Sparse recovery with (non) convex optimization

Rémi Gribonval

METISS project-team (audio signal processing, speech recognition, source separation)

INRIA, Rennes, France

IRISA

UNE UNITÉ DE RECHERCHE À LA POINTE DES SCIENCES
ET DES TECHNOLOGIES DE L'INFORMATION
ET DE LA COMMUNICATION

Ecole d'été en Traitement du Signal de Peyresq

Peyresq, Juillet 2009

Sparsity and Inverse Problems

- More unknowns than equations:

$$\mathbf{A}x_0 = \mathbf{A}x_1 \not\Rightarrow x_0 = x_1$$

- Uniqueness of sparse solutions = identifiability
  - ✦ if x_0, x_1 are "sufficiently sparse",
  - ✦ then $\mathbf{A}x_0 = \mathbf{A}x_1 \Rightarrow x_0 = x_1$

- If x_0, x_1 "sufficiently sparse", identification with
  - ✦ L1-minimization = convex problem
  - ✦ Greedy algorithms

- Robustness to "approximately sparse" and noise

Overview

- Exact recovery
  - greedy algorithms
  - (non)convex Lp-minimization
  - comparisons
- Stability and robustness
  - Instance optimality for Lp-minimization
  - Restricted Isometry Property

Exact recovery

Usual sparsity measures

- L0-norm $$\|x\|_0 := \sum_k |x_k|^0 = \sharp\{k, x_k \neq 0\}$$ $$\underset{\text{support}(x)}{\parallel}$$

- L*p*-norms $$\|x\|_p^p := \sum_k |x_k|^p, 0 \leq p \leq 1$$

- Constrained minimization

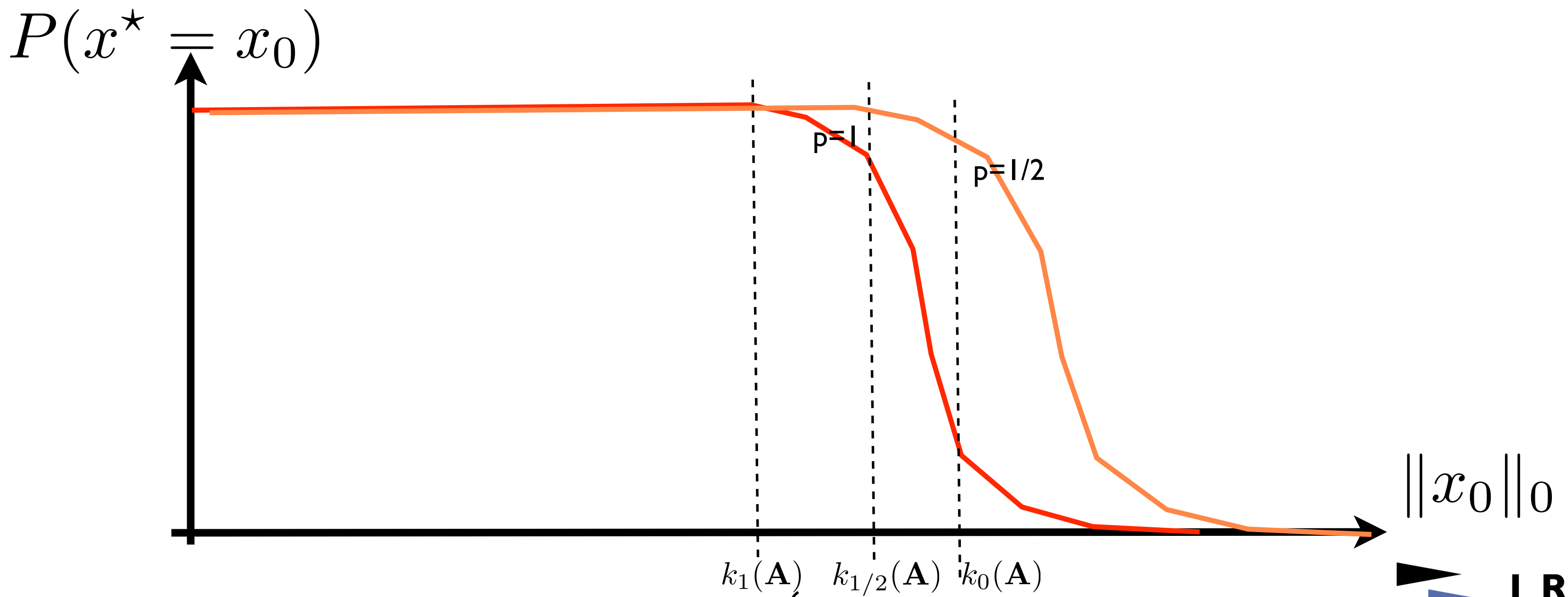
$$x_p^\star \in \arg\min_x \|x\|_p \text{ subject to } \mathbf{b} = \mathbf{A}x$$

Empirical observation : Lp versus L1

$$x_0 \longrightarrow \mathbf{b} := \mathbf{A}x_0 \longrightarrow x_p^\star = \arg \min_{\mathbf{A}x = \mathbf{A}x_0} \|x\|_p$$

reference — direct model — inverse problem

Typical observation *(e.g. Chartrand 2007)* +extrapolation



IRISA

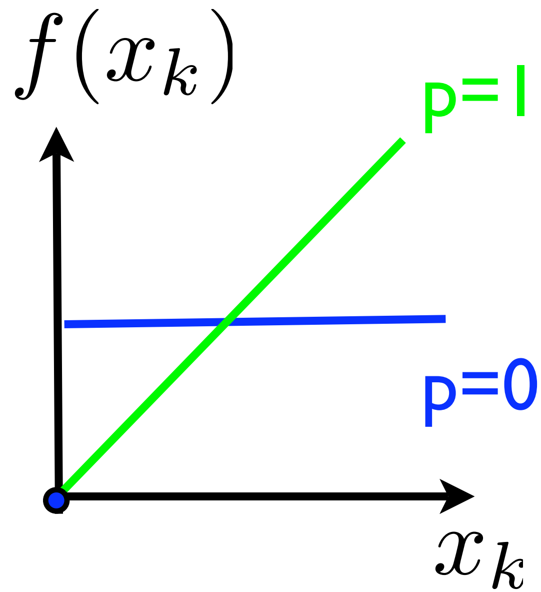
Proved Equivalence between L0 and L1

- "Empty" theorem : assume that $\mathbf{b} = \mathbf{A}x_0$
  - if $\|x_0\|_0 \leq k_0(\mathbf{A})$ then $x_0 = x_0^\star$
  - if $\|x_0\|_0 \leq k_1(\mathbf{A})$ $x_0 = x_1^\star$

- Content = estimation of $k_0(\mathbf{A})$ and $k_1(\mathbf{A})$
  - *Donoho & Huo 2001 : pair of bases, coherence*
  - *Donoho & Elad 2003, Gribonval & Nielsen 2003 : dictionary, coherence*
  - *Candes, Romberg, Tao 2004 : random dictionaries, restricted isometry constants*
  - *Tropp 2004 : idem for Orthonormal Matching Pursuit, cumulative coherence*
- What about $x_p^\star, 0 \leq p \leq 1$?

General sparsity measures

- Lp-norms $\|x\|_p^p := \sum_k |x_k|^p, 0 \leq p \leq 1$
- f-norms! $\|x\|_f := \sum_k f(|x_k|)$
- Constrained minimization



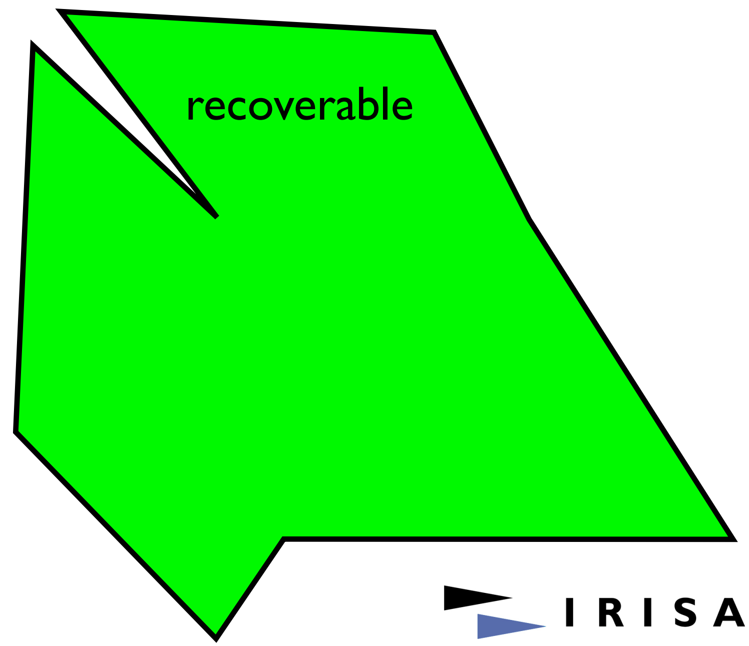
$$x_f^\star = x_f^\star(\mathbf{b}, \mathbf{A}) \in \arg\min_x \|x\|_f \text{ subject to } \mathbf{b} = \mathbf{A}x$$

When do we have $x_f^\star(\mathbf{A}x_0, \mathbf{A}) = x_0$?

Recovery analysis for inverse problem $\mathbf{b} = \mathbf{A}x$

- Holy grail = characterize set of "recoverable" coefficients

$$\{x \in \mathbb{R}^N, x_f^\star(\mathbf{A}x, \mathbf{A}) = x\}$$



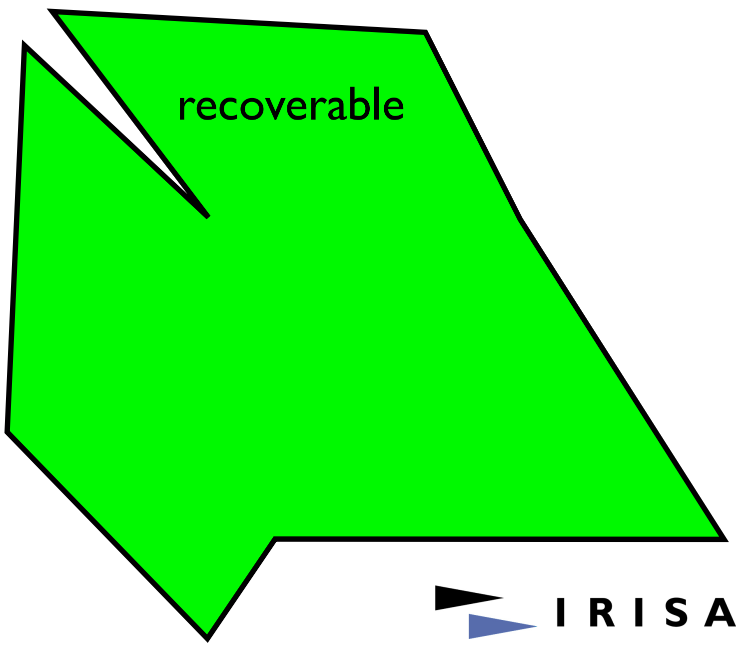
IRISA

Recovery analysis for inverse problem $\mathbf{b} = \mathbf{A}x$

- Holy grail = characterize set of "recoverable" coefficients

$$\{x \in \mathbb{R}^N, x^\star_f(\mathbf{A}x, \mathbf{A}) = x\}$$

- Practice : find "simple" recovery conditions

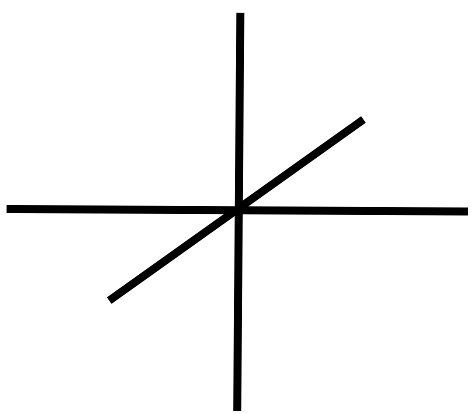


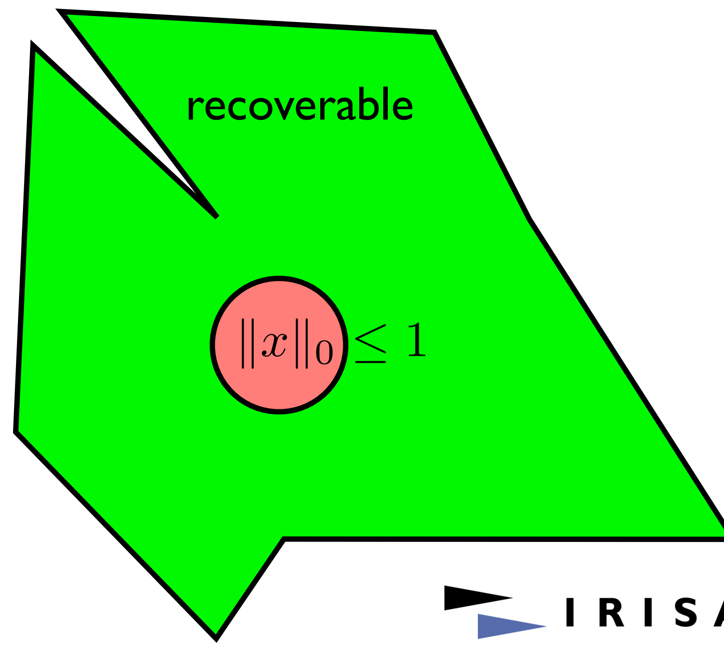
Recovery analysis for inverse problem $\mathbf{b} = \mathbf{A}x$

