Technische Universität München

Direction Finding Based on Measurements of a Sequentially Sampled ULA

Research Internship

Ulrike Höfler, Christoph Kaulich, Michael Würth

Supervisor: Andreas Barthelme, Lorenz Weiland

Department of Electrical Engineering and Information Technology
Professur für Methoden der Signalverarbeitung
Univ.-Prof. Dr.-Ing. Wolfgang Utschick

Contents

Chapter 1

Introduction

There are various approaches to face the problem of direction finding. But they are all based on the general idea of fusing data collected from an antenna array in order to give a good estimate on the direction of the unknown, incident signal. Different approaches demand different detailed knowledge about the stochastic properties of the signal. By exploiting the structure of the antenna array a more precise analysis of the system's performance is possible as well as a reduction of the computational effort.

One important consideration regarding the practical implementation of an antenna array compared to its theoretical counterpart is to keep the costs down. In many practical scenarios this is achieved by implementing fewer receivers than antennas. This means a distinction between the total number of antennas and the number of actual active antennas (RF chains) whose data can be read out simultaneously is introduced. The constellation of the active antennas within the array changes over time to leverage the complete aperture of the array. However, fewer receivers imply a reduced or limited performance. Therefore, it is necessary to find a trade-off between cost reduction and the loss in performance which considers the number of incident signals to be resolved and the accuracy of the direction estimation.

The task of this work is to adjust the ML estimation and the MUSIC algorithm –in general and particularly in case of a ULA– to the scenario of fewer receivers than antennas. In this process, we are especially interested in the relation between the number of resolvable signals and the number of available RF chains. Via simulations we compare the performance of different RF chain constellation and the influence of several system parameters. Finally, we try to confirm the theoretical considerations and simulations by carrying out actual measurements on a small antenna array.

This work has the following structure: In Chapter 2, the system model is introduced. In this connection the parameter denotation is determined and the measurement model is defined.
In Chapter 3, the general condition for resolvable signals and thus unique angle estimation is briefly presented.
In Chapter 4, the DML and SML are defined and introduced as data based estimation methods. Also conditions for the identifiability of the estimation problems are specified.
In Chapter 5, the Cramér-Rao Bound of the DML and SML is calculated in order to obtain a basis of comparison between the introduced methods.
In Chapter 6, the MUSIC algorithm is stated in general and in case of a ULA system. Furthermore, the identifiability condition of the MUSIC is presented.
In Chapter 7, the performed simulations are evaluated and compared for several system parameters.
In Chapter 8, the used hardware for the practical implementation is introduced as well as the measurement setup. The experimental results will be presented and evaluated. Finally, a short outline of the work is given. Pitfalls and findings of this approach will be discussed.

Chapter 2

System Model

2.1 General Definitions

In the following, we state the definitions of some frequently used constants. Thus, we define

- number of antennas M,
- number of impinging wavefronts L,
- number of antenna configurations N,
- number of measurements per configuration T,
- number of RF chains K.

Furthermore, the steering vector of an uniform linear array (ULA) is given by

$$\boldsymbol{a}(\theta) = \left[1\,, \mathrm{e}^{-\mathrm{j}\pi\theta}\,, \cdots\,, \mathrm{e}^{-\mathrm{j}(M-1)\pi\theta}\right]^{\mathrm{T}} \in \mathbb{C}^{M}, \tag{2.1}$$

where $\theta = \cos(\omega) \in [-1; 1)$ parametrizes the angle of arrival (AoA) $\omega \in [0; \pi)$.

2.2 Measurement Model

A single coherent measurement for the nth-antenna configuration at one time snapshot t can be written according to [1] as

$$\boldsymbol{x}_n(t) = \sum_{l=1}^{L} \boldsymbol{a}_n(\theta_l) s_{l,n}(t) + \boldsymbol{n}_n(t) \tag{2.2}$$

$$= \sum_{l=1}^{L} \boldsymbol{G}_n \boldsymbol{a}(\theta_l) s_{l,n}(t) + \boldsymbol{n}_n(t) \tag{2.3}$$

$$= \boldsymbol{G}_n \boldsymbol{A}(\boldsymbol{\theta}) \boldsymbol{s}_n(t) + \boldsymbol{n}_n(t), \tag{2.4}$$

with $\boldsymbol{s}_n(t) \in \mathbb{C}^L$, where $\boldsymbol{n}_n(t) \in \mathbb{C}^K \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_K)\ \forall n$ is additive noise and the angles $\theta_1, \ldots, \theta_L$ are collected in a vector $\boldsymbol{\theta}$. The steering matrix $\boldsymbol{A}(\boldsymbol{\theta}) = \left[\boldsymbol{a}(\theta_1), \cdots, \boldsymbol{a}(\theta_L)\right] \in \mathbb{C}^{M \times L}$ contains the ordinary steering vectors according to (2.1). The appropriate entries are selected via the selection matrix $\boldsymbol{G}_n \in \{0; 1\}^{K \times M}$ as defined in [2]. Each row contains exactly one entry equal to 1.

As an abbreviation we define

$$\boldsymbol{A}_n(\boldsymbol{\theta}) = \boldsymbol{G}_n \boldsymbol{A}(\boldsymbol{\theta}). \tag{2.5}$$

In the following, we consider the case of N non-coherent measurements where each measurement n uses a different selection of antennas. The noise of different measurements is assumed to be uncorrelated. Stacking all measurements in a vector $\boldsymbol{x}(t) = \left[\boldsymbol{x}_1(t)^{\mathrm{T}} \cdots \boldsymbol{x}_N(t)^{\mathrm{T}}\right]^{\mathrm{T}}$, the measurement model can be written as

$$\boldsymbol{x}(t) = \boldsymbol{G}(\mathbf{I}_N \otimes \boldsymbol{A}(\boldsymbol{\theta})) \boldsymbol{s}(t) + \boldsymbol{n}(t) \tag{2.6}$$

$$= \tilde{\boldsymbol{A}}(\boldsymbol{\theta}) \boldsymbol{s}(t) + \boldsymbol{n}(t), \tag{2.7}$$

where $\boldsymbol{G} = \mathrm{blkdiag}(\boldsymbol{G}_1, \ldots, \boldsymbol{G}_N)$ and the vectors $\boldsymbol{n}(t) \in \mathbb{C}^{KN}$, $\boldsymbol{s} \in \mathbb{C}^{LN}$ are defined analogous to $\boldsymbol{x}(t) \in \mathbb{C}^{KN}$. Again, we introduce a matrix $\tilde{\boldsymbol{A}}(\boldsymbol{\theta}) = \boldsymbol{G}(\mathbf{I}_N \otimes \boldsymbol{A}(\boldsymbol{\theta}))$ for simplicity.

Chapter 3

General Identifiability Verification

To ensure the uniqueness of the direction estimates $\hat{\boldsymbol{\theta}}$, the following considerations were made.

In case of a uniform linear array (ULA) of antennas, the covariance matrix $\boldsymbol{C_x}$ of the measured data $\boldsymbol{x}$ exhibits a Toeplitz structure. If in addition $\boldsymbol{C_s}$ is diagonal, a Vandermonde Decomposition of $\boldsymbol{C_x}$ exists and thus the covariance matrix of the measured data can be written as

$$\boldsymbol{C_x} = \boldsymbol{A}\boldsymbol{C_s}\boldsymbol{A}^{\mathrm{H}} + \sigma^2\mathbf{I}. \tag{3.1}$$

As the AoAs $\boldsymbol{\theta}$ are uniquely defined by $\boldsymbol{A}(\boldsymbol{\theta})$, we can identify up to $M-1$ angles in the case of a fully sampled array with M antennas.

Note that this is only a weak result since for a finite number of measurements the real covariance matrix is not available [1].

[1] At this point we refer to [3] which presents tighter bounds for the allowed number of sources with respect to the number of antennas for the case of fully sampled ULA systems.

Chapter 4

Maximum Likelihood Estimation

4.1 DML (Deterministic Maximum Likelihood)

By using T independent measurements $\boldsymbol{x}_n(t)$ per configuration the likelihood function is given as

$$L_n(\boldsymbol{\theta}, \boldsymbol{s}_n; \boldsymbol{x}_n) = \frac{1}{\det(\pi\sigma^2\mathbf{I}_K)^T} \prod_{t=1}^{T} \exp\left(\frac{-\left\|\boldsymbol{x}_n(t) - \boldsymbol{A}_n(\boldsymbol{\theta})\boldsymbol{s}_n(t)\right\|_2^2}{\sigma^2} \right). \quad (4.1)$$

Note that $\det(\pi\sigma^2\mathbf{I}_K)^T$ denotes $\det(\pi\sigma^2\mathbf{I}_K)$ to the power of T and not the transpose. Along with N independent configuration measurements the resulting joint likelihood yields

$$L(\boldsymbol{\theta}, \boldsymbol{s}; \boldsymbol{x}) = \prod_{n=1}^{N} L_n(\boldsymbol{\theta}, \boldsymbol{s}_n; \boldsymbol{x}_n), \quad (4.2)$$

where the SNR is assumed to be known.

Option 1

We write the log-Likelihood function in term of the stacked vectors $\boldsymbol{x}(t)$ and $\boldsymbol{s}(t)$ as

$$\ell(\boldsymbol{\theta}, \boldsymbol{s}; \boldsymbol{x}) = -NTK\ln(\pi\sigma^2) - \sum_{t=1}^{T} \frac{1}{\sigma^2} \left\|\boldsymbol{x}(t) - \tilde{\boldsymbol{A}}(\boldsymbol{\theta})\boldsymbol{s}(t)\right\|_2^2. \quad (4.3)$$

In the following we will consider only the ML-problem at one time snapshot t

$$\left(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{s}}(t)\right) = \underset{\boldsymbol{\theta}, \boldsymbol{s}(t)}{\arg\min} \left\|\boldsymbol{x}(t) - \tilde{\boldsymbol{A}}(\boldsymbol{\theta})\boldsymbol{s}(t)\right\|_2^2. \quad (4.4)$$

Solving for a fixed $\boldsymbol{\theta}$ the Least-Squares solution for the signals $\boldsymbol{s}(t)$ is given by

$$\hat{\boldsymbol{s}}(t) = (\tilde{\boldsymbol{A}}^{\mathrm{H}}\tilde{\boldsymbol{A}})^{-1}\tilde{\boldsymbol{A}}^{\mathrm{H}}\boldsymbol{x}(t) \tag{4.5}$$

where we dropped the argument of $\tilde{\boldsymbol{A}}(\boldsymbol{\theta})$ for convenience. Substituting this result into (4.4) yields

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \left\| (\mathbf{I} - \tilde{\boldsymbol{A}}(\tilde{\boldsymbol{A}}^{\mathrm{H}}\tilde{\boldsymbol{A}})^{-1}\tilde{\boldsymbol{A}}^{\mathrm{H}})\boldsymbol{x}(t) \right\|_2^2 \tag{4.6}$$

$$= \arg\min_{\boldsymbol{\theta}} \left\| \boldsymbol{P}_{\tilde{\boldsymbol{A}}}^{\perp}\boldsymbol{x}(t) \right\|_2^2. \tag{4.7}$$

Thus, the ML estimation problem is equal to minimizing the projection of the measured signal onto the orthogonal complement of the range of $\tilde{\boldsymbol{A}}$. The projection is denoted by $\boldsymbol{P}_{\tilde{\boldsymbol{A}}}^{\perp} = \mathbf{I} - \tilde{\boldsymbol{A}}(\tilde{\boldsymbol{A}}^{\mathrm{H}}\tilde{\boldsymbol{A}})^{-1}\tilde{\boldsymbol{A}}^{\mathrm{H}}$. Instead of minimizing the projection onto the orthogonal complement of the range of $\tilde{\boldsymbol{A}}$, we can equivalently maximize the projection onto the range of $\tilde{\boldsymbol{A}}$. Therefore, the ML problem can be rewritten as

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \left\| \boldsymbol{P}_{\tilde{\boldsymbol{A}}}\boldsymbol{x}(t) \right\|_2^2 \tag{4.8}$$

$$= \arg\max_{\boldsymbol{\theta}} \boldsymbol{x}^{\mathrm{H}}\boldsymbol{P}_{\tilde{\boldsymbol{A}}}\boldsymbol{x}(t) \tag{4.9}$$

with the projector

$$\boldsymbol{P}_{\tilde{\boldsymbol{A}}} = \tilde{\boldsymbol{A}}(\tilde{\boldsymbol{A}}^{\mathrm{H}}\tilde{\boldsymbol{A}})^{-1}\tilde{\boldsymbol{A}}^{\mathrm{H}} \tag{4.10}$$

$$= \boldsymbol{G}(\mathbf{I}_N \otimes \boldsymbol{A}(\boldsymbol{\theta}))((\mathbf{I}_N \otimes \boldsymbol{A}(\boldsymbol{\theta}))^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{G} (\mathbf{I}_N \otimes \boldsymbol{A}(\boldsymbol{\theta})))^{-1}(\mathbf{I}_N \otimes \boldsymbol{A}(\boldsymbol{\theta}))^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}. \tag{4.11}$$

Option 2

Alternatively, we can write the log-Likelihood in terms of the single measurements as

$$\ell_n(\boldsymbol{\theta}, \boldsymbol{s}; \boldsymbol{x}) = -TK\ln(\pi\sigma^2) - \frac{1}{\sigma^2}\sum_{t=1}^{T} \left\| \boldsymbol{x}_n(t) - \boldsymbol{A}_n(\boldsymbol{\theta})\boldsymbol{s}_n(t) \right\|_2^2 \tag{4.12}$$

with the corresponding ML problem

$$\left(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{s}}\right) = \arg\min_{\boldsymbol{\theta}, \boldsymbol{s}_n(t)} \sum_{n=1}^{N}\sum_{t=1}^{T} \left\| \boldsymbol{x}_n(t) - \boldsymbol{A}_n(\boldsymbol{\theta})\boldsymbol{s}_n(t) \right\|_2^2. \tag{4.13}$$

For a fixed $\boldsymbol{\theta}$ this optimization problem decouples into N independent LS-problems w.r.t. $\boldsymbol{s}_n(t), n = 1, \ldots, N$ where the solution for one time snapshot t is given by

$$\hat{\boldsymbol{s}}_n(t) = (\boldsymbol{A}_n^{\mathrm{H}}\boldsymbol{A}_n)^{-1}\boldsymbol{A}_n^{\mathrm{H}}\boldsymbol{x}_n(t). \tag{4.14}$$

Substituting this result into (4.13) yields

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \sum_{n=1}^{N} \sum_{t=1}^{T} \left\| \boldsymbol{P}_{\boldsymbol{A}_n} \boldsymbol{x}_n(t) \right\|_2^2 \tag{4.15}$$

$$= \arg\max_{\boldsymbol{\theta}} \left\| \mathrm{blkdiag}(\boldsymbol{P}_{\boldsymbol{A}_1}, \ldots, \boldsymbol{P}_{\boldsymbol{A}_N}) \boldsymbol{x}(t) \right\|_2^2 . \tag{4.16}$$

The projector onto the range of $\boldsymbol{A}_n$ is defined as

$$\boldsymbol{P}_{\boldsymbol{A}_n} = \boldsymbol{A}_n (\boldsymbol{A}_n^{\mathrm{H}} \boldsymbol{A}_n)^{-1} \boldsymbol{A}_n^{\mathrm{H}} \tag{4.17}$$

$$= \boldsymbol{G}_n \boldsymbol{A} (\boldsymbol{A}^{\mathrm{H}} \boldsymbol{G}_n^{\mathrm{H}} \boldsymbol{G}_n \boldsymbol{A})^{-1} \boldsymbol{A}^{\mathrm{H}} \boldsymbol{G}_n^{\mathrm{H}}. \tag{4.18}$$

Note that by comparing this expression with (4.8), the projector in (4.10) can be rewritten as

$$\boldsymbol{P}_{\tilde{\boldsymbol{A}}} = \mathrm{blkdiag}(\boldsymbol{P}_{\boldsymbol{A}_1}, \ldots, \boldsymbol{P}_{\boldsymbol{A}_N}). \tag{4.19}$$

4.2 DML Specific Identifiability

To proof the identifiability of a system as shown in [2], we simply check if the total number of given data parameters is greater or equal than the number of unknown parameters which have to be estimated.

In the DML case we have NT signal vectors $\boldsymbol{s}_n(t) \in \mathbb{C}^L$ and the angle vector $\boldsymbol{\theta} \in \mathbb{R}^L$ to estimate. This leads to $2LTN + L$ unknown parameters, whereat the factor 2 comes from the imaginary and real part of $\boldsymbol{s}_n(t)$. The number of given data parameter $x_n(t) \in \mathbb{C}^K$ amounts to $2NKT$. This yields the necessary condition

$$2NKT \geq 2LTN + L.$$

4.3 SML (Stochastic Maximum Likelihood)

Compared to the Deterministic ML estimator the stochastic ML estimator introduces the signal $\boldsymbol{s}_n(t), n = 1, \ldots, N$ as a random variable with normal distribution $\boldsymbol{s}_n(t) \sim \mathcal{CN}(\boldsymbol{\mu} = \boldsymbol{0}, \boldsymbol{C}_{\boldsymbol{s}} = \mathbf{I})\ \forall t, N$.

