Manifold Optimization Based Beamforming for DoA and DoD Estimation with a Single Multi-Mode Antenna

Robert Pöhlmann, Siwei Zhang, Armin Dammann

German Aerospace Center (DLR)
Institute of Communications and Navigation
Oberpfaffenhofen, 82234 Wessling, Germany
Email: {Robert.Poehlmann, Siwei.Zhang, Armin.Dammann}@dlr.de

Peter A. Hoeher

University of Kiel
Faculty of Engineering
24143 Kiel, Germany
Email: ph@tf.uni-kiel.de

Abstract—Both direction-of-arrival (DoA) and direction-of-departure (DoD) of a radio signal contain valuable information for localization. Their estimation with antenna arrays is well known. More recently, multi-mode antennas (MMAs), building on the theory of characteristic modes, have been investigated for DoA estimation. This paper introduces joint DoA and DoD estimation with a single MMA on transmitter and receiver side. In general, the polarization of a signal transmitted by an MMA varies with the direction, which makes an appropriate signal model necessary. For best performance, optimized transmit beamforming should be performed. We derive the Cramér-Rao bound (CRB) for DoA and DoD estimation with MMAs, propose an optimized beamformer (OBF), which minimizes the CRB, and evaluate its performance.

I. Introduction

Localization, specifically position and orientation estimation, has been proven useful for many applications, ranging from location information as a service in cellular networks [1] through location aware communications [2] to robotic swarms for extraterrestrial exploration missions [3]. In this context, the signal direction-of-arrival (DoA) and direction-of-departure (DoD) contain valuable information [4], [5].

Classically, DoA is estimated using antenna arrays. More recently, multi-mode antennas (MMAs) have been proposed as a new approach to DoA estimation [6], [7]. MMAs are often realized as planar antennas and build on the theory of characteristic modes. Characteristic modes are mutually orthogonal current distributions on the radiating plane. Due to orthogonality, a mode can (in theory) be excited independently from other modes via its specific antenna port. Because of different current distributions, the modes have different far-field characteristics. An example for such an antenna can be found in [8], and a first experimental validation regarding DoA estimation in [9].

Estimating the DoD with an antenna array is known [10]. The achievable DoD estimation performance depends crucially on the choice of the beamforming or precoding matrix. Several algorithms to obtain optimized beamformers exist [11]–[13]. They are, however, only applicable to antenna arrays with a constant, single polarization. DoD estimation and optimized beamforming with MMAs is not yet covered in the literature.

With this paper we introduce optimized beamformers (OBFs) to enable DoD estimation with a single MMA. For MMAs, polarization has to be considered. The key difference to a diversely polarized antenna array is that for an MMA, the signal polarization is direction dependent. On the one hand this complicates the problem, one the other hand leveraging direction dependency of the signal polarization allows to estimate the DoD without transmitting multiple concurrent beams. An optimized beamformer is still vital for good performance. To this end we derive the Cramér-Rao bound (CRB) for DoA and DoD estimation with MMAs. We then propose an OBF by minimizing the CRB with appropriate constraints. To prove its effectiveness, we show results for an MMA prototype, where we compare the CRBs of the OBF with different constraints to the conventional beamformer (CBF). A significant performance gain can be achieved for DoD estimation.

II. Signal Model

The received, sampled signal of a multiple-input, multiple-output (MIMO) communication system with M-port antennas on TX and RX side can be written as

$$\boldsymbol{r}(n) = \boldsymbol{H}\boldsymbol{f}s(nT - \tau) + \boldsymbol{w}(n), \tag{1}$$

with the channel matrix for line-of-sight propagation $\boldsymbol{H} \in \mathbb{C}^{M \times M}$, the transmit beamforming vector $\boldsymbol{f} \in \mathbb{C}^{M \times 1}$, the delayed transmit signal $s(nT - \tau)$ and white circular symmetric Gaussian noise $\boldsymbol{w}(n)$ with variance σ^2. An extension to multiple concurrent beams, i.e. a beamforming matrix to replace $\boldsymbol{f}$, is possible. Unlike for a conventional antenna array, multiple concurrent beams are not strictly required for DoD estimation with an MMA. The channel matrix,

$$\boldsymbol{H} = \boldsymbol{H}_{\mathrm{co}} + \boldsymbol{H}_{\mathrm{cross}}, \tag{2}$$

is composed of the partial channel matrices for co- and cross-polarization

$$\boldsymbol{H}_{\mathrm{co}} = \boldsymbol{a}_{\mathrm{co}}(\theta)\, h_{\mathrm{co}}\, \boldsymbol{a}_{\mathrm{co}}^H(\psi), \tag{3a}$$

