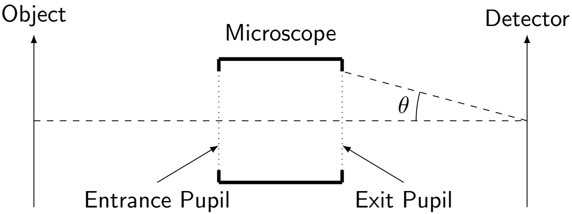
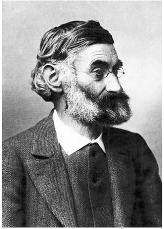
Lecture 12-13: Super-resolution of Positive Sources

V. Morgenshtern

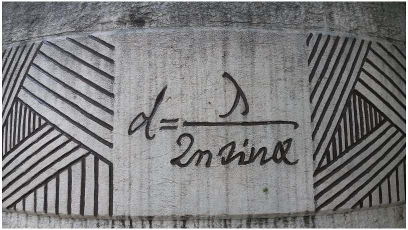
Mathematical Methods in Machine Learning and Signal Processing
SS 2019

Diffraction limits resolution:

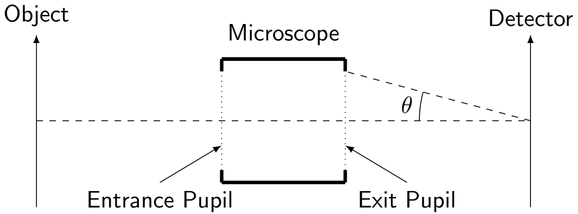


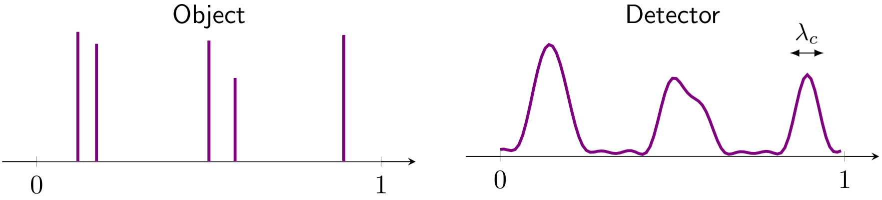


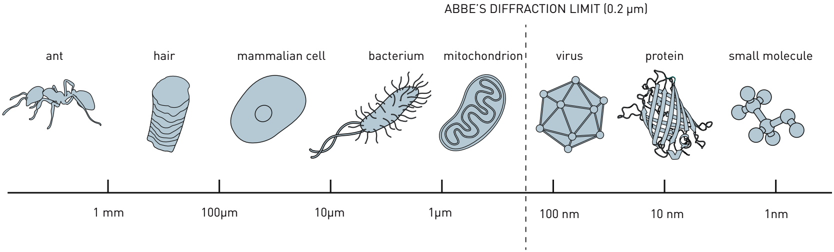
Ernst Abbe



Diffraction limits resolution: $\lambda_c = \frac{\lambda_{\text{LIGHT}}}{2n\sin(\theta)}$

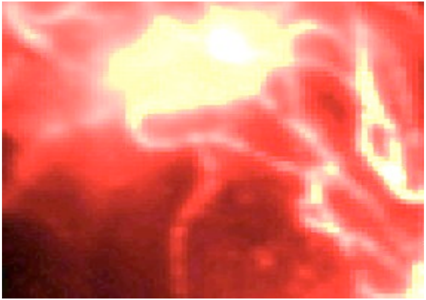






[picture from nobelprize.org]

Looking inside the cell: conventional microscopy



microtubule

Nobel Prize in Chemistry 2014

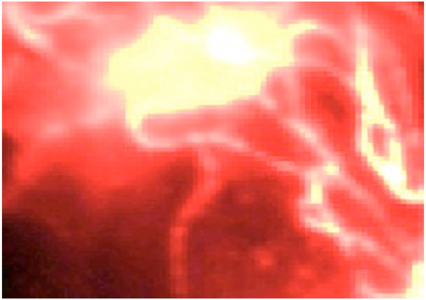
Eric Betzig

Stefan W. Hell

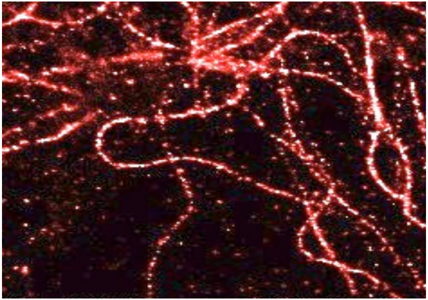
W.E. Moerner

Invention of single-molecule microscopy

Looking inside the cell



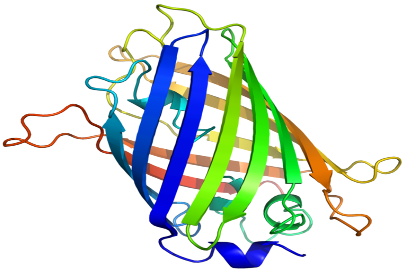
conventional microscopy



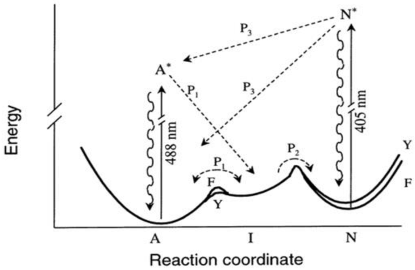
single-molecule microscopy

Single molecule microscopy (basics)

Controlled photoactivation

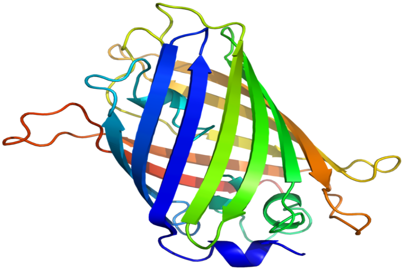


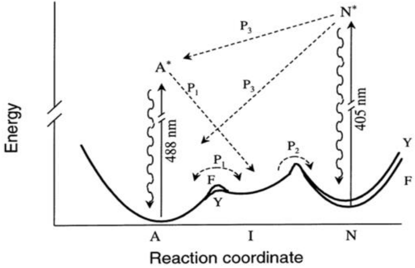
Green fluorescent protein (GFP)



Energy states [Dickson et.al. '97]

Controlled photoactivation



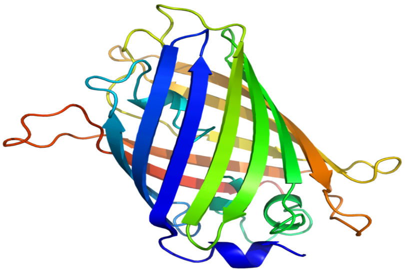


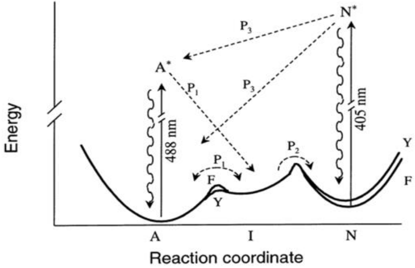
Green fluorescent protein (GFP)

Energy states [Dickson et.al. '97]

- State A is excited to A^* and returns to A upon photon emission

Controlled photoactivation



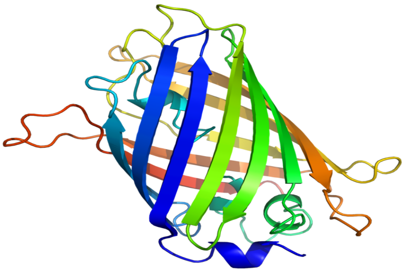


Green fluorescent protein (GFP)

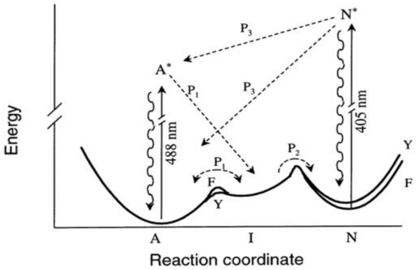
Energy states [Dickson et.al. '97]

- State A is excited to A^* and returns to A upon photon emission
- When I is reached from A there is no fluorescence until I spontaneously moves to A (blinking)

