An Improved DOA Estimator Based On Partial Relaxation Approach

Minh Trinh Hoang, Mats Viberg and Marius Pesavento

TECHNISCHE UNIVERSITÄT DARMSTADT

Communication Systems Group
Darmstadt University of Technology
Darmstadt, Germany

CHALMERS

Department of Electrical Engineering
Chalmers University of Technology
Gothenburg, Sweden

Motivation

- Wide application of DOA estimation

- Multiple families of DOA estimators:
  - Maximum likelihood estimators
  - Subspace-based estimators
  - ...

- Proposal of a DOA estimator under the Partial Relaxation approach
  - Closely related to conventional DOA estimators
  - Efficient implementation for updating eigenvalues

Table of Contents

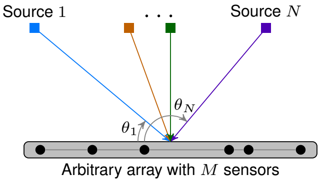
TECHNISCHE UNIVERSITÄT DARMSTADT

Signal Model

TECHNISCHE UNIVERSITÄT DARMSTADT

Multiple Snapshots Model

$$\mathbf{X} = \mathbf{A}(\boldsymbol{\theta})\mathbf{S} + \mathbf{N}$$



- T : Number of available snapshots
- $\mathbf{X} \in \mathbb{C}^{M \times T}$: Received signal matrix
- $\mathbf{S} \in \mathbb{C}^{N \times T}$: Source signal matrix
- $\mathbf{N} \in \mathbb{C}^{M \times T}$: Sensor noise matrix
- $\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), ..., \mathbf{a}(\theta_N)] \in \mathbb{C}^{M \times N}$: Steering matrix

Array Manifold

$$\mathcal{A}_N = \left\{ \mathbf{A} \in \mathbb{C}^{M \times N} | \mathbf{A} = [\mathbf{a}(\vartheta_1), \dots, \mathbf{a}(\vartheta_N)] \text{ with } 0 \leq \vartheta_1 < \dots < \vartheta_N < 180^\circ \right\}$$

Signal Model

TECHNISCHE UNIVERSITÄT DARMSTADT

Covariance Matrix

$$\mathbf{R} = \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \sigma_n^2\mathbf{I}_M$$

- $\mathbf{R} = \mathbb{E}\left\{\mathbf{x}(t)\mathbf{x}(t)^H\right\} \in \mathbb{C}^{M\times M}$: Covariance matrix of the received signal
- $\mathbf{R}_s = \mathbb{E}\left\{\mathbf{s}(t)\mathbf{s}(t)^H\right\} \in \mathbb{C}^{N\times N}$: Covariance matrix of the transmitted signal
- σ_n^2: Noise power at the sensors

Sample Covariance Matrix

$$\hat{\mathbf{R}} = \frac{1}{T}\mathbf{X}\mathbf{X}^H = \hat{\mathbf{U}}_s\hat{\mathbf{\Lambda}}_s\hat{\mathbf{U}}_s^H + \hat{\mathbf{U}}_n\hat{\mathbf{\Lambda}}_n\hat{\mathbf{U}}_n^H$$

- Signal subspace spanned by $\hat{\mathbf{U}}_s$
- N largest eigenvalues $\left\{\hat{\lambda}_1, \dots, \hat{\lambda}_N\right\}$ of $\hat{\mathbf{R}}$ are contained in $\hat{\mathbf{\Lambda}}_s$

- Noise subspace spanned by $\hat{\mathbf{U}}_n$
- $(M - N)$ smallest eigenvalues $\left\{\hat{\lambda}_{N+1}, \dots, \hat{\lambda}_M\right\}$ of $\hat{\mathbf{R}}$ are contained in $\hat{\mathbf{\Lambda}}_n$

Table of Contents

TECHNISCHE UNIVERSITÄT DARMSTADT

Partial Relaxation Approach

Revision of Conventional DOA Estimators

TECHNISCHE UNIVERSITÄT DARMSTADT

General Formulation

$$\left\{\hat{\mathbf{A}}\right\} = \arg\min_{\mathbf{A}\in\mathcal{A}_N} f\left(\mathbf{A}\right)$$

Remarks

- $\mathcal{A}_N$: Highly structured and non-convex set
- $f(\cdot)$: Generally non-convex function with multiple local minima
- Highly computational cost to obtain the global minimum

