Recovering Cosparse Vectors using Convex *vs* Greedy Algorithms


Sangnam Nam & Rémi Gribonval
INRIA Rennes - Bretagne Atlantique, France

Mike Davies
Univ. Edinburgh, UK

Miki Elad
The Technion, Israel

Sparse **analysis** priors

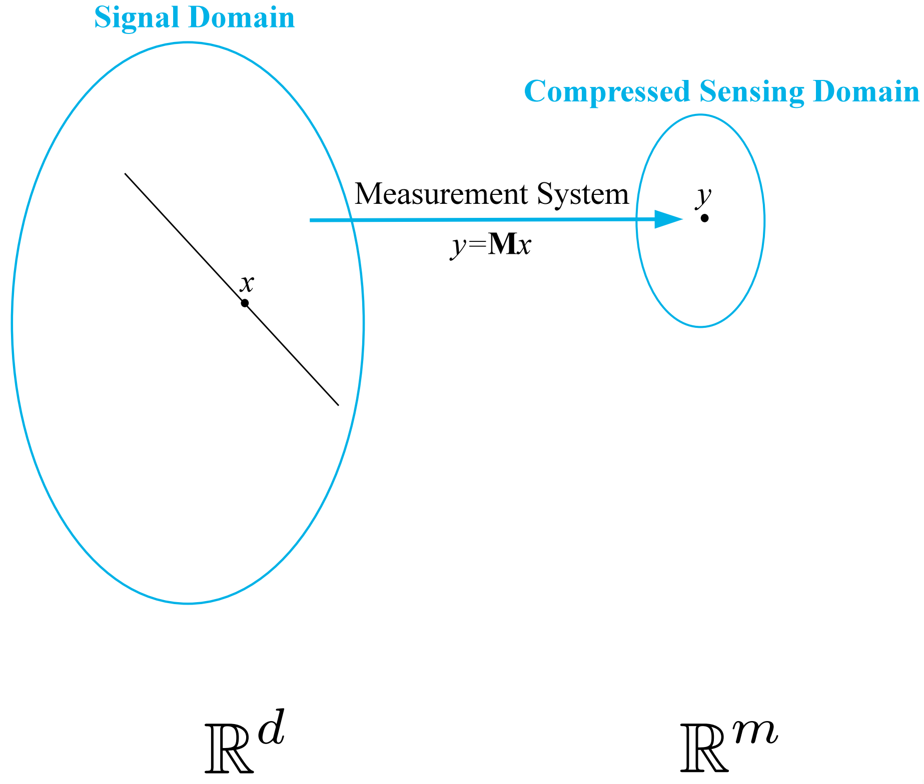
- **Inverse problem** $\mathbf{y} \approx \mathbf{M}\mathbf{x}_0$
- **Estimation by regularization**

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{M}\mathbf{x}\|_2^2 + \lambda\|\boldsymbol{\Omega}\mathbf{x}\|_p$$

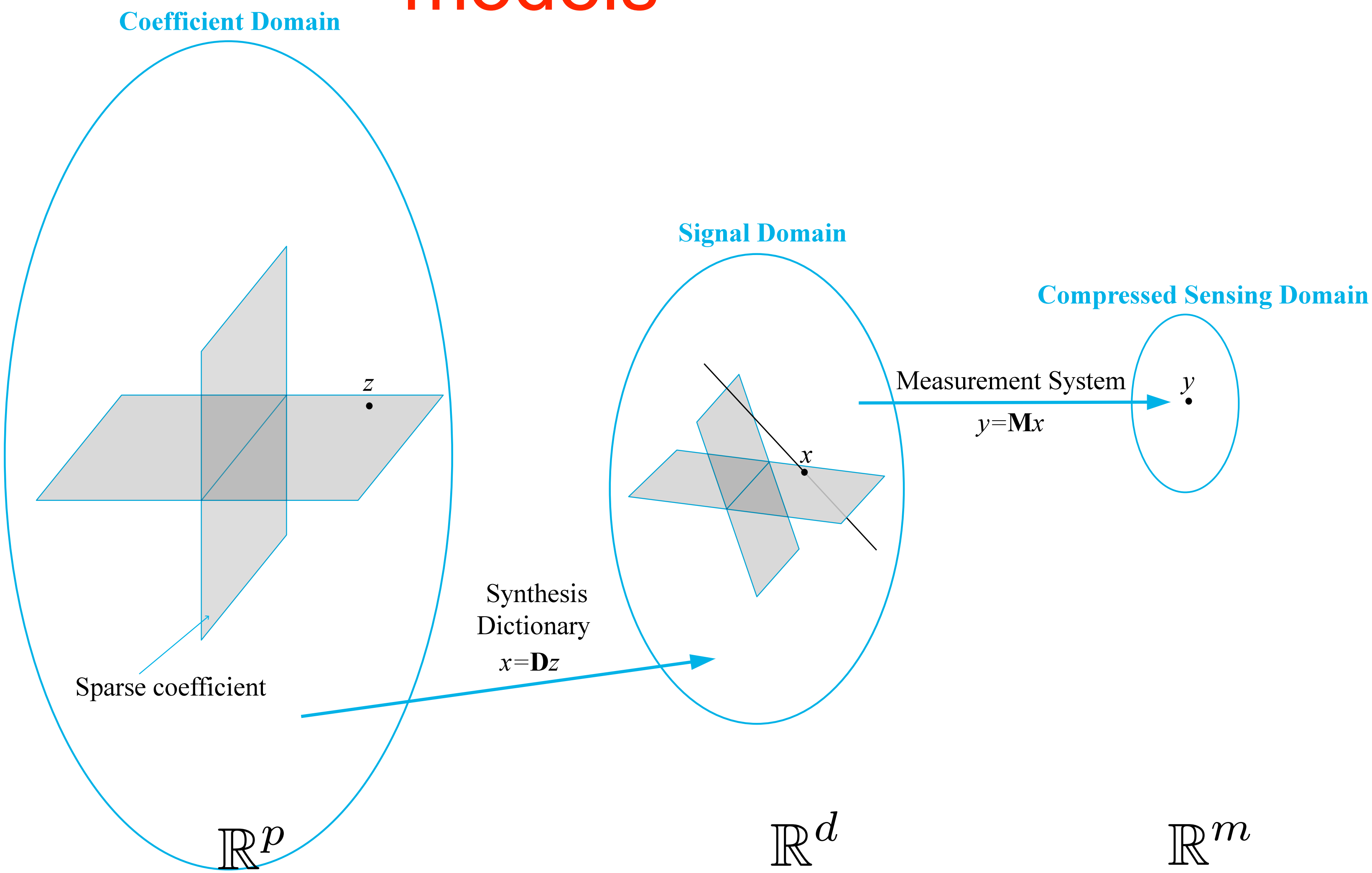
- **Sparsity-inducing effect:**
  - ✓ Analysis coefficients $\boldsymbol{\Omega}\hat{\mathbf{x}}$ of estimate $\hat{\mathbf{x}}$ are sparse
- **Hypotheses to guarantee accurate estimation ?**
  - ✓ Necessary: $\|\mathbf{x}_0 - \hat{\mathbf{x}}\|$ need to be small!
  - ✓ i.e., need true $\mathbf{x}_0$ to be well approximated by some $\hat{\mathbf{x}}$ with **sparse analysis coefficients** $\boldsymbol{\Omega}\hat{\mathbf{x}}$

