Maximum Likelihood-based Gridless DoA Estimation using Structured Covariance Matrix Recovery and SBL with Grid Refinement

Rohan R. Pote and Bhaskar D. Rao

Material Appeared in: Transactions of Signal Processing, 2023

Department of Electrical and Computer Engineering
University of California San Diego

2024 International Conference on Acoustics, Speech and Signal Processing
Apr. 14 - 19, 2024

UC San Diego
JACOBS SCHOOL OF ENGINEERING

Contributions

- Established sparse Bayesian learning (SBL) as correlation-aware technique
  - Effectively utilizes array *geometry* and correlation *prior* information to identify more sources than sensors
- Reformulated SBL as a novel structured matrix recovery problem to perform gridless estimation of DoAs, under MLE framework
- Proposed approach minimizes KL divergence between (true) data distribution and the assumed (*uncorrelated* sources) distribution
- Insights to connect proposed approach with SBL and traditional MLE
- Provided algorithms for both ULA and non-uniform linear arrays in general
- Numerically demonstrated proposed approach very robust
  - Scenarios: single snapshot, correlated sources, small sources separation, more sources than sensors

Outline

Problem of Interest

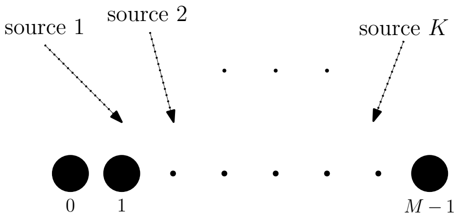


Figure: K narrowband sources impinging on M sensors antenna array

Measurement Model (L snapshots):

$$\mathbf{y}_l = \sum_{k=1}^{K} \phi(\theta_k) x_{l,k} + \mathbf{n}_l = \mathbf{\Phi}_{\boldsymbol{\theta}} \mathbf{x}_l + \mathbf{n}_l, \qquad 0 \leq l < L \tag{1}$$

$\phi(.) \in \mathbb{C}^M$ is a array manifold/response vector, $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_K]^T$,
$\mathbf{x}_l \in \mathbb{C}^K$ and noise $\mathbf{n}_l \in \mathbb{C}^M$ independent of each other, i.i.d. over time,
$\mathbf{n}_l$ is distributed as $\mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I})$.

Aim: To estimate direction of arrivals (DoA), θ_k's, for narrowband sources

Traditional & Modern Approaches

On Traditional Approaches:

- Parametric methods like maximum likelihood estimation (MLE) allow introducing *meaningful parameters*
  - Such parameters may be inferred with a *single snapshot*!
- Resulting cost function may be highly non-linear; model order is unknown

On Modern Approaches:

- Introduce novel *reparameterization* and either explicit or implicit *sparsity*
  - Examples include grid-based and grid-*less* sparse signal recovery algorithms (e.g., matching pursuit, SBL, ANM,SPICE etc.)
- SBL (a grid-based algorithm) formulates the problem under MLE framework, implicit regularization!

To find best of both worlds -
Enhance SBL formulation to perform *gridless DoA estimation under MLE!*

DoA Estimation under Maximum Likelihood Estimation Framework

- We impose a *parametrized* Gaussian prior on $\mathbf{x}_l$ i.e., $\mathbf{x}_l \sim \mathcal{CN}(\mathbf{0}, \mathbf{P})$.
- Explicit knowledge of model order information is needed
- Optimization problem for maximum likelihood DoA estimation

$$\min_{\substack{\boldsymbol{\theta} \in [-\frac{\pi}{2}, \frac{\pi}{2})^K, \\ \mathbf{P} \succ \mathbf{0}, \lambda \geq 0}} \log\det\left(\boldsymbol{\Phi}_{\boldsymbol{\theta}} \mathbf{P} \boldsymbol{\Phi}_{\boldsymbol{\theta}}^H + \lambda \mathbf{I}\right) + \mathrm{tr}\left(\left(\boldsymbol{\Phi}_{\boldsymbol{\theta}} \mathbf{P} \boldsymbol{\Phi}_{\boldsymbol{\theta}}^H + \lambda \mathbf{I}\right)^{-1} \hat{\mathbf{R}}_{\mathbf{y}}\right), \qquad (2)$$

where λ denotes noise variance estimate, and
$\hat{\mathbf{R}}_{\mathbf{y}} = \frac{1}{L} \sum_l \mathbf{y}_l \mathbf{y}_l^H$ denotes the sample covariance matrix.

