Spectrum Sensing – an application of sparse signal recovery

Back projection with sub-sampled DFT matrices
(Maximum Correlation Estimation)

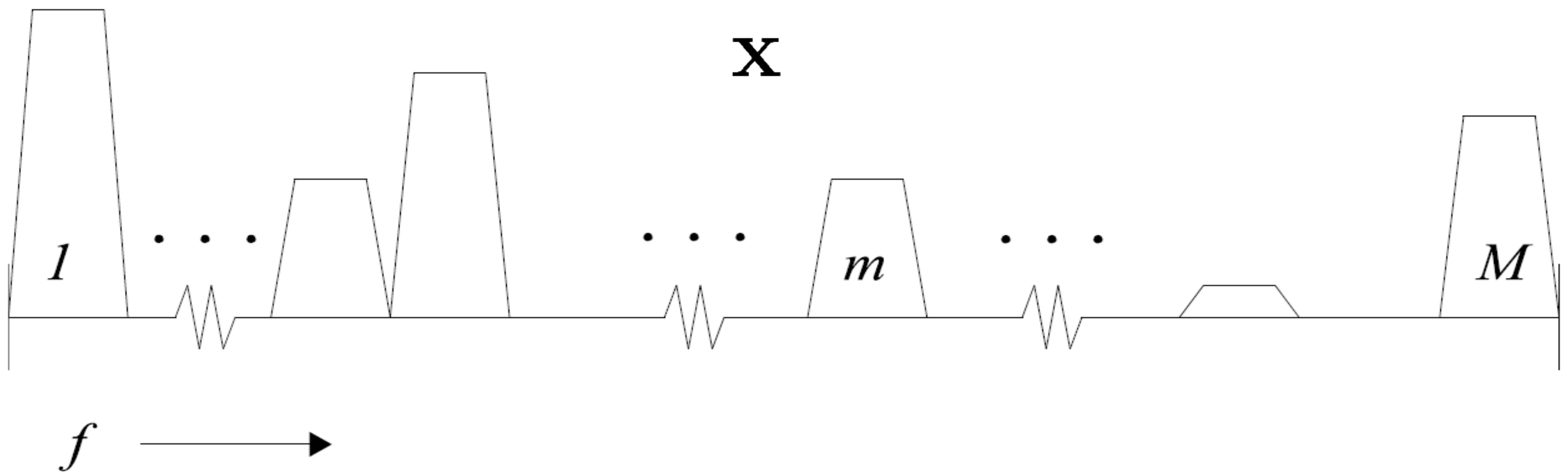
Matt Malloy – CS 838 project

Outline

- Spectrum Sensing
- Sub-sampled DFT matrices
- Block sparsity
  - *On the reconstruction of block-sparse signals with an optimal number of measurements [Stojnic, Parvaresh, Hassibi, '09]*
- Something easier than solving L1 minimization
  - *Back projection + thresholding*
  - Related work
  - Proof: Asymptotic probability of error for 1-sparse signals
  - Proof: Bound for s-sparse signals
- Numerical results comparing back projection to standard CS approach

What is spectrum sensing?

Measurement and classification of the radio spectrum into used and unused bands



Quickly and reliably detect open radio spectrum

Cognitive radio context

- Energy detection
- Coherent detection
- Feature detection

Many practical constraints

- Dynamic range
- Duration of observation

Is spectrum occupation sparse?

Example - Digital TV spectrum

38 x 6 MHz channels

44 MHz – 88 MHz and 470 MHz – 698 MHz

Madison has 6 occupied channels (CBS 3, NBC 15, PBS 21, ABC 27, FOX 47,CW 57)

Digital TV bands in Madison are 6-sparse

No guarantee but much literature either assumes or shows that certain case are.

Traditional approach – sample in frequency

$$y(t) = \sum_{m=1}^{M} s_m(t) + w_m(t)$$

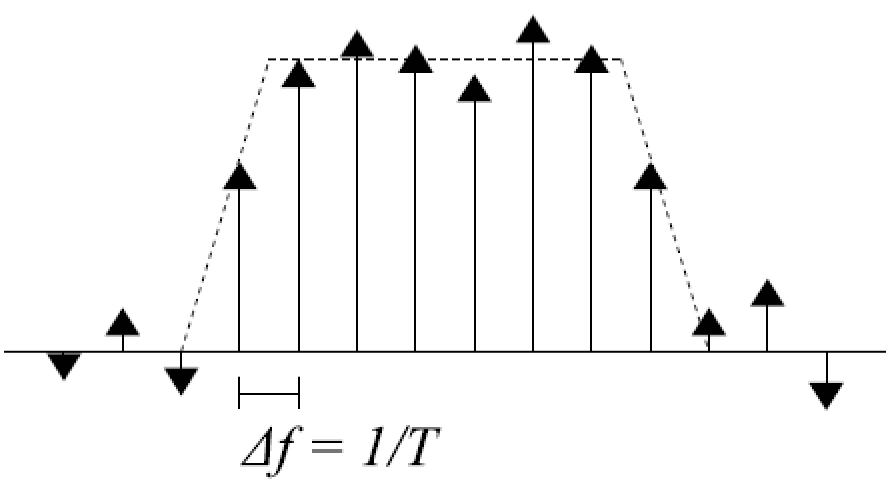
We'll ignore noise

$$\begin{aligned} x_n &= < y(t), \phi_n(t) > \\ &= \int y(t) e^{-j2\pi \frac{n}{T} t} dt \\ &= Y(f)|_{f=n/T} \\ &= \sum_{m=1}^{M} b_m S_m(f)|_{f=n/T} + w \end{aligned}$$

