Lecture 3
Compressive Sensing and Its Application in WSNs

Qi Zhang

(Email: qz@eng.au.dk; office E314)

AARHUS SCHOOL OF ENGINEERING

Wireless Sensor Networks

- Large numbers of sensors
  - image data bases, camera arrays, distributed wireless sensor networks, …
- Some sensor measuring process is costly
- Deluge of data
  - Internet of things, M2M, 24 billion nodes in 2020
  - how to acquire, transmit/store, fuse, process efficiently?

- Facts
  - We are often not dealing with arbitrary signals
  - Most signals of our interests are sparse or highly compressible, i.e., it can be represented by a set of sparse coefficients

ASE | AARHUS SCHOOL OF ENGINEERING

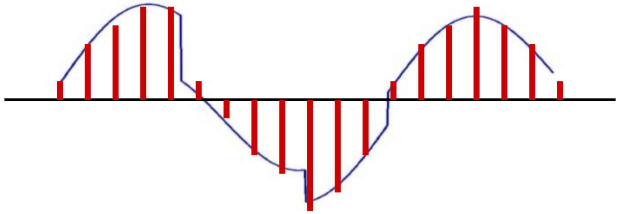
Shannon's Sampling Theorem

- Sampling: The process of reduction of a continuous signal to a discrete signal.
- Shannon's Sampling Theorem:
  - To exactly reconstruct an arbitrary band-limited signal from its samples, the sampling rate needs to be at least twice of the bandwidth,
  - i.e., If a function x(t) contains no frequencies higher than B hertz, it is completely determined by giving its ordinates at a series of points spaced 1/(2B) seconds apart.

Claude Shannon
(1916–2001)

Harry Nyquist
(1889–1976)

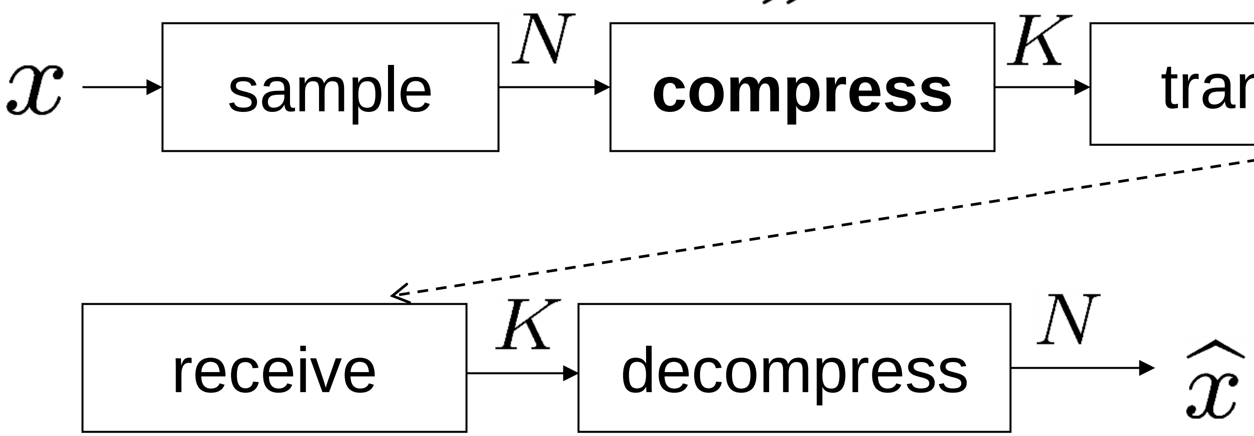
Sufficient condition for perfect reconstruction but not necessary condition!



Sensing by *Sampling (Traditional)*

- Long-established paradigm for digital data acquisition
  - uniformly sample data at Nyquist rate (2x Fourier bandwidth)
  - compress data (signal-dependent, nonlinear)

$$N \gg K$$



AARHUS SCHOOL OF ENGINEERING

Rethink Shannon/Nyquist Theorem

- 2x oversampling Nyquist rate is a worst-case bound for *any* band-limited data
- Sparsity/compressibility irrelevant
- Shannon sampling is a linear process while compression is a nonlinear process

AARHUS SCHOOL OF ENGINEERING

Compressive Sensing/ Compressed Sampling (CS)

- A new sampling theory that leverages compressibility
- Allows a sampling rate significantly lower than the Nyquist rate.
- based on new uncertainty principles
- randomness plays a key role

Summary the entire process of CS

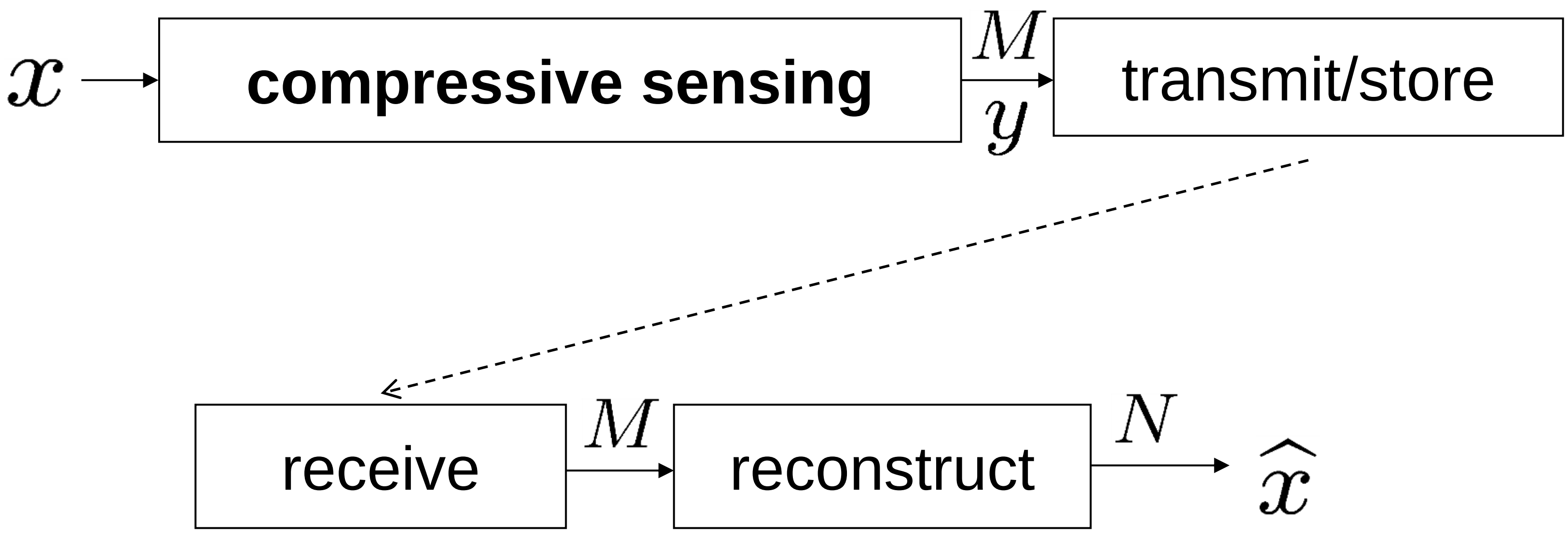
- Steps of CS
  1. Signal sparse representation
  2. Linear encoding and measurement collection
  3. Non-linear decoding (sparse recovery)

AARHUS SCHOOL OF ENGINEERING

Compressive Sensing

- Directly acquire “compressed” data
- Replace samples by more general “measurements”

$$K < \underline{M} \ll N$$

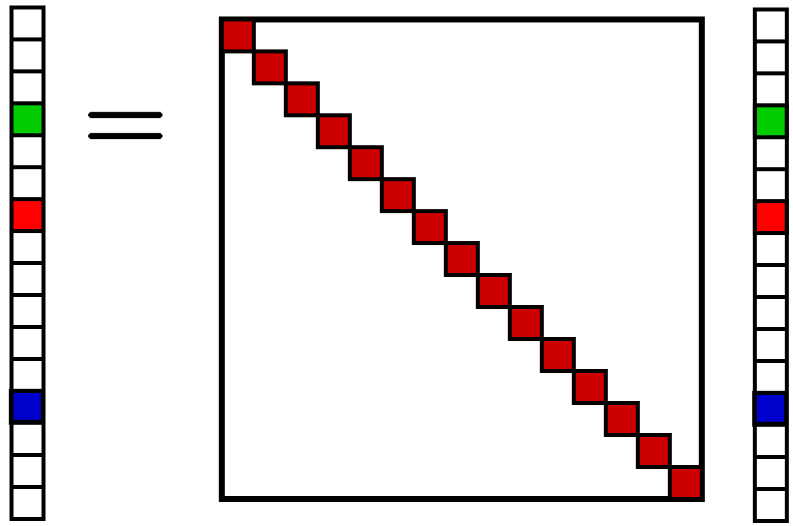


Sampling

- Signal x is K-*sparse* in basis/dictionary Ψ
  - assume sparse in space domain $\Psi = I$

