

Analysis of the Basis Pursuit Algorithm and Applications*

Michael Elad

The Computer Science Department – (SCCM) program
Stanford University

Multiscale Geometric Analysis Meeting

IPAM - UCLA

January 2003

* Joint work with: Alfred M. Bruckstein – CS, Technion

David L. Donoho – Statistics, Stanford



Collaborators

Freddy
Bruckstein

CS
Department
Technion



M. Elad and A.M. Bruckstein, “A Generalized Uncertainty Principle and Sparse Representation in Pairs of Bases”, IEEE Trans. On Information Theory, Vol. 48, pp. 2558-2567, September 2002.



Dave
Donoho

Statistics
Department
Stanford

D. L. Donoho and M. Elad, “Maximal Sparsity Representation via l_1 Minimization”, to appear in Proceedings of the National Academy of Science.



General

- Basis Pursuit algorithm [Chen, Donoho and Saunders, 1995]:
 - Effective for finding sparse over-complete representations,
 - Effective for non-linear filtering of signals.
- Our work (in progress) – better understanding BP and deploying it in signal/image processing and computer vision applications.
- We believe that over-completeness has an important role!
- Today we discuss the analysis of the Basis Pursuit algorithm, giving conditions for its success. We then show some stylized applications exploiting this analysis.



Agenda

1. Introduction

Previous and current work

2. Two Ortho-Bases

Uncertainty \rightarrow Uniqueness \rightarrow Equivalence

3. Arbitrary dictionary

Uniqueness \rightarrow Equivalence

4. Stylized Applications

Separation of point, line, and plane clusters

5. Discussion

Understanding
the BP

Using the BP -
Application



Transforms

- Define the forward and backward transforms by (assume one-to-one mapping)

$$\text{Forward : } \underline{\alpha} = T\{\underline{s}\}$$

$$\text{Backward : } \underline{s} = T^{-1}\{\underline{\alpha}\}$$

\underline{s} – Signal (in the signal space C^N)

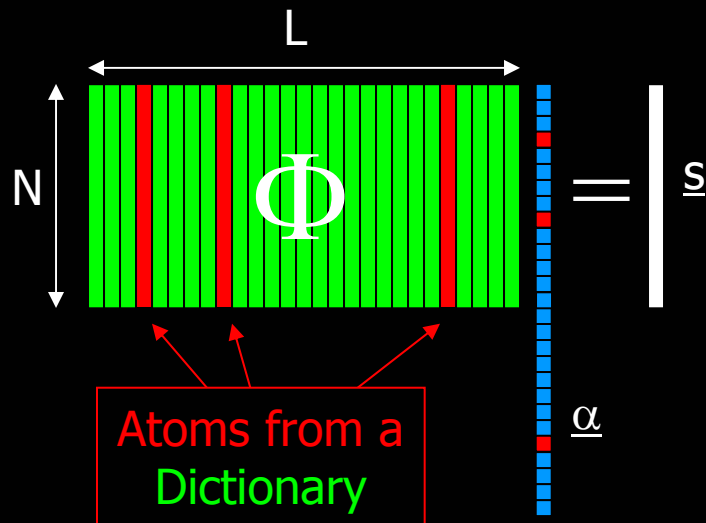
$\underline{\alpha}$ – Representation (in the transform domain C^L , $L \geq N$)

- Transforms T in signal and image processing used for coding, analysis, speed-up processing, feature extraction, filtering, ...

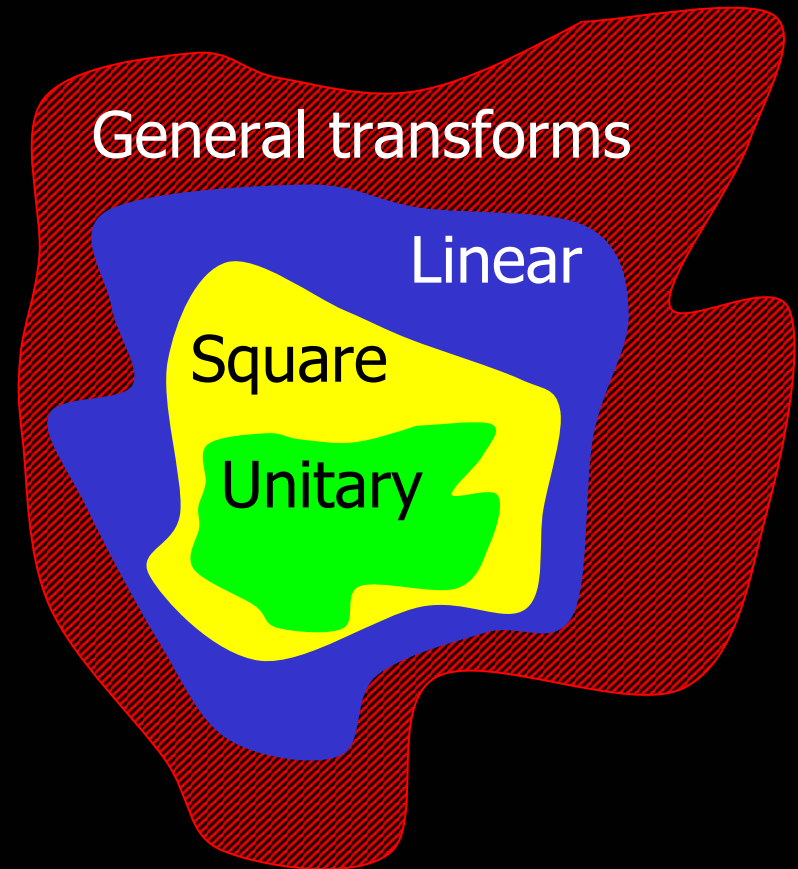


The Linear Transforms

- Special interest - linear transforms (inverse) $\underline{s} = \Phi \underline{\alpha}$



- In square linear transforms, Φ is an N -by- N & non-singular.

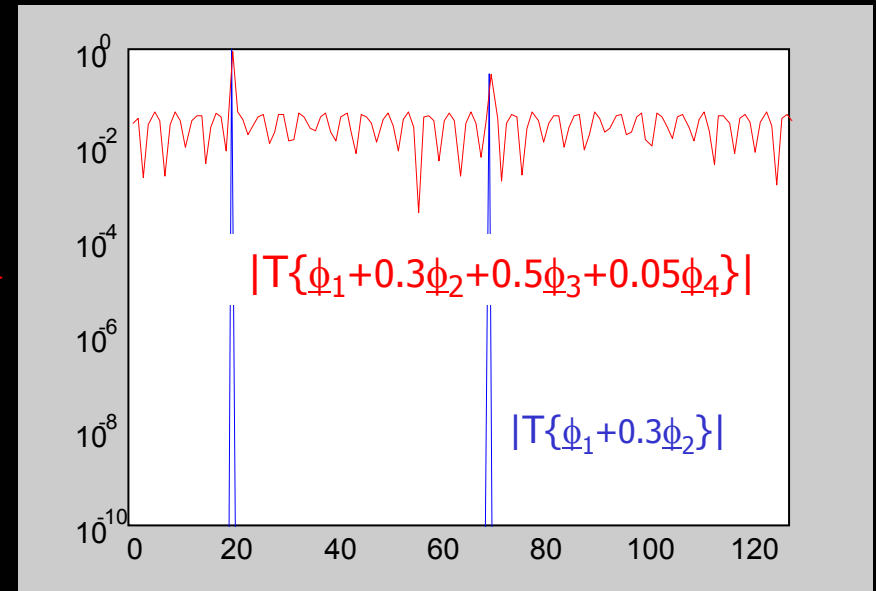
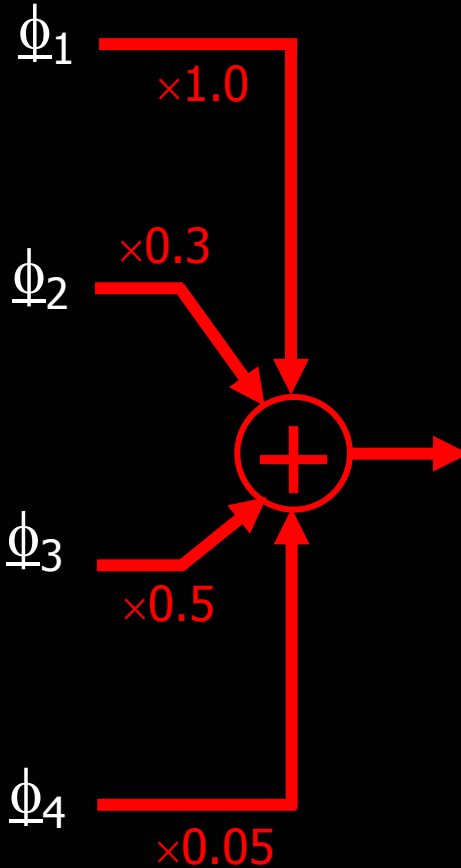
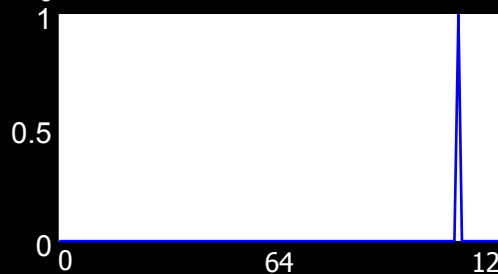
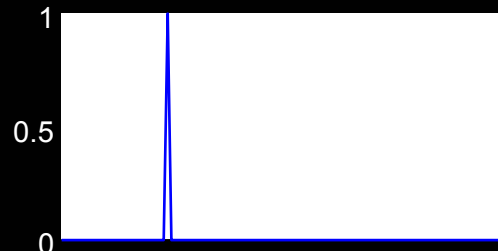
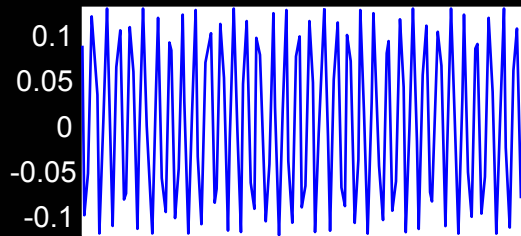
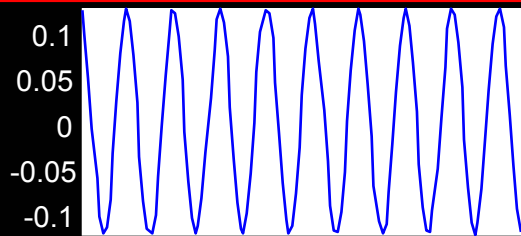


