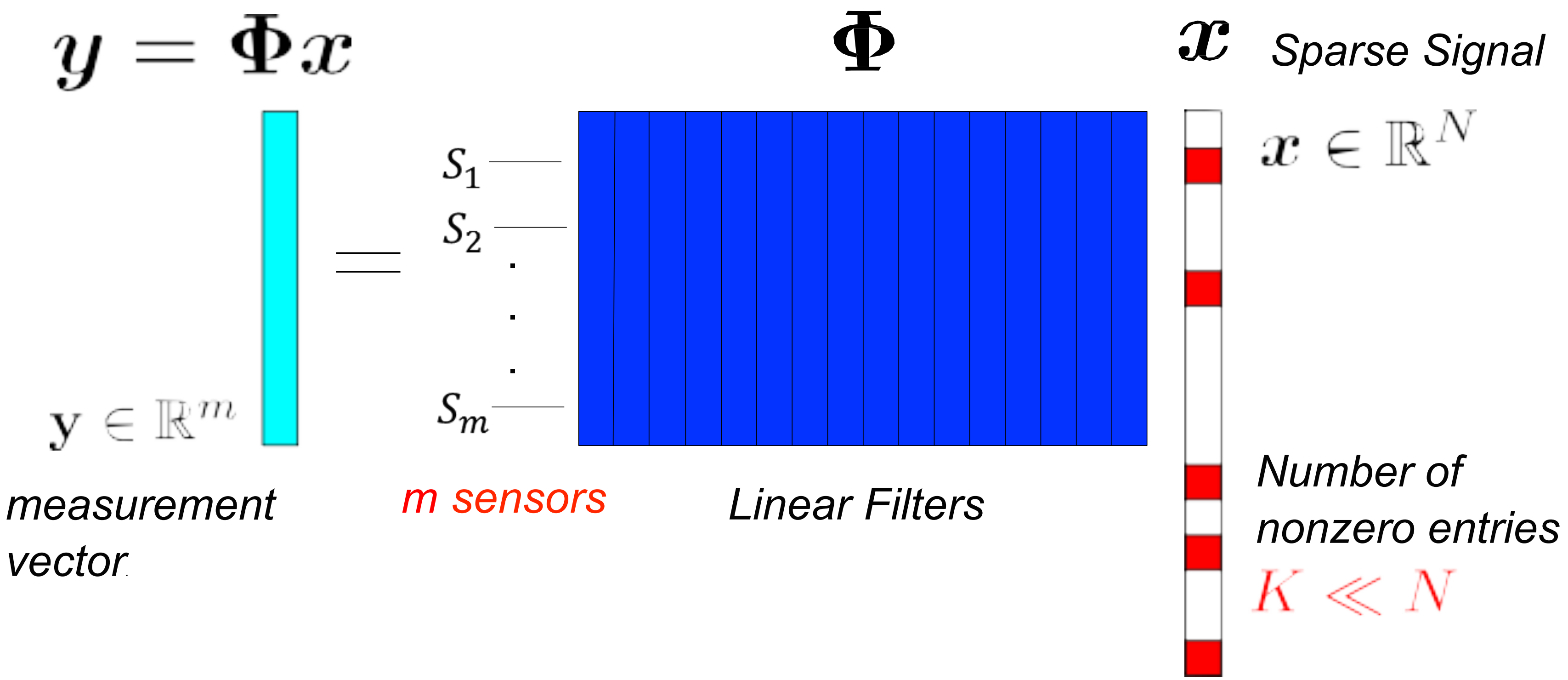
Imperial College London

From Compressed Sensing to Distributed Signal (Data) Processing

Kin K. Leung
Head of Communications and Signal Processing
Electrical & Electronic Engineering, and Computing Departments
Imperial College, London
August 2015

Acknowledgments: Wei Dai, Ervipidis Karseras, Sepideh Nazemi (Imperial College), Ananthram Swami (U.S. Army Research Lab) and Stuart Farquhar (U.K. Dstl)

