Structured Sparsity Models for Compressive Sensing

Chinmay Hegde

Rice University

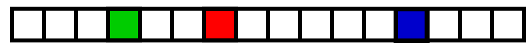
Volkan Cevher

Richard Baraniuk

Marco Duarte

Concise Signal Model: Sparsity

- **Sparse** signal:



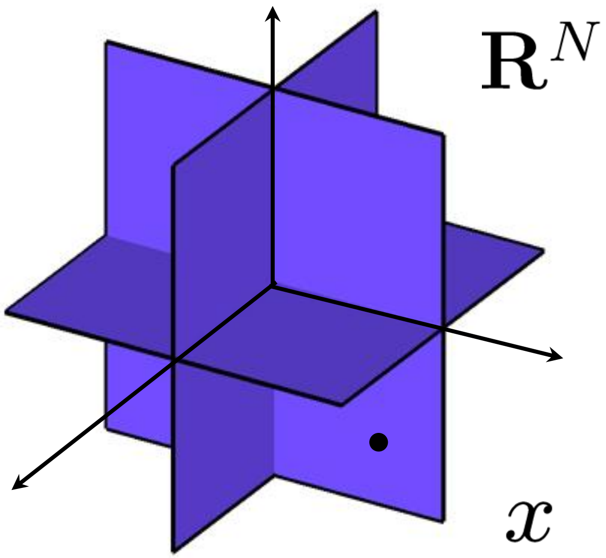
  - only *K* out of *N* coordinates nonzero

Concise Signal Model: Sparsity

- **Sparse** signal:

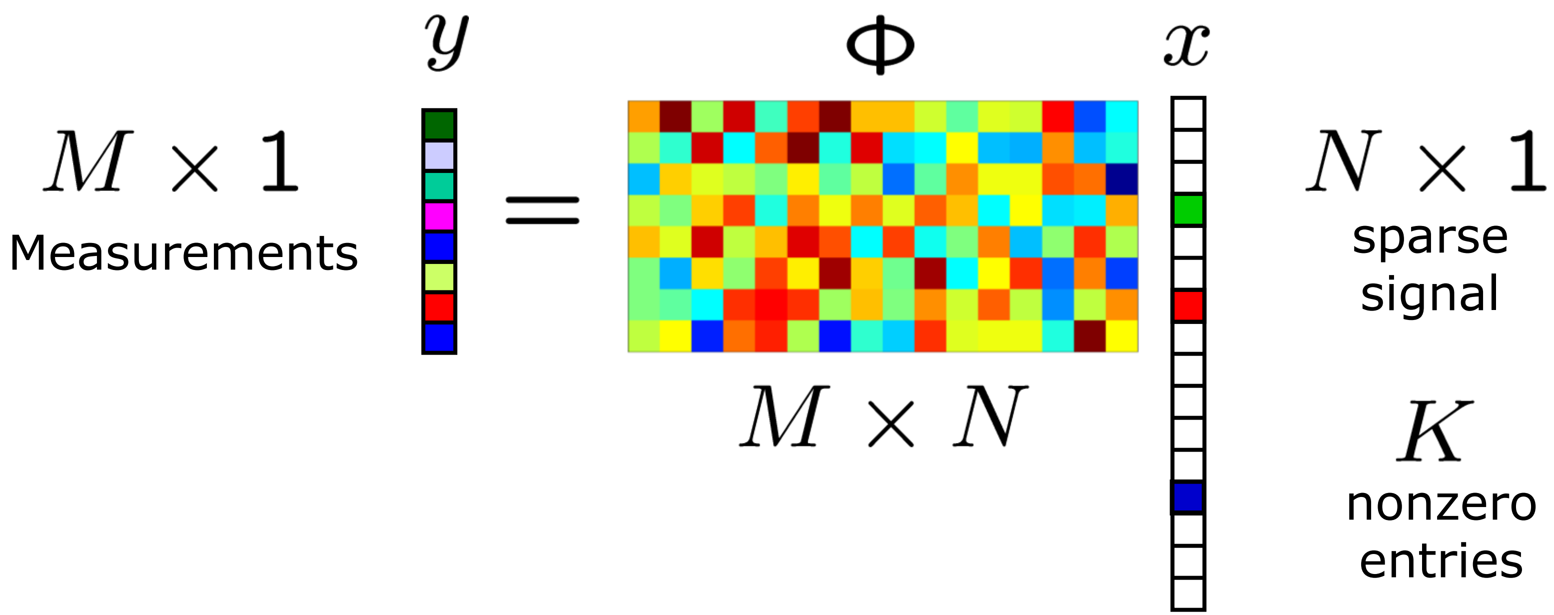
  - only *K* out of *N* coordinates nonzero

- Geometry: *union* of *K*-dimensional subspaces aligned w/ coordinate axes



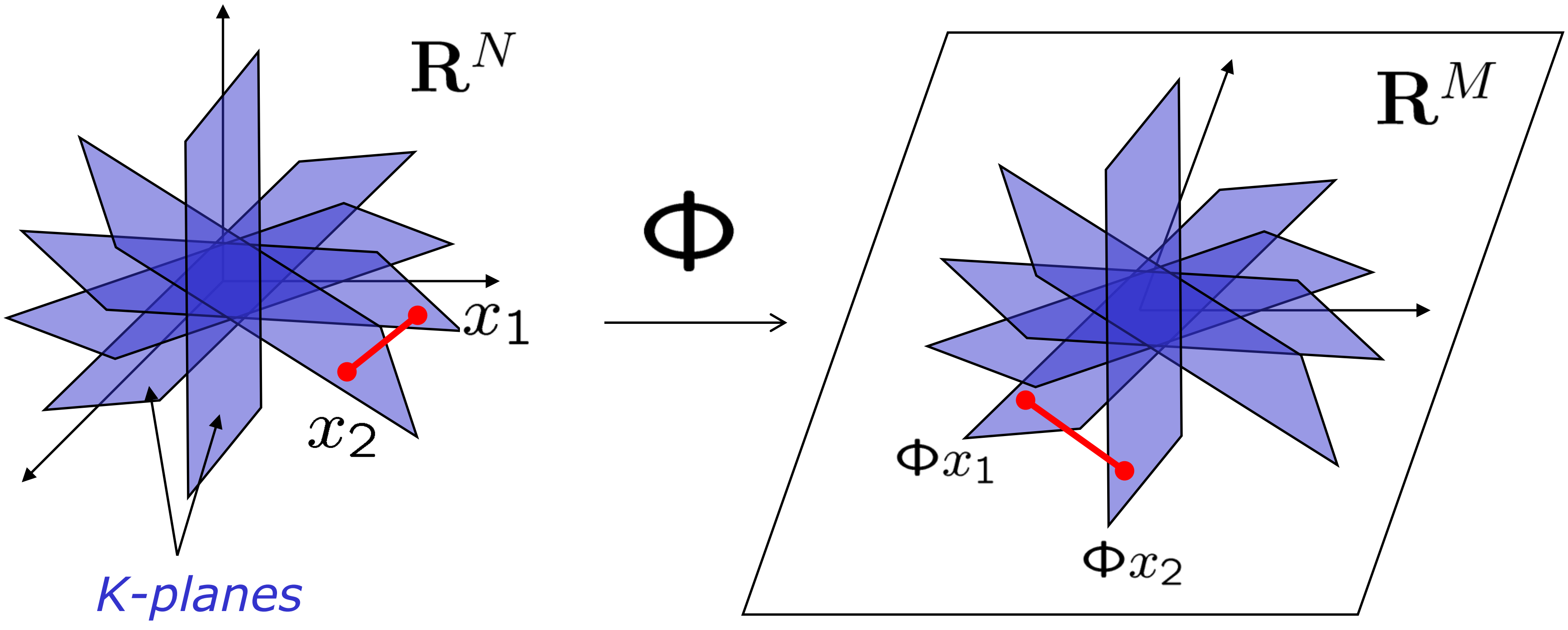
Compressive Sensing

- **Sampling** via dimensionality reduction

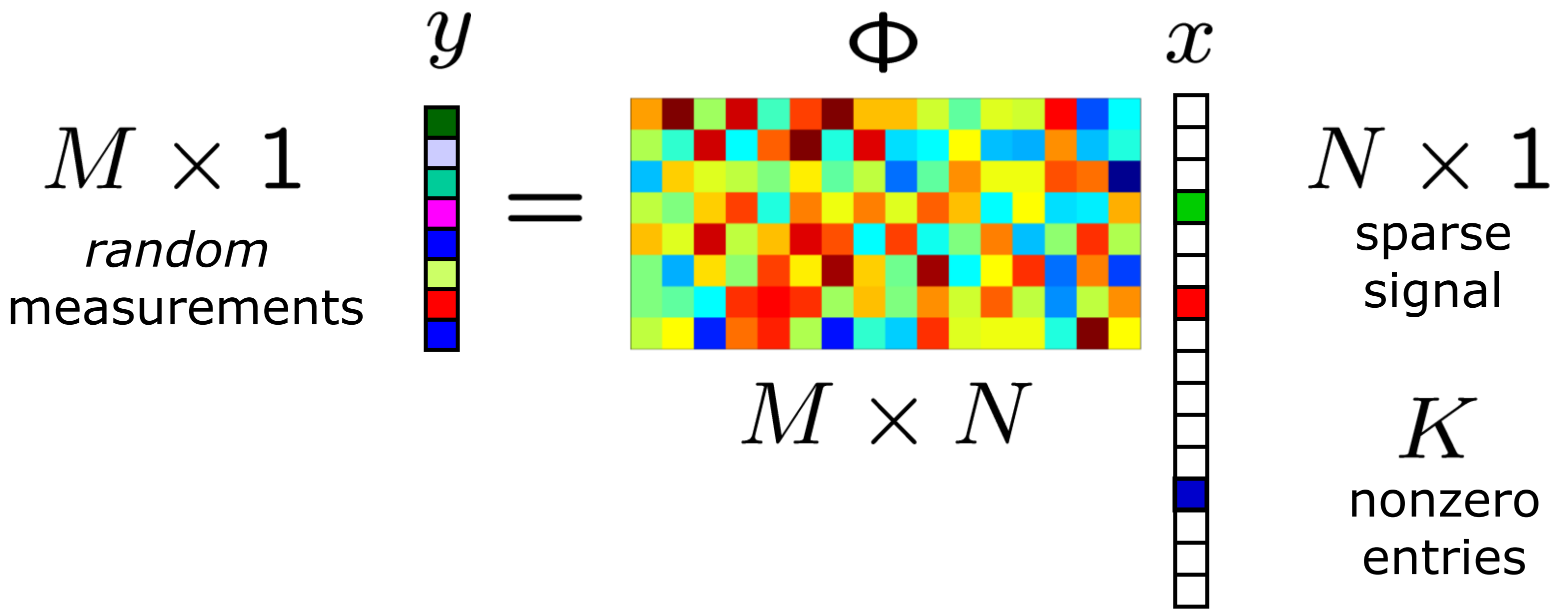


Restricted Isometry Property (RIP)