Signal model

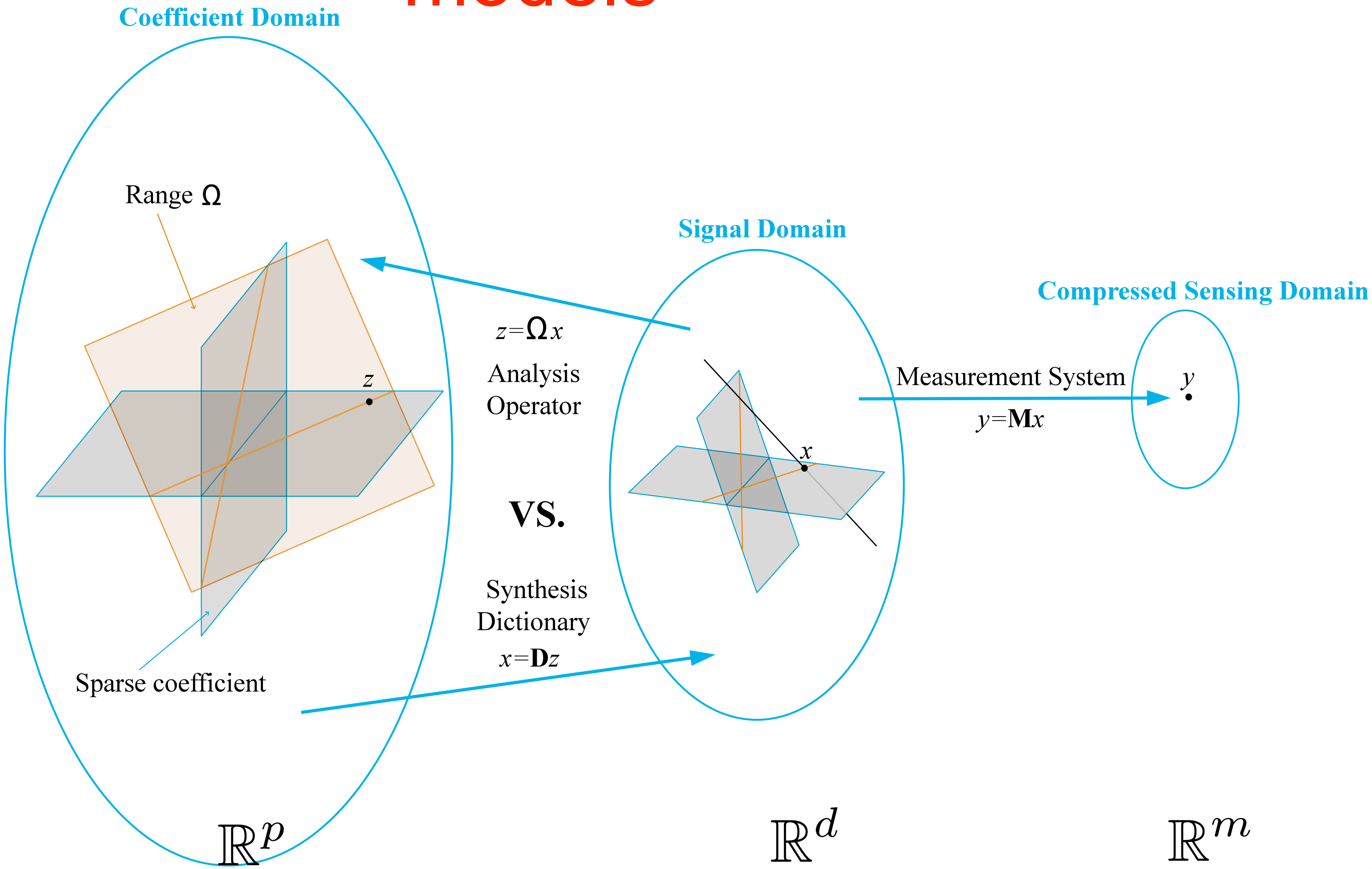
Inverse problems and low-dimensional models



