Joint Localization and Synchronization in Downlink Distributed MIMO

Sauradeep Dey*, Musa Furkan Keskin*, Dario Tagliaferri†, Gonzalo Seco-Granados‡, Henk Wymeersch*

*Department of Electrical Engineering, Chalmers University of Technology, Sweden
†Department of Electronics, Information and Bioengineering, Politecnico di Milano, Italy
‡Universitat Autonoma de Barcelona, Spain

Abstract—We investigate joint localization and synchronization in the downlink of a distributed multiple-input-multiple-output (D-MIMO) system, aiming to estimate the position and phase offset of a single-antenna user equipment (UE) using downlink transmissions of multiple phase-synchronized, multi-antenna access points (APs). We propose two transmission protocols: sequential (*P1*) and simultaneous (*P2*) AP transmissions, together with the ML estimators that either leverage (coherent estimator) or disregard phase information (non-coherent estimator). Simulation results reveal that downlink D-MIMO holds significant potential for high-accuracy localization while showing that *P2* provides superior localization performance and reduced transmission latency.

Index Terms—Distributed MIMO, downlink localization, phase-coherent localization, beamforming, beamfocusing.

I. INTRODUCTION

Localization has been a key research area in wireless communications, evolving from early cellular systems to advanced techniques in 5G and beyond [1]. Traditional methods such as time difference of arrival (TDoA) and angle of arrival (AoA) have been widely used, but their accuracy was limited by narrow bandwidth, hardware timing offsets and small arrays at base stations (BSs) [2]. The advent of massive MIMO in 5G introduced large antenna arrays, improving localization accuracy by leveraging high spatial resolution to estimate AoA and angle-of-departure (AoD) [3]. Looking ahead to 6G, localization is envisioned as a core service with stringent requirements (e.g., sub-cm level for immersive telepresence [4]). To fulfill such demands in 6G, distributed MIMO (D-MIMO), where multiple spatially distributed access points (APs) cooperate to function as a unified system, has emerged as a promising approach for localization [5]. By exploiting near-field propagation effects induced by phase coherence among APs, D-MIMO systems can yield extreme accuracies that can support 6G use cases such as digital twinning and extended reality [6].

A distinguishing feature of emerging D-MIMO localization is the possibility of phase-coherent operation across distributed APs, enabling the entire D-MIMO system to form an extremely large aperture array, which introduces wavefront curvature (i.e., near-field) effects and tremendous gains in resolution [7]. Coherent processing in D-MIMO systems can
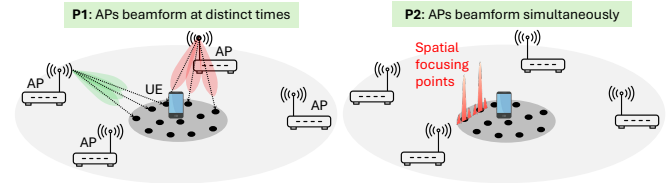
Figure 1: System model of a downlink D-MIMO system for UE localization. The APs are phase-synchronized, leading to two options for downlink beamforming (*P1* and P2).

push the position error to orders of magnitude below the carrier wavelength [8]. Recent studies have mainly focused on *uplink*-based D-MIMO localization [7]–[10], where the network-side phase-coherent processing uses uplink signals received at multiple APs to infer the location of a transmitter user equipment (UE).

Extending this framework to downlink offers advantages and challenges. The key benefit lies in beamfocusing [11], where multiple APs can coordinate to concentrate signals around the UE, unlike uplink, where a UE can only beamform toward a single AP. Downlink also scales efficiently for multiple UEs without additional network overhead, while uplink requires dedicated resources per UE. Additionally, downlink enables decentralized localization, allowing each UE to compute its position locally, reducing fronthaul data load [5]. The primary challenge is that downlink places the processing burden on the UE, which has limited capabilities and must process superposed signals from multiple APs. Despite these trade-offs, phase-coherent downlink D-MIMO localization remains unexplored, leaving a gap in the field.

In this paper, we make progress towards filling this gap by addressing the fundamental problem of joint localization and synchronization of a UE using downlink signals transmitted by spatially distributed phase-coherent APs. The main contributions are as follows: *(i)* For the first time, we propose a downlink phase-coherent D-MIMO localization system, using signals transmitted by multiple phase-synchronized multi-antenna APs. For the transmit signal design, we propose two different transmission protocols: *P1*, where the APs transmit their signals sequentially for conventional far-field beamforming, and *P2*, where all the APs transmit simultaneously for near-field beamfocusing. *(ii)* For both transmission protocols, we design the maximum likelihood (ML) estimators for position and phase offset estimation, considering both coherent and non-coherent processing approaches. A hybrid two-step optimization strategy is developed that capitalizes on non-

coherent processing to provide favorable initial estimates for high-resolution coherent processing. *(iii)* We carry out a performance evaluation of the designed estimators alongside the corresponding performance bounds, showcasing the potential of phase-coherent D-MIMO architectures for high-accuracy localization. Additionally, we observe that *P2* outperforms *P1* in both accuracy and latency.