Lack Of Universality

- Many available square linear transforms – sinusoids, wavelets, packets, ...
- Successful transform – one which leads to sparse representations.
- Observation: Lack of universality - Different bases good for different purposes.
 - Sound = harmonic music (Fourier) + click noise (Wavelet),
 - Image = lines (Ridgelets) + points (Wavelets).
- Proposed solution: Over-complete dictionaries, and possibly **combination of bases**.



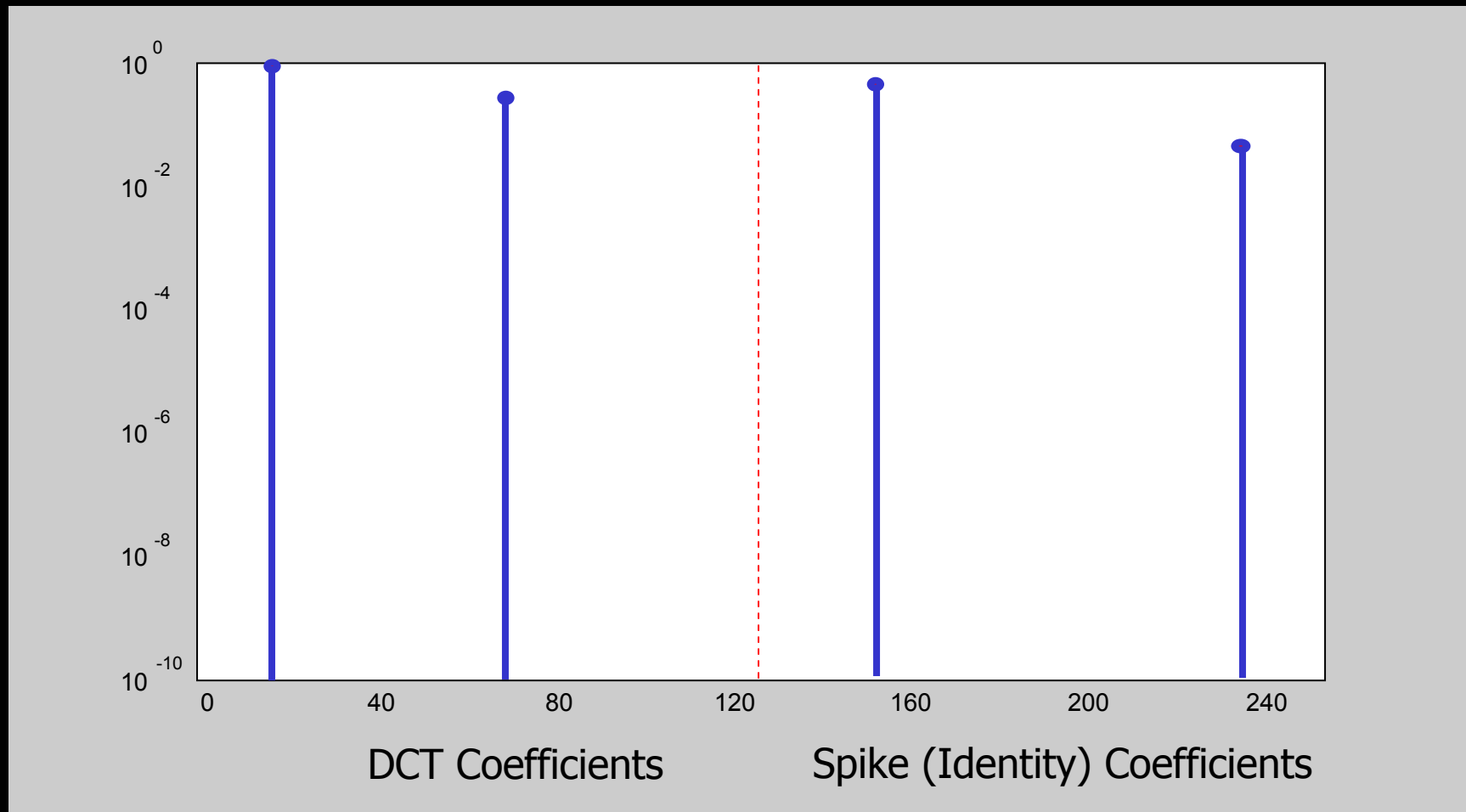
Example – Composed Signal



DCT Coefficients



Example – Desired Decomposition



Matching Pursuit

- Given d unitary matrices $\{\Phi_k, 1 \leq k \leq d\}$, define a dictionary $\Phi = [\Phi_1, \Phi_2, \dots, \Phi_d]$ [Mallat & Zhang (1993)].
- Combined representation per a signal \underline{s} by

$$\underline{s} = \Phi \underline{\alpha}$$

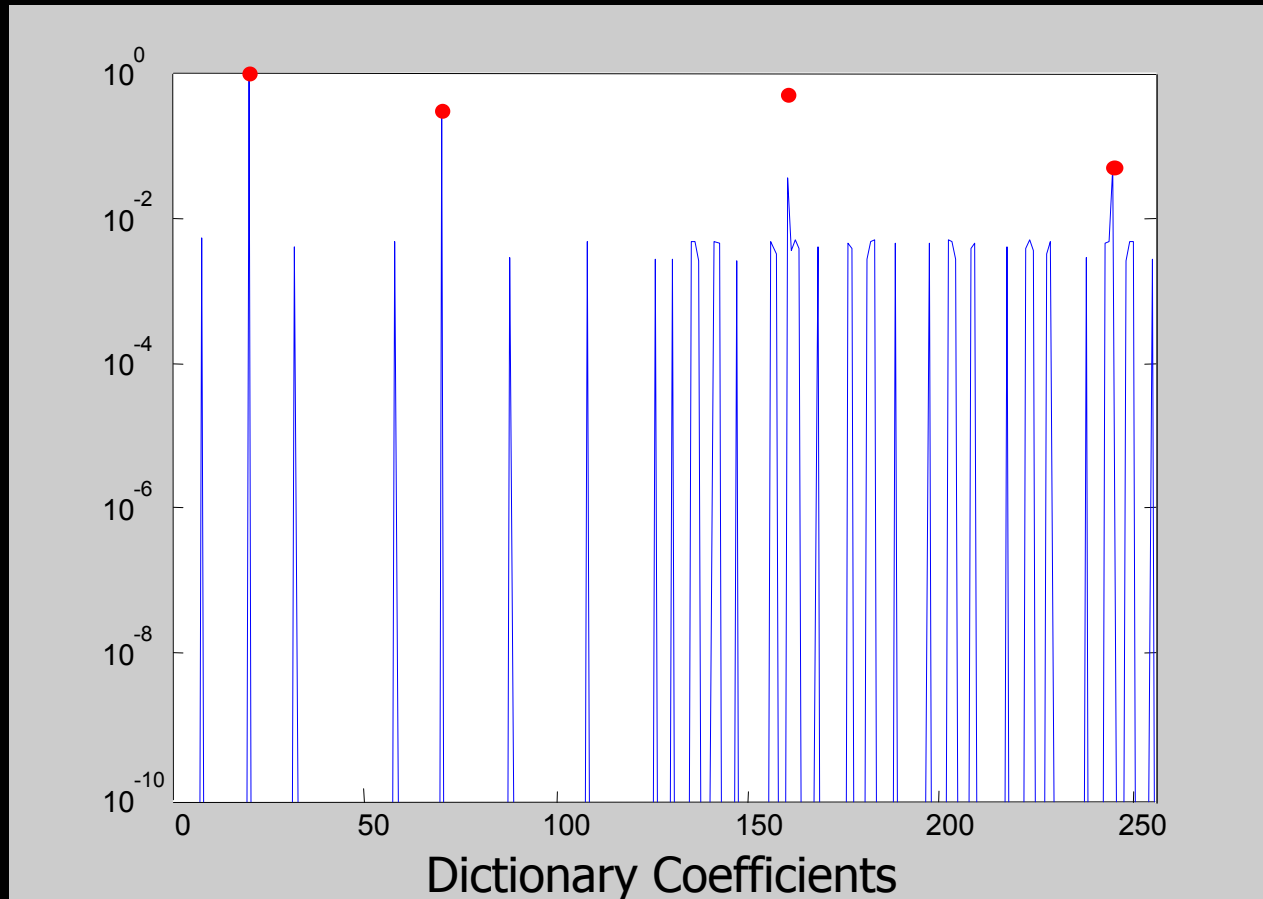
- Non-unique solution $\underline{\alpha}$ - Solve for maximal sparsity