Analogue to the deterministic case, we first take a look at the likelihood function of a single measurement $\boldsymbol{x}_n(t) \sim \mathcal{CN}(\boldsymbol{\mu} = \boldsymbol{0}, \boldsymbol{C}_{\boldsymbol{x}_n} = \boldsymbol{A}_n \boldsymbol{C}_{\boldsymbol{s}} \boldsymbol{A}_n^{\mathrm{H}} + \sigma^2 \mathbf{I}_K)\ \forall t$:

$$L_n(\boldsymbol{\theta}, \boldsymbol{C}_{\boldsymbol{s}}; \boldsymbol{x}_n) = \frac{1}{\det(\pi \boldsymbol{C}_{\boldsymbol{x}_n})^T} \prod_{t=1}^{T} \exp\Big(- \boldsymbol{x}_n(t)^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{x}_n}^{-1} \boldsymbol{x}_n(t) \Big). \tag{4.20}$$

Again, we assume that the SNR is known. The log-likelihood function with N

independent configuration measurements reads as

$$\ell(\boldsymbol{\theta}, \boldsymbol{C_s}; \boldsymbol{x}) = \sum_{n=1}^{N} -T \ln\Big(\det(\pi \boldsymbol{C}_{\boldsymbol{x}_n})\Big) - \sum_{n=1}^{N}\sum_{t=1}^{T} \boldsymbol{x}_n(t)^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{x}_n}^{-1} \boldsymbol{x}_n(t) \tag{4.21}$$

$$= \sum_{n=1}^{N} -T \ln\Big(\det(\pi \boldsymbol{C}_{\boldsymbol{x}_n})\Big) - \sum_{n=1}^{N}\sum_{t=1}^{T} \operatorname{tr}\Big\{\boldsymbol{C}_{\boldsymbol{x}_n}^{-1} \boldsymbol{x}_n(t)\boldsymbol{x}_n(t)^{\mathrm{H}}\Big\} \tag{4.22}$$

$$= \sum_{n=1}^{N} -T \ln\Big(\det(\pi \boldsymbol{C}_{\boldsymbol{x}_n})\Big) - \sum_{n=1}^{N} T \operatorname{tr}\Big\{\boldsymbol{C}_{\boldsymbol{x}_n}^{-1} \hat{\boldsymbol{C}}_{\boldsymbol{x}_n}\Big\}, \tag{4.23}$$

with $\hat{\boldsymbol{C}}_{\boldsymbol{x}_n} = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{x}_n(t)\boldsymbol{x}_n(t)^{\mathrm{H}}$ as the sample covariance matrix.

The corresponding optimization problem reads as

$$\Big(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{C}}_{\boldsymbol{s}}\Big) = \underset{\boldsymbol{\theta}, \boldsymbol{C_s}}{\arg\min} \sum_{n=1}^{N} T \ln\Big(\det(\pi \boldsymbol{C}_{\boldsymbol{x}_n})\Big) + \sum_{n=1}^{N} T \operatorname{tr}\Big\{\boldsymbol{C}_{\boldsymbol{x}_n}^{-1} \hat{\boldsymbol{C}}_{\boldsymbol{x}_n}\Big\}. \tag{4.24}$$

Notice, that we can not easily substitute the covariance matrix $\hat{\boldsymbol{C}}_{\boldsymbol{s}}$ as it was done with $\hat{\boldsymbol{s}}(t)$ in (4.5) in the DML method. Therefore, it is necessary to apply some grid based estimation of $\hat{\boldsymbol{C}}_{\boldsymbol{s}}$. However, this is computationally very expensive. To avoid this additional complexity of the problem and to reduce the computational effort of our simulations, we assumed that the signal covariance matrix is known to the SML estimator (together with the SNR).

4.4 SML Specific Identifiability

For the SML case we have to estimate the signal covariance matrix $\boldsymbol{C_s} \in \mathbb{C}^{L\times L}$ and the angles $\boldsymbol{\theta} \in \mathbb{R}^L$ which adds up to $L^2 + L$ parameters to estimate as described in [2]. The given data per configuration $\boldsymbol{x}_n(t)$ occurs in the sample covariance matrix $\hat{\boldsymbol{C}}_{\boldsymbol{x}_n} \in \mathbb{C}^{K\times K}$. However, the total number of free components within $\hat{\boldsymbol{C}}_{\boldsymbol{x}_n}$ depends on the relation between T and K:

- correlated signals $\boldsymbol{s}_n(t)$
  - $T \geq K$
    If $T \geq K$ holds, there are NK^2 parameters to be considered including all N configurations. This leads to the following inequality:
    $$NK^2 \geq L^2 + L.$$
  - $T < K$
    for $T < K$ the sample covariance matrix $\hat{\boldsymbol{C}}_{\boldsymbol{x}_n}$ is rank deficient which leads to KT free parameters per configuration:
    $$NKT \geq L^2 + L.$$

- uncorrelated signals $\boldsymbol{s}_n(t)$

  Note that if the signal $\boldsymbol{s}_n(t)$ are uncorrelated, the signal covariance matrix $\boldsymbol{C_s}$ is a diagonal matrix with L real entries which means there are – including the L angles – $2L$ parameters to estimate as described in [2].

  - $T \geq K$

    $$NK^2 \geq 2L.$$

  - $T < K$

    $$NKT \geq 2L.$$

Chapter 5

Cramér-Rao Bounds

Next, to get a lower bound on the variance of the already derived ML estimator, the Cramér-Rao Bound will be considered.

The general expression of the Cramér-Rao bound for an unbiased estimator reads as:

$$\boldsymbol{C}_{\hat{\boldsymbol{p}}} = \boldsymbol{F}(\boldsymbol{p})^{-1}, \tag{5.1}$$

where $\boldsymbol{C}_{\hat{\boldsymbol{p}}}$ is the covariance matrix of the deterministic ML estimator and $\boldsymbol{F}(\boldsymbol{\theta})$ the Fisher Information matrix (FIM).

5.1 DML

Since the measured data $\boldsymbol{x}(t) \in \mathbb{R}^{KN}$ is i.i.d with $\boldsymbol{x}(t) \sim \mathcal{CN}(\boldsymbol{\mu}(t) = \tilde{\boldsymbol{A}}\boldsymbol{s}(t), \boldsymbol{C}_{\boldsymbol{x}} = \sigma^2\mathbf{I}_{KN})\ \forall t$, the Slepian-Bangs formula defined in [4] provides an expression for the Fisher Information matrix elements for T time snapshot as follows:

$$\begin{aligned}[\boldsymbol{F}]_{j,k} &= \mathrm{E}\left[\frac{\partial \ell(\boldsymbol{\theta}, \boldsymbol{s}; \boldsymbol{x})}{\partial p_j}\frac{\partial \ell(\boldsymbol{\theta}, \boldsymbol{s}; \boldsymbol{x})}{\partial\, p_k}\right] \\ &= 2\sum_{t=1}^{T} \Re\left\{\frac{\partial \boldsymbol{\mu}(t)^{\mathrm{H}}}{\partial p_j}\boldsymbol{C}_{\boldsymbol{x}}^{-1}\frac{\partial \boldsymbol{\mu}(t)}{\partial p_k}\right\} + T\,\mathrm{tr}\{\boldsymbol{C}_{\boldsymbol{x}}^{-1}\boldsymbol{C}_{\boldsymbol{x}}^{j}\boldsymbol{C}_{\boldsymbol{x}}^{-1}\boldsymbol{C}_{\boldsymbol{x}}^{k}\},\end{aligned} \tag{5.2}$$

where $\boldsymbol{C}_{\boldsymbol{x}}^{j} = \frac{\partial \boldsymbol{C}_{\boldsymbol{x}}^{j}}{\partial p_j}$ denotes the derivative of the covariance matrix $\boldsymbol{C}_{\boldsymbol{x}}$ in terms of the entries of the parameter $\boldsymbol{p} = [\boldsymbol{\theta}^{\mathrm{T}}, \boldsymbol{s}_R^{\mathrm{T}} = \Re\{\boldsymbol{s}^{\mathrm{T}}\}, \boldsymbol{s}_I^{\mathrm{T}} = \Im\{\boldsymbol{s}^{\mathrm{T}}\}, \sigma^2]^{\mathrm{T}}$[1]. This parameter

[1] Although σ^2 is within the vector of unknown parameters, we do not consider σ^2 in the following calculation for the DML case, since it is independent of all other unknown parameters and thereby irrelevant for the calculation of the sub–matrix of interest $\boldsymbol{C}_{\hat{\boldsymbol{p}}}^{\boldsymbol{\theta}\boldsymbol{\theta}}$.

$\boldsymbol{p}$ is a vector, containing all unknown variables which shall be estimated explicitly or implicitly, respectively. Note that the vectors $\boldsymbol{s}_R, \boldsymbol{s}_I \in \mathbb{C}^{LTN}$ contain the stacked signal vectors per configuration N and time snapshot $t = 1, \ldots, T$.

Since, the covariance matrix $\boldsymbol{C}_{\boldsymbol{x}}$ does not depend on $\boldsymbol{\theta}$, (5.2) can be simplified to

$$[\boldsymbol{F}]_{j,k} = 2 \sum_{t=1}^{T} \Re \left\{ \frac{\partial \boldsymbol{\mu}(t)^{\mathrm{H}}}{\partial p_j} \boldsymbol{C}_{\boldsymbol{x}}^{-1} \frac{\partial \boldsymbol{\mu}(t)}{\partial p_k} \right\}. \tag{5.3}$$

	$\theta_1 \cdots \theta_L$	$s_R^1 \quad \cdots \quad s_R^{LTN}$	$s_I^1 \quad \cdots \quad s_I^{LTN}$
θ_1 $\vdots$ θ_L	$\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{\theta}}$	$\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{s}_R}$	$\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{s}_I}$
s_R^1 $\vdots$ s_R^{LTN}	$\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{s}_R}^{\mathrm{H}}$	$\boldsymbol{F}_{\boldsymbol{s}_R,\boldsymbol{s}_R}$	$\boldsymbol{F}_{\boldsymbol{s}_R,\boldsymbol{s}_I}$
s_I^1 $\vdots$ s_I^{LTN}	$\boldsymbol{F}_{\boldsymbol{\theta}\boldsymbol{s}_I}^{\mathrm{H}}$	$\boldsymbol{F}_{\boldsymbol{s}_R,\boldsymbol{s}_I}^{\mathrm{H}}$	$\boldsymbol{F}_{\boldsymbol{s}_I,\boldsymbol{s}_I}$

Figure 5.1: DML Fisher Information matrix $\boldsymbol{F}(\boldsymbol{p})$ sectioned in sub–matrices.

As shown in Figure 5.1 we can section the FIM in several separate sub–matrices as described in [5]. These sub-matrices differ in the combination of elements in $\boldsymbol{p}$ within (5.2).

In the following we will calculate the sub–matrices of the FIM separately.

- $\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{\theta}} \in \mathbb{C}^{L \times L}$

$$[\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{\theta}}]_{j,k} = 2 \sum_{t=1}^{T} \Re \left\{ \frac{\partial \boldsymbol{\mu}(t)^{\mathrm{H}}}{\partial \theta_j} \boldsymbol{C}_{\boldsymbol{x}}^{-1} \frac{\partial \boldsymbol{\mu}(t)}{\partial \theta_k} \right\}$$

  together with the required derivatives

$$\frac{\partial \boldsymbol{\mu}(t)}{\partial \theta_\ell} = \frac{\partial}{\partial \theta_\ell} \boldsymbol{G}(\mathbf{I}_N \otimes \boldsymbol{A}(\boldsymbol{\theta}))\boldsymbol{s}(t) = \boldsymbol{G} \left(\mathbf{I}_N \otimes \frac{\partial}{\partial \theta_\ell} \boldsymbol{A}(\boldsymbol{\theta}) \right) \boldsymbol{s}(t)$$

and

$$\frac{\partial}{\partial \theta_\ell}\boldsymbol{A}(\boldsymbol{\theta}) = \Big[\mathbf{0}\,,\,\cdots,\mathbf{0}\,,\,\tfrac{\partial \boldsymbol{a}(\theta_\ell)}{\partial \theta_\ell}\,,\mathbf{0}\,,\cdots,\mathbf{0}\Big] = \Big[\mathbf{0}\,,\,\cdots,\mathbf{0}\,,\boldsymbol{D}\boldsymbol{a}(\theta_\ell)\,,\mathbf{0}\,,\cdots,\mathbf{0}\Big]\,, \tag{5.4}$$

where

$$\boldsymbol{D} = -j\pi \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & M-1 \end{bmatrix}$$

is the differential operator. For the sake of convenience, $\cos(\theta) \in [-1;1)$ was parameterised by $\theta \in [-1;1)$.

Putting all together, the sub–matrix elements for $(\boldsymbol{\theta},\boldsymbol{\theta})$ part read as

$$[\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{\theta}}]_{j,k} = 2\sum_{t=1}^{T} \Re\Big\{ \Big(\boldsymbol{G}\Big(\mathbf{I}_N \otimes \Big[\mathbf{0}\,,\,\cdots,\mathbf{0}\,,\boldsymbol{D}\boldsymbol{a}(\theta_j)\,,\mathbf{0}\,,\cdots,\mathbf{0}\Big]\Big)\boldsymbol{s}(t)\Big)^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{x}}^{-1} \Big(\boldsymbol{G}\Big(\mathbf{I}_N \otimes \Big[\mathbf{0}\,,\,\cdots,\mathbf{0}\,,\boldsymbol{D}\boldsymbol{a}(\theta_k)\,,\mathbf{0}\,,\cdots,\mathbf{0}\Big]\Big)\boldsymbol{s}(t)\Big)\Big\}.$$

- $\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{s}_R} \in \mathbb{C}^{L\times LTN}$

  The mixture of the derivatives with respect to $\boldsymbol{\theta}$ and $\boldsymbol{s}_R$ yields

$$\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{s}_R} = \begin{bmatrix} 2\sum_{t=1}^{T} \Re\left\{ \frac{\partial \boldsymbol{\mu}(t)^{\mathrm{H}}}{\partial \theta_1} \boldsymbol{C}_{\boldsymbol{x}}^{-1} \frac{\partial \boldsymbol{\mu}(t)}{\partial \boldsymbol{s}_R^{\mathrm{T}}} \right\} \\ \vdots \\ 2\sum_{t=1}^{T} \Re\left\{ \frac{\partial \boldsymbol{\mu}(t)^{\mathrm{H}}}{\partial \theta_L} \boldsymbol{C}_{\boldsymbol{x}}^{-1} \frac{\partial \boldsymbol{\mu}(t)}{\partial \boldsymbol{s}_R^{\mathrm{T}}} \right\} \end{bmatrix}, \tag{5.5}$$

  where we need in addition the derivative

$$\frac{\partial \boldsymbol{\mu}(t)}{\partial \boldsymbol{s}_R^{\mathrm{T}}} = \Big[\mathbf{0}\,,\,\cdots,\mathbf{0}\,,\,\tfrac{\partial \boldsymbol{\mu}(t)}{\partial \boldsymbol{s}_R(t)^{\mathrm{T}}}\,,\mathbf{0}\,,\cdots,\mathbf{0}\Big] = \Big[\mathbf{0}\,,\,\cdots,\mathbf{0}\,,\tilde{\boldsymbol{A}}(\boldsymbol{\theta})\,,\mathbf{0}\,,\cdots,\mathbf{0}\Big]\,.$$

  Here, we defined $\boldsymbol{s}_R(t) \in \mathbb{C}^{LN}$ as the stacking of signal vectors per configuration N at one time snapshot t.

  Another way to include the T time snapshots is to look at the t-th part of the $(\boldsymbol{\theta},\boldsymbol{s}_R)$ sub–matrix and stack the results for $t = 1,\ldots,T$.

$$\begin{aligned} \boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{s}_R} &= \Big[\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{s}_R(1)} \,\ldots\, \boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{s}_R(t)}\Big] \\ &= \begin{bmatrix} 2\Re\left\{ \frac{\partial \boldsymbol{\mu}(1)^{\mathrm{H}}}{\partial \theta_1} \boldsymbol{C}_{\boldsymbol{x}}^{-1} \frac{\partial \boldsymbol{\mu}(1)}{\partial \boldsymbol{s}_R(1)^{\mathrm{T}}} \right\} & \ldots & 2\Re\left\{ \frac{\partial \boldsymbol{\mu}(T)^{\mathrm{H}}}{\partial \theta_1} \boldsymbol{C}_{\boldsymbol{x}}^{-1} \frac{\partial \boldsymbol{\mu}(T)}{\partial \boldsymbol{s}_R(T)^{\mathrm{T}}} \right\} \\ \vdots & & \vdots \\ 2\Re\left\{ \frac{\partial \boldsymbol{\mu}(1)^{\mathrm{H}}}{\partial \theta_L} \boldsymbol{C}_{\boldsymbol{x}}^{-1} \frac{\partial \boldsymbol{\mu}(1)}{\partial \boldsymbol{s}_R(1)^{\mathrm{T}}} \right\} & \ldots & 2\Re\left\{ \frac{\partial \boldsymbol{\mu}(T)^{\mathrm{H}}}{\partial \theta_L} \boldsymbol{C}_{\boldsymbol{x}}^{-1} \frac{\partial \boldsymbol{\mu}(T)}{\partial \boldsymbol{s}_R(T)^{\mathrm{T}}} \right\} \end{bmatrix}, \end{aligned}$$

where $\frac{\partial \boldsymbol{\mu}(t)}{\partial \boldsymbol{s}_R(t)^{\mathrm{T}}} = \tilde{\boldsymbol{A}}(\boldsymbol{\theta})\ \forall t = 1, \ldots, T$.

- $\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{s}_I} \in \mathbb{C}^{L \times LTN}$

  Analogue to the $(\boldsymbol{\theta}, \boldsymbol{s}_R)$ sub–matrix we define $\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{s}_I}$ as

$$\begin{aligned}\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{s}_I} &= \left[\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{s}_I(1)} \ \cdots \ \boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{s}_I(T)}\right] \\ &= \begin{bmatrix} 2\Re\left\{\frac{\partial \boldsymbol{\mu}(1)^{\mathrm{H}}}{\partial \theta_1}\boldsymbol{C}_{\boldsymbol{x}}^{-1}\frac{\partial \boldsymbol{\mu}(1)}{\partial \boldsymbol{s}_I(1)^{\mathrm{T}}}\right\} & \cdots & 2\Re\left\{\frac{\partial \boldsymbol{\mu}(T)^{\mathrm{H}}}{\partial \theta_1}\boldsymbol{C}_{\boldsymbol{x}}^{-1}\frac{\partial \boldsymbol{\mu}(T)}{\partial \boldsymbol{s}_I(T)^{\mathrm{T}}}\right\} \\ \vdots & & \vdots \\ 2\Re\left\{\frac{\partial \boldsymbol{\mu}(1)^{\mathrm{H}}}{\partial \theta_L}\boldsymbol{C}_{\boldsymbol{x}}^{-1}\frac{\partial \boldsymbol{\mu}(1)}{\partial \boldsymbol{s}_I(1)^{\mathrm{T}}}\right\} & \cdots & 2\Re\left\{\frac{\partial \boldsymbol{\mu}(T)^{\mathrm{H}}}{\partial \theta_L}\boldsymbol{C}_{\boldsymbol{x}}^{-1}\frac{\partial \boldsymbol{\mu}(T)}{\partial \boldsymbol{s}_I(T)^{\mathrm{T}}}\right\} \end{bmatrix},\end{aligned}$$

  with $\frac{\partial \boldsymbol{\mu}(t)}{\partial \boldsymbol{s}_I(t)^{\mathrm{T}}} = j\tilde{\boldsymbol{A}}(\boldsymbol{\theta})\ \forall t = 1, \ldots, T$.

