Sensing via Dimensionality Reduction

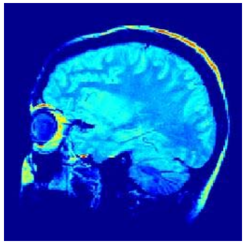
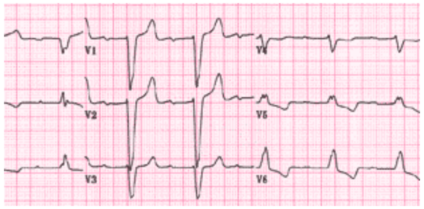
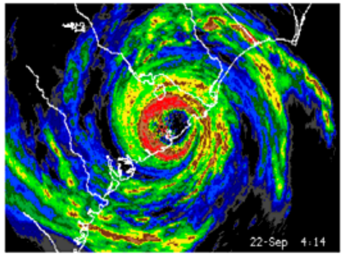
Structured Sparsity Models

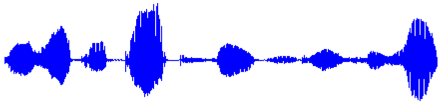
Volkan Cevher

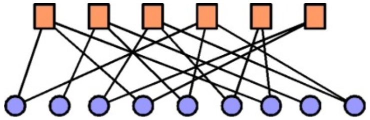
volkan@rice.edu

Rice University DSP





Sensors

1975 -0.08MP

1957 - 30fps

1877 - ?

1977 – 5hours

160MP

200,000fps

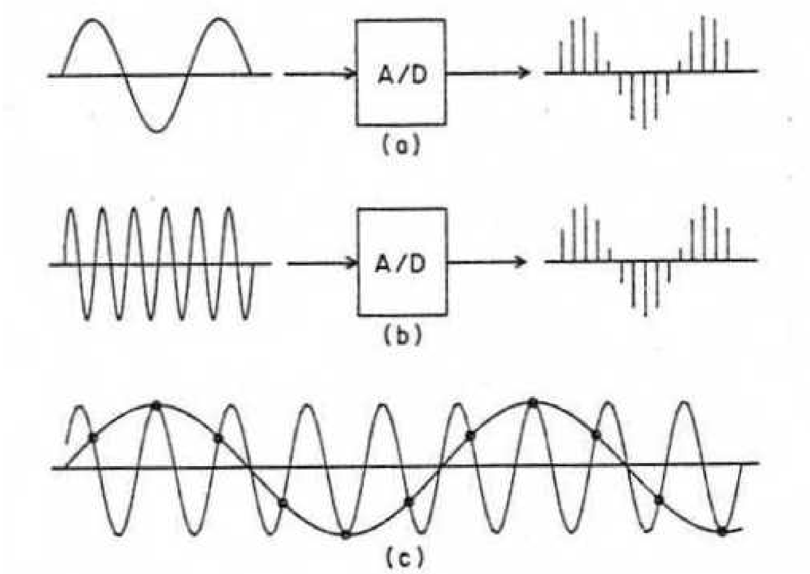
192,000Hz

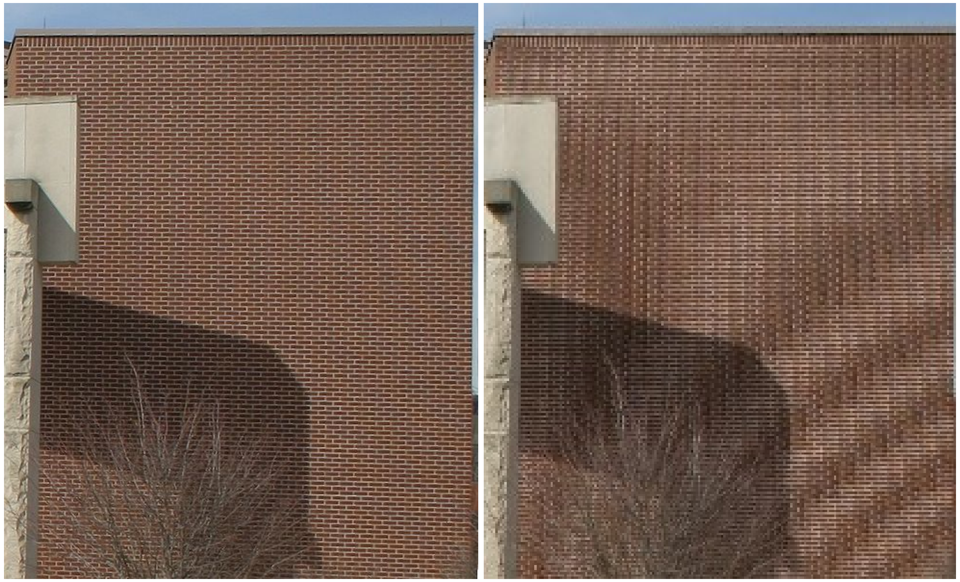
30mins

Digital Data Acquisition

Foundation: *Shannon/Nyquist sampling theorem*

"if you sample densely enough (at the Nyquist rate), you can perfectly reconstruct the original analog data"





time

space

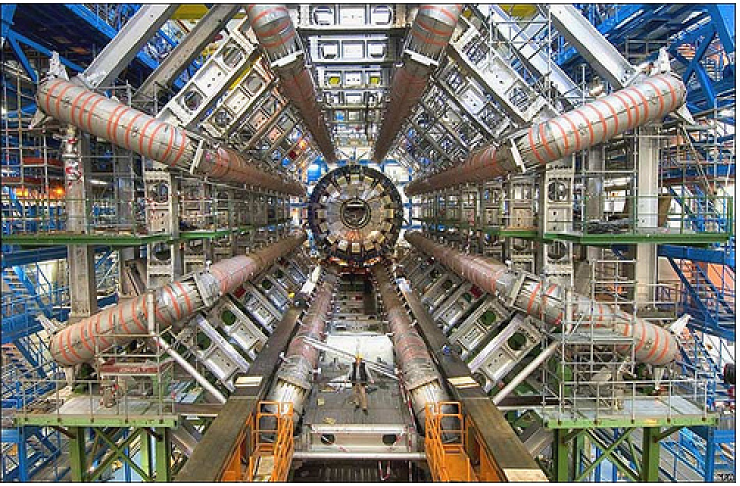
Major Trends in Sensing

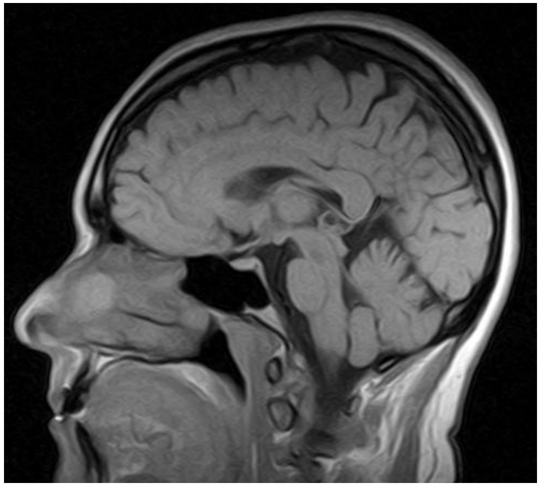
higher resolution / denser sampling

large numbers of sensors

increasing # of modalities / mobility

Major Trends in Sensing





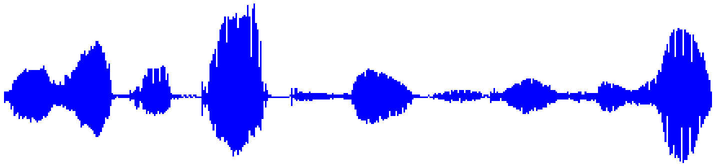
Motivation: solve bigger / more important problems

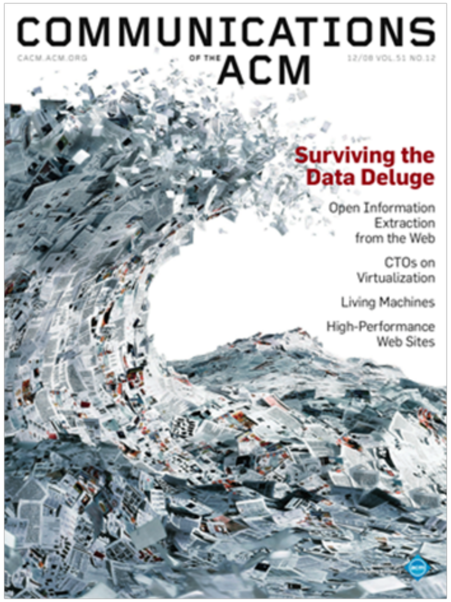
decrease acquisition times / costs

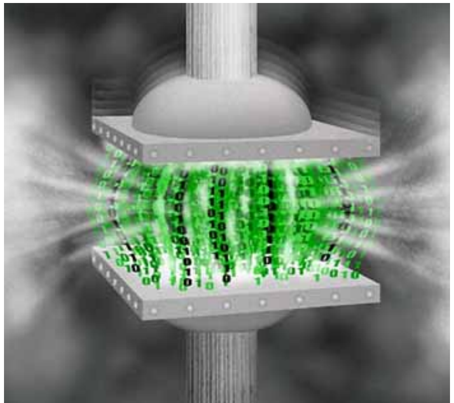
entertainment...

Problems of the Current Paradigm

- **Sampling at Nyquist rate**
  - expensive / difficult
- **Data deluge**
  - communications / storage
- **Sample then compress**
  - not future proof







Approaches

- Do nothing / Ignore

  be content with
  where we are...

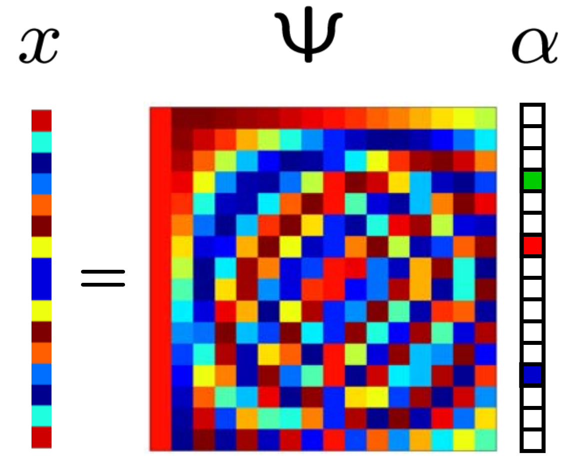
  - generalizes well
  - robust

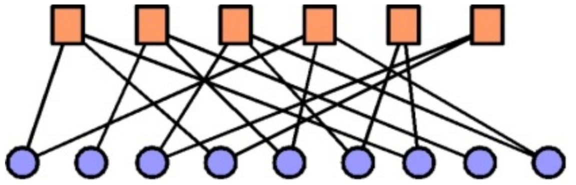
Approaches

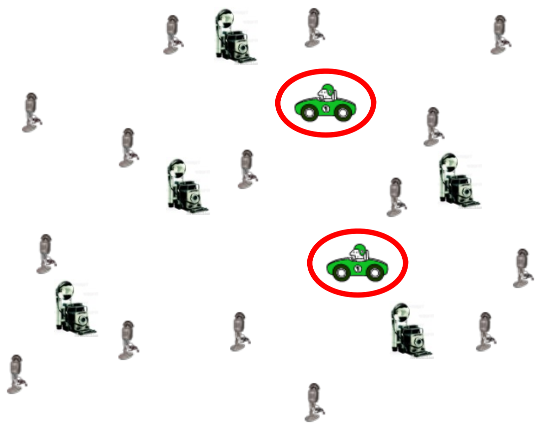
- **Finite Rate of Innovation**

  Sketching / Streaming

  Compressive Sensing







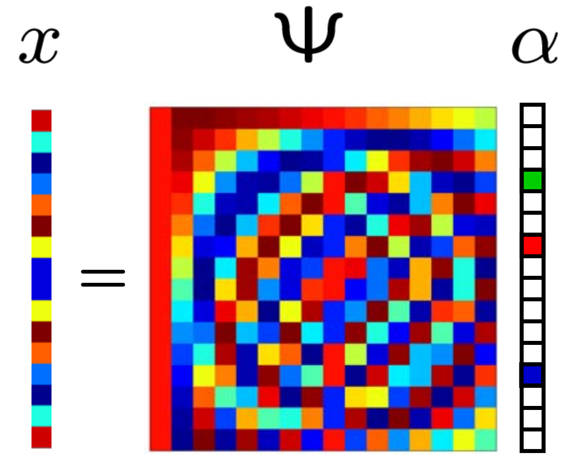
[Vetterli, Marziliano, Blu; Blu, Dragotti, Vetterli, Marziliano, Coulot; Gilbert, Indyk, Strauss, Cormode, Muthukrishnan; Donoho; Candes, Romberg, Tao; Candes, Tao]

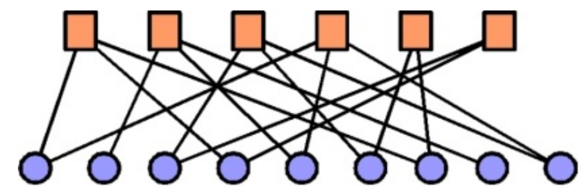
Approaches

- **Finite Rate of Innovation**

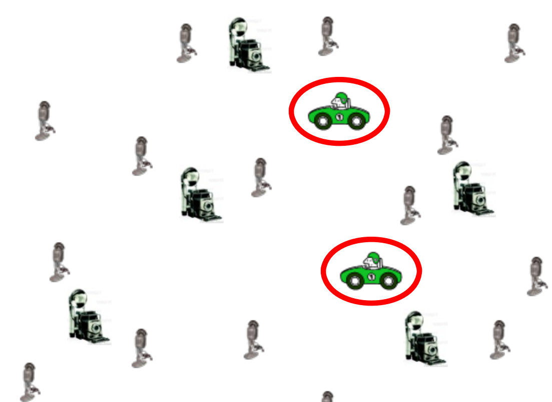
  Sketching / Streaming

  Compressive Sensing





PARSITY



[Vetterli, Marziliano, Blu; Blu, Dragotti, Vetterli, Marziliano, Coulot; Gilbert, Indyk, Strauss, Cormode, Muthukrishnan; Donoho; Candes, Romberg, Tao; Candes, Tao]

Today – Beyond Sparsity

Sensing via dimensionality reduction

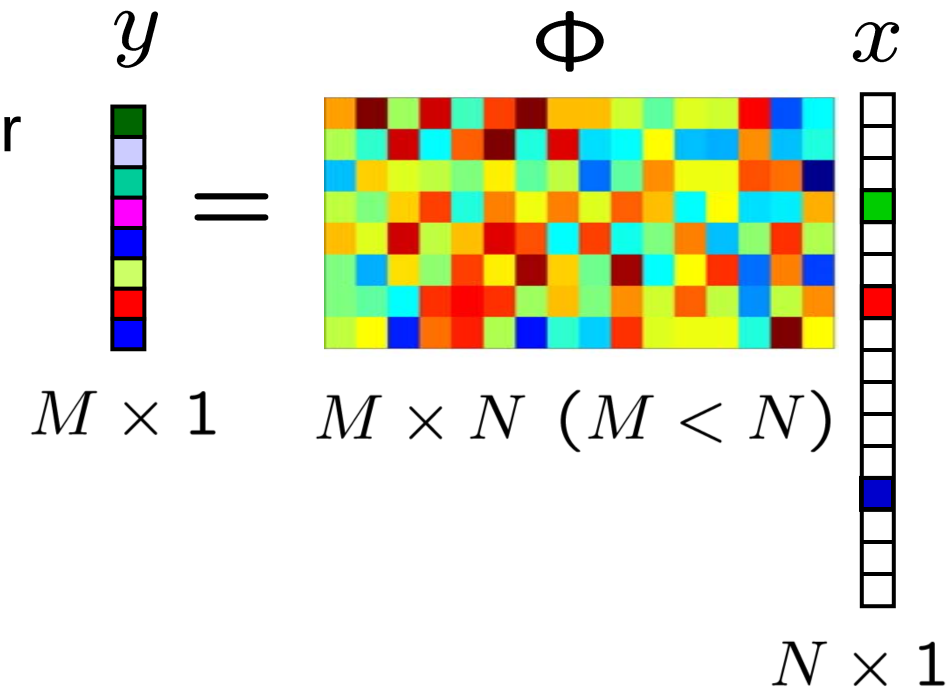
Model-based Compressive Sensing
w/ Structured Sparsity Models

- Reducing sampling / processing / communication costs
- Increasing recovery / processing speed
- Improving robustness / stability

