SPARS 11

Sparsity & Co.:

An Overview of *Analysis* vs *Synthesis* in Low-Dimensional Signal Models

Rémi Gribonval

INRIA, France

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

INRIA

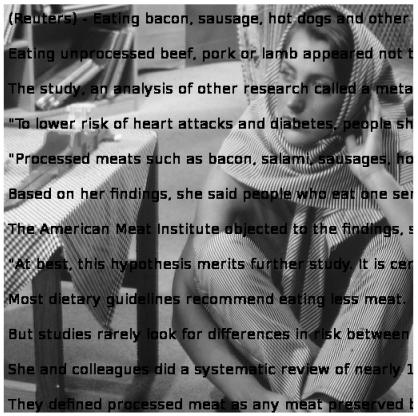
centre de recherche RENNES - BRETAGNE ATLANTIQUE

Outline

- Sparsity: via Analysis or Synthesis ?
- An Lp perspective $0 < p \leq 1$ with *localized frames*
- An L0 perspective: introducing *co-sparsity*
- Cosparse recovery ?
- What's next ?

Sparse Signal / Image Processing

denoising



inpainting

+ Compression,
Source Localization, Separation,
Compressed Sensing ...

Sparse Atomic Decompositions

$$\mathbf{x} \approx \mathbf{D}\mathbf{z}$$

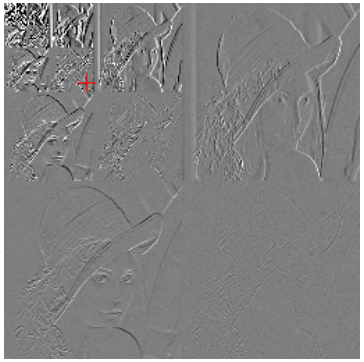
Signal Image	(Overcomplete) dictionary of atoms	Representation Coefficients

Supporting Evidence: Sparsifying *Transforms*

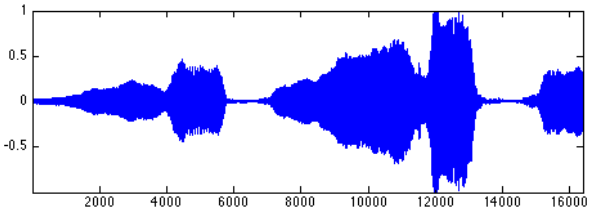
natural image

$\mathbf{x}$

DWT

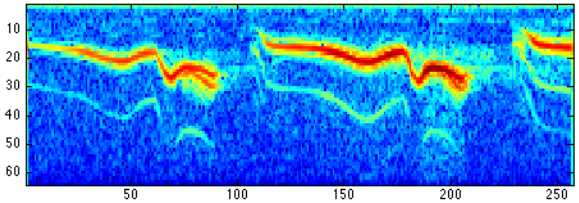


sparse coefficients $\mathbf{z}$

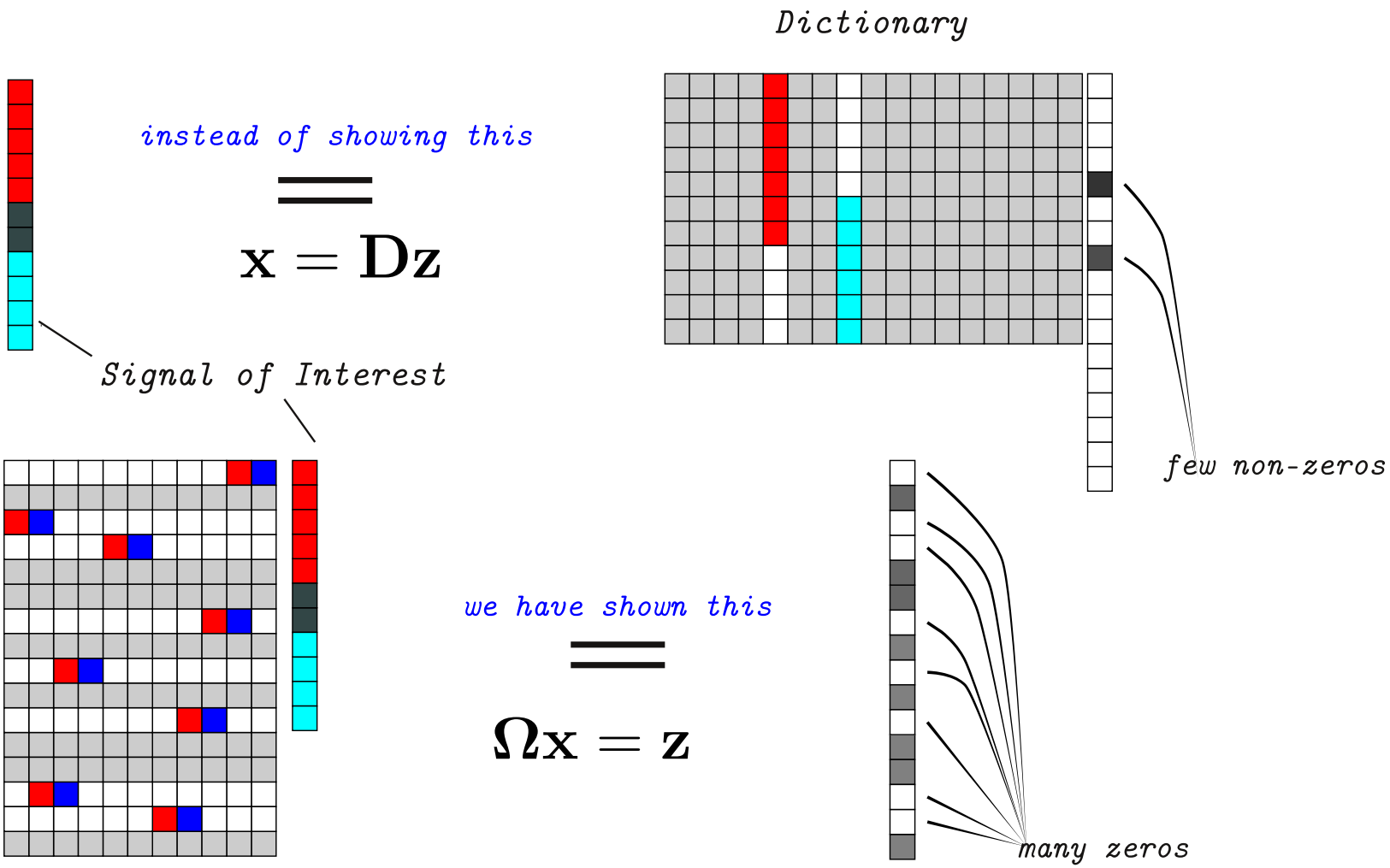


audio signal

STFT

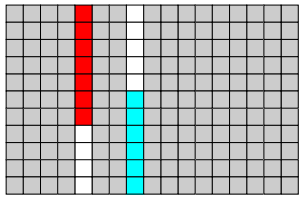


Transforms = Atomic Decompositions?



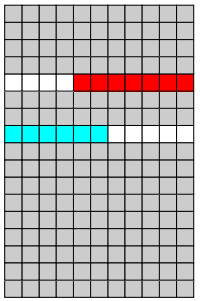
Transforms = Atomic Decompositions ?

Dictionary $\mathbf{D}$



: Tight frame

Operator $\mathbf{\Omega}$



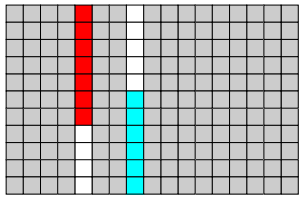
$= \mathbf{D}^T$

$$\mathbf{x} = \mathbf{D}\mathbf{\Omega}\mathbf{x}.$$

$\mathbf{x}$ *: sparse in* $\mathbf{D}$ $\Longleftrightarrow$ $\mathbf{\Omega}\mathbf{x}$ *: sparse*

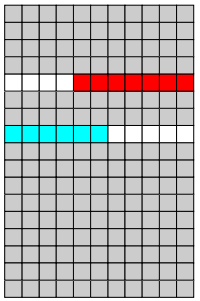
Transforms = Atomic Decompositions ?

Dictionary $\mathbf{D}$



: Tight frame

Operator $\mathbf{\Omega}$



$= \mathbf{D}^T$

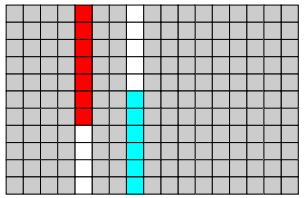
$$\mathbf{x} = \mathbf{D}\mathbf{\Omega}\mathbf{x}.$$

$\mathbf{x}$ *: sparse in* $\mathbf{D}$ $\Longleftrightarrow$ $\mathbf{\Omega}\mathbf{x}$ *: sparse*

- Yes ... but some troubling facts:
  - ✓ infinitely many synthesis representations
  - ✓ only one analysis representation

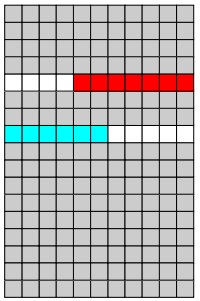
Transforms = Atomic Decompositions ?

Dictionary $\mathbf{D}$



: Tight frame

Operator $\mathbf{\Omega}$



$= \mathbf{D}^T$

$$\mathbf{x} = \mathbf{D}\mathbf{\Omega}\mathbf{x}.$$

$\mathbf{x}$ *: sparse in* $\mathbf{D}$ $\Longleftrightarrow$ $\mathbf{\Omega}\mathbf{x}$ *: sparse*

- Yes ... but some troubling facts:
  - ✓ infinitely many synthesis representations
  - ✓ only one analysis representation
- By the way, what do we mean by «sparse» ?

Analysis *vs* Synthesis: Lp sparsity in frames

with M. Nielsen

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

centre de recherche
RENNES - BRETAGNE ATLANTIQUE

Frames in Hilbert spaces

- Frame = energy preserving analysis transform

$$A \cdot \|\mathbf{x}\|^2 \leq \|\mathbf{D}^T\mathbf{x}\|_2^2 \leq B \cdot \|\mathbf{x}\|_2^2, \quad \forall \mathbf{x} \in \mathcal{H}.$$

- Canonical dual frame $\mathbf{\Omega} = \mathbf{D}^+ := \mathbf{D}^T(\mathbf{D}\mathbf{D}^T)^{-1}$
  - ✓ perfect reconstruction property

$$\mathbf{x} = \mathbf{D}\mathbf{\Omega}\mathbf{x}, \ \forall \mathbf{x} \in \mathcal{H}$$

  - ✓ minimum energy coefficients

$$\|\mathbf{\Omega}\mathbf{x}\|_2 = \min_{\mathbf{z}:\mathbf{D}\mathbf{z}=\mathbf{x}} \|\mathbf{z}\|_2$$

minimum Lp norms?

