# Signal sparsity estimation from compressive noisy projections via sparsified random matrices

41st IEEE ICASSP
20-25 March 2016, Shanghai, China


Chiara Ravazzi

in collaboration with

Sophie M. Fosson    Tiziano Bianchi    Enrico Magli

*Politecnico di Torino - DET*

POLITECNICO DI TORINO
Department of Electronics and Telecommunications


erc

CRISP
Towards Compressive Information Processing Systems

# Motivation and background

POLITECNICO DI TORINO
Department of Electronics and Telecommunications

## Compressed Sensing (CS)

CS system:

- technique for compressed acquisition

$$y = Ax + \eta$$

  - $x \in \mathbb{R}^n \rightsquigarrow$ sparse signal (with at most $k \ll n$ nonzero entries)
  - $A \in \mathbb{R}^{m \times n} \rightsquigarrow$ sensing matrix ($m < n$)
  - $\eta \in \mathbb{R}^m$ additive Gaussian noise $\mathrm{N}(0, \sigma^2)$

- reconstruction from few linear measurements (exploiting signal's sparsity)

Careful design: knowledge of $k$ relevant

- sensing matrix: RIP-$k$, NSP-$k$, coherence-$k$ condition
- number of measurements: $m = O(k \log n/k)$
- recovery algorithms: tuning parameters depends on $k$ (OMP, CoSaMP, Lasso, IRLS, ... )

# Motivation and background

POLITECNICO DI TORINO
Department of Electronics and Telecommunications

## Sparsity estimation

Sparsity: major gap between theory and practice

- CS theory assumption: knowledge of sparsity degree $k$
- In practice: not always true
  - time-varying sparsity (spectrum sensing)
  - spatially-varying sparsity (block-based acquisition of images)
  - is a signal actually sparse in some basis?

Related literature:

- streaming measurements in CS [Romberg&al.2008]
- sparsity estimation via recovery [Wang&al.2012]
- sparsity estimation from measurements [Lopes2013]

**Our contribution: estimate sparsity degree from measurements via sparse random matrices**

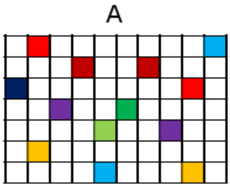
POLITECNICO DI TORINO
Department of Electronics and Telecommunications

## Sparse sensing matrices

CS system: $y = Ax + \eta$

- $x \in \mathbb{R}^n \rightsquigarrow k$-sparse signal (with at most $k$ nonzero entries)
- $\eta$ additive Gaussian noise $\mathrm{N}(0, \sigma^2)$
- $A \rightsquigarrow \gamma$-sparsified random matrix

$$a_{ij} \sim \begin{cases} 0 & \text{with prob. } 1-\gamma \\ \mathcal{N}\left(0, \frac{1}{\gamma}\right) & \text{with prob. } \gamma. \end{cases}$$




Sparse ($\gamma = \Theta(n^{-1})$) vs dense ($\gamma = \Theta(1)$) matrices:

+ low computational complexity and memory requirements
+ enable reconstruction with a slight performance degradation

$$\psi(k) = \gamma k = \tau \Longrightarrow m \geq O\left(\frac{k \log(n/k)}{\tau \log(1+x_{\min}^2 k/\tau)}\right)$$

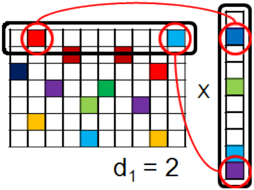
# Sparsity estimation

POLITECNICO DI TORINO
Department of Electronics and Telecommunications

## Noiseless Setting

Maximum Likelihood estimation:

- $y = Ax$, $x$ is $k$-sparse
- $y \sim \mathrm{Ber}(p_k)$, $p_k = 1-(1-\gamma)^k$
- estimate $k$ from $\|y\|_0$ (number of nonzeros in $y$)




$$\implies \hat{k}_{ML} = \left\lfloor \frac{\log\left(1-\frac{\|y\|_0}{m}\right)}{\log(1-\gamma)} \right\rfloor$$

**Thm 1: strong consistency**

- for fixed $k$: $|\widehat{k} - k|/k \leq O(\sqrt{\log m/m})$ w.p.1.
- for large $k, m, n$: many regimes of $\psi(k) = \gamma k$ for strong consistency w.p.1.