$$P_0 : \text{Min}_{\underline{\alpha}} \|\underline{\alpha}\|_0 \text{ s.t. } \underline{s} = \Phi \underline{\alpha}$$

- Hard to solve – a sub-optimal greedy sequential solver: “Matching Pursuit algorithm” .



Example – Matching Pursuit



Basis Pursuit (BP)

- Facing the same problem, and the same optimization task [Chen, Donoho, Saunders (1995)]

$$P_0 : \underset{\underline{\alpha}}{\text{Min}} \|\underline{\alpha}\|_0 \quad \text{s.t.} \quad \underline{s} = \Phi \underline{\alpha}$$

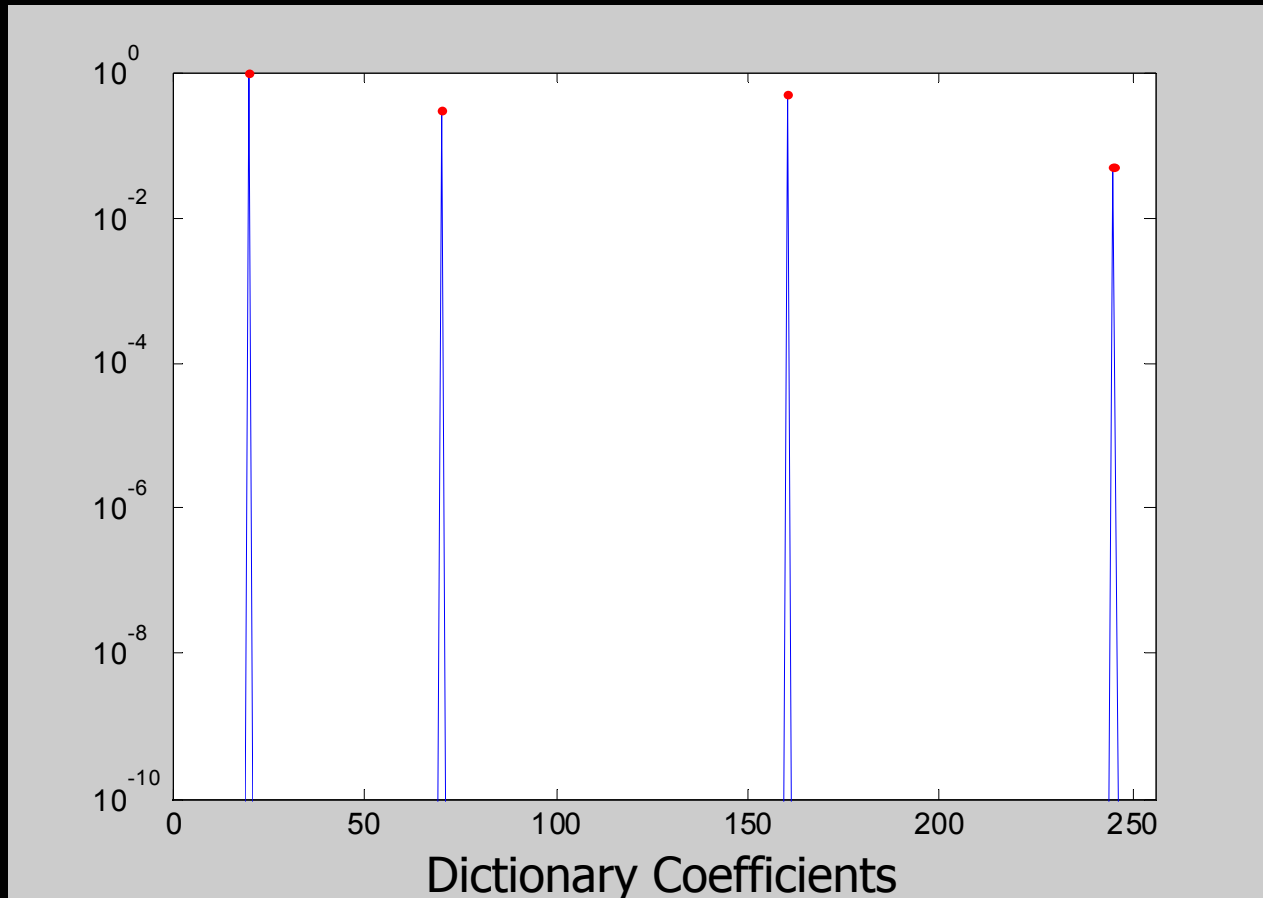
- Hard to solve – replace the l_0 norm by an l_1 :
“Basis Pursuit algorithm”

$$P_1 : \underset{\underline{\alpha}}{\text{Min}} \|\underline{\alpha}\|_1 \quad \text{s.t.} \quad \underline{s} = \Phi \underline{\alpha}$$

- **Interesting observation:** In many cases it successfully finds the sparsest representation.

