CS-541
Wireless Sensor Networks

Lecture 3: Signal Sampling for WSN

Spring Semester 2017-2018

Prof Panagiotis Tsakalides, Dr Athanasia Panousopoulou, Dr Gregory Tsagkatakis

Today's objectives

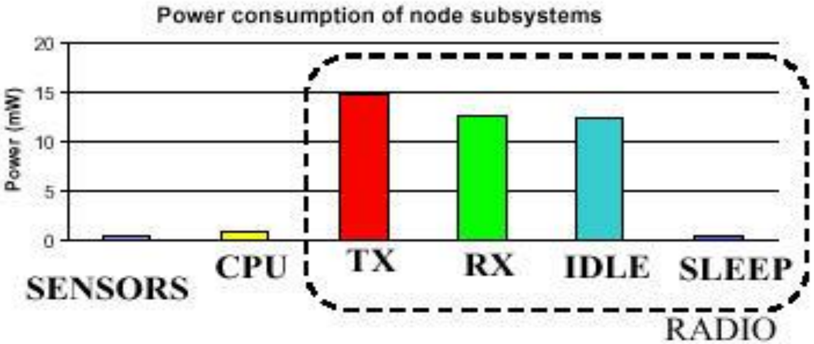
Signal Sampling

Compressed Sensing

Applications in WSN

Sensing in WSNs

Sensing	Quantization	Storage/Processing	Communications
Sensor type	A/D	Size	Route selection
Operations	Bus	Speed	Reliability/Connectivity
Calibration		Complexity	Robustness
Power consumption			



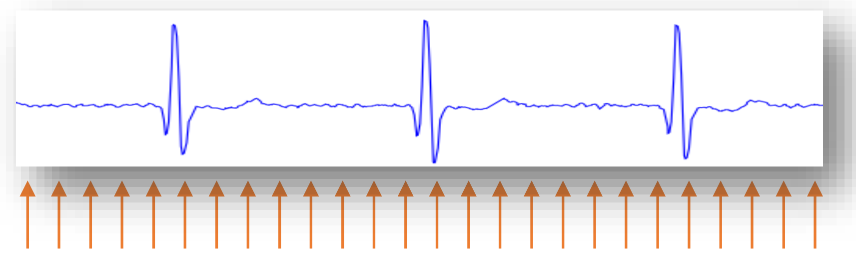
Objectives

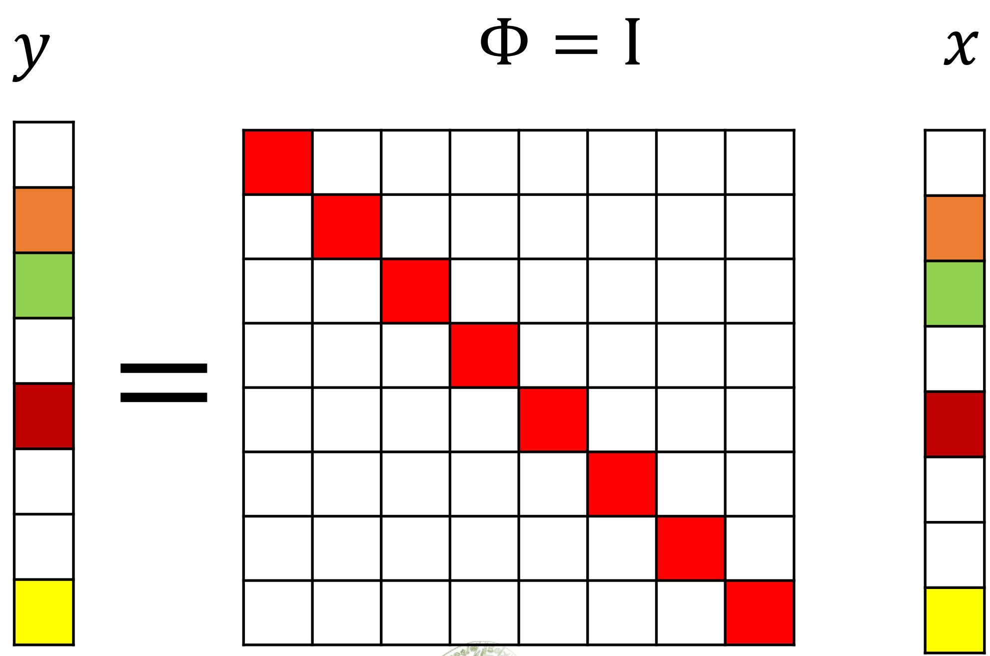
- Efficient data acquisition and gathering
  - Increase life-time of network
  - Reduce communication requirements
  - Handle transmission errors
  - Reduce calibration operations
  - Facilitate data classification
- Prior Knowledge
  - Training data
  - Spatio-temporal correlations

Signal Sensing

Nyquist–Shannon

- limited signal support
- Signal bandwidth B
  Sampling rate $F_s=2B$
  (Nyquist rate)



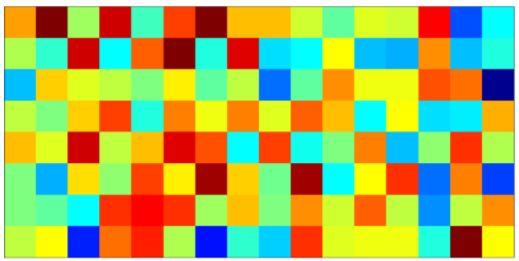


Limitations

- Requirements
- Power/battery
- Storage/Bandwidth
- Calibration

Random Projections
Johnson-Lindenstrauss (JL) lemma

Sensing matrix $M \ll N$



What's wrong with PCA

- Computational complexity
- Universality
- Adaptability
- Robustness

Given $0 < \varepsilon < 1$, a set Q of m points in $\mathbf{R}^N$, and a number $n > 8 \ln(m) / \varepsilon^2$, there is a linear map $f : \mathbf{R}^N \to \mathbf{R}^n$ such that

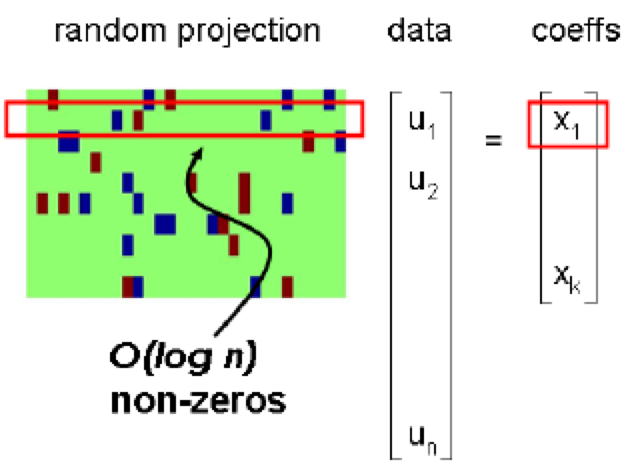
$$(1-\epsilon)\|u-v\|^2 \leq \|f(u)-f(v)\|^2 \leq (1+\epsilon)\|u-v\|^2$$

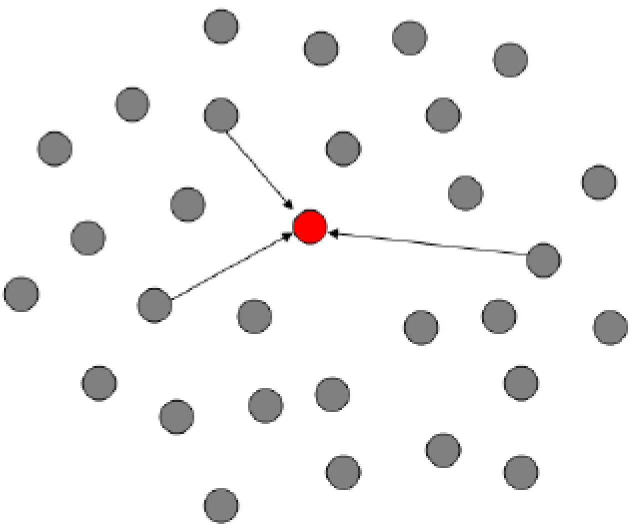
for all $x, y \in Q$.

Simplified Random Projection (SRP)

- f: Random matrix is usually gaussian distributed
  - mean: 0; standart deviation: 1
- f: sparse RP distribution

$$r_{ij} = \sqrt{3} \cdot \begin{cases} +1 & \text{with probability } \frac{1}{6} \\ 0 & \text{with probability } \frac{2}{3} \\ -1 & \text{with probability } \frac{1}{6}. \end{cases}$$

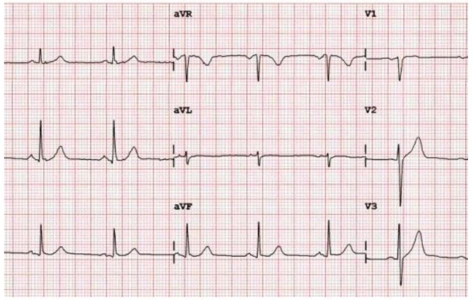




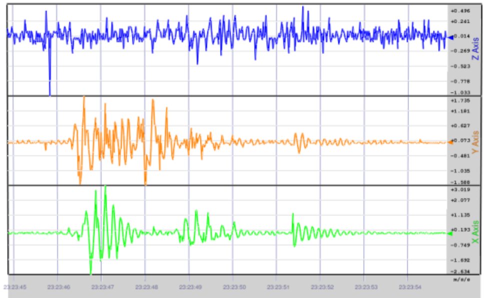
What about recovery?

Can we recovery the original signal from its RP?

YES…. for sparse signal

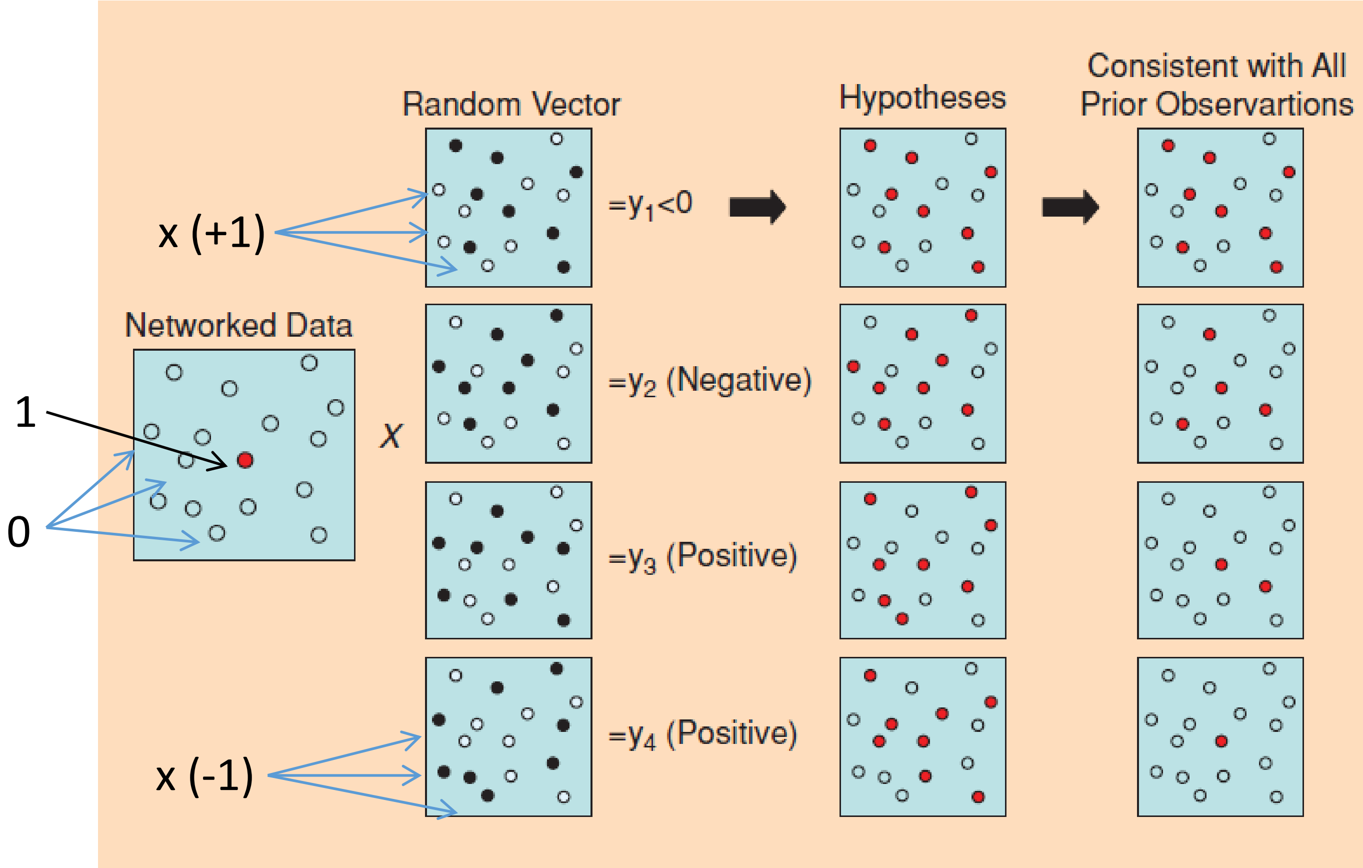


Biological



Environmental

Astronomical



Q: **which sensor is different** (using as few observations as possible) ?

