Demixing Sparse Signals via Convex Optimization

Yi Zhou*, **Yingbin Liang***
*Department of EECS, Syracuse University

Signal Mixture Model

Signal $\mathbf{z}_0 \in \mathbb{R}^n$ is a mixture of $\mathbf{x}_0, \mathbf{y}_0 \in \mathbb{R}^n$:

$$\mathbf{z}_0 := \mathbf{x}_0 + \mathbf{y}_0. \tag{1}$$

1. Image feature decomposition, image denoising, signal separation;
2. How to formulate a proper demixing model?
3. What is the theorectical performance guarantee of the model?

Model Formulation

Sparse signal components: orthonormal basis $\Psi, \Phi \in \mathbb{R}^{n\times n}$

$$\theta_x := \Psi \mathbf{x}_0, \quad \theta_y := \Phi \mathbf{y}_0. \quad \mathrm{card}(\theta_x), \mathrm{card}(\theta_y) \ll n.$$

Convex Optimization Model: for some $\lambda > 0$

$$\min_{\mathbf{x},\mathbf{y}\in\mathbb{R}^n} \|\Psi\mathbf{x}\|_1 + \lambda\|\Phi\mathbf{y}\|_1, \quad \text{s.t.} \quad \mathbf{x}+\mathbf{y}=\mathbf{z}_0. \tag{P}$$

1. Seek for feasible decomposition with minimum ℓ_1 norm (convex);
2. When is **(P)** exact, i.e., $(\mathbf{x}_0, \mathbf{y}_0)$ be its unique solution pair?

Review of Existing Work

Separability relies on the coherence between Ψ, Φ:

Mutual Coherence [1]: $\mu(\Psi, \Phi) := \max_{i,j} |\langle \psi_i, \phi_j\rangle|$.

Let $\mathrm{supp}(\theta_x)$ be fixed and $\mathrm{supp}(\theta_y)$ be uniformly random. Then **(P)** is exact w.h.p provided that $\mathrm{card}(\theta_x) + \mathrm{card}(\theta_y) \le \mathcal{O}\left(1/\mu^2(\Psi,\Phi)\log^6 n\right)$.

- In general, $\mu \in [\frac{1}{\sqrt{n}}, 1]$;
- Mutual coherence barrier: $\mu(\Psi, \Phi) = 1$.

Cluster Coherence [2]: $\mu(\Psi_\Omega, \Phi) := \max_j \sum_{i\in\Omega} |\langle \psi_i, \phi_j\rangle|$.

(P) is exact in the asymptotic regime (i.e., for all $j \to n$) near perfectly provided that the corresponding cluster coherence vanishes.

- Characterize asymptotic exactness;
- We want a quantative local exactness condition.

Local Subspace Coherence

Assumption 1.

The signal components $\mathbf{x}_0, \mathbf{y}_0$ satisfy:
1. $\mathrm{supp}(\theta_x)$ is fixed, while $\mathrm{supp}(\theta_y)$ satisfies $\mathbb{P}(j \in \mathrm{supp}(\theta_y)) \sim \mathrm{Bernoulli}(p_j)$;
2. $\mathrm{sgn}(\theta_x), \mathrm{sgn}(\theta_y)$ take values from $\{+1, -1\}$ with equal probability.

Signal Subspace:

$$\mathcal{X} := \mathrm{span}\{\psi_j, j \in \mathrm{supp}(\theta_x)\}, \quad \mathcal{Y} := \mathrm{span}\{\phi_j, j\in \mathrm{supp}(\theta_y)\}. \tag{2}$$

- $\mathbf{x}_0 \in \mathcal{X}$, and $\mathbf{y}_0 \in \mathcal{Y}$; $\mathrm{P}_\mathcal{X}, \mathrm{P}_\mathcal{Y}$ projections.

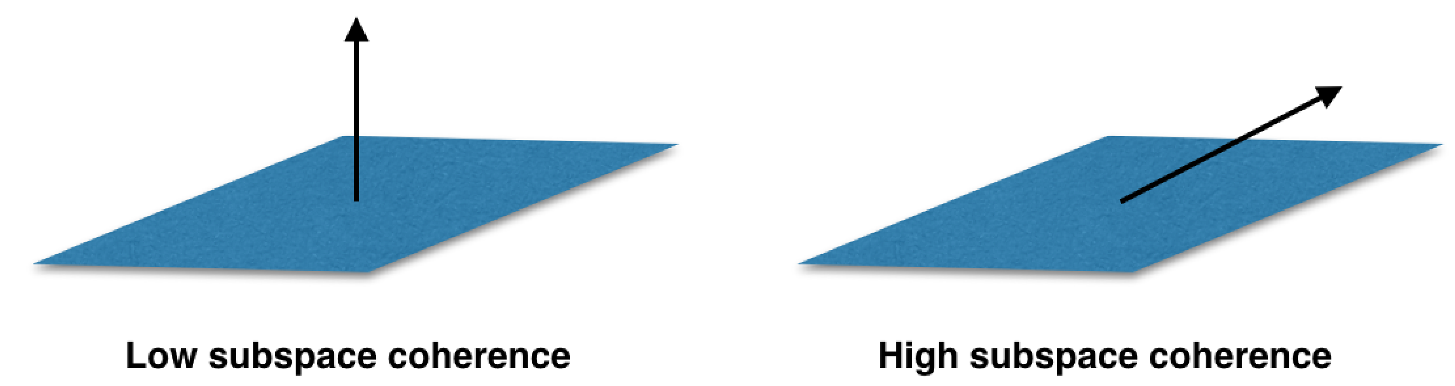
Definition 1. Local Subspace Coherence

The local subspace coherence between basis vectors $\{\phi_j\}_j$ and the subspace $\mathcal{X}$ is

$$\mu(\mathcal{X}, \phi_j) := \|\mathrm{P}_\mathcal{X}\phi_j\|_2, \quad \forall j \in [n]. \tag{3}$$

- Measures how aligned is ϕ_j with subspace $\mathcal{X}$;
- For $\mathcal{X} := \bigoplus_{l=1}^p \mathcal{X}_l$, it holds that $\mu^2(\mathcal{X}, \phi_j) = \sum_{l=1}^p \mu^2(\mathcal{X}_l, \phi_j)$.