- Preserve the structure of sparse signals



Compressive Sensing

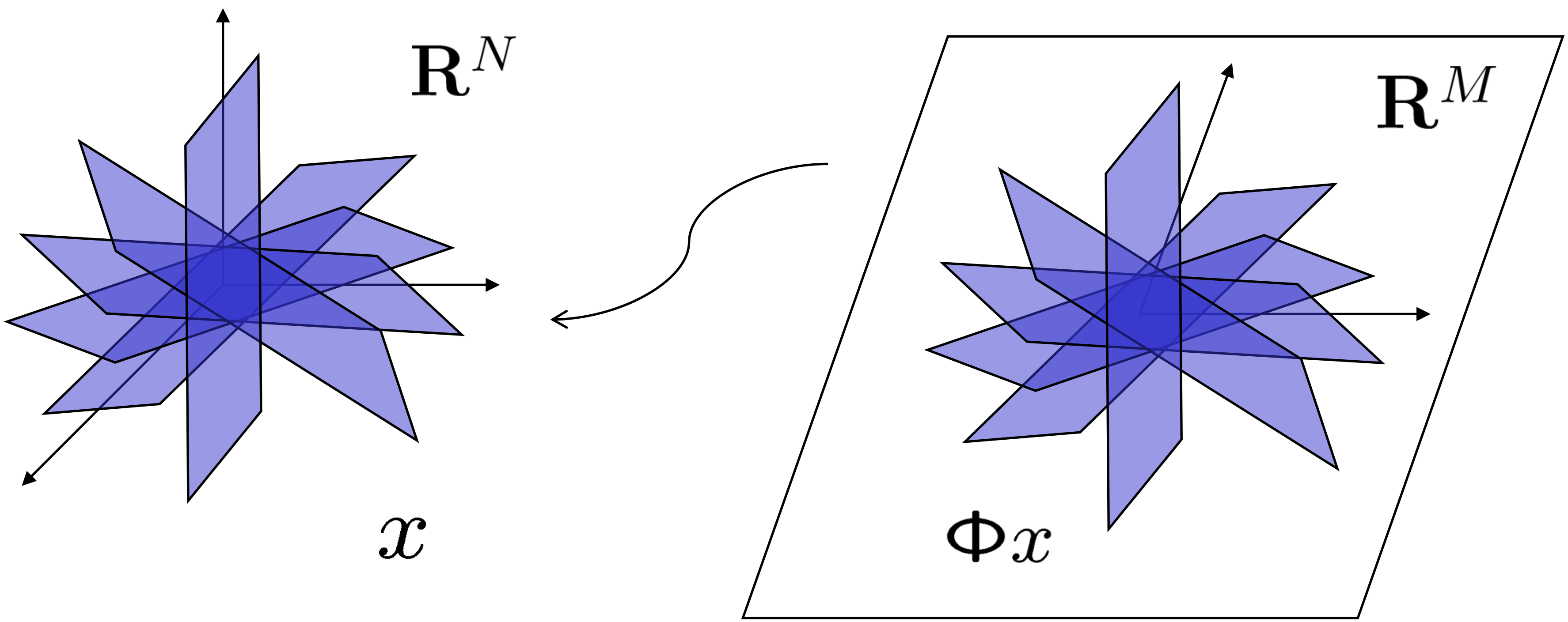


- *Random* subgaussian matrix Φ has the **RIP** whp if

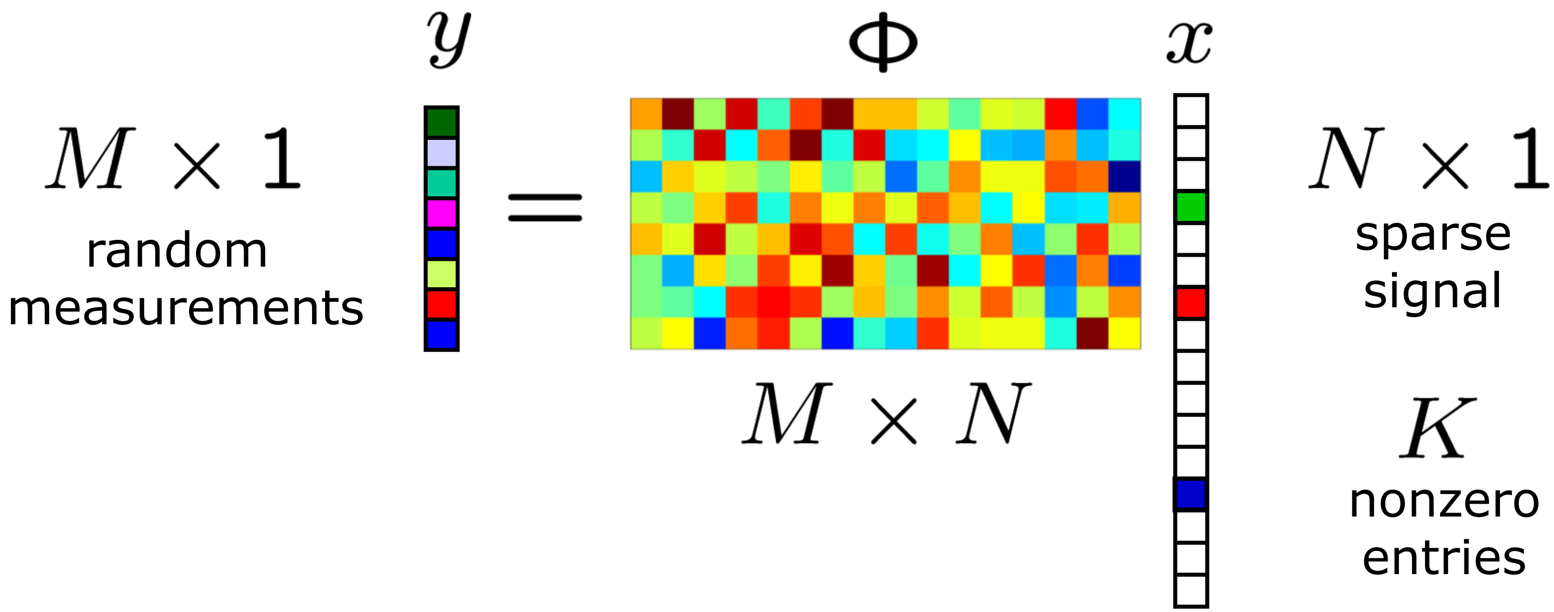
$$M = O(K \log(N/K))$$

Stable Recovery

- Efficient, stable algorithms that give back signal



Compressive Sensing



- ℓ_1-optimization
  [C, R, T]; [D]; [F,W,N]; [H,Y,Z]

- Greedy algorithms
  - OMP [G, T]
  - iterated thresholding [N, F]; [D, D, DeM]; [B, D]
  - CoSaMP [N,T]; Subspace Pursuit [D,M]

Iterated Thresholding

goal: given $y = \Phi x$, recover a sparse x

initialize: $\widehat{x}_0 = 0$, $r = y$, $i = 0$

iteration:

- $i \leftarrow i + 1$
- $b \leftarrow \widehat{x}_{i-1} + \Phi^T r$ **update signal estimate**
- $\widehat{x}_i \leftarrow \mathrm{thresh}(b, K)$ **prune signal estimate** (best *K*-term approx)
- $r \leftarrow y - \Phi \widehat{x}_i$ **update residual**

return: $\widehat{x} \leftarrow \widehat{x}_i$

Performance

- **Sparse signals**
  - noise-free measurements: exact recovery
  - noisy measurements: stable recovery

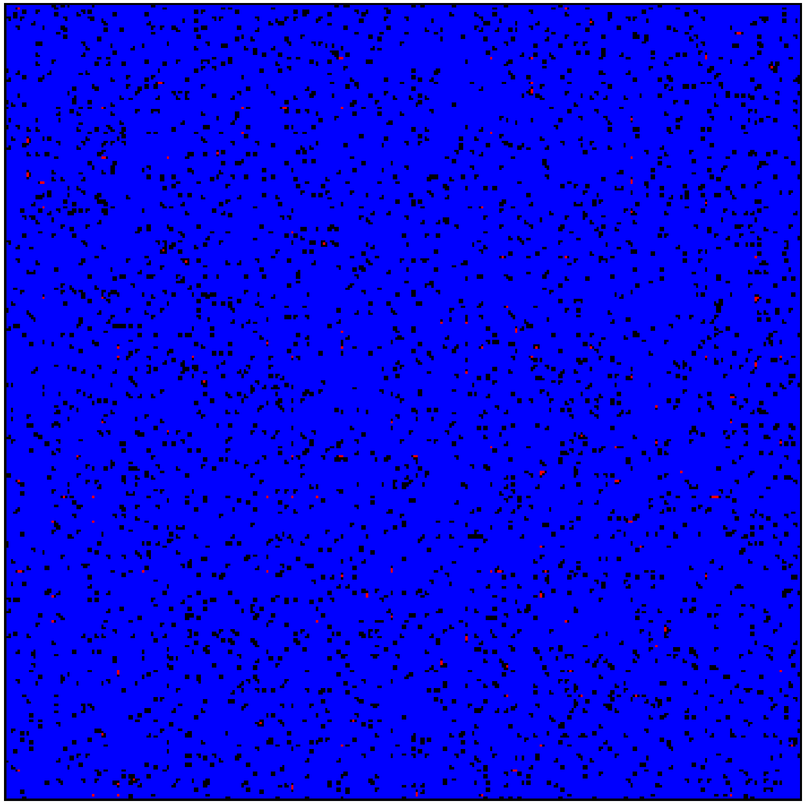
- **Compressible signals**
  - recovery as good as best *K*-sparse approximation

$$\underbrace{\|x - \widehat{x}\|_{\ell_2}}_{\text{CS recovery error}} \leq C_1 \underbrace{\|x - x_K\|_{\ell_2}}_{\text{signal } K\text{-term approx error}} + C_2 \underbrace{\frac{\|x - x_K\|_{\ell_1}}{K^{1/2}}}_{\text{signal } K\text{-term approx error}} + C_3 \underbrace{\epsilon}_{\text{noise}}$$

From Sparsity
to
Structured Sparsity

Beyond Sparse Models

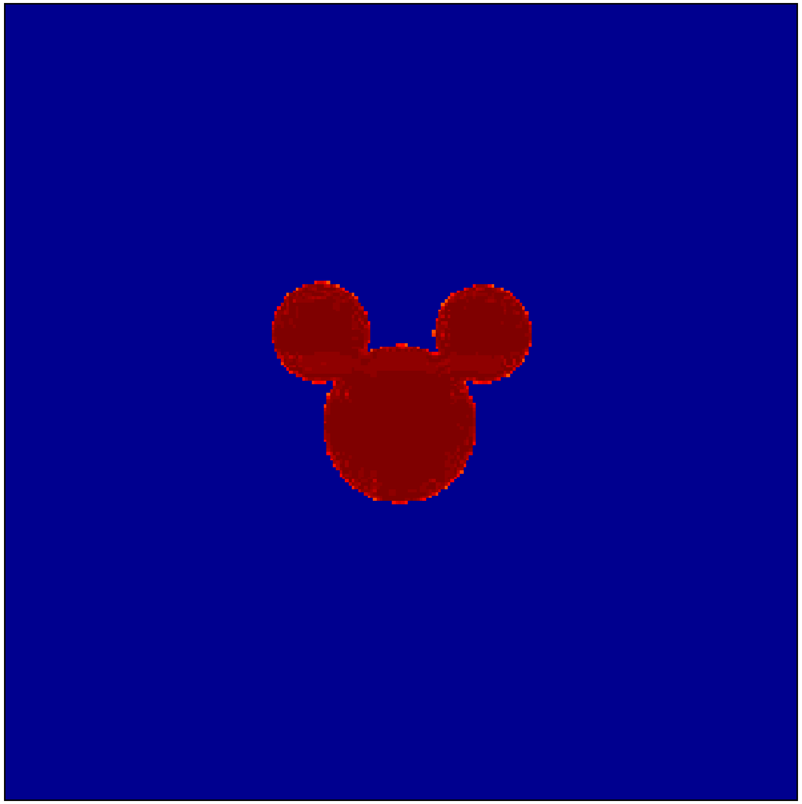
- Sparsity captures **simplistic primary structure**