Controlled photoactivation



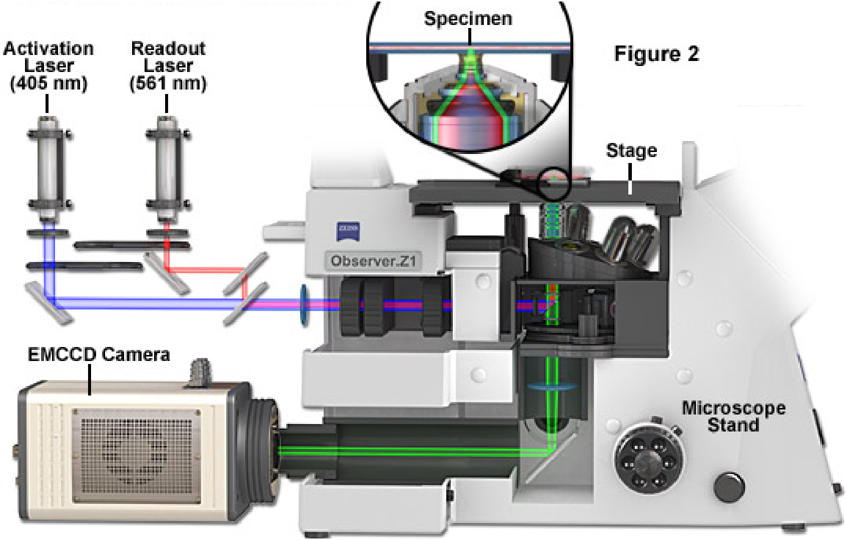
Green fluorescent protein (GFP)



Energy states [Dickson et.al. '97]

- State A is excited to A^* and returns to A upon photon emission
- When I is reached from A there is no fluorescence until I spontaneously moves to A (blinking)
- When I moves to N there is no fluorescence until N is activated by 405nm light and GFP returns to A

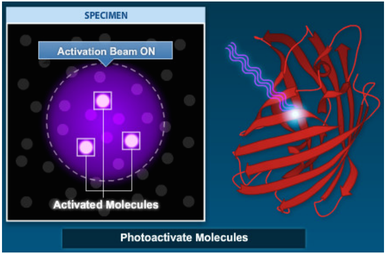
Photoactivated localization microscopy (PALM) Setup



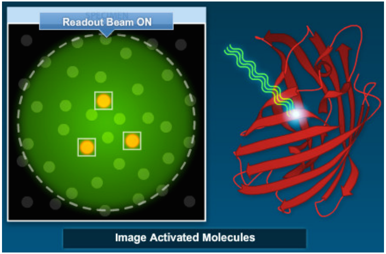
[picture from ZEISS]

PALM Process

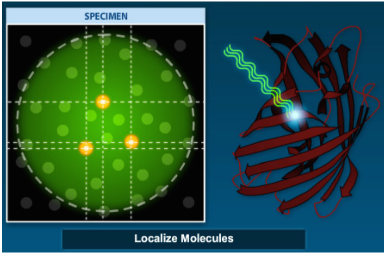
Step 1



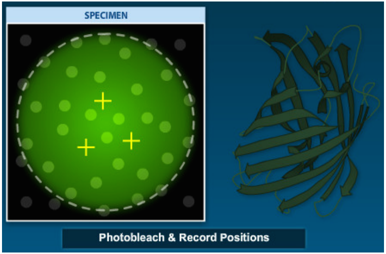
Step 2



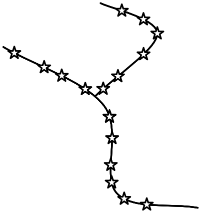
Step 3. Algorithm needed.



Step 4

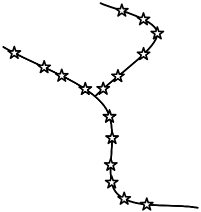


Antibodies: attach fluorescent molecules to the structure

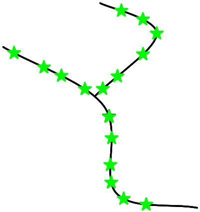


All off

Antibodies: attach fluorescent molecules to the structure

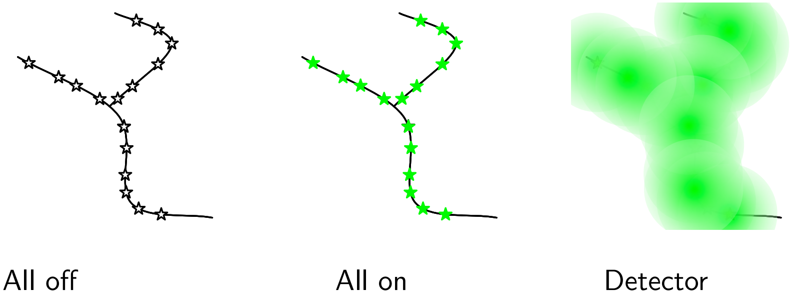


All off



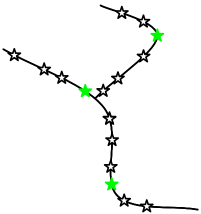
All on

Antibodies: attach fluorescent molecules to the structure



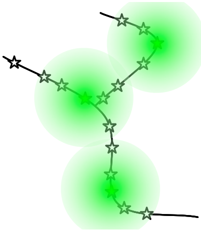
Cannot resolve the structure!

"Blinking" molecules: sparsity



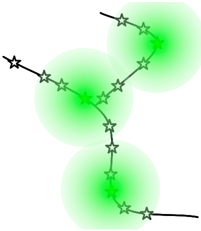
Frame 1

"Blinking" molecules: sparsity



Frame 1

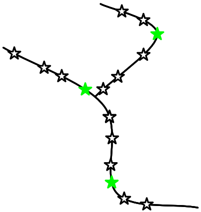
“Blinking” molecules: sparsity



Frame 1

Locate centers of “Gaussian” blobs (parametric estimation)

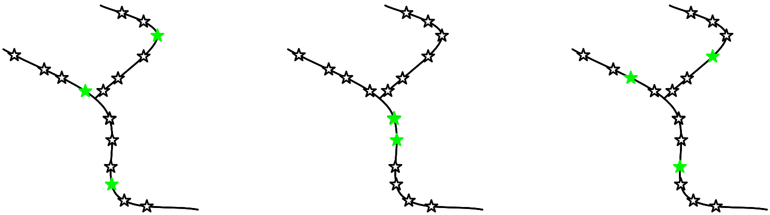
"Blinking" molecules: sparsity



Frame 1

Locate centers of "Gaussian" blobs (parametric estimation)

"Blinking" molecules: sparsity

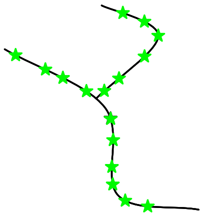


Frame 1 Frame 2 Frame 3

Locate centers of "Gaussian" blobs (parametric estimation)

Combine $\sim$ 10000 frames.

"Blinking" molecules: sparsity



Locate centers of "Gaussian" blobs (parametric estimation)

Combine $\sim$ 10000 frames.

The structure is now resolved!

Next Frontier: image dynamical processes

Imaging ~ 10000 frames is **slow**

Next Frontier: image dynamical processes

Imaging ~ 10000 frames is **slow**

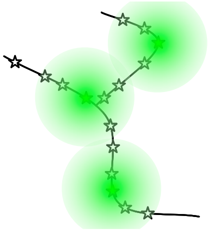
Can we make data acquisition **faster**?

Next Frontier: image dynamical processes

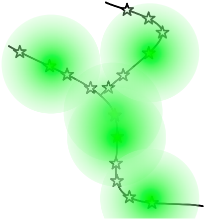
Imaging ~ 10000 frames is **slow**

Can we make data acquisition **faster**?

Image ~ 2500 frames with 4 times more molecules per frame?



parametric estimation works



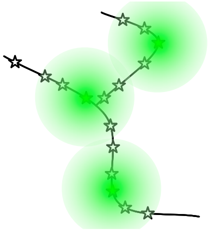
4 times more active molecules
$\Rightarrow$ parametric estimation
does not work

Next Frontier: image dynamical processes

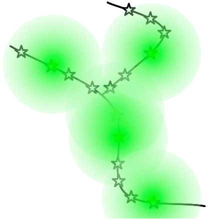
Imaging ~ 10000 frames is **slow**

Can we make data acquisition **faster**?

Image ~ 2500 frames with 4 times more molecules per frame?

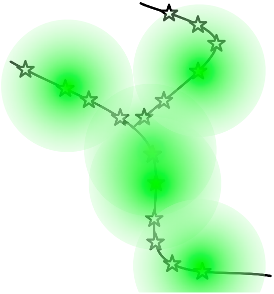


parametric estimation works

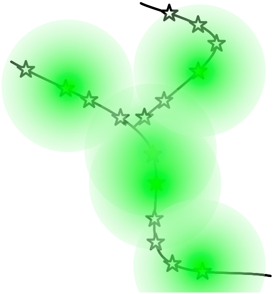


4 times more active molecules
$\Rightarrow$ parametric estimation
does not work

Need powerful super-resolution algorithm!



