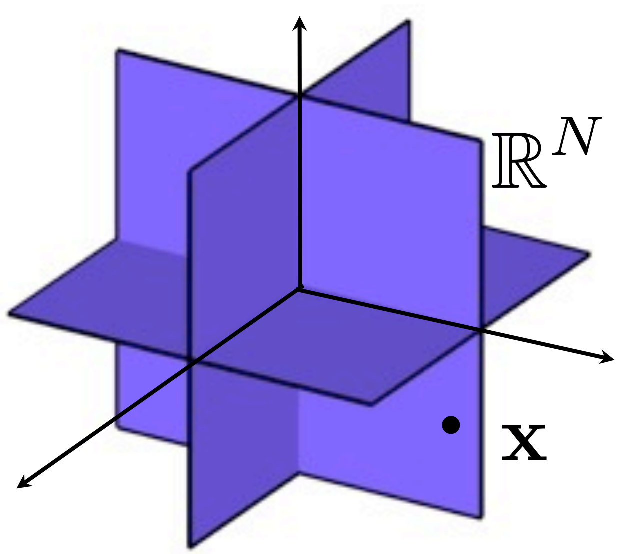
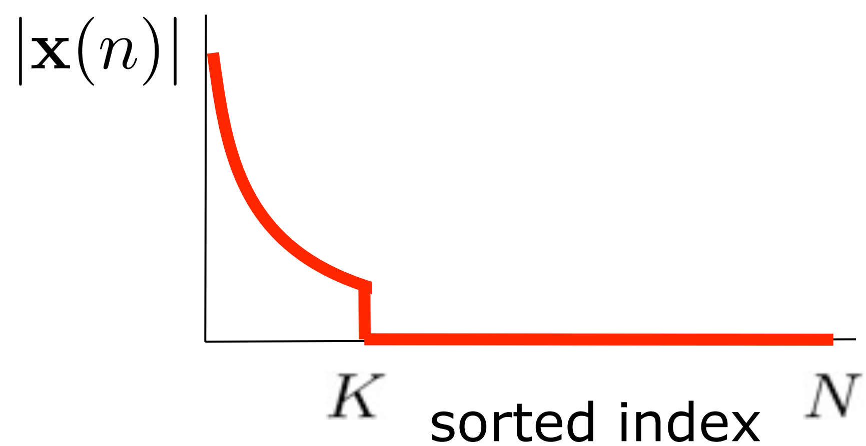
Model-Based Compressive Sensing for Signal Ensembles

Marco F. Duarte
Volkan Cevher
Richard G. Baraniuk

Concise Signal Structure

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

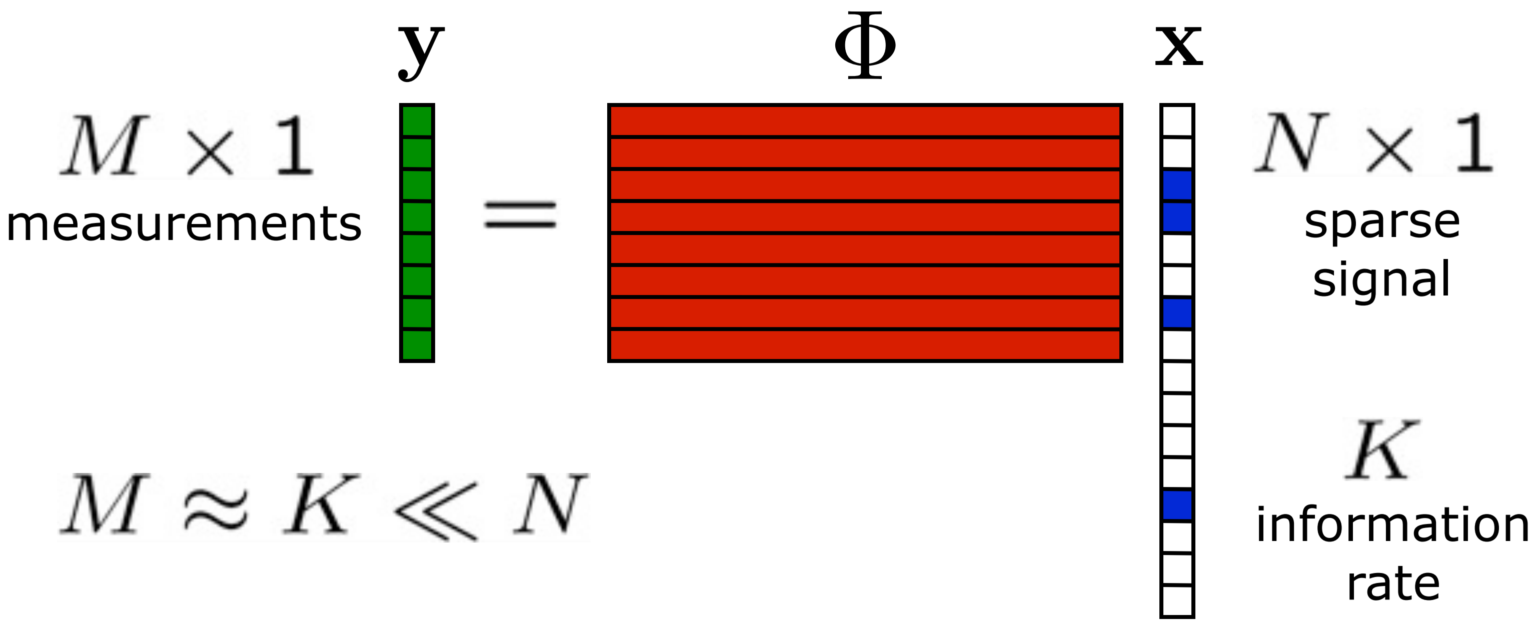




Compressive Sensing

- Replace **samples** by more general **encoder** based on a few linear projections (inner products)
- Recover $\mathbf{x}$ from $\mathbf{y}$ using ***optimization*** (ℓ_1-norm minimization, LPs, QPs) or ***greedy algorithms*** (OMP, CoSaMP, SP, etc.)

$\mathbf{y} = \Phi\mathbf{x}$, $\mathbf{x}$ is sparse

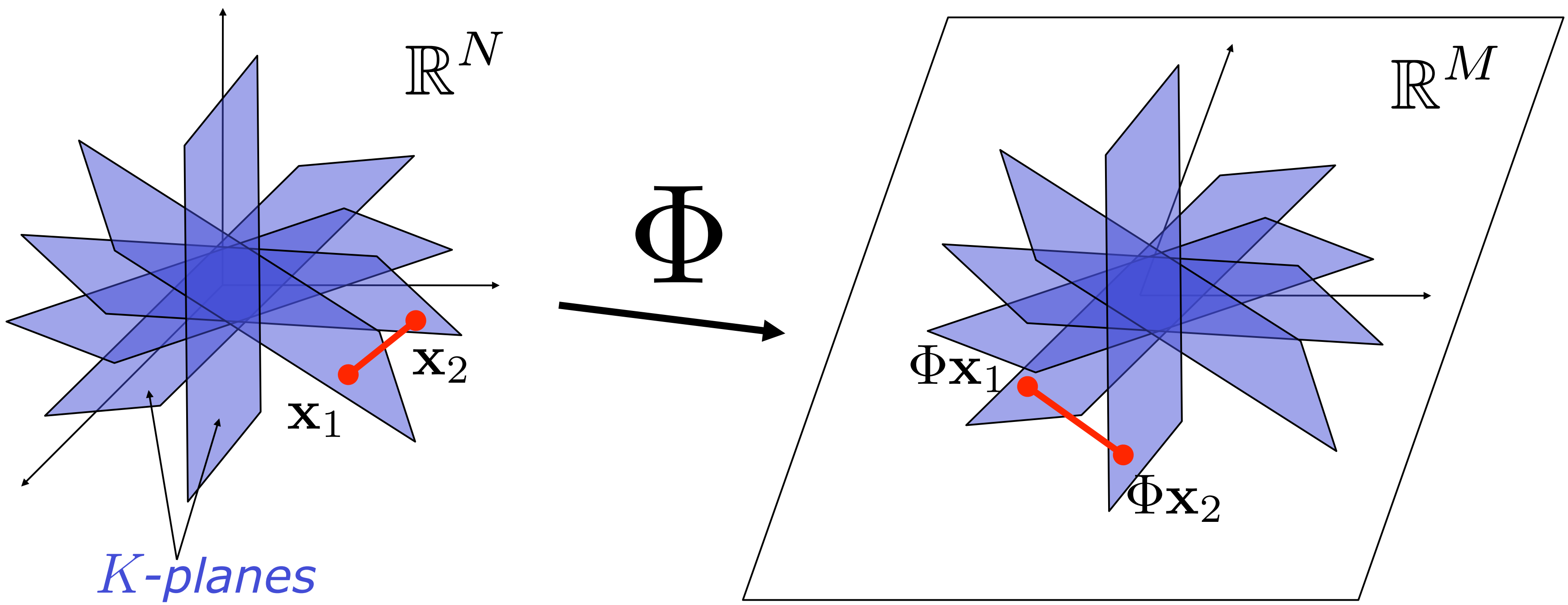


Restricted Isometry Property (RIP)

