Modeling and blind deconvolution via sparse representations

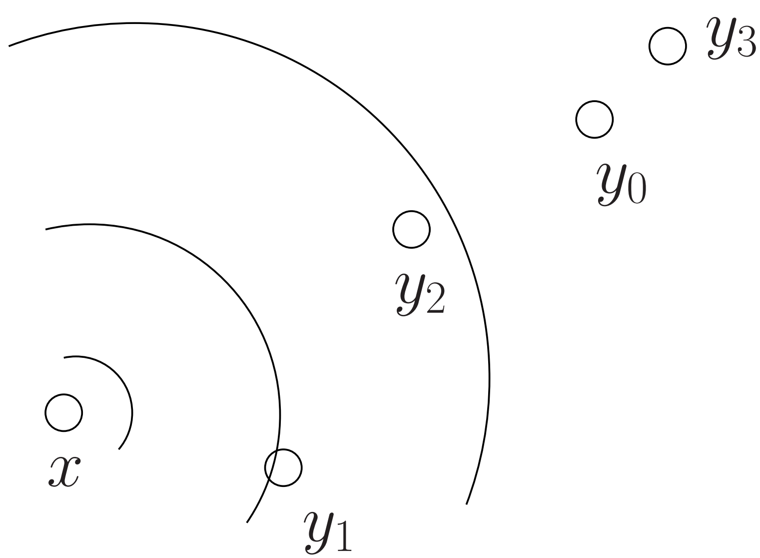
Tryphon Georgiou
University of Minnesota

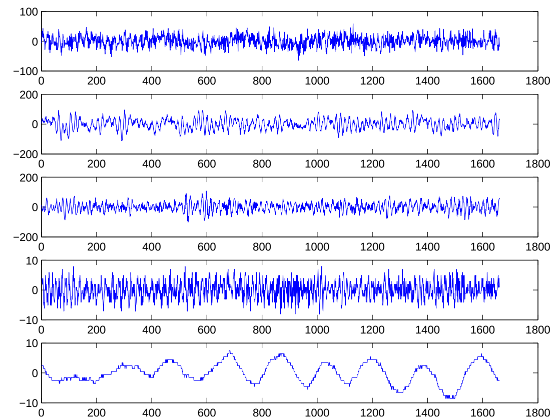
Joint work with
Lipeng Ning and Allen Tannenbaum

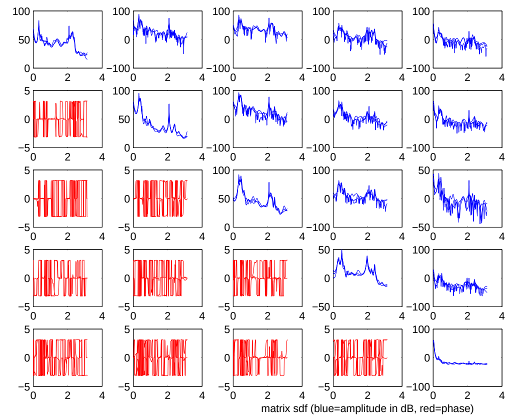
Lund, February 2010

Motivation

disturbance source localization
sinusoids in noise
dynamics







distributed sensor network

Sparsity and L1

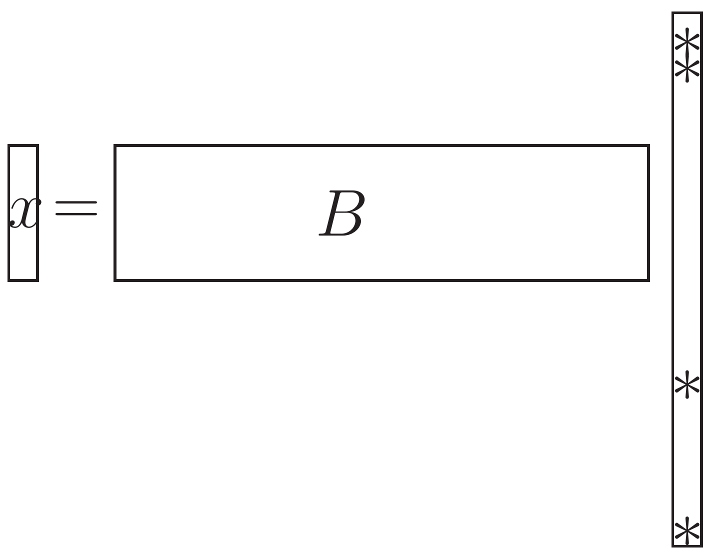
more than 20 years of history. . .

Sparse representations

$\|v\|_0 =$ # or nonzero entries
$\|v\|_1 = \sum_k |v_k|$

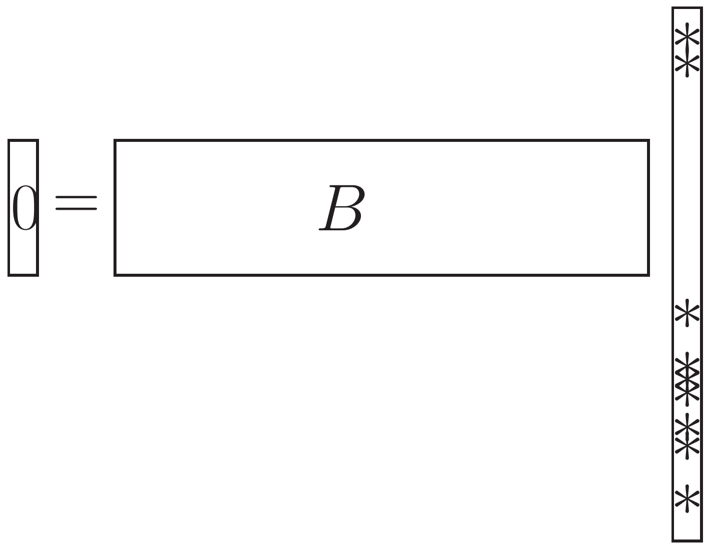
Problem: $\min\{\|v\|_0$ subject to $Bv = x\}$ — a combinatorial problem