- $\boldsymbol{F}_{\boldsymbol{s}_R,\boldsymbol{s}_R} \in \mathbb{C}^{LTN \times LTN}$

$$\begin{aligned}\boldsymbol{F}_{\boldsymbol{s}_R,\boldsymbol{s}_R} &= 2\sum_{t=1}^{T}\Re\left\{\left(\frac{\partial \boldsymbol{\mu}(t)}{\partial \boldsymbol{s}_R^{\mathrm{T}}}\right)^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{x}}^{-1}\frac{\partial \boldsymbol{\mu}(t)}{\partial \boldsymbol{s}_R^{\mathrm{T}}}\right\} \\ &= 2\Re\left\{\begin{bmatrix} \tilde{\boldsymbol{A}}(\boldsymbol{\theta})^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{x}}^{-1}\tilde{\boldsymbol{A}}(\boldsymbol{\theta}) & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \tilde{\boldsymbol{A}}(\boldsymbol{\theta})^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{x}}^{-1}\tilde{\boldsymbol{A}}(\boldsymbol{\theta}) & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \tilde{\boldsymbol{A}}(\boldsymbol{\theta})^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{x}}^{-1}\tilde{\boldsymbol{A}}(\boldsymbol{\theta}) \end{bmatrix}\right\} \\ &= 2\Re\left\{\mathbf{I}_T \otimes (\tilde{\boldsymbol{A}}(\boldsymbol{\theta})^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{x}}^{-1}\tilde{\boldsymbol{A}}(\boldsymbol{\theta}))\right\}.\end{aligned}$$

- $\boldsymbol{F}_{\boldsymbol{s}_I,\boldsymbol{s}_I} \in \mathbb{C}^{LTN \times LTN}$

  Again with the same approach as for the $(\boldsymbol{s}_R, \boldsymbol{s}_R)$ sub–matrix,

$$\begin{aligned}\boldsymbol{F}_{\boldsymbol{s}_I,\boldsymbol{s}_I} &= 2\sum_{t=1}^{T}\Re\left\{\left(\frac{\partial \boldsymbol{\mu}(t)}{\partial \boldsymbol{s}_I^{\mathrm{T}}}\right)^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{x}}^{-1}\frac{\partial \boldsymbol{\mu}(t)}{\partial \boldsymbol{s}_I^{\mathrm{T}}}\right\} \\ &= \boldsymbol{F}_{\boldsymbol{s}_R,\boldsymbol{s}_R}.\end{aligned}$$

- $\boldsymbol{F}_{\boldsymbol{s}_R,\boldsymbol{s}_I} \in \mathbb{C}^{LTN \times LTN}$

The combination of derivatives with respect to $\boldsymbol{s}_R$ and $\boldsymbol{s}_I$ results in

$$\begin{aligned}\boldsymbol{F}_{\boldsymbol{s}_R,\boldsymbol{s}_I} &= 2\sum_{t=1}^{T}\Re\left\{\left(\frac{\partial\boldsymbol{\mu}(t)}{\partial\boldsymbol{s}_R^{\mathrm{T}}}\right)^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{x}}^{-1}\frac{\partial\boldsymbol{\mu}(t)}{\partial\boldsymbol{s}_I^{\mathrm{T}}}\right\}\\ &= 2\Re\left\{\begin{bmatrix}\tilde{\boldsymbol{A}}(\boldsymbol{\theta})^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{x}}^{-1}j\tilde{\boldsymbol{A}}(\boldsymbol{\theta}) & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \tilde{\boldsymbol{A}}(\boldsymbol{\theta})^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{x}}^{-1}j\tilde{\boldsymbol{A}}(\boldsymbol{\theta}) & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \tilde{\boldsymbol{A}}(\boldsymbol{\theta})^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{x}}^{-1}j\tilde{\boldsymbol{A}}(\boldsymbol{\theta})\end{bmatrix}\right\}\\ &= 2\Re\left\{\mathbf{I}_T\otimes(\tilde{\boldsymbol{A}}(\boldsymbol{\theta})^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{x}}^{-1}j\tilde{\boldsymbol{A}}(\boldsymbol{\theta}))\right\}.\end{aligned}$$

Since the FIM is always a hermitian matrix, it is not necessary to calculate the remaining lower triangular part of FIM $\boldsymbol{F}(\boldsymbol{p})$. Finally, to get the overall Cramér-Rao Bound $\boldsymbol{C}_{\hat{\boldsymbol{p}}}$ we invert $\boldsymbol{F}(\boldsymbol{p})$. Since we are only interested in the $(\boldsymbol{\theta},\boldsymbol{\theta})$–section, we just consider the upper left $L\times L$ sub–matrix $\boldsymbol{C}_{\hat{\boldsymbol{p}}}^{\boldsymbol{\theta}\boldsymbol{\theta}}$.

5.2 SML

Again, we calculate the Cramér-Rao Bound with the already introduced Slepian-Bangs formula in (5.2). Since for the stochastic ML the signal $\boldsymbol{s}(t)\in\mathbb{C}^{NL}$ is a random variable with zero mean for all t, the measured data $\boldsymbol{x}(t)\in\mathbb{C}^{KN}$ exhibits another distribution $\boldsymbol{x}(t)\sim\mathcal{CN}\Big(\boldsymbol{\mu}(t)=\mathbf{0},\boldsymbol{C}_{\boldsymbol{x}}=\tilde{\boldsymbol{A}}(\boldsymbol{\theta})(\mathbf{I}_N\otimes\boldsymbol{C}_{\boldsymbol{s}})\tilde{\boldsymbol{A}}(\boldsymbol{\theta})^{\mathrm{H}}+\sigma^2\mathbf{I}_{NK}\Big)\ \forall t$. Note that the mean $\boldsymbol{\mu}(t)$ is equal to $\mathbf{0}\,\forall t$ which is why the first term in the Slepian-Bangs formula is omitted (cf. (5.2)).

$$[\boldsymbol{F}]_{j,k}=T\,\mathrm{tr}\{\boldsymbol{C}_{\boldsymbol{x}}^{-1}\boldsymbol{C}_{\boldsymbol{x}}^{j}\boldsymbol{C}_{\boldsymbol{x}}^{-1}\boldsymbol{C}_{\boldsymbol{x}}^{k}\},\tag{5.6}$$

where $\boldsymbol{C}_{\boldsymbol{x}}^{j}=\frac{\partial\boldsymbol{C}_{\boldsymbol{x}}^{j}}{\partial p_j}$. Again we define a parameter vector

$$\boldsymbol{p}=[\boldsymbol{\theta}^{\mathrm{T}},\underbrace{\boldsymbol{c}_{\mathrm{off}_R}^{\mathrm{T}},\boldsymbol{c}_{\mathrm{diag}}^{\mathrm{T}},\boldsymbol{c}_{\mathrm{off}_I}^{\mathrm{T}}}_{\boldsymbol{c}^{\mathrm{T}}},\sigma^2]^{\mathrm{T}}$$

for the unknown variables[2], with the vector $\boldsymbol{c}\in\mathbb{C}^{L^2}$ containing the entries of the upper triangular matrix of the signal covariance matrix $\boldsymbol{C}_{\boldsymbol{s}}$, sectioned by the off–diagonal real part $\boldsymbol{c}_{\mathrm{off},R}\in\mathbb{R}^{\frac{1}{2}(L^2-L)}$, the diagonal part $\boldsymbol{c}_{\mathrm{diag}}\in\mathbb{R}^{L}$ and the off–diagonal imaginary part $\boldsymbol{c}_{\mathrm{off},I}\in\mathbb{R}^{\frac{1}{2}(L^2-L)}$. Figure 5.2 shows the sub–matrix decomposition of the FIM for the SML case.

[2]Here we included σ^2 in the vector of unknown parameters in contrast to the optimization of the SML likelihood where we considered σ^2 as given. The reason this was done, is that we wanted the Cramér-Rao bound to be less optimistic.

	$\theta_1 \cdots \theta_L$	$\boldsymbol{c}_{\mathrm{off}_R}^{\mathrm{T}} \cdots \boldsymbol{c}_{\mathrm{diag}}^{\mathrm{T}} \cdots \boldsymbol{c}_{\mathrm{off}_I}^{\mathrm{T}}$	σ^2
θ_1 $\vdots$ θ_L	$\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{\theta}}$	$\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{c}}$	$\boldsymbol{F}_{\boldsymbol{\theta},\sigma^2}$
$\boldsymbol{c}_{\mathrm{off}_R}$ $\vdots$ $\boldsymbol{c}_{\mathrm{diag}_R}$ $\vdots$ $\boldsymbol{c}_{\mathrm{off}_I}$	$\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{c}}^{\mathrm{H}}$	$\boldsymbol{F}_{\boldsymbol{c},\boldsymbol{c}}$	$\boldsymbol{F}_{\boldsymbol{c},\sigma^2}$
σ^2	$\boldsymbol{F}_{\boldsymbol{\theta},\sigma^2}^{\mathrm{H}}$	$\boldsymbol{F}_{\boldsymbol{c},\sigma^2}^{\mathrm{H}}$	$\boldsymbol{F}_{\sigma^2,\sigma^2}$

Figure 5.2: SML Fisher Information matrix $\boldsymbol{F}(\boldsymbol{p})$ sectioned in sub–matrices.

- $\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{\theta}} \in \mathbb{C}^{L\times L}$

$$
\begin{aligned}
[\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{\theta}}]_{j,k} &= T\,\mathrm{tr}\{\boldsymbol{C}_{\boldsymbol{x}}^{-1}\boldsymbol{C}_{\boldsymbol{x}}^{j}\boldsymbol{C}_{\boldsymbol{x}}^{-1}\boldsymbol{C}_{\boldsymbol{x}}^{k}\} \\
&= T\,\mathrm{tr}\left\{\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial\theta_j}\boldsymbol{C}_{\boldsymbol{x}}\right)\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial\theta_k}\boldsymbol{C}_{\boldsymbol{x}}\right)\right\},
\end{aligned}
$$

with

$$
\begin{aligned}
\frac{\partial}{\partial\theta_j}\boldsymbol{C}_{\boldsymbol{x}} &= \frac{\partial}{\partial\theta_j}\Big(\tilde{\boldsymbol{A}}(\boldsymbol{\theta})(\mathbf{I}_N\otimes\boldsymbol{C}_{\boldsymbol{s}})\tilde{\boldsymbol{A}}(\boldsymbol{\theta}) + \sigma^2\boldsymbol{G}\mathbf{I}_{NM}\boldsymbol{G}^{\mathrm{H}}\Big) \\
&= \frac{\partial}{\partial\theta_j}\Big(\boldsymbol{G}\left(\mathbf{I}_N\otimes\boldsymbol{A}(\boldsymbol{\theta})\right)(\mathbf{I}_N\otimes\boldsymbol{C}_{\boldsymbol{s}})(\mathbf{I}_N\otimes\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}})\boldsymbol{G}^{\mathrm{H}} \\
&\quad + \sigma^2\boldsymbol{G}\underbrace{\mathbf{I}_{NM}}_{\mathbf{I}_N\otimes\mathbf{I}_M}\boldsymbol{G}^{\mathrm{H}}\Big) \\
&= \frac{\partial}{\partial\theta_j}\Big(\boldsymbol{G}\Big(\mathbf{I}_N\otimes\Big(\underbrace{\boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{C}_{\boldsymbol{s}}\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}}+\sigma^2\mathbf{I}_M}_{\boldsymbol{C}(\boldsymbol{\theta})}\Big)\Big)\boldsymbol{G}^{\mathrm{H}}\Big) \\
&= \frac{\partial}{\partial\theta_j}\Big(\boldsymbol{G}(\mathbf{I}_N\otimes\boldsymbol{C}(\boldsymbol{\theta}))\boldsymbol{G}^{\mathrm{H}}\Big),
\end{aligned}
$$

where $\boldsymbol{C}(\boldsymbol{\theta})$ is the covariance matrix of the fully sampled array. We detect that

solely $\boldsymbol{C}(\boldsymbol{\theta})$ depends on $\boldsymbol{\theta}$. Therefore,

$$\begin{aligned}\frac{\partial}{\partial\theta_j}\boldsymbol{C} &= \frac{\partial\boldsymbol{A}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{C_s}\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}} + \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{C_s}\left(\frac{\partial\boldsymbol{A}(\boldsymbol{\theta})}{\partial\theta_j}\right)^{\mathrm{H}} \\ &= 2\Re\left\{\frac{\partial\boldsymbol{A}(\boldsymbol{\theta})}{\partial\theta_j}\boldsymbol{C_s}\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}}\right\} \\ &= 2\Re\Big\{\Big[\mathbf{0}\cdots\mathbf{0}\boldsymbol{Da}(\theta_j)\mathbf{0}\cdots\mathbf{0}\Big]\boldsymbol{C_s}\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}}\Big\}.\end{aligned}$$

This leads to

$$\frac{\partial}{\partial\theta_j}\boldsymbol{C_x} = \Big(\boldsymbol{G}\Big(\mathbf{I}_N \otimes 2\Re\Big\{\Big[\mathbf{0}\,,\cdots,\mathbf{0}\,,\boldsymbol{Da}(\theta_j)\,,\mathbf{0}\,,\cdots,\mathbf{0}\Big]\boldsymbol{C_s}\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}}\Big\}\Big)\boldsymbol{G}^{\mathrm{H}}\Big). \tag{5.7}$$

Finally we obtain the $(\boldsymbol{\theta\theta})$ sub–matrix

$$\begin{aligned}[\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{\theta}}]_{j,k} = T\,\mathrm{tr}\Big\{\boldsymbol{C_x}^{-1}\Big(\boldsymbol{G}\Big(\mathbf{I}_N \otimes 2\Re\Big\{\Big[\mathbf{0}\,,\cdots,\mathbf{0}\,,\boldsymbol{Da}(\theta_j)\,,\mathbf{0}\,,\cdots,\mathbf{0}\Big]\boldsymbol{C_s}\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}}\Big\}\Big)\boldsymbol{G}^{\mathrm{H}}\Big) \\ \boldsymbol{C_x}^{-1}\Big(\boldsymbol{G}\Big(\mathbf{I}_N \otimes 2\Re\Big\{\Big[\mathbf{0}\,,\cdots,\mathbf{0}\,,\boldsymbol{Da}(\theta_k)\,,\mathbf{0}\,,\cdots,\mathbf{0}\Big]\boldsymbol{C_s}\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}}\Big\}\Big)\boldsymbol{G}^{\mathrm{H}}\Big)\Big\}.\end{aligned}$$

- $\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{c}} \in \mathbb{C}^{L\times L^2}$

$$[\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{c}}]_{j,k_{i,g}} = T\,\mathrm{tr}\left\{\boldsymbol{C_x}^{-1}\left(\frac{\partial}{\partial\theta_j}\boldsymbol{C_x}\right)\boldsymbol{C_x}^{-1}\left(\frac{\partial}{\partial[\boldsymbol{c}]_{k_{i,g}}}\boldsymbol{C_x}\right)\right\},$$

where we defined $[\boldsymbol{c}]_{k_{i,g}}$ as k-th entry of $\boldsymbol{c}$. The subindex i,g denotes the position of the entry $[\boldsymbol{c}]_k$ within the matrix $\boldsymbol{C_s}$. We section again the sub–matrix $(\boldsymbol{\theta},\boldsymbol{c})$ with respect to the defined parts of $\boldsymbol{c}$:

$$\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{c}} = \left[\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{c}_{\mathrm{off}_R}}\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{c}_{\mathrm{diag}}}\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{c}_{\mathrm{off}_I}}\right],$$

with

- $\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{c}_{\mathrm{off}_R}} \in \mathbb{C}^{L\times\frac{1}{2}(L^2-L)}$

$$[\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{c}_{\mathrm{off}_R}}]_{j,k_{i,g}} = T\,\mathrm{tr}\left\{\boldsymbol{C_x}^{-1}\left(\frac{\partial}{\partial\theta_j}\boldsymbol{C_x}\right)\boldsymbol{C_x}^{-1}\left(\frac{\partial}{\partial[\boldsymbol{c}_{\mathrm{off}_R}]_{k_{i,g}}}\boldsymbol{C_x}\right)\right\}.$$