- Holy grail = characterize set of "recoverable" coefficients

$$\{x \in \mathbb{R}^N, x^\star_f(\mathbf{A}x, \mathbf{A}) = x\}$$

- Practice : find "simple" recovery conditions
  - 1-sparse



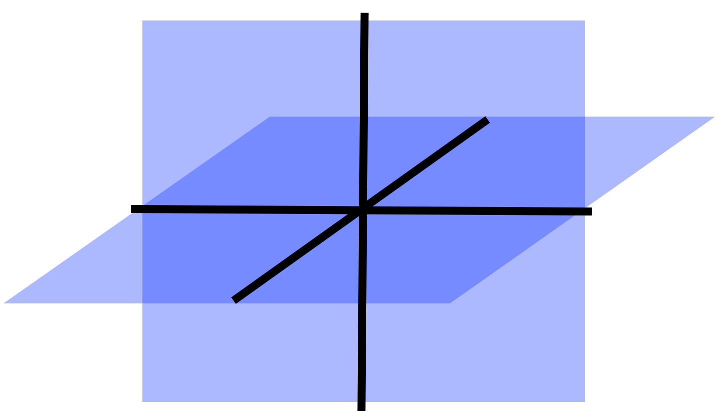


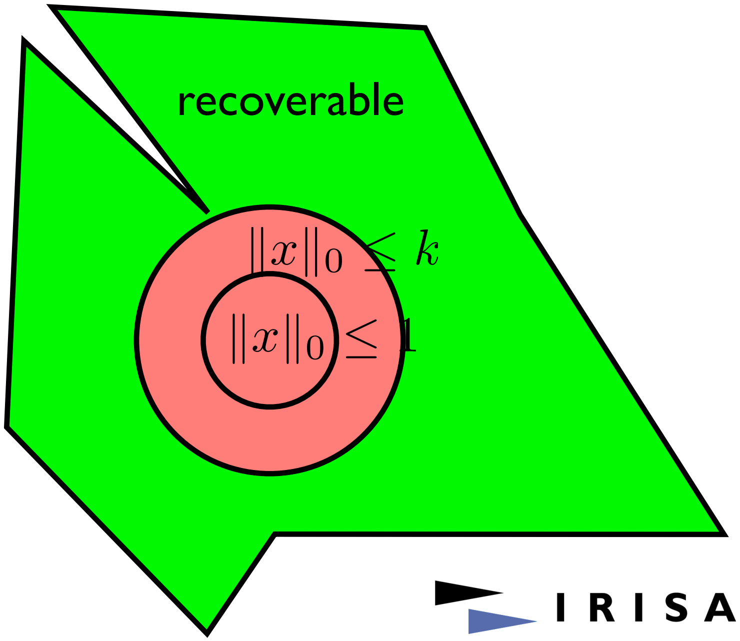
Recovery analysis for inverse problem $\mathbf{b} = \mathbf{A}x$

- Holy grail = characterize set of "recoverable" coefficients

$$\{x \in \mathbb{R}^N, x_f^\star(\mathbf{A}x, \mathbf{A}) = x\}$$

- Practice : find "simple" recovery conditions
  - 1-sparse
  - 2-sparse



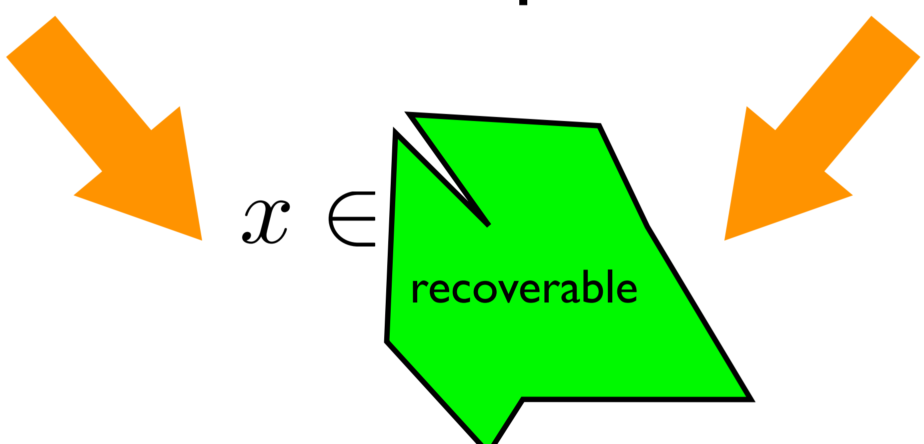


IRISA

Some “simple” recovery conditions

Support	Sparsity level
“recoverable supports” = subsets $I \subset [\![1, N]\!]$ such that $\mathrm{supp}(x) := \{k, x_k \neq 0\} \subset I$	“recoverable sparsity” = integers *k* such that $\|x_0\|_0 \leq k$

$x \in$



Exact recovery: Greedy algorithms

Objective

- Given
  - Dictionary **A**,
  - support set *I*
- Goal = understand when MP (and variants) is guaranteed to only select atoms in *I*, given any input vector $\mathbf{b} = \mathbf{A}_I x_I$
- Analysis: based on *operator norms*

IRISA

Operator norms

- Linear operator $\mathbf{L} : \ell^p \to \ell^q$
- Operator norm $\|\mathbf{L}\|_{p \to q} = \sup_{x \neq 0} \frac{\|\mathbf{L}x\|_q}{\|x\|_p}$
- Adjoint operator $\forall x, y \langle \mathbf{L}x, y \rangle = \langle x, \mathbf{L}^\star y \rangle$
- For real-valued matrices $\mathbf{L}^\star = \mathbf{L}^T$
- Duality relation: for all **L**, $1 \leq p, q \leq \infty$

$$\|\mathbf{L}\|_{p \to q} = \|\mathbf{L}^T\|_{q' \to p'} \qquad \frac{1}{p} + \frac{1}{p'} = 1; \frac{1}{q} + \frac{1}{q'} = 1$$

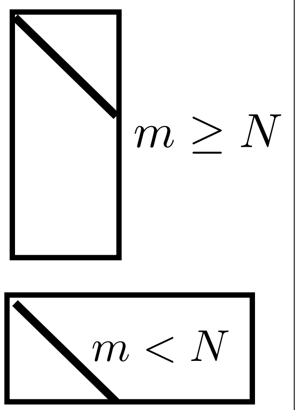
IRISA

Operator norms (ctd)

- When p=q=2
  - Singular Value Decomposition (SVD)

$$\mathbf{L} = U \Sigma V$$

$$UU^T = U^T U = Id_m$$
$$VV^T = V^T V = Id_N$$
$$\Sigma = \mathrm{diag}(\sigma_i),\ \sigma_1 \geq \ldots \geq 0$$



  - Operator norm

$$\|\mathbf{L}\|_{2 \to 2} = \|\Sigma\|_{2 \to 2} = \sigma_{\max(\mathbf{L})}$$

  - Proof: exercise

Operator norms (ctd)

- When p=1, for any q
  - ✦ Columns

$$\mathbf{L} = [\mathbf{L}_1, \ldots, \mathbf{L}_N]$$

  - ✦ Operator norm

$$\|\mathbf{L}\|_{1 \to q} = \max_n \|\mathbf{L}_n\|_q$$

  - ✦ Proof: exercise

Exact Recovery Condition for *MP

- **Theorem:** consider any weak/stagewise greedy algorithm that iterates
  - a **selection** of atoms in a set Γ_i such that

    $$\inf_{l \in \Gamma_i} |\mathbf{A}_l^T \mathbf{r}_{i-1}| \geq t \sup_n |\mathbf{A}_n^T \mathbf{r}_{i-1}|$$

  - an **update** of the residual such that

    $$\mathbf{r}_i \in \mathrm{span}(\mathbf{A}_n, n \in \cup_{j=1}^{i} \Gamma_i)$$

  - Assume = ERC(I) $\sup_{n \notin I} \|\mathbf{A}_I^+ \mathbf{A}_n\|_1 < t$

  - Conclude: given input $\mathbf{b} = \mathbf{A}_I x_I$, the algorithm is guaranteed to only select atoms in *I*: $\forall i, \Gamma_i \subset I$

ERC: proof

- Selection rule implies

$$\inf_{l \in \Gamma_i} |\mathbf{A}_l^T \mathbf{r}_{i-1}| \geq t \sup_n |\mathbf{A}_n^T \mathbf{r}_{i-1}| \quad = t \|\mathbf{A}^T \mathbf{r}_{i-1}\|_\infty$$

- It is sufficient to prove by induction that

$$\sup_{n \notin I} |\mathbf{A}_n^T \mathbf{r}_{i-1}| = \|\mathbf{A}_{I^c}^T \mathbf{r}_{i-1}\|_\infty < t \|\mathbf{A}^T \mathbf{r}_{i-1}\|_\infty$$

since this implies $n \notin I \Rightarrow n \notin \Gamma_i$

- Equivalently, we will just prove by induction

$$\frac{\|\mathbf{A}_{I^c}^T \mathbf{r}_{i-1}\|_\infty}{\|\mathbf{A}_I^T \mathbf{r}_{i-1}\|_\infty} < t$$

IRISA

ERC: proof (ctd)

- **Lemma** $$\sup_{\mathbf{r}\in\mathrm{span}(\mathbf{A}_I)} \frac{\|\mathbf{A}_{I^c}^T\mathbf{r}\|_\infty}{\|\mathbf{A}_I^T\mathbf{r}\|_\infty} = \sup_{n\notin I}\|\mathbf{A}_I^+\mathbf{A}_n\|_1$$
- Proof:

$$\frac{\|\mathbf{A}_{I^c}^T\mathbf{r}\|_\infty}{\|\mathbf{A}_I^T\mathbf{r}\|_\infty} \overset{\mathbf{r}=\mathbf{A}_I c}{=} \frac{\|\mathbf{A}_{I^c}^T\mathbf{A}_I c\|_\infty}{\|\mathbf{A}_I^T\mathbf{A}_I c\|_\infty} \overset{c=(\mathbf{A}_I^T\mathbf{A}_I)^{-1}d}{=} \frac{\|\mathbf{A}_{I^c}^T\mathbf{A}_I(\mathbf{A}_I^T\mathbf{A}_I)^{-1}d\|_\infty}{\|d\|_\infty}$$

$$\leq \|\mathbf{A}_{I^c}^T\mathbf{A}_I(\mathbf{A}_I^T\mathbf{A}_I)^{-1}\|_{\infty\to\infty}$$

$$= \|(\mathbf{A}_I^T\mathbf{A}_I)^{-1}\mathbf{A}_I^T\mathbf{A}_{I^c}\|_{1\to1} = \|\mathbf{A}_I^+\mathbf{A}_{I^c}\|_{1\to1}$$

ERC: proof (end)

- For i=1 $\mathbf{r}_{i-1} \in \mathrm{span}(\mathbf{A}_I)$
- By induction, if this holds true at step i we obtain

$$\frac{\|\mathbf{A}_{I^c}^T \mathbf{r}_{i-1}\|_\infty}{\|\mathbf{A}_I^T \mathbf{r}_{i-1}\|_\infty} \leq \sup_{n \notin I} \|\mathbf{A}_I^+ \mathbf{A}_n\|_1 < t$$

- It follows that $\Gamma_i \subset I$
- By the update rule

$$\mathbf{r}_i \in \mathrm{span}(\mathbf{A}_n, n \in \cup_{j=1}^i \Gamma_i) \subset \mathrm{span}(\mathbf{A}_I)$$

IRISA

Exact recovery:
Lp minimization

Null space

- Null space = kernel

$$z \in \mathcal{N}(\mathbf{A}) \Leftrightarrow \mathbf{A}z = 0$$

- Particular solution *vs* general solution
  - particular solution

$$\mathbf{A}x = \mathbf{b}$$

  - general solution

$$\mathbf{A}x' = \mathbf{b} \Leftrightarrow x' - x \in \mathcal{N}(\mathbf{A})$$

Exact recovery: necessary condition

- Notations
  - index set *I*
  - vector *z*
  - restriction $z_I = (z_i)_{i \in I}$
- Assume there exists $z \in \mathcal{N}(\mathbf{A})$ with

$$\|z_I\|_f > \|z_{I^c}\|_f$$

- Define $\mathbf{b} := Az_I = A(-z_{I^c})$
- The vector z_I is supported in *I* but is *not* the minimum norm representation of $\mathbf{b}$

IRISA

Exact recovery: sufficient condition

- Assume quasi-triangle inequality

$$\forall x, y \|x+y\|_f \leq \|x\|_f + \|y\|_f$$

- Consider *x* with support set *I* and *x'* with $\mathbf{A}x' = \mathbf{A}x$
- Denote $z := x' - x \in \mathcal{N}(\mathbf{A})$ and observe

$$\begin{aligned} \|x'\|_f = \|x+z\|_f &= \|(x+z)_I\|_f + \|(x+z)_{I^c}\|_f \\ &= \|x+z_I\|_f + \|z_{I^c}\|_f \\ &\geq \|x\|_f - \|z_I\|_f + \|z_{I^c}\|_f \end{aligned}$$