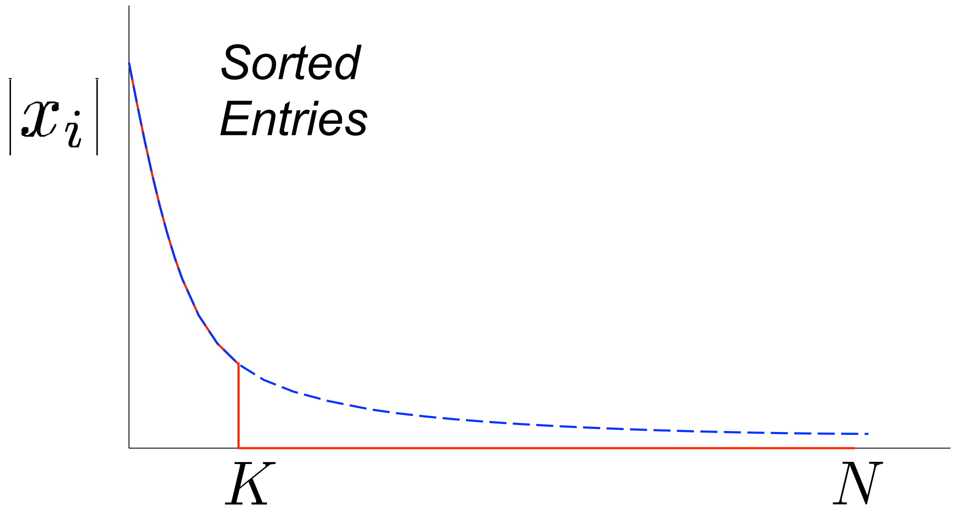
Compressed / Sparse Sensing



$$m = O(K \log \frac{N}{K}) \ll N$$ → Hugh savings

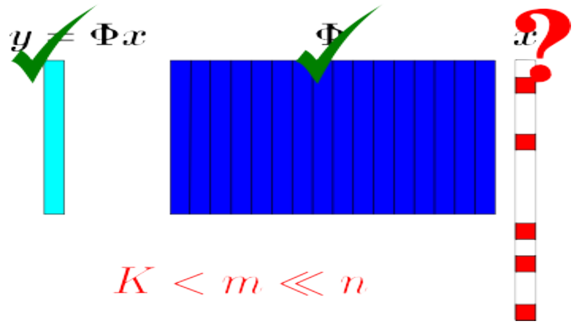
Compressible Signals

Compressible signals: approximated by K -sparse signals

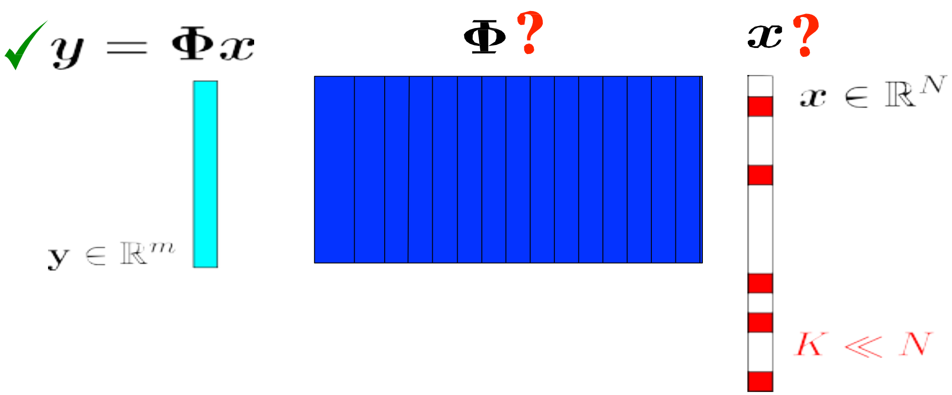


Compressed Sensing (CS) Problem

Problem 1: Recover original signal x, given measurement y and sensing matrix ∅.



Problem 2: Dictionary learning – Identify the sensing matrix ∅ and recover original signal x, given measurement y. (Much tougher!)



CS Reconstruction

Problem: Recover original signal x, given measurement y and sensing matrix $\emptyset$

l_0 – Mininization for CS reconstruction

$$\min_x \ ||x||_0 \quad \text{subject to} \quad ||y - \emptyset x||_2 \leq \varepsilon$$

where $||x||_0$ counts the number of nonzero elements of x

- Solution represents the sparest signal (i.e., with the minimum number of nonzero entries
- Huge complexity: Exhaustive search with complexity C_K^N

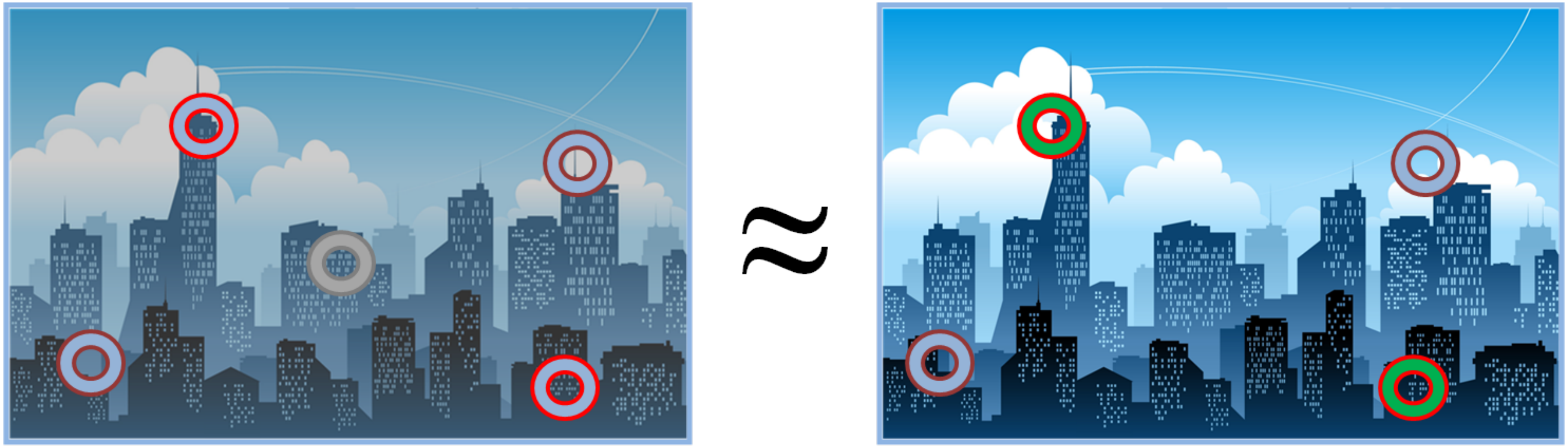
Greedy algorithm

$$\min_x ||y - \emptyset x||_2 \quad \text{subject to} \quad ||x||_0 \leq K$$

- Under the K-restricted isometry property (RIP)
  - ➢ Noiseless case: exact reconstruction
  - ➢ Noisy case: bounded reconstruction distortion