Inverse problems and low-dimensional models

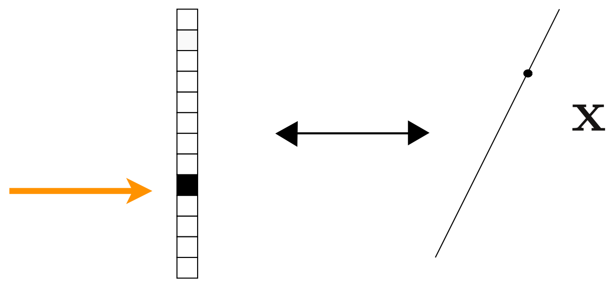


Inverse problems and low-dimensional models



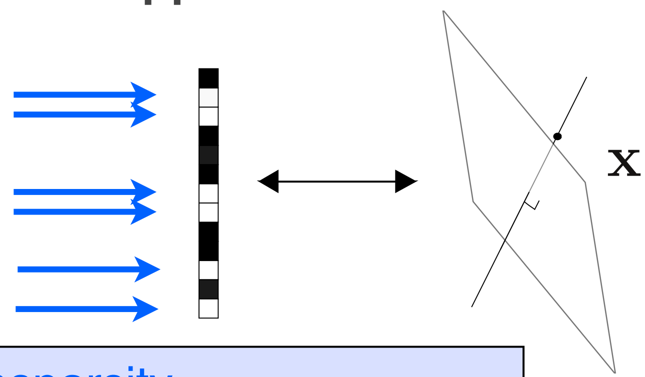
Introducing the cosparse model

- **Sparse synthesis model**
  - ✓ Synthesis dictionary $\mathbf{D}$
  - ✓ Representation $\mathbf{z}$ s.t. $\mathbf{x} = \mathbf{Dz}$
  - ✓ **Support** = location of nonzeroes

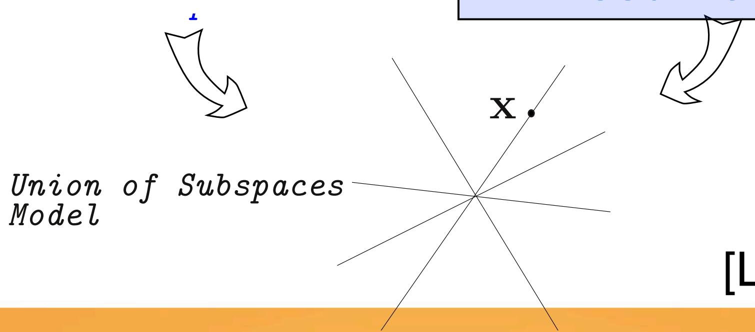


k := sparsity
= dimension of subspace

- **Cosparse analysis model**
  - ✓ Analysis operator $\mathbf{\Omega}$
  - ✓ Representation $\mathbf{\Omega x}$
  - ✓ **Cosupport** = location of zeroes



ℓ := **co**sparsity
= **co**dimension of subspace



Uniqueness results
[Lu & Do 2008] [Nam & al 2011]

(Co)sparse recovery

Algorithms for analysis regularization

- **Many algorithms with empirical success**
  - Starck & al 2003, Elad & al 2007, Portilla 2009, Selesnick & Figueiredo 2009, Afonso, Bioucas-Dias & Figueiredo 2010, Pustelnik & al 2011, Cleju & al 2012, ...
- **Recent recovery guarantees**
  - ✓ **Analysis L1 minimization**:
    - Donoho & Kutyniok 2010, Candès & al 2010, Nam & al 2011, Vaiter & al 2011.

$$\hat{\mathbf{x}} = \arg\min \|\mathbf{\Omega x}\|_1 \ s.t.\ \mathbf{y} = \mathbf{Mx}$$

  - ✓ **Greedy Analysis Pursuit** (GAP) = greedy IRLS
    - Nam & al 2011

$$\hat{\mathbf{x}}_i = \arg\min \|\mathrm{diag}(\mathbf{w}_i)\mathbf{\Omega x}\|_2 \ s.t.\ \mathbf{y} = \mathbf{Mx}$$

$$k_i = \arg\max_k |\langle \omega_k, \hat{\mathbf{x}}_i \rangle|$$

$$\mathbf{w}_{k_i} \leftarrow 0$$

Towards Empirical Phase Transition Diagrams

- Choose
  - ✓ analysis operator $\mathbf{\Omega}$
  - ✓ measurement matrix $\mathbf{M}$
  - ✓ cosparsity ℓ
- Generate random co-sparse vector
  - ✓ draw random Gaussian vector $\mathbf{u} \in \mathbb{R}^d$
  - ✓ draw cosupport Λ = choose ℓ rows of $\mathbf{\Omega}$
  - ✓ project orthogonally to these rows

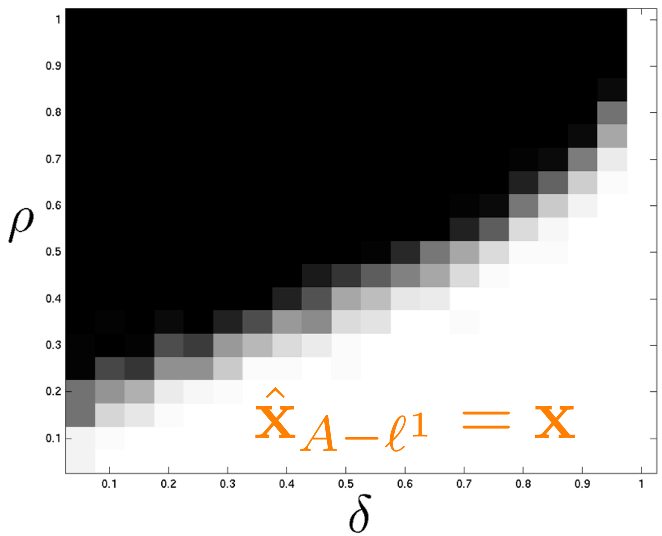
$$\mathbf{x} = (\mathbf{Id} - \mathbf{\Omega}_\Lambda^\dagger \mathbf{\Omega}_\Lambda)\mathbf{u} \in \mathbb{R}^d$$