- Preserve the structure of sparse signals

- RIP of order $2K$ implies: for all K-sparse $\mathbf{x}_1$ and $\mathbf{x}_2$

$$(1-\delta_{2K}) \leq \frac{\|\Phi\mathbf{x}_1 - \Phi\mathbf{x}_2\|_2^2}{\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2} \leq (1+\delta_{2K})$$

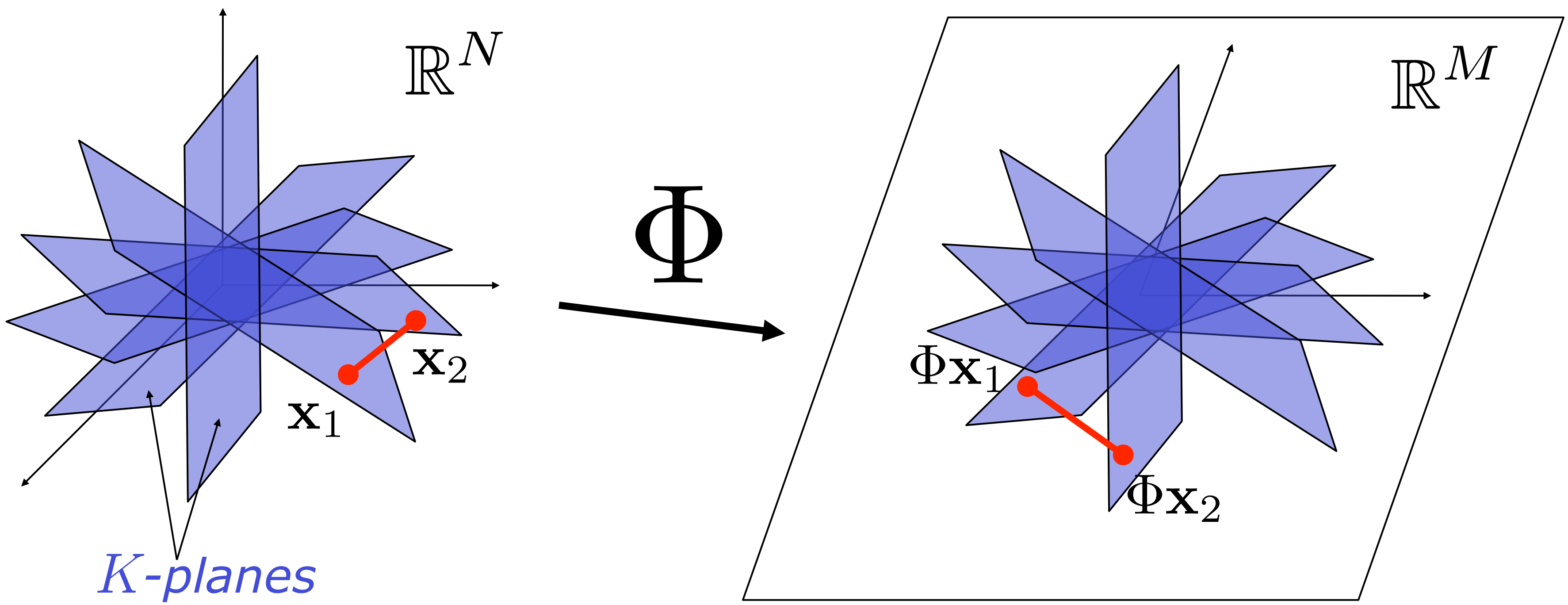


[Candès and Tao]

Restricted Isometry Property (RIP)

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

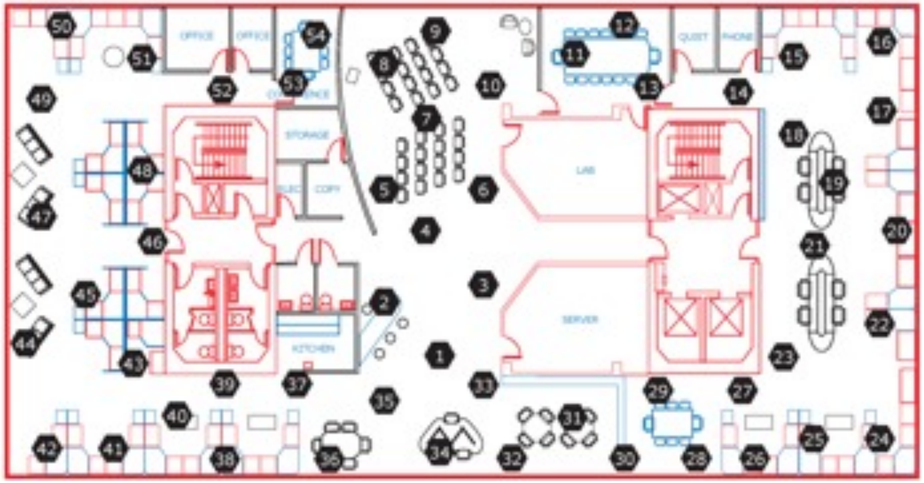
$$M = O(K \log(N/K))$$



[Candès and Tao; Baraniuk, Davenport, DeVore and Wakin]

Sensor Networks

- Networks of many *sensor nodes*
  - sensor, microprocessor for computation, wireless communication, networking, battery
  - sensors observe ***single event***, acquire ***correlated signals***
- Must be energy efficient
  - ***minimize communication*** at expense of off-site computation
  - motivates *distributed compression*



Distributed Compressive Sensing (DCS)

Distributed Sensing

$$
\begin{aligned}
\mathbf{y}_1 &= \Phi_1 \mathbf{x}_1 \\
\mathbf{y}_2 &= \Phi_2 \mathbf{x}_2 \\
&\vdots \\
\mathbf{y}_J &= \Phi_J \mathbf{x}_J
\end{aligned}
$$

[Baron, Duarte, Wakin, Sarvotham, Baraniuk]

Distributed
Compressive Sensing (DCS)

Distributed Sensing

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_J \end{bmatrix} = \begin{bmatrix} \Phi_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \Phi_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \Phi_J \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_J \end{bmatrix}$$

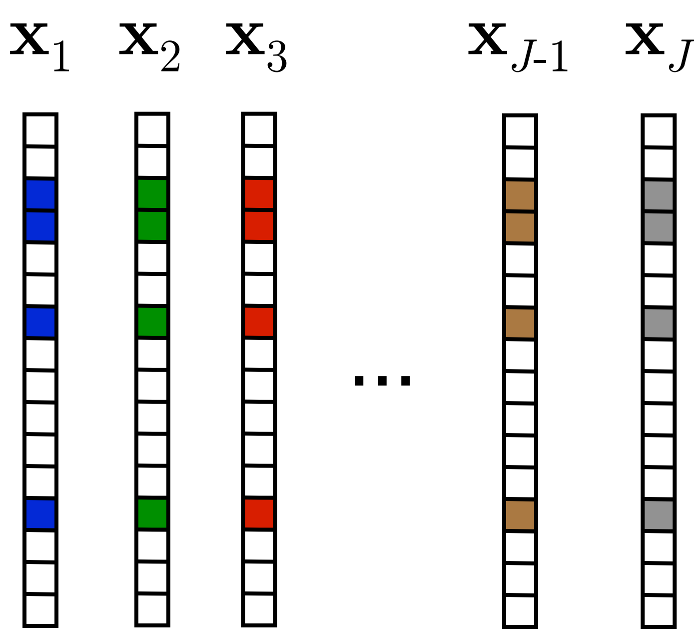
$$\mathbf{Y} \qquad \Phi \qquad \mathbf{X}$$