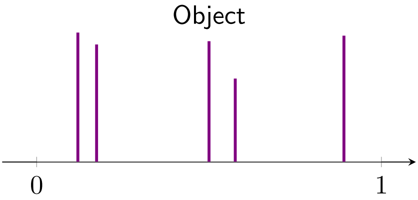
Theory



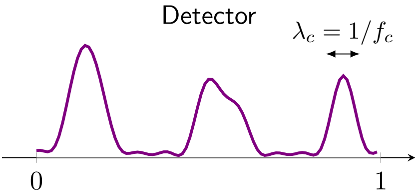
Theory

Which algorithm?
Performance guarantees?
Fundamental limits?

Mathematical model (discrete 1D setup for simplicity)



$$x(t) = \sum_i x_i \delta(t - t_i), \;\; x_i \geq 0$$



$$s(t) = \int f_{\text{low}}(t - t')x(t')dt'$$

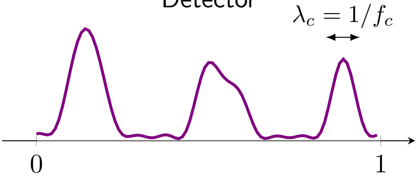
$$f_{\text{low}}(t) = \frac{1}{2f_c}\left(\frac{\sin(2\pi f_c t)}{\pi t}\right)^2$$

Mathematical model (discrete 1D setup for simplicity)

Object

$$x(t) = \sum_i x_i \delta(t - t_i),\ \ x_i \geq 0$$

Detector



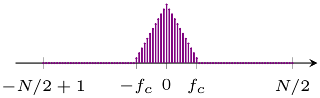
$$s(t) = \int f_{\text{low}}(t - t')x(t')dt'$$

$$\mathbf{x} = [x_0 \cdots x_{N-1}]^{\mathsf{T}} \geq \mathbf{0}$$

$$\mathbf{s} = \mathbf{P}\mathbf{x} + \mathbf{z}$$

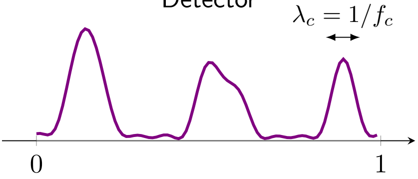
$\mathbf{P} = \mathbf{P}_{\text{tri}}$ is **circulant**

Triangular spectrum



Mathematical model (discrete 1D setup for simplicity)

$$x(t) = \sum_i x_i \delta(t - t_i),\ x_i \geq 0$$



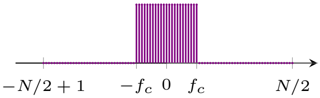
$$s(t) = \int f_{\text{low}}(t - t')x(t')dt'$$

$$\mathbf{x} = [x_0 \cdots x_{N-1}]^\mathsf{T} \geq \mathbf{0}$$

$$\mathbf{s} = \mathbf{P}\mathbf{x} + \mathbf{z}$$

$\mathbf{P} = \mathbf{P}_{\text{flat}}$ is **circulant**

Flat spectrum



Mathematical model (discrete 1D setup for simplicity)

$$\mathbf{P} = \mathbf{F}^{\mathsf{H}}\hat{\mathbf{P}}\mathbf{F}$$

DFT:

$$[\mathbf{F}]_{k,l} = \frac{1}{\sqrt{N}}\, e^{-\mathrm{i}2\pi kl/N}, \quad -N/2+1 \le k \le N/2,\ 0 \le l \le N-1$$

Spectrum:

$$\hat{\mathbf{P}} = \mathrm{diag}([\hat{p}_{-N/2+1} \cdots \hat{p}_{N/2}]^{\mathsf{T}})$$

- **Flat:**

$$\hat{p}_k = \begin{cases} 1, & k = f_c, \ldots, f_c, \\ 0, & \text{otherwise} \end{cases}$$

- **Triangular:**

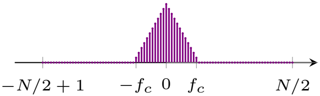
$$\hat{p}_k = \begin{cases} 1 - \frac{|k|}{f_c+1}, & k = -f_c, \ldots, f_c \\ 0, & \text{otherwise} \end{cases}$$

Width of the convolution kernel: $\lambda_c \triangleq 1/f_c$

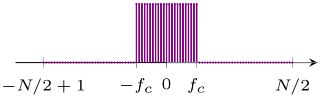
Super-resolution factor and stability

$$\mathbf{x} = [x_0 \cdots x_{N-1}]^\mathsf{T}$$

Triangular spectrum



Flat spectrum



$$\mathbf{s} = \mathbf{P}\mathbf{x} + \mathbf{z}$$

$$\mathrm{SRF} \triangleq N/(2f_c)$$

Super-resolution factor and stability

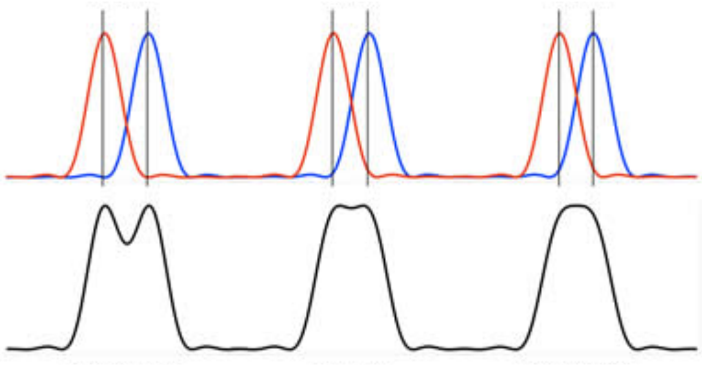
$$\mathbf{x} = [x_0 \cdots x_{N-1}]^\mathsf{T}$$

$$\mathbf{s} = \mathbf{P}\mathbf{x} + \mathbf{z}$$

$$\mathrm{SRF} \triangleq N/(2f_c)$$

Stability: $\|\mathbf{x} - \hat{\mathbf{x}}\| \overset{?}{\leq} \|\mathbf{z}\| \cdot$ (amplification factor)

Classical resolution criteria: separation is about λ_c



$1.22\lambda_c$ Rayleigh criterion

λ_c Abbe criterion

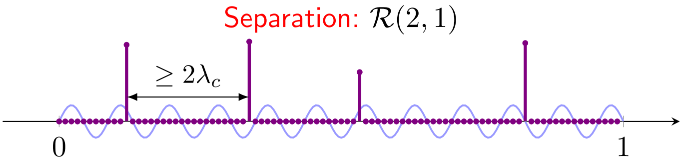
$0.94\lambda_c$ Sparrow criterion

Rayleigh-regularity: $\mathbf{x} \in \mathcal{R}(d, r)$

$\mathbf{x}$ has fewer than r spikes in every $\lambda_c d$ interval $[\lambda_c \triangleq 1/f_c]$

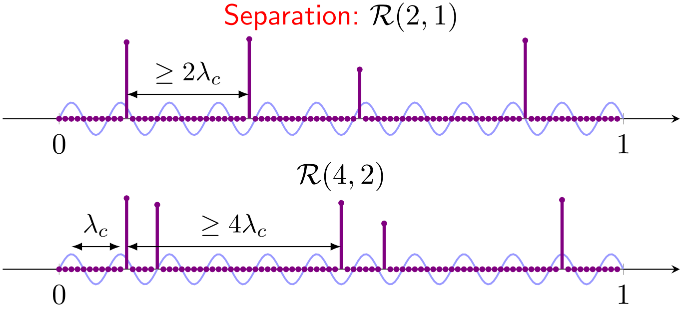
Rayleigh-regularity: $\mathbf{x} \in \mathcal{R}(d, r)$

$\mathbf{x}$ has fewer than r spikes in every $\lambda_c d$ interval $[\lambda_c \triangleq 1/f_c]$



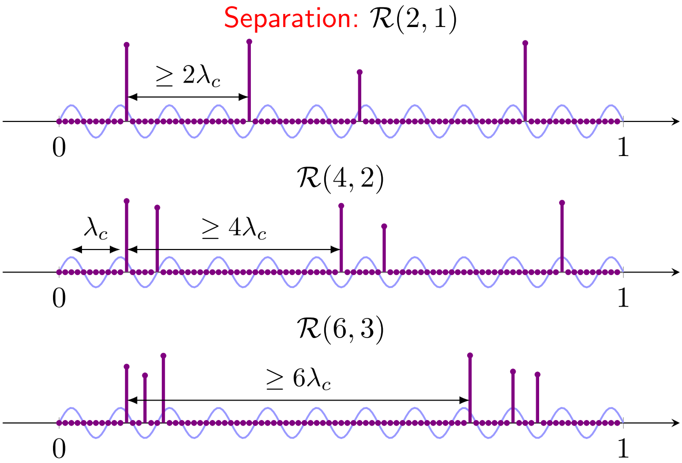
Rayleigh-regularity: $\mathbf{x} \in \mathcal{R}(d, r)$

