Information-Preserving Spatial Filtering for Direction-of-Arrival Estimation

Manuel Stein, Mario Castañeda and Josef A. Nossek
Institute for Circuit Theory and Signal Processing
Technische Universität München, Munich, Germany
Email: {Stein, Castaneda, Nossek}@nws.ei.tum.de

***Abstract*—This work investigates the problem of direction-of-arrival estimation with a large number of antennas. As in practice the number of antennas M is limited by power and hardware constraints, here the possibility to compress the receive signal, prior to the estimation task, to $K < 2M$ real-valued outputs is discussed. Under a Bayesian perspective, we state the problem of finding a linear spatial filter ($2M$ inputs, K outputs) which preserves the information about the direction-of-arrival parameter in an optimum way. In order to attain the lowest possible mean squared error with the compressed data, the filter is designed such that the *Bayesian Cramér-Rao bound* is minimized. An iterative gradient-based filter solution is proposed and the potential estimation performance is investigated for different setups. Simulations of the maximum a posteriori (MAP) estimator show that the accuracy predicted by theory can be attained in practice at low computational cost.**

***Index Terms*—array signal processing, direction-of-arrival estimation, dimensionality reduction, compression, estimation theory**

I. Introduction

Estimating direction-of-arrival parameters is essential in a wide range of signal processing applications including radar, satellite-based positioning and wireless communication. In order to achieve high accuracy, it is attractive to use sensor arrays with a large number of antennas. However, in a variety of situations, scaling up the number of antennas is not possible due to limitations with respect to energy consumption and receiver complexity. Each of the M sensors requires an individual radio front-end with two real-valued analog-to-digital converters (ADC) and the amount of spatial samples processed at each time instance in the digital part of the receiver increases linearly with the number of sensors. The algorithmic complexity of classical direction-of-arrival methods becomes a problem especially with a large number of antennas. For instance for subspace-based estimation techniques, like MUSIC [1], ESPRIT [2] and Unitary ESPRIT [3], an eigenvalue or a singular value decomposition has to be computed or updated, thus requiring $\mathcal{O}(M^3)$ operations [4]. Hence, it is of practical interest, to design a multi-antenna receiver which allows to scale up the number of sensors M with low complexity, moderate energy consumption and without sacrificing the performance of direction-of-arrival estimation. Additionally the architecture of the receiver should allow to trade-off complexity versus estimation performance in a simple way and pose no limitations on the array geometry.

A. Contribution

Here we follow the idea of forming K linear combinations (spatial filter) of the $2M$ real-valued sensor signals with $K < 2M$. Further processing is then performed with the signals of the K virtual sensors at the filter output. For a fixed number of outputs K, this allows to increase the amount of sensors M arbitrarily without changing the structure of the receiver behind the spatial filter. In applications where the data has to be stored for later processing, such an approach can be used to diminish the size of the required memory. Note that such a filter can be implemented in the analog or the digital domain. However, while in such a way the number of spatial samples can be reduced severely, an inappropriate design of the spatial filter will result in significant performance degradation for the estimation of the direction-of-arrival parameter. Therefore, this work discusses the design of the spatial filter matrix. Investigating a theoretic performance characterization, associated with *Bayesian estimation theory* and optimizing the spatial filter with respect to this analytical measure, we show that with an appropriate filter design and an efficient estimation algorithm, high-performance parameter estimation is possible by exclusively using the compressed receive signal with reduced spatial dimension $K < 2M$. A particular strength of our approach is that priori knowledge about the angle-of-arrival parameter in form of a probability density can be taken into account during the filter design.

B. Related Work

The design criterions for linear transformations that have been proposed in the array processing literature can be divided into two frameworks: An energy-based compression method [5] and a sensitivity-based approach [6]. In [5], the linear mapping is chosen such that the average signal-to-noise ratio (SNR) of all impinging signals after the transformation is maximized (energy-based compression). The authors then propose to perform MUSIC on the data set with reduced dimensionality. If the transformation is chosen such that it preserves the signal energy, i.e. the original data set is projected onto the entire signal subspace, the SNR after the transformation is larger than the SNR of the original data set. This is since the noise variance in the subspace is smaller than the noise variance in the original data set. Hence, the computational complexity of the MUSIC estimation algorithm can be reduced with such a transformation and an additional SNR gain can

ISBN 978-3-8007-3584-6

be achieved, leading to higher estimation performance with MUSIC. This SNR gain is a direct consequence of taking into account the a priori knowledge that the angle-of-arrival lies within a given sector[1].

Another interesting approach is discussed in [6]. Here the linear transformation is designed based on the Cramér-Rao lower bound of the angle-of-arrival parameter (or equivalently the Fisher information matrix), which provides a direct performance characterization of efficient estimation algorithms (sensitivity-based compression). The authors show that the minimum number of dimensions of the reduced data set has to be at least equal to twice the number of impinging wavefronts in order to formulate a transformation that is lossless with respect to estimation performance. However, a serious problem within the *Fisher estimation* framework considered in [6] is the fact, that the optimum linear transformation depends on the deterministic angles-of-arrival which are unknown during the compression step. So, formulating the compression problem from a Fisher theoretic perspective (estimation of unknown deterministic parameters) can only lead to suboptimal solutions for the situations encountered in practical array signal processing scenarios. Here we circumvent this problem by stating the problem from a Bayesian perspective (estimation of unknown random parameters).

