Bayesian User Localization and Tracking for Reconfigurable Intelligent Surface Aided MIMO Systems

Boyu Teng, Xiaojun Yuan, Senior Member, IEEE, Rui Wang, Senior Member, IEEE, and Shi Jin, Senior Member, IEEE

Abstract—In this paper, we study the user localization and tracking problem in the reconfigurable intelligent surface (RIS) aided multiple-input multiple-output (MIMO) system, where a multi-antenna base station (BS) and multiple RISs are deployed to assist the localization and tracking of a multi-antenna user. By establishing a probability transition model for user mobility, we develop a message-passing algorithm, termed the Bayesian user localization and tracking (BULT) algorithm, to estimate and track the user position and the angle-of-arrival (AoAs) at the user in an online fashion. We also derive Bayesian Cramér Rao bound (BCRB) to characterize the fundamental performance limit of the considered tracking problem. To improve the tracking performance, we optimize the beamforming design at the BS and the RISs to minimize the derived BCRB. Simulation results show that our BULT algorithm can perform close to the derived BCRB, and significantly outperforms the counterpart algorithms without exploiting the temporal correlation of the user location.

Index Terms—Reconfigurable intelligent surface, user localization, user tracking, MIMO, message passing.

I. INTRODUCTION

Integrated sensing and communication (ISAC) has been identified as a promising technology for the sixth-generation (6G) mobile communications [1]–[3], where the joint design of sensing and communication systems becomes necessary to meet the increasing demand for communication and sensing services [4], [5]. High-precision localization, as a critical issue in sensing, shows great potential needs in the emerging 6G application scenarios such as internet of vehicles (IoV), augmented reality (AR), virtual reality (VR), unmanned aerial vehicle (UAV) communications, etc. [6]–[8]. With large antenna array and broad spectrum bandwidth, 6G communication signals enable high resolution in the angular and delay domains, which makes possible high-precision localization based on radio signals [1]–[3]. [7].

From the electromagnetic propagation theory, the diffraction effect of radio signals becomes weaker as the frequency increases [1]. This phenomenon is particularly severe in 6G since 6G uses much higher frequency bands than earlier generations to support extremely high-speed communications. As a result, 6G communications rely heavily on the existence of line-of-sight (LoS) paths to maintain a sufficient receive power. As for localization, the position information of a target is mostly carried by the LoS path between the target and the radar, which is typically assumed to exist for reliable sensing [9]. Unfortunately, in mobile communication scenarios, transmitters and receivers are usually blocked by obstacles due to highly complicated propagation environments. To address this issue in ISAC tasks, reconfigurable intelligent surface (RIS, a.k.a. intelligent reflecting surface) is introduced as an energy-efficient and cost-effective device for performance enhancement [10]. A RIS is composed of a large number of passive reflecting elements that independently adjust the incident signals by inducing controllable amplitude and additional phase changes, and can be deployed in the wireless environment to create a virtual-line-of-sight (VLoS) path. Moreover, with efficient passive beamforming design, a RIS is able to focus the reflected signals in a desired direction, so as to significantly enhance the communication quality via the VLoS link [11], [12]. While the RIS aided communication system has been extensively studied (see, e.g., [13]–[19] and the references therein), the research on RIS-aided sensing is still in its infancy stage. Early works include RIS aided high-precision user localization [20], [21], mobile target tracking [22], the beamforming design of RIS in sensing tasks [23], etc.

The user localization methods for RIS-aided multiple-input multiple-output (MIMO) systems have been recently studied in [24]–[27]. The authors in [24] proposed a RIS-aided indoor user positioning scheme based on the received signal strength. By iterative optimization of the RIS phase shifts, a high resolution signal strength map is established to locate the target. In [25], a hierarchical-codebook based beam searching method was introduced to estimate the RIS-aided channel parameters which are further utilized to obtain accurate positioning. In [26], the authors estimated the channel angle parameters by maximum likelihood and obtains the user position by combining the channel information of RISs. For the MIMO orthogonal frequency division multiplexing (OFDM) system, the authors in [27] realised environment mapping and user localization by using the twin-RIS structure and exploiting array signal processing to obtain the channel parameters. In the above studies, user/target localization is realised in each time slot.
In this paper, we consider the user localization and tracking problem in a downlink MIMO system, where a multi-antenna base station (BS) and multiple RISs are deployed to assist the localization and tracking of a multi-antenna user. Specifically, the user receives the signals from the RISs and estimates the angle-of-arrivals (AoAs) based on the received signals. Each RIS acts as an anchor with its location known in advance. The user location is estimated by combining the angle information and the RIS location information. For user mobility, we build a probabilistic transition model of user movement by considering the fact that the user position changes continuously over time. Based on a factor-graph representation of the probability model, we develop a novel message-passing algorithm, termed the Bayesian user localization and tracking (BULT) algorithm, to efficiently solve the considered problem. In the algorithm design, various approximations are introduced to reduce the computational complexity involved in message passing.

