A Sampling Theorem for Robust Deconvolution

Carlos Fernandez-Granda

www.cims.nyu.edu/~cfgranda

Institut Poincaré

6/20/2017

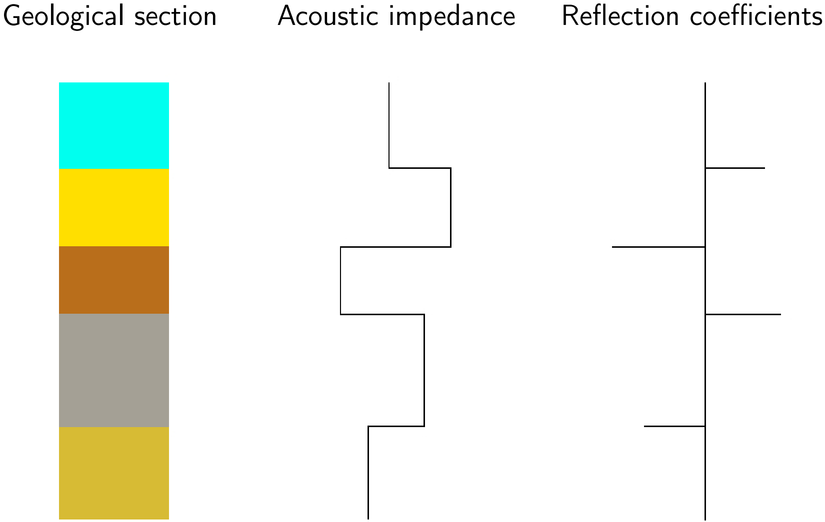
Acknowledgements

Joint work with Brett Bernstein and Emmanuel Candès

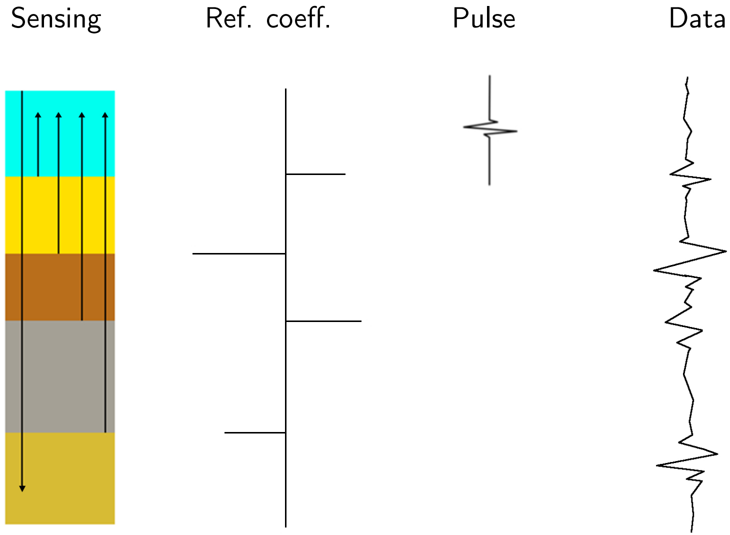
Project funded by NSF award DMS-1616340

Seismology

Reflection seismology

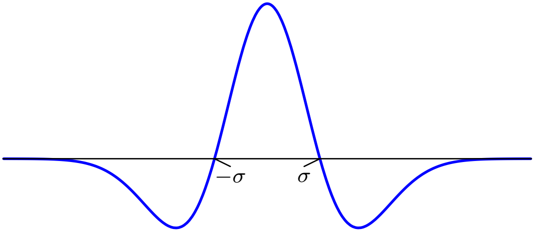


Reflection seismology

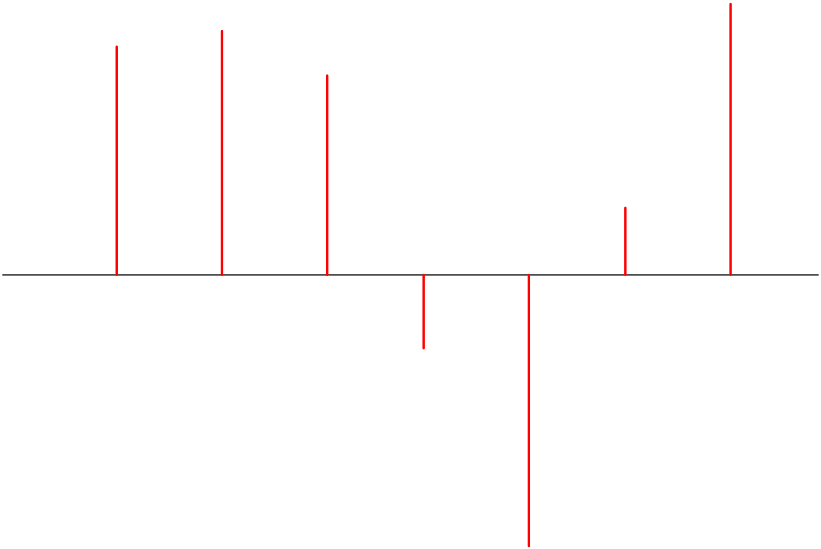


Data ≈ convolution of pulse and reflection coefficients

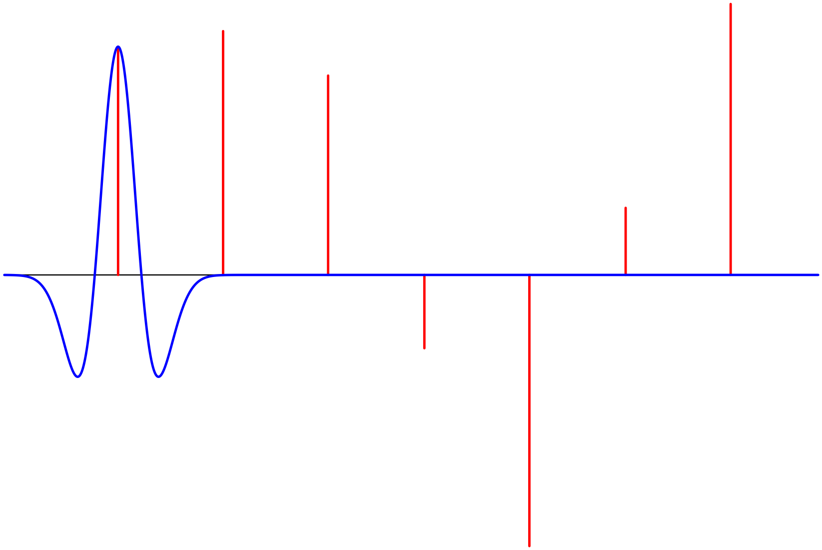
Model for the pulse: Ricker wavelet



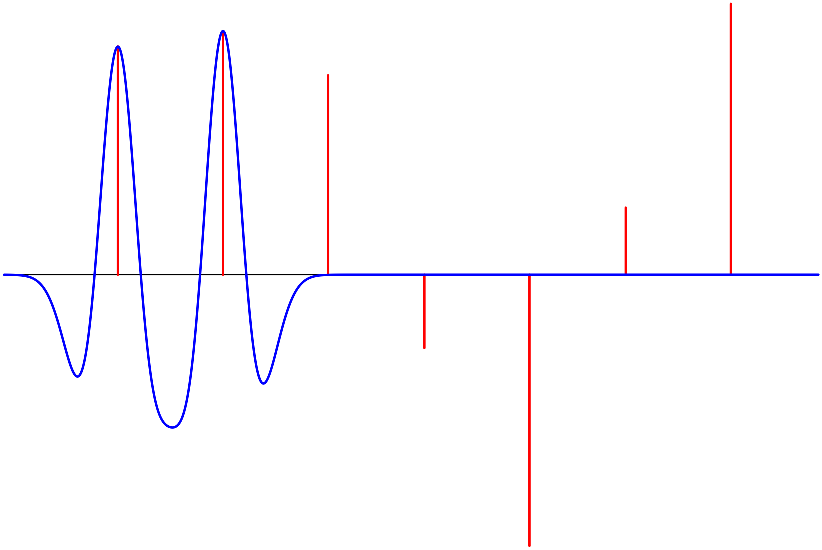
Sensing model for reflection seismology



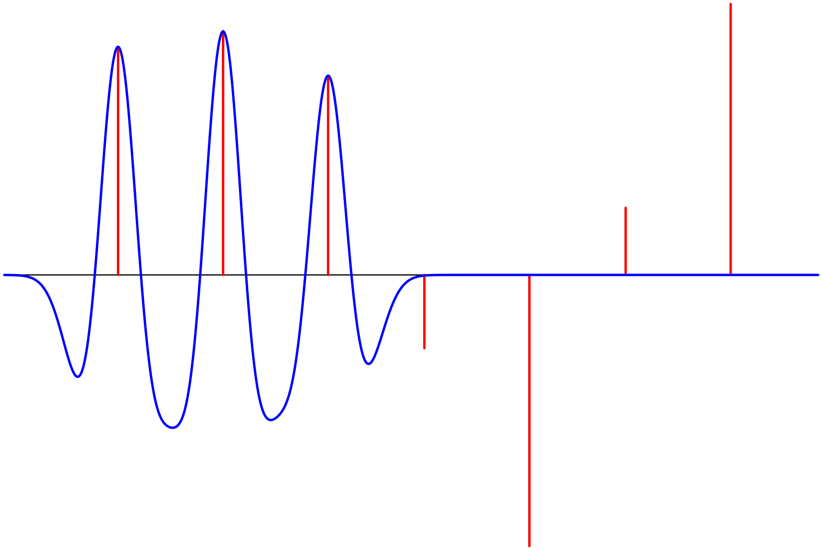
Sensing model for reflection seismology



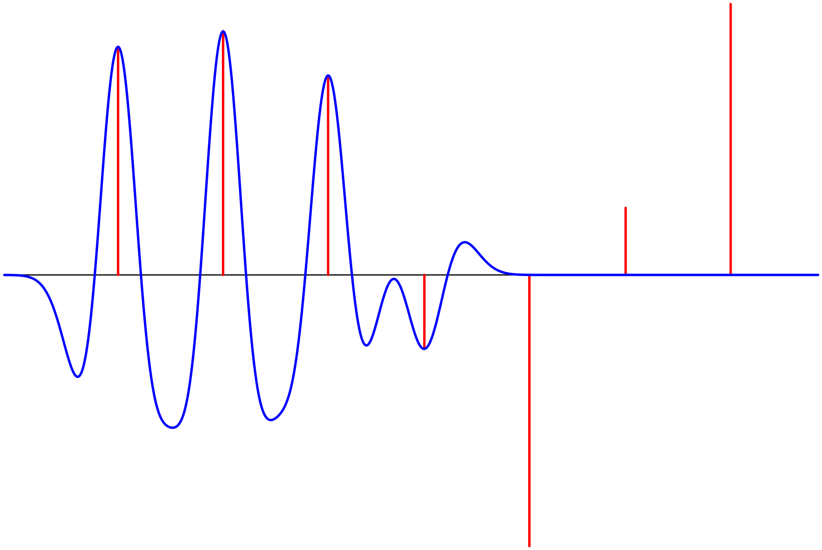
Sensing model for reflection seismology



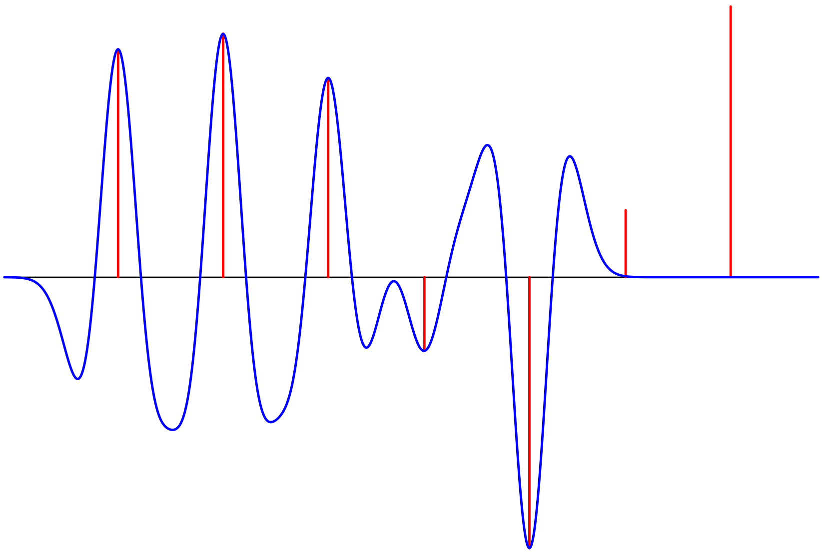
Sensing model for reflection seismology



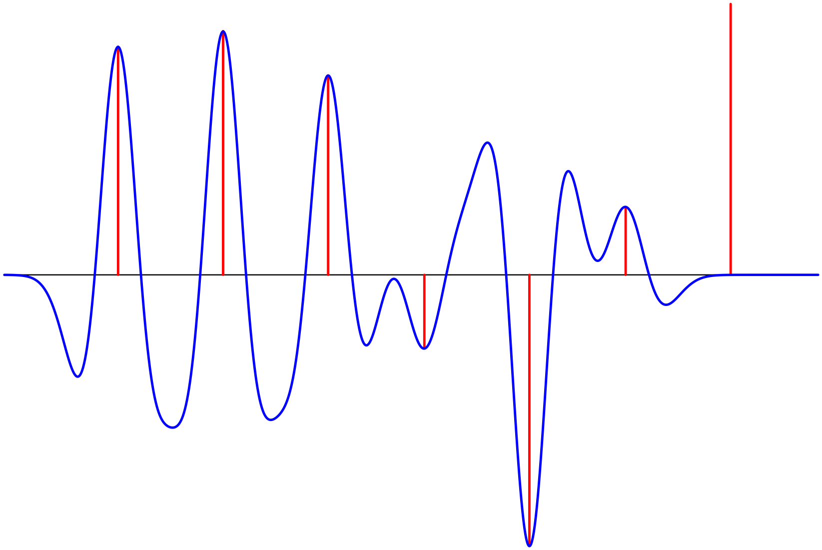
Sensing model for reflection seismology



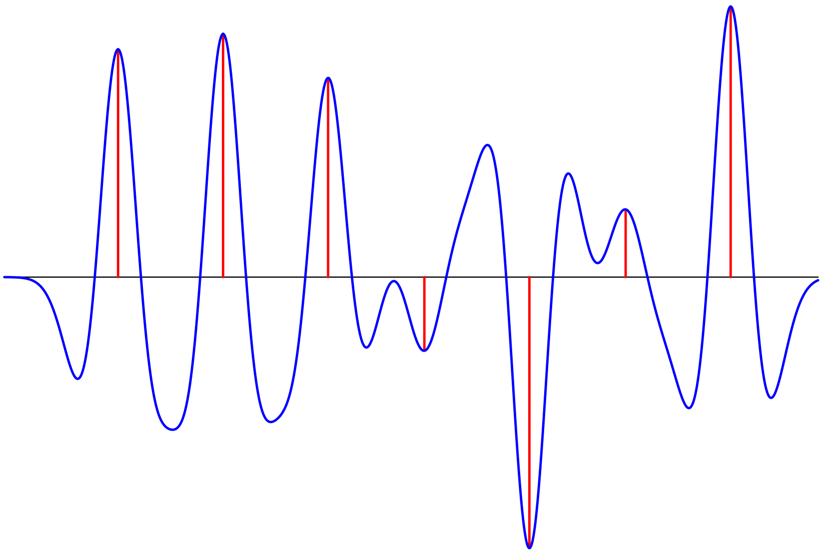
Sensing model for reflection seismology



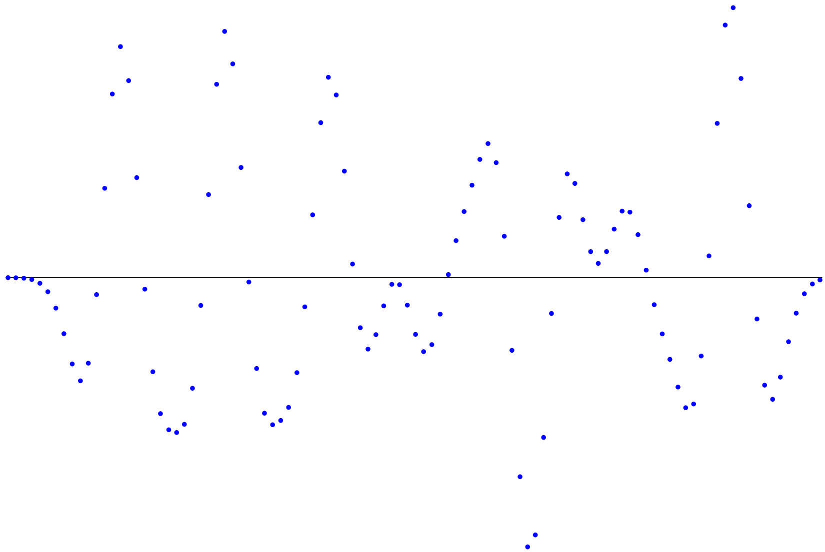
Sensing model for reflection seismology