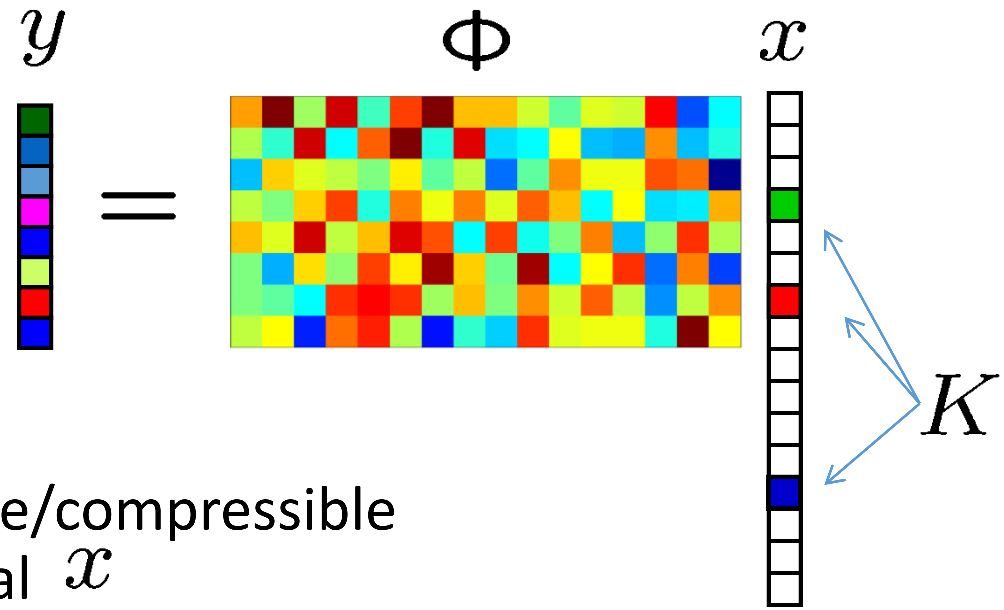
A: **Project the data onto random vectors** (second column)

- Initially: n/2 hypothesis sensors are consistent with each random projection observation
- Exponential decrease of consistence observations

Observations

- Random projections -> binary bisections of the hypothesis space
- Only log n observations are needed to determine which sensor reads the nonzero value.

Compressed Sensing
(or compressive sensing, compressed sampling…)



- **Goal**: Recover signal x from measurements y
- **Problem**: Random projection Φ not full rank (ill-posed inverse problem)
- **Solution**: Exploit the sparse/compressible ***geometry*** of acquired signal x
- **Recovery** via (convex) sparsity penalty or greedy algorithms

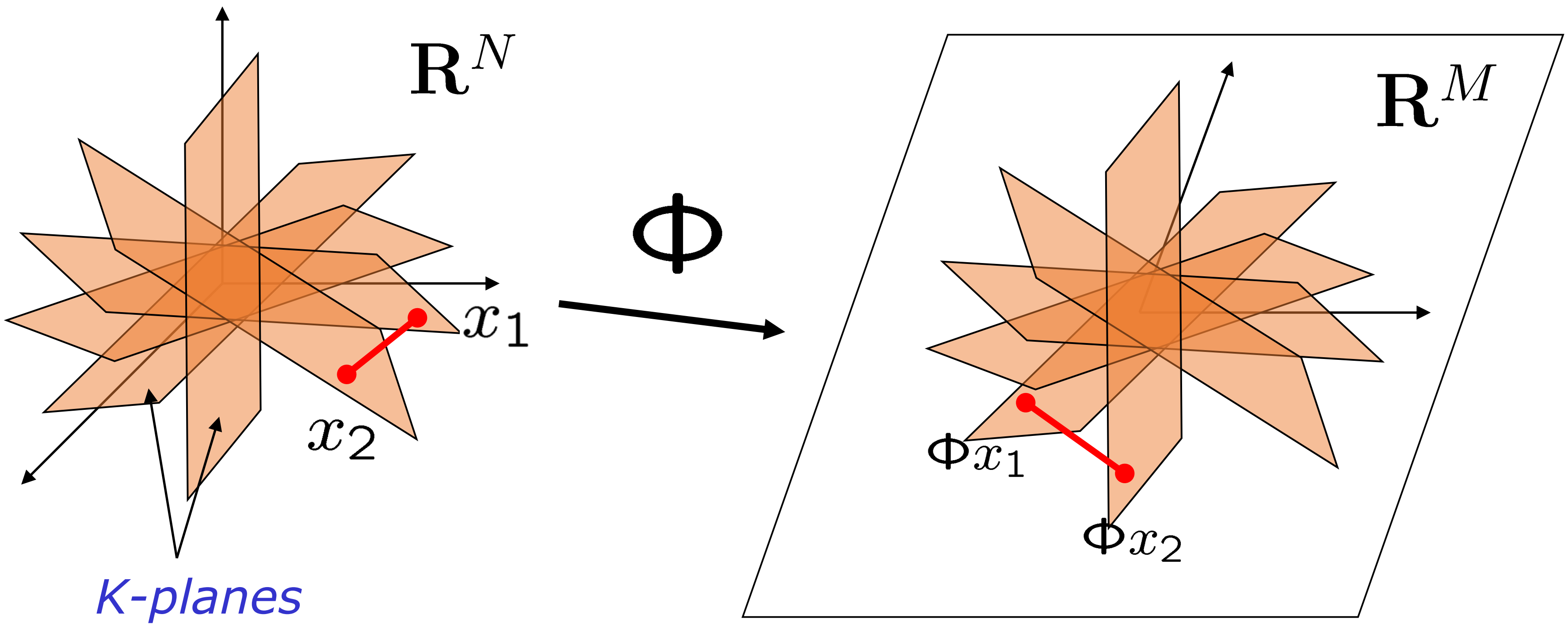
$$\hat{x} = \arg\min_{x} \|x\|_0 \ \text{subject to} \ \Phi x = y$$

NP-hard!

[Donoho; Candes, Romberg, Tao, 2004]

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



How many measurements?

Φ satisfies Restricted Isometry Property (RIP)

For all x that are K sparse

$$(1-\delta_K)||x||_2^2 \leq ||\Phi x||_2^2 \leq (1+\delta_K)||x||_2^2$$

When Φ MxN satisfies RIP of order 2K with δ<√(2)-1,

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

- Random (sub-) Gaussian (iid Gaussian, Bernoulli) satisfy RIP

CS Recovery Algorithm

- **Iterative Thresholding**

Given $y = \Phi x$, recover a sparse x

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

iteration:

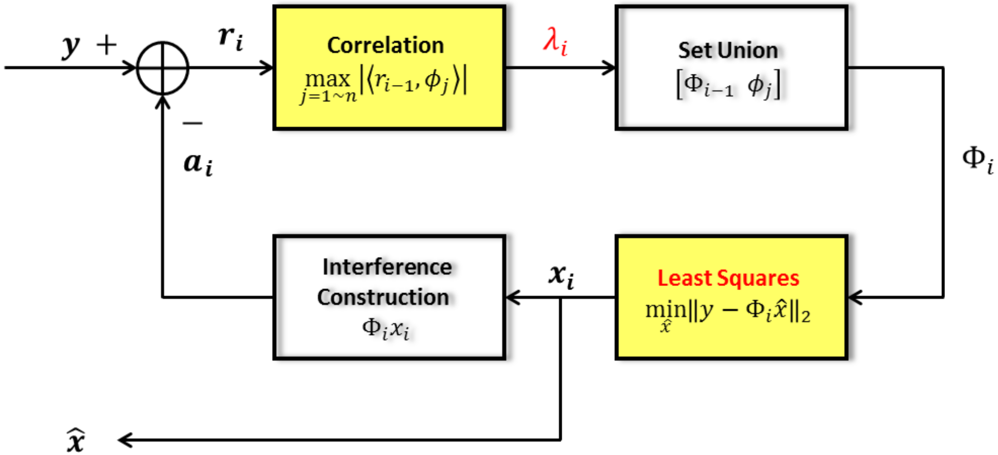
- $i \leftarrow i + 1$
- $b \leftarrow \widehat{x}_{i-1} + \Phi^T r$ **update signal estimate**
- $\widehat{x}_i \leftarrow \text{thresh}(b, K)$ **prune signal estimate** (best *K*-term approx)
- $r \leftarrow y - \Phi \widehat{x}_i$ **update residual**

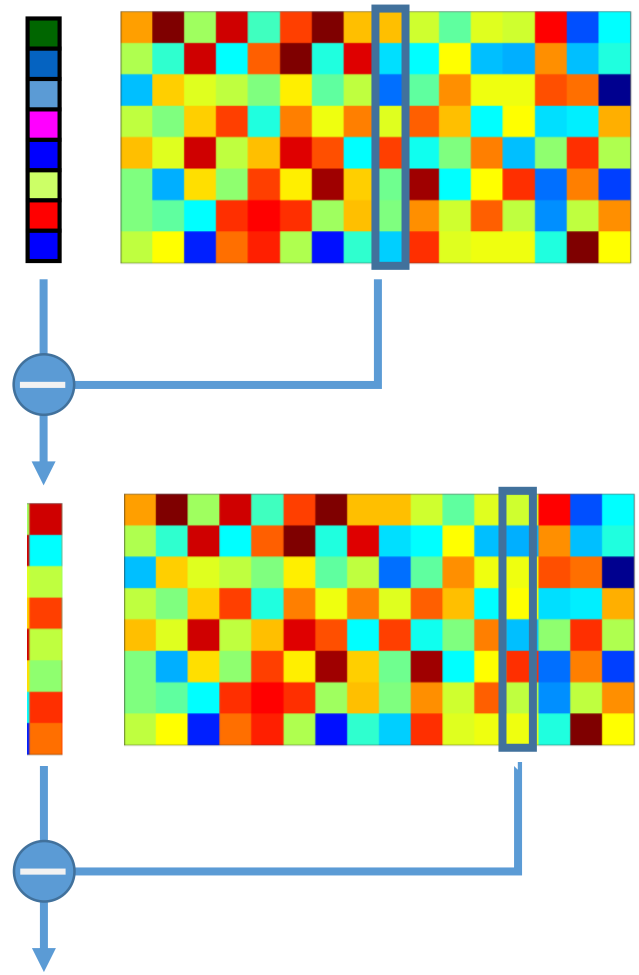
return: $\widehat{x} \leftarrow \widehat{x}_i$