Sparse Bayesian Learning

Remodeling+Implicit sparsity

Grid-based remodeling: discretize the possible values of θ and introduce the measurement matrix $\mathbf{\Phi} \in \mathbb{C}^{M \times G}$, G denotes the grid size

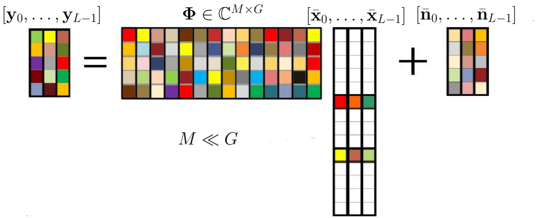


Figure: Multiple measurement vector (MMV) model

Original problem in (1) can be re-written as

$$\mathbf{y}_l = \mathbf{\Phi}\bar{\mathbf{x}}_l + \bar{\mathbf{n}}_l, \qquad 0 \leq l < L, \tag{3}$$

where the i-th column $[\mathbf{\Phi}]_i = \phi(\theta_i)$, for some known $\phi(.)$, $i = 1, \dots, G$.

$\bar{\mathbf{X}} = [\bar{\mathbf{x}}_0, \dots, \bar{\mathbf{x}}_{L-1}]$ is row-sparse i.e., most of the rows are zero.

Sparse Bayesian Learning (2)

Remodeling+Implicit sparsity

Prior on $\bar{\mathbf{x}}_l$:

- SBL imposes a *parameterized, uncorrelated* sources Gaussian prior

  $\bar{\mathbf{x}}_l \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Gamma})$ where $\boldsymbol{\Gamma}$ is a diagonal matrix; let $\mathrm{diag}(\boldsymbol{\Gamma}) = \boldsymbol{\gamma}$
- $\mathbf{y}_l \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Phi}\boldsymbol{\Gamma}\boldsymbol{\Phi}^H + \lambda\mathbf{I})$, λ denotes the estimate for noise variance

uninformative priors $\rightarrow (\boldsymbol{\gamma}, \lambda)$ estimated under MLE framework

MLE optimization problem:

$$\min_{\boldsymbol{\Gamma}\succeq\mathbf{0}, \lambda\geq 0} \log\det\left(\boldsymbol{\Phi}\boldsymbol{\Gamma}\boldsymbol{\Phi}^H + \lambda\mathbf{I}\right) + \mathrm{tr}\left(\left(\boldsymbol{\Phi}\boldsymbol{\Gamma}\boldsymbol{\Phi}^H + \lambda\mathbf{I}\right)^{-1}\hat{\mathbf{R}}_{\mathbf{y}}\right), \qquad (4)$$

where $\hat{\mathbf{R}}_{\mathbf{y}} = \frac{1}{L}\sum_{l=1}^{L}\mathbf{y}_l\mathbf{y}_l^H$ denotes the sample covariance matrix (SCM)

- Many algorithms to solve (4) (e.g., Tipping iterations(Tipping 2001), EM iterations(Wipf and Rao 2004), sequential SBL(Tipping and Faul 2003))

Example: Uniform Linear Array (ULA)

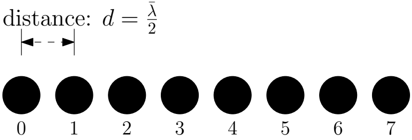


Figure: ULA with M (=8, here) sensors. $\bar{\lambda}$: wavelength of incoming signals

- $\mathbf{\Phi\Gamma\Phi}^H$ is a Toeplitz matrix; consequence of ULA geometry and *uncorrelated* sources prior
- SBL finds the 'best' positive semidefinite (grid-based) Toeplitz matrix approximation to the SCM $\hat{\mathbf{R}}_{\mathbf{y}}$

Idea:

Reparameterize the SBL cost function to directly estimate the entries of the Toeplitz covariance matrix

Proposed (ML) Structured Covariance Matrix Recovery

Let $\mathbf{v}$ denote the first row of a Toeplitz matrix, $\mathrm{Toep}(\mathbf{v})$.