- Conclude:

If $\|z_{I^c}\|_f > \|z_I\|_f$ when $z \in \mathcal{N}(\mathbf{A})$ then I is recoverable

IRISA

Recoverable supports : the "Null Space Property" (I)

- **Theorem 1** [*Donoho & Huo 2001 for L1, G. & Nielsen 2003 for Lp*]
  - Assumption 1: sub-additivity (for quasi-triangle inequality)

$$f(a+b) \leq f(a) + f(b), \forall a, b$$

  - Assumption 2:

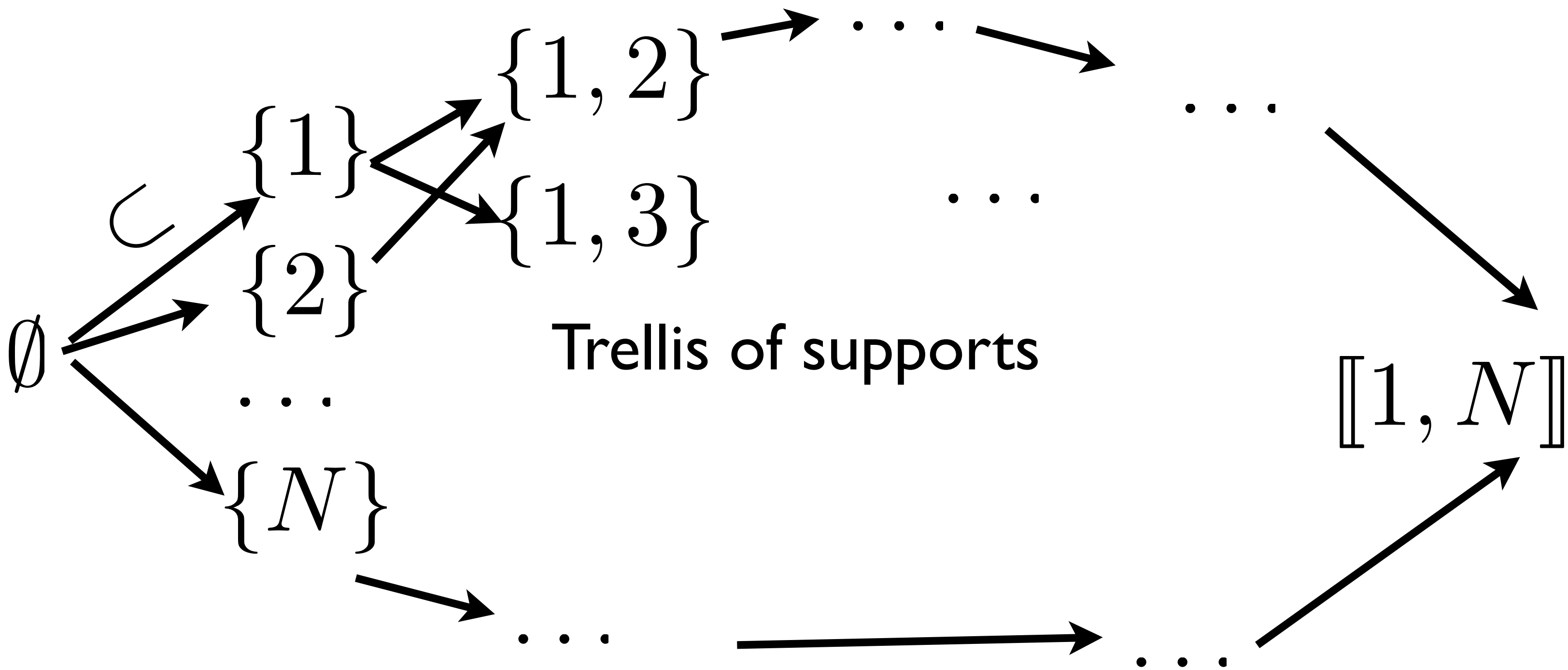
NSP

$$\|z_I\|_f < \|z_{I^c}\|_f \text{ when } z \in \mathcal{N}(\mathbf{A}), z \neq 0$$

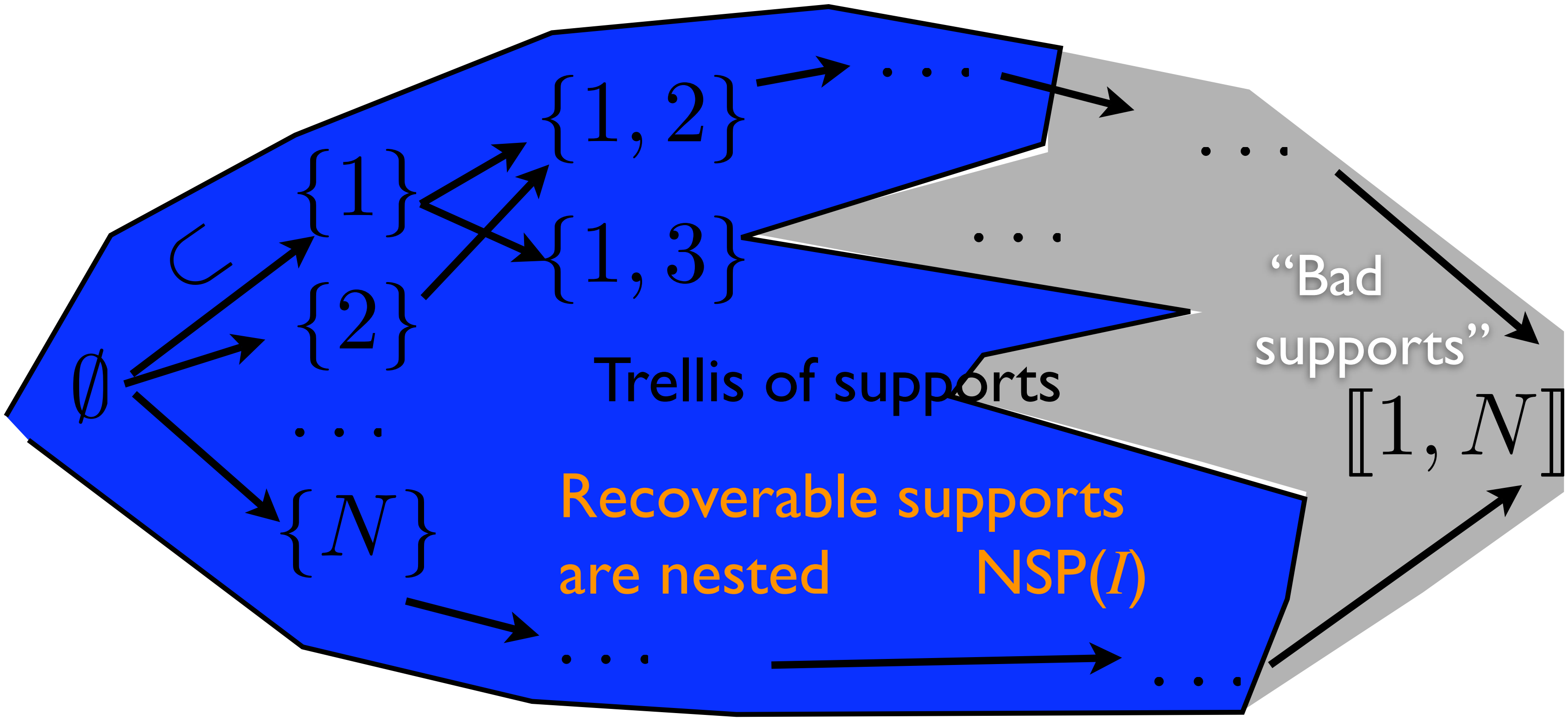
  - Conclusion: $x^\star_f$ recovers every x supported in I
  - The result is sharp: if NSP fails on support *I* there is **at least one failing vector** x supported in *I*