$\mathbf{x}$ has fewer than r spikes in every $\lambda_c d$ interval $[\lambda_c \triangleq 1/f_c]$



Rayleigh-regularity: $\mathbf{x} \in \mathcal{R}(d, r)$

$\mathbf{x}$ has fewer than r spikes in every $\lambda_c d$ interval $[\lambda_c \triangleq 1/f_c]$

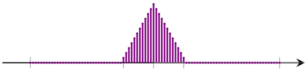


Main results

Recall:

$$\mathbf{s} = \mathbf{P}\mathbf{x} + \mathbf{z}$$

spectrum



Solve:

$$\text{minimize } \|\mathbf{s} - \mathbf{P}\hat{\mathbf{x}}\|_1 \quad \text{subject to} \quad \hat{\mathbf{x}} \geq 0$$

Theorem: (V.Morgenshter, Càndes'14, [1])

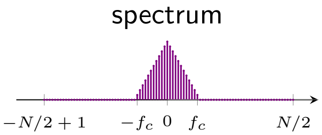
Take $\mathbf{P} = \mathbf{P}_{\text{tri}}$ or $\mathbf{P} = \mathbf{P}_{\text{flat}}$. Assume $\mathbf{x} \geq 0$, $\mathbf{x} \in \mathcal{R}(2r, r)$. Then,

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_1 \leq c \cdot \|\mathbf{z}\|_1 \cdot \left(\frac{N}{2f_c}\right)^{2r}.$$

Main results

Recall:

$$\mathbf{s} = \mathbf{P}\mathbf{x} + \mathbf{z}$$



Solve:

$$\text{minimize } \|\mathbf{s} - \mathbf{P}\hat{\mathbf{x}}\|_1 \quad \text{subject to} \quad \hat{\mathbf{x}} \geq 0$$

Theorem: (V.Morgenshter, Càndes'14, [1])

Take $\mathbf{P} = \mathbf{P}_{\text{tri}}$ or $\mathbf{P} = \mathbf{P}_{\text{flat}}$. Assume $\mathbf{x} \geq 0$, $\mathbf{x} \in \mathcal{R}(2r, r)$. Then,

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_1 \leq c \cdot \|\mathbf{z}\|_1 \cdot \left(\frac{N}{2f_c}\right)^{2r}.$$

Converse: (V.Morgenshter, Càndes'14, [1])

For $\mathbf{P} = \mathbf{P}_{\text{tri}}$, no algorithm can do better than $c \cdot \|\mathbf{z}\|_1 \cdot \left(\frac{N}{2f_c}\right)^{2r-1}$.

Key ideas

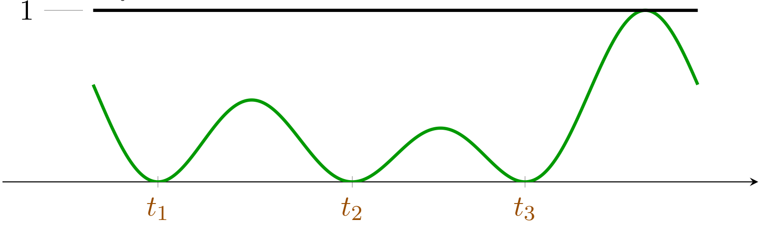
→ **Duality theory:** to prove stability we need a low-frequency trigonometric polynomial that is "curvy"

- **[Dohono, et al.'92, Fuchs'05]** construct trigonometric polynomial that is not "curvy"
- **[Candès and Fernandez-Granda'12]** construct trigonometric polynomial that is "curvy", but construction needs separation
- **New construction:** multiply "curvy" trigonometric polynomials
  - "curvy"
  - construction needs no separation

Dual certificate (noisy case)

- $\mathcal{T}$ is the support of $\mathbf{x}$
- Suppose, we can construct a **low-frequency trig. polynomial**:

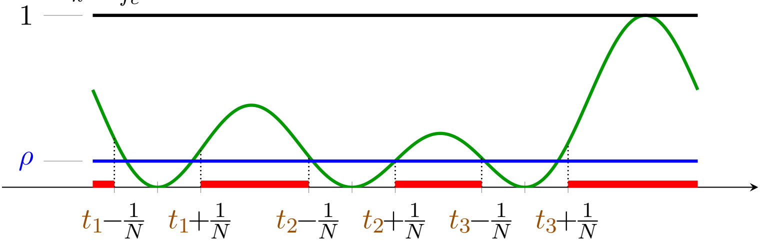
$$q(t) = \sum_{k=-f_c}^{f_c} \hat{q}_k e^{-i2\pi kt}, \quad 0 \leq q(t) \leq 1, \quad q(t_i) = 0 \text{ for all } t_i \in \mathcal{T}.$$



Dual certificate (noisy case)

- $\mathcal{T}$ is the support of $\mathbf{x}$
- Suppose, we can construct a **low-frequency trig. polynomial**:

$$q(t) = \sum_{k=-f_c}^{f_c} \hat{q}_k e^{-i2\pi kt}, \quad 0 \leq q(t) \leq 1, \quad q(t_i) = 0 \text{ for all } t_i \in \mathcal{T}.$$



- Then, $\|\hat{\mathbf{x}} - \mathbf{x}\|_1 \leq 4\|\mathbf{z}\|_1/\rho$.

Proof of Lemma

- Set:
  $\mathbf{h} = [h_0 \cdots h_{N-1}]^\mathsf{T} = \hat{\mathbf{x}} - \mathbf{x}, \quad \mathcal{T} = \{l/N : h_l < 0\}$

Proof of Lemma

- Set:
  $\mathbf{h} = [h_0 \cdots h_{N-1}]^\mathsf{T} = \hat{\mathbf{x}} - \mathbf{x}, \quad \mathcal{T} = \{l/N : h_l < 0\} \subset \mathrm{supp}(\mathbf{x}).$

Proof of Lemma

- Set:
  $\mathbf{h} = [h_0 \cdots h_{N-1}]^\mathsf{T} = \hat{\mathbf{x}} - \mathbf{x}, \quad \mathcal{T} = \{l/N : h_l < 0\} \subset \mathrm{supp}(\mathbf{x}).$
- Dual vector (contains samples) $q_l = q(l/N)$ satisfies:

$$\mathbf{P}_{\mathrm{flat}}\mathbf{q} = \mathbf{q}, \quad \|\mathbf{q}\|_\infty = 1, \quad \text{and} \quad \begin{cases} q_l = 0, & l/N \in \mathcal{T} \\ q_l > \rho, & \text{otherwise.} \end{cases}$$

Proof of Lemma

- Set:
  $\mathbf{h} = [h_0 \cdots h_{N-1}]^\mathsf{T} = \hat{\mathbf{x}} - \mathbf{x}, \quad \mathcal{T} = \{l/N : h_l < 0\} \subset \mathrm{supp}(\mathbf{x}).$
- Dual vector (contains samples) $q_l = q(l/N)$ satisfies:

$$\mathbf{P}_{\text{flat}}\mathbf{q} = \mathbf{q}, \quad \|\mathbf{q}\|_\infty = 1, \quad \text{and} \quad \begin{cases} q_l = 0, & l/N \in \mathcal{T} \\ q_l > \rho, & \text{otherwise.} \end{cases}$$

- On the one hand:

$$\begin{aligned} |\langle \mathbf{q} - \rho/2, \mathbf{h}\rangle| &= |\langle \mathbf{P}(\mathbf{q} - \rho/2), \mathbf{h}\rangle| = |\langle \mathbf{q} - \rho/2, \mathbf{P}\mathbf{h}\rangle| \\ &\le \|\mathbf{q} - \rho/2\|_\infty \|\mathbf{P}\mathbf{h}\|_1 \le \|\mathbf{P}\mathbf{x} - \mathbf{s} + \mathbf{s} - \mathbf{P}\hat{\mathbf{x}}\|_1 \\ &\le \|\mathbf{P}\mathbf{x} - \mathbf{s}\|_1 + \|\mathbf{s} - \mathbf{P}\hat{\mathbf{x}}\|_1 \\ &\le 2\|\mathbf{P}\mathbf{x} - \mathbf{s}\|_1 \le 2\|\mathbf{z}\|_1. \end{aligned}$$