The derivative of $\boldsymbol{C_x}$ with respect to $\boldsymbol{c}_{\mathrm{off}_R}$ reads as

$$\begin{aligned}\frac{\partial}{\partial[\boldsymbol{c}_{\mathrm{off}_R}]_{k_{i,g}}}\boldsymbol{C_x} &= \boldsymbol{G}\left(\mathbf{I}_N \otimes \left(\boldsymbol{A}(\boldsymbol{\theta})\left(\frac{\partial}{\partial[\boldsymbol{c}_{\mathrm{off}_R}]_{k_{i,g}}}\boldsymbol{C_s}\right)\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}}\right)\right)\boldsymbol{G}^{\mathrm{H}} \\ &= \boldsymbol{G}\left(\mathbf{I}_N \otimes \left(\boldsymbol{A}(\boldsymbol{\theta})\begin{bmatrix} 0 & \dots & 0 & 1 & 0 \\ \vdots & \ddots & \vdots & 0 & \vdots \\ 0 & \dots & \ddots & \vdots & \vdots \\ 1 & 0 & \dots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 \end{bmatrix}\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}}\right)\right)\boldsymbol{G}^{\mathrm{H}} \\ &= \boldsymbol{G}\left(\mathbf{I}_N \otimes \left(\boldsymbol{a}(\theta_i)\boldsymbol{a}(\theta_g)^{\mathrm{H}} + \boldsymbol{a}(\theta_g)\boldsymbol{a}(\theta_i)^{\mathrm{H}}\right)\right)\boldsymbol{G}^{\mathrm{H}},\end{aligned} \tag{5.8}$$

where the 1 entries are at the positions i, g and g, i, respectively.
Together we obtain

$$\begin{aligned}[\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{c}_{\mathrm{off}_R}}]_{j,k_{i,g}} = T\,\mathrm{tr}\Big\{&\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial\theta_j}\boldsymbol{C_x}\right)\boldsymbol{C}_{\boldsymbol{x}}^{-1} \\ &\left(\boldsymbol{G}^{\mathrm{H}}\left(\mathbf{I}_N \otimes \left(\boldsymbol{a}(\theta_i)\boldsymbol{a}(\theta_g)^{\mathrm{H}} + \boldsymbol{a}(\theta_g)\boldsymbol{a}(\theta_i)^{\mathrm{H}}\right)\right)\right)\boldsymbol{G}^{\mathrm{H}}\Big\}.\end{aligned}$$

- $\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{c}_{\mathrm{diag}}} \in \mathbb{C}^{L\times L}$

$$[\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{c}_{\mathrm{diag}}}]_{j,k_{i,i}} = T\,\mathrm{tr}\left\{\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial\theta_j}\boldsymbol{C_x}\right)\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial[\boldsymbol{c}_{\mathrm{diag}}]_{k_{i,i}}}\boldsymbol{C_x}\right)\right\},$$

with

$$\begin{aligned}\frac{\partial}{\partial[\boldsymbol{c}_{\mathrm{diag}}]_{k_{i,i}}}\boldsymbol{C_x} &= \boldsymbol{G}\left(\mathbf{I}_N \otimes \left(\boldsymbol{A}(\boldsymbol{\theta})\left(\frac{\partial}{\partial[\boldsymbol{c}_{\mathrm{diag}}]_{k_{i,i}}}\boldsymbol{C_s}\right)\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}}\right)\right)\boldsymbol{G}^{\mathrm{H}} \\ &= \boldsymbol{G}\Big(\mathbf{I}_N \otimes \left(\boldsymbol{a}(\theta_i)\boldsymbol{a}(\theta_i)^{\mathrm{H}}\right)\Big)\boldsymbol{G}^{\mathrm{H}}.\end{aligned} \tag{5.9}$$

- $\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{c}_{\mathrm{off}_I}} \in \mathbb{C}^{L\times\frac{1}{2}(L^2-L)}$

$$[\boldsymbol{F}_{\boldsymbol{\theta},\boldsymbol{c}_{\mathrm{off}_I}}]_{j,k_{i,g}} = T\,\mathrm{tr}\left\{\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial\theta_j}\boldsymbol{C_x}\right)\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial[\boldsymbol{c}_{\mathrm{off}_I}]_{k_{i,g}}}\boldsymbol{C_x}\right)\right\},$$

where

$$\begin{aligned}\frac{\partial}{\partial[\boldsymbol{c}_{\mathrm{off}_I}]_{k_i,g}}\boldsymbol{C}_{\boldsymbol{x}} &= \boldsymbol{G}\left(\mathbf{I}_N \otimes \left(\boldsymbol{A}(\boldsymbol{\theta})\left(\frac{\partial}{\partial[\boldsymbol{c}_{\mathrm{off}_I}]_{k_i,g}}\boldsymbol{C}_{\boldsymbol{s}}\right)\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}}\right)\right)\boldsymbol{G}^{\mathrm{H}} \\ &= \boldsymbol{G}\left(\mathbf{I}_N \otimes \left(\boldsymbol{A}(\boldsymbol{\theta})\begin{bmatrix} 0 & \dots & 0 & j & 0 \\ \vdots & \ddots & \vdots & 0 & \vdots \\ 0 & \dots & \ddots & \vdots & \vdots \\ -j & 0 & \dots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 \end{bmatrix}\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}}\right)\right)\boldsymbol{G}^{\mathrm{H}} \\ &= \boldsymbol{G}\Big(\mathbf{I}_N \otimes \Big(j\boldsymbol{a}(\theta_i)\boldsymbol{a}(\theta_g)^{\mathrm{H}} - j\boldsymbol{a}(\theta_g)\boldsymbol{a}(\theta_i)^{\mathrm{H}}\Big)\Big)\boldsymbol{G}^{\mathrm{H}}.\end{aligned} \tag{5.10}$$

- $\boldsymbol{F}_{\boldsymbol{\theta},\sigma} \in \mathbb{C}^{L\times 1}$

$$[\boldsymbol{F}_{\boldsymbol{\theta},\sigma^2}]_j = T\,\mathrm{tr}\left\{\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial\theta_j}\boldsymbol{C}_{\boldsymbol{x}}\right)\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial\sigma^2}\boldsymbol{C}_{\boldsymbol{x}}\right)\right\},$$

with

$$\begin{aligned}\frac{\partial}{\partial\sigma^2}\boldsymbol{C}_{\boldsymbol{x}} &= \frac{\partial}{\partial\sigma^2}\Big(\boldsymbol{G}\Big(\mathbf{I}_N \otimes \Big(\boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{C}_{\boldsymbol{s}}\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}} + \sigma^2\mathbf{I}_M\Big)\Big)\boldsymbol{G}^{\mathrm{H}}\Big) \\ &= \boldsymbol{G}\mathbf{I}_{NM}\boldsymbol{G}^{\mathrm{H}} = \boldsymbol{G}\boldsymbol{G}^{\mathrm{H}} = \mathbf{I}_{NK}.\end{aligned} \tag{5.11}$$

The $(\boldsymbol{\theta}, \sigma^2)$ sub–vector reads as

$$[\boldsymbol{F}_{\boldsymbol{\theta},\sigma^2}]_j = T\,\mathrm{tr}\left\{\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial\theta_j}\boldsymbol{C}_{\boldsymbol{x}}\right)\boldsymbol{C}_{\boldsymbol{x}}^{-1}\right\},$$

with the derivative $\frac{\partial}{\partial\theta_j}\boldsymbol{C}_{\boldsymbol{x}}$ in (5.7).

- $\boldsymbol{F}_{\boldsymbol{c},\boldsymbol{c}} \in \mathbb{C}^{L^2\times L^2}$

$$[\boldsymbol{F}_{\boldsymbol{c},\boldsymbol{c}}]_{j_{e,q},k_{i,g}} = T\,\mathrm{tr}\left\{\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial[\boldsymbol{c}]_{j_{e,q}}}\boldsymbol{C}_{\boldsymbol{x}}\right)\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial[\boldsymbol{c}]_{k_{i,g}}}\boldsymbol{C}_{\boldsymbol{x}}\right)\right\},$$

As previously, we section the sub–matrix $(\boldsymbol{cc})$ with respect to the defined parts of $\boldsymbol{c}$

$$\boldsymbol{F}_{\boldsymbol{c},\boldsymbol{c}} = \begin{bmatrix} \boldsymbol{F}_{\boldsymbol{c}_{\mathrm{off}_R},\boldsymbol{c}_{\mathrm{off}_R}} & \boldsymbol{F}_{\boldsymbol{c}_{\mathrm{off}_R},\boldsymbol{c}_{\mathrm{diag}}} & \boldsymbol{F}_{\boldsymbol{c}_{\mathrm{off}_R},\boldsymbol{c}_{\mathrm{off}_I}} \\ \boldsymbol{F}^{\mathrm{H}}_{\boldsymbol{c}_{\mathrm{off}_R},\boldsymbol{c}_{\mathrm{diag}}} & \boldsymbol{F}_{\boldsymbol{c}_{\mathrm{diag}},\boldsymbol{c}_{\mathrm{diag}}} & \boldsymbol{F}_{\boldsymbol{c}_{\mathrm{diag}},\boldsymbol{c}_{\mathrm{off}_I}} \\ \boldsymbol{F}^{\mathrm{H}}_{\boldsymbol{c}_{\mathrm{off}_R},\boldsymbol{c}_{\mathrm{off}_R}} & \boldsymbol{F}^{\mathrm{H}}_{\boldsymbol{c}_{\mathrm{diag}},\boldsymbol{c}_{\mathrm{off}_I}} & \boldsymbol{F}_{\boldsymbol{c}_{\mathrm{off}_I},\boldsymbol{c}_{\mathrm{off}_I}} \end{bmatrix},$$

where

- $\boldsymbol{F}_{\boldsymbol{c}_{\text{off}_R},\boldsymbol{c}_{\text{off}_R}} \in \mathbb{C}^{\frac{1}{2}(L^2-L)\times\frac{1}{2}(L^2-L)}$

  $$[\boldsymbol{F}_{\boldsymbol{c}_{\text{off}_R},\boldsymbol{c}_{\text{off}_R}}]_{j_{e,q},k_{i,g}} = T\,\text{tr}\left\{\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial[\boldsymbol{c}_{\text{off}_R}]_{j_{e,q}}}\boldsymbol{C}_{\boldsymbol{x}}\right)\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial[\boldsymbol{c}_{\text{off}_R}]_{k_{i,g}}}\boldsymbol{C}_{\boldsymbol{x}}\right)\right\}$$

  as a representative example. The remaining required derivatives are already introduced in (5.8),(5.9) and (5.10).

- $\boldsymbol{F}_{\boldsymbol{c},\sigma^2} \in \mathbb{C}^{L^2\times 1}$

  $$\begin{aligned}[\boldsymbol{F}_{\boldsymbol{c},\sigma^2}]_{j_{e,q}} &= T\,\text{tr}\left\{\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial[\boldsymbol{c}]_{j_{e,q}}}\boldsymbol{C}_{\boldsymbol{x}}\right)\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial\sigma^2}\boldsymbol{C}_{\boldsymbol{x}}\right)\right\}\\ &= T\,\text{tr}\left\{\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial[\boldsymbol{c}]_{j_{e,q}}}\boldsymbol{C}_{\boldsymbol{x}}\right)\boldsymbol{C}_{\boldsymbol{x}}^{-1}\right\},\end{aligned}$$

  along with (5.11). We decompose $\boldsymbol{F}_{\boldsymbol{c},\sigma^2}$ and get

  $$\boldsymbol{F}_{\boldsymbol{c},\sigma^2} = \begin{bmatrix}\boldsymbol{F}_{\boldsymbol{c}_{\text{off}_R},\sigma^2}\\ \boldsymbol{F}_{\boldsymbol{c}_{\text{diag}},\sigma^2}\\ \boldsymbol{F}_{\boldsymbol{c}_{\text{off}_I},\sigma^2}\end{bmatrix},$$

  with

  $$\begin{aligned}[\boldsymbol{F}_{\boldsymbol{c}_{\text{off}_R},\sigma^2}]_{j_{e,q}} &= T\,\text{tr}\left\{\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial[\boldsymbol{c}_{\text{off}_R}]_{j_{e,q}}}\boldsymbol{C}_{\boldsymbol{x}}\right)\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial\sigma^2}\boldsymbol{C}_{\boldsymbol{x}}\right)\right\}\\ &= T\,\text{tr}\left\{\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial[\boldsymbol{c}_{\text{off}_R}]_{j_{e,q}}}\boldsymbol{C}_{\boldsymbol{x}}\right)\boldsymbol{C}_{\boldsymbol{x}}^{-1}\right\}.\end{aligned}$$

  The other parts of $\boldsymbol{F}_{\boldsymbol{c},\sigma^2}$ have to be calculated analogously.

- $\boldsymbol{F}_{\sigma^2,\sigma^2} \in \mathbb{C}^{1\times 1}$

  $$\begin{aligned}[\boldsymbol{F}_{\sigma^2,\sigma^2}] &= T\,\text{tr}\left\{\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial\sigma^2}\boldsymbol{C}_{\boldsymbol{x}}\right)\boldsymbol{C}_{\boldsymbol{x}}^{-1}\left(\frac{\partial}{\partial\sigma^2}\boldsymbol{C}_{\boldsymbol{x}}\right)\right\}\\ &= T\,\text{tr}\left\{\boldsymbol{C}_{\boldsymbol{x}}^{-1}\boldsymbol{C}_{\boldsymbol{x}}^{-1}\right\}\\ &= T\,\text{tr}\left\{\boldsymbol{C}_{\boldsymbol{x}}^{-2}\right\}.\end{aligned}$$

After assembling the several sub–matrices, we are only interested in the $L \times L$ $(\boldsymbol{\theta},\boldsymbol{\theta})$–section of the Cramèr-Rao Bound $\boldsymbol{C}_{\hat{\boldsymbol{p}}}$, just as in the DML case.

Chapter 6

MUSIC (Multiple Signal Classification)

The MUSIC algorithm belongs in the field of spectral based methods. The eigenstructure of the full covariance matrix which gives further insights into the underlying estimation problem is included in this connection. One big advantage of the MUSIC algorithm compared to ML estimation is the reduction of complexity to a one dimensional optimization problem.

In our case of a sub–sampled antenna array with K antennas per configuration instead of a fully sampled antenna array, it is necessary to choose a slightly different approach than usual.

6.1 Preprocessing

We define the vector $\boldsymbol{z}_n(t) \in \mathbb{C}^M$ referred to [6] which contains K entries unequal zero. For one antenna configuration n and at one time snapshot t, the vector $\boldsymbol{z}_n(t)$ includes all entries of $\boldsymbol{x}_n(t) \in \mathbb{C}^K$:

$$\begin{aligned}\boldsymbol{z}_n(t) &= \boldsymbol{G}_n^{\mathrm{T}} \underbrace{\boldsymbol{G}_n \boldsymbol{x}_M(t)}_{\boldsymbol{x}_n(t)} \\ &= \boldsymbol{T}_n \boldsymbol{x}_M(t),\end{aligned}$$

with $\boldsymbol{x}_M(t) \in \mathbb{C}^M$ as the measured data of the full antenna array at the time snapshot t.

With this approach it is possible to generate – over all N possible configurations –, the sample covariance matrix $\hat{\boldsymbol{C}}_{\boldsymbol{z}} \in \mathbb{C}^{M \times M}$ of the full antenna array based on the new

defined vector $\boldsymbol{z}_n(t)$:

$$\hat{\boldsymbol{C}}_{\boldsymbol{z}} = \frac{1}{T}\sum_{n=1}^{N}\sum_{t=1}^{T}\boldsymbol{z}_n(t)\boldsymbol{z}_n(t)^{\mathrm{H}}.$$

It can be shown that

$$\mathrm{E}[\hat{\boldsymbol{C}}_{\boldsymbol{z}}] = \boldsymbol{C}_{\boldsymbol{z}} = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{C}_{\boldsymbol{s}}\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}} + \sigma^2\mathbf{I}_M + \boldsymbol{\Lambda}$$

applies. The only difference between the assembled covariance matrix $\boldsymbol{C}_{\boldsymbol{z}}$ and the true fully sampled covariance matrix $\boldsymbol{C} = A(\boldsymbol{\theta})\boldsymbol{C}_{\boldsymbol{s}}\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}} + \sigma^2\mathbf{I}_M$, consists in the diagonal matrix $\boldsymbol{\Lambda}$:

$$\boldsymbol{\Lambda} = (M-2)\,\mathrm{Diag}\left(\boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{C}_{\boldsymbol{s}}\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}} + \sigma^2\mathbf{I}_M\right),$$

with the operator $\mathrm{Diag}(\cdot)$ which returns the diagonal elements of the input matrix as a diagonal matrix. The summand $\boldsymbol{\Lambda}$ arises from the fact that in the case of $N = \binom{M}{K}$ possible antenna constellations, the diagonal elements of $\hat{\boldsymbol{C}}_{\boldsymbol{z}}$ were added $(M-1)$ times[1]. To fit the sample covariance matrix $\hat{\boldsymbol{C}}_{\boldsymbol{z}}$ to the desired covariance matrix $\boldsymbol{C}$ it is necessary to carry out a non–linear preprocessing procedure as described in [6]. Therefore, we just divide all diagonal entries of $\hat{\boldsymbol{C}}_{\boldsymbol{z}}$ by the factor $M-1$. This leads us to the adjusted sample covariance matrix:

$$\hat{\hat{\boldsymbol{C}}}_{\boldsymbol{z}} = \hat{\boldsymbol{C}}_{\boldsymbol{z}} \odot \begin{bmatrix} \frac{1}{M-1} & 1 & \cdots & 1 \\ 1 & \frac{1}{M-1} & & 1 \\ \vdots & & \ddots & \vdots \\ 1 & 1 & \cdots & \frac{1}{M-1} \end{bmatrix}, \tag{6.1}$$

where $\odot$ denotes the Schur-Hadamard product. By the use of this non–linear preprocessing it is ensured that $\mathrm{E}[\hat{\hat{\boldsymbol{C}}}_{\boldsymbol{z}}] = \boldsymbol{C}$ holds. Thereby, eigenvector based algorithms, e.g. the MUSIC algorithm can be applied.

6.2 ULA System

However till this point we did not consider the special case of a ULA system as it is existent in our case. For this particular constellation of antennas the covariance matrix exhibits the shape of a hermitian Toeplitz matrix. Accordingly, for a ULA system $\hat{\hat{\boldsymbol{C}}}_{\boldsymbol{z}}$, where we did not consider the Toeplitz structure, is not an adequate estimate to $\boldsymbol{C}$ and further preprocessing on $\hat{\hat{\boldsymbol{C}}}_{\boldsymbol{z}}$ is required.

[1] However, e.g for the case of fixing the first out of K antennas and let the remaining $K-1$ antennas switch their places, only the first entry of $\hat{\boldsymbol{C}}_{\boldsymbol{z}}$ is added $M-1$ times. To fill the gaps, generated by fixing the first antenna, we have to exploit the characteristic Toeplitz structure of the covariance matrix of a ULA.