IRISA

From “recoverable” supports to “sparse” vectors

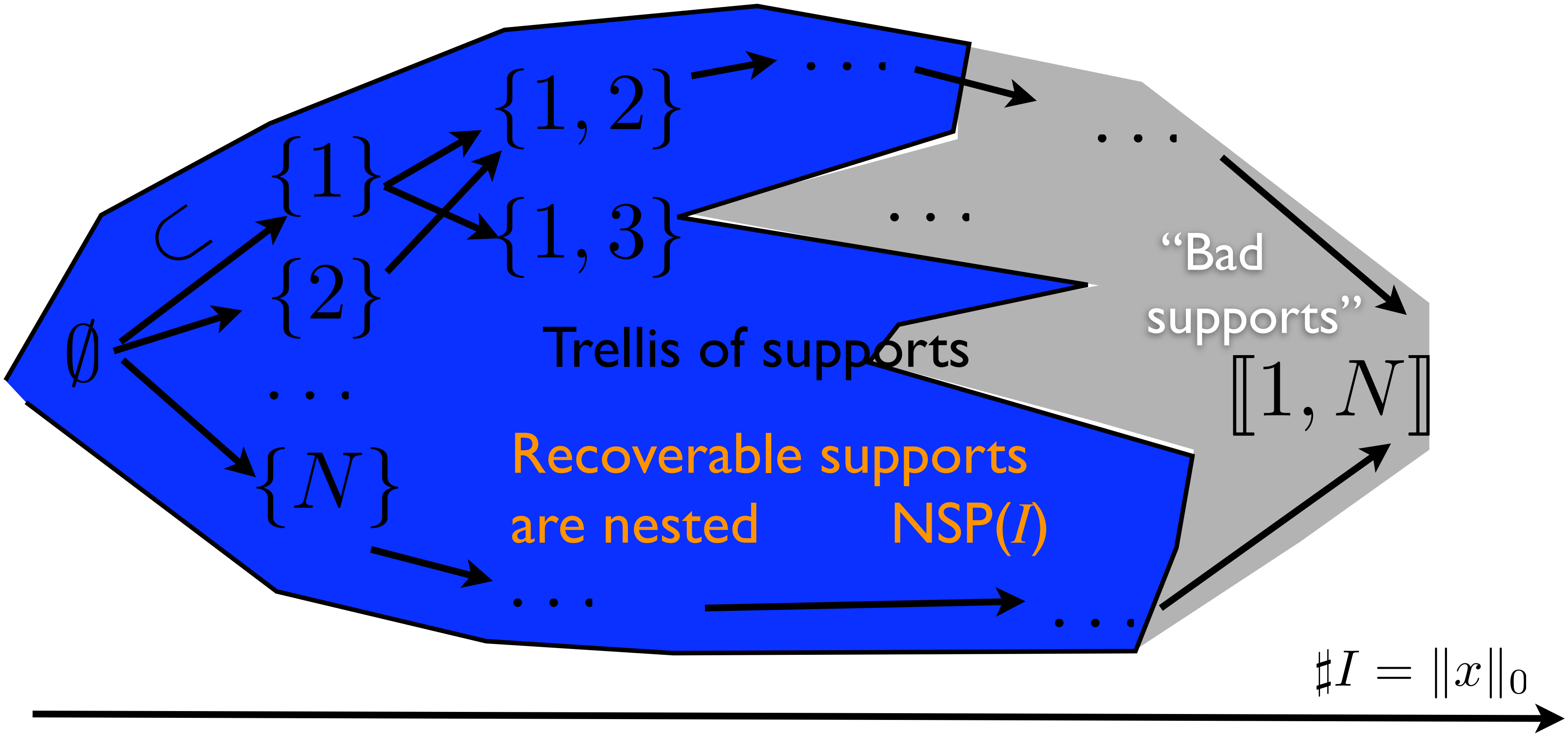


From “recoverable” supports to “sparse” vectors

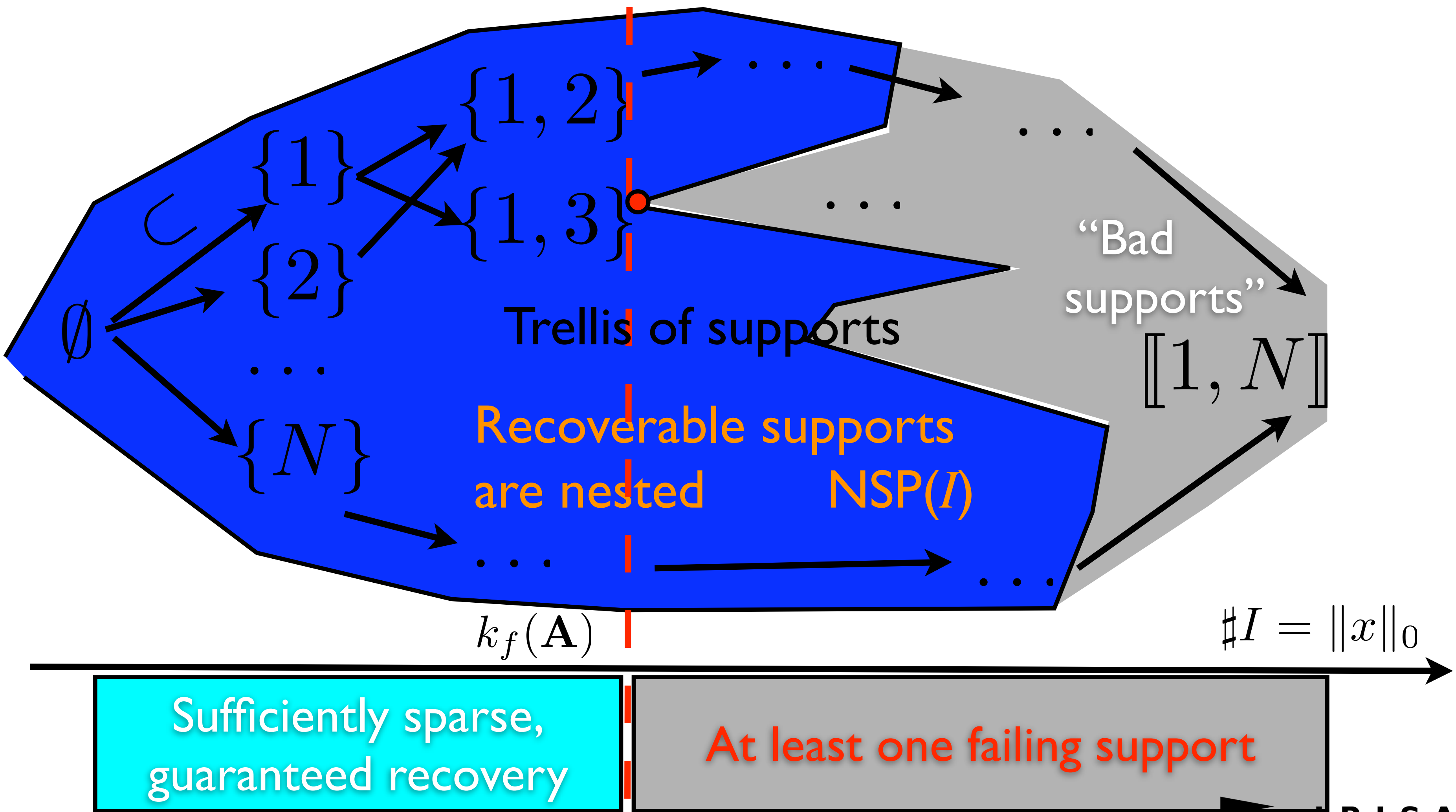


IRISA

From “recoverable” supports to “sparse” vectors



From "recoverable" supports to "sparse" vectors



IRISA

Recoverable sparsity levels: the "Null Space Property" (2)

- Corollary 1 [*Donoho & Huo 2001 for L1, G. Nielsen 2003 for Lp*]
  - Definition :

$$I_k = \quad \text{index of } k \text{ largest components of z}$$

  - Assumption :

NSP

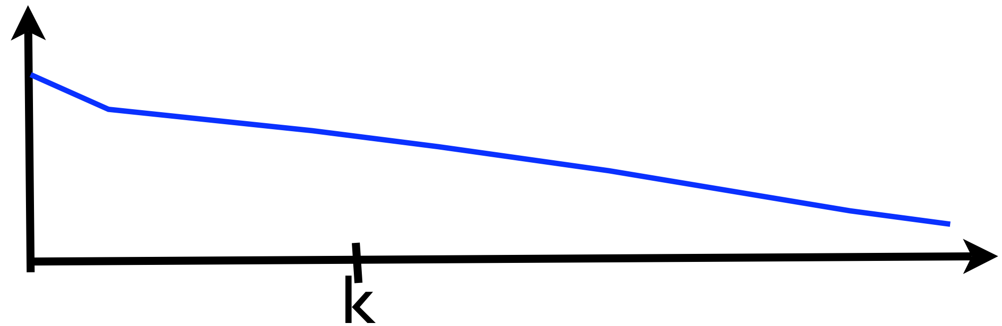
$$\|z_{I_k}\|_f < \|z_{I_k^c}\|_f \quad \text{when } z \in \mathcal{N}(\mathbf{A}), z \neq 0$$

  - Conclusion: $x^\star_f$ recovers every x with $\|x\|_0 \leq k$
  - The result is sharp: if NSP fails there is **at least one failing vector** x with $\|x\|_0 = k$

IRISA

Interpretation of NSP

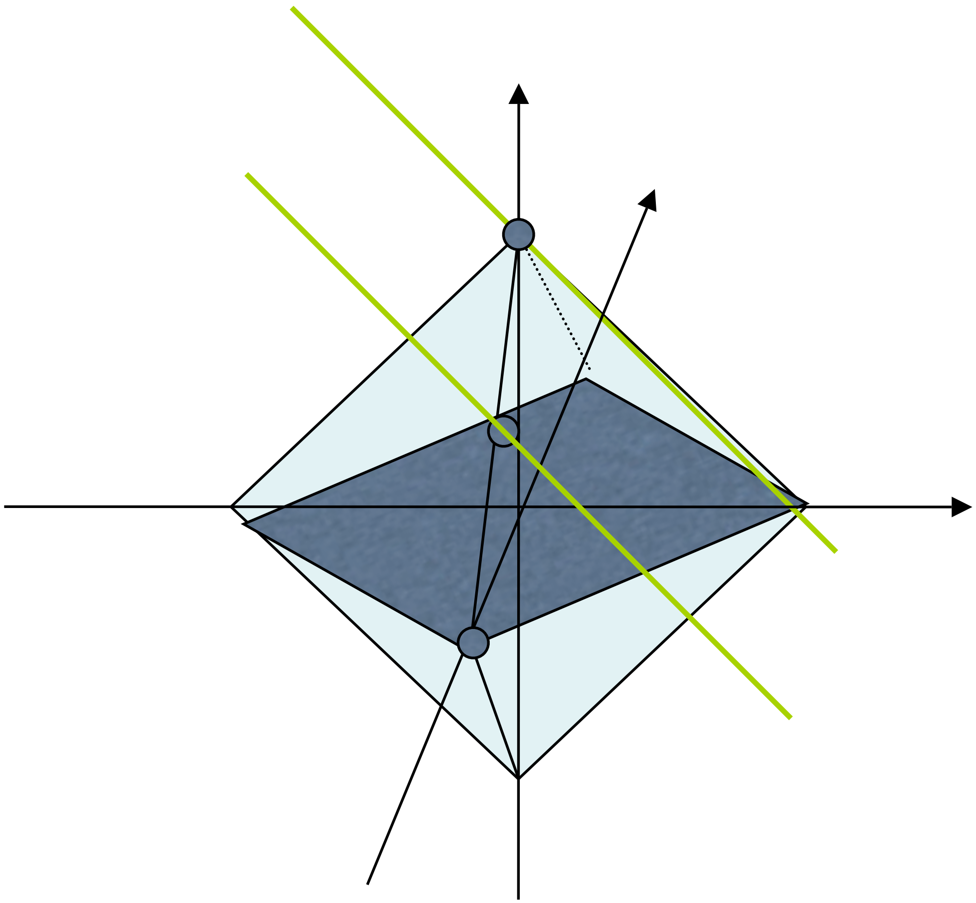
- Geometry in coefficient space:
  - ✦ consider an element z of the Null Space of A
  - ✦ order its entries in decreasing order



  - ✦ the mass of the largest k-terms should not exceed that of the tail $\|z_{I_k}\|_f < \|z_{I_k^c}\|_f$

All elements of the null space must be rather "flat"

Geometric picture



Greedy vs L1

ERC implies NSP

ERC(I) $\sup_{n \notin I} \|\mathbf{A}_I^+ \mathbf{A}_n\|_1 < t$

If columns of $\mathbf{A}_I$
linearly independent

NSP(I, ℓ^1, t)

$$\|z_I\|_1 \leq t \cdot \|z_{I^c}\|_1 \qquad \text{when} \quad z \in \mathcal{N}(\mathbf{A}), z \neq 0$$

ERC implies NSP: proof

- Assume ERC(*I*) + linear independence $\sup_{n \notin I} \|\mathbf{A}_I^+ \mathbf{A}_n\|_1 < t$

- Consider $z \in \mathcal{N}(\mathbf{A})$
  - write $\mathbf{A}_I z_I = -\mathbf{A}_{I^c} z_{I^c}$
  - apply pseudo-inverse $z_I = -\mathbf{A}_I^+ \mathbf{A}_{I^c} z_{I^c}$

$$\|z_I\|_1 \leq \|\mathbf{A}_I^+ \mathbf{A}_{I^c}\|_{1\to 1} \cdot \|z_{I^c}\|_1 \leq t\|z_{I^c}\|_1$$

Greedy vs L1: summary

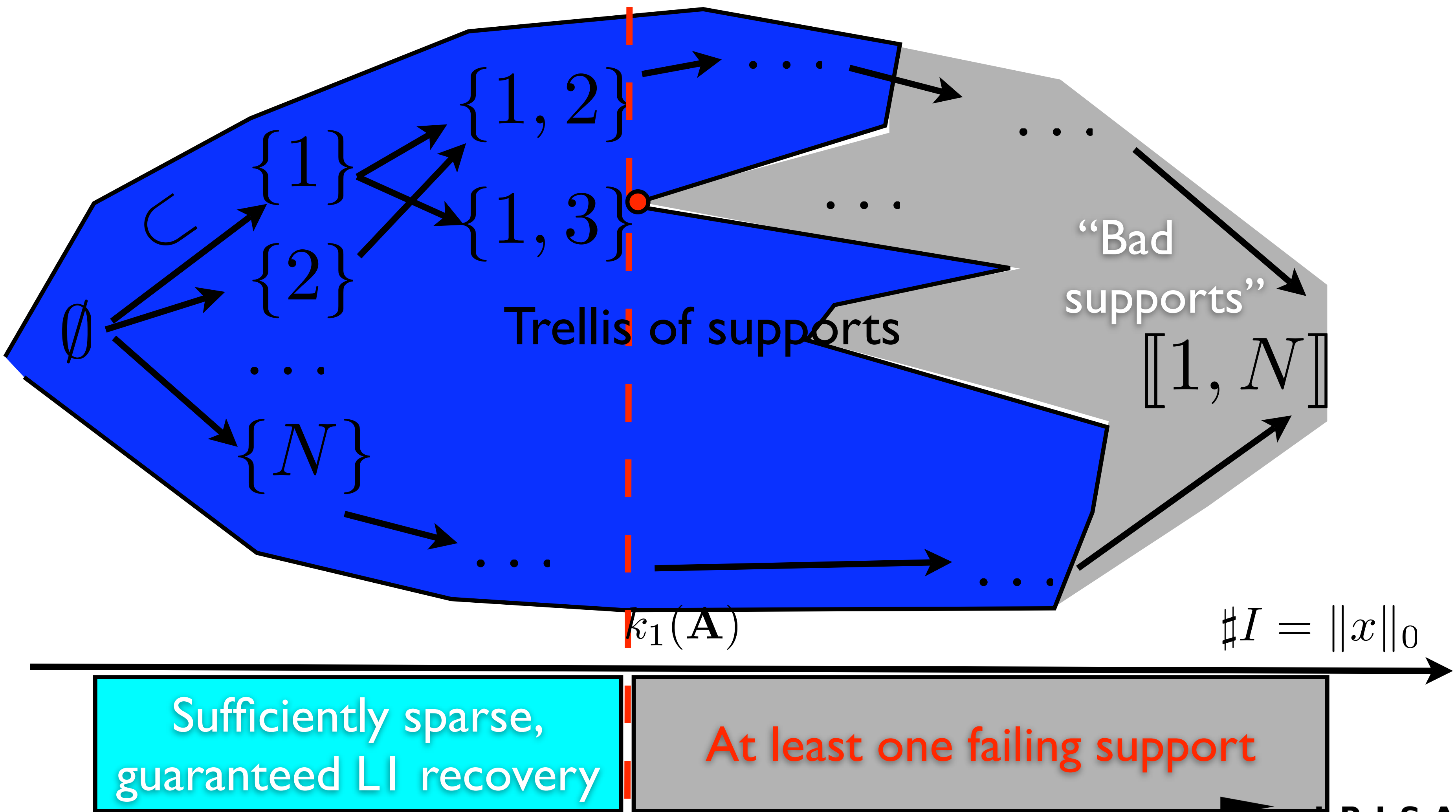
- If *MP is guaranteed to recover all vectors with support *I*, where the atoms in I are linearly independent, then L1 has the same guarantee
- If *MP recovers all k-sparse vectors, then L1 has the same guarantee

$$k_{*\mathrm{MP}}(\mathbf{A}) \leq k_1(\mathbf{A}), \forall \mathbf{A}$$

- Warning: there are support sets I
  - not recovered by L1, while recovered by MP
  - .... but columns of $\mathbf{A}_I$ are linearly dependent.
  - Example: *I* = $[\![1, N]\!]$ when **A** is *m* x *N*

