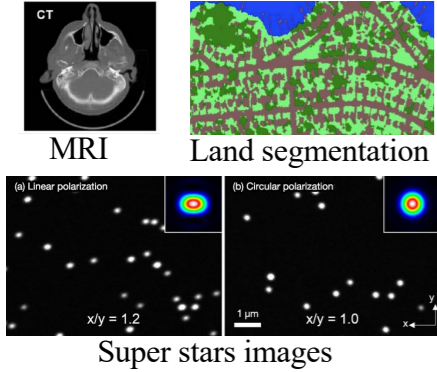


## 1. Application and Goal



**Goal:** apply separation prior in algorithm to improve accuracy

## 2. Motivation: Separation is the Key

**Theory**

**D.L. Donoho, 1990**  
 $\lim_{r \rightarrow \infty} r^{-1} \sup_t \#(S \cap [t, t+r]) < \frac{\Omega}{\pi}$

**Emmanuel J. Candès et al., 2012**  
 $\Delta(T) \geq 2.38/f_c = 2.38\lambda_c$

**Wenjing Liao et al., 2021**  
 $d(\omega_j, \omega_l) > \frac{1}{L} \sqrt{\frac{2}{\pi} \left( \frac{2}{\pi} - \frac{1}{L} - \frac{8s}{\pi L^2} \right)^{-\frac{1}{2}}}$

**Algorithm**

$\min_x \|x\|_1 \text{ s.t. } Ax = y$   
 $\min_{x \in \mathbb{R}^N} \{ \|y - Ax\|_2^2 + \lambda \|x\|_1 \}$   
 ESPRIT  $\hat{\phi} = (\hat{S}_1^* \hat{S}_1)^{-1} \hat{S}_1^* \hat{S}_2$   
 spectral MUSIC  $\frac{1}{a^*(\omega) \hat{G} \hat{G}^* a(\omega)}$

**Separation Blind**

## 5. Estimation by Sampling

**Closed-form:**  $\hat{\lambda}_l^{EW} = \frac{\exp(-M \hat{r}_l / \beta) \pi_l}{\sum_{k=1}^L \exp(-M \hat{r}_k / \beta) \pi_k}$

**Computational Issue** Too Large:  $L \sim O(2^R)$

**Posterior Sampling:**  $\lambda_l \propto \exp(-M \hat{r}_l / \beta) \pi_l, \mathbf{p}_l \in \tilde{\mathcal{P}}$

**Searching Strategy:**

$\tilde{\mathcal{P}}$ : separation prior

## 3. Main Idea: Estimation via Aggregating Sub-Models

A known family  $\mathcal{F}$  of a priori hypotheses on the signal  $f$ :  $\mathcal{F} = \{f_1, f_2, \dots, f_L\}$

**Convex combine:**  $f_\lambda(X) \triangleq \sum_{l=1}^L \lambda_l f_l(X), \forall X \in \mathcal{X}$

**Weight:**  $\Lambda = \left\{ \lambda = (\lambda_1, \dots, \lambda_L)^T \in \mathbb{R}^L: \lambda_j \geq 0, \sum_{j=1}^L \lambda_j = 1 \right\}$

**Sub model:**  $\text{supp} = \mathbf{p} \in \mathcal{P} \triangleq \{0, 1\}^N$

$f_l(X) \triangleq A_l \hat{z}_l \rightarrow \hat{z}_l = \arg \min \|y - A_l z\|_2$   
 $Z_l \triangleq \{z \odot \mathbf{p}_l, z \in \mathbb{C}^N\} \subset \mathbb{C}^N$

$A_l = [A_{\{1\}}^1, 0_{\{2\}}, 0_{\{3\}}, A_{\{5\}}^5, 0_{\{6\}}, \dots, 0_{\{N\}}]$

$\mathbf{p}_l$

1	2	3	4	5	...	N
●	○	○	○	○	...	○
1	0	0	0	1	0	0
0	0	0	0	0	0	0

**Sparsity:  $|\mathbf{p}_l|$ ;**  
**Separation**

## 6. MH-MCMC Algorithm

Fix  $\mathbf{p}_0 = \mathbf{0} \in \mathbb{R}^M$ . For any  $t \geq 0$ , given  $\mathbf{p}_t \in \mathcal{P}$ ,

- 1 Generate a random variable  $Q_t$  with distribution  $q(\cdot | \mathbf{p}_t)$ .
- 2 Generate a random variable  $P_{t+1} = \begin{cases} Q_t, & \text{with probability } r(\mathbf{p}_t, Q_t), \\ \mathbf{p}_t, & \text{with probability } 1 - r(\mathbf{p}_t, Q_t), \end{cases}$

where  $r(p, q) = \min\left(\frac{v_q}{v_p}, 1\right)$ .

3 Compute the least squares estimator  $\hat{\theta}_{P_{t+1}}$ .

## 4. Sparsity-Promoting Aggregation

**Trade-off: MSE and sparsity**

$\hat{\lambda}^{EW} = \arg \min_{\lambda \in \Lambda^L} \left( \sum_{l=1}^L \lambda_l \hat{r}_l + \frac{\beta}{M} \sum_{k=1}^L \lambda_k \log \frac{\lambda_k}{\pi_k} \right)$

**LS error:**  
 $\hat{r}_l = \frac{1}{M} \|y - A \hat{z}_l\|_2^2$

**Sparsity prior:**  
 $\pi_l = \frac{1}{H} \binom{N}{2eL}^{|\mathbf{p}_l|}$

**Consistency with L0-norm**

$\min_{\mathbf{p} \in \mathbb{C}^N} \|y - A\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_0$   
 $\|\mathbf{x}\|_0 \sim |\mathbf{p}_l|$

$\min_{\lambda \in \Lambda^L} \sum_{l=1}^L \lambda_l \hat{r}_l + \lambda_l |\mathbf{p}_l| \log \frac{2eL}{|\mathbf{p}_l|}$

**Structural constraint: separation**

## 7. Simulation