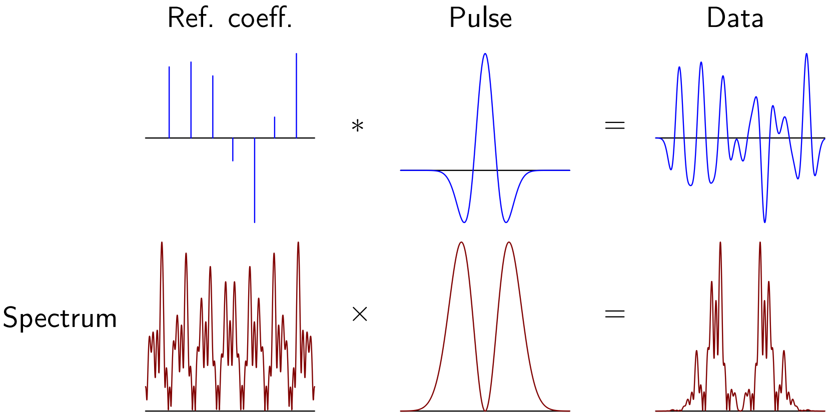
Sensing model for reflection seismology



Sensing model for reflection seismology



Sensing model for reflection seismology

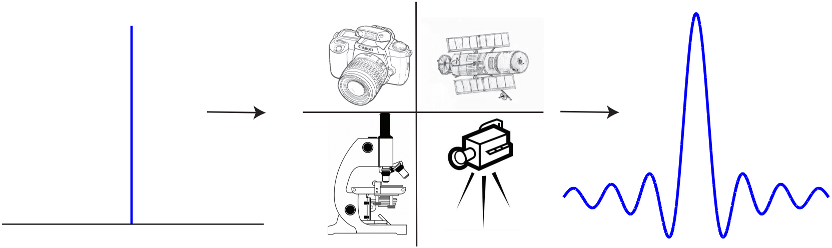


Convolution in time = Pointwise multiplication in frequency

Ill-posed problem! How do we choose between signals consistent with data?

Limits of resolution in imaging

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

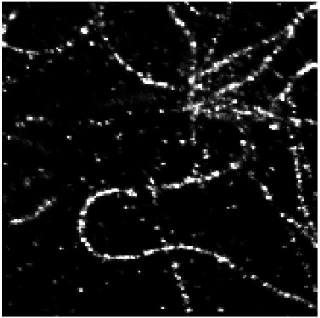


$\delta(t-\tau)$ optical system $h(t-\tau)$

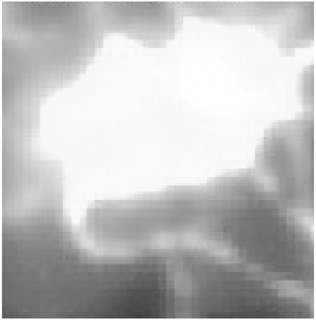
Diffraction imposes a fundamental limit on the resolution of optical systems

Fluorescence microscopy

Data



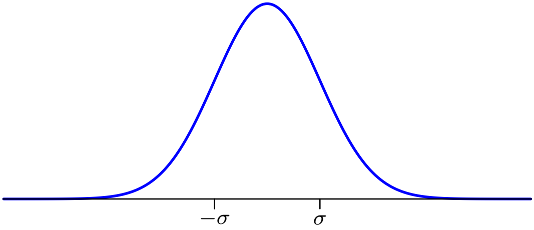
Point sources



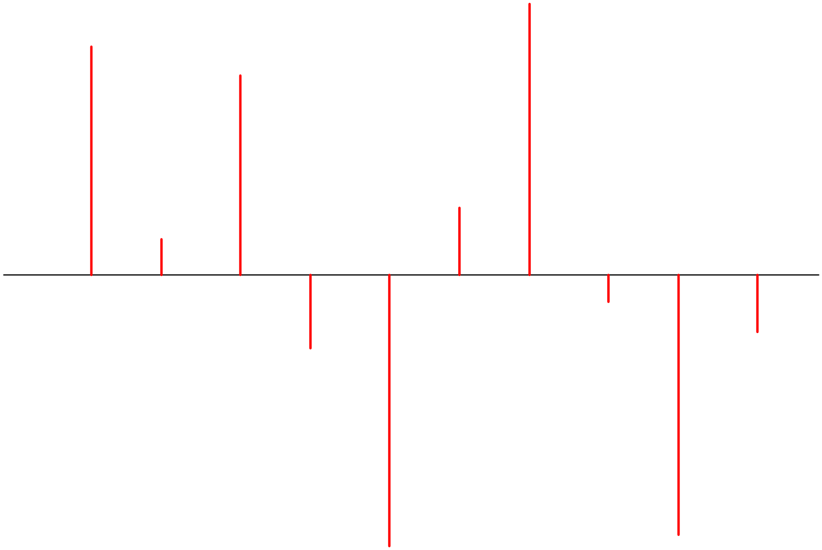
Low-pass blur

(Figures courtesy of V. Morgenshtern)

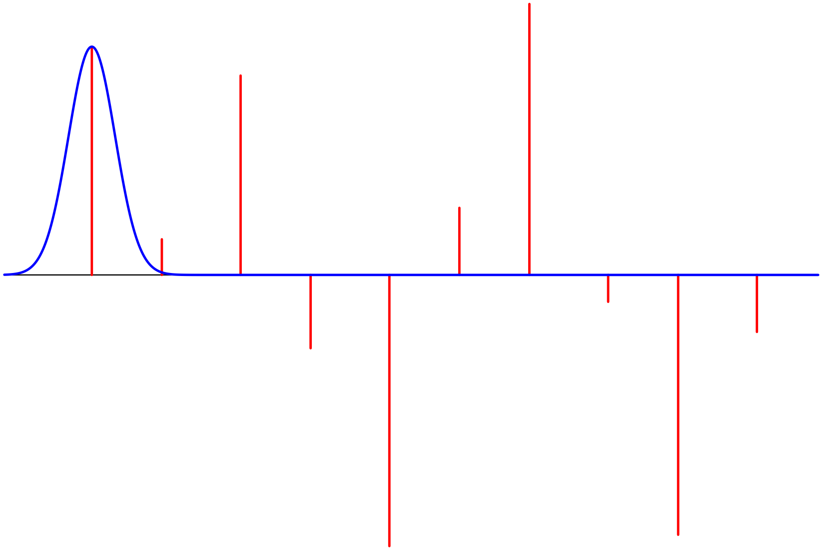
Model for the point-spread function: Gaussian kernel



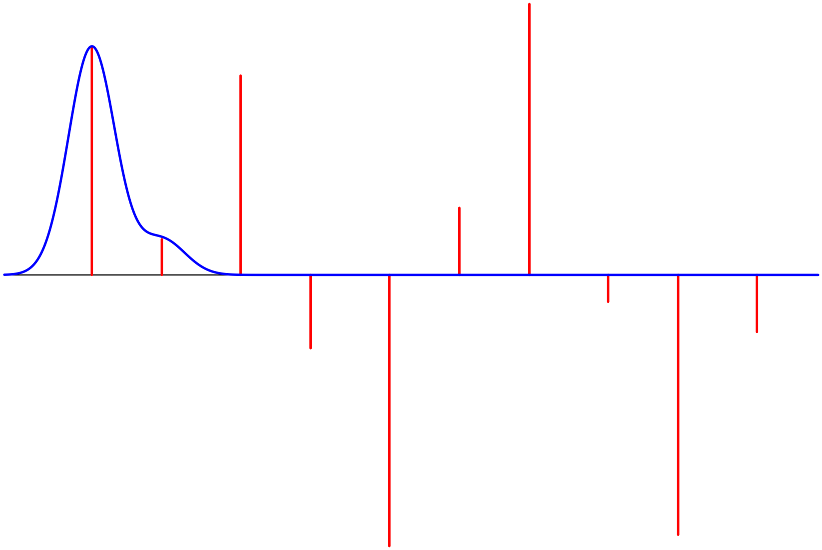
Sensing model for diffraction-limited imaging



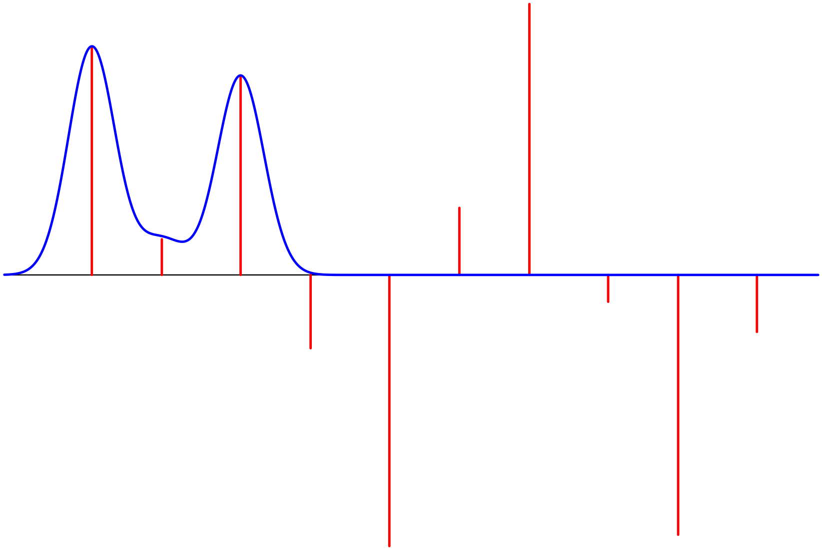
Sensing model for diffraction-limited imaging



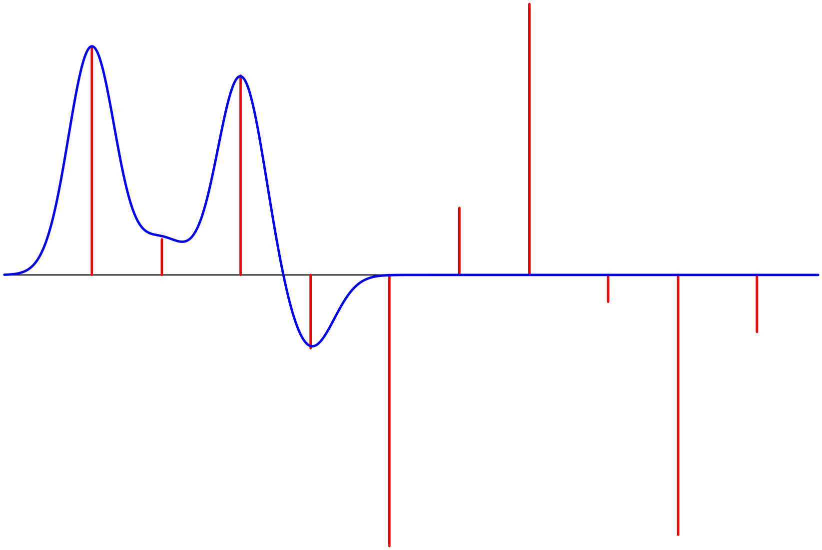
Sensing model for diffraction-limited imaging