Measures of sparsity

- Lp (quasi)-norms
  - ✓ p=0 $$\|\mathbf{z}\|_0 := \sharp\{j : \ z_j \neq 0\}$$
  - ✓ p>0 $$\|\mathbf{z}\|_p^p := \sum_j |z_j|^p$$
- Lp = sparsity-inducing for $0 \leq p \leq 1$

Analysis vs Synthesis sparsity

- For a frame **D** and its canonical dual

$$\|\mathbf{\Omega x}\|_2 = \min_{\mathbf{z}:\mathbf{Dz}=\mathbf{x}} \|\mathbf{z}\|_2$$

- Norm associated to sparsest synthesis coefficients

$$|\mathbf{x}|_p := \inf_{\mathbf{z}:\ \mathbf{Dz}=\mathbf{x}} \|\mathbf{z}\|_p \leq \|\mathbf{\Omega x}\|_p$$

- Converse ? $\|\mathbf{\Omega x}\|_p \leq C|\mathbf{x}|_p$?

Analysis vs Synthesis equivalence: *localized* frames

- Notations
  - Atoms = columns of dictionary $\mathbf{D} = [\mathbf{d}_j]_j$, canonical dual $\mathbf{\Omega}$

- **Theorem**
  - If: $$C_q := \sup_j \|\mathbf{\Omega}\mathbf{d}_j\|_q < \infty$$
  - Then $\forall p,\ q \leq p \leq 2$:
$$|\mathbf{x}|_p \leq \|\mathbf{\Omega}\mathbf{x}\|_p \leq C_q |\mathbf{x}|_p, \quad \forall \mathbf{x}$$

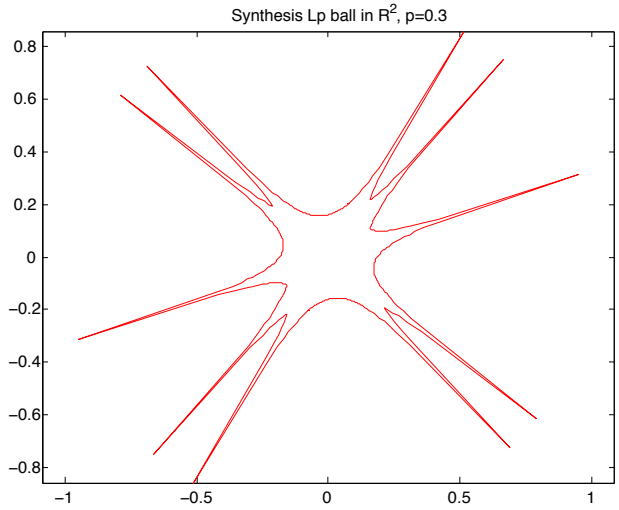
✓ *Minimum L2 norm coefficients = near Lp sparsest!*

- *Gröchenig 2004, Cordero&Gröchenig 2004, Gribonval&Nielsen 2007.*

Geometry of Lp balls

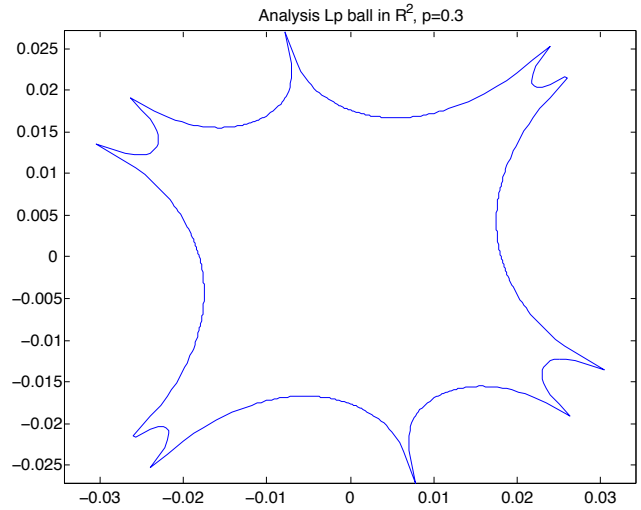
- Synthesis viewpoint

$$\{\mathbf{x} : \ |\mathbf{x}|_p \leq 1\}$$



- Analysis viewpoint

$$\{\mathbf{x} : \ \|\mathbf{\Omega}\mathbf{x}\|_p \leq 1\}$$



- **D** = 5 random unit atoms

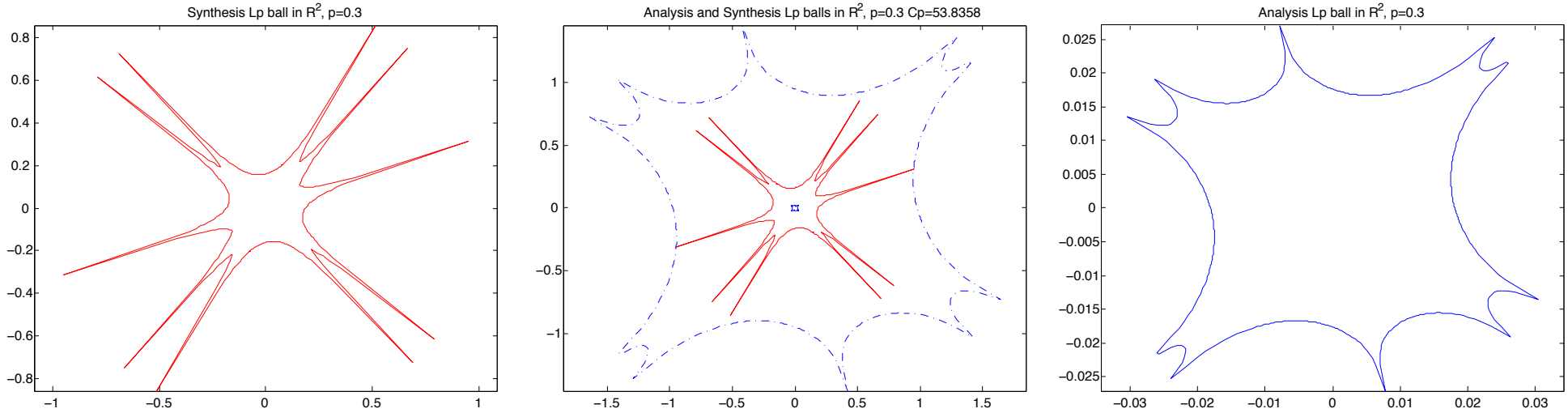
Geometry of Lp balls

- Synthesis viewpoint

$$\{\mathbf{x} : \ |\mathbf{x}|_p \leq 1\}$$

- Analysis viewpoint

$$\{\mathbf{x} : \ \|\mathbf{\Omega}\mathbf{x}\|_p \leq 1\}$$



- **D** = 5 random unit atoms $\quad C_p \approx 50$

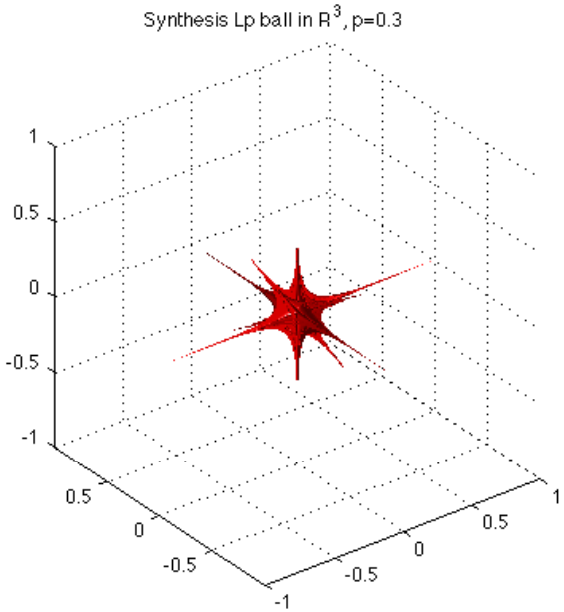
Geometry of Lp balls

- Synthesis viewpoint

$$\{\mathbf{x} : \ |\mathbf{x}|_p \leq 1\}$$

- Analysis viewpoint

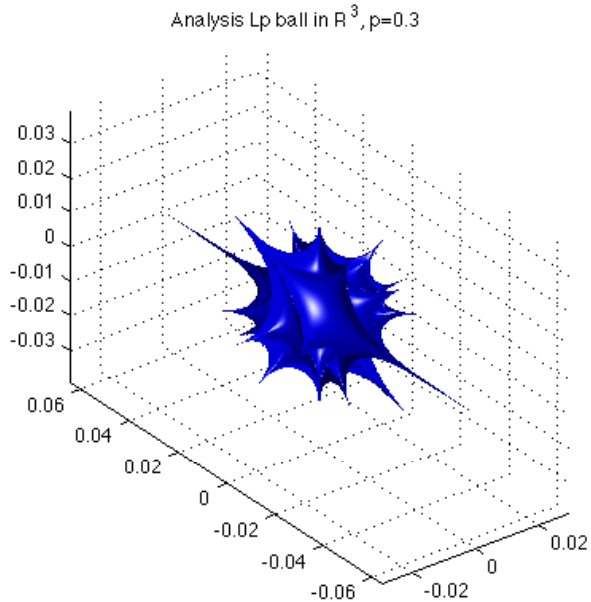
$$\{\mathbf{x} : \ \|\mathbf{\Omega}\mathbf{x}\|_p \leq 1\}$$



✓ *Different sizes*: analysis ball smaller than synthesis one

✓ *Different shapes*: analysis ball has **more peaks** than synthesis one

- For p=1, see Elad & al 2007

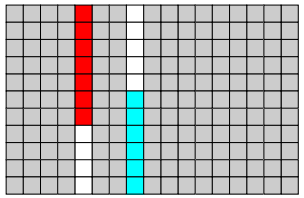


$$C_p \approx 35$$

- **D = Dirac ∪ DCT**

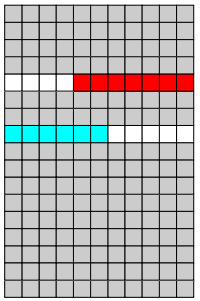
Transforms = Atomic Decompositions ?

