

# Joint Sparse Estimation with Cardinality Constraint via Mixed-Integer Semidefinite Programming



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT



TECHNISCHE UNIVERSITÄT  
ILMENAU

Tianyi Liu<sup>1</sup>, Frederic Matter<sup>2</sup>, Alexander Sorg<sup>1</sup>, Marc E. Pfetsch<sup>2</sup>, Martin Haardt<sup>3</sup>,  
Marius Pesavento<sup>1</sup>

<sup>1</sup>Communication Systems Group, Technical University of Darmstadt, 64283 Darmstadt, Germany

<sup>2</sup>Research Group Optimization, Technical University of Darmstadt, 64283 Darmstadt, Germany

<sup>3</sup>Communications Research Laboratory, Ilmenau University of Technology, 98684 Ilmenau, Germany

## Introduction

- The multiple measurement vectors (MMV) problem aims at jointly estimating multiple signals that share a common sparse support over a known dictionary
- Proposal of a mixed-integer semidefinite programming (MISDP) reformulation for the original formulation of MMV with  $\ell_{2,0}$ -norm constraint

## Mathematical Model and Notations

- Distinct Directions-of-Arrival (DOAs)  $\theta = [\theta_1, \dots, \theta_L]^T$  from  $L$  sources
- Equivalent expression of DOAs in spatial frequencies  $\mu = [\mu_1, \dots, \mu_L]^T$  with  $\mu_l = \pi \cos(\theta_l) \in [-\pi, \pi]$  for  $l = 1, \dots, L$

### Signal Model

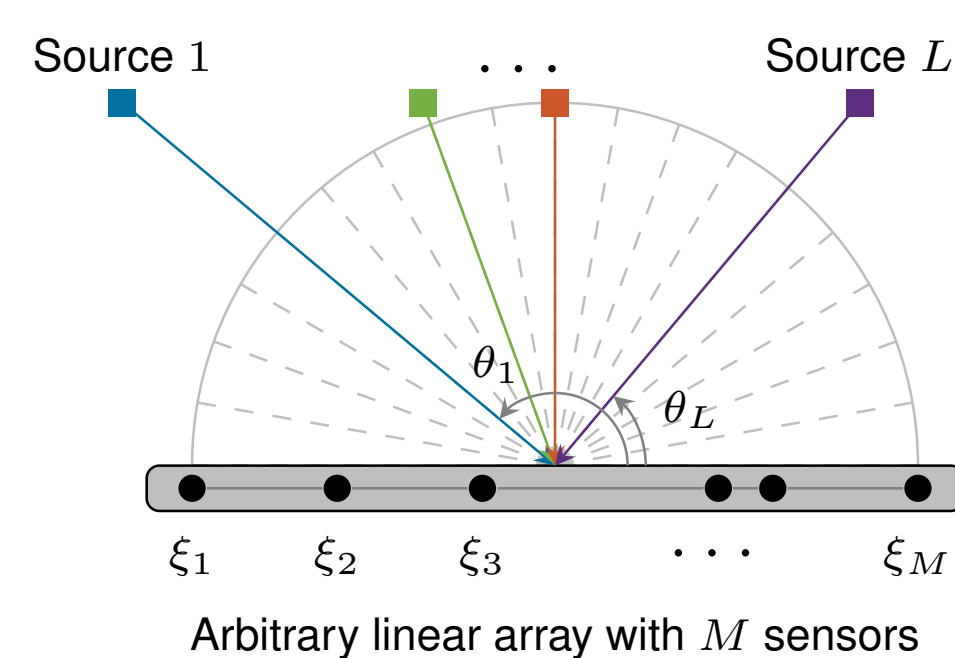
$$\mathbf{Y} = \mathbf{A}(\boldsymbol{\mu})\boldsymbol{\Psi} + \mathbf{N}$$