Really just take time domain measurements and hit with DFT matrix

$$\mathbf{x} \quad = \quad \mathrm{U}^H \mathbf{y}$$

time domain measurements



Signal recovery with sub-sampled Fourier matrices

$$\mathbf{y} = \mathbf{U}\mathbf{x} \qquad [\mathrm{U}]_{\ell,m} = \frac{1}{\sqrt{N}} e^{j2\pi \frac{(\ell-1)(m-1)}{N}}$$

Deterministic, s-sparse

Randomly throw out (*N-p*) time domain samples

$$\mathrm{U} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & \cdots & & 1 \\ 1 & e^{j2\pi\frac{1}{N}} & & & e^{j2\pi\frac{(N-1)}{N}} \\ \vdots & & \ddots & & \\ & & & e^{j2\pi\frac{(\ell-1)(m-1)}{N}} & \\ & & & & \ddots \\ 1 & e^{j2\pi\frac{(N-1)}{N}} & & & e^{j2\pi\frac{(N-1)^2}{N}} \end{bmatrix}_{[N\times N]}$$

$$\mathbf{y}_p = \mathrm{U}_p\mathbf{x}$$

frequency domain measurements

time domain measurements

ℓ_i is chosen at without replacement from *1,...,N*

$$\mathrm{U_p} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & e^{j2\pi\frac{(\ell_1-1)}{N}} & \cdots & e^{j2\pi\frac{(\ell_1-1)(m-1)}{N}} & \cdots & e^{j2\pi\frac{(\ell_1-1)(N-1)}{N}} \\ 1 & e^{j2\pi\frac{(\ell_2-1)}{N}} & \cdots & e^{j2\pi\frac{(\ell_2-1)(m-1)}{N}} & & e^{j2\pi\frac{(\ell_2-1)(N-1)}{N}} \\ \vdots & \vdots & & & & \vdots \\ 1 & e^{j2\pi\frac{(\ell_p-1)}{N}} & \cdots & e^{j2\pi\frac{(\ell_p-1)(m-1)}{N}} & & e^{j2\pi\frac{(\ell_p-1)(N-1)}{N}} \end{bmatrix}_{[p\times N]}$$

Compressed Sensing Approach $\mathbf{y}_p = \mathrm{U}_p\mathbf{x}$

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} ||\mathbf{x}||_1$$
$$\text{subject to} \quad \mathbf{y}_p = \mathrm{U}_p\mathbf{x}$$

Back projection + thresholding

1) Back projection $\tilde{\mathbf{x}} = \mathrm{U}_p^H \mathbf{y}_p = \mathrm{U}_p^H \mathrm{U}_p \mathbf{x}$

2) Threshold
$$\hat{x}_j = \begin{cases} \tilde{x}_j & \textit{if } \tilde{x}_j \textit{ is in s largest elements of } \tilde{\mathbf{x}} \\ 0 & \textit{else} \end{cases}$$

1-sparse signal

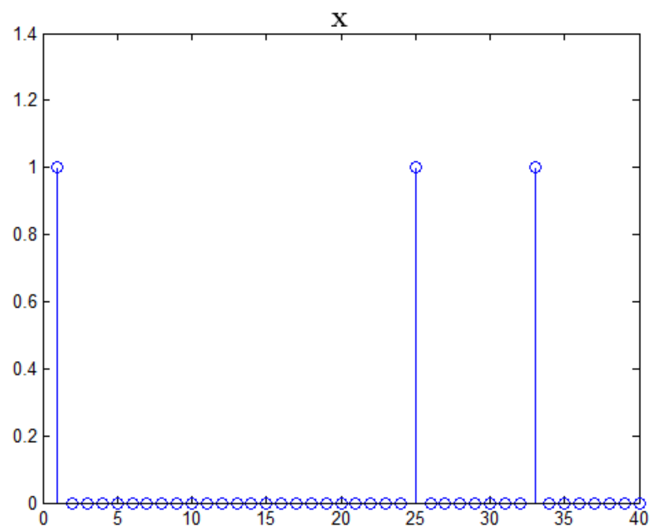
$$Pr(\hat{\mathbf{x}} = \mathbf{x}) = (1 - e^{-p})^{N-1}$$
$$p > \log(N-1)$$

*s-sparse signal**

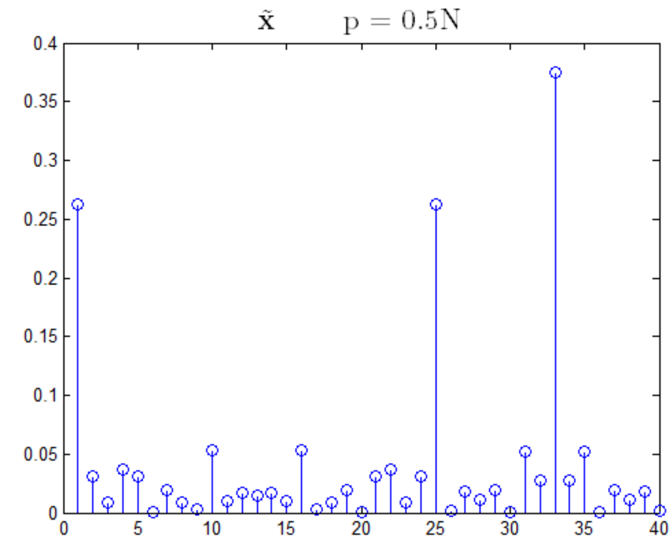
$$Pr(\hat{\mathbf{x}} = \mathbf{x}) \geq (1 - e^{-\frac{p}{4s}})^N$$
$$p > 4s \log N$$