Example: Deterministic Maximum Likelihood (DML) Estimator

$$\left\{\hat{\mathbf{A}}_{\text{DML}}\right\} = \arg\min_{\mathbf{A}\in\mathcal{A}_N} \text{tr}\left\{\left(\mathbf{I}_M - \mathbf{A}\left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}^H\right)\hat{\mathbf{R}}\right\}$$

Partial Relaxation Approach

Revision of Conventional DOA Estimators

TECHNISCHE UNIVERSITÄT DARMSTADT

General Formulation

$$\left\{\hat{\mathbf{A}}\right\} = \arg\min_{\mathbf{A}\in\mathcal{A}_N} f\left(\mathbf{A}\right)$$

Remarks

- $\mathcal{A}_N$: Highly structured and non-convex set
- $f(\cdot)$: Generally non-convex function with multiple local minima
- Highly computational cost to obtain the global minimum

Example: Covariance Fitting (CF) Estimator

$$\left\{\hat{\mathbf{A}}_{\text{CF}}\right\} = \arg\min_{\mathbf{A}\in\mathcal{A}_N} \min_{\mathbf{R_s}\succeq\mathbf{0}} \left|\left|\hat{\mathbf{R}} - \mathbf{A}\mathbf{R}_s\mathbf{A}^H\right|\right|_F^2$$

Partial Relaxation Approach

Revision of Conventional DOA Estimators

TECHNISCHE UNIVERSITÄT DARMSTADT

General Formulation

$$\left\{\hat{\mathbf{A}}\right\} = \arg\min_{\mathbf{A}\in\mathcal{A}_N} f\left(\mathbf{A}\right)$$

Remarks

- $\mathcal{A}_N$: Highly structured and non-convex set
- $f(\cdot)$: Generally non-convex function with multiple local minima
- Highly computational cost to obtain the global minimum

Objective: Find a suboptimal solution without substantial performance degradation

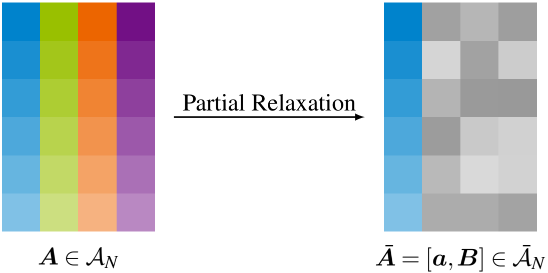
Partial Relaxation Approach
Concept

General Formulation of Conventional Estimators

$$\left\{\hat{\mathbf{A}}\right\} = \arg\min_{\mathbf{A}\in\mathcal{A}_N} f\left(\mathbf{A}\right)$$

Relaxed Array Manifold

$$\bar{\mathcal{A}}_N = \left\{\mathbf{A} \in \mathbb{C}^{M\times N} \middle| \mathbf{A} = \left[\mathbf{a}(\vartheta), \mathbf{B}\right], \mathbf{a}(\vartheta) \in \mathcal{A}_1, \mathbf{B} \in \mathbb{C}^{M\times(N-1)}\right\}$$



Partial Relaxation Approach

Concept

TECHNISCHE UNIVERSITÄT DARMSTADT

General Formulation of Conventional Estimators

$$\left\{\hat{\mathbf{A}}\right\} = \underset{\mathbf{A}\in\mathcal{A}_N}{\arg\min}\, f\left(\mathbf{A}\right)$$

Relaxed Array Manifold

$$\bar{\mathcal{A}}_N = \left\{\mathbf{A} \in \mathbb{C}^{M\times N} \middle| \mathbf{A} = \left[\mathbf{a}(\vartheta), \mathbf{B}\right], \mathbf{a}(\vartheta) \in \mathcal{A}_1, \mathbf{B} \in \mathbb{C}^{M\times(N-1)}\right\}$$

Formulation of the Partial Relaxation (PR) Approach

$$\left\{\hat{\mathbf{a}}_{\mathrm{PR}}\right\} = {}^{N}\underset{\mathbf{a}\in\mathcal{A}_1}{\arg\min}\ \min_{\mathbf{B}\in\mathbb{C}^{M\times(N-1)}} f\left(\left[\mathbf{a}, \mathbf{B}\right]\right)$$

- Relax the manifold structure of the signals from interfering directions
- Grid-search for N-deepest local minima to obtain the estimated DOAs

