

Optimal Precoders for Tracking the AoD and AoA of a mmWave Path

Downloaded from: https://research.chalmers.se, 2024-03-28 21:11 UTC

Citation for the original published paper (version of record):
Garcia, N., Wymeersch, H., Slock, D. (2018). Optimal Precoders for Tracking the AoD and AoA of a mmWave Path. IEEE Transactions on Signal Processing, 66(21): 5718-5729.
http://dx.doi.org/10.1109/TSP.2018.2870368

N.B. When citing this work, cite the original published paper.

© 2018 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, or reuse of any copyrighted component of this work in other works.

This document was downloaded from http://research.chalmers.se, where it is available in accordance with the IEEE PSPB Operations Manual, amended 19 Nov. 2010, Sec, 8.1.9. (http://www.ieee.org/documents/opsmanual.pdf).

(article starts on next page)

Optimal Precoders for Tracking the AoD and AoA of a mmWave Path

Nil Garcia, Henk Wymeersch *Member, IEEE*, Dirk T. M. Slock *Fellow, IEEE*

***Abstract*—In millimeter-wave channels, most of the received energy is carried by a few paths. Traditional precoders sweep the angle-of-departure (AoD) and angle-of-arrival (AoA) space with directional precoders to identify directions with largest power. Such precoders are heuristic and lead to sub-optimal AoD/AoA estimation. We derive optimal precoders, minimizing the Cramér-Rao bound (CRB) of the AoD/AoA under a given uncertainty range, assuming a fully digital architecture at the transmitter and spatial filtering of a single path. The precoders are found by solving a suitable convex optimization problem. We demonstrate that the accuracy can be improved by at least a factor of two over traditional precoders, and show that there is an optimal number of distinct precoders beyond which the CRB does not improve.**

I. Introduction

MILLIMETER-WAVE (mmWave) communication is expected to be one of the key enablers of 5th generation cellular networks [1]. Operating in mmWave frequencies offers a few advantages. First, there are large portions of underutilized bandwidth which could be used for multi-gigabit communications [2]. Second, due to the much shorter wavelength, MIMO systems consisting of many antennas can be compacted into much smaller sizes. However, in order to compensate for the stringent path loss characteristic of millimeter wave, highly directional beamforming is necessary at the transmitter and/or receiver [3]. Since optimal precoding in communications can only be achieved after learning the channel, it is critical that fast and precise channel estimation techniques are developed.

The mmWave channel can be considered parsimonious in the sense that only a few multipath components carry non-negligible energy [4], [5]. This structure inherent to the mmWave channel is leveraged in many techniques for performing quicker and/or finer channel estimation [6]–[10]. Channel estimation can be categorized in (a) initial access, or (b) tracking. Typically, the output of initial access is a set of angle-of-departures (AoD), angle-of-arrivals (AoA) and channel gains of the individual paths. Due to the mobility of the users and the variability of the environment [11], [12], the validity of the initial access estimates come with an expiration time, and so it is necessary to estimate/track the channel periodically. The main idea behind tracking is that channel state information (CSI) is carried over to the next iteration, thus increasing the channel estimation accuracy and/or reducing the channel estimation overhead [13]–[16]. Because the channel parameters can be connected to the environment (such as is the case in geometric models), statistical CSI may also be obtained through extraneous positioning and sensing technologies such as GPS, radars and cameras [12], [17]. For instance, if prior knowledge on the user position is available through GPS, then the set of possible directions between the base station and the user can be considerably reduced [18]–[20].

Whether initial access or tracking, the most common procedure for estimating the mmWave channel consists in sweeping the channel with beams at the transmitter and/or receiver [8], [21], [22]. By detecting the time at which the received power is the largest, the correct pair of beams can be identified and the AoD and/or AoA estimated for each path. While intuitive, such beams may not necessarily yield the best achievable estimation accuracy. To the best of our knowledge, the literature in mmWave precoding has not addressed what are the fundamental limits in terms of AoA and AoD estimation when performing *optimal precoding* at the transmitter.

In this work, we seek to find the best transmit precoders for estimating the AoD and AoA, for a single-path channel, assuming that the AoD and AoA are known to lie within certain ranges of angles. Coarse knowledge on the AoD/AoA may be available, for instance, through initial access or by tracking the AoD/AoA. To abstract the analysis from specific estimators, we use the Cramér-Rao bound (CRB) from [23] as a proxy metric for the variance of the AoD and AoA estimates. Indeed, the CRB is a lower bound on the variance of any unbiased estimator and it is tight at high SNR under some mild conditions [24]. Our main contributions are:

1) Novel formulation of the optimal precoders according to the CRB on the AoD/AoA, with the option to include array gain constraints.
2) Global optimization of the proposed non-convex problems by leveraging tools of convex optimization and majorization theory.
3) Qualitative interpretation of the CRB expressions, and analysis of the optimal beampatterns.

The remainder of this paper is organized as follows. In Section II the system model and problem are described. Section III contains the formulation of the proposed precoders as well as the solution strategy and different extensions. Numerical results are presented in Section IV, followed by conclusions in Section V.

N. Garcia and H. Wymeersch are with the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden. D. T. M. Slock is with the Department of Communication Systems, EURECOM, Biot, France. This research was supported in part, by the EU HIGHTS project (High precision positioning for cooperative ITS applications under grant MG-3.5a-2014-636537) and 5GCAR, as well as the VINNOVA COPPLAR project, funded under Strategic Vehicle Research and Innovation grant No. 2015-04849.

II. System Model

Assume a transmitter (Tx) and a receiver (Rx) with N_{Tx} and N_{Rx} antennas, respectively. The Tx sends M consecutive training sequences (or pilots) consisting of K symbols each, precoded by the vectors $\mathbf{f}_1, \ldots, \mathbf{f}_M \in \mathbb{C}^{N_{\mathrm{Tx}} \times 1}$. The precoders are normalized to $\|\mathbf{f}_m\|_2^2 = 1/M$ for all m, so that increasing its number does not result in increased transmitted energy. We assume a narrow-band model where the signal bandwidth is much smaller than the carrier frequency. The equivalent discrete baseband signal for the k-th symbol of the m-th training sequence at the Rx is

$$\mathbf{y}_{m,k} = \mathbf{W}^{\mathrm{H}} \sum_{p=1}^{P} \alpha_p \, \mathbf{a}_{\mathrm{Rx}}(\phi_p) \mathbf{a}_{\mathrm{Tx}}^{\mathrm{H}}(\theta_p) \, \mathbf{f}_m s_k + \mathbf{n}_{m,k}, \qquad (1)$$

where α_p, ϕ_p, θ_p are the channel gain, AoA and AoD, respectively, of path p, $\mathbf{a}_{\mathrm{Tx}}(\theta) \in \mathbb{C}^{N_{\mathrm{Tx}} \times 1}$ and $\mathbf{a}_{\mathrm{Rx}}(\phi) \in \mathbb{C}^{N_{\mathrm{Rx}} \times 1}$ are the array responses at the Tx and Rx, $\mathbf{n}_{m,k} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$ is white[1] Gaussian noise, and $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_L] \in \mathbb{C}^{N_{\mathrm{Rx}} \times L}$ is a fixed combining matrix. The matrix $\mathbf{W}$ may correspond to a hybrid array with $L < N_{\mathrm{Rx}}$ RF chains, e.g., when the Rx is a low-complexity user terminal. For the Tx we assume a fully digital architecture.

We assume that the P paths are well-separated in terms of AoD and that the Tx knows that the dominant path ($p = 1$) has an AoD and AoA which are known to be within a small range of angles[2]: $\theta_1 \in \mathcal{R}_{\mathrm{Tx}}$ and $\phi_1 \in \mathcal{R}_{\mathrm{Rx}}$. These ranges represent a priori uncertainty regarding the AoD/AoA, originating from mobility of the Rx, or from uncertainty in the tracking algorithm, or from uncertainty in location-aided communications [18]–[20]. This assumption allows us to simplify (1) to

$$\mathbf{y}_{m,k} = \alpha_1 \, \mathbf{W}^{\mathrm{H}} \mathbf{a}_{\mathrm{Rx}}(\phi_1) \, \mathbf{a}_{\mathrm{Tx}}^{\mathrm{H}}(\theta_1) \, \mathbf{f}_m s_k + \mathbf{n}_{m,k}, \qquad (2)$$

provided $\mathbf{a}_{\mathrm{Tx}}^{\mathrm{H}}(\theta_p) \, \mathbf{f}_m \approx 0$ for $p \neq 1$ and all m. So going forward the path index will be omitted. Let x^* be the complex conjugate of x. After coherent combining across time k per precoder $\mathbf{f}_m$, we find $\mathbf{y}_m = \sum_{k=1}^{K} s_k^* \mathbf{y}_{m,k}$, leading to the following matrix observation

$$\begin{aligned} \mathbf{Y} &= \frac{1}{\|\mathbf{s}\|_2} \begin{bmatrix} \mathbf{y}_1 & \cdots & \mathbf{y}_M \end{bmatrix} \\ &= \alpha \|\mathbf{s}\|_2 \mathbf{W}^{\mathrm{H}} \mathbf{a}_{\mathrm{Rx}}(\phi) \, \mathbf{a}_{\mathrm{Tx}}^{\mathrm{H}}(\theta) \, \mathbf{F} + \mathbf{N} \end{aligned} \qquad (3)$$

where $\mathbf{F} = [\mathbf{f}_1 \cdots \mathbf{f}_M]$, $\mathbf{s} = [s_1, \ldots, s_K]$, and the components of $\mathbf{N}$ are i.i.d. Gaussian variables with variance σ^2. For notational convenience, and since we assume $\mathbf{W}$ to be fixed, we introduce $\mathbf{b}_{\mathrm{Rx}}(\phi) = \mathbf{W}^{\mathrm{H}} \mathbf{a}_{\mathrm{Rx}}(\phi)$.

Our goal is to find optimal precoders $\mathbf{F}$ to maximize the quality of the AoD and AoA estimates, given no knowledge of the channel gain α. We tackle this problem by minimizing the CRB[3] on the AoD θ and/or AoA ϕ.

[1]The noise is uncorrelated only if the L combiners are pair-wise orthogonal, i.e., $\mathbf{W}^{\mathrm{H}}\mathbf{W} = \mathbf{I}$. An example of hybrid architecture with pair-wise orthogonal combiners is the case of arrays of subarrays where each RF chain is routed to a disjoint subset of antennas [25]. If the combiners were not pair-wise orthogonal, by whitening the received signal through $(\mathbf{W}^{\mathrm{H}}\mathbf{W})^{-\frac{1}{2}}$, we would reach the same expression (2) but with a different combining matrix $\widetilde{\mathbf{W}} = \mathbf{W}(\mathbf{W}^{\mathrm{H}}\mathbf{W})^{-\frac{1}{2}}$.

[2]If the prior distributions of the AoD/AoA are provided instead, these ranges may be obtained from the confidence intervals.

III. Optimal Precoders

To design the precoders, we choose as metric the CRB, which is a lower bound on the variance of any unbiased estimator. Such bound is well suited to this problem because it generally leads to tractable mathematical expressions, and more importantly, for a sufficiently large SNR and under some mild conditions [24], the variance of the maximum likelihood estimator (MLE) is tight to the CRB. If θ and ϕ are the AoD and AoA, respectively, of the LOS path, then determining θ yields the direction from the Tx to the Rx, and determining ϕ provides the Rx's orientation with respect to the Tx. Thus, for naming purposes and without loss of generality, we refer to the CRB on the θ as direction error bound (DEB), and the CRB on ϕ as orientation error bound (OEB). Once the DEB and OEB are derived, we will proceed with the optimization of the precoder.