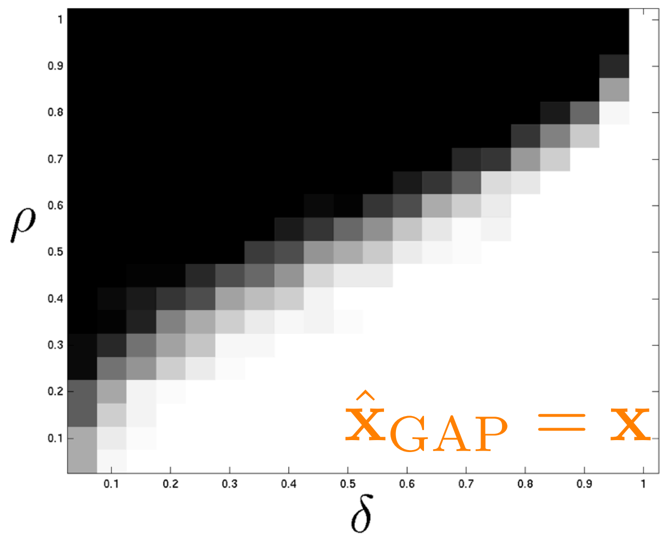
- Generate projection

$$\mathbf{y} = \mathbf{M}\mathbf{x}$$

Example 1: with Generic Analysis Operator

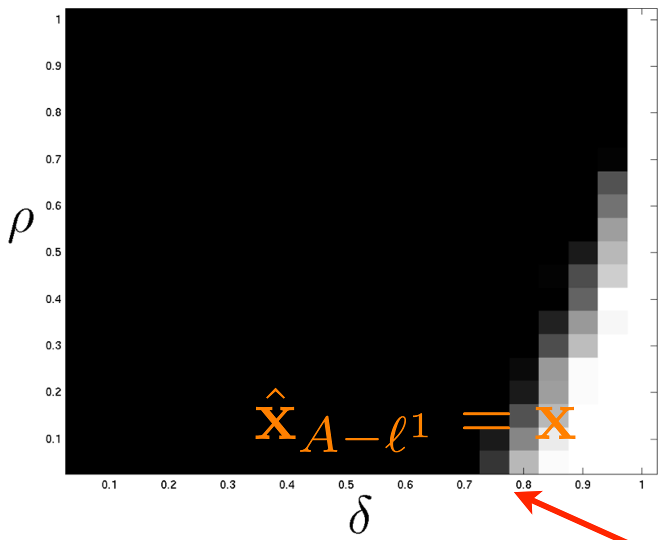
$$\mathbf{M} = \underset{m \times d}{\square} \qquad \delta = \frac{m}{d} \qquad \rho = \frac{d-\ell}{m} \qquad \mathbf{\Omega}\ \square$$

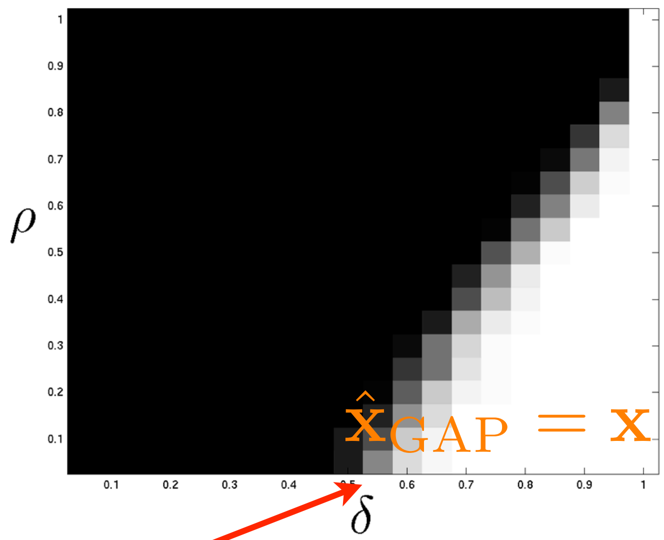




Example 1: with Generic Analysis Operator

$$\mathbf{M} = \underset{m \times d}{\square} \qquad \delta = \frac{m}{d} \qquad \rho = \frac{d-\ell}{m} \qquad \mathbf{\Omega}$$





For GENERIC operators, there is a (high) lower bound on achievable undersampling in a Compressed Sensing Scenario ...