*equal magnitude sparse entries

Example

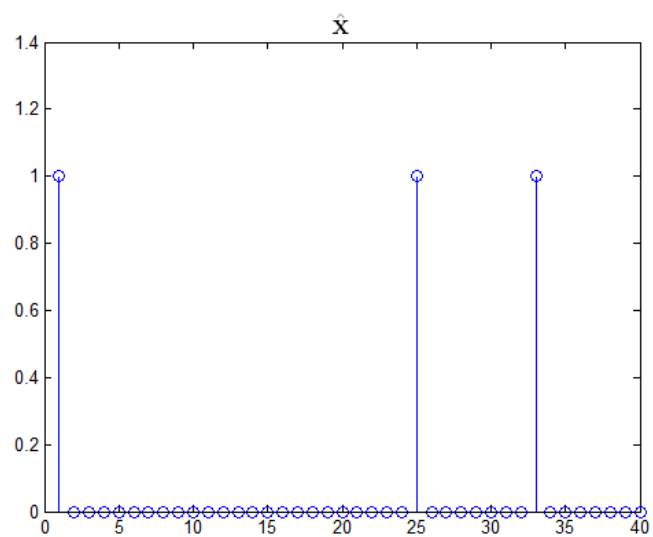


$\rightarrow \mathbf{y}_p = \mathrm{U}_p \mathbf{x} \longrightarrow \tilde{\mathbf{x}} = \mathrm{U}_p^H \mathbf{y}_p$



Potential Issues
- *s* must be known for thresholding
- dynamic range issue
- noise analysis?

Potential Advantages
- Deterministic time



Related work:

[1] Necessary and Sufficient Conditions for Sparsity Pattern Recovery [Fletcher, Rangan, Goyal, 2009]

- Same technique with Gaussian matrices

$$\tilde{\mathbf{x}} = \mathrm{A}_p^T \mathrm{y}_p \qquad [\mathrm{A}]_{i,j} \sim \mathcal{N}\left(0, \frac{1}{N}\right)$$

- *Result* $p > \frac{8(1+||\mathbf{x}||^2)}{x_{min}^2} \log(N-s)$

account dynamic range issue

$$p > 8(1+s)\log(N-s)$$

OMP – Orthogonal matching pursuit algorithms

Uses back projection to find largest element of **x**, remove contribution, and repeat

Trivial case when p = n

$$\tilde{\mathbf{x}} = \mathrm{U}_p^H \mathrm{U}_p \mathbf{x}$$

$$\tilde{\mathbf{x}} = \mathrm{I}\mathbf{x}$$

Interesting when $p < N$

$$\mathrm{U_p}^H \mathrm{U_p} = \begin{bmatrix} \frac{p}{N} & & & small \\ & \frac{p}{N} & & \\ & & \ddots & \\ small & & & \frac{p}{N} \end{bmatrix}$$

$$\tilde{\mathbf{x}} = \mathrm{U_p}^H \mathrm{U_p} \mathbf{x} = \begin{bmatrix} \frac{p}{N} & & & small \\ & \frac{p}{N} & & \\ & & \ddots & \\ small & & & \frac{p}{N} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ x_s \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Elements of $\mathrm{U}_p^H \mathrm{U}_p$

$$[\mathrm{U_p}^H \mathrm{U_p}]_{u,v} = \frac{1}{N}\sum_{i=1}^{p} e^{j2\pi\frac{(\ell_i-1)(u-1)}{N}} e^{-j2\pi\frac{(\ell_i-1)(v-1)}{N}}$$
$$= \frac{1}{N}\sum_{i=1}^{p} e^{j2\pi\frac{(\ell_i-1)(u-v)}{N}}$$

Diagonal elements

$$[\mathrm{U_p}^H \mathrm{U_p}]_{u,u} = \frac{p}{N}$$

Off diagonal elements

Convergence in distribution for large N

$$\frac{\ell_i - 1}{N} \longrightarrow U_i \sim \mathrm{uniform}[0,1)$$

CLT for large~ish p

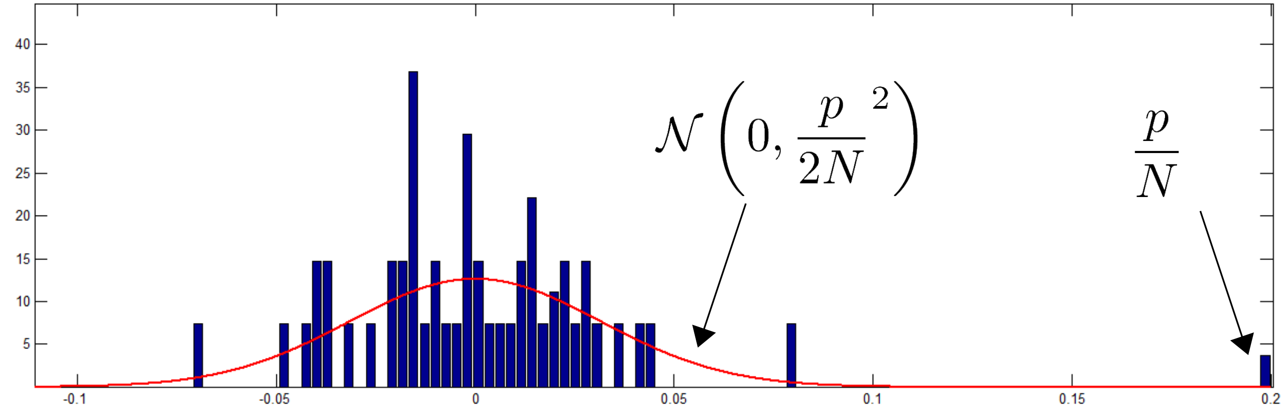
$$[\mathrm{U_p}^H \mathrm{U_p}]_{u,v} \longrightarrow \frac{1}{N}\sum_{i=1}^{p} e^{j2\pi U_i(u-v)} \longrightarrow \mathcal{CN}(0,\sigma^2)$$

$$[\mathrm{U_p}^H \mathrm{U_p}]_{u,v} \longrightarrow \mathcal{CN}\left(0, \frac{p}{N^2}\right)$$

