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

Tianyi Liu[1], Frederic Matter[2], Alexander Sorg[1], Marc E. Pfetsch[2], Martin Haardt[3], Marius Pesavento[1]

[1]Communication Systems Group, Technical University of Darmstadt, 64283 Darmstadt, Germany
[2]Research Group Optimization, Technical University of Darmstadt, 64283 Darmstadt, Germany
[3]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

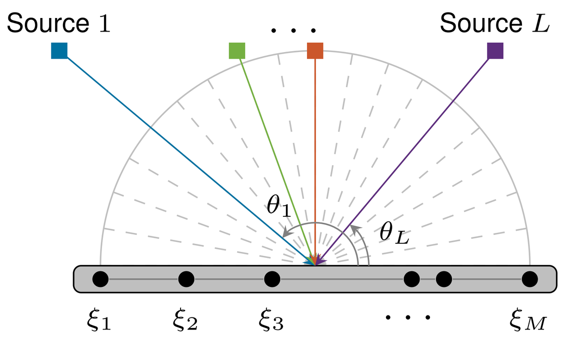
- Distinct Directions-of-Arrival (DOAs) $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_L]^\mathsf{T}$ from L sources
- Equivalent expression of DOAs in spatial frequencies $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_L]^\mathsf{T}$ with

$$\mu_l = \pi \cos(\theta_l) \in [-\pi, \pi) \text{ for } l = 1, \ldots, L$$

- Signal Model

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

$\boldsymbol{Y} \in \mathbb{C}^{M\times N}$: Received signal matrix
$\boldsymbol{\Psi} \in \mathbb{C}^{L\times N}$: Source signal matrix
$\boldsymbol{N} \in \mathbb{C}^{M\times N}$: Sensor noise matrix
N : Number of available snapshots



Arbitrary linear array with M sensors

- Steering matrix $\boldsymbol{A}(\boldsymbol{\mu}) = [\boldsymbol{a}(\mu_1), \ldots, \boldsymbol{a}(\mu_L)] \in \mathbb{C}^{M\times L}$ with

$$\boldsymbol{a}(\mu_l) = [\mathrm{e}^{\mathrm{j}\mu_l\xi_1}, \ldots, \mathrm{e}^{\mathrm{j}\mu_l\xi_M}]^\mathsf{T} \text{ for } l = 1, \ldots, L$$

DML and MAP Estimators

Deterministic Maximum Likelihood (DML) Estimator

- Deterministic source signals and spatio-temporal white Gaussian noise assumption

$$\boldsymbol{n}_i \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2\boldsymbol{I}_M) \quad \Longrightarrow \quad \boldsymbol{y}_i \mid \boldsymbol{\psi}_i \sim \mathcal{CN}(\boldsymbol{A}(\boldsymbol{\mu})\boldsymbol{\psi}_i, \sigma^2\boldsymbol{I}_M) \quad \text{for each snapshot } i$$

- Maximization of the likelihood function

$$\max_{\boldsymbol{\mu}\in[-\pi,\pi)^L,\, \boldsymbol{\Psi}\in\mathbb{C}^{L\times N}} \prod_{n=1}^{N} p(\boldsymbol{y}_n|\boldsymbol{\psi}_n) \quad \Longleftrightarrow \quad \min_{\boldsymbol{\mu}\in[-\pi,\pi)^L,\, \boldsymbol{\Psi}\in\mathbb{C}^{L\times N}} \|\boldsymbol{A}(\boldsymbol{\mu})\boldsymbol{\Psi} - \boldsymbol{Y}\|_\mathsf{F}^2$$

Maximum A Posteriori (MAP) Estimator

- Statistical source signals of uncorrelated Gaussian prior distribution with uniform variances

$$\boldsymbol{\psi}_i \sim \mathcal{CN}(\boldsymbol{0}, \gamma\boldsymbol{I}_L) \quad \text{for each snapshot } i$$

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

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

	DML	MAP
Original	$\min_{\boldsymbol{\mu}\in[-\pi,\pi)^L,\, \boldsymbol{\Psi}\in\mathbb{C}^{L\times N}} \|\boldsymbol{A}(\boldsymbol{\mu})\boldsymbol{\Psi} - \boldsymbol{Y}\|_\mathsf{F}^2$	$\min_{\boldsymbol{\mu}\in[-\pi,\pi)^L,\, \boldsymbol{\Psi}\in\mathbb{C}^{L\times N}} \|\boldsymbol{A}(\boldsymbol{\mu})\boldsymbol{\Psi} - \boldsymbol{Y}\|_\mathsf{F}^2 + \rho\|\boldsymbol{\Psi}\|_\mathsf{F}^2$
Concentrated	$\min_{\boldsymbol{\mu}\in[-\pi,\pi)^L} \operatorname{tr}\left(\boldsymbol{Y}^\mathsf{H}\boldsymbol{\Pi}^\perp_{\boldsymbol{A}(\boldsymbol{\mu})}\boldsymbol{Y}\right)$	$\min_{\boldsymbol{\mu}\in[-\pi,\pi)^L} \operatorname{tr}\left(\boldsymbol{Y}^\mathsf{H}\widetilde{\boldsymbol{\Pi}}^\perp_{\boldsymbol{A}(\boldsymbol{\mu})}\boldsymbol{Y}\right)$