Proof of Lemma

- Set:
  $\mathbf{h} = [h_0 \cdots h_{N-1}]^\mathsf{T} = \hat{\mathbf{x}} - \mathbf{x}, \quad \mathcal{T} = \{l/N : h_l < 0\} \subset \mathrm{supp}(\mathbf{x}).$
- Dual vector (contains samples) $q_l = q(l/N)$ satisfies:

$$\mathbf{P}_{\mathrm{flat}}\mathbf{q} = \mathbf{q}, \quad \|\mathbf{q}\|_\infty = 1, \quad \text{and} \quad \begin{cases} q_l = 0, & l/N \in \mathcal{T} \\ q_l > \rho, & \text{otherwise.} \end{cases}$$

- On the one hand:

$$\begin{aligned} |\langle \mathbf{q} - \rho/2, \mathbf{h} \rangle| &= |\langle \mathbf{P}(\mathbf{q} - \rho/2), \mathbf{h} \rangle| = |\langle \mathbf{q} - \rho/2, \mathbf{P}\mathbf{h} \rangle| \\ &\leq \|\mathbf{q} - \rho/2\|_\infty \|\mathbf{P}\mathbf{h}\|_1 \leq \|\mathbf{P}\mathbf{x} - \mathbf{s} + \mathbf{s} - \mathbf{P}\hat{\mathbf{x}}\|_1 \\ &\leq \|\mathbf{P}\mathbf{x} - \mathbf{s}\|_1 + \|\mathbf{s} - \mathbf{P}\hat{\mathbf{x}}\|_1 \\ &\leq 2\|\mathbf{P}\mathbf{x} - \mathbf{s}\|_1 \leq 2\|\mathbf{z}\|_1. \end{aligned}$$

- On the other hand:

$$|\langle \mathbf{q} - \rho/2, \mathbf{h} \rangle| = \left| \sum_{l=0}^{N-1} (q_l - \rho/2) h_l \right| = \sum_{l=0}^{N-1} (q_l - \rho/2) h_l \geq \rho \|\mathbf{h}\|_1 / 2.$$

Proof of Lemma

- Set:
  $\mathbf{h} = [h_0 \cdots h_{N-1}]^{\mathsf{T}} = \hat{\mathbf{x}} - \mathbf{x}, \quad \mathcal{T} = \{l/N : h_l < 0\} \subset \mathrm{supp}(\mathbf{x}).$
- Dual vector (contains samples) $q_l = q(l/N)$ satisfies:

$$\mathbf{P}_{\text{flat}}\mathbf{q} = \mathbf{q}, \quad \|\mathbf{q}\|_\infty = 1, \quad \text{and} \quad \begin{cases} q_l = 0, & l/N \in \mathcal{T} \\ q_l > \rho, & \text{otherwise.} \end{cases}$$

- On the one hand:

$$\begin{aligned} |\langle \mathbf{q} - \rho/2, \mathbf{h} \rangle| &= |\langle \mathbf{P}(\mathbf{q} - \rho/2), \mathbf{h} \rangle| = |\langle \mathbf{q} - \rho/2, \mathbf{P}\mathbf{h} \rangle| \\ &\le \|\mathbf{q} - \rho/2\|_\infty \|\mathbf{P}\mathbf{h}\|_1 \le \|\mathbf{P}\mathbf{x} - \mathbf{s} + \mathbf{s} - \mathbf{P}\hat{\mathbf{x}}\|_1 \\ &\le \|\mathbf{P}\mathbf{x} - \mathbf{s}\|_1 + \|\mathbf{s} - \mathbf{P}\hat{\mathbf{x}}\|_1 \\ &\le 2\|\mathbf{P}\mathbf{x} - \mathbf{s}\|_1 \le 2\|\mathbf{z}\|_1. \end{aligned}$$

- On the other hand:

$$|\langle \mathbf{q} - \rho/2, \mathbf{h} \rangle| = \left| \sum_{l=0}^{N-1} (q_l - \rho/2) h_l \right| = \sum_{l=0}^{N-1} (q_l - \rho/2) h_l \ge \rho \|\mathbf{h}\|_1 / 2.$$

- Combining: $\|\mathbf{h}\|_1 \le 4\|\mathbf{z}\|_1/\rho$.

Key ideas

- **Duality theory:** to prove stability we need a low-frequency trigonometric polynomial that is "curvy"

→ **[Dohono, et al.'92, Fuchs'05]** construct trigonometric polynomial that is not "curvy"

- **[Candès and Fernandez-Granda'12]** construct trigonometric polynomial that is "curvy", but construction needs separation
- **New construction:** multiply "curvy" trigonometric polynomials
  - "curvy"
  - construction needs no separation

Donoho, et al.'92, Fuchs'05, [2, 3]: "Classical" q(t)

$$q(t) = \prod_{t_0 \in \mathcal{T}} \frac{1}{2} \left[\cos(2\pi(t + 1/2 - t_0)) + 1 \right].$$

Euler's formula:

$$\cos(2\pi t) = \frac{e^{i2\pi t} + e^{-i2\pi t}}{2}$$

Sparsity implies $q(t)$ is low-frequency:

$$q(t) = \sum_{k=-f_c}^{f_c} \hat{q}_k e^{-i2\pi k t} \text{ if } |\mathcal{T}| \leq f_c$$

$s \leq \left|\mathcal{T}\right| \leq f_c = \frac{1}{2} \times$ number of measurements

No square-root bottleneck!

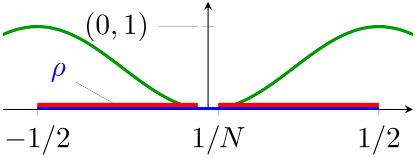
Donoho et al.'92 [2], Fuchs'05 [3]: "Classical" q(t)

$$q(t) = \prod_{t_0 \in \mathcal{T}} \frac{1}{2} \left[\cos(2\pi(t + 1/2 - t_0)) + 1 \right].$$

No separation required

Low curvature!

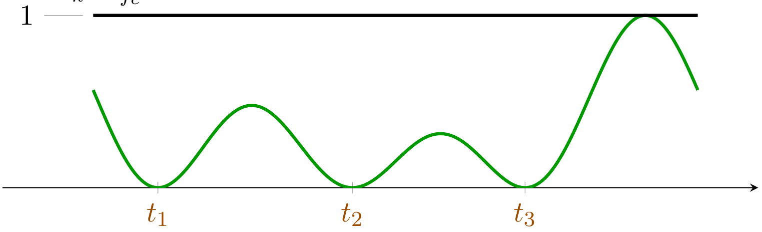
$$q(t - t_0) \approx (t - t_0)^2 \Rightarrow \|\mathbf{x} - \hat{\mathbf{x}}\|_1 \leq \|\mathbf{z}\|_1 \cdot N^2$$



Dual certificate (noiseless case, $\mathbf{z} = \mathbf{0}$)

- $\mathcal{T}$ is the support of $\mathbf{x}$
- Suppose, we can construct a **low-frequency trig. polynomial**:

$$q(t) = \sum_{k=-f_c}^{f_c} \hat{q}_k e^{-i2\pi kt}, \quad 0 \leq q(t) \leq 1, \quad q(t_i) = 0 \text{ for all } t_i \in \mathcal{T}.$$



- Then, $\hat{\mathbf{x}} = \mathbf{x}$.

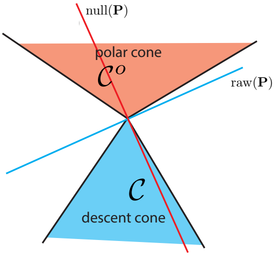
Connection to LASSO ($\mathbf{x}$ can be negative here)

$$\text{minimize} \quad \|\hat{\mathbf{x}}\|_1 \quad \text{subject to} \quad \mathbf{s} = \mathbf{P}\hat{\mathbf{x}}$$

- $\hat{\mathbf{x}} = \mathbf{x}$ iff there exists $\mathbf{q} \perp \text{null}(\mathbf{P})$ and $\mathbf{q} \in \partial\|\mathbf{x}\|_1$

- $\mathbf{P}$ is orthogonal projection onto the set of low-freq. trig. polynomials: $\mathbf{q} \perp \text{null}(\mathbf{P}) \Leftrightarrow$
$$q(t) = \sum_{k=-f_c}^{f_c} \hat{q}_k e^{-i2\pi kt}$$

- $\mathbf{q} \in \partial\|\mathbf{x}\|_1 \Leftrightarrow$
$$\begin{cases} q(t_i) = \text{sign}(x_i) & x_i \neq 0 \\ |q(t_i)| \leq 1 & x_i = 0 \end{cases}$$