$$\boldsymbol{H}_{\mathrm{cross}} = \boldsymbol{a}_{\mathrm{cross}}(\theta)\, h_{\mathrm{cross}}\, \boldsymbol{a}_{\mathrm{cross}}^H(\psi), \tag{3b}$$

for DoA θ and DoD ψ, see Figure 1. For clarity this paper is limited to the 2D case, i.e. azimuth only. An extension to 3D
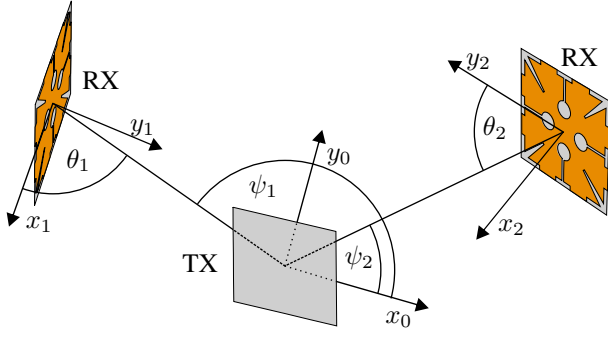
Fig. 1: Transmitting (TX) MMA with DoDs ψ_1, ψ_2 and two receiving (RX) MMAs with DoAs θ_1, θ_2.

| | 1 co | | 3 co | × | 1 cross | × | 3 cross |
| | 2 co | | 4 co | × | 2 cross | × | 4 cross |

Fig. 2: MMA power patterns of the four modes with co- and cross-polarization.

is straightforward. Here $\boldsymbol{a}_{\mathrm{co}}(\theta)$ denotes the complex-valued antenna response at the M antenna ports when the antenna is being illuminated by a wave with reference polarization and DoA θ, while $\boldsymbol{a}_{\mathrm{cross}}(\theta)$ results from illumination with the orthogonal polarization. In order to obtain a continuous, closed-form expression of the MMA response, we use wavefield modeling and manifold separation, see [6] and the references therein. We assume that the antennas on TX and RX side are identical, i.e. $\boldsymbol{a}_{\mathrm{co}}(\psi) = \boldsymbol{a}_{\mathrm{co}}(\theta)$ and $\boldsymbol{a}_{\mathrm{cross}}(\psi) = \boldsymbol{a}_{\mathrm{cross}}(\theta)$. As we consider line-of-sight propagation, the partial complex channel coefficients $h_{\mathrm{co}} = h_{\mathrm{cross}} = \alpha e^{j\varphi}$ are equal.

Figure 2 shows the x-y-plane power patterns for co- and cross-polarization following Ludwig-3 definition [14] of the MMA presented in [8]. This antenna is also assumed for the results presented in Section V. Significant contributions from co- and cross-polarization are apparent, which emphasizes the need to consider polarization diversity for MMAs.

III. CRAMÉR-RAO BOUND

The scenario under investigation is one transmitting MMA and L receiving MMAs, see Figure 1. The unknown parameters to be estimated are the vectors of DoAs,

$$\boldsymbol{\theta} = [\theta_1, ..., \theta_L]^T, \tag{4}$$

and DoDs,

$$\boldsymbol{\psi} = [\psi_1, ..., \psi_L]^T, \tag{5}$$

as well as the nuisance channel parameters $\boldsymbol{\alpha} = [\alpha_1, ..., \alpha_L]^T$, $\boldsymbol{\varphi} = [\varphi_1, ..., \varphi_L]^T$ for L links. Assuming that orthogonal signals $s_l(t)$ are transmitted to the different receivers yields independent links. It is thus sufficient to derive CRB and Fisher information matrix (FIM) for a single link l, hence we drop the subscript l for the remainder of this section. The FIM for the unknown parameter vector $\boldsymbol{\xi} = [\theta, \psi, \alpha, \varphi]^T$ and white Gaussian noise $\boldsymbol{w}(n)$ is then

$$\boldsymbol{J} = \frac{2}{\sigma^2} \operatorname{Re} \left\{ \sum_{n=1}^{N} \frac{\partial E[\boldsymbol{r}(n)]^H}{\partial \boldsymbol{\xi}} \frac{\partial E[\boldsymbol{r}(n)]}{\partial \boldsymbol{\xi}} \right\}, \tag{6}$$

