Spatial Signal Design for Positioning via End-to-End Learning

(article starts on next page)

Spatial Signal Design for Positioning via End-to-End Learning

Steven Rivetti, José Miguel Mateos-Ramos, *Student Member, IEEE*, Yibo Wu, *Student Member, IEEE*, Jinxiang Song, *Student Member, IEEE*, Musa Furkan Keskin, *Member, IEEE*, Vijaya Yajnanarayana, *Senior Member, IEEE*, Christian Häger, *Member, IEEE*, Henk Wymeersch, *Senior Member, IEEE*

Abstract—This letter considers the problem of end-to-end (E2E) learning for joint optimization of transmitter precoding and receiver processing for mmWave downlink positioning. Considering a multiple-input single-output (MISO) scenario, we propose a novel autoencoder (AE) architecture to estimate user equipment (UE) position with multiple base stations (BSs) and demonstrate that E2E learning can match model-based design, both for angle-of-departure (AoD) and position estimation, under ideal conditions without model deficits and outperform it in the presence of hardware impairments.

Index Terms—mmWave positioning, precoder optimization, end-to-end learning.

I. INTRODUCTION

THE combination of high delay resolution at mmWave frequencies thanks to large bandwidth and high angular resolution thanks to large arrays is an important enabler for accurate positioning in 5G [1] and beyond [2]. The estimation of time-of-arrival (ToA), angle-of-arrival (AoA), and angle-of-departure (AoD) is enabled by *designed pilot signals* in time, frequency, and in space (at the base station (BS)) [3]. Such designs, in combination with advanced signal processing, can leverage the physical resources efficiently when suitable models are available. Traditionally, signal designs were optimized for broadcast performance in order to localize all users irrespective of their position [4]. Recently, there has been an increased focus on spatial per-user signal design, leveraging a priori knowledge of the user's location in order to further improve accuracy, both for positioning [5] and sensing [6]. Signal designs can be categorized as *model-based* [5], [7]–[10] or *based on artificial intelligence (AI)* [11]–[14]. Model-based signal designs can be performed based on simple heuristics [7], or on minimizing the Cramér-Rao bound (CRB) on the AoA, AoD, or the position via the position error bound (PEB). After relaxation, the optimization problems can be cast in convex forms, leading to elegant and efficient designs (e.g., [5] for angle estimation and [10] for positioning). From these solutions, online adaptive precoders [8] and robust designs based on predetermined codebooks with power allocation [9] have been considered.

An important limitation of model-based designs is that they require a model of the transmitter, receiver, and propagation channel. Under model mismatch, e.g., hardware impairments (HWIs), model-based approaches may exhibit degraded performance. Moreover, in certain cases, even with perfect model knowledge, finding optimal signal designs can be intractable. To remedy these two shortcomings, end-to-end (E2E) learning
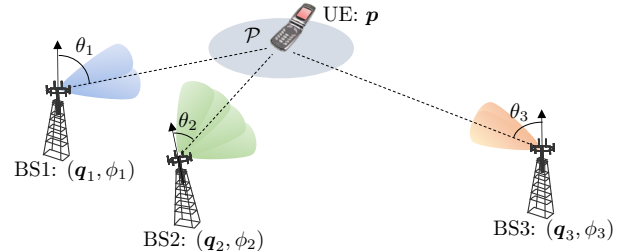
Fig. 1: The MISO downlink scenario comprising $I > 1$ multiple-antenna BSs and a single-antenna user equipment (UE). The UE determines its position based on the AoD estimates with respect to the BSs.

has been gaining interest, first in the context of communication [15] and more recently for sensing [11], but not for positioning. The principle is to model the entire system as an autoencoder (AE) [15] or by a combination or a reinforcement learning transmitter and a supervised learning receiver [16], combined with a suitable loss function (see, e.g., [11]). An application of E2E learning for spatial precoder design can be found in [12], where the probing codebook is implemented by a neural network (NN) module that is jointly trained with the beam predictor in order to predict the optimal narrow beam. Furthermore, [13] extends the learned beamforming to integrated sensing and communication (ISAC) by implementing the transmitter as a convolutional NN able to learn the features of historical channel and predict the next beamforming matrix. E2E learning in the presence of HWIs for ISAC has been proposed in [14]. AI-based solutions have also been applied in other forms to deal with HWIs, e.g., [17] proposes a super-resolution direction of arrival network, implemented as a convolutional NN, that can outperform AoA estimation methods under mutual coupling (MC).

In this paper, E2E learning is applied for the first time in positioning, in order to jointly optimize transmit beamformers and receiver-side algorithms, even in the presence of HWIs. Our contributions are *(i)* a novel AE architecture and loss function for AoD- and positioning-optimized signal design and estimator design; *(ii)* a detailed performance comparison to a state-of-the-art model-based benchmark and corresponding CRBs; *(iii)* an evaluation under different HWIs, namely array element inter-distance perturbation and array MC, demonstrating the robustness of the proposed E2E solution.

