RIS-Aided Joint Localization and Synchronization With a Single-Antenna Receiver: Beamforming Design and Low-Complexity Estimation

Alessio Fascista , *Member, IEEE*, Musa Furkan Keskin , *Member, IEEE*,
Angelo Coluccia , *Senior Member, IEEE*, Henk Wymeersch , *Senior Member, IEEE*,
and Gonzalo Seco-Granados , *Senior Member, IEEE*

Abstract—Reconfigurable intelligent surfaces (RISs) have at-tracted enormous interest thanks to their ability to overcome line-of-sight blockages in mmWave systems, enabling in turn accu-rate localization with minimal infrastructure. Less investigated are however the benefits of exploiting RIS with suitably designed beam-forming strategies for optimized localization and synchronization performance. In this paper, a novel low-complexity method for joint localization and synchronization based on an optimized design of the base station (BS) active precoding and RIS passive phase profiles is proposed, for the challenging case of a single-antenna receiver. The theoretical position error bound is first derived and used as metric to jointly optimize the BS-RIS beamforming, as-suming a priori knowledge of the user position. By exploiting the low-dimensional structure of the solution, a novel codebook-based robust design strategy with optimized beam power allocation is then proposed, which provides low-complexity while taking into account the uncertainty on the user position. Finally, a reduced-complexity maximum-likelihood based estimation procedure is de-vised to jointly recover the user position and the synchronization offset. Extensive numerical analysis shows that the proposed joint BS-RIS beamforming scheme provides enhanced localization and synchronization performance compared to existing solutions, with the proposed estimator attaining the theoretical bounds even at low signal-to-noise-ratio and in the presence of additional uncon-trollable multipath propagation.

I. INTRODUCTION

WITH the introduction of 5G, radio localization has finally been able to support industrial verticals and is no longer limited to emergency call localization [2]–[6]. This ability is enabled by a combination of wideband signals (up to 400 MHz in frequency range 2 (FR2)), higher carrier frequencies (e.g., around 28 GHz), multiple antennas, and a low latency and flexible architecture [3], [7], [8]. Common localization methods rely on time-difference-of-arrival (TDoA) or multi-cell round trip time (multi-RTT) measurements, requiring at least 4 or 3 base stations (BSs), respectively. In order to enable accurate localization with minimal infrastructure, there have been sev-eral studies to further reduce the number of BSs needed for localization. These studies can be broadly grouped in three categories: *(i)* data-driven, based on fingerprinting and deep learning [9], [10]; *(ii)* geometry-driven, based on exploiting passive multipath in the environment [11], [12] (which is itself derived from the multipath-assisted localization [13]); and, more recently, *(iii)* reconfigurable intelligent surface (RIS)-aided ap-proaches [14]–[20]. The latter category extends the concept of multipath-aided localization to RIS, which can actively control the multipath. RISs have attracted enormous interest in the past few years, mainly for their ability to overcome line-of-sight (LoS) blockages in mmWave communications [14], [21], [22]. From the localization point of view, RIS fundamentally offers two benefits: it introduces an extra location reference and provides additional measurements, independent of the pas-sive, uncontrolled multipath [16]. Hence, it avoids the reliance on strong reflectors in the environment, needed by standard multipath-aided localization [18], while also having the potential to low-complexity model based solution, in contrast to deep learning methods.

The use of RIS for localization has only recently been devel-oped, and a number of papers have been dedicated to RIS-aided localization [15]–[20], [23], [24]. Interestingly, RISs allow us to solve very challenging localization problems, such as single-antenna user equipment (UE) localization with a single-antenna BS in LoS [19] and even non-line-of-sight (NLoS) conditions (i.e., where the LoS path is blocked) [23]. While an RIS renders these problems solvable, high propagation losses (especially at

mmWave bands) necessitates long coherent processing intervals to obtain sufficient integrated signal-to-noise ratio (SNR), thus limiting supported mobility. Shorter integration times can be achieved with directional beamforming at the BS side [25], provided it is equipped with many antennas. Such beamforming becomes especially powerful when there exists a priori UE location information [26]. Hence, with the goal of improving localization performance, recent studies have focused on BS precoder optimization in the case of passive multipath [27], [28], while optimization in the presence of RIS involves *joint design of BS precoder and RIS phase profiles*, and thus can provide further accuracy enhancements via additional degrees of freedom. Nevertheless, such studies have been limited to SNR-maximizing heuristics [18], leading to *directional* RIS phase profiles, which may not necessarily lead to localization-optimal solutions [29], [30]. Within the context of *RIS-aided communications*, several works investigate joint design of active transmit precoding at the BS and passive phase shifts at the RIS to optimize various performance objectives, including sum-rate [31]–[34], effective mutual information [35], outage probability [36] and signal-to-interference-noise ratio (SINR) [37]. However, to the best of authors' knowledge, no studies have tackled the problem of joint BS-RIS beamforming to maximize the performance of *RIS-aided localization and synchronization*.

In this paper, we propose a novel joint BS-RIS beamforming design and a low-complexity maximum likelihood (ML) estimator for RIS-aided joint localization and synchronization supported by a single BS, considering the challenging case of a UE equipped with a single-antenna receiver. The optimized design exploits a priori UE location information and considers the BS precoders and RIS phase configurations jointly, in order to minimize the position error bound (PEB). The main contributions are as follows:

- We derive the Fisher Information Matrix (FIM) for localization and synchronization of a UE equipped with a single-antenna receiver, and conduct a theoretical analysis of the achievable performance.
- We formulate the joint design of BS precoder and RIS phase profile as a bi-convex optimization problem for the non-robust case, and propose a solution via alternating optimization. Interestingly, the solution reveals that at both the BS and RIS sides, a certain sequence of beams (namely, *directional* and *derivative* beams [29], [30], [38]) is required to render the problem feasible, in contrast to the corresponding communication problem.
- Based on the optimal solution under perfect knowledge of UE location, we propose a codebook-based design in the robust case, including a set of BS and RIS beams determined by the uncertainty region of UE location, where power optimization across BS beams is formulated as a convex problem.
- Elaborating on the ideas preliminarily introduced in [1], we devise a reduced-complexity estimation procedure based on the ML criterion, which attains the CRLBs even at low SNRs, and exhibits robustness against the presence of uncontrollable multipath.
- We compare the proposed algorithms against different approaches in literature, and show that the proposed designs outperform these benchmarks, not only in terms of PEB and clock error bound (CEB), but also localization and synchronization root-mean-squared errors (RMSEs).
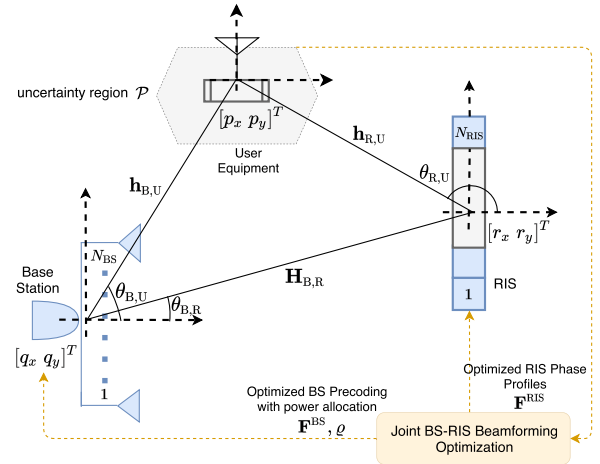

Fig. 1. Considered localization and synchronization scenario with optimized BS active precoding and RIS phase profiles.

II. SYSTEM MODEL AND PROBLEM FORMULATION

In this section, we describe the RIS-aided mmWave downlink (DL) localization scenario including a BS, an RIS and a UE, derive the received signal expression at the UE, and formulate the problem of joint localization and synchronization.

A. RIS-Aided Localization Scenario

We consider a DL localization scenario, as shown in Fig. 1, consisting of a single BS at known location $\boldsymbol{q} = [q_x \ q_y]^{\mathsf{T}}$ equipped with multiple antennas, a single-antenna UE at unknown location $\boldsymbol{p} = [p_x \ p_y]^{\mathsf{T}}$, and an RIS at known location $\boldsymbol{r} = [r_x \ r_y]^{\mathsf{T}}$. The UE has an unknown clock offset Δ with respect to the BS. We assume a two-dimensional (2D) scenario with uniform linear arrays (ULAs) for both the BS and RIS deployments.[1] The numbers of antenna elements at the BS and RIS are N_{BS} and N_{RIS}, respectively. The goal of the UE is to estimate its location and clock offset by exploiting the DL signals it receives through the direct LoS path and through the reflected (controllable) NLoS path generated by the RIS.

B. Signal Model

The BS communicates by transmitting single-stream orthogonal frequency division multiplexing (OFDM) pilots with N subcarriers over G transmissions. Particularly, the g-th transmission uses an OFDM symbol $\boldsymbol{s}_g = [s_g[0] \cdots [s_g[N-1]]^{\mathsf{T}} \in \mathbb{C}^{N \times 1}$ with $\frac{1}{N}\|\boldsymbol{s}_g\|^2 = 1$ and is precoded by the weight vector $\boldsymbol{f}_g \in \mathbb{C}^{N_{\mathrm{BS}} \times 1}$. To keep the transmit energy constant over the entire transmission period, the precoding matrix $\boldsymbol{F} = [\boldsymbol{f}_1 \cdots \boldsymbol{f}_G] \in \mathbb{C}^{N_{\mathrm{BS}} \times G}$ is assumed to satisfy $\mathrm{tr}(\boldsymbol{F}\boldsymbol{F}^H) = 1$. The BS can be

---

[1] We address the joint localization and synchronization problem in 2D, a common choice in the literature because it greatly simplifies the exposition. Moreover, since in mmWave scenarios the distance between transmitter and receiver is large compared to the height of the antennas, considering the projection onto the 2D horizontal plane provides a fairly realistic representation of the dominant propagation phenomena. The proposed methodology can be extended in principle to address the 3D localization setup; we will discuss this possibility after the derivation of the proposed joint BS and RIS beamforming design strategy in Section V and low-complexity estimation algorithm in Section VI, so that the necessary modifications can be described.

equipped with an analog active phased array [39] or a fully digital array.

The DL received signal at the UE associated to the g-th transmission over subcarrier n is given by

$$y^g[n] = \sqrt{P} \boldsymbol{h}^{\mathsf{T}}[n] \boldsymbol{f}_g s_g[n] + \nu^g[n] \qquad (1)$$

for $n = 0, \ldots, N - 1$ and $g = 1, \ldots, G$, with P denoting the transmit power and $\nu^g[n]$ circularly symmetric complex Gaussian noise having zero mean and variance σ^2. In (1), $\boldsymbol{h}[n] \in \mathbb{C}^{N_{\mathrm{BS}} \times 1}$ represents the entire channel, including both the LoS path and the NLoS path (i.e., the reflection path via the RIS), between the BS and the UE for the n-th subcarrier:

$$\boldsymbol{h}^{\mathsf{T}}[n] = \boldsymbol{h}_{\mathrm{B,U}}^{\mathsf{T}}[n] + \boldsymbol{h}_{\mathrm{R,U}}^{\mathsf{T}}[n] \boldsymbol{\Omega}^g \boldsymbol{H}_{\mathrm{B,R}}[n] \qquad (2)$$

where $\boldsymbol{h}_{\mathrm{B,U}}[n] \in \mathbb{C}^{N_{\mathrm{BS}} \times 1}$ is the direct (i.e., LoS) channel between the BS and the UE, $\boldsymbol{H}_{\mathrm{B,R}}[n] \in \mathbb{C}^{N_{\mathrm{RIS}} \times N_{\mathrm{BS}}}$ denotes the channel from the BS to the RIS, $\boldsymbol{\Omega}^g = \mathrm{diag}\left(e^{j\omega_1^g}, \ldots, e^{j\omega_{N_{\mathrm{RIS}}}^g}\right) \in \mathbb{C}^{N_{\mathrm{RIS}} \times N_{\mathrm{RIS}}}$ is the RIS phase control matrix at transmission g, and $\boldsymbol{h}_{\mathrm{R,U}}[n] \in \mathbb{C}^{N_{\mathrm{RIS}} \times 1}$ represents the channel from the RIS to the UE.[2]

The LoS channel in (2) can be expressed as $\boldsymbol{h}_{\mathrm{B,U}}[n] = \alpha_{\mathrm{B,U}} e^{-j2\pi n \frac{\tau_{\mathrm{B,U}}}{NT}} \boldsymbol{a}_{\mathrm{BS}}(\theta_{\mathrm{B,U}})$, where $T = 1/B$ is the sampling period with B the bandwidth, $\alpha_{\mathrm{B,U}} = \rho_{\mathrm{B,U}} e^{j\varphi_{\mathrm{B,U}}}$ with $\rho_{\mathrm{B,U}}$ and $\varphi_{\mathrm{B,U}}$ modulus and phase of the complex amplitude $\alpha_{\mathrm{B,U}}$, $\theta_{\mathrm{B,U}}$ is the angle-of-departure (AoD) from the BS to the UE, and $\tau_{\mathrm{B,U}}$ is the delay between the BS and the UE up to a clock offset Δ, as better specified later. As to $\boldsymbol{a}_{\mathrm{BS}}(\cdot)$, it represents the BS array steering vector whose expression is given by $\boldsymbol{a}_{\mathrm{BS}}(\theta) = \left[1 \; e^{j\frac{2\pi}{\lambda_c} d \sin\theta} \; \cdots \; e^{j(N_{\mathrm{BS}}-1)\frac{2\pi}{\lambda_c} d \sin\theta}\right]^{\mathsf{T}}$ with $\lambda_c = c/f_c$, f_c being the carrier frequency and c the speed of light, and $d = \lambda_c/2$.

In (2), the first tandem channel (i.e., from the BS to the RIS) in the NLoS path through the RIS is defined as

$$\boldsymbol{H}_{\mathrm{B,R}}[n] = \alpha_{\mathrm{B,R}} e^{-j2\pi n \frac{\tau_{\mathrm{B,R}}}{NT}} \boldsymbol{a}_{\mathrm{RIS}}(\phi_{\mathrm{B,R}}) \boldsymbol{a}_{\mathrm{BS}}^{\mathsf{T}}(\theta_{\mathrm{B,R}}) \qquad (3)$$

where $\alpha_{\mathrm{B,R}} = \rho_{\mathrm{B,R}} e^{j\varphi_{\mathrm{B,R}}}$ is the complex gain over the BS-RIS path, $\phi_{\mathrm{B,R}}$ the angle-of-arrival (AoA) and $\theta_{\mathrm{B,R}}$ the AoD from the BS to the RIS, and $\tau_{\mathrm{B,R}}$ the delay between the BS and the RIS. In addition, $\boldsymbol{a}_{\mathrm{RIS}}(\cdot) \in \mathbb{C}^{N_{\mathrm{RIS}} \times 1}$ denotes the array steering vector of the RIS, given by $\boldsymbol{a}_{\mathrm{RIS}}(\theta) = \left[1 \; e^{j\frac{2\pi}{\lambda_c} d \sin\theta} \; \cdots \; e^{j(N_{\mathrm{RIS}}-1)\frac{2\pi}{\lambda_c} d \sin\theta}\right]^{\mathsf{T}}$. Finally, the second tandem channel in (2) is given by

$$\boldsymbol{h}_{\mathrm{R,U}}^{\mathsf{T}}[n] = \alpha_{\mathrm{R,U}} e^{-j2\pi n \frac{\tau_{\mathrm{R,U}}}{NT}} \boldsymbol{a}_{\mathrm{RIS}}^{\mathsf{T}}(\theta_{\mathrm{R,U}}) \qquad (4)$$

with the notations $\alpha_{\mathrm{R,U}} = \rho_{\mathrm{R,U}} e^{j\varphi_{\mathrm{R,U}}}$, $\tau_{\mathrm{R,U}}$, and $\theta_{\mathrm{R,U}}$ having the same meaning as in the BS-to-UE channel model.

The geometric relationships among the BS, RIS, and UE are as follows (assuming for simplicity that the BS is placed at the origin of the reference system, i.e., $\boldsymbol{q} = [0 \; 0]^{\mathsf{T}}$):

$$\tau_{\mathrm{B,U}} = \|\boldsymbol{p}\|/c + \Delta$$

$$\tau_{\mathrm{R}} = \tau_{\mathrm{B,R}} + \tau_{\mathrm{R,U}} = (\|\boldsymbol{r}\| + \|\boldsymbol{r} - \boldsymbol{p}\|)/c + \Delta$$

$$\theta_{\mathrm{B,U}} = \mathrm{atan2}(p_y, p_x), \quad \theta_{\mathrm{R,U}} = \mathrm{atan2}(p_y - r_y, p_x - r_x)$$

$$\theta_{\mathrm{B,R}} = \mathrm{atan2}(r_y, r_x), \quad \phi_{\mathrm{B,R}} = -\pi + \theta_{\mathrm{B,R}}. \qquad (5)$$

---

[2]For simplicity, the model in (2) does not involve uncontrolled multipath components and will be used to derive bounds and algorithms from Section III to Section VI, while the robustness of the algorithms against the presence of uncontrolled NLoS paths will be evaluated through simulations in Section VII-C5.

Notice that $\tau_{\mathrm{B,R}}$, $\theta_{\mathrm{B,R}}$ and $\phi_{\mathrm{B,R}}$ are known quantities being the BS and RIS placed at known positions.

C. Joint Localization and Synchronization Problem

From the DL received signal $\{y^g[n]\}_{\forall n, g}$ in (1) over N subcarriers and G transmissions, the problems of interest are as follows: i) design the BS precoder matrix $\boldsymbol{F}$ and the RIS phase profiles $\{\boldsymbol{\Omega}^g\}_{\forall g}$ to maximize the accuracy of UE location and clock offset estimation; ii) estimate the unknown location $\boldsymbol{p}$ and the unknown clock offset Δ of the UE. To tackle these problems, we first derive a performance metric to quantify the accuracy of localization and synchronization in Section III. Based on this metric, Sections IV-V focus on the joint design of $\boldsymbol{F}$ and $\{\boldsymbol{\Omega}^g\}_{\forall g}$. Finally, Section VI develops an estimator for $\boldsymbol{p}$ and Δ.

III. FISHER INFORMATION ANALYSIS

In this section, we perform a Fisher information analysis to obtain a performance measure for localization and synchronization of the UE, which is needed for the design of $\boldsymbol{F}$ and $\{\boldsymbol{\Omega}^g\}_{\forall g}$ in Section IV and Section V.