$\mathbf{Y} \in \mathbb{C}^{M \times N}$  : Received signal matrix

$\boldsymbol{\Psi} \in \mathbb{C}^{L \times N}$  : Source signal matrix

$\mathbf{N} \in \mathbb{C}^{M \times N}$  : Sensor noise matrix

$N$  : Number of available snapshots



- Steering matrix  $\mathbf{A}(\boldsymbol{\mu}) = [\mathbf{a}(\mu_1), \dots, \mathbf{a}(\mu_L)] \in \mathbb{C}^{M \times L}$  with  $\mathbf{a}(\mu_l) = [e^{j\mu_l \xi_1}, \dots, e^{j\mu_l \xi_M}]^T$  for  $l = 1, \dots, L$

## DML and MAP Estimators

### Deterministic Maximum Likelihood (DML) Estimator

- Deterministic source signals and spatio-temporal white Gaussian noise assumption  $\mathbf{n}_i \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_M) \implies \mathbf{y}_i | \boldsymbol{\psi}_i \sim \mathcal{CN}(\mathbf{A}(\boldsymbol{\mu})\boldsymbol{\psi}_i, \sigma^2 \mathbf{I}_M)$  for each snapshot  $i$

- Maximization of the likelihood function

$$\max_{\boldsymbol{\mu} \in [-\pi, \pi]^L, \boldsymbol{\Psi} \in \mathbb{C}^{L \times N}} \prod_{n=1}^N p(\mathbf{y}_n | \boldsymbol{\psi}_n) \iff \min_{\boldsymbol{\mu} \in [-\pi, \pi]^L, \boldsymbol{\Psi} \in \mathbb{C}^{L \times N}} \|\mathbf{A}(\boldsymbol{\mu})\boldsymbol{\Psi} - \mathbf{Y}\|_F^2$$

### Maximum A Posteriori (MAP) Estimator

- Statistical source signals of uncorrelated Gaussian prior distribution with uniform variances  $\boldsymbol{\psi}_i \sim \mathcal{CN}(\mathbf{0}, \gamma \mathbf{I}_L)$  for each snapshot  $i$

- Maximization of the posterior probability:  $\rho = \sigma^2 / \gamma$

$$\begin{aligned} \max_{\boldsymbol{\mu} \in [-\pi, \pi]^L, \boldsymbol{\Psi} \in \mathbb{C}^{L \times N}} \prod_{n=1}^N p(\boldsymbol{\psi}_n | \mathbf{y}_n) &\iff \max_{\boldsymbol{\mu} \in [-\pi, \pi]^L, \boldsymbol{\Psi} \in \mathbb{C}^{L \times N}} \prod_{n=1}^N p(\mathbf{y}_n | \boldsymbol{\psi}_n) p(\boldsymbol{\psi}_n) \\ &\iff \min_{\boldsymbol{\mu} \in [-\pi, \pi]^L, \boldsymbol{\Psi} \in \mathbb{C}^{L \times N}} \sum_{n=1}^N -\log(p(\mathbf{y}_n | \boldsymbol{\psi}_n)) - \log(p(\boldsymbol{\psi}_n)) \\ &\iff \min_{\boldsymbol{\mu} \in [-\pi, \pi]^L, \boldsymbol{\Psi} \in \mathbb{C}^{L \times N}} \|\mathbf{A}(\boldsymbol{\mu})\boldsymbol{\Psi} - \mathbf{Y}\|_F^2 + \rho \|\boldsymbol{\Psi}\|_F^2 \end{aligned}$$