Sensor networks: Sparse & temporal correlated samples

- Sensed signal is sparse in some domain
- Exact reconstruction is possible at sub-Nyquist rates
- Reduced data volume, while retaining information content



Time $t-1$ $\approx$ Time t

- Consecutive time samples demonstrate high temporal correlation
- Desirable to exploit temporal information for improvements in performance and computational complexity

Observations and question

- Observations
  - Traditionally, Kalman filtering handles dynamic signal recovery well
  - Kalman filtering does not deal with sparse signals well
  - Established compressed sensing (CS) techniques do not handle dynamic, temporally correlated signals, as efficiently as Kalman
  - Bayesian CS handles signal statistics, but not well on temporal characteristics

- Question
  - Can we do better than performing sparse signal reconstruction independently for each frame?
  - Answer: Yes

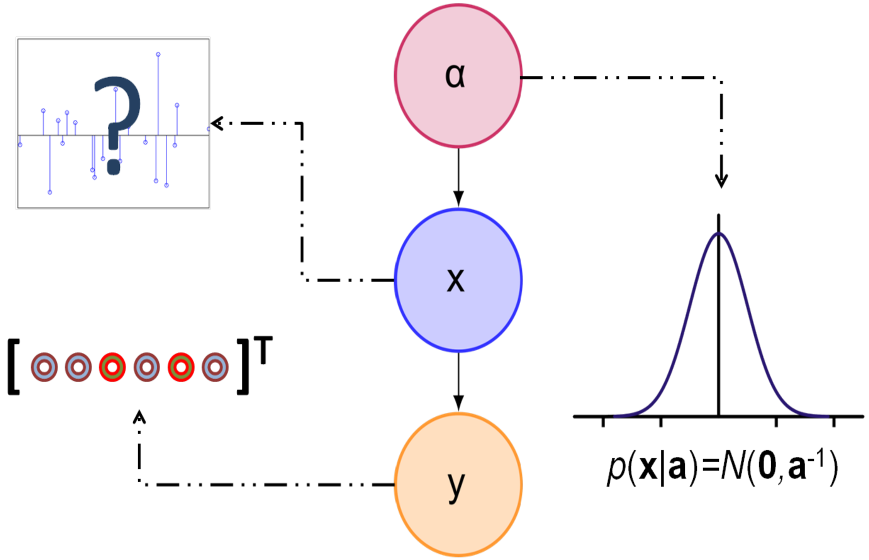
Sparsity from a Bayesian Standpoint

Sparse Bayesian learning allows estimations of signal statistics.

Sparsity of each component x_i is controlled by its variance:

$$p(x_i | \alpha_i) = N(0, \alpha_i).$$

When $\alpha_i = 0$, it is *a-posteriori* certain that $x_i = 0$.



Hierarchical Bayesian Model

Tracking Dynamic Sparse Signals

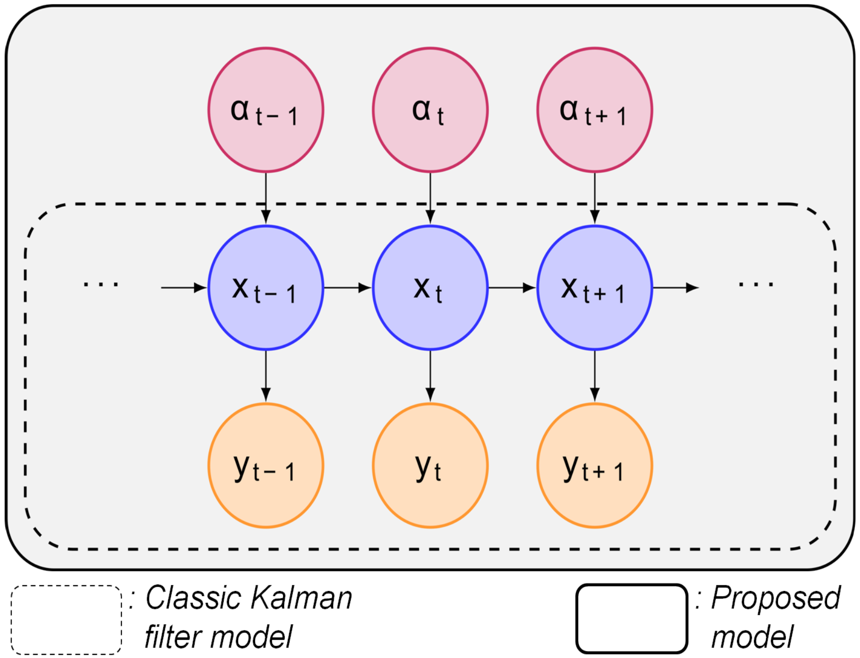
Signals $\boldsymbol{x}_t$ *are sparse in the same domain:*