A. FIM in the Channel Domain

For the estimation problem in Section II-C, we compute the FIM of the unknown channel parameter vector $\boldsymbol{\gamma} = [\tau_{\mathrm{B,U}} \; \theta_{\mathrm{B,U}} \; \rho_{\mathrm{B,U}} \; \varphi_{\mathrm{B,U}} \; \tau_{\mathrm{R}} \; \theta_{\mathrm{R,U}} \; \rho_{\mathrm{R}} \; \varphi_{\mathrm{R}}]^{\mathsf{T}}$ where $\rho_{\mathrm{R}} = \rho_{\mathrm{B,R}} \rho_{\mathrm{R,U}}$ and $\varphi_{\mathrm{R}} = \varphi_{\mathrm{B,R}} + \varphi_{\mathrm{R,U}}$. The FIM $\boldsymbol{J}_{\boldsymbol{\gamma}} \in \mathbb{R}^{8 \times 8}$ satisfies the information inequality [40, Thm. (3.2)]

$$\mathbb{E}\left\{(\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma})(\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma})^{\mathsf{T}}\right\} \succeq \boldsymbol{J}_{\boldsymbol{\gamma}}^{-1} \qquad (6)$$

for any unbiased estimator $\widehat{\boldsymbol{\gamma}}$ of $\boldsymbol{\gamma}$, where $\boldsymbol{A} \succeq \boldsymbol{B}$ means $\boldsymbol{A} - \boldsymbol{B}$ is positive semi-definite. Since the observations in (1) are complex Gaussian, the (h, k)-th FIM entry $[\boldsymbol{J}_{\boldsymbol{\gamma}}]_{h,k} \stackrel{\text{def}}{=} \Lambda(\gamma_h, \gamma_k)$ can be expressed using the Slepian-Bangs formula as [40, Eq. (15.52)]

$$\Lambda(\gamma_h, \gamma_k) = \frac{2}{\sigma^2} \sum_{g=1}^{G} \sum_{n=0}^{N-1} \Re\left\{\left(\frac{\partial m^g[n]}{\partial \gamma_h}\right)^* \frac{\partial m^g[n]}{\partial \gamma_k}\right\} \qquad (7)$$

where x^* denotes the complex conjugate of x and $m^g[n] = \sqrt{P} \boldsymbol{h}^{\mathsf{T}}[n] \boldsymbol{f}_g s_g[n]$ is the noise-free version of the received signal in (1). Using (2)–(4), $m^g[n]$ can be re-written as $m^g[n] = m_{\mathrm{B,U}}^g[n] + m_{\mathrm{R}}^g[n]$, where

$$m_{\mathrm{B,U}}^g[n] \stackrel{\text{def}}{=} \sqrt{P} \rho_{\mathrm{B,U}} e^{j\varphi_{\mathrm{B,U}}} \left[\boldsymbol{c}(\tau_{\mathrm{B,U}})\right]_n \boldsymbol{a}_{\mathrm{BS}}^{\mathsf{T}}(\theta_{\mathrm{B,U}}) \boldsymbol{f}_g s_g[n]$$

$$m_{\mathrm{R}}^g[n] \stackrel{\text{def}}{=} \sqrt{P} \rho_{\mathrm{R}} e^{j\varphi_{\mathrm{R}}} \left[\boldsymbol{c}(\tau_{\mathrm{R}})\right]_n \boldsymbol{b}_{\mathrm{RIS}}^{\mathsf{T}}(\theta_{\mathrm{R,U}}) \boldsymbol{\omega}^g \boldsymbol{a}_{\mathrm{BS}}^{\mathsf{T}}(\theta_{\mathrm{B,R}}) \boldsymbol{f}_g s_g[n]. \qquad (8)$$

In (8), $\tau_{\mathrm{R}} = \tau_{\mathrm{R,U}} + \tau_{\mathrm{B,R}}$ is the delay of the BS-RIS-UE path,

$$\boldsymbol{b}_{\mathrm{RIS}}(\theta) \stackrel{\text{def}}{=} \boldsymbol{a}_{\mathrm{RIS}}(\theta) \odot \boldsymbol{a}_{\mathrm{RIS}}(\phi_{\mathrm{B,R}}) \qquad (9)$$

denotes the combined RIS steering vector including the effect of both the AoD θ and the AoA $\phi_{\mathrm{B,R}}$ as a function of θ,

$$\boldsymbol{c}(\tau) \stackrel{\text{def}}{=} \left[1 \; e^{-j\kappa_1 \tau} \; \cdots \; e^{-j\kappa_{N-1}\tau}\right]^{\mathsf{T}} \qquad (10)$$

represents the frequency-domain steering vector with $\kappa_n = 2\pi \frac{n}{NT}$, and $\boldsymbol{\omega}^g \in \mathbb{C}^{N_{\mathrm{RIS}} \times 1}$ is the vector consisting of the diagonal entries of $\boldsymbol{\Omega}^g$, i.e., $\boldsymbol{\Omega}^g = \mathrm{diag}(\boldsymbol{\omega}^g)$. Here, $\odot$ is the Hadamard (element-wise) product. Hereafter, $\boldsymbol{b}_{\mathrm{RIS}}$ will be used to denote $\boldsymbol{b}_{\mathrm{RIS}}(\theta_{\mathrm{R,U}})$ for the sake of brevity. For the derivative expressions

in (7), we refer the reader to Section S-I-A in the supplemental material.

We now express the FIM elements in (7) as a function of the BS precoder $\boldsymbol{F}$ and the RIS phase profiles $\{\boldsymbol{\omega}^g\}_{g=1}^G$. To that end, the FIM $\boldsymbol{J}_\gamma$ can be written as

$$\boldsymbol{J}_\gamma = \begin{bmatrix} \boldsymbol{J}_{\text{B,U}} & \boldsymbol{J}_{\text{cross}} \\ \boldsymbol{J}_{\text{cross}}^{\mathsf{T}} & \boldsymbol{J}_{\text{R}} \end{bmatrix} \quad (11)$$

where $\boldsymbol{J}_{\text{B,U}} \in \mathbb{R}^{4\times4}$ and $\boldsymbol{J}_{\text{R}} \in \mathbb{R}^{4\times4}$ are the FIM submatrices corresponding to the LoS path and the NLoS (i.e., BS-RIS-UE) path, respectively, and $\boldsymbol{J}_{\text{cross}} \in \mathbb{R}^{4\times4}$ represents the LoS-NLoS path cross-correlation. In addition, let us define

$$\boldsymbol{X}_g \overset{\text{def}}{=} \boldsymbol{f}_g \boldsymbol{f}_g^{\mathsf{H}} \in \mathbb{C}^{N_{\text{BS}}\times N_{\text{BS}}} \quad (12)$$

$$\boldsymbol{\Psi}_g \overset{\text{def}}{=} \boldsymbol{\omega}^g(\boldsymbol{\omega}^g)^{\mathsf{H}} \in \mathbb{C}^{N_{\text{RIS}}\times N_{\text{RIS}}}. \quad (13)$$

The following remark reveals the dependency of the FIM submatrices in (11) on BS precoder and RIS phase profiles using Section S-I-B in the supplemental material.

Remark 1: The dependency of the FIM $\boldsymbol{J}_\gamma$ in (11) on $\boldsymbol{F}$ and $\{\boldsymbol{\omega}^g\}_{g=1}^G$ can be specified as follows:
- $\boldsymbol{J}_{\text{B,U}}$ is a linear function of $\{\boldsymbol{X}_g\}_{g=1}^G$.
- $\boldsymbol{J}_{\text{R}}$ is a bi-linear function of $\{\boldsymbol{X}_g\}_{g=1}^G$ and $\{\boldsymbol{\Psi}_g\}_{g=1}^G$.
- $\boldsymbol{J}_{\text{cross}}$ is a bi-linear function of $\{\boldsymbol{X}_g\}_{g=1}^G$ and $\{\boldsymbol{\omega}^g\}_{g=1}^G$.

B. FIM in the Location Domain

To obtain the location-domain FIM from the channel-domain FIM $\boldsymbol{J}_\gamma$ in (11), we apply a transformation of variables from the vector of the unknown channel parameters $\boldsymbol{\gamma}$ to the vector of location parameters[3]

$$\boldsymbol{\eta} = [p_x \; p_y \; \rho_{\text{B,U}} \; \varphi_{\text{B,U}} \; \rho_{\text{R}} \; \varphi_{\text{R}} \; \Delta]^{\mathsf{T}}. \quad (14)$$

The FIM of $\boldsymbol{\eta}$, denoted as $\boldsymbol{J}_{\boldsymbol{\eta}} \in \mathbb{R}^{7\times7}$, is obtained by means of the transformation matrix $\boldsymbol{T} \overset{\text{def}}{=} \frac{\partial\boldsymbol{\gamma}^{\mathsf{T}}}{\partial\boldsymbol{\eta}} \in \mathbb{R}^{7\times8}$ as

$$\boldsymbol{J}_{\boldsymbol{\eta}} = \boldsymbol{T}\boldsymbol{J}_\gamma\boldsymbol{T}^{\mathsf{T}}, \quad (15)$$

which preserves the linearity and bi-linearity properties of $\boldsymbol{J}_\gamma$ in Remark 1. Please see Section S-II in the supplemental material for the expressions of the elements of $\boldsymbol{T}$.

IV. JOINT TRANSMIT PRECODING AND RIS PHASE PROFILE DESIGN

In this section, assuming *perfect knowledge* of $\boldsymbol{\eta}$ in (14), we tackle the problem of joint design of the transmit BS precoding matrix $\boldsymbol{F}$ and the RIS phase profiles $\{\boldsymbol{\omega}^g\}_{g=1}^G$ to maximize the performance of joint localization and synchronization of the UE. First, we apply convex relaxation and alternating optimization techniques to obtain two convex subproblems to optimize BS precoders for a given RIS phase profile and vice versa. Then, we demonstrate the *low-dimensional structure* of the optimal BS precoders and RIS phase profiles, which will be highly instrumental in Section V in designing codebooks under imperfect knowledge of UE location.

A. Problem Formulation for Joint Optimization

To formulate the joint BS precoding and RIS phase profile optimization problem, we adopt the position error bound (PEB) as metric.[4] From (14)–(15) and Remark 1, the PEB can be obtained as a function of the BS beam covariance matrices, the RIS phase profiles and their covariance matrices $\{\boldsymbol{X}_g, \boldsymbol{\omega}^g, \boldsymbol{\Psi}_g\}_{g=1}^G$, as follows:

$$\mathbb{E}\left\{\|\widehat{\boldsymbol{p}}-\boldsymbol{p}\|^2\right\} \geq \text{tr}\left([\boldsymbol{J}_{\boldsymbol{\eta}}^{-1}]_{1:2,1:2}\right)$$
$$\overset{\text{def}}{=} \text{PEB}\left(\{\boldsymbol{X}_g, \boldsymbol{\omega}^g, \boldsymbol{\Psi}_g\}_{g=1}^G; \boldsymbol{\eta}\right). \quad (16)$$

We note that the PEB depends on the unknown parameters $\boldsymbol{\eta}$ (see Section S-I-B in the supplemental material). In addition, the dependency of the PEB on $\{\boldsymbol{X}_g, \boldsymbol{\omega}^g, \boldsymbol{\Psi}_g\}_{g=1}^G$ can be observed through (15), (16) and Remark 1. Under perfect knowledge of $\boldsymbol{\eta}$, the PEB minimization problem can be formulated as

$$\min_{\{\boldsymbol{X}_g,\boldsymbol{\omega}^g,\boldsymbol{\Psi}_g\}_{g=1}^G} \text{PEB}\left(\{\boldsymbol{X}_g, \boldsymbol{\omega}^g, \boldsymbol{\Psi}_g\}_{g=1}^G; \boldsymbol{\eta}\right) \quad (17a)$$

$$\text{s.t.} \quad \text{tr}\left(\sum_{g=1}^G \boldsymbol{X}_g\right) = 1, \quad (17b)$$

$$\boldsymbol{X}_g \succeq 0, \; \text{rank}(\boldsymbol{X}_g) = 1, \quad (17c)$$

$$\boldsymbol{\Psi}_g = \boldsymbol{\omega}^g(\boldsymbol{\omega}^g)^{\mathsf{H}}, \; |\omega_g[n]| = 1,$$
$$g = 1, \ldots, G, \quad (17d)$$

where the total power constraint in (17b) is due to Section II-B, (17c) results from (12), and (17d) comes from the definition in (13) and the unit-modulus constraints on the elements of the RIS control matrix. We now provide the following lemma on the structure of the objective function in (17a).

Lemma 1: PEB$(\{\boldsymbol{X}_g, \boldsymbol{\omega}^g, \boldsymbol{\Psi}_g\}_{g=1}^G; \boldsymbol{\eta})$ is a multi-convex function of $\{\boldsymbol{X}_g\}_{g=1}^G$, $\{\boldsymbol{\omega}^g\}_{g=1}^G$ and $\{\boldsymbol{\Psi}_g\}_{g=1}^G$.

Proof: From Remark 1 and (15), we see that $\boldsymbol{J}_{\boldsymbol{\eta}}$ is a multi-linear function of $\{\boldsymbol{X}_g\}_{g=1}^G$, $\{\boldsymbol{\omega}^g\}_{g=1}^G$ and $\{\boldsymbol{\Psi}_g\}_{g=1}^G$. Based on the composition rules [42, Ch. 3.2.4], $\text{tr}\left([\boldsymbol{J}_{\boldsymbol{\eta}}^{-1}]_{1:2,1:2}\right)$ is a multi-convex function of $\{\boldsymbol{X}_g\}_{g=1}^G$, $\{\boldsymbol{\omega}^g\}_{g=1}^G$ and $\{\boldsymbol{\Psi}_g\}_{g=1}^G$. ∎

To clarify the use of $\{\boldsymbol{X}_g, \boldsymbol{\omega}^g, \boldsymbol{\Psi}_g\}_{g=1}^G$ instead of $\{\boldsymbol{f}_g, \boldsymbol{\omega}^g\}_{g=1}^G$ in (17) as the optimization variables, the following remark is provided.

Remark 2: While our goal is to optimize the BS precoders $\boldsymbol{f}_g$ and the RIS phase profiles $\boldsymbol{\omega}^g$, we employ the covariances $\boldsymbol{X}_g = \boldsymbol{f}_g\boldsymbol{f}_g^{\mathsf{H}}$ and $\boldsymbol{\Psi}_g = \boldsymbol{\omega}^g(\boldsymbol{\omega}^g)^{\mathsf{H}}$ as the optimization variables in (17). The reason is that the FIM $\boldsymbol{J}_\gamma$ in (11) is *linear* in $\boldsymbol{X}_g$ and $\boldsymbol{\Psi}_g$, but *quadratic* in $\boldsymbol{f}_g$ and $\boldsymbol{\omega}^g$, as seen from Section S-I-B in the supplemental material. In particular,
- all the submatrices in (11) are linear in $\boldsymbol{X}_g$, but quadratic in $\boldsymbol{f}_g$,
- $\boldsymbol{J}_{\text{cross}}$ is linear in $\boldsymbol{\omega}^g$, and $\boldsymbol{J}_{\text{R}}$ is linear in $\boldsymbol{\Psi}_g$, but quadratic in $\boldsymbol{\omega}^g$,

---

[3]Note that the channel gains $\rho_{\text{B,U}}$, $\varphi_{\text{B,U}}$, ρ_{R} and φ_{R} are nuisance parameters that need to be estimated for localization, but do not convey any geometric information that can be useful for localization. Hence, they cannot be expressed as a function of other unknown (geometric) parameters and thus appear in both channel and location domain parameter vectors.

[4]Since positioning and synchronization are tightly coupled [41], considering PEB as the optimization metric would also improve the synchronization performance, which will be verified through simulation results in Section VII. In particular, please see Fig. 6 for an illustration of how PEB-based optimization provides noticeable improvements in the RMSE of both the position and clock offset over the benchmark schemes. For a more detailed comparison between PEB- and CEB-based optimization, we refer the reader to Section S-III in the supplemental material.

as specified in Lemma 1. For PEB minimization, we wish to keep the variables for which the dependencies are linear and discard the remaining ones. This results from the fact that the PEB minimization problem, when written in the epigraph form as will be shown in (27), induces a matrix inequality (MI) constraint that involves the FIM, such as in (27b), which is convex only if the MI is linear [42, Ch. 4.6.2], [43]. Therefore, to have a convex problem, the FIM needs to depend linearly on the optimization variables. This implies that we should keep $\boldsymbol{X}_g$, $\boldsymbol{\Psi}_g$ and $\boldsymbol{\omega}^g$ as the variables in (17), where $\boldsymbol{J}_{\text{cross}}$ is defined as a linear function of $\boldsymbol{\omega}^g$, $\boldsymbol{J}_{\text{R}}$ is defined as a linear function of $\boldsymbol{\Psi}_g$ and the coupling between the two variables $\boldsymbol{\Psi}_g = \boldsymbol{\omega}^g(\boldsymbol{\omega}^g)^{\mathsf{H}}$ is imposed as a constraint in (17d).

B. Relaxed Problem for PEB Minimization

To transform (17) into a tractable form, we will perform two simplifications. First, we approximate the channel-domain FIM in (11) as a block-diagonal matrix, i.e.,

$$\boldsymbol{J}_{\boldsymbol{\gamma}} \approx \boldsymbol{J}_{\boldsymbol{\gamma}}^{\text{bd}} \overset{\text{def}}{=} \begin{bmatrix} \boldsymbol{J}_{\text{B,U}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{J}_{\text{R}} \end{bmatrix}, \qquad (18)$$

by assuming $\boldsymbol{J}_{\text{cross}} \approx \boldsymbol{0}$, which can be justified by the assumption of non-interfering paths under the large bandwidth and large array regime [44]–[46]. Based on Remark 1, this enables removing the dependency of $\boldsymbol{J}_{\boldsymbol{\gamma}}$ and, consequently, the PEB in (16) on $\{\boldsymbol{\omega}^g\}_{g=1}^G$. In this case, the constraint in (17d) should be replaced by

$$\boldsymbol{\Psi}_g \succeq 0 \,, \ \text{rank}(\boldsymbol{\Psi}_g) = 1 \,, \ \text{diag}\,(\boldsymbol{\Psi}_g) = \boldsymbol{1}. \qquad (19)$$

Second, we drop non-convex rank constraints in (17c) and (19).

