ProSparse Denoise: Prony's based Sparse Pattern Recovery in the Presence of Noise

Jon Oñativia[+], Yue M. Lu[*] and Pier Luigi Dragotti[+]

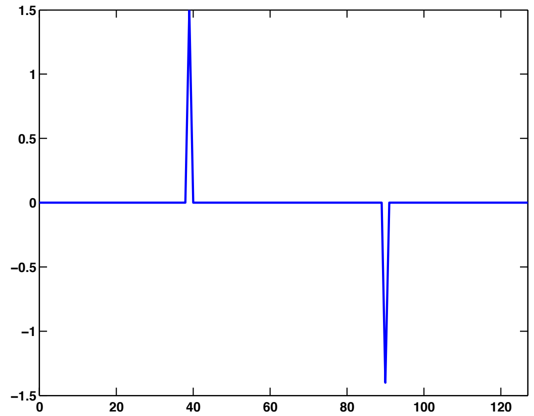
Imperial College London[+]
Harvard University[*]

March 24, 2016

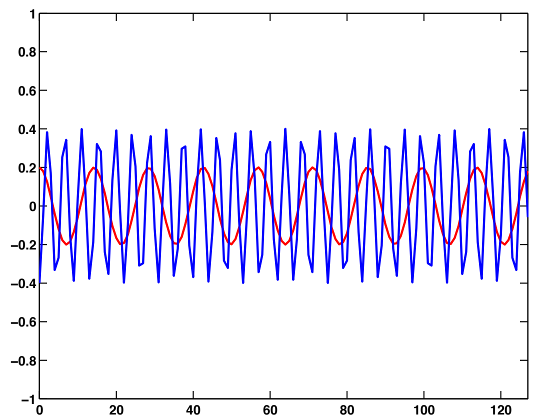
Outline

- Sparse Representation in Pairs of Bases
- ***ProSparse***: A new polynomial time algorithm for sparse signal representation
  - Determinist Bounds
  - Average case performance
  - ProSparse Denoise: signal recovery in the presence of noise

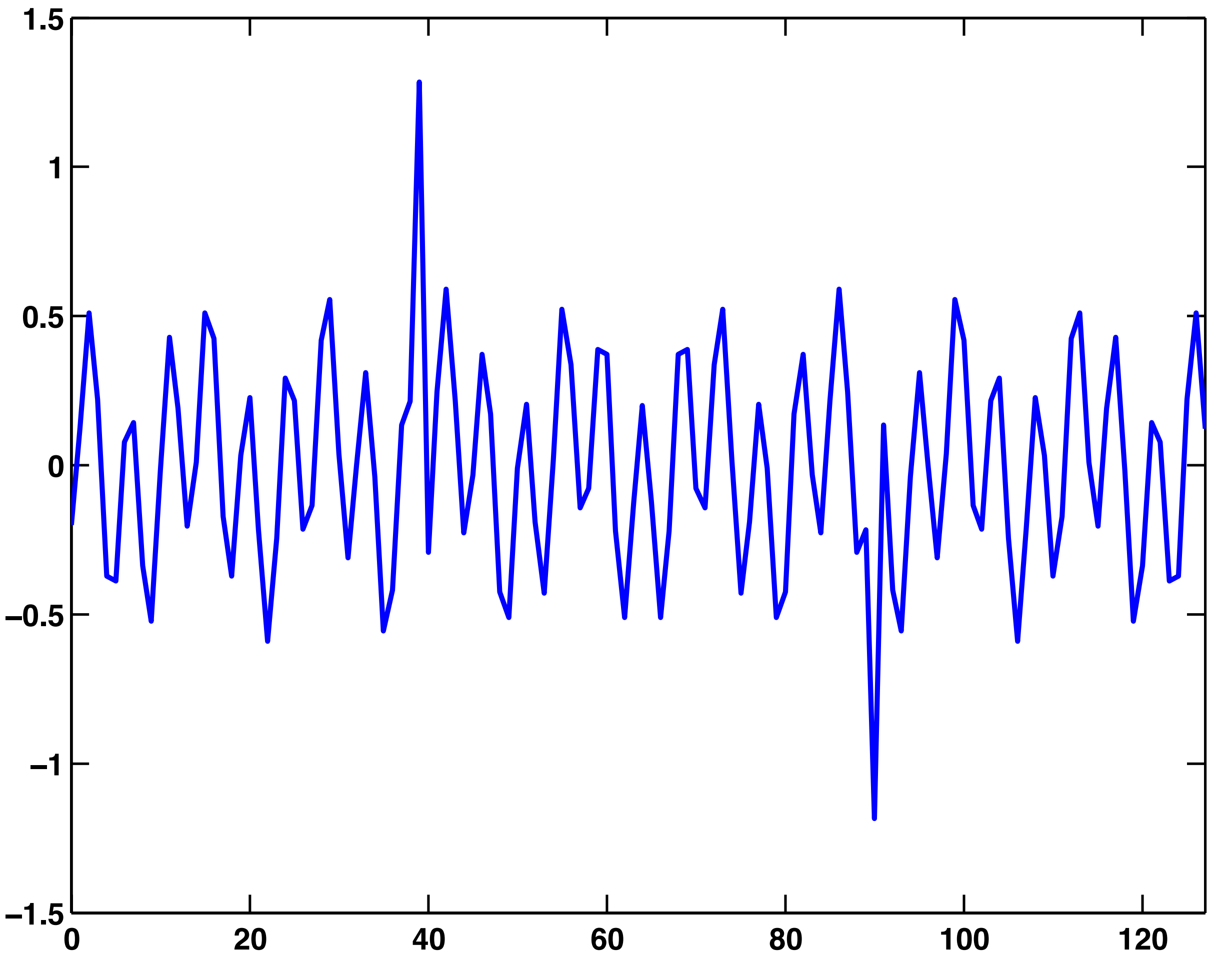
Sparse Representation in Fourier and Canonical Bases



two spikes



two complex exponentials



y (real part plotted)

Goal:

Given y, finds its ***sparse*** representation in Fourier and canonical bases

- Source separation: decompose signals into a smooth part and local innovations

- Prototype for the following problem:

> *Given two bases (or frames)* $D = [\Psi, \Phi]$. *Represent an observed signal as a superposition of a few atoms from* Ψ *and a few atoms from* Φ.

Example: (Curvelets + DCT)

images from [Elad, Starck, Querre, Donoho, 2005]

Sparsity in Fourier and Canonical Bases

Problem formulation:

Assume that

$$\boldsymbol{y} = [\boldsymbol{F}, \boldsymbol{I}] = \boldsymbol{D}\boldsymbol{x}$$

$\boldsymbol{x}$: a (K_p, K_q)-sparse signal. Given $\boldsymbol{y}$, find its sparse representation $\boldsymbol{x}$.