Joint Recovery

[Baron, Duarte, Wakin, Sarvotham, Baraniuk]

JSM-2: Common Sparse Supports Model

- Measure J signals, each K-sparse
- *Signals share sparse components but with different coefficients*
- Recovery using Simultaneous Orthogonal Matching Pursuit (SOMP) algorithm [Tropp, Gilbert, Strauss]



$$\mathbf{x}_j = \sum_{n \in \Omega} \theta_j(n) \psi_n,$$

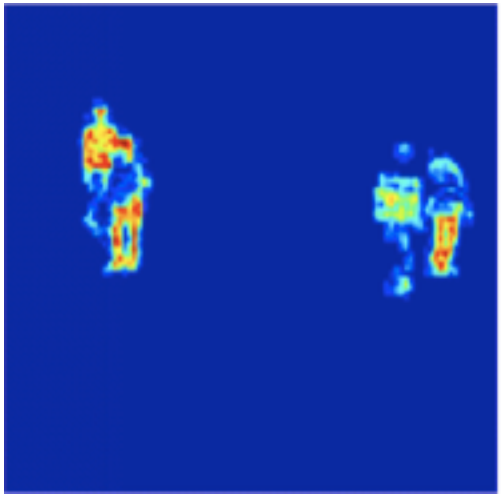
$$|\Omega| = K$$

Beyond Sparse Models

- Sparse signal model captures only **simplistic primary structure**

- For many signal types, location of nonzero coefficients in sparse representation provide ***additional structure***

wavelets:
natural images

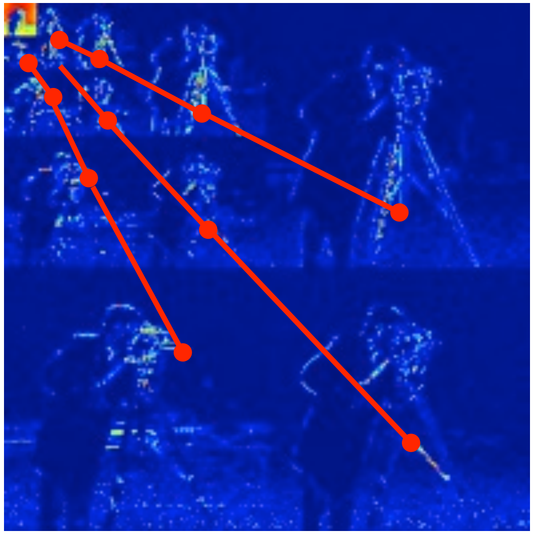


pixels:
background subtracted images

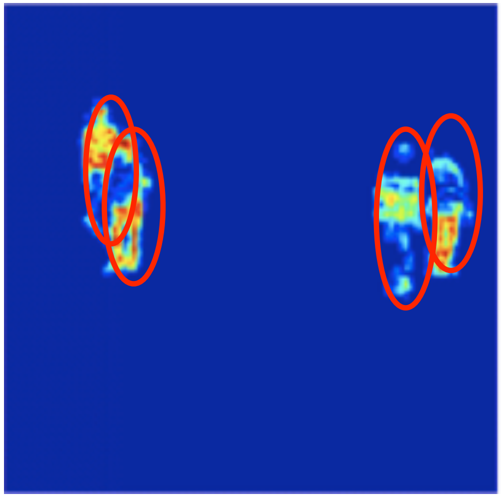
Beyond Sparse Models

- Sparse signal model captures only
  simplistic primary structure

- For many signal types, location of nonzero coefficients in sparse representation provide ***additional structure***



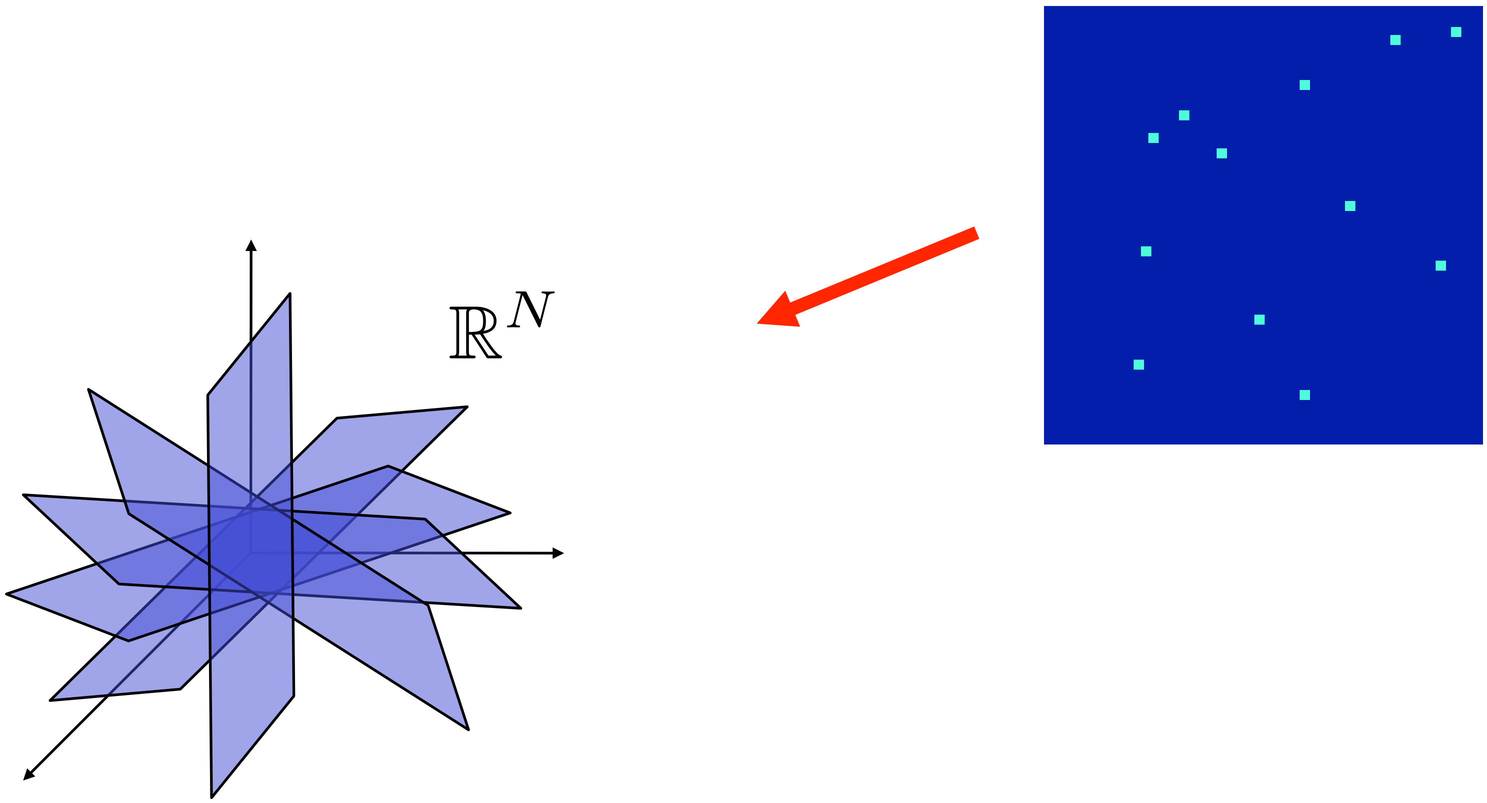
wavelets:
natural images



pixels:
background subtracted
images

Sparse Signals

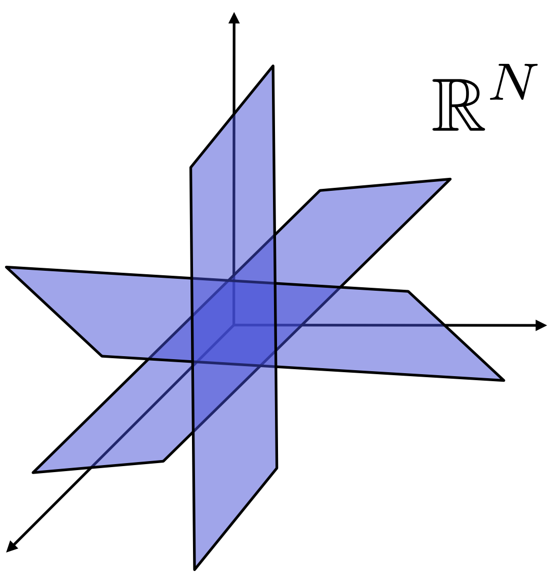
- Defn: A K-**sparse** signal lives on the collection of K-dim subspaces aligned with coord. axes