Sparsity *vs* Co-sparsity

- **Sparsity**
  - ✓ dictionary

    $\mathbf{D} : d \times n$

  - ✓ number of **nonzeroes** = dimension

    $k := \|\mathbf{z}\|_0, \;\; \mathbf{x} = \mathbf{D}\mathbf{z}$

  - ✓ dimension of subspaces

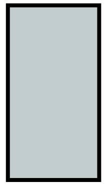
    k

  - ✓ number of subspaces

    $\binom{n}{k}$

- **Cosparsity**
  - ✓ operator

    $\mathbf{\Omega} : p \times d$

  - ✓ number of **zeroes** = co-dimension

    $\ell := p - \|\mathbf{\Omega}\mathbf{x}\|_0$

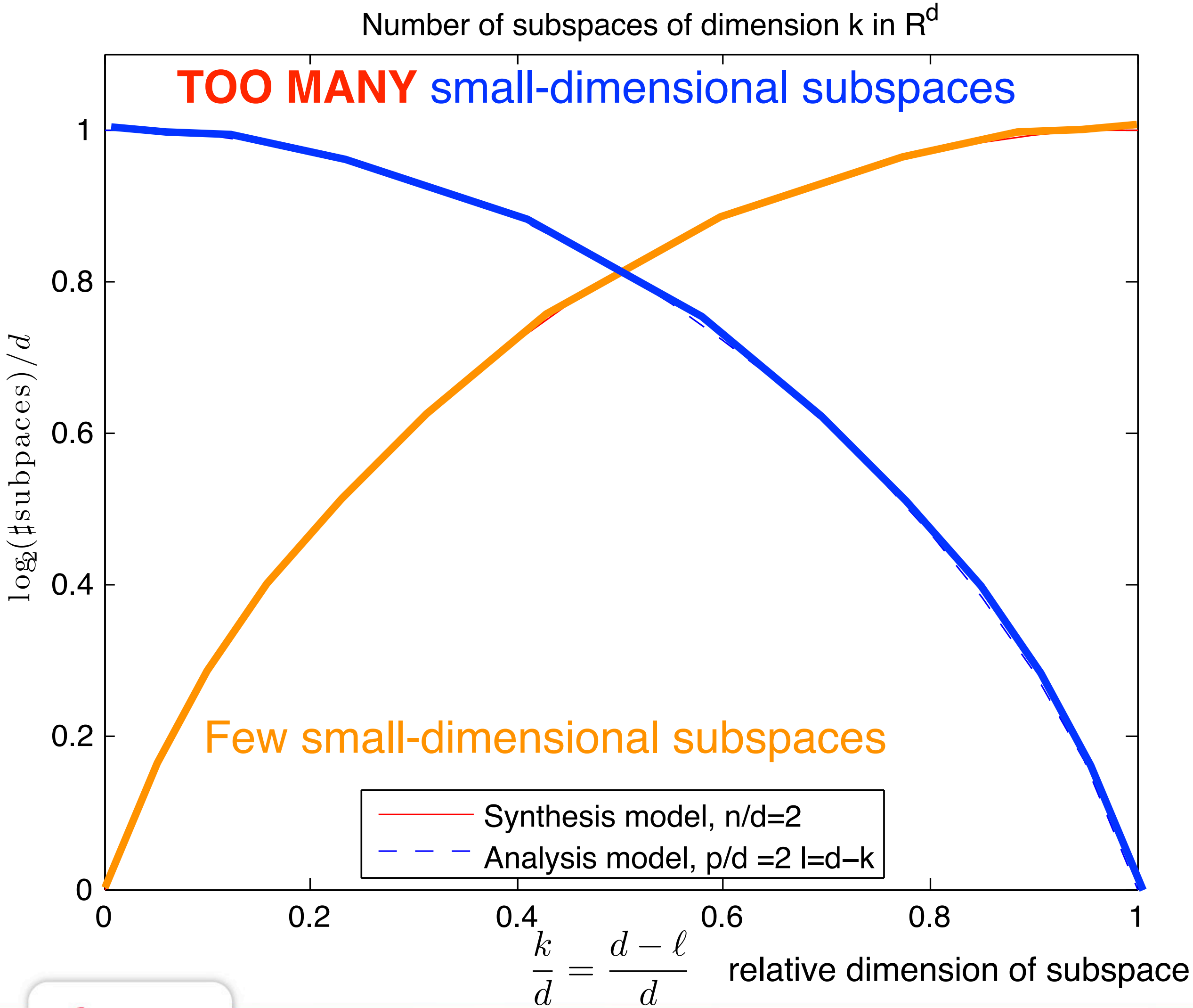
  - ✓ dimension of subspaces

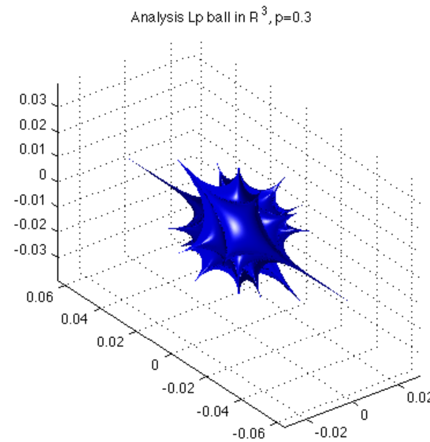
    $d - \ell$

  - ✓ number of subspaces

    $\binom{p}{\ell}$

Counting subspaces: Generic Operators





High cosparsity ?

- Generic analysis operator:
  - ✓ linearly independent rows
  - ✓ If $\mathbf{x} \neq 0$ then orthogonal to at most $\ell < d$ rows

Limited cosparsity for *generic* operators

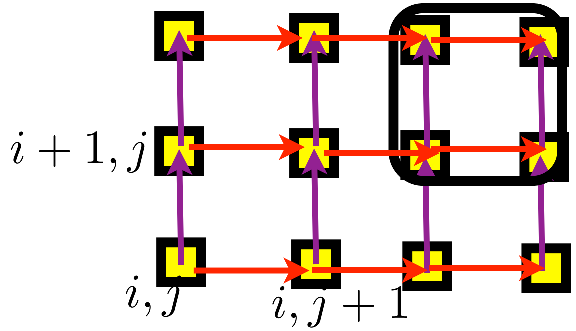
$$\|\mathbf{\Omega x}\|_0 = p - \ell > p - d$$