A. Problem Formulation

From Appendix A, the direction error bound is

$$\mathrm{var}(\hat{\theta}) \geq \mathrm{DEB} = \left[2 \, \mathrm{SNR} \, \|\mathbf{b}_{\mathrm{Rx}}(\phi)\|_2^2 \left(\left\| \mathbf{F}^{\mathrm{H}} \dot{\mathbf{a}}_{\mathrm{Tx}}(\theta) \right\|_2^2 - \frac{\left| \mathbf{a}_{\mathrm{Tx}}^{\mathrm{H}}(\theta) \, \mathbf{F} \mathbf{F}^{\mathrm{H}} \dot{\mathbf{a}}_{\mathrm{Tx}}(\theta) \right|^2}{\|\mathbf{F}^{\mathrm{H}} \mathbf{a}_{\mathrm{Tx}}(\theta)\|_2^2} \right) \right]^{-1}, \qquad (4)$$

and the orientation error bound is

$$\mathrm{var}(\hat{\phi}) \geq \mathrm{OEB} = \left[2 \, \mathrm{SNR} \left\| \mathbf{F}^{\mathrm{H}} \mathbf{a}_{\mathrm{Tx}}(\theta) \right\|_2^2 \left(\left\| \dot{\mathbf{b}}_{\mathrm{Rx}}(\phi) \right\|_2^2 - \frac{\left| \mathbf{b}_{\mathrm{Rx}}^{\mathrm{H}}(\phi) \, \dot{\mathbf{b}}_{\mathrm{Rx}}(\phi) \right|^2}{\|\mathbf{b}_{\mathrm{Rx}}(\phi)\|_2^2} \right) \right]^{-1}, \qquad (5)$$

where $\dot{\mathbf{a}}_{\mathrm{Tx}}(\theta) \triangleq \frac{\mathrm{d}\mathbf{a}_{\mathrm{Tx}}(\theta)}{\mathrm{d}\theta}$, $\dot{\mathbf{b}}_{\mathrm{Rx}}(\phi) \triangleq \frac{\mathrm{d}\mathbf{b}_{\mathrm{Rx}}(\phi)}{\mathrm{d}\phi}$ and

$$\mathrm{SNR} \triangleq \frac{|\alpha|^2 \|\mathbf{s}\|_2^2}{\sigma^2}. \qquad (6)$$

Coarsely speaking, the OEB is inversely proportional to the energy transmitted towards the Rx, $\|\mathbf{F}^{\mathrm{H}} \mathbf{a}_{\mathrm{Tx}}(\theta)\|_2^2$; whereas the term $\mathbf{F}^{\mathrm{H}} \dot{\mathbf{a}}_{\mathrm{Tx}}(\theta)$ that appears in the DEB indicates that it benefits from large variations on the precoders' radiation pattern. For more detailed intuitive interpretations of the DEB and OEB see Appendix B and the examples of Fig. 2.

Both lower bounds depend on the precoders $\mathbf{F}$, but also on the AoD θ and AoA ϕ which are unknown. Next, we propose a min-max approach to the precoders design problem:

Problem 1. *AoD-AoA-optimal precoders. Find the precoders that minimize the worst case DEB or OEB for all possible values of the AoD and AoA.*

$$\min_{\substack{\|\mathbf{f}_m\|_2 = M^{-1} \\ m=1,\ldots,M}} \max_{(\theta,\phi) \in \mathcal{R}_{\mathrm{Tx}} \times \mathcal{R}_{\mathrm{Rx}}} \max \{\mathrm{DEB}, \mathrm{OEB}\} \qquad (7)$$

[3]The use of the CRB requires a high SNR operating condition. While mmWave communication operates under low SNR without beamforming [26], our tracking scenario is congruent with a medium-to-high SNR assumption.

The dependence of the DEB and OEB on θ, ϕ and $\mathbf{F}$ has been omitted for notation clarity. Alternatively, if the only parameter of interest is the AoD, the optimal precoders are the solution to:

Problem 2. *AoD-optimal precoders. Find the precoders that minimize the worst case DEB for all possible values of the AoD and AoA.*

$$\min_{\substack{\|\mathbf{f}_m\|_2=M^{-1}\\ m=1,\ldots,M}} \max_{(\theta,\phi)\in\mathcal{R}_{\mathrm{Tx}}\times\mathcal{R}_{\mathrm{Rx}}} \mathrm{DEB} \tag{8}$$

Conversely, if the only parameter of interest is the AoA, the optimal precoders are obtained from:

Problem 3. *AoA-optimal precoders. Find the precoders that minimize the worst case OEB for all possible values of the AoD and AoA.*

$$\min_{\substack{\|\mathbf{f}_m\|_2=M^{-1}\\ m=1,\ldots,M}} \max_{(\theta,\phi)\in\mathcal{R}_{\mathrm{Tx}}\times\mathcal{R}_{\mathrm{Rx}}} \mathrm{OEB} \tag{9}$$

B. Convex Reformulation

Problems 1–3 are non-convex with respect to $\mathbf{F}$, and therefore, computing their global minimum efficiently is challenging. For a proof of non-convexity, note that the OEB is concave because it is proportional to $\|\mathbf{F}^{\mathrm{H}}\mathbf{a}_{\mathrm{Tx}}(\theta)\|^{-2}$, and the DEB depends on the concave term $\|\mathbf{F}^{\mathrm{H}}\dot{\mathbf{a}}_{\mathrm{Tx}}(\theta)\|^{-2}$. In this section, Problems 1–3 will be reformulated as conic optimization problems, which is a subclass of convex problems. The conditions upon which the original problems and their convex reformulations are equivalent (in the sense that yield the same solution) will be analyzed in the next section. For reference, Fig. 1 presents a flow chart of the steps described in this section and the following one for obtaining the optimal precoders.

The analysis focuses in Problem 1 because the solutions to Problems 2 and 3 will be shown to be subcases. An equivalent optimization problem to Problem 1 is

$$\max_{\substack{\|\mathbf{f}_m\|_2=M^{-1}\\ m=1,\ldots,M}} \min_{(\theta,\phi)\in\mathcal{R}_{\mathrm{Tx}}\times\mathcal{R}_{\mathrm{Rx}}} \min\left\{\mathrm{DEB}^{-1},\mathrm{OEB}^{-1}\right\}. \tag{10}$$

By introducing a slack variable t, the above problem can be expressed in the hypograph form:

$$\begin{aligned}
\max_{\mathbf{F},t}\quad & t && (11a)\\
\text{s.t.}\quad & \min_{(\theta,\phi)\in\mathcal{R}_{\mathrm{Tx}}\times\mathcal{R}_{\mathrm{Rx}}} \mathrm{DEB}^{-1}\geq t && (11b)\\
& \min_{(\theta,\phi)\in\mathcal{R}_{\mathrm{Tx}}\times\mathcal{R}_{\mathrm{Rx}}} \mathrm{OEB}^{-1}\geq t && (11c)\\
& \|\mathbf{f}_m\|_2=\frac{1}{M} \qquad m=1,\ldots,M. && (11d)
\end{aligned}$$

Grid Approximation: The continuous set $\mathcal{R}_{\mathrm{Tx}}\times\mathcal{R}_{\mathrm{Rx}}$ makes optimizing over constraints (11b)–(11c) challenging. Instead we approximate it by a grid $\widetilde{\mathcal{R}}_{\mathrm{Tx}}\times\widetilde{\mathcal{R}}_{\mathrm{Rx}}$ such that

$$\widetilde{\mathcal{R}}_{\mathrm{Tx}}=\left\{\vartheta^1,\ldots,\vartheta^{S_{\mathrm{Tx}}}\right\}\subset\mathcal{R}_{\mathrm{Tx}} \tag{12}$$

$$\widetilde{\mathcal{R}}_{\mathrm{Rx}}=\left\{\varphi^1,\ldots,\varphi^{S_{\mathrm{Rx}}}\right\}\subset\mathcal{R}_{\mathrm{Rx}}, \tag{13}$$

where S_{Tx} and S_{Rx} are the number of discrete angles at the Tx and Rx, respectively, within the prior ranges. Also for notation brevity, define

$$\begin{aligned}
\mathbf{a}_{\mathrm{Tx}}^i&\triangleq\mathbf{a}_{\mathrm{Tx}}(\vartheta^i) & \mathbf{b}_{\mathrm{Rx}}^q&\triangleq\mathbf{b}_{\mathrm{Rx}}(\varphi^q)\\
\dot{\mathbf{a}}_{\mathrm{Tx}}^i&\triangleq\dot{\mathbf{a}}_{\mathrm{Tx}}(\vartheta^i) & \dot{\mathbf{b}}_{\mathrm{Rx}}^q&\triangleq\dot{\mathbf{b}}_{\mathrm{Rx}}(\varphi^q).
\end{aligned} \tag{14}$$

Then, by replacing the continuous set by the grid, and since the DEB/OEB formulas (4)–(5) decouple into a part that depends only on θ and another on ϕ, the left side of (11b)–(11c) can be expressed as

$$\min_{(\theta,\phi)\in\mathcal{R}_{\mathrm{Tx}}\times\mathcal{R}_{\mathrm{Rx}}} \mathrm{DEB}^{-1}\approx$$

$$\approx 2\,\mathrm{SNR}\,K_{\mathrm{D}}\min_{i\in\{1,\ldots,S_{\mathrm{Tx}}\}}\left(\left\|\mathbf{F}^{\mathrm{H}}\dot{\mathbf{a}}_{\mathrm{Tx}}^i\right\|_2^2-\frac{\left|\mathbf{a}_{\mathrm{Tx}}^{i\mathrm{H}}\mathbf{F}\mathbf{F}^{\mathrm{H}}\dot{\mathbf{a}}_{\mathrm{Tx}}^i\right|^2}{\left\|\mathbf{F}^{\mathrm{H}}\mathbf{a}_{\mathrm{Tx}}^i\right\|_2^2}\right) \tag{15}$$

$$\min_{(\theta,\phi)\in\mathcal{R}_{\mathrm{Tx}}\times\mathcal{R}_{\mathrm{Rx}}} \mathrm{OEB}^{-1}\approx 2\,\mathrm{SNR}\,K_{\mathrm{O}}\min_{i\in\{1,\ldots,S_{\mathrm{Tx}}\}}\left\|\mathbf{F}^{\mathrm{H}}\mathbf{a}_{\mathrm{Tx}}^i\right\|_2^2 \tag{16}$$

where K_{D} and K_{O} are the following constants (they do not depend on the optimizing variables $\mathbf{F}$ or t)

$$K_{\mathrm{D}}\triangleq\min_{q\in\{1,\ldots,S_{\mathrm{Rx}}\}}\|\mathbf{b}_{\mathrm{Rx}}^q\|_2^2 \tag{17}$$

$$K_{\mathrm{O}}\triangleq\min_{q\in\{1,\ldots,S_{\mathrm{Rx}}\}}\left(\left\|\dot{\mathbf{b}}_{\mathrm{Rx}}^q\right\|_2^2-\left|\mathbf{b}_{\mathrm{Rx}}^{q\mathrm{H}}\dot{\mathbf{b}}_{\mathrm{Rx}}^q\right|^2\|\mathbf{b}_{\mathrm{Rx}}^q\|_2^{-2}\right). \tag{18}$$

Combining (11) with (15)–(16), results in

$$\begin{aligned}
\max_{\mathbf{F},t}\quad & t && (19a)\\
\text{s.t.}\quad & K_{\mathrm{D}}\left(\left\|\mathbf{F}^{\mathrm{H}}\dot{\mathbf{a}}_{\mathrm{Tx}}^i\right\|_2^2-\frac{\left|\mathbf{a}_{\mathrm{Tx}}^{i\mathrm{H}}\mathbf{F}\mathbf{F}^{\mathrm{H}}\dot{\mathbf{a}}_{\mathrm{Tx}}^i\right|^2}{\left\|\mathbf{F}^{\mathrm{H}}\mathbf{a}_{\mathrm{Tx}}^i\right\|_2^2}\right)\geq t && (19b)\\
& K_{\mathrm{O}}\left\|\mathbf{F}^{\mathrm{H}}\mathbf{a}_{\mathrm{Tx}}^i\right\|_2^2\geq t && (19c)\\
& \|\mathbf{f}_m\|_2=\frac{1}{M} \qquad m=1,\ldots,M. && (19d)
\end{aligned}$$

for $i=1,\ldots,S_{\mathrm{Tx}}$, where (19b) and (19c) are constraints on the DEB and OEB, respectively, and (19d) is a constraint on the total transmit power.

Relaxation on the Energy Constraint: In order to transform (19) into a convex problem, we propose a relaxation of the feasible set. If the solution to the newly relaxed problem is still within the original feasible set, then it is the optimum solution of the original problem. The first proposed relaxation consists of replacing $\|\mathbf{f}_m\|_2=M^{-1}$ for $m=1,\ldots,M$, by $\sum_{m=1}^M\|\mathbf{f}_m\|_2^2=1$. By taking into account that $\mathbf{f}_m$ is the m-th column of $\mathbf{F}$ and that $\|\mathbf{F}\|_{\mathrm{F}}^2=\mathrm{Tr}\,\mathbf{F}\mathbf{F}^{\mathrm{H}}$, the new relaxed constraint can be expressed as

$$\mathrm{Tr}\,\mathbf{F}\mathbf{F}^{\mathrm{H}}=1. \tag{20}$$

Variable Change: Next, we introduce the following variable change

$$\mathbf{X}=\mathbf{F}\mathbf{F}^{\mathrm{H}},\quad \mathrm{rank}\,\mathbf{X}\leq M,\quad \mathbf{X}\succcurlyeq 0, \tag{21}$$

where '$\succcurlyeq 0$' denotes positive semidefinite matrix. The variable change is reversible because $\mathbf{F}$ can always be retrieved from $\mathbf{X}$ by, for instance, a Cholesky decomposition.