with the expectation $E[.]$ and the Hermitian transpose ^{H} [15]. Realizing that $h_{\mathrm{co}}^* h_{\mathrm{co}} = h_{\mathrm{cross}}^* h_{\mathrm{cross}} = \alpha^2$ (* means complex conjugate) and assuming without loss of generality
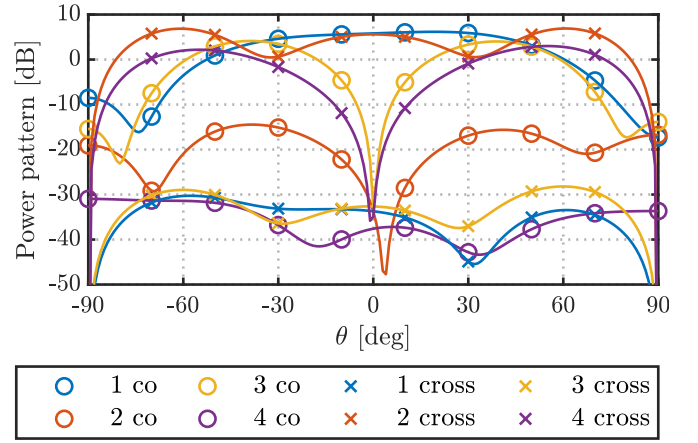
that $\frac{1}{N} \sum_{n=1}^{N} s(n)^* s(n) = 1$, it is convenient to define the signal-to-noise ratio (SNR)

$$\beta = \frac{\alpha^2}{\sigma^2} \tag{7}$$

and the entities

$$\boldsymbol{A}_{\mathrm{co}} = \boldsymbol{a}_{\mathrm{co}}(\theta)\,\boldsymbol{a}_{\mathrm{co}}^H(\psi), \tag{8a}$$

$$\boldsymbol{A}_{\mathrm{cross}} = \boldsymbol{a}_{\mathrm{cross}}(\theta)\,\boldsymbol{a}_{\mathrm{cross}}^H(\psi). \tag{8b}$$

The elements of the FIM for the signal model from Section II are then given by (22a) to (22i) in Appendix A. The DoA estimation CRB is

$$\operatorname{var}\{\hat{\theta}\} \geq \mathrm{CRB}_\theta(\theta, \psi, \boldsymbol{f}, \beta) = [\boldsymbol{J}^{-1}]_{1,1}, \tag{9}$$

and the respective DoD estimation CRB is

$$\operatorname{var}\{\hat{\psi}\} \geq \mathrm{CRB}_\psi(\theta, \psi, \boldsymbol{f}, \beta) = [\boldsymbol{J}^{-1}]_{2,2}, \tag{10}$$

where the square bracket $[.]_{i,j}$ refers to the matrix element in the i-th row and j-th column. Both depend on the actual DoA θ and DoD ψ, the beamforming vector $\boldsymbol{f}$ and the SNR β. The number of samples N is considered fixed for all links and known and is thus not shown explicitly here, but is considered by (22a) to (22i).

IV. OPTIMIZED BEAMFORMING

A. DoA only

First we want to obtain an OBF for DoA estimation at the receiver, which we call DoA-OBF. We assume that for all links $l = 1, ..., L$, we a-priori know that DoA θ_l and DoD ϕ_l lie within a certain range defined by the continuous set $\mathcal{S}_l$, i.e.

$$(\theta_l, \psi_l) \in \mathcal{S}_l. \tag{11}$$

This a-priori information could come e.g. from tracking or other sensors. As it is challenging to optimize over a continuous set, we take an approach similar to [13] and instead optimize over the Q point discrete grids

$$\tilde{\mathcal{S}}_l = \left\{ (\theta_l^1, \psi_l^1), ..., (\theta_l^Q, \psi_l^Q) \right\} \subset \mathcal{S}_l. \tag{12}$$

For large enough Q, this approximation is accurate [13]. We assume that the SNR β_l for each link is known, e.g. from previously received signals. The SNR is implicitly used to weight links in order to minimize the overall CRB. Alternatively we can set $\beta_l = 1$ for $l = 1, ..., L$, which will treat all links equally. The optimization is performed by minimizing the DoA estimation CRB, which leads to the DoA-OBF

$$\min_{\boldsymbol{f} \in \mathcal{M}} \quad \sum_{l=1}^{L} \sum_{q=1}^{Q} \mathrm{CRB}_\theta(\theta_l^q, \psi_l^q, \boldsymbol{f}, \beta_l)$$
$$\text{with} \quad \mathcal{M} = \left\{ \boldsymbol{f} \in \mathbb{C}^M : ||\boldsymbol{f}||_{\mathrm{F}} = M \right\}, \tag{13}$$

where the cost function is a double sum over the independent links and the grid points. Problem (13) is a minimization problem on the complex sphere manifold $\mathcal{M}$. Such optimization problems on manifolds can be solved efficiently [16], [17] with guaranteed convergence [18]. By setting $||\boldsymbol{f}||_{\mathrm{F}} = M$, the total power of the beamformer is constrained.