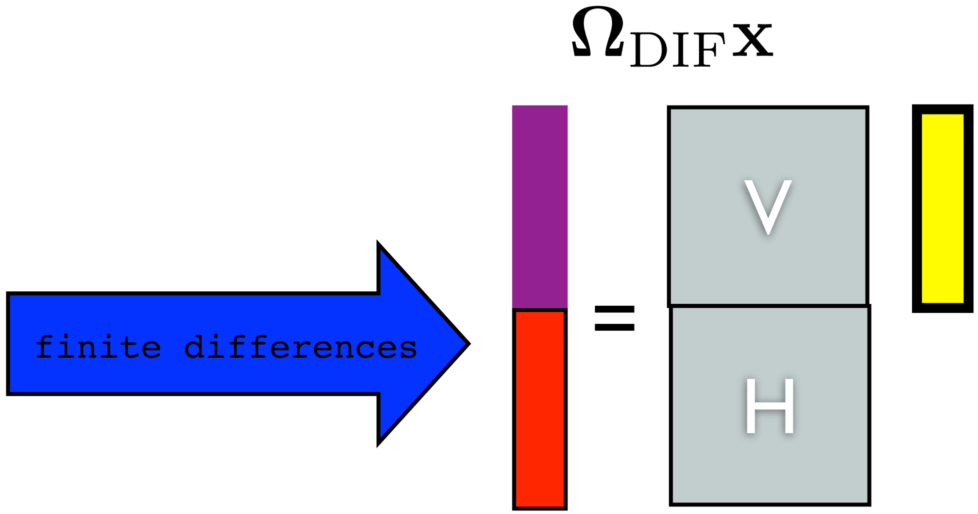
No signal has truely sparse analysis coeffs

Example 2: 2D finite difference operator

- Cousin of TV norm
  - Rudin, Osher, Fatemi 1992

$\mathbf{x} = (x_{ij})$





  - ✓ **cosupport** = edges with equal pixel values
- Operator $\mathbf{\Omega}$ is not a frame!

Dictionaryless! ~~$\mathbf{D}_{\mathrm{DIF}}$~~

- Loops = **linear dependencies** between rows of $\mathbf{\Omega}$ allowing large cosparsity

$$\ell > d$$

An image with high cosparsity

- Shepp-Logan phantom
  - ✓ 11 ellipses
  - ✓ Cosupport defines
    = **22-dimensional subspace**