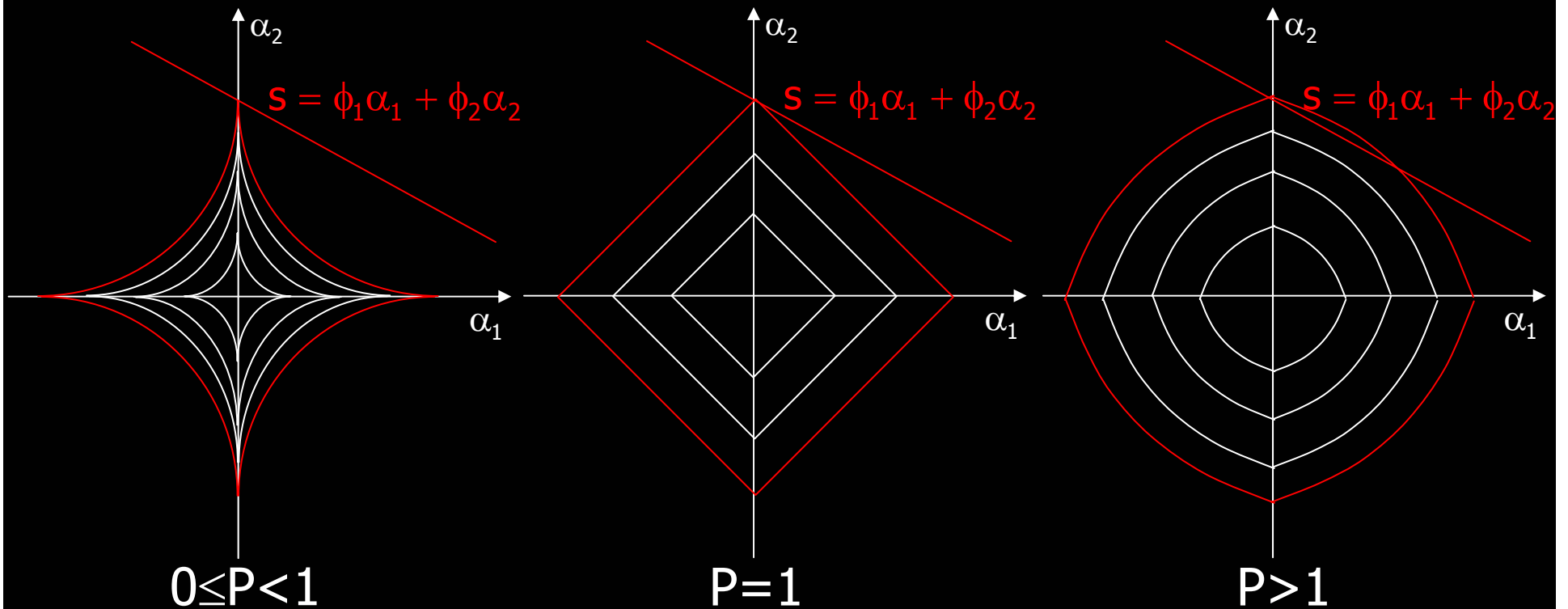


Example – Basis Pursuit



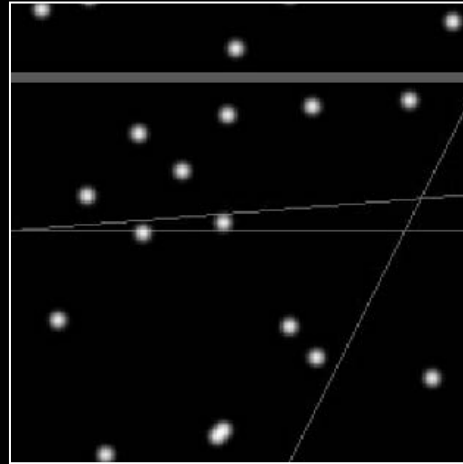
Why ℓ_1 ? 2D-Example

$$\text{Min}_{[\alpha_1, \alpha_2]} |\alpha_1|^p + |\alpha_2|^p \quad \text{s.t.} \quad \mathbf{s} = \phi_1 \alpha_1 + \phi_2 \alpha_2$$

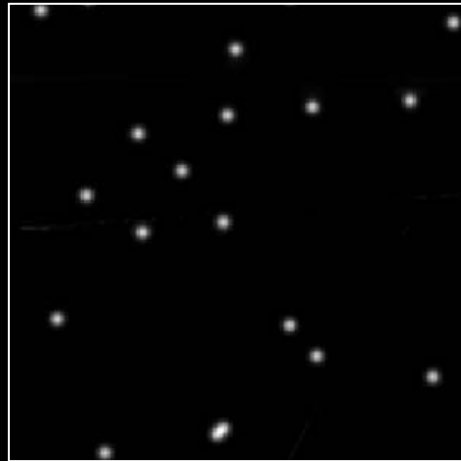


Example – Lines and Points*

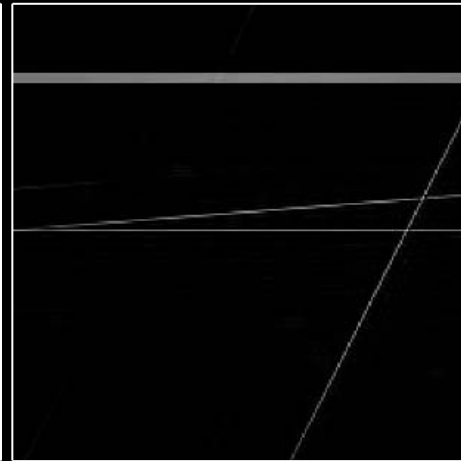
Original image



Wavelet part of the noisy image



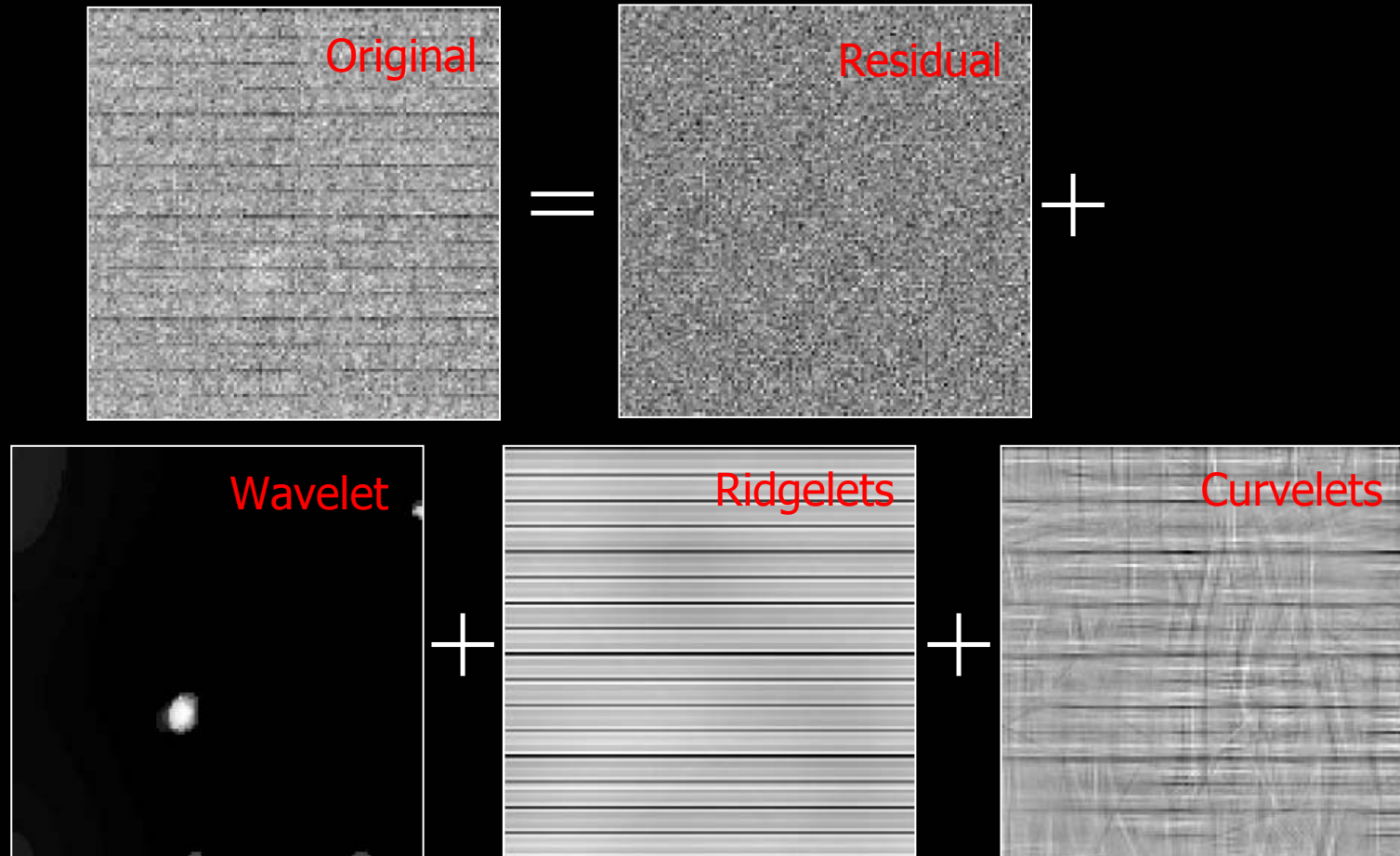
Ridgelets part of the image



* Experiments from [Starck, Donoho, and Candes - Astronomy & Astrophysics 2002.](#)



Example – Galaxy SBS 0335-052*



* Experiments from [Starck, Donoho, and Candes - Astronomy & Astrophysics 2002.](#)