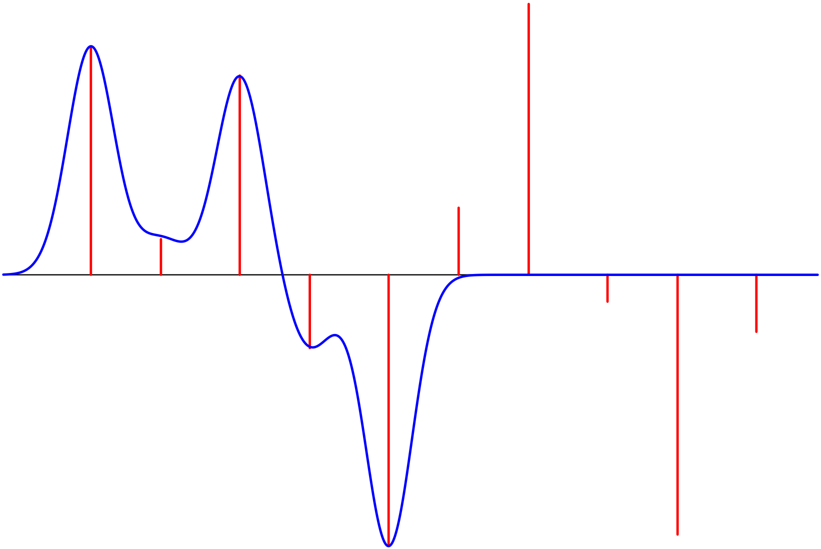
Sensing model for diffraction-limited imaging



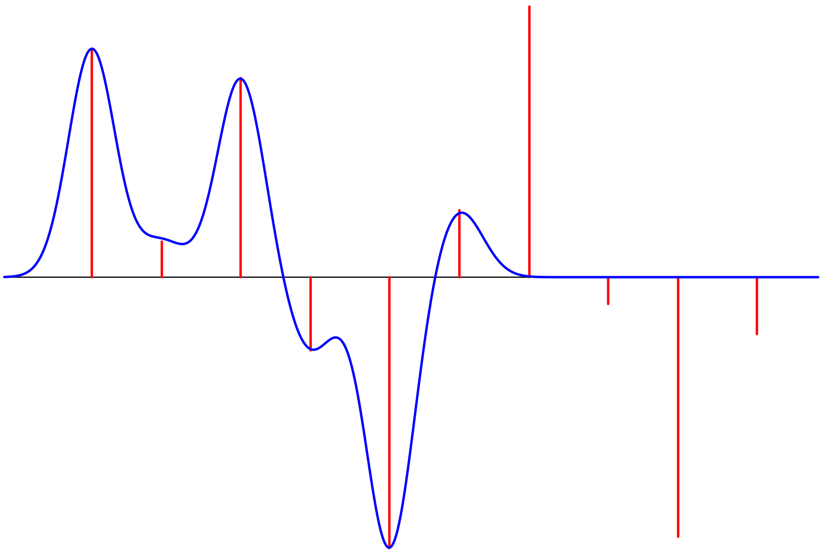
Sensing model for diffraction-limited imaging



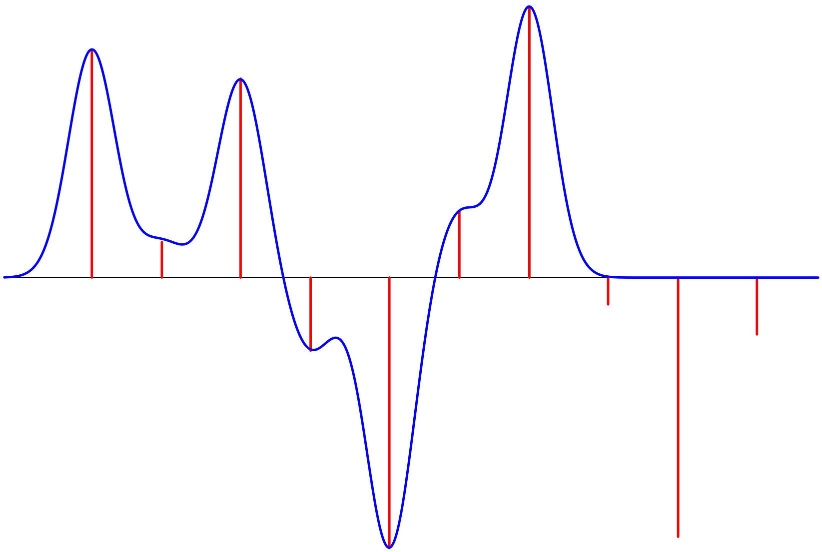
Sensing model for diffraction-limited imaging



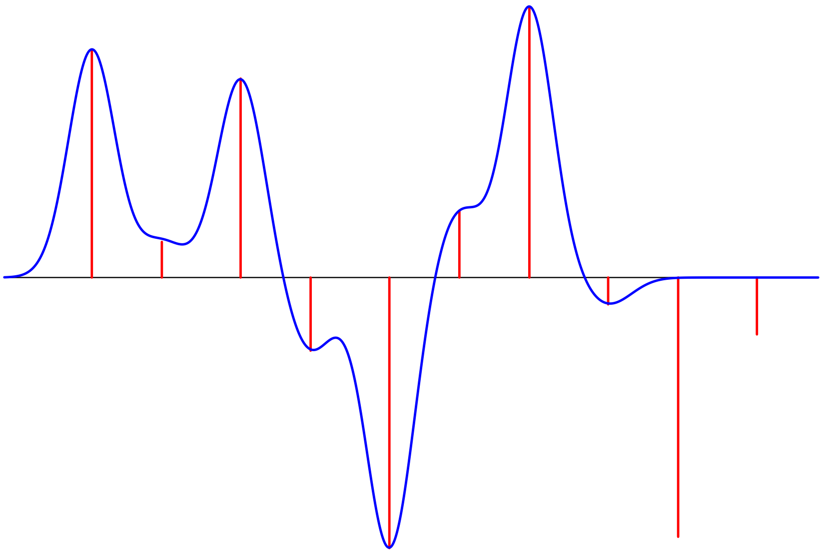
Sensing model for diffraction-limited imaging



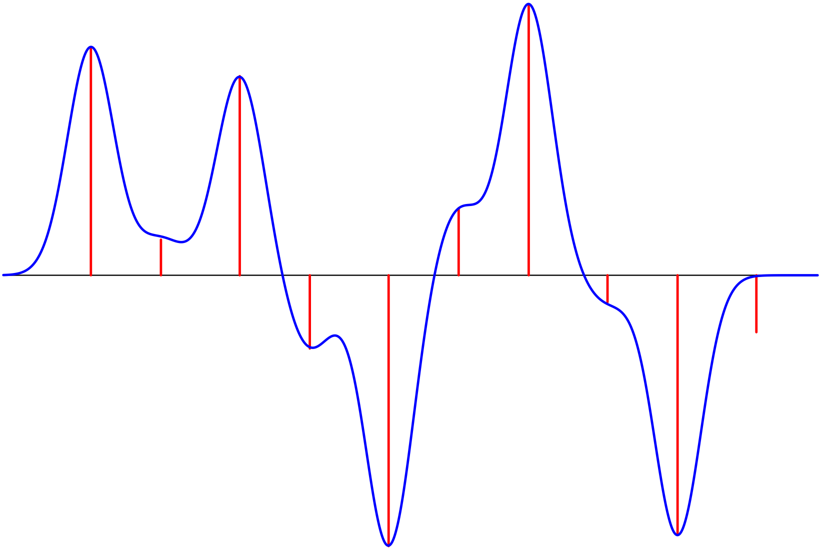
Sensing model for diffraction-limited imaging



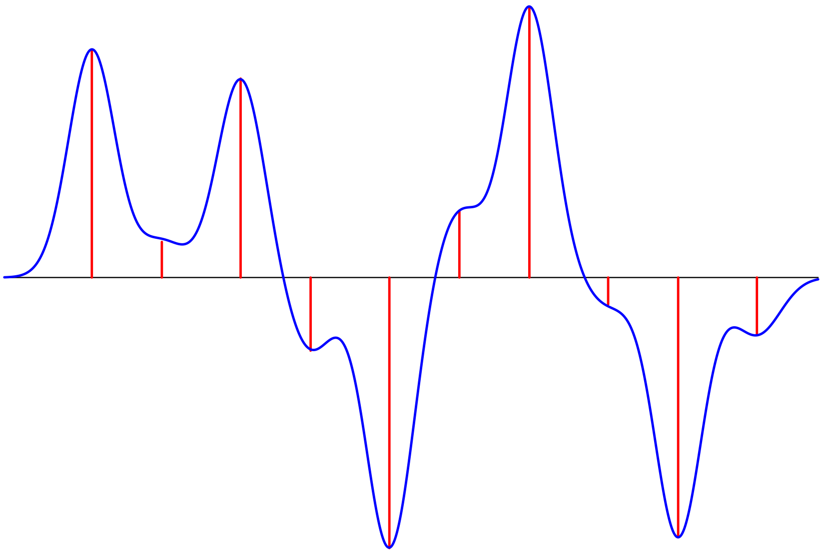
Sensing model for diffraction-limited imaging



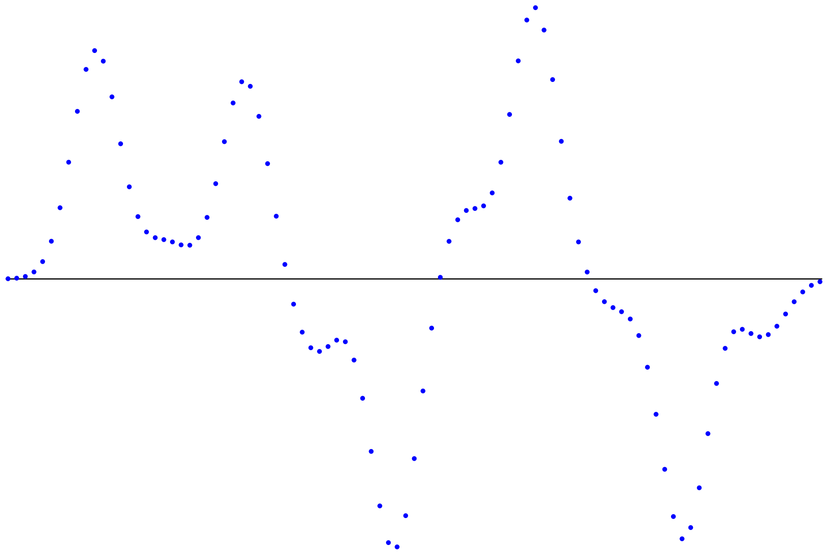
Sensing model for diffraction-limited imaging



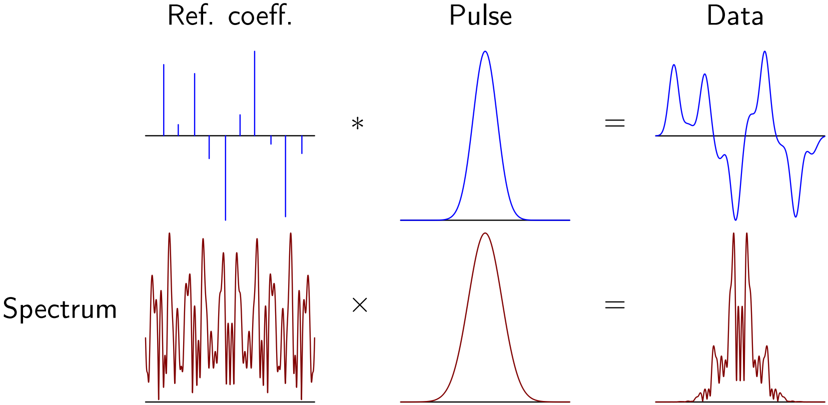
Sensing model for diffraction-limited imaging



Sensing model for diffraction-limited imaging



Sensing model for diffraction-limited imaging



Convolution in time = Pointwise multiplication in frequency

Ill-posed problem! How do we choose between signals consistent with data?

Geophysicists: Minimize ℓ_1 norm

Deconvolution with the ℓ_1 norm

Howard L. Taylor,* Stephen C. Banks,‡ and John F. McCoy§

GEOPHYSICS, VOL. 44, NO. 1 (JANUARY 1979)

LINEAR INVERSION OF BAND-LIMITED REFLECTION SEISMOGRAMS*

FADIL SANTOSA† AND WILLIAM W. SYMES‡

SIAM J. SCI. STAT. COMPUT.
Vol. 7, No. 4, October 1986

Reconstruction of a sparse spike train from a portion of its spectrum and application to high-resolution deconvolution

Shlomo Levy* and Peter K. Fullagar‡

GEOPHYSICS, VOL. 46, NO. 9 (SEPTEMBER 1981)

ROBUST MODELING WITH ERRATIC DATA†

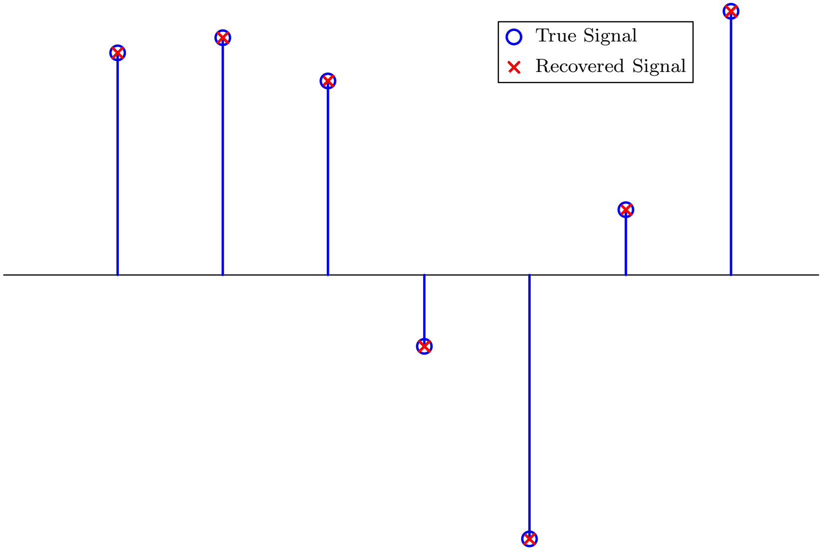
JON F. CLAERBOUT* AND FRANCIS MUIR‡

GEOPHYSICS, VOL. 38, NO. 5 (OCTOBER 1973)