Ideally, solve

$$(P_0): \qquad \arg\min_{\widetilde{\boldsymbol{x}}} \|\widetilde{\boldsymbol{x}}\|_0 \qquad \text{s.t.} \qquad \boldsymbol{y} = \boldsymbol{D}\widetilde{\boldsymbol{x}}$$

Convex relaxation:

$$(P_1): \qquad \arg\min_{\widetilde{\boldsymbol{x}}} \|\widetilde{\boldsymbol{x}}\|_1 \qquad \text{s.t.} \qquad \boldsymbol{y} = \boldsymbol{D}\widetilde{\boldsymbol{x}}$$

[Donoho & Huo, '01]:

- (P_0) has a unique solution when $K = K_p + K_q < \sqrt{N}$
- (P_0) and (P_1) are equivalent when $K < 0.5\sqrt{N}$

Arbitrary Pairs of Orthogonal Bases

[Elad & Bruckstein, 2002]:

Given an arbitrary pair of orthogonal bases $\mathbf{\Psi}$ and $\mathbf{\Phi}$. Define the mutual coherence

$$\mu(\boldsymbol{D}) = \max_{1 \leq k,j \leq M, k \neq j} \frac{|\boldsymbol{d}_k^* \boldsymbol{d}_j|}{\|\boldsymbol{d}_k\|_2 \|\boldsymbol{d}_j\|_2}$$

- (P_0) is unique when $K < 1/\mu(\boldsymbol{D})$
- (P_0) and (P_1) are equivalent when

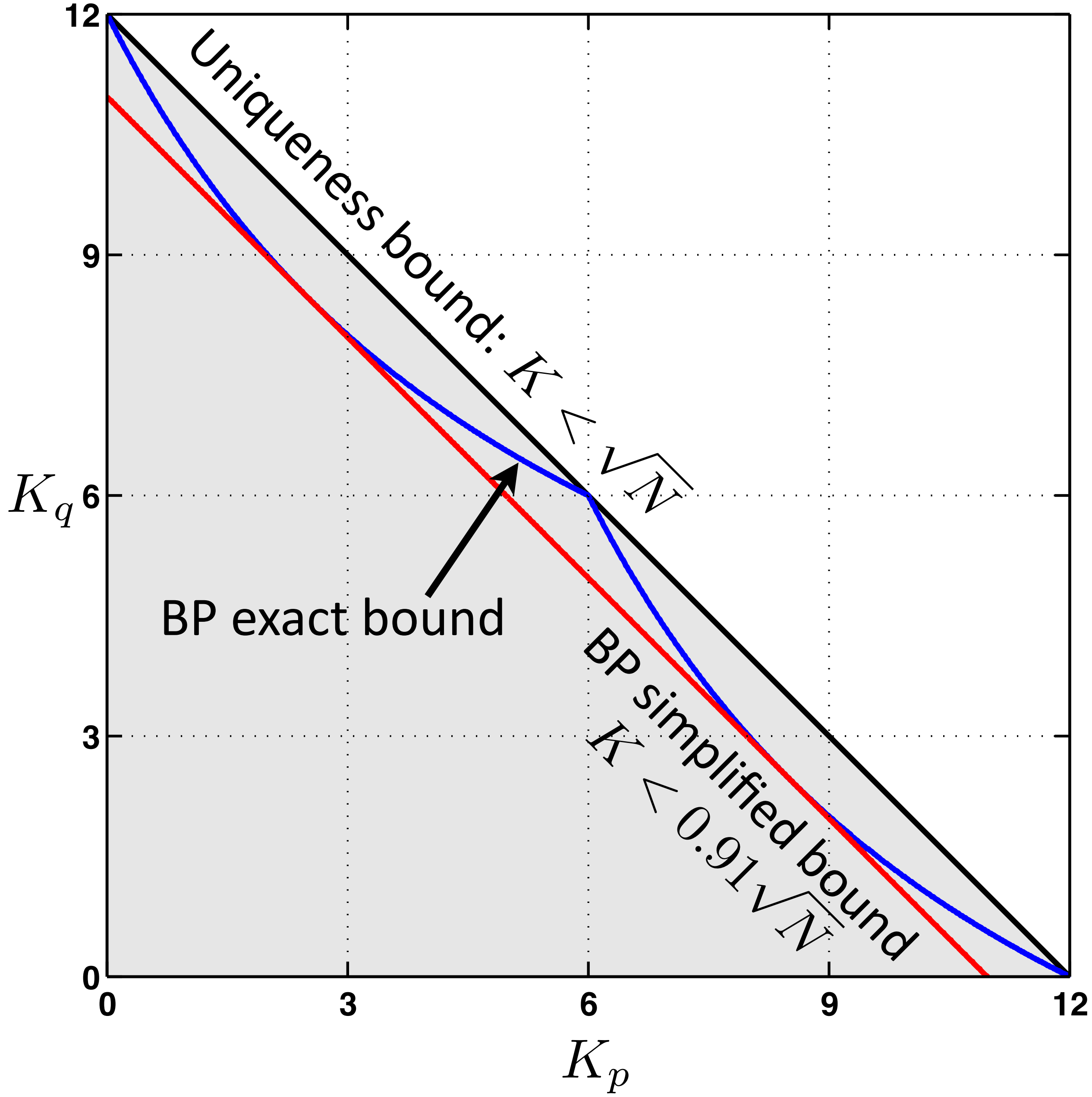
$$2\mu(\boldsymbol{D})^2 K_p K_q + \mu(\boldsymbol{D}) \max\{K_p, K_q\} - 1 < 0 \quad \textit{(Tight Bound)}$$

- Alternatively, (P_0) and (P_1) are equivalent when

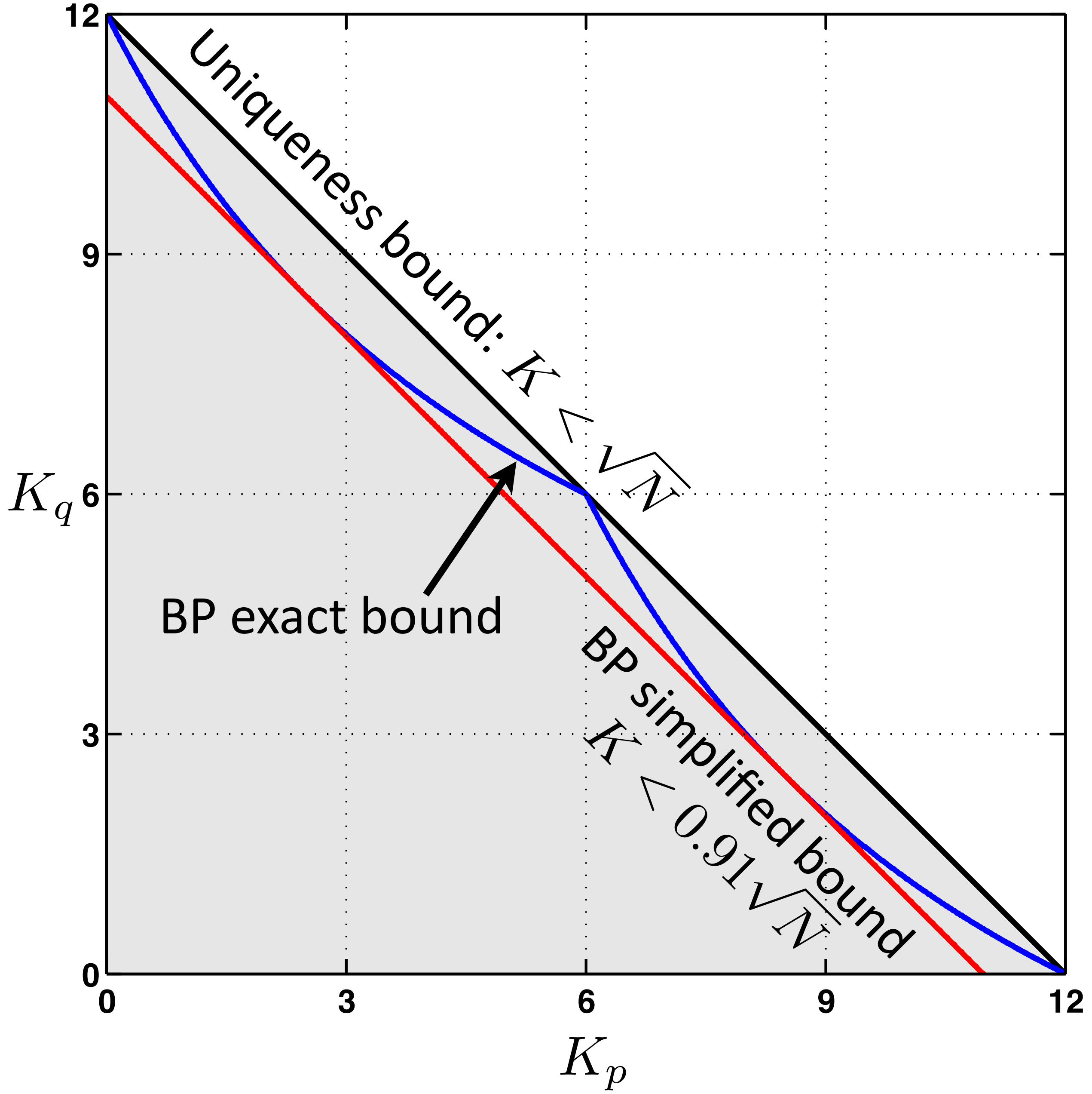
$$K < \sqrt{2} - 0.5/\mu(\boldsymbol{D}) \sim 0.9/\mu(\boldsymbol{D}) \quad \textit{(Weaker Bound)}$$