After these simplifications,[5] a relaxed version of the PEB optimization problem in (17) can be cast as

$$\min_{\{\boldsymbol{X}_g, \boldsymbol{\Psi}_g\}_{g=1}^G} \quad \text{PEB}^{\text{bd}}\left(\{\boldsymbol{X}_g, \boldsymbol{\Psi}_g\}_{g=1}^G; \boldsymbol{\eta}\right) \qquad (20\text{a})$$

$$\text{s.t.} \quad (17\text{b})\,, \ \boldsymbol{X}_g \succeq 0, \qquad (20\text{b})$$

$$\boldsymbol{\Psi}_g \succeq 0 \,, \ \text{diag}\,(\boldsymbol{\Psi}_g) = \boldsymbol{1},$$

$$g = 1, \ldots, G, \qquad (20\text{c})$$

where $\text{PEB}^{\text{bd}}(\{\boldsymbol{X}_g, \boldsymbol{\Psi}_g\}_{g=1}^G; \boldsymbol{\eta}) \overset{\text{def}}{=} \text{tr}([(\boldsymbol{T}\boldsymbol{J}_{\boldsymbol{\gamma}}^{\text{bd}}\boldsymbol{T}^{\mathsf{T}})^{-1}]_{1:2,1:2})$. Using Lemma 1 and the linearity of the constraints in (20b) and (20c), it is observed that the problem (20) is convex in $\{\boldsymbol{X}_g\}_{g=1}^G$ for fixed $\{\boldsymbol{\Psi}_g\}_{g=1}^G$ and convex in $\{\boldsymbol{\Psi}_g\}_{g=1}^G$ for fixed $\{\boldsymbol{X}_g\}_{g=1}^G$. This motivates alternating optimization to solve (20), iterating BS precoders update for fixed RIS phase profiles and RIS phase profiles update for fixed BS precoders.

C. Alternating Optimization to Solve Relaxed Problem

1) *Optimize BS Precoders for Fixed RIS Phase Profiles:* For fixed $\{\boldsymbol{\Psi}_g\}_{g=1}^G$, the subproblem of (20) to optimize $\{\boldsymbol{X}_g\}_{g=1}^G$ can be expressed as

$$\min_{\{\boldsymbol{X}_g\}_{g=1}^G} \quad \text{PEB}^{\text{bd}}\left(\{\boldsymbol{X}_g, \boldsymbol{\Psi}_g\}_{g=1}^G; \boldsymbol{\eta}\right)$$

$$\text{s.t.} \quad (20\text{b}), \qquad (21)$$

---

[5]It should be emphasized that these two relaxations are performed only to reveal the underlying *low-dimensional structure* of the optimal BS and RIS transmission strategies in Section IV-C, which facilitates *codebook design* in Section V-B. For power optimization in Algorithm 1 of Section V-B and for simulation results in Section VII, we employ the true FIM $\boldsymbol{J}_{\boldsymbol{\gamma}}$ instead of the approximated FIM $\boldsymbol{J}_{\boldsymbol{\gamma}}^{\text{bd}}$.

which is a convex problem and can be solved using off-the-shelf solvers [47]. To achieve low-complexity optimization, we can exploit the low-dimensional structure of the optimal precoder covariance matrices, as shown in the following result.

Proposition 1: The optimal BS precoder covariance matrices $\{\boldsymbol{X}_g\}_{g=1}^G$ in (21) can be written as $\boldsymbol{X}_g = \boldsymbol{A}_{\text{BS}}\boldsymbol{\Upsilon}_g\boldsymbol{A}_{\text{BS}}^{\mathsf{H}}$ where

$$\boldsymbol{A}_{\text{BS}} \overset{\text{def}}{=} [\boldsymbol{a}_{\text{BS}}(\theta_{\text{B,R}})\ \boldsymbol{a}_{\text{BS}}(\theta_{\text{B,U}})\ \dot{\boldsymbol{a}}_{\text{BS}}(\theta_{\text{B,U}})]^*, \qquad (22)$$

$\dot{\boldsymbol{a}}_{\text{BS}}(\theta) \overset{def}{=} \partial\boldsymbol{a}_{\text{BS}}(\theta)/\partial\theta$ and $\boldsymbol{\Upsilon}_g \in \mathbb{C}^{3\times 3}$ is a positive semidefinite matrix.

Proof: Please see Appendix A. ∎

2) *Optimize RIS Phase Profiles for Fixed BS Precoders:* For fixed $\{\boldsymbol{X}_g\}_{g=1}^G$, we can formulate the subproblem of (20) to optimize $\{\boldsymbol{\Psi}_g\}_{g=1}^G$ as follows:

$$\min_{\{\boldsymbol{\Psi}_g\}_{g=1}^G} \quad \text{PEB}^{\text{bd}}\left(\{\boldsymbol{X}_g, \boldsymbol{\Psi}_g\}_{g=1}^G; \boldsymbol{\eta}\right)$$

$$\text{s.t.} \quad (20\text{c}), \qquad (23)$$

which is again a convex problem [42]. Similar to (21), the inherent low-dimensional structure of the optimal phase profiles can be exploited to obtain fast solutions to (23), as indicated in the following proposition.

Proposition 2: The optimal RIS phase profile covariance matrices $\{\boldsymbol{\Psi}_g\}_{g=1}^G$ in (23) in the absence of the unit-modulus constraints $\text{diag}(\boldsymbol{\Psi}_g) = \boldsymbol{1}$ can be expressed as $\boldsymbol{\Psi}_g = \boldsymbol{B}_{\text{RIS}}\boldsymbol{\Xi}_g\boldsymbol{B}_{\text{RIS}}^{\mathsf{H}}$, where

$$\boldsymbol{B}_{\text{RIS}} \overset{\text{def}}{=} \left[\boldsymbol{b}_{\text{RIS}}\ \dot{\boldsymbol{b}}_{\text{RIS}}\right]^*, \qquad (24)$$

$\dot{\boldsymbol{b}}_{\text{RIS}}(\theta) \overset{\text{def}}{=} \partial\boldsymbol{b}_{\text{RIS}}(\theta)/\partial\theta$, $\dot{\boldsymbol{b}}_{\text{RIS}} \equiv \dot{\boldsymbol{b}}_{\text{RIS}}(\theta_{\text{R,U}})$ and $\boldsymbol{\Xi}_g \in \mathbb{C}^{2\times 2}$ is a positive semidefinite matrix.

Proof: Please see Appendix B. ∎

Fig. 2 provides a graphical representation of the beams in (22) and (24).

Remark 3: It is worth emphasizing that we never solve the problem (20) to obtain $\boldsymbol{X}_g$ and $\boldsymbol{\Psi}_g$. The sole purpose of the alternating optimization is to formulate the subproblems (21) and (23), and, based on that, to uncover the low-dimensional structure of the optimal BS and RIS transmission strategies, as shown in Prop. 1 and Prop. 2. The derived low-dimensional structure will be exploited in Section V to design the codebooks in (28) under imperfect knowledge of UE location. Hence, the aim of Section IV is not to solve the PEB minimization problem under perfect knowledge of UE location, but to extract analytical insights from the structure of the solution that will be conducive to tackling the more practical problem of PEB optimization under UE location uncertainty in Section V.

D. Interpretation of Proposition 1 and Proposition 2

By focusing on the optimal structure of the precoder covariance matrices obtained in Prop. 1, it emerges that the BS should transmit different beams along the two main directions of the AoDs $\theta_{\text{B,R}}$ and $\theta_{\text{B,U}}$, i.e., the BS should serve both the RIS and the UE. Interestingly, a sort of asymmetry exists in (22): while for the AoD with respect to the RIS, the optimal structure of the precoder includes only the *directional* beam $\boldsymbol{a}_{\text{BS}}(\theta_{\text{B,R}})$, for the AoD with respect to the UE, the BS employs both a *directional* beam $\boldsymbol{a}_{\text{BS}}(\theta_{\text{B,U}})$ and its *derivative* $\dot{\boldsymbol{a}}_{\text{BS}}(\theta_{\text{B,U}})$ [29], [30], [38]. This can be explained by noting that, for positioning purposes, the UE needs to estimate the AoD with respect to the BS, and to do so a certain degree of diversity in the received

(a) Beampatterns of the optimal BS beams $\boldsymbol{A}_{\text{BS}}$ in (22).

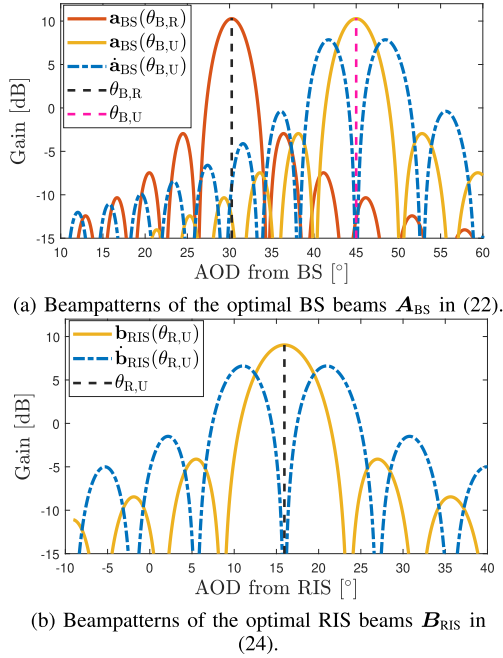(b) Beampatterns of the optimal RIS beams $\boldsymbol{B}_{\text{RIS}}$ in (24).

Fig. 2.   Beampatterns of the localization-optimal BS and RIS beams, including both the directional and derivative beams, obtained for the setup in Section VII-A. The *directional* beams maximize the SNR at the targeted UE location ($\theta_{\text{B,U}}$ and $\theta_{\text{R,U}}$), which serves to provide sufficient SNR for localization, while the *derivative* beams enable the UE to detect small deviations around the nominal direction, similar to monopulse track radars [50], [51], which can be noticed through sharp bending of the beampattern around $\theta_{\text{B,U}}$ and $\theta_{\text{R,U}}$. This sharp curvature around the targeted location allows small deviations in angle to induce large changes in amplitude, thereby facilitating highly accurate mapping from complex amplitude measurements to angles.

V. ROBUST JOINT DESIGN OF BS PRECODER AND RIS PHASE PROFILES UNDER LOCATION UNCERTAINTY

In this section, inspired by Prop. 1 and Prop. 2 in Section IV, we develop *robust joint design* strategies for BS precoder and RIS phase profiles under *imperfect knowledge of UE location* $\boldsymbol{p}$ in (14). To this end, we consider an optimal unconstrained design (without any specific codebook), which turns out to be intractable, and propose a novel codebook-based design with optimized power allocation for joint BS-RIS beamforming.

A. Optimal Unconstrained Design

Solving the PEB minimization problem in (17) requires the knowledge of precise UE location[6] $\boldsymbol{p}$ which, however, may not be available in practice due to measurement noise and tracking errors. Hence, we assume an uncertainty region $\boldsymbol{p} \in \mathcal{P}$ for the UE location and consider the robust design problem that minimizes the worst-case PEB over $\mathcal{P}$ [48], [52]–[54]:

$$\min_{\{\boldsymbol{X}_g, \boldsymbol{\omega}^g, \boldsymbol{\Psi}_g\}_{g=1}^G} \quad \max_{\boldsymbol{p} \in \mathcal{P}} \text{PEB}\left(\{\boldsymbol{X}_g, \boldsymbol{\omega}^g, \boldsymbol{\Psi}_g\}_{g=1}^G; \boldsymbol{\eta}(\boldsymbol{p})\right)$$

$$\text{s.t.} \quad (17\text{b}) - (17\text{d}), \tag{25}$$

where $\boldsymbol{\eta}$ is replaced by $\boldsymbol{\eta}(\boldsymbol{p})$ in the PEB to highlight its dependency on $\boldsymbol{p}$. The epigraph form of (25) can be expressed as

$$\min_{\{\boldsymbol{X}_g, \boldsymbol{\omega}^g, \boldsymbol{\Psi}_g\}_{g=1}^G, t} t \tag{26a}$$

$$\text{s.t.} \quad \text{PEB}\left(\{\boldsymbol{X}_g, \boldsymbol{\omega}^g, \boldsymbol{\Psi}_g\}_{g=1}^G; \boldsymbol{\eta}(\boldsymbol{p})\right) \leq t, \forall \boldsymbol{p} \in \mathcal{P}$$

$$(17\text{b}) - (17\text{d}). \tag{26b}$$

To tackle the semi-infinite optimization problem in (26), we can discretize $\mathcal{P}$ into M grid points $\{\boldsymbol{p}_m\}_{m=0}^{M-1}$ [48] and obtain the following approximated version using (16):

$$\min_{\substack{\{\boldsymbol{X}_g, \boldsymbol{\omega}^g, \boldsymbol{\Psi}_g\}_{g=1}^G \\ t, \{u_{m,k}\}}} t \tag{27a}$$

$$\text{s.t.} \quad \begin{bmatrix} \boldsymbol{J}_{\boldsymbol{\eta}}(\{\boldsymbol{X}_g, \boldsymbol{\omega}^g, \boldsymbol{\Psi}_g\}_{g=1}^G; \boldsymbol{\eta}(\boldsymbol{p}_m)) & \mathbf{e}_k \\ \mathbf{e}_k^T & u_{m,k} \end{bmatrix} \succeq 0 \tag{27b}$$

$$u_{m,0} + u_{m,1} \leq t,$$

$$k = 0, 1, \ m = 0, \ldots, M-1,$$

$$(17\text{b}) - (17\text{d}), \tag{27c}$$

where $\mathbf{e}_k$ is the k-th column of the identity matrix, and the equivalence between (27b), (27c) and the discretized version of (26b) stems from [42, Eq. (7.28)]. In (27b), $\boldsymbol{J}_{\boldsymbol{\eta}}(\{\boldsymbol{X}_g, \boldsymbol{\omega}^g, \boldsymbol{\Psi}_g\}_{g=1}^G; \boldsymbol{\eta}(\boldsymbol{p}_m))$ is the FIM in (15) evaluated at the grid location $\boldsymbol{p}_m$.

Two issues arise that make the problem (27) intractable. First, (17b)–(17d) involve non-convex rank and unit-modulus constraints, which can only be handled via relaxations in Section IV-B. Second, since $\boldsymbol{J}_{\boldsymbol{\eta}}(\{\boldsymbol{X}_g, \boldsymbol{\omega}^g, \boldsymbol{\Psi}_g\}_{g=1}^G; \boldsymbol{\eta}(\boldsymbol{p}_m))$ is not linear with respect to $\{\boldsymbol{X}_g, \boldsymbol{\omega}^g, \boldsymbol{\Psi}_g\}_{g=1}^G$ according to Remark 1, (27b) does not represent a linear matrix inequality (LMI) [43], implying that (27) is not convex [42, Ex. (2.10)]. As a possible remedy, alternating optimization (AO) of $\{\boldsymbol{X}_g\}_{g=1}^G$ and $\{\boldsymbol{\Psi}_g\}_{g=1}^G$ can be performed (after eliminating the dependency of $\boldsymbol{J}_{\boldsymbol{\eta}}$ on $\{\boldsymbol{\omega}^g\}_{g=1}^G$ using the approximation in (18)), where each subproblem becomes convex as bi-linear matrix inequalities (BMIs) degenerate

beams should exist [48], [49]. On the other hand, in the first tandem channel between the BS and the RIS, there is no need to estimate the AoD $\theta_{\text{B,R}}$ (its value is known a priori, given the known positions of both BS and RIS), and from a PEB perspective, the transmitted power should be concentrated in a single directional beam towards the RIS, so as to maximize the received SNR over the whole BS-RIS-UE channel.

Similar conclusions can be derived from Prop. 2. Namely, RIS phase profiles should be steered towards the AoD $\theta_{\text{R,U}}$ with respect to the UE. In addition, both the *directional* beam $\boldsymbol{b}_{\text{RIS}}(\theta_{\text{R,U}})$ and its *derivative* $\dot{\boldsymbol{b}}_{\text{RIS}}(\theta_{\text{R,U}})$ should be employed to maximize the performance of AoD estimation at the UE, which corresponds to the same principle as used in *sum* and *difference* beams of monopulse radar [50].

---

[6]From the viewpoint of joint BS-RIS beamforming, the most essential information required to solve (17) is the UE location (i.e., where to steer the BS and RIS beams). Regarding the other unknown parameters in $\boldsymbol{\eta}$ in (14), we note from Section S-I-B in the supplemental material that the FIM does not depend on a *specific value* of the clock offset Δ (though the FIM depends *functionally* on Δ, as seen from (15) and Section S-II in the supplemental material). Hence, the PEB minimization problem in (17) can be solved without the knowledge of Δ. On the other hand, we assume the channel gains in (14) are perfectly known. As seen from Section S-I-B3 in the supplemental material, the case of uncertain gains leads to intractable PEB expressions due to LoS-NLoS correlations, is therefore left outside the scope of the current work and will be investigated in a future study.

to LMIs when one of the variables is fixed. However, this leads to a high computational complexity roughly given by $\mathcal{O}(N_{\text{BS}}^6)$ and $\mathcal{O}(N_{\text{RIS}}^6)$ [55, Ch. 11] for the BS and RIS subproblems, respectively. To devise a practically implementable solution, we propose a low-complexity codebook-based design strategy, as detailed in Section V-B.

B. Low-Complexity Codebook-Based Design

Motivated by the optimal low-dimensional structure of the BS precoder and the RIS phase profile covariance matrices, derived in Prop. 1 and Prop. 2, we develop a codebook-based low-complexity design approach as a practical alternative to unconstrained design in Section V-A. To this end, let $\{\theta_{\text{B,U}}^{(i)}\}_{i=1}^{L_{\text{BS}}}$ and $\{\theta_{\text{R,U}}^{(i)}\}_{i=1}^{L_{\text{RIS}}}$ denote the uniformly spaced AoDs from the BS to the UE and from the RIS to the UE, respectively, that span the uncertainty region $\mathcal{P}$ of the UE location, where the angular spacing is set to $3\,\text{dB}$ (half-power) beamwidth of the corresponding array [30], [56], [57, Ch. 22.10].