Adapted from "**Model-based Compressive Sensing**", by Volkan Cevher

Matching Pursuit Algorithms

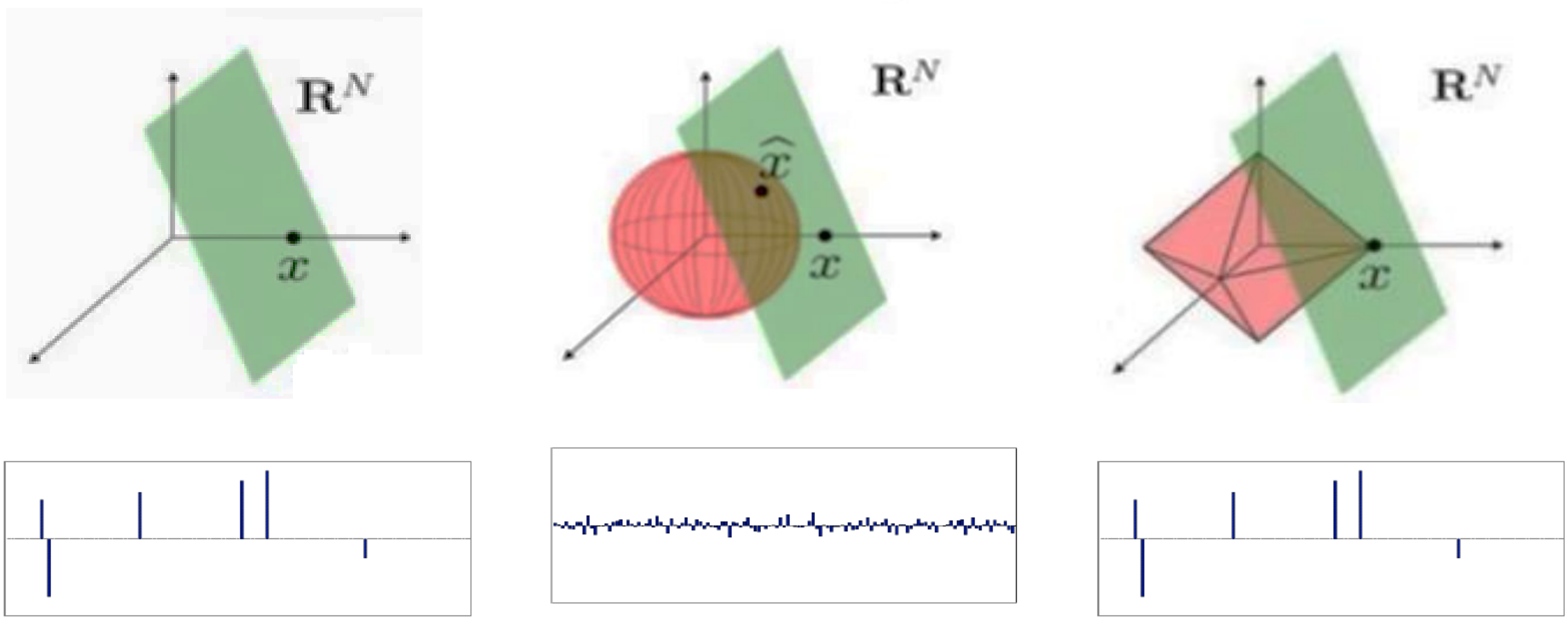
- Use greedy algorithm to iteratively recover sparse signal
- Procedure:
    1. Initialize
    2. Find the column that is most correlated
    3. Set Union (add one col. every iter.)
    4. Solve the least squares
    5. Update data and residual
    6. Back to step 2 or output





A linear programming approach

Replace greedy with convex optimization problem



Compressed Sensing via ℓ_1

- 0-norm is nonconvex ➔ difficult to solve
- 1-norm is convex ➔ Basis Pursuit (Lasso)

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \|\mathbf{y} - \boldsymbol{\Phi}\mathbf{x}\|_2^2 < \epsilon$$

Theorem (Candès, Romberg, Tao 2004 – Candès 2008 – Foucart, Lai 2009 – Foucart 2009)

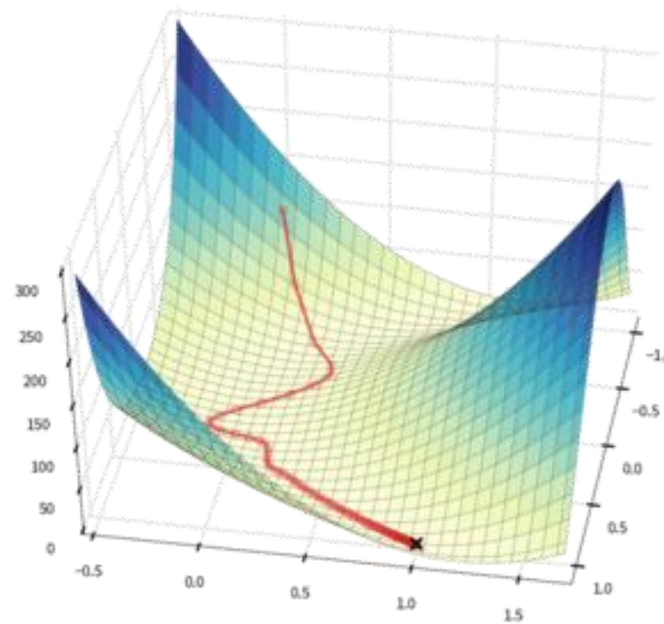
Assume that the restricted isometry constant δ_{2s} of $A \in \mathbb{C}^{m \times N}$ satisfies

$$\delta_{2s} < \frac{2}{3 + \sqrt{7/4}} \approx 0.4627.$$

Then ℓ_1-minimization reconstructs every s-sparse vector $\mathbf{x} \in \mathbb{C}^N$ from $y = Ax$.

Performance of Recovery

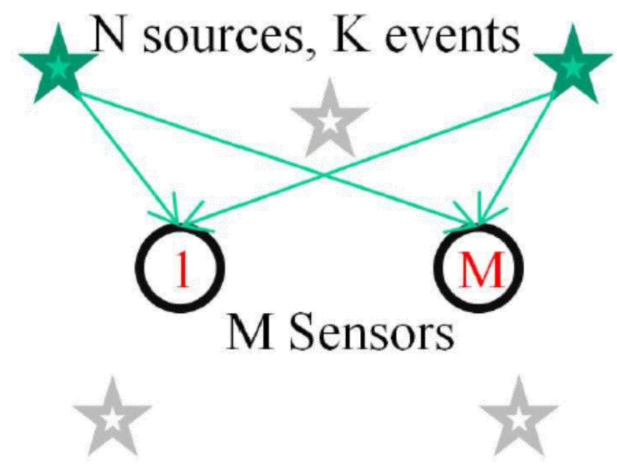
- Using ℓ_1 methods
- **Sparse signals**
  - noise-free : exact recovery
  - noisy : stable recovery
- **Compressible signals**
  - recovery as good as *K*-sparse approximation



$$\|x - \widehat{x}\|_2 \leq C_1 \frac{\|x - x_K\|_1}{K^{1/2}} + C_2 \|n\|_2$$

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

Sparse event detection



- N sources, K events, K<<N, M sensors
- Event vector $\mathbf{X}_{N\times 1}$
- Channel response

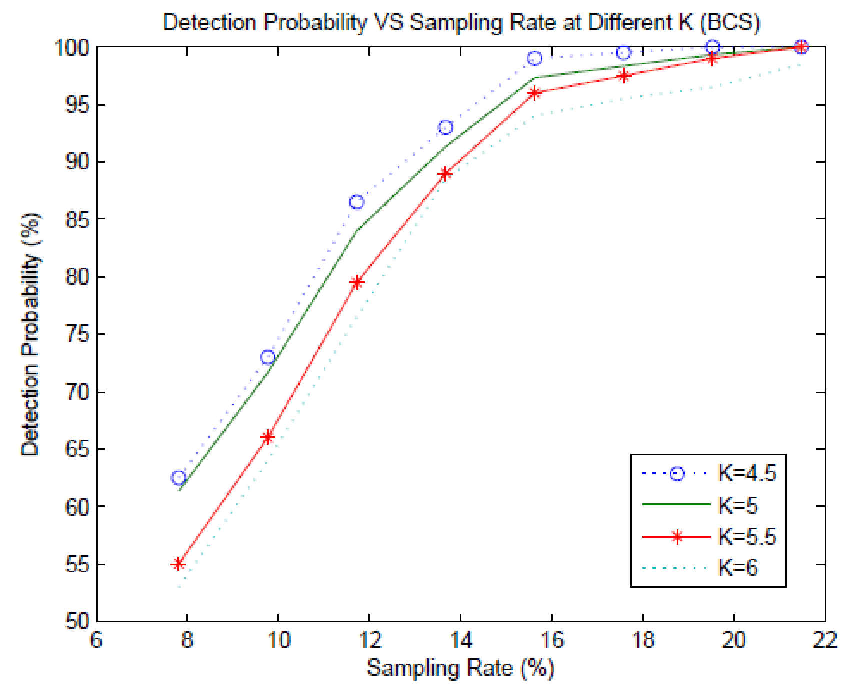
$$G_{m,n} = (d_{m,n})^{-\alpha/2}|h_{m,n}|,$$

- Received signal

$$\mathbf{Y}_{M\times 1} = \mathbf{G}_{M\times N}\mathbf{X}_{N\times 1} + \epsilon_{M\times 1},$$

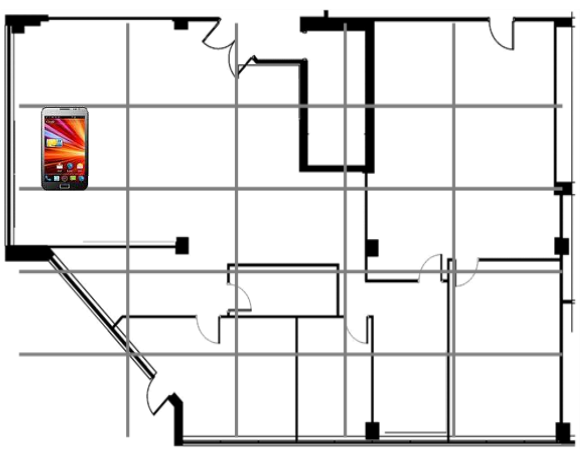
- Formulation

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{Y}=\mathbf{G}\hat{\mathbf{X}}} |\hat{\mathbf{X}}|_1,$$



Sparse location estimation

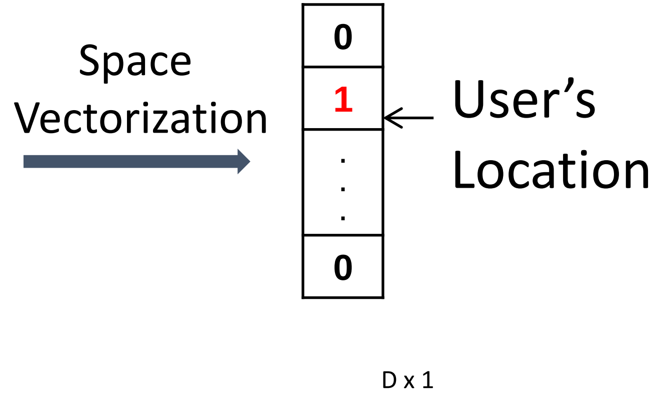
- **The location of mobile device is sparse in space.**