As a benchmark, we derive the Bayesian Cramér Rao bound (BCRB) that serves as a theoretic mean square error (MSE) lower bound of the considered estimation problem. We show by simulation results that our BULT algorithm performs close to the derived BCRB at relatively high signal-to-noise ratio (SNR), and that our algorithm significantly outperforms the counterpart algorithms without exploiting the temporal correlation of the user location.

Furthermore, we consider the design of active beamforming at the BS and passive beamforming at the RISs. We develop an alternating algorithm for the design of active and passive beamforming to minimize the derived BCRB. Alternatively, to reduce computational complexity, we further propose a directional beamforming design at the BS and the RISs based on the estimated user location. We show that both beamforming strategies achieve similar performance, but the latter requires much lower complexity since it does not involve the iteration between active and passive beamforming optimizations.

The remainder of this paper is organized as follows. In Section II, we introduce the RISs aided MIMO system and formulate the user localization and tracking problem. In Section III, we derive the message calculations and develop the Bayesian user localization and tracking algorithm. In Section IV, we analyse the BCRB for the tracking parameters. In Section V, we discuss the beamforming design for the BS and the RISs. Numerical results are presented in Section VI, and the paper concludes in Section VII.

Notations: Throughout, bold lowercase letters and bold capital letters are used to respectively denote vectors and matrices. We use $(\cdot)^T$ and $(\cdot)^H$ to denote the transpose and the conjugate transpose respectively. We use $\text{Tr}(\mathbf{X})$ to denote the trace of $\mathbf{X}$, $\text{diag}(\mathbf{x})$ to denote the diagonal matrix with its diagonal entries given by $\mathbf{x}$, $[\mathbf{X}]_{i,j}$ to denote the $(i, j)$-th term of $\mathbf{X}$, and $\mathbf{I}$ to denote the identity matrix. We use $\mathcal{N}(\mathbf{x}; \mathbf{\mu}, \Sigma)$ and $\mathcal{CN}(\mathbf{x}; \mathbf{\mu}, \Sigma)$ to denote the real Gaussian distribution and the circularly-symmetric Gaussian distribution with mean vector $\mathbf{\mu}$ and covariance matrix $\Sigma$. We use $\mathbb{E}[\cdot]$ to denote the expectation operator, $\odot$ to denote Hadamard product, $\|\cdot\|_p$ to denote the $\ell_p$ norm. $\mathcal{R}[\{\mathbf{X}\}]$ and $\mathcal{I}[\{\mathbf{X}\}]$ are respectively the real part and the imaginary part of $\mathbf{X}$.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Description

We consider a multi-RIS aided MIMO system consisting of a BS, a user, and $K$ RISs, as illustrated in Fig. 1. The BS is equipped with $N_B$ antennas, the user is equipped with $N_U$ antennas, and each RIS consists of $N_R$ reflecting elements. The antenna and reflecting elements are arranged as uniform linear arrays (ULAs) for the BS, the user, and the RISs; and the antenna/element interval is set to half of the carrier frequency wavelength $\lambda$. We assume that the VLoS path reflected by each RIS always exists while the LoS path propagated from the BS to the user is blocked by obstacles. We further assume downlink transmission where the transmitted signals of the BS are reflected by the RISs and then received by the user for localization and tracking.

Assume that a 3D Cartesian coordinate system has been set up appropriately. The position of the BS, the $i$-th RIS and the user in time slot $t$ are denoted by the three-dimension vectors $\mathbf{p}_{B}$, $\mathbf{p}_{R,i}$ and $\mathbf{p}_{U}$, respectively. Further assume that the position of the BS and the RISs can be acquired by the user and the BS accurately in advance since the BS and the RISs are deployed stationarily.