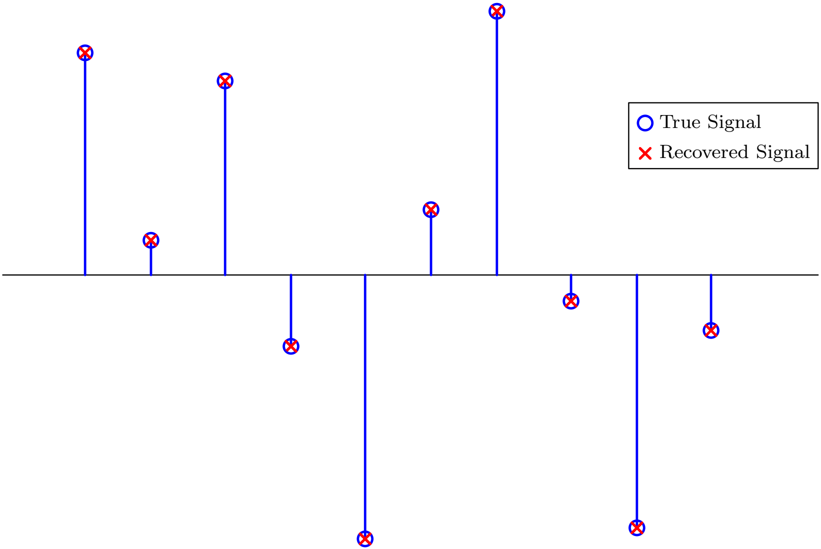
ℓ_1-norm minimization

minimize	$\|\|\text{estimate}\|\|_1$
subject to	samples of convolution with kernel $=$ data

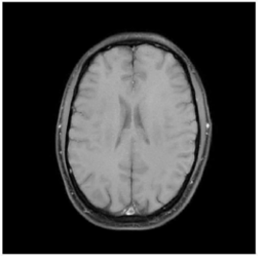
It works

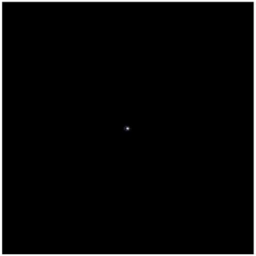


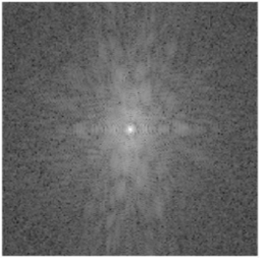
It works



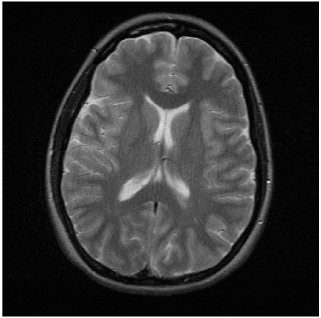
Magnetic resonance imaging



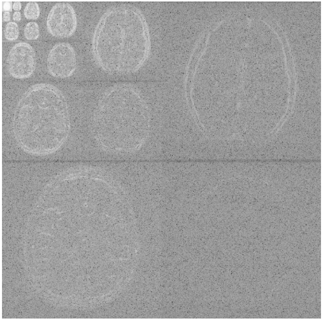




Images are sparse/compressible



Wavelet coefficients



Magnetic resonance imaging

Data: Samples from spectrum

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

Images are compressible ($\approx$ sparse)

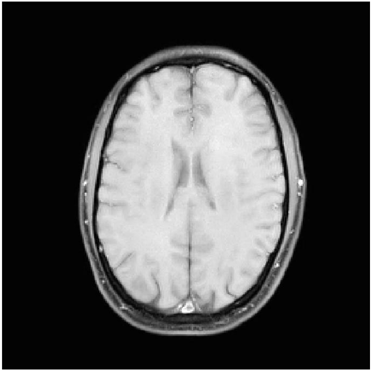
Can we recover compressible signals from less data?

Compressed sensing

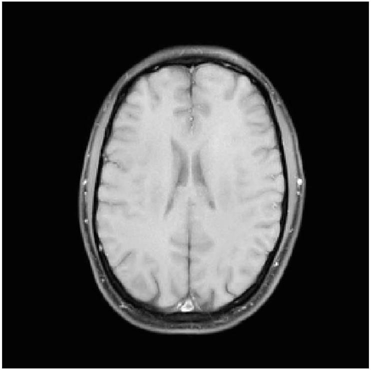
1. Undersample data randomly
2. Solve the optimization problem

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

Compressed sensing in MRI



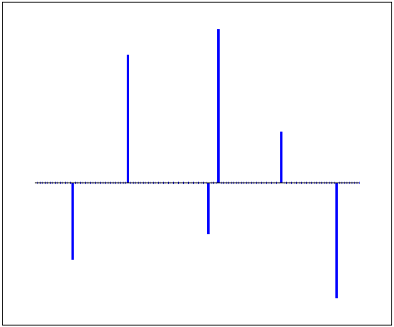
x2 Undersampling



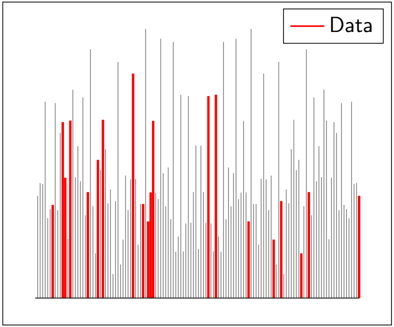
Compressed sensing (basic model)

1. Undersample the spectrum randomly

Signal



Spectrum



Compressed sensing (basic model)

2. Solve the optimization problem

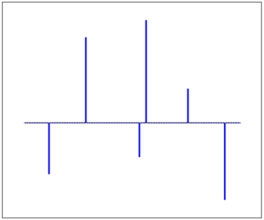
$$\begin{aligned} &\textit{minimize} && ||\text{estimate}||_1 \\ &\textit{subject to} && \text{frequency samples of estimate} = \text{data} \end{aligned}$$

Compressed sensing (basic model)

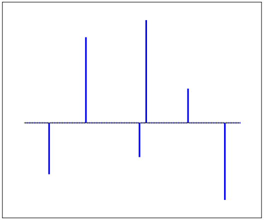
2. Solve the optimization problem

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

Signal



Estimate

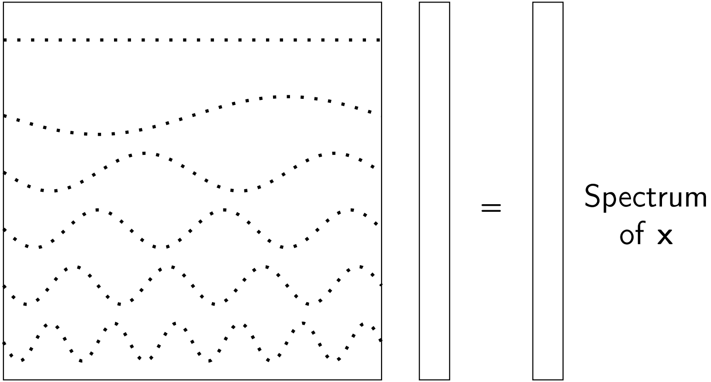


Theoretical questions

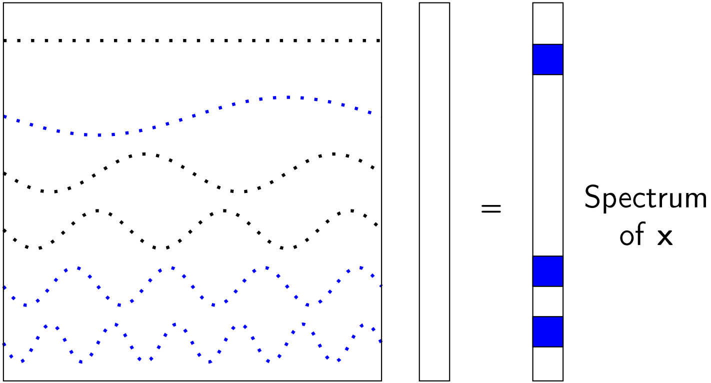
1. Is the problem well posed?

2. Does ℓ_1-norm minimization work?

Is the problem well posed?

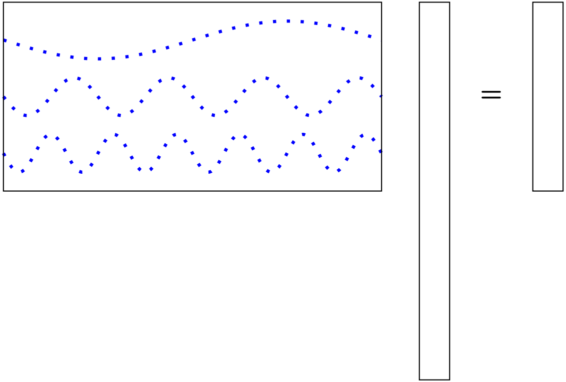


Is the problem well posed?

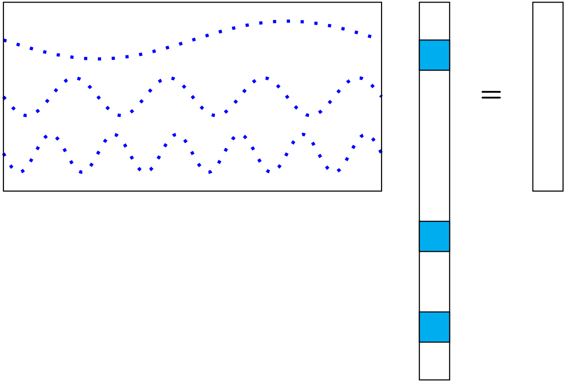


Measurement operator = random frequency samples

Is the problem well posed?

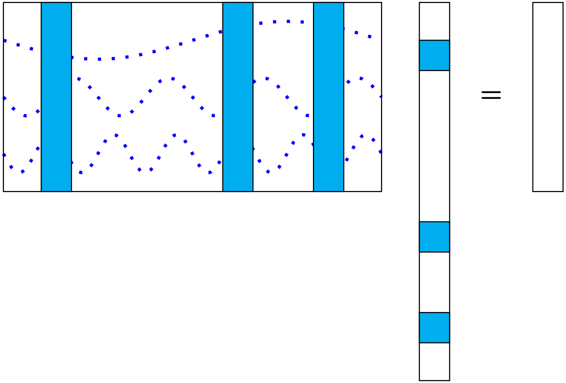


Is the problem well posed?



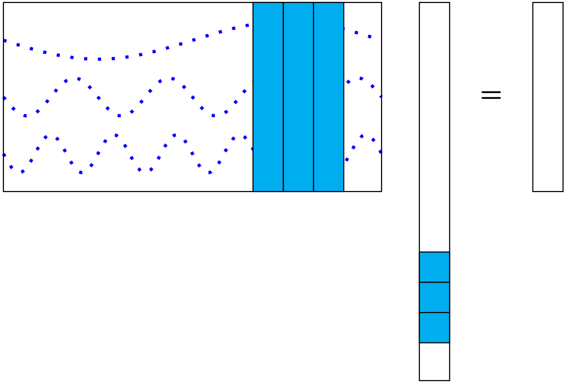
What is the effect of the measurement operator on sparse vectors?

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 $\mathbf{x}$

$$(1 - \delta) \, ||x||_2 \leq ||Ax||_2 \leq (1 + \delta) \, ||x||_2$$

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

2s-RIP implies that for any s-sparse signals x_1, x_2

$$||Ax_2 - Ax_1||_2$$

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 $\mathbf{x}$

$$(1-\delta)\,||x||_2 \leq ||Ax||_2 \leq (1+\delta)\,||x||_2$$

> 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 x_1, x_2

$$||Ax_2 - Ax_1||_2 = ||A\,(x_2 - x_1)||_2$$

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 $\mathbf{x}$

$$(1-\delta)\,||x||_2 \leq ||Ax||_2 \leq (1+\delta)\,||x||_2$$

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

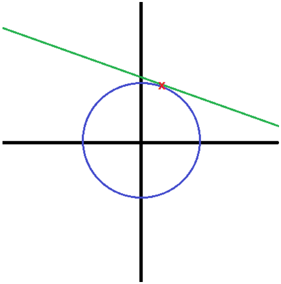
$2s$-RIP implies that for any s-sparse signals x_1, x_2

$$\begin{aligned} ||Ax_2 - Ax_1||_2 &= ||A\,(x_2 - x_1)||_2 \\ &\geq (1-\delta)\,||x_2 - x_1||_2 \end{aligned}$$

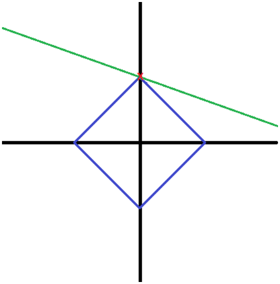
Theoretical questions

1. Is the problem well posed?
2. Does ℓ_1-norm minimization work?

Geometric intuition



$$\begin{array}{ll} \textit{minimize} & \|x'\|_2 \\ \textit{subject to} & Ax' = y \end{array}$$



$$\begin{array}{ll} \textit{minimize} & \|x'\|_1 \\ \textit{subject to} & Ax' = y \end{array}$$

Characterizing the minimum ℓ_1-norm estimate

- Aim: Show that the original signal x is the solution of

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

- This is guaranteed by the existence of a dual certificate

Dual certificate

$v \in \mathbb{R}^m$ is a dual certificate associated to x if

$$q := A^T v$$

satisfies

$$\begin{aligned} q_i &= \text{sign}(x_i) && \text{if } x_i \neq 0 \\ |q_i| &< 1 && \text{if } x_i = 0 \end{aligned}$$

q is a subgradient of the ℓ_1 norm at x

For any vector u

$$||x + u||_1 \geq ||x||_1 + q^T u$$

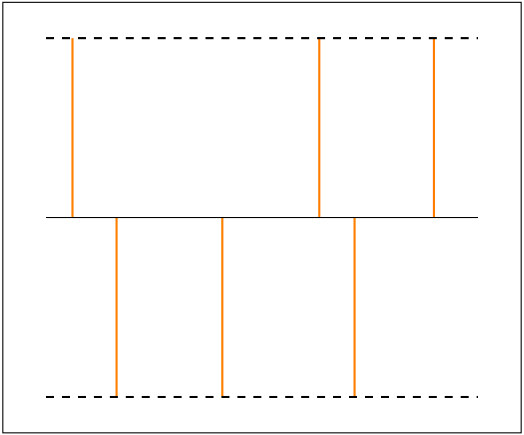
Dual certificate

For any $x + h$ such that $Ah = 0$

$$
\begin{aligned}
||x + h||_1 &\geq ||x||_1 + q^T h && (q \text{ is a subgradient}) \\
&= ||x||_1 + v^T A h && (q = A^T v) \\
&= ||x||_1
\end{aligned}
$$