II. System Model

We consider downlink of a D-MIMO system, as illustrated in Fig. 1, where M randomly deployed APs transmit narrowband signals, to enable the localization and synchronization of a user equipment (UE). The system operates at frequency f_0 with (small) bandwidth W.[1] The 2D location of the m-th AP is given as $\mathbf{p}_m = [x_m, y_m]^\top$, while the location of the UE is given as $\mathbf{u} = [u_x, u_y]^\top$. Each AP is equipped with a uniform linear array (ULA) of N antennas, spaced by $\lambda_0/2$, where $\lambda_0 = c/f_0$ is the carrier wavelength and c is the speed of light, while the UE has a single antenna. The orientation of AP m is denoted by $o_m \in [0, 2\pi)$, measured counterclockwise with respect to the x-axis. The APs are phase-synchronized, thus behaving like a distributed array [10], while the UE is assumed to have an unknown phase offset δ_ϕ, with respect to the AP network. The goal of the system is to estimate both the UE position $\mathbf{u}$ and the phase offset δ_ϕ. Without loss of generality, the UE is assumed to be in the far-field of each individual AP, but in the *near-field* of the complete AP network.

A. Channel Model

The channel between the m-th AP and the UE is represented by the vector $\mathbf{h}_m \in \mathbb{C}^{N \times 1}$, which is assumed herein to have a dominant LoS component, and is given as

$$\mathbf{h}_m = \alpha_m e^{j\delta_\phi} e^{-j2\pi f_0 \tau_m} \mathbf{a}(\theta_m) \triangleq \alpha_m e^{j\phi_m} \mathbf{a}(\theta_m). \quad (1)$$

Here, $\alpha_m \in \mathbb{R}$ denotes the large scale fading coefficient, modeled as $\alpha_m = \lambda_0/(4\pi\|\mathbf{u} - \mathbf{p}_m\|)$, $\tau_m = \|\mathbf{u} - \mathbf{p}_m\|/c$ is the signal time of arrival (TOA) from the m-th AP to the UE, $\mathbf{a}(\theta_m) \in \mathbb{C}^{N \times 1}$ represents the m-th AP array response vector, with angle of departure (AOD) $\theta_m = \arctan((y_m - u_y)/(x_m - u_x)) - o_m$. The n-th element of the array response vector is $[\mathbf{a}(\theta_m)]_n = \exp(j\pi n \sin\theta_m)$. In (1), $\phi_m \triangleq -2\pi f_0 \tau_m + \delta_\phi$ represents the carrier phase term, that accounts for the effects of signal propagation. It depends on the UE position $\mathbf{u}$ through τ_m, and the phase offset δ_ϕ.

B. Signal Transmission Protocols

The UE localization and synchronization problem considered here exploits multiple transmissions over time. The m-th AP transmits the symbol $s_m[k]$ at the k-th time instant, $k = 0, ..., K-1$, using the precoder $\mathbf{f}_m[k] \in \mathbb{C}^{N \times 1}$, where the transmitted symbol and the precoder satisfies the power constraint $\mathbb{E}\{|s_m[k]|^2\} = \mathbb{E}\{\|\mathbf{f}_m[k]\|^2\} = 1$. In this work, we analyze two different transmission protocols (see Fig. 1), referred to as *P1* and *P2*, which dictate how APs coordinate

their operation. These transmission protocols are discussed in detail below. For both protocols, we consider K illumination points, based on the prior knowledge of the UE location. We denote the k-th 2D illumination point by $\mathbf{i}[k]$.

1) Transmission Protocol P1: In this transmission protocol, the APs transmit their signals using time-division multiplexing (TDM), ensuring that at any given time instant, the UE receives the signal from only one AP, which requires K transmissions per AP. Hence, the total transmission duration is $T_{\text{P1}} = MKT$, where $T = 1/W$ represents the duration of a single transmission (or symbol). The signal received by the UE at time k from the m-th AP is given by