B. Channel Model

For the considered user tracking problem, the channel is in general correlated over time. The temporal correlation of the channel is presented in the next subsection by considering the geometric relationship of the user position and the channel parameters. In the following, we establish the geometric channel model for each time slot $t$. Specifically, the multi-path channel from the BS to the $i$-th RIS and from the $i$-th RIS to
the user are denoted by $\mathbf{H}_{RB,i}^{(t)} \in \mathbb{C}^{N_R \times N_b}$ and $\mathbf{H}_{UR,i}^{(t)} \in \mathbb{C}^{N_U \times N_R}$ respectively, modeled as \[ \mathbf{H}_{RB,i}^{(t)} = \frac{\alpha}{1 + \alpha} \mathbf{H}_{RB,i}^{0} + \frac{1}{1 + \alpha} \mathbf{H}_{RB,i}^{1}, \] \[ \mathbf{H}_{UR,i}^{(t)} = \frac{\alpha}{1 + \alpha} \mathbf{H}_{UR,i}^{0} + \frac{1}{1 + \alpha} \mathbf{H}_{UR,i}^{1}, \] \[ (1) \] \[ (2) \]

where $\mathbf{H}_{RB,i}^{0} \in \mathbb{C}^{N_R \times N_b}$ and $\mathbf{H}_{RB,i}^{1} \in \mathbb{C}^{N_R \times N_b}$ are the non-LoS components whose elements follow zero-mean complex Gaussian distributions; $\alpha$ is the Rician K-factor; $\mathbf{H}_{RB,i}^{0} \in \mathbb{C}^{N_R \times N_b}$ and $\mathbf{H}_{UR,i}^{0} \in \mathbb{C}^{N_R \times N_b}$ are the VLoS components given by \[ \mathbf{H}_{RB,i}^{0} = \rho_{RB,i} \mathbf{a}_R(\theta_{RB,i}) \mathbf{a}_B(\tilde{\theta}_{RB,i}), \] \[ \mathbf{H}_{UR,i}^{0} = \rho_{UR,i} \mathbf{a}_U(\theta_{UR,i}) \mathbf{a}_R(\tilde{\theta}_{UR,i}), \] \[ (3) \] \[ (4) \]

where $\rho_{RB,i}$ and $\rho_{UR,i}$ are the complex channel gains for the VLoS path; $\tilde{\theta}_{RB,i}$ and $\tilde{\theta}_{UR,i}$ are respectively the cosine of the angle of arrival (AoA) at the $i$-th RIS, the cosine of the angle of departure (AoD) at the BS, the cosine of the AoA at the user and the cosine of the AoD at the $i$-th RIS, and $t$ is the time index; $\mathbf{a}_R(\tilde{\theta}_{RB,i})$, $\mathbf{a}_U(\theta_{UR,i})$ and $\mathbf{a}_R(\tilde{\theta}_{UR,i})$ are steering vectors in the form of line spectrum as \[ \mathbf{a}_x(\theta) = [1, e^{j\pi\theta}, \ldots, e^{j\pi(N_b-1)\theta}], \quad \text{for } x \in \{R, U\}. \] \[ (5) \]

We note that $\theta_{RB,i}$ and $\theta_{UR,i}$ are invariant over time and are assumed to be known beforehand based on the knowledge of $\mathbf{p}_B$ and $\mathbf{p}_{R,i}$. From (1) and (2), the received signal $\mathbf{y}^{(t)} \in \mathbb{C}^{N_U}$ at the user antennas can be written as \[ \mathbf{y}^{(t)} = \sum_{i=1}^{K} \zeta_i^{(t)} \mathbf{H}_{UR,i}^{(t)} \mathbf{H}_{RB,i}^{(t)} x(t) + \mathbf{w}^{(t)}, \] \[ (6) \]

where the pilot signal $x(t)$ is set by $x(t) = 1$ for the user localization and tracking purpose; $\mathbf{f}^{(t)} \in \mathbb{C}^{N_b}$ is the beamforming vector at BS side; $\mathbf{w}^{(t)} \sim \mathcal{CN}(\mathbf{0}, \sigma_w^2 \mathbf{I})$ is the added Gaussian noise at user antennas; $\zeta_i^{(t)}$ is the reflection coefficient at BS which generally varies in different AoAs and AoDs due to physical imperfection [28]; $\mathbf{H}_{RB,i}^{(t)} = \text{diag}(\mathbf{H}_{RB,i}^{0}, \mathbf{H}_{RB,i}^{1})$ is the diagonal phase shift matrix of the $i$-th RIS with the phase shift vector $\omega_{i}^{(t)} = [\omega_{i,1}, \ldots, \omega_{i,N_b}]$, where $|\omega_{i,n}| = 1$ for $\forall i, n$. We assume that the phase shift of a reflecting element is continuously valued.