Rank Relaxation: The second and final relaxation consists in dropping the rank constraint. Thus, the new optimization problem after including the two relaxations and performing the variable change is

$$\max_{\mathbf{X},t} \quad t \tag{22a}$$

$$\text{s.t.} \quad K_{\mathrm{D}} \left(\dot{\mathbf{a}}_{\mathrm{Tx}}^{i\mathrm{H}} \mathbf{X} \dot{\mathbf{a}}_{\mathrm{Tx}}^{i} - \frac{\left| \mathbf{a}_{\mathrm{Tx}}^{i\mathrm{H}} \mathbf{X} \dot{\mathbf{a}}_{\mathrm{Tx}}^{i} \right|^2}{\mathbf{a}_{\mathrm{Tx}}^{i\mathrm{H}} \mathbf{X} \mathbf{a}_{\mathrm{Tx}}^{i}} \right) \geq t \tag{22b}$$

$$K_{\mathrm{O}}\, \mathbf{a}_{\mathrm{Tx}}^{i\mathrm{H}} \mathbf{X} \mathbf{a}_{\mathrm{Tx}}^{i} \geq t \tag{22c}$$

$$\mathrm{Tr}\, \mathbf{X} = 1 \tag{22d}$$

$$\mathbf{X} \succcurlyeq 0 \tag{22e}$$

for $i = 1, \ldots, S_{\mathrm{Tx}}$. Constraint (22b) can be cast as a second order cone (23b) [27, Chapter 2.3],

$$\max_{\mathbf{X},t} \quad t \tag{23a}$$

$$\left\| \begin{pmatrix} 2\mathbf{a}_{\mathrm{Tx}}^{i\mathrm{H}} \mathbf{X}\, \dot{\mathbf{a}}_{\mathrm{Tx}}^{i} \\ \dot{\mathbf{a}}_{\mathrm{Tx}}^{i\mathrm{H}} \mathbf{X}\, \dot{\mathbf{a}}_{\mathrm{Tx}}^{i} - \frac{t}{K_{\mathrm{D}}} - \mathbf{a}_{\mathrm{Tx}}^{i\mathrm{H}} \mathbf{X}\, \mathbf{a}_{\mathrm{Tx}}^{i} \end{pmatrix} \right\|_2 \leq \dot{\mathbf{a}}_{\mathrm{Tx}}^{i\mathrm{H}} \mathbf{X}\, \dot{\mathbf{a}}_{\mathrm{Tx}}^{i} - \frac{t}{K_{\mathrm{D}}} + \mathbf{a}_{\mathrm{Tx}}^{i\mathrm{H}} \mathbf{X}\, \mathbf{a}_{\mathrm{Tx}}^{i} \tag{23b}$$

$$K_{\mathrm{O}}\, \mathbf{a}_{\mathrm{Tx}}^{i\mathrm{H}} \mathbf{X} \mathbf{a}_{\mathrm{Tx}}^{i} \geq t \tag{23c}$$

$$\mathrm{Tr}\, \mathbf{X} = 1 \tag{23d}$$

$$\mathbf{X} \succcurlyeq 0, \tag{23e}$$

where $\mathbf{X} \in \mathbb{C}^{N_{\mathrm{Tx}} \times N_{\mathrm{Tx}}}$ and $i = 1, \ldots, S_{\mathrm{Tx}}$. Problem (23) is a conic program [28], and consequently convex, because it is composed of linear constraints, second order cones (23b) and a positive semidefinite cone (23e). An advantage of having reformulated our problem as a conic program is that they are well studied in the literature and very efficient solvers exist [29], [30].

Solving Problem 1 requires knowledge of the combining matrix $\mathbf{W}$ since it is needed for computing constants K_{D} and K_{O} (23b), (23c). If the combining matrix of choice $\mathbf{W}$ is more beneficial for AoA estimation than AoD estimation, then the resulting constants, K_{D} and K_{O}, will balance it out by favoring the optimization of the DEB (which is a bound on the AoD accuracy) over the OEB. Vice-versa, if $\mathbf{W}$ is more beneficial for AoD estimation, then the resulting constants will favor the optimization of the OEB. For the solution to Problem 2 simply solve the same problem (23) without constraint (23c). By deleting (23c) and performing the variable change $\frac{t}{K_{\mathrm{D}}} \leftarrow t$, constants K_{O} and K_{D} disappear, making the optimization problem independent of the Rx's array response. To obtain the solution to Problem 3 delete (23b), and following the same reasoning, the optimization problem also becomes independent of the Rx's array response.

C. Recovery of the Precoders

In the previous section, Problem 1 was transformed into a convex problem (23) by

A) approximating the prior ranges with a grid (12), (13),
B) relaxing the feasible set by replacing $\|\mathbf{f}_m\|_2 = M^{-1}$ for $m = 1, \ldots, M$, by $\sum_{m=1}^{M} \|\mathbf{f}_m\|_2^2 = 1$.
C) change of variables $\mathbf{X} = \mathbf{F}\mathbf{F}^{\mathrm{H}}$, $\mathrm{rank}\, \mathbf{X} \leq M$, $\mathbf{X} \succcurlyeq 0$,

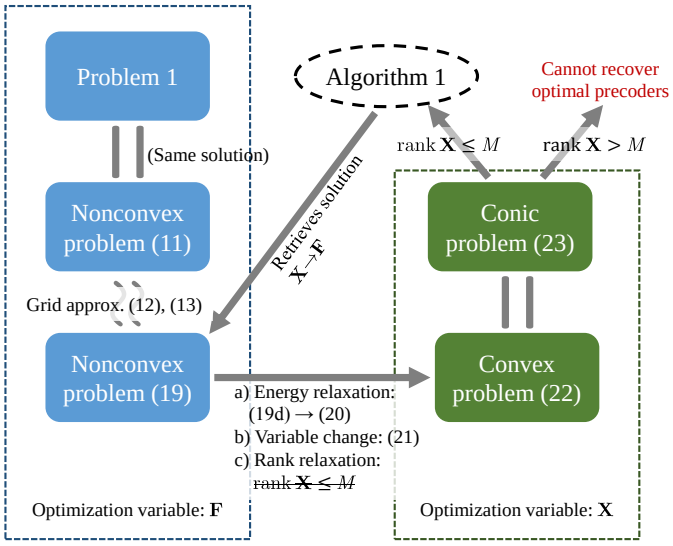


Fig. 1. Flow chart of the multiple steps for obtaining a tractable optimization problem. The nonconvex and convex reformulations are colored in blue and green, respectively. The optimal precoders are obtained by optimizing problem (23) and executing Algorithm 1. We have verified numerically that the grid approximation is accurate when sufficiently dense.

D) relaxing the feasible set by dropping $\mathrm{rank}\, \mathbf{X} \leq M$.

We abuse the notation and call $\mathbf{X}$ the global optimum of (23) instead of the optimizing variable. It turns out that the grid approximation (A) is very accurate when made dense enough. Moreover, the trace constraint in (23d) induces a low-rank solution [31], [32]. Hence, if the number of precoders M is sufficiently large, then it will occur that $\mathrm{rank}\, \mathbf{X} \leq M$, and the relaxation (D) will not affect the solution. Thus, in the remainder of this section we discuss the effect of (B) and (C) assuming (A) is exact and (D) is satisfied. These two later assumptions will be validated numerically in Section IV-A.

Given $\mathbf{X}$, the variable change (C) can be reversed by performing a Cholesky decomposition [33]. However, we are interested in a decomposition from $\mathbf{X}$ to $\mathbf{F}$ that also satisfies $\|\mathbf{f}_m\|_2 = M^{-1}$ for $m = 1, \ldots, M$ (B). If such decomposition existed, then $\mathbf{F}$ could be recovered from $\mathbf{X}$ and satisfy all the constraints to the original Problem 1 in its hypograph form (11). In summary, we need to find $\mathbf{F}$ such that

C1) $\mathbf{F}\mathbf{F}^{\mathrm{H}} = \mathbf{X}$
C2) $\mathrm{diag}\left(\mathbf{F}^{\mathrm{H}}\mathbf{F}\right) = M^{-1}\mathbf{1}$

where $\mathrm{diag}(\mathbf{F}^{\mathrm{H}}\mathbf{F})$ denotes a vertical vector stacking the entries on the main diagonal of $\mathbf{F}^{\mathrm{H}}\mathbf{F}$. Define $R \triangleq \mathrm{rank}\, \mathbf{X} \leq M$ where the latter inequality follows from assumption (D). From C1 we infer that $\mathrm{rank}\, \mathbf{F} = R$. Sufficient conditions for satisfying C1 and C2 can be obtained by expanding $\mathbf{F}$ and $\mathbf{X}$ by their *compact* singular value decomposition (SVD) and eigendecomposition (ED), respectively,

$$\mathbf{F} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{H}} \tag{24}$$

$$\mathbf{X} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{\mathrm{H}}, \tag{25}$$

where $\mathbf{U} = [\mathbf{u}_1 \cdots \mathbf{u}_R] \in \mathbb{C}^{N_{\mathrm{Tx}} \times R}$ is a matrix of left singular vectors, $\mathbf{V} = [\mathbf{v}_1 \cdots \mathbf{v}_R] \in \mathbb{C}^{R \times R}$ is a matrix of right singular vectors, $\mathbf{Q} = [\mathbf{q}_1 \cdots \mathbf{q}_R] \in \mathbb{C}^{N_{\mathrm{Tx}} \times R}$ is a matrix of eigenvectors, and $\boldsymbol{\Sigma}$ and $\boldsymbol{\Lambda}$ are diagonal matrices whose

diagonal elements are the singular values $\mu_1 \cdots \mu_R$ and the eigenvalues $\lambda_1 \cdots \lambda_R$, respectively.

Substituting $\mathbf{F}$ and $\mathbf{X}$ in C1 with (24) and (25), respectively, leads to a new form of C1, $\mathbf{U}\boldsymbol{\Sigma}^2\mathbf{U}^{\mathrm{H}} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{\mathrm{H}}$. This equation is solved by matching the eigenvectors and eigenvalues on both sides of the equation, i.e.,

$$\mathbf{u}_m = \mathbf{q}_m \tag{26}$$

$$\mu_m = \sqrt{\lambda_m}, \tag{27}$$

for $m = 1, \ldots, R$. Similarly, substituting (24) and (25) into C2 yields the new condition

$$\operatorname{diag}\left(\mathbf{V}\boldsymbol{\Sigma}^2\mathbf{V}^{\mathrm{H}}\right) = M^{-1}\mathbf{1}. \tag{28}$$

The (diagonal) elements of $\boldsymbol{\Sigma}$ have been fixed by (27), hence, only $\mathbf{V}$ can be adjusted to meet the new condition. Notice that $\mathbf{Z} \triangleq \mathbf{V}\boldsymbol{\Sigma}^2\mathbf{V}^{\mathrm{H}}$ may be regarded as the eigendecomposition of a matrix whose non-zero eigenvalues are $\mu_1^2 \ldots, \mu_R^2$ and their corresponding eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_R$. Consequently, condition (28) is equivalent to the existence of a matrix $\mathbf{Z}$ with prescribed non-zero eigenvalues $\mu_1^2 \ldots, \mu_R^2$ and M^{-1} in all entries of the main diagonal. Fortunately, for the particular case of constant diagonal, the Schur-Horn theorem [34, B.1. in p. 218] states that $\mathbf{Z}$ always exists, and it can be obtained by the Bendel-Mickey algorithm [35], [36].

In summary, the different pieces of a valid $\mathbf{F}$ satisfying C1 and C2 can be obtained as follows: its left singular vectors by (26), its singular values by (27) and its right singular vectors by finding a matrix satisfying (28) with the Bendel-Mickey algorithm and computing its eigenvectors. All these steps are summarized in Algorithm 1.