II. SYSTEM MODEL

A. Scenario and Signal Model

We consider a mmWave multiple-input single-output (MISO) downlink scenario with $I > 1$ multiple-antenna BSs and a single-antenna UE with unknown location $\boldsymbol{p} = [p_1 \ p_2]^T \in \mathcal{P} \subset \mathbb{R}^2$, where $\mathcal{P}$ is the prior location information. Each BS i has a known location $\boldsymbol{q}_i = [q_{i,1} \ q_{i,2}]^T \in \mathbb{R}^2$ and orientation $\psi_i \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, and is assumed to be equipped with an N_{Tx}-element uniform linear array (ULA) with $\lambda/2$ antenna
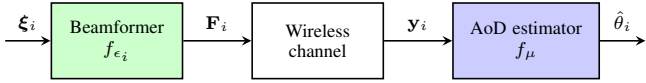
Fig. 2: Block diagram of the proposed AE architecture dedicated to AoD estimation, the green and blue blocks are implemented as trainable feed-forward NNs.
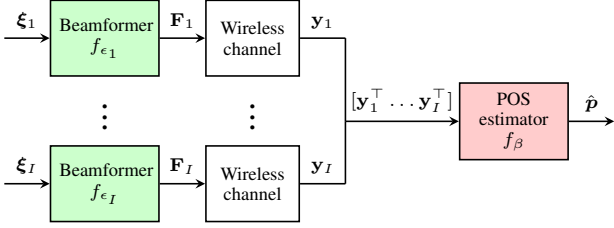

Fig. 3: Block diagram of the proposed AE architecture dedicated to position estimation, the red block is implemented as trainable feed-forward NN.

spacing, where λ denotes the wavelength of the carrier. The scenario is visualized in Fig. 1.

The i-th BS broadcasts a narrowband signal over $T > 1$ successive transmissions. We assume that BS transmissions are orthogonalized in time or frequency [4], leading to the observation at the UE from BS i at transmission t given by

$$y_{i,t} = \alpha_i \mathbf{a}^\top(\theta_i) \mathbf{f}_{i,t} s_{i,t} + n_{i,t} , \qquad (1)$$

where $s_{i,t}$ denotes the pilot signal with an unit power $|s_{i,t}|^2 = 1$, $\alpha_i \in \mathbb{C}$ and $\theta_i \in [-\frac{\pi}{2}, +\frac{\pi}{2}]$ denote, respectively, the complex channel gain and AoD from the i-th BS, $\mathbf{a}(\theta_i) \in \mathbb{C}^{N_{\mathrm{Tx}}}$ is the array steering vector at the BS (ULA of N_{Tx} elements and $\lambda/2$ antenna spacing), $\mathbf{f}_{i,t} \in \mathbb{C}^{N_{\mathrm{Tx}}}$ is the precoder employed by the i-th BS at time t, and $n_{i,t} \sim \mathcal{CN}(0, \sigma^2)$ is the additive white noise with variance σ^2, accounting also for the signal energy. In a more compact form, (1) can be rewritten as

$$\mathbf{y}_i = \alpha_i \left(\mathbf{F}_i^\top \mathbf{a}(\theta_i) \right) \odot \mathbf{s}_i + \mathbf{n}_i , \qquad (2)$$

where $\odot$ represents the Hadamard product, $\mathbf{y}_i = [y_{i,1} \ldots y_{i,T}]^\top$, $\mathbf{F}_i = [\mathbf{f}_{i,1} \ldots \mathbf{f}_{i,T}] \in \mathbb{C}^{N_{\mathrm{Tx}} \times T}$ is the precoder matrix of the i-th BS, $\mathbf{s}_i = [s_{i,1} \ldots s_{i,T}]^\top$, and $\mathbf{n}_i = [n_{i,1} \ldots n_{i,T}]^\top$. From the UE and BS positions, the AoD is computed as

$$\theta_i = \mathrm{atan2}(p_2 - q_{i,2}, p_1 - q_{i,1}) - \psi_i , \qquad (3)$$

which accounts for the BS orientation. We assume that the UE lies in the angular sector $\mathcal{U}_i = [\theta_{i,\min}, \theta_{i,\max}] \in \mathbb{R}^2$ with respect to the BS i, depending on the uncertainty region $\mathcal{P}$.

B. Hardware Impairment Models

Without HWIs, the steering vectors are given by $[\mathbf{a}(\theta)]_k = e^{j\pi k \sin(\theta)}$, $k = 0, \ldots, N_{\mathrm{Tx}} - 1$. We now describe the impact of inter-antenna element spacing perturbations and MC, which lead to an impaired steering vector, denoted by $\tilde{\mathbf{a}}(\theta)$.