Dictionary $\mathbf{D}$



: Tight frame

Operator $\mathbf{\Omega}$



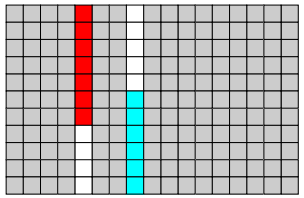
$= \mathbf{D}^T$

$$\mathbf{x} = \mathbf{D}\mathbf{\Omega}\mathbf{x}.$$

$\mathbf{x}$ *: sparse in* $\mathbf{D}$ $\Longleftrightarrow$ $\mathbf{\Omega}\mathbf{x}$ *: sparse*

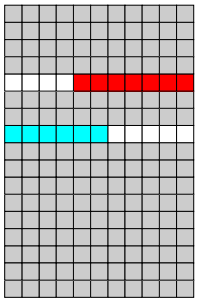
Transforms = Atomic Decompositions ?

Dictionary $\mathbf{D}$



: Tight frame

Operator $\mathbf{\Omega}$



$= \mathbf{D}^T$

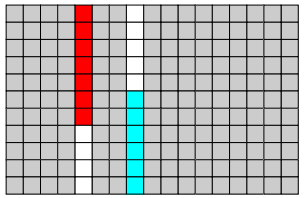
$$\mathbf{x} = \mathbf{D}\mathbf{\Omega}\mathbf{x}.$$

$\mathbf{x}$ *: sparse in* $\mathbf{D}$ $\Longleftrightarrow$ $\mathbf{\Omega}\mathbf{x}$ *: sparse*

- Yes for *localized* frames with Lp norm 0<p<2

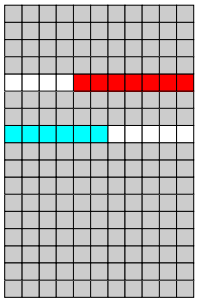
Transforms = Atomic Decompositions ?

Dictionary $\mathbf{D}$



: Tight frame

Operator $\mathbf{\Omega}$



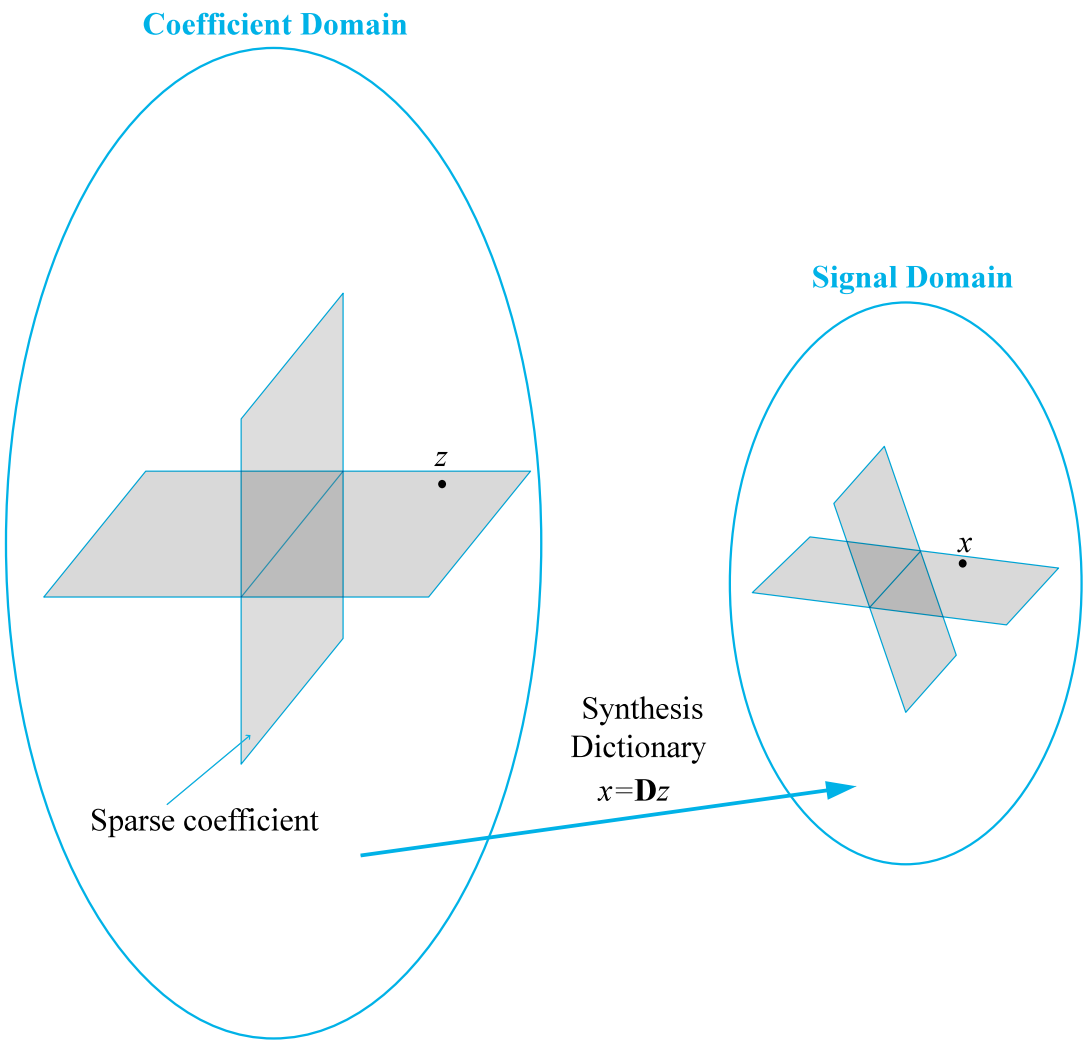
$= \mathbf{D}^T$

$$\mathbf{x} = \mathbf{D}\mathbf{\Omega}\mathbf{x}.$$

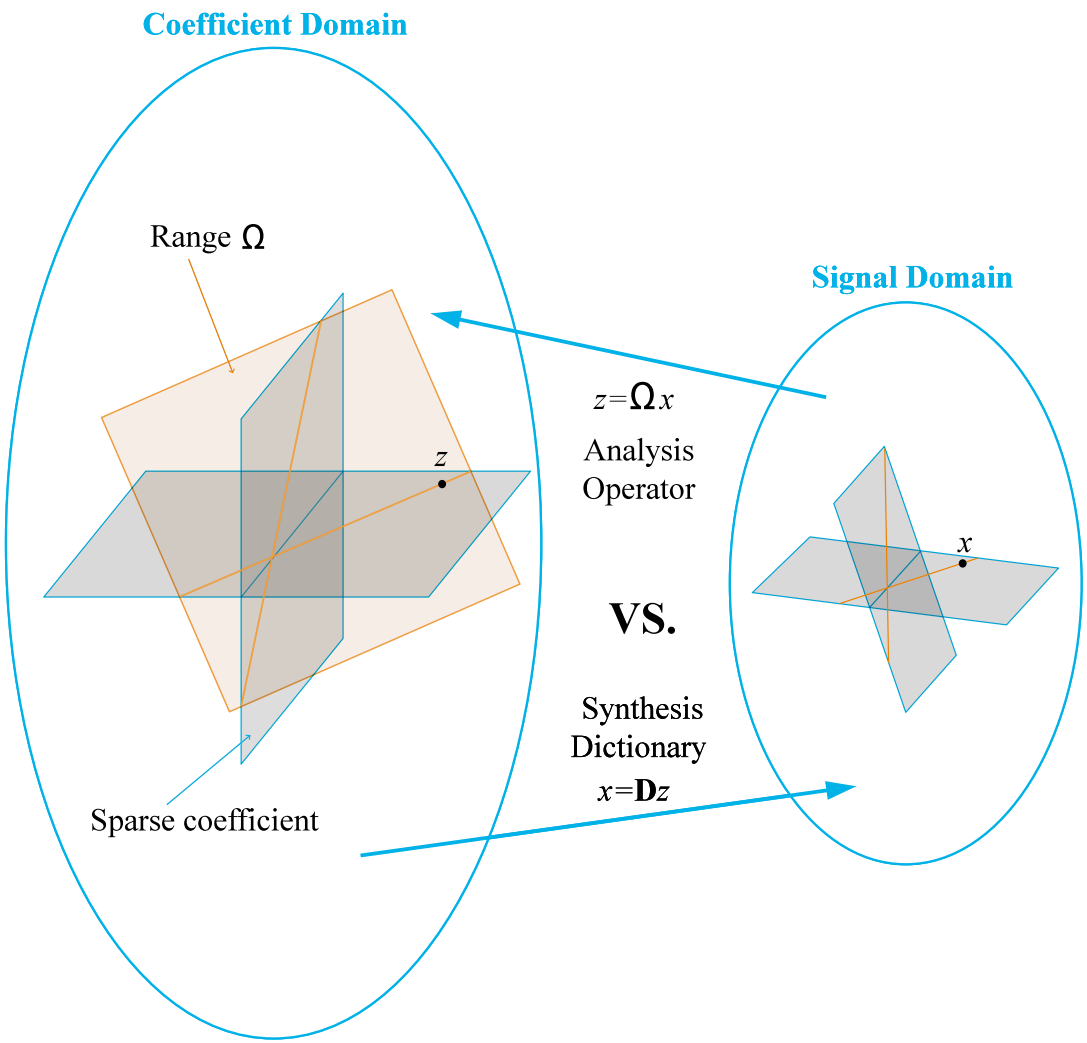
$\mathbf{x}$ *: sparse in* $\mathbf{D}$ $\Longleftrightarrow$ $\mathbf{\Omega}\mathbf{x}$ *: sparse*

- Yes for *localized* frames with Lp norm 0<p<2
- But ...

Geometry of sparse coefficients?



Geometry of sparse coefficients?