Algorithm 1 Retrieval of the optimal precoders $\mathbf{F}$ from $\mathbf{X}$

inputs: $\mathbf{X}, M$
outputs: $\mathbf{F}$

1: compute the *compact* ED of $\mathbf{X} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{\mathrm{H}}$ where the diagonal elements of $\boldsymbol{\Lambda}$ are $\lambda_1, \ldots, \lambda_R$ and $R = \operatorname{rank}\mathbf{X}$
2: execute the Bendel-Mickey algorithm [35], [36] to find an Hermitian matrix $\mathbf{Z}$ of size $M \times M$ with non-zero eigenvalues $\lambda_1, \ldots, \lambda_R$ and all diagonal elements equal to M^{-1}
3: compute the ED of $\mathbf{Z}$ and denote $\mathbf{v}_1, \ldots, \mathbf{v}_R$ the eigenvectors associated to eigenvalues $\lambda_1, \ldots, \lambda_R$
4: set $\mathbf{F} = \mathbf{Q}\boldsymbol{\Lambda}^{\frac{1}{2}}[\mathbf{v}_1 \cdots \mathbf{v}_R]^{\mathrm{H}}$

Remark: The solution to (28) is in general not unique, and so the Bendel-Mickey algorithm is designed to return a random solution in the solution set. Consequently, the solution to Problems 1–3 is random as well.

Examples: Given a set of precoders $\mathbf{F} = [\mathbf{f}_1 \cdots \mathbf{f}_M]$, define the transmit array gain and the aggregated gain (which is the equivalent gain if the energy of the M training sequences are coherently combined at the Rx) as

$$g(\theta, \mathbf{f}_m) \triangleq 10\log_{10}\frac{\left|\mathbf{a}_{\mathrm{Tx}}^{\mathrm{H}}(\theta)\mathbf{f}_m\right|^2}{P_{\mathrm{av}}/(4\pi)} \quad [\mathrm{dB}] \tag{29}$$

$$g_{\mathrm{T}}(\theta, \mathbf{F}) \triangleq 10\log_{10}\frac{\sum_{m=1}^{M}\left|\mathbf{a}_{\mathrm{Tx}}^{\mathrm{H}}(\theta)\mathbf{f}_m\right|^2}{P_{\mathrm{av}}/(4\pi)} \quad [\mathrm{dB}], \tag{30}$$

where $P_{\mathrm{av}}/(4\pi) = M^{-1}\sum_{m=1}^{M} P_m/(4\pi)$ is the average emitted energy per precoder normalized by the radiated power of an isotropic antenna, and P_m is the emitted energy by the m-th precoder computed according to [37, Eq. (33)].

Fig. 2 plots the radiation patterns (gain vs. azimuth) of the AoD-AoA-optimal precoders. The yellow zone in the figures represents the prior range on the AoD. Furthermore, we define the term 'worst aggregated gain', as the lowest aggregated gain within the prior range. As expected from the interpretation in Appendix B, where it was argued that the OEB is optimized when the aggregated gain is maximized, in Fig. 2c the AoA-optimal precoders' worst aggregated gain is 14.1 dB, which is the highest among the three subfigures. The AoD-optimal precoders in Fig. 2b show a large number of ripples, which is also consistent with the interpretation in Section III. The intuitive explanation is that the Rx identifies the AoD by observing the changes of received signal strength and phase changes for the M training sequences. Thus, in order to increase the AoD estimation accuracy, the Tx's precoders gain and phase for closely spaced AoD must be as different as possible. The worst aggregated gain of the AoD-optimal and AoD-AoA-optimal precoders are 11.6 dB and 13.6 dB, respectively. The AoD-AoA-optimal precoders offer a good trade-off in terms of aggregated gain and 'ripples'. Because the AoA estimation accuracy is optimized by maximizing the aggregated gain towards the Rx, the AoA-optimal precoders spill much less energy outside the prior Tx range of angles than the AoD-optimal precoders. To explicitly reduce energy transmitted outside the prior Tx range, Section III-D2, introduces an extra constraint that when added to problem (23) enforces a certain out-of-range attenuation.

D. Additional Constraints

The proposed precoders are purely designed to minimize the CRB of the AoD and/or AoA. However, in practice some additional constraints may be desired, such as null steering towards certain directions in order to mitigate multiuser interference or cancel certain paths. Next, it is shown how some of these constraints can easily be added to the original optimization problems.

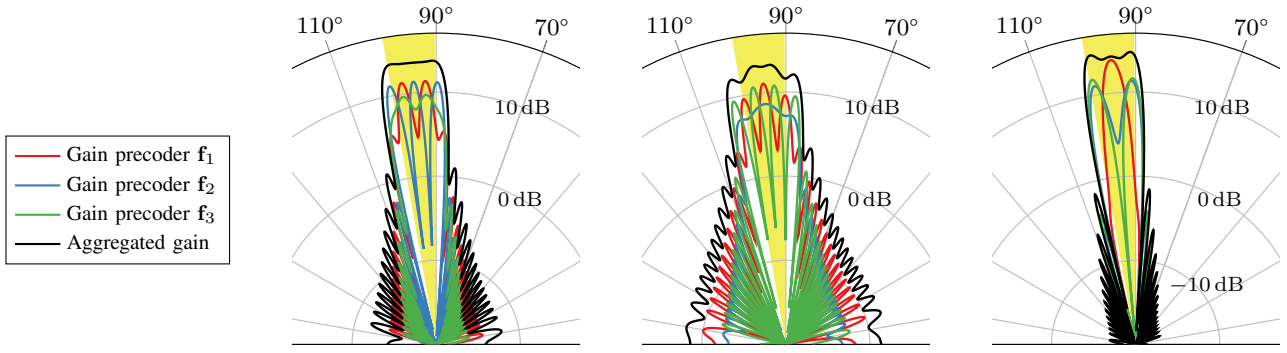
1) Identifiability of the AoD and AoA: A necessary condition for correctly estimating the AoD and/or AoA is that they are identifiable. Otherwise, even in the absence of noise, we may suffer from large estimation errors due to signal ambiguities. The AoA and AoD are identifiable provided that the received signals (3) in absence of noise are different

$$\alpha\,\mathbf{b}_{\mathrm{Rx}}(\phi)\mathbf{a}_{\mathrm{Tx}}^{\mathrm{H}}(\theta)\mathbf{F} \neq \alpha'\,\mathbf{b}_{\mathrm{Rx}}(\phi')\mathbf{a}_{\mathrm{Tx}}^{\mathrm{H}}(\theta')\mathbf{F}, \tag{31}$$

for all possible $\theta \neq \theta'$ in $\mathcal{R}_{\mathrm{Tx}}$, $\phi \neq \phi'$ in $\mathcal{R}_{\mathrm{Rx}}$, or $\alpha \neq \alpha'$. Since the signals are rank 1 matrices, by applying the substitution $\mathbf{b}_{\mathrm{Rx}}(\phi) = \mathbf{W}^{\mathrm{H}}\mathbf{a}_{\mathrm{Rx}}(\phi)$ the condition decouples into pair-wise linear independence between steering vectors at the Rx and Tx:

$$\mathbf{W}^{\mathrm{H}}\mathbf{a}_{\mathrm{Rx}}(\phi) \neq \beta\,\mathbf{W}^{\mathrm{H}}\mathbf{a}_{\mathrm{Rx}}(\phi') \tag{32}$$

$$\mathbf{F}^{\mathrm{H}}\mathbf{a}_{\mathrm{Tx}}(\theta) \neq \gamma\,\mathbf{F}^{\mathrm{H}}\mathbf{a}_{\mathrm{Tx}}(\theta') \tag{33}$$



(a) AoD-AoA-optimal precoders (Problem 1). The rank of $\mathbf{X}$ is 2.

(b) AoD-optimal precoders (Problem 2). The rank of $\mathbf{X}$ is 2.

(c) AoA-optimal precoders (Problem 3). The rank of $\mathbf{X}$ is 3.

Fig. 2. Array gain of the precoders vs. azimuth for all M precoders. The Tx and Rx are equipped with 30-antenna half-wavelength inter antenna spacing uniform linear array (ULA) [38]. The number of training sequences is $M = 3$. The AoD (same for the AoA for simplicity) is known a priori to lie in the range $[90°, 100°]$, indicated by the yellow shaded area. According to Section III-C, the optimality of these precoders is ensured because for all cases $M \geq \operatorname{rank} \mathbf{X}$.

for all complex β and γ, different than zero. From these expressions we can conclude that the combining matrix impacts the identifiability of the AoA, and that the precoding matrix impacts the identifiability of the AoD only. For the particular case of a fully digital Rx ($\mathbf{W} = \mathbf{I}$), the AoA is identifiable for most common types of arrays [39], [40], whereas for arbitrary combiners, the conditions for which the AoA is identifiable are not known in the literature to the best of the authors' knowledge (and also beyond the scope of the current study).

Since this work deals with the design of the precoders, we turn to condition (33) which imposes pair-wise independence and can also be expressed as $|\mathbf{a}_{\mathrm{Tx}}^{\mathrm{H}}(\theta)\mathbf{F}\,\mathbf{F}^{\mathrm{H}}\,\mathbf{a}_{\mathrm{Tx}}(\theta')|^2 \|\mathbf{F}^{\mathrm{H}}\,\mathbf{a}_{\mathrm{Tx}}(\theta)\|_2^{-2} \|\mathbf{F}^{\mathrm{H}}\,\mathbf{a}_{\mathrm{Tx}}(\theta')\|_2^{-2} < 1$. Thus, in order to ensure that the AoD is identifiable we propose to add the following constraints to problem (23):

$$\frac{\left|\mathbf{a}_{\mathrm{Tx}}^{\mathrm{H}}(\vartheta^i)\mathbf{F}\,\mathbf{F}^{\mathrm{H}}\,\mathbf{a}_{\mathrm{Tx}}(\vartheta^{i'})\right|^2}{\|\mathbf{F}^{\mathrm{H}}\,\mathbf{a}_{\mathrm{Tx}}(\vartheta^i)\|_2^2\,\|\mathbf{F}^{\mathrm{H}}\,\mathbf{a}_{\mathrm{Tx}}(\vartheta^{i'})\|_2^2} \leq \rho \tag{34}$$

for all i, i' such that such that $\vartheta^i - \vartheta^{i'} > D (\mathrm{mod}\, 2\pi)$, where $0 \leq \rho < 1$ and D is the angular resolution of the array[4]. By performing the variable change $\mathbf{X} = \mathbf{F}\mathbf{F}^{\mathrm{H}}$ and some algebraic manipulations [27], (34) can be expressed as a second order cone (convex constraint):

$$\left\| \begin{pmatrix} 2\,\mathbf{a}_{\mathrm{Tx}}^{i\mathrm{H}}\mathbf{X}\,\mathbf{a}_{\mathrm{Tx}}^{i'} \\ \sqrt{\rho}\,\mathbf{a}_{\mathrm{Tx}}^{i\mathrm{H}}\mathbf{X}\,\mathbf{a}_{\mathrm{Tx}}^{i} - \sqrt{\rho}\,\mathbf{a}_{\mathrm{Tx}}^{i'\mathrm{H}}\mathbf{X}\,\mathbf{a}_{\mathrm{Tx}}^{i'} \end{pmatrix} \right\|_2 \leq \sqrt{\rho}\,\mathbf{a}_{\mathrm{Tx}}^{i\mathrm{H}}\mathbf{X}\,\mathbf{a}_{\mathrm{Tx}}^{i} + \sqrt{\rho}\,\mathbf{a}_{\mathrm{Tx}}^{i'\mathrm{H}}\mathbf{X}\,\mathbf{a}_{\mathrm{Tx}}^{i'}, \tag{35}$$

where the shortened notation (14) was used. To ensure the AoD can be correctly estimated, this condition must be added to optimization problem (23) for the AoD- and AoD-AoA-optimal precoders.

[4]The resolution of the array is typically related to the antenna aperture and is different for each array configuration. In practice, one could start with an approximate value and then tweak this parameter. For instance, for an N-antennas half-wavelength inter-antenna spaced uniform circular array (UCA) [38], the angle resolution has been checked numerically to be well approximately by $1.6 \sin(\pi/N)$ [radians].

2) Out-of-Range Attenuation: From Fig. 2, we can observe that the precoders radiate some non-negligible energy in directions outside the prior range $\mathcal{R}_{\mathrm{Tx}}$, even though an initial assumption was that the precoders do not illuminate any paths other than the desired one. In addition, there may be some operational constraints such as low sidelobe ratio [41] for improved inter-user interference or the placement of nulls in certain areas of the beampattern.