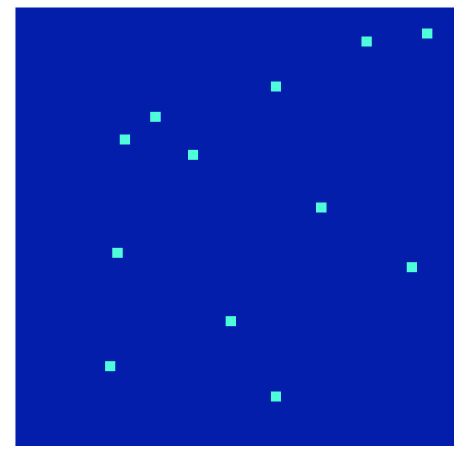


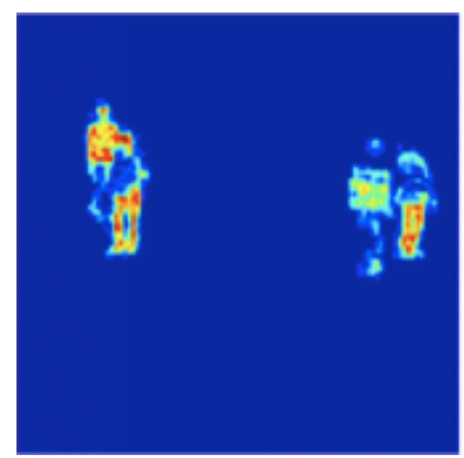
Structured Sparse Signals

- Defn: A K**-structured sparse** signal lives on a particular (reduced) collection of K-dim canonical subspaces

[Lu and Do]
[Blumensath and Davies]