Relaxation: $\min\{\|v\|_1$ subject to $Bv = x\}$ — a convex problem



Primer on sparsity

Def (Donoho & Elad): spark(B) = least number of linearly dependent columns



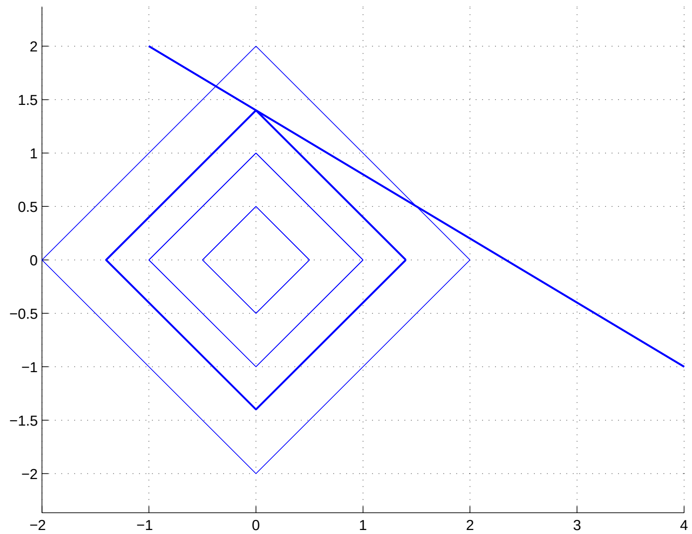
Proposition: $Bv = x$

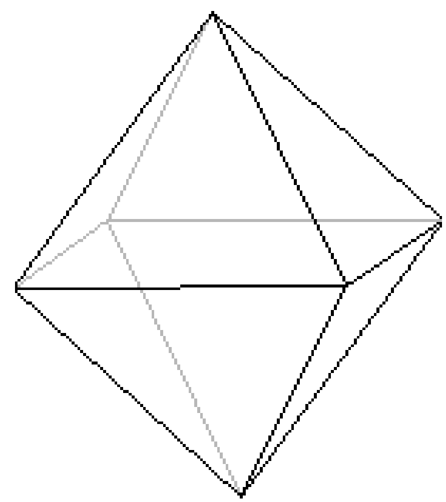
if $\|v\|_0 < \frac{1}{2}$ spark(B), then v is the sparsest solution

Primer (cont.)

$Bv = x$, if B is $m \times n$ (with $m < n$)
for a general x, $\|v_{\text{optimal}}\|_0 = m$

But what if $\|v_{\text{optimal}}\|_0 < m$*?*





Observation: generically $\operatorname{argmin}\{\|v\|_1 \;:\; Bv = x\}$
will lie on a vertex, or edge, etc.

Primer (cont.)

Thm *Donoho, Candes & Tao, Elad, Zhang, . . .*
If B is suitably "well-conditioned",
<u>and</u> there is a sufficiently sparse solution
then:

$$\operatorname{argmin}\{\|v\|_1 \ : \ Bv = x\} = \operatorname{argmin}\{\|v\|_0 \ : \ Bv = x\}$$

Approximate solutions, noisy data

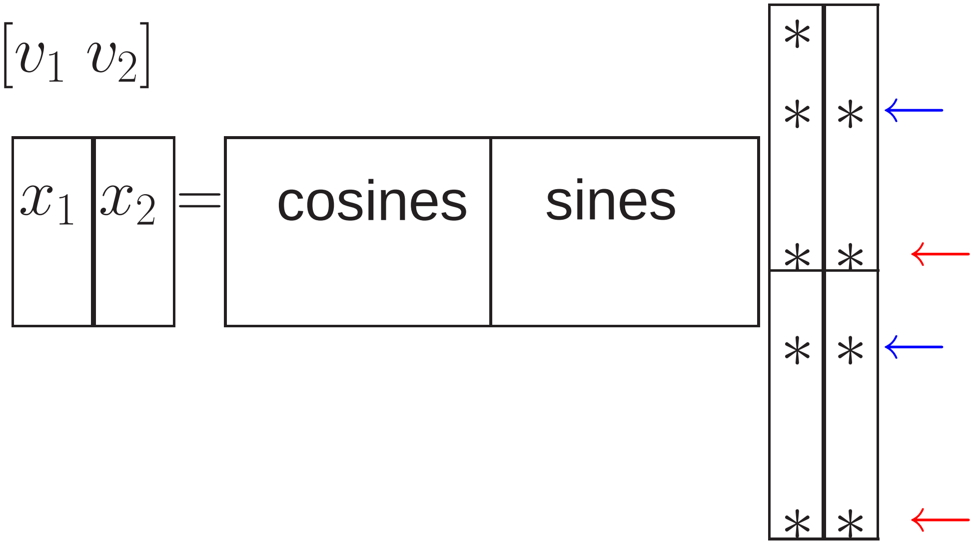
$\min\left\{\|v\|_1 \text{ subject to } \|Bv - x\|_2 \le \epsilon\right\}$ *Basis Pursuit Denoising*

$\min\left\{\|Bv - x\|_2 \text{ subject to } \|v\|_1 \le \sigma\right\}$ *Least Absolute Shrinkage and Selection Operator (LASSO)*

$\min\left\{w\|v\|_1 + \|Bv - x\|_2^2\right\}$ *Relaxed Basis Pursuit*

Joint sparsity, etc.