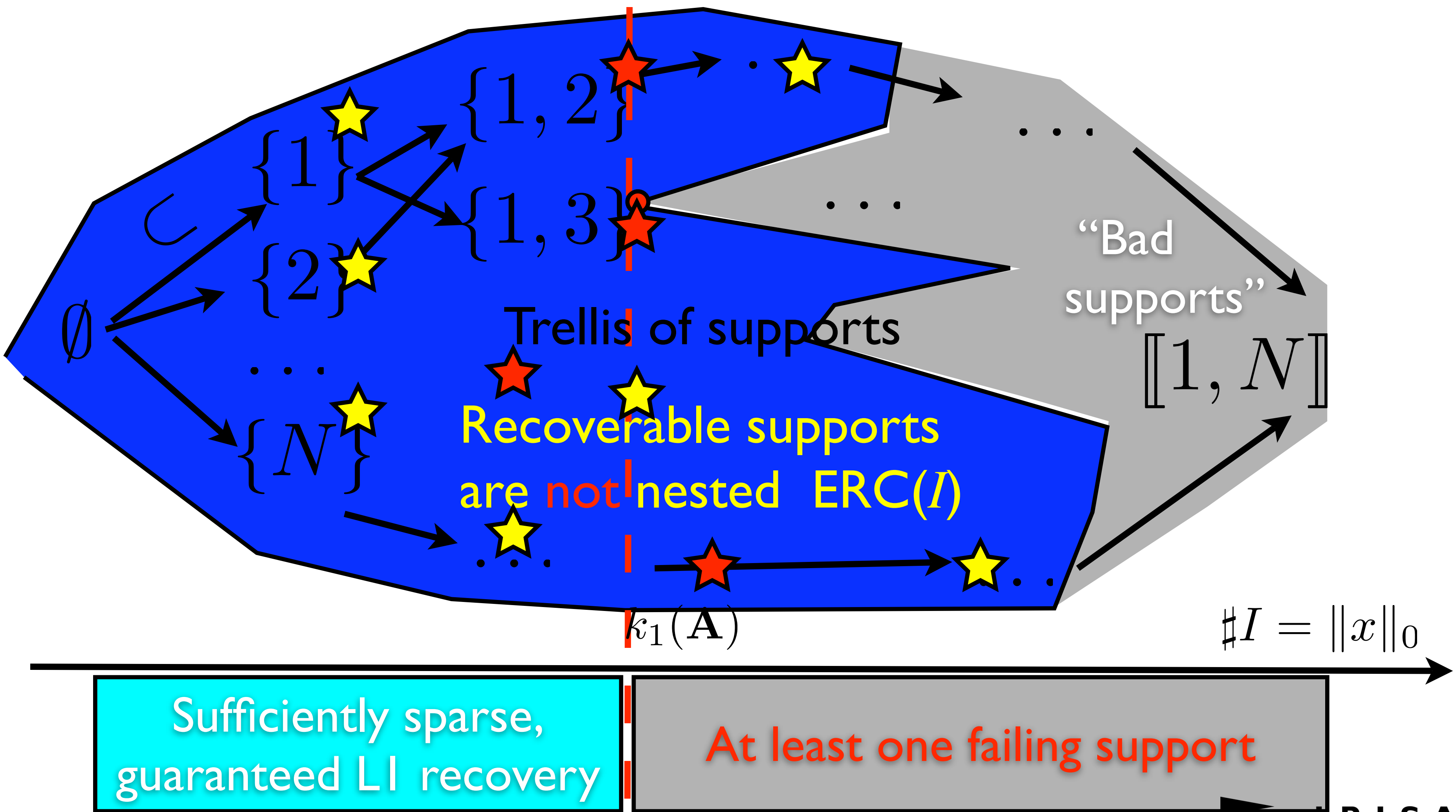
IRISA

Greedy vs L1: summary

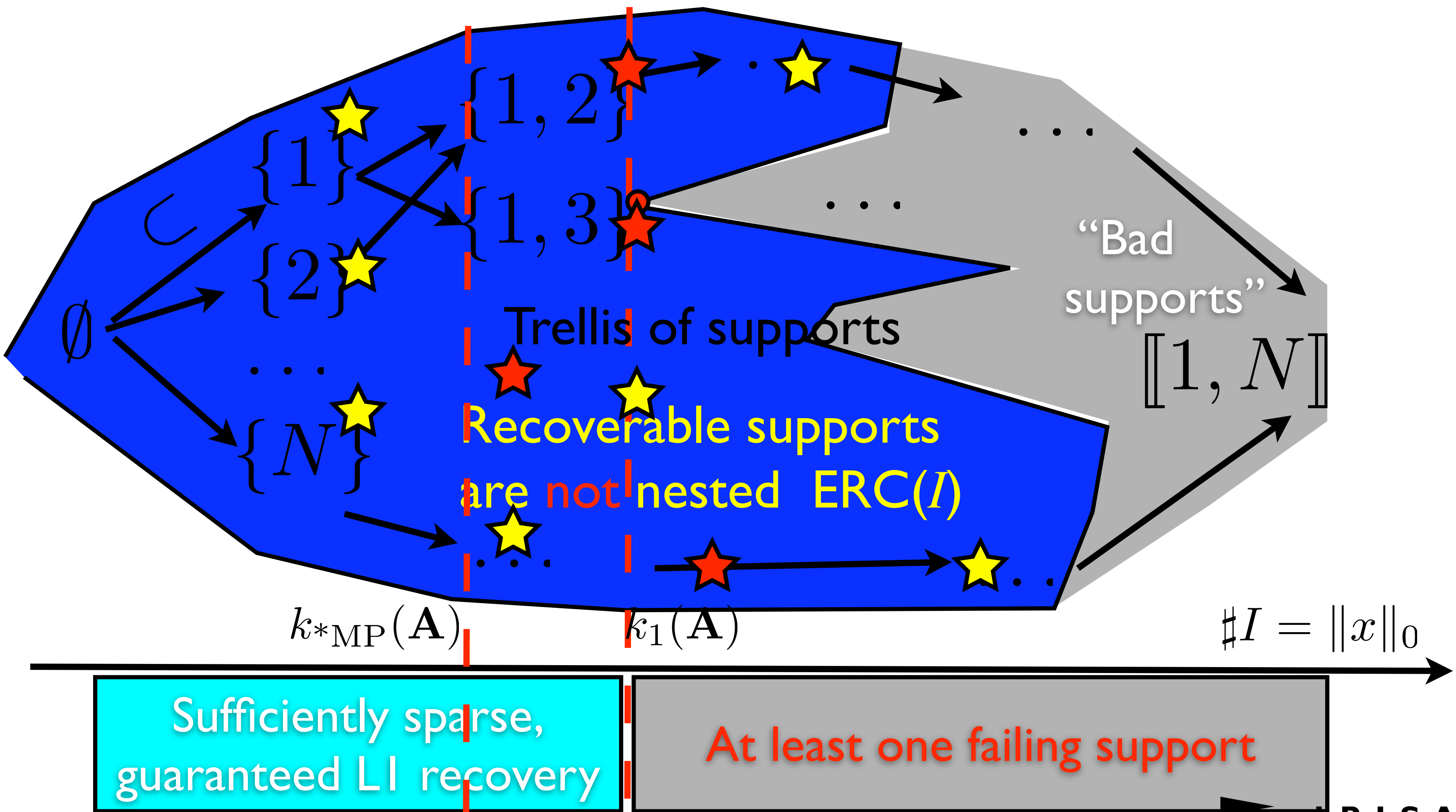


IRISA

Greedy vs L1: summary



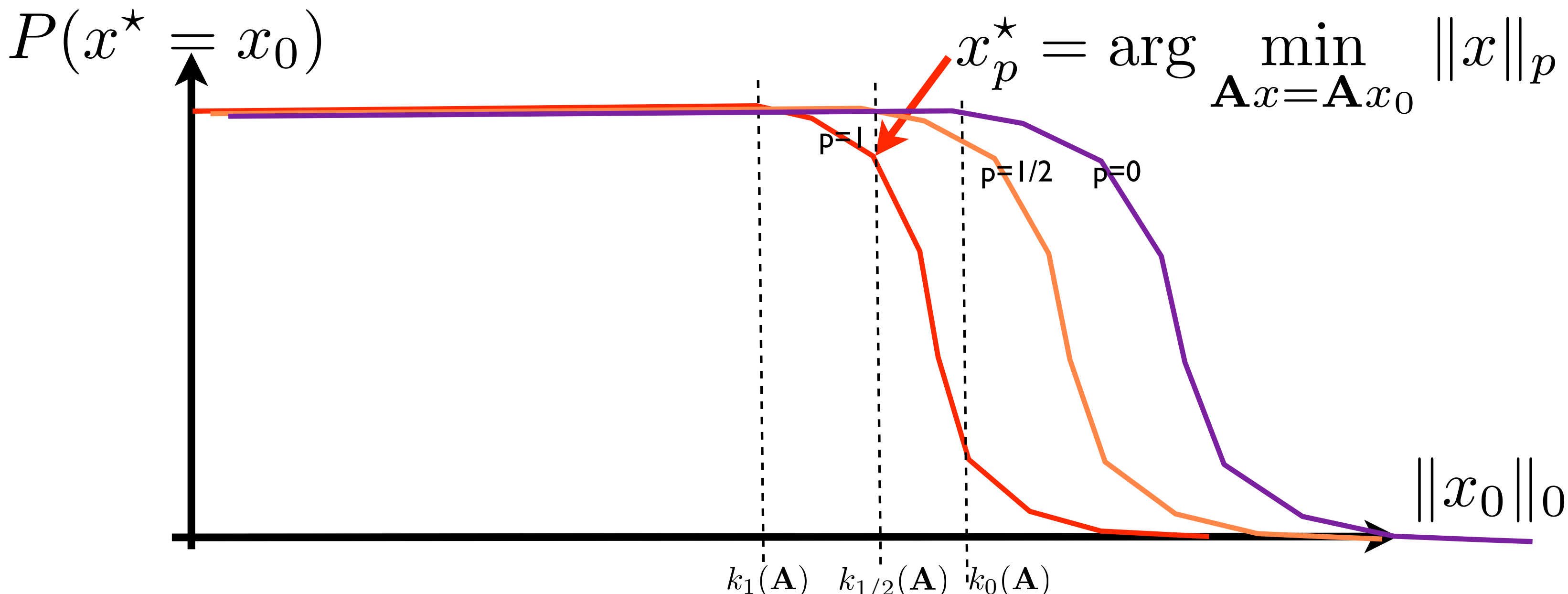
Greedy vs L1: summary



L1 vs Lp

Critical sparsity levels for different Lp norms

Back to empirical observations + extrapolation :



Do we always have $k_1(\mathbf{A}) \leq k_p(\mathbf{A}) \leq k_0(\mathbf{A}), 0 \leq p \leq 1$?

Lp better than L1 (1)

- **Theorem** 2 [*G. Nielsen 2003*]
  - Assumption 1: **sub-additivity** of sparsity measures f, g

$$f(a+b) \leq f(a) + f(b), \forall a, b$$

  - Assumption 2: the function $t \mapsto \frac{f(t)}{g(t)}$ is **non-increasing** $\searrow$
  - Conclusion: $k_g(\mathbf{A}) \leq k_f(\mathbf{A}), \forall \mathbf{A}$

Minimizing $\|x\|_f$ *can recover vectors which are less sparse than required for guaranteed success when minimizing* $\|x\|_g$

IRISA

Lp better than L1 (2)

- **Example**
  - sparsity measures $f(t) = t^p,\ \ g(t) = t^q, 0 \leq p \leq q \leq 1$
  - sub-additivity

$$|a+b|^p \leq |a|^p + |b|^p, \forall a, b, 0 \leq p \leq 1$$

  - function $\dfrac{f(t)}{g(t)} = t^{p-q}$ is non-increasing
  - therefore

$$k_1(\mathbf{A}) \leq k_q(\mathbf{A}) \leq k_p(\mathbf{A}) \leq k_0(\mathbf{A}), \forall \mathbf{A}$$

Lp better than L1: proof

- 1) Since f/g non-decreasing:

$$z_1 \geq z_2 \geq 0 \Rightarrow \frac{f(z_1)}{g(z_1)} \leq \frac{f(z_2)}{g(z_2)}$$

- 2) Similarly

$$z_1 \geq \ldots \geq z_N \geq 0 \Rightarrow \frac{\|z_{1:k}\|_f}{\|z_{1:k}\|_g} \leq \frac{\|z_{k+1:N}\|_f}{\|z_{k+1:N}\|_g}$$

