}$$

Spectral super-resolution in the presence of outliers





$$\boldsymbol{y} = \sum_{j=1}^{k} \boldsymbol{x}_j \begin{bmatrix} \exp\left(i2\pi f_j\right) \\ \exp\left(i2\pi 2 f_j\right) \\ \dots \\ \exp\left(i2\pi n f_j\right) \end{bmatrix} + \sum_{l \in \Omega} \boldsymbol{z}_l \begin{bmatrix} 0 \\ \dots \\ 1 \\ \dots \\ 0 \end{bmatrix}$$

Sinusoidal and spiky atoms

Consider the dictionary

$$\mathcal{D} := \{\boldsymbol{a}(f, 0),\ f \in [0,1]\} \cup \{\boldsymbol{e}(l),\ 1 \leq l \leq n\}$$

where

$$\boldsymbol{a}(f) := \begin{bmatrix} e^{i2\pi f} \\ e^{i2\pi 2f} \\ \cdots \\ e^{i2\pi(n-1)f} \\ e^{i2\pi nf} \end{bmatrix} \qquad \boldsymbol{e}(l) := \begin{bmatrix} 0 \\ \cdots \\ 1 \\ \cdots \\ 0 \end{bmatrix}$$

According to our assumptions

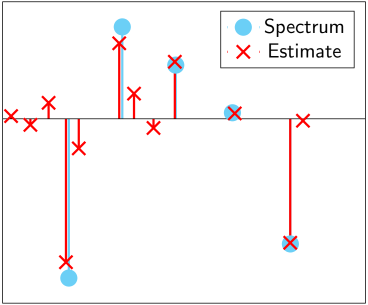
$$\boldsymbol{y} = \sum_{j=1}^{k} \boldsymbol{x}_j \boldsymbol{a}(f_j) + \sum_{l \in \Omega} \boldsymbol{z}_l \boldsymbol{e}(l)$$

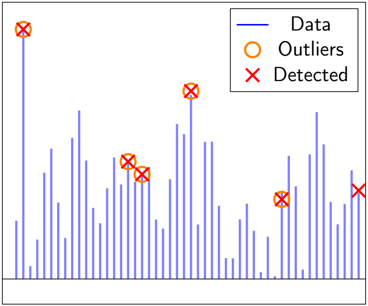
Greedy demixing

Goal: Find sparse decomposition in the dictionary

1. Initialization: Set residual equal to the data vector $\boldsymbol{y}$
2. Selection: Choose atom with higher correlation with residual
3. Pruning: Fit the current atoms to the data and discard any with small contributions, then update the residual

Greedy demixing





Greedy demixing with local optimization

1. Initialization

2. Selection

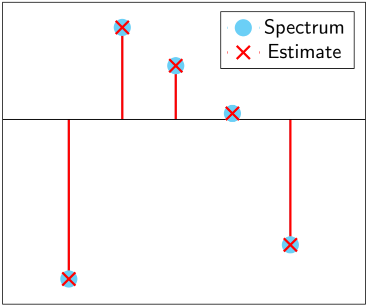
3. Pruning

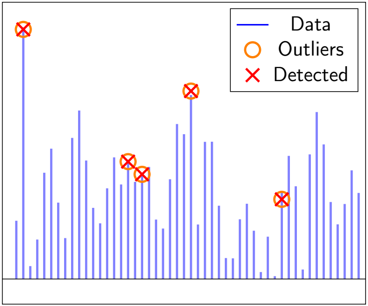
4. Local optimization: Fix the number of sinusoidal atoms $\hat{k}$ and reestimate $f_1, \ldots, f_{\hat{k}}$ by minimizing the function

$$L\left(f_1, \ldots, f_{\hat{k}}\right) := \min_{\hat{\boldsymbol{x}} \in \mathbb{C}^{\hat{k}}, \hat{\boldsymbol{z}} \in \mathbb{C}^{|\widehat{\Omega}|}} \left|\left|\boldsymbol{y} - \sqrt{n} \sum_{j=1}^{\hat{k}} \hat{\boldsymbol{x}}_j \, \boldsymbol{a}\left(f_j, 0\right) - \sum_{l \in \widehat{\Omega}} \hat{\boldsymbol{z}}_l \, \boldsymbol{e}\left(l\right)\right|\right|_2$$

then update the residual

Greedy demixing with local optimization





Conclusion

- Convex programming succeeds beyond compressed sensing if we restrict the class of signals of interest
- A tractable method based on semidefinite programming allows to perform spectral super-resolution in the presence of outliers
- Fast greedy method combined with nonconvex optimization yields promising results

References: Compressed sensing

- *Stable signal recovery from incomplete and inaccurate measurements*. E. J. Candès, J. Romberg and T. Tao. *Comm. Pure Appl. Math.*, 2005
- *Decoding by linear programming*. E. J. Candès and T. Tao. *IEEE Trans. Inform. Theory*, 2004
- *Sparse MRI: The application of compressed sensing for rapid MR imaging*. M. Lustig, D. Donoho and J. M. Pauly. *Magn Reson Med.*, 2007

References: Spectral super-resolution

- *Towards a mathematical theory of super-resolution*. E. J. Candès and C. Fernandez-Granda. *Comm. on Pure and Applied Math.*, 2013
- *Super-resolution of point sources via convex programming*. C. Fernandez-Granda. *Information and Inference*, 2016
- *Compressed Sensing off the grid*. G. Tang, B. Bhaskar, P. Shah, and B. Recht. *IEEE Trans. on Inf. Theory*, 2016
- *Demixing sines and spikes: robust spectral super-resolution in the presence of outliers*. C. Fernandez-Granda, G. Tang, X. Wang, L. Zheng. 2016