1) Antenna Element Spacing Perturbations: We introduce the vector of inter-element distances as $\boldsymbol{d} \in \mathbb{R}^{N_{\mathrm{Tx}}-1}$, where without HWIs, $\boldsymbol{d} = \frac{\lambda}{2} \mathbf{1}_{N_{\mathrm{Tx}}-1}$. Here, $\mathbf{1}_{N_{\mathrm{Tx}}-1}$ denotes a vector of $(N_{\mathrm{Tx}}-1)$ ones. With spacing perturbations caused by HWIs [18], the distance is modeled by

$$\boldsymbol{d} = \frac{\lambda}{2} \mathbf{1}_{N_{\mathrm{Tx}}-1} + \boldsymbol{\gamma}, \gamma_k \sim \mathcal{N}(0, \sigma_\lambda^2), \qquad (4)$$

so that the perturbed steering vector becomes $[\tilde{\mathbf{a}}(\theta)]_k = e^{j2\pi k(d_k/\lambda)\sin(\theta)}$.

2) Mutual Coupling: Following [19], we introduce a coupling matrix $\boldsymbol{B} \in \mathbb{C}^{N_{\mathrm{Tx}} \times N_{\mathrm{Tx}}}$, which is modeled as a banded symmetric Toeplitz matrix whose entries are collected in the vector $\boldsymbol{c} = [1, c_1, \ldots, c_M]^\top$ $(0 < |c_M| < \ldots < |c_1| < 1)$, where M is the number of half-wavelength increments for

which the MC contribution is assumed non-negligible, so that $\tilde{\mathbf{a}}(\theta) = \boldsymbol{B}\mathbf{a}(\theta)$.

III. END-TO-END LEARNING

In this section, we describe the proposed architectures, the associated loss functions, and the model-based benchmark.

A. End-to-End Learning Architecture

We consider two separate AE architectures for AoD and position estimation, as shown in Fig. 2 and Fig. 3, respectively. Fig. 2 shows an E2E architecture to learn BS precoder design (highlighted in green) and UE-side AoD estimation from each BS (highlighted in blue). Fig. 3 shows an E2E architecture to learn BS precoder design (highlighted in green) and UE-side position estimation, based on the combined observation from all BSs (highlighted in red). We assume the wireless channel blocks are instantaneously differentiable.

1) Precoder NN: Each BS has its own precoder. The precoder for BS i is implemented by an NN $f_{\epsilon_i} : \mathbb{R}^3 \to \mathbb{C}^{N_{\mathrm{Tx}}T}$, with learnable parameters ϵ_i. Instead of directly using the AoD uncertainty region $\mathcal{U}_i$ as the NN input, we find it helpful to feed an over-determined parameterization of $\mathcal{U}_i$ as $\boldsymbol{\xi}_i \in \mathbb{R}^3$, with

$$\boldsymbol{\xi}_i = [\theta_{i,\min}, \theta_{i,\max}, (\theta_{i,\max} - \theta_{i,\min})/2]^\top. \qquad (5)$$

The NN output is a real-valued vector with a size $\mathbb{R}^{2N_{\mathrm{Tx}}T}$ that is then converted into the complex-valued precoding matrix $\mathbf{F}_i \in \mathbb{C}^{N_{\mathrm{Tx}} \times T}$. In this conversion, complex numbers are obtained by concatenating the real and imaginary parts, followed by a normalization with its Frobenius norm.

2) AoD Estimation NN: The AoD estimator for BS i is implemented by another NN $f_\mu : \mathbb{C}^T \to \mathbb{R}$, with learnable parameters μ, which takes the observation $\mathbf{y}_i$ as the input and generates an estimate $\hat{\theta}_i$. Since the AoD estimation process is identical for each BS, all I AoD estimators share the same parameters.

3) Position Estimation NN: The position estimator is implemented as $f_\beta : \mathbb{C}^{I \times T} \to \mathbb{R}^2$, with learnable parameters β, which takes as input $\mathbf{y} = [\mathbf{y}_1^\top, \mathbf{y}_2^\top \ldots, \mathbf{y}_I^\top]^\top$ and generates as output the position estimate $\hat{\boldsymbol{p}}$. Since AoD and position are intrinsically related, this direct approach could potentially be replaced with a two-step solution, by leveraging the AoD estimation NNs, at a cost of possible performance loss (due to the data processing theorem), but with possibly lower complexity. A two-step solution also necessitates computing the AoD uncertainties, as in [14].

B. Loss Functions

The E2E AoD estimation and E2E position estimation require two dedicated loss functions:

- *AoD estimation loss:* The loss function is the mean squared error (MSE) between the estimated and true AoDs:

$$\mathcal{L}_{\mathrm{AoD}}(\epsilon_i, \mu) = \mathbb{E}\{|\hat{\theta}_i - \theta_i|^2\} . \qquad (6)$$

The AoD estimators at the UE corresponding to each BS share NN parameters, so there is no need for separate training. Since the AoDs are limited to $[-\pi/2, \pi/2]$, there is no risk of wrapping effects, making the MSE meaningful in this scenario.

- *Positioning loss:* The loss function is set to

$$\mathcal{L}_{\mathrm{position}}(\epsilon_1, \ldots, \epsilon_I, \beta) = \mathbb{E}\{\|\hat{\boldsymbol{p}} - \boldsymbol{p}\|_2^2\} . \qquad (7)$$

TABLE I: NN structures.