$$\mathbf{x} \in \mathbb{R}^{65536}$$

$$\mathbf{\Omega x} \in \mathbb{R}^{130560}$$

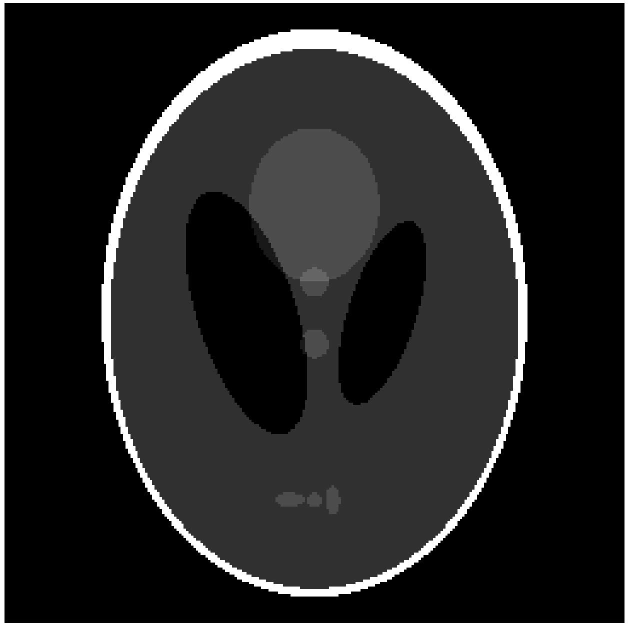
$$\|\mathbf{\Omega x}\|_0 = 2546$$

=length of edges ~ TV norm

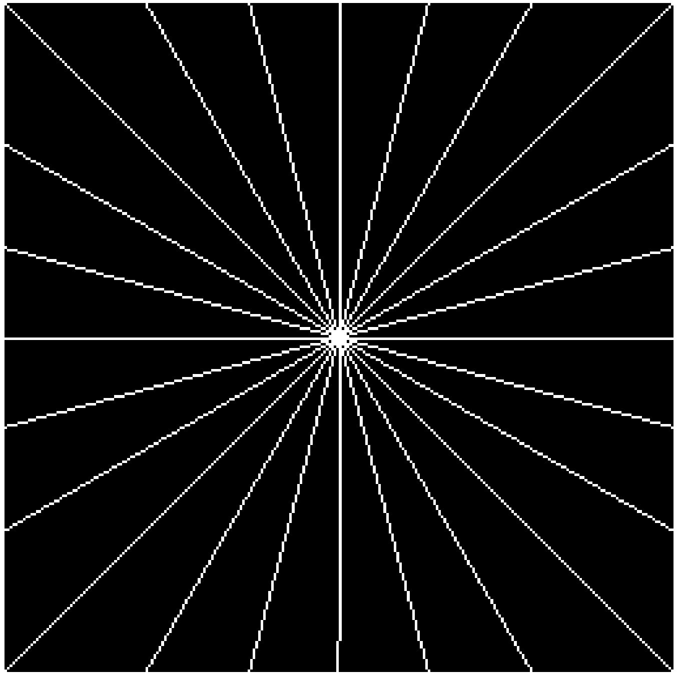
$$\ell = 128014 \quad \approx 2d$$

Uniqueness guarantee

$$m \geq 2554$$



Results: finite difference operator



Fourier subsampling

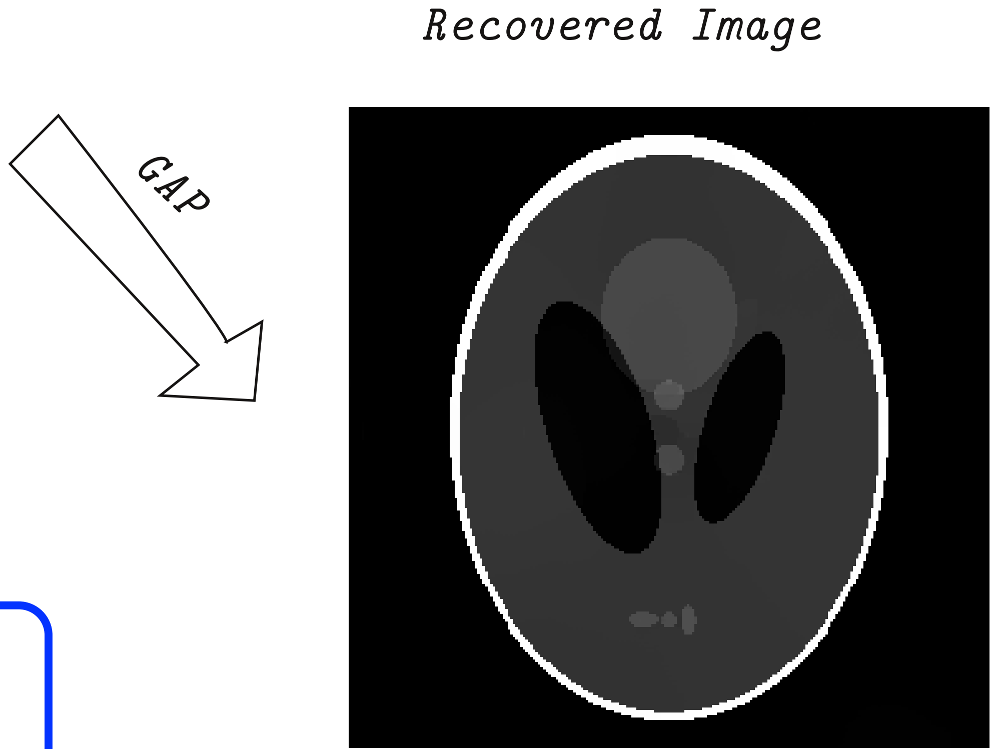
$d = 256 \times 256 = 65536$

12 slices = radial lines

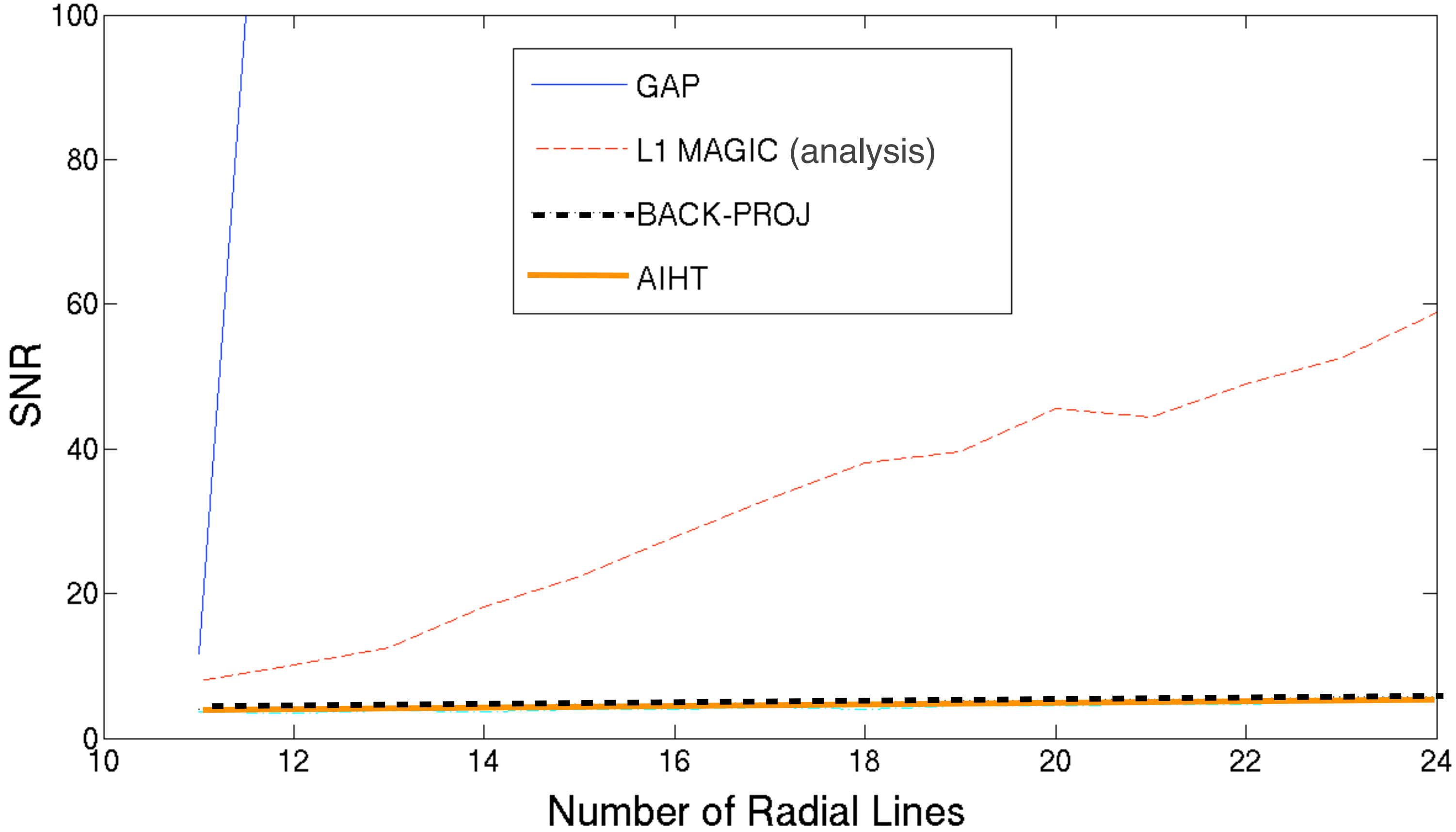
$m = 3032$ = 4.7%

Uniqueness guarantee

$m \geq 2554$



Results: finite difference operator



Linear dependencies = fewer small-dimensional subspaces (links with **matroids**?)