Instead of constraining the total power of the beamformer, a more practical approach is to constrain the power per antenna port, i.e. per transmitter chain. The resulting optimization problem is then

$$\min_{\boldsymbol{f} \in \mathcal{M}} \quad \sum_{l=1}^{L} \sum_{q=1}^{Q} \mathrm{CRB}_\theta(\theta_l^q, \psi_l^q, \boldsymbol{f}, \beta_l)$$
$$\text{with} \quad \mathcal{M} = \left\{ \boldsymbol{f} \in \mathbb{C}^M : |[\boldsymbol{f}]_1| = ... = |[\boldsymbol{f}]_M| = 1 \right\}, \tag{14}$$

which is an optimization problem on the complex circle manifold. $[.]_i$ refers to the i-th element of the vector.

A reference for comparison of the DoA-OBF is the well-known CBF,

$$\boldsymbol{f} = \boldsymbol{a}(\psi) \sqrt{\frac{M}{\boldsymbol{a}(\psi)^H \boldsymbol{a}(\psi)}}, \tag{15}$$

which maximizes the transmit power in one direction. We show a comparison of DoA-OBF and CBF in Section V.

B. Joint DoA and DoD

Building on the basis of the last subsection, we derive the OBF for joint DoA and DoD estimation, which we call DoA-DoD-OBF. Again the optimization minimizes the CRB, in this case for both DoA and DoD estimation. For the DoA-DoD-OBF with total power constraint, we extend (13) to

$$\min_{\boldsymbol{f} \in \mathcal{M}} \quad \sum_{l=1}^{L} \sum_{q=1}^{Q} \mathrm{CRB}_\theta(\theta_l^q, \psi_l^q, \boldsymbol{f}, \beta_l) + \mathrm{CRB}_\psi(\theta_l^q, \psi_l^q, \boldsymbol{f}, \beta_l)$$
$$\text{with} \quad \mathcal{M} = \left\{ \boldsymbol{f} \in \mathbb{C}^M : ||\boldsymbol{f}||_{\mathrm{F}} = M \right\}. \tag{16}$$

For DoA-DoD-OBF with per-port power constraint, (14) is extended to

$$\min_{\boldsymbol{f} \in \mathcal{M}} \quad \sum_{l=1}^{L} \sum_{q=1}^{Q} \mathrm{CRB}_\theta(\theta_l^q, \psi_l^q, \boldsymbol{f}, \beta_l) + \mathrm{CRB}_\psi(\theta_l^q, \psi_l^q, \boldsymbol{f}, \beta_l)$$
$$\text{with} \quad \mathcal{M} = \left\{ \boldsymbol{f} \in \mathbb{C}^M : |[\boldsymbol{f}]_1| = ... = |[\boldsymbol{f}]_M| = 1 \right\}. \tag{17}$$

Here DoA and DoD are equally weighted. Depending on the application, a weighting factor can be introduced to achieve a better trade-off.

C. Manifold Optimization

The manifold optimization problems (13), (14), (16) and (17) can efficiently be solved by the Riemannian trust-region algorithm [17], when the analytical gradient of the cost function is available. We thus derive the Euclidean gradients $\nabla_{\boldsymbol{f}} \mathrm{CRB}_\theta(\theta, \psi, \boldsymbol{f}, \beta)$ and $\nabla_{\boldsymbol{f}} \mathrm{CRB}_\psi(\theta, \psi, \boldsymbol{f}, \beta)$. Conversion to the Riemannian gradient, which is needed for manifold optimization, is provided by [16]. The partial derivative of $\boldsymbol{J}$ w.r.t. element m of $\boldsymbol{f}^*$ is

$$\frac{\partial \boldsymbol{J}}{\partial [\boldsymbol{f}]_m^*} = \boldsymbol{S}_m \boldsymbol{J}', \tag{18}$$

where $\boldsymbol{S}_m \in \mathbb{R}^{4 \times 4M}$ is a selection matrix with all zeros except $[\boldsymbol{S}_m]_{i+1, m+iM}$ with $i = 0, 1, 2, 3$ and $\boldsymbol{J}' \in \mathbb{C}^{4M \times 4}$ denotes the element-wise derivative of $\boldsymbol{J}$ w.r.t. $\boldsymbol{f}^H$, see Appendix B. With the help of [19] we derive the partial derivative of $\boldsymbol{J}_l^{-1}$,

$$\frac{\partial \boldsymbol{J}^{-1}}{\partial [\boldsymbol{f}]_m^*} = -\boldsymbol{J}^{-1} \frac{\partial \boldsymbol{J}}{\partial [\boldsymbol{f}]_m^*} \boldsymbol{J}^{-1} = -\boldsymbol{J}^{-1} \boldsymbol{S}_m \boldsymbol{J}' \boldsymbol{J}^{-1}, \tag{19}$$