In user tracking, we are interested in the channel parameters of the VLoS path, i.e., $\mathbf{H}_{UR,i}$, with its key parameter $\theta_{UR,i}^{(t)}$ related to the user position by the geometric constraint as \[ \theta_{UR,i}^{(t)} = \frac{(\mathbf{p}_{R,i} - \mathbf{p}_U^{(t)})^T \mathbf{e}_U}{\|\mathbf{p}_{R,i} - \mathbf{p}_U^{(t)}\|_2}, \] \[ (7) \]

where $\mathbf{e}_U$ is the unit direction vector of the user antennas and obtained by the built-in sensor in advance. Thus, we separate the VLoS component from the other signal components. Specifically, the received signal in (6) is further expressed as \[ \mathbf{y}^{(t)} = \sum_{i=1}^{K} \zeta_i^{(t)} \mathbf{H}_{UR,i}^{(t)} \mathbf{H}_{RB,i}^{(t)} \mathbf{f}^{(t)} + \mathbf{n}^{(t)}, \] \[ (8) \]

where $\mathbf{n}^{(t)} \in \mathbb{C}^{N_U}$ is the interference-plus-noise term given by \[ \mathbf{n}^{(t)} = \sum_{i=1}^{K} \xi_i^{(t)} \mathbf{H}_{RB,i}^{(t)} \mathbf{H}_{RB,i}^{(t)} \mathbf{f}^{(t)} + \mathbf{H}_{UR,i}^{(t)} \mathbf{H}_{RB,i}^{(t)} \mathbf{f}^{(t)} + \mathbf{w}^{(t)}. \] \[ (9) \]

Considering the severe path loss of the non-LoS path and the beamforming setting for the BS and the RISs provided in Section V, we treat the non-LoS path signal as interference drawn from a complex Gaussian distribution [34], i.e., $\mathbf{n}^{(t)} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I})$. Therefore, the received signal in time slot $t$ can be further simplified by \[ \mathbf{y}^{(t)} = \sum_{i=1}^{K} \xi_i^{(t)} \mathbf{p}_{U,i}^T \mathbf{a}_U(\theta_{U,i}) \mathbf{a}_R(\tilde{\theta}_{U,i}) \mathbf{H}_{RB,i}^{(t)} \mathbf{f}^{(t)} + \mathbf{n}^{(t)}, \] \[ (10a) \]

\[ = \sum_{i=1}^{K} \rho_i^{(t)} \mathbf{a}_U(\theta_{U,i}) + \mathbf{n}^{(t)}, \] \[ (10b) \]

where \[ \rho_i^{(t)} = \mathbf{p}_{U,i}^T \mathbf{a}_U(\theta_{U,i}) \mathbf{a}_R(\tilde{\theta}_{U,i}) \mathbf{H}_{RB,i}^{(t)} \mathbf{f}^{(t)} \] \[ (11) \]

is the equivalent complex path gain of the VLoS path for the $i$-th RIS with $\mathbf{f}^{(t)} = \tilde{\mathbf{f}}_i \mathbf{p}_U^{(t)} \mathbf{p}_{R,i}$. We notice that the estimation of AoA in user $\theta_{U,i}$ falls into the category of line spectrum inference [29]. The estimation of angle parameters can help with user tracking, based on which our user tracking method is developed.

C. Probabilistic Problem Formulation

In this section, we describe the user tracking problem by building a probability model for user position $\theta_{U,i}^{(t)}$, user AoAs $\{\theta_{U,i}^{(t)}\}$ and observation signal $\mathbf{y}^{(t)}$. From (10b), the VLoS channel angle parameters $\theta_{U,i}^{(t)}$ corresponding to different RISs constitute the line spectrum of the received signal. With the geometric constraint in (7), the conditional probability density function (pdf) $p(\mathbf{y}^{(t)} | \mathbf{p}_U^{(t)})$ is represented as \[ p(\mathbf{y}^{(t)} | \mathbf{p}_U^{(t)}) = \delta(\mathbf{y}^{(t)} - (\mathbf{p}_{R,i} - \mathbf{p}_U^{(t)})^T \mathbf{e}_U) \] \[ \|\mathbf{p}_{R,i} - \mathbf{p}_U^{(t)}\|_2, \] \[ (12) \]