$$y_m[k] = \sqrt{P_m}\, \mathbf{h}_m^\top \mathbf{f}_m[k]s_m[k] + n_m[k]\,,$$

$$= \alpha_m e^{j\delta_\phi} x_m[k] + n_m[k] \quad (2)$$

for $k \in \mathcal{T}_m, m = 1, \ldots, M$, where P_m is the power transmitted by the m-th AP and $\mathbf{f}_m[k] = e^{j2\pi f_0 \tau_m[k]} \mathbf{a}^*(\theta_m[k])/\sqrt{N}$, where $\theta_m[k] = \arctan((y_m - i_y[k])/(x_m - i_x[k])) - o_m$ and $\tau_m[k] = \|\mathbf{i}[k] - \mathbf{p}_m\|/c$. The set $\mathcal{T}_m = \{m, m+M, m+2M, ..., m+(K-1)M\}$ denotes the set of time instants when the m-th AP transmits its signal. The scalar $n_m[k] \sim \mathcal{CN}(0, \sigma_n^2)$ is the additive white Gaussian noise (AWGN) at the k-th time instant within $\mathcal{T}_m$ and $x_m[k]$ represents the transmit signal after propagation through the channel, excluding the channel amplitude and the phase offset and is defined as

$$x_m[k] = \sqrt{P_m} e^{-j2\pi f_0 \tau_m} \mathbf{a}^\top(\theta_m)\mathbf{f}_m[k]s_m[k].$$

2) Transmission Protocol P2: In protocol *P2*, the AP network *coherently focuses* the signal on a desired spot within the target region. Over K transmissions, the illumination point is changed to cover the entire area of interest, resulting in a total transmission duration of $T_{\text{P2}} = KT$. This is in contrast to protocol *P1*, where the transmission time scales with the number of APs M, resulting in $T_{\text{P2}} \ll T_{\text{P1}}$, for a fixed K. The downlink signal received by the UE at the k-th time instant is

$$y[k] = \sum_{m=1}^{M} \sqrt{P_m}\, \mathbf{h}_m^\top \mathbf{f}_m[k]s_m[k] + n[k]\,,$$

$$= \sum_{m=1}^{M} \alpha_m e^{j\delta_\phi} x_m[k] + n[k]\,, \quad (3)$$

where the precoders $\mathbf{f}_m[k]$ are identical to those in *P1*.

III. Joint Localization and Synchronization

In this section, we estimate the 2D UE position $\mathbf{u}$ and the phase offset δ_ϕ for both the transmission protocols. The UE receives the signal $\mathbf{y} = [y[0], \ldots, y[S-1]]^\top \in \mathbb{C}^{S \times 1}$ over S transmissions, according to the respective transmission protocols (i.e., $S = MK$ for *P1* and $S = K$ for *P2*). Given the parameter vector to be estimated, $\boldsymbol{\eta} = [\mathbf{u}^\top, \delta_\phi, \boldsymbol{\alpha}^\top]^\top \in \mathbb{R}^{(M+3) \times 1}$, the joint localization and synchronization of the UE is formulated as an ML estimation problem

$$\widehat{\boldsymbol{\eta}} = \operatorname*{argmax}_{\boldsymbol{\eta}} \, \log p(\mathbf{y}|\boldsymbol{\eta}), \quad (4)$$

where $p(\mathbf{y}|\boldsymbol{\eta})$ represents the likelihood function. Notice that estimating the parameters of interest, $\mathbf{u}$ and δ_ϕ, also requires estimating the propagation amplitudes $\boldsymbol{\alpha} = [\alpha_1, \cdots, \alpha_M]^\top \in \mathbb{R}^{M \times 1}$, which act as nuisance parameters. For each trans-

mission protocol, we design two estimators: *(i)* a *coherent* estimator that exploits the carrier phase information; and *(ii)* a *non-coherent* estimator that does not use the carrier phase.

A. Transmission Protocol P1

1) Coherent ML Estimation: We formulate the UE localization and synchronization problem as the joint estimation problem which uses the carrier phase, i.e., the explicit dependence of the carrier phase ϕ_m on the UE position $\mathbf{u}$ through the TOA τ_m and phase offset δ_ϕ. The problem (4) becomes

$$\widehat{\boldsymbol{\eta}} = \arg \min_{\boldsymbol{\eta}} \sum_{m=1}^{M} \sum_{k \in \mathcal{T}_m} |y[k] - e^{j\delta_\phi} \alpha_m x_m[k]|^2. \quad (5)$$