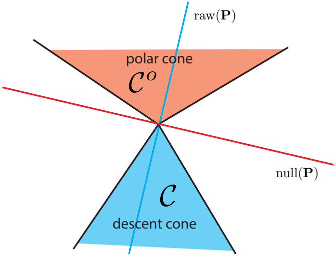
Connection to LASSO ($\mathbf{x}$ can be negative here)

$$\text{minimize} \quad \|\hat{\mathbf{x}}\|_1 \quad \text{subject to} \quad \mathbf{s} = \mathbf{P}\hat{\mathbf{x}}$$

- $\hat{\mathbf{x}} = \mathbf{x}$ iff there exists $\mathbf{q} \perp \mathrm{null}(\mathbf{P})$ and $\mathbf{q} \in \partial\|\mathbf{x}\|_1$

- $\mathbf{P}$ is orthogonal projection onto the set of low-freq. trig. polynomials: $\mathbf{q} \perp \mathrm{null}(\mathbf{P}) \Leftrightarrow$
$$q(t) = \sum_{k=-f_c}^{f_c} \hat{q}_k e^{-i2\pi kt}$$

- $\mathbf{q} \in \partial\|\mathbf{x}\|_1 \Leftrightarrow$
$$\begin{cases} q(t_i) = \mathrm{sign}(x_i) & x_i \neq 0 \\ |q(t_i)| \leq 1 & x_i = 0 \end{cases}$$



Connection to LASSO ($\mathbf{x}$ can be negative here)

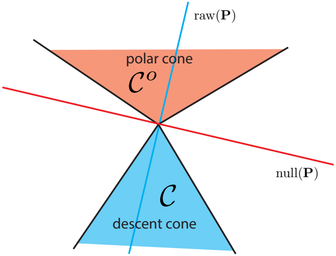
minimize $\|\hat{\mathbf{x}}\|_1$ subject to $\mathbf{s} = \mathbf{P}\hat{\mathbf{x}}$

- $\hat{\mathbf{x}} = \mathbf{x}$ iff there exists $\mathbf{q} \perp \text{null}(\mathbf{P})$ and $\mathbf{q} \in \partial\|\mathbf{x}\|_1$
- $\mathbf{P}$ is orthogonal projection onto the set of low-freq. trig. polynomials: $\mathbf{q} \perp \text{null}(\mathbf{P}) \Leftrightarrow$

$$q(t) = \sum_{k=-f_c}^{f_c} \hat{q}_k e^{-\mathrm{i}2\pi kt}$$

- $\mathbf{q} \in \partial\|\mathbf{x}\|_1 \Leftrightarrow$

$$\begin{cases} q(t_i) = \text{sign}(x_i) & x_i \neq 0 \\ |q(t_i)| \leq 1 & x_i = 0 \end{cases}$$



Connection to LASSO ($\mathbf{x}$ can be negative here)

$$\text{minimize} \quad \|\hat{\mathbf{x}}\|_1 \quad \text{subject to} \quad \mathbf{s} = \mathbf{P}\hat{\mathbf{x}}$$

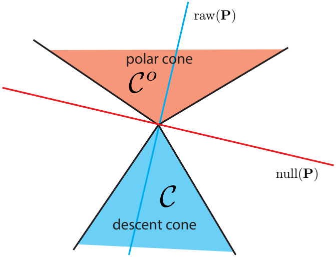
- $\hat{\mathbf{x}} = \mathbf{x}$ iff there exists $\mathbf{q} \perp \mathrm{null}(\mathbf{P})$ and $\mathbf{q} \in \partial\|\mathbf{x}\|_1$

- $\mathbf{P}$ is orthogonal projection onto the set of low-freq. trig. polynomials: $\mathbf{q} \perp \mathrm{null}(\mathbf{P}) \Leftrightarrow$

$$q(t) = \sum_{k=-f_c}^{f_c} \hat{q}_k e^{-i2\pi kt}$$

- $\mathbf{q} \in \partial\|\mathbf{x}\|_1 \Leftrightarrow$

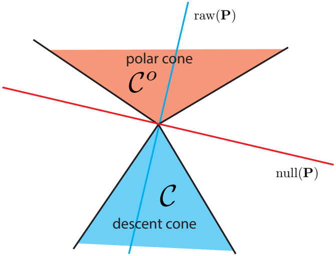
$$\begin{cases} q(t_i) = \mathrm{sign}(x_i) & x_i \neq 0 \\ |q(t_i)| \leq 1 & x_i = 0 \end{cases}$$

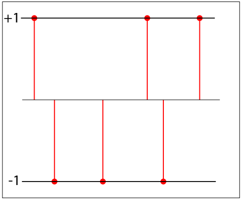


Connection to LASSO ($\mathbf{x}$ can be negative here)

minimize $\quad \|\hat{\mathbf{x}}\|_1 \quad$ subject to $\quad \mathbf{s} = \mathbf{P}\hat{\mathbf{x}}$

- $\hat{\mathbf{x}} = \mathbf{x}$ iff there exists $\mathbf{q} \perp \text{null}(\mathbf{P})$ and $\mathbf{q} \in \partial\|\mathbf{x}\|_1$
- $\mathbf{P}$ is orthogonal projection onto the set of low-freq. trig. polynomials: $\mathbf{q} \perp \text{null}(\mathbf{P}) \Leftrightarrow$
$$q(t) = \sum_{k=-f_c}^{f_c} \hat{q}_k e^{-\mathrm{i}2\pi kt}$$
- $\mathbf{q} \in \partial\|\mathbf{x}\|_1 \Leftrightarrow$
$$\begin{cases} q(t_i) = \text{sign}(x_i) & x_i \neq 0 \\ |q(t_i)| \leq 1 & x_i = 0 \end{cases}$$





$\text{sign}(x) \quad (x \neq 0)$

Key ideas

- **Duality theory:** to prove stability we need a low-frequency trigonometric polynomial that is "curvy"
- **[Dohono, et al.'92, Fuchs'05]** construct trigonometric polynomial that is not "curvy"

→ **[Candès and Fernandez-Granda'12]** construct trigonometric polynomial that is "curvy", but construction needs separation

- **New construction:** multiply "curvy" trigonometric polynomials
  - "curvy"
  - construction needs no separation

Candès, Fernandez-Granda'12, [4]: "Curvy" $q(t)$

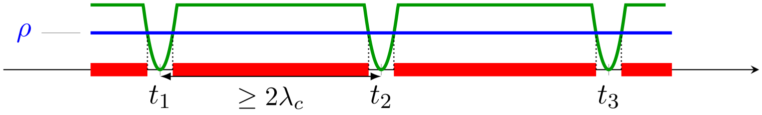
$$q(t) = \sum_{t_j \in \mathcal{T}} a_j K(t - t_j) + \text{corrections},$$

$K(t) \ldots$ low-frequency and "curvy"

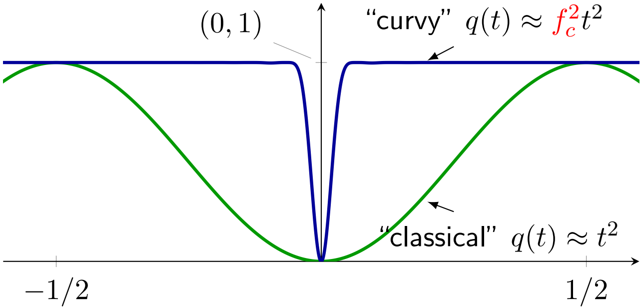
Separation between zeros required: $\mathcal{T} \in \mathcal{R}(2,1)$

High curvature!

$$q(t - t_i) \approx f_c^2 (t - t_i)^2 \Rightarrow \|\mathbf{x} - \hat{\mathbf{x}}\|_1 \leq c \cdot \|\mathbf{z}\|_1 \cdot \left(\frac{N}{2f_c}\right)^2$$



Comparison of Trigonometric Polynomials



Key ideas

- **Duality theory:** to prove stability we need a low-frequency trigonometric polynomial that is "curvy"
- **[Dohono, et al.'92], [Fuchs'05]:** construct trigonometric polynomial that is not "curvy"
- **[Candès and Fernandez-Granda'12]:**, construct trigonometric polynomial that is "curvy", but construction needs separation

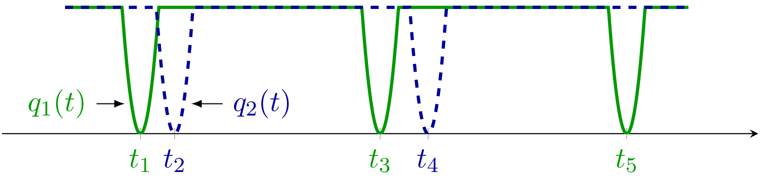
→ **New construction:** multiply "curvy" trigonometric polynomials