Network	Input layer	Hidden layers	Output layer
Beamformer f_ϵ	3	H,H,H,H,H	$N_{\mathrm{Tx}}T$ (linear)
AoD decoder f_μ	2T	H,H,H,H,2H,2H	1 (tanh)
POS decoder f_β	4T	H,H,H,H,2H,2H	2 (linear)

C. Benchmarks

As a comparison, each of the NNs in Fig. 2 and Fig. 3 will be evaluated against a state-of-the-art benchmark.

1) Transmit Precoding Benchmark: The chosen precoder matrix for the BS i is a heuristic solution to the problem of minimization of worst-case CRB on AoD estimation over the uncertainty region $\mathcal{U}_i$. It consists of a hybrid base codebook, comprising both directional beams and their derivatives [10]

$$\mathbf{F}_i^{\mathrm{heur}} = [\mathbf{F}_i^{\mathrm{dir}}, \ \mathbf{F}_i^{\mathrm{der}}] \in \mathbb{C}^{N_{\mathrm{Tx}} \times T} \ , \tag{8}$$

$$\mathbf{F}_i^{\mathrm{dir}} = [\mathbf{a}(\theta_{i,0}), \ldots, \mathbf{a}(\theta_{i,T/2})] \in \mathbb{C}^{N_{\mathrm{Tx}} \times (T/2)} \ , \tag{9}$$

$$\mathbf{F}_i^{\mathrm{der}} = [\dot{\mathbf{a}}(\theta_{i,0}), \ldots, \dot{\mathbf{a}}(\theta_{i,T/2})] \in \mathbb{C}^{N_{\mathrm{Tx}} \times (T/2)} \ , \tag{10}$$

where $\{\theta_{i,g}\}_{g=1}^{T/2}$ represents the evenly spaced angular grid in $\mathcal{U}_i$ and $\dot{\mathbf{a}}(\theta) = \partial\mathbf{a}(\theta)/\partial\theta$. The benchmark precoder $\mathbf{F}_i^{\mathrm{b}}$, defined as $= [\sqrt{\rho_1}\mathbf{f}_{i,1}^{\mathrm{heur}}, \ldots, \sqrt{\rho_T}\mathbf{f}_{i,T}^{\mathrm{heur}}]$ where $\mathbf{f}_{i,t}^{\mathrm{heur}}$ denotes the t-th column of $\mathbf{F}_i^{\mathrm{heur}}$, is obtained by finding the power allocation vector $\boldsymbol{\rho} = [\rho_1 \ldots \rho_T]^\top$ that minimizes the CRB on AoD estimation [10]. Then, $\mathbf{F}_i^{\mathrm{b}}$ is normalized to have unit Frobenius norm; the same operation is implemented by the normalization layer at the output of the beamformer NN, ensuring the usage of the same total power between the two approaches.

2) AoD Estimation Benchmark: The UE implements maximum likelihood (ML) estimation, based on (2), yielding [20]

$$\hat{\theta}_i^{\mathrm{b}} = \arg\min_{\theta_i \in \mathcal{U}_i} \frac{|\mathbf{y}_i^{\mathsf{H}}\mathbf{F}_i^\top\mathbf{a}(\theta_i)|^2}{\|\mathbf{F}_i^\top\mathbf{a}(\theta_i)\|^2} \ . \tag{11}$$

3) Position Estimation Benchmark: Given the AoD estimates from (11), we formulate the measurement likelihood $p(\hat{\theta}_i^{\mathrm{b}}|\theta_i)$ as $p(\hat{\theta}_i^{\mathrm{b}}|\theta_i) \propto \exp(-(\hat{\theta}_i^{\mathrm{b}} - \theta_i)^2/(2\sigma_i^2))$, where σ_i^2 can be obtained from the CRB of the AoD estimator at BS i. Then, it immediately follows that the ML estimator is

$$\hat{\boldsymbol{p}}^{\mathrm{b}} = \arg\min_{\boldsymbol{p} \in \mathcal{P}} \sum_{i=1}^{I} \frac{1}{2\sigma_i^2}(\hat{\theta}_i^{\mathrm{b}} + \psi_i - \mathrm{atan2}(\boldsymbol{q_i}, \boldsymbol{p}))^2 \ . \tag{12}$$

IV. SIMULATION RESULTS

A. Simulation Parameters

We consider a scenario with $I = 2$ BSs, located at $\boldsymbol{q}_1 = [-5, 0]$ and $\boldsymbol{q}_2 = [3, 0]$ with orientations $\boldsymbol{\psi} = [0°, 10°]$, each with $N_{\mathrm{Tx}} = 32$ antenna elements. The number of transmissions is set to $T = 20$ with pilots $s_{i,t} = 1$, and the width of $\mathcal{U}_i$ varies uniformly between $10°$ and $20°$. The channel gains are set based on a target signal-to-noise ratio (SNR), i.e., $\mathrm{SNR}_i = |\alpha_i|^2/\sigma^2$, and the SNRs range from -5 dB to 30 dB. The phase of α_i is uniformly distributed in $[0, 2\pi]$ and the wavelength is set to 10.7 mm (corresponding to a carrier frequency of 28 GHz).