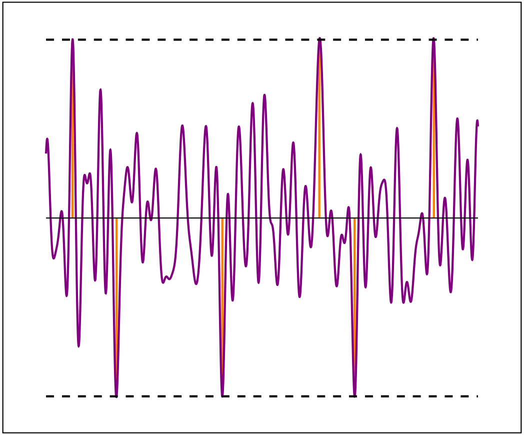
If A_T (where T is the support of x) is injective, x is the unique solution

Dual certificate for compressed sensing



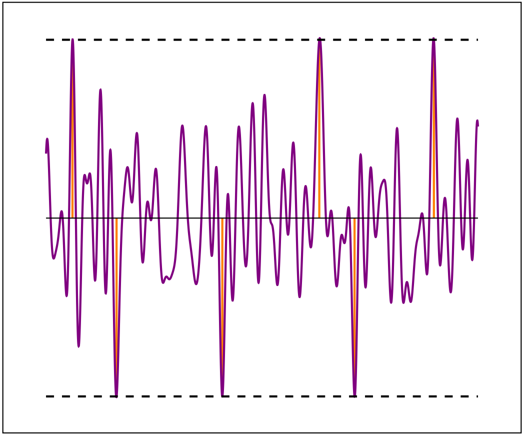
Aim: Show that a dual certificate exists for *any* sparse support and sign pattern

Certificate for compressed sensing



Idea: Minimum-energy interpolator has closed-form solution

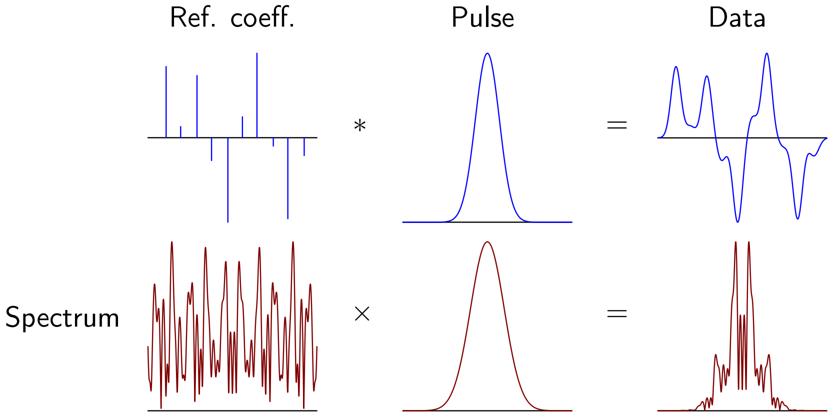
Certificate for compressed sensing



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

(Candès, Romberg, Tao 2006)

Deconvolution in the frequency domain



If kernel is exactly low pass and we have uniform samples at Nyquist rate, equivalent to super-resolution from low-pass data

Mathematical model

- Signal: superposition of Dirac measures with support T

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

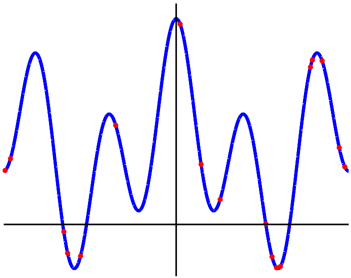
- Data: low-pass Fourier coefficients with cut-off frequency f_c

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

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

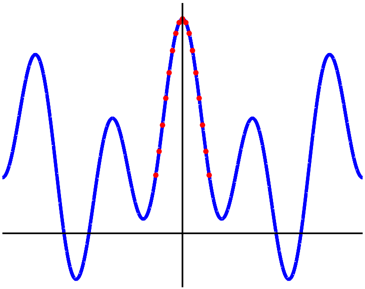
Compressed sensing vs super-resolution

Compressed sensing



spectrum interpolation

Super-resolution



spectrum extrapolation

Total-variation norm

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

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

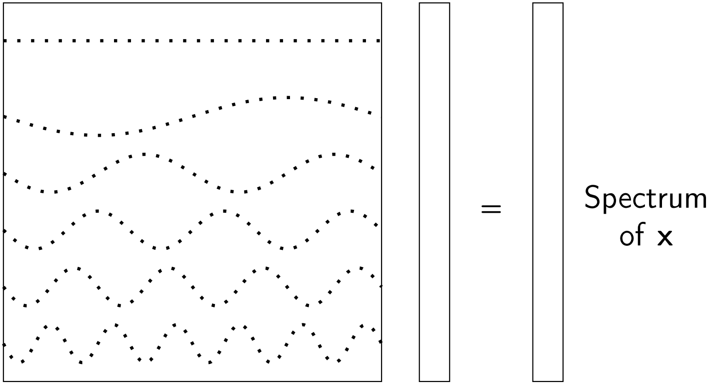
(supremum over all finite partitions B_j of $[0, 1]$)

Theoretical questions

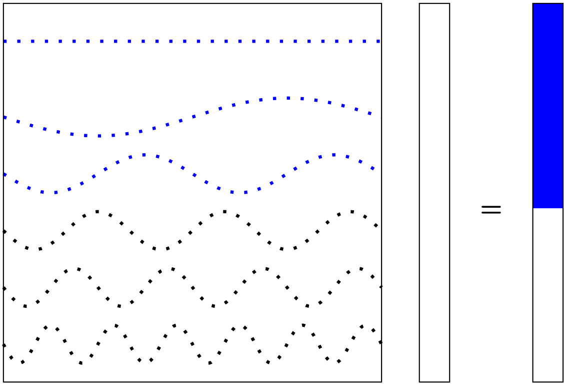
1. Is the problem well posed?

2. Does TV-norm minimization work?

Is the problem well posed?

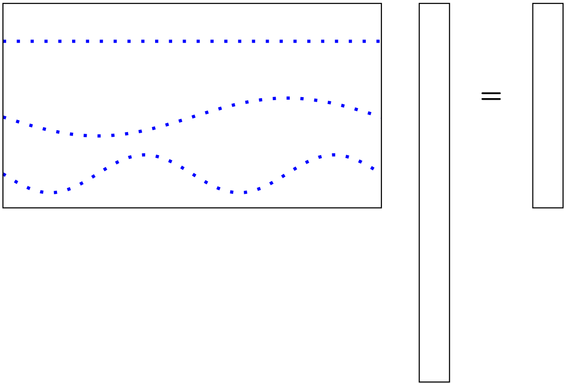


Is the problem well posed?



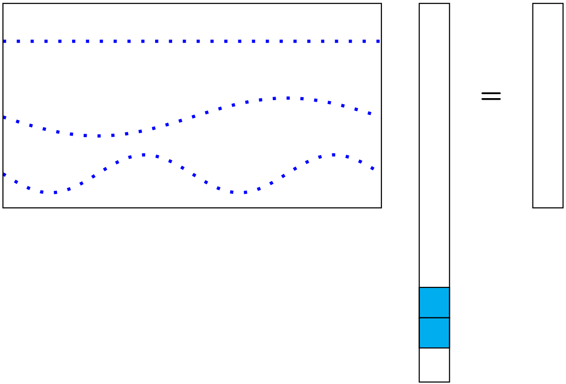
Measurement operator = low-pass samples with cut-off frequency f_c

Is the problem well posed?



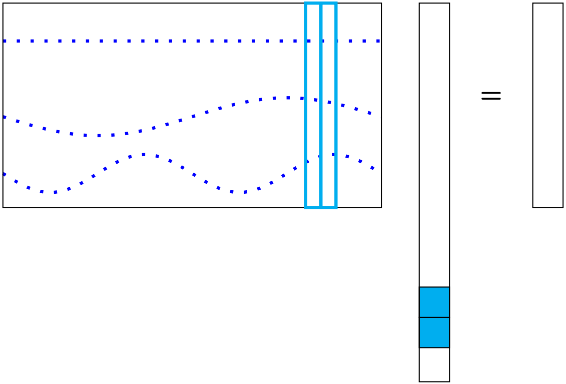
Measurement operator = low-pass samples with cut-off frequency f_c

Is the problem well posed?



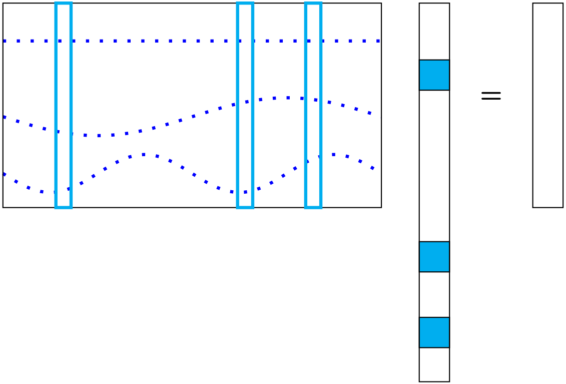
Effect of measurement operator on sparse vectors?

Is the problem well posed?



Submatrix can be very ill conditioned!

Is the problem well posed?

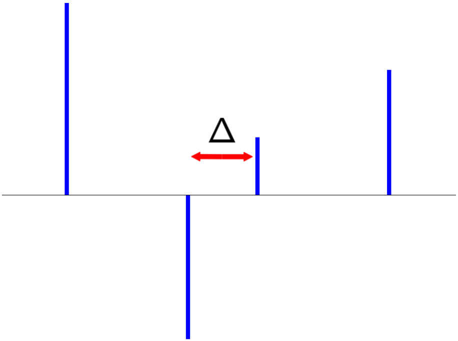


If support is spread out there is hope

Minimum separation

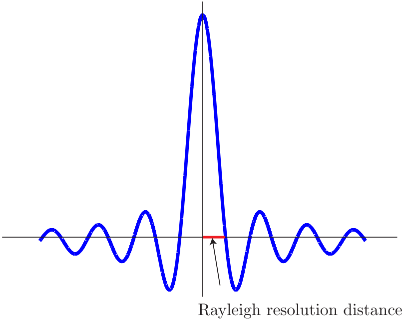
The minimum separation Δ of the support of x is

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



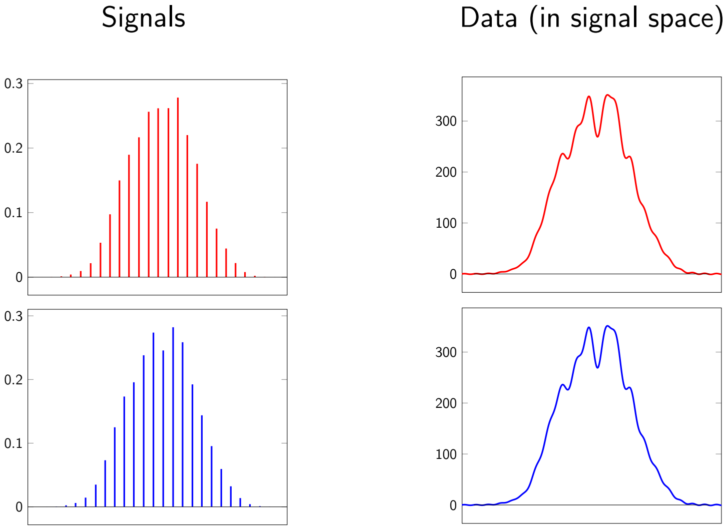
Conditioning of submatrix with respect to Δ

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

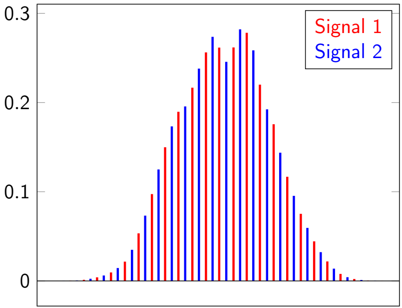


$1/f_c$ is the diameter of the main lobe of the point-spread function (twice the Rayleigh distance)

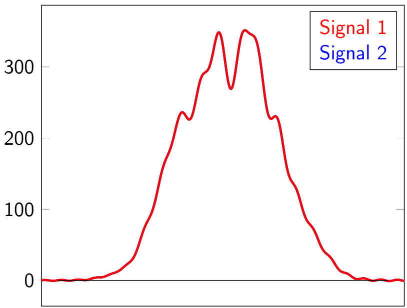
Example: 25 spikes, $f_c = 10^3$, $\Delta = 0.8/f_c$



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



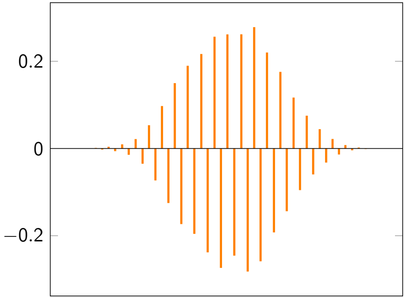
Signals



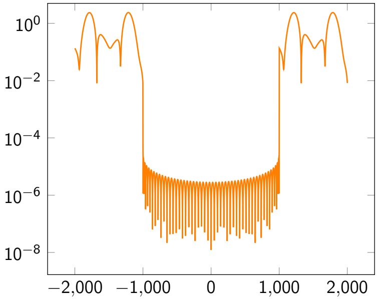
Data (in signal space)