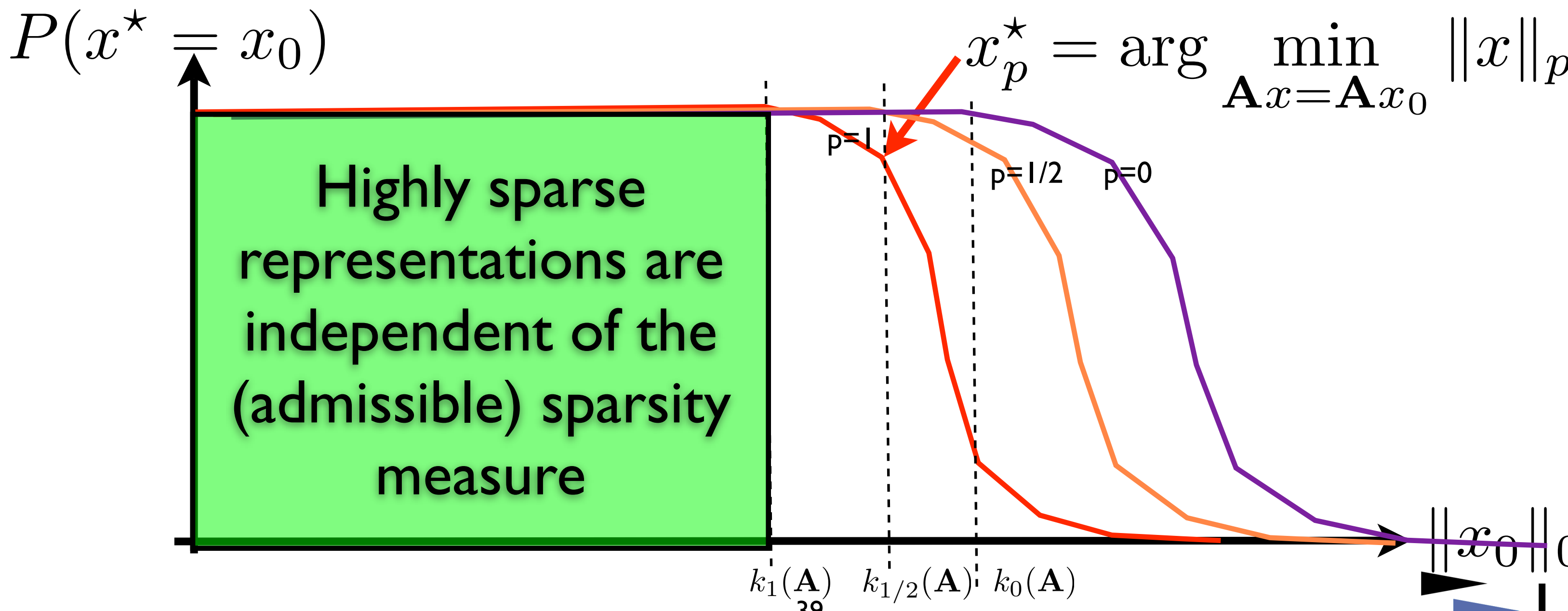
$$I_k = 1:k$$
$$I_k^c = k+1:N$$

$$\frac{\|z_{I_k}\|_f}{\|z_{I_k^c}\|_f} \leq \frac{\|z_{I_k}\|_g}{\|z_{I_k^c}\|_g}$$

- 3) Conclusion : if NSP(g,t,k) then NSP(f,t,k)

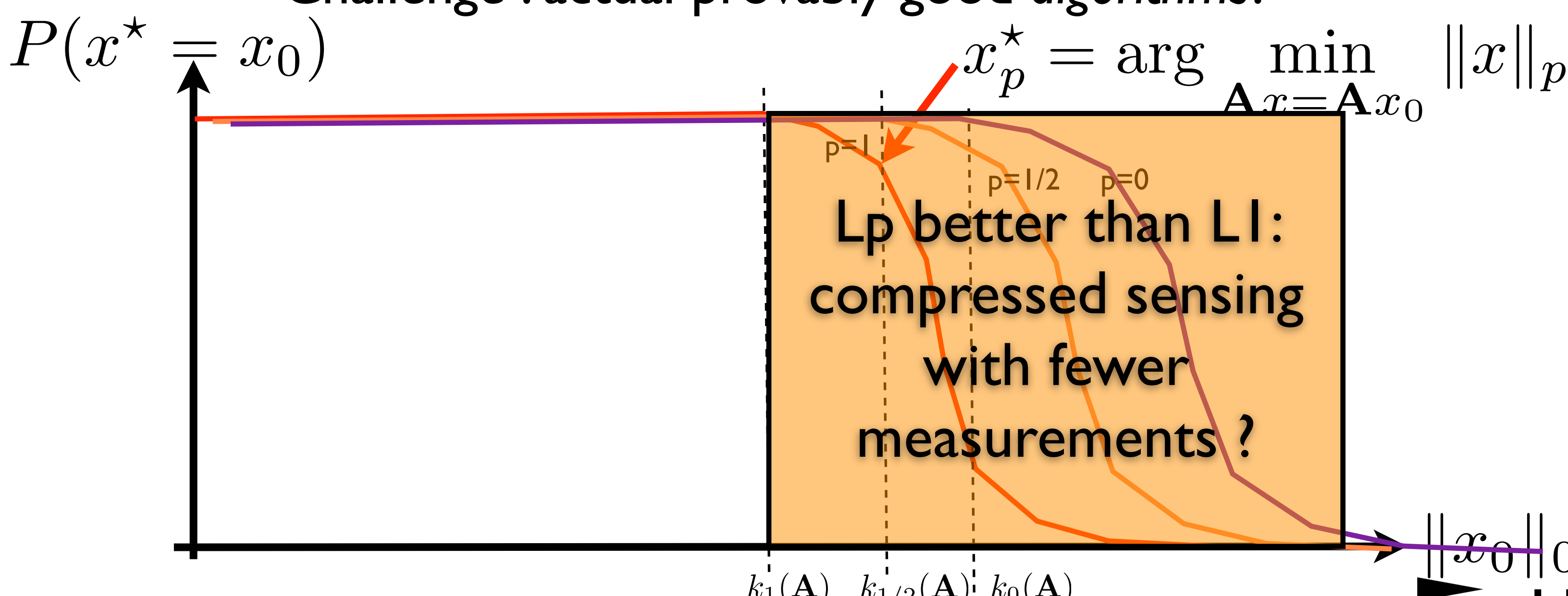
Lp better than L1 (2)

- At sparsity levels where L1 is guaranteed to "succeeds", all Lp p<=1 is also guaranteed to succeed



Lp better than L1 (3)

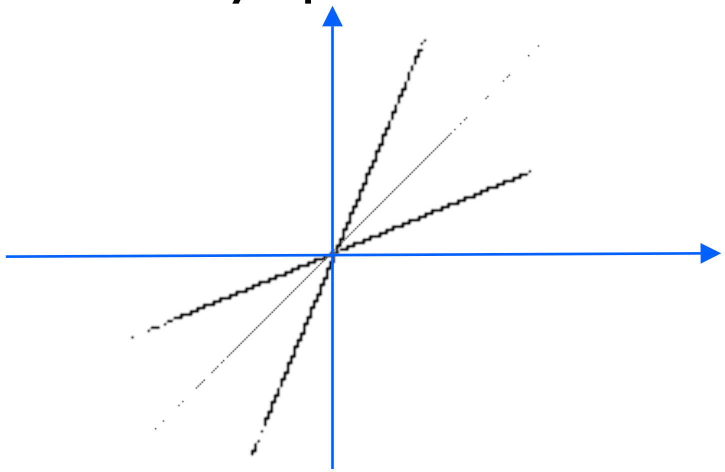
- + Lp p<1 can succeed where L1 fails
  - How much improvement ? Quantify $k_p(\mathbf{A})$?
- - Lp p<1 : nonconvex, has many local mimima
  - Better recovery with Lp *principle*
  - Challenge : actual provably good *algorithms*?



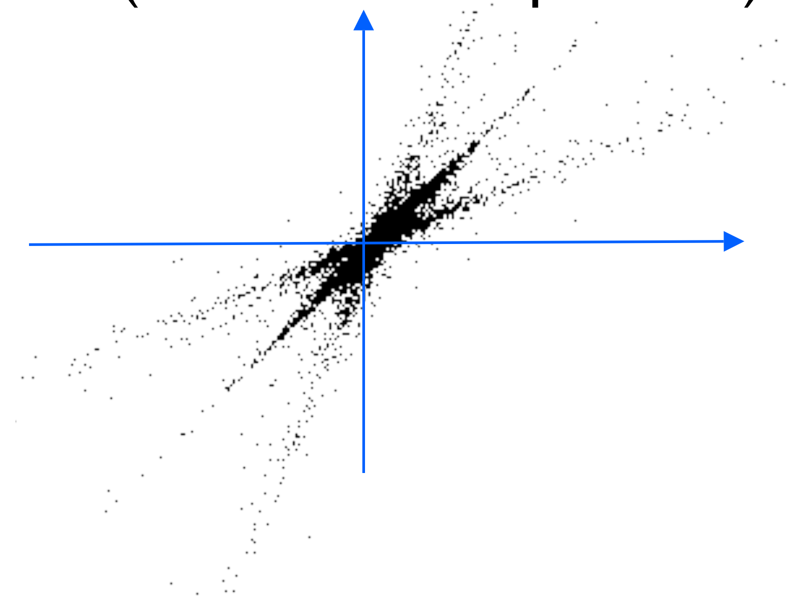
Stability and robustness

Stability

Exactly sparse data



Real data (from source separation)



Stability

- Exact recovery: $\mathbf{b} = \mathbf{A}x$
  - ✦ sparsity assumption $\|x\|_0 \leq k_p(\mathbf{A}) < m$
  - ✦ recovery: $x_p^\star(\mathbf{b}) = x$
- Stability: relax sparsity assumption
  - ✦ best k-term approximation

$$\sigma_k(x) = \inf_{\|y\|_0 \leq k} \|x - y\|$$

  - ✦ goal = stable recovery = instance optimality

$$\|x_p^\star(\mathbf{b}) - x\| \leq C \cdot \sigma_k(x)$$

Instance optimality for Lp minimization

- Assumption:

NSP(k, ℓ^p,t)

$$\|z_{I_k}\|_p^p \leq t \cdot \|z_{I_k^c}\|_p^p \quad \text{when} \quad z \in \mathcal{N}(\mathbf{A}), z \neq 0$$

- Conclusion: instance optimality for all *x*

$$\|x_p^\star(\mathbf{b}) - x\|_p^p \leq C(t) \cdot \sigma_k(x)_p^p$$

$$C(t) := 2\frac{1+t}{1-t}$$

IRISA

Robustness

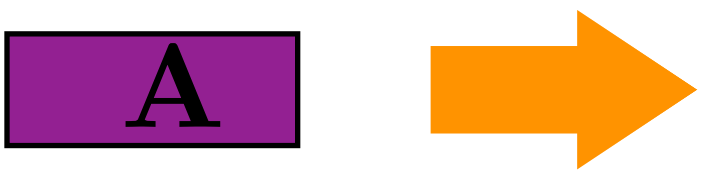
- Noiseless model $\mathbf{b} = \mathbf{A}x$
- Noisy model $\mathbf{b} = \mathbf{A}x + \mathbf{e}$
  - measurement noise
  - modeling error
  - numerical inaccuracies ...
- Goal: robust estimation

$$\|x_p^\star(\mathbf{b}) - x\| \leq C\|e\| + C'\sigma_k(x)$$

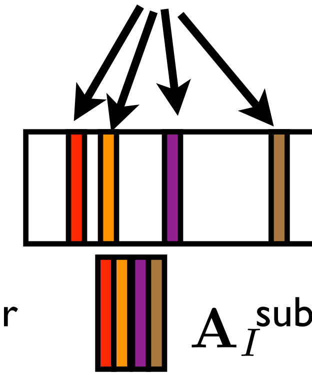
- Tool: restricted isometry property

Restricted Isometry Property

- Definition



N columns



$n \in I, \sharp I \le k$

max over subsets *I* $\mathbf{A}_I$

$\frac{N!}{k!(N-k)!}$

$$\delta_k := \sup_{\sharp I \le k,\ c \in \mathbb{R}^k} \left| \frac{\|\mathbf{A}_I c\|_2^2}{\|c\|_2^2} - 1 \right|$$

- Computation ?
  - naively: combinatorial
  - **open question**: NP ? NP-complete ?