As illustrated in Figure 6.1 we can construct the whole covariance matrix $\hat{\boldsymbol{C}}_{\text{ULA}}$ out of the vector $\hat{\boldsymbol{c}}^{\text{ULA}} = [\hat{c}_{11} \dots \hat{c}_{1M}]$ displayed in yellow. This means in the first step only the estimation of $\hat{\boldsymbol{c}}^{\text{ULA}}$ is required. Without more effort we develop $\hat{\boldsymbol{c}}^{\text{ULA}}$ out of the diagonal and off diagonal elements of $\hat{\hat{\boldsymbol{C}}}_{\boldsymbol{z}}$. Therefore we consider the main and off diagonals of $\hat{\hat{\boldsymbol{C}}}_{\boldsymbol{z}}$ as further samples for the estimation of $\hat{\boldsymbol{c}}^{\text{ULA}}$. By this approach we average over all diagonals separately to obtain a more precise estimation of the entries of $\hat{\boldsymbol{c}}^{\text{ULA}}$, i.e.,

$$\hat{c}_{11} = \frac{1}{M}\left(\hat{\hat{c}}^{\boldsymbol{z}}_{11} + \hat{\hat{c}}^{\boldsymbol{z}}_{22} + \ldots + \hat{\hat{c}}^{\boldsymbol{z}}_{MM}\right),$$

where we representatively calculated $\hat{c}_{11}$ with the diagonal elements of $\hat{\hat{\boldsymbol{C}}}_{\boldsymbol{z}}$. After obtaining all entries of $\hat{\boldsymbol{c}}^{\text{ULA}}$ we construct the upper triangle matrix corresponding to Figure 6.1. Finally we supplement the complex conjugate off diagonal entries and obtain the hermitian Toeplitz sample covariance matrix $\hat{\boldsymbol{C}}_{\text{ULA}}$.

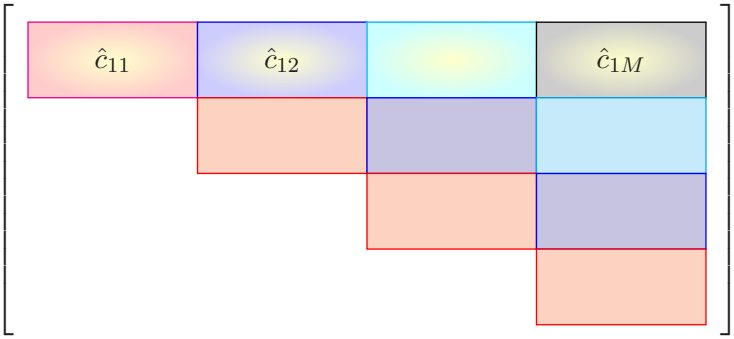

Figure 6.1: Characteristic toeplitz structure of the ULA covariance matrix $\hat{\boldsymbol{C}}_{\text{ULA}}$ in the case of a fully sampled array. The equally colored boxes indicate a identical matrix entry.

6.3 MUSIC Algorithm

The basic approach is to represent the covariance matrix of the fully sampled antenna array $\boldsymbol{C}$ via spectral decomposition:

$$\begin{aligned}\boldsymbol{C} &= \boldsymbol{A}\boldsymbol{C}_{\boldsymbol{s}}\boldsymbol{A}^{\text{H}} + \sigma^2\mathbf{I}_M \\ &= \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\text{H}},\end{aligned}$$

with the eigenvector matrix $\boldsymbol{U} \in \mathbb{C}^{M\times M}$ and the diagonal eigenvalue matrix $\boldsymbol{\Lambda} \in \mathbb{R}^{M\times M}$.

Since the inequality $M > L$ always has to be guaranteed for reasons of identifiability, the signal part $\left(\boldsymbol{A}\boldsymbol{C}_{\boldsymbol{s}}\boldsymbol{A}^{\text{H}}\right) \in \mathbb{C}^{M\times M}$ exhibits $\text{rank}(\boldsymbol{A}\boldsymbol{C}_{\boldsymbol{s}}\boldsymbol{A}^{\text{H}}) = L$. Furthermore,

we assert that the noise part has $\mathrm{rank}(\sigma^2 \mathbf{I}_M) = M$. With these observations it is now possible to specify the structure of the eigenvalue matrix $\boldsymbol{\Lambda}$:

$$\boldsymbol{\Lambda} = \begin{bmatrix} \sigma^2 + \lambda_1 & 0 & \dots & & \dots & 0 \\ 0 & \ddots & & & & \vdots \\ \vdots & & \sigma^2 + \lambda_L & \ddots & & \\ & & \ddots & \sigma^2 & & \vdots \\ \vdots & & & & \ddots & 0 \\ 0 & \dots & & \dots & 0 & \sigma^2 \end{bmatrix} \in \mathbb{C}^{M \times M}, \tag{6.2}$$

where $\lambda_1 \dots \lambda_L$ are the L non–zero eigenvalues of the signal part $\boldsymbol{A}\boldsymbol{C_s}\boldsymbol{A}^{\mathrm{H}}$ and σ^2 the M eigenvalues of the noise part $\sigma^2 \mathbf{I}_M$. Note that all eigenvectors with eigenvalue σ^2 are orthogonal to $\boldsymbol{A}$ and thus within the noise subspace. There are $M - L$ such eigenvector/eigenvalue pairs to find. Hence we structure the covariance matrix $\boldsymbol{C}$ into a noise and a signal part:

$$\boldsymbol{C} = \boldsymbol{U_s}\boldsymbol{\Lambda_s}\boldsymbol{U_s}^{\mathrm{H}} + \sigma^2 \boldsymbol{U_n}\boldsymbol{U_n}^{\mathrm{H}},$$

with the signal and noise eigenvector matrix $\boldsymbol{U_n}, \boldsymbol{U_s}$ and the diagonal signal eigenvalue matrix $\boldsymbol{\Lambda_s}$. The projection operator onto the signal and noise subspace are defined as:

$$\boldsymbol{\Pi} = \boldsymbol{U_s}\boldsymbol{U_s}^{\mathrm{H}}$$

$$\boldsymbol{\Pi}^{\perp} = \boldsymbol{U_n}\boldsymbol{U_n}^{\mathrm{H}}$$

However, in our case the true, fully sampled covariance matrix $\boldsymbol{C}$ is not available but instead only the adjusted sample covariance matrix $\hat{\boldsymbol{C}}_{\boldsymbol{z}}$. As we will see later, for the calculations of the MUSIC spectrum only the noise subspace is of importance. To obtain the estimate of the noise eigenvector matrix $\hat{\boldsymbol{U}}_{\boldsymbol{n}}$ we first spectral factorize $\hat{\boldsymbol{C}}_{\boldsymbol{z}}$:

$$\hat{\boldsymbol{C}}_{\boldsymbol{z}} = \hat{\boldsymbol{U}}\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{U}}^{\mathrm{H}}.$$

With the structure of $\boldsymbol{\Lambda}$ from (6.2) in mind, we conclude that the $M - L$ smallest eigenvalues of $\hat{\boldsymbol{\Lambda}}$ belong to the noise part which means that their corresponding eigenvectors $\hat{\boldsymbol{U}}_{\boldsymbol{n}}$ span the noise subspace.

Finally, we can state the MUSIC spectrum as defined in [7]:

$$P_M(\theta) = \frac{\boldsymbol{a}(\theta)^{\mathrm{H}}\boldsymbol{a}(\theta)}{\boldsymbol{a}(\theta)^{\mathrm{H}}\hat{\boldsymbol{\Pi}}^{\perp}\boldsymbol{a}(\theta)},$$

with $\hat{\boldsymbol{\Pi}}^{\perp} = \hat{\boldsymbol{U}}_{\boldsymbol{n}}\hat{\boldsymbol{U}}_{\boldsymbol{n}}^{\mathrm{H}}$. The peaks of the spectrum determine the angles of the inclined signals.

6.4 MUSIC Specific Identifiability

For a sub–sampled antenna array with M antennas and K active antennas for each configuration n, we are only able to estimate $K - 1$ sources. However, [6] proofs that combining all results from the subarrays, as it was done with $\hat{\hat{\boldsymbol{C}}}_{\boldsymbol{z}}$ from (6.1), it is possible to estimate up to $M - 1$ angles, i.e.,

$$L \leq M - 1. \tag{6.3}$$

Therefore, all L signals can be resolved even in the case $L > K$ as long as (6.3) holds. This result has to be highlighted explicitly, since we come to the same result as if the full array of M antennas would have been used.

Chapter 7

Simulation Results

In the following, we see simulations to compare the considered methods and algorithms. Therefore, we ran several simulations in an SNR range from $-10\,\mathrm{dB}$ to $40\,\mathrm{dB}$ with an ULA of 9 antennas which are placed in an equidistant spacing of $\frac{\lambda}{2}$. For the sake of simplicity, the noise is created according to a scaled identity covariance matrices and the SNR was adjusted only by this scaling factor. In case of multiple impinging wavefronts ($L > 1$), we restrict them to appear at least in a minimum angle distance of 0.1 in the cosine to each other. The reason for this decision is that wavefronts that appear in a small angle distance can hardly be detected if they send at the same frequency without any special signal characteristics which leads to large errors. For smaller angle distances, the CRB we described in Chapter 5 is no longer valid. Instead, you should use the periodic CRB which is described in [8]. To get an appropriate result by the use of Monte Carlo iterations, the results should be preferably free of deviations caused by noise, signal and wavefront[1] angles realizations. Therefore, we used several random noise realizations per signal, and again several signals per angle constellation. To reduce the impact of different angle constellations in the multiple wavefront case, we used 100 Monte Carlo iterations over the angle constellations, where the cosine values of the angles are distributed uniformly in the range $[-1, 1]$.

In the first part of this chapter, we want to consider the case of fewer impinging wavefronts than available RF chains, i.e., $L < K$. In the second part, we regard the case of equal or more impinging wavefronts than available RF chains, i.e., $L \geq K$.

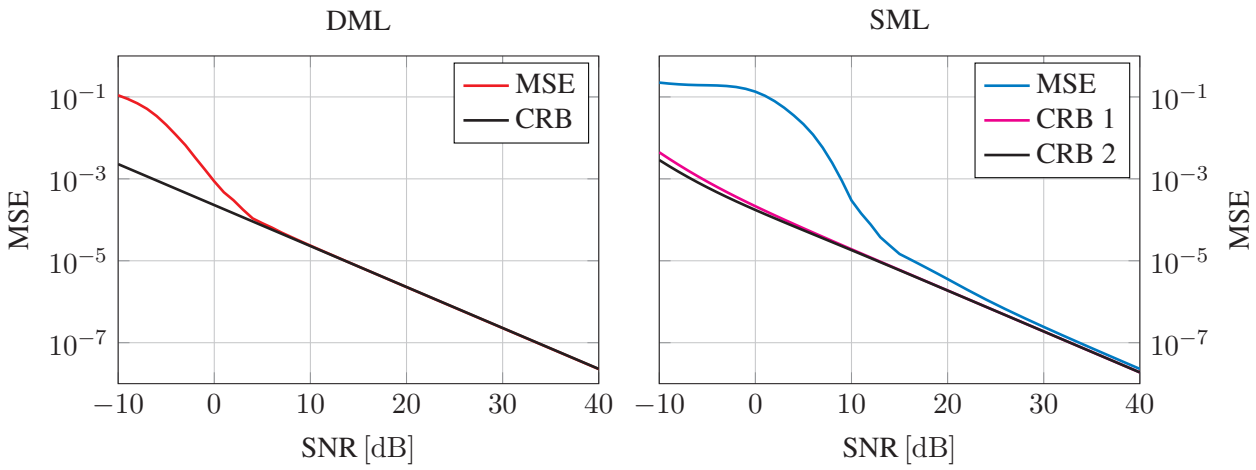

Figure 7.1: CRB and MSE for DML (left) and SML (right) of a full antenna array based on T = 7 snapshots with two random wavefront directions. CRB 1 of SML without knowledge about distributions, CRB 2 of SML with the knowledge of $\boldsymbol{C_s}$ and σ^2

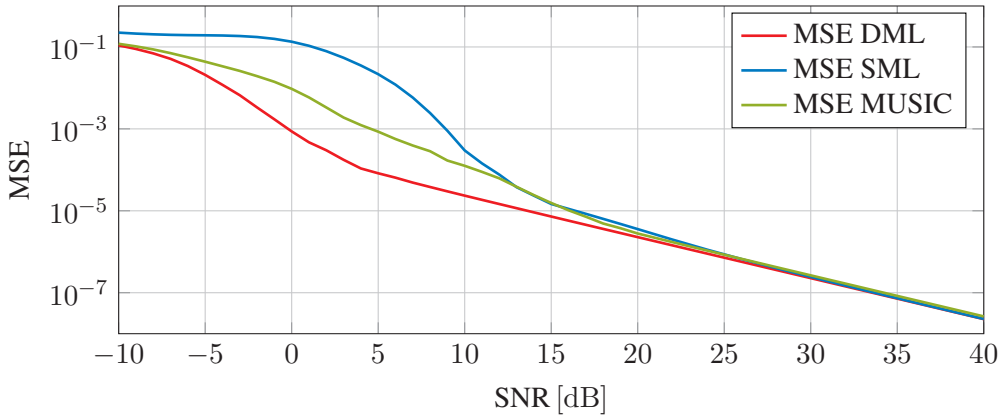

Figure 7.2: MSE of DML, SML and MUSIC of a full antenna array based on T = 7 snapshots with two random wavefront directions

7.1 Fewer number of wavefronts than used RF chains ($L < K$)

In Figure 7.1, we observe the MSE and CRB curves of the DML and SML methods over the SNR for a full antenna array[2]. Each estimation is based on seven snapshots and there are two impinging wavefronts out of randomly chosen directions. In case of the SML estimation, we assumed that the transmit signal and noise distributions are always known to the estimator, whereas we plotted two CRBs – CRB 1 without the knowledge of transmit signal or noise distribution, and CRB 2 with the same knowlege as we give the SML. In this scenario, the SML CRB curves are especially in the higher SNR region very similar. Notice that the MSEs are reaching their CRBs in both cases - DML and SML, whereby the DML reaches the bound much earlier at an SNR of $4\,\mathrm{dB}$ whereas the SML is reaching it much slower and reaches it at about $30\,\mathrm{dB}$.

In Figure 7.2, we see the MSE results of the same simulation setup together with the MSE curve of the MUSIC algorithm. The DML outperforms MUSIC and SML in the lower SNR region from $-10\,\mathrm{dB}$ to $20\,\mathrm{dB}$, where the SML performs worst. The performance of the MUSIC is between that of DML and SML over the whole SNR range.

Figure 7.3 depicts the MSE curves of DML, SML and MUSIC of an estimation of two random directions. As a lower bound we added the CRB of the DML in light gray in the background. The three plots on the left hand side are based on a full antenna array and the three plots on the right hand side are based on a set of subsampled antenna arrays with three RF chains, whereby the first and second antenna are sampled at all timeslots and the the third RF chain switches between the other seven antennas. From top to bottom, the MSE curves are plotted against a changing number of snapshots. As it would be hard to obtain a fair comparison between the two approaches, if we also considered that the switching between the RF chains takes some time, we only compare the approaches according to the total number of snapshots. This means that we use seven times more snapshots for one estimation in the full array case than for one estimation in the subsampled array case.

In the vertical comparison in terms of the different number of total snapshots, we can see that especially the MSE curves are not only scaled according to the number snapshots but they also reach the lower bound (CRB) at a lower SNR. The reason for this is that the snapshots effectively scale the SNR. In the case of a full antenna array, the MSE curves of the MUSIC algorithm estimations are lying always between the DML and the SML and get closer to the DML estimations with an increasing number of snapshots. When we now consider the $holdFirst2$ subarray case, it catches the eye

[1]For some special cases we observe only fixed angles such that there are no Monte Carlo Iterations over the wavefront angles.

[2]A full antenna array is an array which has one RF chain for all available antennas,i.e. 9 RF chains for 9 antennas in our case. So, there is no subsampling in these arrays.