II. System Model

For the discussion, we assume an analog receive signal

$$\boldsymbol{y}(t) = \begin{bmatrix}\boldsymbol{y}_I(t) & \boldsymbol{y}_Q(t)\end{bmatrix}^T \in \mathbb{R}^{2M}, \tag{1}$$

which results from an uniform linear array (distance between sensors equal to half the wavelength) of M sensors with in-phase output $\boldsymbol{y}_I(t) \in \mathbb{R}^M$ and quadrature output $\boldsymbol{y}_Q(t) \in \mathbb{R}^M$. The receive signal model under a narrowband assumption and with a single signal source is

$$\boldsymbol{y}(t) = \gamma \boldsymbol{A}(\zeta)\boldsymbol{\Phi}(\phi)\boldsymbol{x}(t) + \boldsymbol{\eta}(t), \tag{2}$$

where the steering matrix

$$\boldsymbol{A}(\zeta) = \begin{bmatrix}\boldsymbol{A}_I^T(\zeta) & \boldsymbol{A}_Q^T(\zeta)\end{bmatrix}^T \in \mathbb{R}^{2M\times 2}, \tag{3}$$

with the direction-of-arrival parameter $\zeta \in \mathbb{R}$, is determined by the steering matrix of the in-phase sensor outputs

$$\boldsymbol{A}_I(\zeta) = \begin{bmatrix}\alpha_1(\zeta) & \beta_1(\zeta)\\ \alpha_2(\zeta) & \beta_2(\zeta)\\ \vdots & \vdots\\ \alpha_M(\zeta) & \beta_M(\zeta)\end{bmatrix} \in \mathbb{R}^{M\times 2} \tag{4}$$

and the steering matrix of the quadrature outputs

$$\boldsymbol{A}_Q(\zeta) = \begin{bmatrix}-\beta_1(\zeta) & \alpha_1(\zeta)\\ -\beta_2(\zeta) & \alpha_2(\zeta)\\ \vdots & \vdots\\ -\beta_M(\zeta) & \alpha_M(\zeta)\end{bmatrix} \in \mathbb{R}^{M\times 2}, \tag{5}$$

[1]This increase in performance is similar to the one that can be achieved by the Vandermonde invariant transformation proposed in [7], which can be interpreted as a *zoom* into the sector which contains the angles of arrival.

with entries

$$\begin{aligned}\alpha_m(\zeta) &= \cos\left((m-1)\pi \sin(\zeta)\right)\\ \beta_m(\zeta) &= \sin\left((m-1)\pi \sin(\zeta)\right).\end{aligned} \tag{6}$$

The channel phase-offset at the first sensor is modeled by

$$\boldsymbol{\Phi}(\phi) = \begin{bmatrix}\cos(\phi) & -\sin(\phi)\\ \sin(\phi) & \cos(\phi)\end{bmatrix} \in \mathbb{R}^{2\times 2}, \tag{7}$$

while the attenuation is characterized by $\gamma \in \mathbb{R}$. The signal source consists of an in-phase and a quadrature component

$$\boldsymbol{x}(t) = \begin{bmatrix}x_I(t) & x_Q(t)\end{bmatrix}^T \in \mathbb{R}^2, \tag{8}$$

which are assumed to be independent wide-sense stationary random Gaussian processes

$$\int_{-\infty}^{\infty} x_I(\lambda) x_Q(\lambda - t)\mathrm{d}\lambda = 0 \qquad \forall t, \tag{9}$$

with identical auto-correlation function

$$R_x(t) = \int_{-\infty}^{\infty} x_{I/Q}(\lambda) x_{I/Q}(\lambda - t)\mathrm{d}\lambda \qquad \forall t. \tag{10}$$

The signal impinges on the array with a random angle-of-arrival ζ, where all prior knowledge on ζ is incorporated into the prior distribution $p(\zeta)$. The term $\boldsymbol{\eta}(t) \in \mathbb{R}^{2M}$ in (1) stands for additive white sensor noise. Note that we have chosen a notation based on real-valued signals in order to allow to map the signals also on to an odd number K of real-valued outputs. The model used here is based on the complex-valued receive signal model with steering vector

$$\boldsymbol{a}(\zeta) = \begin{bmatrix}1 & e^{-\mathrm{j}\pi\cos\zeta} & \dots & e^{-\mathrm{j}(M-1)\pi\cos\zeta}\end{bmatrix}^T \in \mathbb{C}^M, \tag{11}$$

found in array processing literature.

A. Spatial Filter

The analog receive signal $\boldsymbol{y}(t)$ is processed by a spatial filter $\boldsymbol{B} \in \mathbb{R}^{2M\times K}$, such that a receive signal of reduced dimension $\boldsymbol{r}(t) \in \mathbb{R}^K$ is attained by

$$\begin{aligned}\boldsymbol{r}(t) &= \boldsymbol{B}^T\boldsymbol{y}(t)\\ &= \gamma\boldsymbol{B}^T\boldsymbol{A}(\zeta)\boldsymbol{\Phi}(\phi)\boldsymbol{x}(t) + \boldsymbol{B}^T\boldsymbol{\eta}(t).\end{aligned} \tag{12}$$

Band-limiting the K output signals to one-sided bandwidth B and sampling at a rate of $f_s = 2B$, results in temporally white compressed spatial snapshots of the form

$$\boldsymbol{r}_n = \boldsymbol{B}^T\boldsymbol{y}_n \tag{13}$$

$$= \gamma\boldsymbol{B}^T\boldsymbol{A}(\zeta)\boldsymbol{\Phi}(\phi)\boldsymbol{x}_n + \boldsymbol{B}^T\boldsymbol{\eta}_n, \tag{14}$$

with $\boldsymbol{r}_n \in \mathbb{R}^K$, $\boldsymbol{y}_n \in \mathbb{R}^{2M}$, $\boldsymbol{x}_n \in \mathbb{R}^2$ and $\boldsymbol{\eta}_n \in \mathbb{R}^{2M}$ being the sampled versions of $\boldsymbol{r}(t)$, $\boldsymbol{y}(t)$, $\boldsymbol{x}(t)$ and $\boldsymbol{\eta}(t)$, i.e.