Compressive Sensing 101

- **Goal**: Recover a *sparse* or *compressible* signal x from measurements y

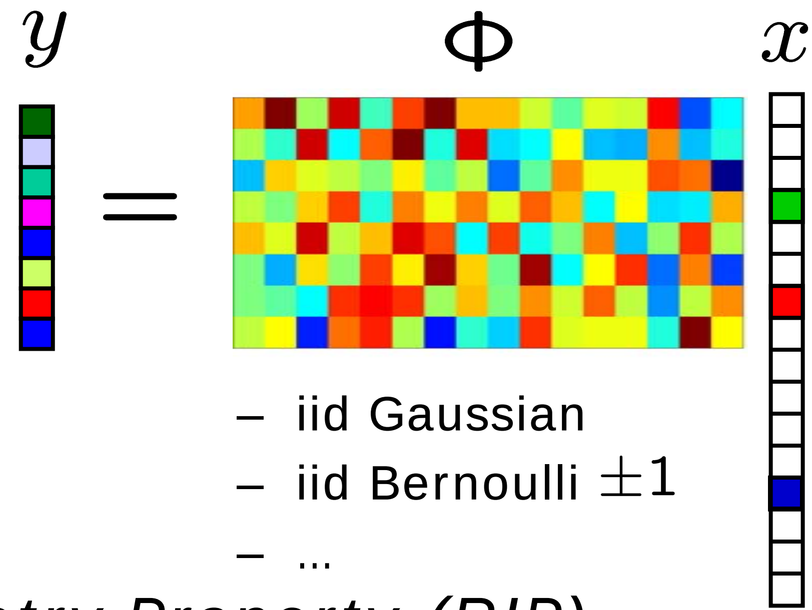
- **Problem**: Random projection Φ not full rank



- **Solution**: Exploit the sparsity/compressibility ***geometry*** of acquired signal x

Compressive Sensing 101

- **Goal**: Recover a *sparse* or *compressible* signal x from measurements y

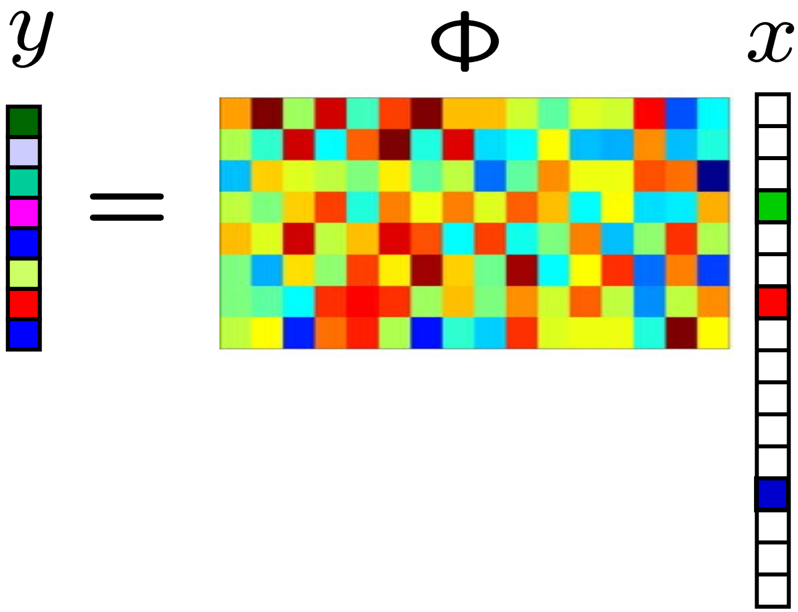


- iid Gaussian
- iid Bernoulli ± 1
- ...

- **Problem**: Random projection Φ not full rank *but satisfies Restricted Isometry Property (RIP)*

- **Solution**: Exploit the sparsity/compressibility ***geometry*** of acquired signal x

Compressive Sensing 101



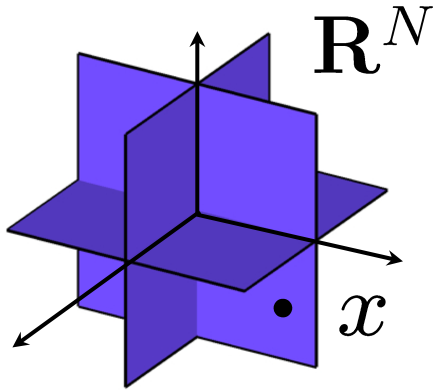
- **Goal**: Recover a *sparse* or *compressible* signal x from measurements y

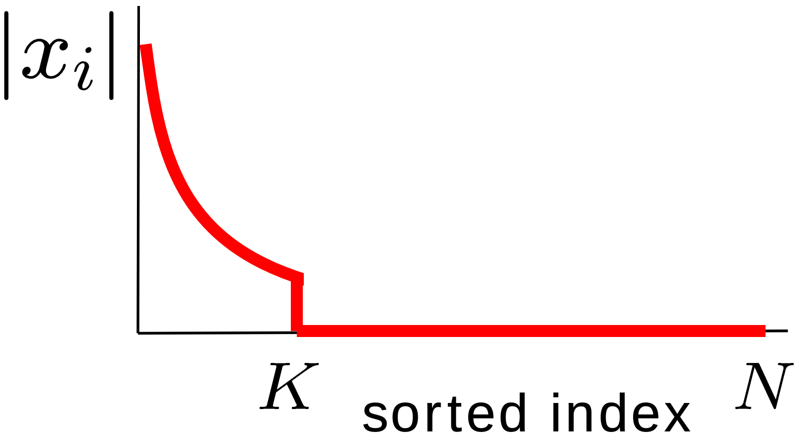
- **Problem**: Random projection Φ not full rank

- **Solution**: Exploit the ***model geometry*** of acquired signal x

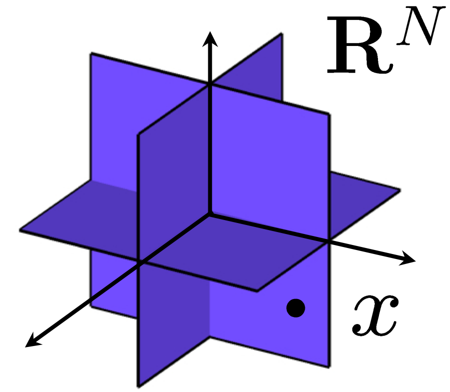
Concise Signal Structure

- **Sparse** signal: only K out of N coordinates nonzero
  - model: union of K-dimensional subspaces aligned w/ coordinate axes



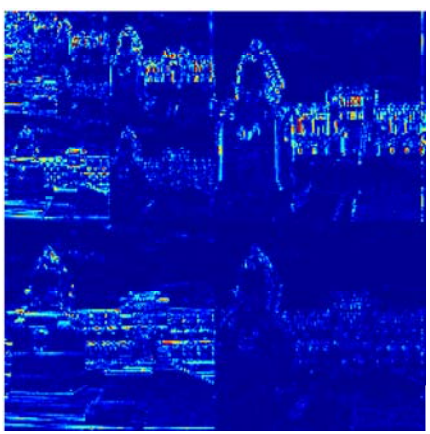


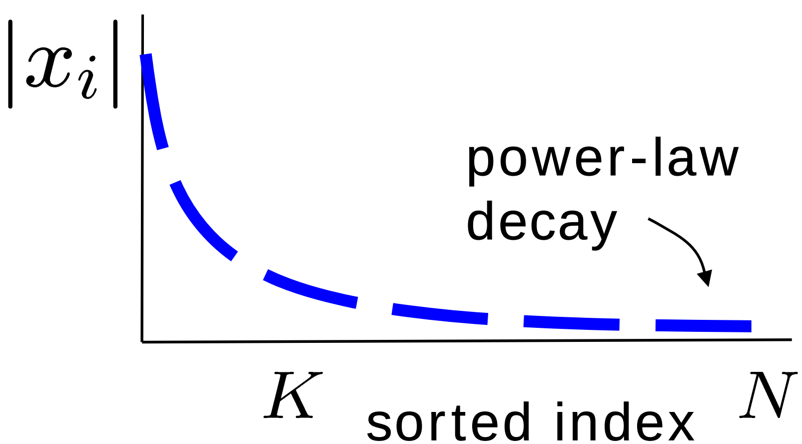
Concise Signal Structure

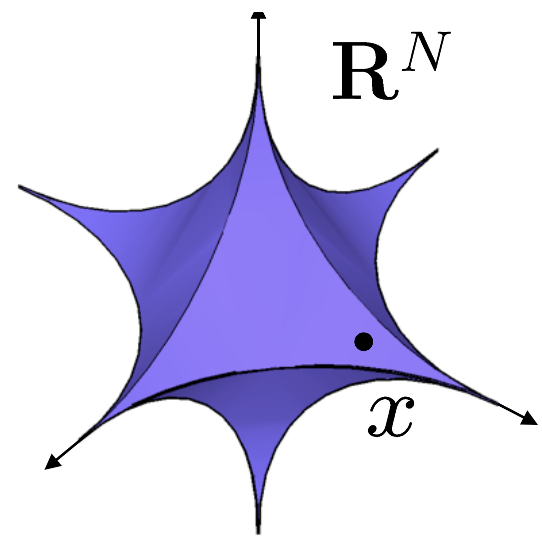


- **Sparse** signal: only K out of N coordinates nonzero
  - model: union of K-dimensional subspaces

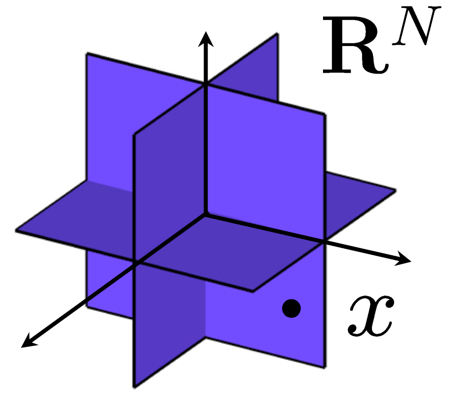
- **Compressible** signal: sorted coordinates decay rapidly to zero
  - model: ℓ_p ball: $\|x\|_p^p = \sum_i |x_i|^p \leq 1, \; p \leq 1$







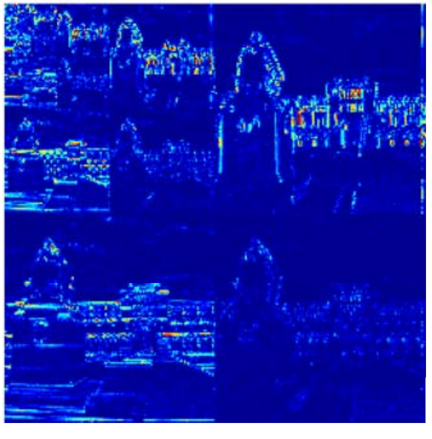
Concise Signal Structure

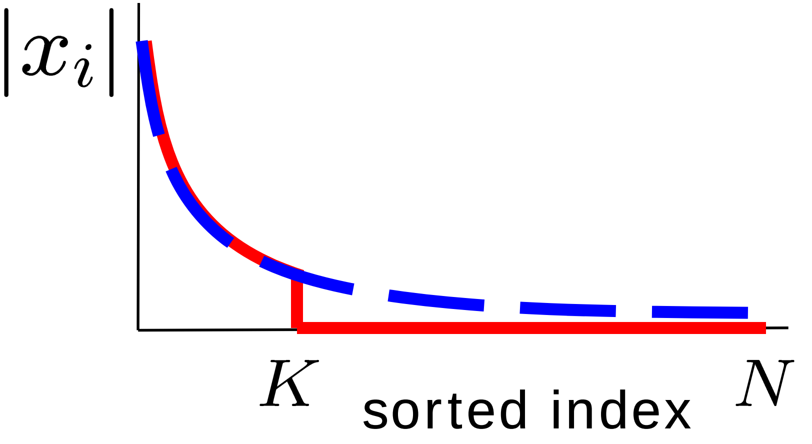


- **Sparse** signal: only K out of N coordinates nonzero
  - model: union of K-dimensional subspaces

- **Compressible** signal: sorted coordinates decay rapidly to zero

well-approximated by a K-sparse signal
(simply by thresholding)

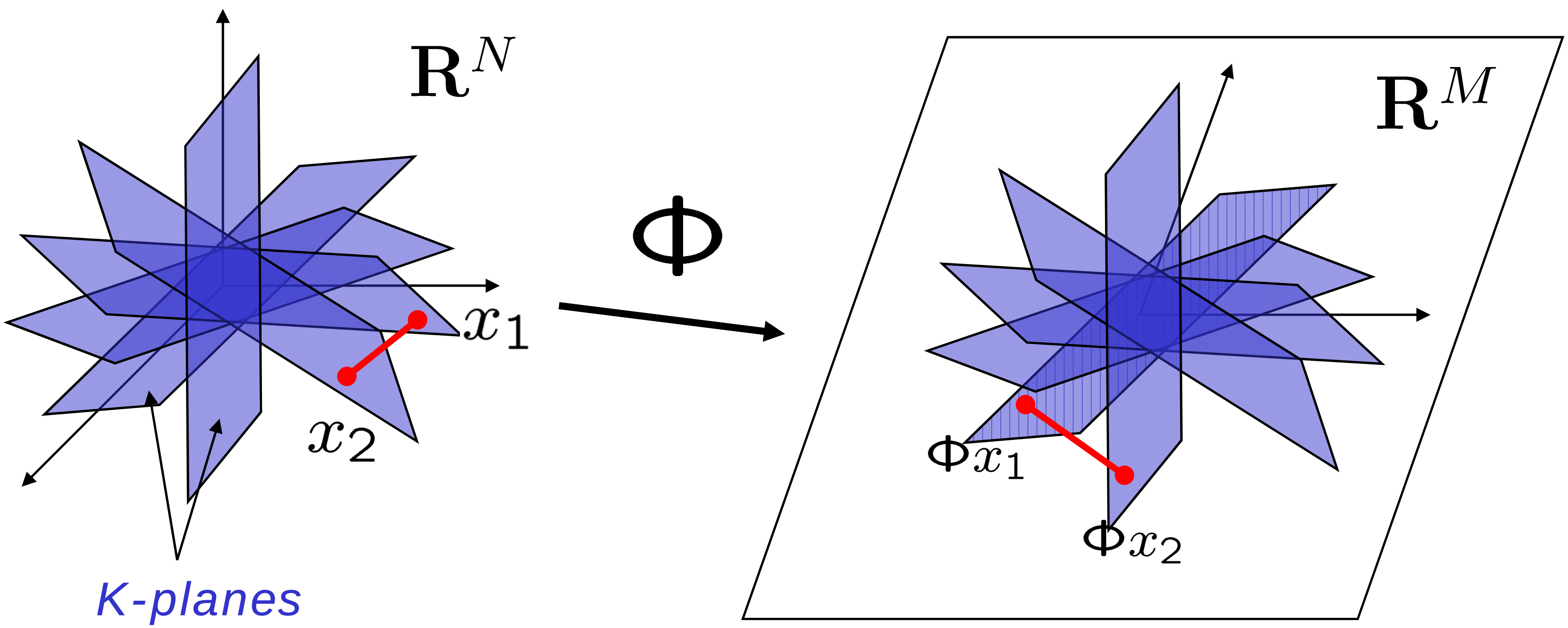




Restricted Isometry Property (RIP)

- Preserve the structure of sparse/compressible signals
- RIP of order *2K* implies: for all *K*-sparse x_1 and x_2

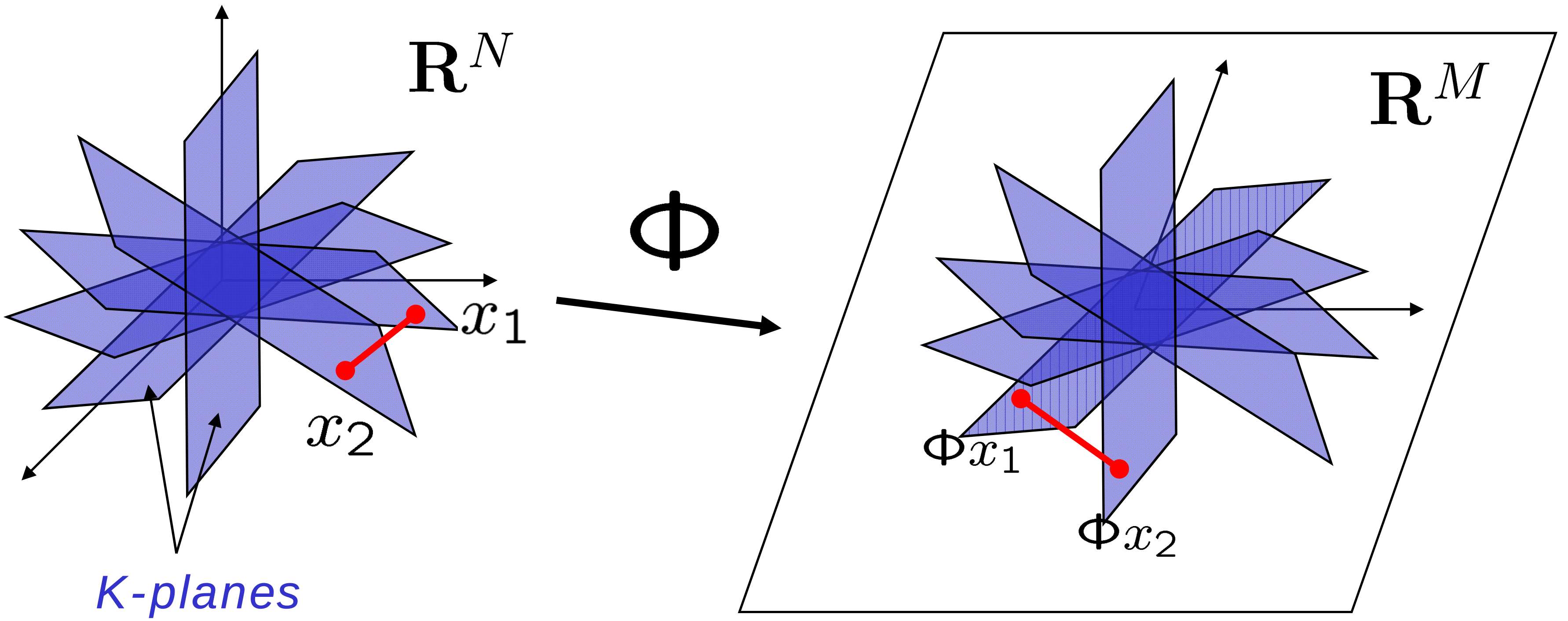
$$(1-\delta_{2K}) \le \frac{\|\Phi x_1 - \Phi x_2\|_2^2}{\|x_1 - x_2\|_2^2} \le (1+\delta_{2K})$$