	DML	MAP
Original	$\min_{\boldsymbol{\mu} \in [-\pi, \pi]^L, \boldsymbol{\Psi} \in \mathbb{C}^{L \times N}} \ \mathbf{A}(\boldsymbol{\mu})\boldsymbol{\Psi} - \mathbf{Y}\ _F^2$	$\min_{\boldsymbol{\mu} \in [-\pi, \pi]^L, \boldsymbol{\Psi} \in \mathbb{C}^{L \times N}} \ \mathbf{A}(\boldsymbol{\mu})\boldsymbol{\Psi} - \mathbf{Y}\ _F^2 + \rho \ \boldsymbol{\Psi}\ _F^2$
Concentrated	$\min_{\boldsymbol{\mu} \in [-\pi, \pi]^L} \text{tr}(\mathbf{Y}^H \tilde{\boldsymbol{\Pi}}_{\mathbf{A}(\boldsymbol{\mu})} \mathbf{Y})$	$\min_{\boldsymbol{\mu} \in [-\pi, \pi]^L} \text{tr}(\mathbf{Y}^H \tilde{\boldsymbol{\Pi}}_{\mathbf{A}(\boldsymbol{\mu})} \mathbf{Y})$

- $\tilde{\boldsymbol{\Pi}}_{\mathbf{A}(\boldsymbol{\mu})} = \mathbf{I}_M - \mathbf{A}(\boldsymbol{\mu}) (\mathbf{A}(\boldsymbol{\mu})^H \mathbf{A}(\boldsymbol{\mu}))^{-1} \mathbf{A}(\boldsymbol{\mu})^H$
- $\tilde{\boldsymbol{\Pi}}_{\mathbf{A}(\boldsymbol{\mu})} = \mathbf{I}_M - \mathbf{A}(\boldsymbol{\mu}) (\mathbf{A}(\boldsymbol{\mu})^H \mathbf{A}(\boldsymbol{\mu}) + \rho \mathbf{I}_L)^{-1} \mathbf{A}(\boldsymbol{\mu})^H = \left( \frac{1}{\rho} \mathbf{A}(\boldsymbol{\mu}) \mathbf{A}(\boldsymbol{\mu})^H + \mathbf{I}_M \right)^{-1}$

## Sparse Representation

- Sample the field-of-view (FOV) in  $K \gg L$  directions with spatial frequencies  $\boldsymbol{\nu} = [\nu_1, \dots, \nu_K]^T$
- On-grid assumption:  $\{\mu_l\}_{l=1}^L \subset \{\nu_k\}_{k=1}^K$
- Sparse signal model

$$\mathbf{Y} = \mathbf{A}(\boldsymbol{\nu})\mathbf{X} + \mathbf{N} \quad \begin{array}{l} - \mathbf{X} \in \mathbb{C}^{K \times N}: \text{Row-sparse representation of } \boldsymbol{\Psi} \\ - \mathbf{A}(\boldsymbol{\nu}) \in \mathbb{C}^{M \times K}: \text{Steering matrix for sampled directions } \boldsymbol{\nu} \end{array}$$

- MAP estimator for the sparse model

$$\min_{\mathbf{X} \in \mathbb{C}^{K \times N}, \|\mathbf{X}\|_{2,0} \leq L} \|\mathbf{A}(\boldsymbol{\nu})\mathbf{X} - \mathbf{Y}\|_F^2 + \rho \|\mathbf{X}\|_F^2$$

## Mixed-Integer Semidefinite Programming Reformulation

- Integer programming reformulation for sparse MAP

$$\min_{\mathbf{u} \in \{0,1\}^K, \mathbf{u}^T \mathbf{1} \leq L} \min_{\mathbf{X} \in \mathbb{C}^{K \times N}} \|\mathbf{A}(\boldsymbol{\nu})\mathbf{D}(\mathbf{u})\mathbf{X} - \mathbf{Y}\|_F^2 + \rho \|\mathbf{X}\|_F^2 \xrightarrow{\text{concentration}} \min_{\mathbf{u} \in \{0,1\}^K, \mathbf{u}^T \mathbf{1} \leq L} \text{tr}(\mathbf{Y}^H \frac{1}{\rho} \mathbf{A}(\boldsymbol{\nu})\mathbf{D}(\mathbf{u})\mathbf{A}(\boldsymbol{\nu})^H + \mathbf{I}_M)^{-1} \mathbf{Y})$$