$$\begin{aligned}\boldsymbol{r}_n &= \boldsymbol{r}\Big(\frac{n-1}{f_s}\Big)\\ \boldsymbol{y}_n &= \boldsymbol{y}\Big(\frac{n-1}{f_s}\Big)\\ \boldsymbol{x}_n &= \boldsymbol{x}\Big(\frac{n-1}{f_s}\Big)\\ \boldsymbol{\eta}_n &= \boldsymbol{\eta}\Big(\frac{n-1}{f_s}\Big),\end{aligned} \tag{15}$$

ISBN 978-3-8007-3584-6

where $n = 1, 2, \dots, N$. The signal covariance $\boldsymbol{R}_x \in \mathbb{R}^{2\times 2}$ is

$$\begin{aligned}\boldsymbol{R}_x &= \mathrm{E}_{\boldsymbol{x}_n}\left[\boldsymbol{x}_n\boldsymbol{x}_n^T\right]\\ &= \boldsymbol{I}_2,\end{aligned} \tag{16}$$

where we use $\mathrm{E}_z\left[\cdot\right]$ to denote the expected value taken with respect to the random variable z. Due to the array model, the noise of the individual array elements can be assumed to be spatially white, i.e. the noise covariance $\boldsymbol{R}_{\boldsymbol{\eta}_n} \in \mathbb{R}^{2M\times 2M}$ is

$$\begin{aligned}\boldsymbol{R}_{\boldsymbol{\eta}_n} &= \mathrm{E}_{\boldsymbol{\eta}_n}\left[\boldsymbol{\eta}_n\boldsymbol{\eta}_n^T\right]\\ &= \boldsymbol{I}_{2M}.\end{aligned} \tag{17}$$

The compressed receive covariance $\boldsymbol{R}_r(\zeta) \in \mathbb{R}^{K\times K}$ is

$$\begin{aligned}\boldsymbol{R}_r(\zeta) &= \mathrm{E}_{\boldsymbol{r}_n}\left[\boldsymbol{r}_n\boldsymbol{r}_n^T\right]\\ &= \boldsymbol{B}^T\boldsymbol{R}_y(\zeta)\boldsymbol{B},\end{aligned} \tag{18}$$

with the original receive covariance $\boldsymbol{R}_y(\zeta) \in \mathbb{R}^{2M\times 2M}$ being

$$\begin{aligned}\boldsymbol{R}_y(\zeta) &= \mathrm{E}_{\boldsymbol{y}_n}\left[\boldsymbol{y}_n\boldsymbol{y}_n^T\right]\\ &= \gamma^2\boldsymbol{A}(\zeta)\boldsymbol{A}^T(\zeta) + \boldsymbol{I}_{2M}.\end{aligned} \tag{19}$$

The probability density (PDF) of a single snapshot $\boldsymbol{r}_n \in \mathbb{R}^K$, within the receive signal consisting of N temporal samples

$$\boldsymbol{r} = \begin{bmatrix}\boldsymbol{r}_1^T & \boldsymbol{r}_2^T & \dots & \boldsymbol{r}_N^T\end{bmatrix}^T \in \mathbb{R}^{NK}, \tag{20}$$

in dependence of the direction-of-arrival parameter ζ is

$$p(\boldsymbol{r}_n|\zeta) = \frac{\exp\left(-\frac{1}{2}\boldsymbol{r}_n^T(\boldsymbol{B}^T\boldsymbol{R}_y(\zeta)\boldsymbol{B})^{-1}\boldsymbol{r}_n\right)}{\sqrt{(2\pi)^K\det\left(\boldsymbol{B}^T\boldsymbol{R}_y(\zeta)\boldsymbol{B}\right)}}. \tag{21}$$

III. Information Measure and Filter Design

Using the prior knowledge $p(\zeta)$ and the signal model (21), the goal is to formulate the problem of finding the optimum spatial filter $\boldsymbol{B}^\star$ with respect to the direction-of-arrival estimation task. Under the assumption that an *efficient* Bayesian estimator $\hat{\zeta}(\boldsymbol{r})$ is used, the estimation performance can be characterized by the *Bayesian Cramér-Rao lower bound* (BCRLB) [8]

$$\begin{aligned}\mathrm{MSE} &= \mathrm{E}_{\boldsymbol{r},\zeta}\left[\left(\hat{\zeta}(\boldsymbol{r}) - \zeta\right)^2\right]\\ &\geq \frac{1}{N\,\mathrm{E}_\zeta\left[F(\boldsymbol{B},\zeta)\right] + \mathrm{E}_\zeta\left[\left(\frac{\partial\ln p(\zeta)}{\partial\zeta}\right)^2\right]}.\end{aligned} \tag{22}$$

In order to minimize the mean squared error (MSE), the filter $\boldsymbol{B}^\star$ has to be designed such that the expected Fisher information measure (EFIM)

$$\begin{aligned}\mathrm{E}_\zeta\left[F(\boldsymbol{B},\zeta)\right] &= \int_{\mathcal{Z}}\int_{\mathcal{R}} p(\zeta)p(\boldsymbol{r}|\zeta)\left(\frac{\partial\ln p(\boldsymbol{r}|\zeta)}{\partial\zeta}\right)^2 \mathrm{d}\boldsymbol{r}\mathrm{d}\zeta\\ &= \bar{F}(\boldsymbol{B}),\end{aligned} \tag{23}$$

with $\mathcal{Z}$ and $\mathcal{R}$ denoting the support of ζ and $\boldsymbol{r}$, is maximized

$$\boldsymbol{B}^\star = \arg\max_{\boldsymbol{B}\in\mathcal{B}} \bar{F}(\boldsymbol{B}) \tag{24}$$

over the allowed set of filters $\mathcal{B}$. Note, that the classical Fisher information (FIM) is defined as [9]

$$F(\boldsymbol{B},\zeta) = \int_{\mathcal{R}} p(\boldsymbol{r}|\zeta)\left(\frac{\partial\ln p(\boldsymbol{r}|\zeta)}{\partial\zeta}\right)^2 \mathrm{d}\boldsymbol{r}. \tag{25}$$