and finally the Euclidean gradients

$$\nabla_{\boldsymbol{f}} \mathrm{CRB}_\theta(\theta, \psi, \boldsymbol{f}, \beta) = 2 \left[\left[\frac{\partial \boldsymbol{J}^{-1}}{\partial [\boldsymbol{f}]_1^*} \right]_{1,1}, ..., \left[\frac{\partial \boldsymbol{J}^{-1}}{\partial [\boldsymbol{f}]_M^*} \right]_{1,1} \right]^T, \tag{20}$$

$$\nabla_{\boldsymbol{f}} \mathrm{CRB}_\psi(\theta, \psi, \boldsymbol{f}, \beta) = 2 \left[\left[\frac{\partial \boldsymbol{J}^{-1}}{\partial [\boldsymbol{f}]_1^*} \right]_{2,2}, ..., \left[\frac{\partial \boldsymbol{J}^{-1}}{\partial [\boldsymbol{f}]_M^*} \right]_{2,2} \right]^T. \tag{21}$$

V. RESULTS

In this section we show simulation results to illustrate the OBF behavior and compare it to the well-known CBF. The simulations are based on the MMA prototype presented in [8], see Figure 2 for the power patterns. DoA estimation with this antenna has been investigated in [6]. We use [16], [17] to solve the manifold optimization problems. To illustrate the OBF behavior in a clear manner, we assume $Q = 1$, i.e. the optimization is performed for one specific direction per link. We assume $\beta_l = 10 \, \mathrm{dB}$ SNR and $N = 1000$ samples.

First we compare the DoA estimation CRB of CBF and DoA-OBF under two different constraints for a single link $L = 1$, fixed DoA $\theta_1 = 0°$ and varying DoD ψ_1. For each value of ψ_1, the beamforming vector $\boldsymbol{f}$ is obtained by either the CBF (15), the DoA-OBF with total power constraint (13) or the DoA-OBF with per-port power constraint (14). Figure 3a shows the square root of the resulting DoA estimation CRB. It can be seen that DoA-OBF with total power constraint offers a slight improvement over the CBF, whereas with per-port constraint the performance is slightly worse. The reason for the small difference to the CBF is that for the single link case, the CBF already maximizes the energy in a certain direction. The only thing the DoA-OBF can do in addition, is to take the sensitivity for different polarizations of the receiving antenna
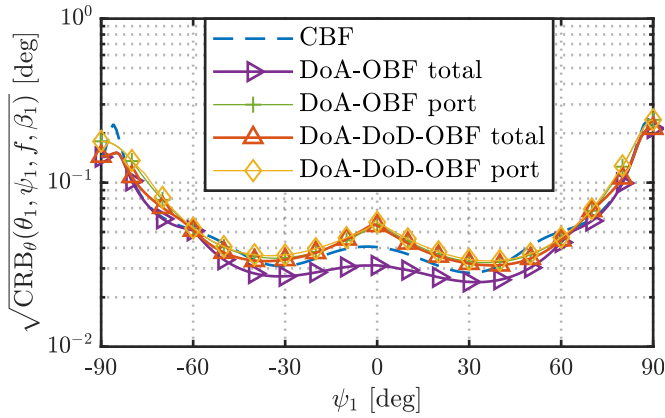
(a) DoA estimation CRB
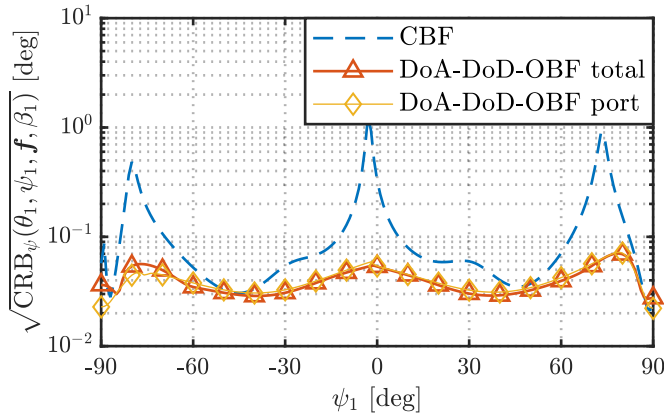
(b) DoD estimation CRB

Fig. 3: DoA and DoD estimation CRB for $\boldsymbol{f}$ obtained by CBF (15), DoA-OBF with total power constraint (13) and per-port power constraint (14), and joint DoA-DoD-OBF with total power constraint (16) and per-port power constraint (17) for fixed DoA $\theta_1 = 0°$, varying DoD ψ_1 and SNR $\beta_1 = 10\,\mathrm{dB}$.

into account, and optimize the polarization of the transmitted signal. Apparently the benefit, at least in this scenario, is small.