Non-Linear Filtering via BP

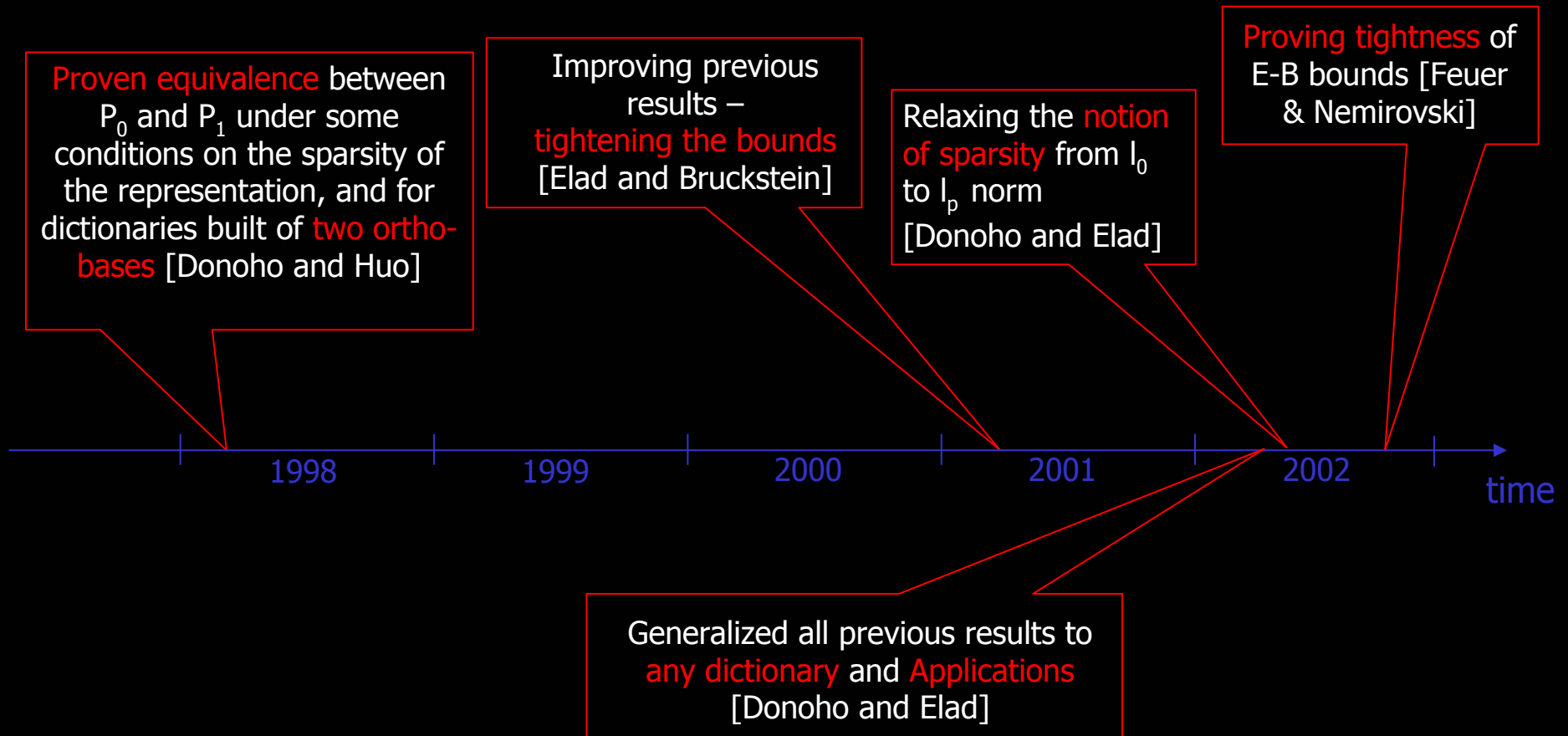
- Through the previous example – Basis Pursuit can be used for non-linear filtering.
- From Transforming to Filtering

$$\text{Min}_{\underline{\alpha}} \|\underline{\alpha}\|_1 \quad \text{s.t.} \quad \underline{s} = \Phi \underline{\alpha} \quad \rightarrow \quad \text{Min}_{\underline{\alpha}} \|\underline{\alpha}\|_1 + \lambda \|\underline{s} - \Phi \underline{\alpha}\|_2^2$$

- What is the relation to alternative non-linear filtering methods, such as PDE based methods (TV, anisotropic diffusion ...), Wavelet denoising?
- What is the role of over-completeness in inverse problems?



(Our) Recent Work



Before we dive ...

- Given a dictionary Φ and a signal \underline{s} , we want to find the sparse “atom decomposition” of the signal.

- Our goal is the solution of $\text{Min}_{\underline{\alpha}} \|\underline{\alpha}\|_0$ s.t. $\underline{s} = \Phi \underline{\alpha}$

- Basis Pursuit alternative is to solve instead

$$\text{Min}_{\underline{\alpha}} \|\underline{\alpha}\|_1 \quad \text{s.t.} \quad \underline{s} = \Phi \underline{\alpha}$$

- Our focus for now: Why should this work?



Agenda

1. Introduction

Previous and current work

2. Two Ortho-Bases

Uncertainty \rightarrow Uniqueness \rightarrow Equivalence

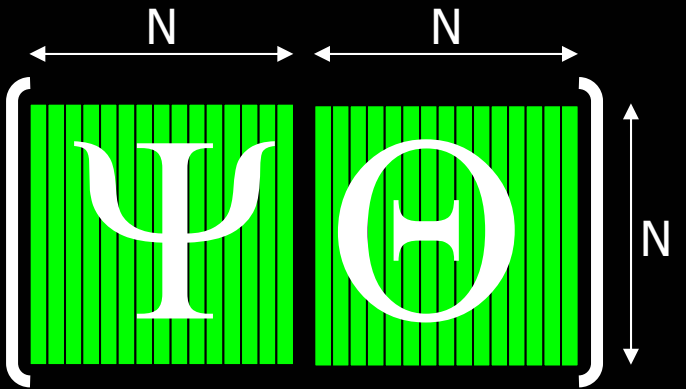
3. Arbitrariness

Uniqueness

4. Stylized

Separation

5. Discussion

$$\Phi = \begin{bmatrix} \Psi & \Theta \end{bmatrix}$$
The diagram shows the matrix Φ as a concatenation of two matrices, Ψ and Θ . Both Ψ and Θ are represented as green rectangles with vertical white lines. Above Ψ and Θ are double-headed arrows labeled 'N', indicating their widths. To the right of the entire structure is a vertical double-headed arrow labeled 'N', indicating the height of the matrix.



Our Objective

Given a signal \underline{s} , and its two representations using Ψ and Θ , what is the lower bound on the sparsity of both?

Our Objective is

$$\begin{array}{l} \underline{s} = \Psi \underline{\alpha} \\ \underline{s} = \Theta \underline{\beta} \end{array} \quad \rightarrow \quad \|\underline{\alpha}\|_0 + \|\underline{\beta}\|_0 \geq \text{Thr}(\Psi, \Theta)$$

We will show that such rule immediately leads to a practical result regarding the solution of the P_0 problem.



Mutual Incoherence

Define $M = \text{Max}_{1 \leq k, j \leq N} \left(\left| \underline{\Psi}_k^H \underline{\Theta}_j \right| \right)$

- M – mutual incoherence between Ψ and Θ .
- M plays an important role in the desired uncertainty rule.
- Properties
 - Generally, $1/\sqrt{N} \leq M \leq 1$.
 - For Fourier+Trivial (identity) matrices $M = 1/\sqrt{N}$.
 - For random pairs of ortho-matrices $M \approx 2\sqrt{\log_e N}/\sqrt{N}$.



Uncertainty Rule

Theorem 1

$$\|\underline{\alpha}\|_0 + \|\underline{\beta}\|_0 \geq 2\sqrt{\|\underline{\alpha}\|_0 \cdot \|\underline{\beta}\|_0} \geq \frac{2}{M}^*$$

Examples:

- $\Psi=\Theta$: $M=1$, leading to $\|\underline{\alpha}\|_0 + \|\underline{\beta}\|_0 \geq 2$.
- $\Psi=I$, $\Theta=F_N$ (DFT): $M = 1/\sqrt{N}$, leading to $\|\underline{\alpha}\|_0 + \|\underline{\beta}\|_0 \geq 2\sqrt{N}$.

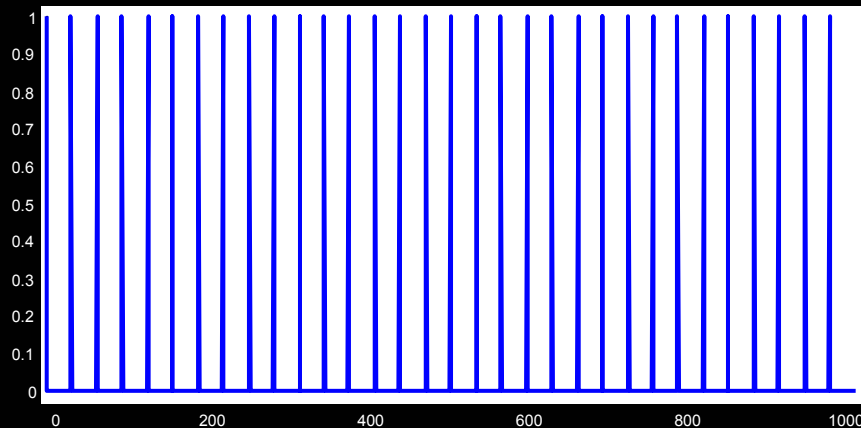
* Donoho & Huo obtained a weaker bound $\|\underline{\alpha}\|_0 + \|\underline{\beta}\|_0 \geq (1 + M^{-1})$



Example

$$\Psi = I, \Theta = F_N \text{ (DFT)} \rightarrow M = 1/\sqrt{N} \rightarrow \|\underline{\alpha}\|_0 + \|\underline{\beta}\|_0 \geq 2\sqrt{N}$$

- For $N=1024$, $\|\underline{s}\|_0 + \|\mathbf{F} \cdot \underline{s}\|_0 \geq 64$.
- The signal satisfying this bound: Picket-fence



Towards Uniqueness

- Given a unit norm signal \underline{s} , assume we hold two different representations for it using Φ

$$\underline{s} = \Phi \underline{\gamma}_1 = \Phi \underline{\gamma}_2$$

- Thus $\underline{0} = \Phi(\underline{\gamma}_1 - \underline{\gamma}_2) = [\Psi, \Theta] \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} \Rightarrow \Psi \underline{x}_1 = -\Theta \underline{x}_2 = \underline{q}$

- Based on the uncertainty theorem we just got:

$$\frac{2}{M} \leq \|\underline{x}_1\|_0 + \|\underline{x}_2\|_0 = \|\underline{\gamma}_1 - \underline{\gamma}_2\|_0 \leq \|\underline{\gamma}_1\|_0 + \|\underline{\gamma}_2\|_0$$



Uniqueness Rule

$$\frac{2}{M} \leq \|\underline{\gamma}_1\|_0 + \|\underline{\gamma}_2\|_0$$

In words: Any two different representations of the same signal CANNOT BE JOINTLY TOO SPARSE.