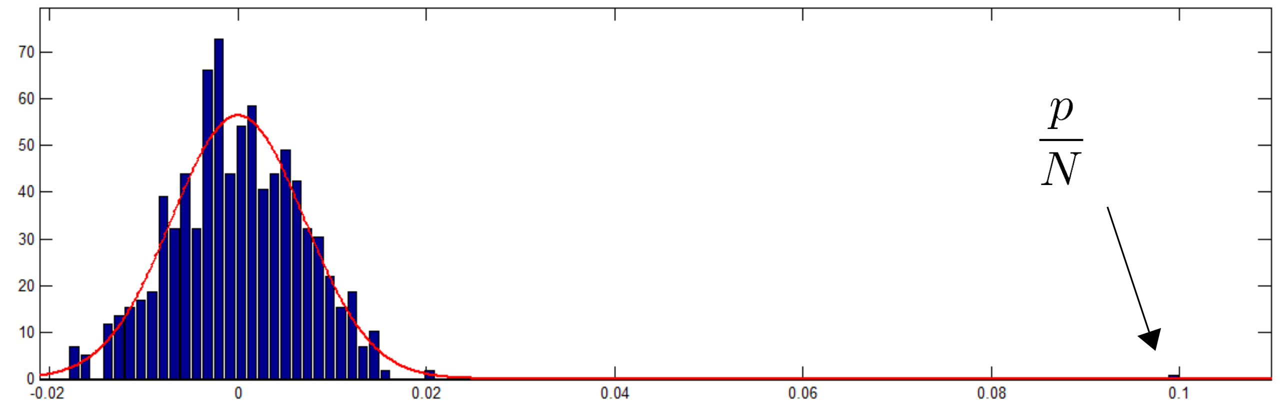
$$\sigma^2 = E\left[\left(\frac{1}{N}\sum_{i=1}^{p} e^{j2\pi U_i}\right)\left(\frac{1}{N}\sum_{i=1}^{p} e^{-j2\pi U_i}\right)\right]$$
$$= \frac{1}{N^2} E\left[\sum_{i=1}^{p} e^{j2\pi U_i} e^{-j2\pi U_i}\right]$$
$$= \frac{p}{N^2}$$

Increasing N

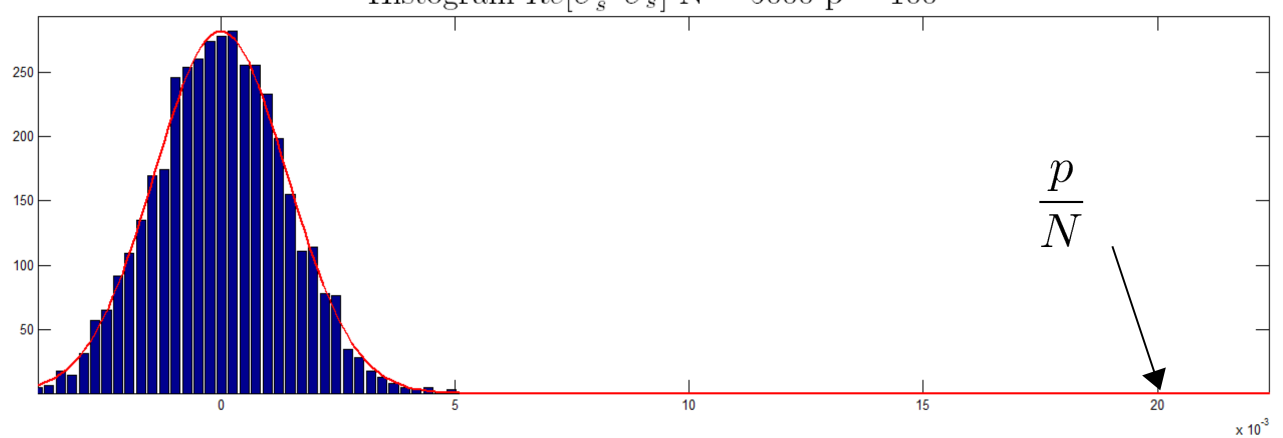
Histogram $Re[U_s^H U_s]$ N = 100 p = 20



Histogram $Re[U_s^H U_s]$ N = 1000 p = 100



Histogram $Re[U_s^H U_s]$ N = 5000 p = 100



Consider 1-sparse case

$$\tilde{\mathbf{x}} = \mathrm{U_p}^H \mathrm{U_p} \mathbf{x} = \begin{bmatrix} \frac{p}{N} & & & \sim \mathcal{CN}(0, \frac{P}{N^2}) \\ & \frac{p}{N} & & \\ & & \ddots & \\ \sim \mathcal{CN}(0, \frac{P}{N^2}) & & & \frac{p}{N} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ x_s \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$\tilde{\mathbf{x}}$ is just s[th] row

$$\tilde{\mathbf{x}} = \begin{bmatrix} x_s \times small \\ \vdots \\ x_s \frac{P}{N} \\ \vdots \\ x_s \times small \end{bmatrix} = \begin{bmatrix} x_s \times \sim \mathcal{CN}(0, \frac{P}{N^2}) \\ \vdots \\ x_s \frac{P}{N} \\ \vdots \\ x_s \times \sim \mathcal{CN}(0, \frac{P}{N^2}) \end{bmatrix}$$