Now we have a look at the DoA-DoD-OBF with total power constraint (16), and the DoA-DoD-OBF with per-port power constraint (17). From Figure 3a we see that the DoA estimation CRB for the DoA-DoD-OBF is, for some angles, slightly worse than the CBF. There is almost no difference between total and per-port power constraint. In terms of DoA estimation CRB, both DoA-DoD-OBF versions are worse than the DoA-OBF with total power constraint, but almost identical to the DoA-OBF with per-port power constraint. Naturally, the DoA performance becomes worse when optimizing for both DoA and DoD estimation instead of optimizing only for DoA.

Having analyzed the impacts on the DoA estimation CRB of both DoA-OBF and DoA-DoD-OBF, we now focus on DoD estimation. Looking at Figure 3b, we can see that the square root of the DoD estimation CRB for the DoA-DoD-OBF is, for most angles, much lower than for the CBF. For some angles, the improvement is more than one order of magnitude. The
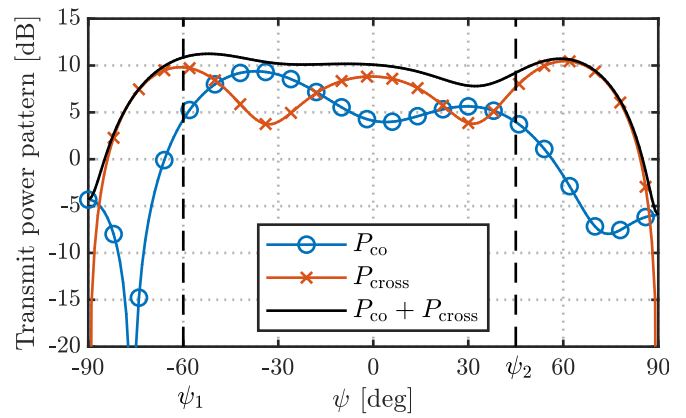

Fig. 4: Transmit power patterns for both polarizations, $P_{\mathrm{co}} = |\boldsymbol{a}_{\mathrm{co}}(\psi)^H \boldsymbol{f}|^2$ and $P_{\mathrm{cross}} = |\boldsymbol{a}_{\mathrm{cross}}(\psi)^H \boldsymbol{f}|^2$, for DoA-DoD-OBF with total power constraint (16) for $L = 2$ links and fixed $\theta_1 = 0°$, $\psi_1 = -60°$, $\theta_2 = 30°$, $\psi_2 = 45°$.

difference between the total and the per-port power constraint is barely visible. To sum up, the DoA-DoD-OBF thus trades a small loss in DoA estimation performance for large gains in DoD estimation performance.

After the single-link case, we finally have a look at joint DoA and DoD with two links. Figure 4 shows the transmit power patterns for DoA-DoD-OBF with total power constraint for two links with $\theta_1 = 0°$, $\psi_1 = -60°$, $\theta_2 = 30°$ and $\psi_2 = 45°$. The maxima of the total transmit power pattern are close to, but not perfectly at the target directions, as we do not maximize transmit power, but minimize the DoA and DoD estimation CRB. Intuitively it can be seen that the power difference (and phase difference, which is not plotted) between polarization components carries information about the DoD.

VI. CONCLUSION

This paper shows how the polarization diversity of MMAs can be exploited to estimate both DoA and DoD without the need for multiple concurrent beams. For this purpose we derive the CRB for joint DoA and DoD estimation with MMAs. We then present an OBF by minimizing the CRB for given directions under power constraints. The presented results show that the DoD estimation CRB using OBF can improve by more than one order of magnitude compared to CBF, whereas the DoA estimation CRB is virtually the same. We also show that the performance difference between using a total versus a per-port power constraint is small for this antenna. An aspect which is beyond the scope of this paper but should be considered in future work is the identifiability of the parameters of interest.

APPENDIX A
ELEMENTS OF $\boldsymbol{J}$

The elements of the FIM for the signal model from Section II are given by (22a) to (22i):

$$[\boldsymbol{J}]_{1,1} = 2\beta N \boldsymbol{f}^H \frac{\partial \boldsymbol{A}^H}{\partial \theta} \frac{\partial \boldsymbol{A}}{\partial \theta} \boldsymbol{f} \qquad (22a)$$

$$[\boldsymbol{J}]_{1,2} = [\boldsymbol{J}]_{2,1} = 2\beta N \operatorname{Re}\left\{\boldsymbol{f}^H \frac{\partial \boldsymbol{A}^H}{\partial \theta}\frac{\partial \boldsymbol{A}}{\partial \psi}\boldsymbol{f}\right\} \quad (22\text{b})$$

$$[\boldsymbol{J}]_{1,3} = [\boldsymbol{J}]_{3,1} = 2\beta N \operatorname{Re}\left\{\boldsymbol{f}^H \frac{\partial \boldsymbol{A}^H}{\partial \theta}\boldsymbol{A}\boldsymbol{f}\right\} \quad (22\text{c})$$