$$\boldsymbol{x}_t = \boldsymbol{x}_{t-1} + \boldsymbol{q}_t$$

$$\boldsymbol{y}_t = \boldsymbol{\Phi}_t \boldsymbol{x}_t + \boldsymbol{n}_t$$

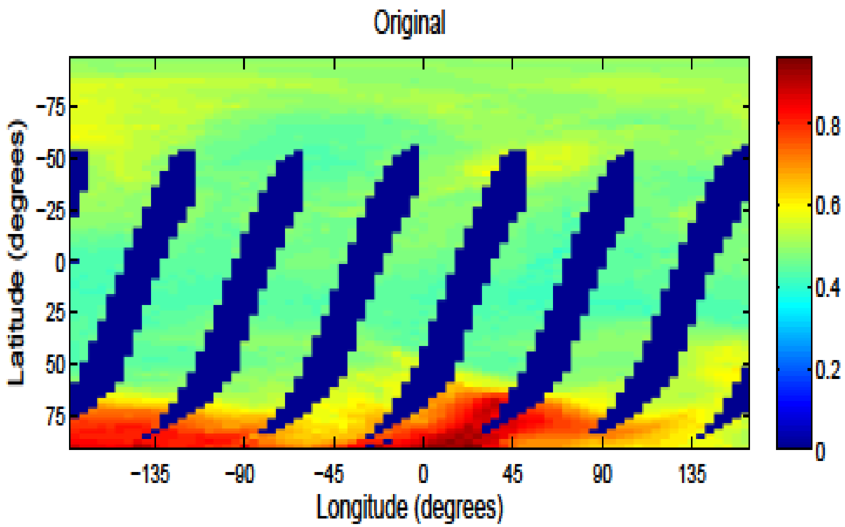
$\boldsymbol{x}_t$ *is sparse* $\rightarrow$ $\boldsymbol{q}_t$ *is sparse:*

$$\boldsymbol{q}_t \sim N(\boldsymbol{0}, \boldsymbol{A}_t)$$
$$\boldsymbol{A}_t = diag(\alpha_1, \cdots, \alpha_N)$$