0	0	0	0	0
1	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0

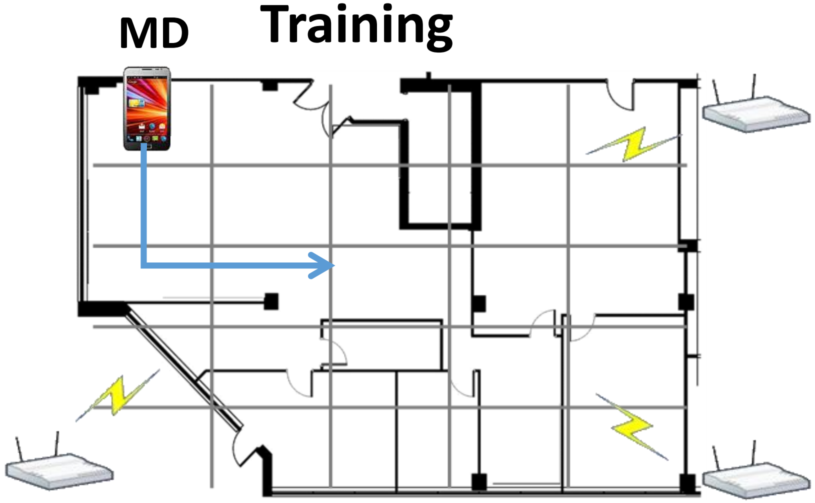
D cells

Sparse vector: b



Sparse location estimation

Training

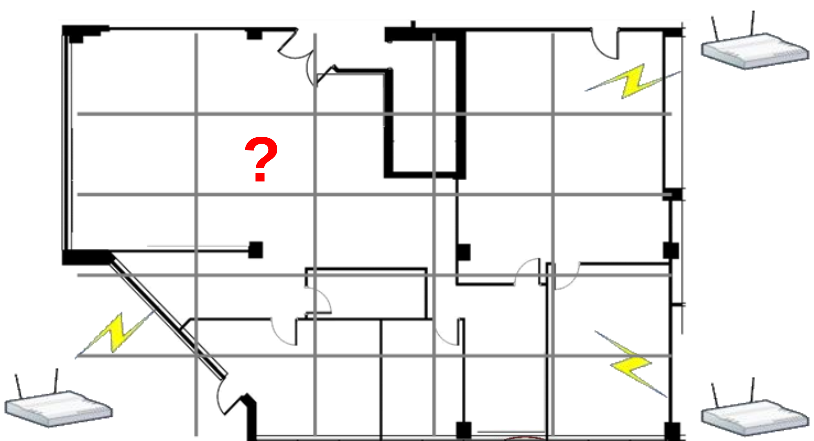


RSS measurements are collected for each position.

Signature map

$T_{AP_1,Cell_1}$	$T_{AP_1,Cell_2}$	…	$T_{AP_1,Cell_D}$
$T_{AP_2,Cell_1}$	$T_{AP_2,Cell_2}$	…	$T_{AP_2,Cell_D}$
$T_{AP3,Cell_1}$	$T_{AP_3,Cell_2}$	…	$T_{AP3,Cell_D}$

Runtime



Runtime measurements

R_{AP_1}	R_{AP_2}	R_{AP3}

Compare

$T_{AP_1,Cell_1}$	$T_{AP_1,Cell_2}$	…	$T_{AP_1,Cell_D}$
$T_{AP_2,Cell_1}$	$T_{AP_2,Cell_2}$	…	$T_{AP_2,Cell_D}$
$T_{AP_3,Cell_1}$	$T_{AP_3,Cell_2}$	…	$T_{AP_3,Cell_D}$

Localization Server (LS)

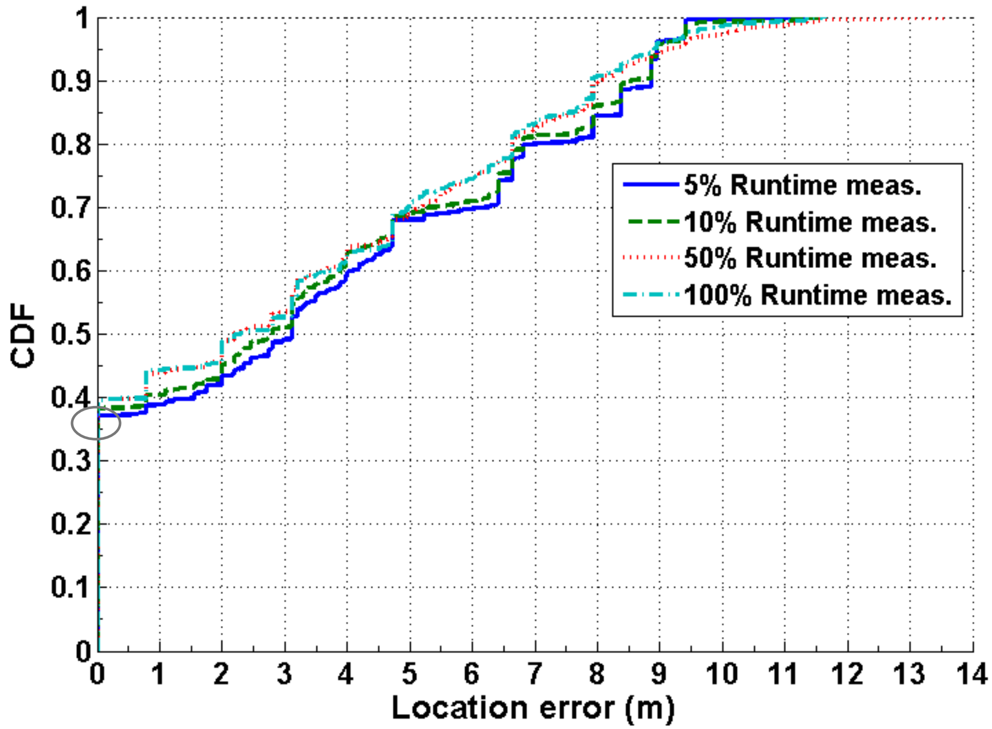
Location Estimation

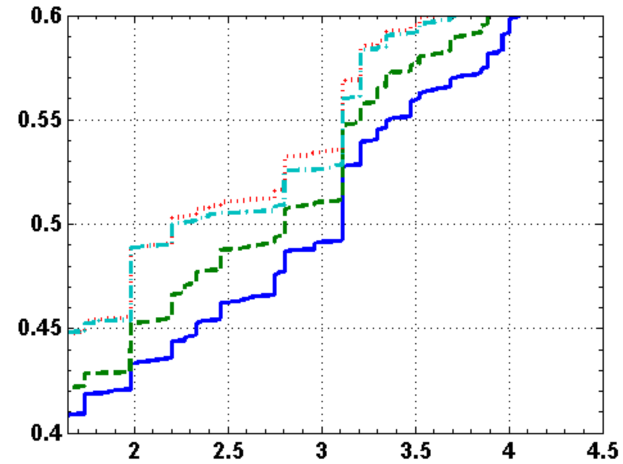
Sparse location estimation

Active laboratory area of 8.5 by 14 meters

5 APs, 135 training cells, cell size: 0.55 x 0.55 m

Online observations: 30 distinct cells, Performance metric: Location Error (m)





Empirical CDF as a function of CS measurements

Key Insights from the Compressive Sensing

1. Sparse or compressible

➢ ***not sufficient alone***

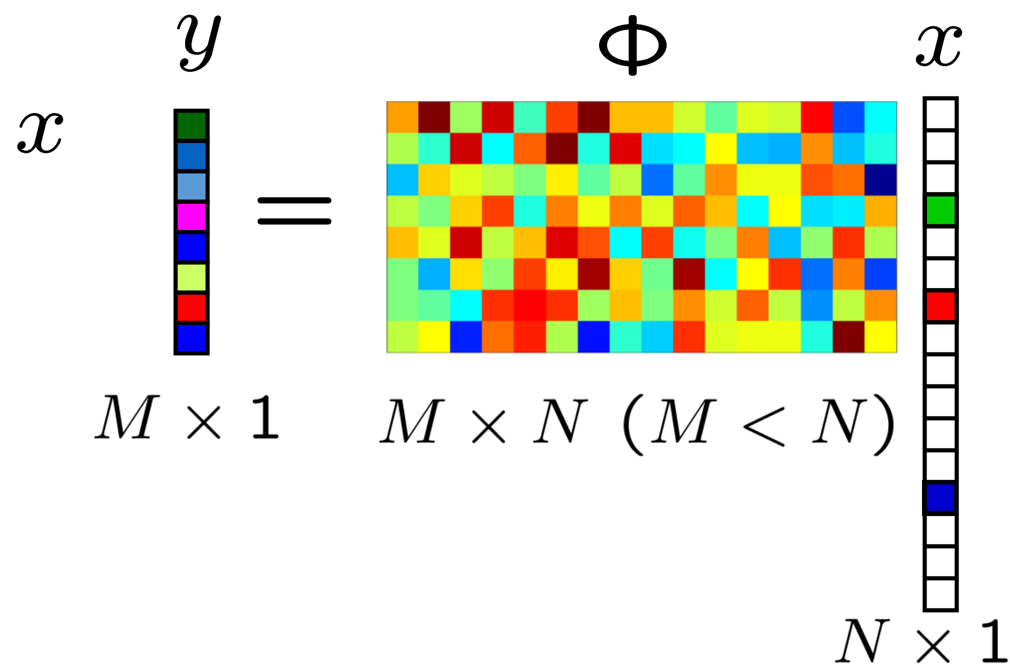
2. Projection Φ

➢ ***information preserving***

(restricted isometry property - RIP)