5% sparse image

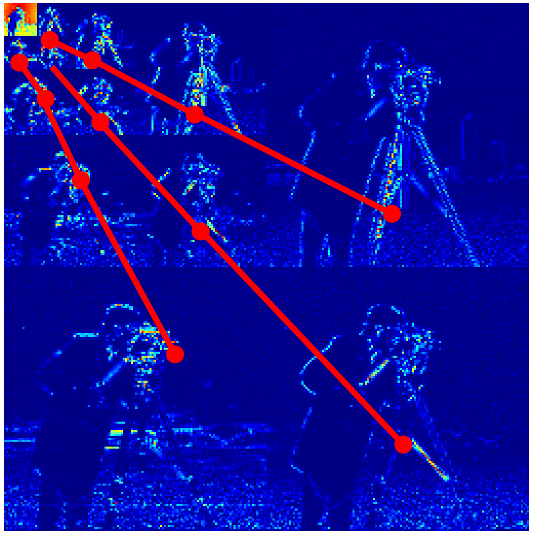
Beyond Sparse Models

- Most real-world apps exhibit **additional structure**

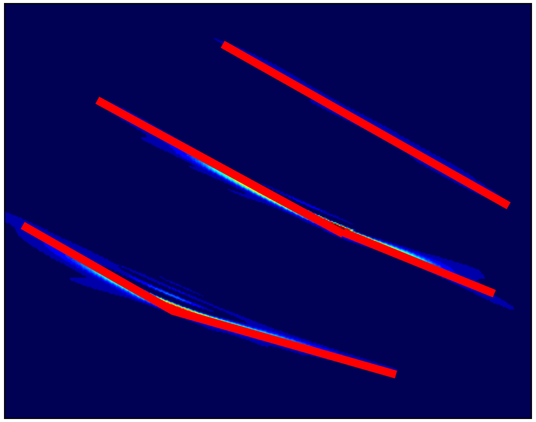


5% sparse image

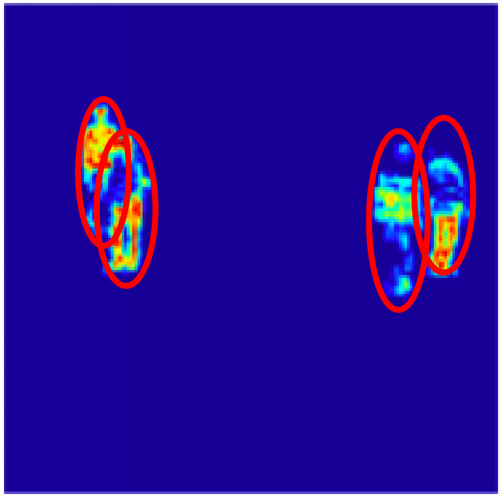
Beyond Sparse Models



wavelets:
natural images



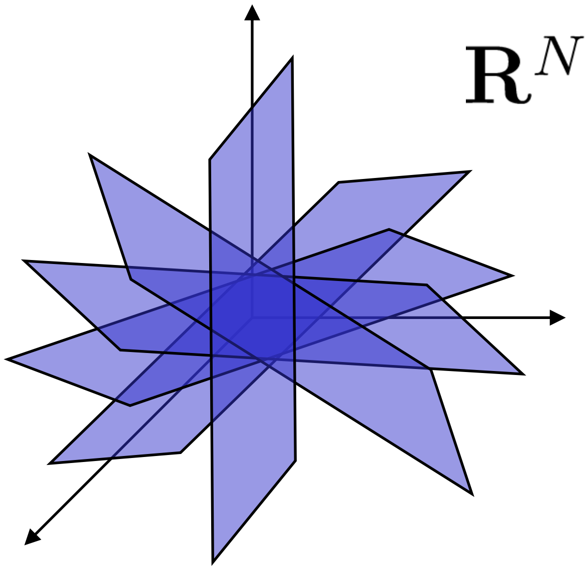
Gabor atoms:
chirps/tones

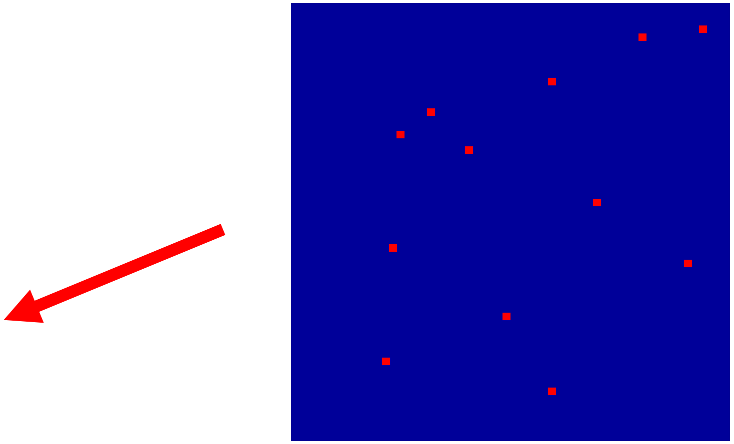


pixels:
background subtracted
images

Sparse Signals

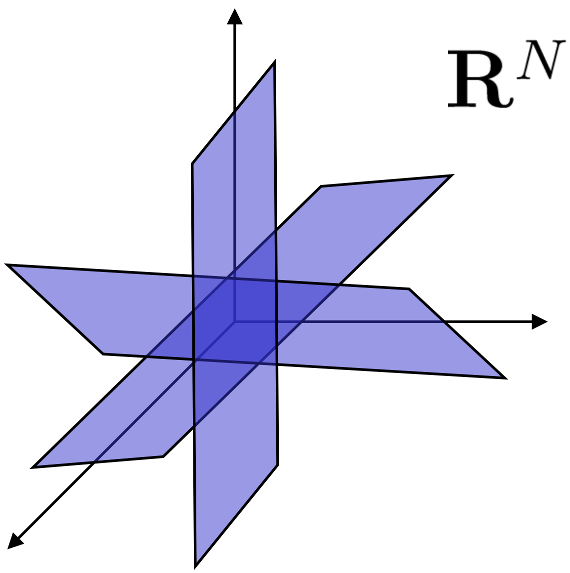
- Defn: ***K*-sparse signals** comprise *all* *K*-dimensional canonical subspaces

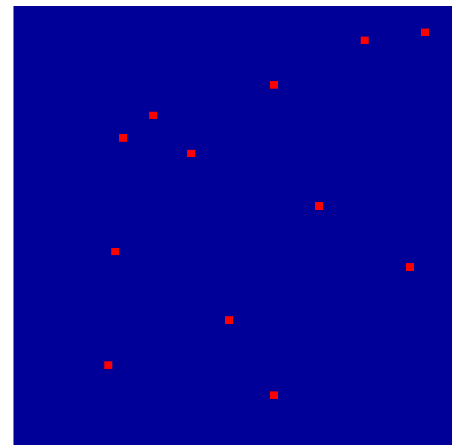


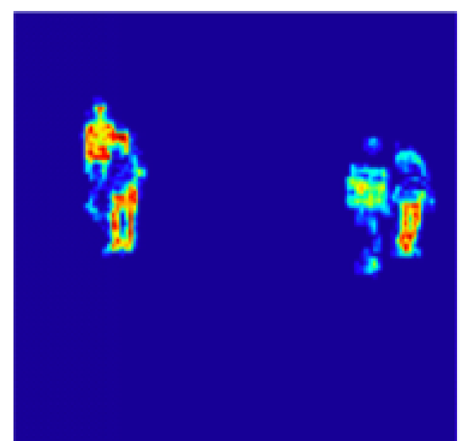


Model-Sparse Signals

- Defn: A ***K*-sparse signal *model*** comprises a particular (*reduced*) set of K-dim canonical subspaces [B, D], [L, D]

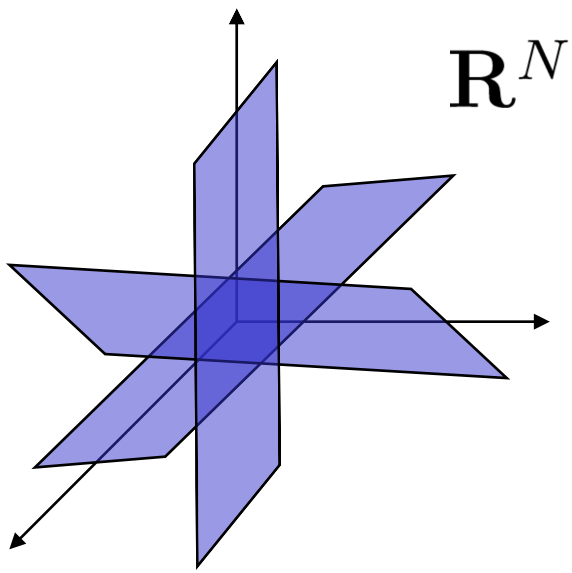




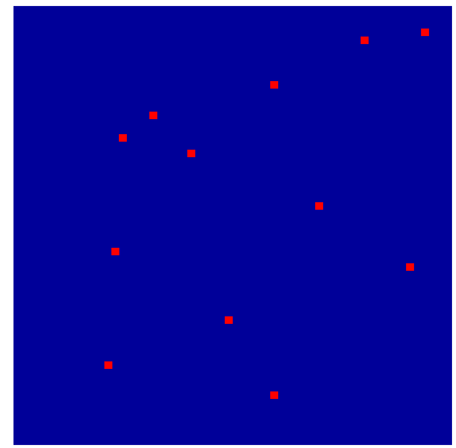


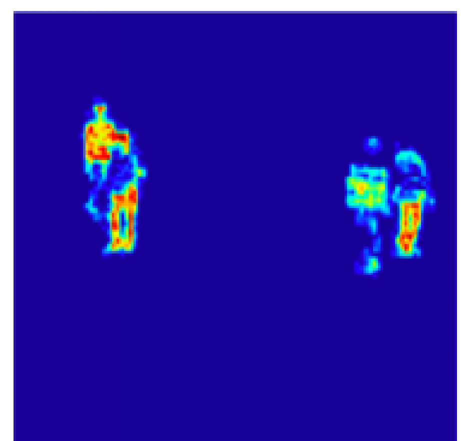
Model-Sparse Signals

- Defn: A ***K*-sparse signal *model*** comprises a particular (*reduced*) set of *K*-dim canonical subspaces [B, D], [L, D]



$$x^* \leftarrow \mathcal{M}(x, K)$$

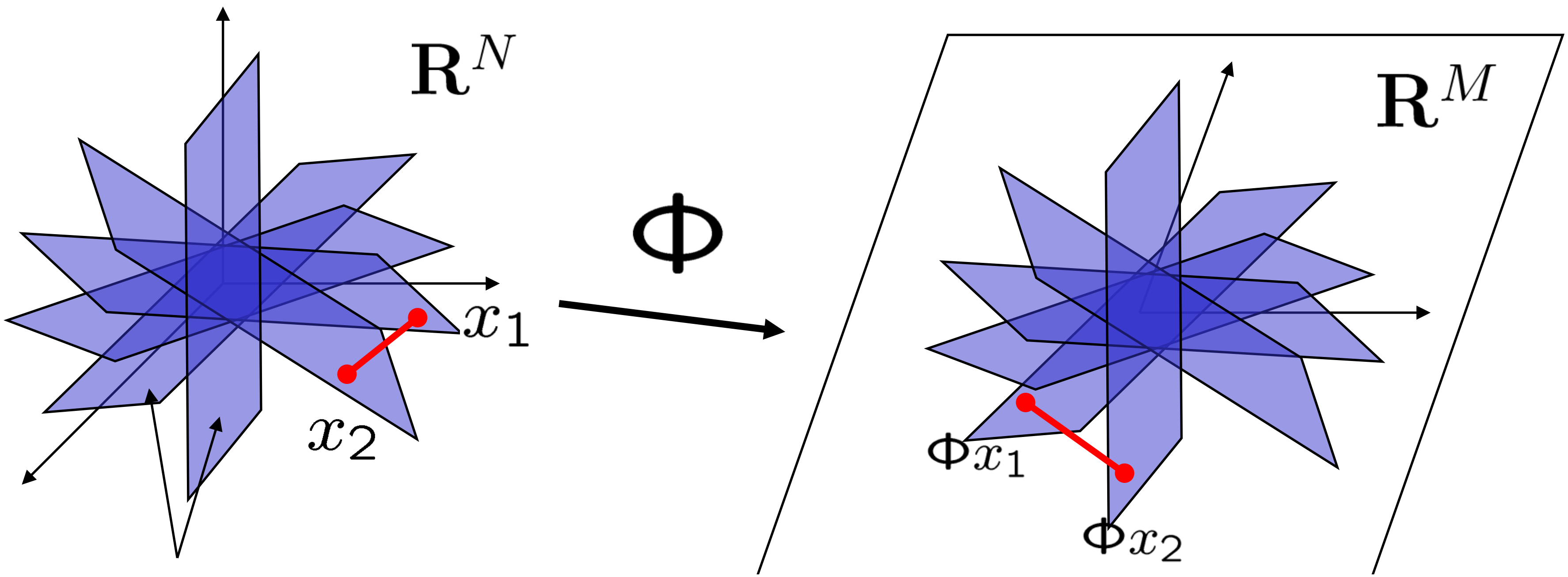




Model-based CS

Sampling Bounds

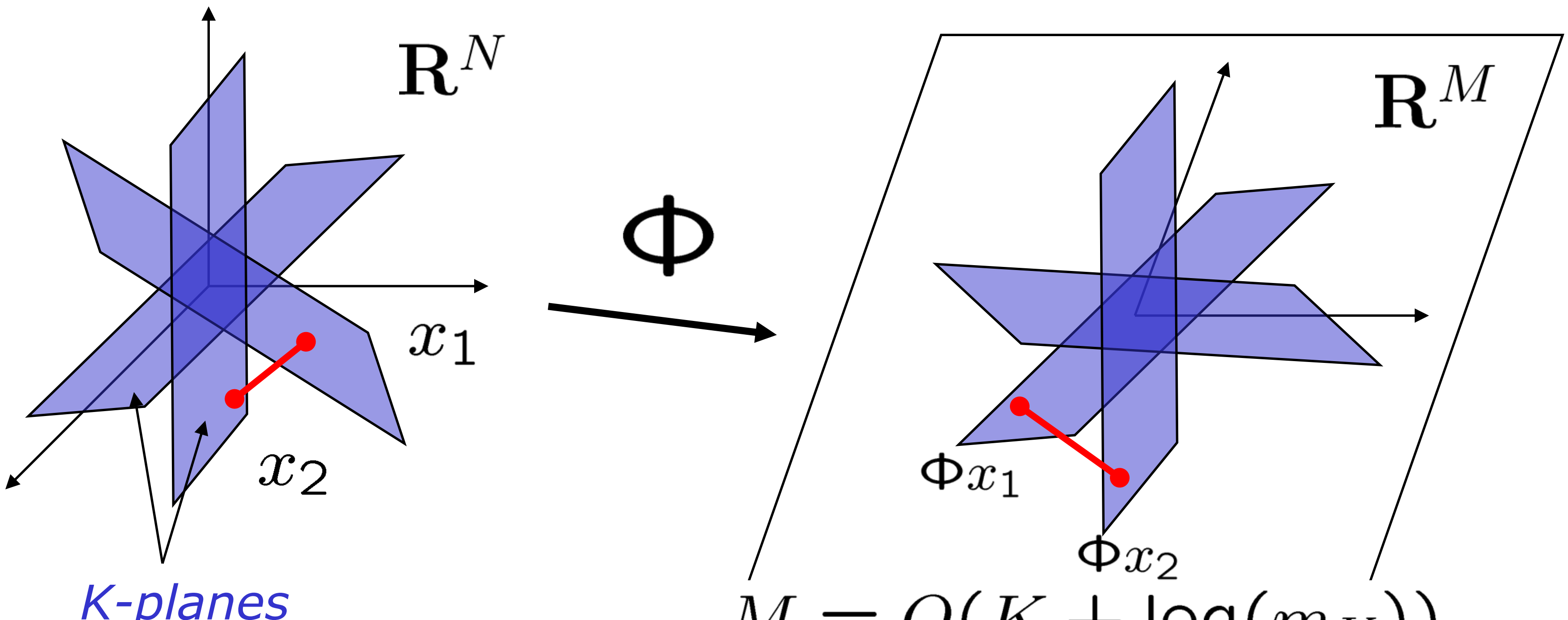
- **RIP:** stable embedding



$$M = O(K \log(N/K))$$

Sampling Bounds

- **Model-RIP:** stable embedding
  [B, D]; [B,D,DeV,W]



$$M = O(K + \log(m_K))$$

Iterated Thresholding

goal: given $y = \Phi x$, recover a sparse x

initialize: $\widehat{x}_0 = 0$, $r = y$, $i = 0$

iteration:

- $i \leftarrow i + 1$
- $b \leftarrow \widehat{x}_{i-1} + \Phi^T r$ **update signal estimate**
- $\widehat{x}_i \leftarrow \text{thresh}(b, K)$ **prune signal estimate** (best *K*-term approx)
- $r \leftarrow y - \Phi \widehat{x}_i$ **update residual**

return: $\widehat{x} \leftarrow \widehat{x}_i$

Iterated **Model** Thresholding

goal: given $y = \Phi x$, recover model-sparse x

initialize: $\widehat{x}_0 = 0$, $r = y$, $i = 0$

iteration:

- $i \leftarrow i + 1$
- $b \leftarrow \widehat{x}_{i-1} + \Phi^T r$ **update signal estimate**
- $\widehat{x}_i \leftarrow \underline{\mathcal{M}(b, K)}$ **prune signal estimate** (best *K*-term model approx)
- $r \leftarrow y - \Phi \widehat{x}_i$ **update residual**

return: $\widehat{x} \leftarrow \widehat{x}_i$

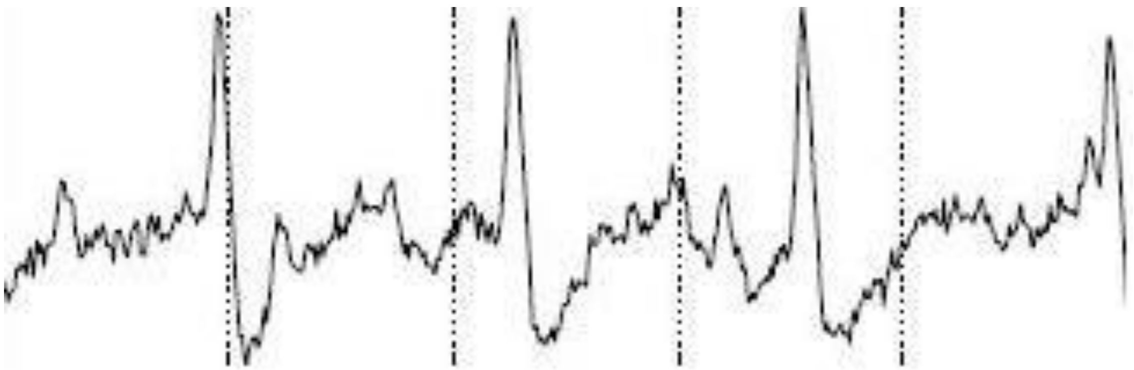
Recipe for CS recovery algorithm

Ingredients

- Model
- Sampling bound M
- Signal approximation algorithm $\mathcal{M}(\cdot, K)$

Application: 1D-signals

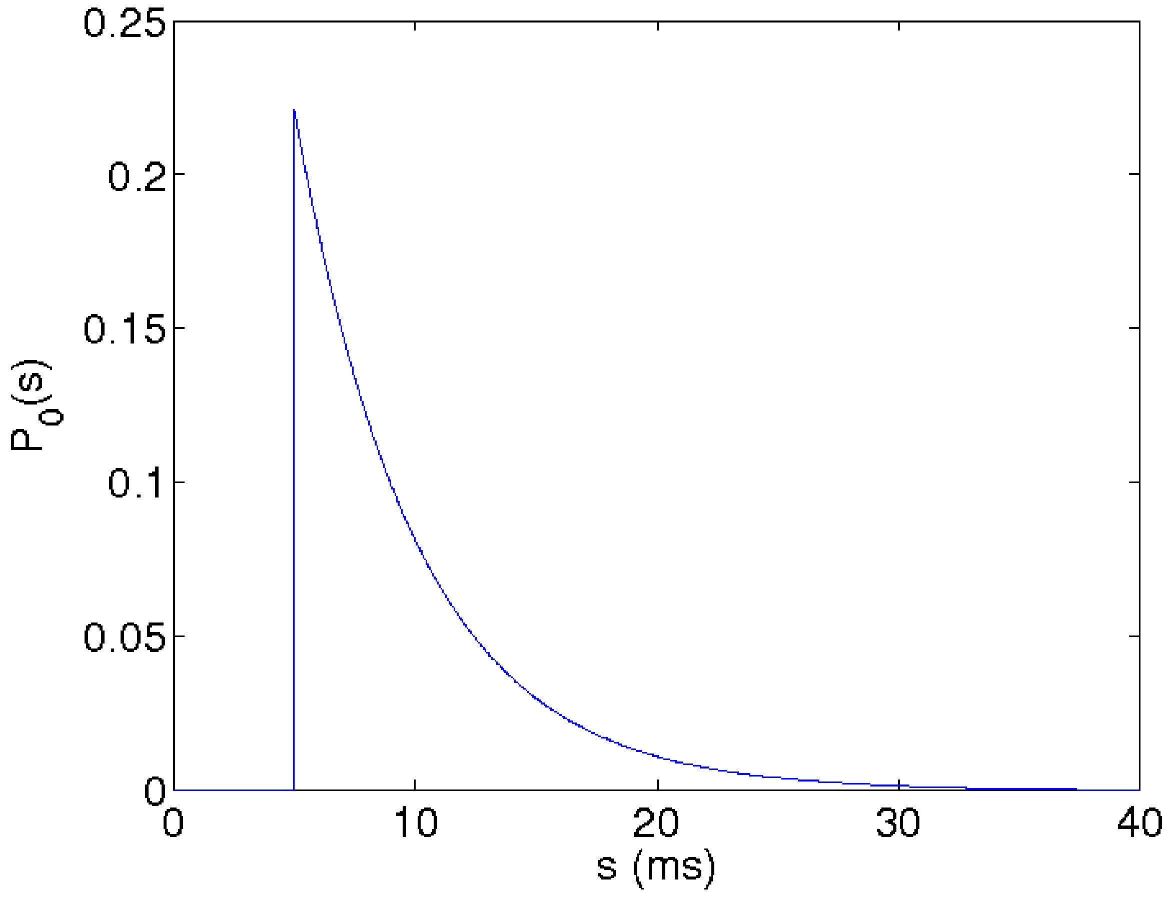
- Eg. Neuronal spike trains



[Lewicki, ’98]

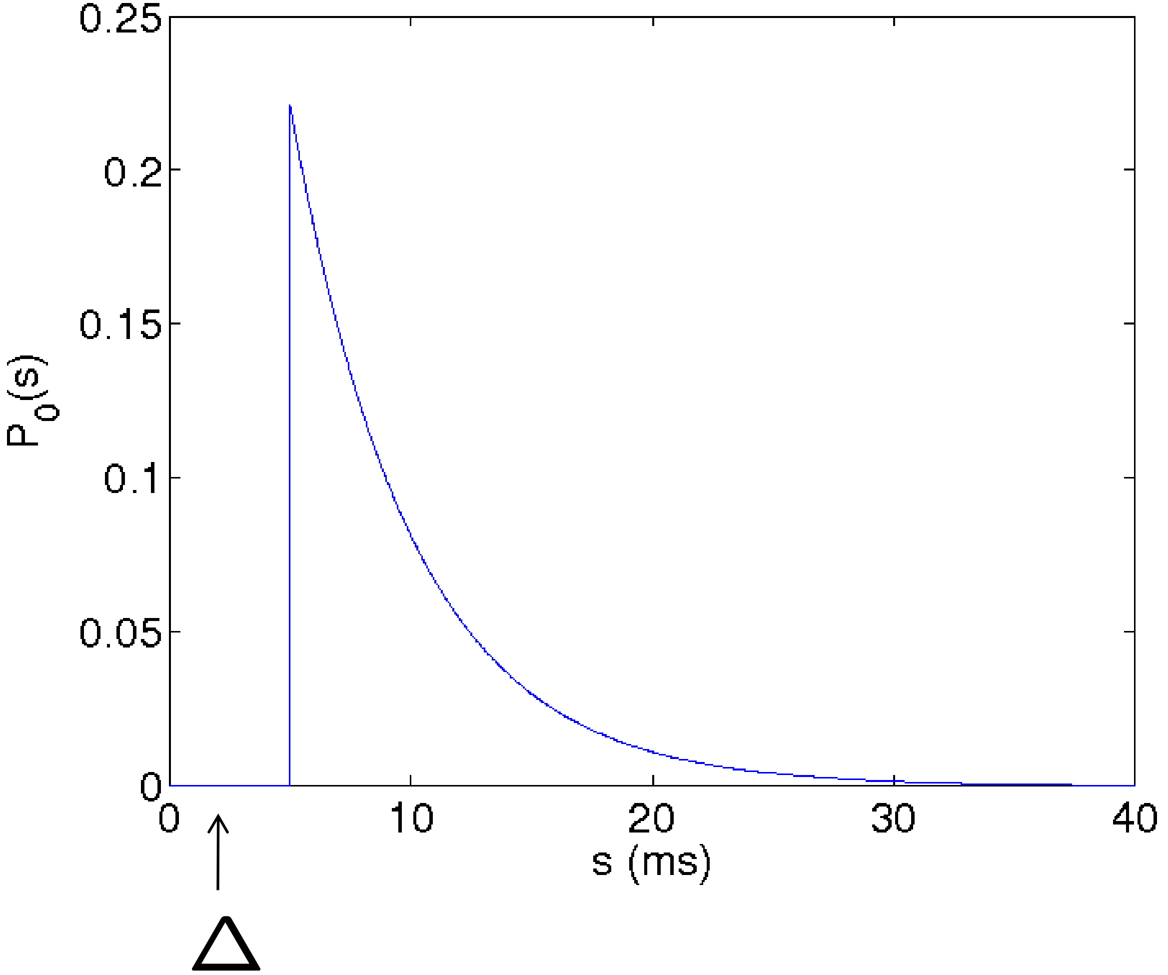
Application: spike trains

- Absolute refractoriness



Application: spike trains

- Absolute refractoriness

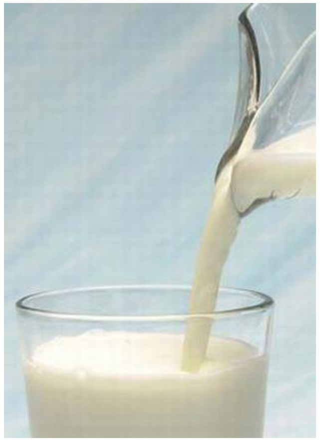


Ingredient 1: model

$x \in \mathcal{M}(K, \Delta)$ if

- $\|x\|_0 = K$
- no pair of consecutive nonzeros occur within Δ locations of each other.