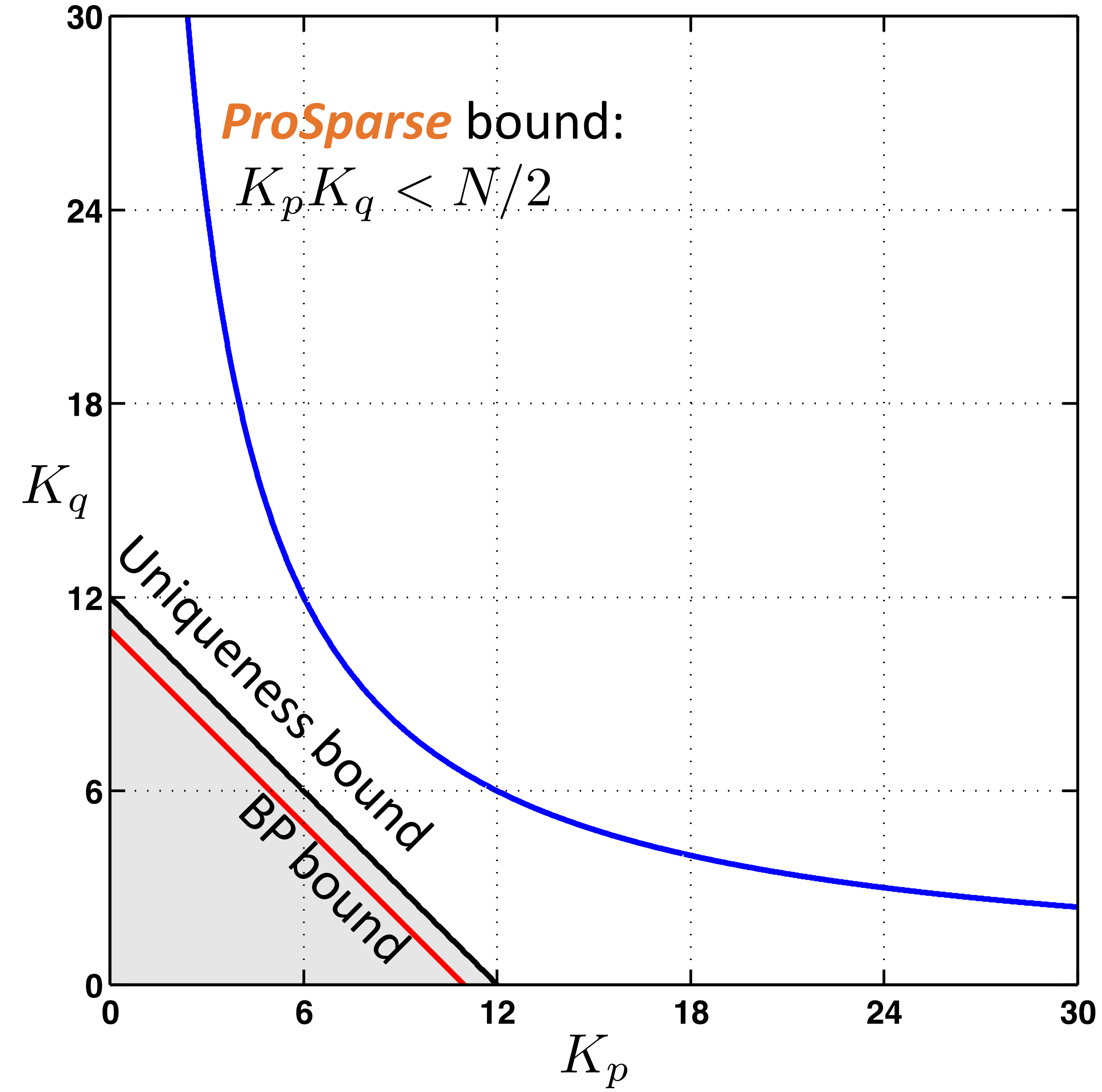
Note: when $\mathbf{\Psi} = \boldsymbol{F}$ and $\mathbf{\Phi} = \boldsymbol{I}$, then $\mu(D) = 1/\sqrt{N}$

Fourier and canonical bases: $N = 144$



Fourier and canonical bases: $N = 144$





ProSparse: Prony's based sparse signal recovery

Consider the case when the signal $\boldsymbol{y} = \boldsymbol{F}_N \boldsymbol{c}$, for some K-sparse vector $\boldsymbol{c} \in \mathbb{R}^N$

The sparse vector $\boldsymbol{c}$ can be reconstructed from ***any*** $2K$ ***consecutive*** entries of $\boldsymbol{y}$

Prony's Method

G. C. F. M. R. de Prony

- The nth entry of $\boldsymbol{y}$:

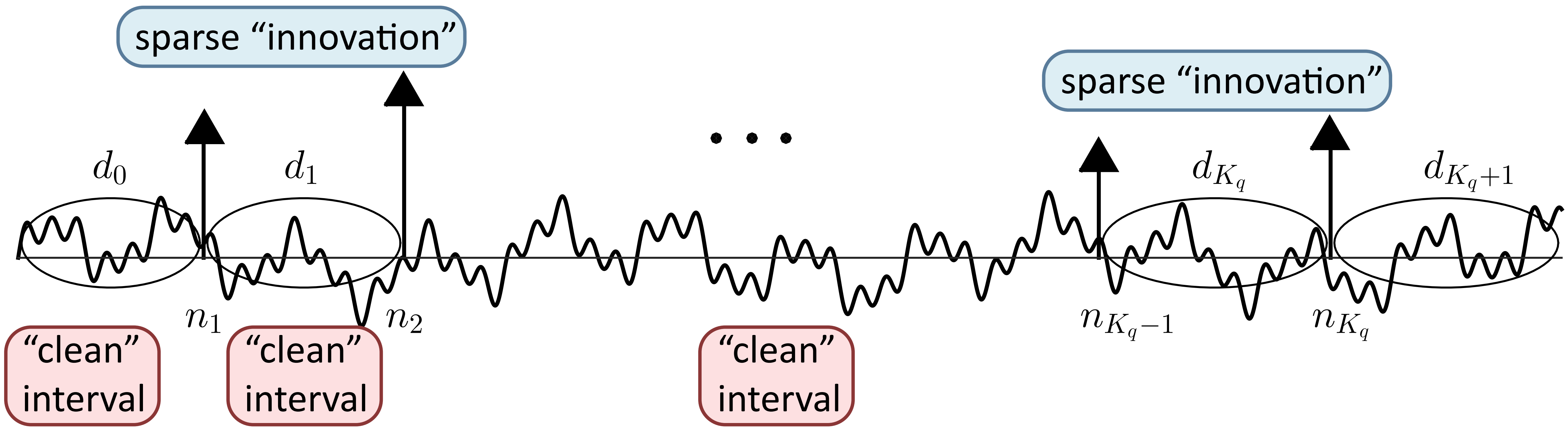
$$y_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{K-1} c_{m_k} \, e^{j2\pi m_k n/N} = \sum_{k=0}^{K-1} \alpha_k u_k^n$$