Restricted Isometry Property (RIP)

- Preserve the structure of sparse/compressible signals
- Random subGaussian (iid Gaussian, Bernoulli) matrix has the RIP with high probability if

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



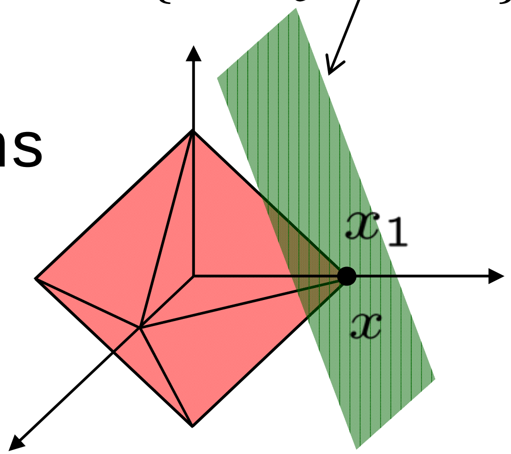
Recovery Algorithms

- **Goal:** given $y = \Phi x + e$

  recover x

- ℓ_1 and convex optimization formulations
  - basis pursuit, Dantzig selector, Lasso, ...

$$\widehat{x} = \arg\min \|x\|_1 \text{ s.t. } y = \Phi x$$



- Greedy algorithms
  - orthogonal matching pursuit,
    iterative thresholding (IT),
    compressive sensing matching pursuit (CoSaMP)
  - at their core: **iterative sparse approximation**

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

Performance of Recovery

- Using ℓ_1 methods, IT, CoSaMP
- **Sparse signals**
  - noise-free measurements: exact recovery
  - noisy measurements: stable recovery
- **Compressible signals**
  - recovery as good as K-sparse approximation

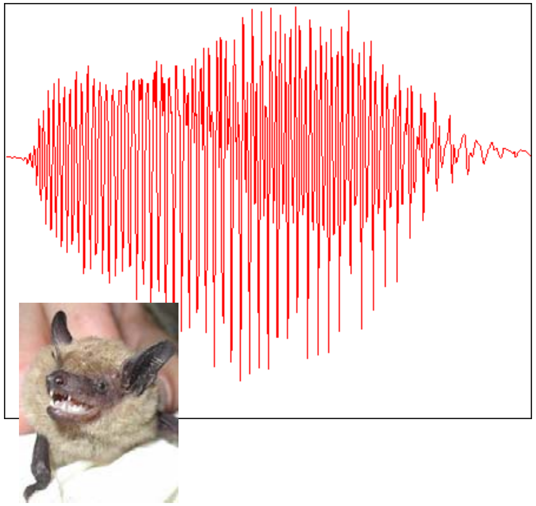
$$\|x - \widehat{x}\|_{\ell_2} \leq C_1 \|x - x_K\|_{\ell_2} + C_2 \frac{\|x - x_K\|_{\ell_1}}{K^{1/2}} + C_3 \epsilon$$

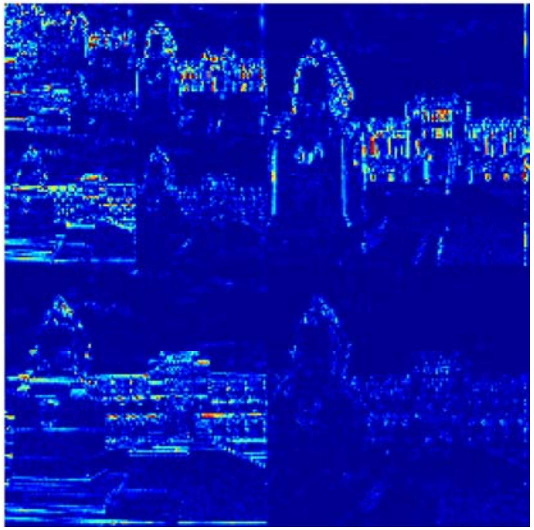
CS recovery error — signal K-term approx error — noise

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

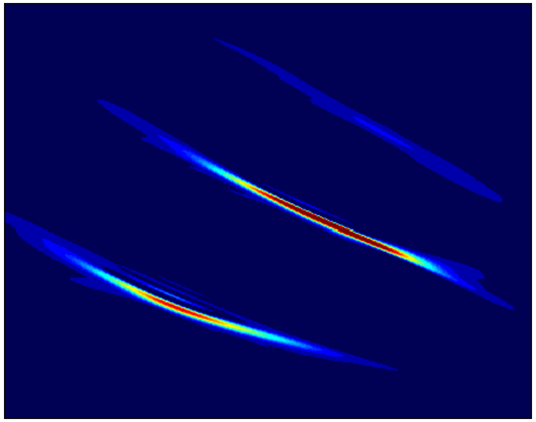
From Sparsity
to
Model-based (*structured*) Sparsity

Sparse Models

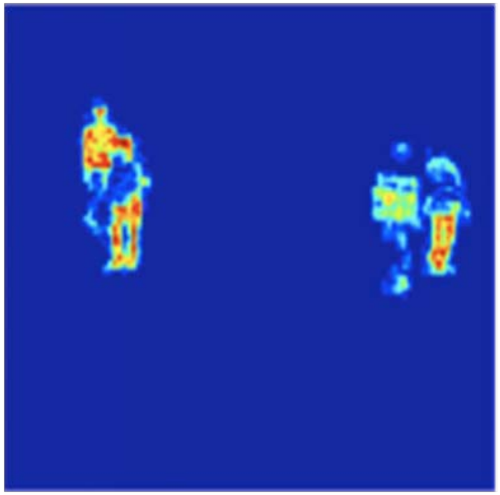




wavelets:
natural images



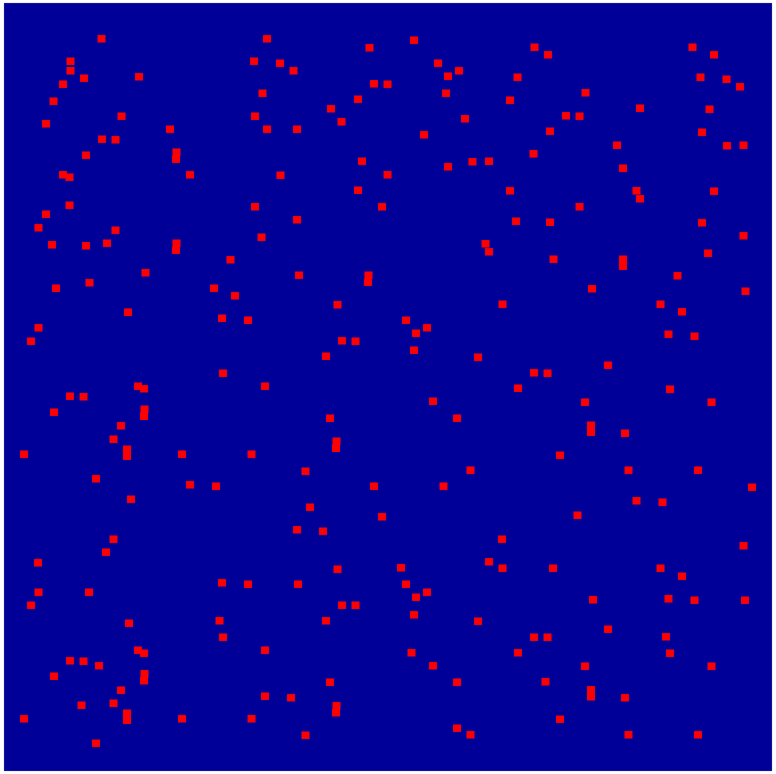
Gabor atoms:
chirps/tones



pixels:
background subtracted
images

Sparse Models

- Sparse/compressible signal model captures **simplistic primary structure**

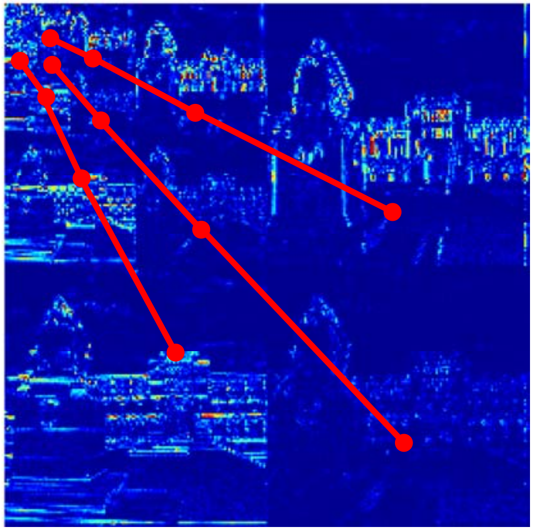


sparse image

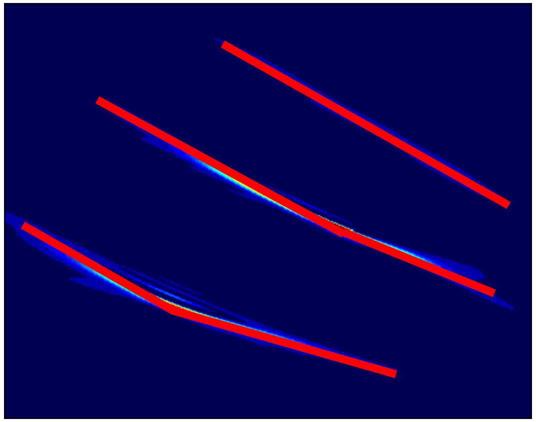
Beyond Sparse Models

- Sparse/compressible signal model captures
  simplistic primary structure

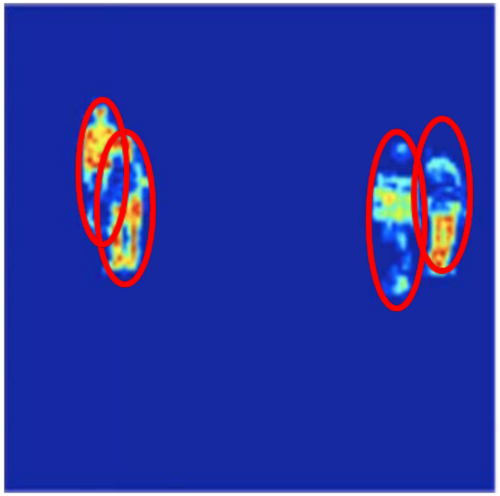
- Modern compression/processing algorithms capture
  richer secondary coefficient structure



wavelets:
natural images



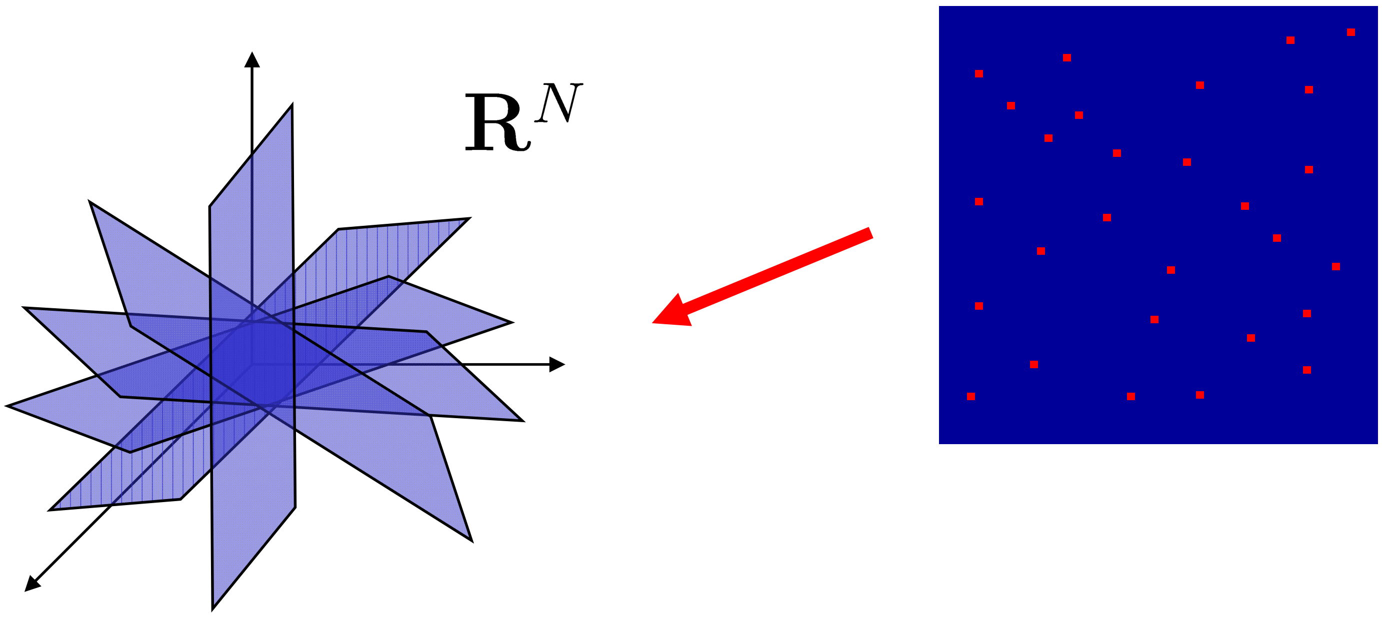
Gabor atoms:
chirps/tones



pixels:
background subtracted images

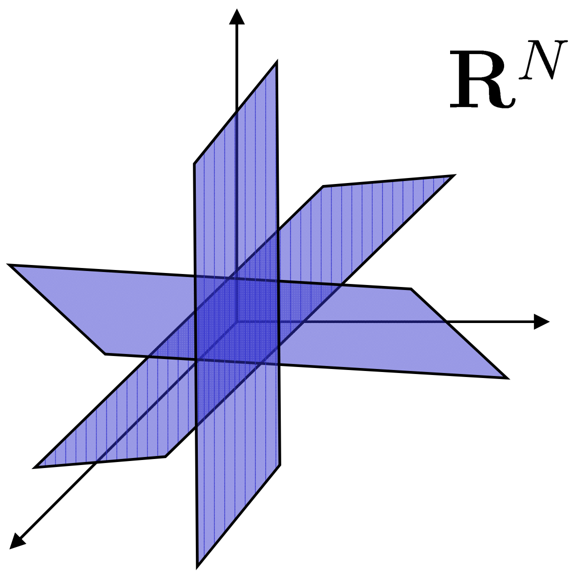
Sparse Signals

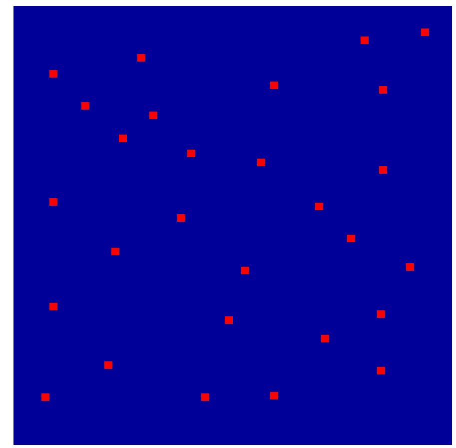
- Defn: ***K*-sparse signals** comprise a particular set of *K*-dim canonical subspaces

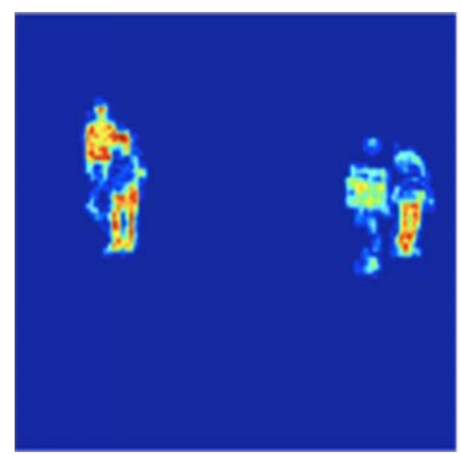


Model-Sparse Signals

- Defn: A ***K*-sparse signal *model*** comprises a particular (*reduced*) set of *K*-dim canonical subspaces