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

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



Difference



Spectrum

Theoretical questions

1. Is the problem well posed?

2. Does TV-norm minimization work?

Estimation via convex programming

For data of the form $y = \mathcal{F}_c \, x$, we solve

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

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

Dual certificate

A dual certificate of the TV norm at

$$x = \sum_j a_j \delta_{t_j} \qquad a_j \in \mathbb{C},\ t_j \in T$$

guarantees that x is the unique solution if

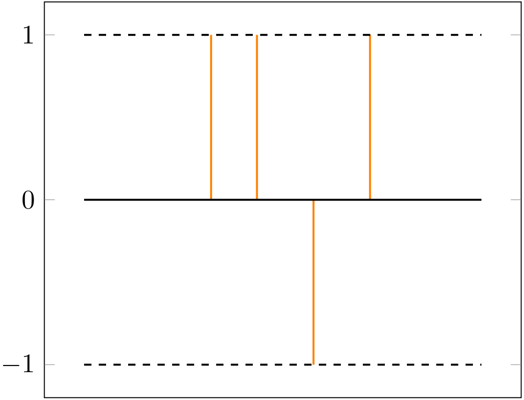
$$q := \mathcal{F}_c^* \, v = \sum_{k \leq |f_c|} v_k e^{i2\pi kt}$$

$$q(t_j) = \operatorname{sign}(a_j) \qquad \text{if } t_j \in T$$

$$|q(t)| < 1 \qquad \text{if } t \notin T$$

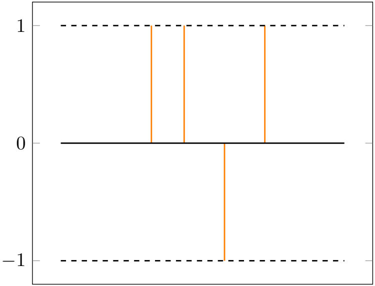
Range of $\mathcal{F}_c^*$ is spanned by low pass sinusoids instead of random sinusoids

Certificate for super-resolution



Aim: Interpolate sign pattern

Certificate for super-resolution



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

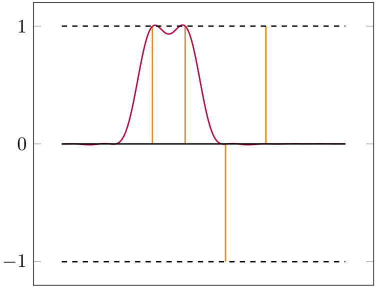
$$q(t) = \sum_{t_j \in T} \alpha_j \, F\,(t - t_j)$$

Certificate for super-resolution

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

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

Certificate for super-resolution



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

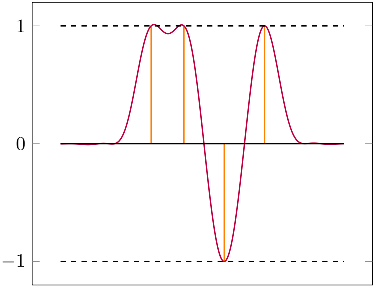
$$q(t) = \sum_{t_j \in T} \alpha_j \, F(t - t_j)$$

Certificate for super-resolution

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

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

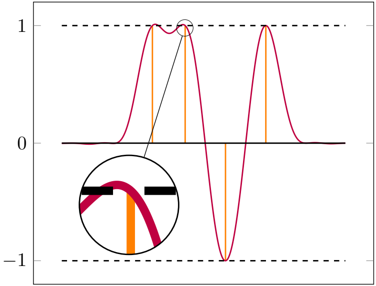
Certificate for super-resolution



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

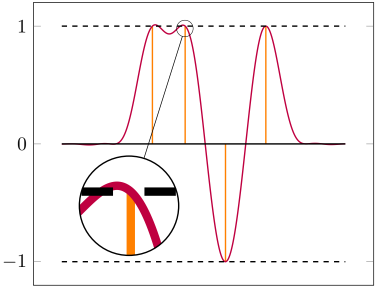
$$q(t) = \sum_{t_j \in T} \alpha_j \, F(t - t_j)$$

Certificate for super-resolution



Problem: Magnitude of certificate locally exceeds 1

Certificate for super-resolution

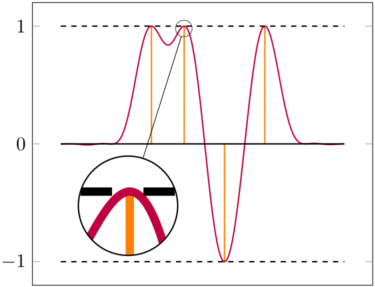


Problem: Magnitude of certificate locally exceeds 1

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

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

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

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

Guarantees for super-resolution

Theorem [Candès, F. 2012]

If the minimum separation of the signal support obeys

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

then recovery via convex programming is exact

Theorem [Candès, F. 2012]

In 2D convex programming super-resolves point sources with a minimum separation of

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

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

Guarantees for super-resolution

Theorem [F. 2016]

If the minimum separation of the signal support obeys

$$\Delta \geq 1.26\,/f_c,$$

then recovery via convex programming is exact

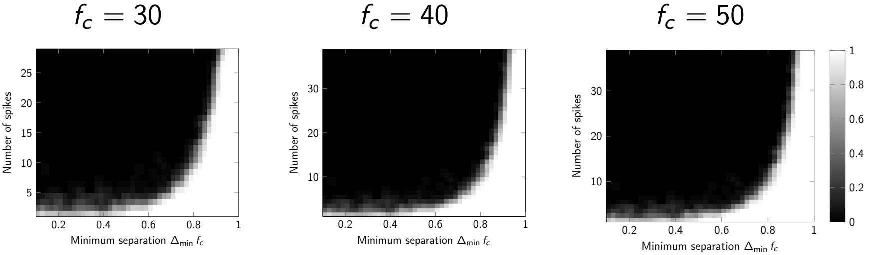
Theorem [Candès, F. 2012]

In 2D convex programming super-resolves point sources with a minimum separation of

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

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

Numerical evaluation of minimum separation



Numerically TV-norm minimization succeeds if $\Delta \geq \frac{1}{f_c}$

Motivation

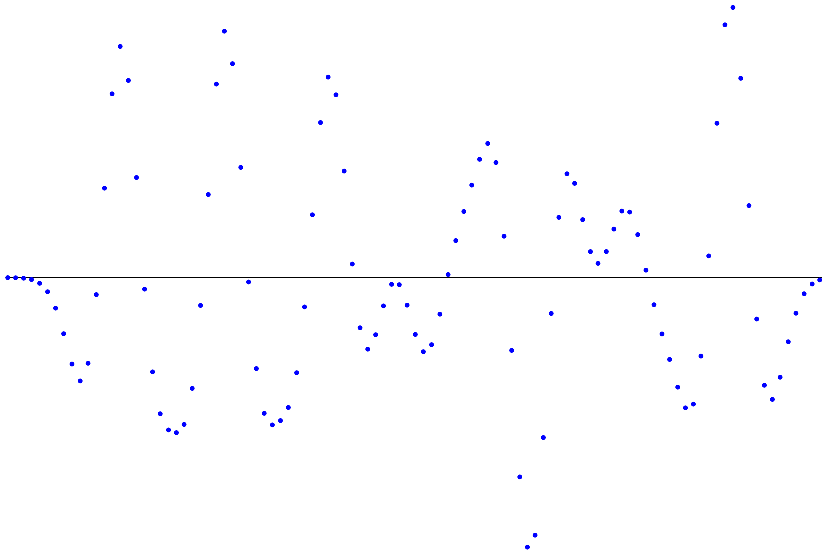
Compressed Sensing

Deconvolution in the Frequency Domain

A Sampling Theorem for Deconvolution

Robustness to Noise

Deconvolution from sampled data



Mathematical model

- Signal: superposition of Dirac measures with support T

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

- Data: n samples of convolution with Gaussian/Ricker kernel K

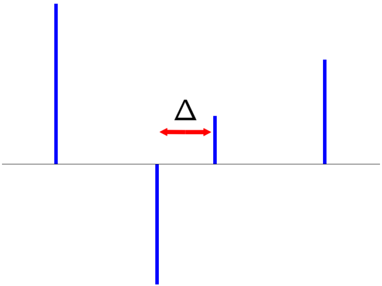
$$\begin{aligned} y &:= \mathcal{K}\, x \\ y_i &:= (K * x)(s_i), \quad i = 1, 2, \ldots, n \end{aligned}$$

Theoretical questions

1. Is the problem well posed?

2. Does TV-norm minimization work?

Minimum separation



Kernels are approximately low-pass

The support cannot be too clustered

Sampling proximity

We need two samples per spike

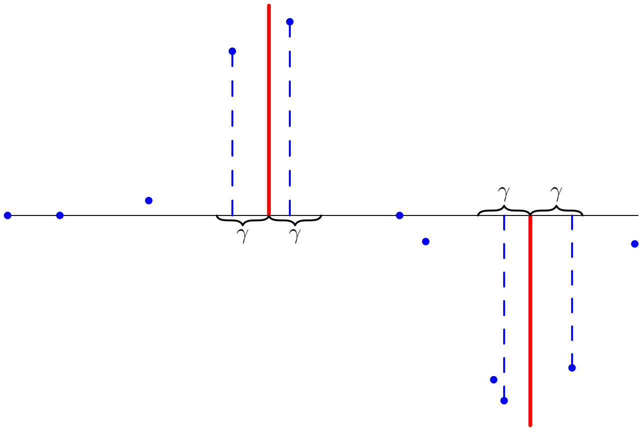
Convolution kernel decays: at least two samples close to each spike

Samples S and support T have sample proximity γ if for every $t_i \in T$ there exist $s, s' \in S$ such that

$$|t_i - s| \leq \gamma \quad \text{and} \quad \left|t_j - s'\right| \leq \gamma$$

We consider arbitrary non-uniform sampling patterns with fixed γ

Sampling proximity



Theoretical questions

1. Is the problem well posed?

2. Does TV-norm minimization work?

Estimation via convex programming

Optimization over finite real measures $\tilde{x}$

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

$$(\mathcal{K}\,\tilde{x})_j := (K * \tilde{x})\,(s_j)\,, \quad j = 1, 2, \ldots, n$$

Dual certificate

A dual certificate of the TV norm at

$$x = \sum_i a_i \delta_{t_i} \qquad a_i \in \mathbb{R},\ t_i \in T$$

guarantees that x is the unique solution if

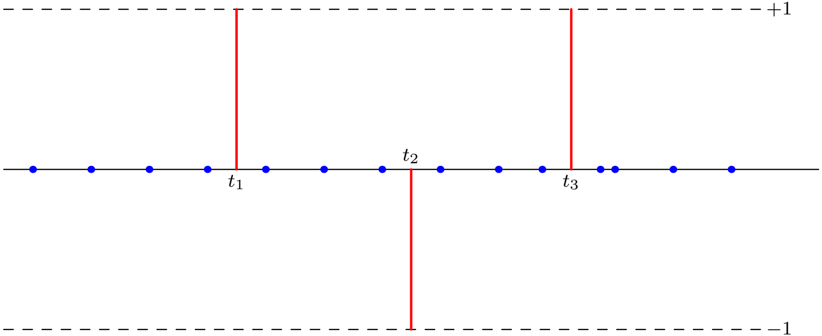
$$q(t) := (\mathcal{K}^T v)(t) = \sum_{j=1}^{n} v_j K(s_j - t)$$

$$q(t_i) = \operatorname{sign}(a_i) \qquad \text{if } t_i \in T$$

$$|q(t)| < 1 \qquad \text{if } t \notin T$$

Range of $\mathcal{K}^T$ is spanned by shifted copies of K fixed at the samples

Certificate for deconvolution



Certificate construction

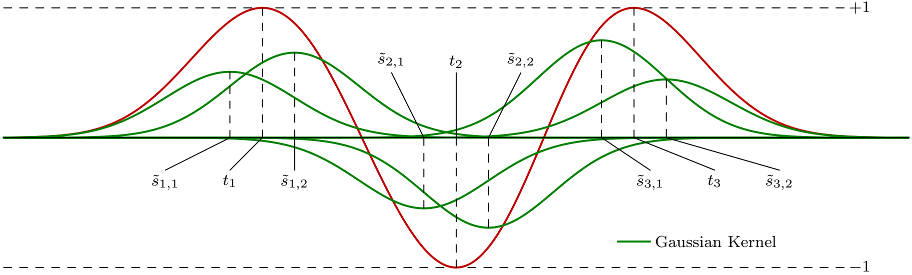
Only use subset $\widetilde{S}$ containing 2 samples close to each spike

$$q(t) = \sum_{s_j \in \widetilde{S}} v_j K\left(s_j - t\right)$$

Fit v so that for all $t_i \in T$

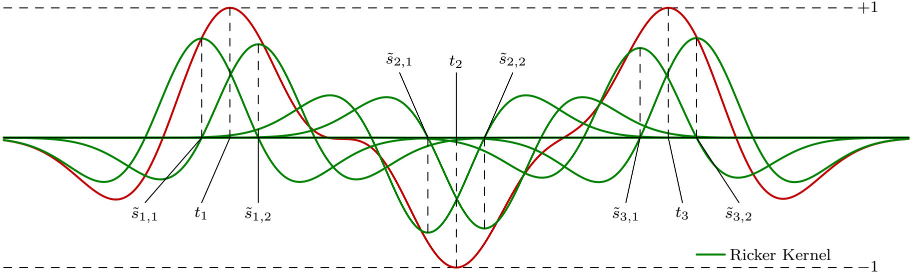
$$\begin{aligned} q\left(t_i\right) &= \operatorname{sign}\left(a_i\right) \\ q'\left(t_i\right) &= 0 \end{aligned}$$

It works!



+1
−1
$\tilde{s}_{2,1}$
t_2
$\tilde{s}_{2,2}$
$\tilde{s}_{1,1}$
t_1
$\tilde{s}_{1,2}$
$\tilde{s}_{3,1}$
t_3
$\tilde{s}_{3,2}$
Gaussian Kernel

It works!