where $\alpha_k \stackrel{\text{def}}{=} c_{m_k}/\sqrt{N}$ and $u_k \stackrel{\text{def}}{=} e^{j2\pi m_k/N}$

- ***Sparse recovery*** $\longrightarrow$ ***harmonic retrieval***

Applications: harmonic retrieval, ECC, finite rate of innovation sampling, ...

Given $\boldsymbol{y} = [\boldsymbol{F}, \boldsymbol{I}] = \boldsymbol{D}\boldsymbol{x}$, where $\boldsymbol{x}$ is (K_p, K_q)-sparse

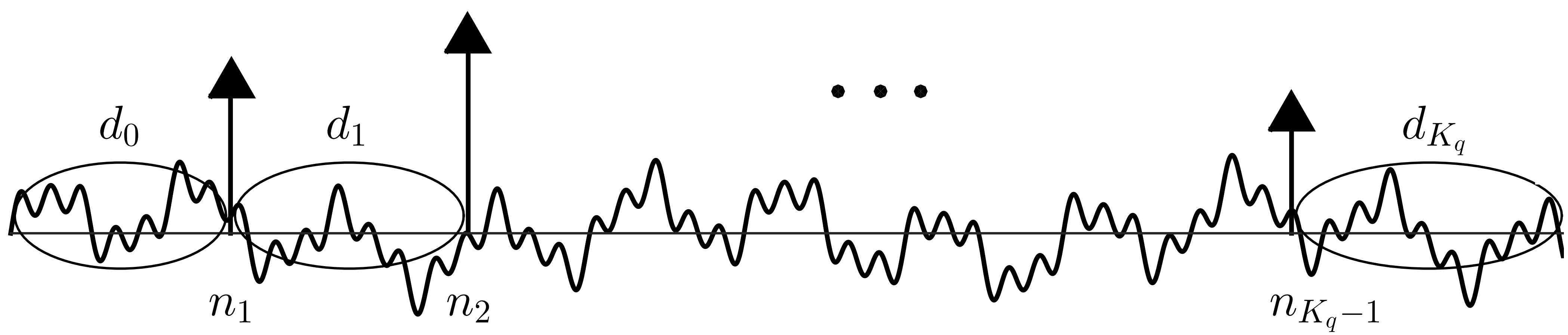


- K_p Fourier atoms —> need a "clean" interval of length $2K_p$
- For sufficiently sparse signals, such intervals always exist
- Sequential search and test: polynomial complexity

Theorem [Dragotti & Lu, 2014]: Let $\boldsymbol{D} = [\boldsymbol{F}, \boldsymbol{I}]$ and $\boldsymbol{y} \in \mathbb{C}^N$ an arbitrary signal.

There exists an algorithm, with a worst-case complexity of $\mathcal{O}(N^3)$, that finds ***all*** (K_p, K_q)-sparse signal $\boldsymbol{x}$ such that

$$\boldsymbol{y} = \boldsymbol{D}\boldsymbol{x} \quad \text{and} \quad K_p K_q < N/2.$$



$$\sum_{0 \le i \le K_q} d_i = N - K_q \quad \Rightarrow \quad \max\{d_i\} \ge \frac{N - K_q}{K_q + 1} \ge 2K_p$$

Works for $\boldsymbol{D} = [\boldsymbol{\Psi}, \boldsymbol{\Phi}]$ if the columns of $\boldsymbol{F}\boldsymbol{\Psi}^*\boldsymbol{\Phi}$ have ***localized supports***

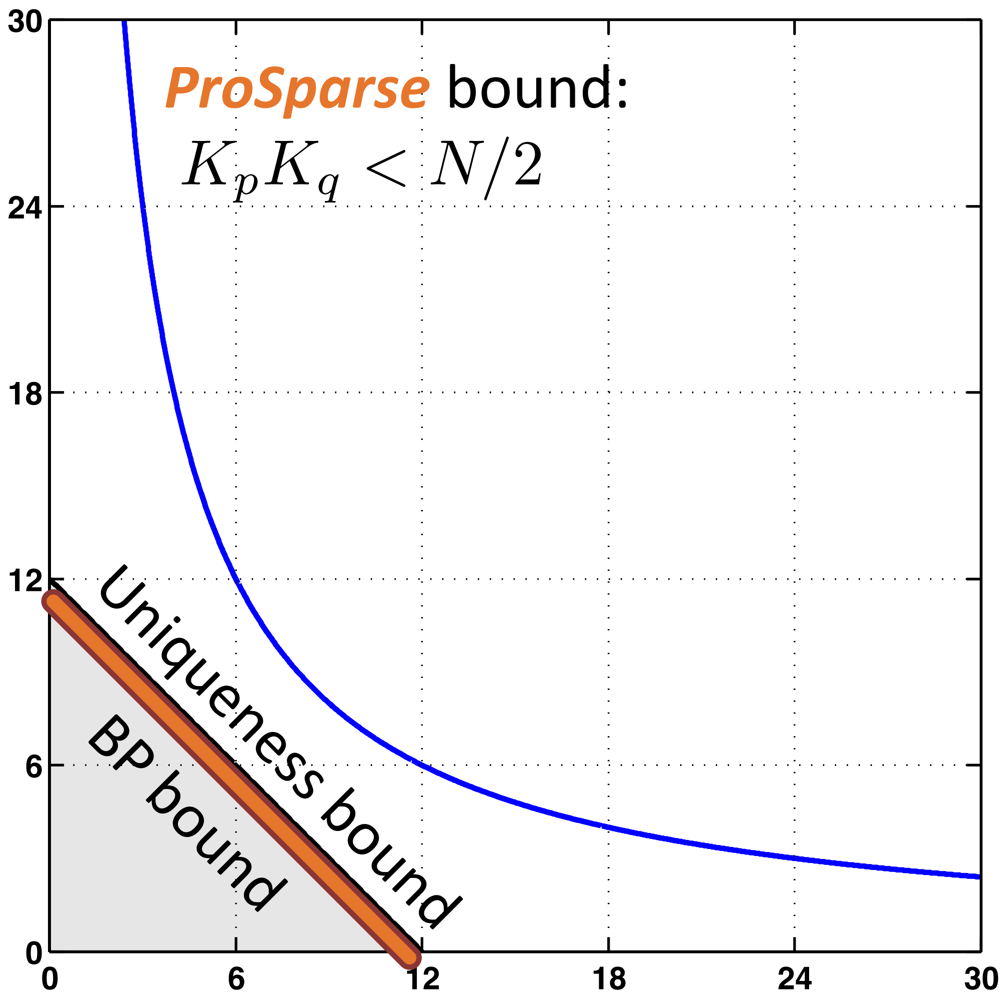
More generally:

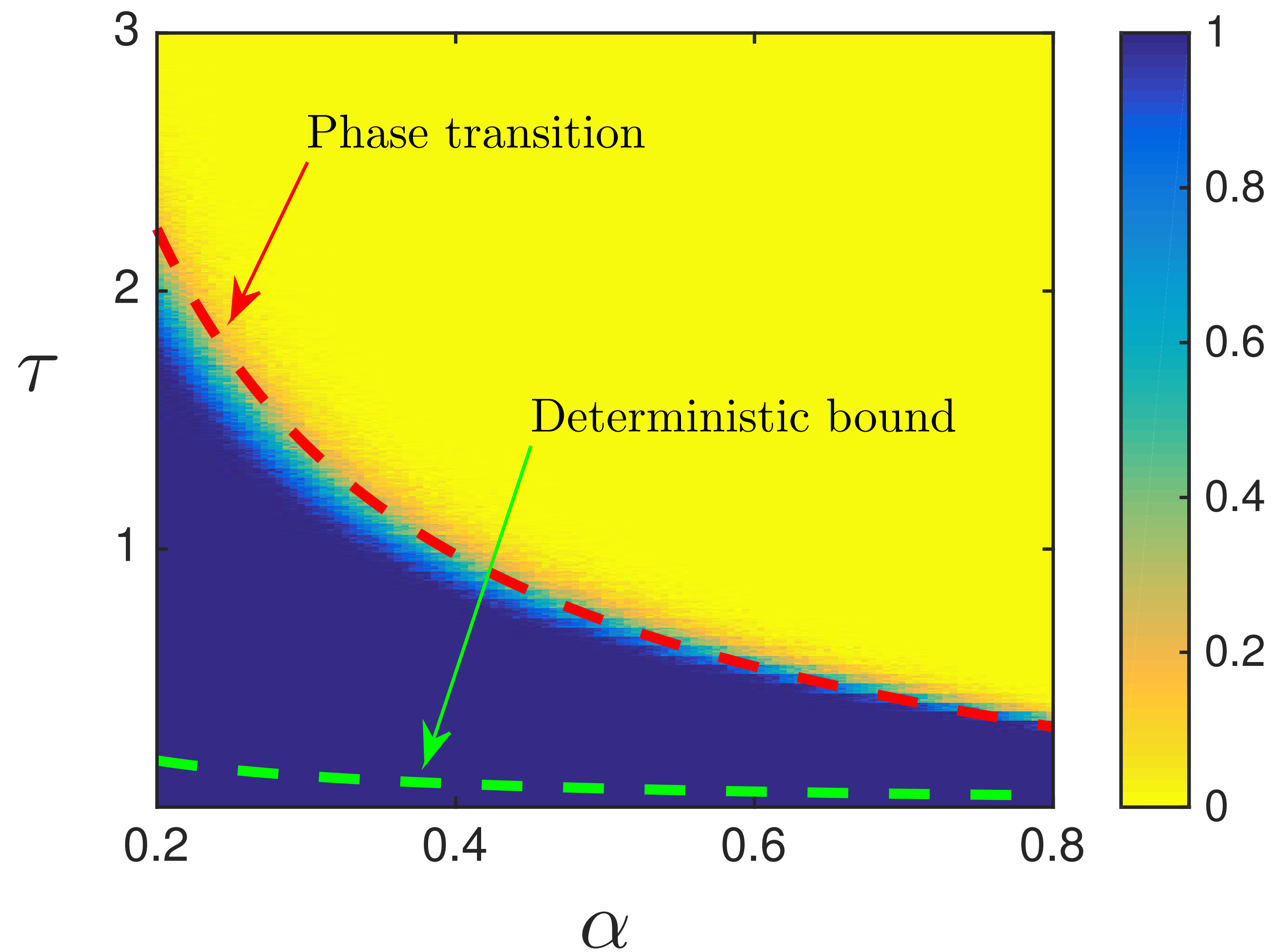
$$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{s}$$

- $\boldsymbol{s}$: noise with local "footprints"
- $\boldsymbol{x}$: a "***locally reconstructable***" signal

Examples:

- Sparse in Fourier, DCT, random bases or frames …
- Continuous sparse sinusoids: $x_n = \sum_k c_k e^{j\omega_k n}$
- Low-dimensional subspace





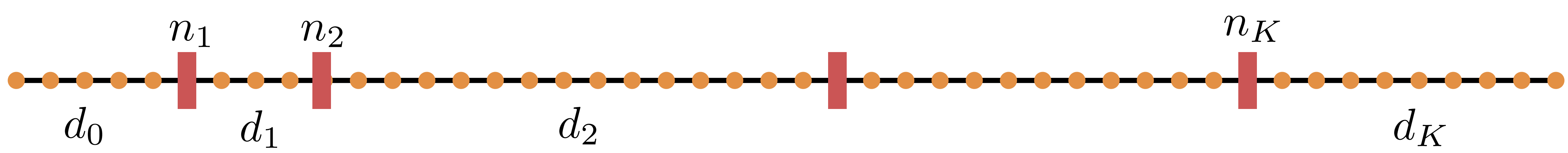
$K_q = \alpha N$ for $0 < \alpha < 1$

Bound: $K_p < 1/(2\alpha)$

In practice: $K_p < \tau(\alpha) \log N$

N consecutive integers

Randomly select K integers (sampling w/o replacement)



Joint distribution of the interval lengths:

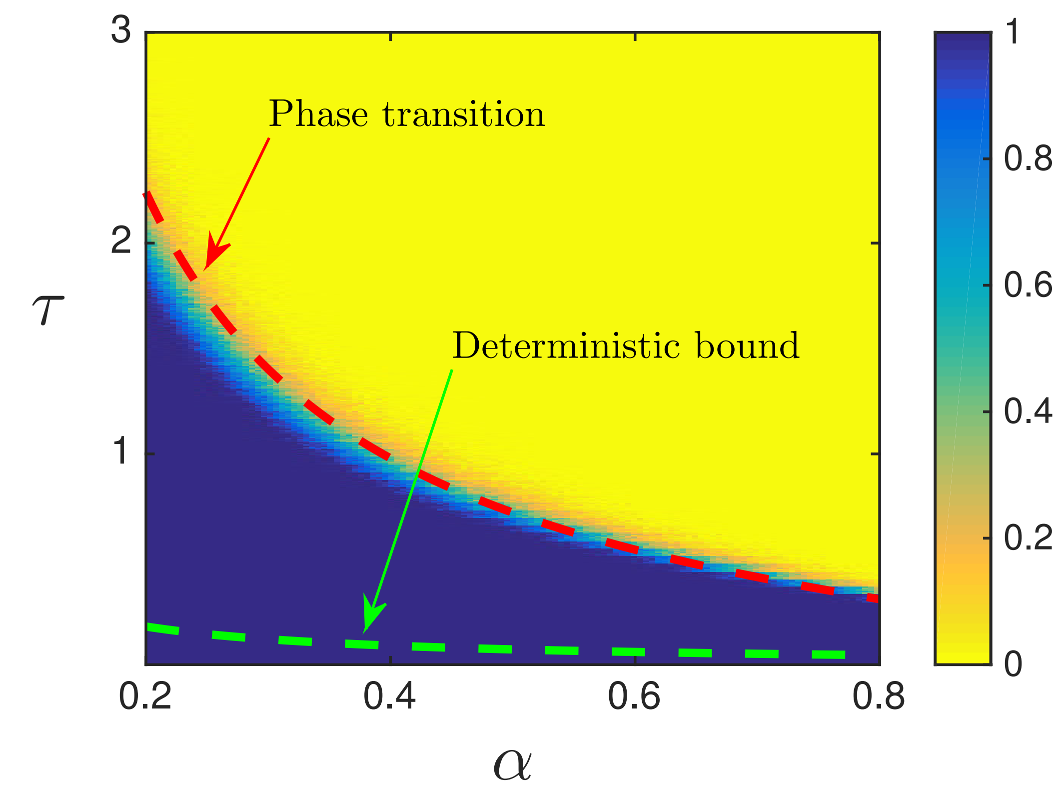
$$\mathbb{P}(d_0, d_1, \ldots, d_K) = \frac{1}{\binom{N}{K}} \mathbb{1}\left(\sum_k d_k = N - K\right) \qquad \text{for } d_k = 0, 1, 2, \ldots$$