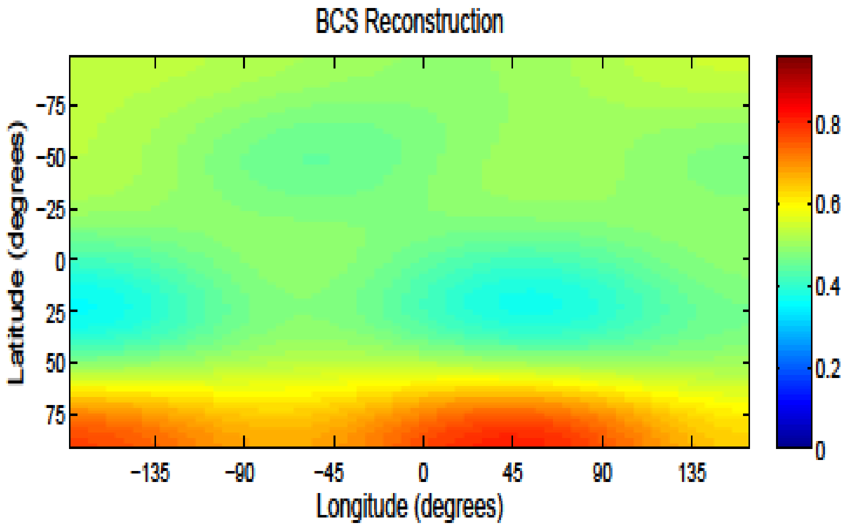


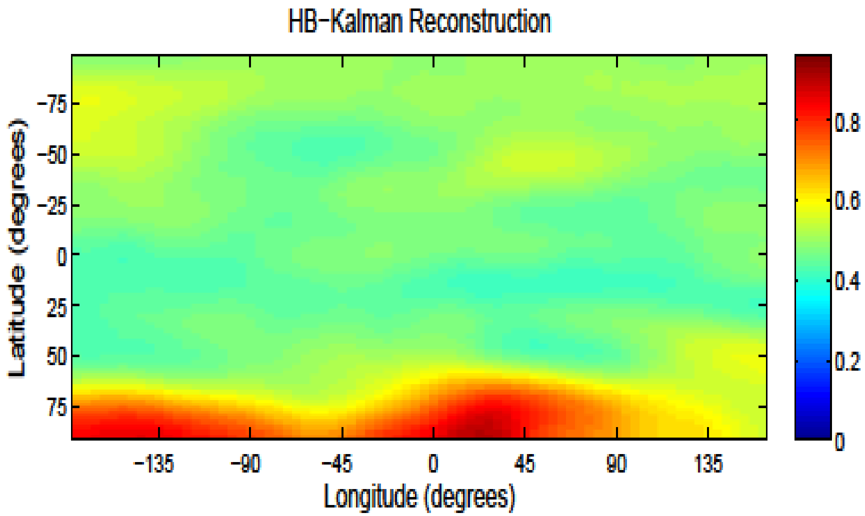
Imperial College London

NASA ozone measurements and reconstruction

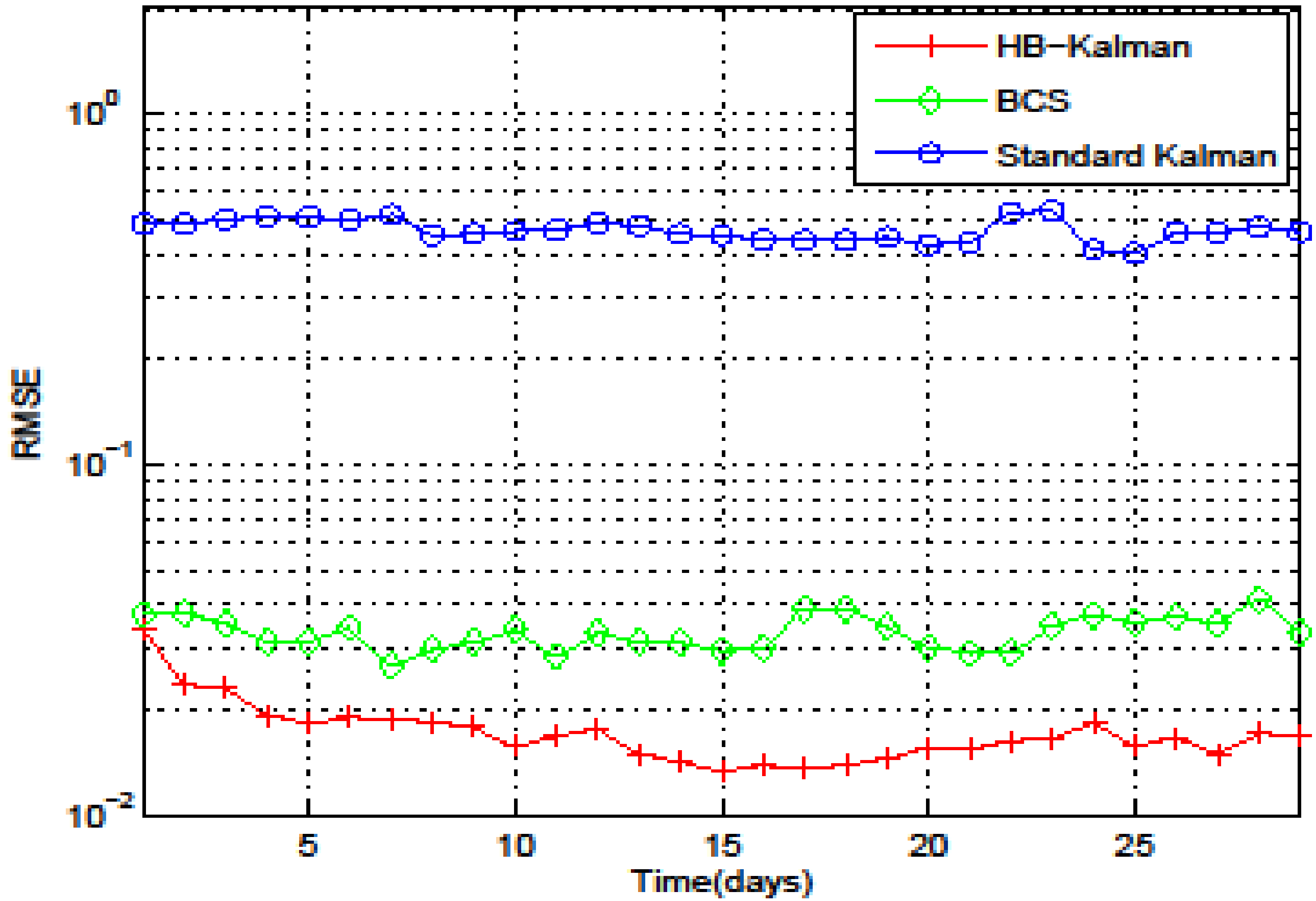


- Global ozone distribution
- Blue strips represent missing data
- Signal is sparse in DCT domain
- Hybrid-Bayesian Kalman approach yields lower reconstruction errors than Bayesian CS