With (21) and the substitution

$$\boldsymbol{D} = \boldsymbol{B}\left(\boldsymbol{B}^T\boldsymbol{R}_y(\zeta)\boldsymbol{B}\right)^{-1}\boldsymbol{B}^T, \tag{26}$$

the FIM is found to be given by

$$F(\boldsymbol{B},\zeta) = \frac{1}{2}\,\mathrm{tr}\left(\frac{\partial\boldsymbol{R}_y(\zeta)}{\partial\zeta}\boldsymbol{D}\frac{\partial\boldsymbol{R}_y(\zeta)}{\partial\zeta}\boldsymbol{D}\right), \tag{27}$$

where $\mathrm{tr}\left(\cdot\right)$ is the trace operator. The derivative of the receive covariance matrix $\boldsymbol{R}_y(\zeta)$ is given by

$$\frac{\partial\boldsymbol{R}_y(\zeta)}{\partial\zeta} = \gamma^2\left(\frac{\partial\boldsymbol{A}(\zeta)}{\partial\zeta}\boldsymbol{A}^T(\zeta) + \boldsymbol{A}(\zeta)\frac{\partial\boldsymbol{A}^T(\zeta)}{\partial\zeta}\right) \in \mathbb{R}^{2M\times 2M}, \tag{28}$$

where the derivative of the steering matrix is

$$\frac{\partial\boldsymbol{A}(\zeta)}{\partial\zeta} = \begin{bmatrix}\frac{\partial\boldsymbol{A}_I^T(\zeta)}{\partial\zeta} & \frac{\partial\boldsymbol{A}_Q^T(\zeta)}{\partial\zeta}\end{bmatrix}^T \in \mathbb{R}^{2M\times 2}, \tag{29}$$

with in-phase component

$$\frac{\partial\boldsymbol{A}_I(\zeta)}{\partial\zeta} = \begin{bmatrix}\frac{\partial\alpha_1(\zeta)}{\partial\zeta} & \frac{\partial\beta_1(\zeta)}{\partial\zeta}\\ \frac{\partial\alpha_2(\zeta)}{\partial\zeta} & \frac{\partial\beta_2(\zeta)}{\partial\zeta}\\ \vdots & \vdots\\ \frac{\partial\alpha_M(\zeta)}{\partial\zeta} & \frac{\partial\beta_M(\zeta)}{\partial\zeta}\end{bmatrix} \in \mathbb{R}^{M\times 2} \tag{30}$$

and quadrature component

$$\frac{\partial\boldsymbol{A}_Q(\zeta)}{\partial\zeta} = \begin{bmatrix}-\frac{\partial\beta_1(\zeta)}{\partial\zeta} & \frac{\partial\alpha_1(\zeta)}{\partial\zeta}\\ -\frac{\partial\beta_2(\zeta)}{\partial\zeta} & \frac{\partial\alpha_2(\zeta)}{\partial\zeta}\\ \vdots & \vdots\\ -\frac{\partial\beta_M(\zeta)}{\partial\zeta} & \frac{\partial\alpha_M(\zeta)}{\partial\zeta}\end{bmatrix} \in \mathbb{R}^{M\times 2}, \tag{31}$$

while the individual entries of the derivatives are

$$\begin{aligned}\frac{\partial\alpha_m(\zeta)}{\partial\zeta} &= -(m-1)\pi\cos\left(\zeta\right)\sin\left((m-1)\pi\sin\left(\zeta\right)\right)\\ \frac{\partial\beta_m(\zeta)}{\partial\zeta} &= (m-1)\pi\cos\left(\zeta\right)\cos\left((m-1)\pi\sin\left(\zeta\right)\right).\end{aligned} \tag{32}$$

IV. Filter Solution through Gradient Ascent

In order to investigate the potential estimation performance which can be attained with the compressed data $\boldsymbol{r}$, we propose an iterative gradient method to determine an approximate solution $\boldsymbol{B}^\star$. In each step l the previous solution $\boldsymbol{B}^{(l-1)}$ is updated by the gradient weighted by the step-size $\kappa^{(l)}$

$$\begin{aligned}\boldsymbol{B}^{(l)} &= \boldsymbol{B}^{(l-1)} + \kappa^{(l)}\frac{\partial\bar{F}(\boldsymbol{B})}{\partial\boldsymbol{B}}\\ &= \boldsymbol{B}^{(l-1)} + \kappa^{(l)}\,\mathrm{E}_\zeta\left[\frac{\partial F(\boldsymbol{B},\zeta)}{\partial\boldsymbol{B}}\right].\end{aligned} \tag{33}$$