Related to ***Bose-Einstein distribution*** in statistical physics

Proposition [Oñativia, Dragotti & Lu, 2015]:

Let $K = \lfloor \alpha N \rfloor$ for some $0 < \alpha < 1$

$$\lim_{N\to\infty} \frac{\max_k d_k}{\log N} = \frac{-1}{\log(1-\alpha)} \stackrel{\text{def}}{=} \tau^*(\alpha) \quad \textbf{\textit{in prob.}}$$



Corollary: Let $\boldsymbol{y} \in \mathbb{C}^N$ be a linear combination of $K_p = \tau \log N$ Fourier atoms and $K_q = \lfloor \alpha N \rfloor$ spikes. If the locations of the spikes are sampled uniformly at random, then

$$\lim_{N\to\infty} \mathbb{P}(\text{ProSparse succeeds}) = \begin{cases} 1, & \text{if } \tau < \tau^*(\alpha) \\ 0, & \text{if } \tau > \tau^*(\alpha) \end{cases}$$

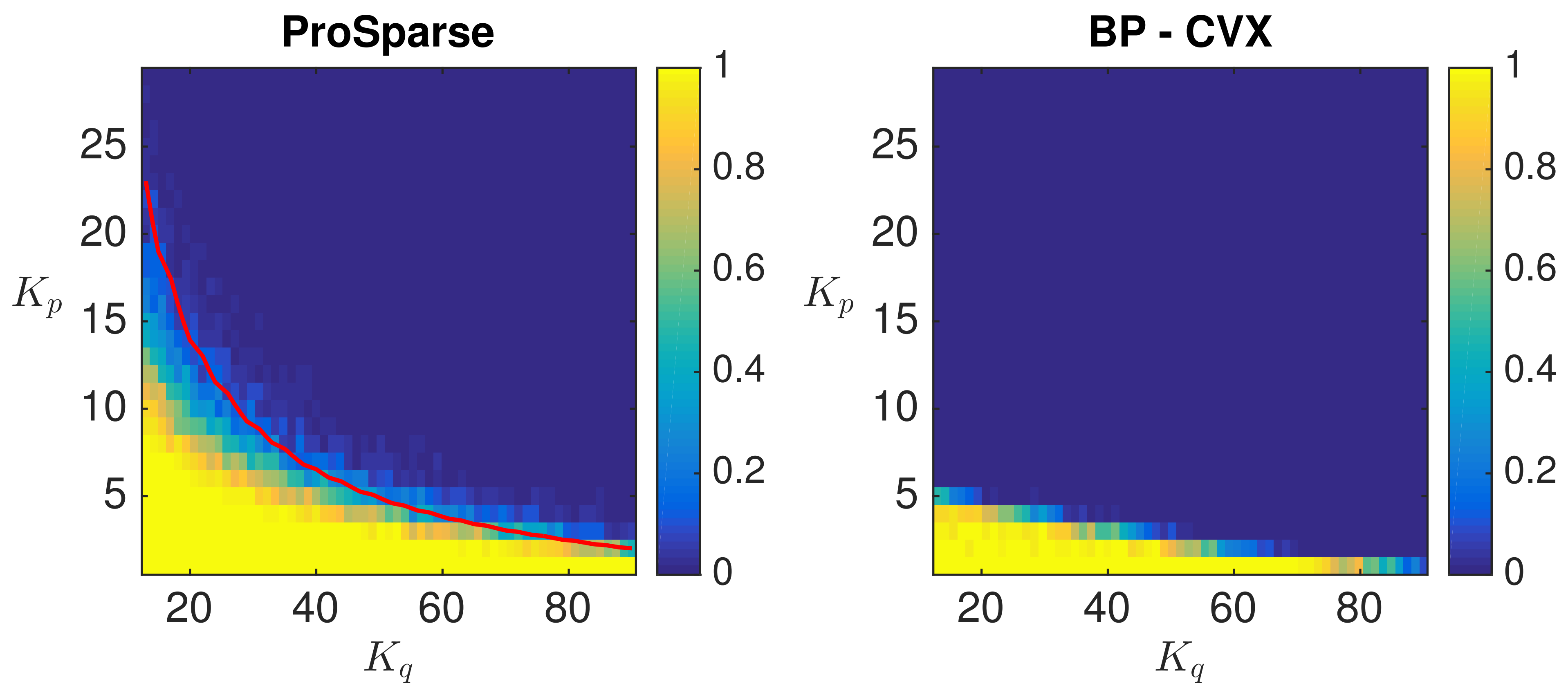
Comparing with BP (Average-Performance)

BP: $K_p + K_q \doteq cN/\sqrt{\log N}$ [Candes & Romberg, 2006]

ProSparse: $K_p = \tau(\alpha)\log N, K_q = \alpha N$

But:

- ProSparse only depends on the distribution of spike locations;
- Fourier frames
- Arbitrary coefficient distributions