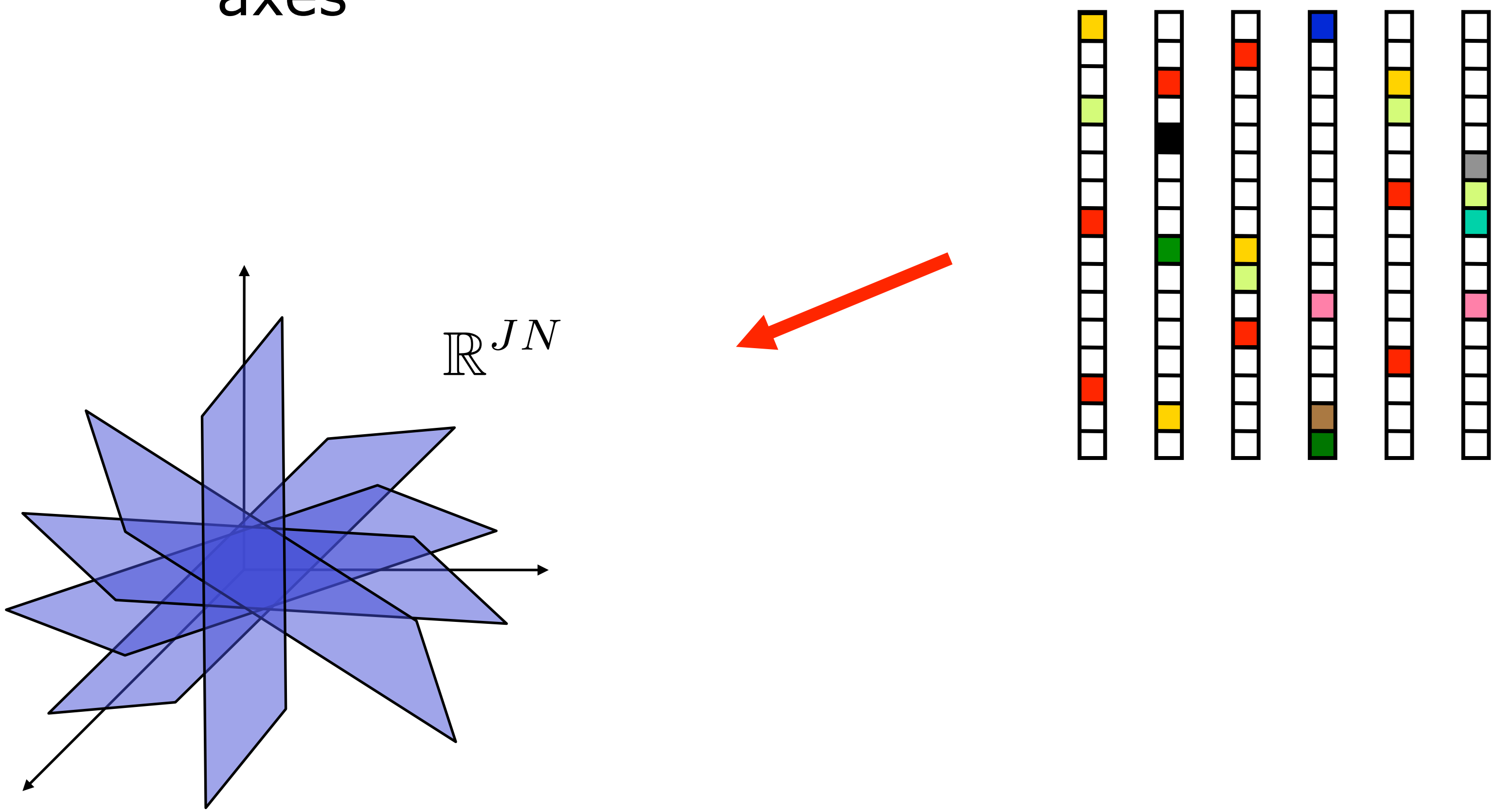






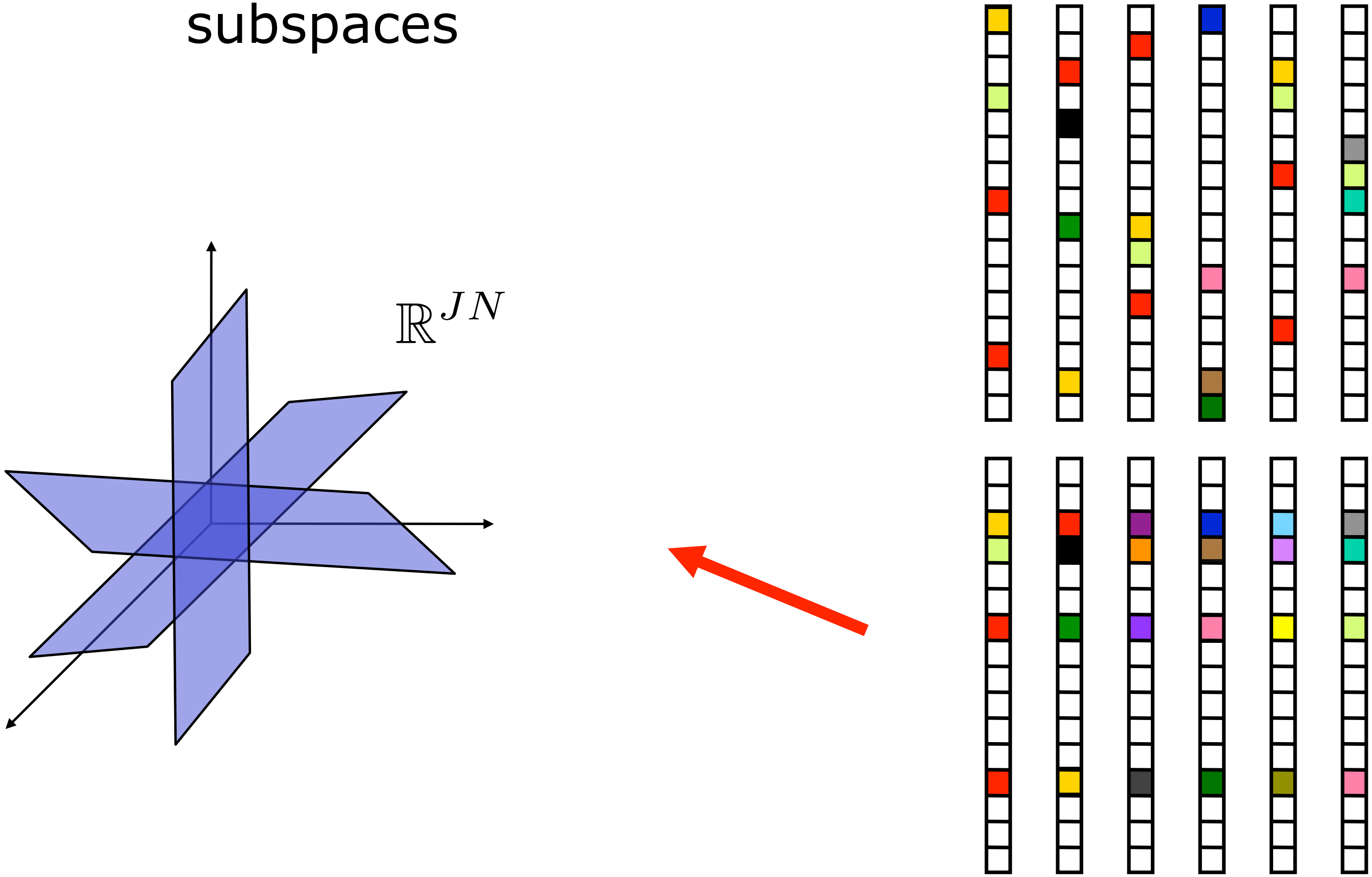
Sparse Signal Ensemble

- Defn: An ensemble of J K-**sparse** signal lives on a collection of JK-dim subspaces aligned with coord. axes



Structured Sparse Signal Ensemble

- Defn: An ***structured ensemble*** of J K-sparse signals with ***common sparse support*** lives on a particular (reduced) collection of JK-dim canonical subspaces



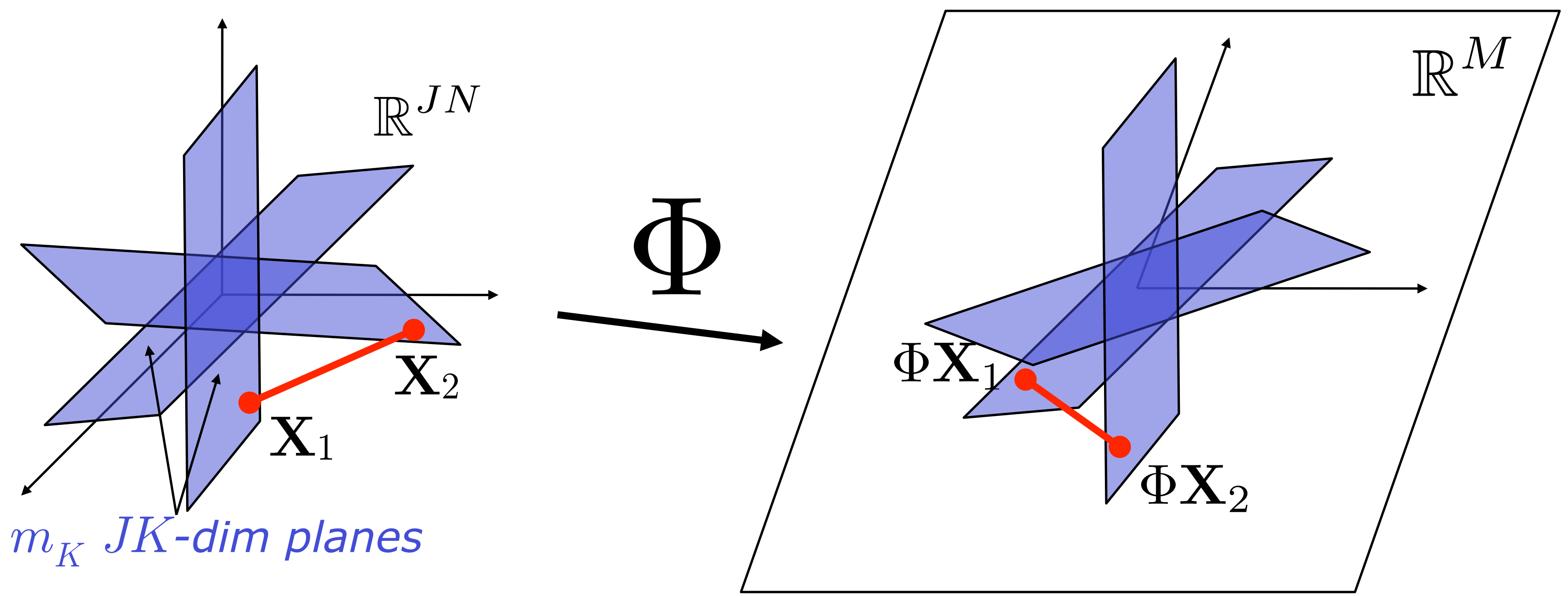
RIP for Structured Sparsity Model

- Preserve the structure ***only*** of sparse signals that *follow the structure*

- Random (i.i.d. Gaussian, Bernoulli) matrix has the JSM-2 RIP with high probability if

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

[Blumensath and Davies]