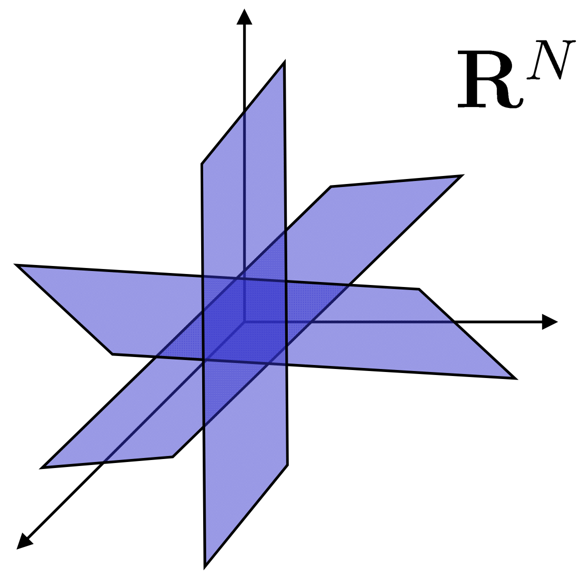






Model-Sparse Signals

- Defn: A ***K*-sparse signal *model*** comprises a particular (*reduced*) set of *K*-dim canonical subspaces

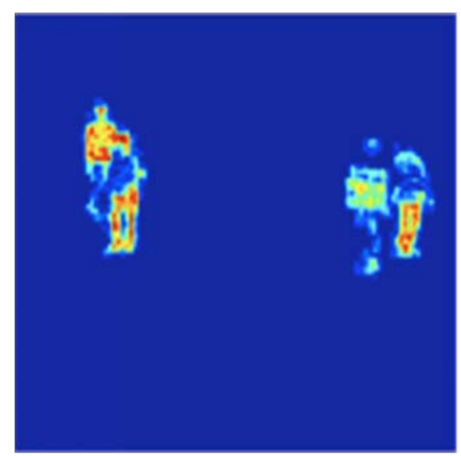


- **Structured subspaces**

< > fewer subspaces

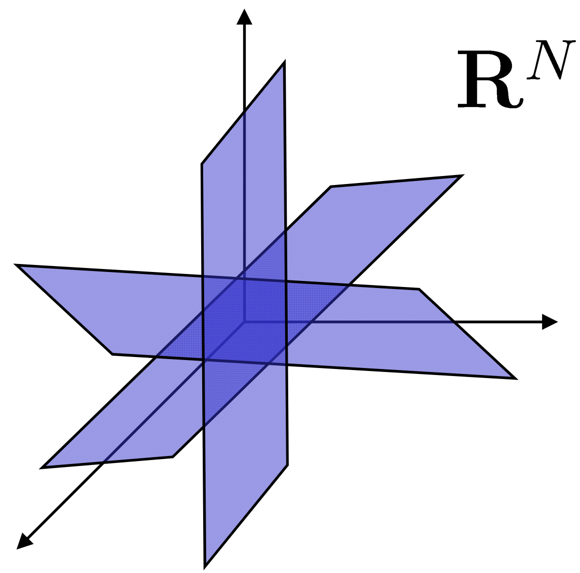
< > relaxed RIP

< > ***fewer measurements***



Model-Sparse Signals

- Defn: A ***K*-sparse signal *model*** comprises a particular (*reduced*) set of *K*-dim canonical subspaces

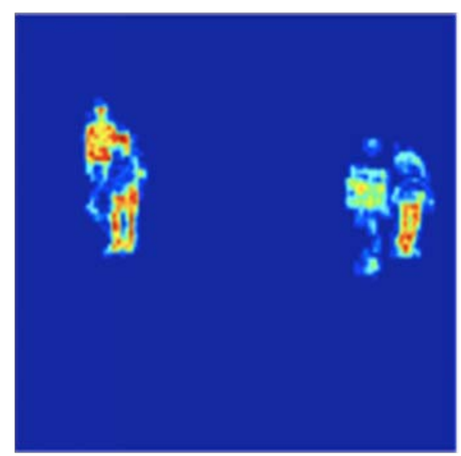


- **Structured subspaces**

< > increased signal discrimination

< > ***improved recovery perf.***

< > ***faster recovery***

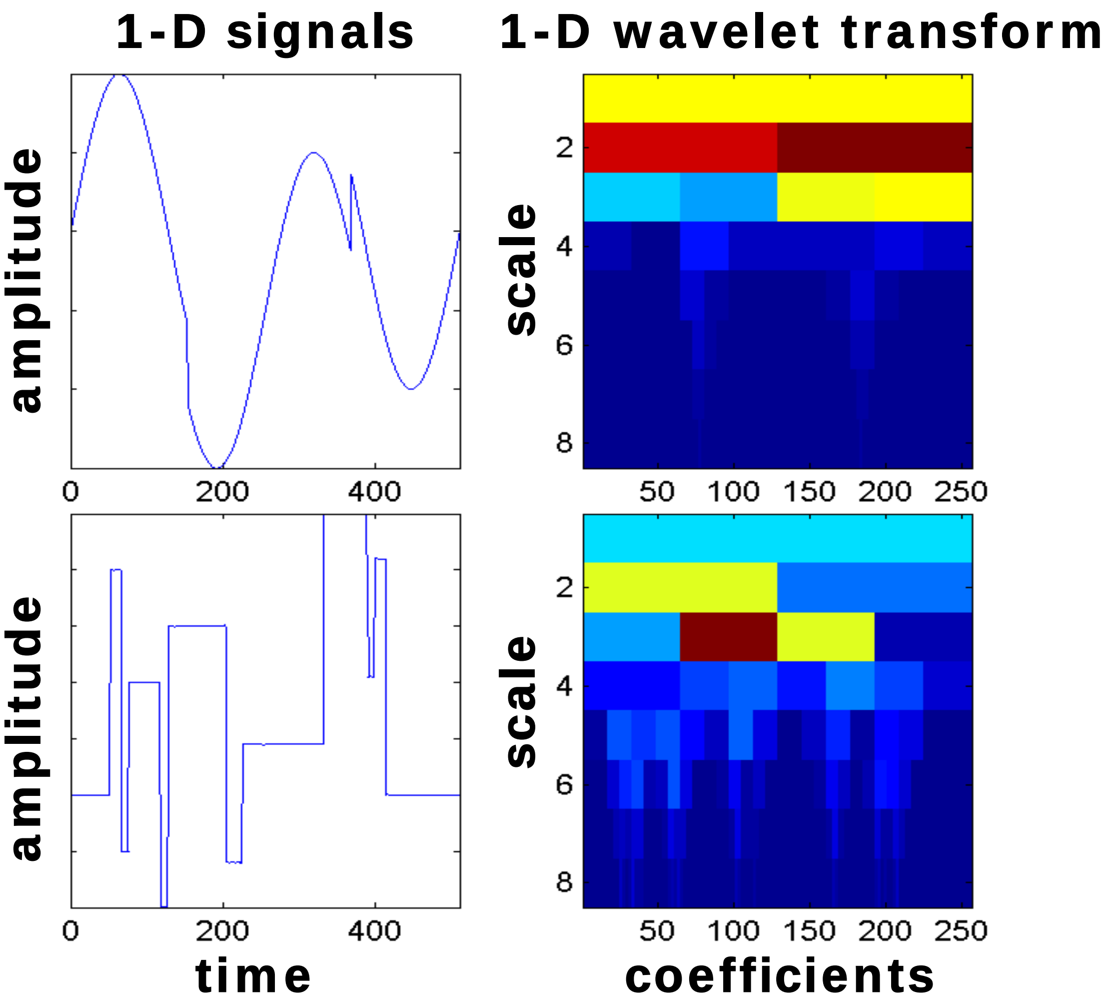


Model-based CS

Running Example:
Tree-Sparse Signals

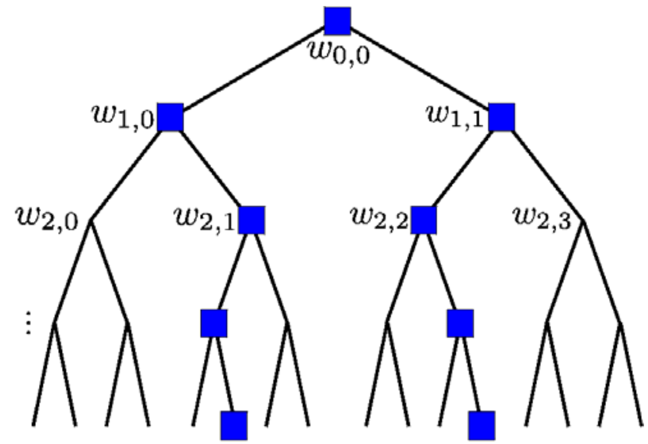
[Baraniuk, **VC**, Duarte, Hegde]

Wavelet Sparse

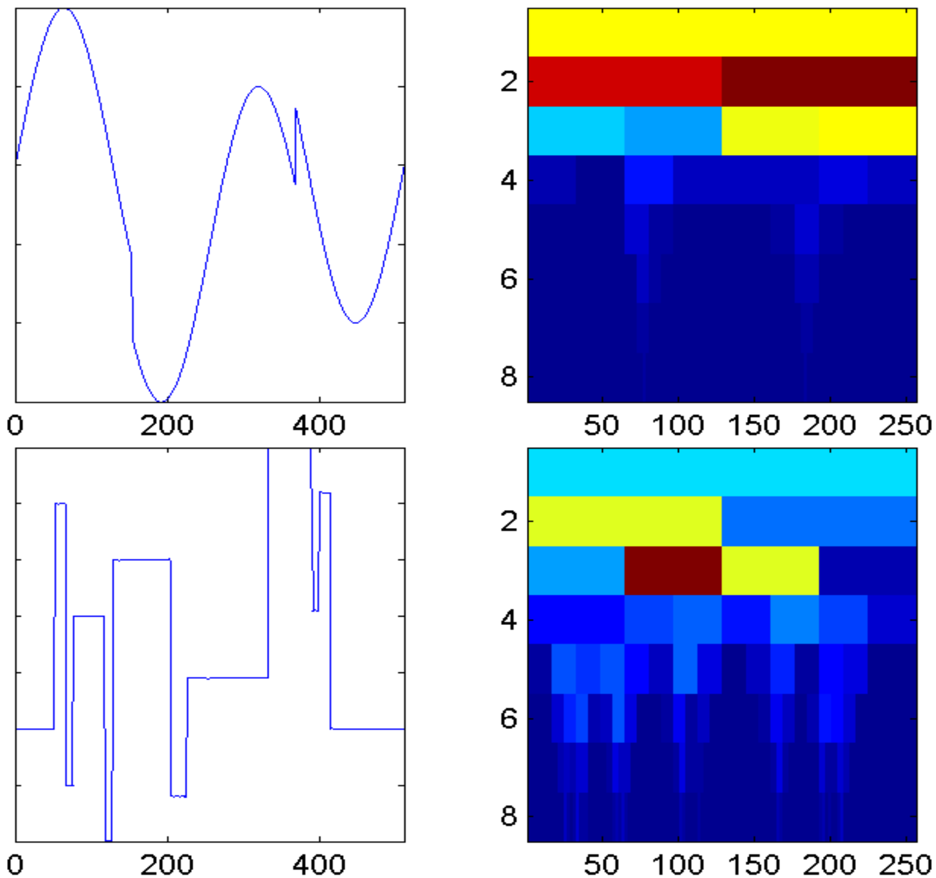


- Typical of wavelet transforms of natural signals and images (piecewise smooth)

Tree-Sparse

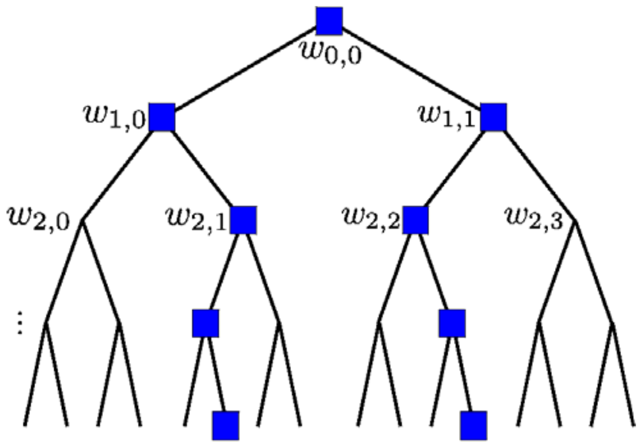


- **Model:** *K*-sparse coefficients
  + significant coefficients
  lie on a **rooted subtree**



- Typical of wavelet transforms
  of natural signals
  and images
  (piecewise smooth)

Tree-Sparse



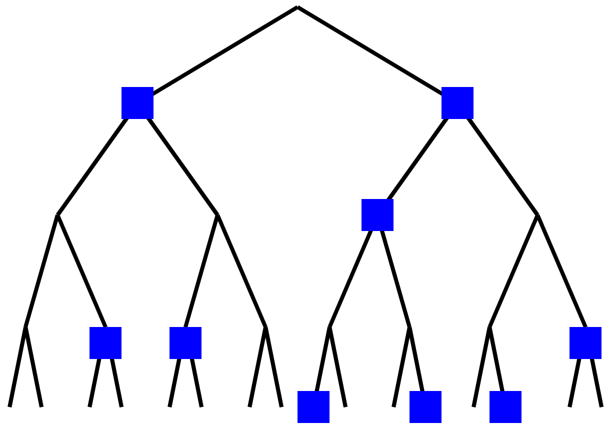
- **Model:** K-sparse coefficients
  + significant coefficients lie on a rooted subtree

- **Sparse approx:** find best set of coefficients
  - sorting
  - hard thresholding

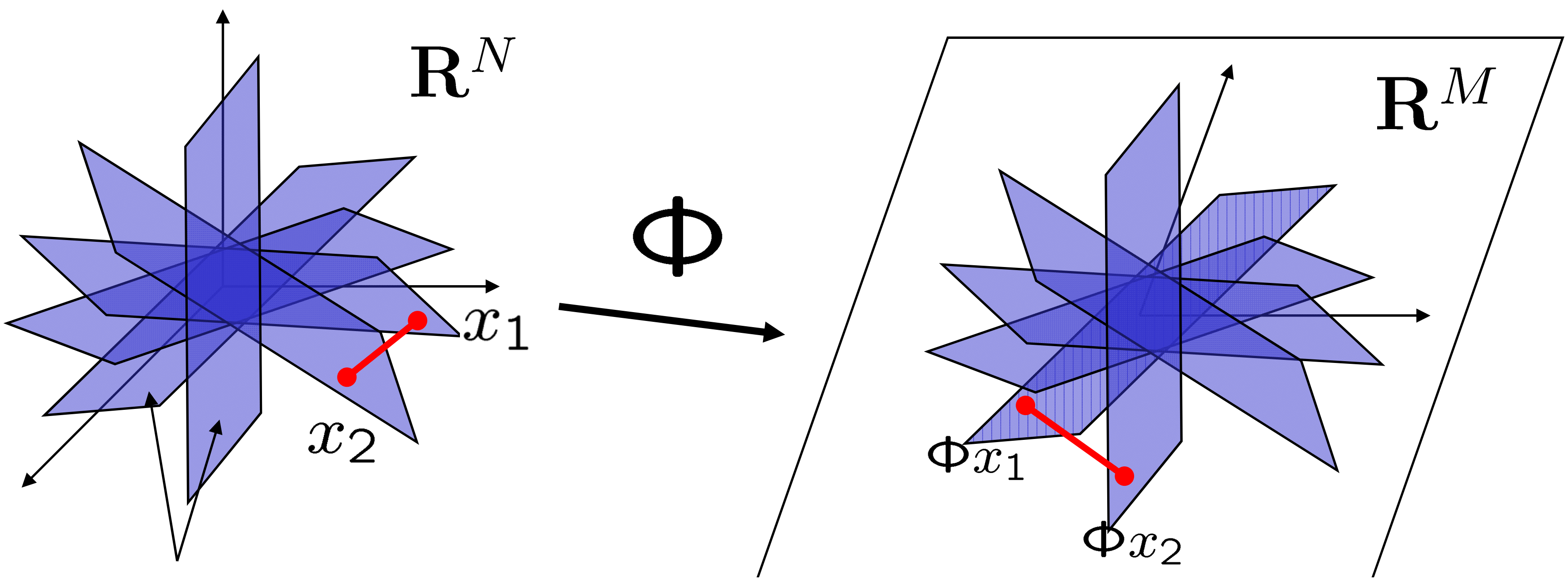
- **Tree-sparse approx:** find best rooted subtree of coefficients
  - CSSA [Baraniuk]
  - dynamic programming [Donoho]