- **Samples**

$$N \times 1 \qquad y = \Phi x, \quad \Phi = I \qquad N \times 1$$



K

ASE AARHUS SCHOOL OF ENGINEERING

Basis representation

- We can take the signal apart, writing it as a discrete linear combination of “atoms”:

$$x(t) = \sum_{i \in N} \alpha(i)\psi_i(t)$$

for some fixed set of *basis* signal $\{\psi_i(t)\}_{i \in N}$

- E.g. Fourier series:

$$x(t) = \sum_{k \epsilon Z} \alpha(k)e^{j2\pi kt}$$

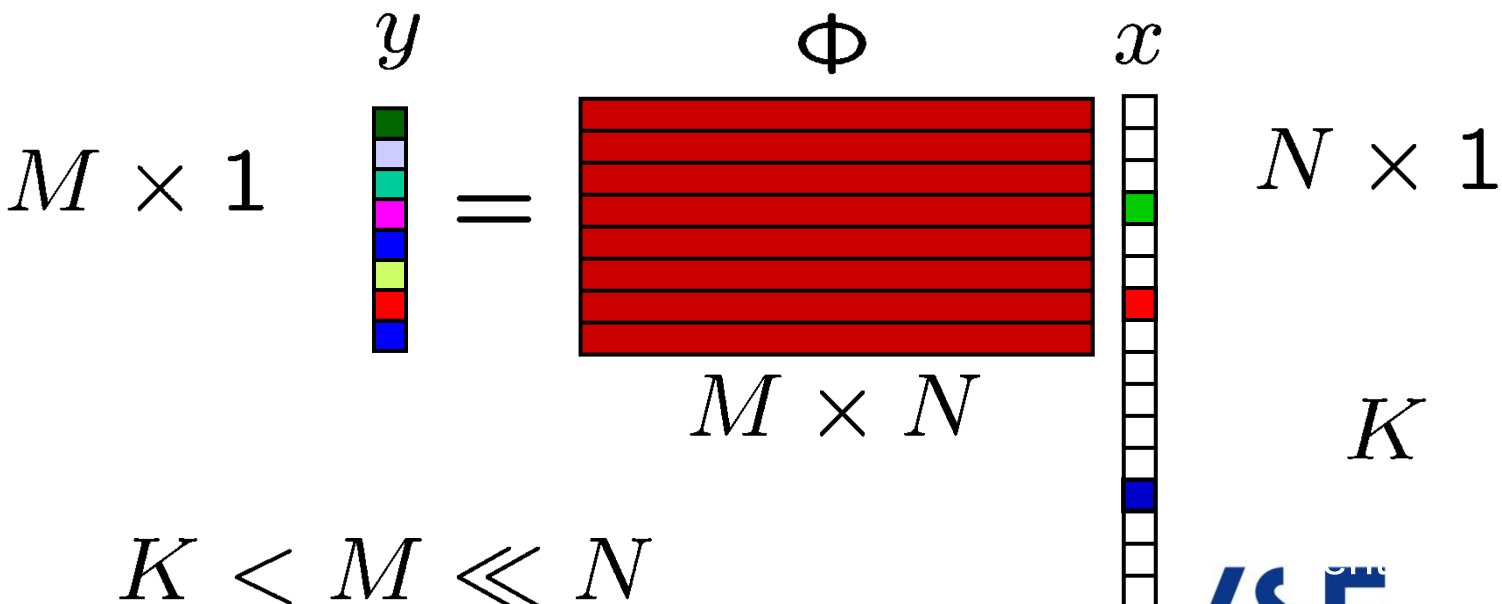
- Orthonormal basis (Orthobasis): if inner product

$$\langle \psi_i, \psi_j \rangle = \begin{cases} 1 \; i = j \\ 0 \; i \neq j \end{cases}$$

Compressive Data Acquisition

- When data is sparse/compressible, can directly acquire a ***condensed representation*** with no/little information loss through ***dimensionality reduction***

$$y = \Phi x$$



$$K < M \ll N$$

CS Signal Recovery

- Reconstruction/decoding: given $y = \Phi x$
  (ill-posed inverse problem) find x

- Clearly $y = \Phi x$ is an underdetermined equation system
  - An infinite number of solutions
  - Sparsity of x plays its role!

Underdetermined Linear Systems

- A linear system with fewer equations (M) than unknowns (N). (M<N) An example:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 3 & 7 & 1 & 1 & 3 & -2 \\ 2 & 5 & 6 & -9 & 4 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}$$

$\boldsymbol{y}_{M*1} = \Phi_{M*N} \boldsymbol{x}_{N*1}$ under-determined , M<N

- Without additional constraints, it would have infinite many possible solutions!

Sparse Solutions

- If $\boldsymbol{x} = [x_1 \quad 0 \quad 0 \quad x_4 \quad 0 \quad 0]^T$, then the previous underdetermined system becomes

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 3 & 7 & 1 & 1 & 3 & -2 \\ 2 & 5 & 6 & -9 & 4 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ 0 \\ 0 \\ x_4 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 2 & -9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_4 \end{bmatrix}$$

- $\boldsymbol{x}$ is a 2-sparse (has 2 non-zero elements) in this case. There is a unique solution.
- Questions: if $\boldsymbol{x}$ is K-sparse, K<=M will there be always a unique solution?

Solving for K-sparse Solution

- If the non-zero positions of $\boldsymbol{x}$ is known, the original underdetermined linear system Φx = y will reduce to a $M * K$ linear system $\Phi_r \boldsymbol{x}_r = \boldsymbol{y}$
  - where $\boldsymbol{x}_r$ is a reduced vector of non-zero elements of $\boldsymbol{x}$.

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \boxed{\begin{matrix} 3 \\ 2 \end{matrix}} & \begin{matrix} 7 \\ 5 \end{matrix} & \begin{matrix} 1 \\ 6 \end{matrix} & \boxed{\begin{matrix} 1 \\ -9 \end{matrix}} & \begin{matrix} 3 \\ 4 \end{matrix} & \begin{matrix} -2 \\ 7 \end{matrix} \end{bmatrix} \begin{bmatrix} x_1 \\ 0 \\ 0 \\ x_4 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 2 & -9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_4 \end{bmatrix}$$

$$\boldsymbol{y} = \Phi \boldsymbol{x} = \Phi_r \boldsymbol{x}_r$$

- Issue: Non-zero positions of a *K*-sparse vector solution are often unknown!

K-Sparse Solution with Unknown Positions

- An optimization formulation:
  - Find $\boldsymbol{x}$ that minimizes $\|x\|_0$ subject to $\boldsymbol{y} = \Phi\boldsymbol{x}$
    - Where $\|x\|_0 = \sum_{n=1}^{N}|x_n|^0$ = number of non-zero elements of $\boldsymbol{x}$

- A Subset selection problem:
  - Select a subset of K columns of the Φ matrix to form a $M * K$ matrix Φ_r such that $\boldsymbol{x}_r$=$\Phi_r{}^{\dagger}\boldsymbol{y}$ is a K-sparse solution to the underdetermined system $\boldsymbol{y} = \Phi\boldsymbol{x}$ where $\Phi_r{}^{\dagger}$is the pseudo-inverse of the matrix Φ_r.
- An NP hard question with $\binom{N}{K} = \frac{N!}{K!(N-K)!}$ possible choices!

AARHUS SCHOOL OF ENGINEERING

CS Signal Recovery

- Direct solution to "Find **x** that minimizes $\|x\|_0$ subject to $\boldsymbol{y} = \Phi \boldsymbol{x}$ " is hard (solving a subset selection problem which is NP-hard)
- A tractable decoder is needed!
- Replacing the l_0 "norm" by the l_1 norm, i.e., relaxing the objective function of "minimizing $\|x\|_0$" to "minimizing $\|x\|_1$"
  - **Find $\boldsymbol{x}$ that minimizes $\|x\|_1$ subject to $\boldsymbol{y} = \Phi \boldsymbol{x}$.**
  - Where $\|x\|_1 = \sum_{n=1}^{N} |x_n|$
- It is a convex optimization problem that can be solved using linear programming efficiently
- Also known as "Basis Pursuit"

AARHUS SCHOOL OF ENGINEERING

CS Signal Recovery

- Ideally, to recover $\boldsymbol{x}$ when the number of measurement M=2K
  - The amount of information that $\boldsymbol{x}$ carries, as $\boldsymbol{x}$ is uniquely determined by the K indices and the K values of its non-zero entries
- However, we must pay a reasonable price for not knowing the support of $\boldsymbol{x}$ (i.e., the position of the non-zero entries. There are totally $\binom{N}{K} = \frac{N!}{K!(N-K)!}$ possiblilies!)