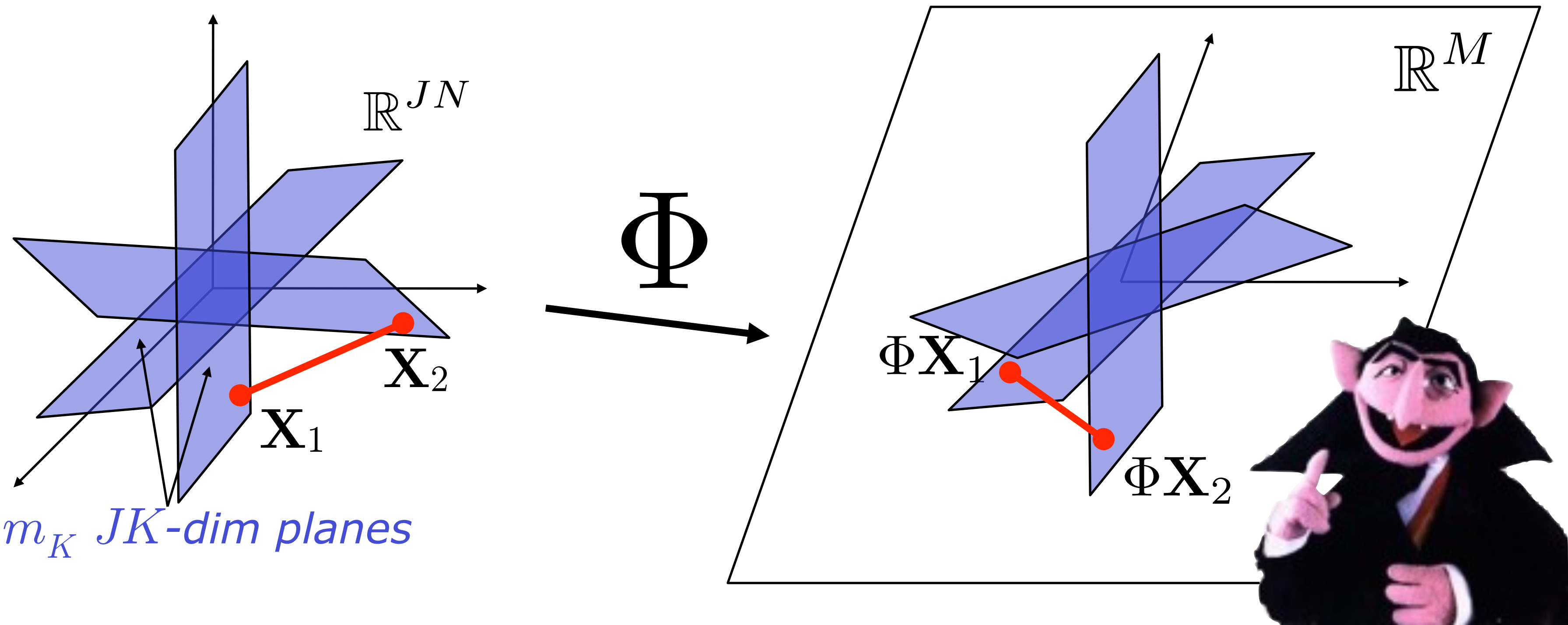
RIP for Structured Sparsity Model

- Preserve the structure ***only*** of sparse signals that *follow the structure*

- Random (i.i.d. Gaussian, Bernoulli) matrix has the JSM-2 RIP with high probability if

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

[Blumensath and Davies]

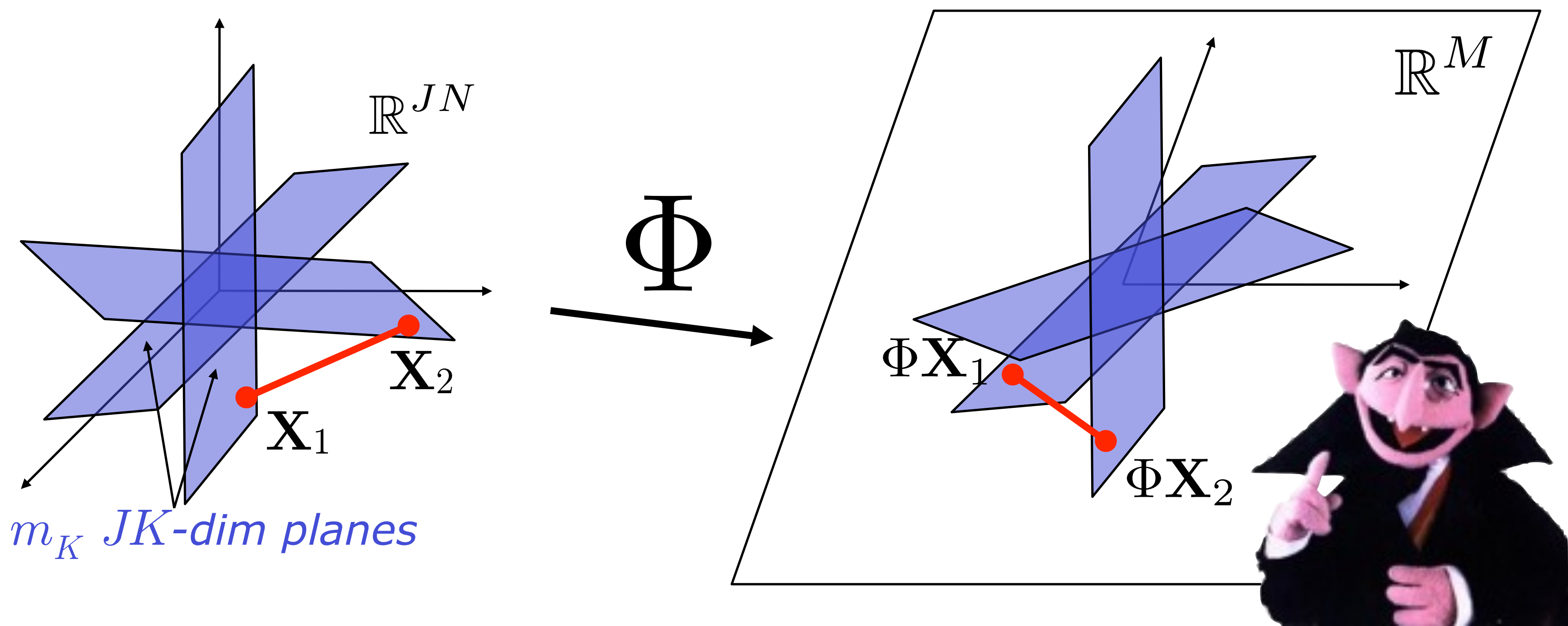


RIP for Common Sparse Support Model

- Random (i.i.d. Gaussian, Bernoulli) matrix has the model-based RIP with high probability if

$$M = O\left(KJ + K\log(N/K)\right)$$

- Distributed settings: measurements from different sensors can be ***added together*** to effectively obtain dense measurement matrix.



Standard CS Recovery

CoSaMP

[Needell and Tropp]

– calculate current residual $\mathbf{r} = \mathbf{y} - \Phi\widehat{\mathbf{x}}$

– form residual signal estimate $\mathbf{e} = \Phi^T\mathbf{r}$

– ***calculate enlarged support*** $\Omega = \mathrm{supp}(\widehat{\mathbf{x}}) \cup \mathrm{supp}(\mathfrak{T}(\mathbf{e}, 2K))$

– estimate signal for enlarged support $\mathbf{b}|_\Omega = \Phi|_\Omega^\dagger \mathbf{y}, \ \mathbf{b}|_{\Omega^C} = 0$

– ***shrink support*** $\widehat{\mathbf{x}} = \mathfrak{T}(\mathbf{b}, K)$

Model-Based CS Recovery

Model-based CoSaMP

$\mathcal{M}_K$: K-term ***structured sparse approximation algorithm***

- calculate current residual $\quad \mathbf{r} = \mathbf{y} - \Phi\widehat{\mathbf{x}}$
- form residual signal estimate $\quad \mathbf{e} = \Phi^T\mathbf{r}$
- ***calculate enlarged support*** $\quad \Omega = \mathrm{supp}(\widehat{\mathbf{x}}) \cup \mathrm{supp}(\mathcal{M}_{2K}(\mathbf{e}))$
- estimate signal for enlarged support $\quad \mathbf{b}|_\Omega = \Phi|_\Omega^\dagger \mathbf{y},\ \mathbf{b}|_{\Omega^C} = 0$
- ***shrink support*** $\quad \widehat{\mathbf{x}} = \mathcal{M}_K(\mathbf{b})$

[Baraniuk, Cevher, Duarte, Hegde 2008]

Model-Based Recovery for JSM-2

Model-based Distributed CoSaMP → CoSOMP

- calculate current residual ***at each sensor*** $\mathbf{r}_j = \mathbf{y}_j - \Phi_j \widehat{\mathbf{x}}_j$
- form residual signal estimate ***at each sensor*** $\mathbf{e}_j = \Phi_j^T \mathbf{r}_j$
- ***merge sensor estimates*** $\mathbf{e} = \sum_{j=1}^{J} (\mathbf{e}_j \cdot \mathbf{e}_j)$
- ***calculate enlarged support*** $\Omega = \text{supp}(\widehat{\mathbf{x}}) \cup \text{supp}(\mathfrak{T}(\mathbf{e}, 2K))$
- estimate signal proxy ***at each sensor*** $\mathbf{b}_j|_{\omega} = \Phi_j|_{\Omega}^{\dagger} \mathbf{y}_j, \ \mathbf{b}_j|_{\omega^C} = 0$
- ***merge sensor estimates*** $\mathbf{b} = \sum_{j=1}^{J} (\mathbf{b}_j \cdot \mathbf{b}_j)$
- ***shrink estimate support*** $\Lambda = \text{supp}(\mathfrak{T}(\mathbf{b}, K))$
- update signal estimates ***at each sensor*** $\widehat{\mathbf{x}}_j|_{\Lambda} = \mathbf{b}_j|_{\Lambda}, \ \widehat{\mathbf{x}}_j|_{\Lambda^C} = 0$