Transforms = Atomic Decompositions ?
Generic Analysis Operators

- For **generic** tight frame with n=2d we have

$$\|\mathbf{\Omega x}\|_0 \leq d \quad \Longrightarrow \quad \mathbf{x} = 0$$

✓ *No signal has truely sparse analysis coefficients*

- But:

✓ *The fact that* $\|\mathbf{\Omega x}\|_0 < n$ *is a* ***model*** *on signal* $\mathbf{x}$

✓ *Many analysis operators of interest are* ***not generic***

- *ex: Casazza, Heinecke, Krahmer, and Kutyniok. Optimally Sparse Frames. 2010.* **(Session #1)**

Analysis *vs* Synthesis: *Cosparsity*

with S. Nam, M. Davies, M. Elad

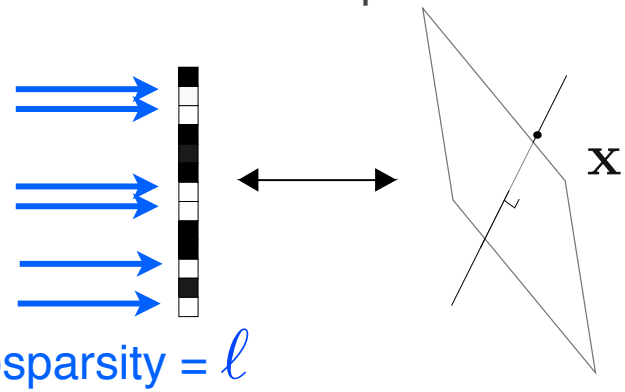
Introducing the cosparse model

- **Cosparse analysis model**
  - ✓ Analysis operator $\mathbf{\Omega}$
  - ✓ Representation $\mathbf{\Omega x}$
  - ✓ Zeroes of the representation

- **Sparse synthesis model**
  - ✓ Synthesis dictionary $\mathbf{D}$
  - ✓ Representation $\mathbf{z}$ s.t. $\mathbf{x} = \mathbf{Dz}$
  - ✓ Nonzeroes of the representation

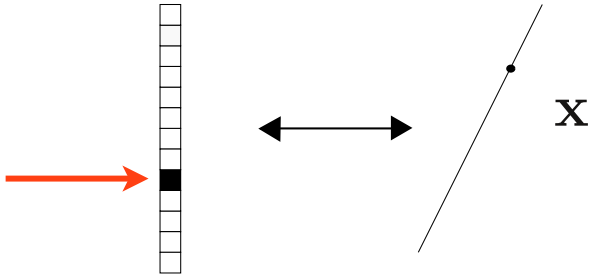
Introducing the cosparse model

- **Cosparse analysis model**
  - ✓ Analysis operator $\mathbf{\Omega}$
  - ✓ Representation $\mathbf{\Omega x}$
  - ✓ Zeroes of the representation



cosparsity = ℓ
= codimension of subspace

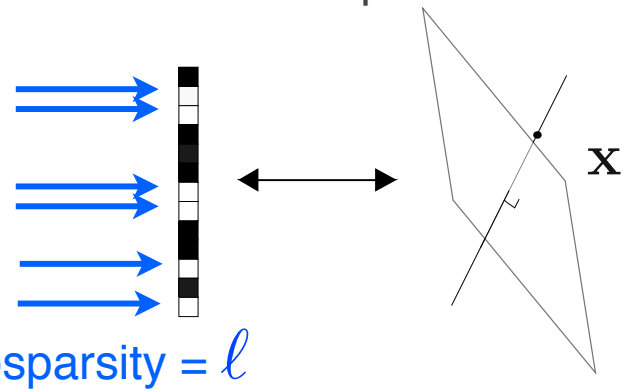
- **Sparse synthesis model**
  - ✓ Synthesis dictionary $\mathbf{D}$
  - ✓ Representation $\mathbf{z}$ s.t. $\mathbf{x} = \mathbf{Dz}$
  - ✓ Nonzeroes of the representation



sparsity = *k*
= dimension of subspace

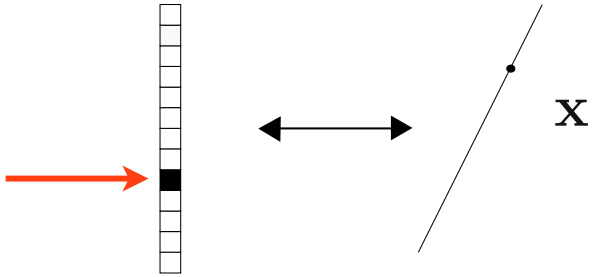
Introducing the cosparse model

- **Cosparse analysis model**
  - ✓ Analysis operator $\mathbf{\Omega}$
  - ✓ Representation $\mathbf{\Omega x}$
  - ✓ Zeroes of the representation

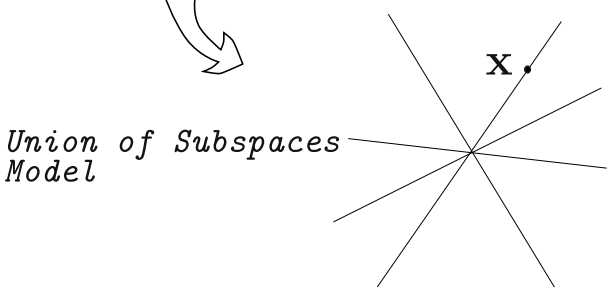


cosparsity = ℓ
= codimension of subspace

- **Sparse synthesis model**
  - ✓ Synthesis dictionary $\mathbf{D}$
  - ✓ Representation $\mathbf{z}$ s.t. $\mathbf{x} = \mathbf{Dz}$
  - ✓ Nonzeroes of the representation



sparsity = k
= dimension of subspace



Co-sparsity *vs* Sparsity

- **Cosparsity**
  - ✓ operator

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

  - ✓ number of zeroes = co-dimension

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

  - ✓ dimension of subspace

    $$d - \ell$$

  - ✓ number of subspaces

    $$\binom{n}{\ell}$$

- **Sparsity**
  - ✓ dictionary

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

  - ✓ number of nonzeroes = dimension

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

  - ✓ dimension of subspace

    $$k$$

  - ✓ number of subspaces

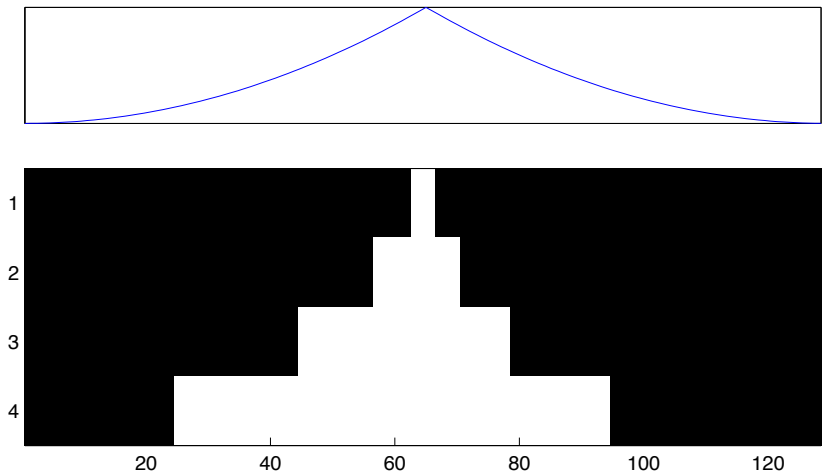
    $$\binom{n}{k}$$

Example 1: Undecimated wavelets

- Sparse model: wavelet expansions $\mathbf{x} = \sum_{j,k} z_{j,k} \psi_{j,k}$
  - *support* = location of significant wavelet coefficients
  - a single singularity = a large *footprint*
    - Dragotti&Vetterli 2003

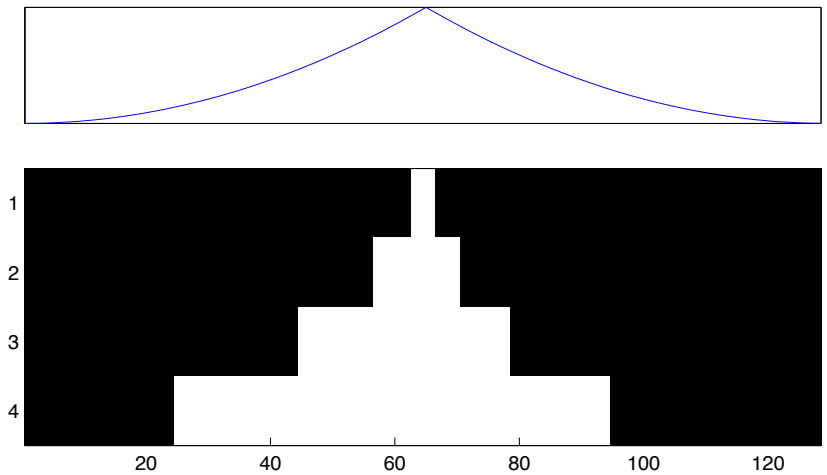
- Cosparse model ?
  - *cosupport* = zero-crossings
    - Logan 1977, Mallat 1991

- Two-scale relations, etc.
  - linear dependencies
  - allows larger cosparsity
    - Selesnick & Figueiredo 2009



Example 1: Undecimated wavelets

- Sparse model: wavelet expansions $\mathbf{x} = \sum_{j,k} z_{j,k} \psi_{j,k}$
  - *support* = location of significant wavelet coefficients
  - a single singularity = a large *footprint*
    - Dragotti&Vetterli 2003



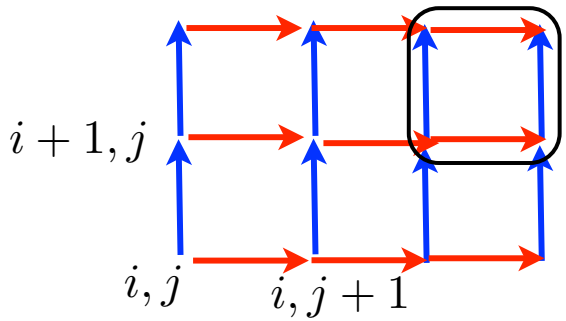
- Cosparse model ?
  - *cosupport* = zero-crossings
    - Logan 1977, Mallat 1991

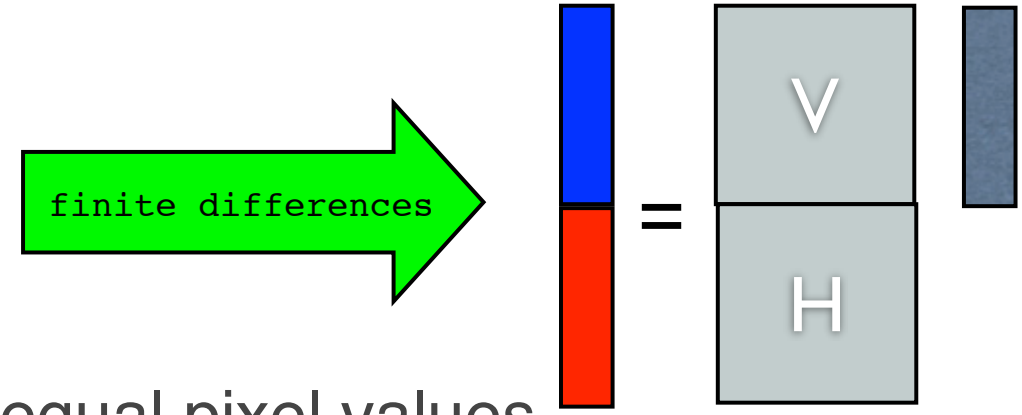
- Two-scale relations, etc.
  - linear dependencies
  - allows larger cosparsity $\ell > d$
    - Selesnick & Figueiredo 2009

Example 2: finite difference operator

- Finite-difference operator = cousin of TV norm
  - Rudin, Osher, Fatemi 1992

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





$\Omega_{\mathrm{DIF}}\mathbf{x}$

✓ *cosupport* = edges with equal pixel values

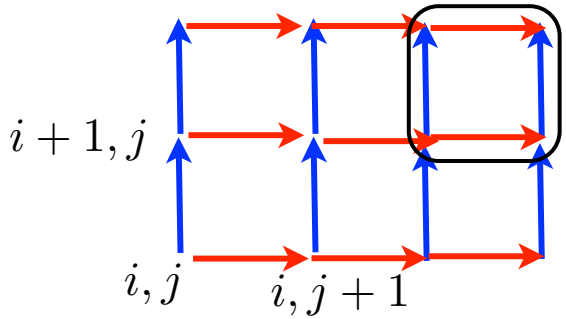
✓ not a frame!