Main Result



Theorem 1. Exactness of (P)

Suppose Assumption 1 hold and set $\lambda = \frac{1}{\log n}$ in **(P)**. Then $(\mathbf{x}_0, \mathbf{y}_0)$ is the unique minimizer of **(P)** with probability at least $1 - n^{-\sqrt{C_0}}$, provided that for all $j \in [n]$

$$1 - p_j \ge C_0 \mu(\mathcal{X}, \phi_j)\log^2 n, \tag{4}$$

where C_0 is a universal positive constant.

- essentially, $\mathbb{P}(\phi_j \notin \mathcal{Y}) \propto \mu(\mathcal{X}, \phi_j)$;
- makes $\mathcal{X}, \mathcal{Y}$ be incoherent and hence distinguishable.

Illustrative Examples

Example 1: $\Psi = \mathcal{F}, \Phi = \mathcal{I}$.

- Coherence pattern: $|\langle \psi_i, \phi_j\rangle| = \frac{1}{\sqrt{n}}$;
- local subspace coherence $\mu(\mathcal{X}, \phi_j) \equiv \sqrt{\frac{\mathrm{card}(\theta_x)}{n}}$,

$$\mathbb{E}\mathrm{card}(\theta_y) + \mathrm{card}(\theta_x) \le \mathcal{O}(\frac{n}{\log^4 n}). \tag{5}$$

Example 2: $\Psi = \mathcal{H}, \Phi = \mathcal{F}$.

- Coherence pattern: $|\langle \psi_i, \phi_j\rangle| \le 2^{-\frac{k(j)+|k(j)-k(i)|}{2}}$.;
- local subspace coherence $\sum_{j=1}^n \mu(\mathcal{X}, \phi_j) \le \frac{2\sqrt{2}}{\sqrt{2}-1}\mathrm{card}(\theta_x)$,

$$\mathbb{E}\mathrm{card}(\theta_y) + \mathrm{card}(\theta_x) \le \mathcal{O}(\frac{n}{\log^2 n}). \tag{6}$$

Numerical Experiment

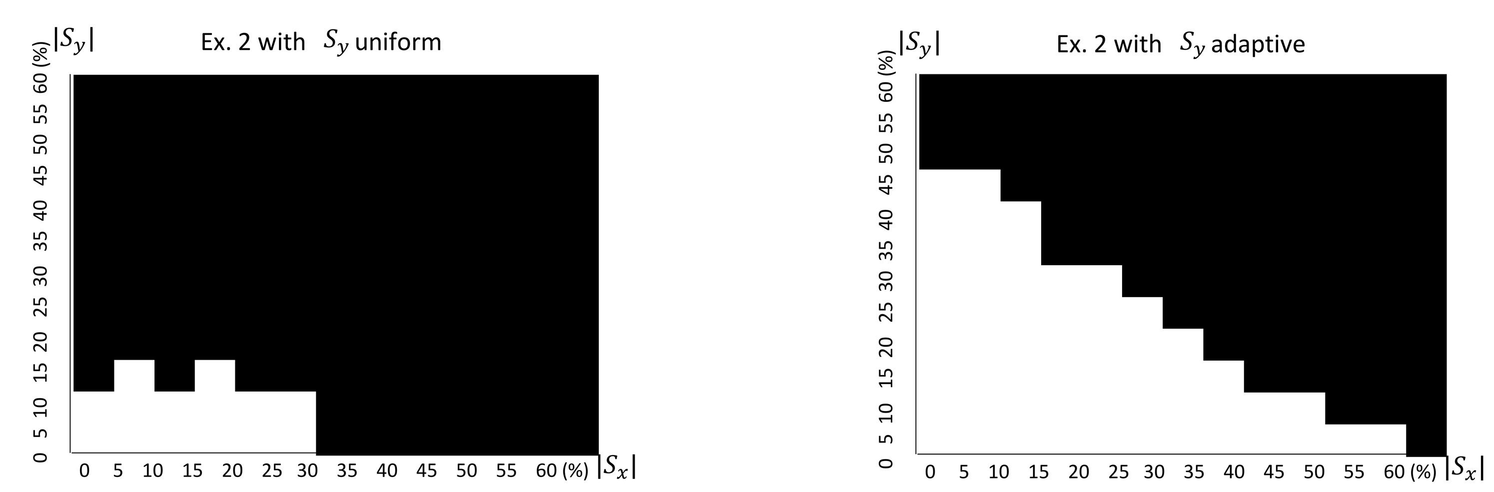


Fig. 2: Comparison of success region between (a) θ_y be uniformly at random and (b) θ_y be adapted to local subspace coherence.

References

[1] E. J. Candès and J. Romberg. Quantitative robust uncertainty principles and optimally sparse decompositions. *Foundations of Computational Mathematics*, 6(2):227–254, 2006.

[2] D. L. Donoho and G. Kutyniok. Microlocal analysis of the geometric separation problem. *Communications on Pure and Applied Mathematics*, 66(1):1–47, 2013.