ISBN 978-3-8007-3584-6

The Fisher information gradient $\frac{\partial F(\boldsymbol{B},\zeta)}{\partial \boldsymbol{B}} \in \mathbb{R}^{2M \times K}$ is

$$\begin{aligned}\frac{\partial F(\boldsymbol{B},\zeta)}{\partial \boldsymbol{B}} &= \frac{1}{2}\sum_{i,j} \frac{\partial \operatorname{tr}\left(\frac{\partial \boldsymbol{R}_y(\zeta)}{\partial \zeta}\boldsymbol{D}\frac{\partial \boldsymbol{R}_y(\zeta)}{\partial \zeta}\boldsymbol{D}\right)}{\partial [\boldsymbol{D}]_{ij}} \frac{\partial [\boldsymbol{D}]_{ij}}{\partial \boldsymbol{B}} \\ &= \frac{1}{2}\sum_{i,j} \left[\frac{\partial \operatorname{tr}\left(\frac{\partial \boldsymbol{R}_y(\zeta)}{\partial \zeta}\boldsymbol{D}\frac{\partial \boldsymbol{R}_y(\zeta)}{\partial \zeta}\boldsymbol{D}\right)}{\partial \boldsymbol{D}}\right]_{ij} \frac{\partial [\boldsymbol{D}]_{ij}}{\partial \boldsymbol{B}} \\ &= \sum_{i,j} \left[\left(\frac{\partial \boldsymbol{R}_y(\zeta)}{\partial \zeta}\right)^T \boldsymbol{D} \left(\frac{\partial \boldsymbol{R}_y(\zeta)}{\partial \zeta}\right)^T\right]_{ij} \frac{\partial [\boldsymbol{D}]_{ij}}{\partial \boldsymbol{B}}.\end{aligned} \tag{34}$$

In the following we define the matrix

$$\Delta_{ij} = (\boldsymbol{e}_j \boldsymbol{e}_i^T + \boldsymbol{e}_i \boldsymbol{e}_j^T) \in \mathbb{R}^{2M \times 2M}, \tag{35}$$

where $\boldsymbol{e}_j\, \mathbb{R}^{2M}$ is a vector with $2M$ elements set to zero except the j-th element which is set to one. With this convention

$$\begin{aligned}\frac{\partial [\boldsymbol{D}]_{ij}}{\partial \boldsymbol{B}} &= \frac{\partial \boldsymbol{e}_i^T \boldsymbol{B}\big(\boldsymbol{B}^T \boldsymbol{R}_y(\zeta)\boldsymbol{B}\big)^{-1}\boldsymbol{B}^T \boldsymbol{e}_j}{\partial \boldsymbol{B}} \\ &= \frac{\partial \operatorname{tr}\left(\boldsymbol{B}^T \boldsymbol{e}_j \boldsymbol{e}_i^T \boldsymbol{B}\big(\boldsymbol{B}^T \boldsymbol{R}_y(\zeta)\boldsymbol{B}\big)^{-1}\right)}{\partial \boldsymbol{B}} \\ &= \frac{1}{2}\frac{\partial \operatorname{tr}\left(\boldsymbol{B}^T \Delta_{ij} \boldsymbol{B}\big(\boldsymbol{B}^T \boldsymbol{R}_y(\zeta)\boldsymbol{B}\big)^{-1}\right)}{\partial \boldsymbol{B}} \\ &= \big(\boldsymbol{I} - \boldsymbol{R}_y(\zeta)\boldsymbol{D}\big)\Delta_{ij}\boldsymbol{B}\big(\boldsymbol{B}^T \boldsymbol{R}_y(\zeta)\boldsymbol{B}\big)^{-1},\end{aligned} \tag{36}$$

the information gradient is found to be given by

$$\begin{aligned}\frac{\partial F(\boldsymbol{B},\zeta)}{\partial \boldsymbol{B}} &= \sum_{i,j} \left[\left(\frac{\partial \boldsymbol{R}_y(\zeta)}{\partial \zeta}\right)^T \boldsymbol{D} \left(\frac{\partial \boldsymbol{R}_y(\zeta)}{\partial \zeta}\right)^T\right]_{ij} \cdot \\ &\quad \cdot \big(\boldsymbol{I} - \boldsymbol{R}_y(\zeta)\boldsymbol{D}\big)\Delta_{ij}\boldsymbol{B}\big(\boldsymbol{B}^T \boldsymbol{R}_y(\zeta)\boldsymbol{B}\big)^{-1}\end{aligned} \tag{37}$$

and can finally be simplified to

$$\frac{\partial F(\boldsymbol{B},\zeta)}{\partial \boldsymbol{B}} = \big(\boldsymbol{I} - \boldsymbol{R}_y(\zeta)\boldsymbol{D}\big)\boldsymbol{G}\boldsymbol{B}\big(\boldsymbol{B}^T \boldsymbol{R}_y(\zeta)\boldsymbol{B}\big)^{-1}, \tag{38}$$

where

$$\boldsymbol{G} = \frac{\partial \boldsymbol{R}_y^T(\zeta)}{\partial \zeta}\boldsymbol{D}\frac{\partial \boldsymbol{R}_y^T(\zeta)}{\partial \zeta} + \frac{\partial \boldsymbol{R}_y(\zeta)}{\partial \zeta}\boldsymbol{D}^T\frac{\partial \boldsymbol{R}_y(\zeta)}{\partial \zeta}. \tag{39}$$

V. Theoretic Performance Analysis

For an initial performance analysis we assume that the direction-of-arrival parameter ζ has zero mean and is distributed according to a symmetric beta distribution

$$p(\zeta) = \frac{1}{\pi}\frac{\Gamma(2\rho)}{\Gamma^2(\rho)}\left(\frac{\frac{\pi}{2}+\zeta}{\pi}\right)^{(\rho-1)}\left(\frac{\frac{\pi}{2}-\zeta}{\pi}\right)^{(\rho-1)}, \tag{40}$$

with support $[-\frac{\pi}{2};\frac{\pi}{2}]$ and $\rho > 2$, where

$$\Gamma(x) = \int_0^\infty \lambda^{x-1} e^{-\lambda} \mathrm{d}\lambda \tag{41}$$

is the gamma function. In this setting the variance is

$$\begin{aligned}\sigma_\zeta^2 &= \mathrm{E}_\zeta\left[(\zeta - \mu_\zeta)^2\right] \\ &= \frac{\pi^2}{4(2\rho+1)}.\end{aligned} \tag{42}$$

Larger values of ρ indicate that the variance of the angle-of-arrival ζ is smaller. In Fig. 1 the distribution $p(\zeta)$ is shown for different ρ. Note that the presented technique is not restricted to scenarios with the distribution $p(\zeta)$ used in this example.