where $\delta(\cdot)$ is the Dirac delta function. Based on (10b), the likelihood function of $\mathbf{\theta}^{(t)} = [\theta_{U,1}, \ldots, \theta_{U,K}]^T$ and $\mathbf{\rho}^{(t)} = [\rho_{i,1}, \ldots, \rho_{i,K}]^T$ at time slot $t$ is given by \[ p(\mathbf{y}^{(t)} | \mathbf{\theta}^{(t)}, \mathbf{\rho}^{(t)}) = \mathcal{CN}(\mathbf{y}^{(t)}; \sum_{i=1}^{K} \rho_i^{(t)} \mathbf{a}_U(\theta_{U,i}), \sigma_n^2 \mathbf{I}). \] \[ (13) \]

Due to the mobility of the user and the uncertainty of the reflection coefficient $\zeta_i^{(t)}$, we model the prior distribution of $\rho_i^{(t)}$ as a complex Gaussian distribution as \[ p(\rho_i^{(t)}) = \mathcal{CN}(\rho_i^{(t)}; 0, \sigma_\rho^2). \] \[ (14) \]

We assume that $\rho_i^{(t)}$ is independent of $\rho_{i'}^{(t)}$ for any $i \neq i'$ or $t \neq t'$. 
We now introduce a Markov probability transition model to help with the successive user tracking problem. Following \cite{30} \cite{31}, we model the difference between the positions at any two adjacent time slots as an independent Gaussian noise, i.e.,

$$p^{(t)} = p^{(t-1)} + q^{(t-1)}, \tag{15}$$

where $q^{(t-1)} \sim N(0, C_q)$ is the Gaussian transition noise. Thus, the conditional probability of $p^{(t)}$ given $p^{(t-1)}$ is given by

$$p(p^{(t)} | p^{(t-1)}) = N(p^{(t)}; p^{(t-1)}, C_q). \tag{16}$$

We assume that $q^{(t)}$ and $q^{(t')} \sim N(0, C_q)$ are independent for any $t \neq t'$, i.e., the user positions $\{p^{(t)}\}$ form a Markov chain satisfying

$$p(p^{(t)} | p^{(t-1)}, y^{(t-1)}) = p(p^{(t)} | p^{(t-1)}), \tag{17}$$

where $p^{(t)} = \left[p^{(t)}_u \ldots p^{(t)}_u \right]^T$ is the collection of user positions from time slot $t$ to time slot $t'$. Similar notation applies to $y^{(1:t)}$, $\theta^{(1:t)}$ and $\rho^{(1:t)}$ for the received signals, the AoAs and the equivalent path gains.

Based on the above discussions, the joint pdf of $p^{(0:t)}_u$, $\theta^{(1:t)}$, $\rho^{(1:t)}$ and $y^{(1:t)}$ is given by

$$p(p^{(0:t)}_u, \theta^{(1:t)}, \rho^{(1:t)}, y^{(1:t)}) = \prod_{t=1}^T p(p^{(t)}_u | p^{(t-1)}_u, y^{(1:t-1)}, \theta^{(1:t-1)}, \rho^{(1:t-1)}) \cdot p(\theta^{(1:t)} | \theta^{(0:t-1)}) \cdot p(\rho^{(1:t)} | \rho^{(0:t-1)}) \cdot p(y^{(1:t)} | \rho^{(1:t)}), \tag{18}$$

where $p(\theta^{(1:t)} | \theta^{(0:t-1)}) = \prod_{t=1}^T p(\theta^{(t)} | \theta^{(t-1)})$, $p(\rho^{(1:t)} | \rho^{(0:t-1)}) = \prod_{t=1}^T p(\rho^{(t)} | \rho^{(t-1)})$, and $p(p^{(0:t)}_u)$ is the user position distribution at the initial time slot. Then, our localization and tracking problem is to estimate $p^{(t)}_u$ in each time slot $t$ given the historical received signal $y^{(1:t)}$ and the initial position $p^{(0)}_u$. With the joint pdf given in \cite{18}, following the Bayes' theorem, the posterior distribution of user position $p^{(t)}_u$ is given by

$$p(p^{(t)}_u | y^{(1:t)}, p^{(0:t)}_u) = \frac{p(p^{(0:t)}_u, \theta^{(1:t)}, \rho^{(1:t)}, y^{(1:t)})}{p(y^{(1:t)} | \rho^{(1:t)})} \cdot p(\theta^{(1:t)} | \theta^{(0:t-1)}) dp^{(0:t)}_u \cdot p(\theta^{(1:t)}, \rho^{(1:t)}, y^{(1:t)}) \cdot p(\theta^{(0:t-1)}) dp^{(0:t-1)} \cdot p(\rho^{(1:t)} | \rho^{(0:t-1)}) dp^{(0:t-1)}, \tag{19}$$