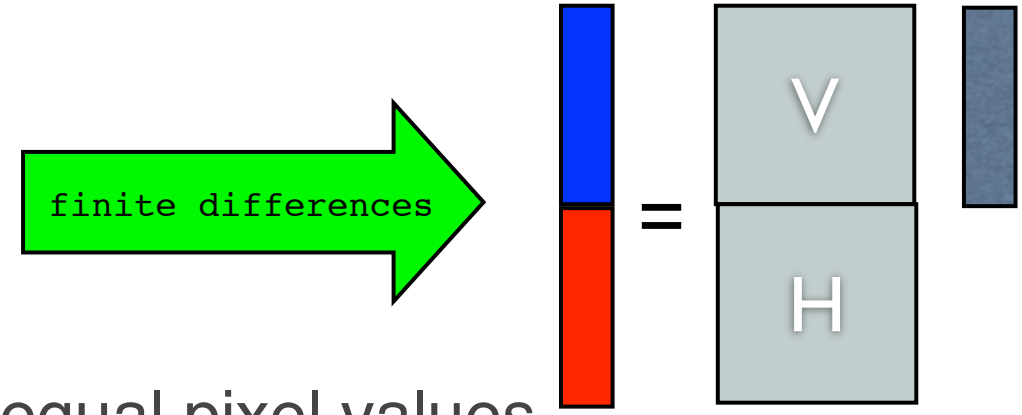
- Loops
  - ✦ linear dependencies between rows
  - ✦ allows larger cosparsity

Example 2: finite difference operator

- Finite-difference operator = cousin of TV norm
  - Rudin, Osher, Fatemi 1992

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





$\mathbf{\Omega}_{\mathrm{DIF}}\mathbf{x}$

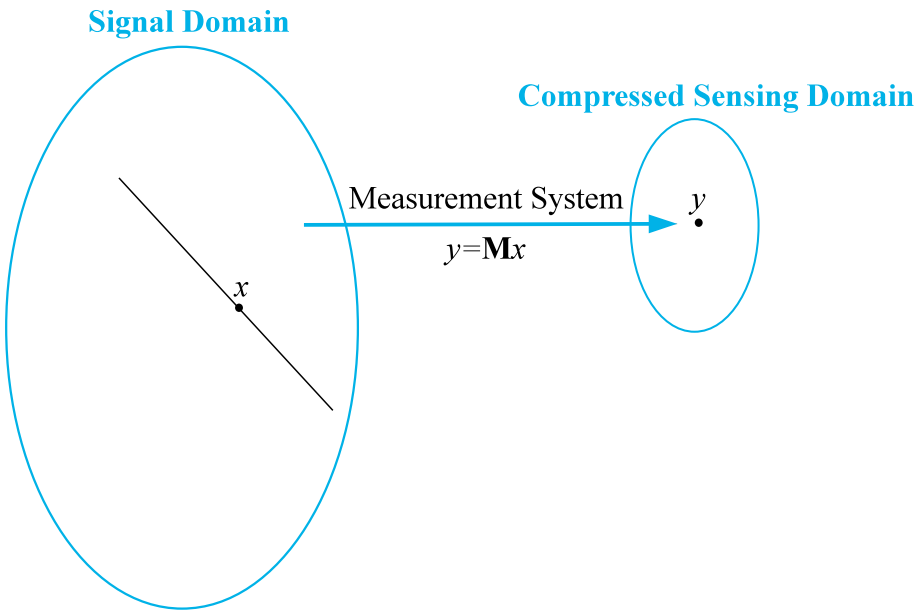
✓ *cosupport* = edges with equal pixel values

✓ not a frame!

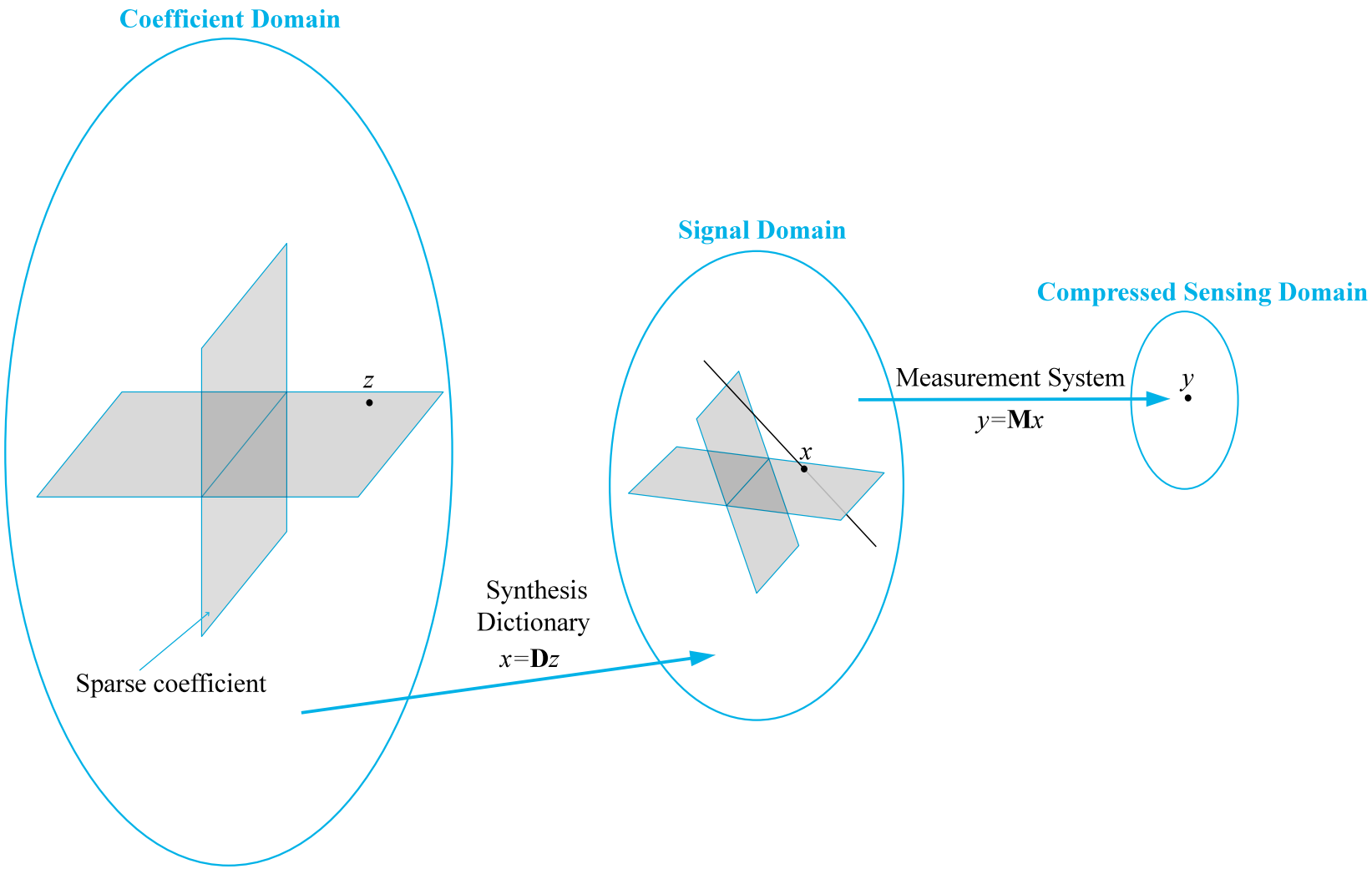
- Loops
  - ✦ linear dependencies between rows
  - ✦ allows larger cosparsity $\ell > d$

Cosparse recovery ?