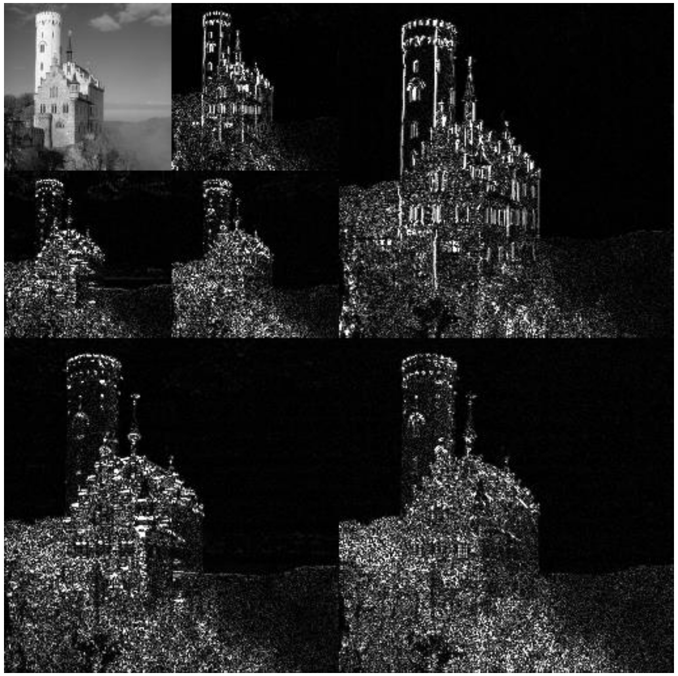
3. Decoding algorithms

➢ ***tractable***



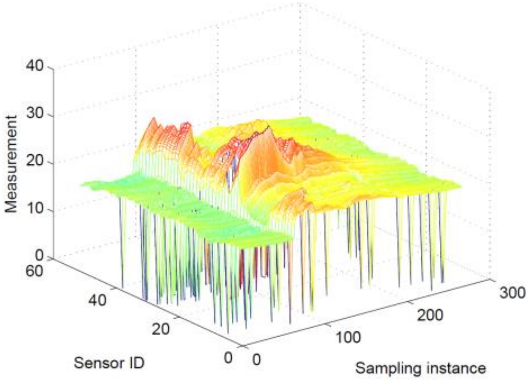
Slide by *Volkan Cevher*

Sparsity in **a basis:** Dictionaries



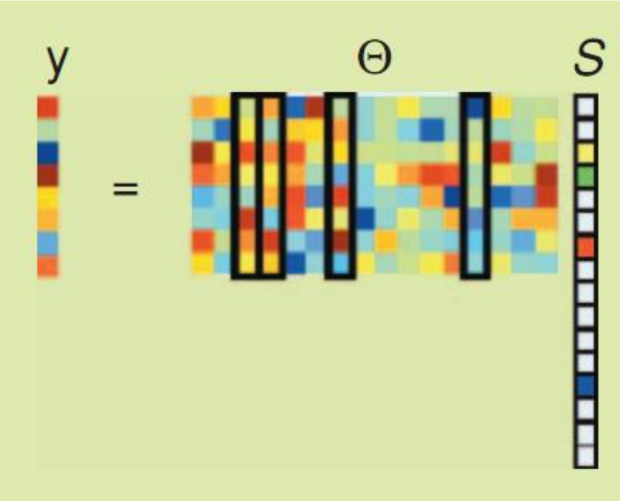
Images

Sound

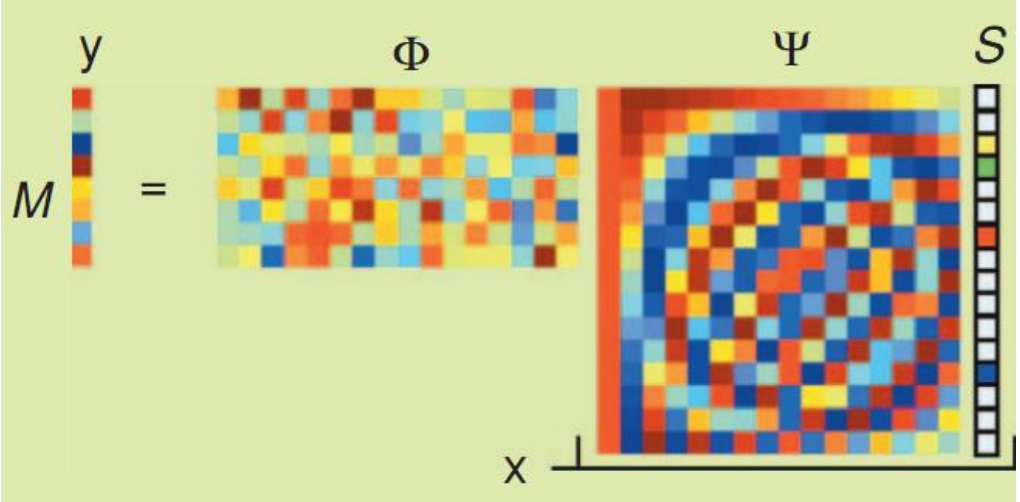


WSN

Sparsity on dictionaries



$$\min_{\mathbf{s}} \|\mathbf{s}\|_1 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{\Theta s}$$

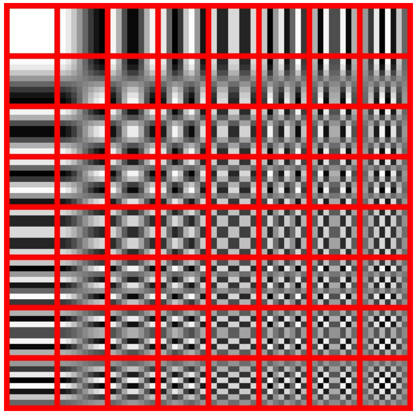


$$\min_{\mathbf{s}} \|\mathbf{s}\|_1 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{\Phi \Psi s}$$

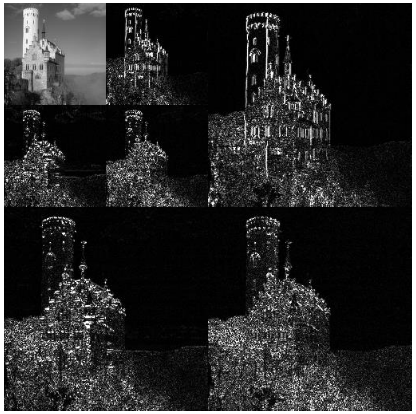
Sparse Modeling: Approach 1

- Step 1: Choose a signal model with structure
  - e.g. bandlimited, smooth with *r* vanishing moments, etc.
- Step 2: *Analytically* design a sparsifying basis/frame that exploits this structure
  - e.g. DCT, wavelets, Gabor, etc.

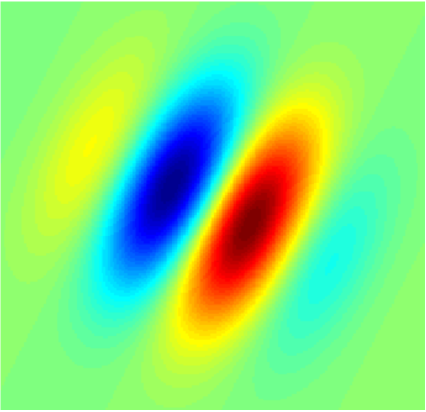
DCT



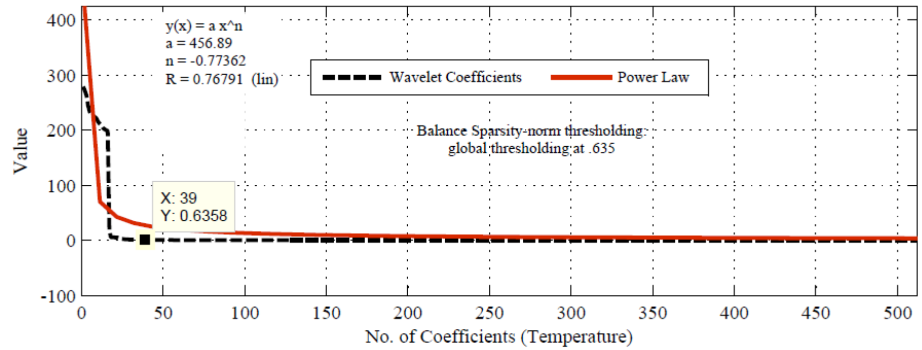
Wavelets

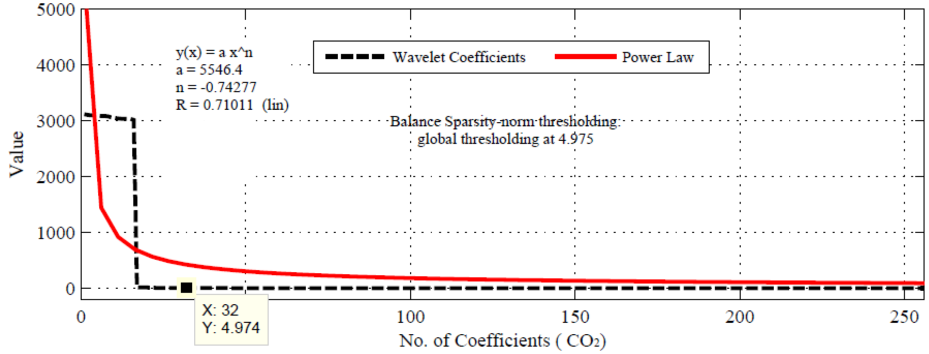


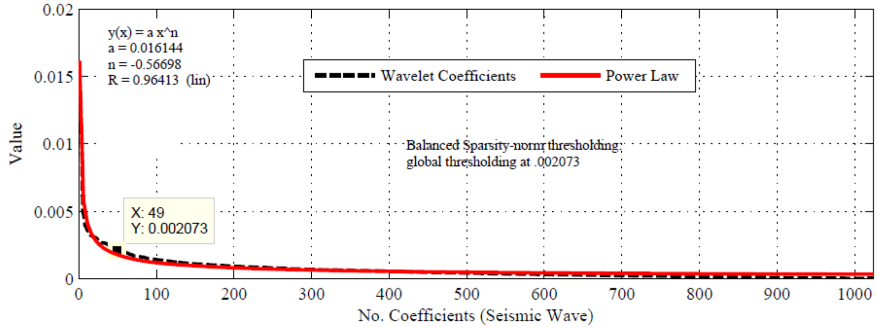
Gabor



WSN data on dictionaries

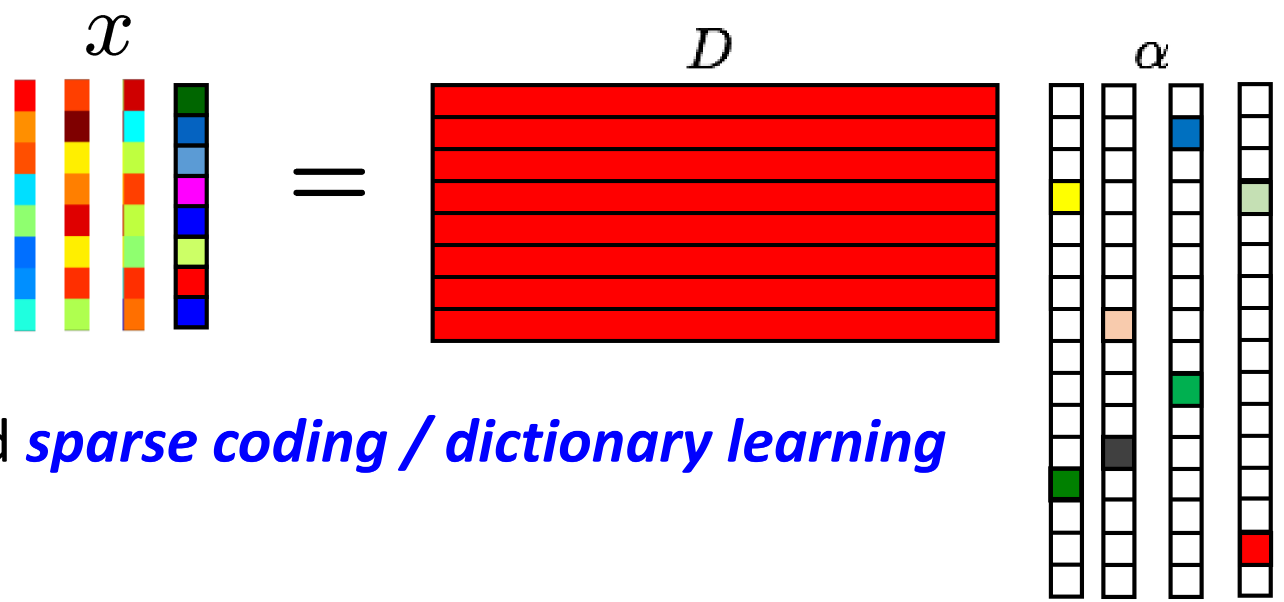






Sparse Modeling: Approach 2

- Learn the sparsifying basis/frame from ***training data***
- Problem formulation: given a large number of training signals, design a **dictionary** *D* that **simultaneously** sparsifies the training data



- Called ***sparse coding / dictionary learning***

Dictionary Learning

- Requirement: Incoherence: correlation between Φ and D

$$\mu(\Phi, d) = \sqrt{N} \cdot \max_{1 \le i, j \le N} |\langle \phi_i, d_j \rangle|$$

- Typical formulation: Given training data

$$X = \begin{pmatrix} | & | & & | \\ x_1 & x_2 & \dots & x_T \\ | & | & & | \end{pmatrix} \Rightarrow \min_{\substack{\alpha_i \in \mathbb{R}^Q \\ D \in \mathbb{R}^{N \times Q}}} \sum_i \|x_i - D\alpha_i\|_2^2 + \lambda \|\alpha_i\|_0$$