SBL optimization problem *reparameterization*:

$$\min_{\substack{\mathbf{v}\in\mathbb{C}^M \text{ s.t.} \\ \mathrm{Toep}(\mathbf{v})\succeq\mathbf{0},\lambda\geq 0}} \log\det\left(\mathrm{Toep}(\mathbf{v})+\lambda\mathbf{I}\right)+\mathrm{tr}\left(\left(\mathrm{Toep}(\mathbf{v})+\lambda\mathbf{I}\right)^{-1}\hat{\mathbf{R}}_{\mathbf{y}}\right). \quad (5)$$

- DoAs can be estimated by decomposing the solution $\mathrm{Toep}(\mathbf{v}^*)$; root-MUSIC used in simulations
- Low-rank solution is encouraged by the $\log\det$ term[1], while its effect is moderated by the noise variance term, '$+\lambda\mathbf{I}$'

Problem is non-convex in $(\mathbf{v},\lambda)$; solved using majorization-minimization

[1] M. Fazel, H. Hindi, and S. P. Boyd (2003). "Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices". In: *Proceedings of the 2003 American Control Conference, 2003.* Vol. 3, 2156–2162 vol.3.

ULA with missing sensors case

- For ULA with missing sensors, $\mathbf{T}(\mathbf{v}) = \mathbf{S}\mathrm{Toep}(\mathbf{v})\mathbf{S}^T$ ($\mathbf{S}$: sampling matrix)
- General formulation:

$$\min_{\substack{\mathbf{v}\in\mathbb{C}^{M_{\mathrm{apt}}} \text{ s.t.} \\ \mathrm{Toep}(\mathbf{v})\succeq\mathbf{0},\lambda\geq 0}} \log\det\left(\mathbf{T}(\mathbf{v})+\lambda\mathbf{I}\right)+\mathrm{tr}\left(\left(\mathbf{T}(\mathbf{v})+\lambda\mathbf{I}\right)^{-1}\hat{\mathbf{R}}_{\mathbf{y}}\right). \tag{6}$$

Resulting SDP:

$$\begin{aligned} \min_{\mathbf{v}\in\mathbb{C}^{M_{\mathrm{apt}}},\, \mathbf{U}\in\mathbb{C}^{M\times M}} \quad & \mathrm{tr}\left(\left(\mathbf{T}(\mathbf{v}^{(k)})+\lambda\mathbf{I}\right)^{-1}\mathbf{T}(\mathbf{v})\right)+\mathrm{tr}\left(\mathbf{U}\,\hat{\mathbf{R}}_{\mathbf{y}}\right) \\ \text{subject to} \quad & \begin{bmatrix} \mathbf{U} & \mathbf{I}_M \\ \mathbf{I}_M & \mathbf{T}(\mathbf{v})+\lambda\mathbf{I} \end{bmatrix} \succeq \mathbf{0},\, \mathrm{Toep}(\mathbf{v})\succeq\mathbf{0}, \end{aligned} \tag{7}$$

(7) is solved iteratively with identity matrix initialization.

Note: $\mathrm{Toep}(\mathbf{v})\succeq\mathbf{0}$ imposed instead of just $\mathbf{T}(\mathbf{v})\succeq\mathbf{0}$ $\rightarrow$ pertinent model

DoAs can be estimated by decomposing the solution:

- $\mathbf{T}(\mathbf{v}^*)$ when fewer sources than sensors
- $\mathrm{Toep}(\mathbf{v}^*)$ otherwise

Comparison with SBL

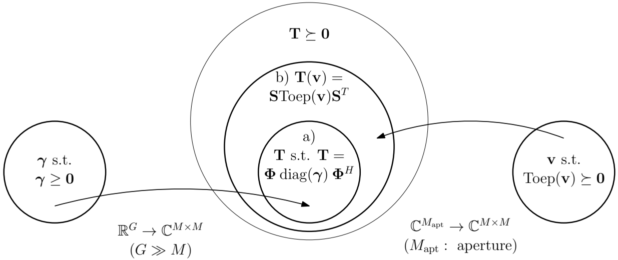


Figure: a) Positive semidefinite (PSD) cone used by SBL to fit to measurements b) Proposed PSD cone that includes SBL's search region

Consider the following *updated* SBL optimization problem:

$$\min_{\mathbf{\Phi}} \min_{\mathbf{\Gamma}\succeq \mathbf{0}, \lambda \geq 0} \log\det\left(\mathbf{\Phi}\mathbf{\Gamma}\mathbf{\Phi}^H + \lambda\mathbf{I}\right) + \mathrm{tr}\left(\left(\mathbf{\Phi}\mathbf{\Gamma}\mathbf{\Phi}^H + \lambda\mathbf{I}\right)^{-1}\hat{\mathbf{R}}_{\mathbf{y}}\right). \quad (8)$$

Theorem

The proposed problem in (6) and in (8) are equivalent, in that they achieve the same globally minimum cost.