# Sparsity estimation

## Noisy Setting

Sparsity estimation: high computational complexity

- $y = Ax + \eta$, $x$ is $k$-sparse, $\eta \sim \mathcal{N}(0, \sigma^2)$
- $y \sim f_k$ is a mixture of up to $2^k$ Gaussians: intractable

Our approach:

- approximate $f_k$ as 2-component Gaussian mixture (2-GMM)

$$f_k^{2-GMM}(y) = (1 - p_k)\phi\left(y|\sigma^2\right) + p_k\phi\left(y|\sigma^2 + \frac{\|x\|_2^2}{p_k}\right)$$

- estimate 2-GMM parameters via Expectation-Maximization
- compute

$$\hat{k}_{EM} = \left\lfloor \frac{\log(1 - p_k)}{\log(1 - \gamma)} \right\rfloor$$

# Sparsity estimation

POLITECNICO DI TORINO
Department of Electronics and Telecommunications

## Noisy Setting

**Thm 2: 2-GMM approximation error**

2-GMM approximates real distribution for large $k$:
$x_{\min}^2 k = \Theta(1) \implies$

$$\|f_k - f_k^{2-GMM}\|_{\mathrm{Kol}} \leq C(\psi(k) + \psi(k)^2), C \in \mathbb{R}, \psi(k) = \gamma k$$

Summary: Sparse matrices are good

- for recovery [Wang&Wainwright2010]
- for sparsity estimation in noiseless setting
- for sparsity estimation in noisy setting (2-GMM approximation error is bounded)

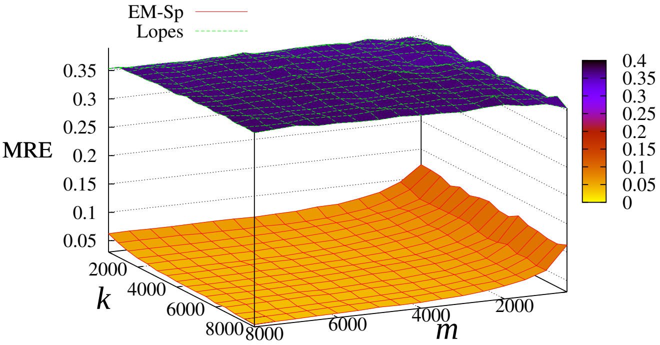
# Numerical experiments

POLITECNICO DI TORINO
Department of Electronics and Telecommunications

## Synthetic signals

Effect of noise: $\psi(k) = \Theta(1)$

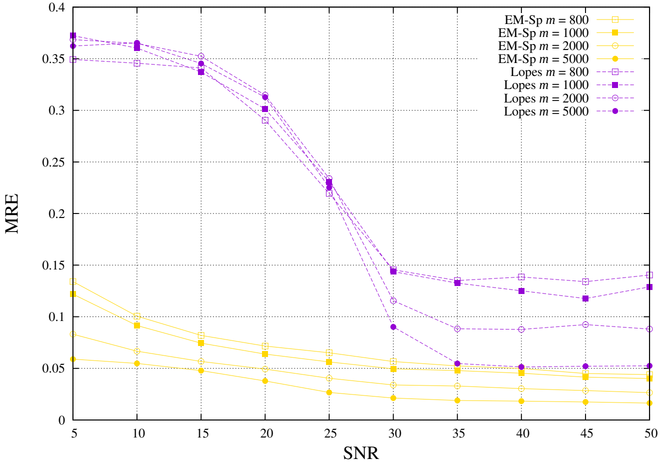
- minimal $\mathrm{SNR} = x_{\min}^2 k/\sigma^2 = 10\mathrm{dB}$
- mean relative error (MRE) averaged over 400 runs