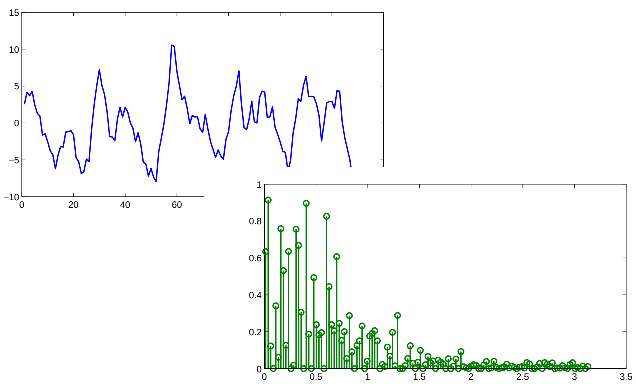
$[x_1\ x_2] = [\text{cosines}, \text{sines}][v_1\ v_2]$



“promote” coherent choices of cosines and sines

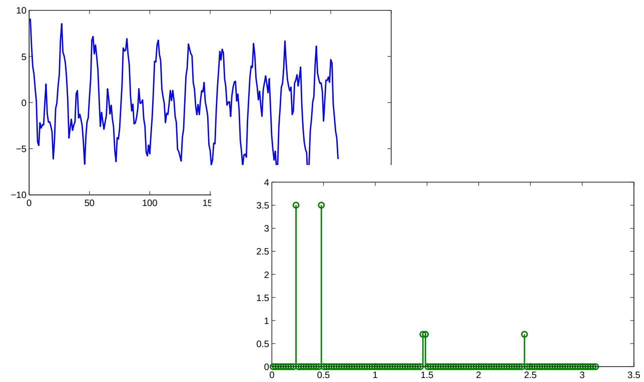
Identification: signals + dynamics

Blind deconvolution, sinusoids in colored noise, etc.

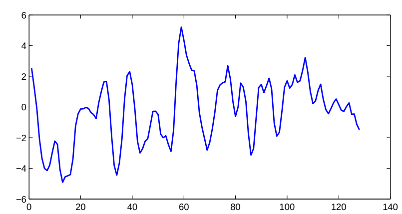


$\Rightarrow$

Signal



Noise



Sparsity vs. modeling error

$$v^* = \arg\min_v \{w\|v\|_1 + \frac{1}{2}\|y - Bv\|_2^2\}$$

Dual Problem:

$$\min_v \frac{1}{2}\|Bv\|_2^2$$
$$\text{s.t. } |B^T(y - Bv)|_i \leq w$$

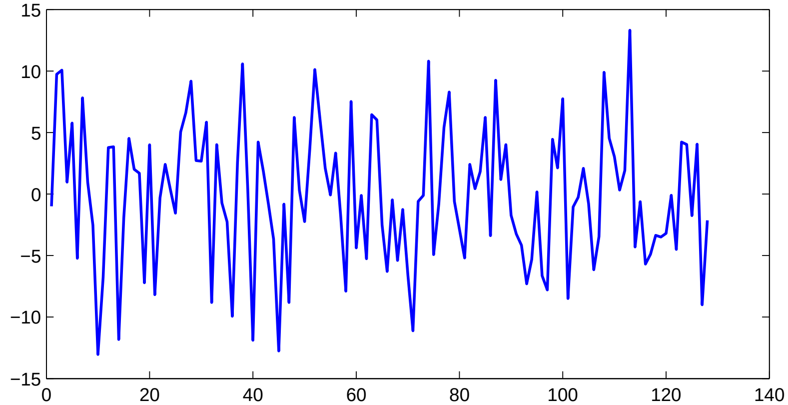
— minimizer $v^* = \lambda + n$, λ multiplier and $n \in$ Null(B).