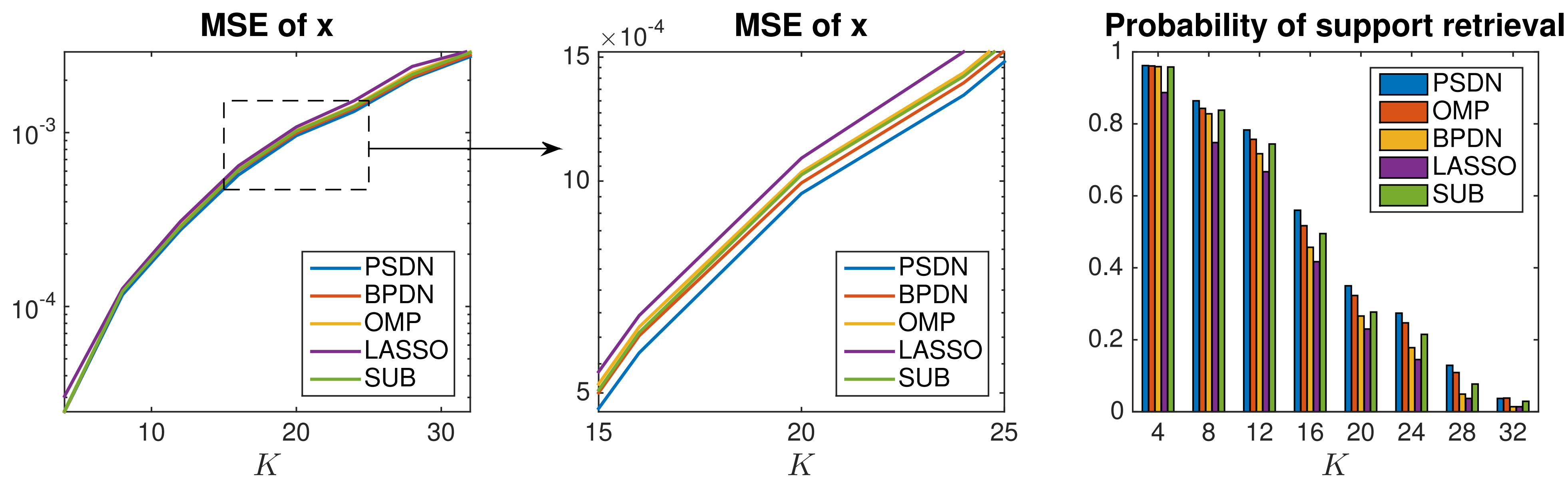


ProSparse Denoise

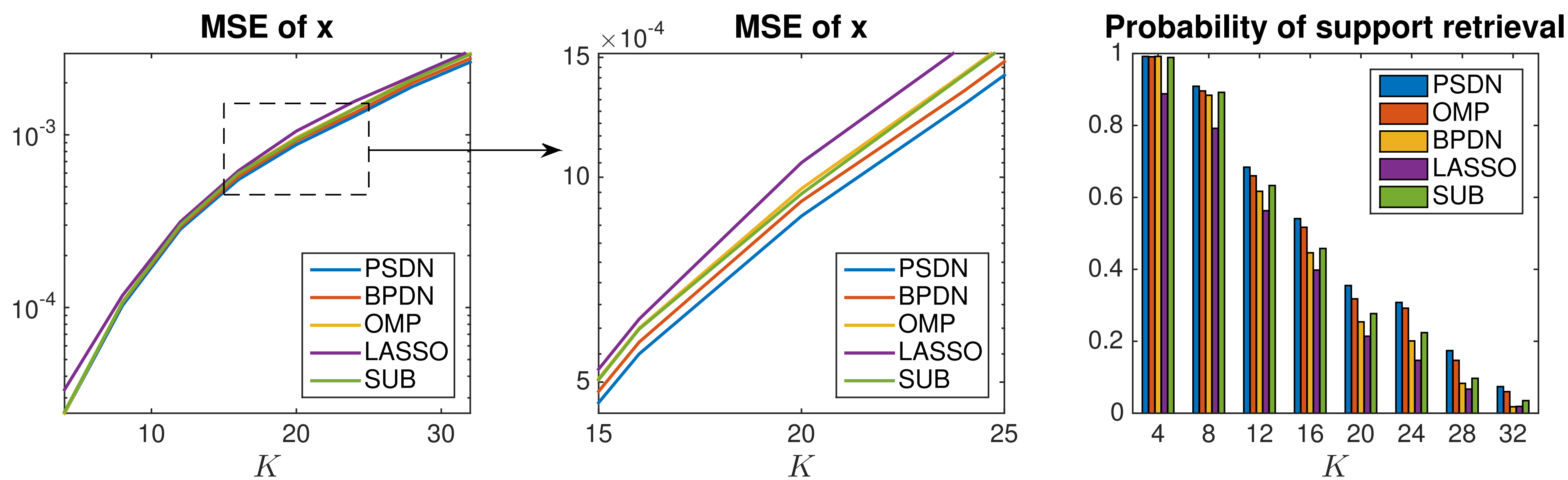
- ***Setting***: $\mathbf{y} = [\mathbf{F}, \mathbf{I}][\mathbf{x}_p^T, \mathbf{x}_q^T] + \epsilon$ where *the noise* is i.i.d. Gaussian
- ***Key Ingredients of ProSparse Denoise***:
  - Replace Prony's with a noise resilient version: Cadzow algorithm
  - Treat Spikes as noise
- ***Algorithm:***

1. Estimate the K_p Fourier atoms using Cadzow
2. Remove this contribution from ***y***, estimate the largest spike from the residual and remove it from ***y***
3. Repeat steps 1 and 2, K_q times.
4. Estimate the spikes using duality

Simulation Results

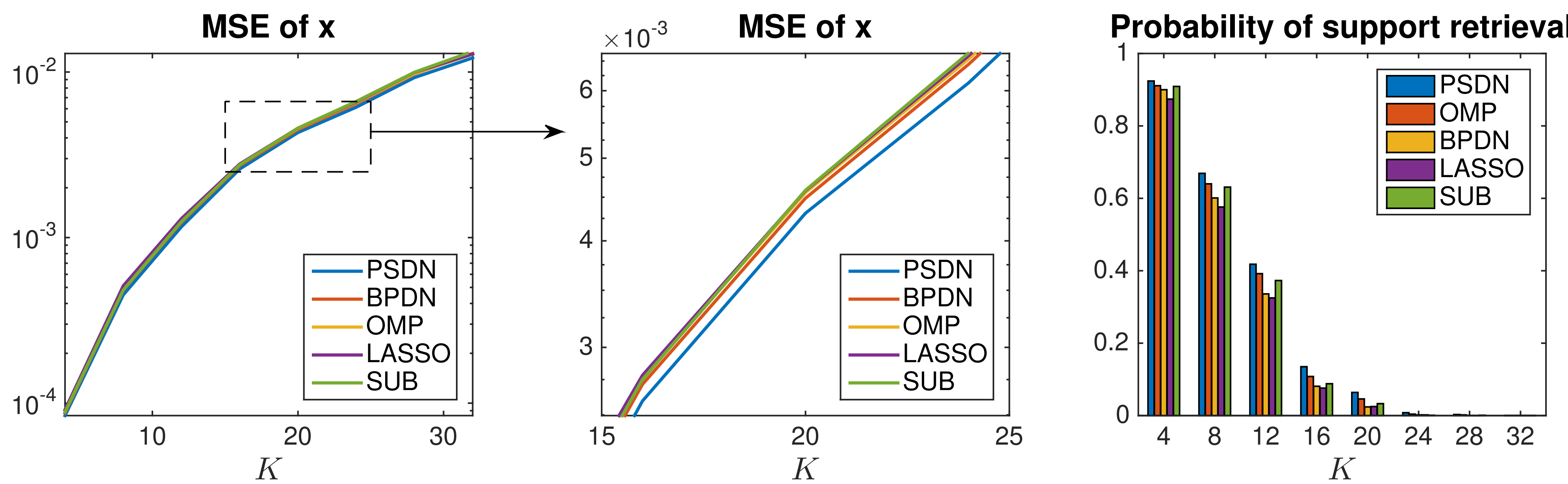


(a) SNR = 10 dB, bias = 50%.