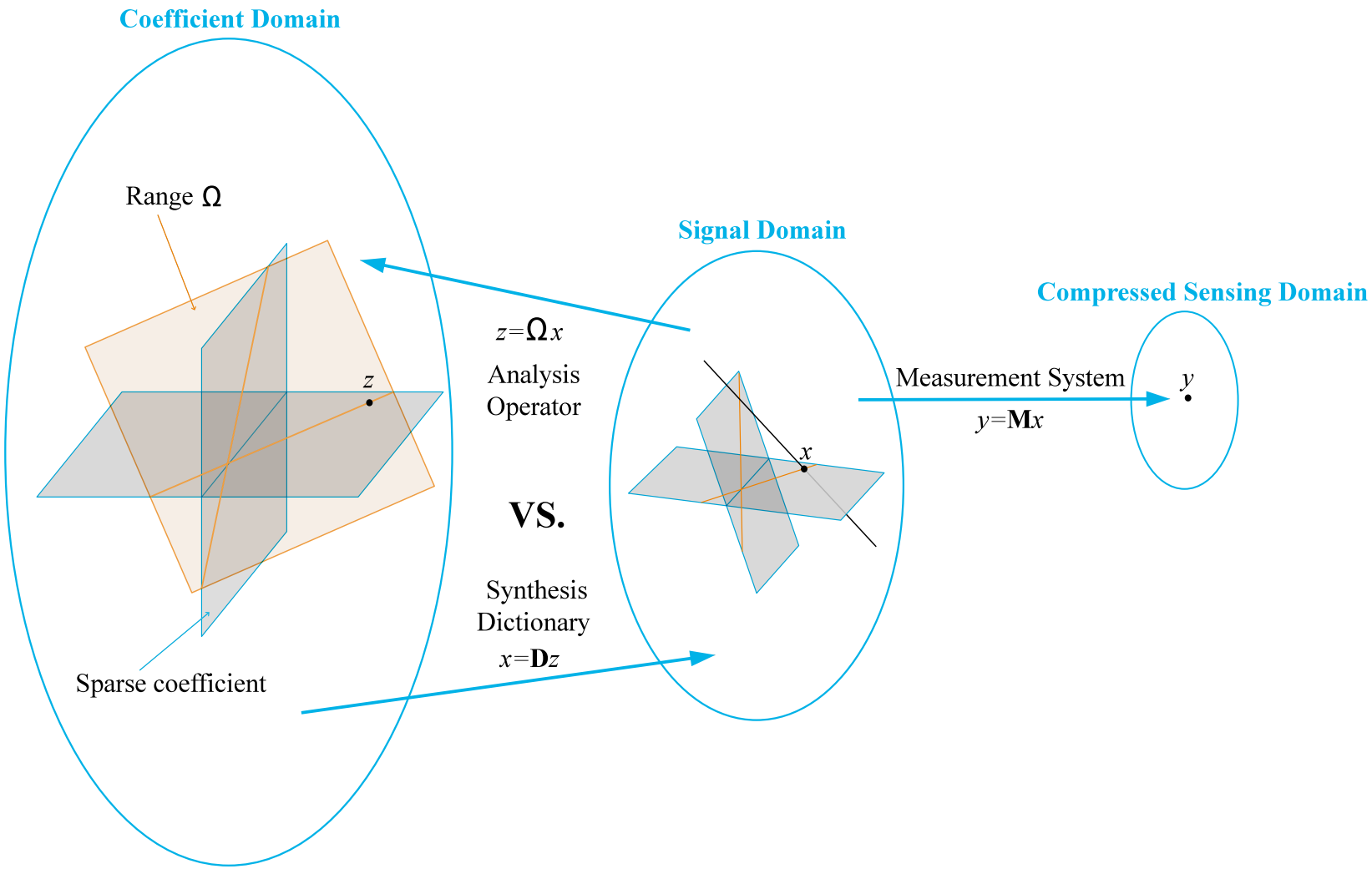
Cosparse models and inverse problems



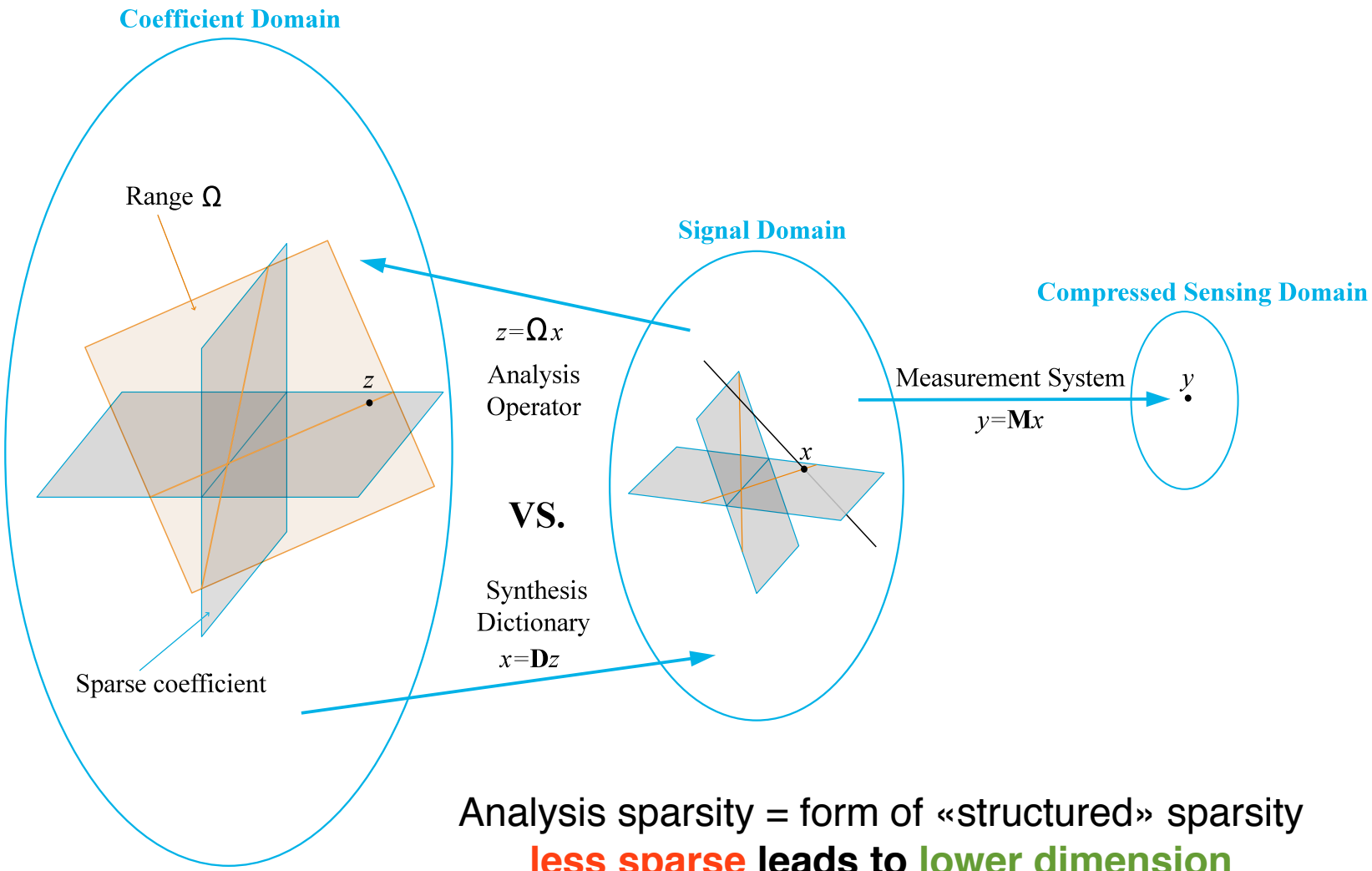
Cosparse models and inverse problems



Cosparse models and inverse problems



Cosparse models and inverse problems



Analysis sparsity = form of «structured» sparsity

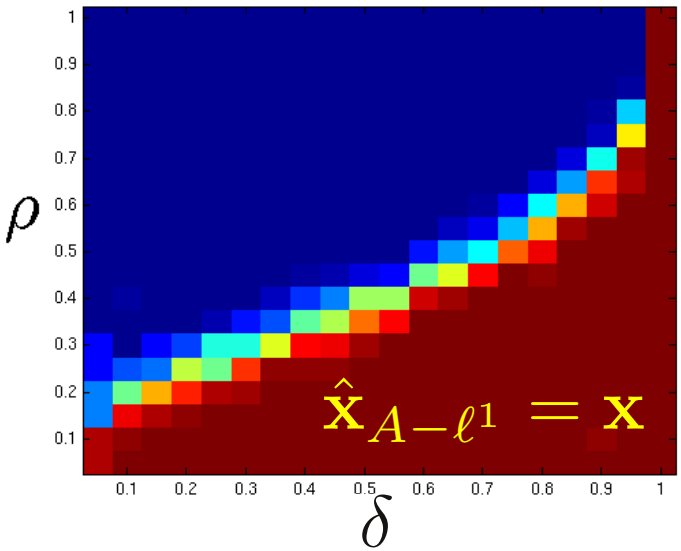
less sparse leads to lower dimension

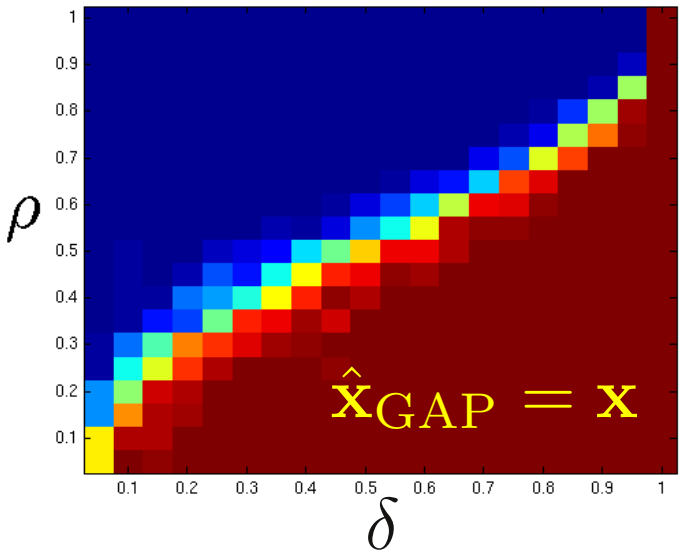
Optimization principles / algorithms

- Idealized problem $\hat{\mathbf{x}}_{A-L0} := \arg \min_{\mathbf{x}:\mathbf{y}=\mathbf{M}\mathbf{x}} \|\mathbf{\Omega}\mathbf{x}\|_0$
  - Starck & al 2003
  - Portilla 2009
  - Selesnick & Figueiredo 2009
  - Afonso, Bioucas-Dias & Figueiredo 2010 **(Session #9)**
- Convex relaxation $\hat{\mathbf{x}}_{A-L1} := \arg \min_{\mathbf{x}:\mathbf{y}=\mathbf{M}\mathbf{x}} \|\mathbf{\Omega}\mathbf{x}\|_1$
  - Elad & al 2007
  - Candès & al 2010
- Greedy analysis pursuit (GAP) ~ analysis-OMP
  - Nam & al 2011 **(Session #9)**
- Iterative cosparse projections ~ analysis-IHT
  - Gyries & al 2011

Results: Generic Analysis Operator

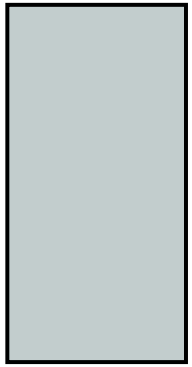
$$\mathbf{M} = \overset{m \times d}{\boxed{}} \qquad \delta = \frac{m}{d} \qquad \rho = \frac{d-\ell}{m} \qquad \mathbf{\Omega} \boxed{}$$