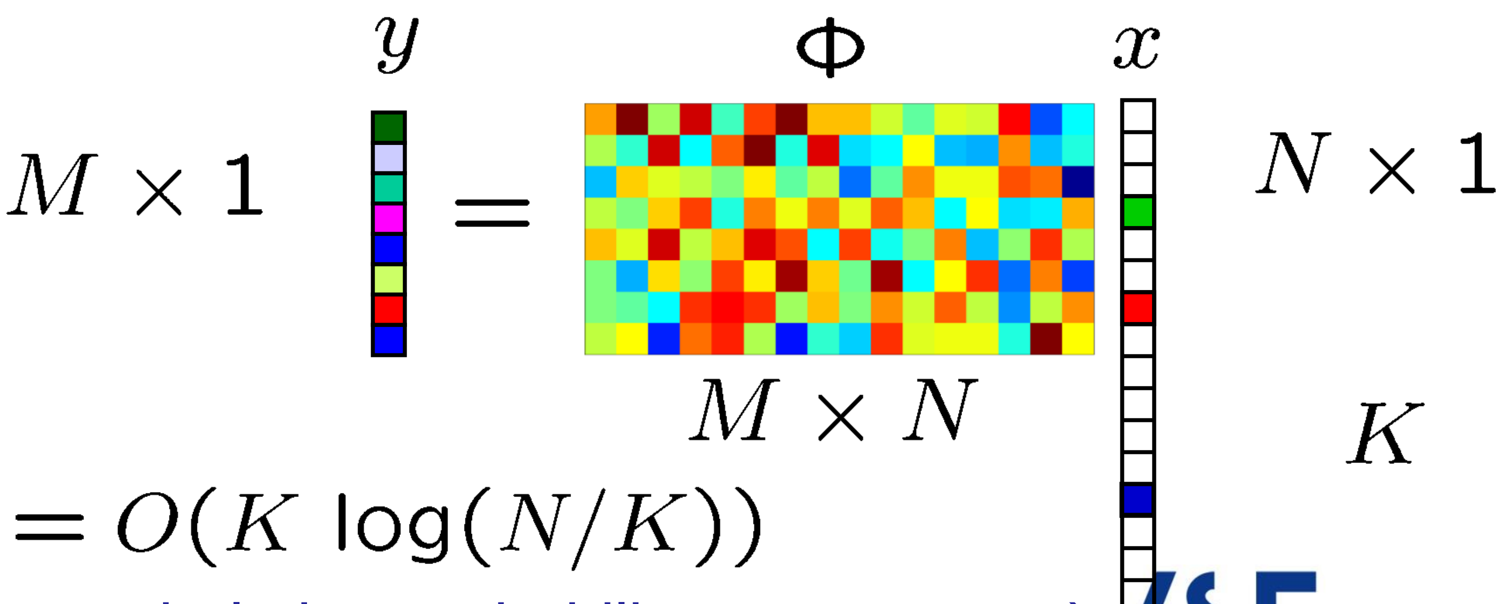
AARHUS SCHOOL OF ENGINEERING

CS Signal Recovery

- When data is sparse/compressible, we can directly acquire a ***condensed representation*** with no/little information loss

$$y = \Phi x$$

- ***Random projection*** will work



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

(with overwhelming probability to recover x)

AARHUS SCHOOL OF ENGINEERING

Incoherence

- Definition
  - The coherence between the sensing/sampling basis Φ and the representation basis Ψ is

$$\mu(\Phi, \Psi) = \sqrt{n} \cdot \max_{1 \le k,j \le n} \left| \langle \phi_k, \psi_j \rangle \right|$$

  The coherence measures the largest correlation between any two elements of Φ and Ψ. $\mu(\Phi, \Psi) \in [1, \sqrt{n}]$

- Compressive sensing needs a pair with low coherence, i.e., maximum incoherence $\mu(\Phi, \Psi)$=1
  - E.g., One such pair is the time-frequency pair. A narrow rectangle function in the time domain corresponds to the wide-spread function in the frequency domain.
  - E.g., random matrices are largely incoherent with any fixed basis Ψ.

Incoherence (cont…)

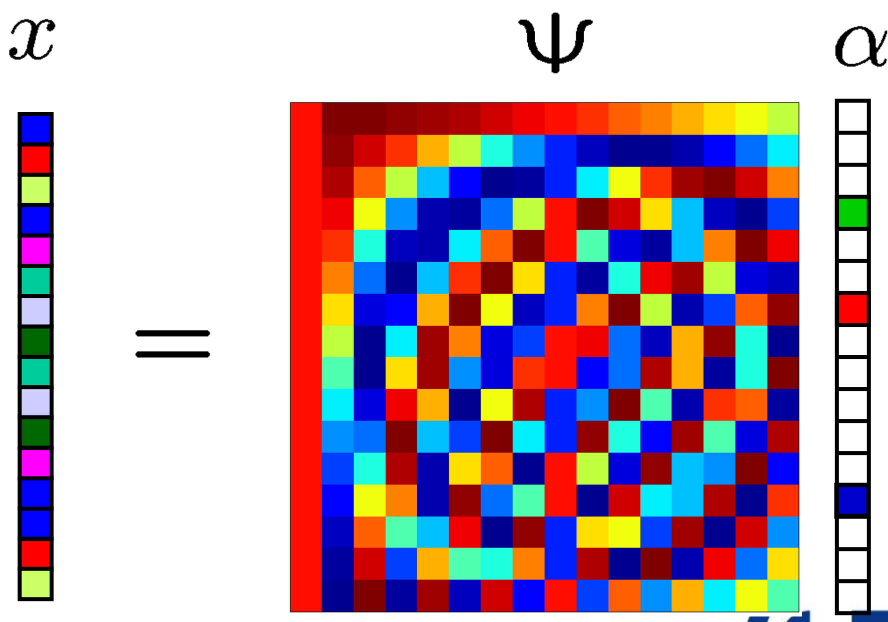
- By extension

Random waveform $\varphi_i(t)$ with independent identical distribution (i.i.d.) entries, e.g., Gaussian or ± 1 binary entries will exhibit a very low coherence with any fixed representation Ψ .