Let $\{\pi^q\}_{q=1}^{S_{\mathrm{a}}}$ be the AoDs for which we wish to enforce a lower transmitting power, and recall $\{\vartheta^i\}_{i=1}^{S_{\mathrm{Tx}}}$ is the grid of angles within the prior range (12). The total transmitted energy in direction π^q over the M training sequences is $\|\mathbf{F}^{\mathrm{H}}\mathbf{a}_{\mathrm{Tx}}(\pi^q)\|_2^2$. Let $A_q < 1$ be the desired attenuation factor, then, we propose to add the following constraint to problem (23),

$$\left\|\mathbf{F}^{\mathrm{H}}\mathbf{a}_{\mathrm{Tx}}\left(\pi^q\right)\right\|_2^2 \leq A_q \left\|\mathbf{F}^{\mathrm{H}}\mathbf{a}_{\mathrm{Tx}}\left(\vartheta^i\right)\right\|_2^2 \tag{36}$$

for $q = 1, \ldots, S_{\mathrm{a}}$ and $i = 1, \ldots, S_{\mathrm{Tx}}$, which can be transformed to linear constraints after performing the variable change $\mathbf{X} = \mathbf{F}\mathbf{F}^{\mathrm{H}}$,

$$\mathbf{a}_{\mathrm{Tx}}^{\mathrm{H}}\left(\pi^q\right)\mathbf{X}\,\mathbf{a}_{\mathrm{Tx}}\left(\pi^q\right) \leq A_q\,\mathbf{a}_{\mathrm{Tx}}^{\mathrm{H}}\left(\vartheta^i\right)\mathbf{X}\,\mathbf{a}_{\mathrm{Tx}}\left(\vartheta^i\right). \tag{37}$$

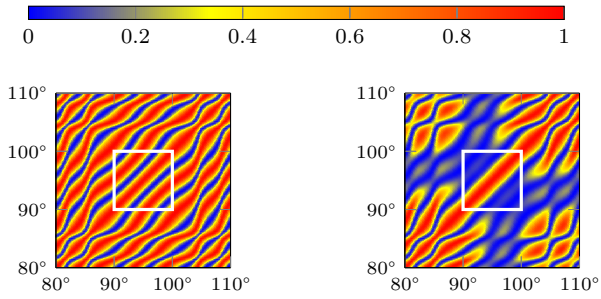
These $S_{\mathrm{a}}S_{\mathrm{Tx}}$ constraints can be reduced to $S_{\mathrm{a}} + S_{\mathrm{Tx}}$ by incorporating a dummy variable z,

$$\mathbf{a}_{\mathrm{Tx}}^{\mathrm{H}}\left(\pi^q\right)\mathbf{X}\,\mathbf{a}_{\mathrm{Tx}}\left(\pi^q\right) \leq A_q\,z \qquad q = 1, \ldots, S_{\mathrm{a}} \tag{38}$$

$$\mathbf{a}_{\mathrm{Tx}}^{\mathrm{H}}\left(\vartheta^i\right)\mathbf{X}\,\mathbf{a}_{\mathrm{Tx}}\left(\vartheta^i\right) \geq z \qquad i = 1, \ldots, S_{\mathrm{Tx}}. \tag{39}$$

Examples: Fig. 3 plots the left hand side of (34) when the precoders are obtained with or without the identifiability constraint. Ideally, only the anti-diagonal across the white square should be red as is the case in the right figure. The left figure, has two red stripes, and therefore, there are pairs of AoDs within the range $\mathcal{R}_{\mathrm{Tx}}$ which result in the same signals, and consequently are not identifiable even in the absence of noise.

In Fig. 4 two sets of AoD-AoA-optimal precoders are generated. The first set of precoders has no constraint on the aggregated gain (30) towards directions outside of the range of interest $\mathcal{R}_{\mathrm{Tx}}$. The second set of precoders imposes a 20 dB



(a) Without identifiability constraints (34).

(b) With identifiability constraints (34) and $\rho = 0.1$.

Fig. 3. Colormap of left hand side of (34) versus θ and θ' when employing AoD-AoA-optimal precoders. The Tx and the Rx have, each, a 30-antennas uniform linear array (ULA) [38]. The number of training sequences $M = 4$ and the rank of $\mathbf{X}$ is 4. It is known a priori that the AoD lies in the interal $\mathcal{R}_{\mathrm{Tx}} = [90°, 100°]$. The white square indicates the region $\mathcal{R}_{\mathrm{Tx}} \times \mathcal{R}_{\mathrm{Tx}}$.

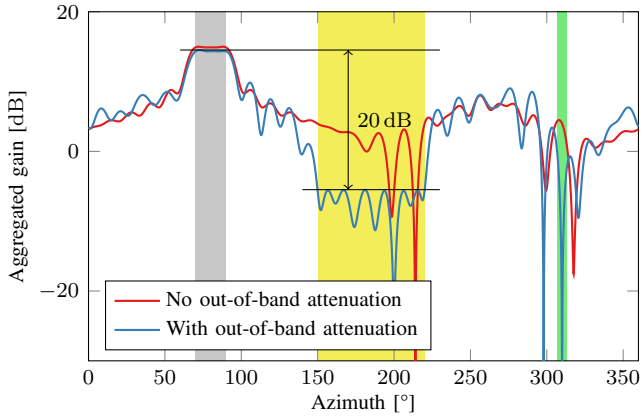


Fig. 4. Aggregated gain (30) of the AoD-AoA-optimal precoders vs. azimuth. The Tx and Rx are equipped with 30-antenna half-wavelength inter antenna spacing UCA. The number of training sequences is $M = 3$ and the rank $\mathbf{X}$ is 2, ensuring the precoders are optimal. The AoD (same for the AoA for simplicity) is known a priori to lie in the range $[70°, 90°]$, indicated by the grey shaded bin. Two different sets of precoders are used. First set is the optimum solution to Problem 1. The second set is the optimum solution to Problem 1 while imposing a 20 dB attenuation in the range [150°,220°] indicated by the yellow shaded bin, and a null at 310° indicated by the green shaded bin.

attenuation in the range [150°,220°] and a null at 310°. Note that the aggregated gain within the range is virtually the same for both sets of precoders.

IV. Numerical Results

The performance of the optimal precoders is illustrated next. To this end, we define the following figures of merit which are in fact the square roots of the objective functions of Problems 1–3, respectively,

$$\text{Worst case rEB} = \max_{(\theta,\phi)\in\mathcal{R}_{\mathrm{Tx}}\times\mathcal{R}_{\mathrm{Rx}}} \sqrt{\mathrm{EB}} \tag{40}$$

$$\text{Worst case rDEB} = \max_{(\theta,\phi)\in\mathcal{R}_{\mathrm{Tx}}\times\mathcal{R}_{\mathrm{Rx}}} \sqrt{\mathrm{DEB}} \tag{41}$$

$$\text{Worst case rOEB} = \max_{(\theta,\phi)\in\mathcal{R}_{\mathrm{Tx}}\times\mathcal{R}_{\mathrm{Rx}}} \sqrt{\mathrm{OEB}}, \tag{42}$$

where $\mathrm{EB} = \max\{\mathrm{DEB}, \mathrm{OEB}\}$ and the 'r' denotes square root. The following experiments simulate the signal model (3)

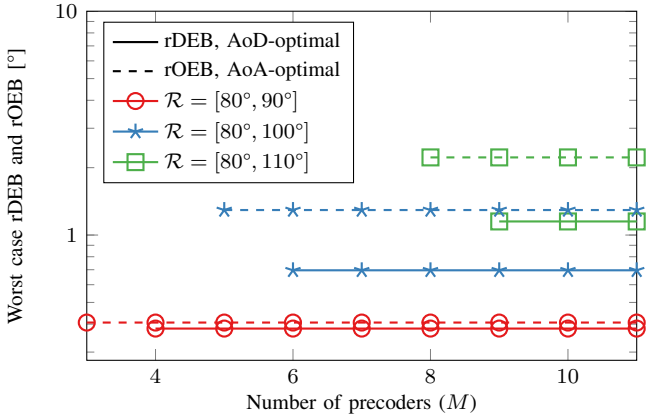


Fig. 5. Lower bounds on the estimation accuracy vs. the number of different precoders used at the Tx. The total training time is fixed, and consequently the duration of each training sequence (precoder) decreases inversely proportional to M. Both, the Tx and the Rx, equip a ULA. For simplicity, the prior range on the AoD and AoA are set equal, i.e., $\mathcal{R} \triangleq \mathcal{R}_{\mathrm{Rx}} = \mathcal{R}_{\mathrm{Tx}}$.

and use the above metrics to evaluate the performance of the proposed set of precoders. We assume 30-antenna uniform linear arrays (ULA) [38] with an inter-antenna spacing of half wavelength at the Tx and Rx. The Tx has a fully digital architecture and the Rx equips an array of subarrays [25] with 5 RF chains. Since this work did not deal with the design of the combining matrix, we implement a simple strategy which consists in associating each RF chain to a combiner with maximum gain towards equispaced directions within the prior range of angles at the Rx. This is achieved by matching each combining vector to the array response in that direction, i.e.,

$$\mathbf{W} = \begin{pmatrix} \mathbf{a}_{\mathrm{Rx}}(\tilde{\phi}_1) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{a}_{\mathrm{Rx}}(\tilde{\phi}_L) \end{pmatrix}, \tag{43}$$

where $\tilde{\phi}_1, \ldots, \tilde{\phi}_L$ are the equispaced directions within the interval $\mathcal{R}_{\mathrm{Rx}}$, and $\mathbf{a}_{\mathrm{Rx}}(\phi) \in \mathbb{C}^{\frac{N_{\mathrm{Rx}}}{L}\times 1}$ because each RF chain is routed to $\frac{N_{\mathrm{Rx}}}{L}$ antennas only. The SNR (6) is fixed to −5 dB. The number of transmitted training sequences, each precoded differently, is $M = 5$. Parameter ρ defined in Section III-D1, which ensures identifiability of the AoD, is set to $\rho = 0.6$ for the AoD- and AoD-AoA-optimal precoders. The prior ranges on the AoD and AoA are, both, $\mathcal{R} \triangleq \mathcal{R}_{\mathrm{Tx}} = \mathcal{R}_{\mathrm{Rx}} = [90°, 100°]$.

A. Impact of the Number of Precoders

The convex reformulation in Section III-B of our original Problems 1–3 led to optimization problem (23) which does not depend on M. Thus, the optimal DEB and/or OEB do not change as a function of the number of precoders M provided that a feasible precoding matrix exists, i.e., $\operatorname{rank}\mathbf{X} \leq M$ as explained in Section III-C. To showcase this, Fig. 5 plots the worst case (w.c.) rDEB and rOEB for the AoD- and AoA-optimal precoders, respectively, versus the number of precoders for different widths of the prior ranges. For the cases where $M < \operatorname{rank}\mathbf{X}$, no values of the w.c. rDEB

and rOEB are plotted because the optimal precoders are unknown. The smallest value of M in every curve is the rank of $\mathbf{X}$ (e.g., $\operatorname{rank}\mathbf{X} = 5$ for the AoA-optimal precoders with $\mathcal{R} = [80°, 100°]$ in Fig. 5). Therefore, the rank of $\mathbf{X}$ can be regarded as the minimum precoding diversity M necessary for achieving optimal estimation accuracy. Increasing M beyond this optimum value does not translate into better estimation accuracy[5]. Naturally, the required precoding diversity increases with the size of the ranges of angles $\mathcal{R}$. The accuracy difference in terms of rDEB and rOEB is due to the fact that the Rx's hybrid array projects the received signal from an N_{Rx}-dimensional space to an L-dimensional subspace through a non-optimal combining matrix, hence, lowering the quality of the observations. While not plotted here, we have observed that the rDEB and rOEB have similar accuracy when the Rx array is fully digital.

B. Estimators and Ambiguities

Let $\hat{\theta}$ be an estimate on the AoD. Then, the worst case root mean square error on the AoD is defined as

$$\text{Worst case rMSE} = \max_{(\theta,\phi)\in\mathcal{R}_{\mathrm{Tx}}\times\mathcal{R}_{\mathrm{Rx}}} \sqrt{\mathbb{E}\left(\hat{\theta}-\theta\right)^2}. \quad (44)$$

If the MSE is tight to the DEB, such as is the case of the ML estimator, then the w.c. rMSE is tight to the w.c. rDEB. To evaluate its tightness and the effect of the parameter ρ (introduced in Section III-D1), we run a Monte Carlo simulation with 100 random experiments. For each experiment we generate a random received signal (3) where the channel phase is distributed randomly over the interval $[0, 2\pi]$, the noise samples are drawn from a complex Gaussian random variable with variance σ^2, and the AoD/AoA are drawn from a uniform distribution over the range $\mathcal{R} = [90°, 100°]$. The channel gain and training sequence energy are fixed to $|\alpha| = \|\mathbf{s}\|_2 = 1$, and consequently, in order to obtain a -5 dB SNR at the Rx, the noise variance is fixed to $\sigma^2 = \mathrm{SNR}^{-1}$.