$$[\boldsymbol{J}]_{1,4} = [\boldsymbol{J}]_{4,1} = -2\beta N \operatorname{Im}\left\{\boldsymbol{f}^H \frac{\partial \boldsymbol{A}^H}{\partial \theta}\boldsymbol{A}\boldsymbol{f}\right\} \quad (22\text{d})$$

$$[\boldsymbol{J}]_{2,2} = 2\beta N \boldsymbol{f}^H \frac{\partial \boldsymbol{A}^H}{\partial \psi}\frac{\partial \boldsymbol{A}}{\partial \psi}\boldsymbol{f} \quad (22\text{e})$$

$$[\boldsymbol{J}]_{2,3} = [\boldsymbol{J}]_{3,2} = 2\beta N \operatorname{Re}\left\{\boldsymbol{f}^H \frac{\partial \boldsymbol{A}^H}{\partial \psi}\boldsymbol{A}\boldsymbol{f}\right\} \quad (22\text{f})$$

$$[\boldsymbol{J}]_{2,4} = [\boldsymbol{J}]_{4,2} = -2\beta N \operatorname{Im}\left\{\boldsymbol{f}^H \frac{\partial \boldsymbol{A}^H}{\partial \psi}\boldsymbol{A}\boldsymbol{f}\right\} \quad (22\text{g})$$

$$[\boldsymbol{J}]_{3,3} = [\boldsymbol{J}]_{4,4} = 2\beta N \boldsymbol{f}^H \boldsymbol{A}^H \boldsymbol{A}\boldsymbol{f} \quad (22\text{h})$$

$$[\boldsymbol{J}]_{3,4} = [\boldsymbol{J}]_{4,3} = 0. \quad (22\text{i})$$

APPENDIX B
PARTIAL DERIVATIVES OF $\boldsymbol{J}$

The element-wise partial derivatives of $\boldsymbol{J}$ w.r.t. $\boldsymbol{f}^H$ are given by (23a) to (23i):

$$[\boldsymbol{J}']_{1:M,1} = 8\beta N \frac{\partial \boldsymbol{A}^H}{\partial \theta}\frac{\partial \boldsymbol{A}}{\partial \theta}\boldsymbol{f} \quad (23\text{a})$$

$$[\boldsymbol{J}']_{1:M,2} = [\boldsymbol{J}']_{M+1:2M,1} = \\ 4\beta N \left(\frac{\partial \boldsymbol{A}^H}{\partial \theta}\frac{\partial \boldsymbol{A}}{\partial \psi} + \frac{\partial \boldsymbol{A}^H}{\partial \psi}\frac{\partial \boldsymbol{A}}{\partial \theta}\right)\boldsymbol{f} \quad (23\text{b})$$

$$[\boldsymbol{J}']_{1:M,3} = [\boldsymbol{J}']_{2M+1:3M,1} = \\ 4\beta N \left(\frac{\partial \boldsymbol{A}^H}{\partial \theta}\boldsymbol{A} + \boldsymbol{A}^H\frac{\partial \boldsymbol{A}}{\partial \theta}\right)\boldsymbol{f} \quad (23\text{c})$$

$$[\boldsymbol{J}']_{1:M,4} = [\boldsymbol{J}']_{3M+1:4M,1} = \\ \text{j}4\beta N \left(\frac{\partial \boldsymbol{A}^H}{\partial \theta}\boldsymbol{A} - \boldsymbol{A}^H\frac{\partial \boldsymbol{A}}{\partial \theta}\right)\boldsymbol{f} \quad (23\text{d})$$

$$[\boldsymbol{J}']_{M+1:2M,2} = 8\beta N \frac{\partial \boldsymbol{A}^H}{\partial \psi}\frac{\partial \boldsymbol{A}}{\partial \psi}\boldsymbol{f} \quad (23\text{e})$$

$$[\boldsymbol{J}']_{M+1:2M,3} = [\boldsymbol{J}']_{2M+1:3M,2} = \\ 4\beta N \left(\frac{\partial \boldsymbol{A}^H}{\partial \psi}\boldsymbol{A} + \boldsymbol{A}^H\frac{\partial \boldsymbol{A}}{\partial \psi}\right)\boldsymbol{f} \quad (23\text{f})$$

$$[\boldsymbol{J}']_{M+1:2M,4} = [\boldsymbol{J}']_{3M+1:4M,2} = \\ \text{j}4\beta N \left(\frac{\partial \boldsymbol{A}^H}{\partial \psi}\boldsymbol{A} - \boldsymbol{A}^H\frac{\partial \boldsymbol{A}}{\partial \psi}\right)\boldsymbol{f} \quad (23\text{g})$$

$$[\boldsymbol{J}']_{2M+1:3M,3} = [\boldsymbol{J}']_{3M+1:4M,4} = 8\beta N \boldsymbol{A}^H \boldsymbol{A}\boldsymbol{f} \quad (23\text{h})$$

$$[\boldsymbol{J}']_{2M+1:3M,4} = [\boldsymbol{J}']_{3M+1:4M,3} = \boldsymbol{0}_M. \quad (23\text{i})$$

Here $[.]_{i:j,k}$ refers to the subvector which is formed by taking the matrix elements in rows i to j and column k. $\boldsymbol{0}_M$ is an all-zero vector of length M.