Ingredient 2: sampling bound



- # subspaces = # sparsity patterns

$$m_K = \binom{N-(K-1)(\Delta-1)}{K}$$

- Number of measurements

$$M \geq CK \log(N/K - \Delta)$$

Ingredient 3: approximation

- Given arbitrary $x \in \mathbf{R}^N$, find closest $x^* \in \mathcal{M}(K, \Delta)$

- Equivalent to finding optimal *binary* support pattern

$$s = (s_1, \ldots, s_N) \in \mathcal{M}(K, \Delta)$$

- Portion of signal lying within a given support pattern:

$$x_{|s} := (s_1 x_1, s_2 x_2, \ldots, s_N x_N)$$

- Problem

$$\min_s \|x - x_{|s}\|_2, s \in \mathcal{M}$$

Ingredient 3: approximation

- Can be transformed into a integer program

$$\begin{aligned} s^* &= \arg\min c^\top s, \\ Ws &\leq u. \end{aligned}$$

- Can be relaxed into a linear program

Ingredient 3: approximation

- Can be transformed into a integer program

$$\begin{aligned} s^* &= \arg\min c^\top s, \\ Ws &\leq u. \end{aligned}$$

- Can be relaxed into a linear program

- *W* is totally unimodular, so polytope corners are integer, so solution is always integer

Ingredient 3: approximation

- Can be transformed into a integer program

$$\begin{aligned} s^* &= \arg\min c^\top s, \\ Ws &\leq u. \end{aligned}$$

- Can be relaxed into a linear program

- *W* is totally unimodular, so polytope corners are integer, so solution is always integer

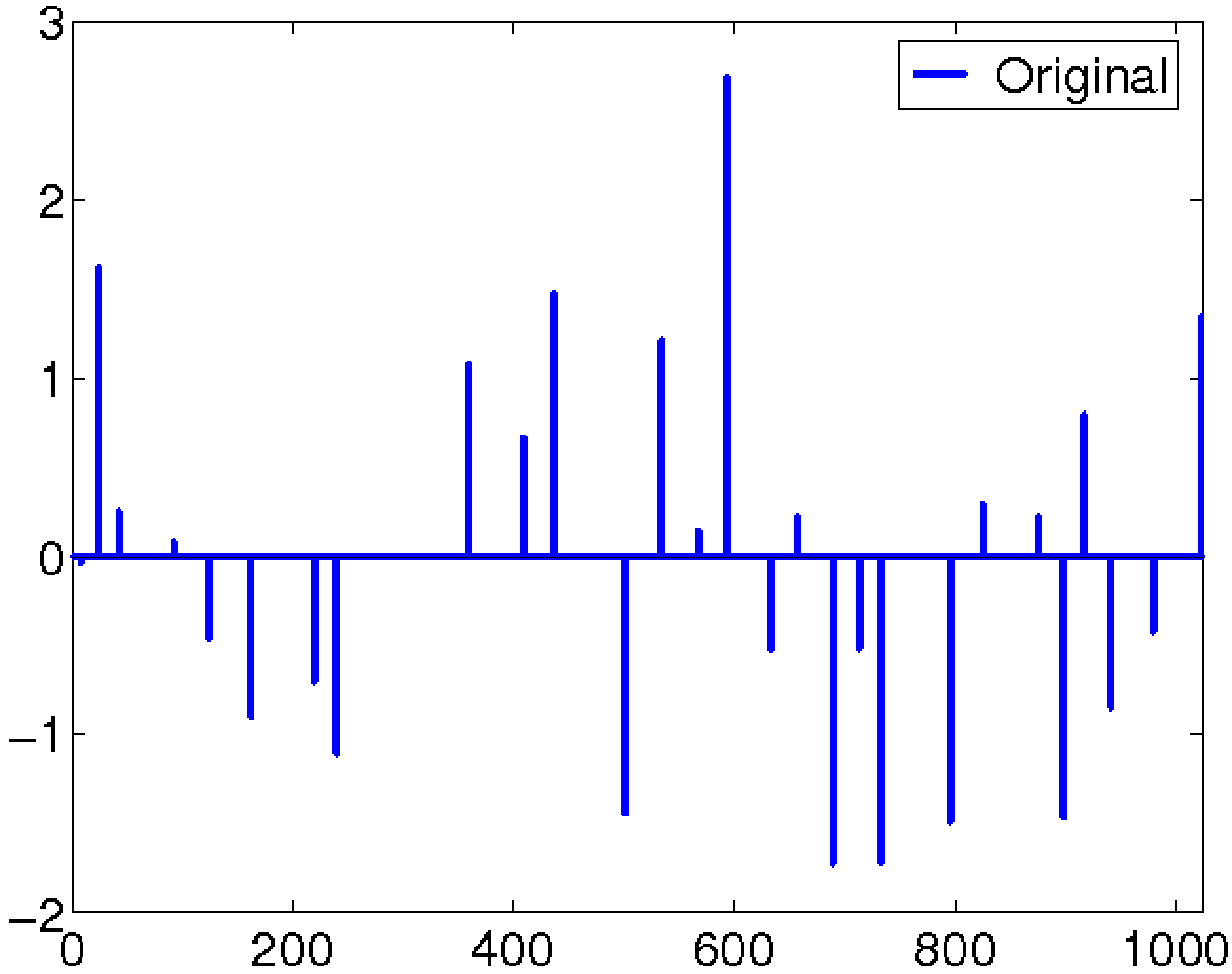
Voila

Theorem

If $y = \Phi x + n, x \in \mathcal{M}(K, \Delta)$ and Φ satisfies (K, Δ)-RIP, then the algorithm converges to an estimate $\widehat{x}$, such that