Thresholding step – pick largest entry (in 1-sparse case)

$$\hat{x}_j = \begin{cases} \frac{N}{p}\tilde{x}_j & |\tilde{x}_j|^2 > |\tilde{x}_i|^2 \quad \forall \quad i \neq j \\ 0 & else \end{cases}$$

1-sparse error analysis

Probability of error is just probability that any of *N-1* 'noise' elements exceed *p/N*

independent of magnitude of non-zero entry

$$Pr(\hat{x}_i = x_i) = Pr\left(|x_s|^2 \underbrace{|[\mathrm{U_p}^H \mathrm{U_p}]_{s,i}|^2}_{Z} < |x_s|^2 \left(\frac{p}{N}\right)^2 \right) \qquad i \neq s$$

Error probability for 1 entry

$$Pr(\hat{x}_i = x_i) = Pr\left(Z < \left(\frac{p}{N}\right)^2 \right)$$

$$Z \sim \chi_2^2 \sim \exp\left(\frac{N^2}{p}\right)$$

$$F_Z(z) = 1 - e^{-\frac{N^2}{p} z}$$

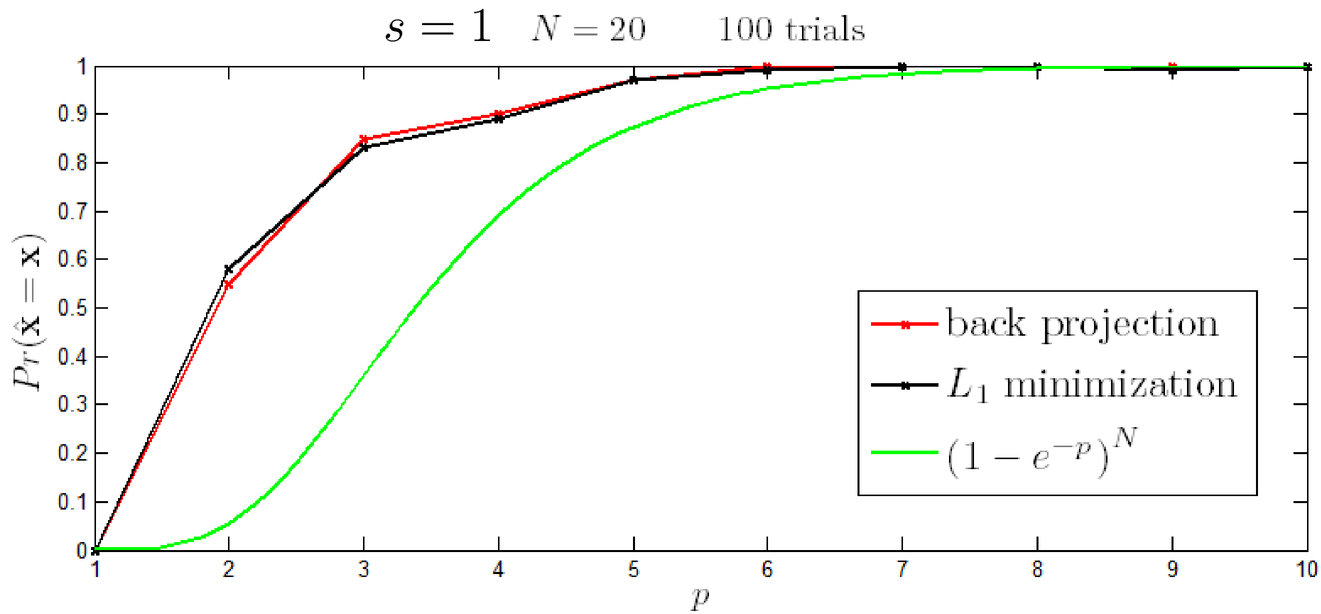
$$\begin{aligned} Pr(\hat{x}_i = x_i) &= 1 - e^{-\left(\frac{N^2}{p}\right)\left(\frac{p}{N}\right)^2} \\ Pr(\hat{x}_i = x_i) &= 1 - e^{-p} \end{aligned}$$