Reconstruction error of ozone measurements

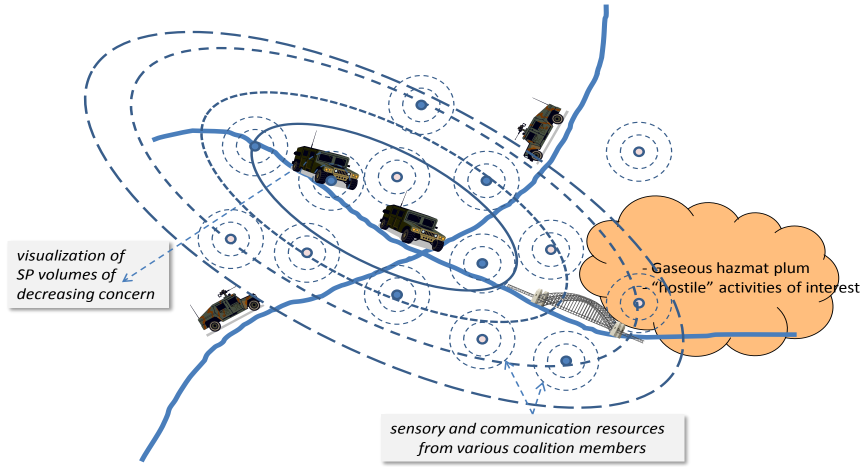


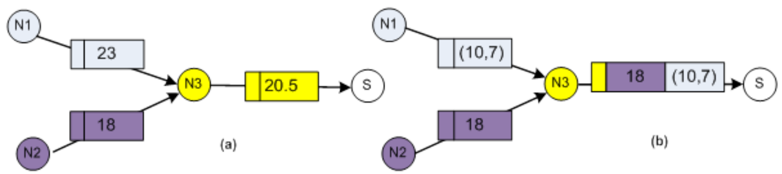
Imperial College London

From Centralized Compressed Sensing to Distributed Signal (Data) Processing

Distributed Signal Processing (In-Network Data Processing)

- Dynamic, Distributed Signal Processing
  - Multiple, distributed signal/data sources
  - Huge data volume!
  - Complex temporal-spatial correlations
  - Limited communication or computing resources
  - Dynamic info requirements (e.g. user location)
  - Applications e.g., situation awareness





- Solution Techniques
  - Process signal/data while being transferred hop-by-hop toward the user destination (In-Network Data Processing, INDP)
  - Lossy vs. lossless processing
  - Optimize use of bandwidth and computing resources while providing satisfactory quality of information (QoI)

Current Network Model

- Operating scenario
  - A user sends a query to request for information
  - An aggregation tree is formed to transfer and process requested information

- Parameters of concern
  - Energy consumption at each node: receiving, computation and transmission
  - Data (fusion, compression, aggregation) reduction rate: $0 \leq \delta_i \leq 1$ for each node *i*

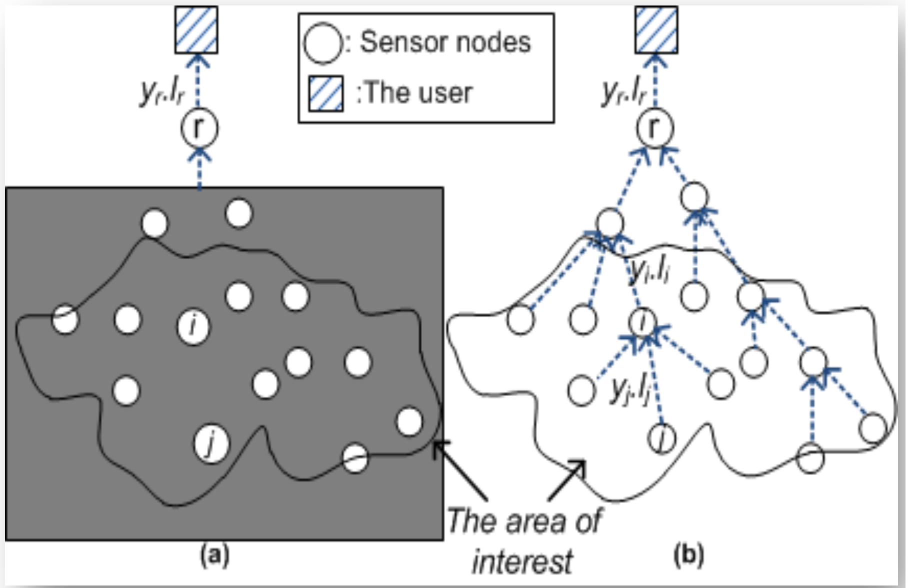
- Distributed approach for data aggregation
  - To achieve the optimal trade-off between energy consumption and QoI (e.g., amount of data received at the end user)

Problem Formulation: Distributed Signal Processing

- The global optimization (GO) problem

$$min_{\{\delta_i\}} \sum_{i=1}^{r} P_i(\delta_i, y_i)$$

$$s.t. \quad y_r\delta_r \geq \gamma$$



where $P_i(..)$ = energy consumption
y_i= amount of data input
δ_i= reduction rate at node i

- Constraint represents QoI requirements
- Possible to extend formulation for other settings/applications
- Unfortunately, it is an NP-hard problem!