Universality

- Random measurements can be used for signals sparse in *any* basis

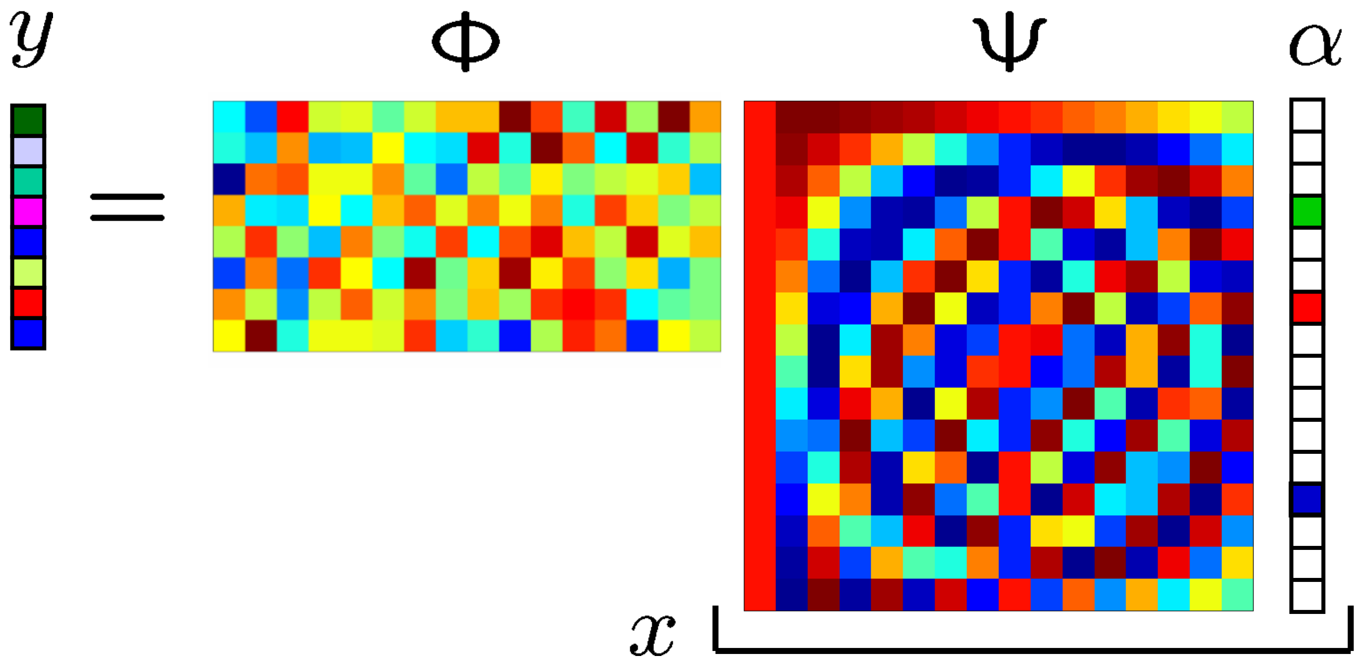
$$x = \Psi \alpha$$



Universality

- Random measurements can be used for signals sparse in *any* basis

$$y = \Phi x = \Phi \Psi \alpha$$

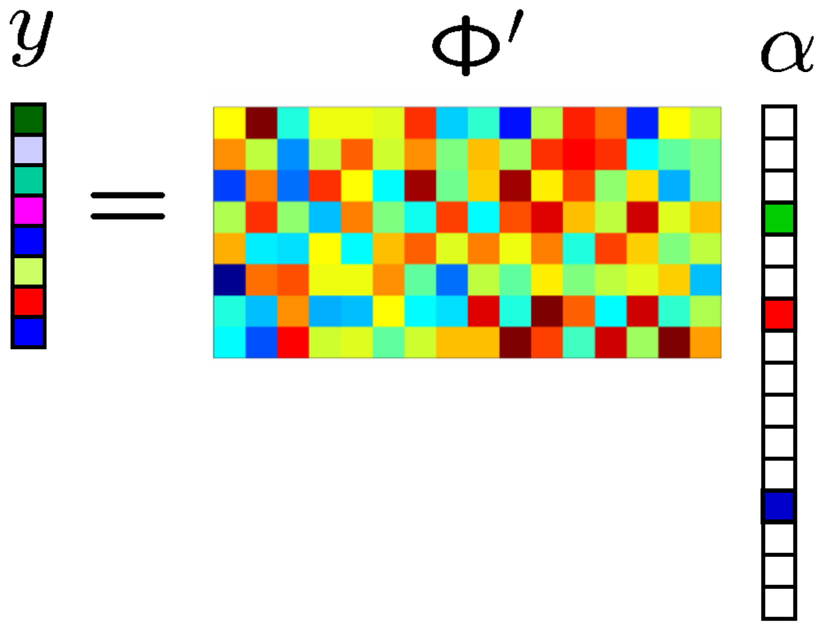


AARHUS SCHOOL OF ENGINEERING

Universality

- Random measurements can be used for signals sparse in *any* basis

$$y = \Phi x = \Phi \Psi \alpha = \Phi' \alpha$$



$N \times 1$

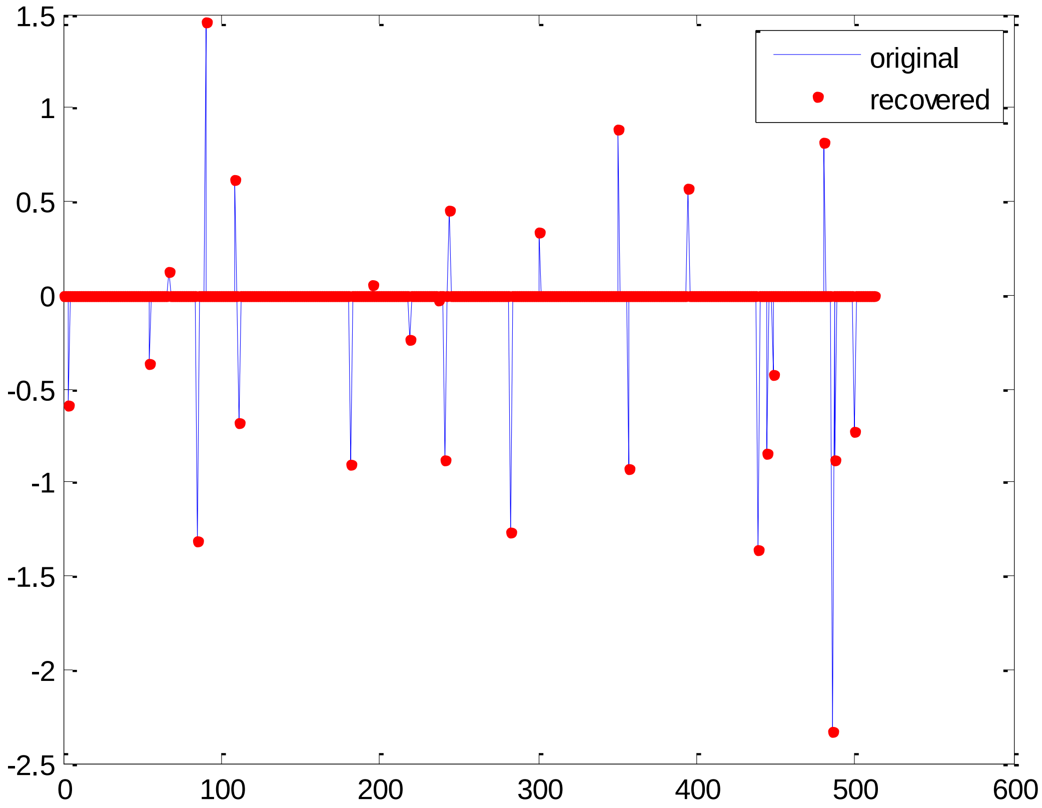
sparse coefficient vector

K

nonzero entries

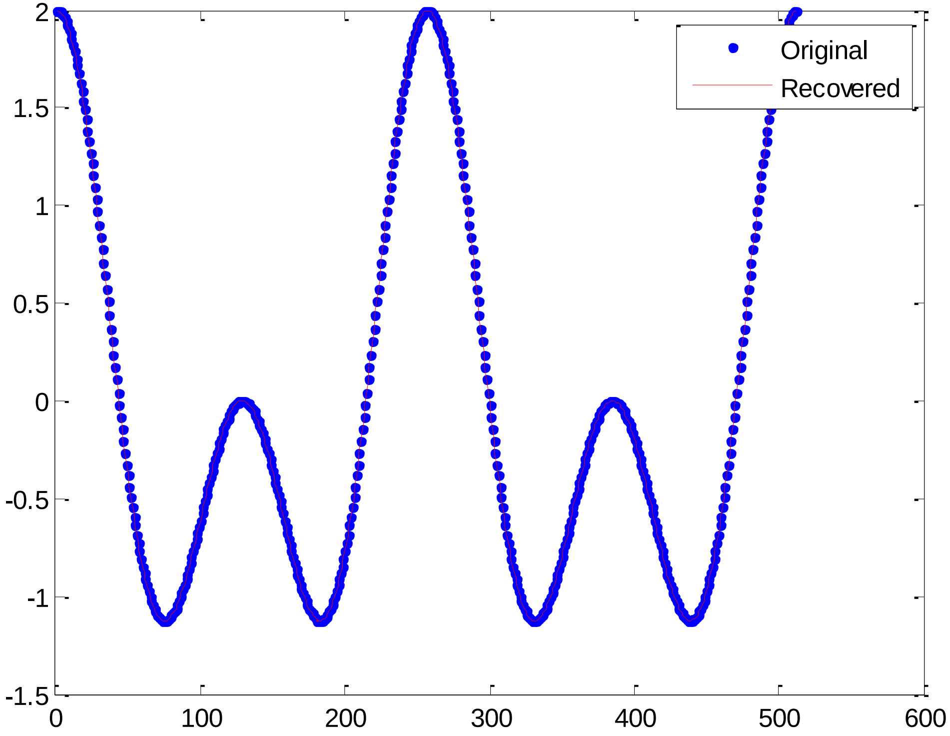
Sparse recovery example 1

- N=512;
- K=25;
- M=6*K;