Fig. 6 plots the w.c. rMSE for the MLE[6] for different values of ρ. As expected, the MLE's w.c. rMSE converges to the w.c. rDEB for a sufficient large SNR. For decreasing values of ρ, the SNR threshold (which is the SNR value at which the estimators' accuracy matches the lower bound) is shifted to the left. For $\rho = 1$, the identifiability constraint is removed and the MLE fails to always correctly estimate the AoD for any given SNR. Coarsely speaking, ρ controls the degree of maximum *similarity* between the received signals for any two pair of AoD, where $\rho = 0$ imposes signal orthogonality and $\rho = 1$ removes the constraint. Thus, as ρ decreases from 1 to 0, the received signals for any two AoD become less similar and more resilient to noise.

[5]Note that Problems 1–3 are set up such that the total transmitted energy is constant regardless of M. If the transmitted energy scaled up with the number of precoders, certainly the accuracy would improve.

[6]The MLE requires maximizing the likelihood function (45) over the unknown parameters (α, θ, ϕ). The ML estimate of the channel gain requires solving a least squares whose solution can be computed analytically as a function of θ and ϕ. Then, the ML estimate of the channel gain is plugged back to the log-likelihood function, and perform en exhaustive search over the two remaining variables (θ, ϕ) using a discrete grid.

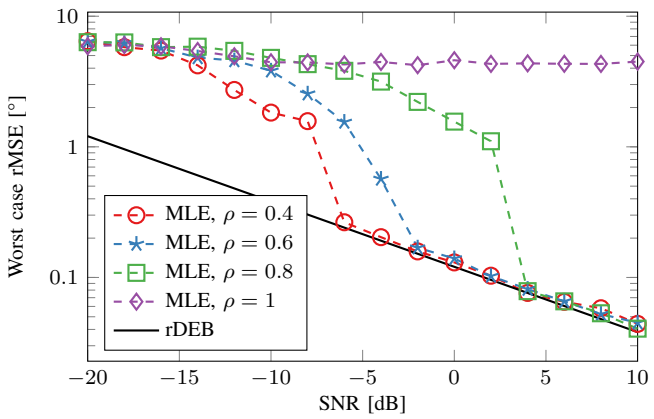


Fig. 6. AoD estimation accuracy vs. SNR for different values of ρ (defined in Section III-D1). $\mathcal{R}_{\mathrm{Rx}} = \mathcal{R}_{\mathrm{Tx}} = [90°, 100°]$.

C. Comparison with Traditional Beams

Traditionally, in mmWave channel estimation, the precoders and combiners are designed such that they steer energy towards few directions. The channel is sounded sequentially in time with all possible pairs of precoders and combiners, and by detecting the pair leading to the largest power at the Rx, one can estimate the AoD and/or AoA with beamwidth accuracy. Its widespread use is due to the simplicity of the approach, despite not necessary being optimal in the sense of minimizing the mean error. Two of the most common approaches are sector beams [6], [16], [26] and maximum-gain beams [22], [42], [43], which we refer to "sectors" and "beams", respectively, for short. Since this work does not deal with the design of optimal combiners, we compare the proposed optimal precoders with sectors and beams at the Tx only. Sectors split the ranges $\mathcal{R}_{\mathrm{Tx}}$, evenly, into M subregions; each sector has a large constant array gain within one subregion and low array gain outside. Beams maximize the energy towards each direction without seeking to create a flat beampattern within each subrange. To design the sectors we use classical tools from filter design [44, Chapter 1.5.1], whereas the beams are designed by matching the precoder to the steering vector in that direction.

Fig 7 plots the w.c. rDEB for the AoD-optimal precoders, beams and sectors. As outlined in Section IV-A, the accuracy of the AoD-optimal precoders remains constant for any M larger than $\operatorname{rank}\mathbf{X}$, which in this particular case is 4. The AoD-optimal precoders is two to four times more accurate than the beams and sectors, respectively.

D. AoA Versus AoD Estimation

A common belief is that when estimating the angle of a path at an array, operating as a receiver and estimating the AoA of the incoming path is more precise than operating as a transmitter and estimating the AoD of the outgoing path. For instance, in [45] the authors argue that if a base station has more antennas than the user, then for estimating the direction from the base station to the user, uplink AoA estimation is more accurate than downlink AoD estimation. While it may be true for non-optimal precoders, we show that if both arrays

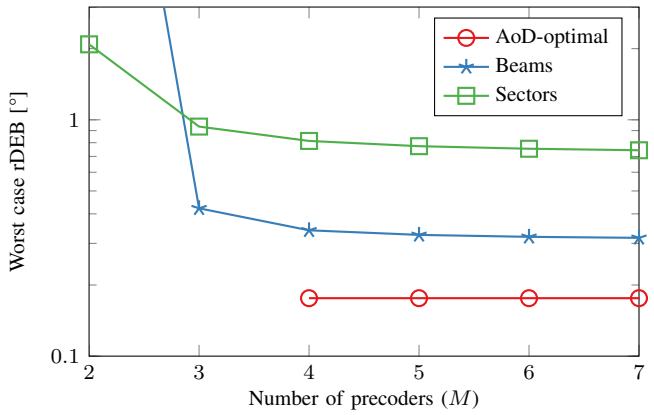


Fig. 7. Performance comparison between AoD-optimal precoders, beams and sectors. $\mathcal{R}_{\mathrm{Rx}} = \mathcal{R}_{\mathrm{Tx}} = [90°, 100°]$.

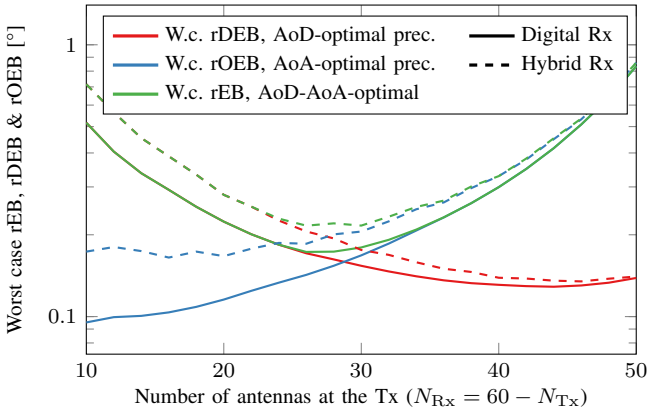


Fig. 8. Comparison between AoD and AoA estimation. The sum of the number of antennas at the Tx and Rx are kept constant to 60 antennas; so as the number of Tx antennas increases, the number of Rx antennas decreases, and vice-versa. $\mathcal{R}_{\mathrm{Rx}} = \mathcal{R}_{\mathrm{Tx}} = [90°, 100°]$.

are fully digital and the Tx is performing optimal precoding, both approaches yield approximately the same accuracy.

Fig. 8 plots the bounds on the AoD/AoA estimation accuracy (40)–(42) versus the number of antennas at the Tx and Rx, assuming the total number of antennas is 60. The goal is to observe the effect of accumulating more antennas at the Tx or Rx on the estimation accuracy of the AoD and AoA. For a fair comparison, we begin by assuming that both arrays have fully digital architectures (represented as solid lines in the figure). For the case where the Tx and Rx have the same number of antennas (i.e., $N_{\mathrm{Tx}} = N_{\mathrm{Rx}} = 30$), the accuracy is very similar whether estimating the AoD or AoA ($\sim 0.2°$). However, for the two extreme cases where the Rx has 50 antennas and the Tx only 10 ($\sim 0.1°$ at the left side of the figure), the AoA estimation accuracy slightly outperforms the AoD accuracy for the case the Tx has 50 antennas and the Rx only 10 ($\sim 0.2°$ at the right side of the figure). Jumping to the case where the Rx equips a hybrid array with 5 RF chains, we notice that the AoD and AoA estimation accuracy worsens, in particular for the AoA. This loss of accuracy caused by the hybrid Rx is expected because, mathematically, the Rx is compressing the received signal onto

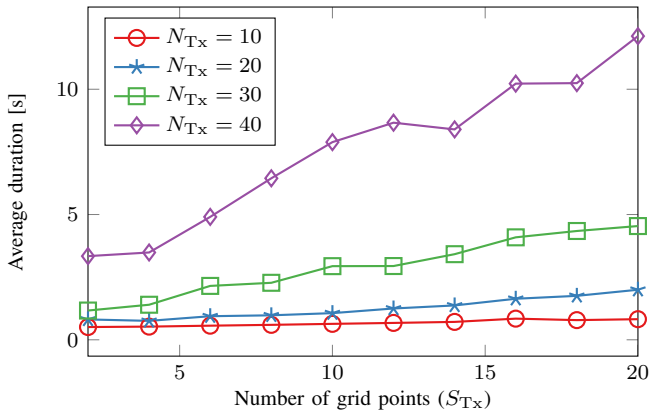


Fig. 9. Average duration for obtaining the AoD-AoA-optimal precoders which involves solving problem (23) and running Algorithm 1. The two main parameters affecting the number of variables and constraints are the number of antennas at the Tx, N_{Tx}, and the number of grid points S_{Tx}.

a L-dimensional subspace. Thus, in general, the AoD and AoA estimation accuracy will differ. However, in the particular case where both arrays have the same number of antennas, are fully digital, and the Tx performs optimal precoding, operating as a Tx and estimating the AoD, or operating as a Rx and estimating the AoA, result in similar accuracy.

E. Complexity Analysis

Obtaining the optimal precoders requires solving problem (23). This is a conic (hence, convex) optimization problem [46] because the constraints are composed of second-order cones (23b) and a positive semidefinite cone (23e). Conic problems can be efficiently solved by the interior-points method [30]. Second-order cone programs (SOCPs) and semidefinite programs (SDPs) are particular instances of conic problems where all constraints are second-order cones or positive semidefinite cones, respectively. These categories are important because, while both types of problems have polynomial complexity, for equal number of variables and constraints, SDPs are generally substantially more computationally complex than SOCPs. By leveraging the Schur complement, the proposed conic problem (23) could be reformulated as an SDP by transforming all second-order cone into positive semidefinite cones. However, it would not be wise in terms of computational complexity. Unfortunately, to the best of the author's knowledge, the computational complexity of conic problems with mixed cones is not well understood in the literature, and an analysis on its SDP form would lead to too pessimistic bounds. Thus, we resolve to a numerical analysis on the computational complexity.

Fig. 9 plots the average duration of solving problem (23) and executing Algorithm 1 with MATLAB in a laptop computer with 2 GHz clock speed. From the figure we observe that the duration appears to be linear with the number of grid points and superlinear with the number of antennas.

V. Conclusions

This paper dealt with the design of precoders for mmWave communication, optimized for estimating the AoD and AoA

under a given uncertainty range. Our focus was on a single propagation path, assuming that such a path has been identified for precoding and combining so that weaker paths can be ignored. The design is based on minimizing the worst-case CRB of the AoD and AoA, which is a tight lower bound to the variance of MLE at medium-to-high SNRs. Through a convex reformulation, optimal precoders are recovered, leading to a two-fold or larger improvement on the estimation accuracy with respect to traditional schemes. We found that beyond a certain number of precoders at the Tx, the CRB on the AoD/AoA does not improve and that the minimum number of precoders for optimal performance is a side-product of the optimization procedure. Finally, numerical evidence shows an array can estimate the direction of a path in transmit (AoD) or receive (AoA) mode with similar accuracy when performing optimal precoding.