Model-Based CS Recovery Guarantees

Theorem:
Assume we obtain noisy CS measurements of a signal ensemble $\mathbf{Y} = \Phi\mathbf{X} + \mathbf{n}$. If Φ has the model-based RIP with $\delta_K < 0.1$, then we have

$$\|\mathbf{X} - \widehat{\mathbf{X}}\|_2 \leq C_1\|\mathbf{X} - \mathcal{M}_K(\mathbf{X})\|_2 + \frac{C_2}{\sqrt{K}}\|\mathbf{X} - \mathcal{M}_K(\mathbf{X})\|_1 + C_3\|\mathbf{n}\|_2$$

CS recovery error

signal *K*-term **structured sparse approximation** error

noise

In words, ***instance optimality*** based on ***structured sparse approximation***

Model-Based CS Recovery Guarantees

Theorem:
Assume we obtain noisy CS measurements of a signal ensemble $\mathbf{Y} = \Phi\mathbf{X} + \mathbf{n}$. If each Φ_j has the RIP with $\delta_K < 0.1$, then we have

$$\|\mathbf{X} - \widehat{\mathbf{X}}\|_2 \leq C_1 \|\mathbf{X} - \mathcal{M}_K(\mathbf{X})\|_2 + \frac{C_2}{\sqrt{K}} \|\mathbf{X} - \mathcal{M}_K(\mathbf{X})\|_1 + C_3 \|\mathbf{n}\|_2$$

CS recovery error

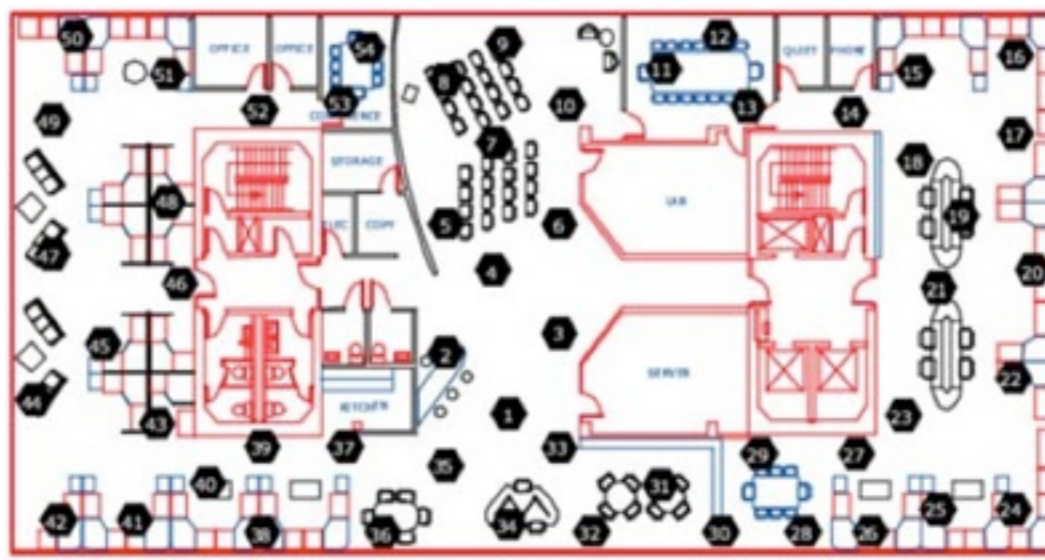
signal *K*-term **structured sparse approximation** error

noise

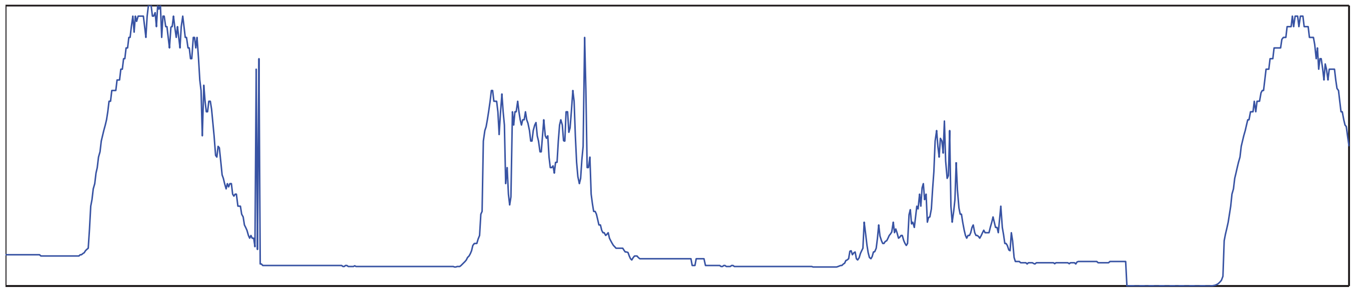
In words, ***instance optimality*** based on ***structured sparse approximation***

Real Data Example

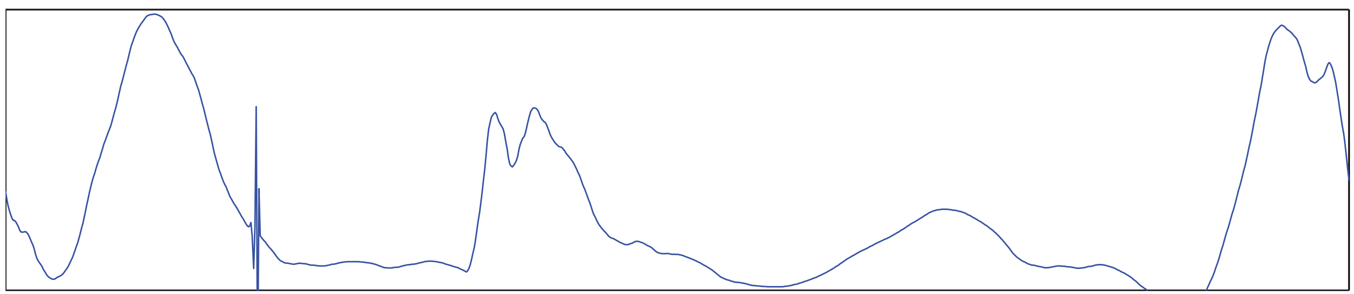
- Environmental Sensing in Intel Berkeley Lab
- $J = 49$ sensors, $N = 1024$ samples each
- Compare:
  - independent recovery CoSaMP
  - existing joint recovery SOMP
  - model-based joint recovery CoSOMP