Relying on Prop. 1, Prop. 2 and their interpretation in Section IV-D, we propose the following codebooks for the BS precoder and the RIS phase profiles [30] consisting of both *directional* and *derivative* beams (please refer to Section S-IV in the supplemental material for additional details on how to obtain these codebooks):

$$\boldsymbol{F}^{\text{BS}} = \left[\boldsymbol{a}_{\text{BS}}(\theta_{\text{B,R}}) \; \overline{\boldsymbol{F}}^{\text{BS}} \; \dot{\boldsymbol{F}}^{\text{BS}}\right]^* \in \mathbb{C}^{N_{\text{BS}} \times (2L_{\text{BS}}+1)}, \quad (28\text{a})$$

$$\boldsymbol{F}^{\text{RIS}} = \left[\overline{\boldsymbol{F}}^{\text{RIS}} \; \dot{\boldsymbol{F}}^{\text{RIS}}\right]^* \in \mathbb{C}^{N_{\text{RIS}} \times 2L_{\text{RIS}}}, \quad (28\text{b})$$

where $\overline{\boldsymbol{F}}^{\text{BS}} \stackrel{\text{def}}{=} \left[\boldsymbol{a}_{\text{BS}}\left(\theta_{\text{B,U}}^{(1)}\right) \cdots \boldsymbol{a}_{\text{BS}}\left(\theta_{\text{B,U}}^{(L_{\text{BS}})}\right)\right]$, $\dot{\boldsymbol{F}}^{\text{BS}} \stackrel{\text{def}}{=} \left[\dot{\boldsymbol{a}}_{\text{BS}}\left(\theta_{\text{B,U}}^{(1)}\right) \cdots \dot{\boldsymbol{a}}_{\text{BS}}\left(\theta_{\text{B,U}}^{(L_{\text{BS}})}\right)\right]$ and

$$\overline{\boldsymbol{F}}^{\text{RIS}} \stackrel{\text{def}}{=} \left[\boldsymbol{b}_{\text{RIS}}\left(\theta_{\text{R,U}}^{(1)}\right) \cdots \boldsymbol{b}_{\text{RIS}}\left(\theta_{\text{R,U}}^{(L_{\text{RIS}})}\right)\right], \quad (29)$$

$$\dot{\boldsymbol{F}}^{\text{RIS}} \stackrel{\text{def}}{=} \left[\widetilde{\boldsymbol{b}}_{\text{RIS}}\left(\theta_{\text{R,U}}^{(1)}\right) \cdots \widetilde{\boldsymbol{b}}_{\text{RIS}}\left(\theta_{\text{R,U}}^{(L_{\text{RIS}})}\right)\right]. \quad (30)$$

In (30), due to phase-only control of RIS profiles, we employ $\widetilde{\boldsymbol{b}}_{\text{RIS}}(\theta)$, which is the best approximation with unit-modulus entries to $\dot{\boldsymbol{b}}_{\text{RIS}}(\theta)$ in (24). To obtain $\widetilde{\boldsymbol{b}}_{\text{RIS}}(\theta)$ from $\dot{\boldsymbol{b}}_{\text{RIS}}(\theta)$, the projected gradient descent algorithm in [58, Alg. 1] is used.

For each transmission, we choose a BS-RIS signal pair $\{\boldsymbol{F}_{:,i}^{\text{BS}}, \boldsymbol{F}_{:,j}^{\text{RIS}}\}$, corresponding to the ith beam in $\boldsymbol{F}^{\text{BS}}$ and the jth beam in $\boldsymbol{F}^{\text{RIS}}$, which leads to $G = (2L_{\text{BS}}+1)2L_{\text{RIS}}$ transmissions in total.[7] To minimize the worst-case PEB using this codebook-based approach, we formulate a beam power allocation problem that finds the optimal power $\boldsymbol{\varrho} = [\varrho_1 \ldots \varrho_G]^T$ of BS beams in each transmission under total power constraint:[8]

Algorithm 1: Joint BS Precoder and RIS Phase Profile Design With Power Optimized Codebooks.

1: **Input:** Uncertainty region $\mathcal{P}$ of the UE location $\boldsymbol{p}$ in (14).
2: **Output:** Optimized BS-RIS signal pairs $\{\sqrt{\varrho_g^\star}\boldsymbol{F}_{:,i}^{\text{BS}}, \boldsymbol{F}_{:,j}^{\text{RIS}}\}_{\forall i,j}$ with the optimal powers $\{\varrho_g^\star\}_{\forall g}$.
a) Determine the uniformly spaced AoDs from the BS to the UE $\{\theta_{\text{B,U}}^{(i)}\}_{i=1}^{L_{\text{BS}}}$ and those from the RIS to the UE $\{\theta_{\text{R,U}}^{(i)}\}_{i=1}^{L_{\text{RIS}}}$ based on $\mathcal{P}$.
b) Construct the BS and RIS codebooks in (28).
c) Perform power allocation across $G = (2L_{\text{BS}}+1)2L_{\text{RIS}}$ transmissions, each employing a different BS-RIS signal pair $\{\boldsymbol{F}_{:,i}^{\text{BS}}, \boldsymbol{F}_{:,j}^{\text{RIS}}\}$, by solving the problem in (31).

$$\min_{\substack{\boldsymbol{\varrho},t \\ \{u_{m,k}\}}} \quad t \quad (31\text{a})$$

$$\text{s.t.} \begin{bmatrix} \boldsymbol{J}_{\boldsymbol{\eta}}(\{\boldsymbol{X}_g, \boldsymbol{\omega}^g, \boldsymbol{\Psi}_g\}_{g=1}^G; \boldsymbol{\eta}(\boldsymbol{p}_m)) & \mathbf{e}_k \\ \mathbf{e}_k^T & u_{m,k} \end{bmatrix} \succeq 0,$$

$$u_{m,0} + u_{m,1} \leq t, \; k = 0,1, \; m = 0,\ldots,M-1,$$

$$\text{tr}\left(\sum_{g=1}^G \boldsymbol{X}_g\right) = 1, \; \boldsymbol{\varrho} \succeq \boldsymbol{0}, \; \boldsymbol{X}_g = \varrho_g \boldsymbol{F}_{:,i}^{\text{BS}}(\boldsymbol{F}_{:,i}^{\text{BS}})^{\mathsf{H}},$$

$$\boldsymbol{\omega}^g = \boldsymbol{F}_{:,j}^{\text{RIS}}, \; \boldsymbol{\Psi}_g = \boldsymbol{\omega}^g(\boldsymbol{\omega}^g)^{\mathsf{H}}, \; g = 1,\ldots,G, \quad (31\text{b})$$

where the mapping between the transmission index g and the BS-RIS beam index pair (i,j) is performed according to $g = i + (2L_{\text{BS}}+1)(j-1)$ for $i = 1,\ldots,2L_{\text{BS}}+1$ and $j = 1,\ldots,2L_{\text{RIS}}$. As (31b) is LMI in $\boldsymbol{\varrho}$ and $\{u_{m,k}\}$ (see Remark 1), the problem (31) is convex. After obtaining the optimal power allocation vector $\boldsymbol{\varrho}^\star = [\varrho_1^\star \ldots \varrho_G^\star]^T$ as the solution to (31), the optimized codebook is given by the collection of the BS-RIS signal pairs $\left\{\sqrt{\varrho_g^\star}\boldsymbol{F}_{:,i}^{\text{BS}}, \boldsymbol{F}_{:,j}^{\text{RIS}}\right\}_{\forall i,j}$. The overall BS-RIS signal design algorithm is summarized in Algorithm 1. The computational complexity of (31) is approximately given by $\mathcal{O}(M^3)$ [55, Ch. 11], [30], under the assumption that M is on the same order as G. Since $M < N_{\text{BS}}^2$ and $M < N_{\text{RIS}}^2$ in practice (see Section VII-A), the proposed (non-iterative) design strategy in Algorithm 1 is more efficient than even the individual iterations of an AO approach in Section V-A.

As anticipated, the proposed robust joint design of BS precoders and RIS phase profiles can be in principle extended to the 3D case, using a 2D array (e.g., a URA) in place of the ULA. In this case, three types of beams need to be employed, namely, *directional* beams, *azimuth derivative* beams and *elevation derivative* beams, in contrast to only directional and derivative beams as in the 2D scenario.

VI. MAXIMUM LIKELIHOOD JOINT LOCALIZATION AND SYNCHRONIZATION

In this section, we first derive the joint ML estimator of the desired position $\boldsymbol{p}$ and clock offset Δ. To overcome the need of an exhaustive 3D grid-based optimization of the resulting compressed log-likelihood function, we then provide a reduced-complexity estimator that leverage a suitable reparameterization

---

[7]Due to the dependence of L_{BS} and L_{RIS} on the $3\,\text{dB}$ beamwidth of the respective arrays at the BS and RIS, G is a function of the number of elements at the BS and RIS as well as the size of the uncertainty region $\mathcal{P}$. In addition, depending on whether the SNR is sufficient using a single slot of G transmissions, the slot can be repeated multiple times to reach the desired level of SNR.

[8]Each beam in $\boldsymbol{F}^{\text{BS}}$ and $\boldsymbol{F}^{\text{RIS}}$ is normalized to have unit norm prior to power optimization.

of the signal model to decouple the dependencies on the delays and AoDs, enabling a separate though accurate initial estimation of both p and Δ. Such estimated values are subsequently used as initialization for an iterative low-complexity optimization of the joint ML cost function, which provides the refined position and clock offset estimates.

A. Joint Position and Clock Offset Maximum Likelihood Estimation

To formulate the joint ML estimation problem, let $\boldsymbol{\Theta} = [p_x \; p_y \; \Delta]^{\mathsf{T}}$ denote the vector containing the desired UE position and clock offset parameters. By parameterizing the unknown AoDs ($\theta_{\mathrm{B,U}}$ and $\theta_{\mathrm{R,U}}$) and delays ($\tau_{\mathrm{B,U}}$ and τ_{R}) as a function of the sought $\boldsymbol{\Theta}$ through (5), and stacking all the N signals received over each transmission g, we obtain the more compact expression

$$\boldsymbol{y}_g = \sqrt{P}\boldsymbol{B}_g\boldsymbol{\alpha} + \boldsymbol{\nu}_g \qquad (32)$$

with

$$\boldsymbol{y}_g = [y_g[0] \; \cdots \; y_g[N-1]]^{\mathsf{T}},$$

$$\boldsymbol{\alpha} = [\alpha_{\mathrm{B,U}} \; \alpha_{\mathrm{R}}]^{\mathsf{T}},$$

$$\boldsymbol{B}_g = \left[\left(\tilde{\boldsymbol{S}}_{\mathrm{B,U}}^g\right)^{\mathsf{T}} \boldsymbol{a}_{\mathrm{BS}}(\theta_{\mathrm{B,U}}), \; \left(\tilde{\boldsymbol{S}}_{\mathrm{R}}^g\right)^{\mathsf{T}} \boldsymbol{A}^{\mathsf{T}}(\boldsymbol{\Omega}^g)^{\mathsf{T}} \boldsymbol{a}_{\mathrm{RIS}}(\theta_{\mathrm{R,U}}) \right],$$

$$\tilde{\boldsymbol{S}}_{\mathrm{B,U}}^g = [\boldsymbol{s}_g[0] \; \cdots \; e^{-j\kappa_{N-1}\tau_{\mathrm{B,U}}}\boldsymbol{s}_g[N-1]],$$

where $\boldsymbol{s}_g[n] = \boldsymbol{f}_g s_g[n]$, $\tilde{\boldsymbol{S}}_{\mathrm{R}}^g$ is defined as $\tilde{\boldsymbol{S}}_{\mathrm{B,U}}^g$ but with τ_{R} in place of $\tau_{\mathrm{B,U}}$, $\boldsymbol{A} = \boldsymbol{a}_{\mathrm{RIS}}(\phi_{\mathrm{B,R}})\boldsymbol{a}_{\mathrm{BS}}^{\mathsf{T}}(\theta_{\mathrm{B,R}})$, and $\alpha_{\mathrm{R}} = \alpha_{\mathrm{B,R}}\alpha_{\mathrm{R,U}}$. Without loss of generality, we assume that σ^2 is already known (its estimate can be straightforwardly obtained as $\hat{\sigma}^2 = \sum_{g=1}^{G} \|\boldsymbol{y}_g - \sqrt{P}\boldsymbol{B}_g\boldsymbol{\alpha}\|^2/(NG)$ once the rest of parameters have been estimated), so leaving $\boldsymbol{\alpha}$ as the sole vector of unknown nuisance parameters. Following the ML criterion, the estimation problem can be thus formulated as

$$\hat{\boldsymbol{\Theta}}^{\mathrm{ML}} = \arg\min_{\boldsymbol{\Theta}} \left[\min_{\boldsymbol{\alpha}} L(\boldsymbol{\Theta}, \boldsymbol{\alpha}) \right] \qquad (33)$$

where

$$L(\boldsymbol{\Theta}, \boldsymbol{\alpha}) = \sum_{g=1}^{G} \|\boldsymbol{y}_g - \sqrt{P}\boldsymbol{B}_g\boldsymbol{\alpha}\|^2 \qquad (34)$$

represents the likelihood function. It is not difficult to show that the value of the complex vector $\boldsymbol{\alpha} \in \mathbb{C}^{2\times 1}$ minimizing (34) is given by $\hat{\boldsymbol{\alpha}}^{\mathrm{ML}} = \frac{1}{\sqrt{P}}\boldsymbol{B}^{-1}\sum_{g=1}^{G}\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{y}_g$ where $\boldsymbol{B} = \sum_{g=1}^{G}\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{B}_g$. Substituting $\hat{\boldsymbol{\alpha}}^{\mathrm{ML}}$ back into the likelihood function (34) leads to

$$L(\boldsymbol{\Theta}) = \sum_{g=1}^{G} \|\boldsymbol{y}_g - \sqrt{P}\boldsymbol{B}_g(\boldsymbol{\Theta})\hat{\boldsymbol{\alpha}}^{\mathrm{ML}}(\boldsymbol{\Theta})\|^2 \qquad (35)$$

where we explicitly highlighted the remaining dependency on the sole desired parameter vector $\boldsymbol{\Theta}$. Accordingly, the final joint ML (JML) estimator of UE position and clock offset is

$$\hat{\boldsymbol{\Theta}}^{\mathrm{ML}} = \arg\min_{\boldsymbol{\Theta}} L(\boldsymbol{\Theta}). \qquad (36)$$

Unfortunately, $\hat{\boldsymbol{\Theta}}^{\mathrm{ML}}$ cannot be effortlessly retrieved being $L(\boldsymbol{\Theta})$ a highly non-linear function with multiple potential local minima. A more practical solution consists in finding a good initial estimate of $\boldsymbol{\Theta}$ and use it to compute $\hat{\boldsymbol{\Theta}}^{\mathrm{ML}}$ by means of a

low-complexity iterative optimization. The latter consists in adopting a numerical optimization approach such as the Nelder-Mead algorithm to iteratively optimize the JML cost function in (36) starting from a more accurate initial estimate $\hat{\boldsymbol{\Theta}}$. As well-known, the Nelder-Mead procedure does not require any derivative information, which makes it suitable for problems with non-smooth functions like (36), and is recognized to be extremely fast to converge (in all our trials, the number of required iterations was always less than 30). A direct way to obtain such initialization is to perform an exhaustive grid search over the 3D space of the unknown p and Δ. To overcome the burden of a full-dimensional optimization, in the next section we present a relaxed ML estimator of the position and the clock offset, able to provide a good initialization for the iterative optimization of (36), but at a considerably lower computational complexity.

B. Proposed Reduced-Complexity Estimator

1) Relaxed Maximum Likelihood Position Estimation: We start by stacking all the observations collected over the G transmissions and by further manipulating the resulting model, obtaining the new expression

$$\underbrace{\begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_G \end{bmatrix}}_{\boldsymbol{y} \in \mathbb{C}^{GN \times 1}} = \underbrace{\begin{bmatrix} \boldsymbol{\Phi}_{\mathrm{B,U}}^1(\theta_{\mathrm{B,U}}(\boldsymbol{p})) & \boldsymbol{\Phi}_{\mathrm{R,U}}^1(\theta_{\mathrm{R,U}}(\boldsymbol{p})) \\ \vdots & \vdots \\ \boldsymbol{\Phi}_{\mathrm{B,U}}^G(\theta_{\mathrm{B,U}}(\boldsymbol{p})) & \boldsymbol{\Phi}_{\mathrm{R,U}}^G(\theta_{\mathrm{R,U}}(\boldsymbol{p})) \end{bmatrix}}_{\boldsymbol{\Phi}(\theta_{\mathrm{B,U}}(\boldsymbol{p}), \theta_{\mathrm{R,U}}(\boldsymbol{p})) \overset{\mathrm{def}}{=} \boldsymbol{\Phi}(\boldsymbol{p}) \in \mathbb{C}^{GN \times 2N}} \underbrace{\begin{bmatrix} \boldsymbol{e}_{\mathrm{B,U}} \\ \boldsymbol{e}_{\mathrm{R}} \end{bmatrix}}_{\boldsymbol{e} \in \mathbb{C}^{2N \times 1}} + \begin{bmatrix} \boldsymbol{\nu}_1 \\ \vdots \\ \boldsymbol{\nu}_G \end{bmatrix}$$

$$(37)$$

where $\boldsymbol{\Phi}_{\mathrm{B,U}}^g(\theta_{\mathrm{B,U}}(\boldsymbol{p})) = \mathrm{diag}(\boldsymbol{a}_{\mathrm{BS}}^{\mathsf{T}}(\theta_{\mathrm{B,U}}(\boldsymbol{p}))\boldsymbol{S}^g)$, $\boldsymbol{\Phi}_{\mathrm{R,U}}^g(\theta_{\mathrm{R,U}}(\boldsymbol{p}))$ $= \mathrm{diag}(\boldsymbol{a}_{\mathrm{RIS}}^{\mathsf{T}}(\theta_{\mathrm{R,U}}(\boldsymbol{p}))\boldsymbol{\Omega}^g \boldsymbol{A}\boldsymbol{S}^g)$, $\boldsymbol{S}^g = [\boldsymbol{s}_g[0] \; \cdots \; \boldsymbol{s}_g[N-1]]$, $g = 1, \ldots, G$, and

$$\boldsymbol{e}_{\mathrm{B,U}} = \sqrt{P}\alpha_{\mathrm{B,U}} \begin{bmatrix} 1 \\ e^{-j\kappa_1 \tau_{\mathrm{B,U}}} \\ \vdots \\ e^{-j\kappa_{N-1}\tau_{\mathrm{B,U}}} \end{bmatrix}, \; \boldsymbol{e}_{\mathrm{R}} = \sqrt{P}\alpha_{\mathrm{R}} \begin{bmatrix} 1 \\ e^{-j\kappa_1 \tau_{\mathrm{R}}} \\ \vdots \\ e^{-j\kappa_{N-1}\tau_{\mathrm{R}}} \end{bmatrix}.$$

$$(38)$$

We now observe that (37) allows us to decouple the dependencies on the delays and AoDs in (32), with the new matrix $\boldsymbol{\Phi}$ that depends only on the desired $\boldsymbol{p}$ through the geometric relationships with the corresponding AoDs $\theta_{\mathrm{B,U}}(\boldsymbol{p})$ and $\theta_{\mathrm{R,U}}(\boldsymbol{p})$. By relaxing the dependency of $\boldsymbol{e}$ on the delays $\tau_{\mathrm{B,U}}$ and τ_{R}, and considering it as a generic unstructured $2N$-dimensional vector, a relaxed ML-based estimator (RML) of $\boldsymbol{p}$ can be derived as

$$\hat{\boldsymbol{p}}^{\mathrm{RML}} = \arg\min_{\boldsymbol{p}} \left[\min_{\boldsymbol{e}} \|\boldsymbol{y} - \boldsymbol{\Phi}(\boldsymbol{p})\boldsymbol{e}\|^2 \right]. \qquad (39)$$