+1
−1
$\tilde{s}_{1,1}$
t_1
$\tilde{s}_{1,2}$
$\tilde{s}_{2,1}$
t_2
$\tilde{s}_{2,2}$
$\tilde{s}_{3,1}$
t_3
$\tilde{s}_{3,2}$
Ricker Kernel

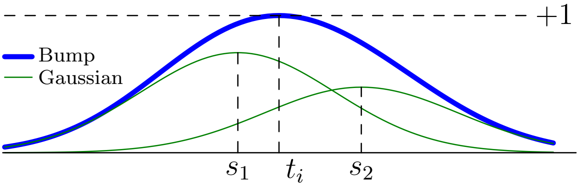
Certificate construction

Problem: The construction is difficult to analyze (coefficients vary)

Solution: Reparametrization into *bumps* and *waves*

$$\begin{aligned} q(t) &= \sum_{s_j \in \widetilde{S}} v_j K\left(s_j - t\right) \\ &= \sum_{t_i \in T} \alpha_i B_{t_i}(t, \tilde{s}_{i,1}, \tilde{s}_{i,2}) + \beta_i W_{t_i}(t, \tilde{s}_{i,1}, \tilde{s}_{i,2}), \end{aligned}$$

Bump function (Gaussian kernel)

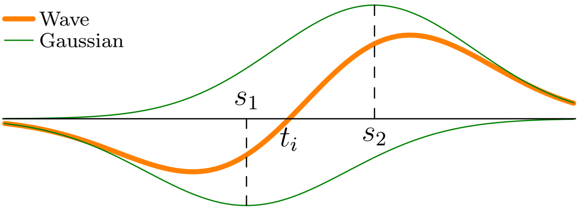


$$B_{t_i}(t, \tilde{s}_{i,1}, \tilde{s}_{i,2}) := b_{i,1}K(\tilde{s}_{i,1} - t) + b_{i,2}K(\tilde{s}_{i,2} - t)$$

$$B_{t_i}(t_i, \tilde{s}_{i,1}, \tilde{s}_{i,2}) = 1$$

$$\frac{\partial}{\partial t} B_{t_i}(t_i, \tilde{s}_{i,1}, \tilde{s}_{i,2}) = 0$$

Wave function (Gaussian kernel)

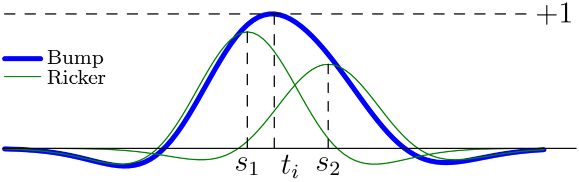


$$W_{t_i}(t, \tilde{s}_{i,1}, \tilde{s}_{i,2}) = w_{i,1}K(\tilde{s}_{i,1} - t) + w_{i,2}K(\tilde{s}_{i,2} - t)$$

$$W_{t_i}(t_i, \tilde{s}_{i,1}, \tilde{s}_{i,2}) = 0$$

$$\frac{\partial}{\partial t}W_{t_i}(t_i, \tilde{s}_{i,1}, \tilde{s}_{i,2}) = 1$$

Bump function (Ricker wavelet)

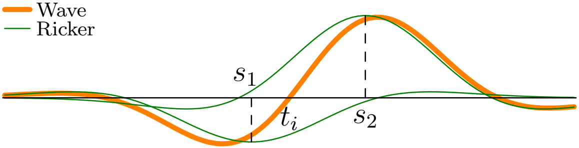


$$B_{t_i}(t, \tilde{s}_{i,1}, \tilde{s}_{i,2}) := b_{i,1} K(\tilde{s}_{i,1} - t) + b_{i,2} K(\tilde{s}_{i,2} - t)$$

$$B_{t_i}(t_i, \tilde{s}_{i,1}, \tilde{s}_{i,2}) = 1$$

$$\frac{\partial}{\partial t} B_{t_i}(t_i, \tilde{s}_{i,1}, \tilde{s}_{i,2}) = 0$$

Wave function (Ricker wavelet)



$$W_{t_i}(t, \tilde{s}_{i,1}, \tilde{s}_{i,2}) = w_{i,1}K(\tilde{s}_{i,1} - t) + w_{i,2}K(\tilde{s}_{i,2} - t)$$

$$W_{t_i}(t_i, \tilde{s}_{i,1}, \tilde{s}_{i,2}) = 0$$

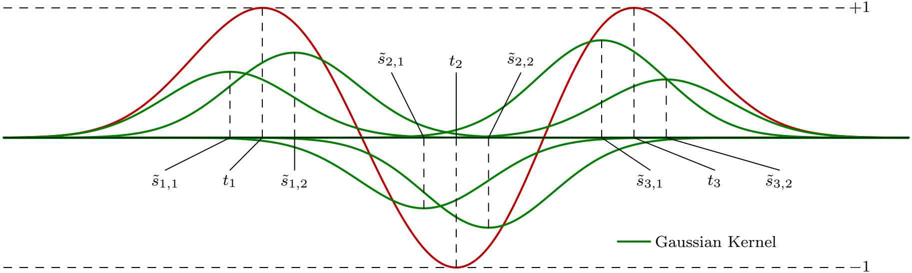
$$\frac{\partial}{\partial t}W_{t_i}(t_i, \tilde{s}_{i,1}, \tilde{s}_{i,2}) = 1$$

Certificate construction

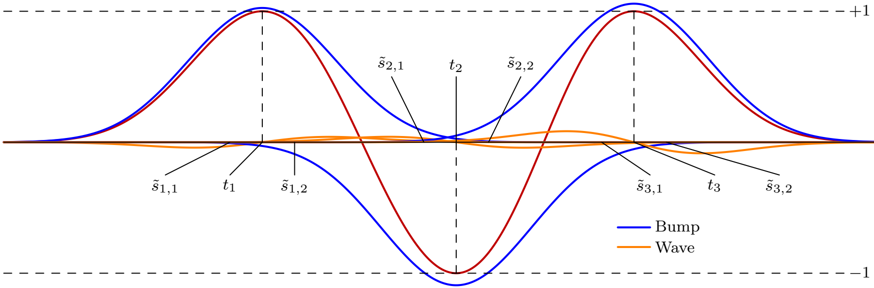
Reparametrization decouples the coefficients

$$\begin{aligned} q(t) &= \sum_{s_j \in \widetilde{S}} v_j K\left(s_j - t\right) \\ &= \sum_{t_i \in T} \alpha_i B_{t_i}(t, \tilde{s}_{i,1}, \tilde{s}_{i,2}) + \beta_i W_{t_i}(t, \tilde{s}_{i,1}, \tilde{s}_{i,2}) \\ &\approx \sum_{t_i \in T} \operatorname{sign}\left(a_i\right) B_{t_i}(t, \tilde{s}_{i,1}, \tilde{s}_{i,2}) \end{aligned}$$

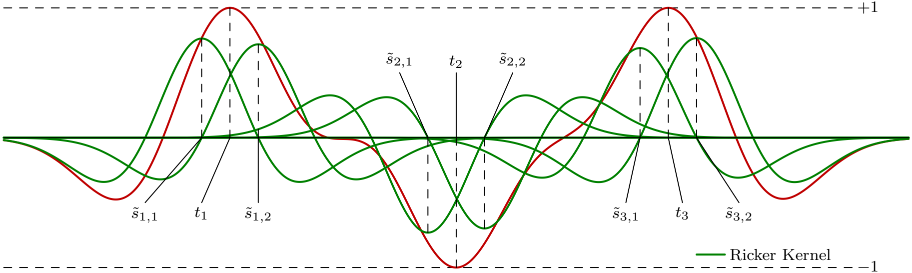
Certificate for deconvolution (Gaussian kernel)



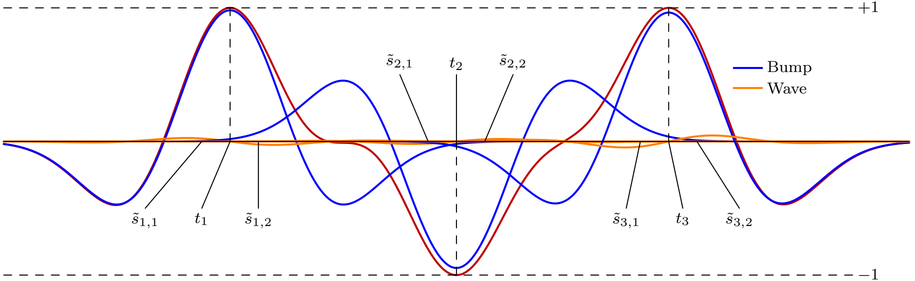
Certificate for deconvolution (Gaussian kernel)



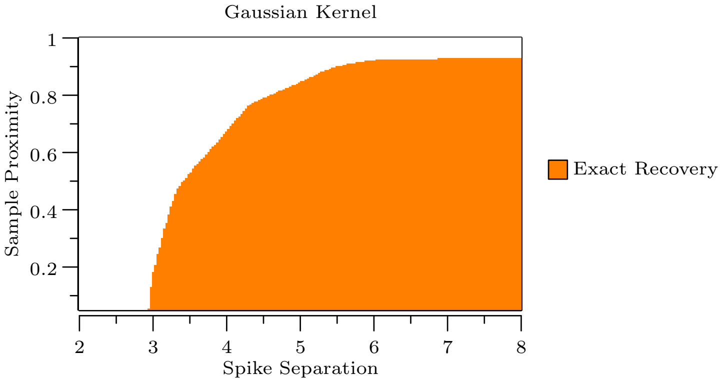
Certificate for deconvolution (Ricker wavelet)



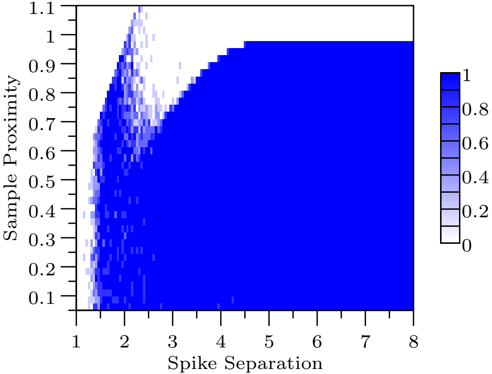
Certificate for deconvolution (Ricker wavelet)



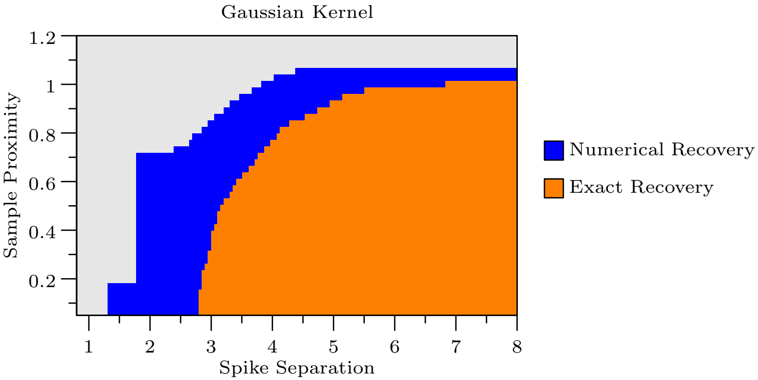
Exact recovery guarantees [Bernstein, F. 2017]



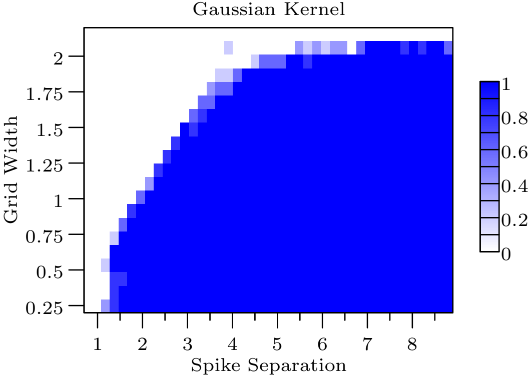
Numerical experiments (Gaussian kernel)



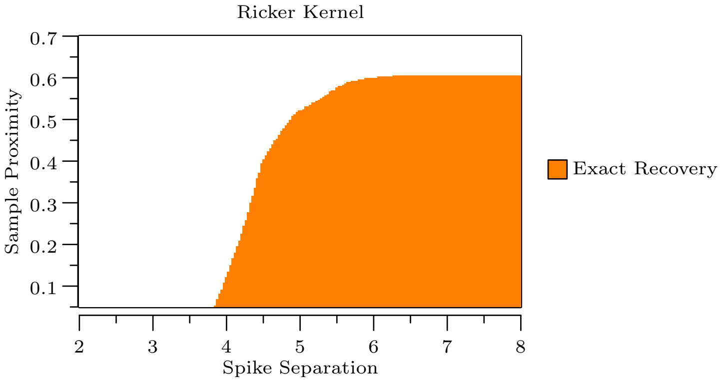
Guarantees vs numerical experiments (Gaussian kernel)



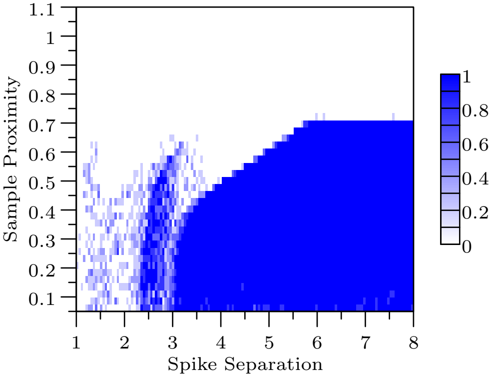
Numerical experiments (Gaussian kernel)



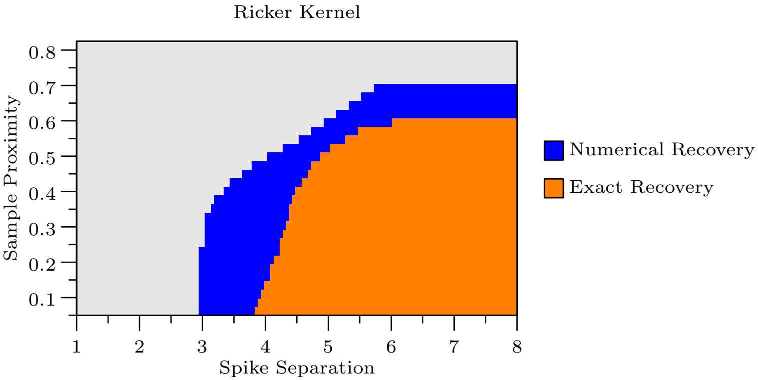
Exact recovery guarantees [Bernstein, F. 2017]



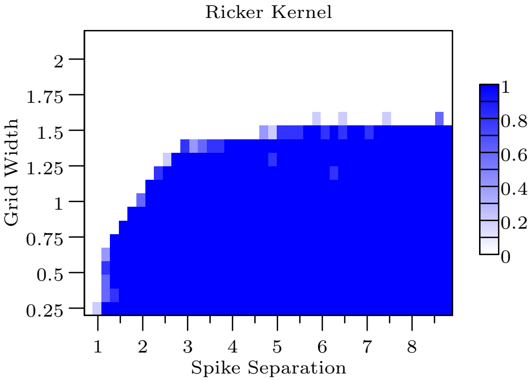
Numerical experiments (Ricker wavelet)



Guarantees vs numerical experiments (Ricker wavelet)



Numerical experiments (Ricker wavelet)