Error probability for entire signal

$$Pr(\hat{\mathbf{x}} = \mathbf{x}) = (1 - e^{-p})^{N-1}$$

$$p > \log(N-1)$$

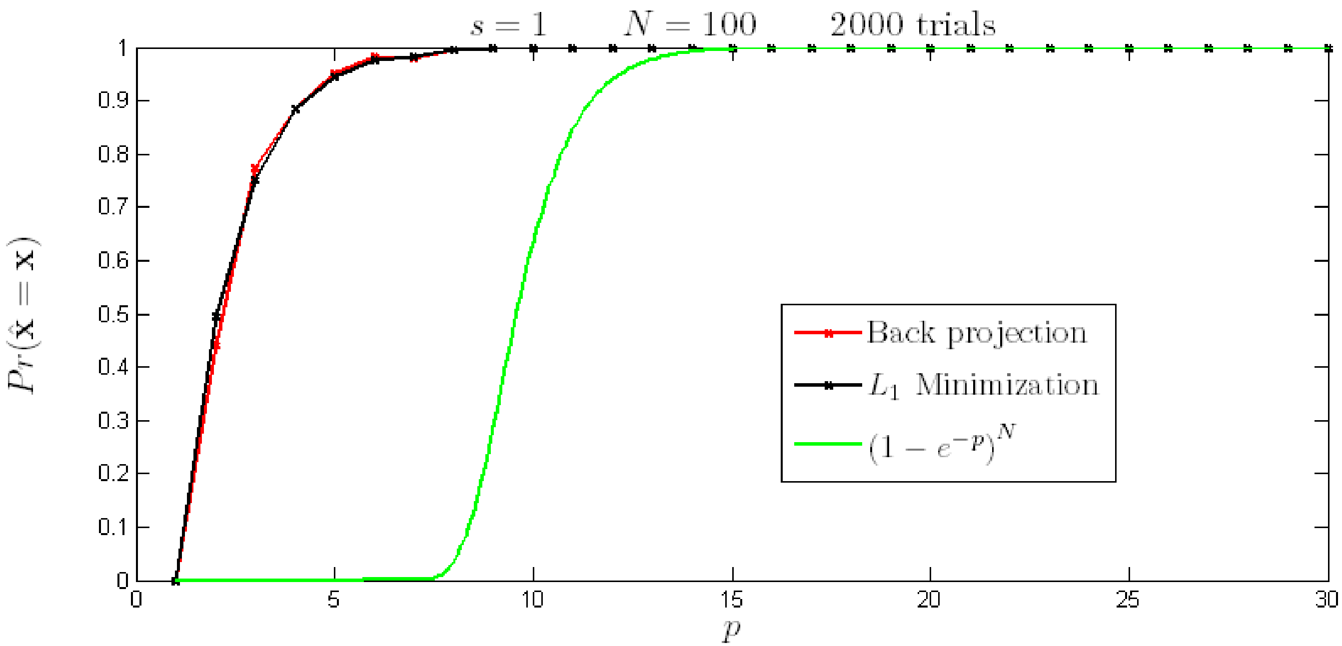
Numerical results – $s = 1$



Average L1 CPU time

Average back projection CPU time

$$\bar{t}_{cs} = 0.1136s$$
$$\bar{t}_{bp} = 4.56 \times 10^{-5}s$$



$$\bar{t}_{cs} = 0.3976s$$
$$\bar{t}_{bp} = 1.3807 \times 10^{-4}s$$

s-sparse error analysis

Set of indices where x lives

$$I_s = \{i : x_i \neq 0\}$$

After thresholding, error occurs if worst 'no signal' entry exceeds magnitude of worst 'signal' entry

$$\tilde{x}_i \quad \sim \quad \begin{cases} \mathcal{CN}\left(0, \frac{sP}{N^2}\right) & i \notin I_s \\ \mathcal{CN}\left(\frac{P}{N}, \frac{sP}{N^2}\right) & i \in I_s \end{cases}$$

$$P_E = Pr(\hat{\mathbf{x}} \neq \mathbf{x}) = Pr\left(\max_{i \notin I_s} |\tilde{x}_i|^2 > \min_{i \in I_s} |\tilde{x}_i|^2\right)$$

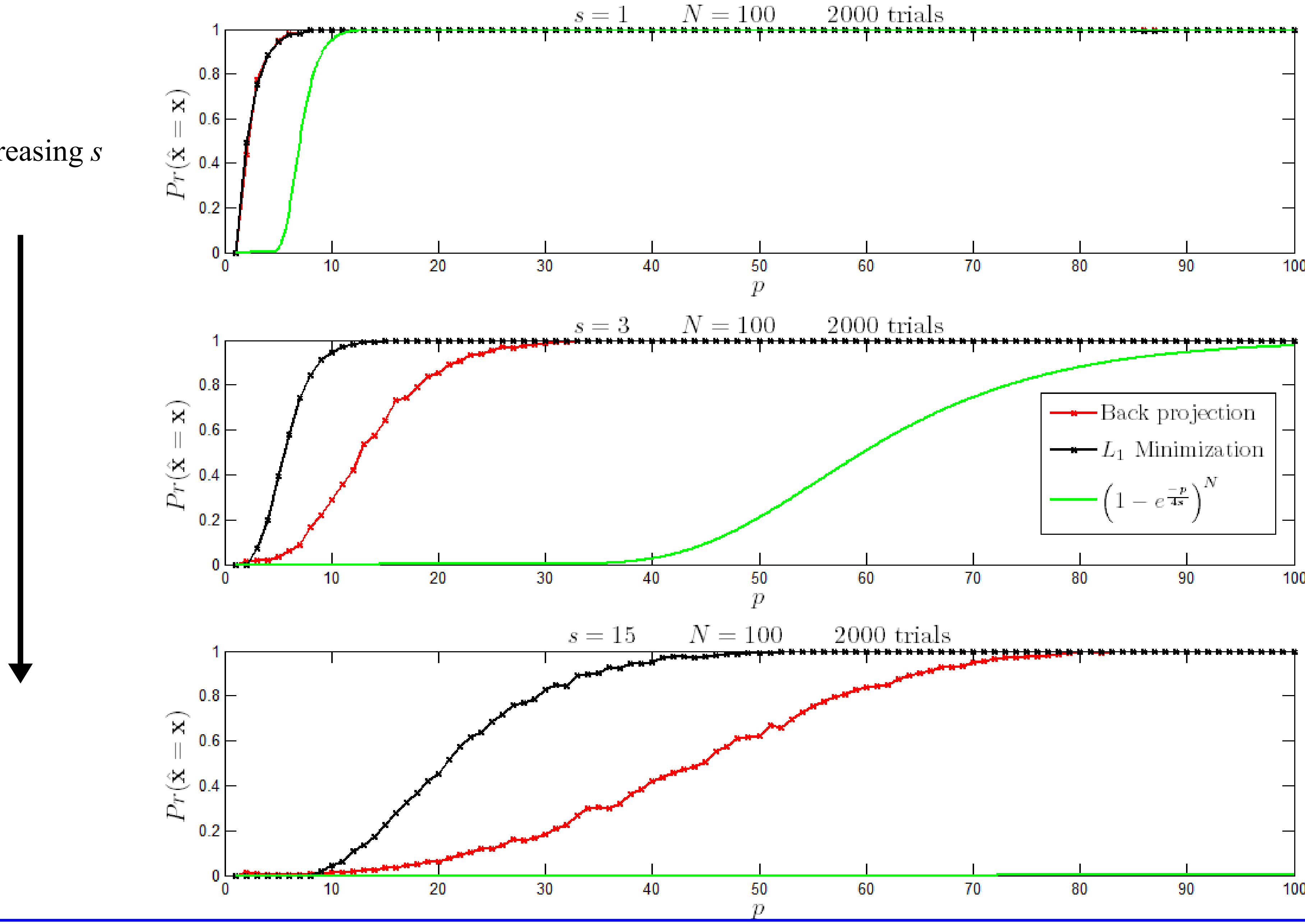
Define a suboptimal threshold at $\dfrac{P}{2N}$