The inner minimization of (39) can be more easily solved by decomposing it over the different N subcarriers as

$$\min_{\boldsymbol{e}} \|\boldsymbol{y} - \boldsymbol{\Phi}(\boldsymbol{p})\boldsymbol{e}\|^2 = \min_{\boldsymbol{e}_0, \ldots, \boldsymbol{e}_{N-1}} \sum_{n=0}^{N-1} \|\boldsymbol{y}_n - \boldsymbol{\Phi}_n\boldsymbol{e}_n\|^2 \qquad (40)$$

where we exploited the peculiar structure of $\boldsymbol{\Phi}(\boldsymbol{p})$, which consists of blocks of $N \times N$ diagonal matrices, with $\boldsymbol{y}_n = $

$[y_1[n] \; \cdots \; y_G[n]]^\mathsf{T},$

$$\boldsymbol{\Phi}_n(\boldsymbol{p}) = \begin{bmatrix} \phi^1_{\mathrm{B,U},n}(\boldsymbol{p}) & \phi^1_{\mathrm{R,U},n}(\boldsymbol{p}) \\ \vdots & \vdots \\ \phi^G_{\mathrm{B,U},n}(\boldsymbol{p}) & \phi^G_{\mathrm{R,U},n}(\boldsymbol{p}) \end{bmatrix} \in \mathbb{C}^{G \times 2} \qquad (41)$$

$\boldsymbol{e}_n = [e_{\mathrm{B,U}}[n] \; e_{\mathrm{R}}[n]]^\mathsf{T} \in \mathbb{C}^{2\times 1}$, $\phi^g_{\mathrm{B,U},n}(\boldsymbol{p}) = \boldsymbol{a}^\mathsf{T}_{\mathrm{BS}}(\boldsymbol{p})s_g[n]$ and $\phi^g_{\mathrm{R,U},n}(\boldsymbol{p}) = \boldsymbol{a}^\mathsf{T}_{\mathrm{RIS}}(\boldsymbol{p})\boldsymbol{\Omega}^g \boldsymbol{A}s_g[n]$, for $n = 0, \ldots, N-1$, $g = 1, \ldots, G$. Each unknown vector $\boldsymbol{e}_n$ minimizing (40) can be separately obtained as

$$\hat{\boldsymbol{e}}^{\mathrm{RML}}_n(\boldsymbol{p}) = \left(\boldsymbol{\Phi}^\mathsf{H}_n(\boldsymbol{p})\boldsymbol{\Phi}_n(\boldsymbol{p})\right)^{-1} \boldsymbol{\Phi}^\mathsf{H}_n(\boldsymbol{p})\boldsymbol{y}_n \qquad (42)$$

that is, each $\boldsymbol{e}_n$ is estimated by pseudo-inverting the corresponding matrix $\boldsymbol{\Phi}_n(\boldsymbol{p})$. The inverse in (42) can be computed in closed-form

$$\left(\boldsymbol{\Phi}^\mathsf{H}_n\boldsymbol{\Phi}_n\right)^{-1} = \frac{1}{u_n z_n - v_n w_n} \begin{bmatrix} z_n & -v_n \\ -w_n & u_n \end{bmatrix} \qquad (43)$$

where $u_n = \sum^G_{g=1} |\phi^g_{\mathrm{B,U},n}|^2$, $v_n = \sum^G_{g=1} (\phi^g_{\mathrm{B,U},n})^* \phi^g_{\mathrm{R,U},n}$, $w_n = \sum^G_{g=1} (\phi^g_{\mathrm{R,U},n})^* \phi^g_{\mathrm{B,U},n}$, and $z_n = \sum^G_{g=1} |\phi^g_{\mathrm{R,U},n}|^2$, and we omitted the dependency on $\boldsymbol{p}$ for brevity. Accordingly, the RML estimator can be more conveniently obtained as

$$\hat{\boldsymbol{p}}^{\mathrm{RML}} = \arg \min_{\boldsymbol{p}} \sum^{N-1}_{n=0} \|\boldsymbol{y}_n - \boldsymbol{l}_n(\boldsymbol{p})\|^2 \qquad (44)$$

with the elements of the vector $\boldsymbol{l}_n(\boldsymbol{p})$ given by

$$\begin{aligned} l_g[n](\boldsymbol{p}) = \frac{1}{u_n z_n - v_n w_n} &\Bigg[(\phi^g_{\mathrm{B,U},n} z_n - \phi^g_{\mathrm{R,U},n} w_n) \\ &\times \sum^G_{\ell=1} (\phi^\ell_{\mathrm{B,U},n})^* y_\ell[n] \\ &+ (\phi^g_{\mathrm{R,U},n} u_n - \phi^g_{\mathrm{B,U},n} v_n) \sum^G_{\ell=1} (\phi^\ell_{\mathrm{R,U},n})^* y_\ell[n] \Bigg]. \end{aligned} \qquad (45)$$

A 2D grid search is then performed on the RML cost function provided in (44) to obtain the initial UE position estimate $\hat{\boldsymbol{p}}^{\mathrm{RML}}$, which will be used together with the clock offset estimate obtained in the next section as initial point to iteratively optimize the 3D plain JML cost function given in (36).

2) FFT-Based Clock Offset Estimation: As a byproduct of the above estimation of $\boldsymbol{p}$, it is possible to derive an efficient estimator of the unknown delays $\tau_{\mathrm{B,U}}$ and τ_{R}, which in turn will be used to retrieve a closed-form estimate of the sought Δ. Specifically, we first plug $\hat{\boldsymbol{p}}^{\mathrm{RML}}$ back in (42) to obtain an estimate of the vectors $\boldsymbol{e}_n$, $n = 0, \ldots, N-1$. The elements of the estimated vectors $\hat{\boldsymbol{e}}^{\mathrm{RML}}_n$ can be then merged according to (38) to obtain an estimate of the two vectors $\hat{\boldsymbol{e}}_{\mathrm{B,U}}(\hat{\boldsymbol{p}}^{\mathrm{RML}})$ and $\hat{\boldsymbol{e}}_{\mathrm{R}}(\hat{\boldsymbol{p}}^{\mathrm{RML}})$, respectively. The key observation consists in the fact that the elements of both $\hat{\boldsymbol{e}}_{\mathrm{B,U}}(\hat{\boldsymbol{p}}^{\mathrm{RML}})$ and $\hat{\boldsymbol{e}}_{\mathrm{R}}(\hat{\boldsymbol{p}}^{\mathrm{RML}})$ can be interpreted as discrete samples of complex exponentials having normalized frequencies $\nu_{\mathrm{B,U}} = -\frac{\tau_{\mathrm{B,U}}}{NT_s}$ and $\nu_{\mathrm{R}} = -\frac{\tau_{\mathrm{R}}}{NT_s}$, respectively. This allows to estimate the delays $\tau_{\mathrm{B,U}}$ and τ_{R} by searching for the dominant peaks in the FFT of the corresponding vectors $\hat{\boldsymbol{e}}_{\mathrm{B,U}}(\hat{\boldsymbol{p}}^{\mathrm{RML}})$ and $\hat{\boldsymbol{e}}_{\mathrm{R}}(\hat{\boldsymbol{p}}^{\mathrm{RML}})$. By defining $\boldsymbol{f}_h(\hat{\boldsymbol{p}}^{\mathrm{RML}}) = \mathrm{FFT}(\hat{\boldsymbol{e}}_h(\hat{\boldsymbol{p}}^{\mathrm{RML}}))$ as the FFT of the vector $\hat{\boldsymbol{e}}_h(\hat{\boldsymbol{p}}^{\mathrm{RML}})$ (with either $h = B, U$ or $h = R$) computed on N_F points, we first seek for

---

Algorithm 2: Low-Complexity Joint Localization and Synchronization Algorithm.

1: **Input:** Received signals $\{y^g[n]\}_{\forall n,g}$, optimized BS-RIS precoders $\{\sqrt{\varrho^\star_g}\boldsymbol{F}^{\mathrm{BS}}_{:,i}, \boldsymbol{F}^{\mathrm{RIS}}_{:,j}\}_{\forall i,j}$.
2: **Output:** UE position $\hat{\boldsymbol{p}}^{\mathrm{ML}}$ and clock offset $\hat{\Delta}^{\mathrm{ML}}$.
  a) Perform a coarse 2D search to obtain an initial estimate $\hat{\boldsymbol{p}}^{\mathrm{RML}}$ via RML in (44).
  b) Use $\hat{\boldsymbol{p}}^{\mathrm{RML}}$ to reconstruct the two vectors $\hat{\boldsymbol{e}}_{\mathrm{B,U}}(\hat{\boldsymbol{p}}^{\mathrm{RML}})$ and $\hat{\boldsymbol{e}}_{\mathrm{R}}(\hat{\boldsymbol{p}}^{\mathrm{RML}})$ based on (42) and (38).
  c) Search for the dominant peaks in the FFT-transformed vectors $(\boldsymbol{f}_{\mathrm{B,U}}(\hat{\boldsymbol{p}}^{\mathrm{RML}}), \boldsymbol{f}_{\mathrm{R}}(\hat{\boldsymbol{p}}^{\mathrm{RML}}))$ and compute the corresponding delays estimates $(\hat{\tau}^{\mathrm{FFT}}_{\mathrm{B,U}}, \hat{\tau}^{\mathrm{FFT}}_{\mathrm{R}})$.
  d) Compute the initial estimate $\hat{\Delta}^{\mathrm{FFT}}$ using (48).
  e) Use $\hat{\boldsymbol{\theta}}^{\mathrm{RML}} = [\hat{\boldsymbol{p}}^{\mathrm{RML}} \; \hat{\Delta}^{\mathrm{FFT}}]^\mathsf{T}$ as initialization to iteratively solve the JML in (36) and obtain the final estimates $\hat{\boldsymbol{p}}^{\mathrm{ML}}$ and $\hat{\Delta}^{\mathrm{ML}}$.

---

the index corresponding to the maximum element in $\boldsymbol{f}_h(\hat{\boldsymbol{p}}^{\mathrm{RML}})$

$$\hat{k}_h(\hat{\boldsymbol{p}}^{\mathrm{RML}}) = \arg \max_k \left[\left| \boldsymbol{f}_h(\hat{\boldsymbol{p}}^{\mathrm{RML}})[k] \right| : 0 \le k \le N_F - 1 \right] \qquad (46)$$

with $|\boldsymbol{f}_h(\hat{\boldsymbol{p}}^{\mathrm{RML}})[k]|$ denoting the absolute value of the k-th element of $\boldsymbol{f}_h(\hat{\boldsymbol{p}}^{\mathrm{RML}})$. Since the first $N_F/2 + 1$ elements correspond to positive values of the normalized frequency $\nu_o \in [0, 1/2]$, while the remaining $N_F/2 - 1$ are associated to the negative part of the spectrum, i.e., $\nu_h \in (-1/2, 0)$, the estimate of the delays can be obtained by mapping the corresponding $\hat{k}_h(\hat{\boldsymbol{p}}^{\mathrm{RML}})$ as

$$\hat{\tau}^{\mathrm{FFT}}_h = \begin{cases} -\frac{\hat{k}_h}{N_F}NT_S & \text{if } 0 \le \hat{k}_h \le N_F/2 \\ (1/2 - \frac{\hat{k}_h}{N_F})NT_S & \text{if } N_F/2 + 1 \le \hat{k}_h \le N_F - 1 \end{cases} \qquad (47)$$

where we omitted the dependency on $\hat{\boldsymbol{p}}^{\mathrm{RML}}$ for conciseness. Once the two delays have been estimated, the sought clock offset Δ can be obtained in closed-form as

$$\hat{\Delta}^{\mathrm{FFT}} = \frac{1}{2} \left[\hat{\tau}^{\mathrm{FFT}}_{\mathrm{B,U}} - \|\hat{\boldsymbol{p}}^{\mathrm{RML}}\|/c + \hat{\tau}^{\mathrm{FFT}}_{\mathrm{R}} - (\|\boldsymbol{r}\| + \|\boldsymbol{r} - \hat{\boldsymbol{p}}^{\mathrm{RML}}\|)/c \right]. \qquad (48)$$

The obtained estimate $\hat{\boldsymbol{\theta}}^{\mathrm{RML}} = [\hat{\boldsymbol{p}}^{\mathrm{RML}} \; \hat{\Delta}^{\mathrm{FFT}}]^\mathsf{T}$ is then used to initialize an iterative optimization procedure (e.g., Nelder-Mead) to efficiently solve the JML estimation problem in (36). The main steps of the proposed reduced-complexity estimation algorithm are summarized in Algorithm 2.

It is worth noting that also the proposed joint localization and estimation algorithm can be extended to the 3D case, in which also elevation angles are considered. In fact, the properties used to obtain the relaxation of the ML cost function and to estimate the delays via FFT are fulfilled not only by ULAs but also by uniform rectangular arrays (URAs). The final position estimation in the RML approach would be then performed on a 3D grid instead of a 2D one. The computational complexity of the procedure, of course, would be higher as in any higher-dimensional problem, but no additional theoretical issues arise.

C. Complexity Analysis

In this section, we analyze the computational complexity of the joint localization and synchronization algorithm proposed

in Section VI-B, also in comparison to the plain 3D JML estimator derived in Section VI-A. Asymptotically speaking, we observe that the complexity in performing the 3D optimization required by the plain JML estimator in (36) is on the order of $\mathcal{O}(Q^3 G N N_\mathrm{E})$, where Q denotes the number of evaluation points per dimension (either p_x coordinate, p_y coordinate of the UE position, or clock offset Δ), assumed to be the same for all the three dimensions for the sake of exposition, and $N_\mathrm{E} = N_\mathrm{BS} + N_\mathrm{BS} N_\mathrm{RIS} + N_\mathrm{RIS}^2$ a term related to the number of elements at both BS and RIS. On the other hand, by analyzing the different steps involved in the proposed joint localization and synchronization algorithm (Algorithm 2), it emerges that the overall complexity is given by the sum of three terms

$$\mathcal{O}(Q^2 G N N_\mathrm{E}) + \mathcal{O}(N_F \log N_F) + \mathcal{O}(N_\mathrm{I} G N N_\mathrm{E}). \quad (49)$$

The first term $\mathcal{O}(Q^2 G N N_\mathrm{E})$ corresponds to the two-dimensional optimization required to obtain the initial UE position estimate $\hat{\boldsymbol{p}}^\mathrm{RML}$ according to (44). The second term $\mathcal{O}(N_F \log N_F)$ denotes the complexity required to compute the FFT of the two vectors $\hat{\boldsymbol{e}}_\mathrm{B,U}(\hat{\boldsymbol{p}}^\mathrm{RML})$ and $\hat{\boldsymbol{e}}_\mathrm{R}(\hat{\boldsymbol{p}}^\mathrm{RML})$ and to search for the dominant peaks yielding the clock offset estimate $\hat{\Delta}^\mathrm{FFT}$ in (48). The third and last term represents instead the complexity required by the Nelder-Mead procedure to iteratively optimize the JML cost function starting from the initial point $\hat{\boldsymbol{\Theta}}^\mathrm{RML}$, with N_I denoting the number of total iterations. This contribution is practically negligible compared to the first term in (49) being the Nelder-Mead procedure extremely fast and typically converging in a few iterations (in all our trials, the number of iterations was always $N_\mathrm{I} < 30$).

Considering that the minimum number of points required to compute the FFT is equal to the length of the involved vectors, i.e. $N_F \geq N$ (in our simulations we set $N_F = 512$) and that the FFT step is performed just once, it is apparent that the overall complexity is practically dominated by the first term $\mathcal{O}(Q^2 G N N_\mathrm{E})$, namely by the two-dimensional search required to obtain an initial estimate of $\boldsymbol{p}$. In this respect, the proposed joint localization and synchronization algorithm is able to reduce the complexity required by the plain JML estimator, which is cubic in Q, to a quadratic cost in Q plus two very low-cost subsequent estimation steps (FFT and iterative optimization).

VII. SIMULATION ANALYSIS AND RESULTS

In this section, we conduct a numerical analysis to assess the performance of the low-complexity localization and synchronization algorithm presented in Section VI, when fed with the robust strategy for joint design of BS precoding and RIS phase profiles proposed in Section V. The performance of the proposed approach is compared with the theoretical lower bounds derived in Section III, as well as against other state-of-the-art strategies for the design of BS and RIS precoding matrices, under different values of the main system parameters. We consider the root mean squared error (RMSE) as performance metric, estimated on 1000 independent Monte Carlo trials.

A. Simulation Setup

The analyzed scenario consists of a single BS placed at known position $\boldsymbol{q} = [0 \; 0]^\mathsf{T}$ m, a RIS placed at $\boldsymbol{r} = [12 \; 7]^\mathsf{T}$ m, and a UE with unknown location $\boldsymbol{p} = [5 \; 5]^\mathsf{T}$ m. The numerical evaluations are conducted assuming the transmission of $G = (2L_\mathrm{BS} + 1)2L_\mathrm{RIS}$ OFDM pilot signals in DL over a

typical mmWave carrier frequency $f_c = 28\,\mathrm{GHz}$ with bandwidth $B = 100\,\mathrm{MHz}$, along N subcarriers equally spaced in frequency by $\Delta f = 240\,\mathrm{kHz}$. The BS is equipped with $N_\mathrm{BS} = 16$ antennas, while the RIS has $N_\mathrm{R} = 32$ elements. The channel amplitudes are generated according to the common path loss model in free space, i.e., $\rho_\mathrm{B,R} = \lambda_c/(4\pi\|\boldsymbol{r}\|)$, $\rho_\mathrm{B,U} = \lambda_c/(4\pi\|\boldsymbol{p}\|)$, and $\rho_\mathrm{R,U} = \lambda_c/(4\pi\|\boldsymbol{p} - \boldsymbol{r}\|)$, respectively, while the phases $\varphi_\mathrm{B,U}$ and φ_R are assumed to be uniformly distributed over $[-\pi, \pi]$. We set the clock offset to $\Delta = \frac{1}{8} N T_s$, while the transmitted power P is varied in order to obtain different ranges of the received SNR over the LoS path, defined as $\mathrm{SNR} = 10 \log_{10}(P \rho_\mathrm{B,U}^2/(N_0 B))$, where N_0 is the noise power spectral density and $\sigma^2 = N_0 B$. In the following, we consider an uncertainty region $\mathcal{P}$ for the UE position having an extent of 3 m along both x and y directions. For this setup, using the typical 3 dB beamwidth angular spacing of an ULA (about $1.8/N_\mathrm{BS}$) leads to $L_\mathrm{BS} = 7$ and $L_\mathrm{RIS} = 6$, which in turn correspond to $G = 180$ OFDM symbols. In the supplemental material, we report additional performance analyses also for the case in which the uncertainty is increased to 5 m. The number of discrete UE positions $\{\boldsymbol{p}_m\}_{m=1}^M$ used to solve (31) is set to $M = 9$. For more details on the setting of M, we refer the reader to Section S-VI in the supplemental material.