We tackle this optimization problem by first solving for the optimal values of α_m for $m = 1, \ldots, M$. Differentiating the cost function in (5) with respect to α_m and equating it to zero, we get

$$\hat{\alpha}_m = \frac{\sum_{k \in \mathcal{T}_m} \mathcal{R}\{e^{j\delta_\phi} x_m[k] y^*[k]\}}{\sum_{k \in \mathcal{T}_m} |x_m[k]|^2} \triangleq \frac{\mathcal{R}\{e^{j\delta_\phi} \mathbf{y}_m^{\mathsf{H}} \mathbf{x}_m\}}{\|\mathbf{x}_m\|_2^2}, \quad (6)$$

where, the vectors $\mathbf{y}_m = [y[k]]_{k \in \mathcal{T}_m}$ and $\mathbf{x}_m = [x_m[k]]_{k \in \mathcal{T}_m}$ gathers the samples of the observation and signal model for the m-th AP, respectively. Note that $\hat{\alpha}_m$ is *conditional* on $\mathbf{u}$ and δ_ϕ. Substituting $\hat{\alpha}_m$ into (5), we obtain

$$\sum_{m=1}^{M} \sum_{k \in \mathcal{T}_m} \left| y[k] - e^{j\delta_\phi} \frac{\mathcal{R}\{e^{j\delta_\phi} \mathbf{y}_m^{\mathsf{H}} \mathbf{x}_m\}}{\|\mathbf{x}_m\|^2} x_m[k] \right|^2$$

$$= \sum_{m=1}^{M} \sum_{k \in \mathcal{T}_m} \left| \underbrace{y[k] - \frac{(\mathbf{x}_m^{\mathsf{H}} \mathbf{y}_m) x_m[k]}{2\|\mathbf{x}_m\|^2}}_{\widetilde{y}_m[k]} - e^{j2\delta_\phi} \underbrace{\frac{(\mathbf{y}_m^{\mathsf{H}} \mathbf{x}_m) x_m[k]}{2\|\mathbf{x}_m\|^2}}_{c_m[k]} \right|^2$$

$$= \sum_{m=1}^{M} \|\widetilde{\mathbf{y}}_m - e^{j2\delta_\phi} \mathbf{c}_m\|^2, \quad (7)$$

where $\widetilde{\mathbf{y}}_m = [\widetilde{y}_m[k]]_{k \in \mathcal{T}_m}$, and $\mathbf{c}_m = [c_m[k]]_{k \in \mathcal{T}_m}$. The value of δ_ϕ that minimizes the objective in (7) is given as $\hat{\delta}_\phi = \frac{1}{2} \angle \sum_{m=1}^{M} \mathbf{c}_m^{\mathsf{H}} \widetilde{\mathbf{y}}_m$, which remains conditional on $\mathbf{u}$. By substituting $\hat{\delta}_\phi$ in (7), the ML optimization problem becomes

$$\hat{\mathbf{u}} = \arg \min_{\mathbf{u}} \sum_{m=1}^{M} \|\widetilde{\mathbf{y}}_m(\mathbf{u}) - e^{j2\hat{\delta}_\phi(\mathbf{u})} \mathbf{c}_m(\mathbf{u})\|^2. \quad (8)$$

The optimal UE position cannot be obtained in closed form. A practical approach to solve (8) is to perform a 2D grid search.

2) Non-coherent ML Estimation: An alternative approach to solve the joint localization and synchronization problem is the non-coherent one, where the carrier phase ϕ_m is assumed to be an unknown (nuisance) parameter that does not depend on the UE position $\mathbf{u}$. In this case the channel amplitudes are treated as unstructured complex variables $\beta_m = \alpha_m e^{j\phi_m}$. The corresponding ML optimization problem is formulated as

$$[\hat{\mathbf{u}}, \hat{\boldsymbol{\beta}}] = \arg \min_{\{\mathbf{u}, \boldsymbol{\beta}\}} \sum_{m=1}^{M} \sum_{k \in \mathcal{T}_m} |y[k] - \beta_m z_m[k]|^2. \quad (9)$$

where $z_m[k] = \sqrt{P_m} \mathbf{a}^{\top}(\theta_m) \mathbf{f}_m[k] s_m[k]$ is the signal model excluding the complex amplitudes. The estimate of β_m conditional on $\mathbf{u}$ is given by $\hat{\beta}_m = \mathbf{z}_m^{\mathsf{H}} \mathbf{y}_m / \|\mathbf{z}_m\|^2$, where $\mathbf{z}_m = [z_m[k]]_{k \in \mathcal{T}_m}$ gathers the samples of the signal model for the m-th AP. By substituting $\hat{\beta}_m$ in (9), we obtain the
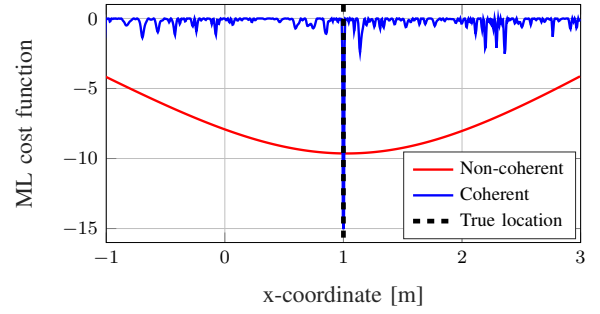

Figure 2: ML cost functions with respect to the x-coordinate of UE position in non-coherent and coherent processing for transmission protocol *P1*.

non-coherent ML estimator

$$\hat{\mathbf{u}} = \arg\min_{\mathbf{u}} \sum_{m=1}^{M} \left\| \mathbf{y}_m - \frac{\mathbf{z}_m(\mathbf{u})^{\mathsf{H}} \mathbf{y}_m}{\|\mathbf{z}_m(\mathbf{u})\|^2} \mathbf{z}_m(\mathbf{u}) \right\|^2, \quad (10)$$

which can be solved by a 2D grid search.