RIP, stability, robustness

RIP(k, δ)

$$\delta_{2k}(\mathbf{A}) \leq \delta$$

[Candès 2008] $t := \sqrt{2}\delta/(1-\delta)$

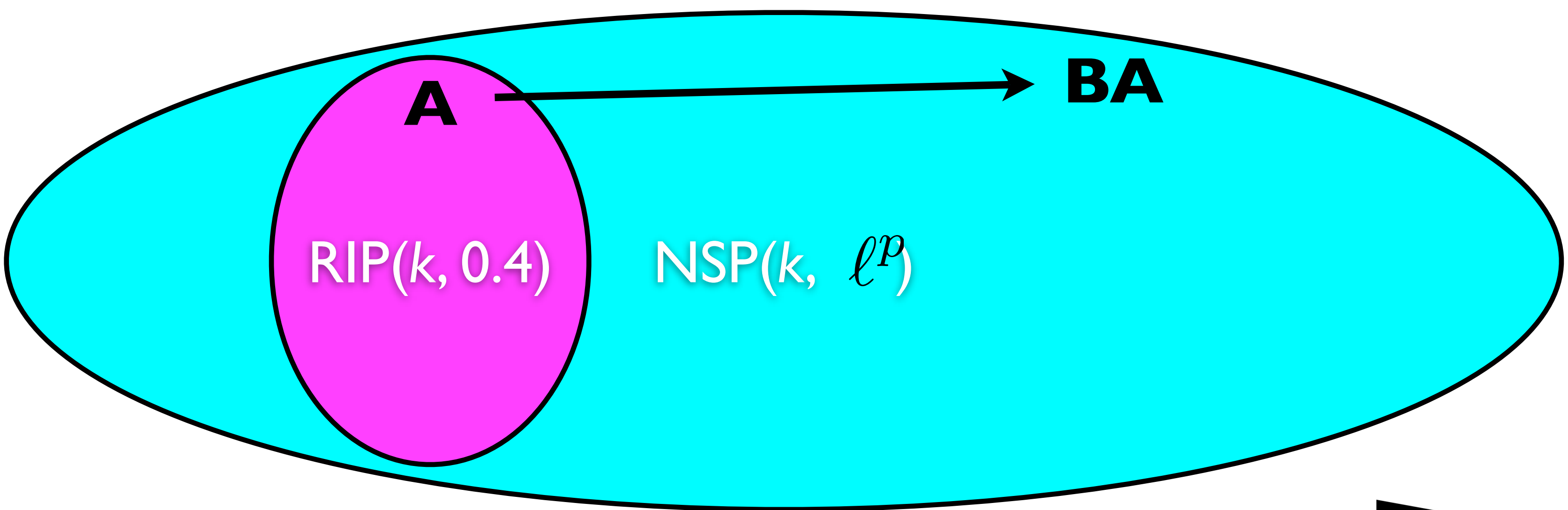
NSP(k, ℓ^1,t)

$$\|z_{I_k}\|_1 \leq t \cdot \|z_{I_k^c}\|_1 \quad \text{when} \quad z \in \mathcal{N}(\mathbf{A}), z \neq 0$$

- L1-recovery of *k*-sparse vectors, **robust to noise** when $\mathbf{b} = \mathbf{A}x + \mathbf{n}$ $\delta_{2k}(\mathbf{A}) < \sqrt{2} - 1 \approx 0.414$
  - *Foucart-Lai 2008*: Lp with p<1, and $\delta_{2k}(\mathbf{A}) < 0.4531$
  - *Chartrand 2007, Saab & Yilmaz 2008*: other RIP condition for p<1
  - *G., Figueras & Vandergheynst 2006*: robustness with *f*-norms

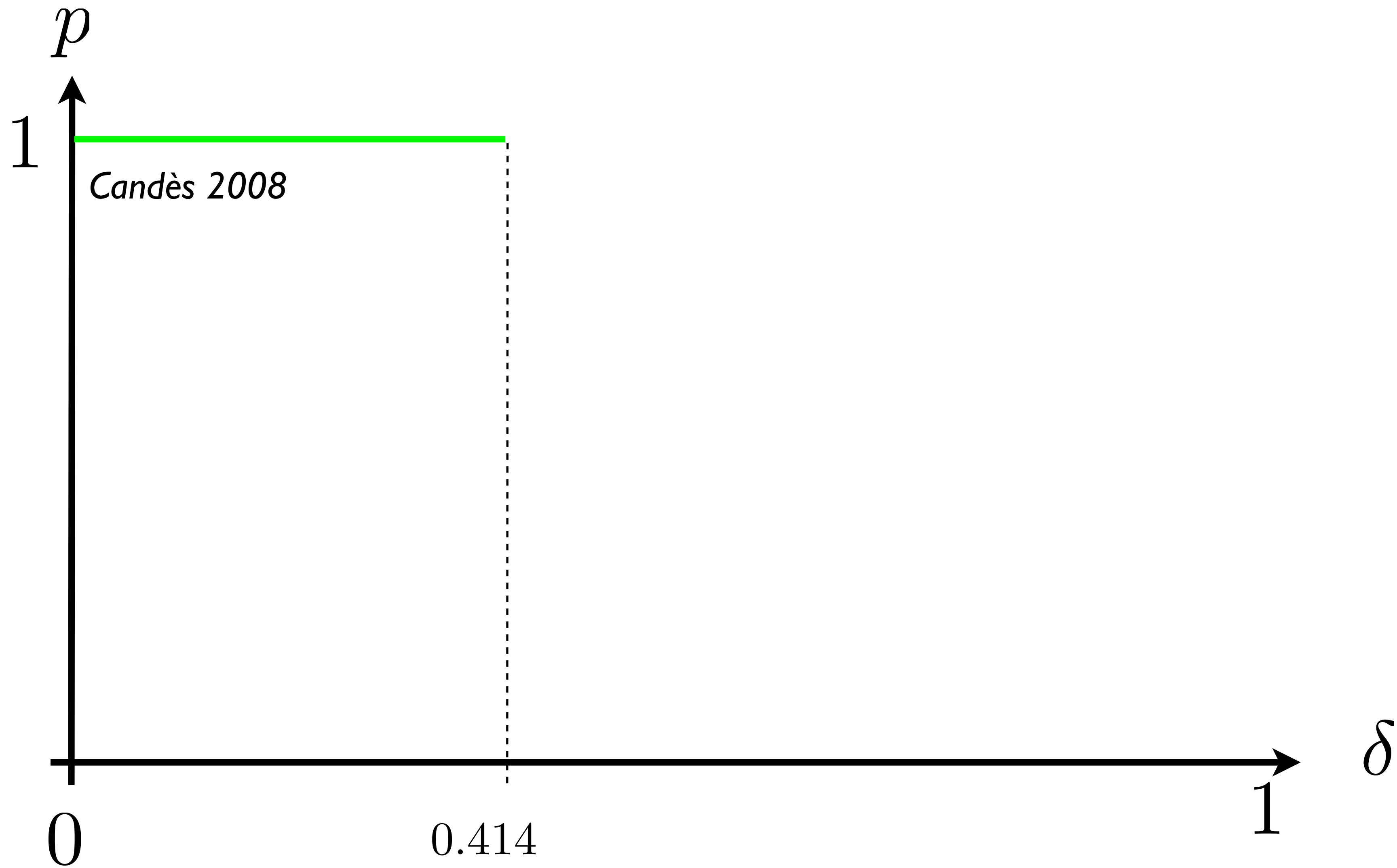
How sharp is the RIP condition ?

- The Null Space Property for Lp
  - "algebraic" + sharp property, only depends on $\mathcal{N}(\mathbf{A})$
  - invariant by linear transforms $\mathbf{A} \rightarrow \mathbf{BA}$
- The RIP(k, δ) condition is "metric" ... and not invariant
  - even with "rescaled" RIP



When does $\delta_{2k}(\mathbf{A}) < \delta$ imply $k \leq k_p(\mathbf{A})$?

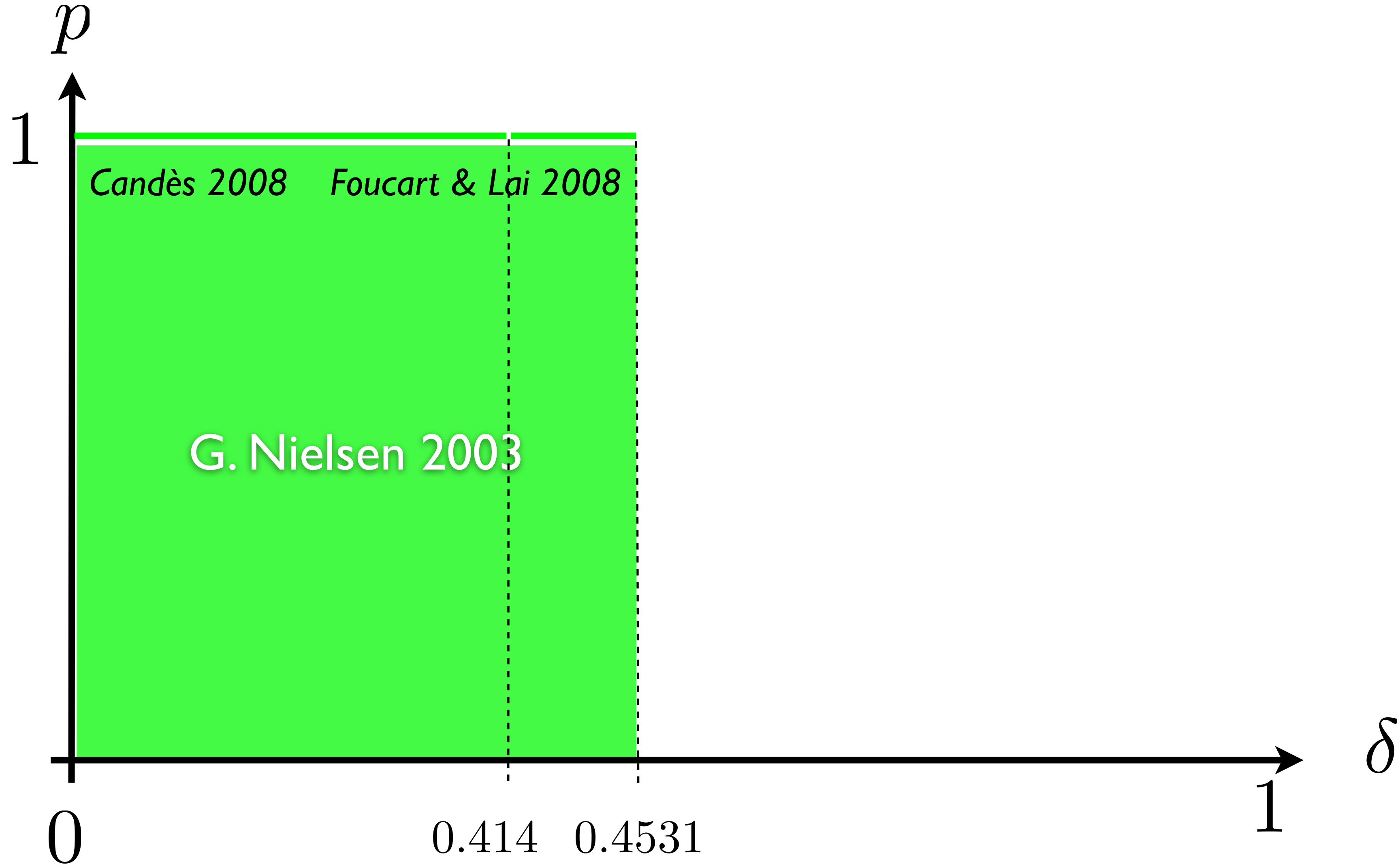
When does $\delta_{2k}(\mathbf{A}) < \delta$ imply $k \leq k_p(\mathbf{A})$?



IRISA

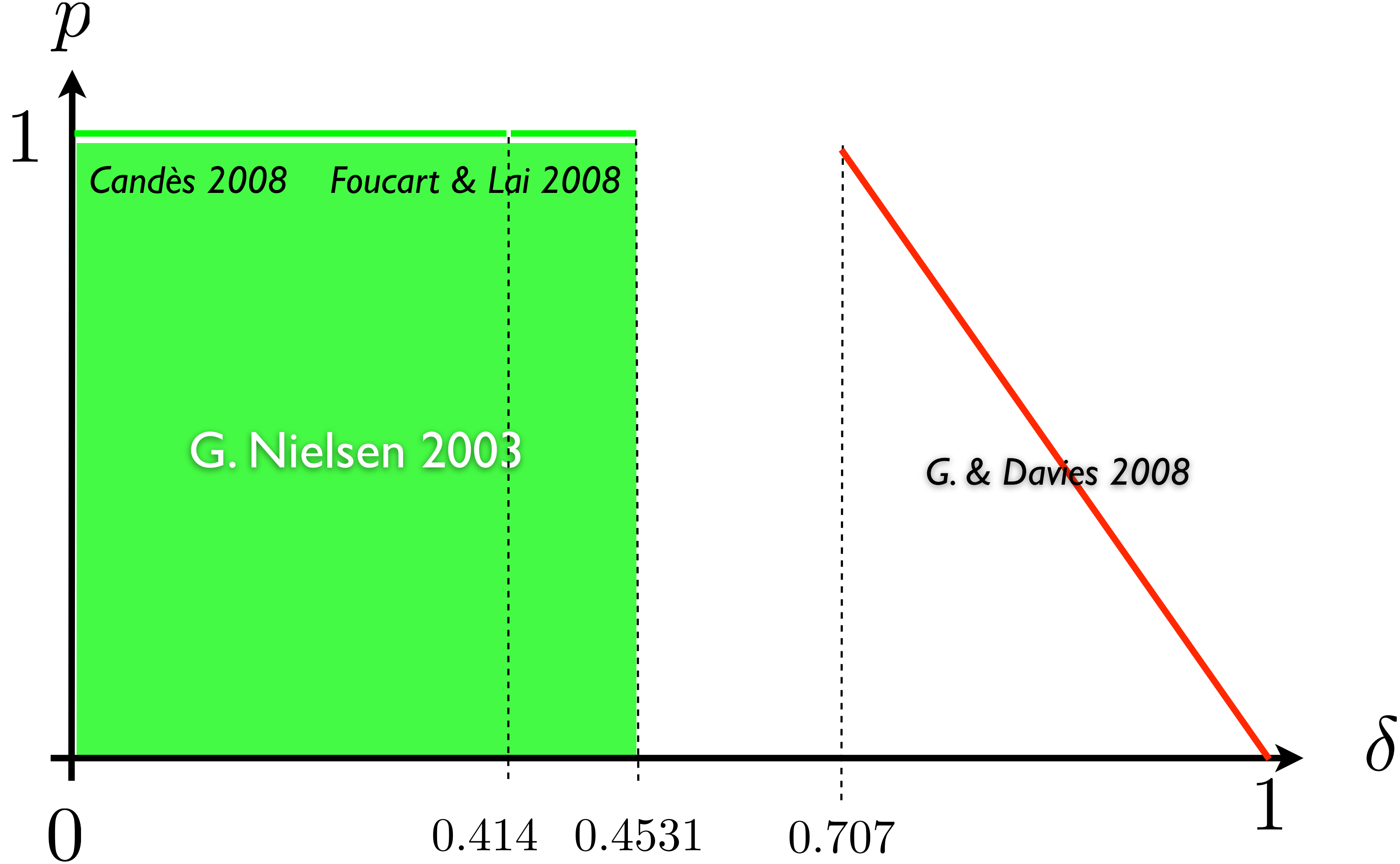
When does $\delta_{2k}(\mathbf{A}) < \delta$ imply $k \leq k_p(\mathbf{A})$?