REFERENCES

[1] J. A. del Peral-Rosado, R. Raulefs, J. A. López-Salcedo, and G. Seco-Granados, "Survey of cellular mobile radio localization methods: From 1G to 5G," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 1124–1148, Secondquarter 2018.

[2] R. D. Taranto, S. Muppirisetty, R. Raulefs, D. Slock, T. Svensson, and H. Wymeersch, "Location-aware communications for 5G networks: How location information can improve scalability, latency, and robustness of 5G," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 102–112, Nov. 2014.

[3] E. Staudinger, D. Shutin, C. Manss, A. Viseras, and S. Zhang, "Swarm technologies for future space exploration missions," in *Proc. I-SAIRAS '18: Int. Symp. Artificial Intelligence, Robotics and Automation in Space*, Madrid, Spanien, Jun. 2018.

[4] J. Talvitie, M. Valkama, G. Destino, and H. Wymeersch, "Novel algorithms for high-accuracy joint position and orientation estimation in 5G mmWave systems," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2017.

[5] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch, "Position and orientation estimation through millimeter-wave MIMO in 5G systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 1822–1835, Mar. 2018.

[6] R. Pöhlmann, S. A. Almasri, S. Zhang, T. Jost, A. Dammann, and P. A. Hoeher, "On the potential of multi-mode antennas for direction-of-arrival estimation," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 5, pp. 3374–3386, May 2019.

[7] S. A. Almasri, R. Pöhlmann, N. Doose, P. A. Hoeher, and A. Dammann, "Modeling aspects of planar multi-mode antennas for direction-of-arrival estimation," *IEEE Sensors Journal*, vol. 19, no. 12, pp. 4585–4597, Jun. 2019.

[8] D. Manteuffel and R. Martens, "Compact multimode multielement antenna for indoor UWB massive MIMO," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 7, pp. 2689–2697, Jul. 2016.

[9] R. Pöhlmann, G. Pedregosa, S. Caizzone, E. Staudinger, and P. A. Hoeher, "Multi-mode antenna enabled direction-of-arrival estimation for swarm navigation," in *Proc. 16th Workshop Positioning, Navigation and Communications (WPNC)*, Bremen, Germany, Oct. 2019.

[10] M. Jin, G. Liao, and J. Li, "Joint DOD and DOA estimation for bistatic MIMO radar," *Signal Processing*, vol. 89, no. 2, pp. 244–251, Feb. 2009.

[11] R. Koirala, B. Denis, D. Dardari, and B. Uguen, "Localization bound based beamforming optimization for multicarrier mmWave MIMO," in *Proc. 14th Workshop Positioning, Navigation and Communications (WPNC)*, Oct. 2017.

[12] R. Koirala, B. Denis, B. Uguen, D. Dardari, and H. Wymeersch, "Localization optimal multi-user beamforming with multi-carrier mmWave MIMO," in *Proc. IEEE 29th Annu. Int. Symp. Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2018.

[13] N. Garcia, H. Wymeersch, and D. T. M. Slock, "Optimal precoders for tracking the AoD and AoA of a mmWave path," *IEEE Transactions on Signal Processing*, vol. 66, no. 21, pp. 5718–5729, Nov. 2018.

[14] A. Ludwig, "The definition of cross polarization," *IEEE Transactions on Antennas and Propagation*, vol. 21, no. 1, pp. 116–119, Jan. 1973.

[15] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, 1st ed. Englewood Cliffs, N.J: Prentice Hall, 1993.

[16] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre, "Manopt, a Matlab toolbox for optimization on manifolds," *Journal of Machine Learning Research*, vol. 15, pp. 1455–1459, Apr. 2014.

[17] P.-A. Absil, C. Baker, and K. Gallivan, "Trust-region methods on Riemannian manifolds," *Foundations of Computational Mathematics*, vol. 7, no. 3, pp. 303–330, Jul. 2007.

[18] N. Boumal, P.-A. Absil, and C. Cartis, "Global rates of convergence for nonconvex optimization on manifolds," *IMA Journal of Numerical Analysis*, vol. 39, no. 1, Feb. 2018.

[19] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," Nov. 2012.