- Efficient algorithms, MOD, ***K-SVD***

K–SVD – An Overview

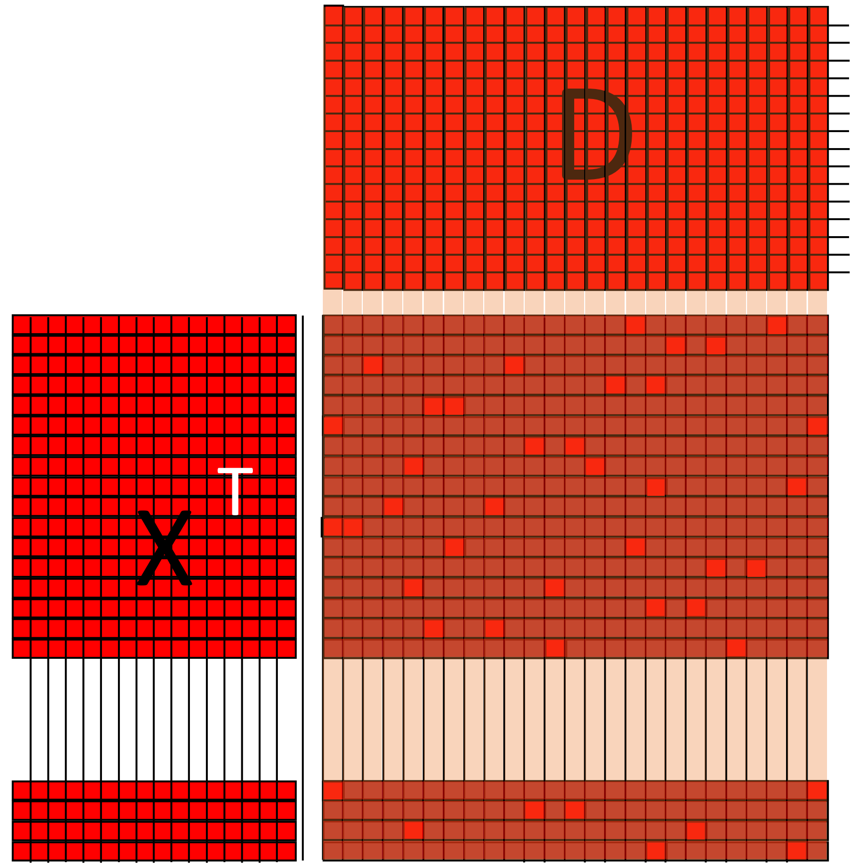
Initialize D

Sparse Coding

Use MP or BP

Dictionary Update

Column-by-Column by SVD computation



Aharon, Elad & Bruckstein ('04)

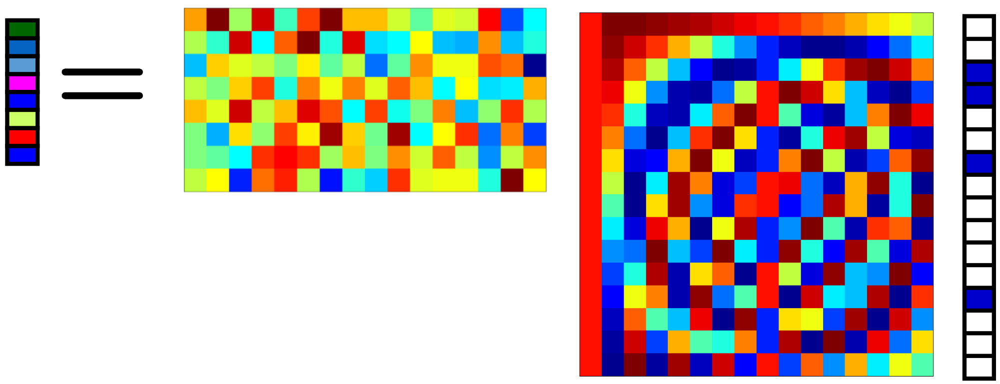
Complex Signal Reconstruction

Possible if signal is **sparse** in dictionary

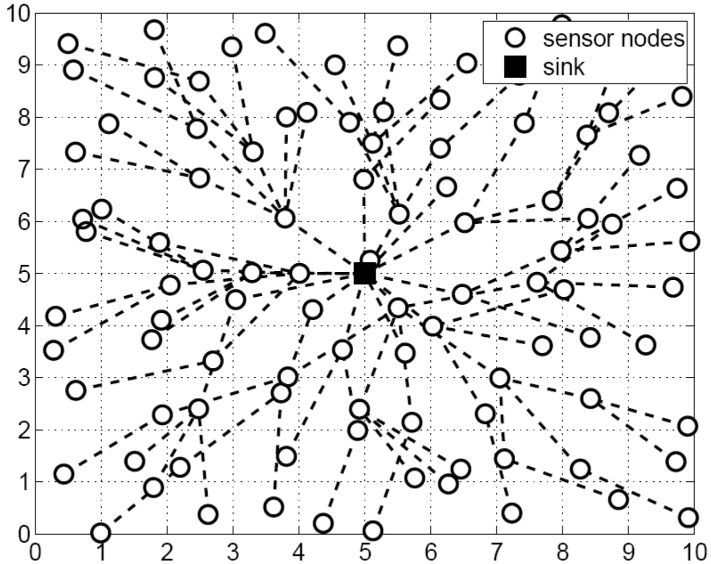
$y = Ds$ where $||s||_0 < K$

Reconstruction based on L_1 minimization

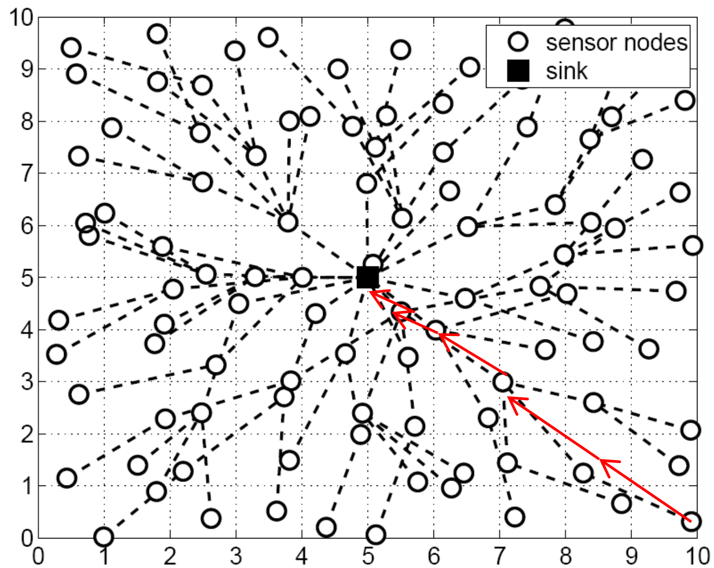
$$\min_{\mathbf{s}} ||\mathbf{s}||_1 \text{ subject to } ||\mathbf{y} - \mathbf{\Phi D s}||_2^2 < \epsilon$$

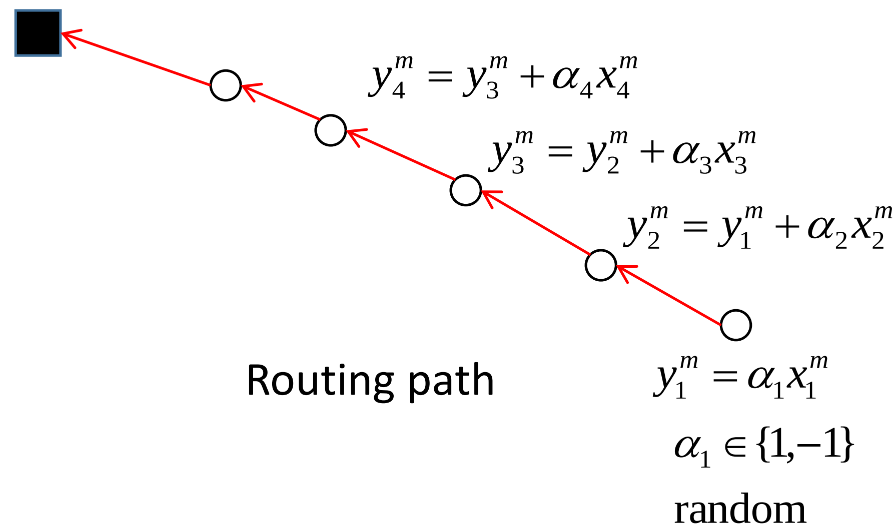


Compressive Data Gathering in WSN



Compressive Data Gathering in WSN





$$\begin{pmatrix} \cdots\cdots \\ \cdots\cdots \\ \cdots\cdots \\ y_{l_m}^m \end{pmatrix} = \begin{pmatrix} \cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots \\ \alpha_1 \;\; \alpha_2 \;\; \alpha_3 \ldots \alpha_{l_m}, \ldots \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ . \\ x_n \end{pmatrix}$$

Compressive Data Gathering

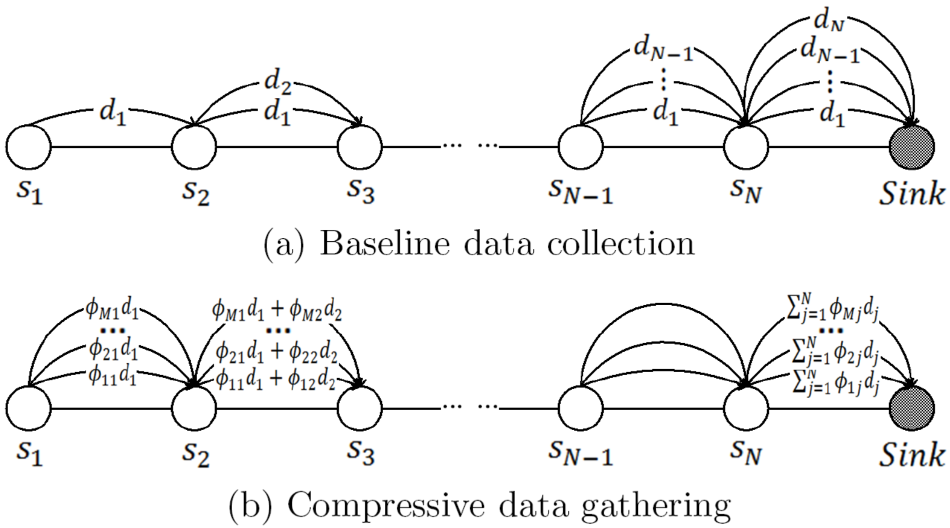


Figure 2: Comparing baseline data collection and compressive data gathering in a multi-hop route