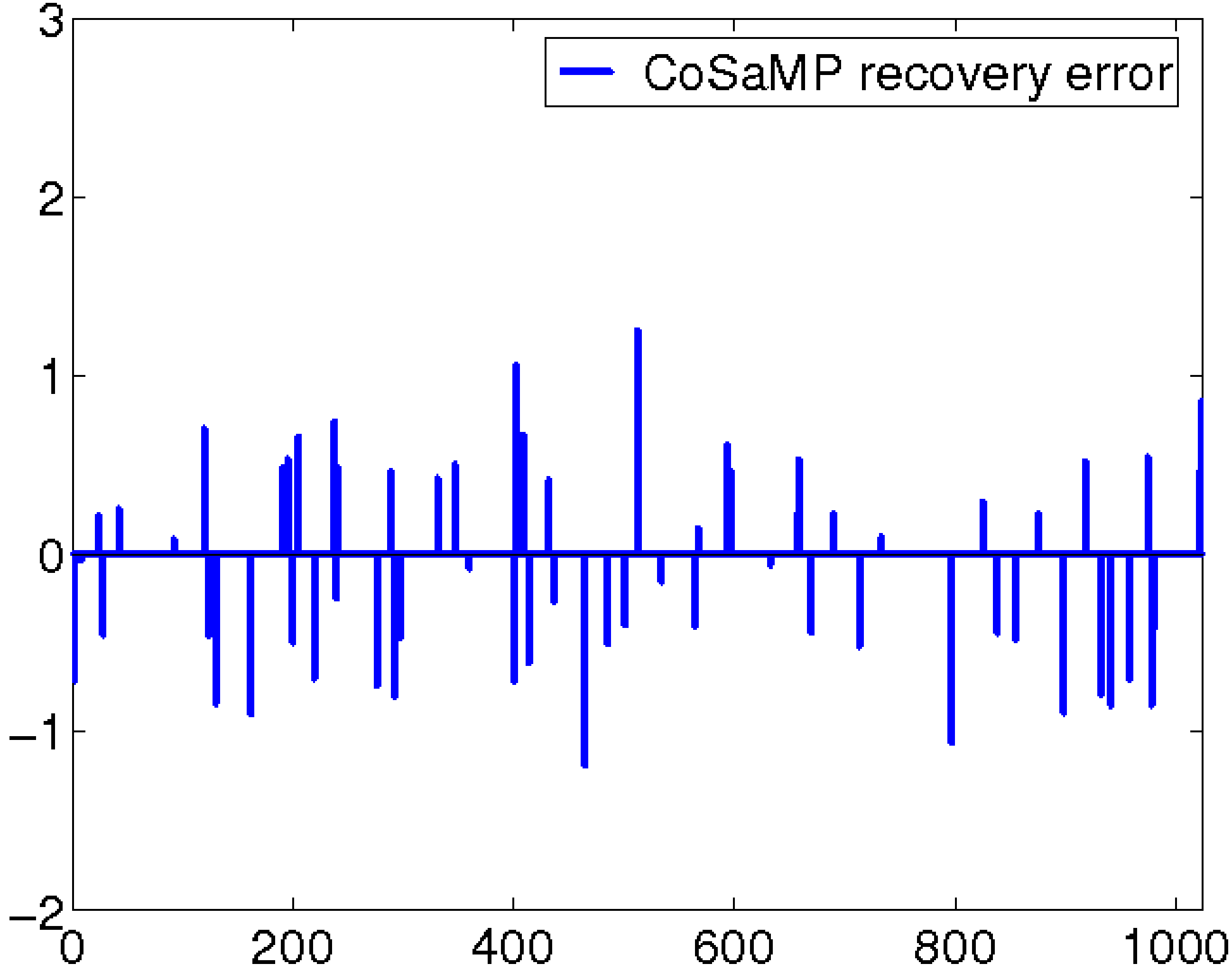
$$\|x - \widehat{x}\|_2 \leq C\|n\|_2$$

Results



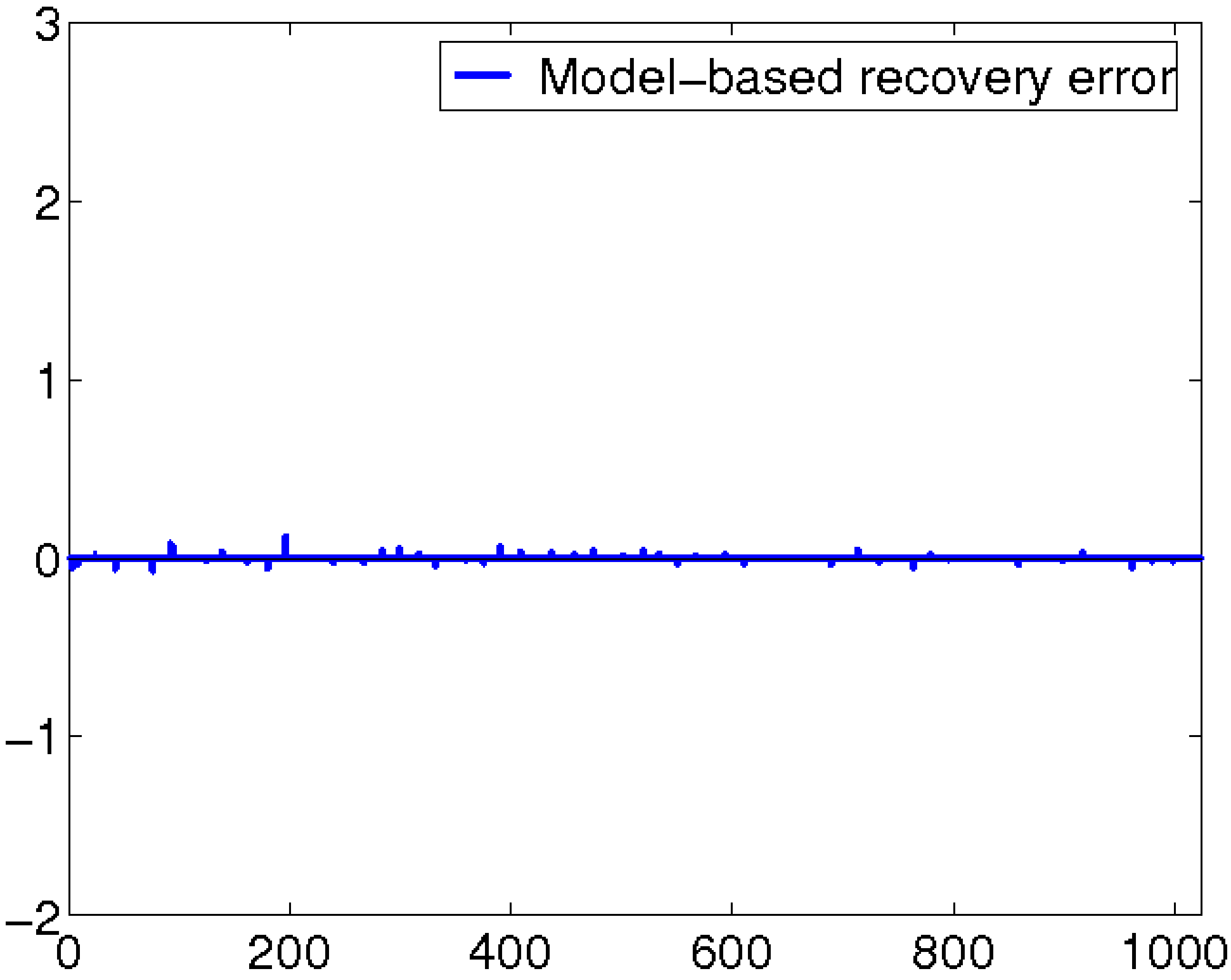
$N = 1024, K = 50, \Delta = 10$

Results



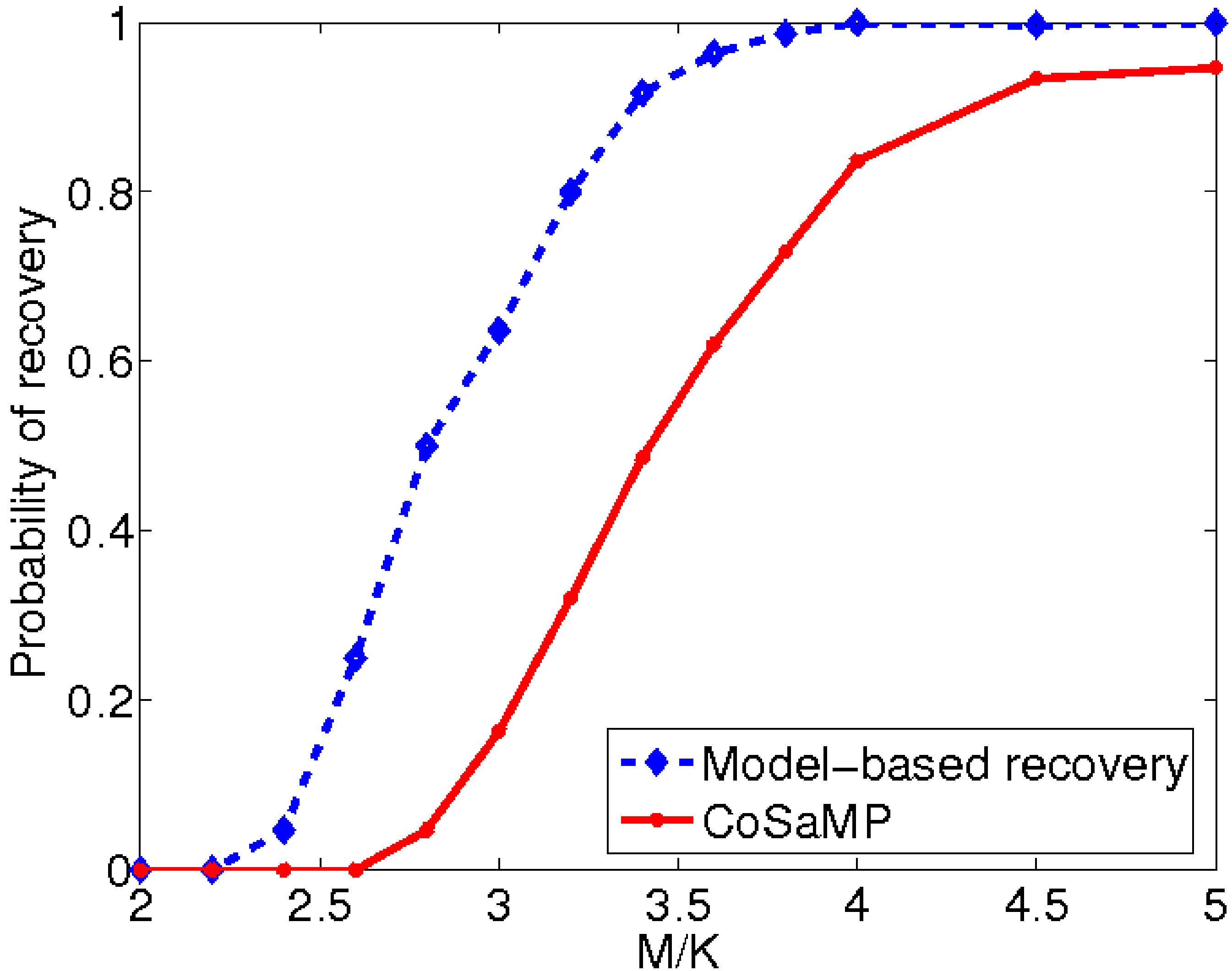
M = 150, Distortion = 1.76dB

Results



M = 150, Distortion = 25.53dB

Results



Probability of recovery
0
0.2
0.4
0.6
0.8
1
2
2.5
3
3.5
4
4.5
5
M/K
Model−based recovery
CoSaMP

Related Work

- Ex: wavelet-trees
  [Duarte , Wakin, Baraniuk], [La, Do], [Baraniuk, Cevher, Duarte, H]

- Ex: block sparsity / signal ensembles
  [Tropp, Gilbert, Strauss], [Stojnic, Parvaresh, Hassibi], [Eldar, Mishali], [Baron, Duarte et al], [B, C, D, H]

- Ex: clustered signals
  [C, D, H, B], [C, Indyk, H, B]

Related Work

- Ex: wavelet-trees

  [Duarte , Wakin, Baraniuk], [La, Do], [Baraniuk, Cevher, Duarte, H]

- Ex: block sparsity / signal ensembles
  [Tropp, Gilbert, Strauss], [Stojnic, Parvaresh, Hassibi],
  [Eldar, Mishali], [Baron, Duarte et al], [B, C, D, H]

- Ex: clustered signals
  [C, D, H, B], [C, Indyk, H, B]

- Model-compressible signals
  - Restricted Amplification Property (RAmP)
  - Instance-optimal guarantees in some cases

    [Baraniuk, Cevher, Duarte, H]

Summary

- Why CS works: stable embedding for signals with concise geometric structure

- Sparse signals >> **model-sparse signals**

- Model-based **recovery algorithms**

Advantages: provably fewer measurements
flexible framework for algorithm design
stable recovery

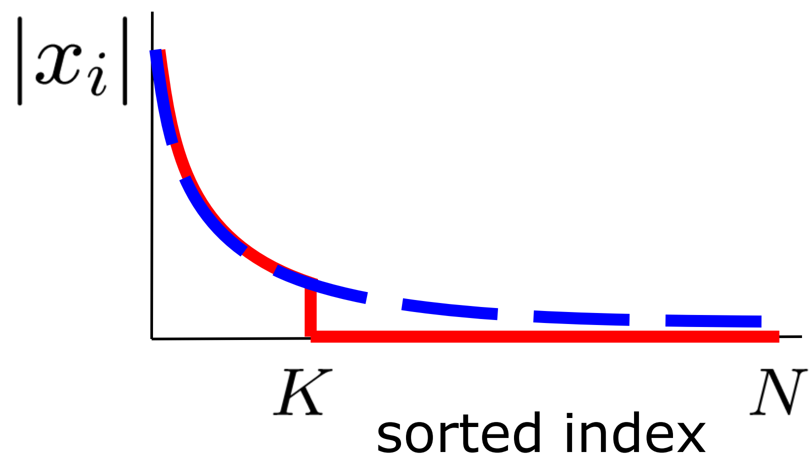
www.dsp.rice.edu/cs

- Blank slide

- Another blank slide

Compressible Signals

- Real-world signals are compressible, not sparse
- Recall: **compressible** <> approximable by sparse
  - compressible signals lie close to a union of subspaces
  - ie: approximation error decays rapidly as $K \rightarrow \infty$
  - nested in that $\mathrm{supp}\{x_K\} \subset \mathrm{supp}\{x'_K\}, \ K < K'$



- If Φ has RIP, then both sparse and compressible signals are stably recoverable via LP or greedy alg

Model-Compressible Signals

- **Model-compressible** <> approximable by model-sparse
  - model-compressible signals lie close to a reduced union of subspaces
  - ie: model-approx error decays rapidly as $K \to \infty$

- Nested approximation property (**NAP**): model-approximations nested in that

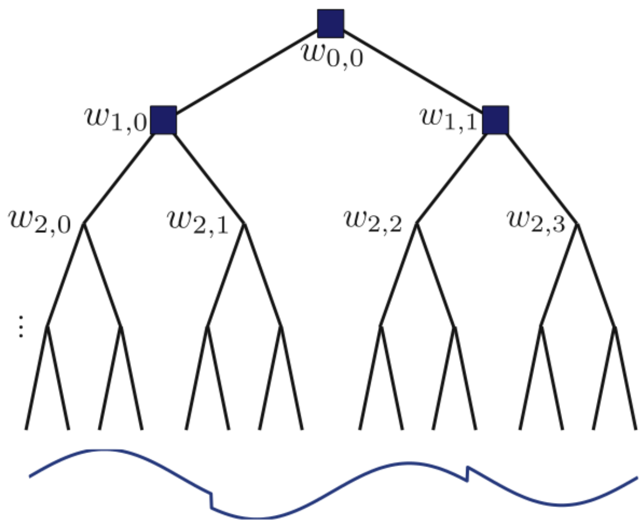
$$\text{supp}\{x_K\} \subset \text{supp}\{x'_K\}, \; K < K'$$

Model-Compressible Signals

- **Model-compressible** <> approximable by model-sparse
  - model-compressible signals lie close to a reduced union of subspaces
  - ie: model-approx error decays rapidly as $K \to \infty$

- Nested approximation property (**NAP**): model-approximations nested in that

$\text{supp}\{x_K\} \subset \text{supp}\{x'_K\}, \; K < K'$

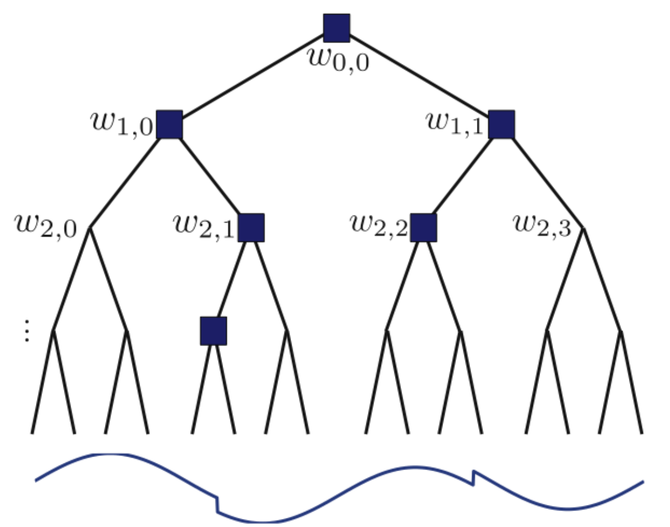


Model-Compressible Signals

- **Model-compressible** <> approximable by model-sparse
  - model-compressible signals lie close to a reduced union of subspaces
  - ie: model-approx error decays rapidly as $K \to \infty$

- Nested approximation property (**NAP**): model-approximations nested in that

$$\mathrm{supp}\{x_K\} \subset \mathrm{supp}\{x'_K\}, \ \ K < K'$$

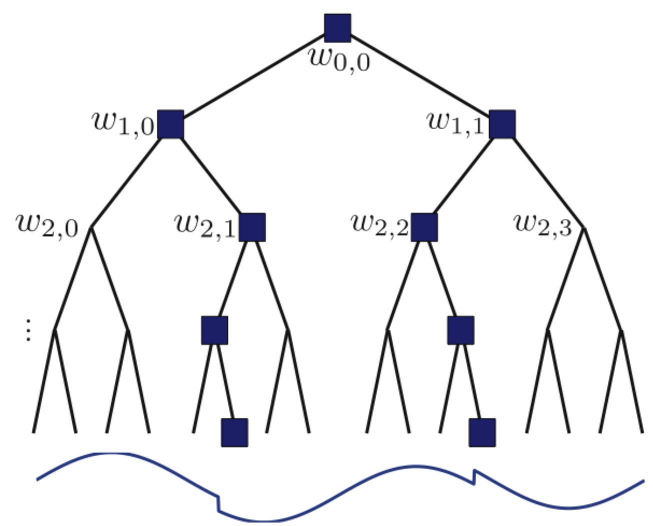


Model-Compressible Signals

- **Model-compressible** <> approximable by model-sparse
  - model-compressible signals lie close to a reduced union of subspaces
  - ie: model-approx error decays rapidly as $K \to \infty$

- Nested approximation property (**NAP**): model-approximations nested in that

$$\text{supp}\{x_K\} \subset \text{supp}\{x'_K\}, \;\; K < K'$$



Model-Compressible Signals

- **Model-compressible** <> approximable by model-sparse
  - model-compressible signals lie close to a reduced union of subspaces
  - ie: model-approx error decays rapidly as $K \to \infty$

- Nested approximation property (**NAP**): model-approximations nested in that

  $\text{supp}\{x_K\} \subset \text{supp}\{x'_K\},\ K < K'$

- **New result:** while model-RIP enables stable model-sparse recovery, model-RIP is *not sufficient* for stable model-compressible recovery!