- $\boldsymbol{\Pi}^\perp_{\boldsymbol{A}(\boldsymbol{\mu})} = \boldsymbol{I}_M - \boldsymbol{A}(\boldsymbol{\mu})\left(\boldsymbol{A}(\boldsymbol{\mu})^\mathsf{H}\boldsymbol{A}(\boldsymbol{\mu})\right)^{-1}\boldsymbol{A}(\boldsymbol{\mu})^\mathsf{H}$
- $\widetilde{\boldsymbol{\Pi}}^\perp_{\boldsymbol{A}(\boldsymbol{\mu})} = \boldsymbol{I}_M - \boldsymbol{A}(\boldsymbol{\mu})\left(\boldsymbol{A}(\boldsymbol{\mu})^\mathsf{H}\boldsymbol{A}(\boldsymbol{\mu}) + \rho\boldsymbol{I}_L\right)^{-1}\boldsymbol{A}(\boldsymbol{\mu})^\mathsf{H} = \left(\frac{1}{\rho}\boldsymbol{A}(\boldsymbol{\mu})\boldsymbol{A}(\boldsymbol{\mu})^\mathsf{H} + \boldsymbol{I}_M\right)^{-1}$

Sparse Representation

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

$$\boldsymbol{Y} = \boldsymbol{A}(\boldsymbol{\nu})\boldsymbol{X} + \boldsymbol{N}$$

  - $\boldsymbol{X} \in \mathbb{C}^{K\times N}$: Row-sparse representation of $\boldsymbol{\Psi}$
  - $\boldsymbol{A}(\boldsymbol{\nu}) \in \mathbb{C}^{M\times K}$: Steering matrix for sampled directions $\boldsymbol{\nu}$
- MAP estimator for the sparse model

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

Mixed-Integer Semidefinite Programming Reformulation

- Integer programming reformulation for sparse MAP

$$\min_{\boldsymbol{u}\in\{0,1\}^K,\, \boldsymbol{u}^\mathsf{T}\boldsymbol{1}\le L} \; \min_{\boldsymbol{X}\in\mathbb{C}^{K\times N}} \|\boldsymbol{A}(\boldsymbol{\nu})\boldsymbol{D}(\boldsymbol{u})\boldsymbol{X} - \boldsymbol{Y}\|_\mathsf{F}^2 + \rho\|\boldsymbol{X}\|_\mathsf{F}^2$$

$$\xRightarrow{\text{concentration}} \quad \min_{\boldsymbol{u}\in\{0,1\}^K,\, \boldsymbol{u}^\mathsf{T}\boldsymbol{1}\le L} \operatorname{tr}\left(\boldsymbol{Y}^\mathsf{H}(\tfrac{1}{\rho}\boldsymbol{A}(\boldsymbol{\nu})\boldsymbol{D}(\boldsymbol{u})\boldsymbol{A}(\boldsymbol{\nu})^\mathsf{H} + \boldsymbol{I}_M)^{-1}\boldsymbol{Y}\right)$$

  - Binary vector $\boldsymbol{u}$ determines the active directions
- Mixed-Integer semidefinte programming (MISDP) reformulation

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

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

$$\begin{bmatrix} \frac{1}{\rho}\boldsymbol{A}(\boldsymbol{\nu})\boldsymbol{D}(\boldsymbol{u})\boldsymbol{A}(\boldsymbol{\nu})^\mathsf{H} + \boldsymbol{I}_M & \boldsymbol{Y} \\ \boldsymbol{Y}^\mathsf{H} & \boldsymbol{T} \end{bmatrix} \succeq 0 \quad \Longleftrightarrow \quad \boldsymbol{T} - \boldsymbol{Y}^\mathsf{H}(\tfrac{1}{\rho}\boldsymbol{A}(\boldsymbol{\nu})\boldsymbol{D}(\boldsymbol{u})\boldsymbol{A}(\boldsymbol{\nu})^\mathsf{H} + \boldsymbol{I}_M)^{-1}\boldsymbol{Y} \succeq 0$$