- "curvy"
- construction needs no separation

New construction: curvature without separation

Partition support: $\mathcal{T} = \mathcal{T}_1 \cup \mathcal{T}_2, \quad r = 2$

Regularity: $\mathcal{T} \in \mathcal{R}(2 \cdot 2, 2) \Rightarrow \mathcal{T}_i \in \mathcal{R}(4, 1)$

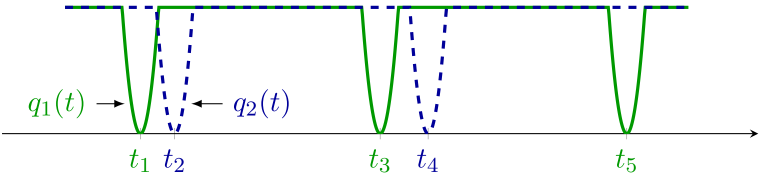


$$q(t; f_c) = q_1(t; f_c/2) \times q_2(t; f_c/2)$$

New construction: curvature without separation

Partition support: $\mathcal{T} = \mathcal{T}_1 \cup \mathcal{T}_2, \quad r = 2$

Regularity: $\mathcal{T} \in \mathcal{R}(2 \cdot 2, 2) \Rightarrow \mathcal{T}_i \in \mathcal{R}(4, 1)$



$$q(t; f_c) = q_1(t; f_c/2) \times q_2(t; f_c/2)$$

High curvature!

$$q(t - t_i) \approx \frac{f_c^{2r}}{r^{2r}}(t - t_i)^{2r} \Rightarrow \|\mathbf{x} - \hat{\mathbf{x}}\|_1 \leq c \cdot \|\mathbf{z}\|_1 \cdot \left(\frac{N}{2f_c}\right)^{2r}$$

Summation vs. multiplication

Remember: $q(t)$ must be frequency-limited to f_c!

Summation vs. multiplication

Remember: $q(t)$ must be frequency-limited to f_c!

[Donoho et.al.'92, Fuchs'05]:

$$q(t) = \prod_{t_j \in \mathcal{T}} \underbrace{\frac{1}{2}\left[\cos(2\pi(t + 1/2 - t_j)) + 1\right]}_{\text{frequency one}}$$

Summation vs. multiplication

Remember: $q(t)$ must be frequency-limited to f_c!

[Donoho et.al.'92, Fuchs'05]:

$$q(t) = \prod_{t_j \in \mathcal{T}} \underbrace{\frac{1}{2}\left[\cos(2\pi(t + 1/2 - t_j)) + 1\right]}_{\text{frequency one}}$$

[Candès, Fernandez-Granda,'12]:

$$q(t) = \sum_{t_j \in \mathcal{T}} \underbrace{a_j K(t - t_j)}_{\text{frequency } f_c}$$

Summation vs. multiplication

Remember: $q(t)$ must be frequency-limited to f_c!

[Donoho et.al.'92, Fuchs'05]:

$$q(t) = \prod_{t_j \in \mathcal{T}} \underbrace{\frac{1}{2}\left[\cos(2\pi(t + 1/2 - t_j)) + 1\right]}_{\text{frequency one}}$$

[Candès, Fernandez-Granda,'12]:

$$q(t) = \sum_{t_j \in \mathcal{T}} \underbrace{a_j K(t - t_j)}_{\text{frequency } f_c}$$

This work:

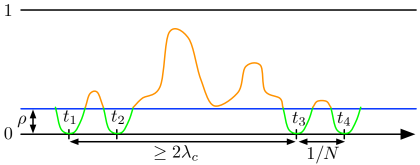
$$q(t) = \prod_{k=1}^{r} \sum_{t_{jk} \in \mathcal{T}_k} \underbrace{a_{jk} K(t - t_{jk})}_{\text{frequency } f_c/r}$$

Complex vs. positive signals

Why do we need $\mathbf{x} \geq \mathbf{0}$?

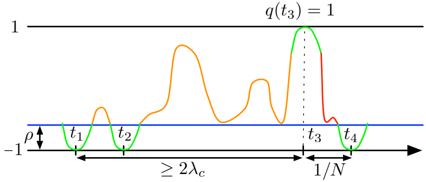
$\mathbf{x} \geq \mathbf{0}$

Interpolate **zero** on supp. of $\mathbf{x}$



$\mathbf{x} \in \mathbb{C}^N$

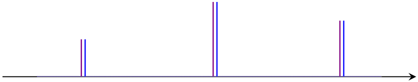
Interpolate $\operatorname{sign}(\mathbf{x})$ on supp. of $\mathbf{x}$



Does not exist! (Bernstein Th.)

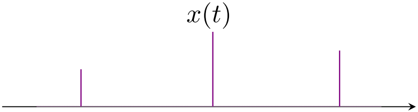
Continuous setup

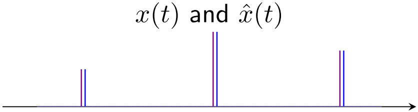
f_c fixed, $N \to \infty \Rightarrow \mathrm{SRF}_{\mathrm{OLD}} \to \infty$



Is the problem hopeless?

f_c fixed, $N \to \infty \Rightarrow \mathrm{SRF}_{\mathrm{OLD}} \to \infty$





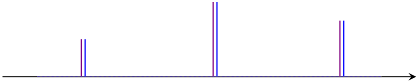
Is the problem hopeless?

No: we need to be less ambitions!

f_c fixed, $N \to \infty \Rightarrow \mathrm{SRF}_{\mathrm{OLD}} \to \infty$

$x(t)$

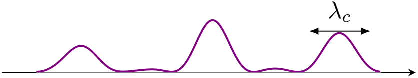
$x(t)$ and $\hat{x}(t)$



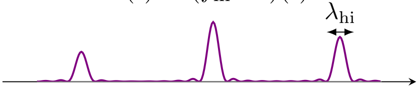
Is the problem hopeless?

No: we need to be less ambitions!

$s(t) = (f_{\mathrm{low}} \star x)(t)$



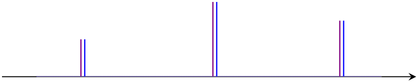
$\hat{x}(t) = (f_{\mathrm{hi}} \star x)(t)$



f_c fixed, $N \to \infty \Rightarrow \mathrm{SRF}_{\mathrm{OLD}} \to \infty$

$x(t)$

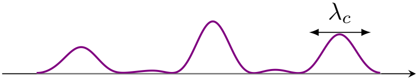
$x(t)$ and $\hat{x}(t)$



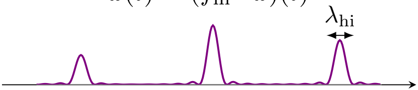
Is the problem hopeless?

No: we need to be less ambitions!

$s(t) = (f_{\mathrm{low}} \star x)(t)$



$\hat{x}(t) = (f_{\mathrm{hi}} \star x)(t)$



Error$= \| f_{\mathrm{hi}} \star (x - \hat{x}) \|_1$

$$\mathrm{SRF}_{\mathrm{NEW}} = \lambda_c / \lambda_{\mathrm{hi}}$$

Need new tools

Theorem: (V. Morgenshtern, 2019, [5])

Assume $x(t) \geq 0$, $x(t) \in \mathcal{R}(2r, r)$. Then,

$$\|f_{\mathrm{hi}} \star (x - \hat{x})\|_1 \leq c \cdot \left(\frac{\lambda_c}{\lambda_{\mathrm{hi}}}\right)^{2r} \cdot \|z(t)\|_1.$$

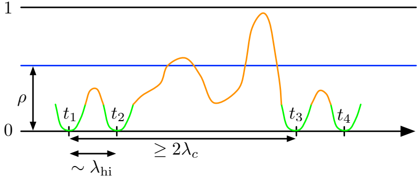
Need new tools

Theorem: (V. Morgenshtern, 2019, [5])

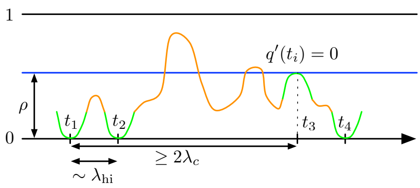
Assume $x(t) \geq 0$, $x(t) \in \mathcal{R}(2r, r)$. Then,

$$\|f_{\text{hi}} \star (x - \hat{x})\|_1 \leq c \cdot \left(\frac{\lambda_c}{\lambda_{\text{hi}}}\right)^{2r} \cdot \|z(t)\|_1.$$