$$\begin{aligned} Pr(\hat{\mathbf{x}} = \mathbf{x}) = 1 - P_E = \quad & \geq \quad Pr\left(\bigcap_{i \notin I_s} \left[|\tilde{x}_i|^2 < \left(\frac{P}{2N}\right)^2\right] \cap \bigcap_{i \in I_s} \left[|\tilde{x}_i|^2 > \left(\frac{P}{2N}\right)^2\right]\right) \\ & = \quad Pr\left(\left|\underbrace{Y}_{\mathcal{CN}\left(0, \frac{sP}{N^2}\right)}\right|^2 < \left(\frac{P}{2N}\right)^2\right)^N \\ & = \quad \left(1 - e^{-\frac{p}{4s}}\right)^N \end{aligned}$$

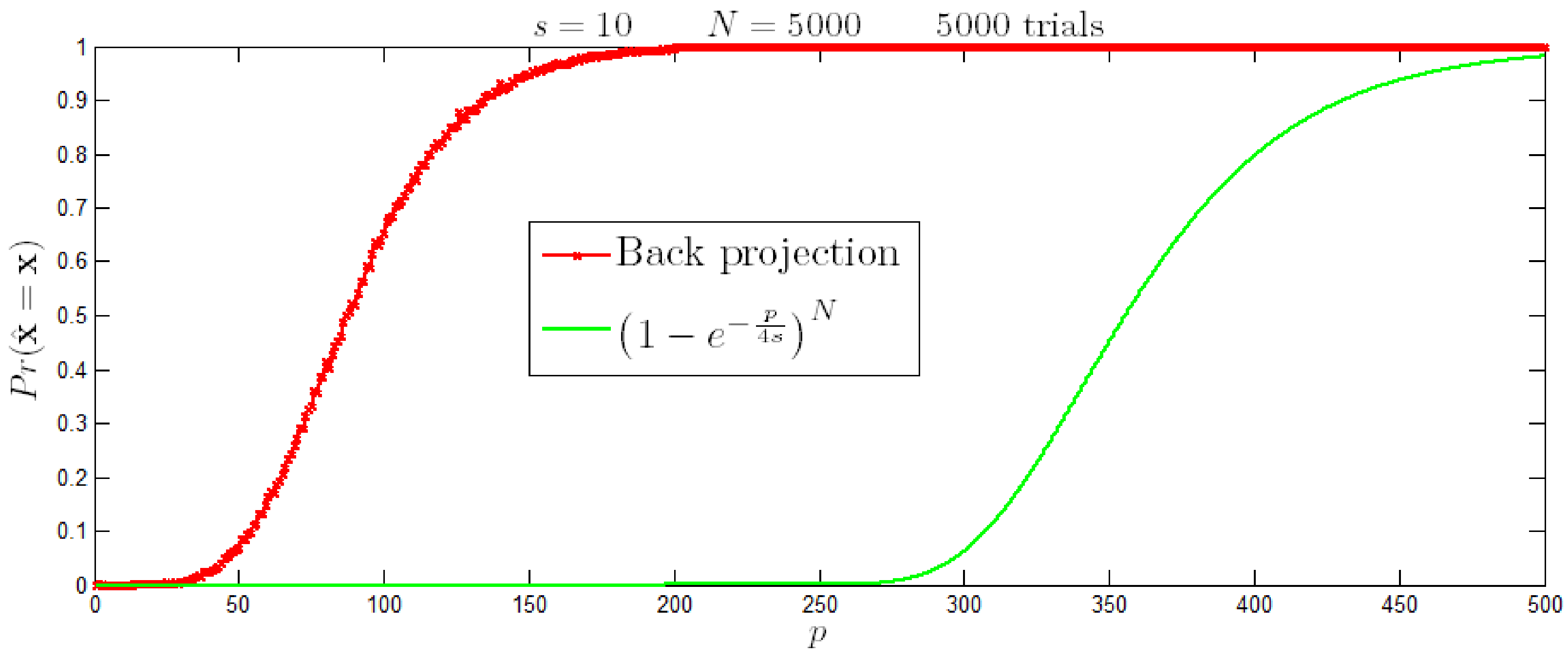
$$Pr(\hat{\mathbf{x}} = \mathbf{x}) \geq \left(1 - e^{-\frac{p}{4s}}\right)^N$$