Imperial College London

Local Constrained Optimization (LCO) Formulation

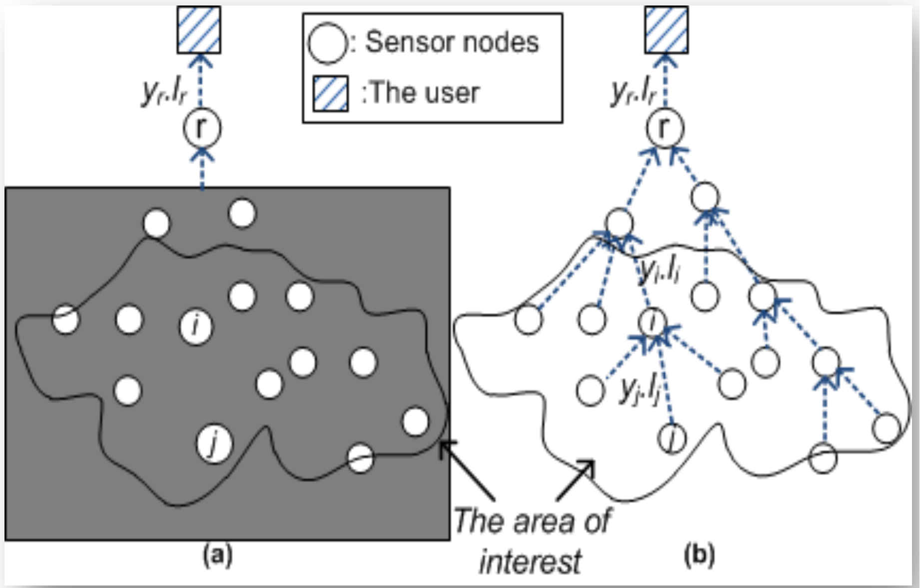
- The local constrained optimization (LCO) problem

$$min_{\{\delta_i\}} \sum_{i=1}^{r} P_i(\delta_i, y_i)$$
$$s.t. \quad y_i \delta_i \geq \gamma \ for \ \forall i$$

where $P_i(..)$ = power consumption

y_i= amount of data input

δ_i= reduction rate at node i



- Additional constraints impose QoI requirement at each node

Imperial College London

Distributed solution to the LCO problem

- Assumptions
  - Energy consumption as separable functions: $P_i(\delta_i, y_i) = f_i(\delta_i) g_i(y_i)$
  - Communication energy consumptions for receiving and transmitting are proportional to the amount of data involved

- Theorem: Under the assumptions, the LCO problem is equivalent to a distributed one as follows:

$$min_{\{\delta_i\}} \sum_{i=1}^{r} P_i(\delta_i, y_i) \quad \Longleftrightarrow \quad \sum_{i=1}^{r} min_{\{\delta_i\}} P_i(\delta_i, y_i)$$

$$s.t. \quad y_i \delta_i \geq \gamma \ for\ \forall i \qquad\qquad s.t. \quad y_i \delta_i \geq \gamma \ for\ \forall i$$

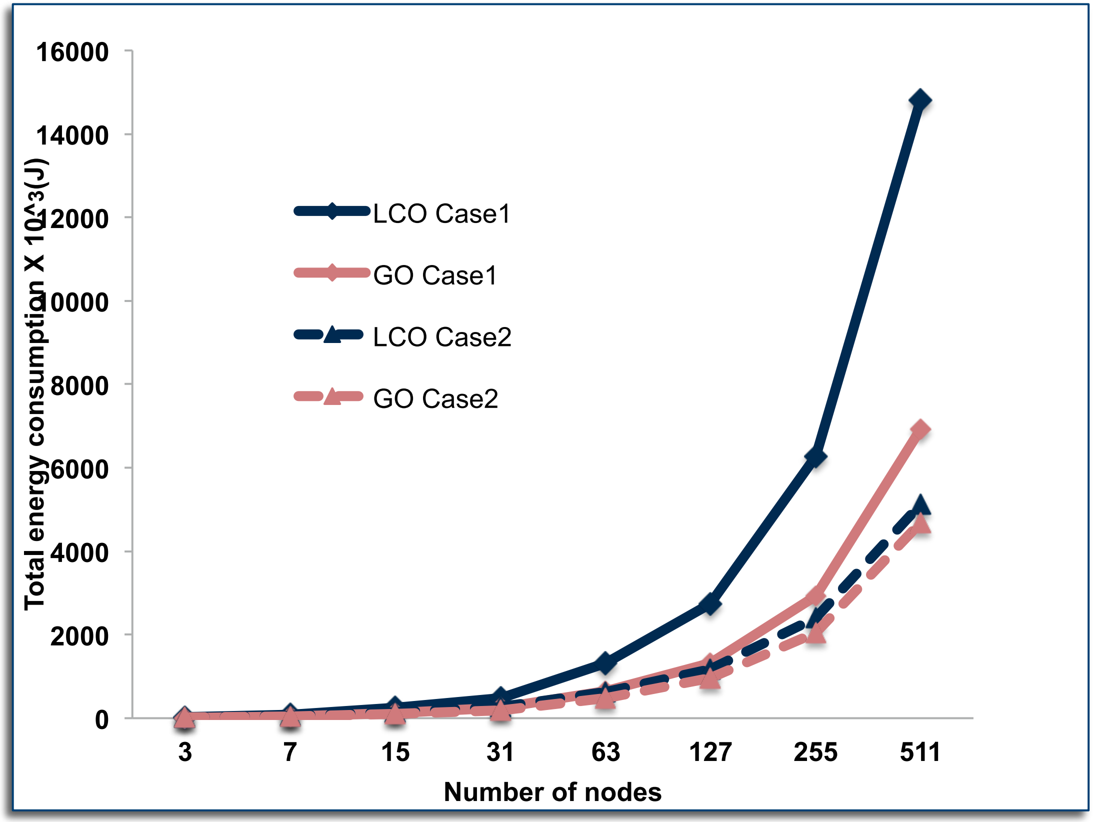
Each node optimizes on its own.
Fully distributed solution!