For modeling the HWIs, we generate the MC matrix $\boldsymbol{B}$ as a banded symmetric Toeplitz matrix built from the coefficients vector $\boldsymbol{c} = [1, 0.9e^{-j\pi/3}, 0.75e^{j\pi/4}, 0.55e^{-j\pi/10}, 0.25e^{-j\pi/6}]^\top$, while for generating the antenna element spacing perturbations, we set $\sigma_\lambda = \lambda/100$.

B. Autoencoder Training

The mini-batch size S is set to 10000 and we train with mean AoD uniformly distributed in $[-60°, 60°]$ and $\mathcal{U}_i$'s width uniformly distributed within $[10°, 20°]$. In terms of positioning AE, the training follows a similar rationale: each minibatch's sample is associated to a true position $\boldsymbol{p}$, modelled as a 2-D uniform random variable within a 10 m^2 area in front of the BSs. The observations $\mathbf{y}_i$ are then generated by calculating θ_i according to (3). Then, the mean of $\mathcal{U}_i$ is set to $\theta_{\mathrm{mid},i} = \theta_i + \nu_i$, where ν is a random variable varying uniformly within the interval $[-15°, 15°]$, as the a priori information induces a $30°$ wide $\mathcal{U}_i$ on both BS. The beam former NN input is then defined as $\boldsymbol{\xi}_i = [\theta_{\mathrm{mid},i} - 15°, \theta_{\mathrm{mid},i} + 15°, 15°]^\top$.

Based on a hyper-parameter search, which aimed to determine the smallest NN with the best possible performance, the number of hidden neurons 'H' is set to 256 and each layer uses a rectified linear unit (ReLU) activation function. Further details are provided in Table I. We also note that in practical applications, it may be of interest to use NN architectures with less complexity (i.e., fewer layers and/or neurons per layer) by sacrificing some accuracy. In terms of optimizer, we use the Adam optimizer [21] with a learning rate controlled by a scheduler whose starting value is 0.001 and lower bound is at 10^{-8}. We have found that re-training the systems with fixed SNR ranging from -5 dB to 30 dB yields better results than using a different SNR in every batch or sample.
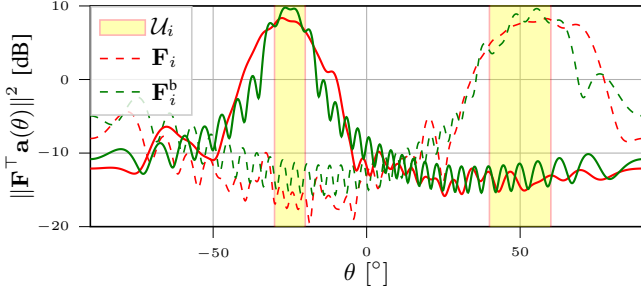
C. Results

1) Without Hardware Impairments: Fig. 4-(a) shows the aggregated response $\|\mathbf{F}^\top\mathbf{a}(\theta)\|^2$ of the AoD-optimized learned precoder $\mathbf{F} = \mathbf{F}_i$ for the two BSs for angle uncertainty intervals $\mathcal{U}_1 = [40°, 60°]$ and $\mathcal{U}_2 = [-30°, -20°]$, along with that of the benchmark precoder $\mathbf{F} = \mathbf{F}_i^{\mathrm{b}}$. Despite the AE having no knowledge of the benchmark precoder, the learned precoder has a strong similarity in terms of the aggregate response[1]. Fig. 4-(b) shows the AoD root mean squared error (RMSE) performance vs. SNR for BS 1, along with the corresponding CRBs[2]. The implicit power allocation process carried out by the AE in finding $\mathbf{F}_i$ is able to achieve the same performance bounds obtained through the explicit optimization process to determine $\mathbf{F}_i^{\mathrm{b}}$. Furthermore, both approaches are able to attain the CRB at an SNR around 10 dB. This trend is confirmed for positioning as well: Fig. 4-(c) shows that the E2E solution can reach the same PEB as its model-based counterpart, attaining it around an SNR of 5 dB.

2) Results under Hardware Impairments: Next, we show the impact of model mismatch caused by HWIs on the AoD and position estimation, while revealing the capability of the proposed AE to compensate for the resulting performance degradation.

First, we consider array element spacing perturbations, shown in Fig. 5 and Fig. 6. The observations are generated using the model from Section II-B1, while the model-based benchmark is unaware of this impairment. From Fig. 5, we observe that the AE precoder responses are less similar to the benchmark, compared to the case without HWI: this difference in the precoders can be interpreted as an active adaptation to $\tilde{\mathbf{a}}(\theta)$. This is also seen in Fig. 6, which shows the AoD and positioning RMSE. In particular, at medium and high SNR

---

[1]The position-optimized precoders exhibit similar trends (results not shown for space reasons).