Partial Relaxation Approach

Proposed Estimators

TECHNISCHE UNIVERSITÄT DARMSTADT

Formulation of PR-Constrained Covariance Fitting (PR-CCF)

$$\left\{\hat{\mathbf{a}}_{\text{PR-CCF}}\right\} = {}^{N}\underset{\mathbf{a}\in\mathcal{A}_1}{\arg\min}\; \min_{\sigma_s^2\geq 0,\mathbf{E}} \left|\left|\hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H - \mathbf{E}\mathbf{E}^H\right|\right|_F^2$$

$$\text{subject to } \hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H - \mathbf{E}\mathbf{E}^H \succeq \mathbf{0}$$

$$\text{rank}(\mathbf{E}) \leq N-1$$

PR-CCF Estimator

$$\left\{\hat{\mathbf{a}}_{\text{PR-CCF}}\right\} = {}^{N}\underset{\mathbf{a}\in\mathcal{A}_1}{\arg\min}\; \sum_{k=N}^{M}\lambda_k^2\left(\hat{\mathbf{R}} - \frac{1}{\mathbf{a}^H\hat{\mathbf{R}}^{-1}\mathbf{a}}\mathbf{a}\mathbf{a}^H\right)$$

Remarks

- Not applicable if $\hat{\mathbf{R}}$ is singular
- Eigenvalues are extensively required

Partial Relaxation Approach

Proposed Estimators

TECHNISCHE UNIVERSITÄT DARMSTADT

PR-Unconstrained Covariance Fitting (PR-UCF)

$$\left\{\hat{\mathbf{a}}_{\text{PR-UCF}}\right\} = {}^{N}\underset{\mathbf{a}\in\mathcal{A}_1}{\arg\min}\ \min_{\sigma_s^2\geq 0,\mathbf{E}} \left|\left|\hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H - \mathbf{E}\mathbf{E}^H\right|\right|_F^2$$

$$\text{subject to rank}(\mathbf{E}) \leq N-1$$

Equivalent formulation of the inner optimization

$$\min_{\sigma_s^2\geq 0} \sum_{k=N}^{M} \lambda_k^2\left(\hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H\right)$$

- No closed-form solution for the minimizer $\hat{\sigma}_{s,\text{U}}^2$
- $\lambda_k^2\left(\hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H\right)$ is continuously differentiable with respect to σ_s^2

Minimization by Bisection Search or Newton's Method possible

Table of Contents

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Computational Aspects

Core Numerical Problem

TECHNISCHE UNIVERSITÄT DARMSTADT

Objective: Efficient computation of the eigenvalue decomposition

$$\mathbf{D} - \rho \mathbf{z}\mathbf{z}^H = \bar{\mathbf{U}}\bar{\mathbf{D}}\bar{\mathbf{U}}^H$$

- $\mathbf{D} = \text{diag}(d_1, \ldots, d_K) \in \mathbb{R}^{K \times K}$ with $d_1 > \ldots > d_K$
- $\rho > 0$
- $\mathbf{z} = [z_1, \ldots, z_K]^T \in \mathbb{C}^{K \times 1}$ has no zero component

- $\bar{\mathbf{D}} = \text{diag}(\bar{d}_1, \ldots, \bar{d}_K) \in \mathbb{R}^{K \times K}$ with $\bar{d}_1 > \ldots > \bar{d}_K$
- $\bar{\mathbf{U}} = [\bar{\mathbf{u}}_1, \ldots, \bar{\mathbf{u}}_K] \in \mathbb{C}^{K \times K}$ contains the normalized eigenvectors $\bar{\mathbf{u}}_k$ associated with the eigenvalues $\bar{d}_k$

Interlacing Property

a) The modified eigenvalues $\bar{d}_k$ satisfy $h(\bar{d}_k) = 0$ where the secular function $h(x)$ is defined as:

$$h(x) = 1 - \rho \mathbf{z}^H (\mathbf{D} - x\mathbf{I})^{-1} \mathbf{z}$$
$$= 1 - \rho \sum_{k=1}^{K} \frac{|z_k|^2}{d_k - x}$$

b) The modified eigenvalues $\bar{d}_k$ down-interlace with the initial eigenvalues d_k