Interval Relaxation and Randomized Rounding (IRRR) [1]

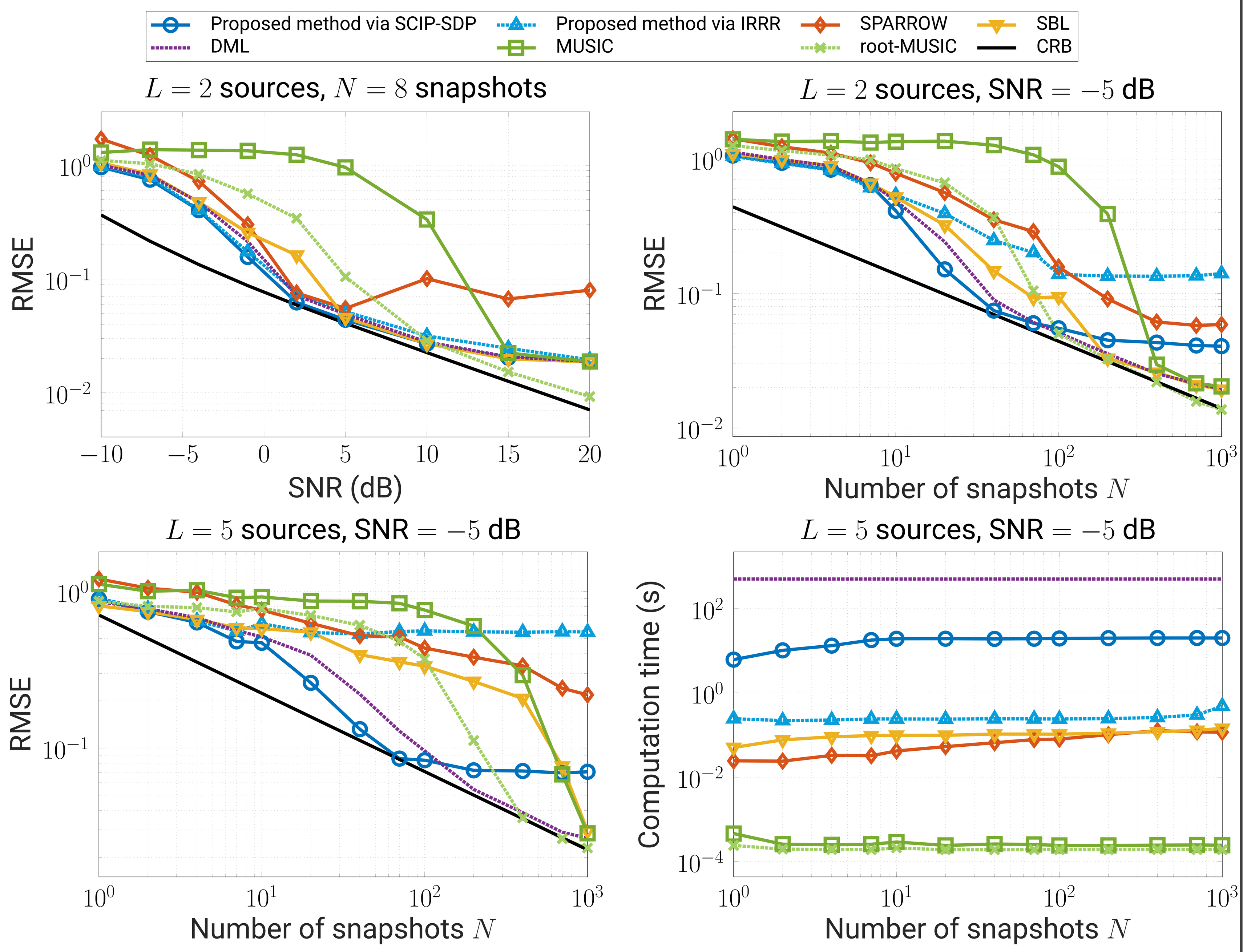
Step 1 **Interval Relaxation**: Solve the convex continuous relaxation

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

Step 2 **Randomized Rounding**: Randomly generate $T > 0$ binary solutions $\tilde{\boldsymbol{u}} \in \{0,1\}^K$ 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 SCI-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

[1] M. Pilanci, M. J. Wainwright, and L. El Ghaoui, "Sparse learning via Boolean relaxations," *Mathematical Programming*, vol. 151, pp. 63–87, June 2015.

[2] 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.

This work was supported in part by the DFG PRIDE Project PE 2080/2-1 and in part by the EXPRESS II Project within the DFG priority program CoSIP (DFG-SPP 1798).
2023 9th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)