— *if $w > |B^T y|_\infty$, v is zero*

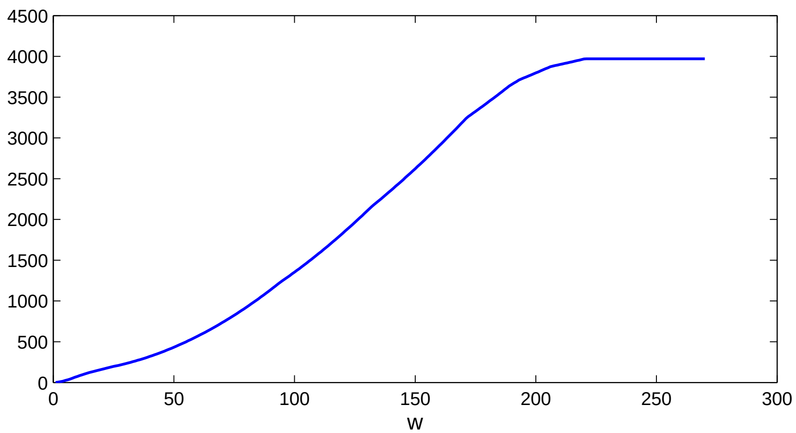
Sparsity vs. weight

$$\min_v w\|v\|_1 + \frac{1}{2}\|y - Bv\|_2^2$$

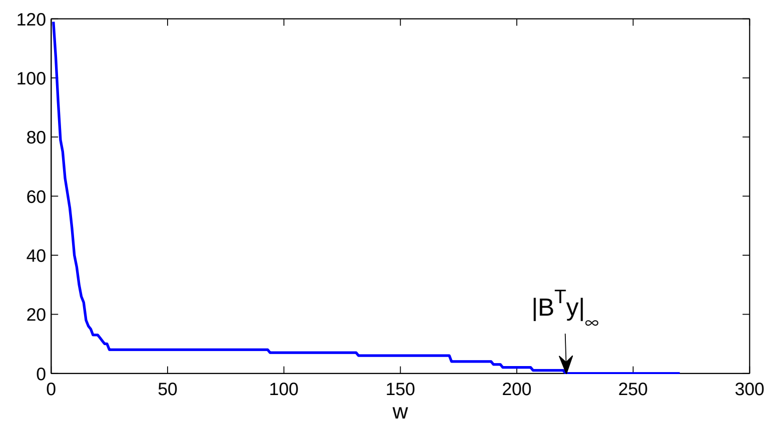
sinusoids in white noise



$\|y - Bv\|^2$ *vs. weight*

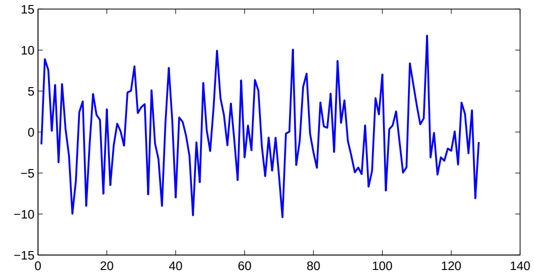


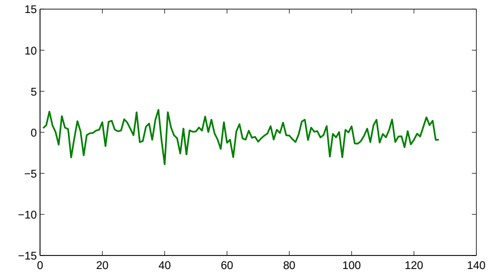
sparsity vs. weight



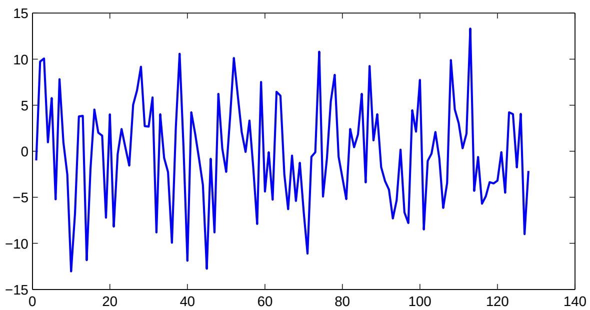
Signal recovery

recovered

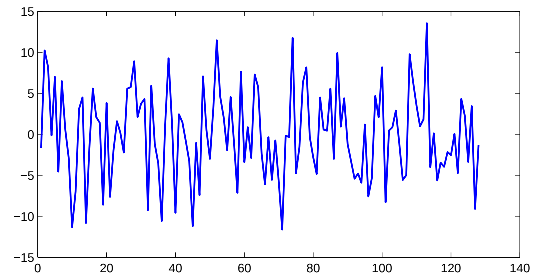


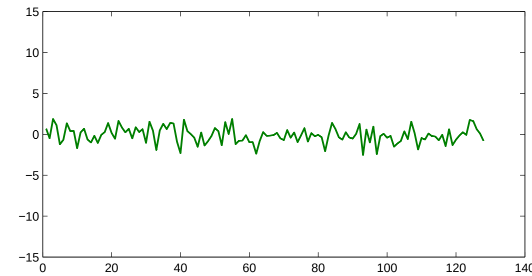


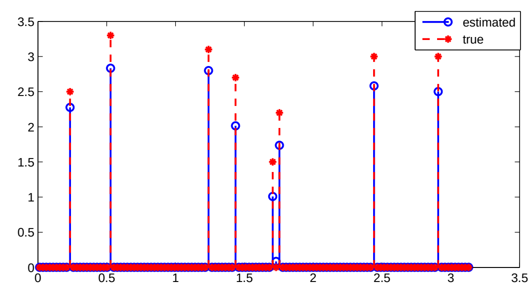
signal + noise



“true”







Iterative re-weighting

Candes, Wakin, Boyd

$$\min_v \|Wv\|_1 + \frac{1}{2}\|y - Bv\|_2^2$$

with $W = \mathrm{diag}(w_i)$, and update

$$w_i^{k+1} = \frac{1}{|v_i^k| + \epsilon}$$

in the limit... $\frac{|v_i|}{|v_i|+\epsilon} \approx \begin{cases} 0, & |v_i| \ll \epsilon \\ 1, & |v_i| \gg \epsilon \end{cases}$

Insight

Candes, Wakin, Boyd

— iterative minimization of a surrogate function “interpolating” $\|v\|_0$ and $\|v\|_1$

looking at duality

$$\min_v \frac{1}{2}\|Bv\|_2^2$$

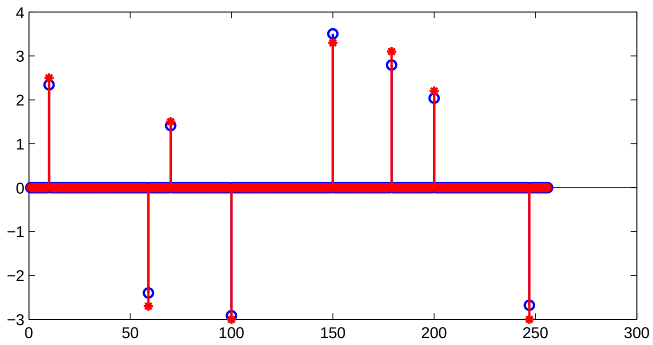
$$\text{s.t. } |B^T(y - Bv)|_i \leq w_i$$

How well does it do?