B. Transmission Protocol P2

1) Coherent ML Estimation: Similar to Section III-A1, we design the coherent ML estimator for UE localization and synchronization, by exploiting the carrier phase for transmission protocol *P2*. By gathering the signal models $x_m[k]$ $\forall k, m$ from (3) into matrix $\mathbf{X} \in \mathbb{C}^{K \times M}$ and defining $\mathbf{y} = [y[1], \cdots, y[K]]^{\top}$ and $\boldsymbol{\alpha} = [\alpha_1, \cdots, \alpha_M]^{\top}$, the ML estimation problem is

$$\widehat{\boldsymbol{\eta}} = \arg \min_{\boldsymbol{\eta}} \|\mathbf{y} - e^{j\delta_\phi} \mathbf{X} \boldsymbol{\alpha}\|_2^2,$$

$$= \arg \min_{\boldsymbol{\eta}} -2\mathcal{R}\{e^{-j\delta_\phi} \boldsymbol{\alpha}^{\top} \mathbf{X}^{\mathsf{H}} \mathbf{y}\} + \boldsymbol{\alpha}^{\top} \mathbf{X}^{\mathsf{H}} \mathbf{X} \boldsymbol{\alpha}. \quad (11)$$

By applying [12, Lemma S1], the second term in the objective of (11) simplifies to $\boldsymbol{\alpha}^{\top} \mathcal{R}\{\mathbf{X}^{\mathsf{H}} \mathbf{X}\} \boldsymbol{\alpha}$. Substituting this and using the fact that $\boldsymbol{\alpha}$ is a real vector, we rewrite (11) as

$$\widehat{\boldsymbol{\eta}} = \arg \min_{\boldsymbol{\eta}} \boldsymbol{\alpha}^{\top} \mathcal{R}\{\mathbf{X}^{\mathsf{H}} \mathbf{X}\} \boldsymbol{\alpha} - 2\mathcal{R}\{e^{-j\delta_\phi} \boldsymbol{\alpha}^{\top} \mathbf{X}^{\mathsf{H}} \mathbf{y}\}. \quad (12)$$

To obtain the optimal $\boldsymbol{\alpha}$, we differentiate the objective function with respect to $\boldsymbol{\alpha}$ and set it to zero. For a given δ_ϕ value, the optimal $\boldsymbol{\alpha}$ is $\hat{\boldsymbol{\alpha}} = \mathcal{R}\{\mathbf{X}^{\mathsf{H}} \mathbf{X}\}^{-1} \mathcal{R}\{e^{-j\delta_\phi} \mathbf{X}^{\mathsf{H}} \mathbf{y}\}$. Substituting $\hat{\boldsymbol{\alpha}}$ in the objective function in (11), we get

$$[\hat{\mathbf{u}}, \hat{\delta}_\phi] = \arg \min_{\{\mathbf{u}, \delta_\phi\}} \|\mathbf{v} - e^{j2\delta_\phi} \mathbf{w}\|^2, \quad (13)$$

where, $\mathbf{v} = \mathbf{y} - \mathbf{X} \mathcal{R}\{\mathbf{X}^H \mathbf{X}\}^{-1} \mathbf{X}^H \mathbf{y}/2$, and $\mathbf{w} = \mathbf{X} \mathcal{R}\{\mathbf{X}^H \mathbf{X}\}^{-1} \mathbf{X}^T \mathbf{y}^*/2$. The optimal δ_ϕ is obtained in closed form, and is $\hat{\delta}_\phi = \frac{\angle \mathbf{w}^H \mathbf{v}}{2}$. Substituting $\hat{\delta}_\phi$ in (13), we get

$$\hat{\mathbf{u}} = \arg \min_{\mathbf{u}} \|\mathbf{v} - e^{j2\hat{\delta}_\phi} \mathbf{w}\|^2 \quad (14)$$

To solve (14), a 2D exhaustive search is performed.