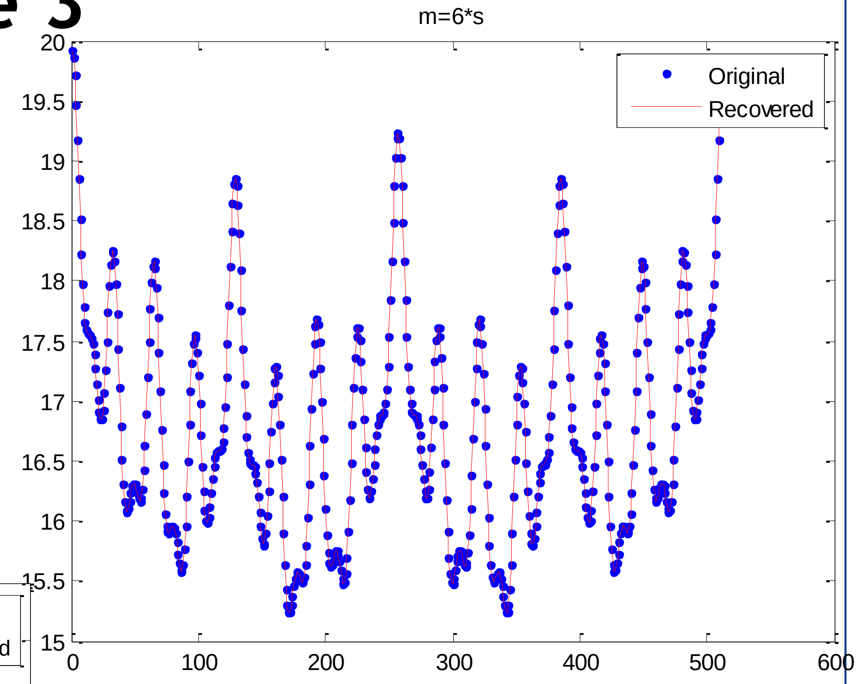
Sparse recovery example 2

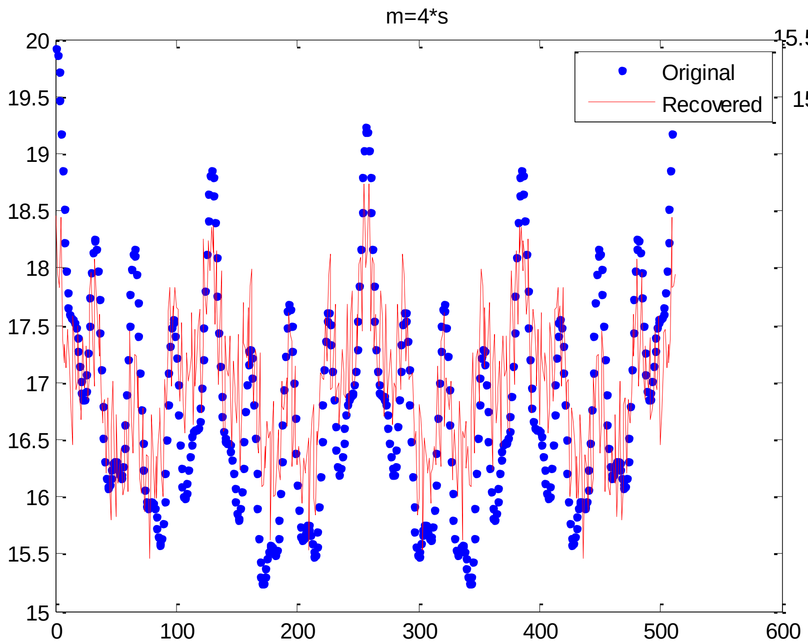
- N=512;
- K= 4 (nearly sparse)
- M=4*K



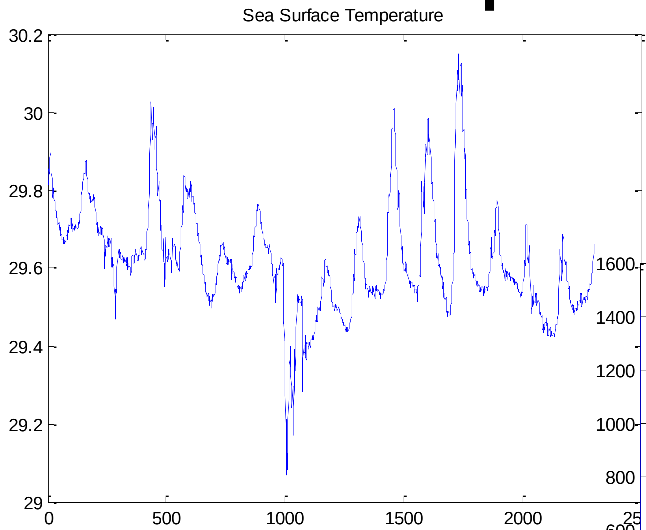
Sparse recovery example 3

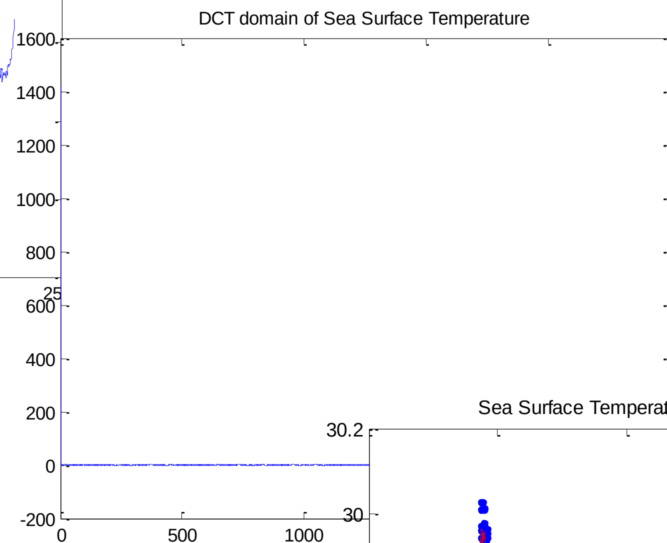
- N=512
- K=?
- M=6*K or M=4*K;

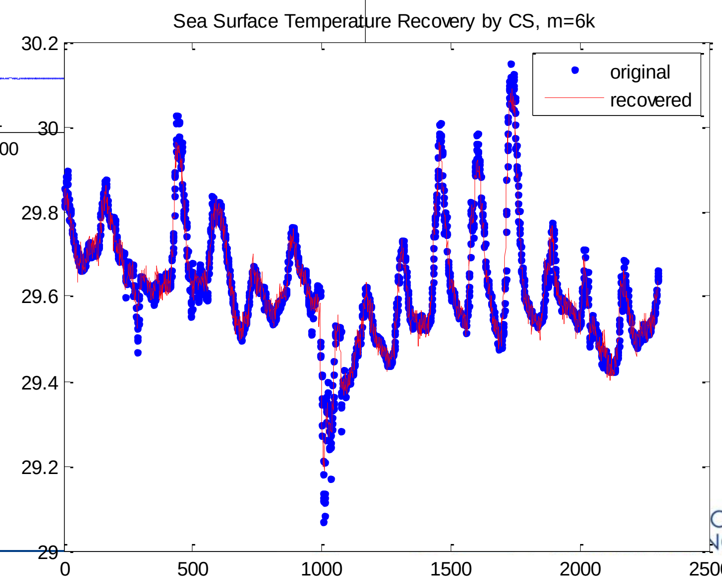




Sea surface temperature by CS



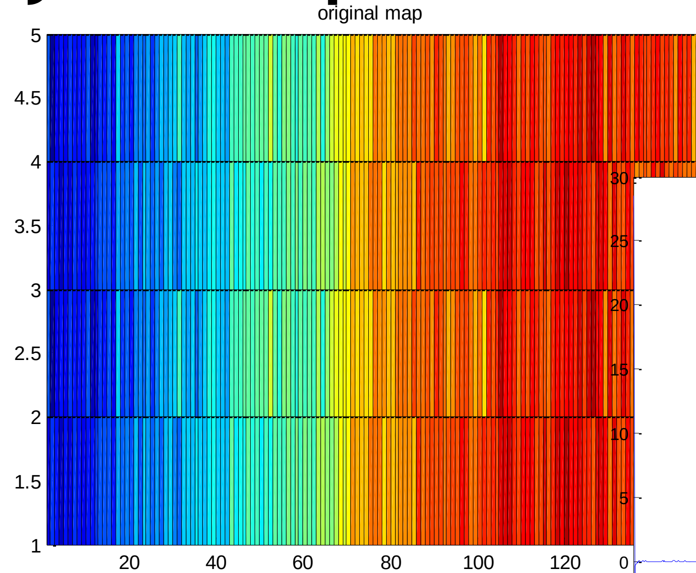


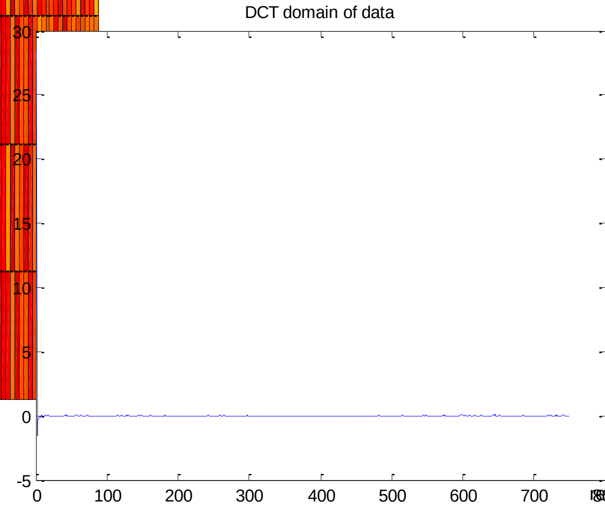


- N=2304
- M=696
- Normalized MSE=10^{-6}

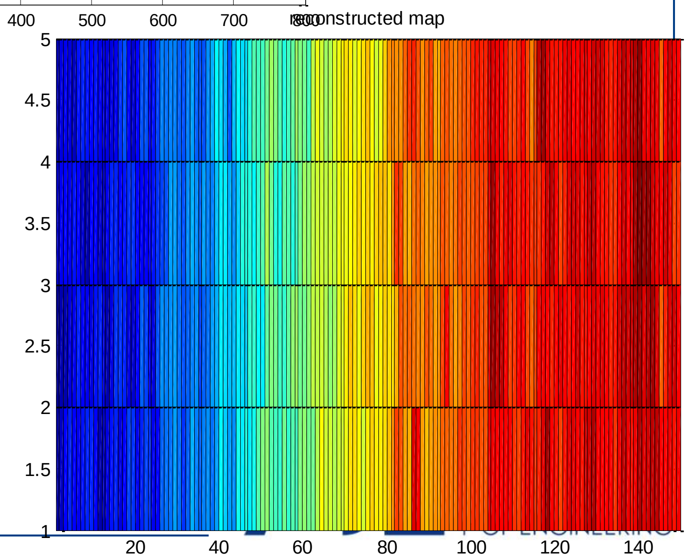
- http://tao.ndbc.noaa.gov/tao/data_download/process_results.php?type=sst&reso=hres&year1=2009&mon1=10&day1=19&year2=2009&mon2=11&day2=03&format=ascii&compression=none&stations=T8N180W