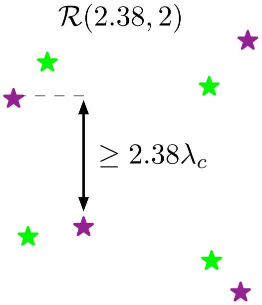
Can do: all zeros



Need: arbitrary pattern $\{0, +\rho\}$



2D Super-resolution



Theorem: (V. Morgenshtern and E. Candès, 2016, [1])

Take $\mathbf{P} = \mathbf{P}_{\mathrm{tri,2D}}$ or $\mathbf{P} = \mathbf{P}_{\mathrm{flat,2D}}$. Assume $\mathbf{x} \geq 0$, $\mathbf{x} \in \mathcal{R}(2.38r, r)$. Then,

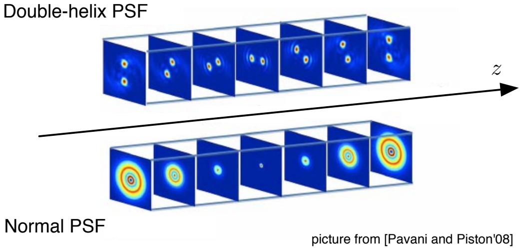
$$\|\hat{\mathbf{x}} - \mathbf{x}\|_1 \leq c \cdot \left(\frac{N}{2f_c}\right)^{2r} \delta.$$

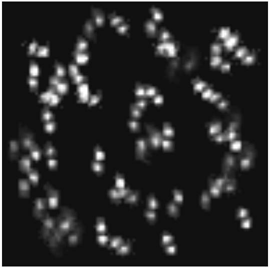
New: number of spikes is linear in the number of observations

Improving microscopes

Collaboration with Moerner Lab, C.A. Sing-Long, E. Candès

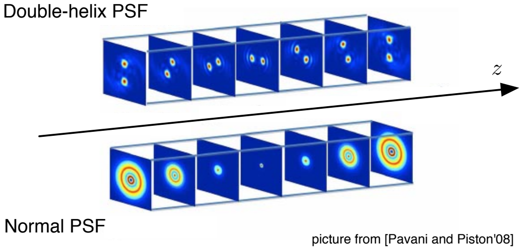
Reconstruction of 3D signals from 2D data



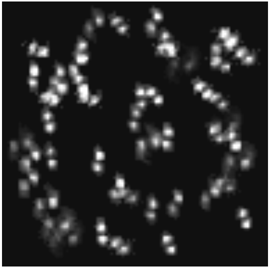


2D double-helix data

Reconstruction of 3D signals from 2D data



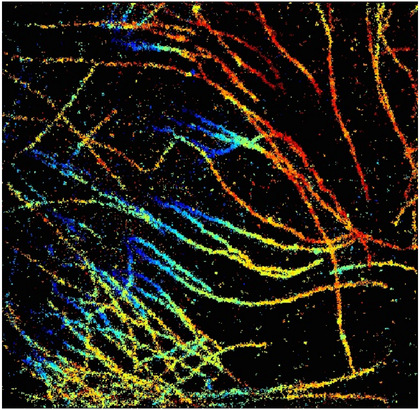
picture from [Pavani and Piston'08]



2D double-helix data

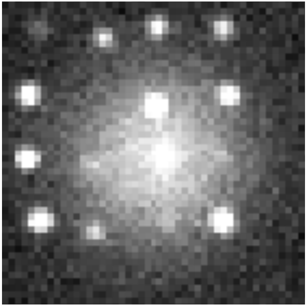
$$\begin{aligned} &\text{minimize} \quad \frac{1}{2}\|\mathbf{s} - \mathbf{P}\hat{\mathbf{x}}\|_2^2 + \lambda\sigma\|\operatorname{diag}(\mathbf{w})\hat{\mathbf{x}}\|_1 \\ &\text{subject to} \quad \hat{\mathbf{x}} \geq 0 \end{aligned}$$

Preliminary result: 4 times faster than state-of-the-art



10000 CVX problems solved
TFOCS first order solver
millions of variables

Flexible framework: smooth background separation



minimize	$\frac{1}{2}\|\mathbf{s} - \mathbf{P}(\hat{\mathbf{x}} + \mathbf{b})\|_2^2 + \lambda\sigma\|\hat{\mathbf{x}}\|_1$
subject to	$\hat{\mathbf{x}} \geq 0$
	$\mathbf{b}$ low freq. trig. polynomial (background)

Conclusion

Convex optimization is a near-optimal method
for super-resolution of positive sources

- Flexibility and good practical performance
- Non-asymptotic precise stability bounds
- Rayleigh-regularity is fundamental: separation between spikes is only one part of the picture

Backup slides

Connection to Bernstein theorem

Consider: $q(t) = \sum_{k=-f_c}^{f_c} \hat{q}_k e^{-\mathrm{i}2\pi kt}$ with $\|q\|_\infty \leq 1$
Then: $\|q'\|_\infty \leq 2f_c$

Connection to Bernstein theorem

Consider: $q(t) = \sum_{k=-f_c}^{f_c} \hat{q}_k e^{-i2\pi kt}$ with $\|q\|_\infty \leq 1$
Then: $\|q'\|_\infty \leq 2f_c$

"Curvy" $q(t)$ has best possible curvature!

Connection to Bernstein theorem

Consider: $q(t) = \sum_{k=-f_c}^{f_c} \hat{q}_k e^{-i2\pi kt}$ with $\|q\|_\infty \leq 1$
Then: $\|q'\|_\infty \leq 2f_c$

"Curvy" $q(t)$ has best possible curvature!

Since

$$q(t_i) = 0$$
$$q'(t_i) = 0$$
$$\|q\|_\infty \leq 1$$

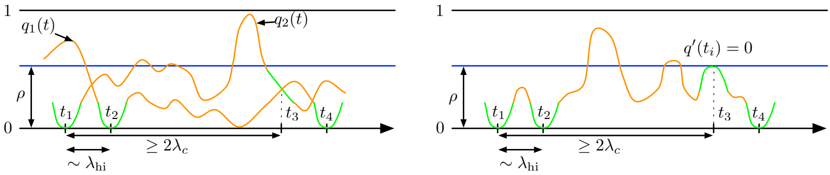
We conclude:

$$\begin{aligned}\|q'\|_\infty \leq 2f_c &\Rightarrow \|q''\|_\infty \leq (2f_c)^2 \\ &\Rightarrow q(t-t_i) \leq (2f_c)^2(t-t_i)^2 \\ &\Rightarrow q(t_i + 1/N) \leq \frac{(2f_c)^2}{N^2} = \frac{1}{\mathrm{SRF}^2}\end{aligned}$$

New tools

1. Control behavior on separated set
2. **Multiply**

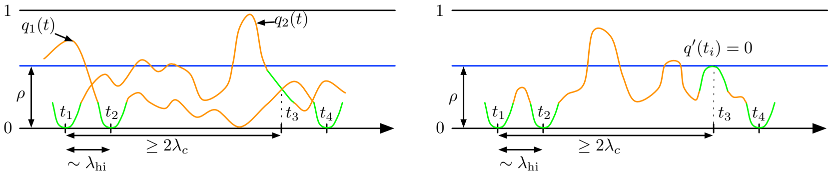
$$q(t) = q_1(t) \times q_2(t)$$
$$0 = q'(t_3) = q_1'(t_3)q_2(t_3) + q_1(t_3)q_2'(t_3)$$



New tools

1. Control behavior on separated set
2. **Multiply**

$$q(t) = q_1(t) \times q_2(t)$$
$$0 = q'(t_3) = q_1'(t_3)q_2(t_3) + q_1(t_3)q_2'(t_3)$$



3. **Sum**

$$q(t) = \sum_{r} \prod_{k=1}^{r} \sum_{t_{jk} \in \mathcal{T}_k} \underbrace{a_{jk} K(t - t_{jk})}_{\text{frequency } f_c/r}$$

V. I. Morgenshtern and E. J. Candès, "Super-resolution of positive sources: the discrete setup," *SIAM J. Imaging Sci.*, vol. 9, pp. 412—444, Mar. 2016.

D. L. Donoho, I. M. Johnstone, J. C. Hoch, and A. S. Stern, "Maximum entropy and the nearly black object," *J. Roy. Statist. Soc. Ser. B*, vol. 54, pp. 41–81, June 1992.

J.-J. Fuchs, "Sparsity and uniqueness for some specific under-determined linear systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 5, pp. v/729–v/732, 2005.

E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Commun. Pure Appl. Math.*, vol. 67, pp. 906–956, June 2014.

V. I. Morgenshtern, "Super-resolution of positive sources: the continuous setup," *J. of Fourier Analysis and Appl.*, 2019.
To be submitted.