Then, an online estimate of $p^{(t)}_u$ can be obtained by following the minimum mean-square error (MMSE) or maximum a posteriori (MAP) principle. However, exact posterior estimation is generally intractable due to prohibitively high computational complexity caused by the integral in \cite{19}. Thus, we resort to a low-complexity solution by following the message-passing principle, as detailed in the subsequent section.

### III. BAYESIAN USER LOCALIZATION AND TRACKING ALGORITHM

#### A. Factor Graph Representation

In this section, we introduce our BULT algorithm based on the message passing principle. The factor graph corresponding to \cite{18} is constructed as shown in Fig. 2, where each blank circle represents a variable node, and each black rectangle represents a factor node. A variable node is connected to a factor node if the factor contains the variable. The variable node $y^{(t)}$ is omitted since $y^{(t)}$ is observed. In the factor graph, the position tracking module and AoA estimation module are introduced to illustrate the BULT algorithm. Specifically, the position tracking module includes the probability factors for the Markov chain of $\{p^{(t)}_u\}$; the AoA estimation module includes the probability factors for the receive signal $y^{(t)}$ and the user AoAs $\{\theta^{(t)}_u\}$. These two modules are connected by factor node $p(\theta^{(t)}_u | p^{(t)}_u)$ and variable node $\theta^{(t)}_u$ according to the geometric constraint. For notational brevity, we represent the factor node $p(\theta^{(t)}_u | p^{(t)}_u)$ by $\varphi^{(t)}_i$ and the factor node $p(p^{(t)}_u | p^{(t-1)}_u)$ by $\psi^{(t)}_i$ in the subscript of a message. Denote by $\Delta_{a \rightarrow b}(\cdot)$ the message from node $a$ to $b$, and by $m_{a \rightarrow b}$ and $C_{a \rightarrow b}$ the mean vector and the covariance matrix of message $\Delta_{a \rightarrow b}(\cdot)$, respectively. Denote by $\Delta_a(\cdot)$ the message of variable node $a$.

#### B. AoA Estimation Module

The AoA estimation module is designed to provide the estimations of user AoAs $\{\theta^{(t)}_u\}$ which are further utilized in the position tracking module. The messages in the AoA estimation module are described below.

1) Messages from $\theta^{(t)}_{U,P}$ to $\varphi^{(t)}_{i}$: Before the further discussion on the message calculation, we first give the pdf of a Von Mises (VM) distribution

$$M(\theta; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp(\kappa \cos(\theta - \mu)), \tag{20}$$

where $I_0$ denotes the modified Bessel function of the first kind in order $0$, $\mu$ and $\kappa$ are the mean direction and concentration parameters respectively.

For $\forall t$, $1 \leq i \leq K$, following the sum-product rule, the message from variable node $\theta^{(t)}_u$ to factor node $\varphi^{(t)}_i$ is given by

$$\Delta_{\theta^{(t)}_{U,i} \rightarrow \varphi^{(t)}_{i}} (\theta^{(t)}_u)$$

Fig. 2: Factor graph representation of \cite{18} for an example in the time slot $t$ and $K = 4$. This factor graph is divided into two modules: the position tracking module and the AoA estimation module.
\[ \alpha \int_{\theta_i^{(t)}} \int_{\rho^{(t)}} p(y^{(t)}|\theta^{(t)}, \rho^{(t)}) p(\rho^{(t)}) \prod_{j \neq i} \Delta_{\psi_i^{(t)} \rightarrow \psi_j^{(t)}} (\theta_i^{(t)}), \]

where \( \theta_i^{(t)} \) denotes the set includes all the entries in \( \theta^{(t)} \) except the \( i \)-th one, the message \( \Delta_{\psi_i^{(t)} \rightarrow \psi_j^{(t)}} (\theta_i^{(t)}) \) can be treated as a prior knowledge of \( \theta_i^{(t)} \) and is approximated by a VM distribution as discussed in next subsection. We further express \( \Delta_{\theta_i^{(t)} \rightarrow \psi_j^{(t)}} (\theta_i^{(t)}) \) as