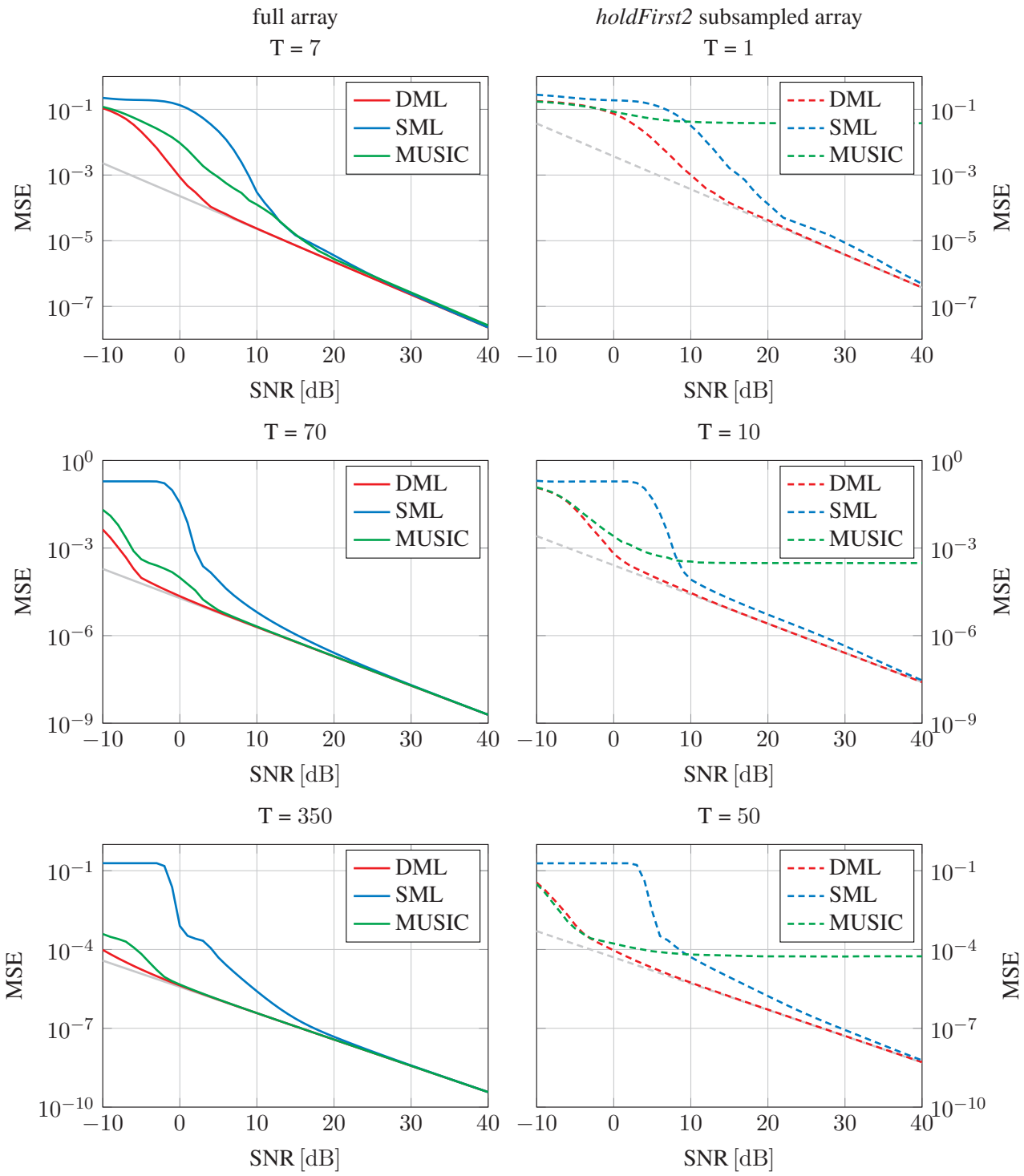

Figure 7.3: MSEs of DML, SML and MUSIC of a full antenna array (left) and *holdFirst2* subsampled array (right) based on T snapshots per (sub-) array with two random wavefront directions

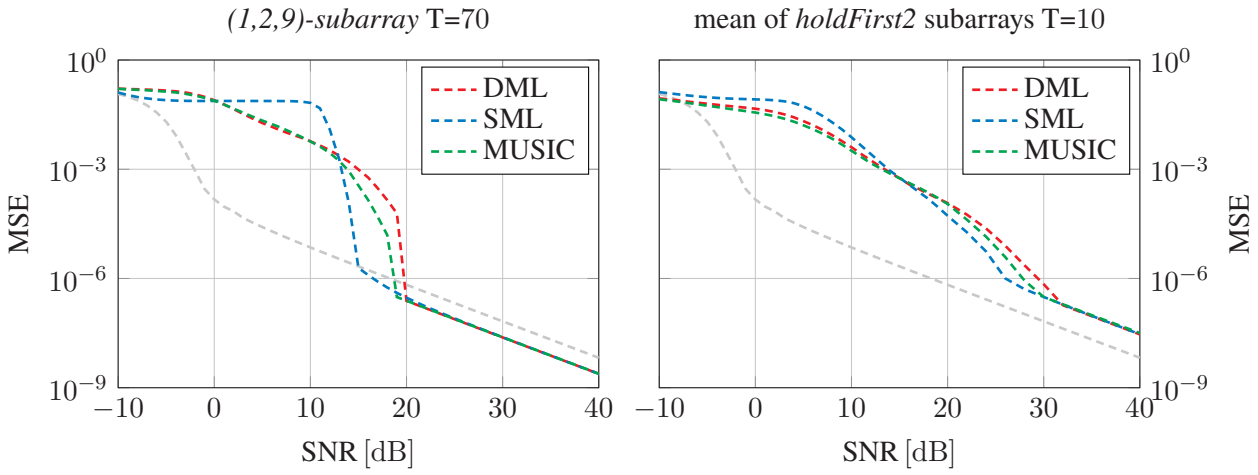

Figure 7.4: Subarray-based approaches to overcome the saturation of the MUSIC algorithm with the reference MSE of the DML full array with T = 70 snapshots and fixed impinging wavefront angles of $[0, 0.5]$ in the cosine

that the MUSIC algorithm curves saturate for all different numbers of snapshots. At the first glance, this is maybe a bit confusing but it gets clear when we take a closer look onto the way how the sampled covariance matrix $\hat{\boldsymbol{C}}_{\boldsymbol{z}}$ is built. In Section 6.1, we have shown that the sampled sub-covariance matrices are assembled to a full-sized matrix. However, these sub-covariance matrices are based on different signal realizations. Thus, all the sub-covarianve matrices show different deviations from the real covariance matrix. This leads to the saturation regarding the SNR by using such sub-matrices fo a finite number of snapshots.

Two alternative ways to use subsampled arrays, in which the MUSIC algorithm does not lead to a saturation in the case of $L < K$, are shown in Figure 7.4. In the first case, we only use the first, the second and the last antenna of the array and do not switch between the antennas. The first and second antennas are chosen to exhibit the $\frac{\lambda}{2}$ distance within each subarray such that we do not lose the ability to localize all wavefronts from -1 to 1 with this single subarray. Another reason for this choice is we can be sure that the steering matrix $\tilde{\boldsymbol{A}}$ has two independent vectors for all possible impinging wavefronts. So, the probability that the subsampled steering matrix has rank 3 is higher[3]. The last one is chosen to get the maximum distance in the setup, which offers the highest possible array sensitivity. We call this array simply *(1,2,9)-subarray*. In the second case we use the same subarrays as in the standard *holdFirst2* case but we do not stick them together in one estimation but calculate one estimation per subarray.

[3]We need a full-ranked steering matrix to be able to find the Least-Squares solution in (4.5), where we have to invert $(\tilde{\boldsymbol{A}}^{\mathrm{H}}\tilde{\boldsymbol{A}})$

The final estimation is then obtained by averaging over the single estimations. In both cases we only use three RF chains and the total number of T = 70 snapshots. Therefor, we use the DML of *holdFirst2* with T = 10 as reference.

You can easily see that using the *(1,2,9)-subarray* leads to a better estimation result in the SNR region from $20\,\mathrm{dB}$ on for DML, SML and MUSIC compared to the reference *holdFirst2* DML MSE curve. This is caused by the high sensitivity gained using the biggest distance in the array at all snapshots. As a drawback, the estimates get a lot worse in the lower SNR region. When we now consider the approach of taking the mean of single subarrays in the right plot, we come to a completely different result. The DML and SML methods can no longer profit from getting all information in one estimation. So they perform worse than in the case of using all *holdFirst2* subarrays for one estimation. Only the MUSIC algorithm benefits from this in the higher SNR region. When we compare the two approaches, we see that the *(1,2,9)-subarray* approach performs clearly better than the approach of taking the mean of single subarray estimations at high SNR. But notice that there is no longer the possibility to use MUSIC and SML also in the case of $L \geq K$ using these approaches, as we perform one estimations for each subarray so that the estimators get only the signals of exactly K antennas.

7.2 Equal or higher number of wavefronts than used RF chains ($L \geq K$)

Next, we want to consider the case of equal or fewer RF chains than impinging wavefronts. In Figure 7.5, we can see three spectrums of the MUSIC algorithm caused by five wavefronts at an SNR of $10\,\mathrm{dB}$ based on one realization. The total number of snapshots is similar whereas the number of used RF chains is different. In case of the *holdFirst* subarrays, there are only two RF chains whereby one RF chain always samples the first antenna and the other RF chain is switched between the remaining antennas. As one might expect, the full array of nine antennas can determine the directions very good and the spectrum shows highly defined peaks. Although that there are only three RF chains in the *holdFirst2* subarrarray case we observe that the spectrum also shows five peaks which are not that well defined any more. But they are still located close to the actual directions. In case of the *holdFirst* subarrays we observe a more tremendous degradation in the sharpness. But even more evident is the fact that the wavefront impinging from direction -0.7 can not be detected anymore.

We have already discussed that the usage of subsampled arrays leads to a saturation in combination with the MUSIC algorithm. Now, we want to consider what happens when we use a different preprocessing method for the subsampled covariance matrix. Therefor, we regard a simulation with two RF chains and three wavefront directions impinging from the fixed directions $[-0.3, 0.1, 0.4]$, which is plotted in Figure 7.6. On the left side, we consider the usage of *holdFirst* subarrays and on the right side the

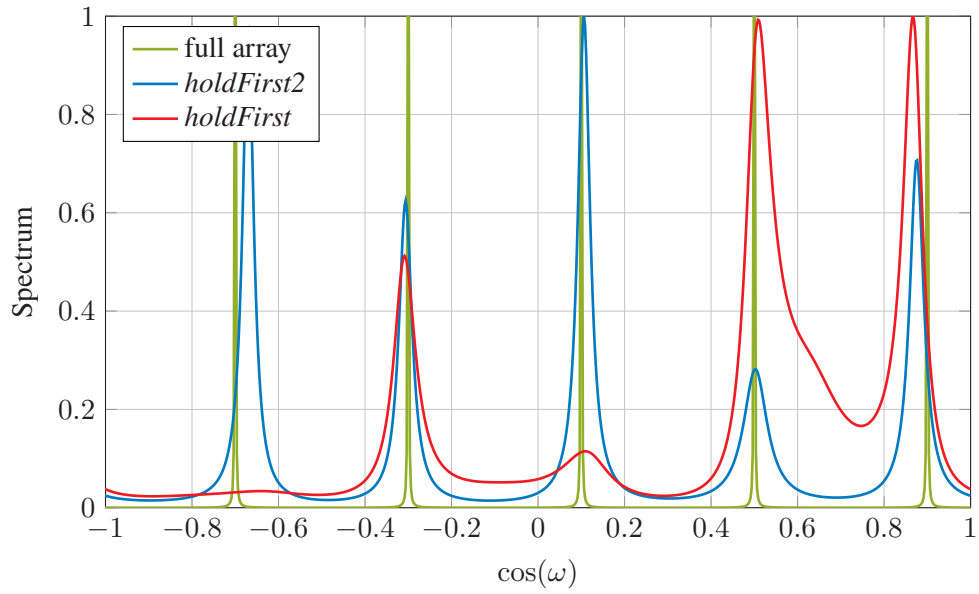

Figure 7.5: Spectra of different antenna array constellations with similar number of totaled snapshots (70) and five fixed impinging wavefront angles of $[-0.7, -0.3, 0.1, 0.5, 0.9]$ in the cosine at an SNR of $10\,\mathrm{dB}$

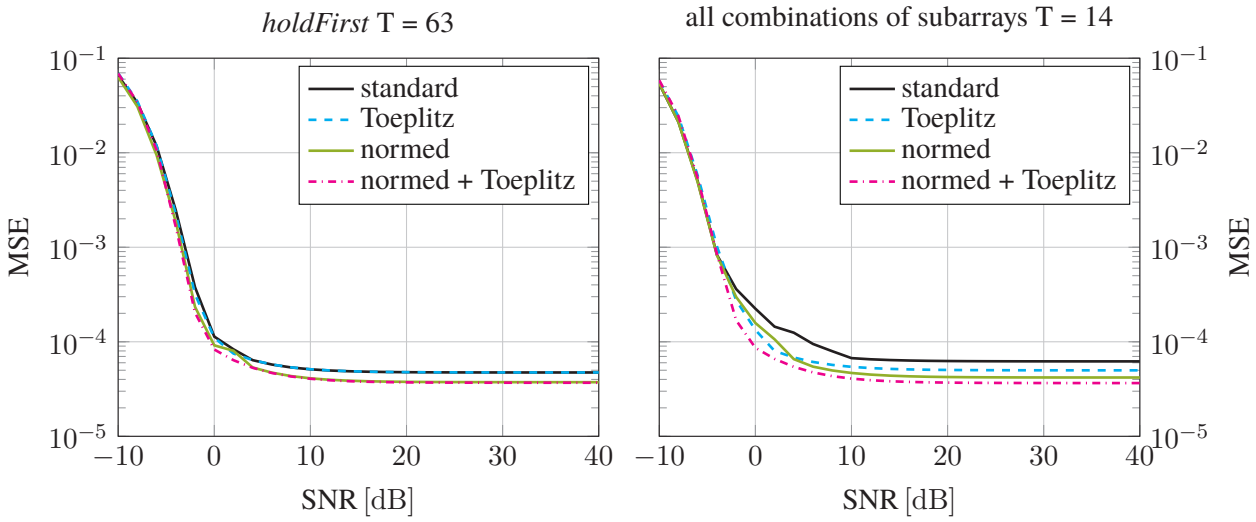

Figure 7.6: MSE of MUSIC for different covariance matrices built of the same received signal using two RF chains and fixed impinging wavefront angles of $[-0.3, 0.1, 0.4]$ in the cosine

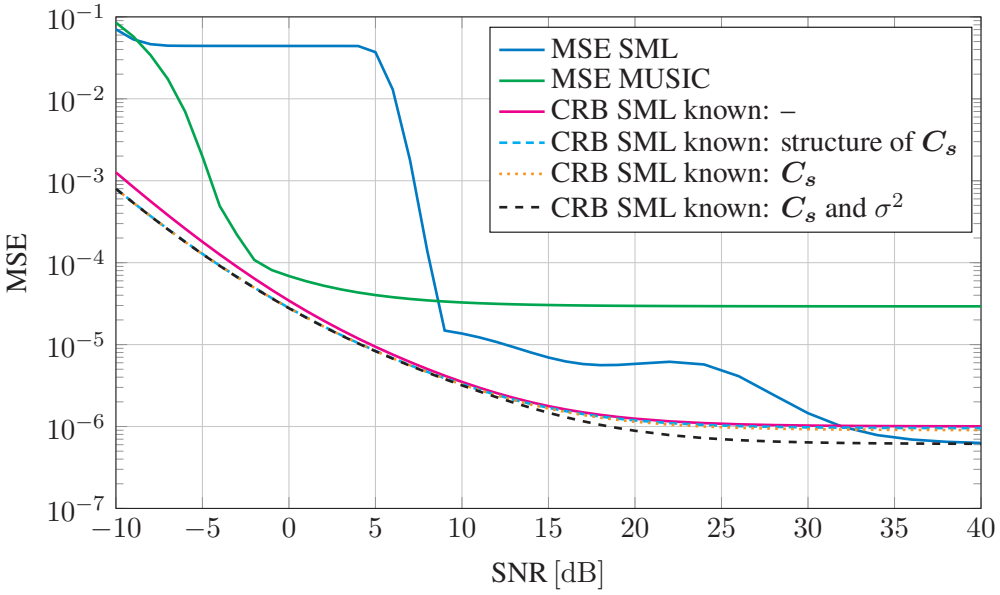

Figure 7.7: MSE of SML and MUSIC using two RF chains with *holdFirst* subarrays and four SML CRBs with different assumptions, all is simulated with two fixed impinging wavefront angles of $[-0.3, 0]$ in the cosine and T = 50 snapshots per subarray

use of all combinations of subarrays. In case of standard preprocessing as described in Section 6.1, we only used the Toeplitz structure of the covariance matrix to fill the gaps, generated by fixing the first antenna (*holdFirst*). Since these gaps are only located in the off-diagonals of the covariance matrix, solely the main-diagonal is not included in the Toeplitz structure. As a consequence, the Toeplitz case introduced in Section 6.2 for *holdFirst* does not have a big effect, as only the diagonal has to be adjusted. Different to this, the ULA specific preprocessing including the Toeplitz structure for all combinations of subarrays, which was already described in Section 6.2, leads to a better performance such that the MSE curve saturates at a slightly lower MSE level. Next, we want to consider the case when we preprocess the single subsampled covariance matrices before we assemble them to the full covariance matrix. Here, we normalize the sub covariance matrices by division of all elements by the sum of its diagonal elements. By observing the plots, we come to the conclusion that the normalization leads to a benefit regarding the saturation, whereas there is no difference for low SNR. The combination of Toeplitz structure and normalization thus leads to the best result especially in case of using all combinations of subarrays.

Next, we consider the case of two RF chains and an equal number of wavefronts ($K = L$), which is plotted in Figure 7.7. The simulation is based on *holdFirst* subarrays each with 50 snapshots and two fixed wavefront directions of $[-0.3, 0]$. Since we

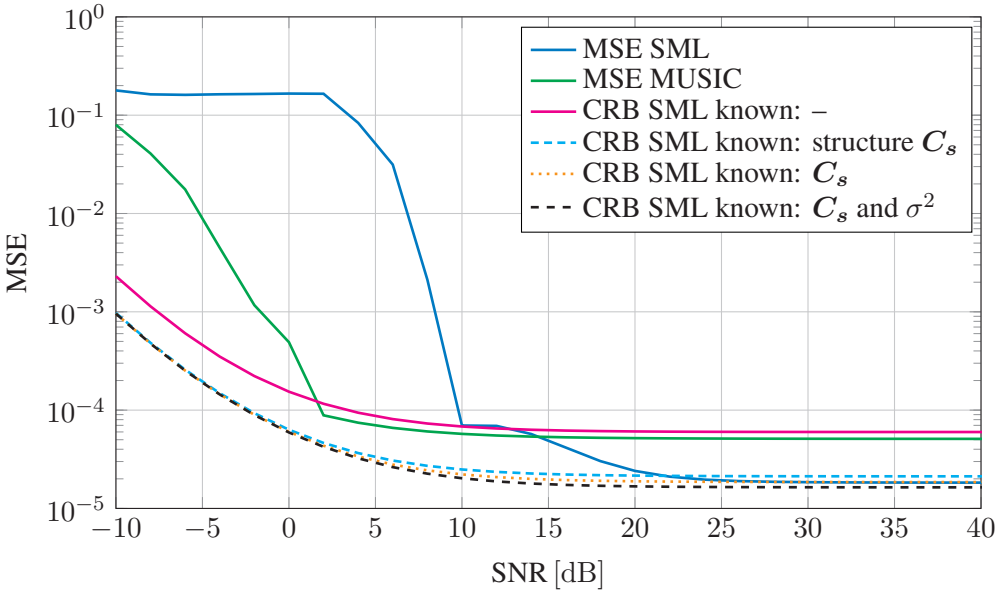

Figure 7.8: MSE of SML and MUSIC using two RF chains with *holdFirst* subarrays and four SML CRBs with different assumptions, all is simulated with three fixed impinging wavefront angles of $[-0.3, 0.1, 0.4]$ in the cosine and T = 50 snapshots per subarray

use subarrays, the MUSIC algorithm leads to a saturation of the MSE curve which was already described before. At this point we highlight that also the MSE and CRB curves of the SML saturate, now. The reason for the saturation of the SML method in the case $L \geq K$ is explained by Sheinvald et al. in [9]. Also remarkable is the wavy characteristic of the SML MSE curve between 15 and $30\,\mathrm{dB}$. Unfortunately, we did not found a profound reason for this. Compared to the full antenna array case in Figure 7.1, we notice that the variation of the CRB of the SML is bigger, here. The controlling factor of the spread in the lower SNR region is basically down to the knowledge of the statistical characteristics of the signal. Due to the bad performance of the SML in the lower SNR region, this seems to be not important for the considered number of snapshots, as the SML performs bad in the lower SNR region, anyway. However, the spread in the higher SNR region is mainly based on the knowledge of the variance of the noise. Therefor, it becomes important which statistical knowledge is given to the SML estimator. We can see that especially the knowledge of the noise variance becomes valuable, here. When we compare the MSE curves of MUSIC and SML, we recognize that the MUSIC algorithm performs clearly better in the lower SNR region until about $8\,\mathrm{dB}$. For a higher SNR the SML outperforms the MUSIC method, as the MUSIC algorithm has already run into saturation.