- Similar connection exists between classical MLE formulation in (2) and (8)

More General Non-Uniform Linear Array

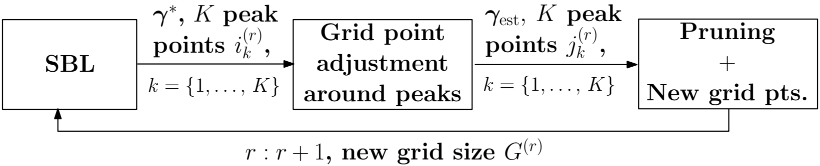


Fig. Proposed SBL with likelihood-based grid refinement procedure.
At $r = 0$, SBL is run with a uniform grid.

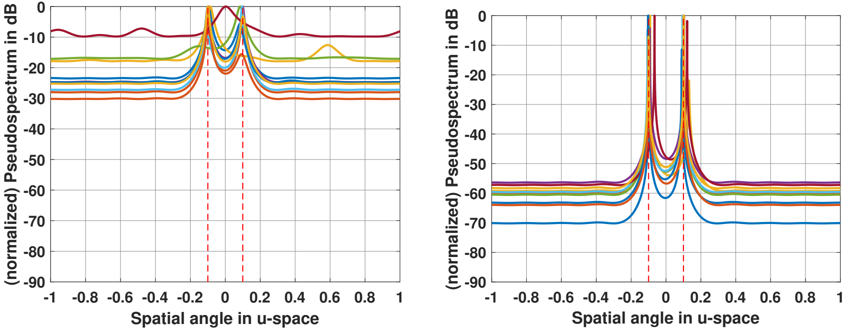
- Consider sensors placed arbitrarily on a linear aperture
- Structured received signal covariance matrix is neither Toeplitz nor sampled from Toeplitz
- Proposed idea extends SBL by improving upon the initial grid in two steps (see figure)

Simulations

MATLAB code available

- Performance compared with: ANM (Tang et al. 2013), Reweighted ANM (RAM) (Yang and Xie 2016), gridless SPARse ROW-norm reconstruction (SPARROW) (Steffens, Pesavento, and Pfetsch 2018), and gridless SPICE(Yang and Xie 2015)
- Proposed 'StructCovMLE' algorithm and RAM are iterative; run 20 iterations
- Provide noise variance, σ_n^2, to all algorithms, except gridless SPICE; set $\lambda = \sigma_n^2$ in 'StructCovMLE'
- Scenarios considered:
    - resolution (+ regularization-free 'StructCovMLE' vs. RAM)
    - number of snapshots
    - source identifiability

Single Snapshot Scenario

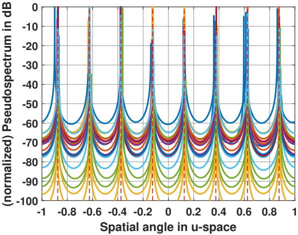


(a) FB averaged SCM (b) StructCovMLE (Proposed)

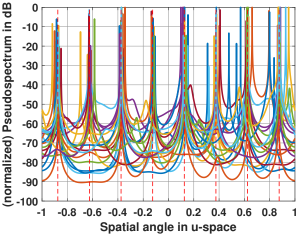
Parameters: ULA with $M = 10$ sensors, $K = 2$ sources, SNR $= 20$ dB, 10 random realizations.

- True DoA locations marked in red vertical dashed lines
- Proposed 'StructCovMLE' approach can identify two sources, unlike SCM-based and even with forward-backward averaging in some cases

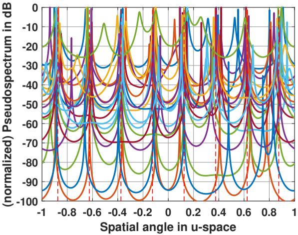
More Sources than Sensors Scenario



(a) StructCovMLE
(RMSE = 0.005)



(b) RAM
(RMSE = 0.106)



(d) GL-SPICE
(RMSE = 0.072)

Parameters: Nested array(Pal and Vaidyanathan 2010) with $M = 6$ sensors, locations$=\{0, 1, 2, 3, 7, 11\}$, $K = 8$ sources, SNR $= 20$ dB, $L = 4$ snapshots, 20 random realizations.