2) Non-coherent ML Estimation: For non-coherent ML estimation, similar to Section III-A2, the carrier phase ϕ_m is treated as an unknown parameter independent of the UE position, and the channel amplitudes are considered as unstructured complex scalar quantities $\beta_m = \alpha_m e^{j\phi_m}$. The corresponding ML optimization problem is $\min_{\mathbf{u}, \boldsymbol{\beta}} \|\mathbf{y} - \mathbf{Z}(\mathbf{u}) \boldsymbol{\beta}\|^2$ where $\boldsymbol{\beta} = [\beta_1, \cdots, \beta_M]^{\top} \in \mathbb{C}^{M \times 1}$, and matrix $\mathbf{Z}(\mathbf{u}) \in \mathbb{C}^{K \times M}$ gathers the model quantities $z_m[k]$. The optimal $\boldsymbol{\beta}$ can be readily derived as $\hat{\boldsymbol{\beta}} = (\mathbf{Z}^{\mathsf{H}} \mathbf{Z})^{-1} \mathbf{Z}^{\mathsf{H}} \mathbf{y}$. Substituting the optimal $\boldsymbol{\beta}$ in the ML cost function in (15), we obtain

$$\hat{\mathbf{u}} = \arg \min_{\mathbf{u}} \|\mathbf{y} - \mathbf{Z} (\mathbf{Z}^{\mathsf{H}} \mathbf{Z})^{-1} \mathbf{Z}^{\mathsf{H}} \mathbf{y}\|^2, \quad (15)$$

which can be solved by a 2D grid search.

C. Coherent vs. Non-Coherent ML Cost-Functions

In this section, we first analyze the differences between the coherent and non-coherent ML cost functions for transmission protocol *P1* (conclusions are similar for P2). Based on this analysis, we develop a low-complexity algorithm to estimate the UE position, combining the two cost functions.

1) ML Cost Function Analysis: Fig. 2 shows a 1D cut of coherent and non-coherent ML cost functions for protocol *P1*. As expected, leveraging the carrier phase yields a cost function with a more pronounced global minimum at the true UE location, as compared to its non-coherent counterpart. However, the coherent ML cost function features several local minima that prevents the use of conventional gradient descent techniques, whereas the non-coherent ML cost function is smoother, with a distinct global minimum but yield less accurate UE position estimates.

2) Proposed Hybrid Approach: To harness the strengths of both cost functions, we propose a two-stage localization algorithm. In the first stage, we search for the minimum of the non-coherent cost function. Since this cost-function is smooth, efficient gradient descent with a coarse initial guess is sufficient. This provides a good initial estimate, close to the UE position. In the second stage, we refine the estimate by searching for the minimum of the coherent ML cost function. We perform this search only within a small region around the initial estimate, using a finer grid.

IV. SIMULATION RESULTS

In this section, we perform numerical simulations to evaluate performances under both transmission protocols.

A. Scenario

We consider a network of $M = 20$ APs, each equipped with $N = 4$ antennas. The APs are placed uniformly along the edges of a 40 m × 40 m square centered at the origin, as shown in Fig. 3, with their antennas oriented towards the center. The APs precoded unity pilots (i.e., $s_m[k] = 1 \; \forall m, k$) to illuminate $K = 36$ illumination points. These illumination points are uniformly distributed over a 5 m × 5 m square region around the origin. The total transmit energy of the AP network remains fixed across both transmission protocols. The UE is positioned at $\mathbf{u} = [1, 1]^\top$ m and has a phase offset of $\delta_\phi = 10°$ relative to the AP network. The system operates at a carrier frequency of $f_0 = 3.5$ GHz. The total transmit energy of the network is fixed for both transmission protocols. The noise power is evaluated over a reference bandwidth of $W = 120$ kHz, corresponding to a single OFDM subcarrier spacing, and is given by $\sigma_n^2 = k_B T_{\text{tn}} W$, where $k_B = 1.38 \times 10^{-23}$ J/K is the Boltzmann constant, and $T_{\text{tn}} = 290$ K is the standard thermal noise temperature.

B. Algorithm Performance Assessment

We begin by evaluating the performance of the designed estimators for transmission protocols *P1* and *P2*. Fig. 4 presents the root mean squared error (RMSE) of the UE position estimate $\mathbf{u}$ for both protocols, along with their corresponding
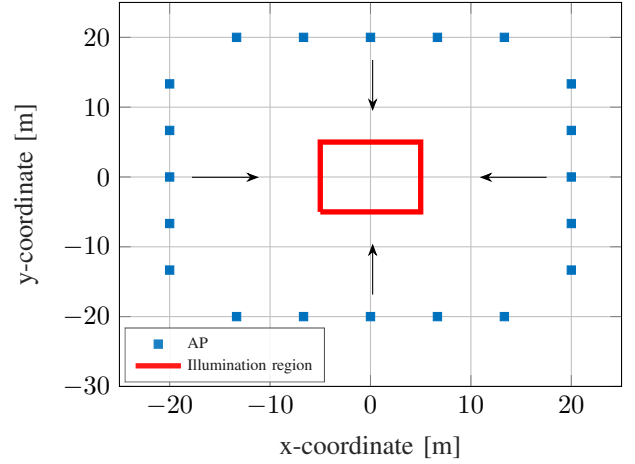