$$d_1 > \bar{d}_1 > d_2 > \bar{d}_2 > ... > d_K > \bar{d}_K$$

Objective: Determine $\bar{d}_k$ which satisfies $h(\bar{d}_k) = 0$

Computational Aspects
Iterative Algorithm

TECHNISCHE UNIVERSITÄT DARMSTADT

Rewriting the secular function for determining $\bar{d}_k$

$$
\begin{aligned}
0 &= 1 - \sum_{i=1}^{k} \frac{|z_i|^2}{d_i - x} - \sum_{i=k+1}^{K} \frac{|z_i|^2}{d_i - x} \\
\iff & \sum_{i=1}^{k} \frac{|z_i|^2}{d_i - x} = 1 - \sum_{i=k+1}^{K} \frac{|z_i|^2}{d_i - x} \\
\iff & -\psi_k(x) = 1 + \phi_k(x)
\end{aligned}
$$

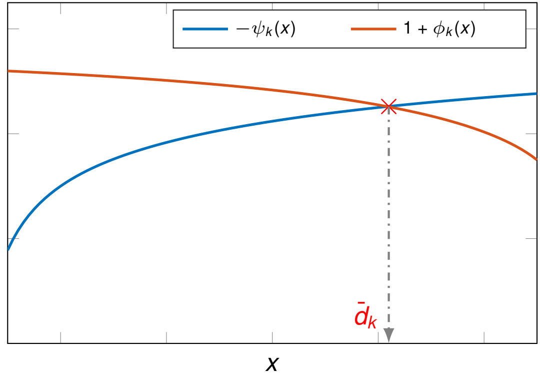
Idea: Approximate ψ_k and ϕ_k with a rational function of first degree

$$
R_{k;p,q}(x) = \begin{cases} p + \dfrac{q}{d_{k+1} - x} & \text{if } 0 \leq k \leq K-1 \\ 0 & \text{if } k = K, \end{cases}
$$