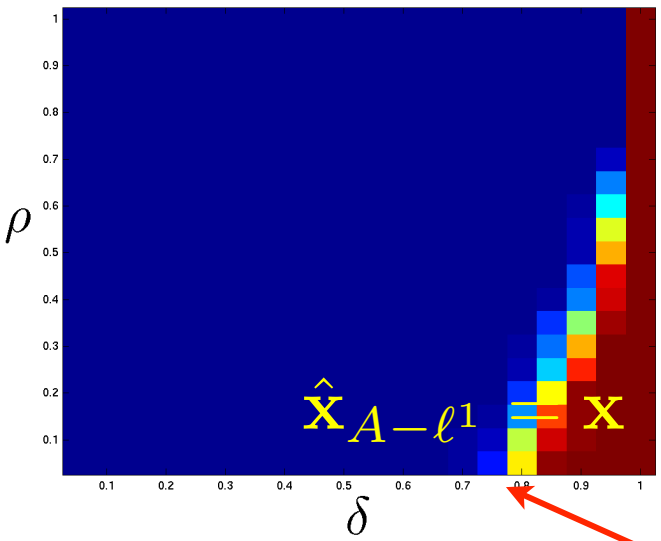


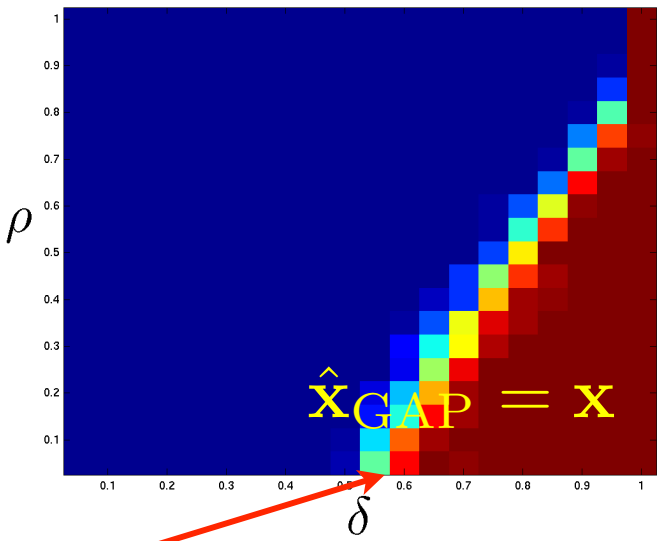


Results: Generic Analysis Operator

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







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

Co-sparsity *vs* Sparsity

- **Cosparsity**
  - ✓ operator

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

  - ✓ number of zeroes = co-dimension

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

  - ✓ dimension of subspace

    $$d - \ell$$

  - ✓ number of subspaces

    $$\binom{n}{\ell}$$

- **Sparsity**
  - ✓ dictionary

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

  - ✓ number of nonzeroes = dimension

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

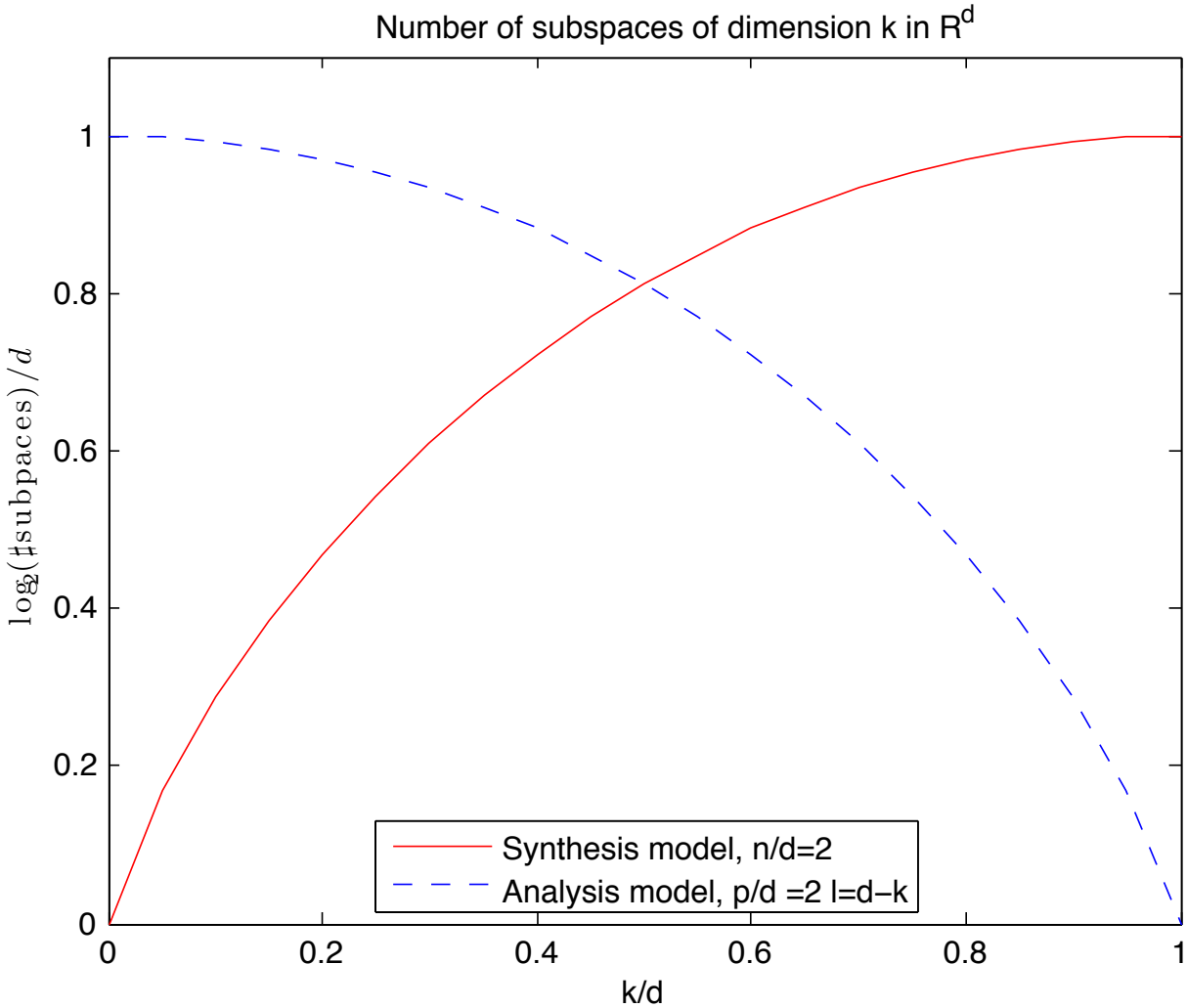
  - ✓ dimension of subspace

    $$k$$

  - ✓ number of subspaces

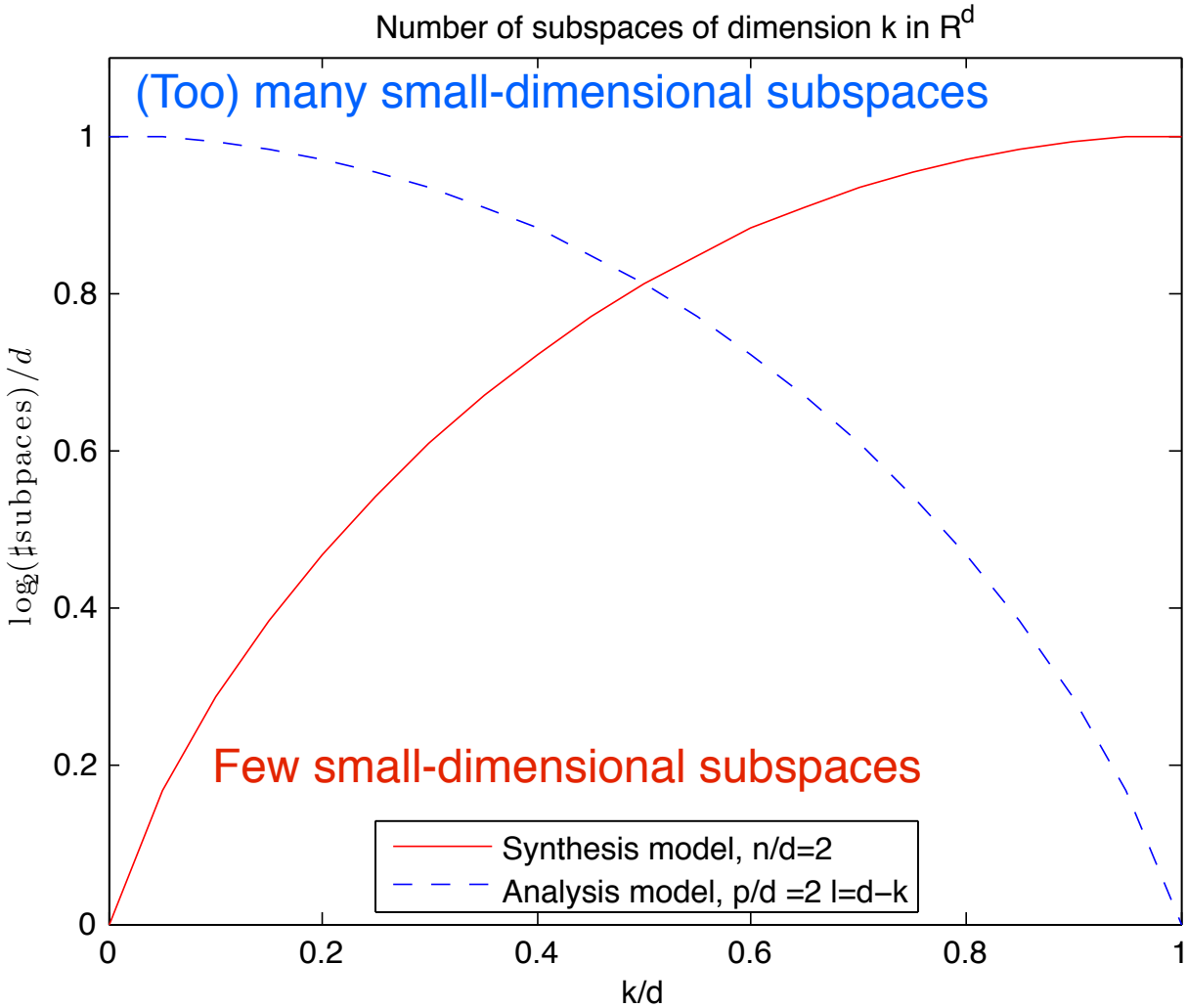
    $$\binom{n}{k}$$

Counting subspaces: Generic Operators



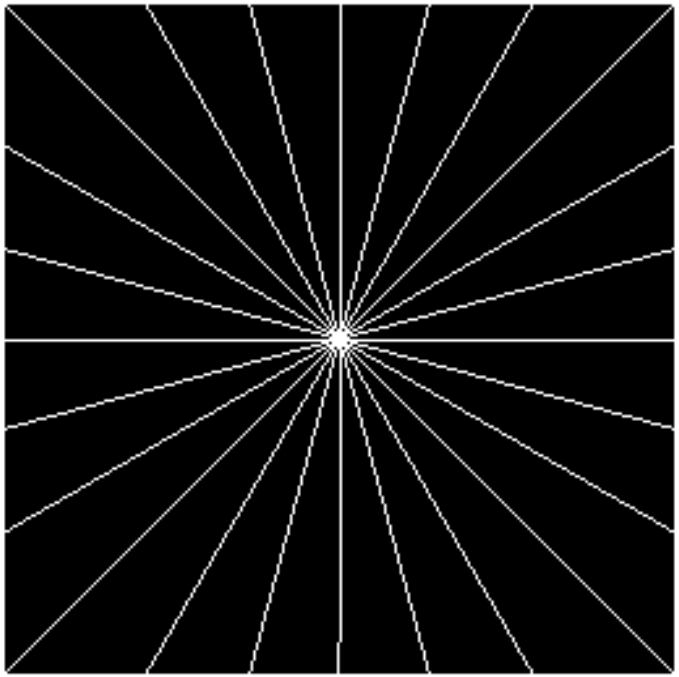
Remark:
sparse / cosparse models describe combinatorially many subspaces with only $n \times d$ parameters