Apply CS to Spatial correlated sensor data





- N=750;
- Sample =75;
- Normalized MSE 0.000237



CS Hallmarks

- CS changes the rules of the data acquisition game
  - exploits a priori signal *sparsity* information
- **Universal**
  - same random projections / hardware can be used for *any* compressible signal class (*generic*)
- **Democratic**
  - each measurement carries the same amount of information
  - simple encoding
  - robust to measurement loss and quantization
- **Asymmetrical** (most processing at decoder)
- Random projections weakly **encrypted**

Toolbox for Sparse Recovery

- Sparselab

http://sparselab.stanford.edu/

- L1magic

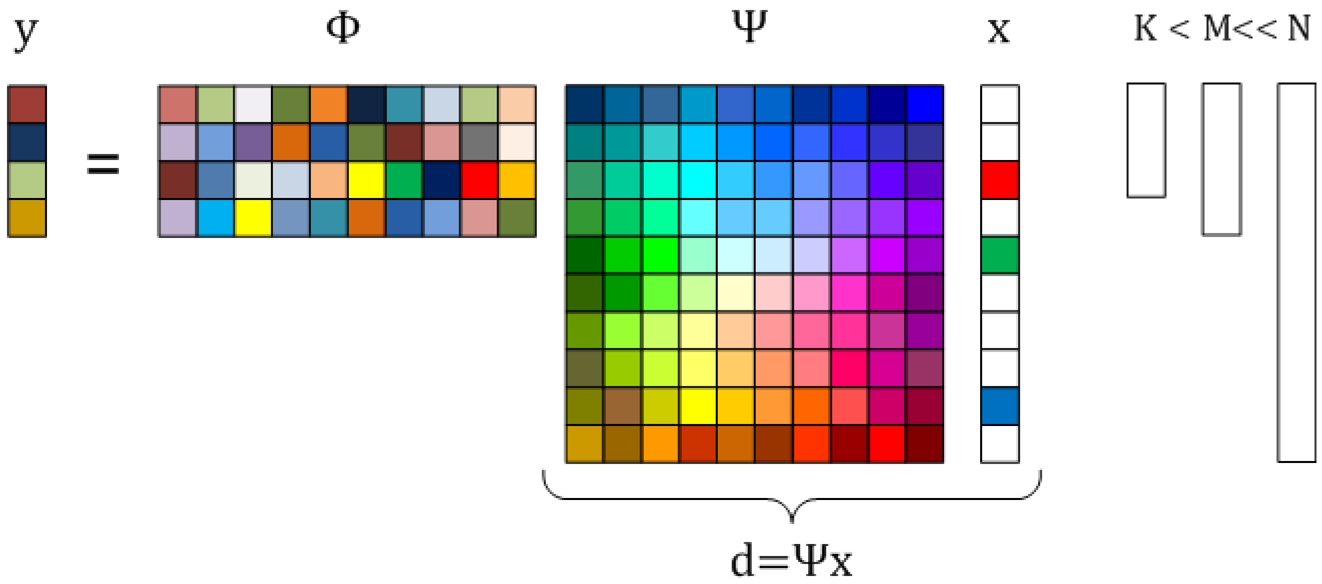
http://users.ece.gatech.edu/~justin/l1magic/

AARHUS SCHOOL OF ENGINEERING

Chong Luo et. Al, “Compressive Data Gathering for Large-Scale Wireless Sensor Networks”, MobiCom’09, September 20–25, 2009, Beijing, China.

AARHUS SCHOOL OF ENGINEERING

Compressive Sensing



- If an N-dimensional signal is K-sparse in a known domain Ψ, it can be recovered from M random measurements by:

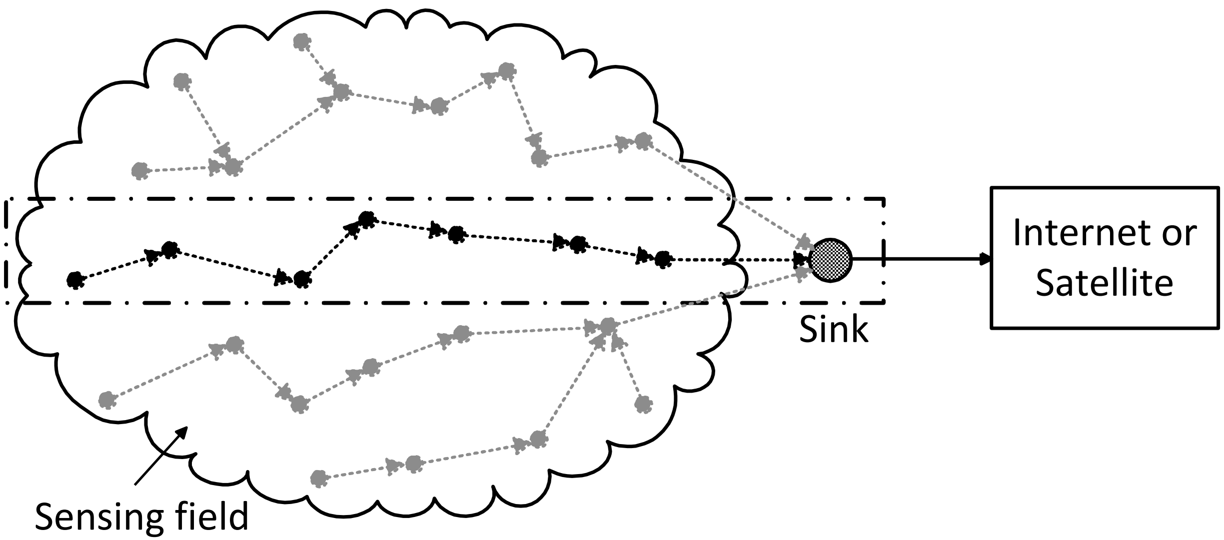
$$\min_{\mathbf{x}\in\mathbb{R}^N} \|\mathbf{x}\|_{l_1} \quad s.t. \quad \mathbf{y} = \Phi\mathbf{d},\ \mathbf{d} = \Psi\mathbf{x}$$

From Compressive Sensing to Compressive Data Gathering

- The asymmetrical property makes CS a perfect match for wireless sensor networks

Compressive Sensing	Compressive Data Gathering
Sample-then-compress ⇩ Sample-with-compression	Compress-then-transmit ⇩ Compress-with-transmission

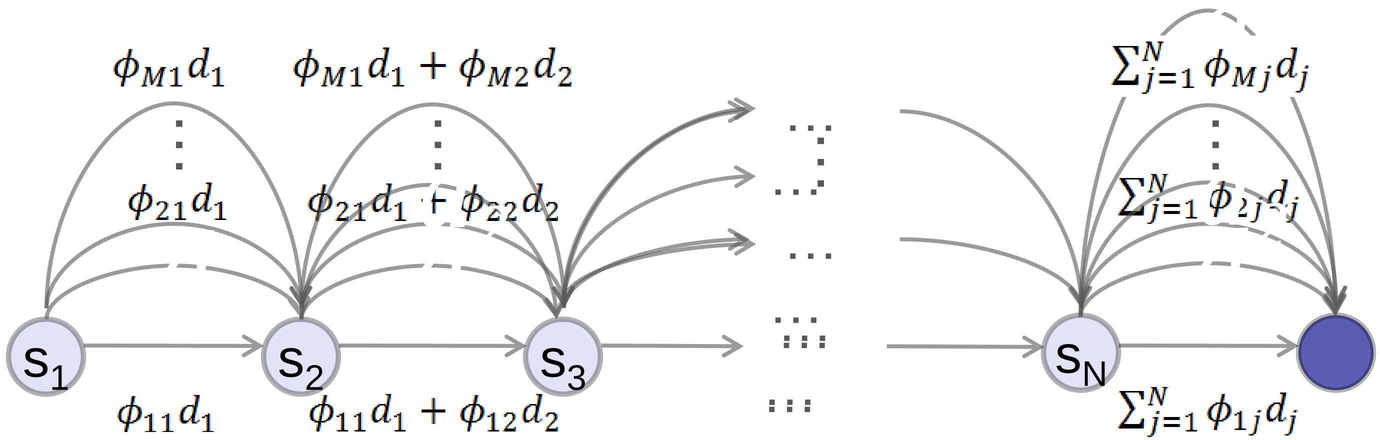
Data Gathering in WSNs



- Challenges
  - Global communication cost reduction
  - Energy consumption load balancing

Basic Idea

- A simple chain topology



	Global comm. cost	Bottleneck load
Baseline	N(N+1)/2	N
Proposed CDG	NM	M

Is Reconstruction Possible?