- Binary vector  $\mathbf{u}$  determines the active directions

- Mixed-Integer semidefinite programming (MISDP) reformulation

$$\min_{\mathbf{u} \in \{0,1\}^K, \mathbf{T} \in \mathbb{S}_M^+} \text{tr}(\mathbf{T}) \quad \text{s.t.} \quad \begin{bmatrix} \frac{1}{\rho} \mathbf{A}(\boldsymbol{\nu})\mathbf{D}(\mathbf{u})\mathbf{A}(\boldsymbol{\nu})^H + \mathbf{I}_M & \mathbf{Y} \\ \mathbf{Y}^H & \mathbf{T} \end{bmatrix} \succeq 0, \quad \mathbf{u}^T \mathbf{1} \leq L$$

- Since  $\frac{1}{\rho} \mathbf{A}(\boldsymbol{\nu})\mathbf{D}(\mathbf{u})\mathbf{A}(\boldsymbol{\nu})^H \succ 0$ , by Schur complement formula

$$\begin{bmatrix} \frac{1}{\rho} \mathbf{A}(\boldsymbol{\nu})\mathbf{D}(\mathbf{u})\mathbf{A}(\boldsymbol{\nu})^H + \mathbf{I}_M & \mathbf{Y} \\ \mathbf{Y}^H & \mathbf{T} \end{bmatrix} \succeq 0 \iff \mathbf{T} - \mathbf{Y}^H \left( \frac{1}{\rho} \mathbf{A}(\boldsymbol{\nu})\mathbf{D}(\mathbf{u})\mathbf{A}(\boldsymbol{\nu})^H + \mathbf{I}_M \right)^{-1} \mathbf{Y} \succeq 0$$

## Interval Relaxation and Randomized Rounding (IRRR) [1]

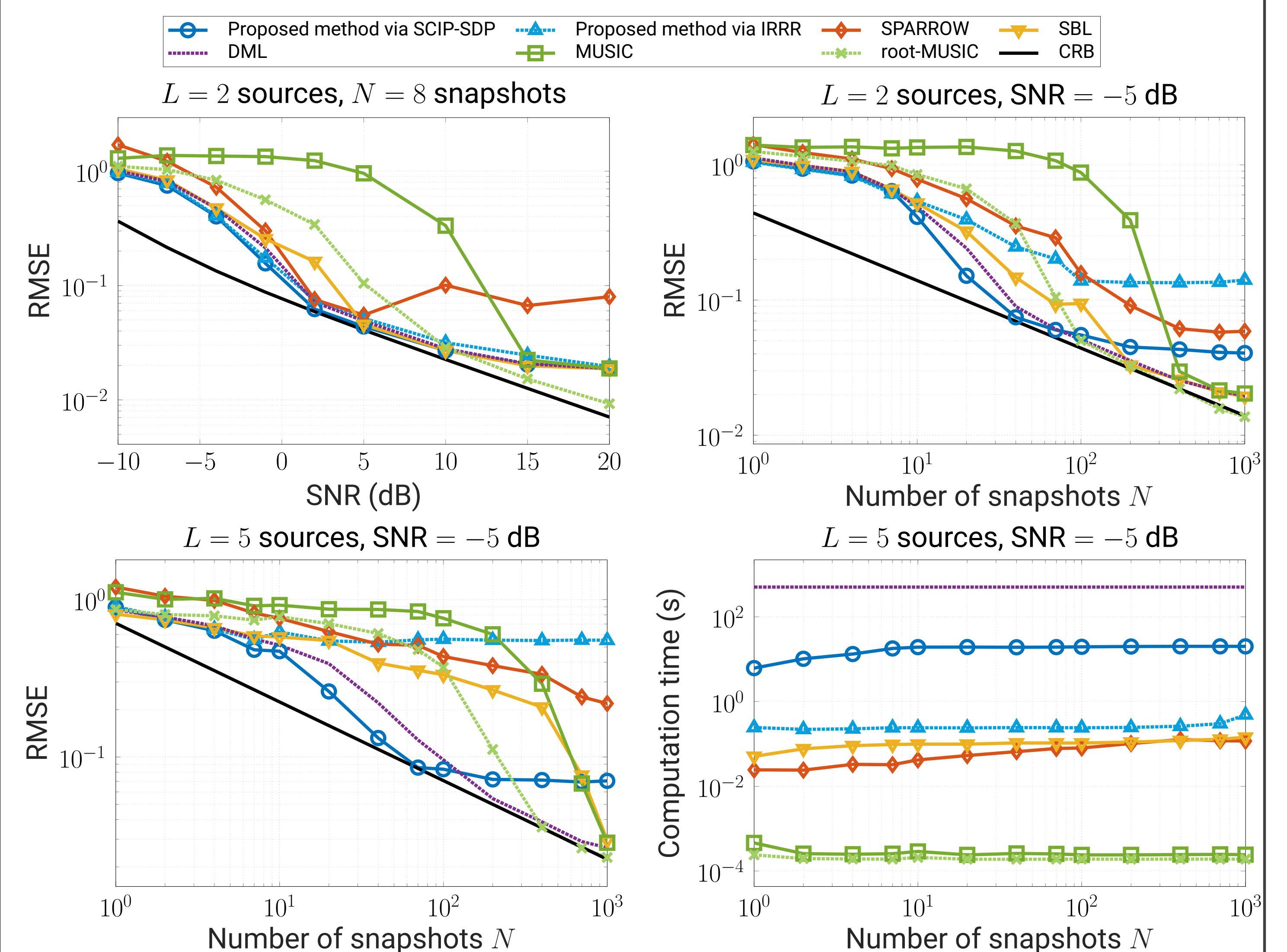
Step 1 **Interval Relaxation**: Solve the convex continuous relaxation