Sparse

- **Model:** *K*-sparse coefficients



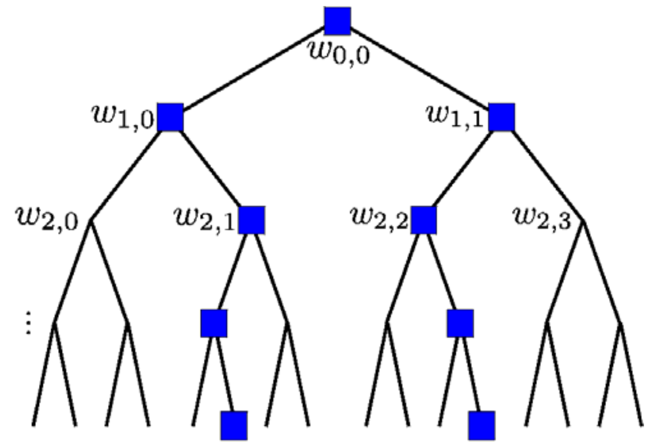
- **RIP:** stable embedding



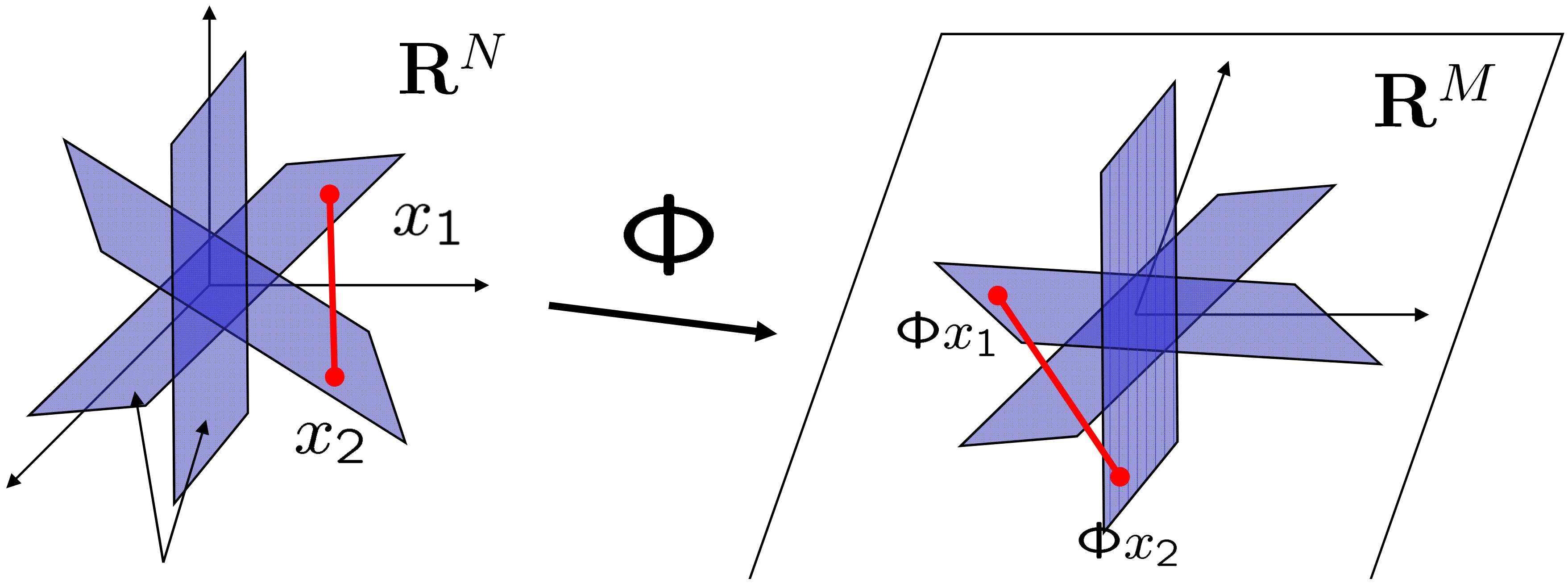
$$M = O(K \log(N/K))$$

Tree-Sparse

- **Model:** *K*-sparse coefficients
  + significant coefficients
  lie on a rooted subtree

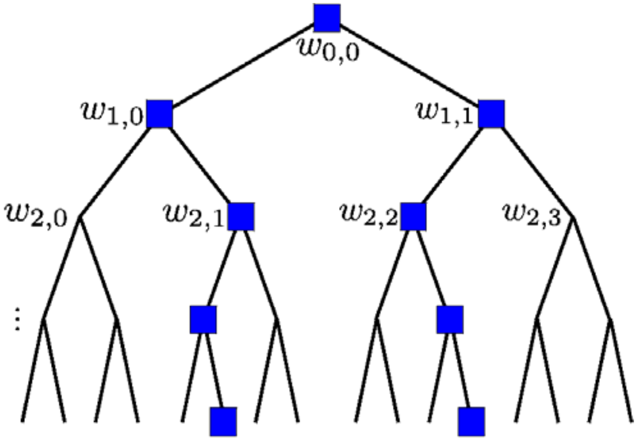


- **Tree-RIP:** stable embedding



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

Tree-Sparse



- **Model:** *K*-sparse coefficients
  + significant coefficients lie on a rooted subtree

- **Tree-RIP:** stable embedding

- **Recovery:** new model based algorithms

[**VC**, Duarte, Hegde, Baraniuk; Baraniuk, **VC**, Duarte, Hegde]

Standard CS Recovery

- **Iterative Thresholding**

[Nowak, Figueiredo; Kingsbury, Reeves; Daubechies, Defrise, De Mol; Blumensath, Davies; ...]

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

Model-based CS Recovery

- **Iterative Model Thresholding**

[**VC**, Duarte, Hegde, Baraniuk; Baraniuk, **VC**, Duarte, Hegde]

Given $y = \Phi x$, recover a model sparse $x \in \mathcal{M}$

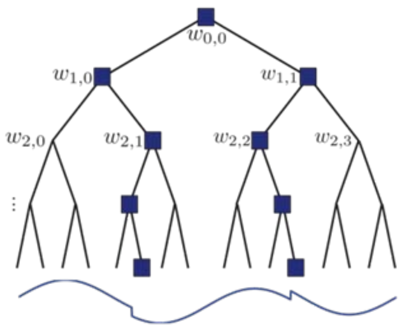
initialize: $\widehat{x}_0 = 0$, $r = y$, $i = 0$

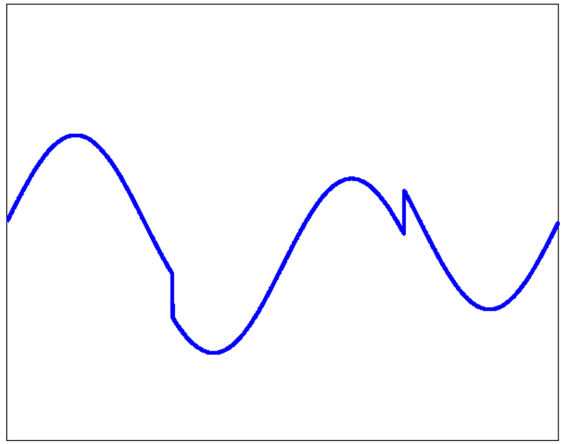
iteration:

- $i \leftarrow i + 1$
- $b \leftarrow \widehat{x}_{i-1} + \Phi^T r$ **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**

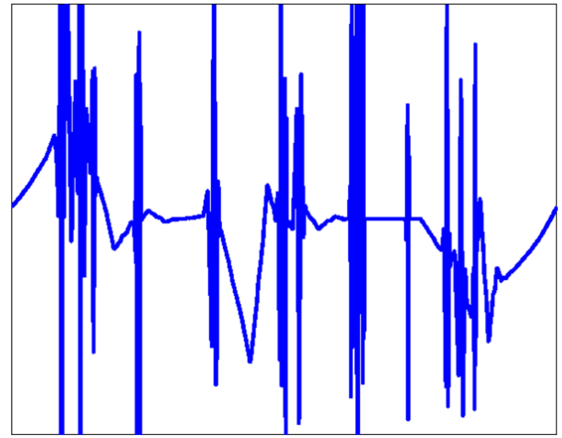
return: $\widehat{x} \leftarrow \widehat{x}_i$

Tree-Sparse Signal Recovery



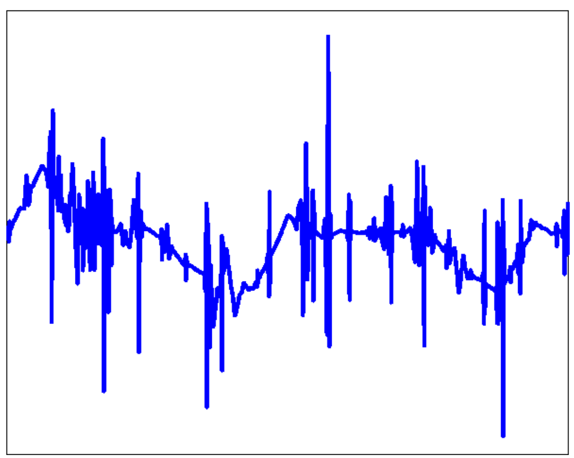


target signal

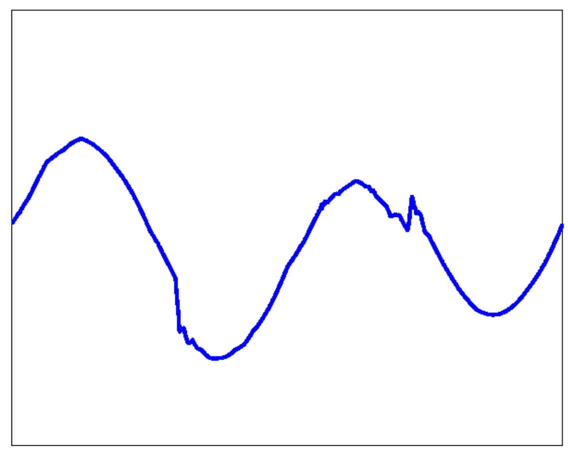


CoSaMP,
(MSE=1.12)

N=1024
M=80



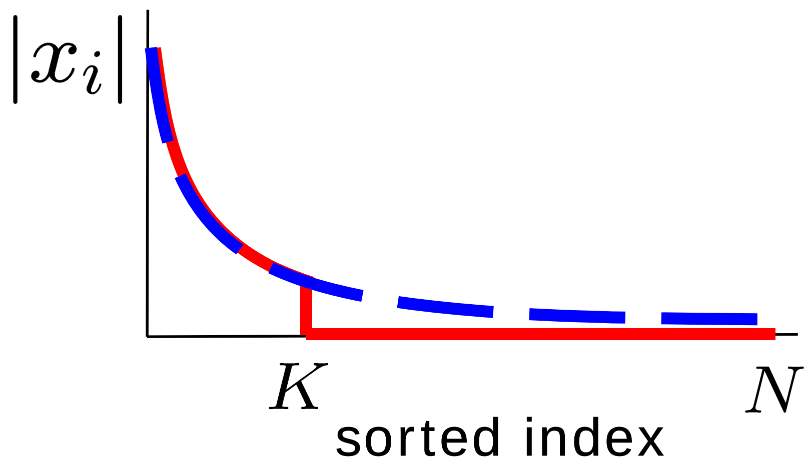
L1-minimization
(MSE=0.751)



Tree-sparse CoSaMP
(MSE=0.037)

Compressible Signals

- Real-world signals are compressible, not sparse
- Recall: **compressible** <> well approximated by sparse
  - compressible signals lie close to a union of subspaces
  - ie: approximation error decays rapidly as $K \to \infty$
- If Φ has RIP, then both sparse and compressible signals are stably recoverable



Model-Compressible Signals

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

Model-Compressible Signals

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

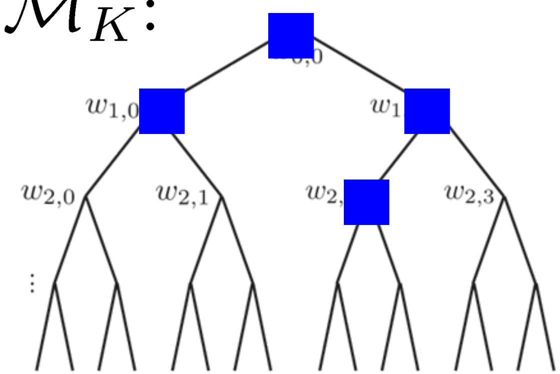
- While model-RIP enables stable model-sparse recovery, **model-RIP is *not* sufficient for stable model-compressible recovery at $\mathcal{O}(K)$!**

Stable Recovery

- Stable model-compressible signal recovery at $\mathcal{O}(K)$ requires that Φ have both:
  - RIP + **Restricted Amplification Property**

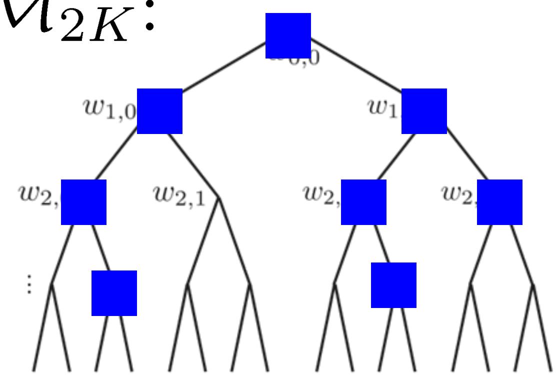
- **RAmP:** controls nonisometry of Φ in the approximation's **residual subspaces**

$\mathcal{M}_K$:



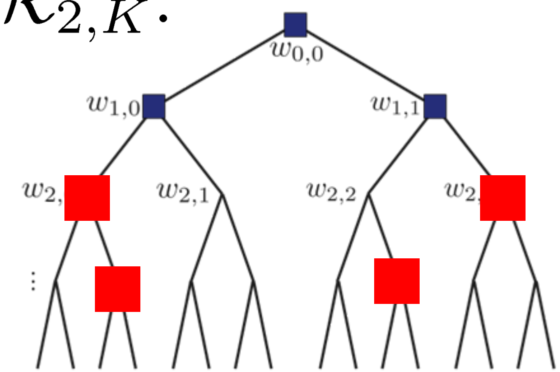
optimal K-term model recovery (error controlled by Φ RIP)

$\mathcal{M}_{2K}$:



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

$\mathcal{R}_{2,K}$:



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

Tree-RIP, Tree-RAmP

Theorem: An *M*x*N* iid subgaussian random matrix has the Tree(*K*)-RIP if

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

Theorem: An *M*x*N* iid subgaussian random matrix has the Tree(*K*)-RAmP if

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

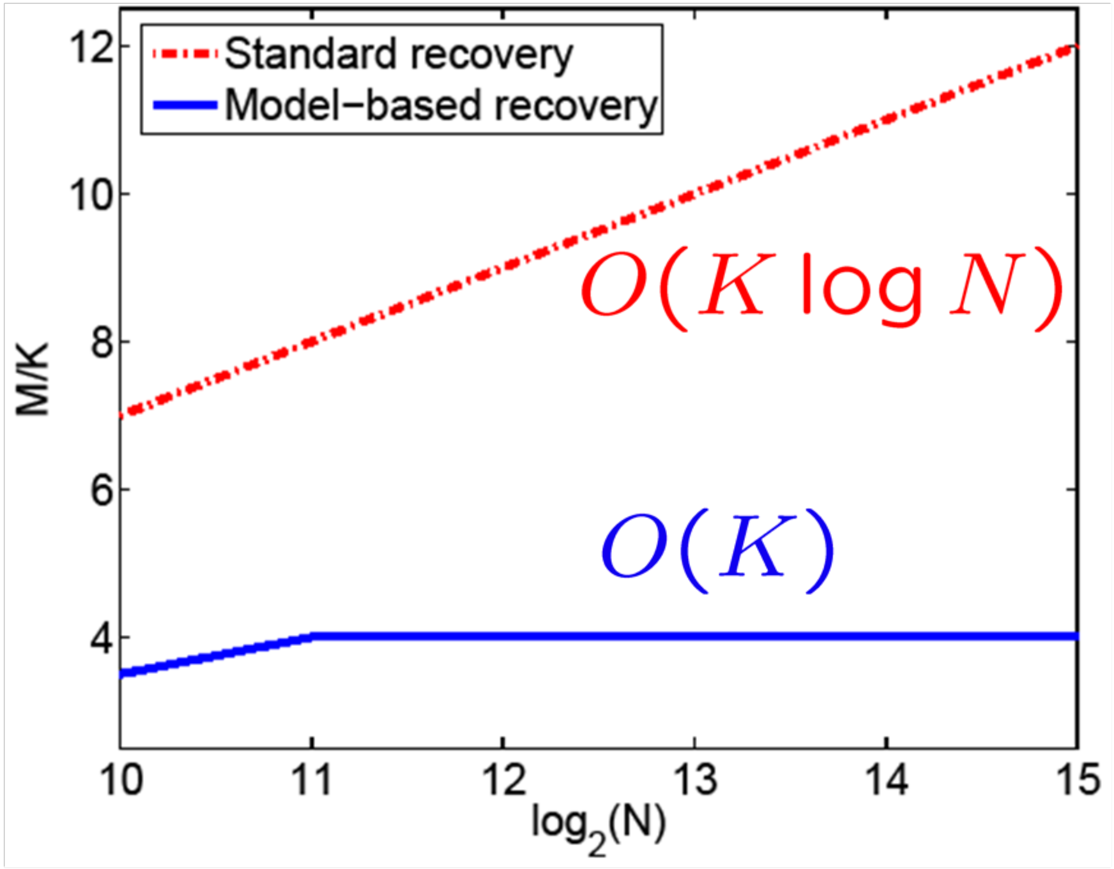
with probability $1 - \exp(-t)$.