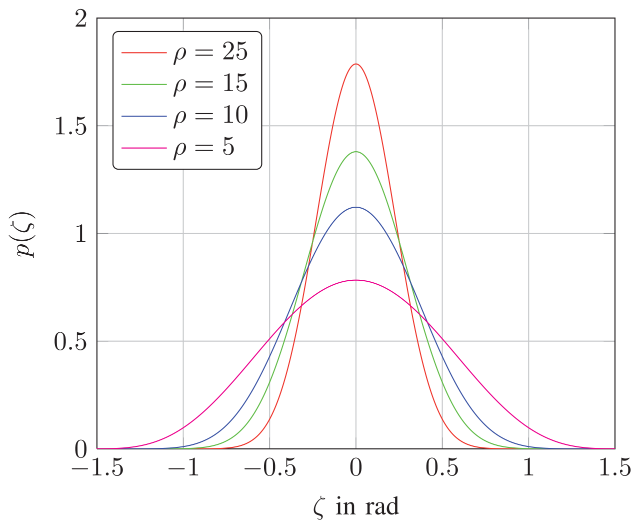

Fig. 1. Beta distribution $p(\zeta)$

A. Results - Setup Gradient Method

With the prior (40), we run the algorithm derived in section IV with random initial filter $\boldsymbol{B}^{(0)}$ and a diminishing step size

$$\kappa^{(l)} = \frac{1}{l+\beta}, \tag{43}$$

where $\beta = 10$ to guarantee convergence [10]. The algorithm is terminated when the information measure $\bar{F}\big(\boldsymbol{B}^{(l)}\big)$ improves less than 10^{-4} percent with respect to the absolute value of $\bar{F}\big(\boldsymbol{B}^{(l-1)}\big)$. Recall that the EFIM $\bar{F}(\boldsymbol{B})$ is defined in (23).

B. Results - Performance and Number of Sensors M

In order to evaluate the performance of the proposed dimension reduction technique, in Fig. 2 the square root of $\bar{F}(\boldsymbol{B}^\star)$ is depicted as a function of the amount of antennas M for different number of filter outputs K. For the plot we assume SNR $= 0$ dB with $\rho = 25$ to characterize the a priori information about the angle-of-arrival. It is observed that for each setting K, the information measure $\sqrt{\bar{F}(\boldsymbol{B}^\star)}$ scales faster than linear with the number of antennas M. While with $K = 1$ the performance is far from optimum, $K = 16$ outputs are sufficient to perform estimation without considerable accuracy-loss for the depicted range of M.

ISBN 978-3-8007-3584-6

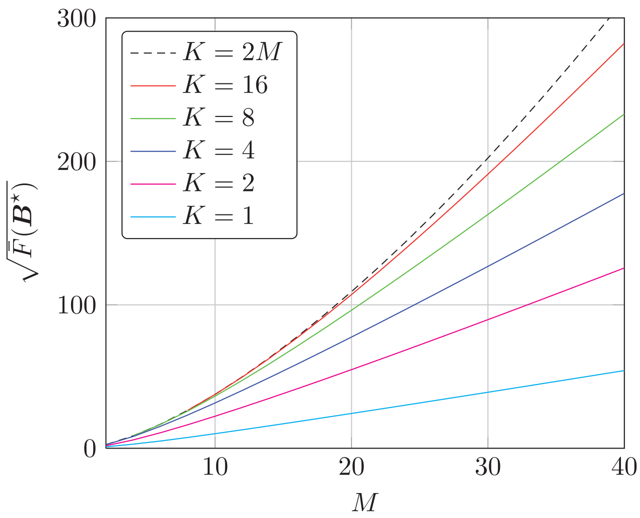

Fig. 2. EFIM vs. number of antennas (SNR $= 0$ dB, $\rho = 25$)

C. Results - Information-loss and Filter Output Size K

Fig. 3 shows the information loss

$$\chi = 10 \log \left(\frac{\bar{F}(\boldsymbol{B}^{\star})}{\bar{F}(\boldsymbol{I}_{2M})} \right) \tag{44}$$

in dB, for the scenario described in Fig. 2. The loss in (44) compares the estimation error achieved with the compressed sensor data and the original signal. It is observed that with $K = 16$ and $M = 40$ the loss is less than 1.0 dB.