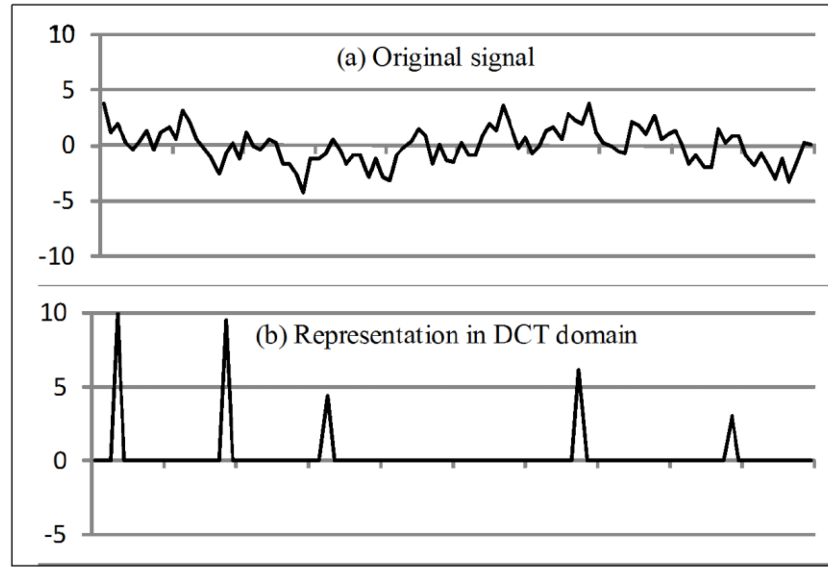
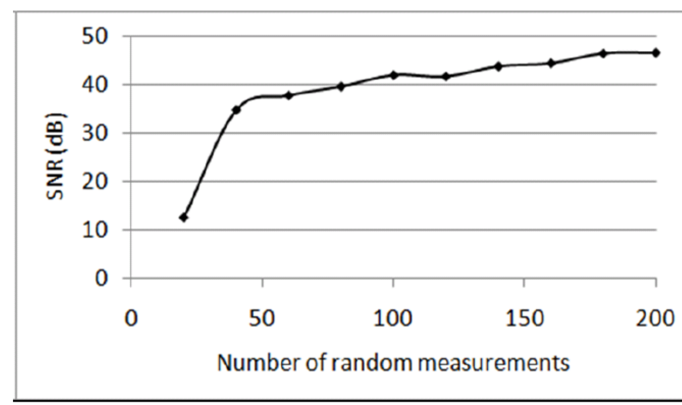
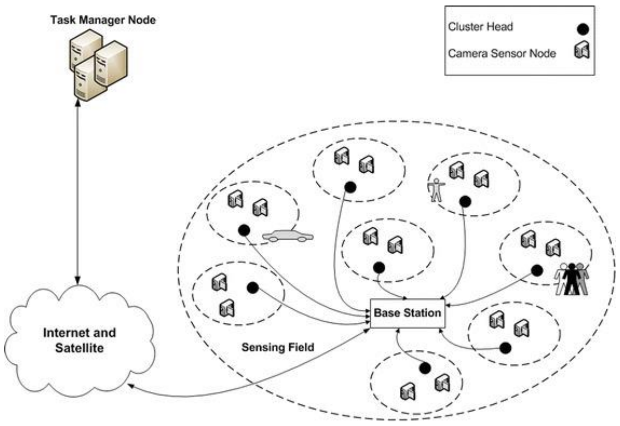
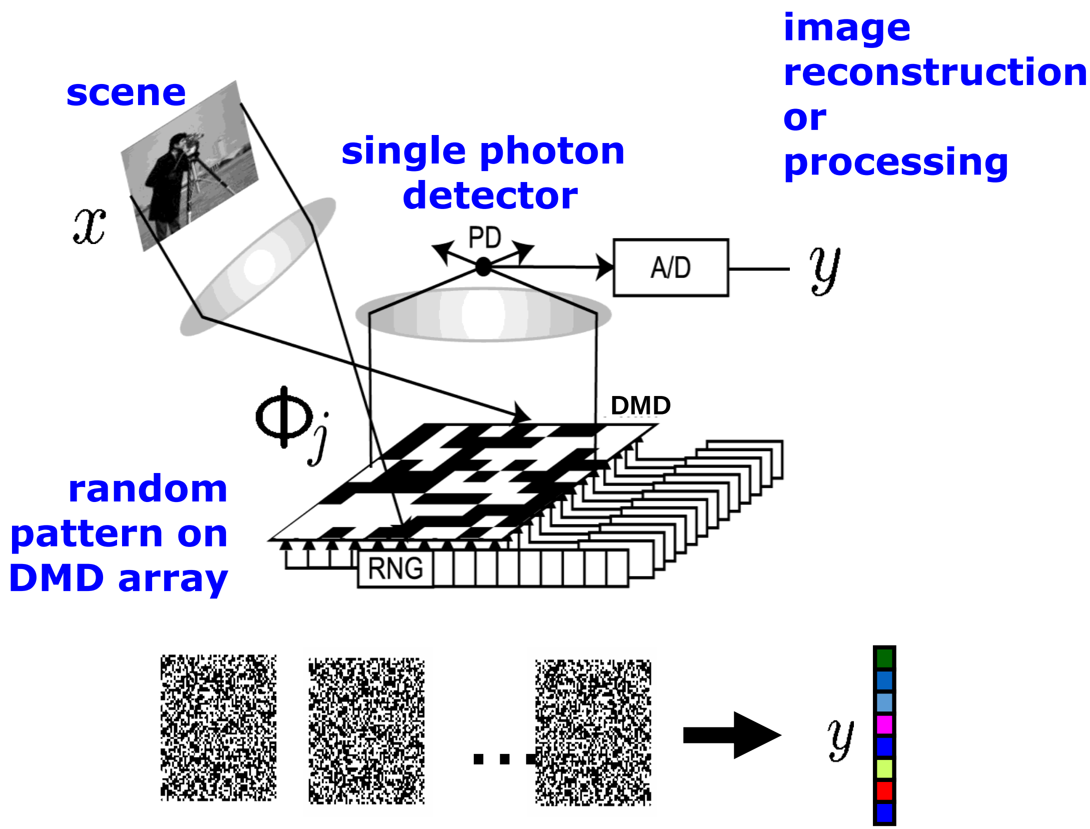


Figure 5: A 5-sparse signal in DCT domain



CS in Wireless Video Sensor Networks

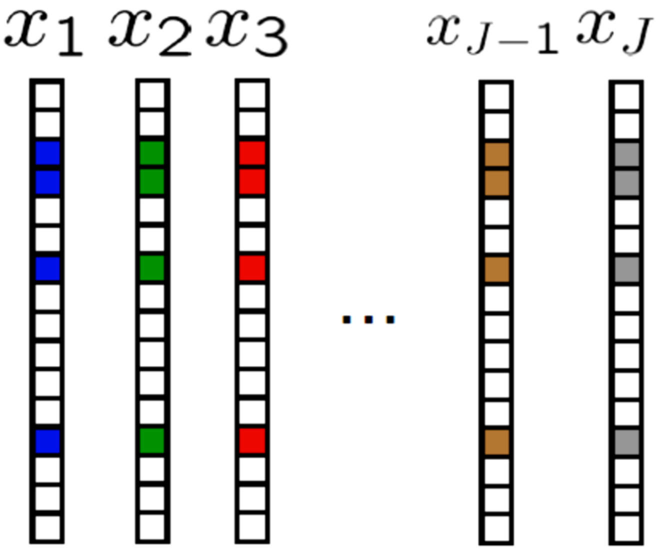




Extending CS

Joint sparsity

- *share sparse components*
- *different coefficients*



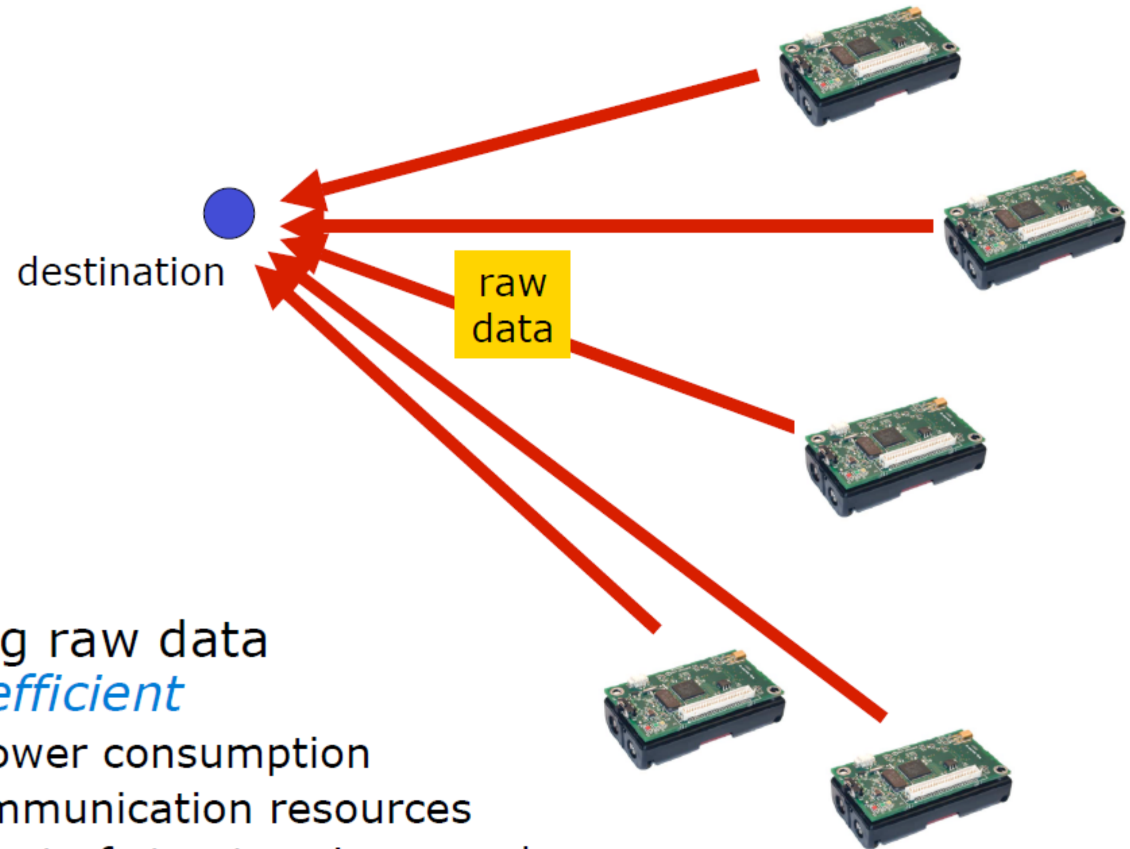
Mixed ℓ_2/ℓ_1 -norm solutions

$$\widehat{X} = \arg\min \|X\|_{2,1} \text{ s.t. } Y = \Phi X$$

$$\widehat{X} = \arg\min \|X\|_{2,1} \text{ s.t. } \|Y - \Phi X\|_2 \leq \epsilon$$

Greedy solutions: simultaneous orthogonal matching pursuit

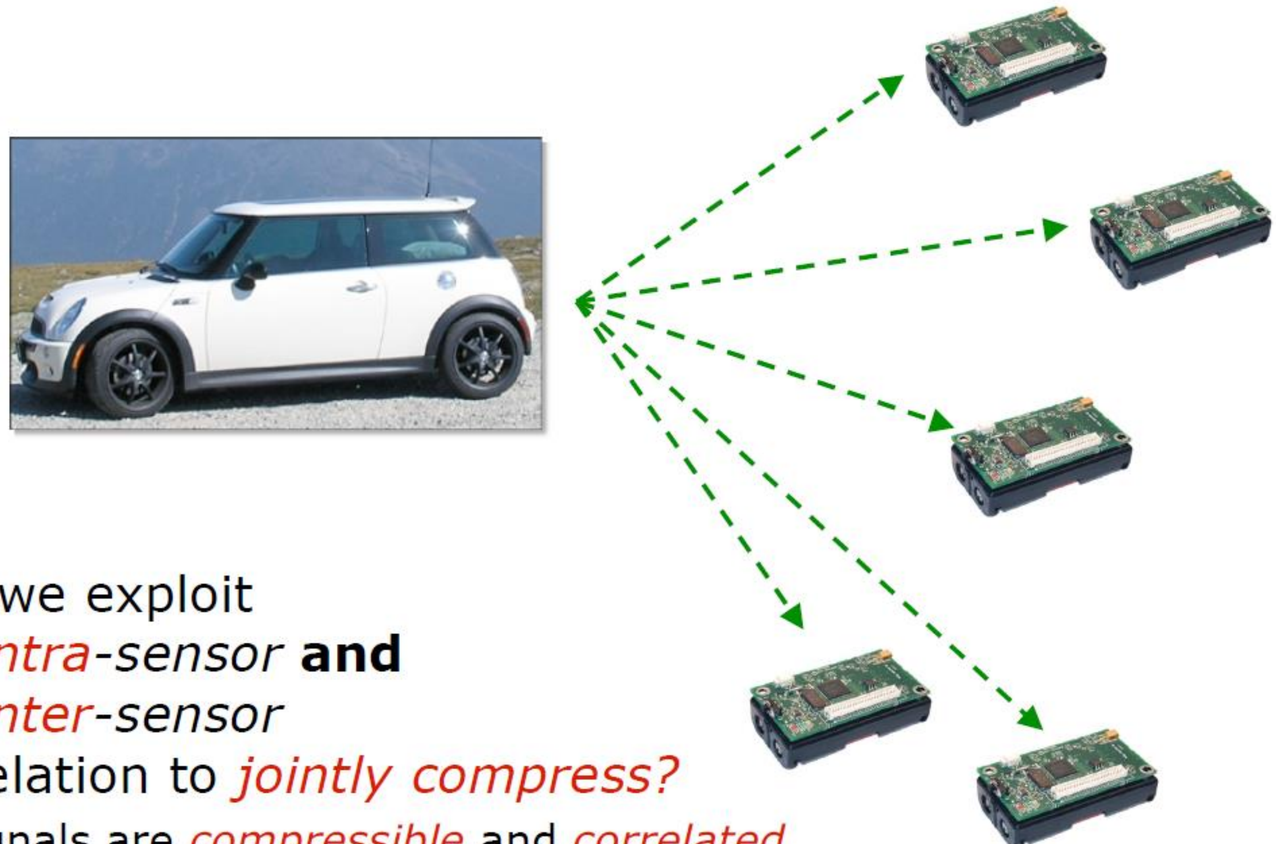
The Need for Compression



- Transmitting raw data typically *inefficient*
  - reduced power consumption
  - limited communication resources
  - large amount of structure in sensed signals