Simulation

- Number samples for correct recovery

- Piecewise cubic signals + wavelets

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



Performance of Recovery

- Using model-based IT, 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

$$\|x - \widehat{x}\|_{\ell_2} \leq C_1 \|x - x_{\mathcal{M}_K}\|_{\ell_2} + C_2 \frac{\|x - x_{\mathcal{M}_K}\|_{\ell_1}}{K^{1/2}} + C_3 \epsilon$$

CS recovery error

signal *K*-term model approx error

noise

[Baraniuk, **VC**, Duarte, Hegde]

Other Useful Models

- When the model-based framework makes sense:
  - **model** with
    - fast approximation algorithm
  - sensing **matrix** Φ with
    - model-RIP
    - model-RAmP

Other Useful Models

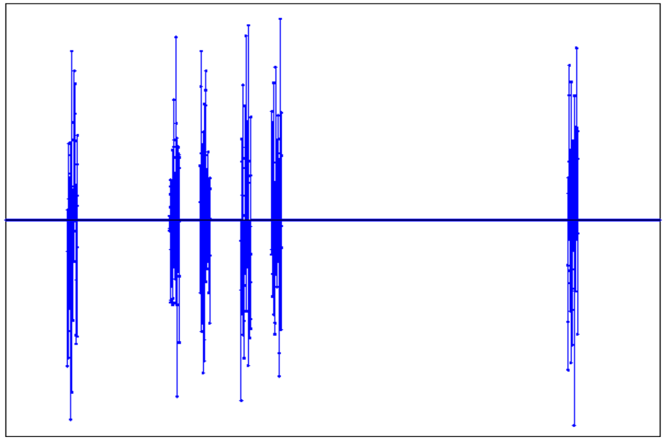
- When the model-based framework makes sense:
  - **model** with
    - fast approximation algorithm
  - sensing **matrix** Φ with
    - model-RIP
    - model-RAmP

- **Ex: block sparsity / signal ensembles**
  [Tropp, Gilbert, Strauss], [Stojnic, Parvaresh, Hassibi], [Eldar, Mishali], [Baron, Duarte et al], [Baraniuk, **VC**, Duarte, Hegde]

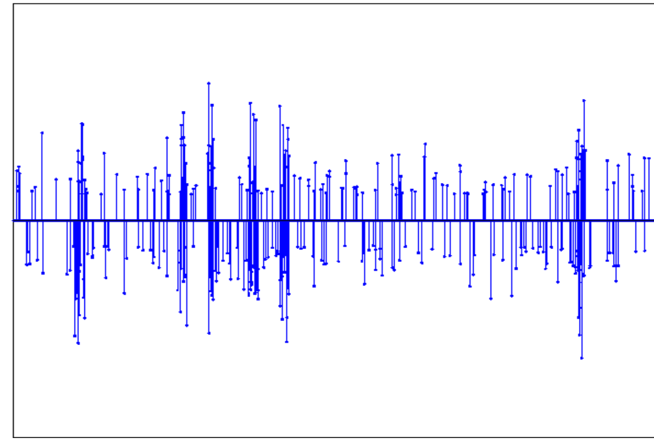
- **Ex: clustered signals**
  [**VC**, Duarte, Hegde, Baraniuk], [**VC**, Indyk, Hegde, Baraniuk]

- **Ex: neuronal spike trains**
  [Hegde, Duarte, **VC**] – Best paper award at SPARS'09

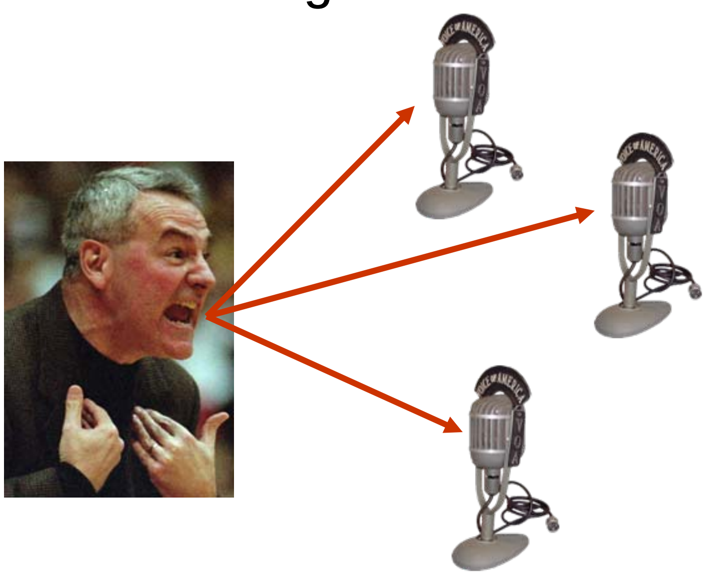
Block-Sparse Signal

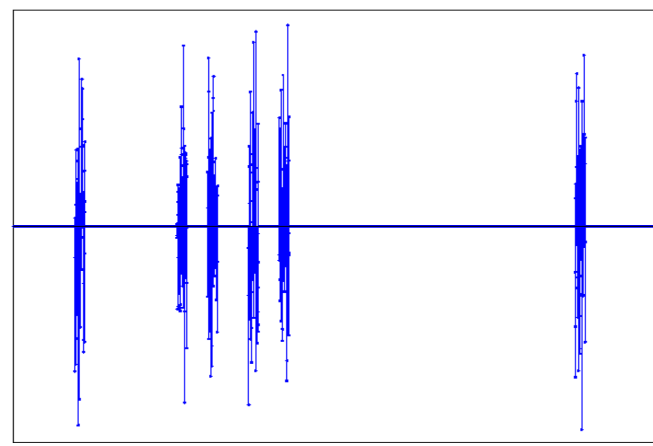


target



CoSaMP (MSE = 0.723)

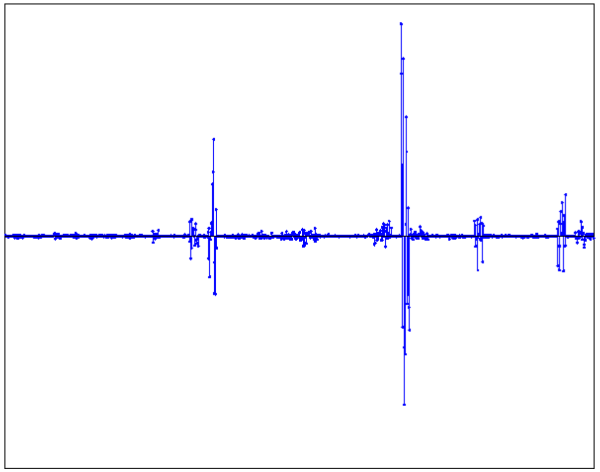




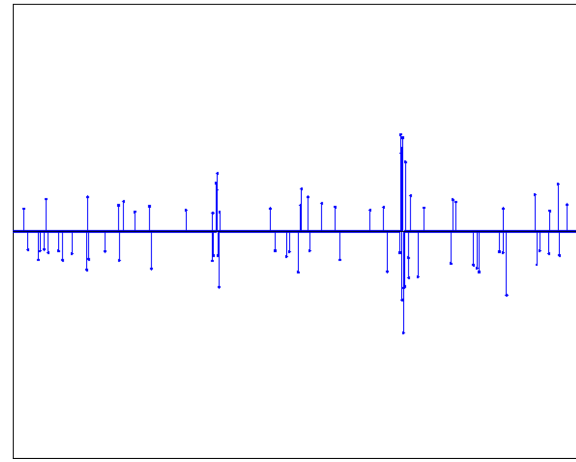
block-sparse model recovery
(MSE=0.015)

Blocks are pre-specified.

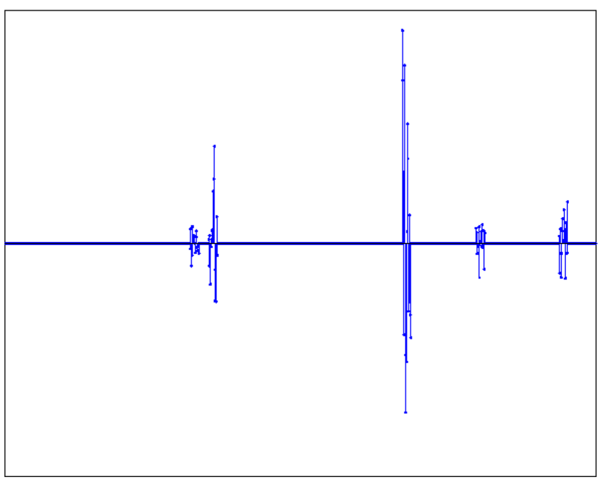
Block-Compressible Signal



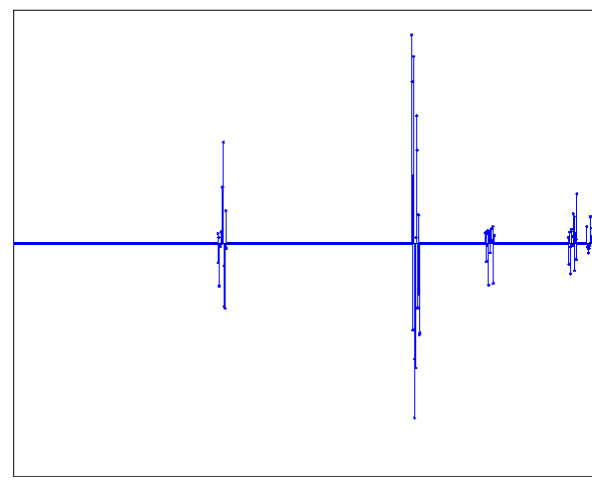
target



CoSaMP (MSE=0.711)



best 5-block approximation
(MSE=0.116)



block-sparse recovery
(MSE=0.195)

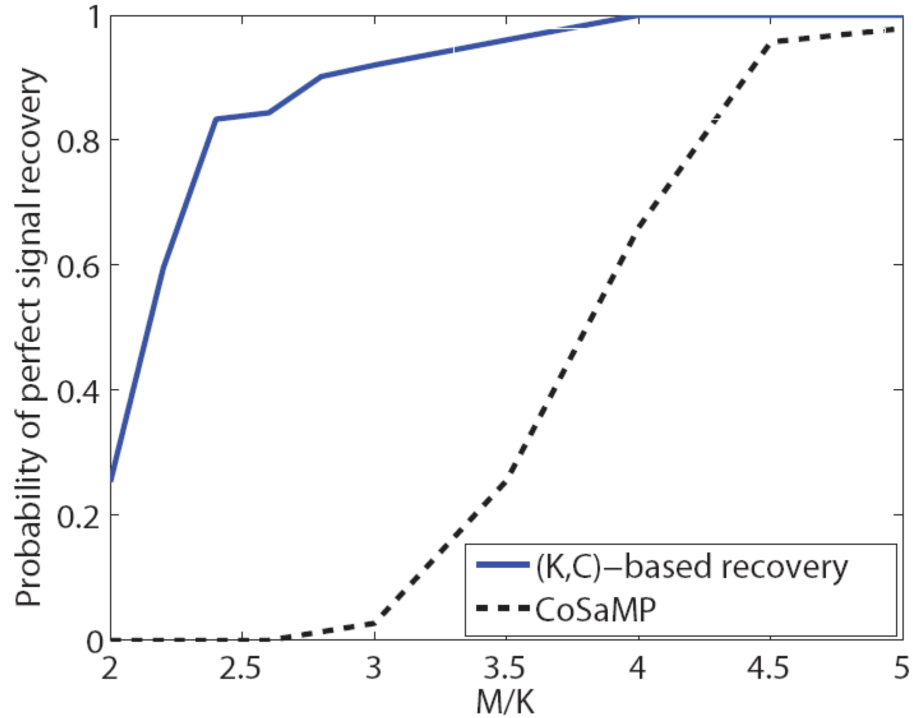
Clustered Sparsity

- **(K,C) sparse signals** (1-D)
  - K-sparse within at most C clusters

- For stable recovery (model-RIP + RAmP) $M = O(K + C\log(N/C))$

[**VC**, Indyk, Hedge, Baraniuk]

- Model approximation using **dynamic programming**

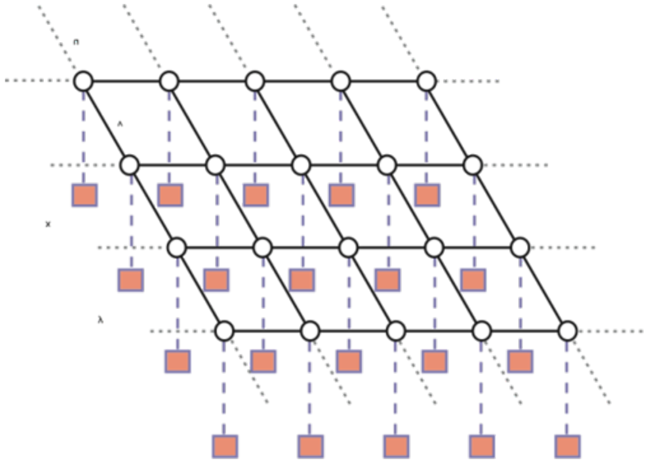


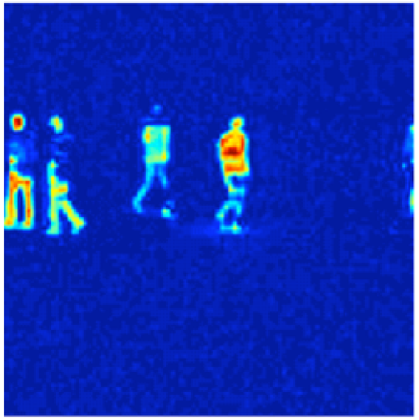
- Includes **block sparsity** as a special case

Clustered Sparsity

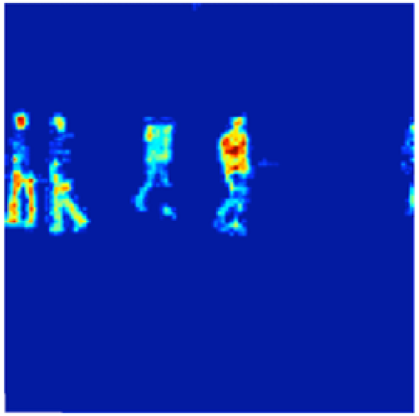
- Model clustering of significant pixels in space domain using **graphical model** (MRF)
- Ising model approximation via **graph cuts**

[**VC**, Duarte, Hedge, Baraniuk]

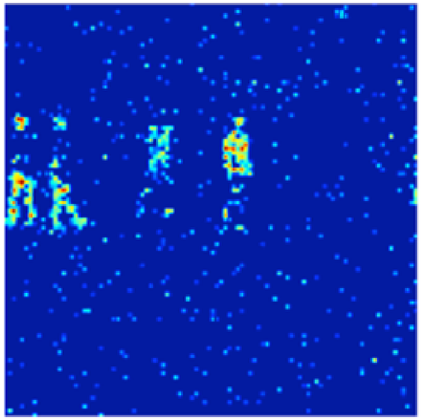




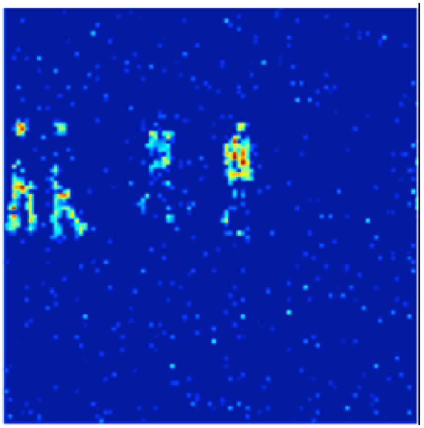
target



Ising-model recovery



CoSaMP recovery



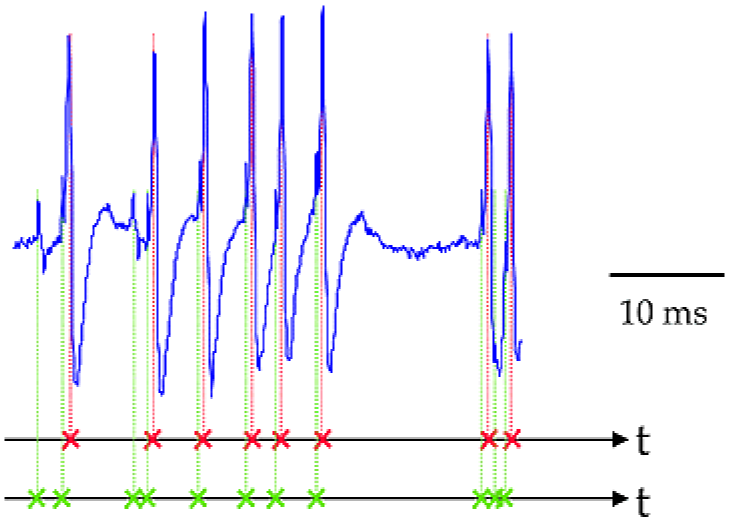
LP (FPC) recovery

Neuronal Spike Trains

- Model the firing process of a single neuron via 1D Poisson process with spike trains