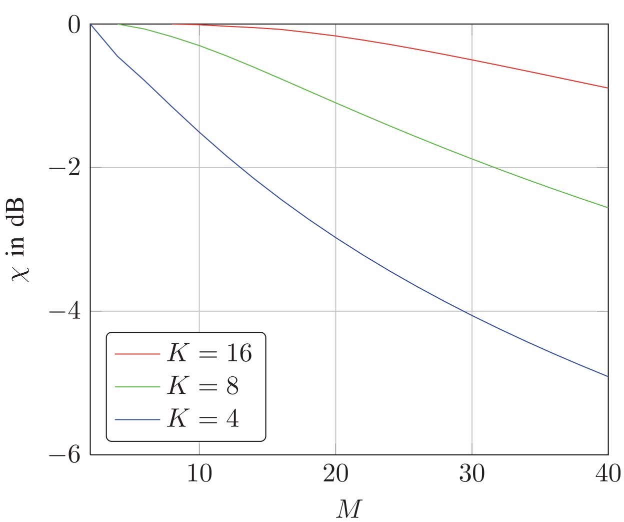

Fig. 3. Information-loss vs. number of antennas (SNR $= 0$ dB, $\rho = 25$)

D. Results - Information-loss and Prior Knowledge ρ

In order to visualize the impact of the prior knowledge $p(\zeta)$, in Fig. 4 the performance-loss χ is depicted for different prior qualities ρ for a SNR $= 0$ dB and for $M = 50$ antennas. It becomes obvious, that the prior $p(\zeta)$ plays an important role for the design of the compression matrix. If the variability of the angle of arrival parameter decreases ($\rho \to \infty$), the signal can be compressed with smaller performance-loss χ.

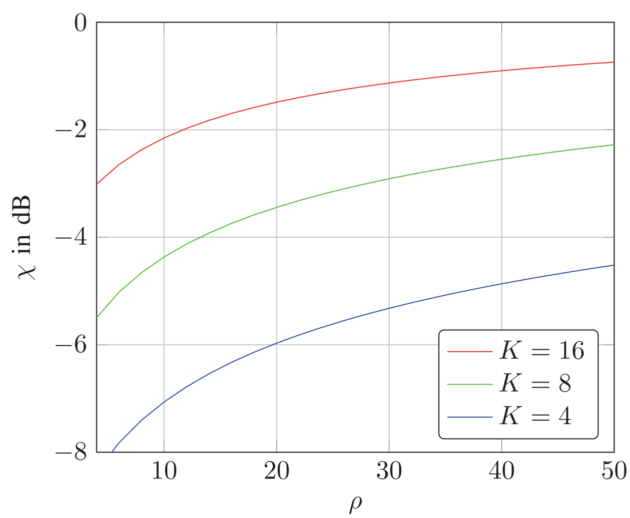

Fig. 4. Information-loss vs. prior knowledge (SNR $= 0$ dB, $M = 50$)

E. Results - Information-loss and Signal-to-Noise Ratio

The information-loss χ for $\rho = 10$ and with $M = 50$ antennas is depicted in Fig. 5 as a function of SNR. For medium SNR, where the first term of the covariance matrix $\boldsymbol{R}_y(\zeta)$ (see eq. 19) is dominant, the loss χ is smaller than at low SNR values.

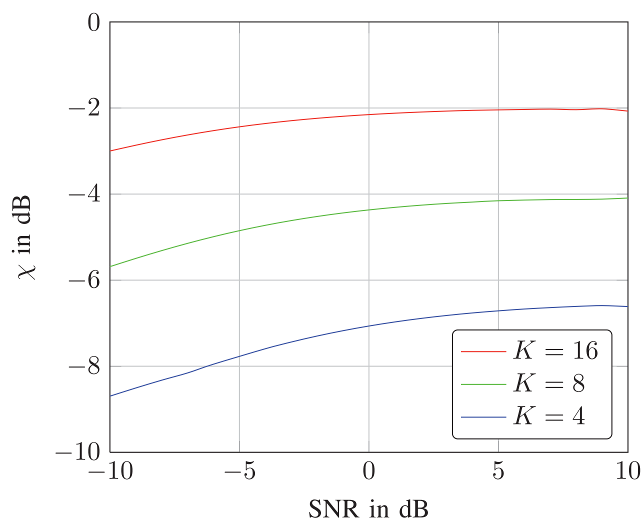

Fig. 5. Information-loss vs. SNR ($\rho = 10$, $M = 50$)

VI. Practical Performance Analysis

As the presented results exclusively rely on theoretic measures, the last part of the discussion is devoted to demonstrate the practical impact of the proposed approach. Therefore, the performance of a real estimation algorithm is investigated. We

ISBN 978-3-8007-3584-6

use the filter matrices $\boldsymbol{B}^\star$ found with the proposed method and calculate the maximum a posteriori estimator (MAP)

$$\hat{\zeta}_{\mathrm{MAP}}(\boldsymbol{r}) = \arg\max_{\zeta} \Big(\ln p(\boldsymbol{r}|\zeta) + \ln p(\zeta) \Big) = \arg\max_{\zeta} \left(\sum_{n=1}^{N} \ln p(\boldsymbol{r}_n|\zeta) + \ln p(\zeta) \right). \quad (45)$$

With the PDF of a single snapshot given by (21), we obtain

$$\ln p(\boldsymbol{r}_n|\zeta) = -\frac{K}{2}\ln(2\pi) - \frac{1}{2}\ln\Big(\det\big(\boldsymbol{B}^T\boldsymbol{R}_y(\zeta)\boldsymbol{B}\big)\Big) - \frac{1}{2}\boldsymbol{r}_n^T(\boldsymbol{B}^T\boldsymbol{R}_y(\zeta)\boldsymbol{B})^{-1}\boldsymbol{r}_n, \quad (46)$$

and the MAP estimator can be written as

$$\hat{\zeta}_{\mathrm{MAP}}(\boldsymbol{r}) = \arg\min_{\zeta} \ln\Big(\det\big(\boldsymbol{B}^T\boldsymbol{R}_y(\zeta)\boldsymbol{B}\big)\Big) + \mathrm{tr}\Big(\tilde{\boldsymbol{R}}_r(\boldsymbol{B}^T\boldsymbol{R}_y(\zeta)\boldsymbol{B})^{-1}\Big) - \frac{2}{N}\ln p(\zeta), \quad (47)$$

where the covariance matrix

$$\tilde{\boldsymbol{R}}_r = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{r}_n\boldsymbol{r}_n^T \quad (48)$$

represents an approximation of the covariance matrix $\boldsymbol{R}_r$.