Counting subspaces: Generic Operators



Remark:
sparse / cosparse models describe combinatorially many subspaces with only $n \times d$ parameters

Results: finite difference operator

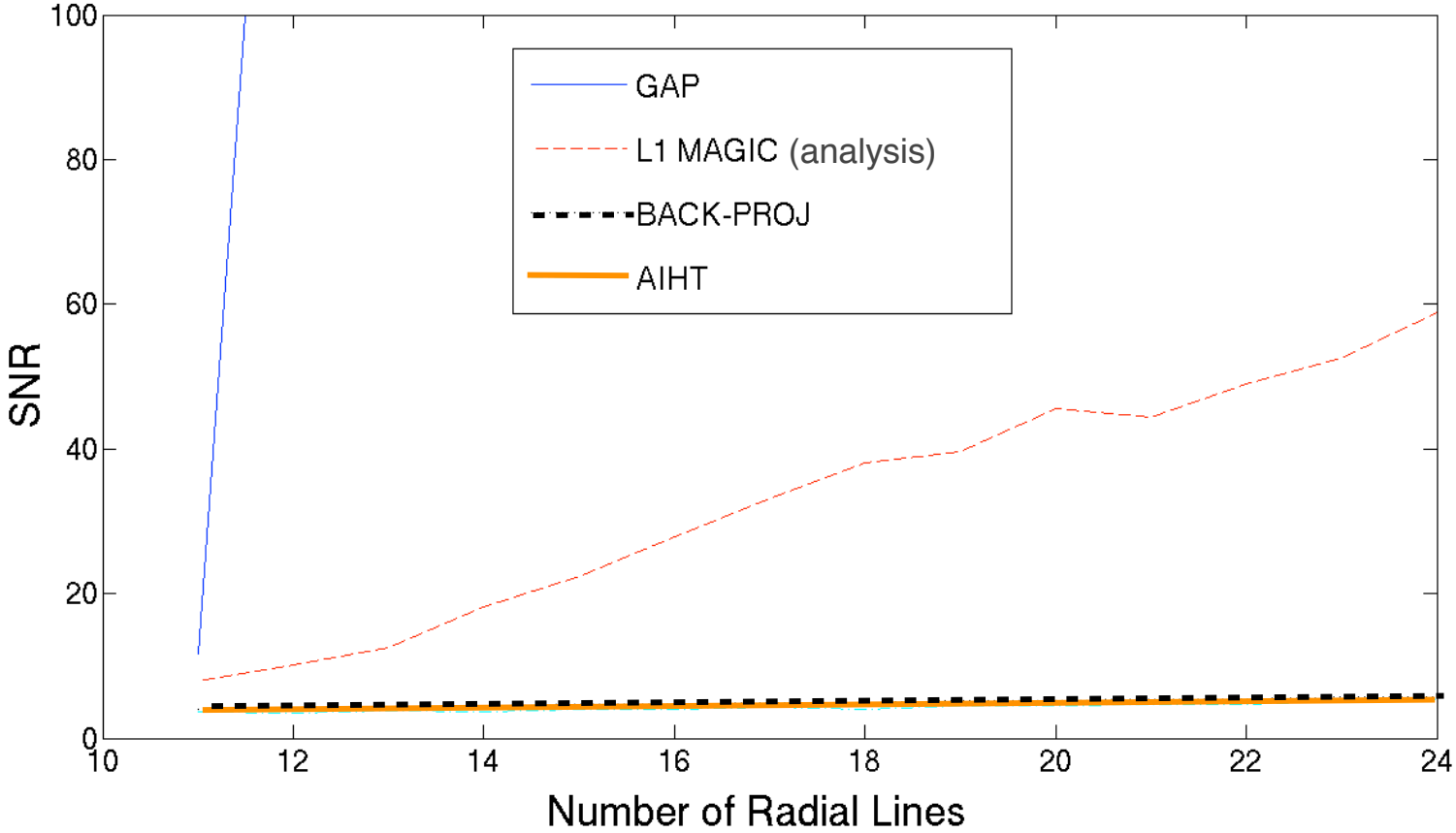


Sampling locations of Fourier transform of 256x256 image (4.63% of total)

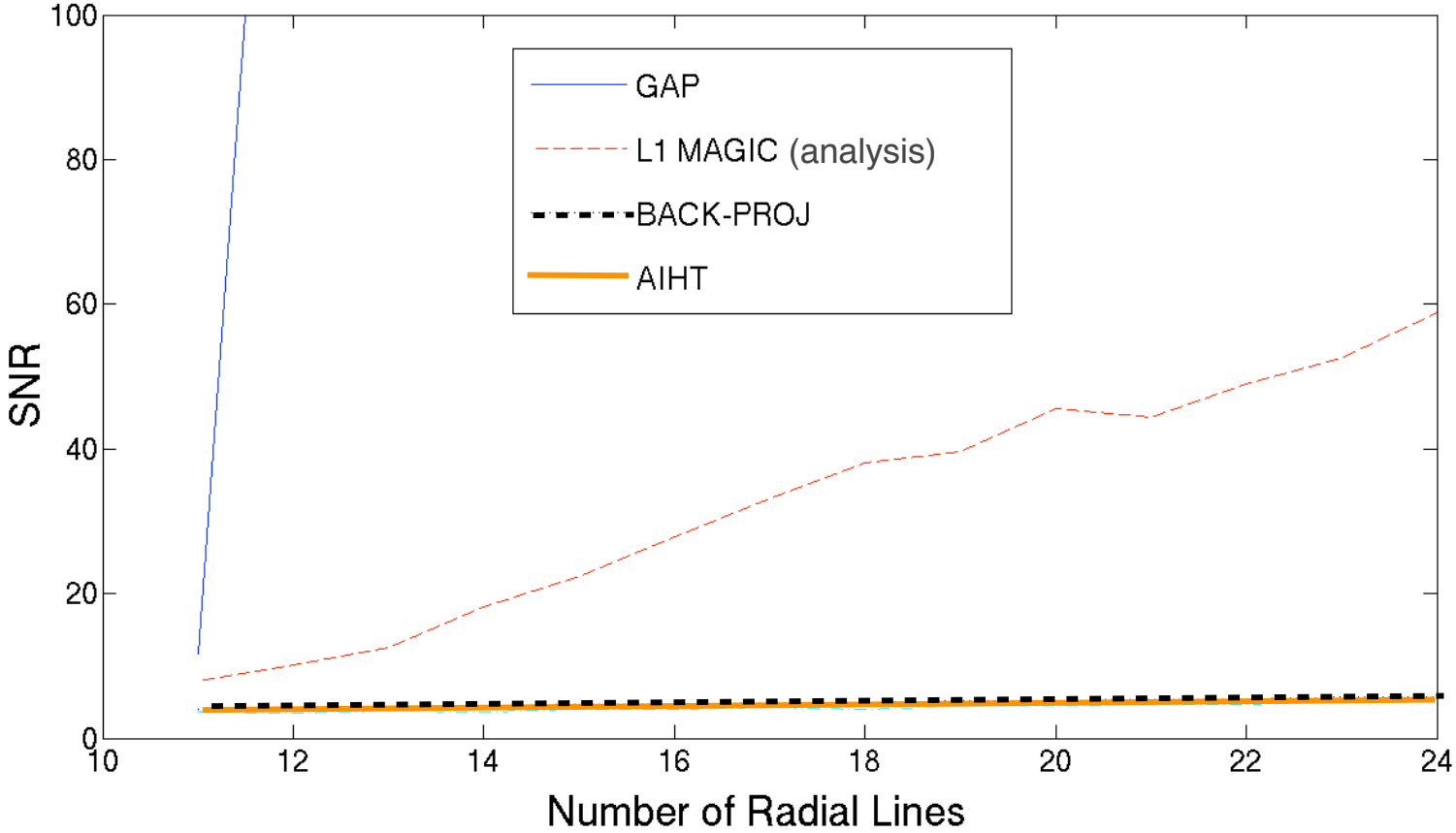
GAP

Recovered Image

Results with finite difference operator



Results with finite difference operator



Linear dependencies = fewer small-dimensional subspaces

Conclusions

Summary

- Traditional **Sparse 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

- **Cosparse Analysis Model**
  - ✓ **Analysis operator**

$\langle \omega_i, \mathbf{x} \rangle = 0$ for many indices

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

  - ✓ «Carving out» model: constraints

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

$$(\mathbf{\Delta}\mathbf{x})_{|\dot{\Omega}} = 0$$

Take-home message

- Revisited viewpoint on «transforms *vs* dictionaries»
  - ✓ Concept of *cosparsity,* contrasted with sparsity
    - ✦ Union of subspace models
    - ✦ Different relations between number of subspaces and dimension
    - ✦ Different role of linear dependencies: seem desirable for inverse problems
  - ✓ Co-sparse recovery guarantees with inverse problems
    - ✦ Cosparse model more naturally fits «sparse analysis» algorithms
    - ✦ New algorithms, recovery guarantees, empirically outperform analysis-L1
- References
  - ✦ G. & Nielsen. *Highly sparse representations from dictionaries are unique and independent of the sparseness measure*. ACHA, May 2007.
  - ✦ Nam, Davies, Elad & Gribonval. *The Cosparse Analysis Model and Algorithms*. arXiv:1106.4987, 2011. Submitted.

What's next ?

- @ SPARS 11

✓ Recovery guarantees for GAP & Analysis-L1
  - **(Nam & al, Session #9)**

✓ Learning/designing analysis operators
  - **(Rubinstein & Elad, Yaghoobi & al, Session #19)**
  - **Fadili & Peyré 2011, Ophir & al 2011**

✓ Hybrid sparse/cosparse models
  - **(Afonso & al, Session #9)**

What's next ?

@ SPARS ?

What's next ?

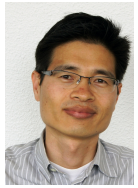
@ SPARS ?

What's next ?

(co)SPARS !

THANKS

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

- Design:
  - Jules Espiau (INRIA, France)

- Funding:
  - EU FET-Open

SMALL small-project.eu

- Join the team for a postdoc! remi.gribonval@inria.fr

(ERC StG 2011 «PLEASE»)

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.