Figure 3: D-MIMO system deployment. The orientation of the AP antenna arrays are indicated by the arrows.
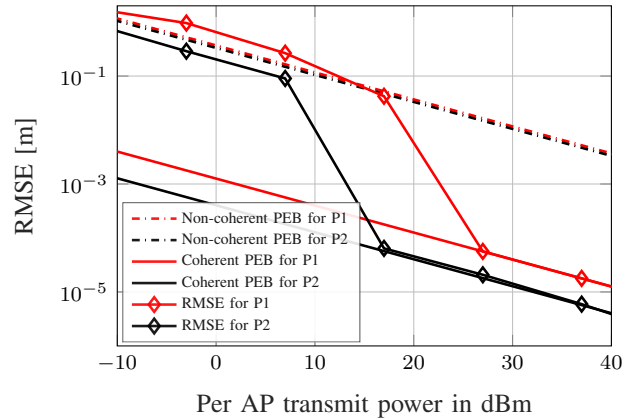

Figure 4: RMSE of the UE position estimate, and PEB for transmission protocols *P1* and *P2*.

position error bounds (PEBs), as a function of the per-AP transmit power. We see that both *P1* and *P2* achieve sub-meter accuracy even at low transmit power levels, with centimeter-level (or better) precision attained as the SNR increases. We further observe that protocol *P2*, where APs transmit simultaneously—resulting in lower latency—outperforms *P1* in both RMSE and PEB. Specifically, the PEB with *P2* is approximately 6 dB lower than that of *P1*. However, this improvement depends on the number of illumination points, which is set to $K = 36$ in this case. A more detailed analysis of the PEB as a function of the number of illumination points is provided later in Fig. 6. Additionally, as expected, the PEB for non-coherent processing is significantly higher than that for coherent processing—by approximately 24 dB for *P1* and 30 dB for *P2*. Furthermore, for non-coherent processing, the PEBs for both *P1* and *P2* are nearly identical for $K = 36$; however, for smaller K values, *P2* performs substantially worse than *P1*, as will be investigated next.

C. PEB Analysis

Fig. 5 shows the 1D PEB as a function of the x-coordinate of the UE position, for $y = 1$ m, considering both coherent and non-coherent processing for protocols *P1* and *P2*. The illumination region, which is indicated by the vertical dotted
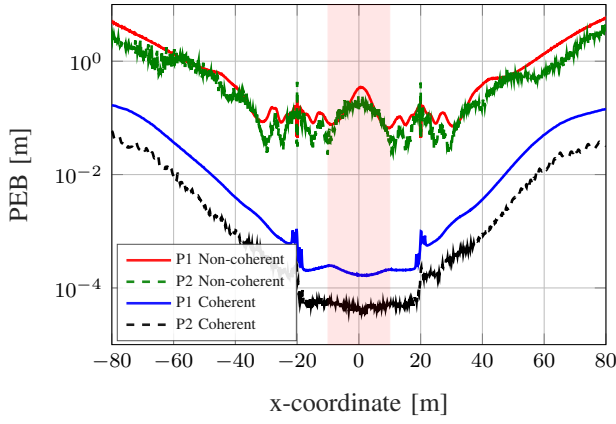
Figure 5: PEB values versus the x-coordinate of the UE position for transmission protocols *P1* and *P2* with coherent and non-coherent processing. The illumination region is shown as red shaded area.
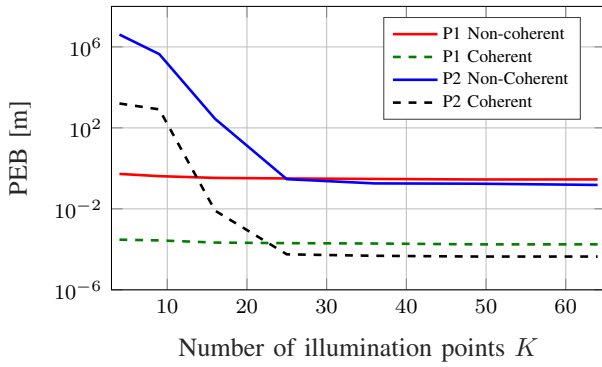

Figure 6: PEB for different number of illumination points (K).

lines in Fig. 5, represents the portion of space where APs focus (*P2*) or point (*P1*) their beams. We observe that for coherent processing, the PEB values for both *P1* and *P2* reach their minima at the center of the illumination region. In contrast, for non-coherent processing, the PEB at the center is higher than at the edges. This occurs because the beams formed by APs in non-coherent operation are much wider than those in coherent operation, resulting in limited spatial diversity at the center (illuminated by the mainlobe, unlike the edges illuminated by the sidelobes), which degrades positioning performance [13]. As expected, the PEB monotonically increases as the UE moves beyond the illumination region, since the available signal power decreases moving far away from the illumination region. Outside the illumination region, the received signal primarily originates from the side lobes of the beams, whose energy decreases with increasing distance.