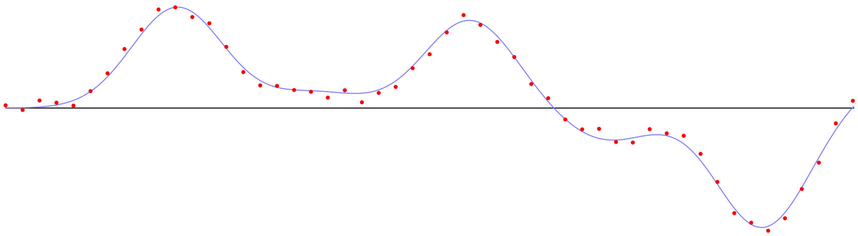


Dense additive noise

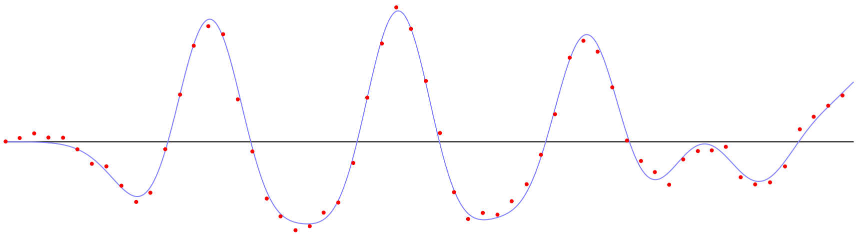
Noise with bounded ℓ_2 norm, i.e. $||z||_2 < \xi$

$$y_i := (K * x)(s_i) + z_i \quad i = 1, 2, \ldots, n.$$

Robustness to dense noise



Robustness to dense noise

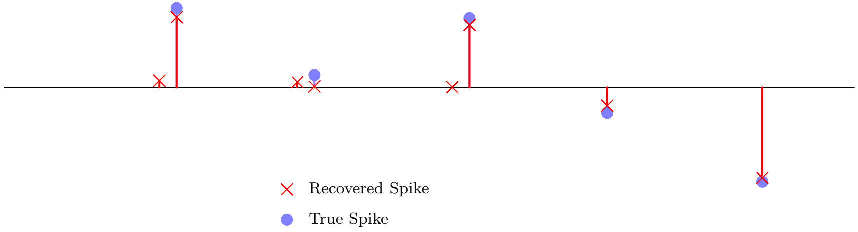


Robust deconvolution via convex programming

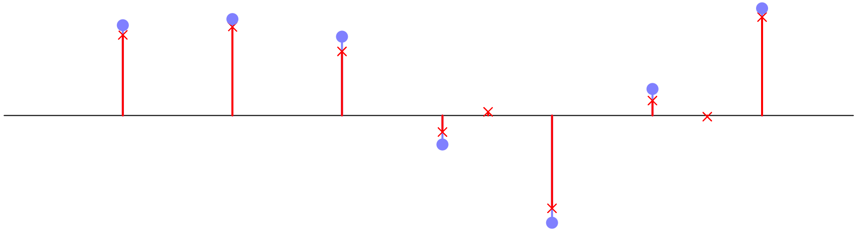
Noise level ξ is assumed known

$$\begin{aligned} \underset{\tilde{x}}{\text{minimize}} \quad & ||\tilde{x}||_{\text{TV}} \\ \text{subject to} \quad & \sum_{i=1}^{m} (y_i - (K * \tilde{x})(s_i))^2 \leq \xi^2 \end{aligned}$$

Robustness to dense noise



Robustness to dense noise



Support-detection accuracy

Original signal, support T

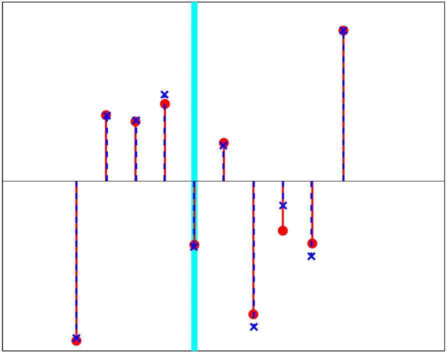
$$x = \sum_j a_j \delta_{t_j} \qquad a_j \in \mathbb{R},\ t_j \in T$$

Estimated signal, support $\hat{T}$

$$\hat{x} = \sum_j \hat{a}_j \delta_{\hat{t}_j} \qquad \hat{a}_j \in \mathbb{R},\ \hat{t}_j \in \hat{T}$$

Spike detection [Bernstein, F. 2017]

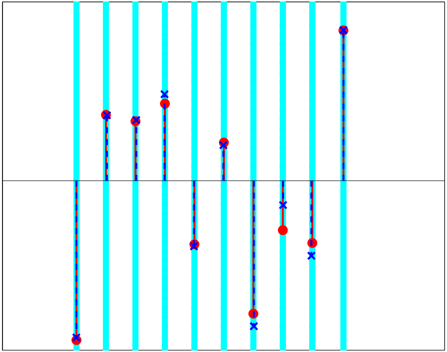
Under the same assumptions as for exact recovery



$$\left| a_j - \sum_{\left\{ \hat{t}_l \in \hat{T} : |\hat{t}_l - t_j| \leq \eta\sigma \right\}} \hat{a}_l \right| \leq C_1 \xi \sqrt{|T|} \quad \text{for all } t_j \in T, \quad \eta \leq 0.15\sigma$$

Support-detection accuracy [Bernstein, F. 2017]

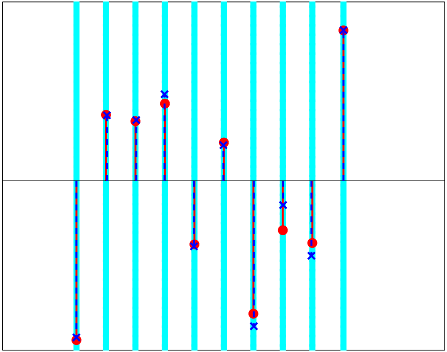
Under the same assumptions as for exact recovery



$$\sum_{\left\{\hat{t}_l \in \hat{T}, t_j \in T : |\hat{t}_l - t_j| \leq \eta\sigma\right\}} |\hat{a}_l| \left(\hat{t}_l - t_j\right)^2 \leq C_2 \xi \sqrt{|T|}, \quad \eta \leq 0.15\sigma$$

False positives [Bernstein, F. 2017]

Under the same assumptions as for exact recovery



$$\sum_{\left\{\hat{t}_l \in \hat{T} : |\hat{t}_l - t_j| > \eta\sigma\right\}} |\hat{a}_l| \leq C_3 \xi \sqrt{|T|}, \quad \eta \leq 0.15\sigma$$

Support-detection accuracy

Corollary

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

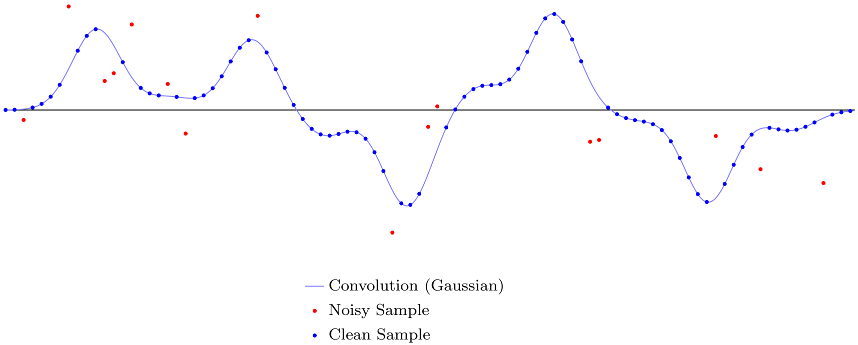
$$|t_i - \hat{t}_i| \leq \sqrt{\frac{C_2 \xi}{|a_i| - C_1 \xi}}$$

Sparse additive noise

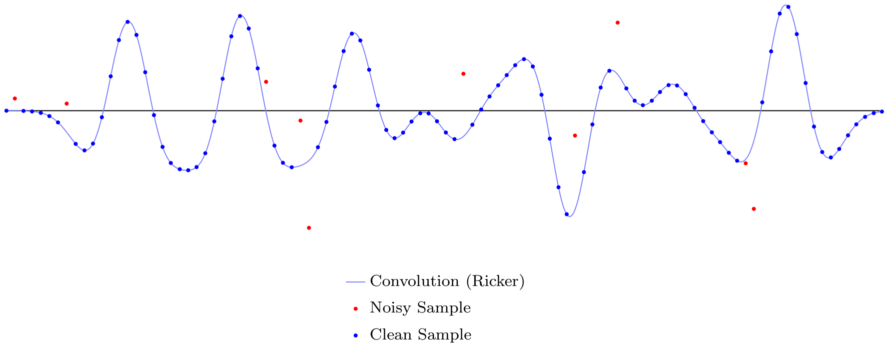
Impulsive noise $w \in \mathbb{R}^n$ with arbitrary amplitude

$$y_i := (K * x)(s_i) + w_i \quad i = 1, 2, \ldots, n.$$

Robustness to sparse noise



Robustness to sparse noise

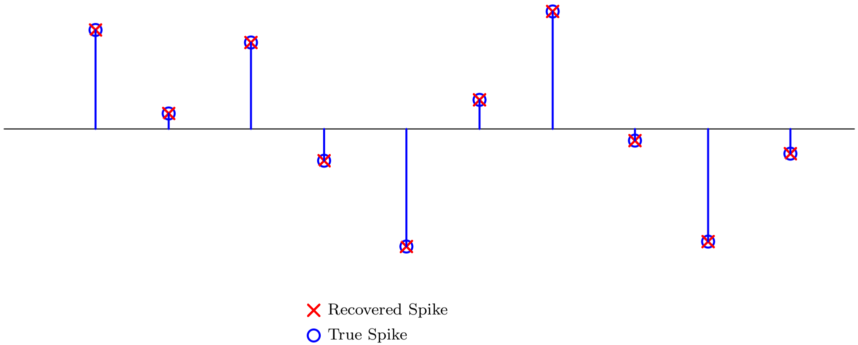


Robust deconvolution via convex programming

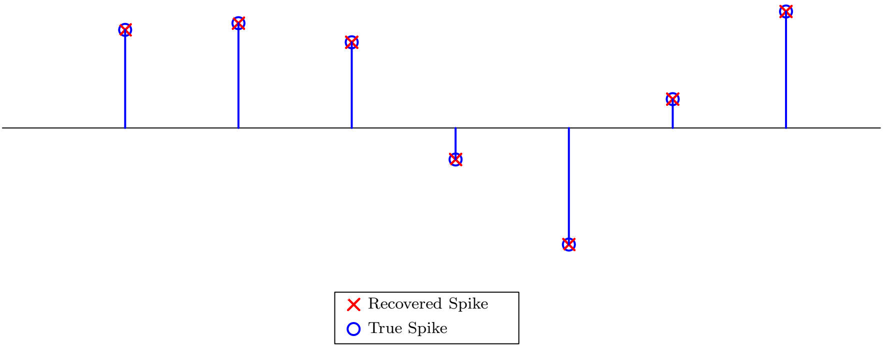
We incorporate an additional variable to model sparse noise

$$\begin{aligned} \underset{\tilde{x},\, \tilde{w}}{\text{minimize}} \quad & \|\tilde{x}\|_{\text{TV}} + \lambda \|\tilde{w}\|_1 \\ \text{subject to} \quad & (K * \tilde{x})(s_i) + \tilde{w}_i = y_i, \quad i = 1, \ldots, n, \end{aligned}$$

Robustness to sparse noise



Robustness to sparse noise



Theoretical guarantees [Bernstein, F. 2017]

Exact recovery occurs for $\lambda = 2$, as long as

- The samples lie on a *grid* with step size $0.065\,\sigma \leq \tau \leq 0.2375\,\sigma$ (Gaussian) or $0.0775\,\sigma \leq \tau \leq 0.165\,\sigma$ (Ricker)
- The signal has a *minimum separation* of $\Delta(T) \geq 3.751\sigma$ (Gaussian) or $\Delta(T) \geq 5.056\sigma$ (Ricker)
- The noisy samples are also *separated* by the same distance
- There are 2 *clean* samples surrounding each noisy sample
- There are 2 *clean* samples surrounding each spike

Conclusion

Geophysicists proposed ℓ_1-norm based deconvolution in the 1970s

Compressed-sensing intuition / tools for randomized measurements do not apply directly

Conditions beyond sparsity are necessary to make the problem well posed

Under such conditions the method achieves exact recovery and is robust to dense and sparse noise

Related work

- *Exact reconstruction using Beurling minimal extrapolation*. Y. De Castro and F. Gamboa. *J. of Math. Analysis and App.*, 2012
- *Exact support recovery for sparse spikes deconvolution*. V. Duval and G. Peyré. *Found. of Comp. Math.*, 2015
- *Robust recovery of stream of pulses using convex optimization*. T. Bendory, S. Dekel and A. Feuer. *J. of Math. Analysis and App.*, 2016
- *Super-resolution without separation*. G. Schiebinger, E. Robeva and B. Recht. *Information and Inference*, 2016
- *Towards Generalized FRI Sampling With an Application to Source Resolution in Radioastronomy*. H. Pan, T. Blu and M. Vetterli. *IEEE Trans. on Signal Proc.*, 2016

References: Reflection seismology

- *Robust modeling with erratic data*. J. F. Claerbout and F. Muir. *Geophysics*, 1973
- *Deconvolution with the ℓ_1 norm*. H. L. Taylor, S. C. Banks and J. F. McCoy. *Geophysics*, 1979
- *Reconstruction of a sparse spike train from a portion of its spectrum and application to high-resolution deconvolution*. S. Levy and P. K. Fullagar. *Geophysics*, 1981
- *Linear inversion of band-limited reflection seismograms*. F. Santosa and W. W. Symes. *SIAM J. Sci. Stat. Comp.*, 1986

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: Super-resolution

- *Prolate spheroidal wave functions, Fourier analysis, and uncertainty V V - The discrete case*. D. Slepian. *Bell System Technical Journal*, 1978
- *Super-resolution, extremal functions and the condition number of Vandermonde matrices*. A. Moitra. *Symposium on Theory of Computing (STOC)*, 2015
- *Towards a mathematical theory of super-resolution*. E. J. Candès and C. Fernandez-Granda. *Comm. on Pure and Applied Math.*, 2013

References: Deconvolution

- *Deconvolution of point sources: A sampling theorem and robustness guarantees*. B. Bernstein, C. Fernandez-Granda