Computational Aspects

Iterative Algorithm

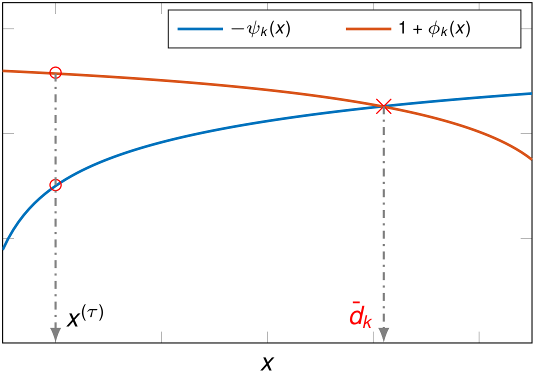
Rational Approximation



Computational Aspects

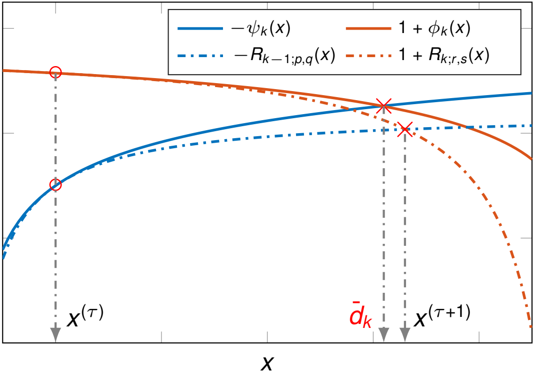
Iterative Algorithm

Rational Approximation



TECHNISCHE UNIVERSITÄT DARMSTADT

Rational Approximation

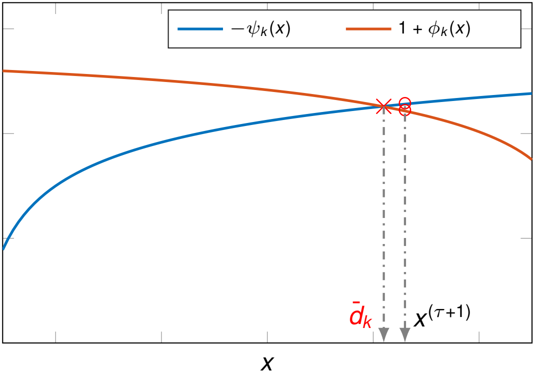


Computational Aspects

Iterative Algorithm

TECHNISCHE UNIVERSITÄT DARMSTADT

Rational Approximation



- Closed-form update
- Quadratic rate of convergence

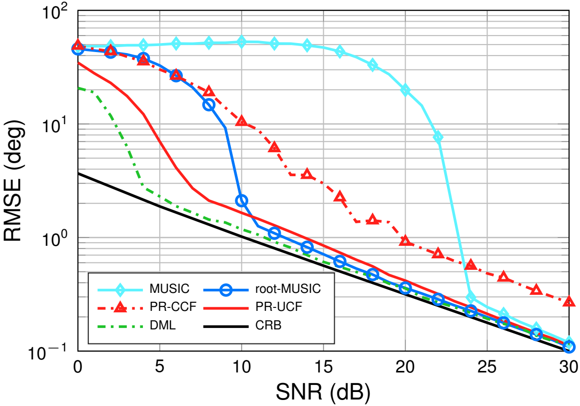
Table of Contents

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Simulation Results
Influence of SNR

TECHNISCHE UNIVERSITÄT DARMSTADT

$M = 10, \theta = [45^\circ, 50^\circ]^T, T = 8$



Simulation Results

Execution Time

$M = 10, \theta = [45^\circ, 50^\circ]^T, T = 100, N_G = 1800$

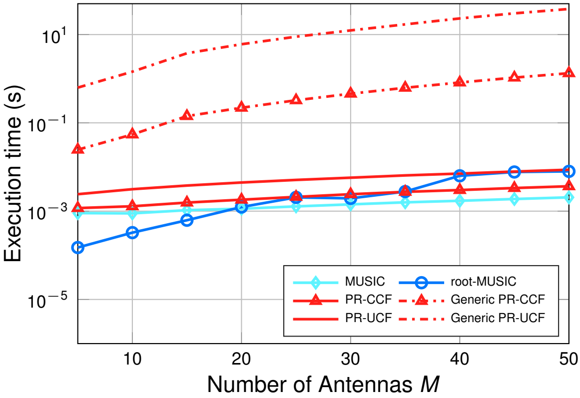


Table of Contents

TECHNISCHE UNIVERSITÄT DARMSTADT

Conclusions and Outlook

TECHNISCHE UNIVERSITÄT DARMSTADT

Conclusions

- Structure of interfering directions is relaxed
- Proposed estimator based on the covariance fitting problem
- Improved non-asymptotic behavior without exploiting any special structure of the sensor array
- Efficient implementation using rank-one update

Outlook

- Statistical properties of the proposed DOA estimator
- Generalization to multidimensional parameter estimation

TECHNISCHE UNIVERSITÄT DARMSTADT

Thank you for your attention!

Appendix

Numerical Solution for PR-UCF using Bisection Search

TECHNISCHE UNIVERSITÄT DARMSTADT

Define

$$g(\sigma_s^2) = \sum_{k=N}^{M} \bar{\lambda}_k(\sigma_s^2) = \sum_{k=N}^{M} \lambda_k^2 \left(\hat{\mathbf{R}} - \sigma_s^2 \mathbf{a}\mathbf{a}^H\right)$$

Asymptotic analysis of $g'(\sigma_s^2)$

$$g'(\sigma_s^2) = -\sum_{k=N}^{M} \frac{2\bar{\lambda}_k(\sigma_s^2)}{\sigma_s^4\ \mathbf{a}^H \left(\hat{\mathbf{R}} - \bar{\lambda}_k(\sigma_s^2)\mathbf{I}_M\right)^{-2} \mathbf{a}}$$

- If $\sigma_s^2 \to 0 \Rightarrow g'(\sigma_s^2) < 0$
- If $\sigma_{s,0}^2 \to \infty$:

$$g\left(\sigma_s^2\right) \approx \sigma_s^4 \, ||\mathbf{a}||_2^4$$
$$g'(\sigma_s^2) \to +\infty$$

Minimization by finding an interval where $g'(\sigma_s^2)$ changes sign and performing bisection search

Appendix

Numerical Solution for PR-UCF using Newton's Method

TECHNISCHE UNIVERSITÄT DARMSTADT

Define

$$A = \sum_{j=1}^{N} \frac{\left|z_j\right|^2}{\left(\hat{\lambda}_j - \bar{\lambda}_k\left(\sigma_s^2\right)\right)^2}$$

$$B = \sum_{j=1}^{N} \frac{\left|z_j\right|^2}{\left(\hat{\lambda}_j - \bar{\lambda}_k\left(\sigma_s^2\right)\right)^3}$$