[2]The benchmark CRB and the AE CRB are computed by employing $\mathbf{F}_i^{\mathrm{b}}$ and $\mathbf{F}_i$ as the precoding matrices, respectively [22, Ch. 3]. Since the CRB depends on the transmit signal, not on the receiver processing, different precoders may lead to different CRB values. Additionally, we note that under HWIs the steering vector model used in the CRB computation is the true one, i.e., $\tilde{\mathbf{a}}(\theta)$.

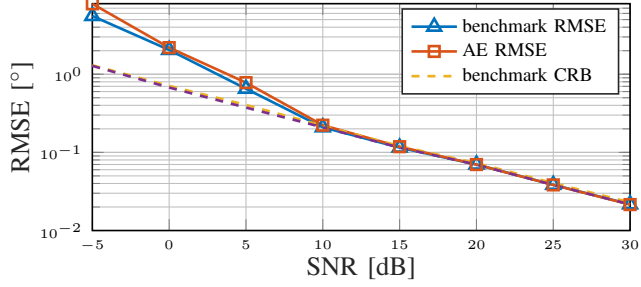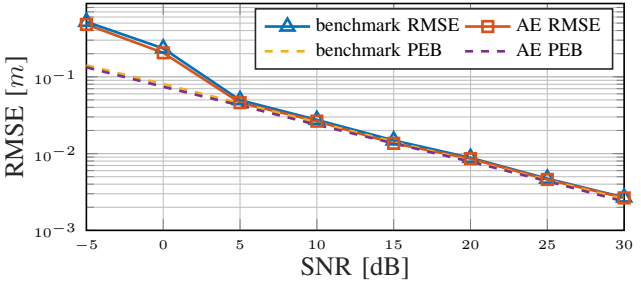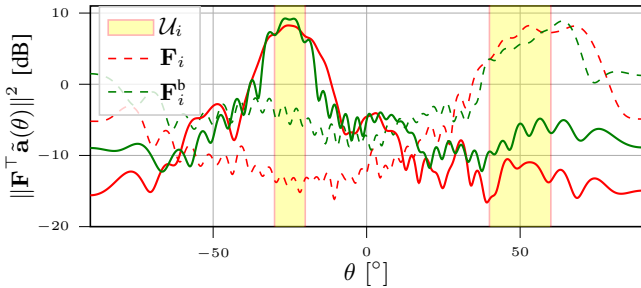(a) Precoder responses with SNR=10 dB

(b) AoD RMSE for $\mathcal{U}_1 = [40°, 60°]$

(c) Positioning RMSE for $\boldsymbol{p} = [0.5, 5]$ m

Fig. 4: Results without HWIs: Performance comparison of the AEs with the benchmark.

Fig. 5: Aggregate response of the precoders $\mathbf{F}_i$ compared against the respective benchmarks $\mathbf{F}_i^{\mathrm{b}}$ under array element spacing perturbations with $\sigma_\lambda = \lambda/100$ and for SNR = 10 dB.

(a) AoD RMSE for $\mathcal{U}_1 = [40°, 60°]$

(b) Positioning RMSE for $\boldsymbol{p} = [0.5, 5]$ m

Fig. 6: RMSE performance assessment in the presence of array element spacing perturbations.
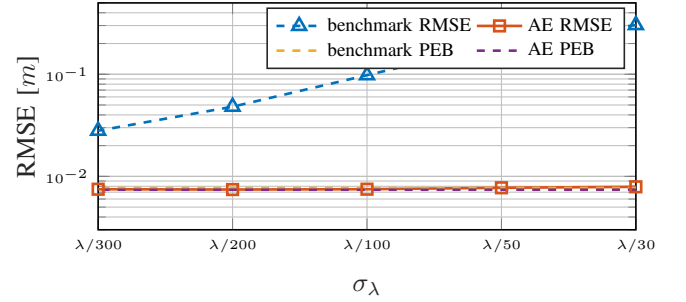
Fig. 7: Positioning RMSE performances at an SNR of 20 dB for increasingly large array element spacing perturbations σ_λ.

values, the benchmark suffers from significant performance penalties due to mismatch between the true model $\tilde{\mathbf{a}}(\theta)$ and the employed model $\mathbf{a}(\theta)$, in line with the theoretical results from [23]. The AE is able to attain its CRB in both AoD and position estimation, verifying the effectiveness of the proposed architecture under model imperfections. Moreover, Fig. 7 plots the position RMSE with respect to σ_λ for a fixed SNR of 20 dB, which further confirms the robustness of the E2E solution. Specifically, the positioning AE can achieve the PEB regardless of σ_λ, whereas the model-based approach leads to a performance penalty that increases with σ_λ.