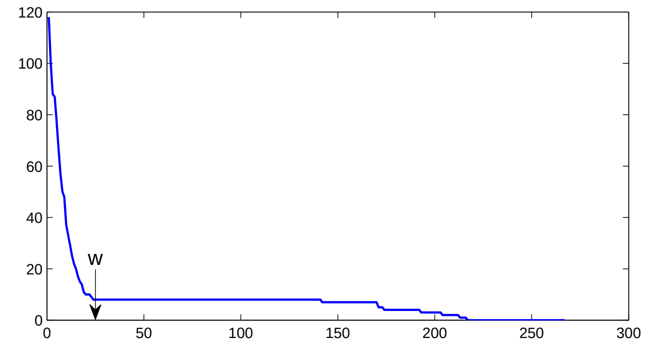
for sinusoids in white noise. . . *very well*

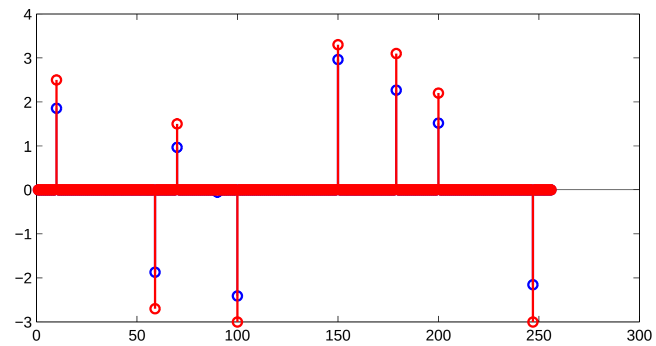
Candes, Wakin, Boyd

. . . in two iterations. . .

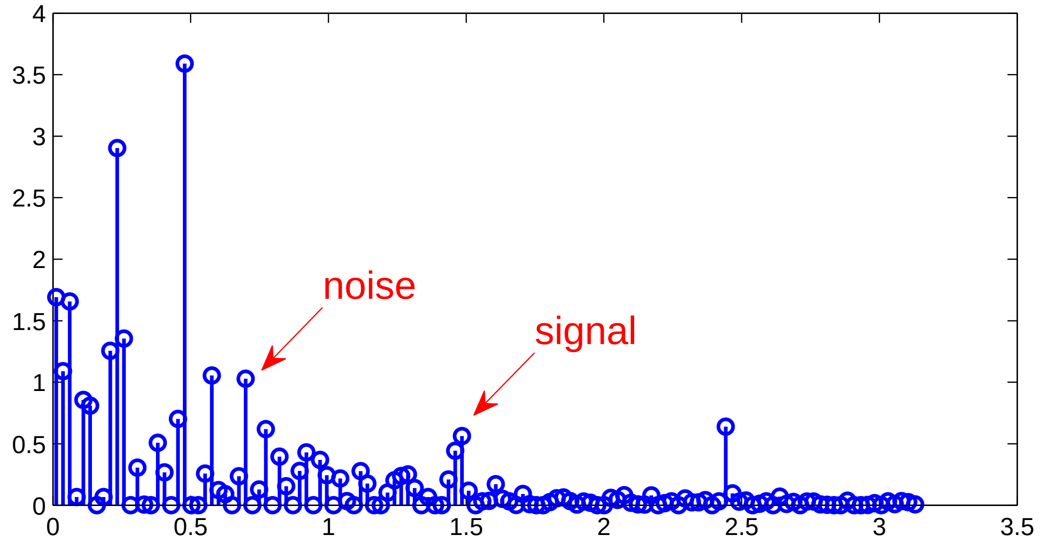


vs.

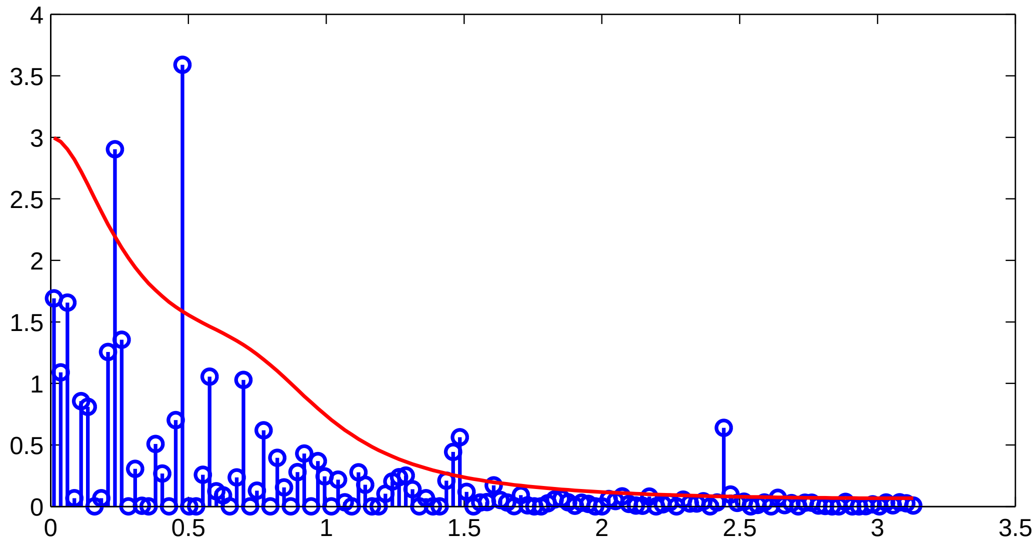




What if noise is colored?