B. Benchmark Precoding Schemes

To benchmark the proposed joint BS-RIS signal design algorithm proposed in Algorithm 1, we consider the following state-of-the-art schemes.

1) Directional Codebook (Uniform): This scheme considers only directional beams in the codebook and uses uniform (equal) power allocation among them, i.e, the BS does not implement the optimal power allocation provided in (31). To guarantee a fair comparison, we double the angular sampling rate of the uncertainty region of the UE, so obtaining the same number of transmissions G used by proposed codebooks in (28). This leads to a set of AoDs from the BS to the UE $\{\widetilde{\theta}_\mathrm{B,U}^{(i)}\}_{i=1}^{2L_\mathrm{BS}}$ and of AoDs from the RIS to the UE $\{\widetilde{\theta}_\mathrm{R,U}^{(i)}\}_{i=1}^{2L_\mathrm{RIS}}$. Accordingly, we consider the following directional codebooks for the BS and RIS transmissions:

$$\boldsymbol{F}^\mathrm{BS} = \left[\boldsymbol{a}_\mathrm{BS}(\theta_\mathrm{B,R}) \; \widetilde{\boldsymbol{F}}^\mathrm{BS} \right]^* \in \mathbb{C}^{N_\mathrm{BS} \times (2L_\mathrm{BS}+1)}, \quad (50\mathrm{a})$$

$$\boldsymbol{F}^\mathrm{RIS} = \left[\boldsymbol{b}_\mathrm{RIS}\left(\widetilde{\theta}_\mathrm{R,U}^{(1)}\right) \cdots \boldsymbol{b}_\mathrm{RIS}\left(\widetilde{\theta}_\mathrm{R,U}^{(2L_\mathrm{RIS})}\right) \right]^* \in \mathbb{C}^{N_\mathrm{RIS} \times 2L_\mathrm{RIS}}, \quad (50\mathrm{b})$$

where $\widetilde{\boldsymbol{F}}^\mathrm{BS} \stackrel{\mathrm{def}}{=} [\boldsymbol{a}_\mathrm{BS}(\widetilde{\theta}_\mathrm{B,U}^{(1)}) \cdots \boldsymbol{a}_\mathrm{BS}(\widetilde{\theta}_\mathrm{B,U}^{(2L_\mathrm{BS})})]$.

2) Directional Codebook (Optimized): This scheme uses the same directional codebook in (50) and performs the optimal power allocation for the BS beams in (50a) according to (31).

3) DFT Codebook (Optimized): Let $\boldsymbol{G}^N \in \mathbb{C}^{N \times N}$ denote a DFT matrix. In addition, denote by $\theta_\mathrm{B,U}^c$ and $\theta_\mathrm{R,U}^c$, respectively, the AoD from BS to UE and the AoD from RIS to UE, corresponding to the center of the two AoDs uncertainty regions computed from $\mathcal{P}$. This scheme selects the columns from the corresponding DFT matrices that are closest to the center AoDs as:

$$\boldsymbol{F}^\mathrm{BS,DFT} = \boldsymbol{G}_{:,(\ell_\mathrm{B,U}-L_\mathrm{BS}+1):(\ell_\mathrm{B,U}+L_\mathrm{BS})}^{N_\mathrm{BS}} \in \mathbb{C}^{N_\mathrm{BS} \times 2L_\mathrm{BS}}, \quad (51\mathrm{a})$$

$$\boldsymbol{F}^\mathrm{RIS,DFT} = \boldsymbol{G}_{:,(\ell_\mathrm{R,U}-L_\mathrm{RIS}+1):(\ell_\mathrm{R,U}+L_\mathrm{RIS})}^{N_\mathrm{RIS}} \in \mathbb{C}^{N_\mathrm{RIS} \times 2L_\mathrm{RIS}}, \quad (51\mathrm{b})$$

Fig. 3. RMSEs on the estimation of $\boldsymbol{p}$ as a function of the SNR for the directional codebook, DFT codebook, and proposed codebook.
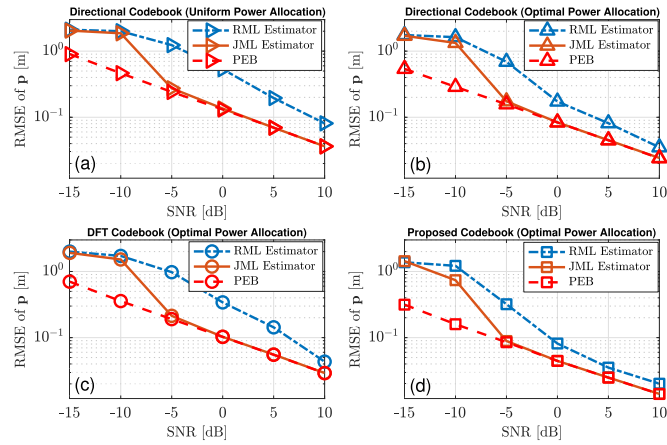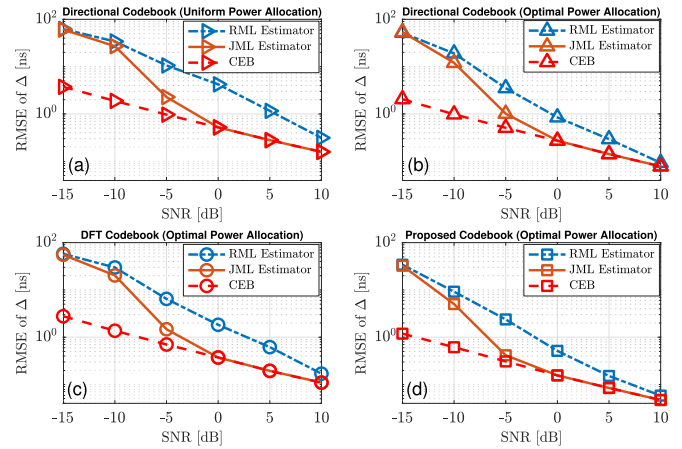

Fig. 4. RMSE on the estimation of Δ as a function of the SNR for the directional codebook, DFT codebook, and proposed codebook.

where $\quad \ell_{\mathrm{B,U}} \stackrel{\text{def}}{=} \arg\min_\ell \|\boldsymbol{G}_{:,\ell}^{N_{\mathrm{BS}}} - \boldsymbol{a}_{\mathrm{BS}}^*(\theta_{\mathrm{B,U}}^c)\| \quad$ and $\quad \ell_{\mathrm{R,U}} \stackrel{\text{def}}{=}$ $\arg\min_\ell \|\boldsymbol{G}_{:,\ell}^{N_{\mathrm{RIS}}} - \boldsymbol{b}_{\mathrm{RIS}}^*(\theta_{\mathrm{R,U}}^c)\|$. Based on (51), the DFT codebooks [28], [59], [60] for the BS and RIS transmission can be expressed as follows:

$$\boldsymbol{F}^{\mathrm{BS}} = \left[\boldsymbol{a}_{\mathrm{BS}}^*(\theta_{\mathrm{B,R}}) \ \boldsymbol{F}^{\mathrm{BS,DFT}} \right] \in \mathbb{C}^{N_{\mathrm{BS}} \times (2L_{\mathrm{BS}}+1)}, \quad (52\mathrm{a})$$

$$\boldsymbol{F}^{\mathrm{RIS}} = \boldsymbol{F}^{\mathrm{RIS,DFT}} \in \mathbb{C}^{N_{\mathrm{RIS}} \times 2L_{\mathrm{RIS}}}. \quad (52\mathrm{b})$$

Also for this scheme, we perform power allocation for the BS-RIS beams in (52) using (31).

C. Results and Discussion

1) Comparison Between Uniform and Proposed Beam Power Allocation: We start the numerical analysis by assessing the validity of the beam power allocation strategy proposed in Section V. In this respect, we select as a precoding scheme the directional codebook given by (50) and perform a direct comparison between the case in which all the G beams share the total transmitted power uniformly, i.e., each beam is transmitted with a power equal to P/G, with the case in which at the g-th beam is allocated a fraction of the total power given by the corresponding g-th element of the allocation vector $\boldsymbol{\varrho}^*$, the latter obtained as a solution of the optimal power allocation problem in (31). Figs. (3a)-(3b) and Figs. (4a)-(4b) show the RMSEs on the estimation of the UE position $\boldsymbol{p}$ and clock offset Δ, respectively, as a function of the SNR, for the directional codebook with uniform and optimal power allocation, also in comparison to the theoretical lower bounds (PEB and CEB[9]) derived in Section III. The proposed low-complexity estimation algorithm is denoted as "JML" and it is implemented as described in Algorithm 2, using the power allocation strategy proposed in Algorithm 1. For completeness, we also report the performance of the RML estimator which is used to obtain an initial estimate of the vector $\boldsymbol{\Theta}$. By comparing the RMSEs in the Fig. (3a) and Fig. (3b) (analogously Fig. (4a) and Fig. (4b)), it clearly emerges that the proposed power allocation strategy yields more accurate estimates of the UE parameters compared to the uniform power allocation, as also confirmed by the gap between the corresponding lower bounds. Such results demonstrate that adopting

[9] The CEB is given by $[\boldsymbol{J}_{\boldsymbol{\eta}}^{-1}]_{7,7}$, where $\boldsymbol{J}_{\boldsymbol{\eta}}$ is the FIM in (15).

a simple uncontrolled (uniform) power allocation scheme at the BS side likely leads to a waste of energy towards directions that do not provide useful contributions for the estimation process. The inefficient use of the transmitted power becomes even more critical in an RIS-assisted localization scenario, being the NLoS channel linking the BS and the UE through the RIS subject to a more severe path loss resulting from the product of two separated tandem channels (ref. (3) and (4)). In light of these considerations, the subsequent comparisons will be conducted assuming optimal power allocation.

2) Comparison Between State-of-The-Art and Proposed Joint BS-RIS Signal Design: The set of figures reported in Fig. 3 and Fig. 4 show a detailed performance comparison between the proposed joint BS-RIS precoding scheme and the state-of-the-art approaches listed in Section VII-B, when used within the proposed low-complexity localization and synchronization algorithm. On the one hand, the obtained results demonstrate the effectiveness of the proposed estimation approaches: despite its intrinsic suboptimality, the RML algorithm (dash-dot curves) provides satisfactory initial estimates of both $\boldsymbol{p}$ and Δ parameters for all the considered precoding schemes, with an accuracy that tend to increase with the SNR and with a complexity reduced to a 2D search in the estimation process. Accordingly, the RMSEs of the RML and JML estimators are close when the SNR is small because in that regime the initial estimates $\hat{\boldsymbol{\Theta}}^{\mathrm{RML}} = [\hat{\boldsymbol{p}}^{\mathrm{RML}} \ \hat{\Delta}^{\mathrm{FFT}}]^{\mathsf{T}}$ provided by the RML estimator are quite inaccurate. As a result, the iterative optimization procedure gets trapped into local wrong minima and leads to solutions (in terms of position and clock offset estimates) for the JML that are very close to that of the RML estimator. Remarkably, the RMSEs of the JML estimator (solid curves) immediately attain the corresponding lower bounds as soon as the initialization provided by the RML becomes sufficiently accurate, providing excellent localization and synchronization performance already at $\mathrm{SNR} = -5$ dB for all the considered precoding schemes. To better highlight the necessity of adopting the more accurate initialization $\hat{\boldsymbol{\Theta}}^{\mathrm{RML}}$ obtained via the proposed RML estimator, we also evaluate the performance of the JML estimator initialized with a random realization of the vector $\boldsymbol{\Theta}$. More specifically, we conduct an additional simulation analysis to directly
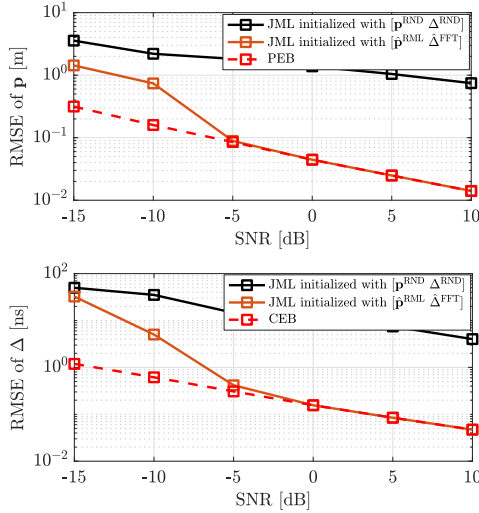
Fig. 5. Comparison between the JML initialized with the estimates $\hat{\Theta}^{\mathrm{RML}}$ provided by the proposed RML estimator and with random values Θ^{RND}.
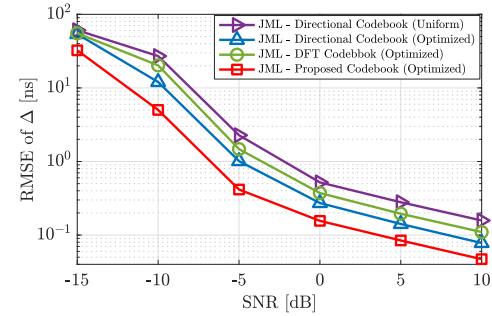

(a) RMSE on the estimation of $\boldsymbol{p}$.


(b) RMSE on the estimation of Δ.

Fig. 6. Comparison between the RMSEs of (a) $\boldsymbol{p}$ and (b) Δ using the proposed JML estimator for different precoding schemes, as a function of the SNR.

compare the two versions of the JML estimator, using for the randomly-initialized JML an initial point $\Theta^{\mathrm{RND}} = [\boldsymbol{p}^{\mathrm{RND}} \; \Delta^{\mathrm{RND}}]^{\mathsf{T}}$ obtained by selecting $\boldsymbol{p}^{\mathrm{RND}}$ as a random position within the assumed uncertainty region $\mathcal{P}$ and by generating $\Delta^{\mathrm{RND}} \sim \mathcal{U}(0, 2\Delta)$. To make the comparison fair, we re-generated the value of Θ^{RND} for each independent Monte Carlo trial. The results in terms of RMSEs reported in Fig. 5 demonstrate that the performance of the JML estimator significantly worsen when Θ^{RND} is used as initial point. This behavior is due to the fact that the iterative minimization of a highly non-linear cost function such as that of the JML estimator gets trapped into local erroneous minima when a random initialization is used, and in turn produces wrong position and clock offset estimates. This confirms the need to seek for a more accurate initialization as the one proposed in Section VI-B, which allows the JML estimator to attain the theoretical lower bounds.

On the other hand, a direct comparison among the PEBs and CEBs in Figs. 3–4 (dashed curves) reveals that the proposed robust joint BS-RIS precoding strategy offers the best localization and synchronization performance among all the considered schemes. To better highlight the advantages of the proposed approach, in Fig. 6 we report an explicit comparison among the RMSEs on the estimation of $\boldsymbol{p}$ and Δ for the JML estimator fed with different precoding schemes. As it can be noticed, the proposed robust joint BS-RIS precoding scheme significantly outperforms both the directional and DFT codebooks. Interestingly, the values assumed by the corresponding RMSEs in Figs. 6(a)–(b) (solid curves with □ marker) demonstrate that the UE can be localized with an error lower than 10 cm and, at the same time, synchronization can be recovered with a subnanosecond precision, for SNR ≥ -5 dB. From this analysis, we can conclude that combining the proposed codebooks in (28a)–(28b) with a power allocation strategy that aims at minimizing the worst-case PEB allows us to achieve a better coverage of the uncertainty region $\mathcal{P}$, while properly taking into account the different directional and derivative beams transmitted towards the UE and the RIS.

3) Robustness Analysis: We now conduct an analysis aimed at assessing the effective robustness of the proposed joint BS
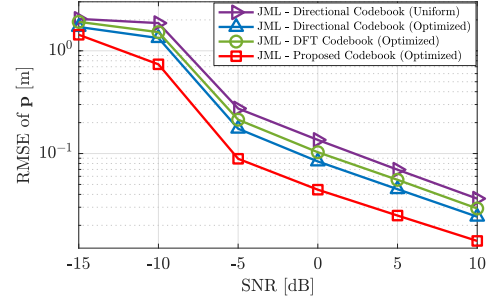
and RIS beamforming design strategy to different UE positions falling within the assumed uncertainty region $\mathcal{P}$. More specifically, we test the values assumed by the PEB when the UE spans different locations around the nominal one $\mathbf{p}$, considering both the proposed robust design approach and its corresponding non-robust version, the latter obtained by simply shrinking the extent of the uncertainty region to a very small area of 0.1 m around the nominal UE position. The results reported in Fig. 7(a), obtained for SNR $= 0$ dB, show that the PEB exhibits quite similar values within the whole uncertainty region, confirming the robustness of the proposed joint active BS and passive RIS beamforming design strategy. Interestingly, the PEB keeps reasonable values even when the UE falls slightly outside the considered region $\mathcal{P}$. Conversely, the values assumed by the PEB in Fig. 7(b) clearly indicate an evident position accuracy degradation for UE locations different from the nominal one, leading to errors that are almost three times those experienced in Fig. 7(a) with the proposed robust joint design strategy.