Imperial College London

Numerical results

- Cases with balanced binary aggregation trees & computable optimal solutions

Cases	Parameter settings
Case1	$e_T = e_R = e_C = 0.00024$ $\gamma = 5$
Case2	$e_T = e_R = 0.00024$ $e_C = 0.00012$ $\gamma = 5$



LCO closely approximates GO when

- Communication costs higher than computation
- Number of nodes increases

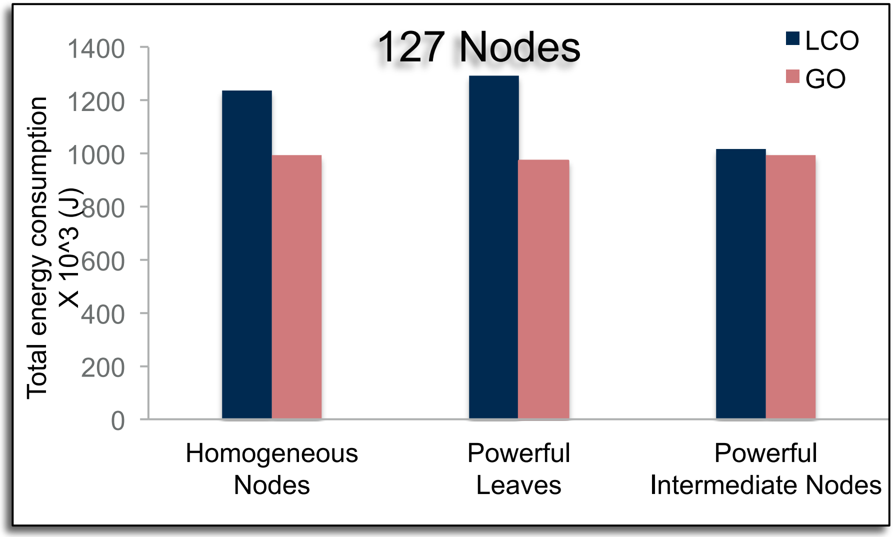
Imperial College London

Numerical results (continued)

Balanced aggregation tree with 127 node:

- Homogeneous Nodes (HN)
- Powerful Leaves (PL)
- Powerful Intermediate Nodes (PI)

Cases	Parameter settings
HN	$e_{T_i} = e_{R_i} = 0.00024$ for all i e_{C_i}=0.00012 for all i $\gamma = 5$
PL	$e_{T_i} = e_{R_i} = 0.00024$ for all i e_{C_i}=0.00012 for $i \in \{intermediate\ nodes\}$ e_{C_i}=0.00006 for $i \in \{leaf\ nodes\}$ $\gamma = 5$
PI	$e_{T_i} = e_{R_i} = 0.00024$ for all i e_{C_i}=0.00006 for $i \in \{intermediate\ nodes\}$ e_{C_i}=0.00012 for $i \in \{leaf\ nodes\}$ $\gamma = 5$



LCO closely approximates GO when intermediate nodes have powerful computation capability

Imperial College London

Concluding remarks and future work

- Concluding remarks
  - Signals are often sparse in some domain, thus compress sensing (CS) techniques are applicable
  - CS techniques have been developed to treat sparse signals with time dynamics
  - Distributed signal processing (DSP) is useful, but open issues exist
  - DSP (e.g., sensor networks) has to be considered with communication constraints for optimal performance
  - Globally optimal DSP is hard to achieve, but suboptimal distributed solutions may be possible
- Future work
  - Generalize the conditions under which the local-constrained optimization problem can lead to fully distributed solutions
  - How can we perform the CS in distributed ways?
  - Incorporate other aspects of signal processing (e.g., image, detection) into the DSP framework

Imperial College London

Acknowledgments and references

- Acknowledgments
  - Wei Dai, Ervipidis Karseras, Sepideh Nazemi (Imperial College), Ananthram Swami (U.S. Army Research Lab) and Stuart Farquhar (U.K. Dstl)
  - Research funding: EU SmartEN Project, U.K. UDRC in Signal Processing, U.S./U.K. ITA Project

- Publications

  - E. Karseras, K.K. Leung and W. Dai, "Tracking dynamic sparse signals using hierarchical Bayesian Kalman filters," presented at the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Vancouver, Canada, May 2013.
  - E. Karseras, K.K. Leung and W. Dai, “Tracking Dynamic Sparse Signals with Kalman Filters: Framework and Improved Inference,” presented at the International Conference on Sampling Theory and Applications (SampTA), Bremen, Germany, July 2013.
  - S. Nazemi, K.K. Leung and A. Swami, “A Distributed, Energy-Efficient and QoI-Aware Framework for In-Network Processing,” presented at the IEEE PIMRC 2014.