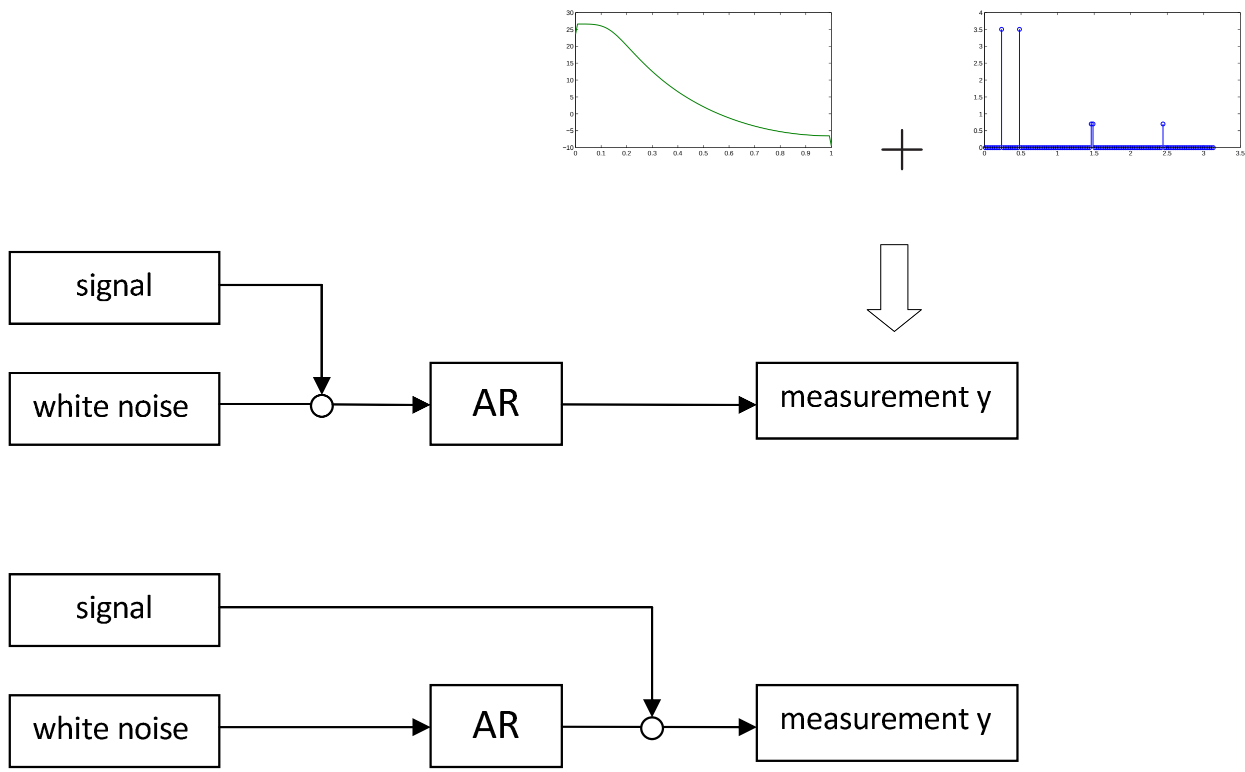


Insight

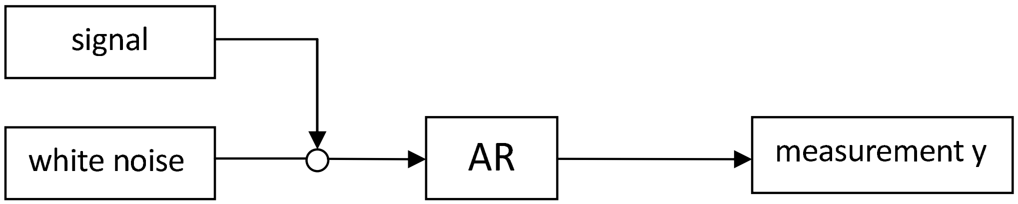


e.g., choose W accordingly. . .

System identification



System identification (cont.)

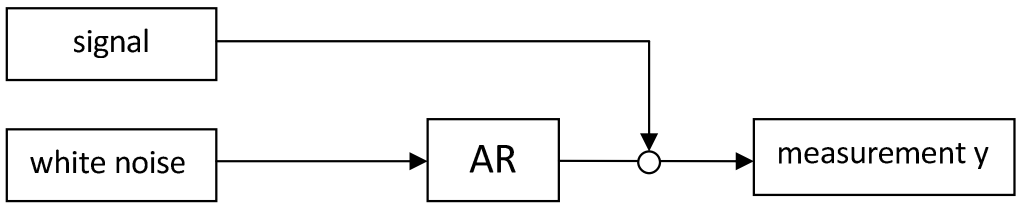


$$\begin{bmatrix} y_k \\ y_{k+1} \\ \vdots \end{bmatrix} = \begin{bmatrix} y_{k-1} & y_{k-2} & \cdots & y_{k-l} \\ y_k & y_{k-1} & \cdots & y_{k-l-1} \\ \vdots & \ddots & \ddots & \vdots \end{bmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_l \end{pmatrix} + \begin{pmatrix} x_k \\ x_{k+1} \\ \vdots \end{pmatrix} + \text{noise}$$

$$y = H_y a + Bv + \text{noise}$$

$$\min_{a,v} w\|v\|_1 + \frac{1}{2}\|y - H_y a - Bv\|_2^2$$

System identification (cont.)