Second derivative of $g(\sigma_s^2) = \sum_{k=N}^{M} \bar{\lambda}_k^2\left(\sigma_s^2\right)$

$$g''(\sigma_s^2) = \sum_{k=N}^{M} \frac{4\sigma_s^2\bar{\lambda}_k\left(\sigma_s^2\right)A^2 - 4\bar{\lambda}_k\left(\sigma_s^2\right)B + A}{\sigma_s^8 A^3}$$

Appendix

Eigenvalue Computation

TECHNISCHE UNIVERSITÄT DARMSTADT

Algorithm 1 Determining the k-th eigenvalue $\bar{d}_k$

1: **Initial.**: Iteration index $\tau = 0$, initial point $x^{(0)} \in (d_{k+1}, d_k)$, tolerance $\epsilon = 10^{-9}$

2: **repeat**

3: Approximate $\psi_k(x)$ by determining the parameters p and q such that:

$$R_{k-1;p,q}(x^{(\tau)}) = \psi_k(x^{(\tau)}) \text{ and } R'_{k-1;p,q}(x^{(\tau)}) = \psi'_k(x^{(\tau)})$$

4: Approximate $\phi_k(x)$ by determining the parameters r and s such that:

$$R_{k;r,s}(x^{(\tau)}) = \phi_k(x^{(\tau)}) \text{ and } R'_{k;r,s}(x^{(\tau)}) = \phi'_k(x^{(\tau)})$$

5: Determine $x^{(\tau+1)} \in (d_{k+1}, d_k)$ which satisfies:

$$-R_{k-1;p,q}(x^{(\tau+1)}) = 1 + R_{k;r,s}(x^{(\tau+1)})$$

6: $\tau \leftarrow \tau + 1$

7: **until** $\left| x^{(\tau+1)} - x^{(\tau)} \right| < \epsilon$

8: **return** $\bar{d}_k = x^{(\tau+1)}$

Appendix

Eigenvalue Computation

TECHNISCHE UNIVERSITÄT DARMSTADT

Remarks

- Eigenvalues of the previous direction are reused for initializations
- Reduced execution time by using properties of the trace operator

Application to PR-UCF

$$\bar{\lambda}_k(\sigma_s^2) = \lambda_k\left(\hat{\mathbf{R}} - \sigma_s^2 \mathbf{a}\mathbf{a}^H\right)$$

$$g'\left(\sigma_s^2\right) = -2\mathbf{a}^H\hat{\mathbf{R}}\mathbf{a} + 2\sigma_s^2 \left|\left|\mathbf{a}\right|\right|_2^4 + \sum_{k=1}^{N-1} \frac{2\bar{\lambda}_k\left(\sigma_s^2\right)}{\sigma_{s,0}^4 \mathbf{a}^H\left(\hat{\mathbf{R}} - \bar{\lambda}_k(\sigma_{s,0}^2)\mathbf{I}_M\right)^{-2}\mathbf{a}}$$

Appendix
Eigenvalue Computation

Remarks

- Eigenvalues of the previous direction are reused for initializations
- Reduced execution time by using properties of the trace operator

Application to PR-UCF

$$\bar{\lambda}_k(\sigma_s^2) = \lambda_k\left(\hat{\mathbf{\Lambda}} - \sigma_s^2 \mathbf{z}\mathbf{z}^H\right)$$

$$g'\left(\sigma_s^2\right) = -2\mathbf{z}^H\hat{\mathbf{\Lambda}}\mathbf{z} + 2\sigma_s^2 \left|\left|\mathbf{z}\right|\right|_2^4 + \sum_{k=1}^{N-1} \frac{2\bar{\lambda}_k\left(\sigma_s^2\right)}{\sigma_s^4 \sum_{j=1}^{M} \frac{\left|z_j\right|^2}{\left(\hat{\lambda}_j - \bar{\lambda}_k\left(\sigma_s^2\right)\right)^2}}$$

with $\hat{\mathbf{R}} = \hat{\mathbf{U}}\hat{\mathbf{\Lambda}}\hat{\mathbf{U}}^H$, $\mathbf{z} = \hat{\mathbf{U}}^H\mathbf{a}$

Appendix
Complexity Summary

TECHNISCHE UNIVERSITÄT DARMSTADT

Total computational complexity (including overhead)