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} & \dots & \phi_{1N} \\ \phi_{21} & \phi_{22} & \dots & \phi_{2N} \\ \vdots & & \vdots & \\ \phi_{M1} & \phi_{M2} & \dots & \phi_{MN} \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{pmatrix}$$

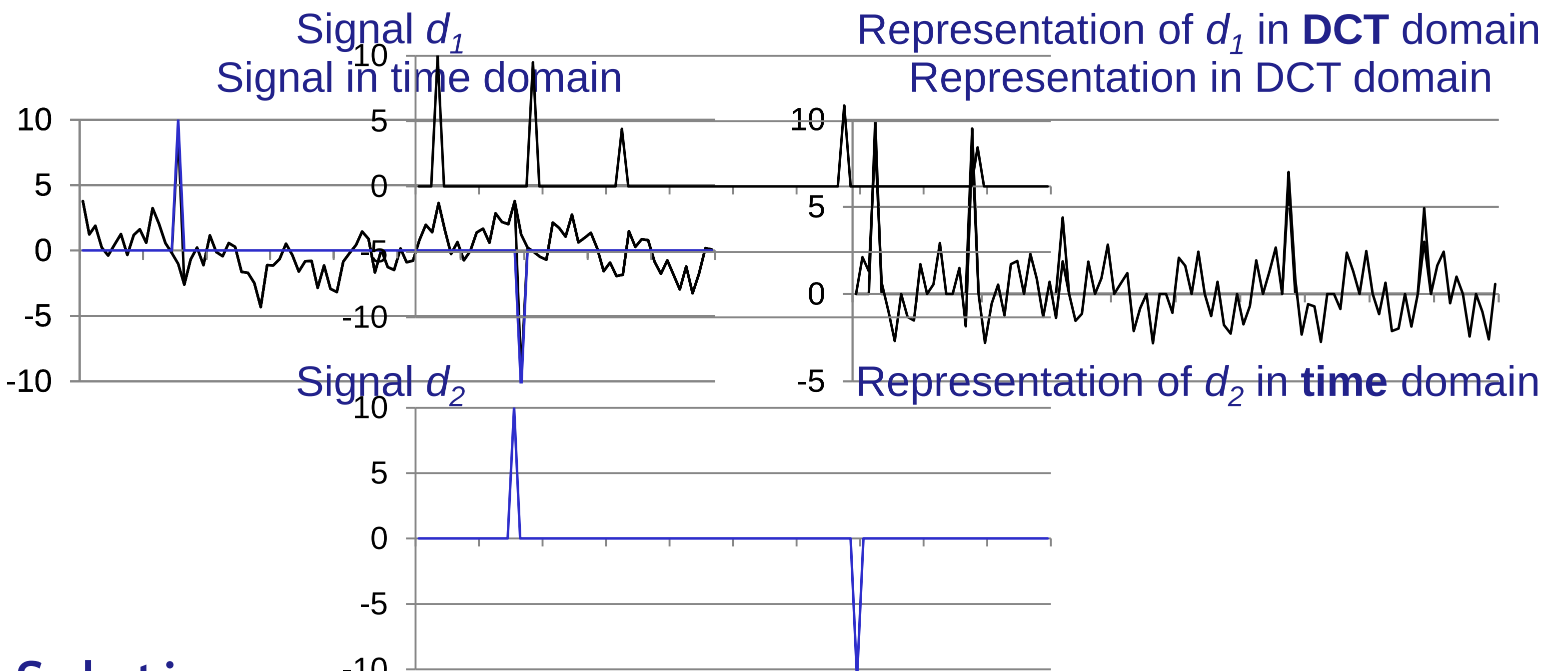
M<<N

- Facts
  - Sensor readings exhibit strong spatial correlations
- According to CS theory
  - Reconstruction can be achieved in a noisy setting by solving: $\min_{\mathbf{x}\in\mathbb{R}^N} \|\mathbf{x}\|_{l_1} \quad s.t. \quad \|\mathbf{y} - \Phi\mathbf{d}\|_{l_2} < \epsilon,\ \mathbf{d} = \Psi\mathbf{x}$

AARHUS SCHOOL OF ENGINEERING

Practical Problem 1

- Abnormal readings compromise data sparsity



- Solution:

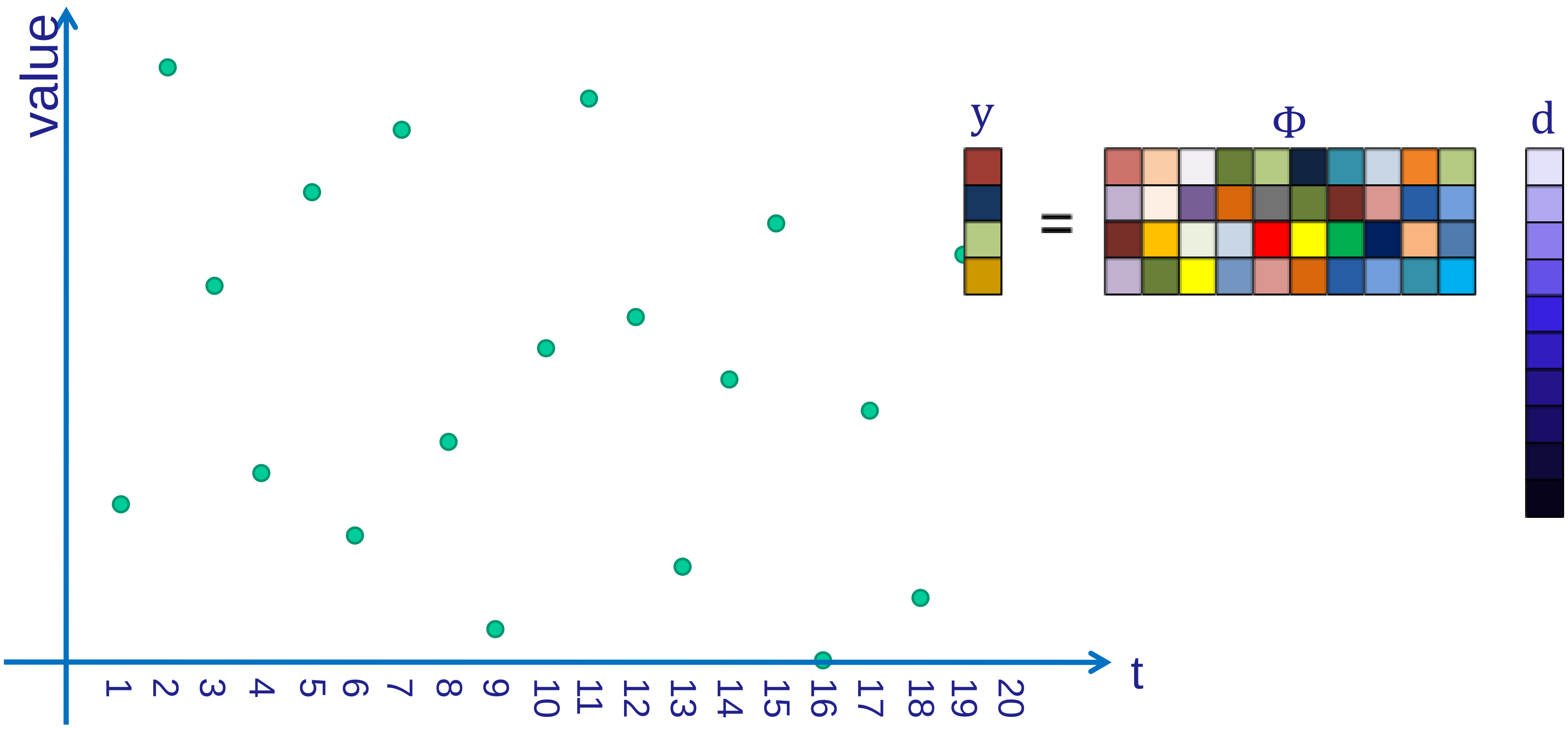
$$d = d_1 + d_2 = \Psi x_1 + I x_2 = [\Psi \; I] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

7-sparse

Overcomplete basis

Practical Problem 2

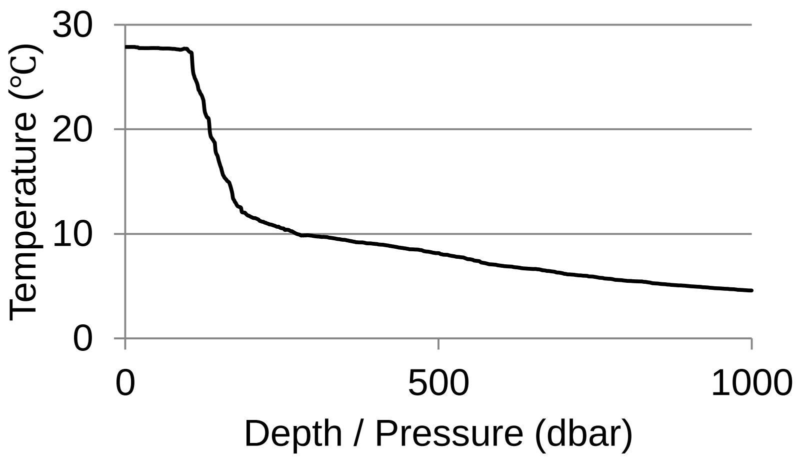
- If a signal is not sparse in any intuitively known domain