4) Performance Assessment for Reduced Number of Transmitted Beams: To corroborate the above results, we investigate the possibility to adopt an ad-hoc heuristic that allows to reduce the total number of transmitted beams G. The main idea originates from observing that, when the BS is transmitting a beam towards the UE, all the different configurations of the RIS phase profiles should not have a significant impact onto the ultimate localization and synchronization performance, being the BS-RIS path likely illuminated with a negligible amount of transmitted power. In other words, when $\boldsymbol{F}_{:,i}^{\mathrm{BS}} = \boldsymbol{a}_{\mathrm{BS}}^{*}\left(\theta_{\mathrm{B,U}}^{(i)}\right)$ or $\boldsymbol{F}_{:,i}^{\mathrm{BS}} = \dot{\boldsymbol{a}}_{\mathrm{BS}}^{*}\left(\theta_{\mathrm{B,U}}^{(i)}\right)$, we propose to neglect the transmission of the $2L_{\mathrm{RIS}}$ different RIS beams $\boldsymbol{F}_{:,j}^{\mathrm{RIS}} = \boldsymbol{b}_{\mathrm{RIS}}^{*}\left(\theta_{\mathrm{R,U}}^{(j)}\right)$ and $\boldsymbol{F}_{:,j+L_{\mathrm{RIS}}}^{\mathrm{RIS}} = \widetilde{\dot{\boldsymbol{b}}}_{\mathrm{RIS}}^{*}\left(\theta_{\mathrm{R,U}}^{(j)}\right)$, for $j = 1, \ldots, L_{\mathrm{RIS}}$, and use for the
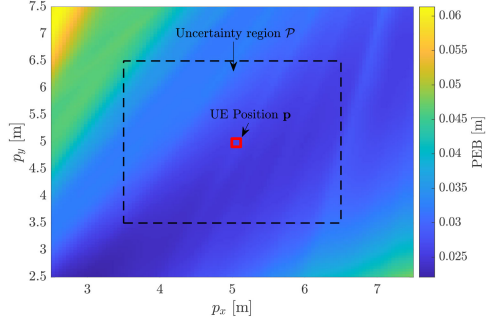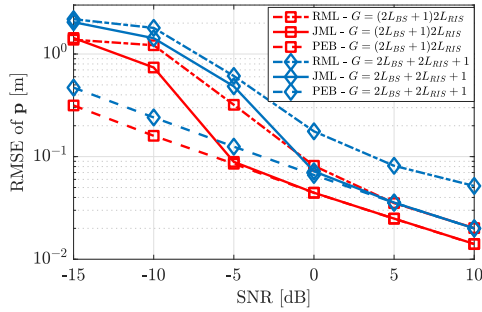
(a) PEB evaluation for the proposed robust joint BS-RIS design strategy.
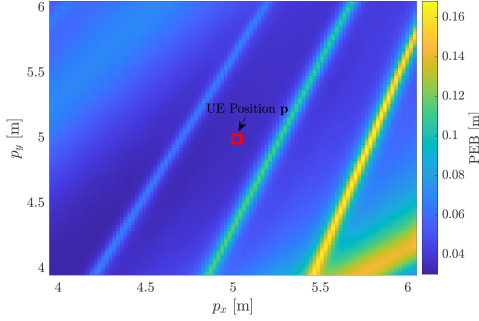
(b) PEB evaluation for non-robust joint BS-RIS design strategy.

Fig. 7. Comparison between the PEBs with robust and non-robust joint BS-RIS beamforming design strategies.

Fig. 8. Performance comparison between the case of full number of transmitted beams $G = (2L_{\mathrm{BS}} + 1)2L_{\mathrm{RIS}}$ and proposed heuristic using a reduced $G = 2L_{\mathrm{BS}} + 2L_{\mathrm{RIS}} + 1$.

corresponding pairs $\{\boldsymbol{F}_{:,i}^{\mathrm{BS}}, \boldsymbol{F}_{:,j}^{\mathrm{RIS}}\}$ a single configuration of the RIS phase profile given by $\boldsymbol{F}_{:,j}^{\mathrm{RIS}} = \boldsymbol{b}_{\mathrm{RIS}}(\theta_{\mathrm{R,U}}^c)$. Conversely, when the BS is transmitting the beam towards the RIS, that is, the precoding vector is set to $\boldsymbol{F}_{:,i}^{\mathrm{BS}} = \boldsymbol{a}_{\mathrm{BS}}^*(\theta_{\mathrm{B,R}})$, we consider for $\boldsymbol{F}_{:,j}^{\mathrm{RIS}}$ all the $2L_{\mathrm{RIS}}$ possible configurations of the RIS phase profile. In doing so, the signal received by the UE in (8) will be always observed for different RIS phase profiles, providing the necessary information to estimate the AoD $\theta_{\mathrm{R,U}}$. This procedure allows us to reduce the total number of transmitted beams to $G = 2L_{\mathrm{BS}} + 2L_{\mathrm{RIS}} + 1$.

To validate such an intuition, in Fig. 8 we compare the RMSEs on the estimation of $\boldsymbol{p}$ and the related PEBs as a function of the SNR, for both cases of full and reduced number of transmissions G. By comparing the PEBs in Fig. 8, we appreciate a slight degradation of the theoretical accuracy achievable in case of reduced G (analogous behavior is obtained for Δ).
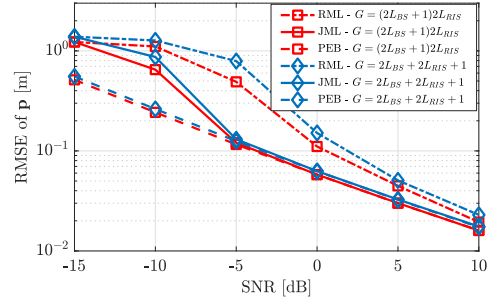
Fig. 9. Performance comparison between the case of full number of transmitted beams $G = (2L_{\mathrm{BS}} + 1)2L_{\mathrm{RIS}}$ and proposed heuristic using a reduced $G = 2L_{\mathrm{BS}} + 2L_{\mathrm{RIS}} + 1$ for increased AoD separation between UE and RIS.

Interestingly, despite the more challenging scenario, the JML estimator combined with the proposed power allocation strategy (solid curves with $\diamond$ marker) is still able to provide very accurate localization performance, though attaining the bounds at higher SNR of 0 dB. In this respect, an important trade-off between the estimation accuracy and the total number of transmission tends to emerge: for this specific case, the proposed heuristic leads to a 85% reduction in the number of involved transmissions (and, consequently, in the time needed to localize and synchronize the UE), but at the price of slightly increased values of RMSEs and related bounds. In Section S-V of the supplemental material, we have conducted a similar analysis for the case in which the uncertainty region $\mathcal{P}$ has been increased to 5 m along each direction. The obtained results reveal that the gaps between the estimation performance in cases of full and reduced G tend to increase as the uncertainty increases. This behavior can be explained by noting that the proposed heuristic is based on the underling assumption that almost no power is received by the RIS when the BS is transmitting a beam in the directions of the UE. However, when the uncertainty region $\mathcal{P}$ grows, the corresponding set of AoDs from the BS to the UE $\{\widetilde{\theta}_{\mathrm{B,U}}^{(i)}\}_{i=1}^{2L_{\mathrm{BS}}}$ progressively spans an increased area and, consequently, some of the beams directed towards the UE could likely illuminate the RIS path with a non-negligible amount of power. In these cases, the different configurations of the RIS phase profiles start to have a noticeable effect on the resulting estimation accuracy, thus preventing the possibility to reduce G without experiencing evident performance losses. Overall, an interesting insight can be derived from this analysis: the smaller the initial uncertainty about the UE position, the shorter the time required to localize and synchronize it accurately.

To further corroborate these insights, we consider a second different scenario in which the UE position is moved to $\boldsymbol{p} = [3 \; -1]^{\mathsf{T}}$ m, so that the angular separation $|\theta_{\mathrm{B,U}} - \theta_{\mathrm{R,U}}|$ between the UE and the RIS increases from $14.7°$ to $48.7°$ in this new configuration, while the rest of the parameters are kept the same as for Fig. 8. The effect of this change on the gap between the estimation performances of the full-G and reduced-G cases can be observed in Fig. 9. More specifically, the gap between the RMSEs on the estimation of $\boldsymbol{p}$ of the full-G and reduced-G cases is significantly reduced by moving the UE to a location where the AoD difference becomes much larger (analogous behavior is obtained for Δ). This behavior is perfectly in line with our previous findings and can be explained by noting that the RIS practically receives a negligible amount of power when the BS is
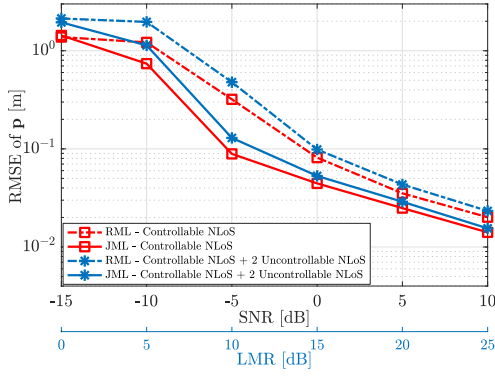
Fig. 10.   Performance comparison between the matched scenario with a single controllable NLoS path through RIS and the mismatched scenario with two additional uncontrollable NLoS paths.

illuminating the UE, being the extent of the uncertainty region $\mathcal{P}$ not sufficiently large to include beams that illuminate the RIS. Hence, it can be concluded that, for scenarios with widely separated AoDs and sufficiently small uncertainty regions, it is reasonable to employ the codebook with reduced G since it provides almost the same performance as in the case of full G using smaller number of transmissions. In this respect, it is worth noting that the gap between the two cases will similarly reduce also in the cases in which the AoDs from BS to RIS and UE are close, but the uncertainty region is small enough to guarantee that the beams do not illuminate the RIS path. Overall, we can conclude that the performances in cases of full and reduced G are mainly related to both the UE location and the extent of the uncertainty region $\mathcal{P}$.

5) Performance Assessment in Presence of Uncontrollable Multipath: To further challenge the proposed approach, we also investigate a scenario accounting for the simultaneous presence of the controllable NLoS path through the RIS, as well as of two additional uncontrollable NLoS paths generated by two local scatterers in the surrounding environment, located at unknown positions $\boldsymbol{m}_1 = [2 \ 7]^\mathsf{T}$ m and $\boldsymbol{m}_2 = [6 \ 2]^\mathsf{T}$, respectively. In doing so, we can test the robustness of the algorithms in a propagation environment that is mismatched with respect to the model considered at design stage. Assuming that each uncontrollable NLoS path is characterized by a single dominant ray, we generate the absolute value of the complex amplitudes as $|\alpha_{\mathrm{NLOS},i}| = \Gamma\lambda_c/(4\pi[\|\boldsymbol{m}_i\| + \|\boldsymbol{m}_i - \boldsymbol{p}\|)])$, with $\Gamma = 0.7$ reflection coefficient [45]. To analyze the robustness under different multipath conditions, we keep fixed the power of the uncontrollable NLoS paths (denoted by P_{NLoS}^i, $i = 1, 2$) and increase only the power along the LoS path (denoted by P_{LoS}) and the controllable RIS path, using the LoS-to-multipath ratio (LMR) defined as LMR $= P_{\mathrm{LOS}}/\sum_{i=1}^2 P_{\mathrm{NLoS}}^i$. For the considered setup, varying the SNR in the range from -15 dB up to 10 dB corresponds to a LMR varying from 0 dB up to 25 dB, with 5 dB steps. In Fig. 10, we show the evolution of the RMSEs on the estimation of $\boldsymbol{p}$ as a function of the SNR, for both cases with and without uncontrollable NLoS paths. The obtained results reveal that the proposed approach is effective also in this more challenging scenario, with both the RML and JML algorithms that exhibit a slight degradation of the achieved localization performance only for small values of the SNR (similar considerations hold

true for the clock offset Δ), that is, when the multipath in terms of LMR is more severe.

VIII. CONCLUSION

In this paper, we have considered the problem of joint localization and synchronization of a single-antenna UE served by a single BS in the presence of a RIS, assuming the existence of a LoS path and a controllable NLoS path through the RIS. To maximize the performance of localization and synchronization under UE location uncertainty, a novel codebook-based low-complexity design strategy for joint optimization of active BS precoding and passive RIS phase shifts has been proposed, based on the derived low-dimensional structure of precoders and phase profiles. In addition, we have developed a reduced-complexity ML-based estimator by exploiting the special signal structure that enables decoupled estimation of UE location and clock offset. Extensive simulations showed that the proposed joint BS-RIS beamforming approach provides significant improvements in both localization and synchronization performance (on the order of meters and nanoseconds, respectively, at SNR $= -10$ dB) over the state-of-the-art benchmarks. Moreover, the proposed estimator is able to attain the corresponding theoretical limits at a relatively low SNR (around -5 dB) and found to be resilient against uncontrolled multipath. As a future direction, we plan to evaluate the impact of discrete RIS phase shifts on the estimation performance.

APPENDIX A
PROOF OF PROPOSITION 1

Following similar arguments as in [38, App. C], we represent the covariance matrix in (21) for the g-th transmission as

$$\boldsymbol{X}_g = \boldsymbol{\Gamma}_g \boldsymbol{\Gamma}_g^\mathsf{H}, \tag{53}$$

where $\boldsymbol{\Gamma}_g$ admits a decomposition

$$\boldsymbol{\Gamma}_g = \boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}} \boldsymbol{\Gamma}_g + \boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}^\perp \boldsymbol{\Gamma}_g, \tag{54}$$

with $\boldsymbol{\Pi}_{\boldsymbol{X}} \stackrel{\text{def}}{=} \boldsymbol{X} \left(\boldsymbol{X}^H \boldsymbol{X}\right)^{-1} \boldsymbol{X}^H$ denoting the orthogonal projector onto the columns of $\boldsymbol{X}$ and $\boldsymbol{\Pi}_{\boldsymbol{X}}^\perp \stackrel{\text{def}}{=} \boldsymbol{I} - \boldsymbol{\Pi}_{\boldsymbol{X}}$. Then, $\boldsymbol{X}_g$ in (53) can be re-written using (54) as

$$\boldsymbol{X}_g = \overline{\boldsymbol{X}}_g + \widetilde{\boldsymbol{X}}_g, \tag{55}$$

where

$$\overline{\boldsymbol{X}}_g \stackrel{\text{def}}{=} \boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}} \boldsymbol{\Gamma}_g \boldsymbol{\Gamma}_g^\mathsf{H} \boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}} \tag{56}$$

$$\widetilde{\boldsymbol{X}}_g \stackrel{\text{def}}{=} \boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}} \boldsymbol{\Gamma}_g \boldsymbol{\Gamma}_g^\mathsf{H} \boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}^\perp + \boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}^\perp \boldsymbol{\Gamma}_g \boldsymbol{\Gamma}_g^\mathsf{H} \boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}$$
$$+ \boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}^\perp \boldsymbol{\Gamma}_g \boldsymbol{\Gamma}_g^\mathsf{H} \boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}^\perp. \tag{57}$$

Since $\boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}^\perp \boldsymbol{A}_{\mathrm{BS}} = \boldsymbol{0}$ by definition, we have

$$\boldsymbol{A}_{\mathrm{BS}}^\mathsf{H} \widetilde{\boldsymbol{X}}_g \boldsymbol{A}_{\mathrm{BS}} = \boldsymbol{0}. \tag{58}$$

We now provide three lemmas to facilitate the proof of Prop. 1.

Lemma 2: The FIM $\boldsymbol{J}_\gamma$ in (7) does not depend on the component $\widetilde{\boldsymbol{X}}_g$ of $\boldsymbol{X}_g$ in (55).

Proof: Based on the definition of $\boldsymbol{A}_{\mathrm{BS}}$ in (22) and the FIM elements in Section S-I-B in the supplemental material, we observe that the dependence of the FIM $\boldsymbol{J}_\gamma$ on $\boldsymbol{X}_g$ is only through the elements of $\boldsymbol{A}_{\mathrm{BS}}^\mathsf{H} \boldsymbol{X}_g \boldsymbol{A}_{\mathrm{BS}} \in \mathbb{C}^{3\times3}$. Then, it follows

from (55) and (58) that the FIM does not depend on $\widetilde{\boldsymbol{X}}_g$, i.e., the dependence of the FIM on $\boldsymbol{X}_g$ is only through $\overline{\boldsymbol{X}}_g$ in (55). ∎

Remark 4: The component $\widetilde{\boldsymbol{X}}_g$ of $\boldsymbol{X}_g$ in (55) contributes non-negatively to the total power consumption, i.e., $\mathrm{tr}(\widetilde{\boldsymbol{X}}_g) \geq 0$.

Proof: Opening up the terms in $\widetilde{\boldsymbol{X}}_g$ in (57), we have

$$
\begin{aligned}
\mathrm{tr}(\widetilde{\boldsymbol{X}}_g) &= \mathrm{tr}\left(\boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}\boldsymbol{\Gamma}_g\boldsymbol{\Gamma}_g^{\mathsf{H}}\boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}^{\perp}\right) + \mathrm{tr}\left(\boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}^{\perp}\boldsymbol{\Gamma}_g\boldsymbol{\Gamma}_g^{\mathsf{H}}\boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}\right) \\
&\quad + \mathrm{tr}\left(\boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}^{\perp}\boldsymbol{\Gamma}_g\boldsymbol{\Gamma}_g^{\mathsf{H}}\boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}^{\perp}\right) \\
&= \mathrm{tr}\left(\boldsymbol{\Gamma}_g\boldsymbol{\Gamma}_g^{\mathsf{H}}\boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}^{\perp}\boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}\right) + \mathrm{tr}\left(\boldsymbol{\Gamma}_g\boldsymbol{\Gamma}_g^{\mathsf{H}}\boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}\boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}^{\perp}\right) \\
&\quad + \left\|\boldsymbol{\Gamma}_g^{\mathsf{H}}\boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}^{\perp}\right\|_{\mathsf{F}}^2 \\
&= \left\|\boldsymbol{\Gamma}_g^{\mathsf{H}}\boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}^{\perp}\right\|_{\mathsf{F}}^2 \geq 0,
\end{aligned}
\tag{59}
$$

where $\|\cdot\|_{\mathsf{F}}$ represents the matrix Frobenius norm. ∎

Lemma 3: The component $\widetilde{\boldsymbol{X}}_g^{\star}$ in (55) of an optimal $\boldsymbol{X}_g^{\star}$ obtained as the solution to (21) satisfies $\mathrm{tr}(\widetilde{\boldsymbol{X}}_g^{\star}) = 0$.