When does $\delta_{2k}(\mathbf{A}) < \delta$ imply $k \leq k_p(\mathbf{A})$?

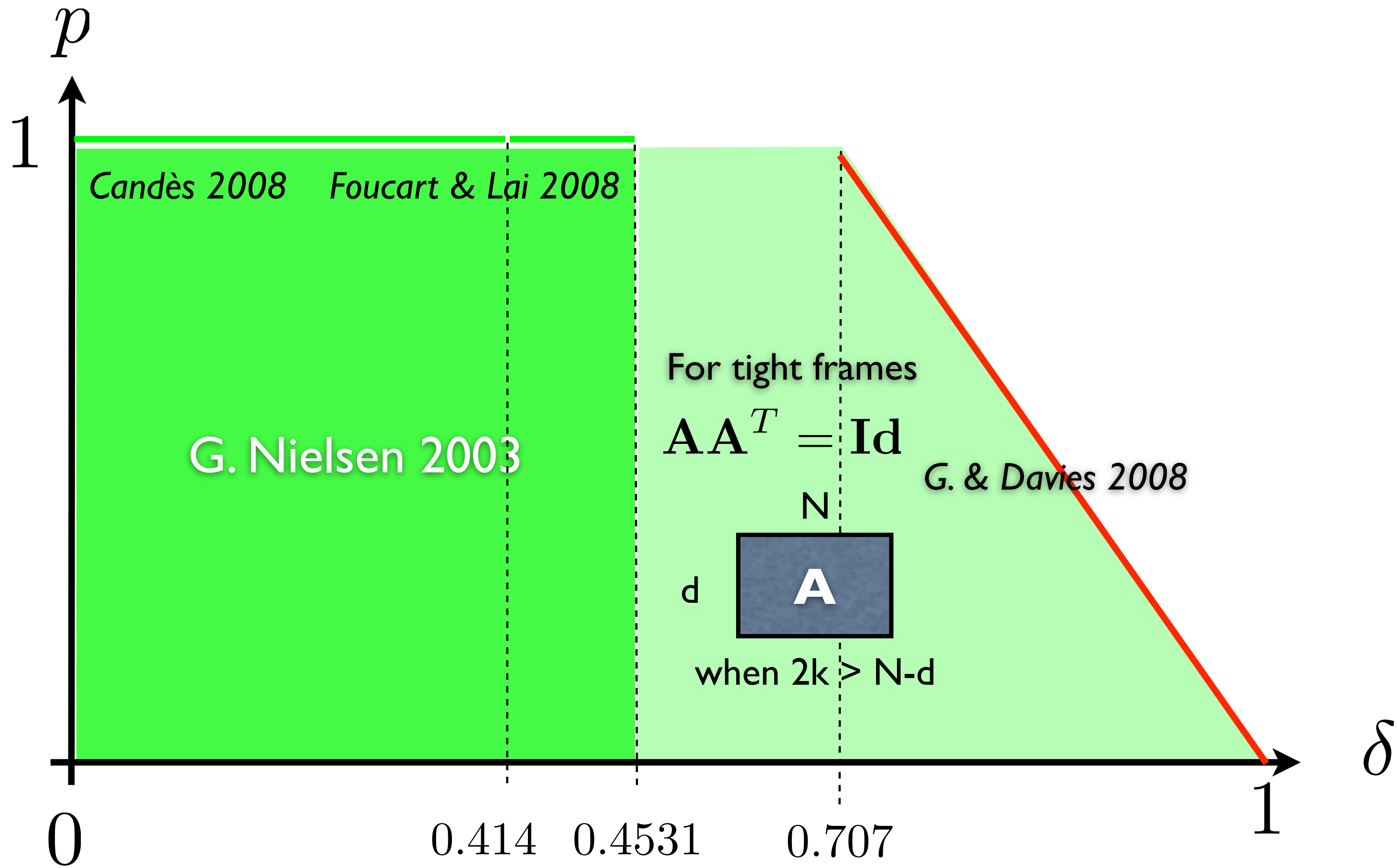


IRISA

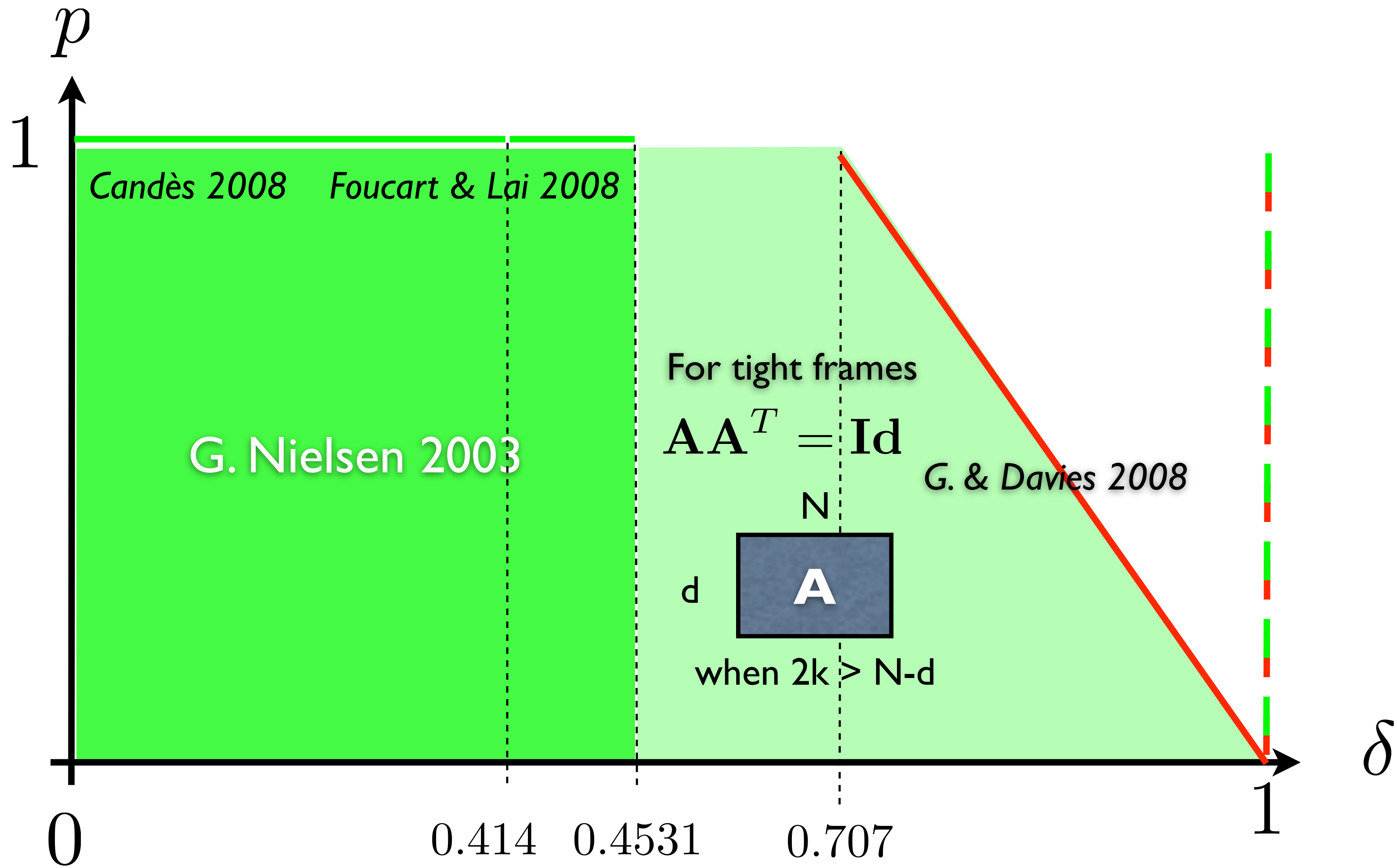
When does $\delta_{2k}(\mathbf{A}) < \delta$ imply $k \leq k_p(\mathbf{A})$?



When does $\delta_{2k}(\mathbf{A}) < \delta$ imply $k \leq k_p(\mathbf{A})$?

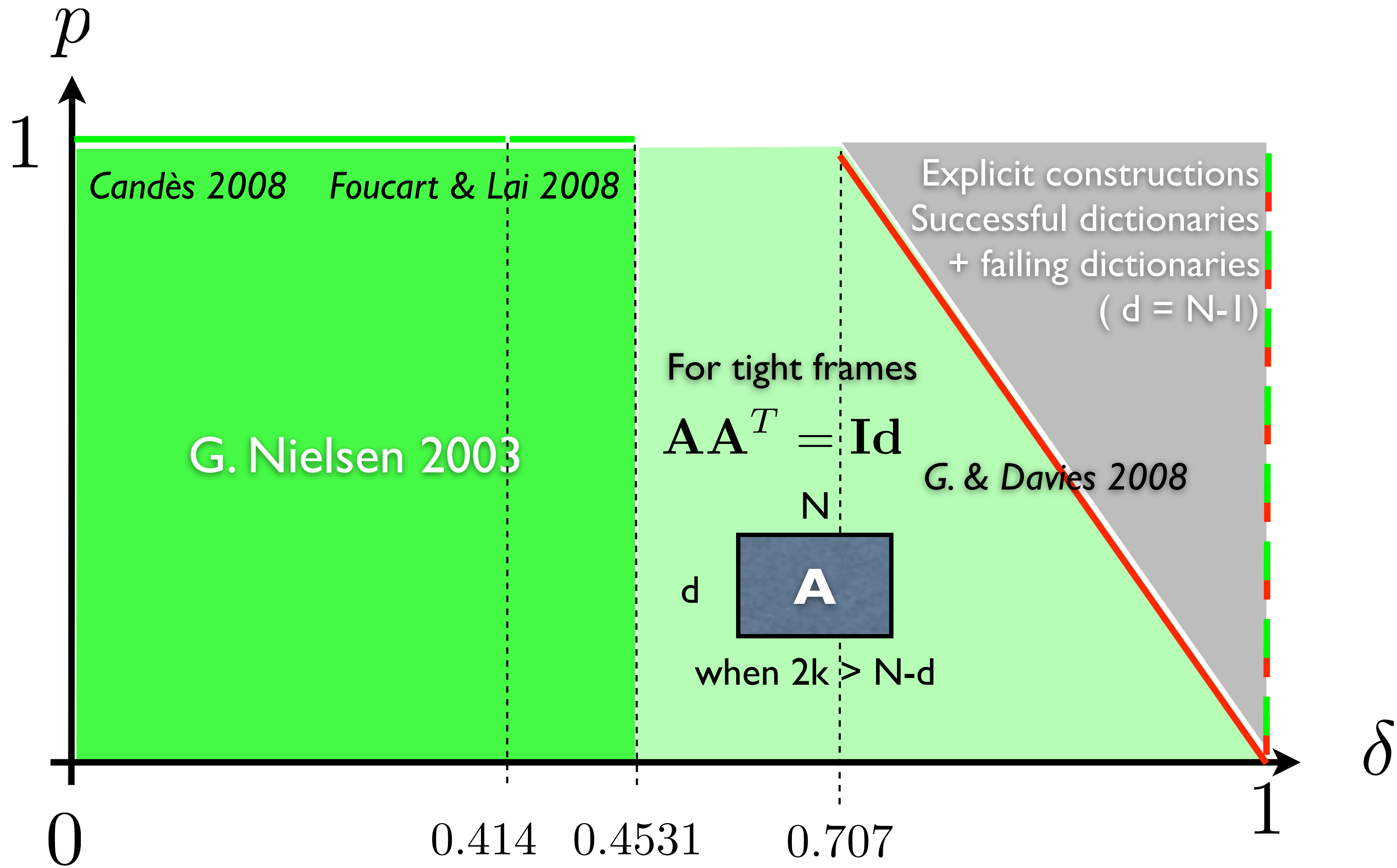


When does $\delta_{2k}(\mathbf{A}) < \delta$ imply $k \leq k_p(\mathbf{A})$?



IRISA

When does $\delta_{2k}(\mathbf{A}) < \delta$ imply $k \leq k_p(\mathbf{A})$?



Summary

- Recovery conditions based on number of nonzero components $\|x\|_0$

$$k_{*\mathrm{MP}}(\mathbf{A}) \leq k_1(\mathbf{A}) \leq k_p(\mathbf{A}) \leq k_q(\mathbf{A}) \leq k_0(\mathbf{A}), \forall \mathbf{A}$$

- **Warning**:
  - there often exists vectors beyond these critical sparsity levels, which are recovered
  - there often exists vectors beyond these critical sparsity levels, where the successful algorithm is not the one we would expect