For more wavefronts than RF chains ($L > K$), we observe the saturation of SML and MUSIC MSE curves again in Figure 7.8 with the same number of 50 snapshots per subarray. In difference to the two wavefronts case before, we see that the spread between the CRBs is more pronounced, here. Additionally, the main influence regarding the width of the spread is no longer determined by the knowledge of the noise variance but the knowledge of the right structure of the covariance matrix of the received signal. Of course, we must not forget that the assumption to have a diagonal structure only improves the estimation when the underlying signals are really uncorrelated. In comparison to the previous subsampled MUSIC simulations with 50 snapshots, we recognize that the algorithm performs almost identically for two and three wavefronts. Even for the *holdFirst2* subarrays containing three RF chains in Figure 7.3 we see a saturation in the same MSE magnitude order.

Chapter 8

Implementation via Software Defined Radio

Software Defined Radio provides a powerful framework for developing radio applications. The main advantage of this concept is that most of the development can be done in software. This offers a flexibility which can be used to evaluate the performance of DoA estimation methods in real-world scenarios. In this chapter we describe the used hardware and the measurement setup. Additionally, the implementation of the MUSIC method is shown and the measurement results are presented.

8.1 Hardware

The measurements are done with Universal Software Radio Peripherals (USRPs). The used NI USRP RIO 2940 consists of two full duplex RF channels for transmitting and receiving. The internal structure of one channel is shown in Figure 8.1 which is taken from [10]. Each channel has two antenna ports from which one is for transmission and receiving operations (TX1/RX1) and one only for receiving (RX2). Looking at the demodulation process, the received signal passes several filter and amplification stages and gets then demodulated in up to two mixing steps. For this purpose the signal is mixed with the signal from a reference oscillator and a version shifted by 90 degrees to obtain the in-phase and quadrature component of the corresponding complex baseband signal. After applying an anti-aliasing lowpass filter the signal is sampled at a rate of $200\,\mathrm{MHz}$. Depending on the settings made by the software, the digital signal is down sampled in a next step. Note that the USRP can work not only as a homodyne receiver but also as an IF receiver. The integrated FPGA of the device also implements a numerical controlled oscillator (NCO) so that an intermediate frequency can be

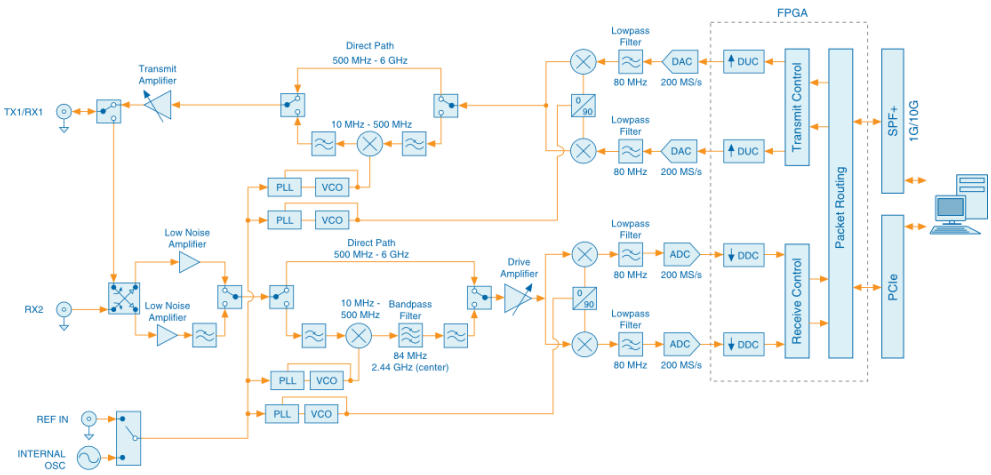

Figure 8.1: Block diagram of the USRP RIO 2944

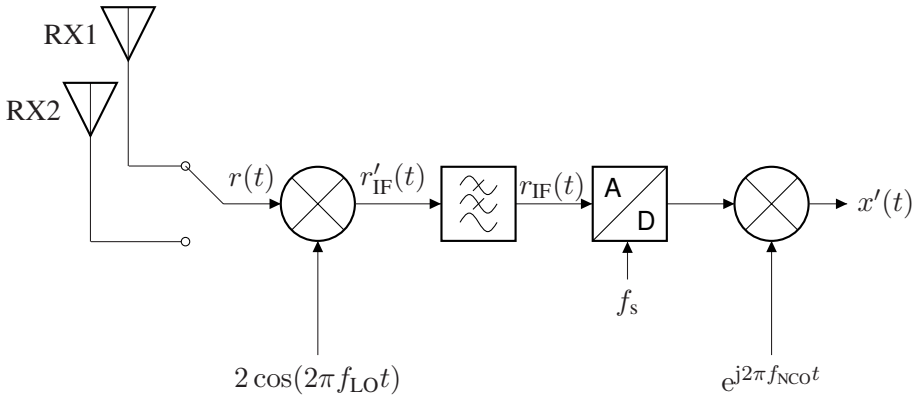

Figure 8.2: Reduced block diagram of a receive channel

chosen.

The transmit chain is implemented in the same way, only the signal travels in the reverse direction.

8.2 IF Receiver

As a detailed knowledge about the hardware structure is not necessary for most implementation tasks, we reduce the receiver model in Figure 8.1 by focusing only on the main receiving steps. A simplified receiver model is given in Figure 8.2. Note that this model does not exactly resemble the block diagram in Figure 8.1 as the hardware device divides the incoming signal in the in-phase and quadrature component before the analog/digital conversion and not afterwards. Nevertheless, this model is sufficient

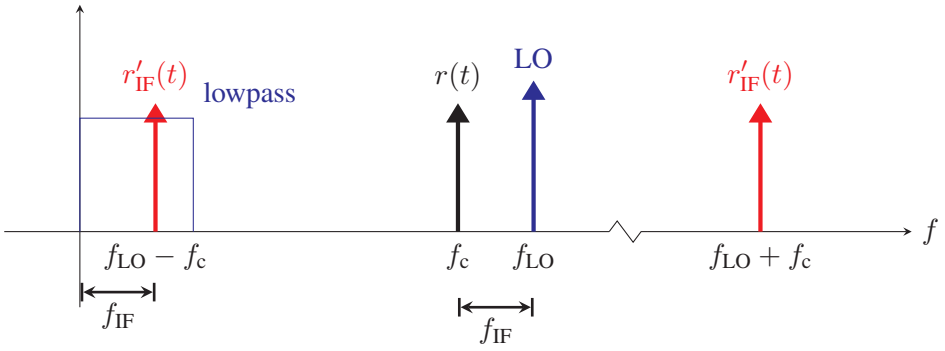

Figure 8.3: Power spectral density of the signal components in the receiving process

for programming the device as all tunable parameters are included. The oscillator frequencies f_{LO} and f_{NCO} as well as the sampling rate f_{s} can be set by programming.

To illustrate the demodulation process, consider a signal $r(t) = A\cos(2\pi f_{\mathrm{c}} t + \varphi)$ at the carrier frequency f_{c} with amplitude A. After the first mixing stage the received signal is given by

$$\begin{aligned} r'_{\mathrm{IF}}(t) &= r(t) \cdot 2\cos(2\pi f_{\mathrm{LO}} t) \\ &= A\cos(2\pi(f_{\mathrm{LO}} - f_{\mathrm{c}})t - \varphi) + A\cos(2\pi(f_{\mathrm{LO}} + f_{\mathrm{c}})t + \varphi). \end{aligned} \tag{8.1}$$

To remove the spectral component at the sum frequency $f_{\mathrm{LO}} + f_{\mathrm{c}}$, the lowpass filter is applied. This results in the IF signal

$$r_{\mathrm{IF}}(t) = A\cos(2\pi f_{\mathrm{IF}} t - \varphi). \tag{8.2}$$

where we have introduced the intermediate frequency $f_{\mathrm{IF}} = f_{\mathrm{LO}} - f_{\mathrm{c}}$. The situation is depicted in Figure 8.3 for the case $f_{\mathrm{LO}} > f_{\mathrm{c}}$.

The second, digital mixing stage yields

$$x'(t) = r_{\mathrm{IF}}(t)\mathrm{e}^{\mathrm{j}2\pi f_{\mathrm{NCO}} t} \tag{8.3}$$

$$\begin{aligned} &= A\cos(2\pi f_{\mathrm{IF}} t - \varphi)\mathrm{e}^{\mathrm{j}2\pi f_{\mathrm{NCO}} t} \\ &= \frac{A}{2}\mathrm{e}^{\mathrm{j}(2\pi(f_{\mathrm{IF}} + f_{\mathrm{NCO}})t - \varphi)} + \frac{A}{2}\mathrm{e}^{-\mathrm{j}(2\pi(f_{\mathrm{IF}} - f_{\mathrm{NCO}})t - \varphi)} \end{aligned} \tag{8.4}$$

where we have used the identity $\cos(\alpha) = \frac{\mathrm{e}^{\mathrm{j}\alpha} + \mathrm{e}^{-\mathrm{j}\alpha}}{2}$. Next, we choose $f_{\mathrm{NCO}} = -f_{\mathrm{IF}}$ which results in

$$x'(t) = \frac{A}{2}\mathrm{e}^{-\mathrm{j}\varphi} + \frac{A}{2}\mathrm{e}^{-\mathrm{j}(4\pi f_{\mathrm{IF}} t - \varphi)}. \tag{8.5}$$

The equivalent complex baseband signal is then obtained after applying an additional, digital lowpass filter which removes the higher order demodulation term. The resulting signal is given by

$$x(t) = \frac{A}{2}\mathrm{e}^{-\mathrm{j}\varphi}. \tag{8.6}$$

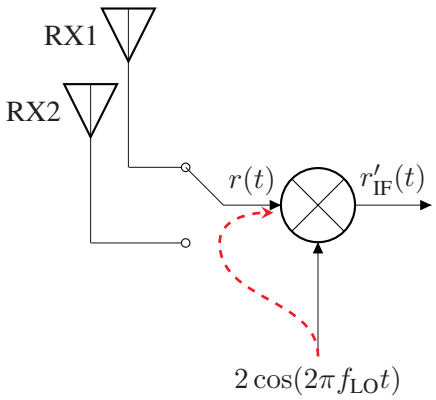

Figure 8.4: LO feedthrough

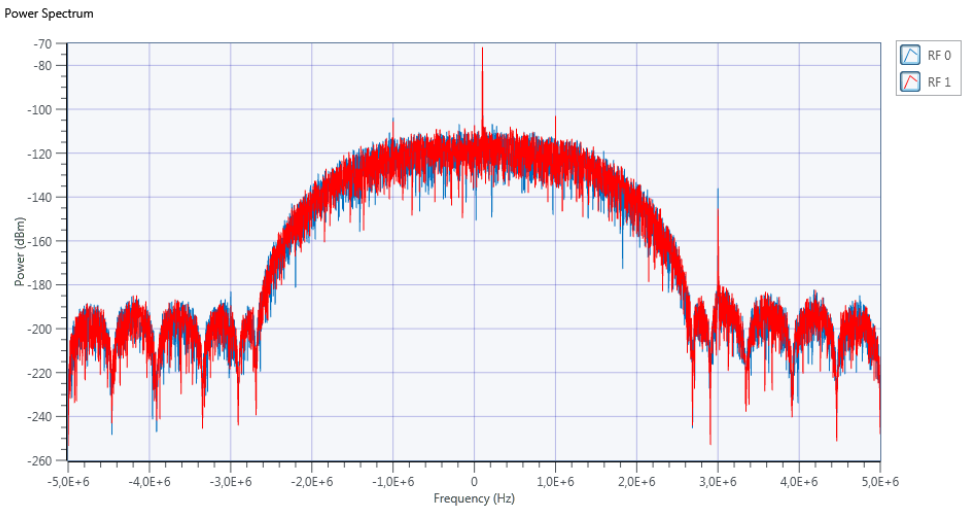

Figure 8.5: Filtered LO leakage at $f_{\mathrm{IF}} = 3\,\mathrm{MHz}$ in spectral domain

	RF0		RF1	
	RX1	RX2	RX1	RX2
Configuration 1	•		•	
Configuration 2	•			•
Configuration 3		•	•	
Configuration 4		•		•

Table 8.1: Different antenna configurations for the measurements

At this point, one physical effect has to be considered. Due to imperfect electromagnetic shielding, the LO signal leaks into the receive path of the RF front-end as illustrated in Figure 8.4 and gets mixed down itself [11]. There, we obtain a spectral component at f_{IF} from the baseband signal. This spectral component can be removed by filtering the baseband signal with a FIR filter. In our measurements we used a blackman-harris window filter[1] of order ten. due to its strong decay in time domain and frequency domain as well. An example of a filtered signal is shown in Figure 8.5. It can be observed that the LO feedthrough which is $3\,\mathrm{MHz}$ above the carrier frequency is attenuated to a level that is in the order of the noise.

8.3 Measurement Setup and Programming

The measurements are done at two different carrier frequencies that are freely usable for short range devices. Firstly, we choose $f_{\mathrm{c}} = 633.75\,\mathrm{MHz}$ as a carrier frequency which belongs to the so called PMSE band and is originally intended for wireless microphones [13]. Secondly, the measurements are done in the LTE duplex gap ($1.785\,\mathrm{GHz} - 1.805\,\mathrm{GHz}$) at a carrier frequency $f_{\mathrm{c}} = 1.795\,\mathrm{GHz}$. This frequency band is freely usable as well [14].

The measurement setup consists of two USRP. One USRP is equipped with on antenna for the transmission of the signal to be detected. The antenna is connected to the TX1 port of the RF chain RF0. In order to localize the transmitter, the receiver USRP is equipped with an ULA comprising four antennas. As the used USRP has only two receiver chains switching of the antennas is necessary to utilize the complete aperture of the array.

With the used hardware one can switch between the antennas connected to the receiver chain RF0 and the antennas connected to RF1. Thus, four different antenna constellations are possible which are summarized in Table 8.1. To resolve the maximum range of angles the distance between all pairs of antennas that are measured at the

[1] According to [12], the structure of the discrete time blackman-harris window is given by $w(n) = a_0 - a_1 \cos\left(\frac{2\pi n}{\mathbb{N}-1}\right) + a_2 \cos\left(\frac{4\pi n}{\mathbb{N}-1}\right) - a_3 \cos\left(\frac{6\pi n}{\mathbb{N}-1}\right)$.

	Frequency	Wavelength	antenna separation d
PMSE band	$633.75\,\mathrm{MHz}$	$47.3\,\mathrm{cm}$	$7.9\,\mathrm{cm}$
LTE duplex gap	$1.795\,\mathrm{GHz}$	$16.7\,\mathrm{cm}$	$5.6\,\mathrm{cm}$

Table 8.2: Antenna separation for an ULA of size $\lambda/2$

same time should not be larger than half the wavelength. This ensures that there do not occur any ambiguities in the parametrization of the steering vector defined in (2.1). Thus, we choose the whole size of the antenna array to be equal to $\frac{\lambda}{2}$. The resulting antenna separation for the used frequencies is listed in Table 8.2

As the investigated DoA estimation methods are based on a far-field approximation, one has to ensure that the distance R between transmitter and receiver is large enough. For electrically small antennas that have a size smaller or equal to the wavelength, one commonly says that the receiver is located in the transmitters far-field if the distance between transmitter and receiver is one order of magnitude larger compared to the wavelength [15], i.e.

$$R \gg \lambda. \tag{8.7}$$

The programming is done with the *LabVIEW Communications System Design Suite 2.0*. As the transmit USRP does not have to perform a complex estimation task but only to transmit a single tone, the programming can be completely done using the sample project *Wireless Prototyping Fundamentals* without any modification. The sample project establishes a connection to the USRP RIO and provides the complete functionality needed. The used settings for the carrier frequency of $633.75\,\mathrm{MHz}$ are summarized in Table 8.3. Note that the the button *Enable RF 0* in the Rx Configuration tab has to be deactivated in order to guarantee a continuous transmission. Otherwise the USRP would perform time division duplexing to transmit and receive over the same channel. The setting for the other RF channel do not matter for the purpose of transmission.