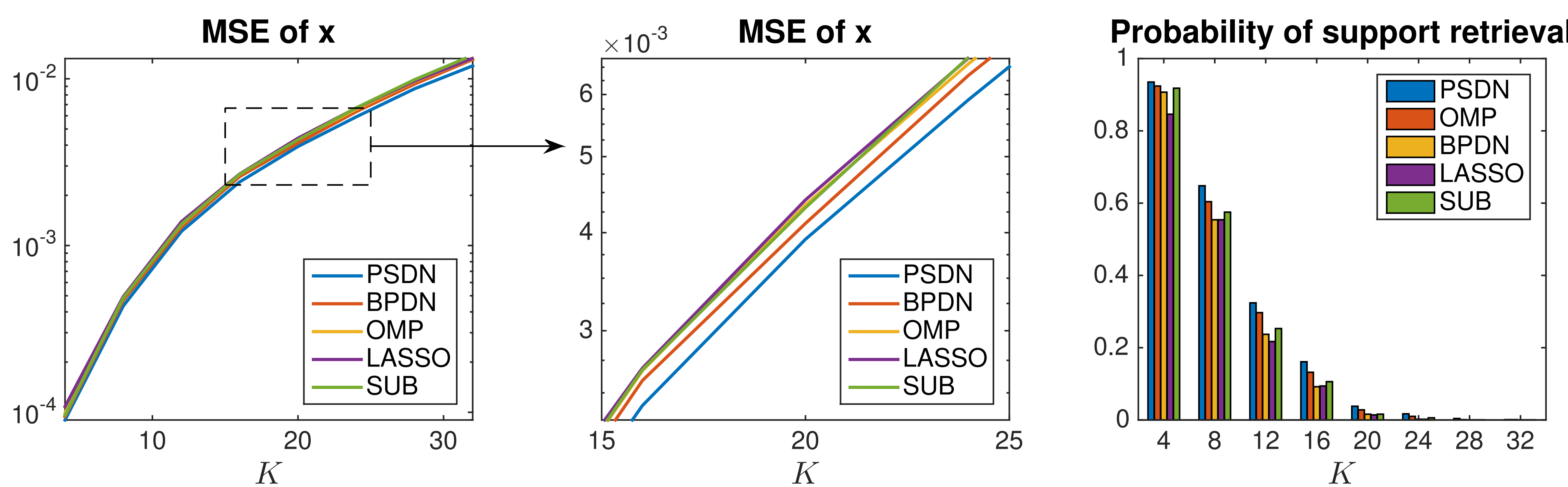


(c) SNR = 10 dB, bias = 25%.

Simulation Results (cont'd)



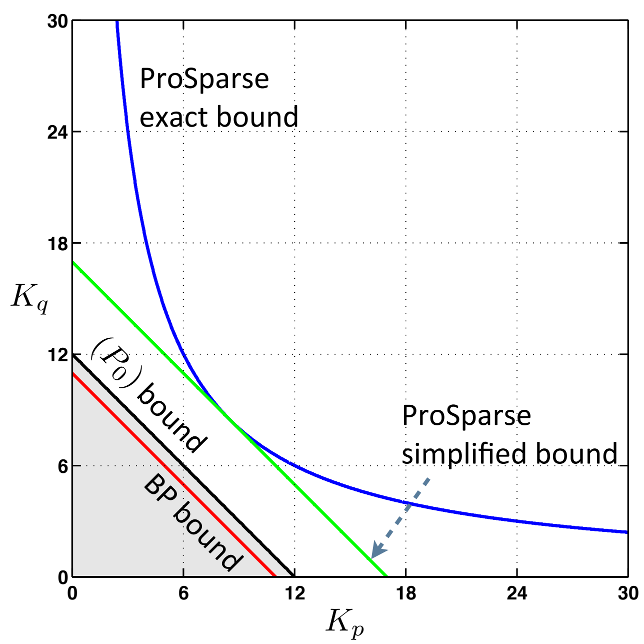
(b) SNR = 5 dB, bias = 50%.

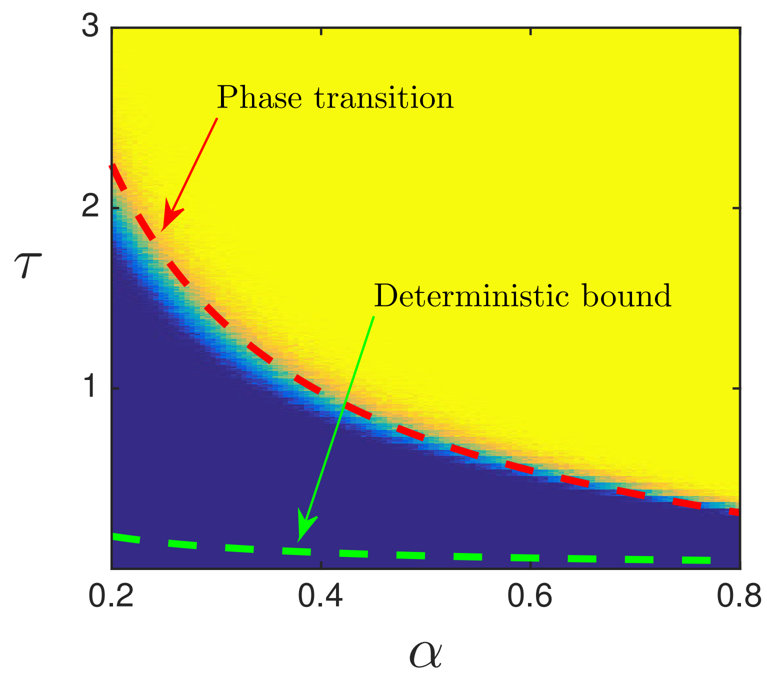


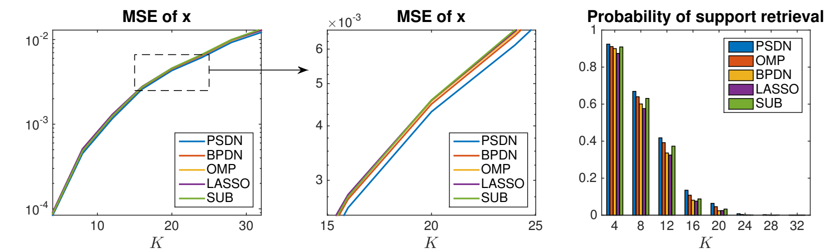
(d) SNR = 5 dB, bias = 25%.

Summary

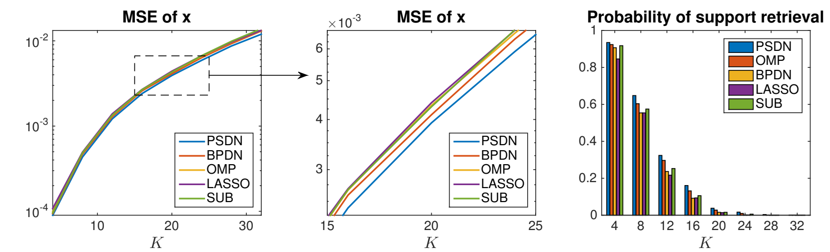
- ***ProSparse***: a polynomial time algorithm that decomposes a signal into a sum of a sparse signals and a locally-reconstructable signal
- ProSparse is based on mapping sparse representation problem with Structured Least Squares Methods
- For Fourier + Identity, deterministic bound is better than BP and unicity bounds
- Tight Bound on Average case performance
- Promising denoising results
- How far can the basic ideas behind ProSparse be extended?







(**b**) SNR = 5 dB, bias = 50%.



(**d**) SNR = 5 dB, bias = 25%.

Deterministic results:

P. L. Dragotti and Y. M. Lu, "On Sparse Representation in Fourier and Local Bases," IEEE Transactions on Information Theory, vol. 60, no. 12, pp. 7888-7899, 2014.

Average-case performance:

Paper with full proofs will be posted on arXiv soon.

ProSparse Denoise:

J. Onativia, Y.M. Lu and P.L. Dragotti, ICASSP 2016