An important insight into the performance of the two protocols emerges from examining how the PEB varies with the number of illumination points, K. Fig. 6 illustrates the average PEB in the illumination region as a function of the number of illumination points. We observe that, for *P1*, the PEB values remain nearly constant for any number of illumination points. Indeed, in *P1*, the APs transmit sequentially, generating broad beams. This covers the entire illumination region, even when the number of illumination points is small. In contrast, for *P2*, the PEB is significantly higher when the number of

illumination points is low. In this case, when the UE is far away from an illumination point, it receives a signal that is due to the sidelobes of the *coherent power pattern* generated by the AP network, that rapidly decrease around each illumination point. This is equivalent of having illumination gaps that, on average, increase the PEB. As the number of illumination points increases, these gaps progressively reduce and the PEB consequently decreases, stabilizing around a slightly smaller value compared to P1. The trend of the PEB with respect to the number of illumination points allows selecting the minimum value of K for which protocol *P2* is advantageous compared to *P1*, yielding similar PEB performance at a lower latency.

V. CONCLUSION

This paper investigated joint localization and synchronization in the downlink of a distributed MIMO system with phase-coherent access points. We developed maximum likelihood estimators for two transmission protocols—sequential (*P1*) and simultaneous (*P2*)—and analyzed their performance under coherent and non-coherent processing. Our results demonstrate that *P2* achieves superior localization accuracy and lower latency. Furthermore, we introduced a hybrid optimization approach to enhance estimation robustness. These findings highlight the potential of phase-coherent downlink D-MIMO for high-precision localization in future 6G networks.

REFERENCES

[1] J. A. del Peral-Rosado *et al.*, "Survey of cellular mobile radio localization methods: From 1G to 5G," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 1124–1148, 2018.

[2] F. Mogyorósi *et al.*, "Positioning in 5G and 6G networks—a survey," *Sensors*, vol. 22, no. 13, p. 4757, 2022.

[3] X. Li *et al.*, "Massive MIMO-based localization and mapping exploiting phase information of multipath components," *IEEE Transactions on Wireless Communications*, vol. 18, no. 9, pp. 4254–4267, 2019.

[4] M. A. Uusitalo *et al.*, "6G vision, value, use cases and technologies from european 6G flagship project Hexa-X," *IEEE Access*, vol. 9, pp. 160 004–160 020, 2021.

[5] H. Guo *et al.*, "Integrated communication, localization, and sensing in 6G D-MIMO networks," *arXiv preprint arXiv:2403.19785*, 2024.

[6] A. Behravan *et al.*, "Positioning and sensing in 6G: Gaps, challenges, and opportunities," *IEEE Vehicular Technology Magazine*, vol. 18, no. 1, pp. 40–48, 2023.

[7] N. Vukmirović *et al.*, "Position estimation with a millimeter-wave massive MIMO system based on distributed steerable phased antenna arrays," *EURASIP journal on advances in signal processing*, vol. 2018, pp. 1–17, 2018.

[8] ——, "Performance limits of direct wideband coherent 3D localization in distributed massive MIMO systems," *Sensors*, vol. 21, no. 10, p. 3401, 2021.

[9] ——, "Direct wideband coherent localization by distributed antenna arrays," *Sensors*, vol. 19, no. 20, p. 4582, 2019.

[10] A. Fascista *et al.*, "Joint localization, synchronization and mapping via phase-coherent distributed arrays," *IEEE Journal of Selected Topics in Signal Processing*, pp. 1–16, 2025.

[11] Ö. T. Demir *et al.*, "Foundations of user-centric cell-free massive MIMO," *Foundations and Trends® in Signal Processing*, vol. 14, no. 3-4, pp. 162–472, 2021.

[12] M. F. Keskin *et al.*, "Monostatic sensing with OFDM under phase noise: From mitigation to exploitation," *IEEE Transactions on Signal Processing*, vol. 71, pp. 1363–1378, 2023.

[13] Y. Zhang *et al.*, "Efficient indoor localization via switched-beam antennas," *IEEE Transactions on Mobile Computing*, vol. 19, no. 9, pp. 2101–2115, 2020.