Model-Compressible Signals

- **Model-compressible** <> approximable by model-sparse
  - model-compressible signals lie close to a reduced union of subspaces
  - ie: model-approx error decays rapidly as $K \rightarrow \infty$

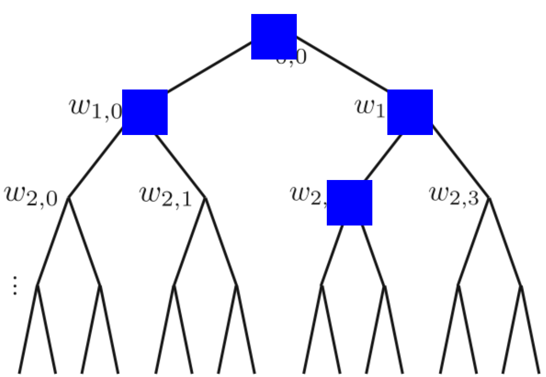
- **New result:** while model-RIP enables stable model-sparse recovery, model-RIP is *not sufficient* for stable model-compressible recovery!

- Ex: If Φ has the Tree-RIP, then tree-sparse signals can be stably recovered with $M = O(K)$
- However, tree-compressible signals cannot be stably recovered

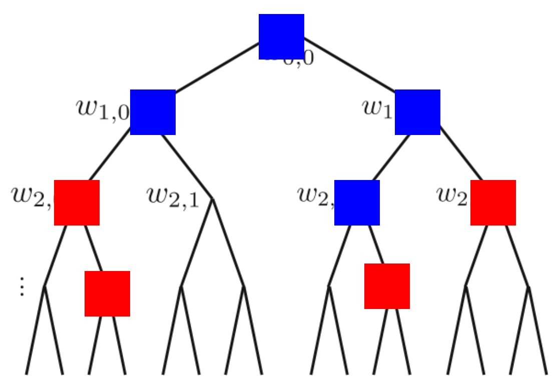
Stable Recovery

- **Result:** Stable model-compressible signal recovery requires that Φ have both:
  - RIP + **Restricted Amplification Property**

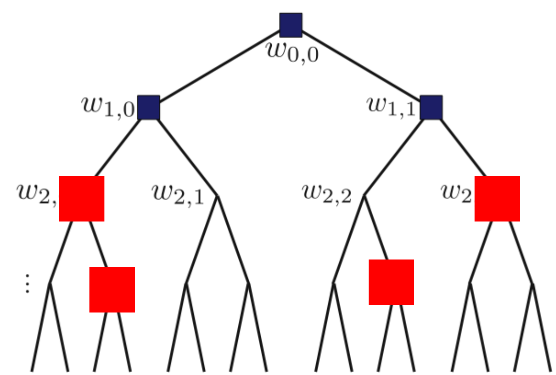
- **RAmP:** controls nonisometry of Φ in the approximation's **residual subspaces**



optimal K-term model recovery (error controlled by Φ RIP)



optimal $2K$-term model recovery (error controlled by Φ RIP)



residual subspace (error *not* controlled by Φ RIP)

Tree-RIP, Tree-RAmP

Theorem: An MxN iid subgaussian random matrix has the Tree(K)-RIP if

$$M \geq \begin{cases} \frac{2}{c\delta_{\mathcal{T}_K}^2}\left(K \ln \frac{48}{\delta_{\mathcal{T}_K}} + \ln \frac{512}{Ke^2} + t\right) & \text{if } K < \log_2 N \\ \frac{2}{c\delta_{\mathcal{T}_K}^2}\left(K \ln \frac{24e}{\delta_{\mathcal{T}_K}} + \ln \frac{2}{K+1} + t\right) & \text{if } K \geq \log_2 N \end{cases}$$

Theorem: An MxN iid subgaussian random matrix has the Tree(K)-RAmP if

$$M \geq \begin{cases} \frac{2}{\left(\sqrt{1+\epsilon_K}-1\right)^2}\left(10K + 2\ln \frac{N}{K(K+1)(2K+1)} + t\right) & \text{if } K \leq \log_2 N \\ \frac{2}{\left(\sqrt{1+\epsilon_K}-1\right)^2}\left(10K + 2\ln \frac{601N}{K^3} + t\right) & \text{if } K > \log_2 N \end{cases}$$

Performance

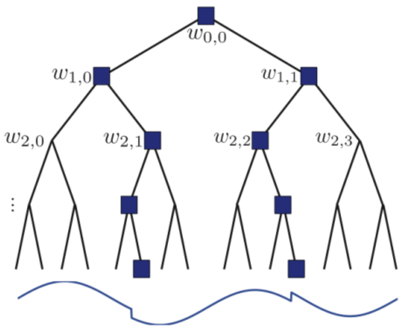
- Using model-based IHT, CoSaMP with RIP and RAmP

- **Model-sparse signals**
  - noise-free measurements: exact recovery
  - noisy measurements: stable recovery

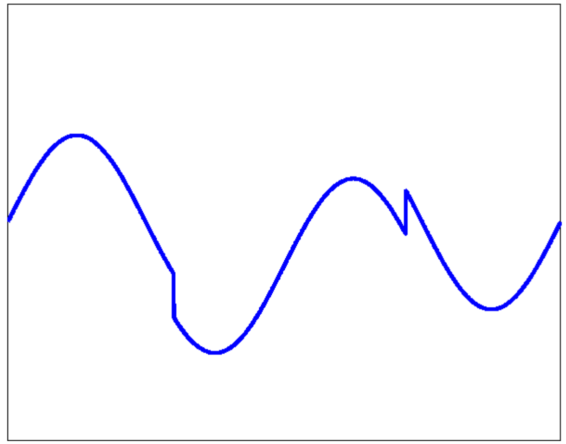
- **Model-compressible signals**
  - recovery as good as *K*-model-sparse approximation

$$\underbrace{\|x - \widehat{x}\|_{\ell_2}}_{\text{CS recovery error}} \leq C_1 \underbrace{\|x - x_K\|_{\ell_2}}_{\text{signal model } K\text{-term approx error}} + C_2 \underbrace{\frac{\|x - x_K\|_{\ell_1}}{K^{1/2}}}_{\text{signal model } K\text{-term approx error}} + C_3 \underbrace{\epsilon}_{\text{noise}}$$

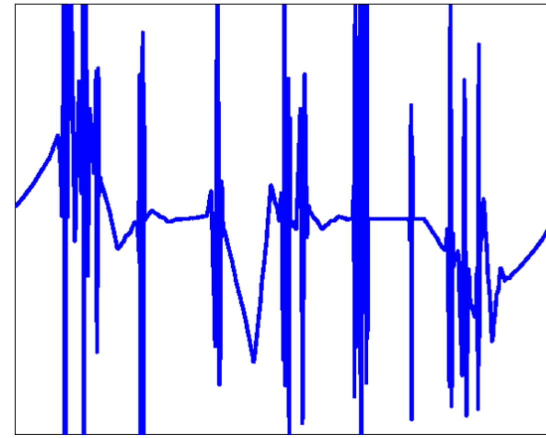
Tree-Sparse Signal Recovery



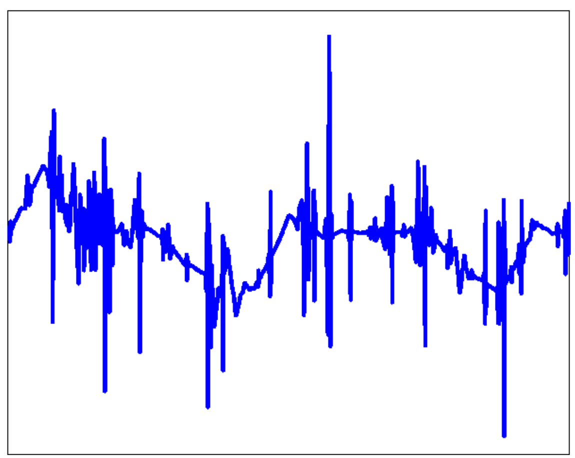
N=1024
M=80



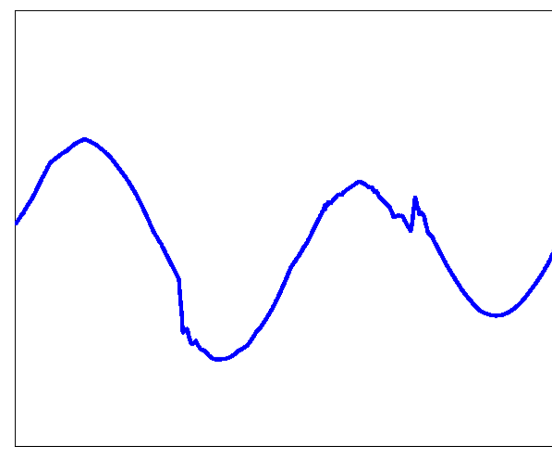
target signal



CoSaMP,
(MSE=1.12)



L1-minimization
(MSE=0.751)



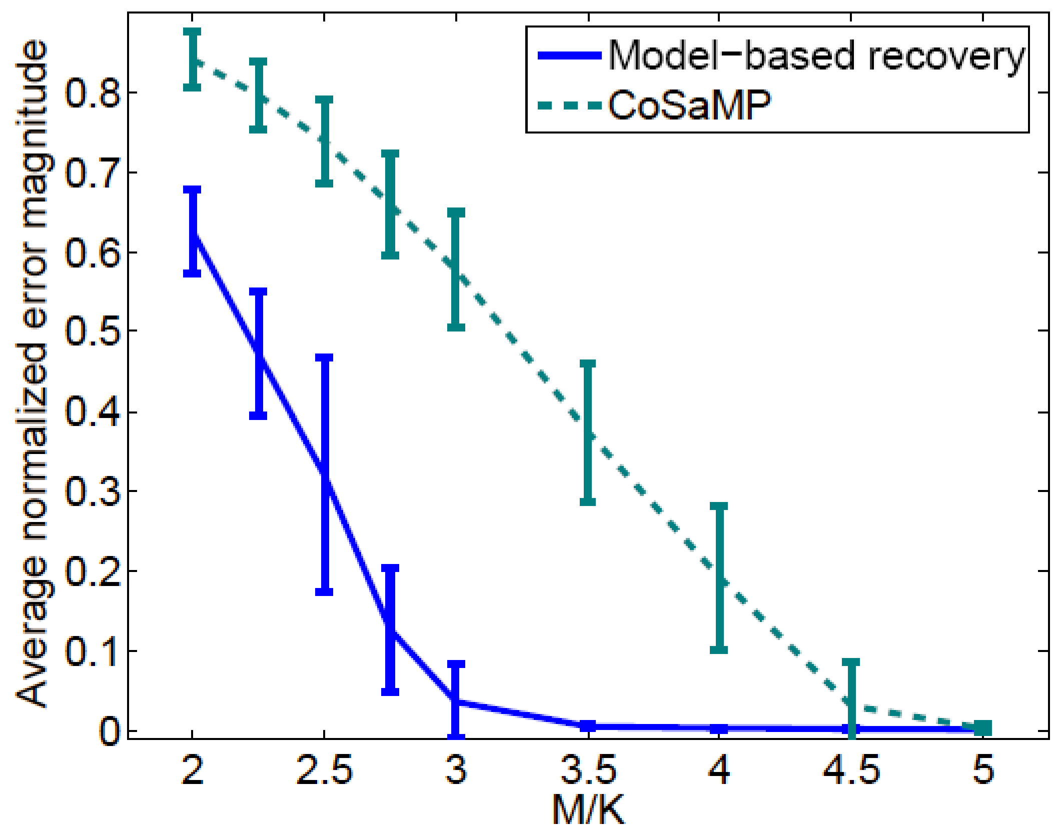
Tree-sparse CoSaMP
(MSE=0.037)

Simulation

- Recovery performance (MSE) vs. number of measurements

- Piecewise cubic signals + wavelets

- Models/algorithms:
  - sparse (CoSaMP)
  - tree-sparse (tree-CoSaMP)