A. Results - MAP Performance and Signal-to-Noise Ratio

In Fig. 6 the root mean squared error

$$\mathrm{RMSE} = \sqrt{\mathrm{E}_{\boldsymbol{r},\zeta}\left[(\hat{\zeta}_{\mathrm{MAP}}(\boldsymbol{r}) - \zeta)^2\right]} \quad (49)$$

of the MAP estimator is plotted for different values K as a function of SNR. We compare the RMSE to the theoretic performance measure (BRCLB), while for all simulations we assume $M = 16$, $\rho = 25$ and $N = 2500$ and use 10000 realizations to approximate the expectation in (49). Results in Fig. 6 show that the performance predicted in theory by the Bayesian Cramér-Rao lower bound (BCRLB) is in indeed achieved in practice if the observation length N and the SNR are chosen sufficiently large. Since we consider an array with $M = 16$ antennas, the case $K = 32$ represents the performance with the original data set. In the example the dimension of the original data ($2M = 32$ real-valued outputs) can be reduced by 50 percent ($K = 16$ real-valued outputs) without significant loss in estimation accuracy.[2]

VII. Conclusion

The problem of performing direction-of-arrival estimation from observation data of reduced dimension has been discussed. In order to compress the array data without substantial loss in the information about the direction-of-arrival parameter ζ, an optimization problem for the filter matrix $\boldsymbol{B}$ was established which leads to the smallest mean squared error with fixed output dimensionality K. Since a closed solution $\boldsymbol{B}^\star$ seems difficult to obtain, we have used a gradient-based method. Although this choice of $\boldsymbol{B}^\star$ might be suboptimal, results show that the dimension of the receive data (and therefore the analog or digital complexity of the receiver) can be reduced without loosing the capability of high-accuracy direction-of-arrival estimation. In particular the possibility to trade-off complexity versus accuracy, through the design parameter K, is an interesting feature of the proposed approach. Future works will focus on generalizing the presented results to scenarios with multiple signal sources and investigate the performance of classical direction-of-arrival methods with compressed data.

[2]For a subspace-based estimation algorithm like MUSIC/ESPRIT, the complexity reduces from $\mathcal{O}(16^3)$ to $\mathcal{O}(8^3)$. Therefore, at the expense of a negligible performance loss, the number of operations is diminished to 12.5 percent (factor 8) in this example by using the proposed compression method.

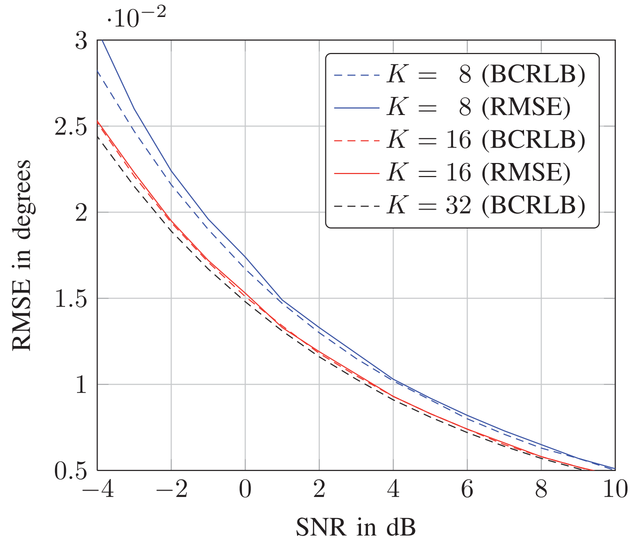

Fig. 6. RMSE with MAP estimator vs. SNR ($\rho = 25$, $M = 16$)

References

[1] R. O. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," in *IEEE Trans. Antennas Propagation*, Vol. 34, No. 3, pp. 276-280, March 1986.

[2] R. Roy, and T. Kailath, "Esprit - Estimation Of Signal Parameters Via Rotational Invariance Techniques", in *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol. 37, No. 7, pp. 984–995, July 1989.

[3] M. Haardt and J. A. Nossek, "Unitary ESPRIT: How to Obtain Increased Estimation Accuracy with a Reduced Computational Burden", in *IEEE Trans. on Signal Processing*, Vol. 43, No. 5, pp. 1232-1242, May 1995.

[4] G. H. Golub and C. F. van Loan, *Matrix Computations*, Third Edition, The John Hopkins University Press, Baltimore, MD, 1996.

[5] B. D. van Veen and B. G. Williams, "Dimensionality Reduction In High Resolution Direction Of Arrival Estimation," in *Asilomar Conference on Signals, Systems and Computers*, Vol. 2, pp. 588-592, November 1988.

[6] S. Anderson, "Optimal Dimension Reduction for Array Processing-Generalized", in *Asilomar Conference on Signals, Systems and Computers*, Vol. 2, pp. 918-922, November 1991.

[7] T. Kurpjuhn, M. T. Ivrlač and J. A. Nossek, "Vandermonde Invariance Transformation", in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vol 5, pp. 2929-2932, May 2001.

[8] H. L. Van Trees, K. L. Bell, "Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking", *Wiley-Interscience/IEEE Press*, 2007.

[9] S. M. Kay, "Fundamentals of Statistical Signal Processing: Estimation Theory", *Pretice Hall*, 1993.

[10] D. P. Bertsekas, A. Nedic and A. E. Ozdaglar: Convex Analysis and Optimization, *Athena Scientific*, 2003.

ISBN 978-3-8007-3584-6