\[ \Delta_{\theta_i^{(t)} \rightarrow \psi_j^{(t)}} (\theta_i^{(t)}) \approx \frac{p(\theta_i^{(t)}|y^{(t)})}{\Delta_{\psi_j^{(t)} \rightarrow \theta_i^{(t)}} (\theta_i^{(t)})}, \]

where the integral in (22a) is approximated by the posterior estimation of \( \theta_i^{(t)} \) as \( p(\theta_i^{(t)}|y^{(t)}) \). We resort to the variational Bayesian method to obtain \( p(\theta_i^{(t)}|y^{(t)}) \). By modeling the pdf of each AoA as a VM distribution, the VALSE algorithm is developed in [29] by following the variational inference principle to obtain the approximated posterior estimates of the AoAs \( \{l_i^{(t)}\} \) and the equivalent path gains \( \{\rho_i^{(t)}\} \). Also, the model parameters \( \sigma_\theta^2 \) and \( \sigma_\rho^2 \) are iteratively estimated by the expectation maximization method. We denote the mentioned VM distributions as

\[ p(\theta_i^{(t)}|y^{(t)}) = \mathcal{M}\left(\pi \theta_i^{(t)}, \mu_{\theta_i^{(t)}}, \kappa_{\theta_i^{(t)}}\right), \]

\[ \Delta_{\theta_i^{(t)} \rightarrow \psi_j^{(t)}} (\theta_i^{(t)}) = \mathcal{M}\left(\pi \theta_i^{(t)}, \mu_{\psi_j^{(t)} \rightarrow \theta_i^{(t)}}, \kappa_{\psi_j^{(t)} \rightarrow \theta_i^{(t)}}\right), \]

where the variable \( \theta_i^{(t)} \) is multiplied by a constant \( \pi \) to obey the VM distribution. From [32], the family of the VM distribution is closed under multiplication. Then, the message \( \Delta_{\theta_i^{(t)} \rightarrow \psi_j^{(t)}} (\theta_i^{(t)}) \) is further approximated by

\[ \Delta_{\theta_i^{(t)} \rightarrow \psi_j^{(t)}} (\theta_i^{(t)}) \approx \mathcal{M}\left(\pi \theta_i^{(t)}, \mu_{\psi_j^{(t)} \rightarrow \theta_i^{(t)}}, \kappa_{\psi_j^{(t)} \rightarrow \theta_i^{(t)}}\right), \]

where \( \mu_{\theta_i^{(t)} \rightarrow \psi_j^{(t)}} \) and \( \kappa_{\theta_i^{(t)} \rightarrow \psi_j^{(t)}} \) have a relationship of

\[ \kappa_{\theta_i^{(t)} \rightarrow \psi_j^{(t)}} \exp \left(j \mu_{\theta_i^{(t)} \rightarrow \psi_j^{(t)}} \right) = \kappa_{\psi_j^{(t)} \rightarrow \theta_i^{(t)}} \exp \left(-j \mu_{\psi_j^{(t)} \rightarrow \theta_i^{(t)}} \right). \]

The calculation of message \( \Delta_{\theta_i^{(t)} \rightarrow \psi_j^{(t)}} \) in (25) can be regarded as the extrinsic message calculation of the AoA estimation module; see more discussions on extrinsic messages, e.g., in [33].

C. Position Tracking Module

The position tracking module aims to estimate the user position \( \mathbf{p}_U^{(t)} \) based on the estimated AoAs. The messages in

the position tracking module are described below.

1) Messages from \( \phi_i^{(t)} \) to \( \mathbf{p}_U^{(t)} \): For \( t, 1 \leq i \leq K \), the message from factor node \( \phi_i^{(t)} \) to variable node \( \mathbf{p}_U^{(t)} \) is given by

\[ \Delta_{\phi_i^{(t)} \rightarrow \mathbf{p}_U^{(t)}} \propto \int \Delta_{\phi_i^{(t)} \rightarrow \mathbf{p}_U^{(t)}} \Delta_{\theta_i^{(t)} \rightarrow \phi_i^{(t)}} (\theta_i^{(t)}), \]

where \( \Delta_{\theta_i^{(t)} \rightarrow \phi_i^{(t)}} (\theta_i^{(t)}) \) is approximated as the VM distribution given in (24). Therefore, (26) can be expressed as

\[ \Delta_{\phi_i^{(t)} \rightarrow \mathbf{p}_U^{(t)}} \propto \exp \left( \kappa_{\theta_i^{(t)} \rightarrow \phi_i^{(t)}} \cos \left( \frac{\pi}{\| \mathbf{p}_{R,i} - \mathbf{p}_U^{(t)} \|} - \mu_{\theta_i^{(t)} \rightarrow \phi_i^{(t)}} \right) \right). \]