$$\hat{\mathbf{u}} = \arg \min_{\mathbf{u} \in [0,1]^K, \mathbf{T} \in \mathbb{S}_M^+} \text{tr}(\mathbf{T}) \quad \text{s.t.} \quad \begin{bmatrix} \frac{1}{\rho} \mathbf{A}(\boldsymbol{\nu})\mathbf{D}(\mathbf{u})\mathbf{A}(\boldsymbol{\nu})^H + \mathbf{I}_M & \mathbf{Y} \\ \mathbf{Y}^H & \mathbf{T} \end{bmatrix} \succeq 0, \quad \mathbf{u}^T \mathbf{1} \leq L$$

Step 2 **Randomized Rounding**: Randomly generate  $T > 0$  binary solutions  $\tilde{\mathbf{u}} \in \{0,1\}^K$  where each entry  $\tilde{u}_k$  independently follows the Bernoulli distribution  $\mathbb{P}[\tilde{u}_k = 1] = \hat{u}_k$  and  $\mathbb{P}[\tilde{u}_k = 0] = 1 - \hat{u}_k$ . Choose the best feasible solution among them.

## Simulation Results

- The MISDP reformulation exactly solved by the SCIP-SDP solver [2] based on branch-and-bound
- Continuous SDP problems solved by MOSEK solver
- Uniform linear array with  $M = 8$  sensors; uncorrelated sources;  $K = 100$  grid points
- Comparison sparsity-based DOA estimation methods:
  - SPARROW: Convex method with  $\ell_{2,1}$ -norm regularization
  - Sparse Bayesian Learning (SBL): Similarly uncorrelated Gaussian prior assumption but source variances are estimated



## References

- M. Pilanci, M. J. Wainwright, and L. El Ghaoui, "Sparse learning via Boolean relaxations," *Mathematical Programming*, vol. 151, pp. 63–87, June 2015.
- T. Gally, M. E. Pfetsch, and S. Ulbrich, "A framework for solving mixed-integer semidefinite programs," *Optimization Methods and Software*, vol. 33, pp. 594–632, May 2018.