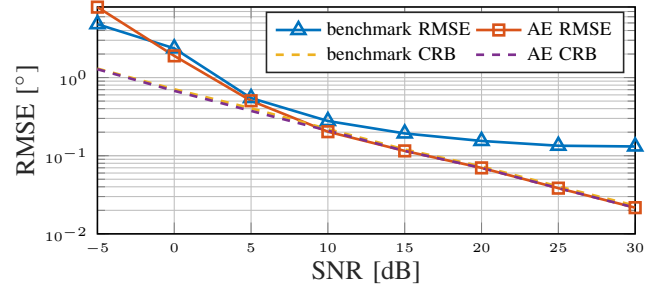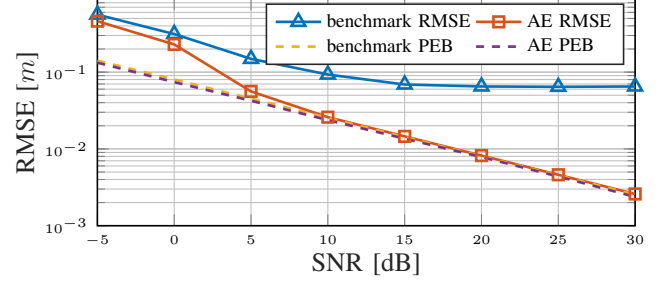Second, we evaluate the impact of MC, where the observations are now generated according to the model from Section II-B2. In Fig. 8, the precoder responses are shown, which suggests that the proposed learning-based approach can naturally adapt its precoder to deal with HWIs, leading to a beampattern that is different from that of the model-based approach. Fig. 9 illustrates the RMSEs and the CRBs of the considered strategies under the impact of MC. It is seen that the precoder generated by the AE can achieve the same performance bound as the model-based benchmark. In terms of RMSE, the MC prevents the benchmark estimator from attaining its bound, while the proposed AE can successfully reach the theoretical limits. For positioning with the benchmark estimator, the MC induces an error floor effect
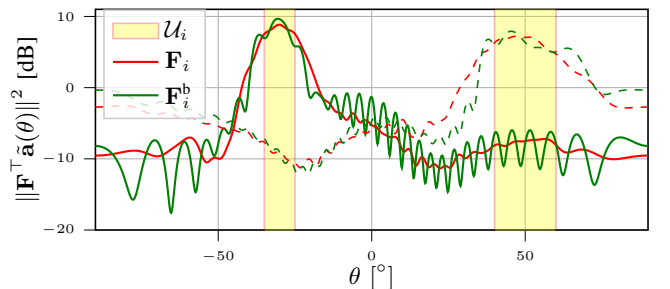
Fig. 8: Aggregate response of the precoders $\mathbf{F}_i$ compared against the respective benchmarks $\mathbf{F}_i^{\mathrm{b}}$ under MC for SNR = 10 dB.
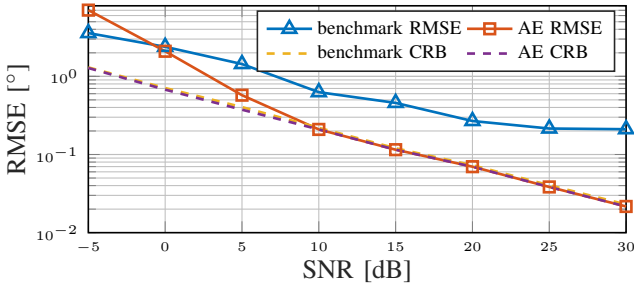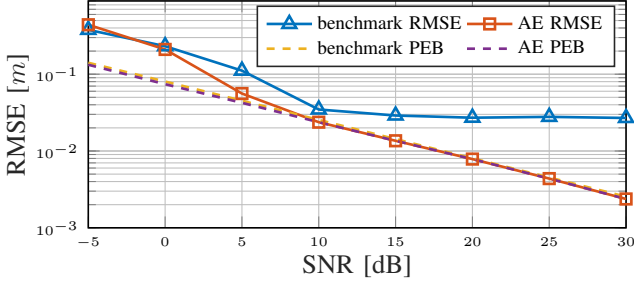
(a) AoD RMSE for $\mathcal{U}_1 = [40°, 60°]$

(b) Positioning RMSE for $\boldsymbol{p} = [0.5, 5]$ m

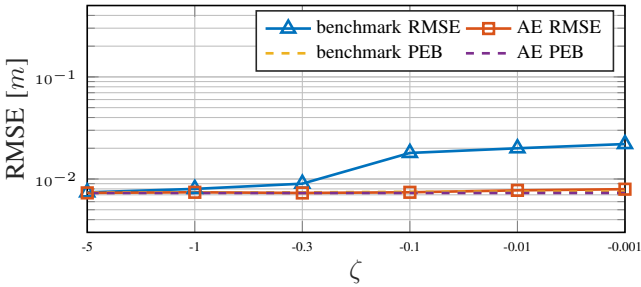Fig. 9: RMSE performance assessment in presence of MC.

Fig. 10: Positioning RMSE performances at an SNR of 20 dB for different magnitudes of MC.

beyond 10 dB, as expected from [23]. Further insights into the effects of MC are provided in Fig. 10 (using the same RMSE scale as Fig. 7), where we model the MC coupling coefficients vector as $|c_k| = \exp(\zeta k), k \in \{0, \ldots, 4\}$ and retain the phase of the original $\boldsymbol{c}$ reported in Section IV-A. The resulting matrix $\boldsymbol{B}$ is normalized to have the same Frobenius norm as the matrix $\boldsymbol{B}$ built from the vector $\boldsymbol{c}$ reported in Section IV-A. Similar to Fig. 7, the E2E solution is able to attain its performance bound, whereas the model-based solution shows a performance penalty inversely proportional to the decay parameter ζ. Comparing with Fig. 7, we do however note that the impact of MC is less severe than array spacing perturbations.