Appendix A
Fisher Information Matrix and Cramér-Rao Bound

The log-likelihood function of the observations (3) after neglecting constant terms is proportional to

$$\log \rho\left(\mathbf{Y}|\theta,\phi,\alpha\right) \propto -\frac{1}{\sigma^2}\left\|\mathbf{Y}-\alpha\|\mathbf{s}\|_2\mathbf{b}_{\mathrm{Rx}}\left(\phi\right)\mathbf{a}_{\mathrm{Tx}}^{\mathrm{H}}(\theta)\mathbf{F}\right\|_{\mathrm{F}}^2. \tag{45}$$

According to the analysis of [23], the Fisher information matrix of the unknown parameters turns out to be

$$\mathbf{J}=\begin{bmatrix}\Phi_{\theta,\theta} & \Phi_{\theta,\phi} & \Phi_{\theta,\Re\alpha} & \Phi_{\theta,\Im\alpha}\\ \Phi_{\theta,\phi} & \Phi_{\phi,\phi} & \Phi_{\phi,\Re\alpha} & \Phi_{\phi,\Im\alpha}\\ \Phi_{\theta,\Re\alpha} & \Phi_{\phi,\Re\alpha} & \Phi_{\Re\alpha,\Re\alpha} & 0\\ \Phi_{\theta,\Im\alpha} & \Phi_{\phi,\Im\alpha} & 0 & \Phi_{\Im\alpha,\Im\alpha}\end{bmatrix}, \tag{46}$$

where[7]

$$\Phi_{\theta,\theta}=2\,\mathrm{SNR}\,\|\mathbf{b}_{\mathrm{Rx}}(\phi)\|_2^2\left\|\dot{\mathbf{b}}_{\mathrm{Tx}}(\theta)\right\|_2^2 \tag{47a}$$

$$\Phi_{\phi,\phi}=2\,\mathrm{SNR}\left\|\dot{\mathbf{b}}_{\mathrm{Rx}}(\phi)\right\|_2^2\|\mathbf{b}_{\mathrm{Tx}}(\theta)\|_2^2 \tag{47b}$$

$$\Phi_{\Re\alpha,\Re\alpha}=\Phi_{\Im\alpha,\Im\alpha}=\frac{2\|s\|_2^2}{\sigma^2}\|\mathbf{b}_{\mathrm{Rx}}(\phi)\|_2^2\|\mathbf{b}_{\mathrm{Tx}}(\theta)\|_2^2 \tag{47c}$$

$$\Phi_{\theta,\phi}=2\,\mathrm{SNR}\,\Re\left[\mathbf{b}_{\mathrm{Rx}}^{\mathrm{H}}(\phi)\dot{\mathbf{b}}_{\mathrm{Rx}}(\phi)\mathbf{b}_{\mathrm{Tx}}^{\mathrm{H}}(\theta)\dot{\mathbf{b}}_{\mathrm{Tx}}(\theta)\right] \tag{47d}$$

$$\Phi_{\theta,\Re\alpha}=\frac{2\|s\|_2^2}{\sigma^2}\|\mathbf{b}_{\mathrm{Rx}}(\phi)\|_2^2\,\Re\left[\alpha\,\dot{\mathbf{b}}_{\mathrm{Tx}}^{\mathrm{H}}(\theta)\mathbf{b}_{\mathrm{Tx}}(\theta)\right] \tag{47e}$$

$$\Phi_{\theta,\Im\alpha}=\frac{2\|s\|_2^2}{\sigma^2}\|\mathbf{b}_{\mathrm{Rx}}(\phi)\|_2^2\,\Im\left[\alpha\,\dot{\mathbf{b}}_{\mathrm{Tx}}^{\mathrm{H}}(\theta)\mathbf{b}_{\mathrm{Tx}}(\theta)\right] \tag{47f}$$

$$\Phi_{\phi,\Re\alpha}=\frac{2\|s\|_2^2}{\sigma^2}\Re\left[\alpha\,\mathbf{b}_{\mathrm{Rx}}^{\mathrm{H}}(\phi)\dot{\mathbf{b}}_{\mathrm{Rx}}(\phi)\right]\|\mathbf{b}_{\mathrm{Tx}}(\theta)\|_2^2 \tag{47g}$$

$$\Phi_{\phi,\Im\alpha}=\frac{2\|s\|_2^2}{\sigma^2}\Im\left[\alpha\,\mathbf{b}_{\mathrm{Rx}}^{\mathrm{H}}(\phi)\dot{\mathbf{b}}_{\mathrm{Rx}}(\phi)\right]\|\mathbf{b}_{\mathrm{Tx}}(\theta)\|_2^2, \tag{47h}$$

$\mathbf{b}_{\mathrm{Tx}}(\theta)\triangleq\mathbf{F}^{\mathrm{H}}\mathbf{a}_{\mathrm{Tx}}(\theta)$, $\dot{\mathbf{b}}_{\mathrm{Tx}}(\theta)\triangleq\frac{\mathrm{d}}{\mathrm{d}\theta}\mathbf{F}^{\mathrm{H}}\mathbf{a}_{\mathrm{Tx}}(\theta)$ and $\mathrm{SNR}=|\alpha|^2\|\mathbf{s}\|_2^2\sigma^{-2}$. The CRB requires inverting the FIM. Define

$$\mathbf{J}_{11}=\begin{bmatrix}\Phi_{\theta,\theta} & \Phi_{\theta,\phi}\\ \Phi_{\theta,\phi} & \Phi_{\phi,\phi}\end{bmatrix} \tag{48}$$

$$\mathbf{J}_{12}=\mathbf{J}_{21}^T=\begin{bmatrix}\Phi_{\theta,\Re\alpha} & \Phi_{\theta,\Im\alpha}\\ \Phi_{\phi,\Re\alpha} & \Phi_{\phi,\Im\alpha}\end{bmatrix} \tag{49}$$

$$\mathbf{J}_{22}=\begin{bmatrix}\Phi_{\Re\alpha,\Re\alpha} & 0\\ 0 & \Phi_{\Im\alpha,\Im\alpha}\end{bmatrix}. \tag{50}$$

Then, by the block matrix inversion formula, the CRB on the AoD and AoA is

$$\left(\mathbf{J}_{11}-\mathbf{J}_{12}\mathbf{J}_{22}^{-1}\mathbf{J}_{21}\right)^{-1}=\begin{bmatrix}\mathrm{DEB} & 0\\ 0 & \mathrm{OEB}\end{bmatrix} \tag{51}$$

where

$$\mathrm{DEB}^{-1}=2\,\mathrm{SNR}\,\|\mathbf{b}_{\mathrm{Rx}}(\phi)\|_2^2\left(\left\|\dot{\mathbf{b}}_{\mathrm{Tx}}(\theta)\right\|_2^2-\frac{\left|\dot{\mathbf{a}}_{\mathrm{Tx}}^{\mathrm{H}}(\theta)\dot{\mathbf{b}}_{\mathrm{Tx}}(\theta)\right|^2}{\|\mathbf{b}_{\mathrm{Tx}}(\theta)\|_2^2}\right) \tag{52}$$

$$\mathrm{OEB}^{-1}=2\,\mathrm{SNR}\,\|\mathbf{b}_{\mathrm{Tx}}(\theta)\|_2^2\left(\left\|\dot{\mathbf{b}}_{\mathrm{Rx}}(\phi)\right\|_2^2-\frac{\left|\mathbf{b}_{\mathrm{Rx}}^{\mathrm{H}}(\phi)\dot{\mathbf{b}}_{\mathrm{Rx}}(\phi)\right|^2}{\|\mathbf{b}_{\mathrm{Rx}}(\phi)\|_2^2}\right). \tag{53}$$

When computing the DEB (or OEB) for uniform linear arrays (ULA), to avoid dealing with zero-information points [47], we recommend the use of the 'spatial frequency' $\omega\triangleq\cos(\theta)/2$ (or $\omega\triangleq\cos(\phi)/2$) as unknown parameter instead of θ (or ϕ) as done in [48].

Appendix B
Interpretation of the DEB and OEB

The intuition behind DEB's expression (4), (52) is a bit cumbersome since it includes multiple terms that depend on $\mathbf{F}$. For notation convenience, define $\mathbf{b}_{\mathrm{Tx}}(\theta,\mathbf{F})\triangleq\mathbf{F}^{\mathrm{H}}\mathbf{a}_{\mathrm{Tx}}(\theta)$ and its derivative $\dot{\mathbf{b}}_{\mathrm{Tx}}(\theta,\mathbf{F})\triangleq\frac{\mathrm{d}}{\mathrm{d}\theta}\mathbf{b}_{\mathrm{Tx}}(\theta,\mathbf{F})$. By omitting all factors that do not depend on $\mathbf{F}$, the DEB may be expressed as

$$\mathrm{DEB}\propto\left(\left\|\dot{\mathbf{b}}_{\mathrm{Tx}}(\theta,\mathbf{F})\right\|_2^2-\left|\frac{\mathbf{b}_{\mathrm{Tx}}^{\mathrm{H}}(\theta,\mathbf{F})}{\|\mathbf{b}_{\mathrm{Tx}}(\theta,\mathbf{F})\|_2}\dot{\mathbf{b}}_{\mathrm{Tx}}(\theta,\mathbf{F})\right|^2\right)^{-1} \tag{54}$$

$$=\left\|\dot{\mathbf{b}}_{\mathrm{Tx}}^{\perp}(\theta,\mathbf{F})\right\|_2^{-2}, \tag{55}$$

where $\dot{\mathbf{b}}_{\mathrm{Tx}}^{\perp}(\theta,\mathbf{F})$ is the projection of $\dot{\mathbf{b}}_{\mathrm{Tx}}(\theta,\mathbf{F})$ onto the orthogonal complement of $\mathbf{b}_{\mathrm{Tx}}(\theta,\mathbf{F})$. The function $\mathbf{b}_{\mathrm{Tx}}(\theta,\mathbf{F})$ not only appears in the CRB, but careful attention reveals that it also appears in the signal model (3),

$$\mathbf{Y}=\alpha\|\mathbf{s}\|_2\mathbf{b}_{\mathrm{Rx}}(\phi)\mathbf{b}_{\mathrm{Tx}}^{\mathrm{H}}(\theta,\mathbf{F})+\mathbf{N}, \tag{56}$$

and it carries all information about the AoD. Since the CRB is a tight lower bound on the variance when the noise tends to zero, the signal $\mathbf{b}_{\mathrm{Tx}}(\theta,\mathbf{F})$ should be analyzed for the case where a small perturbation is applied to the AoD, in which case a good approximation is its first order Taylor polynomial

$$\mathbf{b}_{\mathrm{Tx}}(\theta',\mathbf{F})\approx\mathbf{b}_{\mathrm{Tx}}(\theta,\mathbf{F})+(\theta'-\theta)\dot{\mathbf{b}}_{\mathrm{Tx}}(\theta,\mathbf{F}), \tag{57}$$

where θ' and θ represent two closely spaced AoD. Let $\beta\mathbf{b}_{\mathrm{Tx}}(\theta,\mathbf{F})$ be the projection of $\dot{\mathbf{b}}_{\mathrm{Tx}}(\theta,\mathbf{F})$ onto $\mathbf{b}_{\mathrm{Tx}}(\theta,\mathbf{F})$, then,

$$\mathbf{b}_{\mathrm{Tx}}(\theta',\mathbf{F})\approx\left[1+(\theta'-\theta)\beta\right]\mathbf{b}_{\mathrm{Tx}}(\theta,\mathbf{F})+(\theta'-\theta)\dot{\mathbf{b}}_{\mathrm{Tx}}^{\perp}(\theta,\mathbf{F}). \tag{58}$$

From (57) we can infer that if $\dot{\mathbf{b}}_{\mathrm{Tx}}(\theta,\mathbf{F})$ is large, $\mathbf{b}_{\mathrm{Tx}}(\theta',\mathbf{F})$ and $\mathbf{b}_{\mathrm{Tx}}(\theta,\mathbf{F})$ will differ more. However, if $\dot{\mathbf{b}}_{\mathrm{Tx}}(\theta,\mathbf{F})$ and $\mathbf{b}_{\mathrm{Tx}}(\theta,\mathbf{F})$ are collinear (only differ by a scaling factor), then

[7]Contrary to [23], we have not assumed that $\|\mathbf{a}_{\mathrm{Tx}}(\theta)\|=\|\mathbf{a}_{\mathrm{Rx}}(\phi)\|=1$.

$\dot{\mathbf{b}}_{\mathrm{Tx}}^{\perp}(\theta, \mathbf{F}) = 0$, and by (58) the two signals will also be collinear. Since the channel gains are unknown, two collinear signals will be indistinguishable at the Rx, making the AoD unidentifiable. Thus, as asserted by the DEB, the accuracy in estimating the AoD depends on the magnitude of the part of $\dot{\mathbf{b}}_{\mathrm{Tx}}(\theta, \mathbf{F})$ orthogonal to $\mathbf{b}_{\mathrm{Tx}}(\theta, \mathbf{F})$.

Regarding the OEB (5), (53), from the Tx's perspective, its minimization boils down to maximizing $\|\mathbf{F}^{\mathrm{H}}\mathbf{a}_{\mathrm{Tx}}(\theta)\|_2^2$, i.e., maximizing the aggregated array gain (30) towards direction θ, which makes sense because maximizing the received energy helps combat the noise.