For the estimation, we build the LabVIEW project *DoA Estimation*. It is based on the sample project *Wireless Prototyping Fundamentals* and uses its resources to establish a connection to the receiver USRP and retrieve the measured data. Furthermore, the methods for displaying the power spectrum as well as the IQ data is kept for testing purposes. The settings we used for a carrier frequency of $633.75\,\mathrm{MHz}$ are summarized in Table 8.4. Note that the LO frequency is chosen larger than the carrier frequency by $3\,\mathrm{MHz}$. Thus, the receiver works as an IF receiver with high side LO injection as described in Section 8.2. The resulting IF frequency is $f_{\mathrm{IF}} = 3\,\mathrm{MHz}$ and the correct NCO frequency f_{NCO} is automatically chosen by the device based on the given carrier frequency. If one uses the default setting of $f_{\mathrm{LO}} = -1$, the device automatically sets the LO frequency as close to the carrier frequency as possible, i.e.

Enable RF 0	ON
Antenna RF 0	TX1
Frequency RF 0	$633.75\,\mathrm{MHz}$
LO Frequency RF 0	$636.75\,\mathrm{MHz}$
Output Power RF 0	$10\,\mathrm{dBm}$
Sample Rate	$1\,\mathrm{M/s}$
Number of Samples	8192
Tone Frequency	$100\,\mathrm{kHz}$
Start Trigger	Immediate
Number of Samples is Finite	OFF

Table 8.3: Tx configuration for the transmitter USRP

Enable RF 0 / RF 1	ON
Antenna RF 0 / RF 1	RX1 / RX2
Frequency	$633.75\,\mathrm{MHz}$
LO Frequency	$636.75\,\mathrm{MHz}$
Reference Level	$0\,\mathrm{dBm}$
Sample Rate	$10\,\mathrm{M/s}$
Number of Samples	8192
Tone Frequency	$100\,\mathrm{kHz}$
Start Trigger	Immediate
Number of Samples is Finite	ON

Table 8.4: Rx configuration for the receiver USRP

$f_c \approx f_{LO}$, which results in an homodyne receiver with zero IF.

In the following subsections we describe some effects that make the implementation difficult and present some ideas how to circumvent these issues to improve the performance of the overall system.

8.3.1 Phase Synchronization Using Multiple Receivers

A key for a good performance of the system is a coherent demodulation of the RF signal. Thus, there are two issues one has to face.

Firstly, the LOs should be phase-locked to exactly the same frequency [16]. This conditions ensures that all signals are coherent in phase at the end of the analog stage. Using only one USRP RIO, this condition is not a problem, since the phased-locked loops of both LOs are connected to the same reference oscillator as it can be seen in Figure 8.1. Nevertheless, using more than one USRP an additional phase synchronization has to be done. To synchronize multiple URSPs, the devices are able to share their reference oscillator signal with other USRP devices. It is important to note, that the shared signal has to be amplified before it is fed into the other devices. Furthermore it is crucial that all cables which connect the devices have the same length to ensure an equal propagation time of the reference signal. According to [17] such a system can be build using an *OctoClock CDA-2990* device that takes the *Ref Out* signal of one USRP, amplifies it and distributes the signal to all devices. Also the devices which provides the reference signal should use for the demodulation the reference signal form the OctoClock in order to make sure that the signal has the same propagation time for all devices.

Secondly, a similar synchronization has to be performed for the analog-digital converters, since a synchronous sampling is necessary to obtain a signal snapshot form all antennas at the same time instant. Again, this step is only important if one uses multiple USRPs. The synchronization can again be done in the same way as the LO synchronization using the *OctoClock CDA-2990*. As a reference signal the USRPs provide a PPS trigger signal at the *PPS Trig Out* port. Fore more information about the synchronization we refer to [18].

As no OctoClock device was available, we restrict ourselves to only one USRP device in the following.

8.3.2 Phase Alignment

Even if the receiver chains are fully synchronized, there remains a phase difference between the different receiver paths as shown in Figure 8.6. This offset results from the fact that the synchronization as described in the previous subsection only refers to the frequency of the LOs but not to the phase. A way to align the received signals in phase is to measure the phase offset between the channels and compensate it digitally.

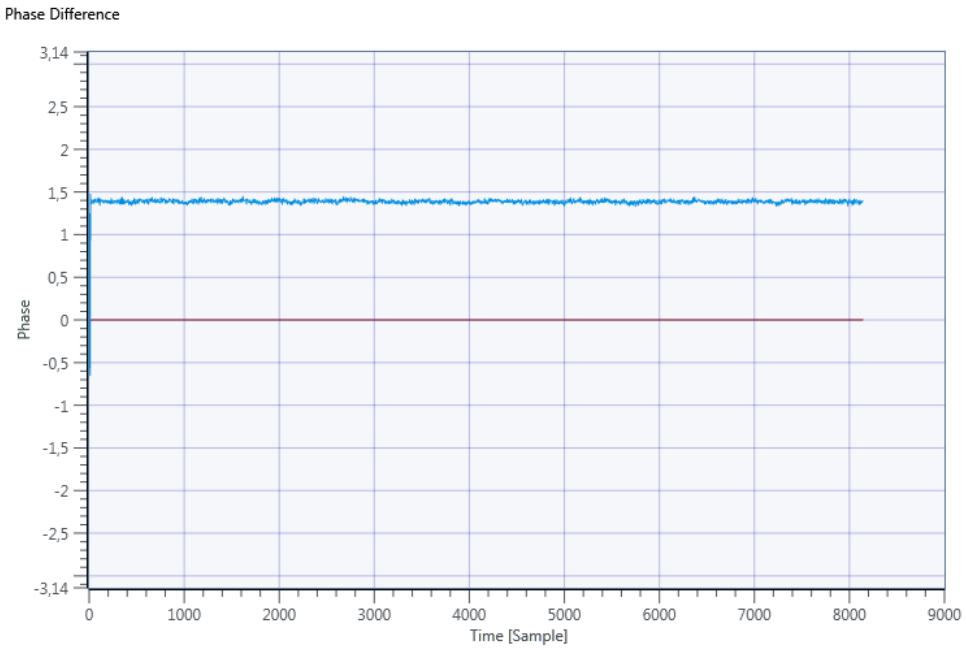

Figure 8.6: Phase offset at $f_c = 633.75\,\mathrm{MHz}$

In [19] a method is described that uses another USRP for transmitting a reference tone over a cable to all receivers. If all cables have equal length, the phase difference of the received signal stems completely from the LO phase offset. The phase offset between the received signals, is then equal to the LO phase difference. For different cable lengths, one has to take into account an additional phase offset due to the different propagation times of the reference signal.

As in our setup no cable that branches out was available, we tried to leverage the LO leakage. In general, the LO feedthrough is an undesired leakage effect but if we set the LO frequency close to the carrier frequency, the LO feedthrough can be used as a reference tone. The phase alignment is done by estimating the phase difference between the LO leakage signals.

With a LO frequency only slightly different from the carrier frequency the USRPs work as an homodyne receiver. Thus, no filter can be applied to attenuate the LO feedthrough without attenuating the useful signal as well. This is especially important, if the signal power of the received signal is at a similar level as the power of the LO leakage. Not removing the influence of the LO leakage by a filter results in strong interference and makes an accurate estimation more difficult. Unfortunately, the transmit power of the USRP is in a range where the received signal power is usually smaller than the LO feedthrough, this method of phase alignment is not applicable. A scenario where this method does not lead to a considerable disturbance of the useful

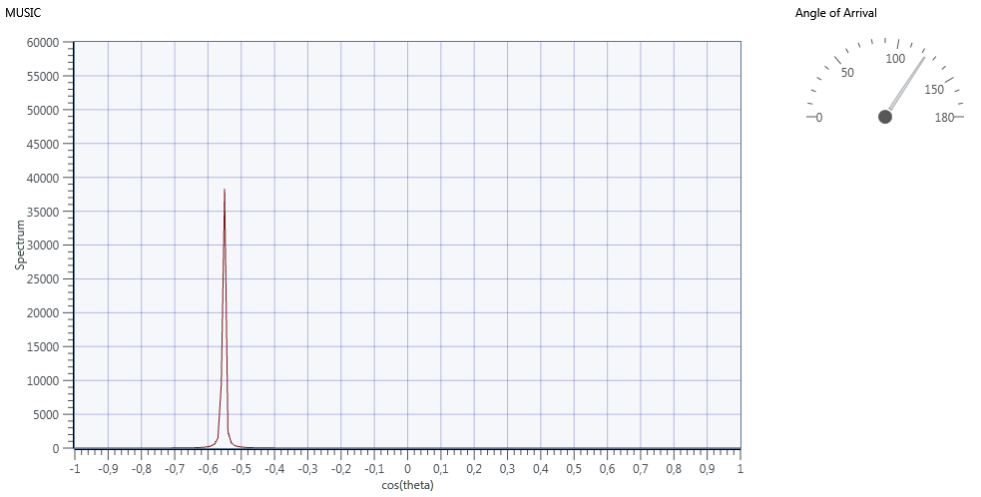

Figure 8.7: MUSIC spectrum (2 antennas, $f_c = 633.75\,\mathrm{MHz}$, approximately 100 degrees)

signal is when the microphone transmitter for lecture halls is used for transmission.

8.3.3 Power Normalization

It can be observed that the received power sometimes varies widely from receiver to receiver. This seems to result at least partly from losses due to imperfect cables. To overcome this effect, we normalize the power of the received signal to unit power before further processing.

8.4 Measurement Results

The measurements are performed with two antennas and four antennas with switching, respectively. In Figure 8.7 and Figure 8.8 are shown two examples of a typical MUSIC spectrum. As there was no possibility for a proper synchronization of the LOs, the estimated angle is quite far away from the true angle. Nevertheless, the peak is much larger using four antennas. This suggests, that one can expect a gain using more antennas even if not all sensors are measured at the same time. A MUSIC spectrum at the larger carrier frequency is depicted in Figure 8.9.

Furthermore, we observed that the measurements are strongly affected by reflections. If position of objects in the surrounding changes, also the MUSIC spectrum changes notably.

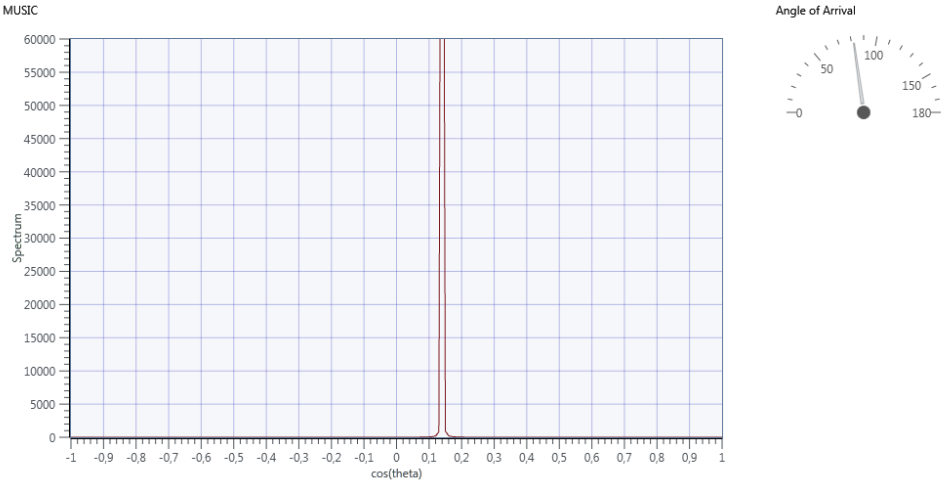

Figure 8.8: MUSIC spectrum (4 antennas, $f_c = 633.75\,\mathrm{MHz}$, approximately 100 degrees)

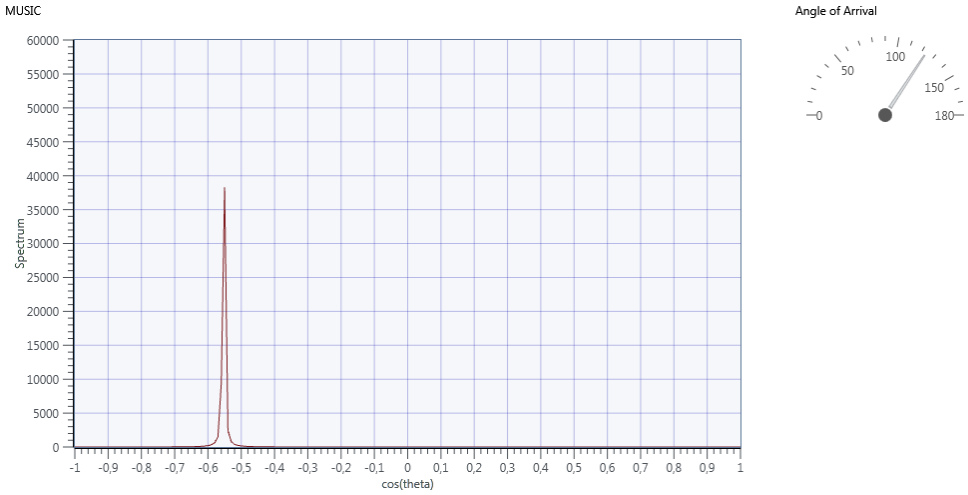

Figure 8.9: MUSIC spectrum (2 antennas, $f_c = 1.795\,\mathrm{GHz}$, approximately 100 degrees)

Chapter 9

Conclusion

In this work, we have discussed a general system model for direction finding considering also the special case of a subsampled ULA. For DoA estimation we considered the SML and DML methods along with the MUSIC algorithm. In order to obtain an impression of the achievable performance of the estimators, we derived the CRB as lower bound for the MSE. To ensure the identifiability of the estimation problem in the case of sumbsampled antenna arrays, we introduced a bound for the maximum number of impinging wavefronts for given system parameters. Our simulations have shown that in the case of fewer available RF chains than incident wavefronts the DML method performs clearly best. The problem of saturation occurring in the MUSIC algorithm was tackled by the two special subarray configurations (*(1,2,9)-subarray* and mean of *holdFirst2*) at the cost of the identifiability in case of more wavefronts. Additionally, we have been able to improve the MUSIC algorithm heuristically by the normalization of the sub-covariance matrices regarding their diagonal elements and by exploiting the ULA specific Toeplitz structure of the signal covariance matrix. For more impinging wavefronts our simulations have shown that also the SML estimation is underlying a saturation.

In the next step, it would be interesting to find a reason for the wavy structure of the SML MSE curve for $L \geq K$. Additional effort can also be put in the reduction of computational complexity and how all of this can be used in real time systems, e.g. by the use of compressed sensing. Another task is to investigate the performance of our methods in case of correlated signals.

In order to analyze the performance of the DoA estimation methods in a real world scenario, Software Defined Radio can be used. We investigated the receiver structure of the hardware devices and discussed several implementation issues. It seems crucially to synchronize the receivers accurately. Moreover, our measurements

showed that reflections due to the environment are a key problem in the used setup.

Bibliography

[1] J. Sheinvald and M. Wax, "Localization of multiple signals using subarrays data," in *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, vol. 3. IEEE, 1995, pp. 2112–2115.

[2] ——, "Direction finding with fewer receivers via time-varying preprocessing," *IEEE transactions on signal processing*, vol. 47, no. 1, pp. 2–9, 1999.

[3] M. Wax and I. Ziskind, "On unique localization of multiple sources by passive sensor arrays," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 996–1000, 1989.

[4] O. Besson and Y. I. Abramovich, "On the fisher information matrix for multivariate elliptically contoured distributions," *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1130–1133, 2013.

[5] A. Zeira and B. Friedlander, "Direction finding with time-varying arrays," *IEEE transactions on signal processing*, vol. 43, no. 4, pp. 927–937, 1995.

[6] E. Fishler and H. Messer, "Multiple source direction finding with an array of m sensors using two receivers," in *Statistical Signal and Array Processing, 2000. Proceedings of the Tenth IEEE Workshop on*. IEEE, 2000, pp. 86–89.

[7] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE signal processing magazine*, vol. 13, no. 4, pp. 67–94, 1996.

[8] T. Routtenberg and J. Tabrikian, "Non-bayesian periodic cramér-rao bound," *IEEE Transactions on Signal Processing*, vol. 61, no. 4, pp. 1019–1032, 2013.

[9] J. Sheinvald, M. Wax, and A. J. Weiss, "On the achievable localization accuracy of multiple sources at high snr," *IEEE Transactions on Signal Processing*, vol. 45, no. 7, pp. 1795–1799, 1997.

[10] National Instruments, "What is NI USRP hardware?" White Paper, Aug. 2018. [Online]. Available: http://www.ni.com/white-paper/12985/en/

[11] A. Niknejad, “Integrated cicuits for communication,” Lecture Notes, 2005.

[12] Mathworks, “Matlab 2018b. Documentation.” [Online]. Available: https://de.mathworks.com/help/signal/ref/blackmanharris.html

[13] Bundesnetzagentur, “PMSE - programme making and special events.” [Online]. Available: https://www.bundesnetzagentur.de/DE/Sachgebiete/Telekommunikation/Unternehmen_Institutionen/Frequenzen/Firmennetze/PMSE/PMSE-node.html

[14] ——, “Frequenzplan,” Mar. 2018. [Online]. Available: https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Sachgebiete/Telekommunikation/Unternehmen_Institutionen/Frequenzen/Frequenzplan.pdf?__blob=publicationFile&v=8

[15] C. Balanis, *Antenna Theory*. Wiley, 1997.

[16] National Instruments, “Introduction to the USRP-2945 and USRP-2955 software defined radios,” White Paper, Feb. 2017. [Online]. Available: http://www.ni.com/white-paper/53624/en/

[17] ——, “Angle of arrival detection with NI USRP and LabVIEW Communication,” Forum Contribution, 2017. [Online]. Available: https://forums.ni.com/t5/Software-Defined-Radio/Angle-of-Arrival-Detection-with-NI-USRP-and-LabVIEW/ta-p/3534214

[18] ——, “Synchronization overview for USRP RIO devices using LabVIEW FPGA,” Online Manual, Mar. 2018. [Online]. Available: http://zone.ni.com/reference/en-XX/help/373380J-01/usrphelp/synchronization/

[19] M. Willerton, D. Yates, V. Goverdovsky, and C. Papavassiliou, “Experimental characterization of a large aperture array localization technique using an SDR testbench,” *Wireless Innovation Forum Conference on Communications Technologies and Software Defined Radio (SDR’11-WInnComm)*, 2011.