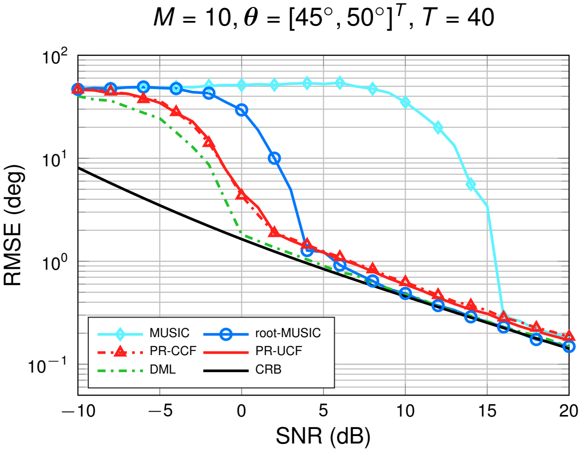
Estimator	**Generic**	**Rank-one Update**
PR-CCF	$O(M^3 N_G)$	$O(M^2 N_G)$
PR-UCF	$O(M^3 N_G N_I)$	$O(M^2 N_G N_I)$
MUSIC	$O(MNN_G)$	

Table: Complexity for computing the null-spectra

- M : Number of sensors
- N : Number of sources
- N_G : Number of look-directions of the complete angle-of-view
- N_I : Number of bisection steps

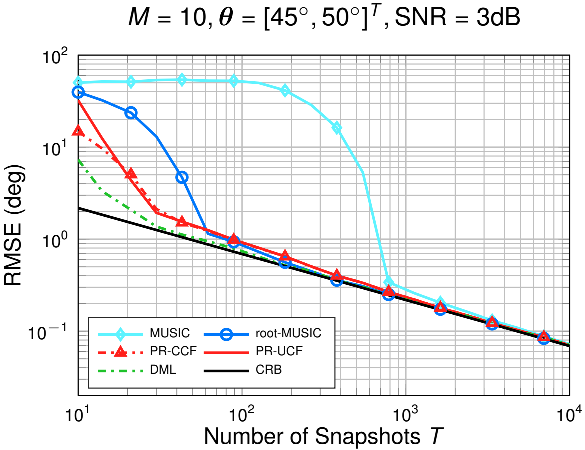
Appendix

Influence of SNR



Appendix

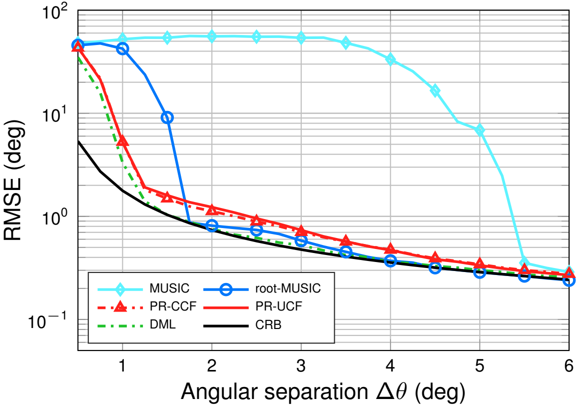
Influence of Number of Snapshots



Appendix

Influence of Angular Separation

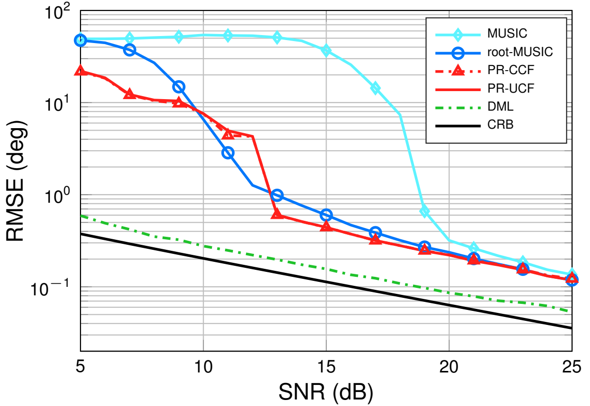
$M = 10, \boldsymbol{\theta} = [45^\circ, 45^\circ + \Delta\theta]^T$, SNR = 10dB, $T = 100$



Appendix

Correlated Source Signals

$M = 10, \boldsymbol{\theta} = [45^\circ, 50^\circ]^T, \rho = 0.95, T = 200$



Appendix
Bisection Method on PR-UCF

$M = 10, \boldsymbol{\theta} = [45^\circ, 50^\circ]^T, T = 100, N_G = 1800$