REFERENCES

[1] W. Roh, J.-Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 106–113, 2014.

[2] P. Wang, Y. Li, L. Song, and B. Vucetic, "Multi-gigabit millimeter wave wireless communications for 5G: From fixed access to cellular networks," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 168–178, 2015.

[3] S. Kutty and D. Sen, "Beamforming for millimeter wave communications: An inclusive survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 949–973, 2016.

[4] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Transactions on Communications*, vol. 63, no. 9, pp. 3029–3056, 2015.

[5] W. Keusgen, R. J. Weiler, M. Peter, M. Wisotzki, and B. Göktepe, "Propagation measurements and simulations for millimeter-wave mobile access in a busy urban environment," in *39th International Conference on Infrared, Millimeter, and Terahertz waves*. IEEE, 2014, pp. 1–3.

[6] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, 2014.

[7] T. E. Bogale, L. B. Le, and X. Wang, "Hybrid analog-digital channel estimation and beamforming: Training-throughput tradeoff," *IEEE Transactions on Communications*, vol. 63, no. 12, pp. 5235–5249, 2015.

[8] Z. Xiao, T. He, P. Xia, and X.-G. Xia, "Hierarchical codebook design for beamforming training in millimeter-wave communication," *IEEE Transactions on Wireless Communications*, vol. 15, no. 5, pp. 3380–3392, 2016.

[9] K. Venugopal, A. Alkhateeb, N. G. Prelcic, and R. W. Heath Jr, "Channel estimation for hybrid architecture based wideband millimeter wave systems," *arXiv preprint arXiv:1611.03046*, 2016.

[10] J. Rodríguez-Fernández, N. González-Prelcic, and R. W. Heath, "Channel estimation in mixed hybrid-low resolution MIMO architectures for mmWave communication," in *Signals, Systems and Computers, 2016 50th Asilomar Conference on*. IEEE, 2016, pp. 768–773.

[11] V. Va, T. Shimizu, G. Bansal, R. W. Heath Jr *et al.*, "Millimeter wave vehicular communications: A survey," *Foundations and Trends® in Networking*, vol. 10, no. 1, pp. 1–113, 2016.

[12] J. Choi, V. Va, N. González-Prelcic, R. Daniels, C. R. Bhat, and R. W. Heath Jr, "Millimeter wave vehicular communication to support massive automotive sensing," *IEEE Communications Magazine*, vol. 54, no. 12, pp. 160–167, 2016.

[13] P. Sanchis, J. Martinez, J. Herrera, V. Polo, J. Corral, and J. Marti, "A novel simultaneous tracking and direction of arrival estimation algorithm for beam-switched base station antennas in millimeter-wave wireless broadband access networks," in *IEEE Antennas and Propagation Society International Symposium*, vol. 1, 2002, pp. 594–597.

[14] J. Seo, Y. Sung, G. Lee, and D. Kim, "Training beam sequence design for millimeter-wave MIMO systems: A POMDP framework," *submitted for publication*, 2015.

[15] J. He, T. Kim, H. Ghauch, K. Liu, and G. Wang, "Millimeter wave MIMO channel tracking systems," in *Globecom Workshops*. IEEE, 2014, pp. 416–421.

[16] M. Kokshoorn, P. Wang, Y. Li, and B. Vucetic, "Fast channel estimation for millimetre wave wireless systems using overlapped beam patterns," in *IEEE International Conference on Communications (ICC)*, 2015, pp. 1304–1309.

[17] R. Di 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, 2014.

[18] N. Garcia, H. Wymeersch, E. G. Ström, and D. Slock, "Location-aided mm-wave channel estimation for vehicular communication," in *17th International Workshop on Signal Processing Advances in Wireless Communications*. IEEE, 2016, pp. 1–5.

[19] J. C. Aviles and A. Kouki, "Position-aided mm-wave beam training under NLOS conditions," *IEEE Access*, vol. 4, pp. 8703–8714, 2016.

[20] F. Maschietti, D. Gesbert, P. de Kerret, and H. Wymeersch, "Robust location-aided beam alignment in millimeter wave massive MIMO," in *IEEE GLOBECOM*, 2017.

[21] T. Nitsche, C. Cordeiro, A. B. Flores, E. W. Knightly, E. Perahia, and J. C. Widmer, "IEEE 802.11 ad: directional 60 Ghz communication for multi-Gigabit-per-second wi-fi," *IEEE Communications Magazine*, vol. 52, no. 12, pp. 132–141, 2014.

[22] J. Wang, Z. Lan, C.-S. Sum, C.-W. Pyo, J. Gao, T. Baykas, A. Rahman, R. Funada, F. Kojima, I. Lakkis *et al.*, "Beamforming codebook design and performance evaluation for 60 GHz wideband WPANs," in *IEEE 70th Vehicular Technology Conference Fall*, 2009, pp. 1–6.

[23] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch, "5G position and orientation estimation through millimeter wave MIMO," in *GLOBECOM 2015*. IEEE, 2015, pp. 1–6.

[24] H. L. Van Trees, *Detection, estimation, and modulation theory*. John Wiley & Sons, 2004.

[25] R. Méndez-Rial, C. Rusu, N. González-Prelcic, A. Alkhateeb, and R. W. Heath, "Hybrid mimo architectures for millimeter wave communications: Phase shifters or switches?" *IEEE Access*, vol. 4, pp. 247–267, 2016.

[26] J. G. Andrews, T. Bai, M. N. Kulkarni, A. Alkhateeb, A. K. Gupta, and R. W. Heath, "Modeling and analyzing millimeter wave cellular systems," *IEEE Transactions on Communications*, vol. 65, no. 1, pp. 403–430, 2017.

[27] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear algebra and its applications*, vol. 284, no. 1, pp. 193–228, 1998.

[28] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[29] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[30] MOSEK ApS, "The MOSEK optimization toolbox for MATLAB manual, version 7.1 (revision 51)," http://mosek.com, (accessed on March 20, 2016).

[31] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM review*, vol. 52, no. 3, pp. 471–501, 2010.

[32] M. Fazel, H. Hindi, and S. P. Boyd, "A rank minimization heuristic with application to minimum order system approximation," in *American Control Conference, 2001*, vol. 6. IEEE, 2001, pp. 4734–4739.

[33] G. H. Golub and C. F. Van Loan, *Matrix computations*. JHU Press, 2012, vol. 3.

[34] I. Olkin and A. W. Marshall, *Inequalities: theory of majorization and its applications*. Academic press, 1979, vol. 143.

[35] R. B. Bendel and M. R. Mickey, "Population correlation matrices for sampling experiments," *Communications in Statistics-Simulation and Computation*, vol. 7, no. 2, pp. 163–182, 1978.

[36] P. I. Davies and N. J. Higham, "Numerically stable generation of correlation matrices and their factors," *BIT Numerical Mathematics*, vol. 40, no. 4, pp. 640–651, 2000.

[37] M. T. Ivrlač and J. A. Nossek, "Toward a circuit theory of communication," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 7, pp. 1663–1683, 2010.

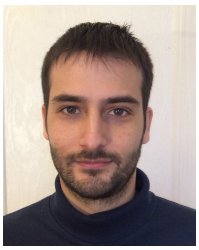
[38] J.-A. Tsai, R. M. Buehrer, and B. D. Woerner, "BER performance of a uniform circular array versus a uniform linear array in a mobile radio environment," *IEEE Transactions on Wireless Communications*, vol. 3, no. 3, pp. 695–700, 2004.

[39] L. C. Godara and A. Cantoni, "Uniqueness and linear independence of steering vectors in array space," *The Journal of the Acoustical Society of America*, vol. 70, no. 2, pp. 467–475, 1981.

[40] H. Gazzah and K. Abed-Meraim, "Optimum ambiguity-free directional and omnidirectional planar antenna arrays for DOA estimation," *IEEE Transactions on signal processing*, vol. 57, no. 10, pp. 3942–3953, 2009.

[41] V. Venkateswaran, F. Pivit, and L. Guan, "Hybrid RF and digital beamformer for cellular networks: Algorithms, microwave architectures,

and measurements," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 7, pp. 2226–2243, 2016.
[42] S. Han, I. Chih-Lin, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 186–194, 2015.
[43] G. E. Garcia, G. Seco-Granados, E. Karipidis, and H. Wymeersch, "Transmitter beam selection in millimeter-wave mimo with in-band position-aiding," *arXiv preprint arXiv:1705.05668*, 2017.
[44] D. G. Manolakis, V. K. Ingle, and S. M. Kogon, *Statistical and adaptive signal processing: spectral estimation, signal modeling, adaptive filtering, and array processing*. McGraw-Hill Boston, 2000.
[45] Z. Abu-Shaban, X. Zhou, T. Abhayapala, G. Seco-Granados, and H. Wymeersch, "Error bounds for uplink and downlink 3D localization in 5G millimeter wave systems," *IEEE Transactions on Wireless Communications*, 2018.
[46] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. Siam, 2001, vol. 2.
[47] E. Bashan, A. J. Weiss, and B.-S. Yaakov, "Estimation near "zero information" points: angle-of-arrival near the endfire," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 4, pp. 1250–1264, 2007.
[48] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 5, pp. 720–741, 1989.



Nil Garcia (S'14–M'16) received the Telecommunications Engineer degree from the Polytechnic University of Catalonia (UPC), Barcelona, Spain, in 2008; and the double Ph.D. degree in electrical engineering from the New Jersey Institute of Technology, Newark, NJ, USA, and from the National Polytechnic Institute of Toulouse, Toulouse, France, in 2015.

He is currently a postdoctoral researcher of Communication Systems with the Department of Signals and Systems at Chalmers University of Technology, Sweden. In 2009, he worked was an engineer in the Centre National d'Études Spatiales (CNES). In 2008 and 2009, he had internships in CNES and NASA. Hi research interests are in the areas of localization, intelligent transportation systems and 5G.



Henk Wymeersch (S'01–M'05) obtained the Ph.D. degree in Electrical Engineering/Applied Sciences in 2005 from Ghent University, Belgium.

He is currently a Professor of Communication Systems with the Department of Signals and Systems at Chalmers University of Technology, Sweden. Prior to joining Chalmers, he was a postdoctoral researcher from 2005 until 2009 with the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology.

Prof. Wymeersch served as Associate Editor for IEEE Communication Letters (2009–2013), IEEE Transactions on Wireless Communications (since 2013), and IEEE Transactions on Communications (since 2016). His current research interests include cooperative systems and intelligent transportation.



Dirk T. M. Slock received an EE degree from Ghent University, Belgium in 1982. In 1984 he was awarded a Fulbright scholarship for Stanford University, USA, where he received the MSEE, MS in Statistics, and PhD in EE in 1986, 1989 and 1989 resp. While at Stanford, he developed new fast recursive least-squares algorithms for adaptive filtering. In 1989-91, he was a member of the research staff at the Philips Research Laboratory Belgium. In 1991, he joined EURECOM where he is now professor. At EURECOM, he teaches statistical signal processing (SSP) and signal processing techniques for wireless communications. He invented semi-blind channel estimation, the chip equalizer-correlator receiver used by 3G HSDPA mobile terminals, spatial multiplexing cyclic delay diversity (MIMO-CDD) now part of LTE, and his work led to the Single Antenna Interference Cancellation (SAIC) integrated in the GSM standard in 2006. Recent keywords are MIMO interference channel, multi-cell, distributed resource allocation, variational and empirical Bayesian techniques, large system analysis, audio source separation, location estimation and exploitation. In 1992 he received one best journal paper award from IEEE-SP and one from EURASIP. He is the coauthor of two IEEE Globecom'98, one IEEE SIU'04, one IEEE SPAWC'05, one IEEE WPNC'16 and one IEEE SPAWC'18 best student paper award, and a honorary mention (finalist in best student paper contest) at IEEE SSP'05, IWAENC'06, IEEE Asilomar'06 and IEEE ICASSP'17. He was an associate editor for the IEEE-SP Transactions in 1994-96 and the IEEE Signal Processing Letters in 2009-10. He was the General Chair of the IEEE-SP SPAWC'06 and IWAENC'14 workshops, and EUSIPCO'15. He cofounded the start-ups SigTone in 2000 (music signal processing products) and Nestwave in 2014 (Ultra Low-Power Indoor and Outdoor Mobile Positioning). He has also been active as a consultant on xDSL, DVB-T and 3G systems. He is a Fellow of IEEE and EURASIP. In 2018 he received the URSI France medal.