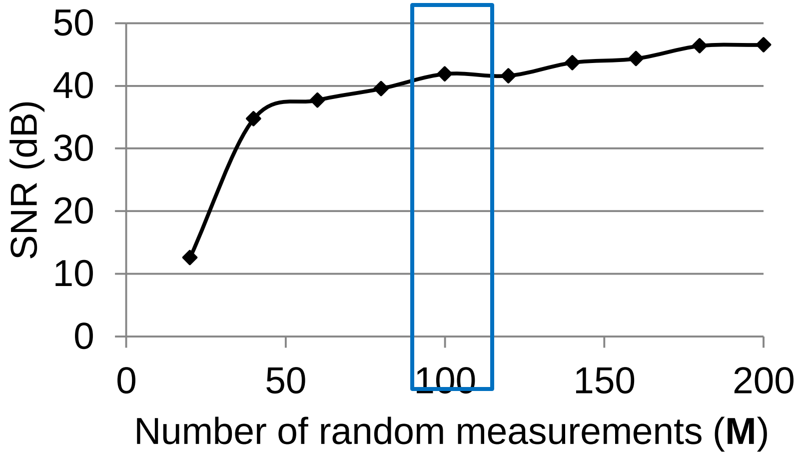
AARHUS SCHOOL OF ENGINEERING

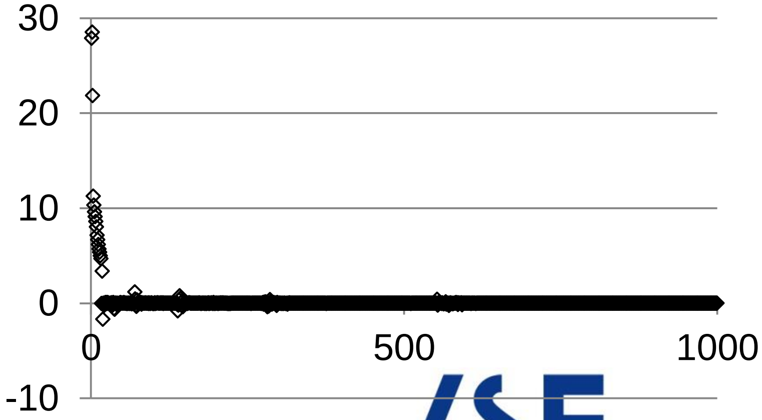
Example 1

- CTD (Conductivity, temperature and depth) data from NOAA (National Oceanic and Atmospheric Adminstration)
  - N=1000, K≈40



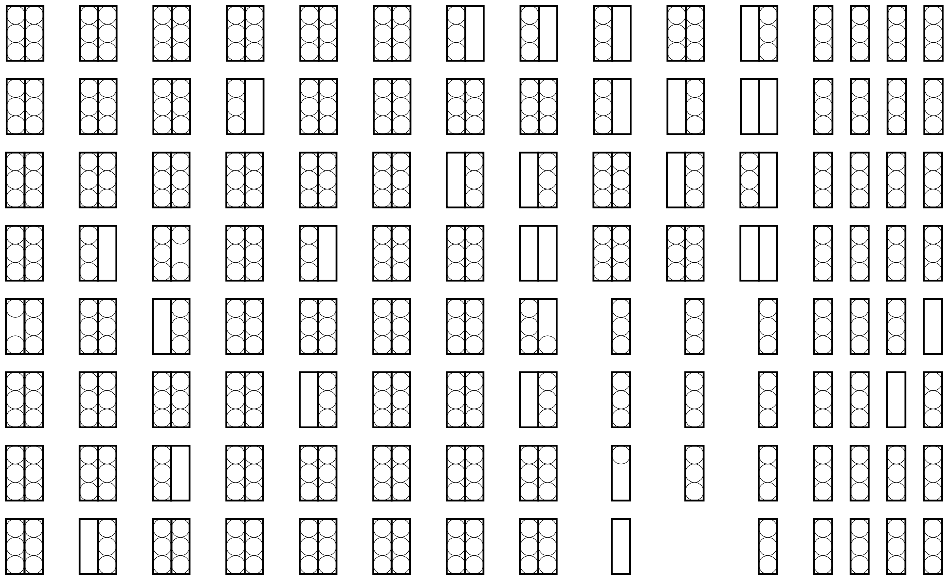
M=100	
Recon. Precision	99.2%
Comm. Reduction	5
Capacity Gain	10





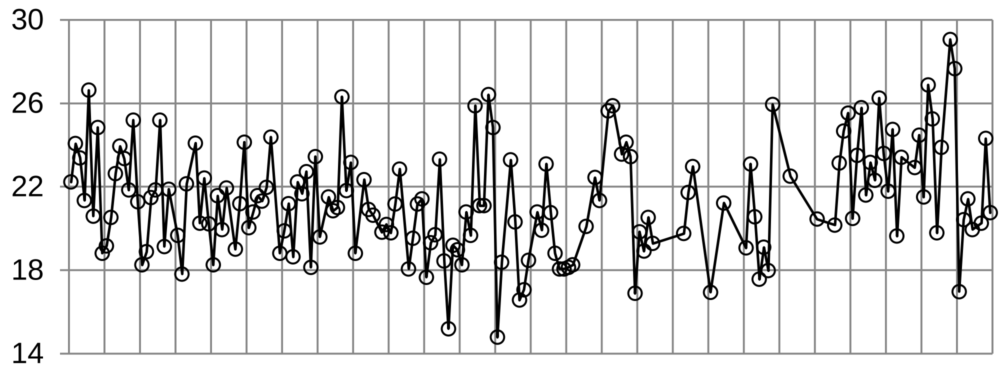
Example 2

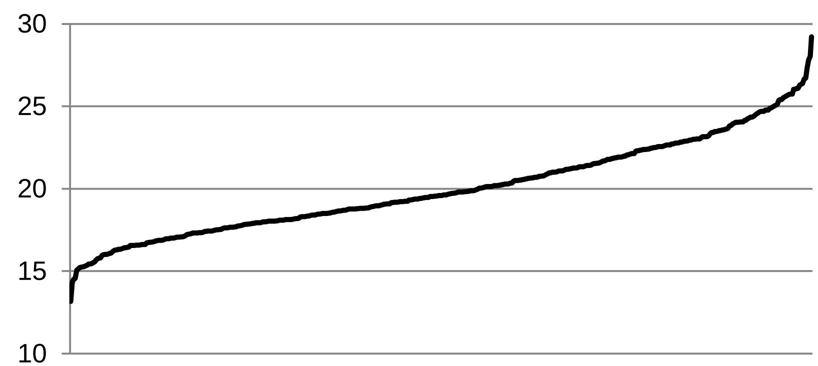
- Temperature data from data center
  - 498 temperature sensors



Example 2 cont…

- sensor readings exhibit little spatial correlations

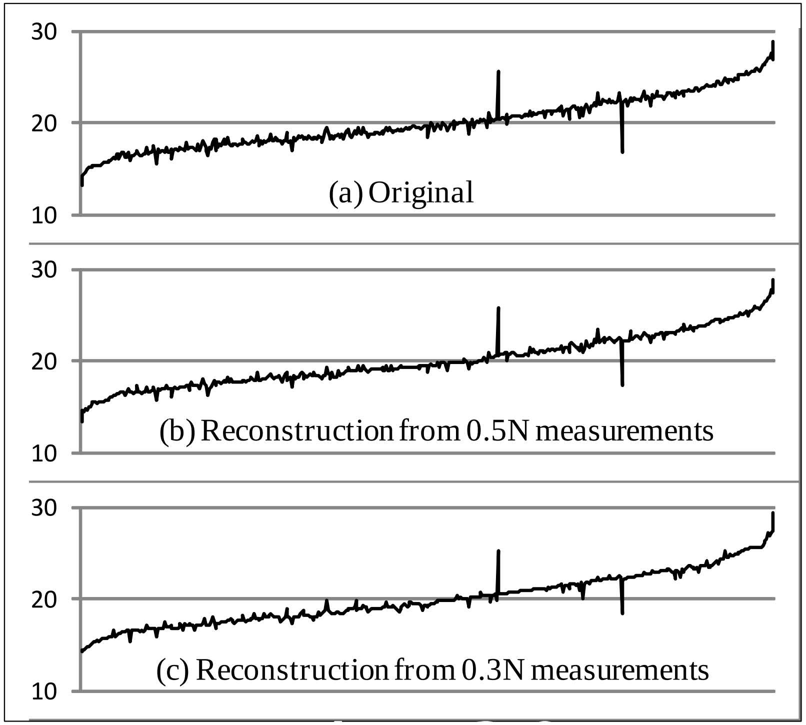




- Reorder sensors according to their readings at t_0

Example 2 cont…

- Utilizing Temporal Correlation
- Sensor readings at $t_0 + \Delta t$ are sparse as well
  - Temperatures do not change violently with time



AARHUS SCHOOL OF ENGINEERING

Conclusion

- Compressive Sensing is an emerging field which may bring fundamental changes to networking and data communications research
- The Contributions of Compressive Data Gathering (CDG)
  - The first complete design to apply CS theory to sensor data gathering
  - CDG exploits “universal sparsity”
  - CDG improves network capacity
- Note: CDG is not suitable for small scale sensor networks when the signal sparsity may not be prominent enough and the potential capacity gain may be too small.