- Proposed 'StructCovMLE' algorithm able to localize all the 8 sources
- Superior identifiability performance also evident from the lower RMSE value (in u-space), as compared to the other techniques

Conclusion & Future Work

- Proposed a *novel reformulation of the SBL optimization problem* to recover DoAs in gridless manner
- Approach naturally leads to *estimating a structured covariance matrix* in the *MLE* sense
- Optimized the cost function iteratively using majorization-minimization; SDP in each iteration
- Provided perspectives to *connect* the new approach with the *traditional MLE* framework and the *modern SBL* formulation

Future Work:

- Lower complexity implementations to solve proposed 'StructCovMLE' problem
- Grid-based discretization may not be necessary when parametric dimension is small, scope to improve SBL with grid refinement

Thank you for your time!
Let me know if you have any questions.

References I

Fazel, M., H. Hindi, and S. P. Boyd (2003). "Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices". In: *Proceedings of the 2003 American Control Conference, 2003*. Vol. 3, 2156–2162 vol.3.

Krim, H. and M. Viberg (1996). "Two decades of array signal processing research: the parametric approach". In: *IEEE Signal Processing Magazine* 13.4, pp. 67–94.

Pal, P. and P. P. Vaidyanathan (2010). "Nested Arrays: A Novel Approach to Array Processing With Enhanced Degrees of Freedom". In: *IEEE Transactions on Signal Processing* 58.8, pp. 4167–4181.

Steffens, C., M. Pesavento, and M. E. Pfetsch (2018). "A Compact Formulation for the $\ell_{2,1}$ Mixed-Norm Minimization Problem". In: *IEEE Transactions on Signal Processing* 66.6, pp. 1483–1497.

Tang, G. et al. (2013). "Compressed Sensing Off the Grid". In: *IEEE Transactions on Information Theory* 59.11, pp. 7465–7490.

Tipping, M. (2001). "Sparse Bayesian learning and the relevance vector machine". In: *Machine Learning Research* 1, pp. 211–244.

Tipping, Michael E. and Anita C. Faul (2003). "Fast Marginal Likelihood Maximisation for Sparse Bayesian Models". In: *Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics*. Ed. by Christopher M. Bishop and Brendan J. Frey. Vol. R4. Proceedings of Machine Learning Research. PMLR, pp. 276–283.

Wipf, D. and B. D. Rao (2004). "Sparse Bayesian learning for basis selection". In: *IEEE Transactions on Signal Processing* 52.8, pp. 2153–2164.

Yang, Z. and L. Xie (2015). "On Gridless Sparse Methods for Line Spectral Estimation From Complete and Incomplete Data". In: *IEEE Transactions on Signal Processing* 63.12, pp. 3139–3153.

— (2016). "Enhancing Sparsity and Resolution via Reweighted Atomic Norm Minimization". In: *IEEE Transactions on Signal Processing* 64.4, pp. 995–1006.

Review: Traditional Approach (2)

	Methods	Primary Bottleneck[2]
(a)	i. Spatial filtering (beamforming)	Aperture/ degrees of freedom
	ii. Subspace based methods	*Number of snapshots*
(b)	Deterministic/ Stochastic MLE	Model & computational complexity

(a) Spectral based methods (b) Parametric methods

Table: Summary of traditional approaches

[2] H. Krim and M. Viberg (1996). "Two decades of array signal processing research: the parametric approach". In: *IEEE Signal Processing Magazine* 13.4, pp. 67–94.

Comparison with SBL

- SBL not only finds a structured matrix fit to the measurements, it also *factorizes* it (not a major problem!)
- SBL solution (which it factorizes) also in the search region of proposed optimization problem
- Root-MUSIC helps in this quest of factorization
- Proposed approach finds a structured covariance matrix in MLE sense, hence called ‘StructCovMLE’

‘StructCovMLE’ goes beyond SBL, provides *gridless* DoA estimation (How?)

Comparison with SBL (2)

Consider the following *updated* SBL optimization problem:

$$\min_{\mathbf{\Phi}} \min_{\mathbf{\Gamma}\succeq\mathbf{0},\lambda\geq 0} \log\det\left(\mathbf{\Phi}\mathbf{\Gamma}\mathbf{\Phi}^H + \lambda\mathbf{I}\right) + \mathrm{tr}\left(\left(\mathbf{\Phi}\mathbf{\Gamma}\mathbf{\Phi}^H + \lambda\mathbf{I}\right)^{-1}\hat{\mathbf{R}}_{\mathbf{y}}\right). \tag{9}$$

Theorem

The proposed problem in (6) and in (8) are equivalent, in that they achieve the same globally minimum cost.

In other words, the proposed approach estimates a structured covariance matrix fit to the measurements in the MLE sense over

- *all appropriate (i.e., array manifolds as columns) dictionaries for SBL*
- (OR) *all model orders for classical MLE* (not discussed here, shown in paper).