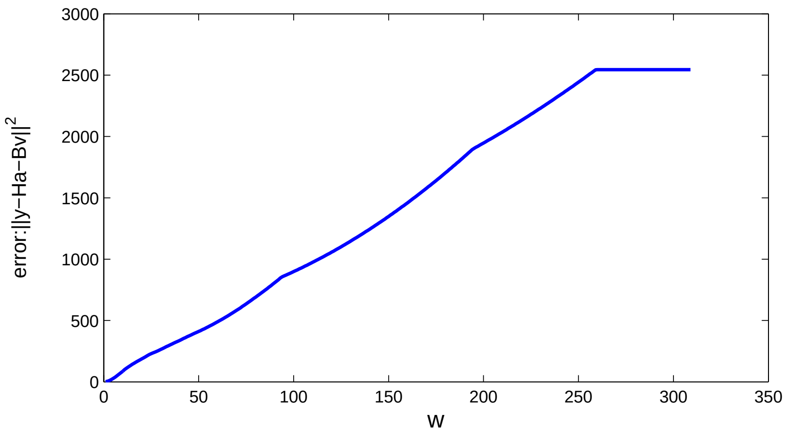
$$\text{noise} = \begin{pmatrix} 1 & -a_1 & \dots & -a_l & 0 & \dots & \dots \\ 0 & 1 & -a_1 & \dots & -a_l & 0 & \dots \\ \vdots & \ddots & \ddots & & & \ddots & \vdots \end{pmatrix} \left(\begin{bmatrix} y_k \\ y_{k+1} \\ \vdots \end{bmatrix} - Bv \right)$$

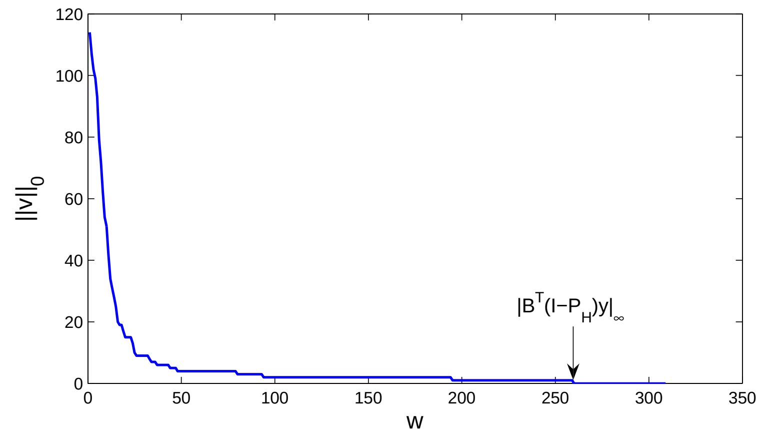
$$\text{noise} = T_a(y - Bv)$$

$$\min_{a,v} \|Wv\|_1 + \frac{1}{2}\|T_a(y - Bv)\|_2^2$$

Sparsity vs. weight

$$\min_{a,v} w\|v\|_1 + \frac{1}{2}\|y - H_y a - Bv\|_2^2$$





— *if* $w > \|B^T(I - P_H)y\|_\infty$, v *is zero*

Iterative re-weighting *a la Candes etal.*

$$\min_{a,v} \|Wv\|_1 + \frac{1}{2}\|y - H_y a - Bv\|_2^2$$

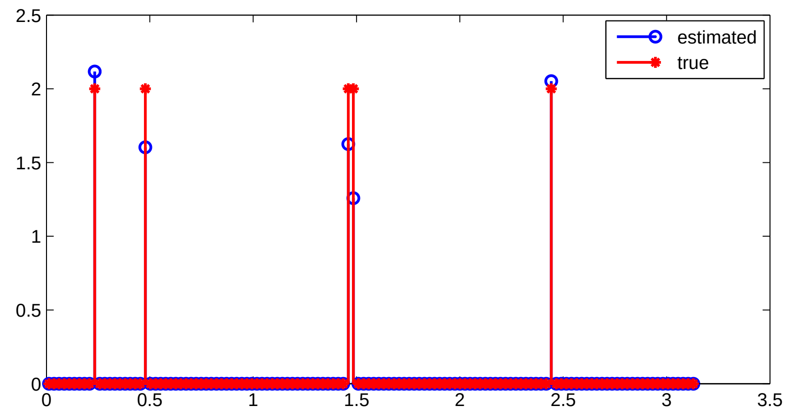
— update

$$w_i = \frac{1}{SNNR(i) + \epsilon}$$

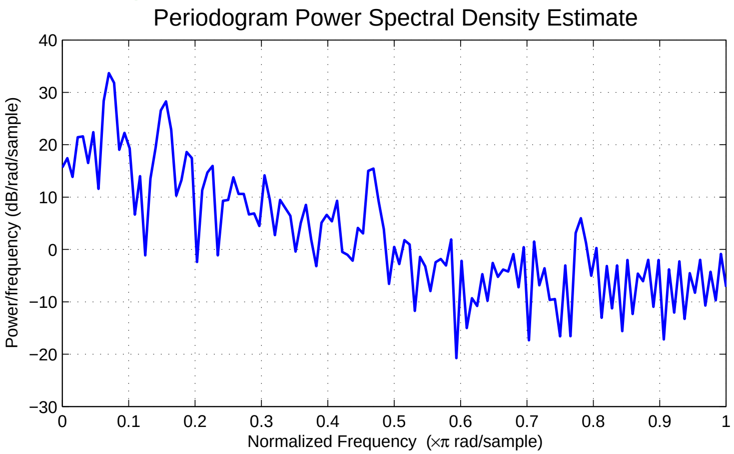
e.g. $w_i \sim 1/(\|(v_{\sin}, v_{\cos})\| + \epsilon)$
. . . perio/AR-spectrum