V. CONCLUSIONS

We have addressed the problem of positioning and AoD estimation at a UE, based on downlink MISO transmission. To this end, we propose a novel AE architecture with judiciously designed inputs and loss functions, which jointly learns optimized precoders and receivers under UE location uncertainty. We have compared the AE performance against model-based precoder designs and ML estimators. Through numerical simulations, the learned precoders are shown to yield the same bounds as their model-based counterparts. Without model imperfections, the learned receiver can attain the same RMSE level as the ML estimator. In the presence

of HWIs, the learned receiver can significantly outperform the ML estimator, especially at high SNRs and large degree of inter-element perturbations and MC, showcasing the robustness of the proposed AE architecture against model deficits. Possible future work include extension to 3D scenarios and investigation of two-step architectures that exploit the relation between AoD and position (i.e., (12)) to jointly design their corresponding NN estimators for reduced complexity.

REFERENCES

[1] 3rd Generation Partnership Project (3GPP), "Study on NR positioning support TR 38.855," *Technical Specification Group Radio Access Network*, 2019.

[2] S. Bartoletti *et al.*, "Positioning and sensing for vehicular safety applications in 5G and beyond," *IEEE Communications Magazine*, vol. 59, no. 11, pp. 15–21, 2021.

[3] R. Keating *et al.*, "Overview of positioning in 5G new radio," in *IEEE International Symposium on Wireless Communication Systems (ISWCS)*, 2019, pp. 320–324.

[4] S. Dwivedi *et al.*, "Positioning in 5G networks," *IEEE Communications Magazine*, vol. 59, no. 11, pp. 38–44, 2021.

[5] N. Garcia *et al.*, "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.

[6] F. Liu *et al.*, "Toward dual-functional radar-communication systems: Optimal waveform design," *IEEE Transactions on Signal Processing*, vol. 66, no. 16, pp. 4264–4279, 2018.

[7] A. Fascista *et al.*, "Low-complexity accurate mmwave positioning for single-antenna users based on angle-of-departure and adaptive beamforming," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 4866–4870.

[8] B. Zhou *et al.*, "Successive localization and beamforming in 5G mmwave MIMO communication systems," *IEEE Transactions on Signal Processing*, vol. 67, no. 6, pp. 1620–1635, 2019.

[9] A. Kakkavas *et al.*, "Power allocation and parameter estimation for multipath-based 5G positioning," *IEEE Transactions on Wireless Communications*, vol. 20, no. 11, pp. 7302–7316, 2021.

[10] M. F. Keskin *et al.*, "Optimal spatial signal design for mmwave positioning under imperfect synchronization," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 5, pp. 5558–5563, 2022.

[11] W. Jiang *et al.*, "Joint design of radar waveform and detector via end-to-end learning with waveform constraints," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 1, pp. 552–567, 2022.

[12] Y. Heng *et al.*, "Learning site-specific probing beams for fast mmWave beam alignment," *IEEE Transactions on Wireless Communications*, 2022.

[13] C. Liu *et al.*, "Learning-based predictive beamforming for integrated sensing and communication in vehicular networks," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 8, pp. 2317–2334, 2022.

[14] J. M. Mateos-Ramos *et al.*, "End-to-end learning for integrated sensing and communication," in *IEEE International Conference on Communications (ICC)*, 2022, pp. 1942–1947.

[15] T. O'Shea *et al.*, "An introduction to deep learning for the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 563–575, 2017.

[16] F. A. Aoudia *et al.*, "End-to-end learning of communications systems without a channel model," in *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, 2018, pp. 298–303.

[17] P. Chen *et al.*, "SDOAnet: An efficient deep learning-based DOA estimation network for imperfect array," *arXiv preprint arXiv:2203.10231*, 2022.

[18] T. Yassine *et al.*, "mpNet: Variable depth unfolded neural network for massive MIMO channel estimation," *IEEE Transactions on Wireless Communications*, vol. 21, no. 7, pp. 5703–5714, 2022.

[19] Z. Zheng *et al.*, "Robust adaptive beamforming against mutual coupling based on mutual coupling coefficients estimation," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 10, pp. 9124–9133, 2017.

[20] A. Fascista *et al.*, "Millimeter-wave downlink positioning with a single-antenna receiver," *IEEE Transactions on Wireless Communications*, vol. 18, no. 9, pp. 4479–4489, 2019.

[21] D. P. Kingma *et al.*, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[22] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, Inc., 1993.

[23] H. Chen *et al.*, "MCRB-based performance analysis of 6G localization under hardware impairments," in *IEEE International Conference on Communications Workshops (ICC Workshops)*, 2022, pp. 115–120.