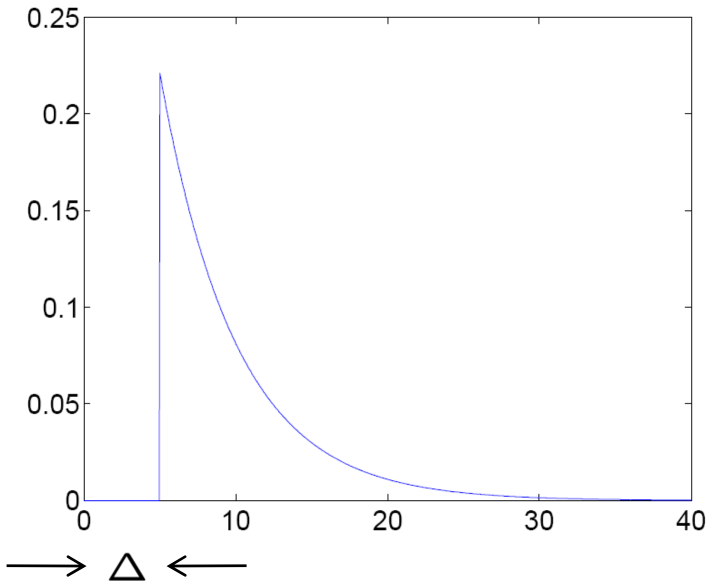
$$x_n = \sum_{k=1}^{K} \alpha_k \delta[n - n_k]$$

  - Exploit the **refractory period** of neurons



- Model approximation problem:

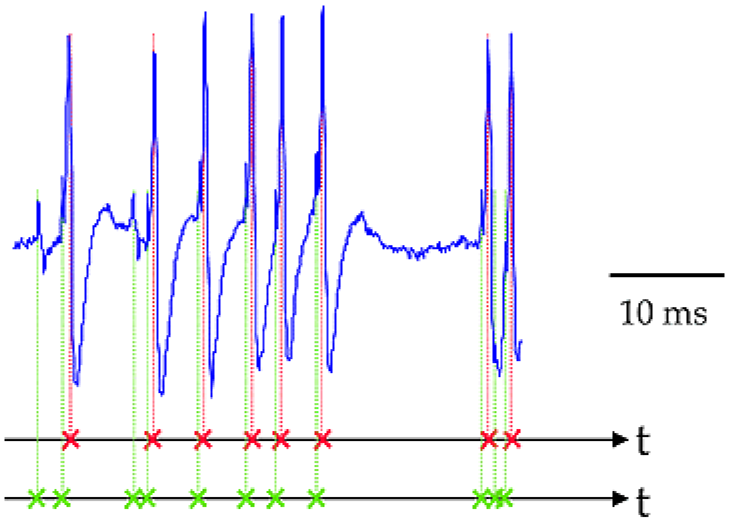
  - Find **a *K*-sparse signal** such that **its coefficients are separated by at least Δ**



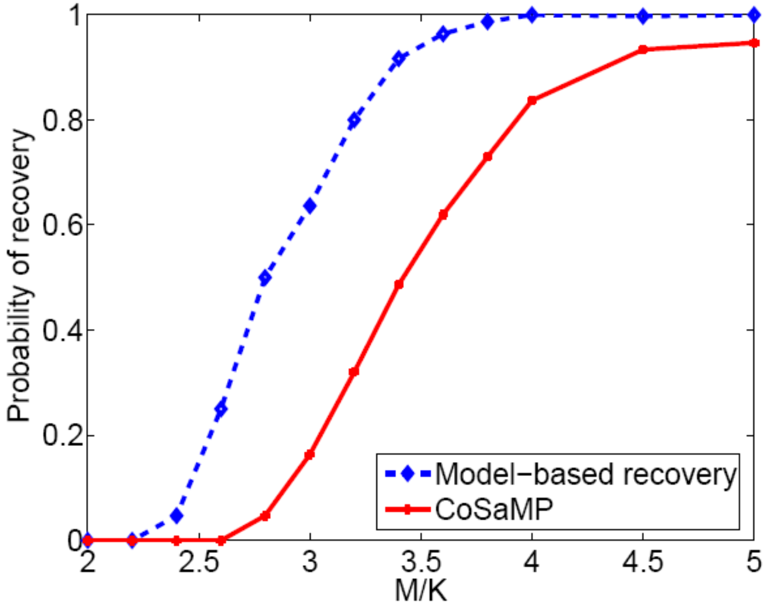
Neuronal Spike Trains

- Model the firing process of a single neuron via 1D Poisson process with spike trains
  - Stable recovery

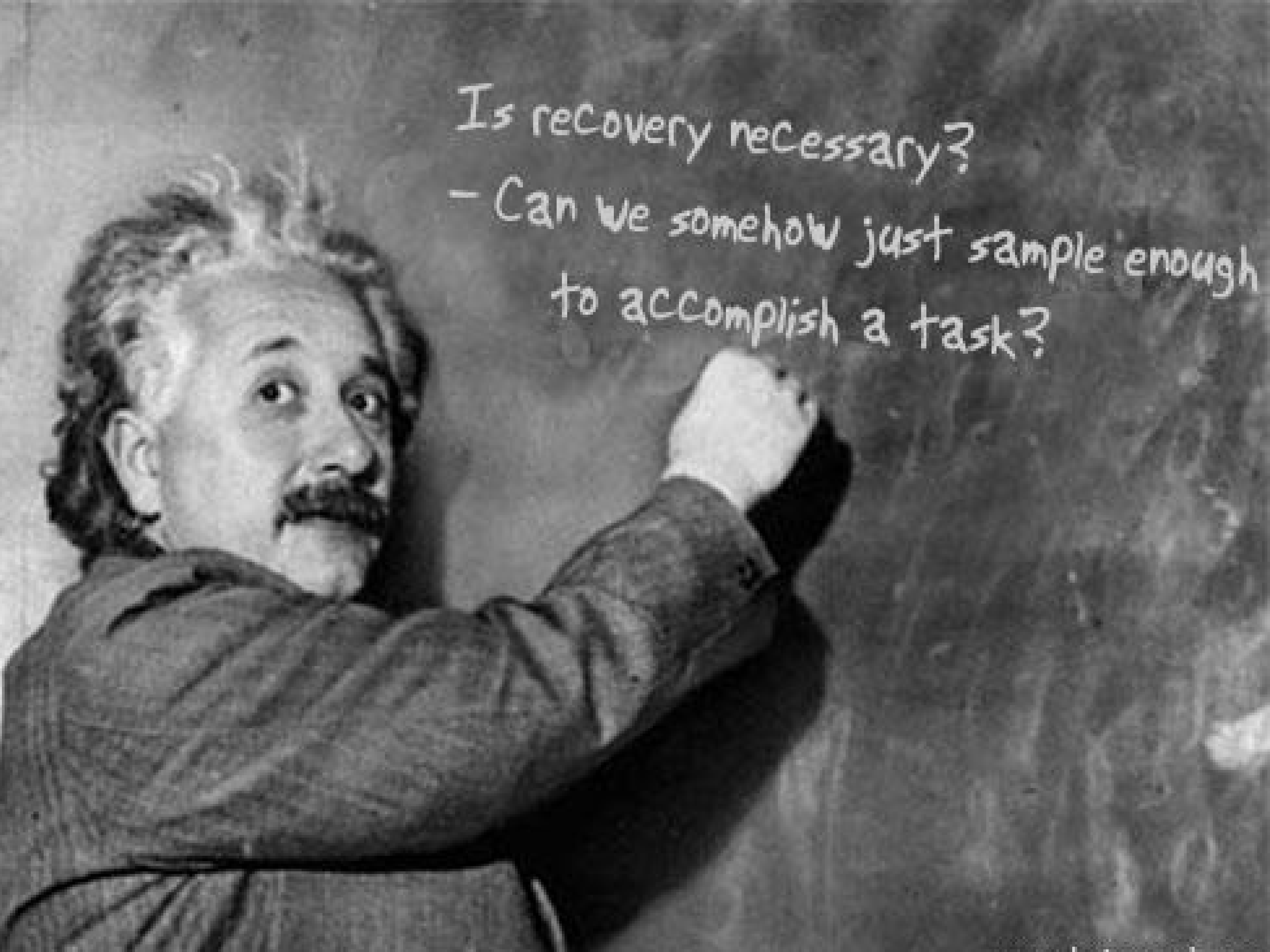
$$M = O(K \log(N/K - \Delta))$$



- Model approximation solution:
  - Integer program
  - **Efficient & provable solution** due to total unimodularity of linear constraint



[Hedge, Duarte, **VC**; SPARS'09]


Is recovery necessary?
- Can we somehow just sample enough
to accomplish a task?

Signal recovery is not **always** required.

ELVIS:

Enhanced
Localization
via
Incoherence and
Sparsity

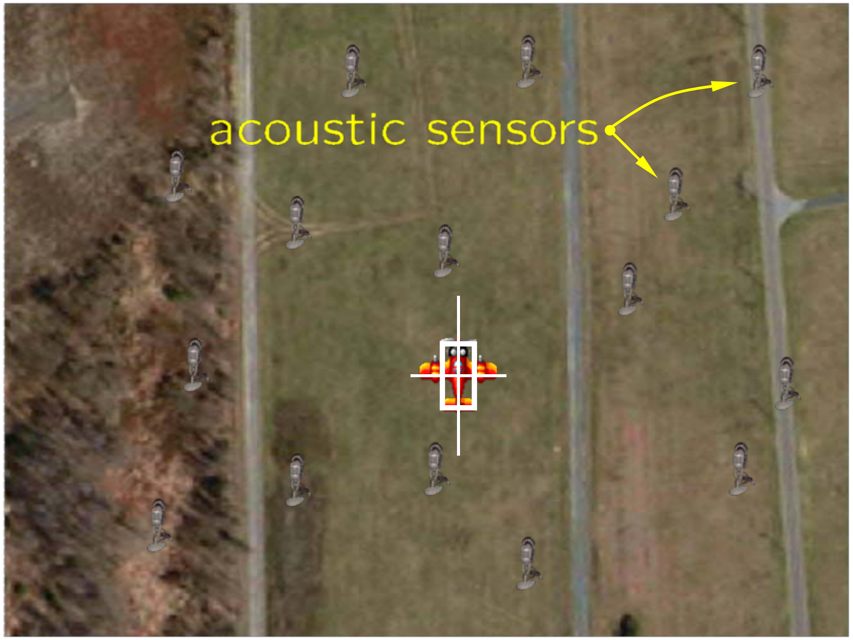
Localization Problem

- **Goal:** Localize targets by fusing measurements from a network of sensors

[**VC**, Duarte, Baraniuk; Model and Zibulevsky; **VC**, Gurbuz, McClellan, Chellappa; Malioutov, Cetin, and Willsky; Chen et al.]

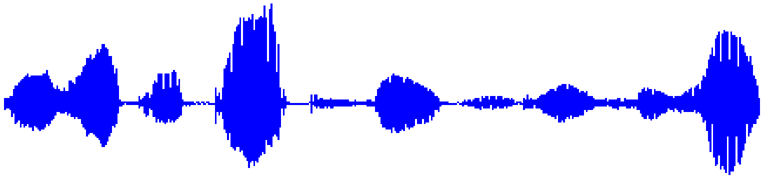
Localization Problem

- **Goal:** Localize targets by fusing measurements from a network of sensors
  - collect time signal data
  - communicate signals across the network
  - solve an optimization problem

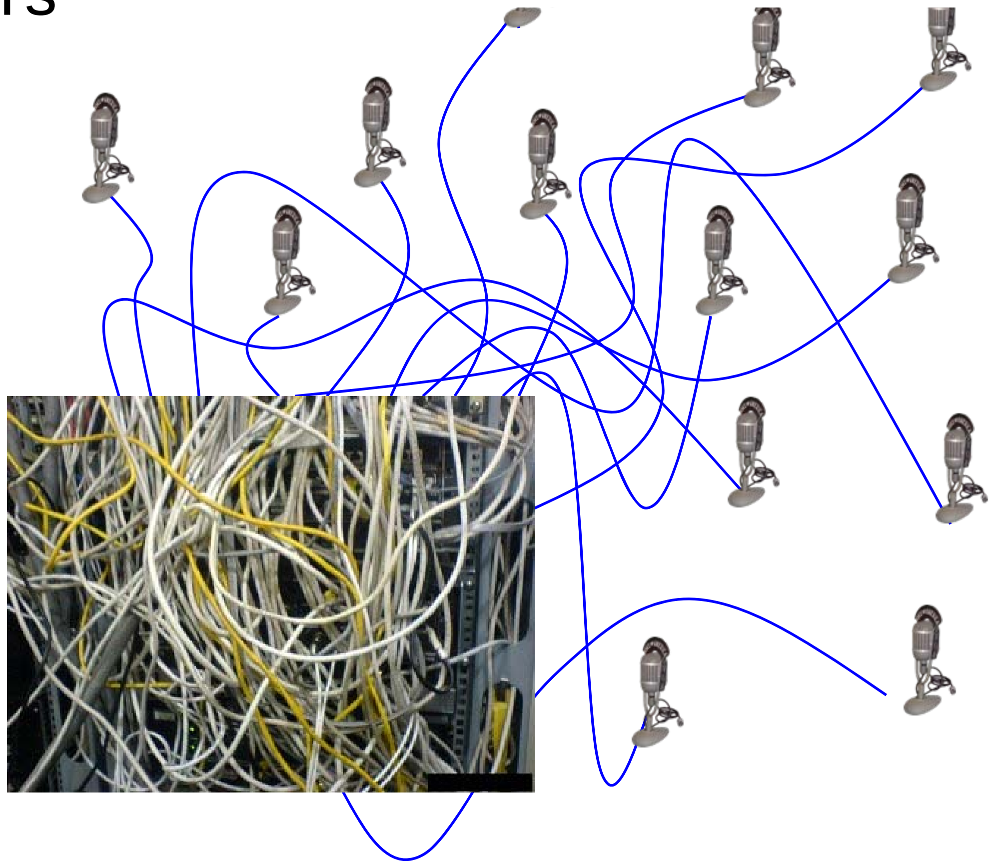


Bottlenecks

- **Goal:** Localize targets by fusing measurements from a network of sensors
  - collect time signal data
    - requires potentially high-rate (Nyquist) sampling
  - communicate signals across the network
    - potentially large communication burden
  - solve an optimization problem



Need compression



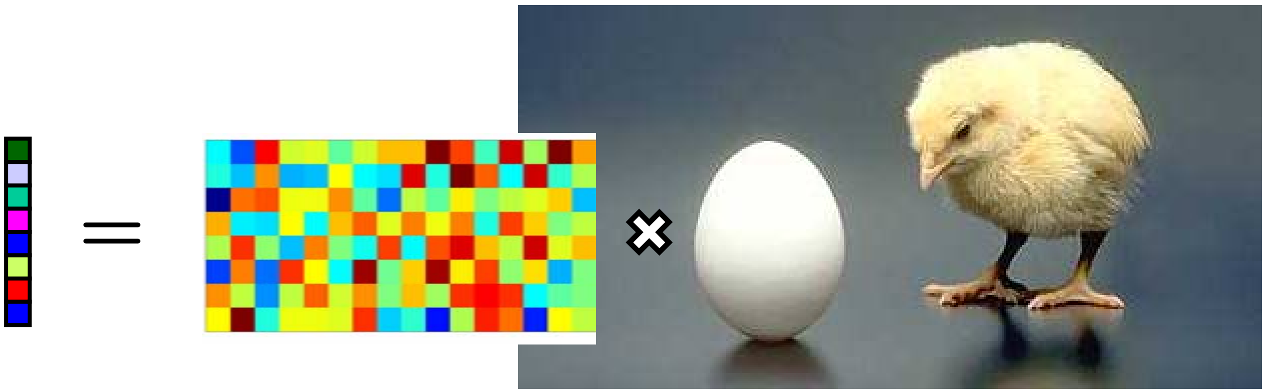
An Important Detail

- **Solve two entangled problems for localization**
  - Estimate source locations
  - Estimate source signals

ELVIS

- **Instead, solve one localization problem**
  - Estimate source locations
  - ~~Estimate source signals~~

by exploiting **random projections** of **observed signals**



ELVIS

- **Instead, solve one localization problem**
  - Estimate source locations by exploiting **random projections** of **observed signals**
  - ~~Estimate source signals~~

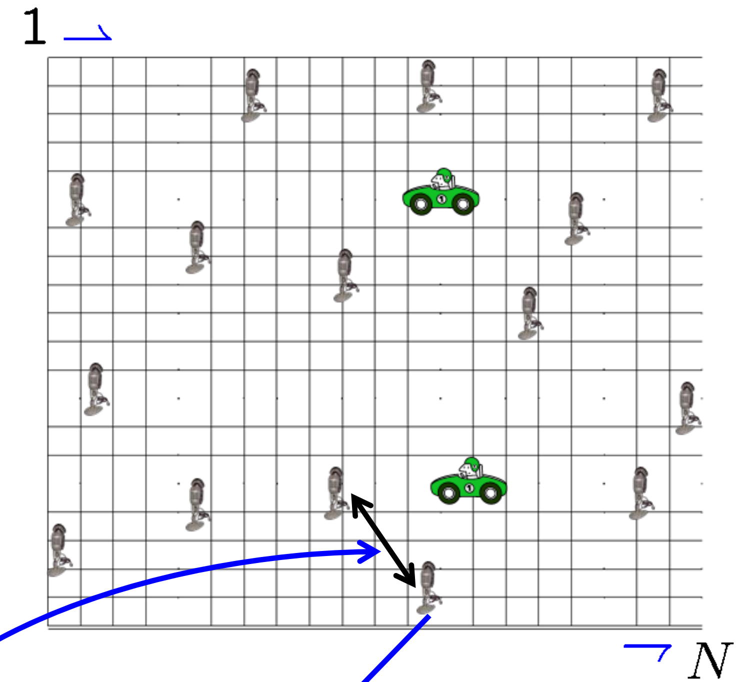
[**VC**, Boufounos, Baraniuk, Gilbert, Strauss]