Example

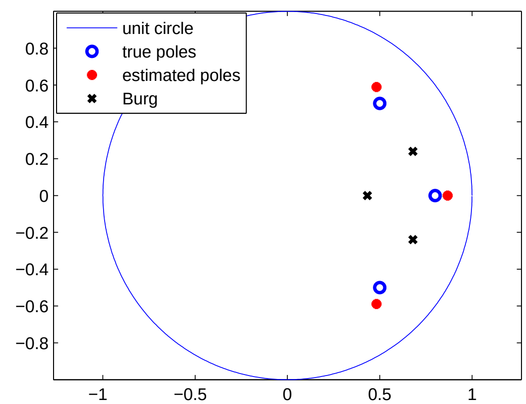
spectral lines



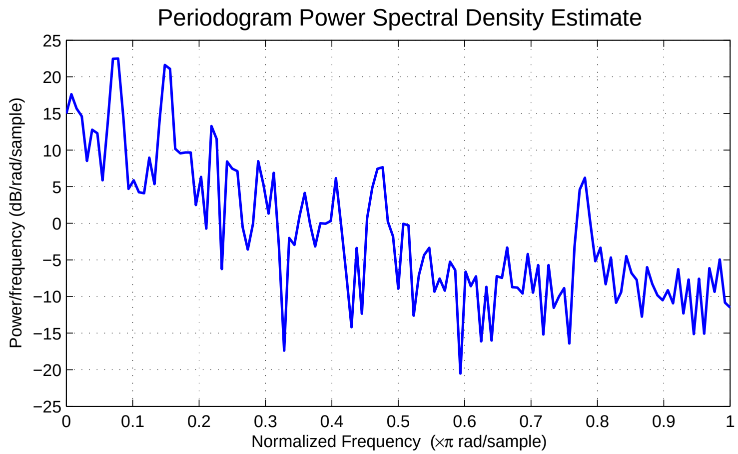
periodogram

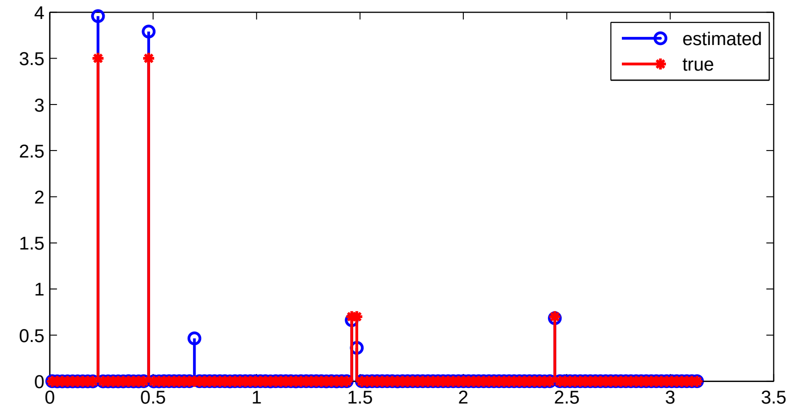


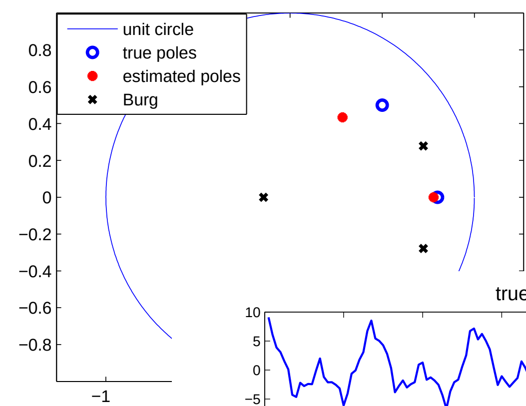
dynamics

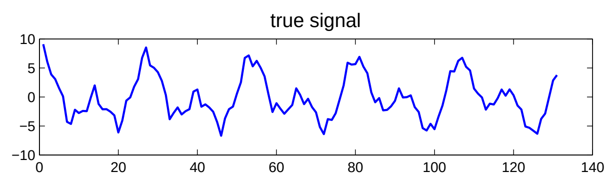


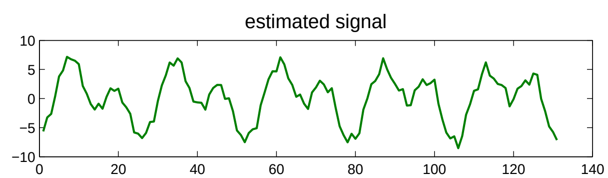
Example


Periodogram Power Spectral Density Estimate
Power/frequency (dB/rad/sample)
Normalized Frequency (×π rad/sample)


estimated
true


unit circle
true poles
estimated poles
Burg


true signal


estimated signal

Recap

— sparse representations in system identification

resolution (limits?)

— interplay between dynamics and sparsity?

if $(v = 0, a)$ satisfy conditions of the dual, then v_{opt} is "small".
is there a "uniqueness" result?
stability of the AR model?