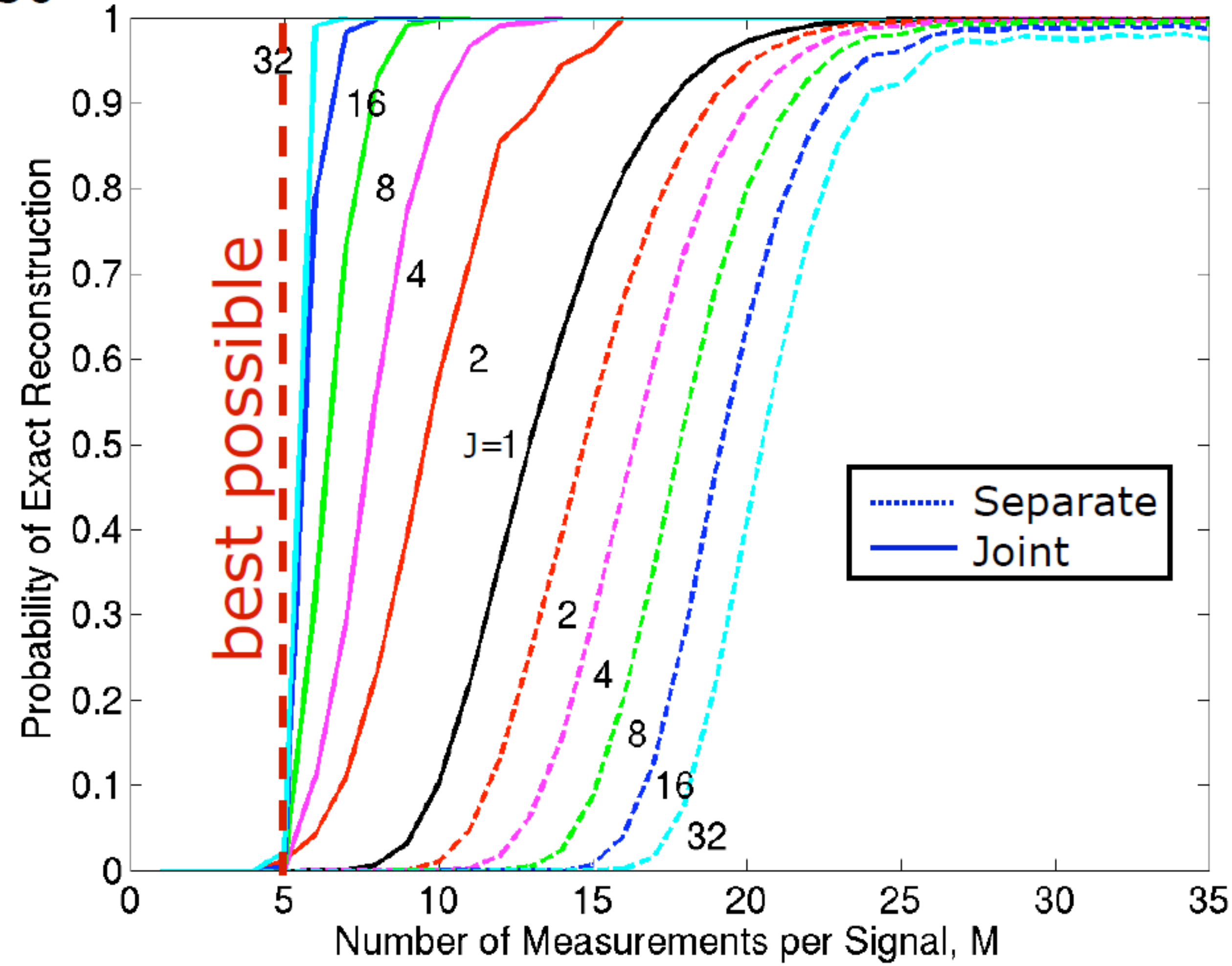
Universal Distributed Sensing via Random Projections
M. F. Duarte, M. B. Wakin, D. Baron, and R. G. Baraniuk

Correlation



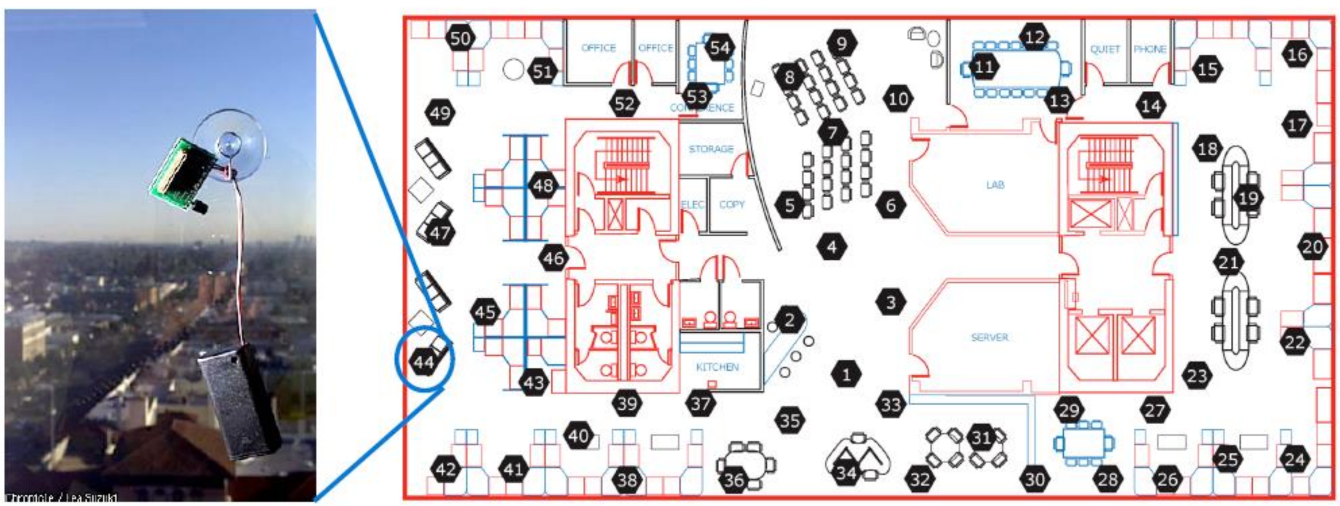
- Can we exploit
  intra-sensor **and**
  inter-sensor
  correlation to *jointly compress?*
  - signals are *compressible* and *correlated*
- *Distributed source coding* problem

K=5 *N*=50 Common Sparse Support Results

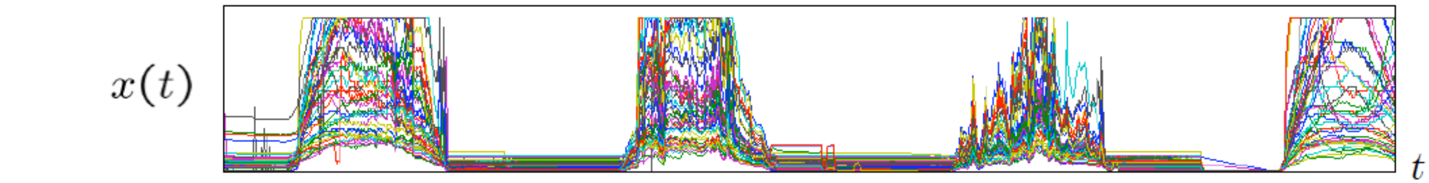


Real Data Example

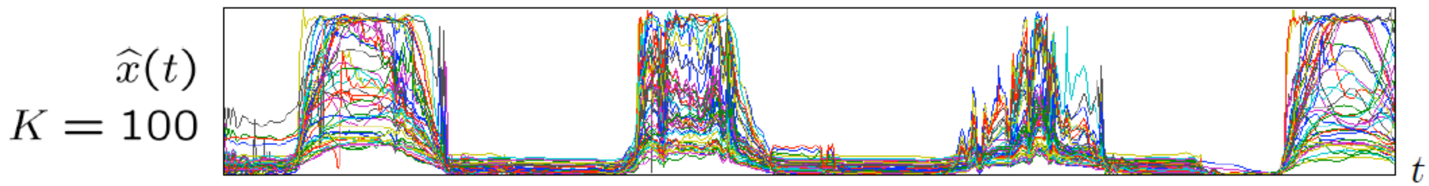
- Dataset: Indoor Environmental Sensing
- $J = 49$ sensors, $N = 1024$ samples each
- Compare compression using:
  - transform coding approx — K largest terms per sensor
  - independent CS — $4K$ measurements per sensor
  - DCS: common sparse supports — $4K$ measurements per sensor



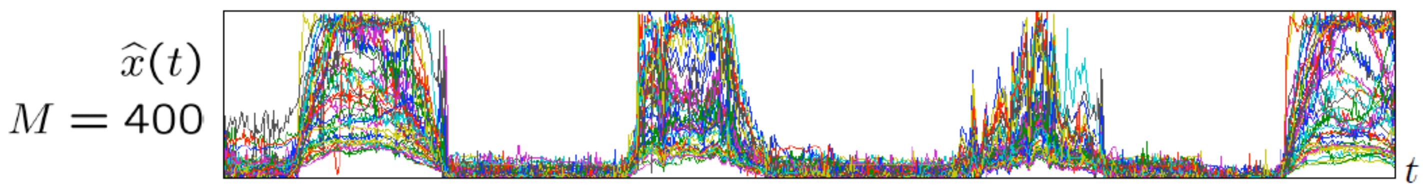
Light Intensity - Wavelets



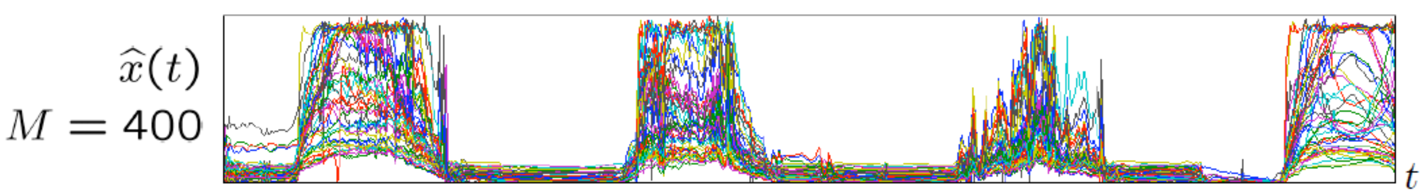
(a) Original



(b) Transform Coding, SNR = 26.4842 dB



(c) Compressed Sensing, SNR = 21.6426 dB



(d) Distributed Compressed Sensing, SNR = 27.1906 dB

Reading material

Haupt, J., Bajwa, W. U., Rabbat, M., & Nowak, R. "Compressed sensing for networked data." IEEE Signal Processing Magazine, vol. 25(2), 92-101, 2008.

Qaisar, Saad, Rana Muhammad Bilal, Wafa Iqbal, Muqaddas Naureen, and Sungyoung Lee. "Compressive sensing: From theory to applications, a survey." *Journal of Communications and networks* 15, no. 5 (2013): 443-456.

Useful links

http://dsp.rice.edu/cs

http://nuit-blanche.blogspot.gr/