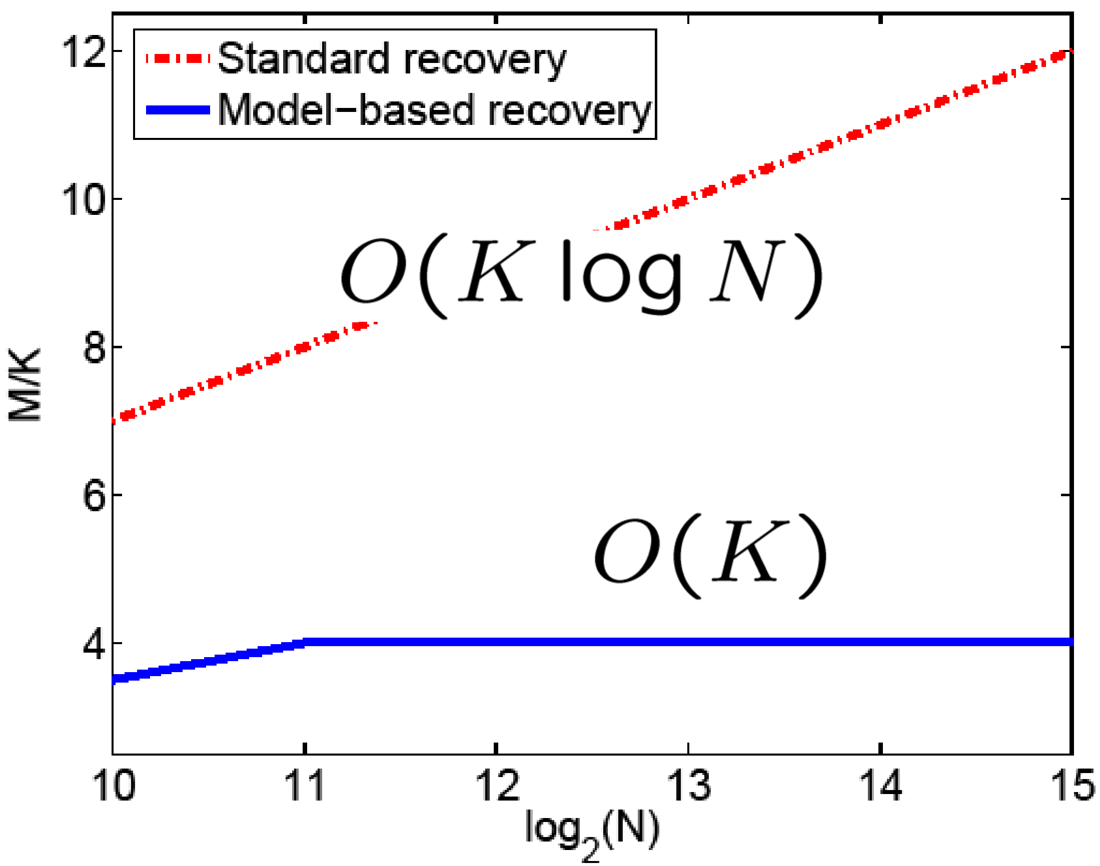


Simulation

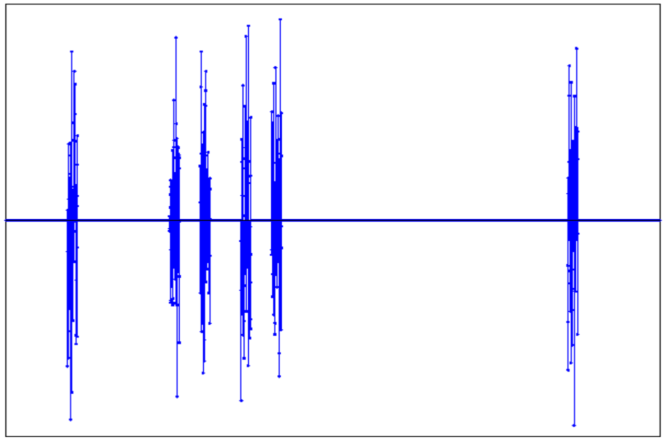
- Number samples for correct recovery

- Piecewise cubic signals + wavelets

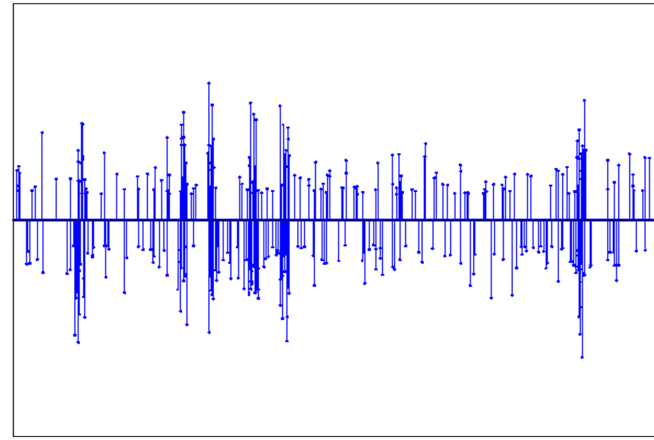
- Models/algorithms:
  - sparse (CoSaMP)
  - tree-sparse (tree-CoSaMP)



Block-Sparse Signal

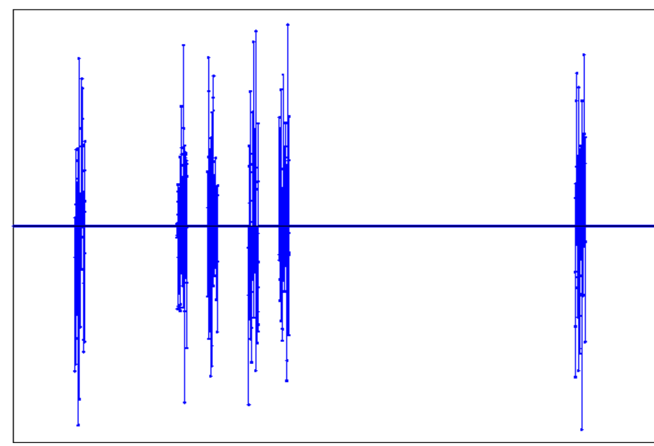


target



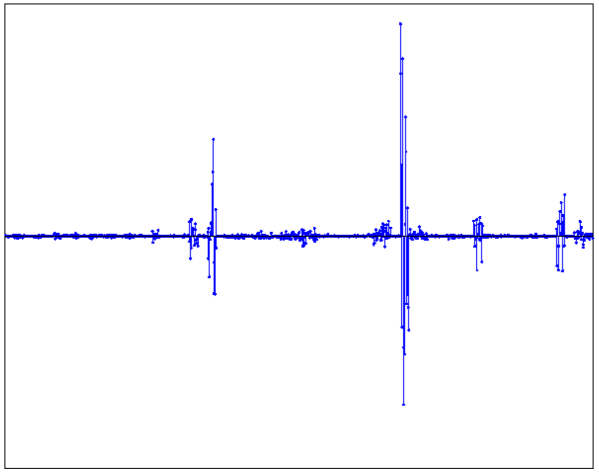
CoSaMP (MSE = 0.723)

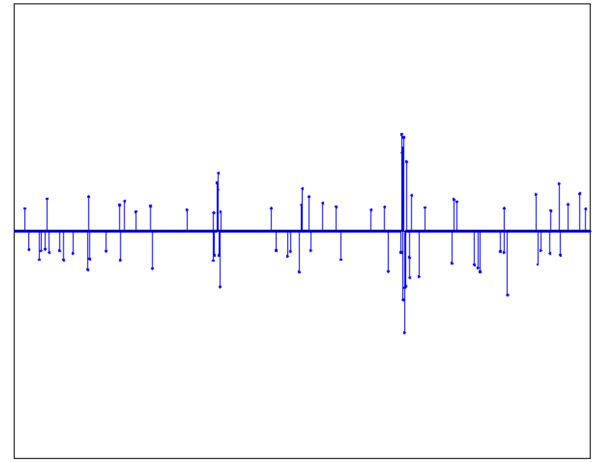
N = 4096
K = 6 active blocks
J = block length = 64
M = 2.5JK = 960 msnts

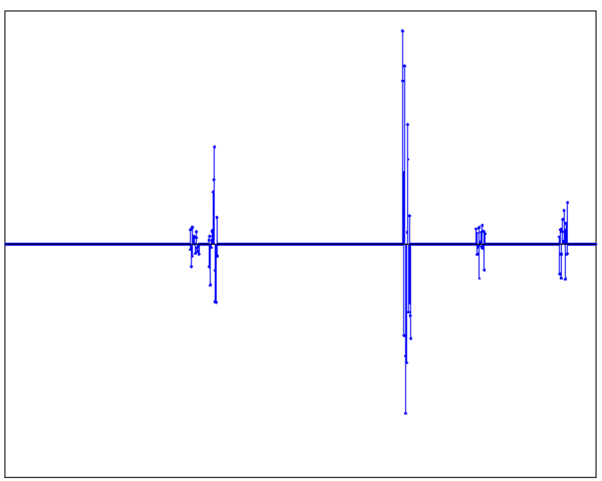


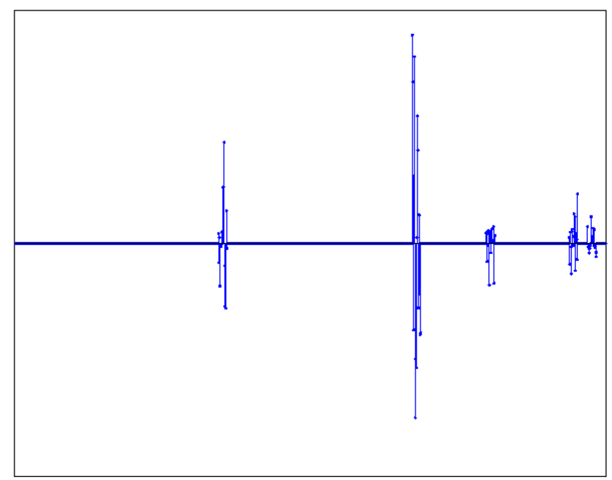
block-sparse model recovery
(MSE=0.015)

Block-Compressible Signal


target

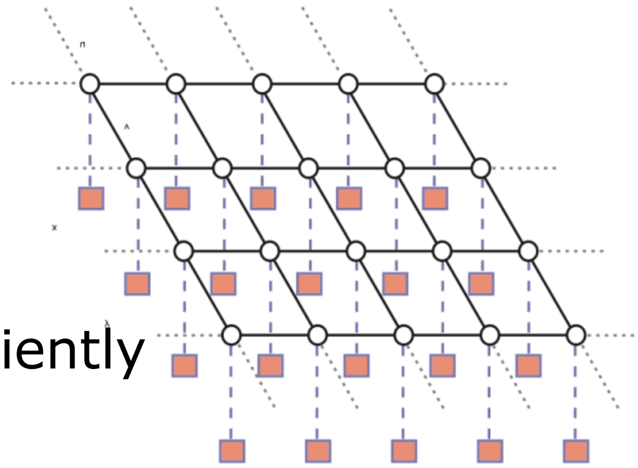

CoSaMP (MSE=0.711)

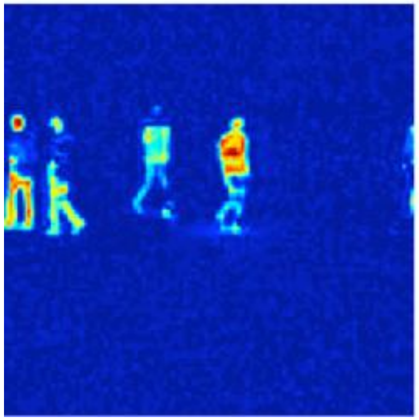

best 5-block approximation
(MSE=0.116)


block-sparse recovery
(MSE=0.195)

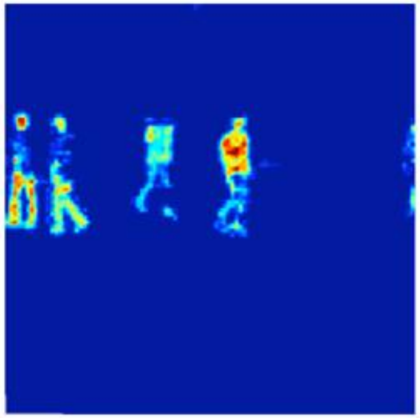
Clustered Signals

- **Graphical model**
- Model **clustering of significant pixels** in space domain using Ising MRF
- Ising model approximation performed efficiently using **graph cuts**

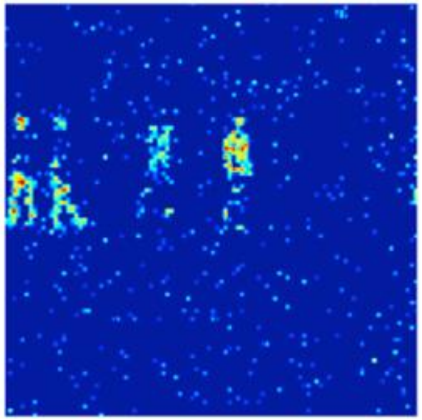




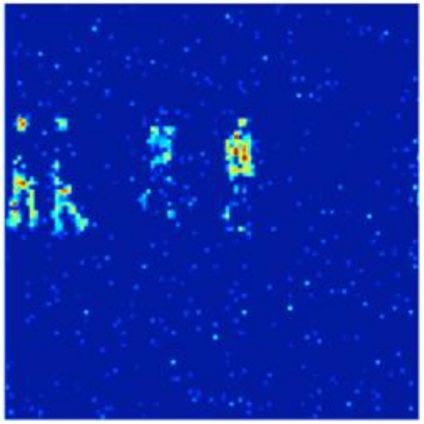
target



Ising-model recovery



CoSaMP recovery



LP (FPC) recovery