# Numerical experiments

## Synthetic signals

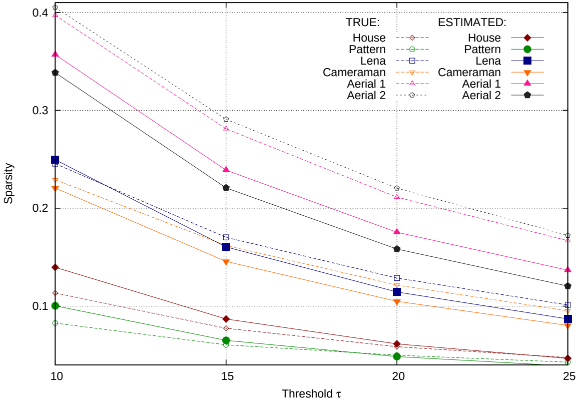
Mean relative error (MRE) of estimated sparsity ($k = 1000$)



# Numerical experiments

## Non-exactly sparse signals

Sparsity defined as the fraction of DCT components above $\tau$




Sensing parameters: $\gamma = \frac{2\tau}{3n}, \quad m = 5\frac{n}{\tau}$

# Numerical experiments

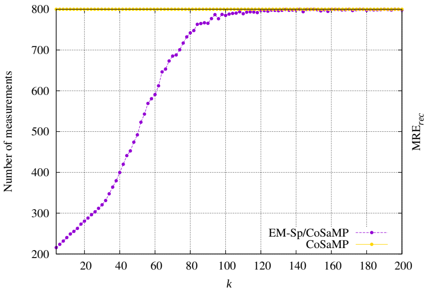
POLITECNICO
DI TORINO
Department
of Electronics and
Telecommunications

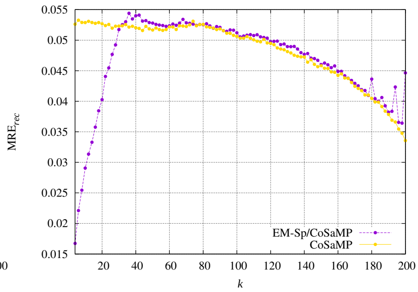
## Sparsity estimation for signal recovery

$\mathcal{S} = \{x \in \mathbb{R}^{1600}, k \in \{4, \ldots, 200\}\}$, SNR=30dB

- CoSaMP: dense matrix ($\gamma = 1$), $m = 4k_{\max}$
- EM-Sp/CoSaMP:
  - Sparse matrix ($\gamma = 8/k_{\max}$), $k_{\max}$ measurements, compute $\widehat{k}$;
  - Dense matrix ($\gamma = 1$), acquire new measurements
  - recover via CoSaMP.

Total measurements $m = \min\{4k_{\max}, \max\{10\widehat{k}, k_{\max} + 4\widehat{k}\}\}$



# Concluding remarks

POLITECNICO DI TORINO
Department of Electronics and Telecommunications

Sparsity estimation:

- noise-free setting: asymptotic behavior of ML-estimator for different regimes of CS system parameters;
- noisy setting: sparsity estimation via EM algorithm.
- numerical experiments: synthetic and real data

Future developments: useful tool in several applications

- Adaptive acquisition and sequential recovery
- Model based compressed sensing
- Estimation of support overlap between correlated signals
  - distributed compressed sensing (JSM-1, JSM-2)
  - embeddings of Jaccard coefficients for near-duplicates detection

# Recommended reading

POLITECNICO DI TORINO
Department of Electronics and Telecommunications

This paper and companion papers:

- C. Ravazzi, S. M. Fosson, T. Bianchi, E. Magli, Signal sparsity estimation from compressive noisy projections via sparsified random matrices, Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing 2016.
- C. Ravazzi, S. M. Fosson, T. Bianchi, E. Magli, Sparsity Estimation from Compressive Projections via Sparse Random Matrices, submitted to IEEE Transactions on Signal Processing, March 2016.
- D. Valsesia, S. M. Fosson, C. Ravazzi, T. Bianchi, E. Magli, SparseHash: Embedding Jaccard coefficient between support of signals, submitted to IEEE International Conference on Multimedia and Expo 2016