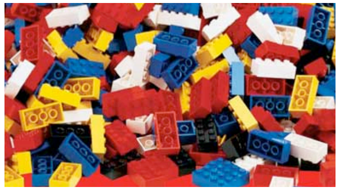
Conclusions

Take Home Message

- Traditional **Synthesis Model**
  - ✓ **Synthesis dictionary** of *atoms*

$$\mathbf{x} = \mathbf{D}\mathbf{z} = \sum_i z_i \mathbf{d}_i \qquad \|\mathbf{z}\|_0 \ll \text{dimension}$$

  - ✓ «Lego» model: building blocks



  - ✓ Low-dimension = few atoms
    - ✦ Ex: man-made codes in communications

- Alternate **Analysis Model**
  - ✓ **Analysis operator**

$\langle \omega_i, \mathbf{x} \rangle = 0$ for many *rows* of $\boldsymbol{\Omega}$

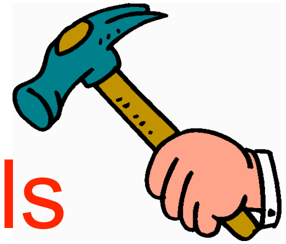
$$\|\boldsymbol{\Omega}\mathbf{x}\|_0 \ll \text{dimension}$$

  - ✓ «Carving out» model: constraints

  - ✓ Low-dimension = many constraints
    - ✦ Ex: coupling with laws of physics

$$\left(\boldsymbol{\Delta}\mathbf{x} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{x}\right)_{|\dot{\mathcal{D}}} = 0$$

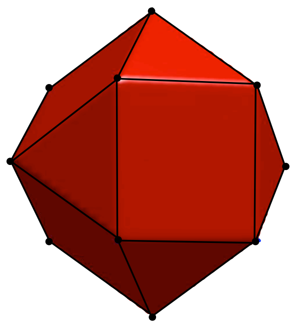
Recent results / Hanging fruits / Nails



- Uniqueness wrt inverse problems (U of subspaces) [Nam & al 2011]
- Cosparse recovery **algorithms** with guarantees
  - Analysis-L1 [Donoho & Kutyniok 2010, Candès & al 2010, Nam & al 2011, Vaiter & al 2011, ...]
  - Greedy Analysis Pursuit [Nam & al 2011]
  - Analysis Iterative Hard Thresholding, Hard Thresholding Pursuit, CoSAMP, Subspace Pursuit [Gyries & al 2011, 2012]
- **Hybrid** sparse/cosparse models [Figueiredo & al 2011]
- **Group cosparsity** [Nam & al 2012]
- **Learning/designing** analysis operators [Ophir & al 2011, Yaghoobi & al 2011, 2012, Peyré & Fadili 2011, Rubinstein & Elad 2012]
- Connections with PDEs ... [Nam & al 2012]

A sampler of more open questions

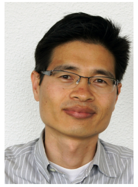
- Stability and robustness guarantees
- Hardness of projection on closest cosparse set ?
- Analysis equivalent of coherence ?
- Characterization of «good» analysis operators ?
  - ✓ linear dependencies are desirable
  - ✓ non-generic geometry of analysis ball is needed
    - ✦ ex: TV-ball (courtesy G. Peyré)



- Analysis equivalent of low-rank ?
- Equivalent of other atomic norms ?

THANKS

- Joint work with
  - **Sangnam Nam** (INRIA, France)
  - Mike Davies (University of Edinburgh, UK)
  - Miki Elad (The Technion, Israel)



- Funding

small-project.eu

SMALL

SEVENTH FRAMEWORK PROGRAMME

Bibliography (1)

- Logan, Jr. *Information in the zero crossings of bandpass signals*. AT&T Technical Journal, 56:487– 510, Apr. 1977.
- Mallat. *Zero-crossings of a wavelet transform*. Information Theory, IEEE Transactions on, 37(4):1019 –1033, jul 1991.
- Rudin, Osher, and Fatemi. *Nonlinear total variation based noise removal algorithms*. Physica D, 60:259–268, 1992.
- Dragotti and Vetterli. *Wavelet footprints: Theory, algorithms, and applications*. IEEE Transactions on Signal Processing, 51(5):1306–1323, May 2003.
- Starck, Elad, and Donoho. *Image decomposition : Separation of textures from piecewise smooth content*. Wavelets X, SPIE, 2003.
- Cordero and Gröchenig. *Localization of frames ii*. ACHA, 17(1):29 – 47, 2004. Special Issue: Frames in Harmonic Analysis, Part 1.
- Gröchenig. *Localization of frames, Banach frames, and the invertibility of the frame operator*. J. Fourier Anal. Appl., 10(2):105–132, Mar. 2004.

Bibliography (2)

- Selesnick and Figueiredo. *Signal restoration with overcomplete wavelet transforms: Comparison of analysis and synthesis priors*. SPIE (Wavelets XIII), August 2009.
- Elad, Milanfar, and Rubinstein. *Analysis versus synthesis in signal priors*. Inverse problems, 23:947–968, 2007.
- Portilla. *Image restoration through l0 analysis-based sparse optimization in tight frames*. ICIP, pages 3865–3868, 2009.
- Candès, Eldar, Needell, and Randall. *Compressed sensing with coherent and redundant dictionaries*. ACHA, Corrected Proof:–, 2010.
- Kutyniok. *Data separation by sparse representations*. In Compressed Sensing: Theory and Applications. Cambridge University Press, to appear.