Numerical Results – Increasing *s*

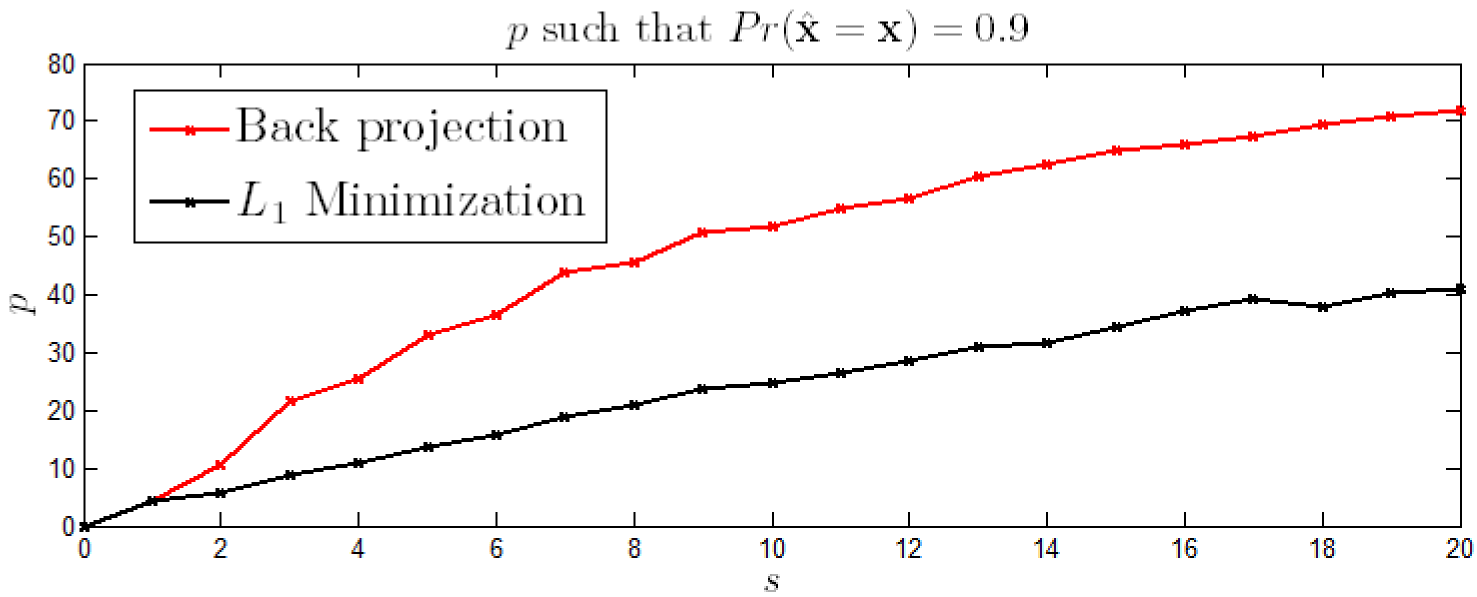
Increasing *s*



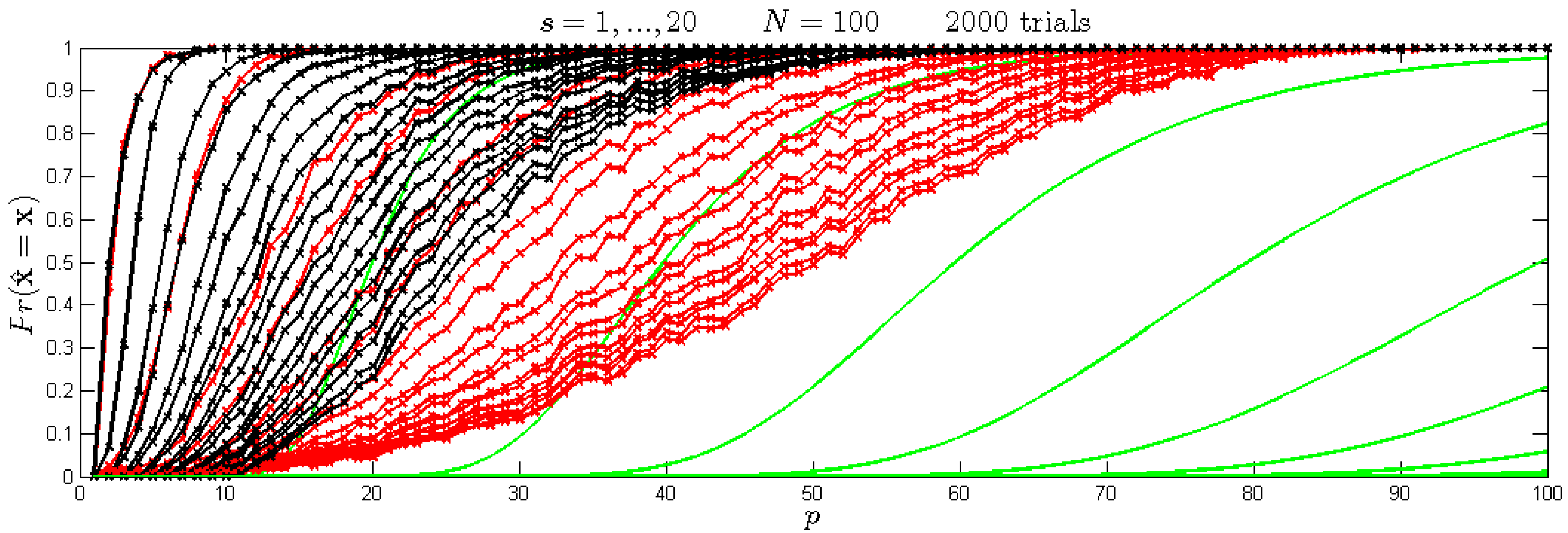
Numerical Results – Large N



Questions?



p such that $Pr(\hat{\mathbf{x}} = \mathbf{x}) = 0.9$
Back projection
L_1 Minimization
p
s
0 10 20 30 40 50 60 70 80
0 2 4 6 8 10 12 14 16 18 20



$s = 1, \ldots, 20$ $N = 100$ 2000 trials
$Pr(\hat{\mathbf{x}} = \mathbf{x})$
p
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
0 10 20 30 40 50 60 70 80 90 100