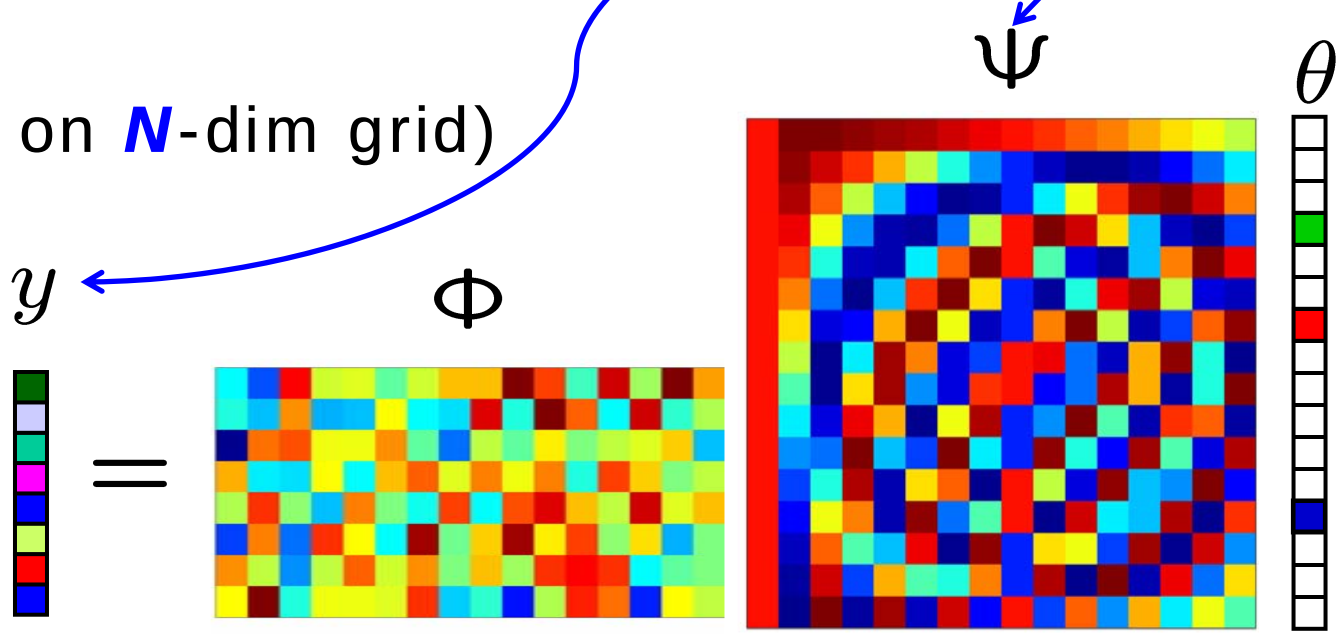
- Bayesian model order selection & MAP estimation results in a **decentralized sparse approximation framework** that leverages
  - Source sparsity
  - Incoherence of sources
  - Spatial sparsity of sources

ELVIS

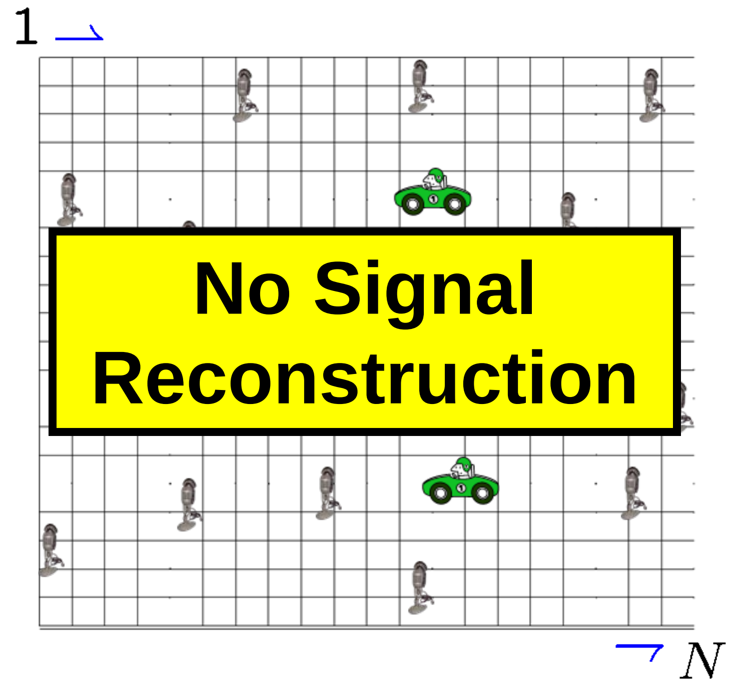
- **Use random projections** of observed signals **two ways**:
  - Create local sensor dictionaries that sparsify source locations
  - Create intersensor communication messages

(***K*** targets on ***N***-dim grid)





ELVIS

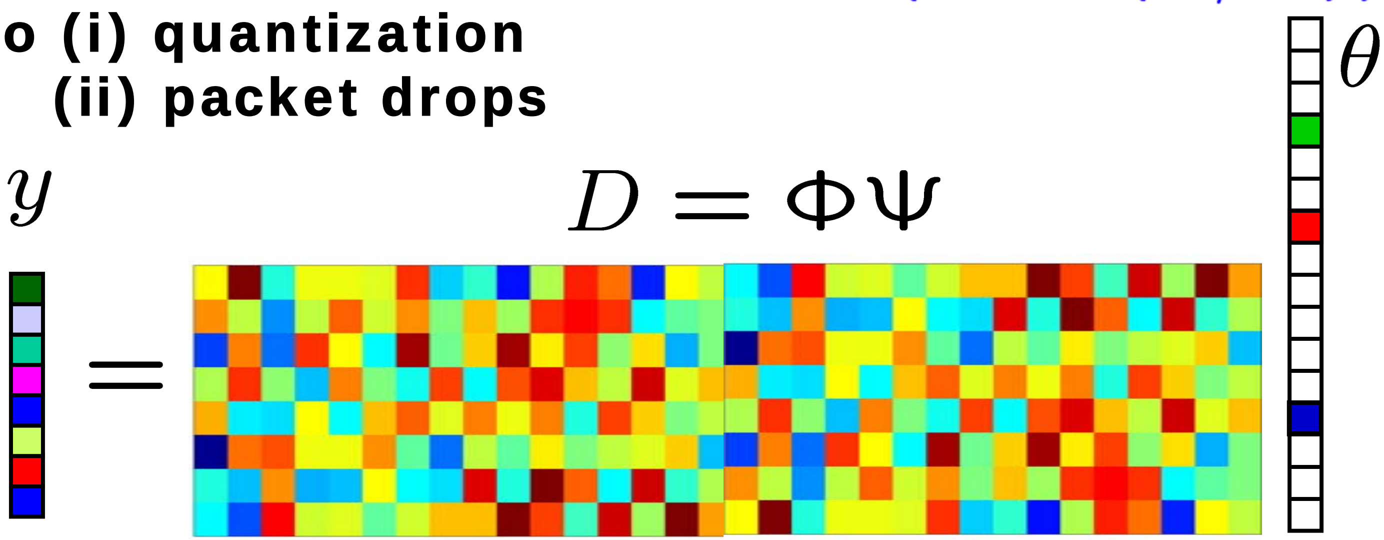


- **Use random projections** of observed signals **two ways**:
  - Create local sensor dictionaries that sparsify source locations

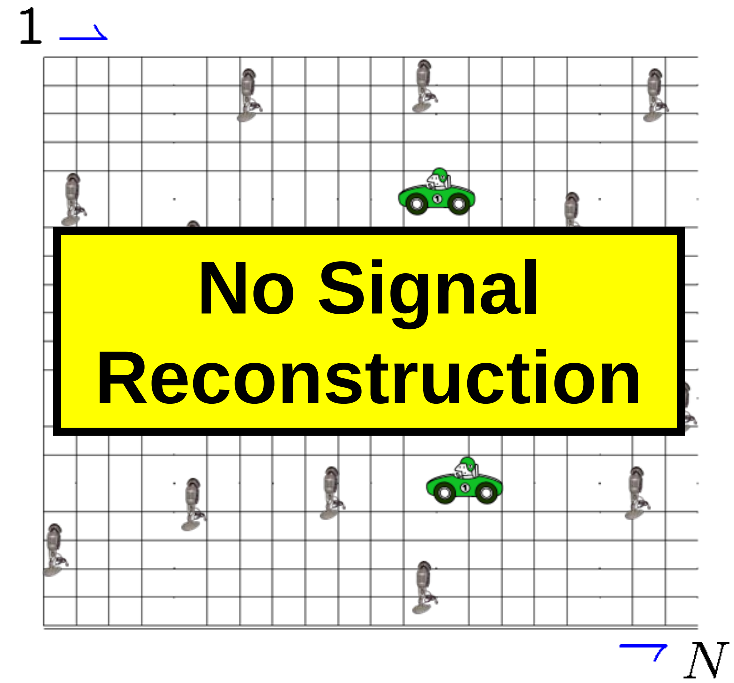
    sample at source sparsity
  - Create intersensor communication messages

    communicate at spatial sparsity $\sim O(K \log(N/K))$

    robust to (i) quantization
    (ii) packet drops



ELVIS



- **Use random projections** of observed signals **two ways**:
  - Create local sensor dictionaries that sparsify source locations

    sample at source sparsity
  - Create intersensor communication messages

    communicate at spatial sparsity $\sim O(K \log(N/K))$

    robust to (i) quantization

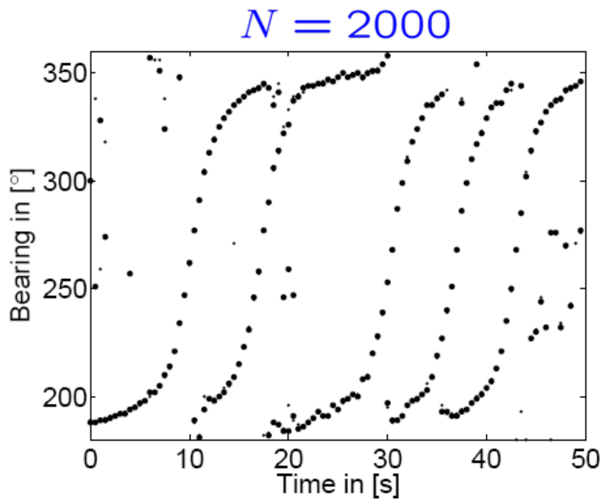
    (ii) packet drops

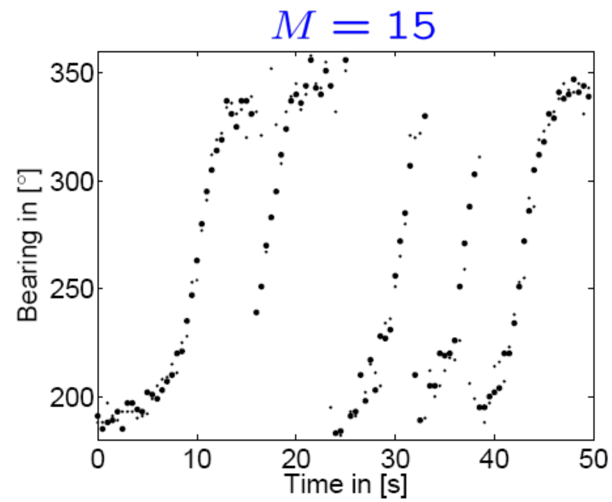
- Provable greedy estimation for ELVIS dictionaries

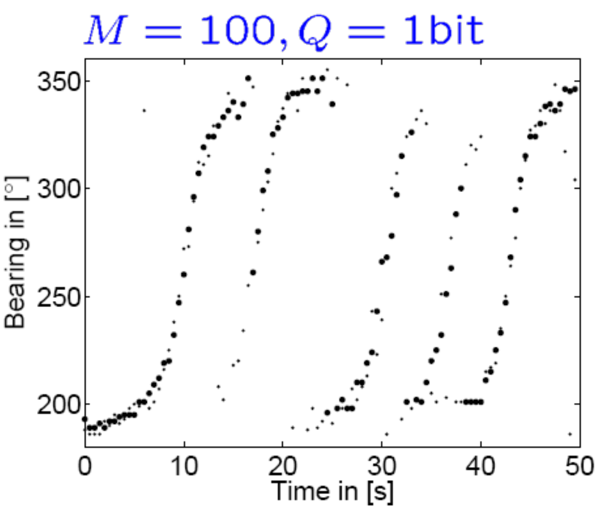
  Bearing pursuit

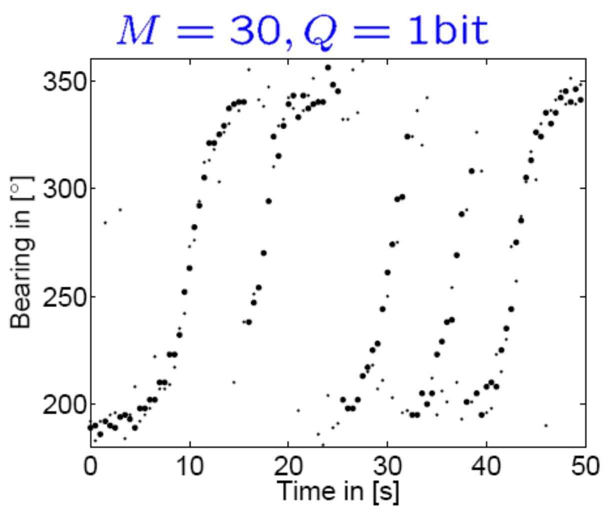
Field Data Results

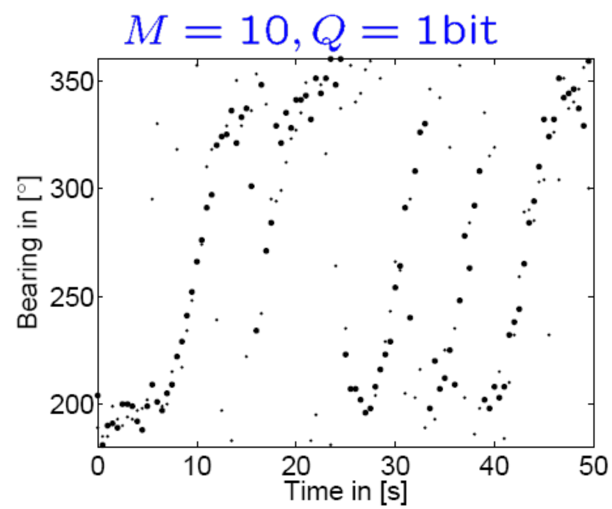
5 vehicle convoy





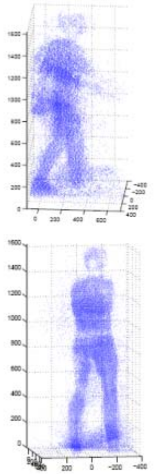






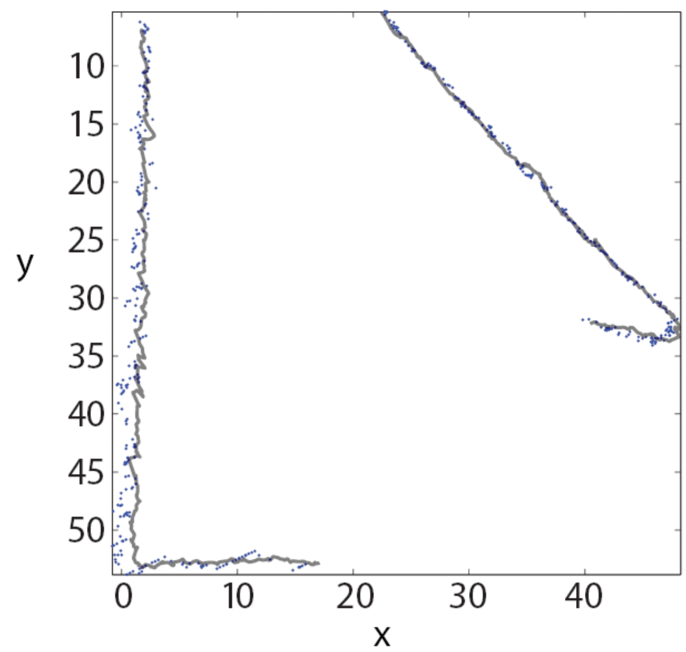
>100 × sub-Nyquist

Yet Another Application



20% Compression

No performance loss in tracking



Conclusions

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

- Sparse signals >> model-sparse signals

- Compressible signals >> model-compressible signals

upshot: fewer measurements
faster and more stable recovery

new concept: RAmP

Volkan Cevher / volkan@rice.edu