Theorem 2

If we found a representation that satisfy

$$\frac{1}{M} > \|\underline{\gamma}\|_0$$

Then necessarily it is unique (the sparsest).

*

* Donoho & Huo obtained a weaker bound $\|\underline{\gamma}\|_0 < 0.5(1 + M^{-1})$



Uniqueness Implication

- We are interested in solving

$$P_0 : \underset{\underline{\gamma}}{\text{Min}} \|\underline{\gamma}\|_0 \text{ s.t. } \underline{s} = [\Psi, \Theta]\underline{\gamma}.$$

- Somehow we obtain a candidate solution $\hat{\underline{\gamma}}$.
- The uniqueness theorem tells us that a simple test on $\hat{\underline{\gamma}}$ ($M \cdot \|\hat{\underline{\gamma}}\|_0 < 1$) could tell us if it is the solution of P_0 .
- However:
 - If the test is negative, it says nothing.
 - This does not help in solving P_0 .
 - This does not explain why P_1 may be a good replacement.



Equivalence - Goal

- We are going to solve the following problem

$$P_1 : \underset{\underline{\alpha}}{\text{Min}} \|\underline{\gamma}\|_1 \quad \text{s.t.} \quad \underline{s} = [\Psi, \Theta]\underline{\gamma}.$$

- The questions we ask are:
 - Will the P_1 solution coincide with the P_0 one?
 - What are the conditions for such success?
- We show that if indeed the P_0 solution is sparse **enough**, then P_1 solver finds it exactly.



Equivalence - Result

Given a signal \underline{s} with a representation $\underline{s} = [\Psi, \Theta] \underline{\gamma}$,

Assuming a sparsity on $\underline{\gamma}$ such that (assume $k_1 < k_2$)

$$\underline{\gamma} = \left[\underbrace{\gamma_1 \ \gamma_2 \ \cdots \ \gamma_N}_{k_1 \text{ non-zeros}}, \underbrace{\gamma_{N+1} \ \gamma_{N+2} \ \cdots \ \gamma_{2N}}_{k_2 \text{ non-zeros}} \right]$$

Theorem 3

If k_1 and k_2 satisfy $2M^2k_1k_2 + Mk_2 - 1 < 0$
then P_1 will find the correct solution.

A weaker requirement is given by $k_1 + k_2 < \frac{\sqrt{2}-0.5}{M}$ *

* Donoho & Huo obtained a weaker bound $\|\underline{\gamma}\|_0 < 0.5(1 + M^{-1})$



The Various Bounds

Signal dimension: $N=1024$,

Dictionary: $\Psi=I$, $\Theta=F_N$,

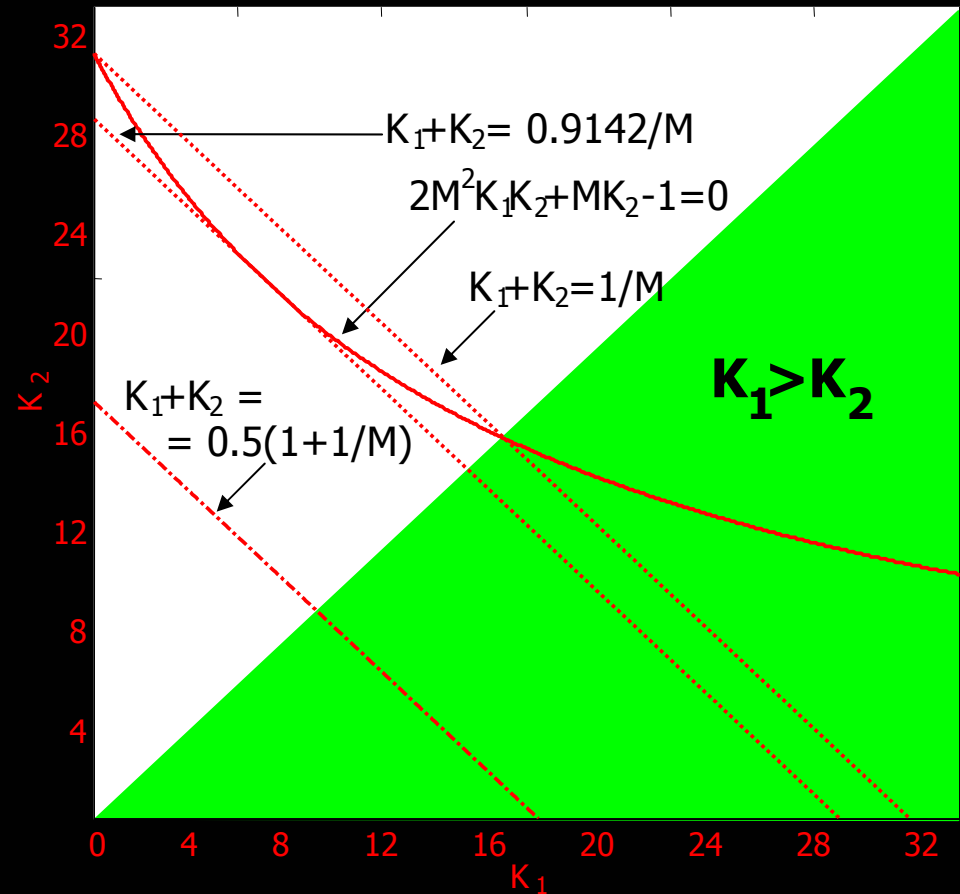
Mutual incoherence $M=1/32$.

Results

Uniqueness: 31 entries and below,

Equivalence:

- 16 entries and below (D-H),
- 29 entries and below (E-B).



Equivalence – Uniqueness Gap

- For uniqueness we got the requirement $\|\underline{\gamma}\|_0 < \frac{1}{M}$
- For equivalence we got the requirement $\|\underline{\gamma}\|_0 < \frac{\sqrt{2}-0.5}{M}$
- Is this gap due to careless bounding?
- Answer [by Feuer and Nemirovski, to appear in *IEEE Transactions On Information Theory*]: No, both bounds are indeed tight.



Agenda

1. Introduction

Previous and current work

2. Two Ortho-Bases

Uncertainty \rightarrow Uniqueness \rightarrow B

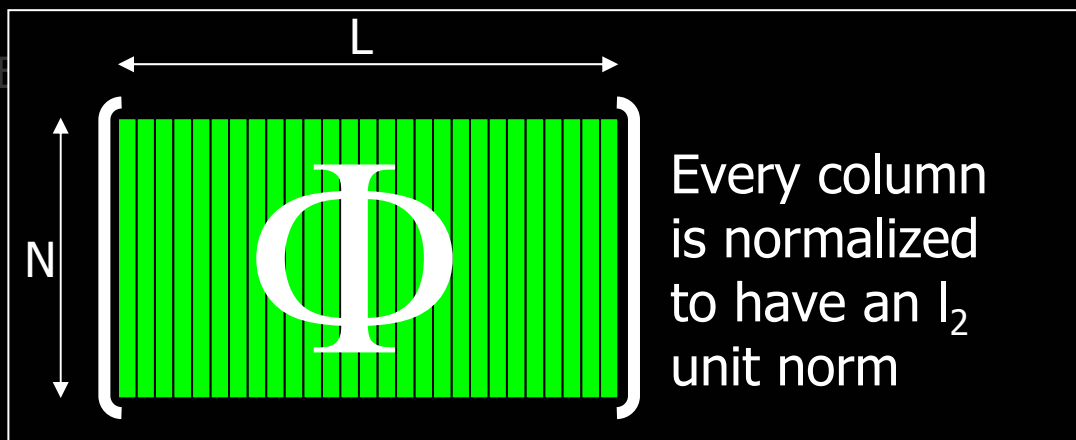
3. Arbitrary dictionary

Uniqueness \rightarrow Equivalence

4. Stylized Applications

Separation of point, line, and plane clusters

5. Discussion



Why General Dictionaries?

- In many situations
 - We would like to use **more than just two ortho-bases** (e.g. Wavelet, Fourier, and ridgelets);
 - We would like to use **non-ortho bases** (pseudo-polar FFT, Gabor transform, ...),
 - In many situations we would like to use **non-square transforms** as our building blocks (Laplacian pyramid, shift-invariant Wavelet, ...).
- In the following analysis we assume **ARBITRARY DICTIONARY** (frame). We show that BP is successful over such dictionaries as well.