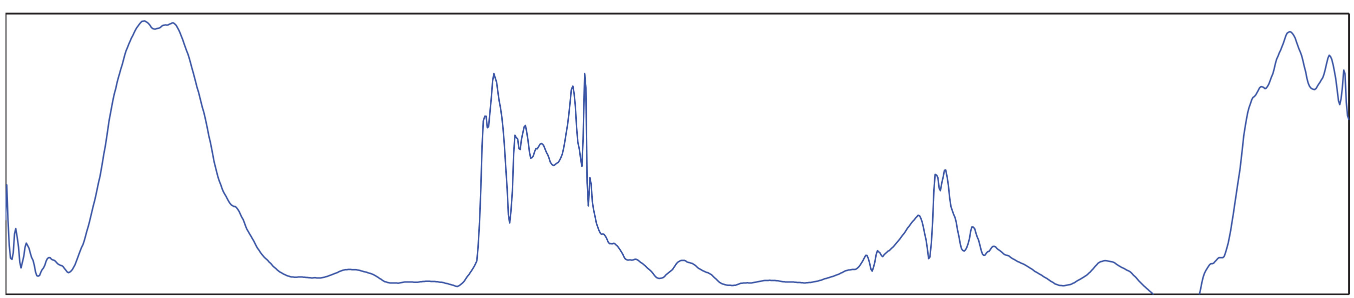
Experimental Results - Brightness Data



(a) Original signal



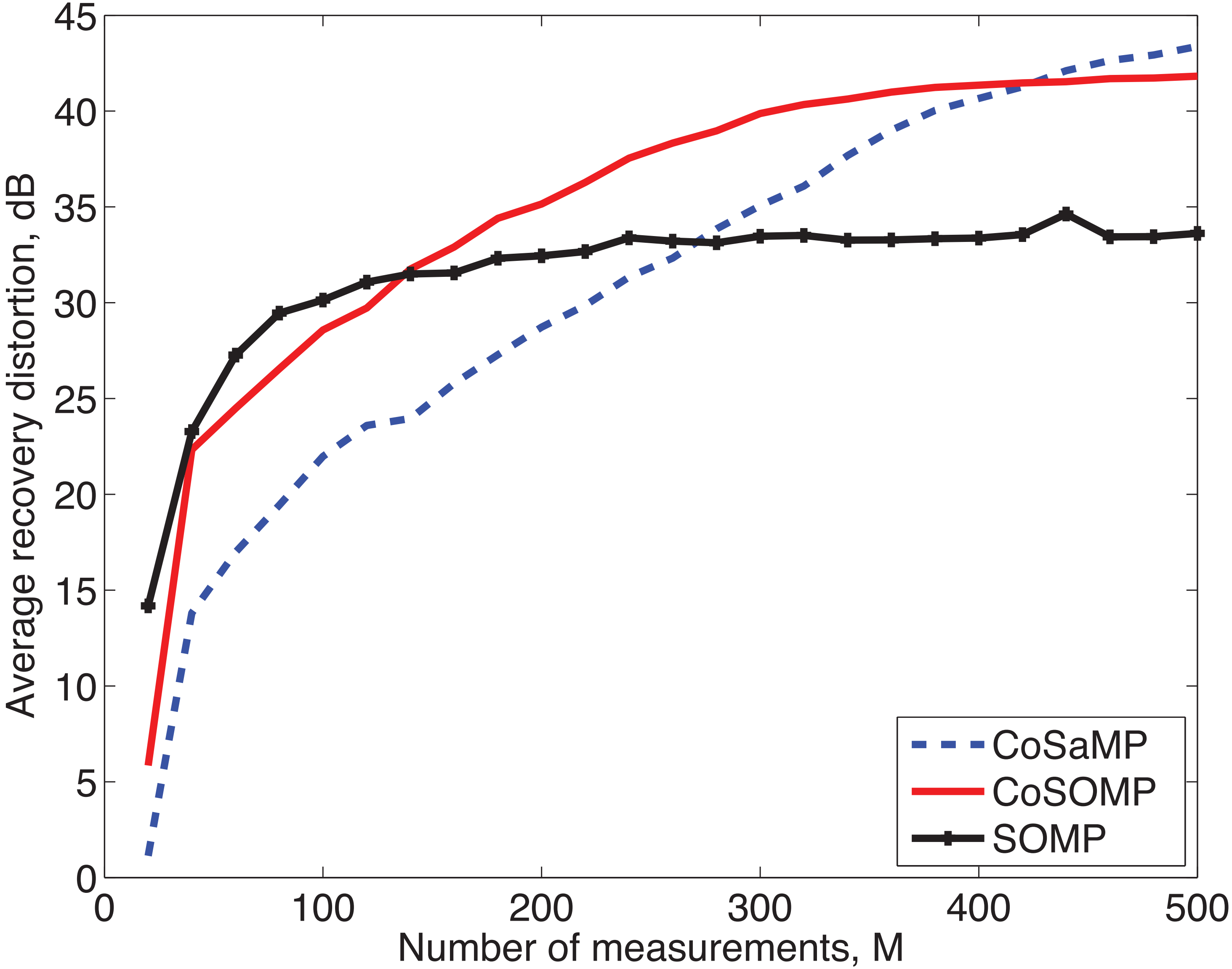
(c) CoSaMP recovery, distortion = 15.1733 dB



(d) CoSOMP recovery, distortion = 16.3197 dB

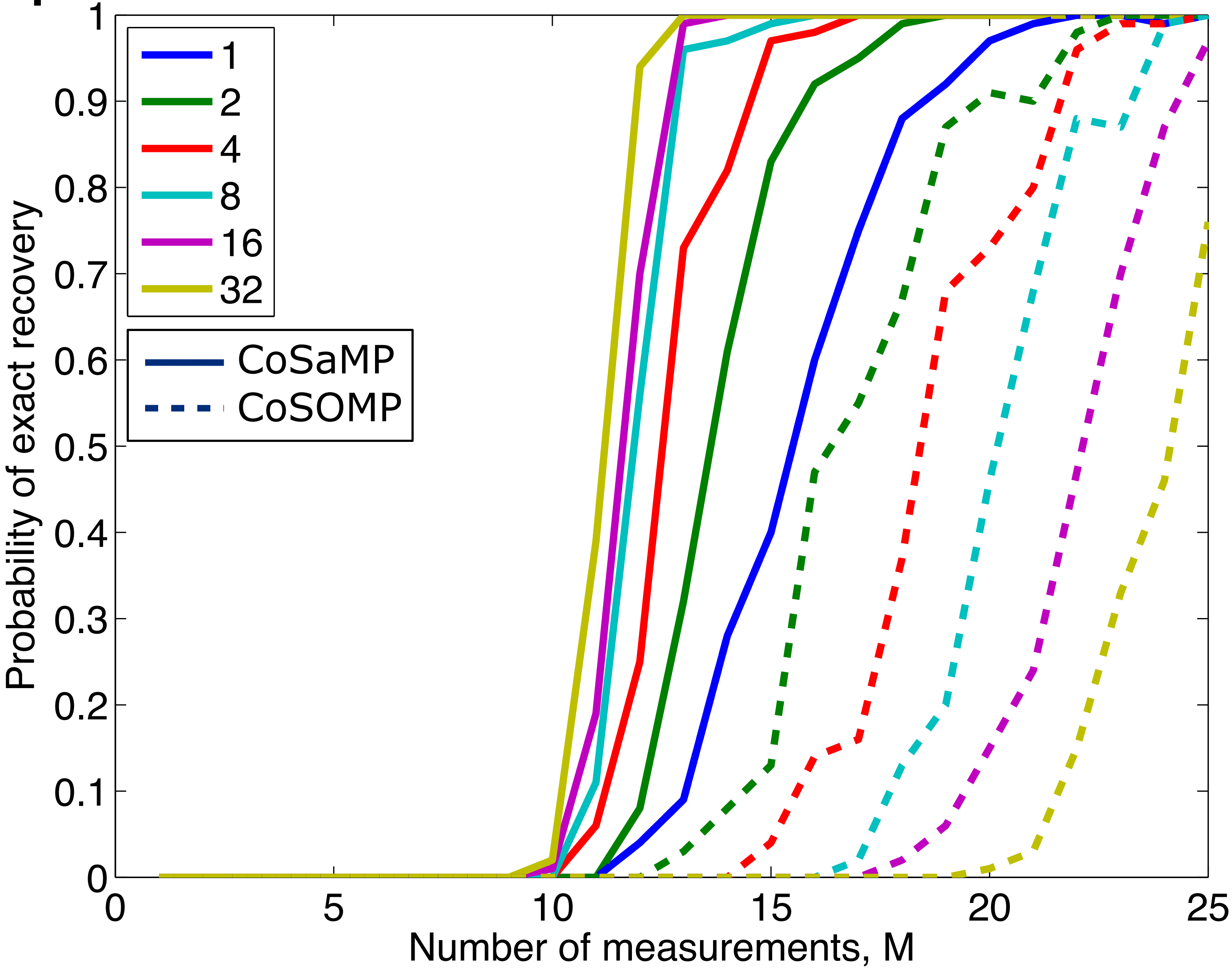
$$N = 1024,\ J = 48,\ M = 400$$

Experimental Results - Humidity Data



$$N = 1024,\ J = 48$$

Experimental Results - Network Size



$N = 1024$

Conclusions

- ***Intuitive union-of-subspaces model*** to encode structure of jointly sparse signal ensembles
- Structure enables ***reduction*** in number of measurements required for recovery
- Signal recovery algorithms are ***easily adapted*** to leverage additional structure
- ***New structure-based recovery guarantees*** such as instance optimality

Rice University DSP

dsp.rice.edu/cs