Proof: To prove the lemma, we resort to proof by contradiction. For a given optimal solution

$$
\boldsymbol{X}_g^{\star} = \overline{\boldsymbol{X}}_g^{\star} + \widetilde{\boldsymbol{X}}_g^{\star}, g = 1, \dots, G
\tag{60}
$$

with $\mathrm{tr}(\widetilde{\boldsymbol{X}}_g^{\star}) > 0$ for some g, consider an alternative solution

$$
\boldsymbol{X}_g^{\star\star} = \overline{\boldsymbol{X}}_g^{\star\star} + \widetilde{\boldsymbol{X}}_g^{\star\star}, g = 1, \dots, G
\tag{61}
$$

where

$$
\overline{\boldsymbol{X}}_g^{\star\star} \overset{\mathrm{def}}{=} \overline{\boldsymbol{X}}_g^{\star}\left(1 + \frac{\mathrm{tr}\left(\sum_{g=1}^{G}\widetilde{\boldsymbol{X}}_g^{\star}\right)}{\mathrm{tr}\left(\sum_{g=1}^{G}\overline{\boldsymbol{X}}_g^{\star}\right)}\right)
\tag{62}
$$

$$
\widetilde{\boldsymbol{X}}_g^{\star\star} \overset{\mathrm{def}}{=} \boldsymbol{0}.
\tag{63}
$$

It can be readily verified from (60)–(63) that

$$
\mathrm{tr}\left(\sum_{g=1}^{G}\boldsymbol{X}_g^{\star\star}\right) = \mathrm{tr}\left(\sum_{g=1}^{G}\boldsymbol{X}_g^{\star}\right).
\tag{64}
$$

In addition, from Lemma 2, we note that the FIM obtained for $\boldsymbol{X}_g^{\star}$ in (60) and $\boldsymbol{X}_g^{\star\star}$ in (61) depend only on $\boldsymbol{A}_{\mathrm{BS}}^{\mathsf{H}}\overline{\boldsymbol{X}}_g^{\star}\boldsymbol{A}_{\mathrm{BS}}$ and $\boldsymbol{A}_{\mathrm{BS}}^{\mathsf{H}}\overline{\boldsymbol{X}}_g^{\star\star}\boldsymbol{A}_{\mathrm{BS}}$, respectively. Since $\boldsymbol{A}_{\mathrm{BS}}^{\mathsf{H}}\overline{\boldsymbol{X}}_g^{\star\star}\boldsymbol{A}_{\mathrm{BS}} = \zeta\boldsymbol{A}_{\mathrm{BS}}^{\mathsf{H}}\overline{\boldsymbol{X}}_g^{\star}\boldsymbol{A}_{\mathrm{BS}}$ for some $\zeta > 1$ according to (62), the alternative solution $\boldsymbol{X}_g^{\star\star}$ in (61) would achieve smaller PEB than the optimal solution $\boldsymbol{X}_g^{\star}$ in (60) (due to scaling of the FIM by $\zeta > 1$). Combining this with (64) shows that $\boldsymbol{X}_g^{\star}$ cannot be an optimal solution of (21), which completes the proof. ∎

Based on (59) in Remark 4, it can be observed that $\mathrm{tr}(\widetilde{\boldsymbol{X}}_g) = 0$ implies $\boldsymbol{\Gamma}_g^{\mathsf{H}}\boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}^{\perp} = \boldsymbol{0}$, which in turn yields $\widetilde{\boldsymbol{X}}_g = \boldsymbol{0}$ using (57). Hence, from Remark 4 and Lemma 3, we infer that $\widetilde{\boldsymbol{X}}_g^{\star}$ of an optimal $\boldsymbol{X}_g^{\star}$ should satisfy $\widetilde{\boldsymbol{X}}_g^{\star} = \boldsymbol{0}$. Finally, from (55) and (56), an optimal $\boldsymbol{X}_g^{\star}$ obtained as the solution to (21) can be expressed as

$$
\begin{aligned}
\boldsymbol{X}_g^{\star} &= \boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}}\boldsymbol{\Gamma}_g\boldsymbol{\Gamma}_g^{\mathsf{H}}\boldsymbol{\Pi}_{\boldsymbol{A}_{\mathrm{BS}}} \\
&= \boldsymbol{A}_{\mathrm{BS}}\boldsymbol{\Upsilon}_g\boldsymbol{A}_{\mathrm{BS}}^{\mathsf{H}},
\end{aligned}
\tag{65}
$$

where $\boldsymbol{\Upsilon}_g \overset{\mathrm{def}}{=} (\boldsymbol{A}_{\mathrm{BS}}^{\mathsf{H}}\boldsymbol{A}_{\mathrm{BS}})^{-1}\boldsymbol{A}_{\mathrm{BS}}^{\mathsf{H}}\boldsymbol{\Gamma}_g\boldsymbol{\Gamma}_g^{\mathsf{H}}\boldsymbol{A}_{\mathrm{BS}}(\boldsymbol{A}_{\mathrm{BS}}^{\mathsf{H}}\boldsymbol{A}_{\mathrm{BS}})^{-1}$, which completes the proof of Proposition 1.

We note that there exists an equivalent orthogonal solution $\boldsymbol{\Upsilon}_g$ (corresponding to the full-rank version of $\boldsymbol{A}_{\mathrm{BS}}$) in (65), leading to the same covariance $\boldsymbol{X}_g$, as shown in Section S-V in the supplemental material. Moreover, it is worth highlighting that $\boldsymbol{X}_g$ and $\boldsymbol{\Upsilon}_g$ are two identical solutions (having different dimensions) of the problem (21), and thus one can always be obtained from the other using (65) and

$$
\boldsymbol{\Upsilon}_g = (\boldsymbol{A}_{\mathrm{BS}}^{\mathsf{H}}\boldsymbol{A}_{\mathrm{BS}})^{-1}\boldsymbol{A}_{\mathrm{BS}}^{\mathsf{H}}\boldsymbol{X}_g\boldsymbol{A}_{\mathrm{BS}}(\boldsymbol{A}_{\mathrm{BS}}^{\mathsf{H}}\boldsymbol{A}_{\mathrm{BS}})^{-1}.
\tag{66}
$$

In other words, *(i)* one can either solve (21) directly with respect to $\boldsymbol{X}_g \in \mathbb{C}^{N_{\mathrm{BS}} \times N_{\mathrm{BS}}}$, or, *(ii)* one can solve (21) with respect to $\boldsymbol{\Upsilon}_g \in \mathbb{C}^{3 \times 3}$ by inserting the relation $\boldsymbol{X}_g = \boldsymbol{A}_{\mathrm{BS}}\boldsymbol{\Upsilon}_g\boldsymbol{A}_{\mathrm{BS}}^{\mathsf{H}}$ into both the objective (21) and the constraint (20b), and find the corresponding $\boldsymbol{X}_g$ through $\boldsymbol{X}_g = \boldsymbol{A}_{\mathrm{BS}}\boldsymbol{\Upsilon}_g\boldsymbol{A}_{\mathrm{BS}}^{\mathsf{H}}$.

APPENDIX B
PROOF OF PROPOSITION 2

From Section S-I-B2 in the supplemental material, we observe that the FIM $\boldsymbol{J}_\gamma^{\mathrm{bd}}$ in (18) depends on $\boldsymbol{\Psi}_g$ only through the elements of the matrix $\boldsymbol{B}_{\mathrm{RIS}}^{\mathsf{H}}\boldsymbol{\Psi}_g\boldsymbol{B}_{\mathrm{RIS}} \in \mathbb{C}^{2 \times 2}$. Then, the claim in the proposition can easily be proved by employing similar arguments to those in Appendix A.

REFERENCES

[1] A. Fascista, A. Coluccia, H. Wymeersch, and G. Seco-Granados, "RIS-aided joint localization and synchronization with a single-antenna mmWave receiver," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2021, pp. 4455–4459.

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

[3] H. Wymeersch, G. Seco-Granados, G. Destino, D. Dardari, and F. Tufvesson, "5G mmWave positioning for vehicular networks," *IEEE Wireless Commun.*, vol. 24, no. 6, pp. 80–86, Dec. 2017.

[4] J. A. del Peral-Rosado, G. Seco-Granados, S. Kim, and J. A. López-Salcedo, "Network design for accurate vehicle localization," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4316–4327, May 2019.

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

[6] A. Fascista, A. Coluccia, H. Wymeersch, and G. ranados, "Low-complexity accurate mmwave positioning for single-antenna users based on angle-of-departure and adaptive beamforming," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2020, pp. 4866–4870.

[7] S. Dwivedi *et al.*, "Positioning in 5G networks," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 38–44, Nov. 2021, doi: 10.1109/MCOM.011.2100091.

[8] A. Fascista, A. De Monte, A. Coluccia, H. Wymeersch, and G. Seco-Granados, "Low-complexity downlink channel estimation in mmwave multiple-input single-output systems," *IEEE Wireless Commun. Lett.*, vol. 11, no. 3, pp. 518–522, Mar. 2022.

[9] J. Gante, L. Sousa, and G. Falcao, "Dethroning GPS: Low-power accurate 5G positioning systems using machine learning," *IEEE Trans. Emerg. Sel. Top. Circuits Syst.*, vol. 10, no. 2, pp. 240–252, Jun. 2020.

[10] J. Gante *et al.*, "Deep learning architectures for accurate millimeter wave positioning in 5G," *Neural Proc. Lett.*, vol. 51, no. 1, pp. 487–514, 2020.

[11] Y. Ge *et al.*, "5G SLAM using the clustering and assignment approach with diffuse multipath," *Sensors*, vol. 20, no. 16, 2020, Art. no. 4656.

[12] A. Fascista, A. Coluccia, H. Wymeersch, and G. Seco-Granados, "Downlink single-snapshot localization and mapping with a single-antenna receiver," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4672–4684, Jul. 2021.

[13] K. Witrisal *et al.*, "High-accuracy localization for assisted living: 5G systems will turn multipath channels from foe to friend," *IEEE Signnal Process. Mag.*, vol. 33, no. 2, pp. 59–70, Mar. 2016.

[14] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M. -S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116753–116773, 2019.

[15] S. Hu, F. Rusek, and O. Edfors, "Beyond massive MIMO: The potential of positioning with large intelligent surfaces," *IEEE Trans. Signal Process.*, vol. 66, no. 7, pp. 1761–1774, Apr. 2018.

[16] H. Wymeersch, J. He, B. Denis, A. Clemente, and M. Juntti, "Radio localization and mapping with reconfigurable intelligent surfaces: Challenges, opportunities, and research directions," *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 52–61, Dec. 2020.

[17] W. Wang and W. Zhang, "Joint beam training and positioning for intelligent reflecting surfaces assisted millimeter wave communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6282–6297, Oct. 2021.

[18] A. Elzanaty, A. Guerra, F. Guidi, and M.-S. Alouini, "Reconfigurable intelligent surfaces for localization: Position and orientation error bounds," *IEEE Trans. Signal Process.*, vol. 69, pp. 5386–5402, 2021.

[19] K. Keykhosravi, M. F. Keskin, G. Seco-Granados, and H. Wymeersch, "SISO RIS-enabled joint 3D downlink localization and synchronization," in *Proc. IEEE Int. Conf. Commun.*, 2021, pp. 1–6.

[20] H. Zhang, H. Zhang, B. Di, K. Bian, Z. Han, and L. Song, "Towards ubiquitous positioning by leveraging reconfigurable intelligent surface," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 284–288, Jan. 2021.

[21] M. Di Renzo *et al.*, "Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 1, pp. 1–20, 2019.

[22] C. Pan *et al.*, "Reconfigurable intelligent surfaces for 6G systems: Principles, applications, and research directions," *IEEE Comm. Mag.*, vol. 59, no. 6, pp. 14–20, Jun. 2021.

[23] M. Rahal, B. Denis, K. Keykhosravi, B. Uguen, and H. Wymeersch, "RIS-enabled localization continuity under near-field conditions," in *Proc. IEEE 22nd Int. Workshop Signal Process. Adv. Wireless Commun.*, 2021, pp. 436–440, doi: 10.1109/SPAWC51858.2021.9593200.

[24] J. He, H. Wymeersch, L. Kong, O. Silvén, and M. Juntti, "Large intelligent surface for positioning in millimeter wave MIMO systems," in *Proc. IEEE 91st Veh. Technol. Conf.*, 2020, pp. 1–5.

[25] W. Roh *et al.*, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 106–113, Feb. 2014.

[26] 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 Signnal Process. Mag.*, vol. 31, no. 6, pp. 102–112, Nov. 2014.

[27] B. Zhou, A. Liu, and V. Lau, "Successive localization and beamforming in 5G mmwave MIMO communication systems," *IEEE Trans. Signal Process.*, vol. 67, no. 6, pp. 1620–1635, Mar. 2019.

[28] A. Kakkavas, H. Wymeersch, G. Seco-Granados, M. H. C. García, R. A. Stirling-Gallacher, and J. A. Nossek, "Power allocation and parameter estimation for multipath-based 5G positioning," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7302–7316, Nov. 2021.

[29] A. Kakkavas, G. Seco-Granados, H. Wymeersch, M. H. C. Garcia, R. A. Stirling-Gallacher, and J. A. Nossek, "5G downlink multi-beam signal design for LOS positioning," in *Proc. IEEE Glob. Commun. Conf.*, 2019, pp. 1–6, doi: 10.1109/GLOBECOM38437.2019.9013374.

[30] M. F. Keskin, F. Jiang, F. Munier, G. Seco-Granados, and H. Wymeersch, "Optimal spatial signal design for mmwave positioning under imperfect synchronization," *IEEE Trans. Veh. Technol.*, vol. 71, no. 5, pp. 5558–5563, May 2022.

[31] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, and H. V. Poor, "Hybrid beamforming for reconfigurable intelligent surface based multi-user communications: Achievable rates with limited discrete phase shifts," *IEEE J. Sel. Areas Commun*, vol. 38, no. 8, pp. 1809–1822, Aug. 2020.

[32] X. Ma, S. Guo, H. Zhang, Y. Fang, and D. Yuan, "Joint beamforming and reflecting design in reconfigurable intelligent surface-aided multi-user communication systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 3269–3283, May 2021.

[33] C. Huang, R. Mo, and C. Yuen, "Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1839–1850, Aug. 2020.

[34] H. Guo, Y. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3064–3076, May 2020.

[35] Z. Zhou, N. Ge, Z. Wang, and L. Hanzo, "Joint transmit precoding and reconfigurable intelligent surface phase adjustment: A decomposition-aided channel estimation approach," *IEEE Trans. Commun.*, vol. 69, no. 2, pp. 1228–1243, Feb. 2021.

[36] M.-M. Zhao, A. Liu, and R. Zhang, "Outage-constrained robust beamforming for intelligent reflecting surface aided wireless communication," *IEEE Trans. Signal Process.*, vol. 69, pp. 1301–1316, 2021.

[37] M. Najafi, V. Jamali, R. Schober, and H. V. Poor, "Physics-based modeling and scalable optimization of large intelligent reflecting surfaces," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2673–2691, Apr. 2021.

[38] J. Li, L. Xu, P. Stoica, K. W. Forsythe, and D. W. Bliss, "Range compression and waveform optimization for MIMO radar: A Cramér–Rao bound based study," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 218–232, Jan. 2008.

[39] S. H. Talisa, K. W. O'Haver, T. M. Comberiate, M. D. Sharp, and O. F. Somerlock, "Benefits of digital phased array radars," *Proc. IEEE IRE*, vol. 104, no. 3, pp. 530–543, Mar. 2016.

[40] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, Inc., 1993.

[41] R. Mendrzik, F. Meyer, G. Bauch, and M. Z. Win, "Enabling situational awareness in millimeter wave massive MIMO systems," *IEEE J. Sel. Top. Signal Process.*, vol. 13, no. 5, pp. 1196–1211, Sep. 2019.

[42] S. Boyd *et al.*, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[43] M. Fukuda *et al.*, "Branch-and-cut algorithms for the bilinear matrix inequality eigenvalue problem," *Comput. Optim. Appl.*, vol. 19, no. 1, pp. 79–105, 2001.

[44] R. Mendrzik, H. Wymeersch, G. Bauch, and Z. Abu-Shaban, "Harnessing NLOS components for position and orientation estimation in 5G millimeter wave MIMO," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 93–107, Jan. 2019.

[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 Trans. Wireless Commun.*, vol. 17, no. 8, pp. 4939–4954, Aug. 2018.

[46] A. Kakkavas, M. H. C. García, R. A. Stirling-Gallacher, and J. A. Nossek, "Performance limits of single-anchor millimeter-wave positioning," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5196–5210, Nov. 2019.

[47] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," *Recent Adv. Learn. Control*, V. Blondel, S. Boyd, and H. Kimura, Eds. 2008, pp. 95–110, Lecture Notes in Control and Information Sciences. [Online]. Available: http://stanford.edu/~boyd/graph_dcp.html

[48] H. Zhao, N. Zhang, and Y. Shen, "Beamspace direct localization for large-scale antenna array systems," *IEEE Trans. Signal Process.*, vol. 68, pp. 3529–3544, 2020.

[49] A. Fascista, A. Coluccia, H. Wymeersch, and G. Seco-Granados, "Millimeter-wave downlink positioning with a single-antenna receiver," *IEEE Trans. Wireless Commun.*, vol. 18, no. 9, pp. 4479–4490, Sep. 2019.

[50] U. Nickel, "Overview of generalized monopulse estimation," *IEEE Aerosp. Elect. Syst. Mag.*, vol. 21, no. 6, pp. 27–56, Jun. 2006.

[51] X. Zhang, P. K. Willett, and Y. Bar-Shalom, "Monopulse radar detection and localization of multiple unresolved targets via joint bin processing," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1225–1236, Apr. 2005.

[52] W. W.-L. Li, Y. Shen, Y. J. Zhang, and M. Z. Win, "Robust power allocation for energy-efficient location-aware networks," *IEEE/ACM Trans. Netw.*, vol. 21, no. 6, pp. 1918–1930, Dec. 2013.

[53] M. Z. Win, W. Dai, Y. Shen, G. Chrisikos, and H. Vincent Poor, "Network operation strategies for efficient localization and navigation," *Proc. IEEE*, vol. 106, no. 7, pp. 1224–1254, Jul. 2018.

[54] M. F. Keskin, A. D. Sezer, and S. Gezici, "Optimal and robust power allocation for visible light positioning systems under illumination constraints," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 527–542, Jan. 2019.

[55] A. Nemirovski, "Interior point polynomial time methods in convex programming," *Lecture Notes*, vol. 42, no. 16, pp. 3215–3224, 2004.

[56] J. A. Zhang, X. Huang, Y. J. Guo, J. Yuan, and R. W. Heath, "Multibeam for joint communication and radar sensing using steerable analog antenna arrays," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 671–685, Jan. 2019.

[57] S. J. Orfanidis, *Electromagnetic Waves and Antennas*. New Brunswick, NJ, USA: Rutgers Univ., 2002.

[58] J. Tranter, N. D. Sidiropoulos, X. Fu, and A. Swami, "Fast unit-modulus least squares with applications in beamforming," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2875–2887, Jun. 2017.

[59] S. Noh, M. D. Zoltowski, and D. J. Love, "Multi-resolution codebook and adaptive beamforming sequence design for millimeter wave beam alignment," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5689–5701, Sep. 2017.

[60] D. Love and R. W. Heath, "Equal gain transmission in multiple-input multiple-output wireless systems," *IEEE Trans. Commun.*, vol. 51, no. 7, pp. 1102–1110, Jul. 2003.