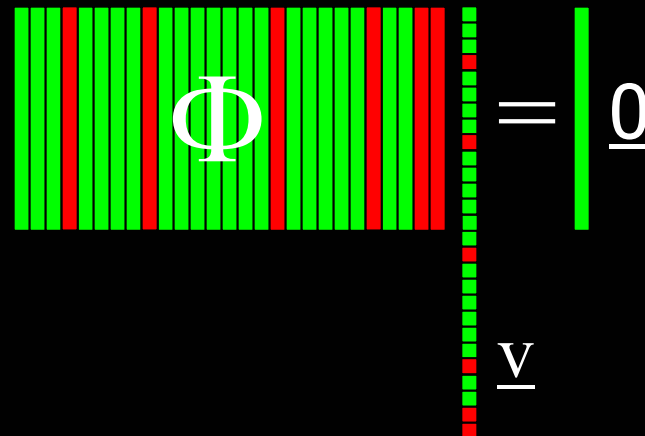


Uniqueness - Basics

- Given a unit norm signal \underline{s} , assume we hold two different representations for it using Φ

$$\underline{s} = \Phi \underline{\gamma}_1 = \Phi \underline{\gamma}_2 \Rightarrow \Phi (\underline{\gamma}_1 - \underline{\gamma}_2) = \underline{0}$$

- In the two-ortho case - simple splitting and use of the uncertainty rule – here there is no such splitting !!
- The equation $\Phi \underline{v} = \underline{0}$ implies a linear combination of columns from Φ that are linearly dependent. What is the smallest such group?



The diagram shows a matrix Φ represented by a grid of vertical bars. The bars are colored green and red. The matrix is multiplied by a vector \underline{v} , which is also represented by a vertical bar with green and red segments. The result is the zero vector $\underline{0}$.



Uniqueness – Matrix “Spark”

Definition: Given a matrix Φ , define $\sigma = \text{Spark}\{\Phi\}$ as the smallest integer such that there exists at least one group of σ columns from Φ that is linearly dependent.

Examples:

$$\text{Spark} \left\{ \begin{bmatrix} 1 & 0 & \dots & 0 & | & 1 \\ 0 & 1 & \dots & 0 & | & 1 \\ \vdots & \vdots & \ddots & \vdots & | & \vdots \\ 0 & 0 & \dots & 1 & | & 1 \end{bmatrix} \right\} = N+1; \quad \text{Spark} \left\{ \begin{bmatrix} 1 & 0 & \dots & 0 & | & 1 \\ 0 & 1 & \dots & 0 & | & 0 \\ \vdots & \vdots & \ddots & \vdots & | & \vdots \\ 0 & 0 & \dots & 1 & | & 0 \end{bmatrix} \right\} = 2$$



“Spark” versus “Rank”

The notion of spark is confusing – here is an attempt to compare it to the notion of rank

Rank	Spark
Definition: Maximal # of columns that are linearly independent	Definition: Minimal # of columns that are linearly dependent
Computation: Sequential - Take the first column, and add one column at a time, performing Gram-Schmidt orthogonalization. After L steps, count the number of non-zero vectors – This is the rank.	Computation: Combinatorial - sweep through 2^L combinations of columns to check linear dependence - the smallest group of linearly dependent vectors is the Spark.

Generally: $2 \leq \sigma = \text{Spark}\{\Phi\} \leq \text{Rank}\{\Phi\} + 1.$



Uniqueness – Using the “Spark”

- Assume that we know the spark of Φ , denoted by σ .
- For any pair of representations of \underline{s} we have

$$\underline{s} = \Phi \underline{\gamma}_1 = \Phi \underline{\gamma}_2 \Rightarrow \Phi(\underline{\gamma}_1 - \underline{\gamma}_2) = \underline{0}$$

- By the definition of the spark we know that if $\Phi \underline{v} = \underline{0}$ then $\|\underline{v}\|_0 \geq \sigma$. Thus

$$\|\underline{\gamma}_1 - \underline{\gamma}_2\|_0 \geq \sigma$$

- From here we obtain the relationship

$$\sigma \leq \|\underline{\gamma}_1 - \underline{\gamma}_2\|_0 \leq \|\underline{\gamma}_1\|_0 + \|\underline{\gamma}_2\|_0$$



Uniqueness Rule – 1

$$\sigma \leq \|\gamma_1\|_0 + \|\gamma_2\|_0$$

Any two different representations of the same signal using an **arbitrary dictionary** cannot be jointly sparse.

Theorem 4

If we found a representation that satisfy

$$\frac{\sigma}{2} > \|\gamma\|_0$$

Then necessarily it is unique (the sparsest).



Lower bound on the “Spark”

- Define $0(?) < M = \text{Max}_{\substack{1 \leq k, j \leq L \\ k \neq j}} \left\{ \left| \phi_k^H \phi_j \right| \right\} \leq 1$

(notice the resemblance to the previous definition of M).

- We can show (based on Geršgorin disks theorem) that a lower-bound on the spark is obtained by

$$\sigma \geq 1 + \frac{1}{M}.$$

- Since the Geršgorin theorem is un-tight, this lower bound on the Spark is too pessimistic.



Uniqueness Rule – 2

$$1 + \frac{1}{M} \leq \sigma \leq \|\gamma_1\|_0 + \|\gamma_2\|_0$$

Any two different representations of the same signal using an **arbitrary dictionary** cannot be jointly sparse.

Theorem 5

If we found a representation that satisfy *

$$\frac{\sigma}{2} \geq \frac{1}{2} \left(1 + \frac{1}{M} \right) > \|\gamma\|_0$$

Then necessarily it is unique (the sparsest).

* This is the same as Donoho and Huo's bound! Have we lost tightness?



"Spark" Upper bound

- The Spark can be found by solving

$$\left\{ S_k : \underset{\underline{\gamma}}{\text{Min}} \|\underline{\gamma}\|_0 \quad \text{s.t.} \quad \Phi \underline{\gamma} = \underline{0} \quad \& \quad \gamma_k = 1 \right\}_{k=1}^L \quad \longrightarrow \quad \left\{ \underline{\gamma}_{-k}^S \right\}_{k=1}^L$$

$$\longrightarrow \quad \sigma = \underset{1 \leq k \leq L}{\text{Min}} \left\| \underline{\gamma}_{-k}^S \right\|_0$$

- Use Basis Pursuit

$$\left\{ Q_k : \underset{\underline{\gamma}}{\text{Min}} \|\underline{\gamma}\|_1 \quad \text{s.t.} \quad \Phi \underline{\gamma} = \underline{0} \quad \& \quad \gamma_k = 1 \right\}_{k=1}^L \quad \longrightarrow \quad \left\{ \underline{\gamma}_{-k}^Q \right\}_{k=1}^L$$

- Clearly $\left\| \underline{\gamma}_{-k}^Q \right\|_0 \geq \left\| \underline{\gamma}_{-k}^S \right\|_0$. Thus $\sigma = \underset{1 \leq k \leq L}{\text{Min}} \left\| \underline{\gamma}_{-k}^S \right\|_0 \leq \underset{1 \leq k \leq L}{\text{Min}} \left\| \underline{\gamma}_{-k}^Q \right\|_0$.



Equivalence – The Result

Following the same path as shown before for the equivalence theorem in the two-ortho case, and adopting the new definition of M we obtain the following result:

Theorem 6

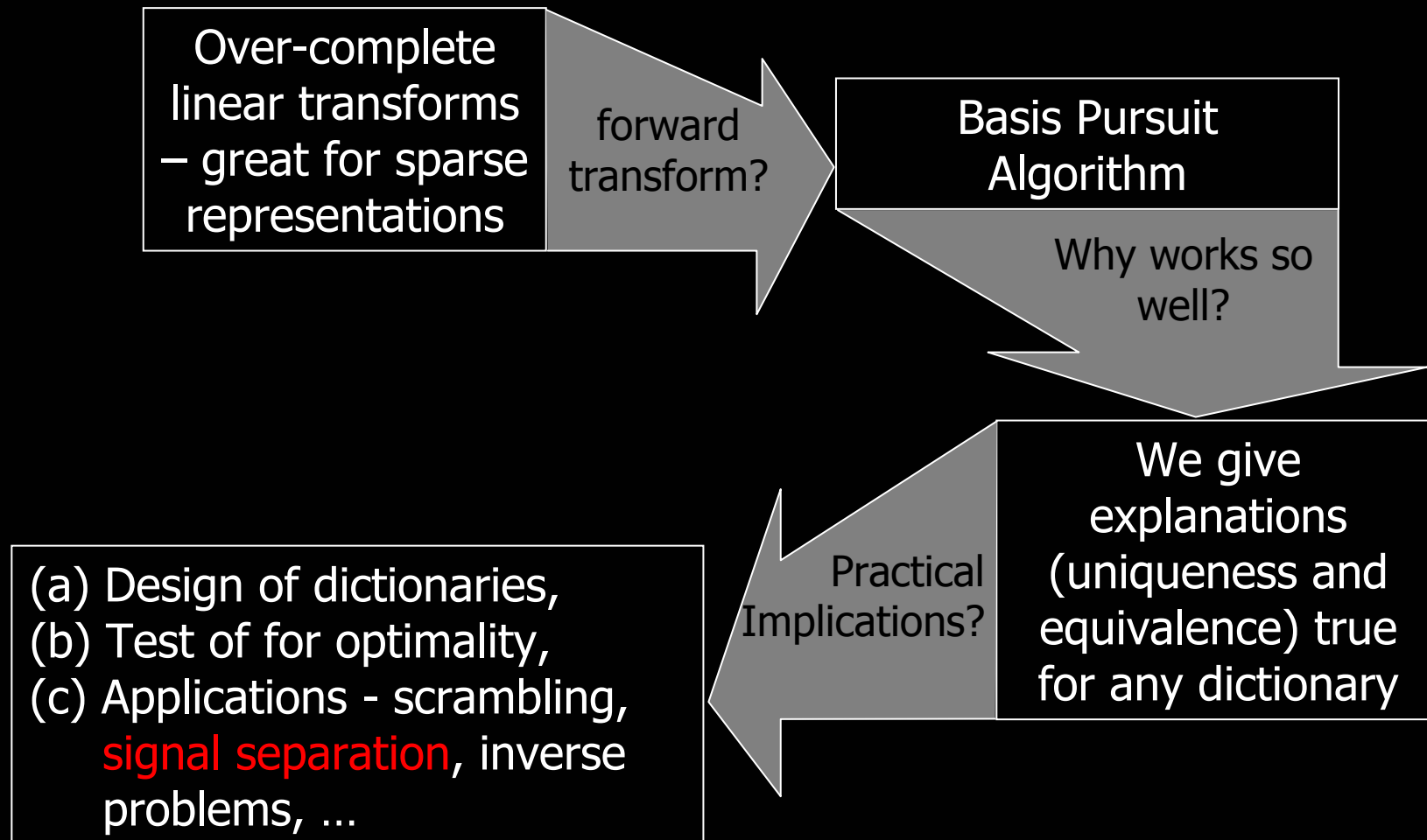
Given a signal \underline{s} with a representation $\underline{s} = \Phi \underline{\gamma}$,
Assuming that $\|\underline{\gamma}\|_0 < 0.5(1 + 1/M)$, P_1 (BP) is
Guaranteed to find the sparsest solution.

*

* This is the same as Donoho and Huo's bound! Is it non-tight?



To Summarize so far ...



Agenda

1. Introduction

Previous and current work

2. Two Ortho-Bases

Uncertainty \rightarrow Uniqueness \rightarrow Equivalence

3. Arbitrary dictionary

Uniqueness \rightarrow Equivalence

4. Stylized Applications

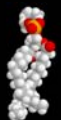
Separation of point, line, and plane clusters

5. Discussion



Problem Definition

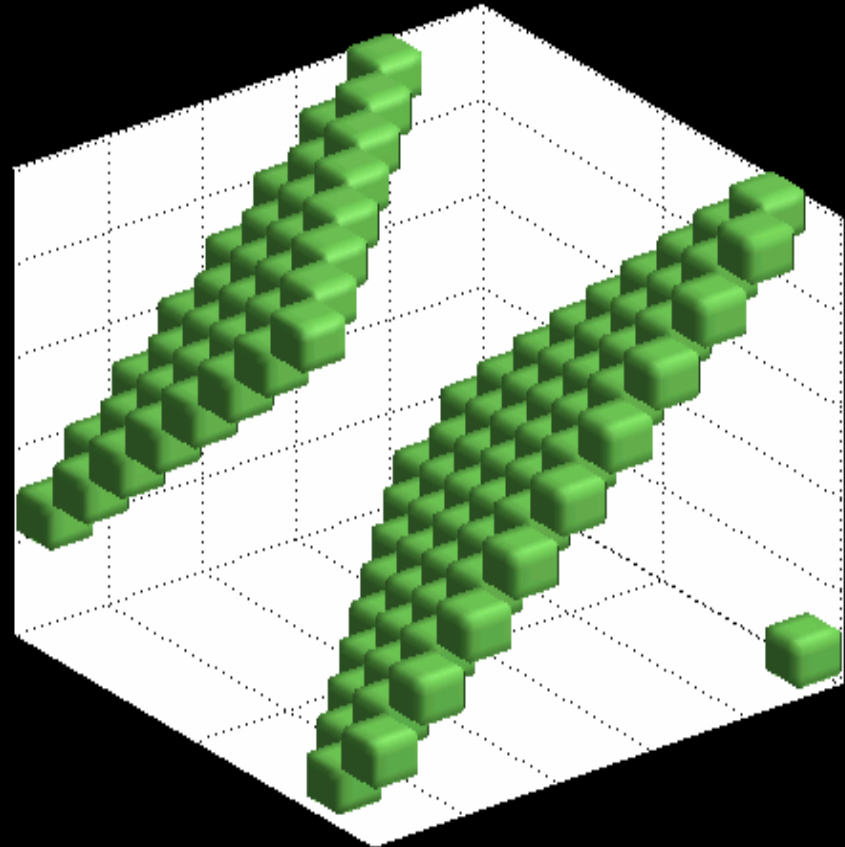
- Assume a 3D volume $S[n_1, n_2, n_3] = S[\underline{n}]$ of size $p \times p \times p$ Voxels.
- S contains **digital** points, lines, and planes, defined **algebraically** (\mathbf{z}_p^3):
 - Point – one Voxel
 - Line – p Voxels (defined by a starting Voxel on the cube's face, and a slope – wrap around is permitted),
 - Plane – p^2 Voxels (defined by a height Voxel in the cube, and 2 slope parameters – wrap around is permitted).
- Task: Decompose S into the point, line, and plane atoms – sparse decomposition is desired.



Properties To Use

Digital lines and planes are constructed such that:

- Any two lines are disjoint, or intersect in a single Voxel.
- Any two planes are disjoint, or intersect in a line of p Voxels.
- Any digital line intersect a with a digital plane not at all, in a single Voxel, or along the p Voxels being the line itself.



Our Dictionary

- We construct a dictionary to contain all occurrences of points, lines, and planes ($O(p^4)$ columns).
- Every atom in our set is reorganized lexicographically as a column vector of size p^3 -by-1.
- The spark of this dictionary is $\sigma=p+1$ (p points and one line, or p lines and one plane).
- The mutual incoherence is given by $M=p^{-0.5}$ (normalized inner product of a plane with a line, or line with a point).

$$\langle \text{Line, Point} \rangle = \frac{1}{1 \cdot \sqrt{p}}$$



Thus we can say that ...

Due to
Theorem 4

If we found a representation that satisfies

$$\|\underline{\gamma}\|_0 < \frac{p+1}{2}$$

then necessarily it is sparsest (P_0 solution).

Due to
Theorem 6

Given a volume \underline{s} with a representation
 $\underline{s} = \Phi \underline{\gamma}$ such that $\|\underline{\gamma}\|_0 < 0.5(1 + \sqrt{p})$, (BP)
is guaranteed to find it.



Implications

- Assume $p=1024$:
 - A representation with fewer than 513 atoms IS KNOWN TO BE THE SPARSEST.
 - A representation with fewer than 17 atoms WILL BE FOUND BY THE BP.
- What happens in between? We know that the second bound is very loose! BP will succeed far above 17 atoms. Further work is needed to establish this fact.
- What happens above 513 atoms? No theoretical answer yet!



Towards Practical Application

- The algebraic definition of the points, lines, and planes atoms was chosen to enable the computation of σ and M .
- Although wrap around makes this definition non-intuitive, these results are suggestive of what may be the case for geometrically intuitive definitions of points, lines, and planes.
- Further work is required to establish this belief.



Agenda

1. Introduction

Previous and current work

2. Two Ortho-Bases

Uncertainty \rightarrow Uniqueness \rightarrow Equivalence

3. Arbitrary dictionary

Uniqueness \rightarrow Equivalence

4. Stylized Applications

Separation of point, line, and plane clusters

5. Discussion



Summary

- The Basis Pursuit can be used as a
 - Forward transform, leading to sparse representation.
 - Way to achieve non-linear filtering.
- The dream: the over-completeness idea is highly effective, and should be used in modern methods in representation and inverse-problems.
- We would like to contribute to this development by
 - Supplying clear(er) explanations about the BP behavior,
 - Improve the involved numerical tools, and then
 - Deploy it to applications.



Future Work

- What dictionary to use? Relation to learning?
- BP beyond the bounds – Can we say more?
- Relaxed notion of sparsity? When zero is really zero?
- How to speed-up BP solver (both accurate and approximate)?
- Theory behind approximate BP?
- Applications – Demonstrating the concept for practical problems, such as denoising, coding, restoration, signal separation ...