2) Messages along the Markov chain \( \{\phi_i^{(t)}\} \): For \( t \), the message from variable node \( \phi_i^{(t)} \) to factor node \( \phi_i^{(t+1)} \) is given by

\[ \Delta_{\phi_i^{(t)} \rightarrow \phi_i^{(t+1)}} (\phi_i^{(t)}) \propto \Delta_{\phi_i^{(t)} \rightarrow \phi_i^{(t+1)}} \Delta_{\phi_i^{(t)} \rightarrow \phi_i^{(t)}} \mathcal{G}^{(t)} (\phi_i^{(t)}), \]

where

\[ \mathcal{G}^{(t)} (\phi_i^{(t)}) = \prod_{j=1}^K \Delta_{\phi_j^{(t)} \rightarrow \phi_i^{(t)}} (\phi_j^{(t)}). \]

Based on the central limit theorem [34], the message \( \mathcal{G}^{(t)} (\phi_i^{(t)}) \) can be approximated as a Gaussian distribution as

\[ \mathcal{G}^{(t)} (\phi_i^{(t)}) = \mathcal{N}(\phi_i^{(t)}; m_{\phi_i^{(t)}}, C_{\phi_i^{(t)}}), \]

where the mean vector \( m_{\phi_i^{(t)}} \) and the covariance matrix \( C_{\phi_i^{(t)}} \) are obtained by using the gradient descent method (GDM) and the Taylor series expansion. The detailed derivation can be found in Appendix A.

For \( t \), the message from factor node \( \phi_i^{(t)} \) to variable node \( \mathbf{p}_U^{(t)} \) is given by

\[ \Delta_{\phi_i^{(t)} \rightarrow \mathbf{p}_U^{(t)}} (\mathbf{p}_U^{(t)}) \propto \int_{\mathbf{p}_U^{(t)}} p_{\mathbf{p}_U^{(t+1)} \rightarrow \phi_i^{(t)}} (\mathbf{p}_U^{(t+1)}) p(\mathbf{p}_U^{(t)}) |\mathbf{p}_U^{(t+1)}|. \]

We note that the factor nodes \( \{\phi_i^{(t)}\} \) are in the form of Gaussian distribution as given in (16). By assuming that a Gaussian message is provided in the initial time slot, all the messages passing along the Markov chain are Gaussian since the product of two Gaussian messages is still Gaussian. Therefore, the mean vector and covariance matrix of message \( \Delta_{\phi_i^{(t)} \rightarrow \phi_i^{(t+1)}} (\mathbf{p}_U^{(t)}) \) are respectively given by

\[ m_{\mathbf{p}_U^{(t)} \rightarrow \phi_i^{(t+1)}} = C_{\mathbf{p}_U^{(t)} \rightarrow \phi_i^{(t+1)}} \left( C_{\mathbf{p}_U^{(t)} \rightarrow \phi_i^{(t+1)}}^{-1} m_{\phi_i^{(t)} \rightarrow \phi_i^{(t+1)}} + C_{\phi_i^{(t)} \rightarrow \phi_i^{(t+1)}}^{-1} m_{\phi_i^{(t)} \rightarrow \phi_i^{(t+1)}} \right), \]

\[ C_{\mathbf{p}_U^{(t)} \rightarrow \phi_i^{(t+1)}} = C_{\mathbf{p}_U^{(t)} \rightarrow \phi_i^{(t+1)}}^{-1} + C_{\phi_i^{(t)} \rightarrow \phi_i^{(t+1)}}^{-1}, \]

where according to (16) and (31) we have

\[ m_{\phi_i^{(t)} \rightarrow \mathbf{p}_U^{(t)}} = m_{\mathbf{p}_U^{(t)} \rightarrow \phi_i^{(t)}}, \]

\[ C_{\phi_i^{(t)} \rightarrow \mathbf{p}_U^{(t)}} = C_{\mathbf{p}_U^{(t)} \rightarrow \phi_i^{(t)}} + C_q, \]
Based on (32) and (33), the mean vectors and covariance matrices of the Gaussian messages along the Markov chain can be recursively calculated.

3) Messages from \( \mathbf{p}_U^{(t)} \) to \( \varphi_i^{(t)} \): For \( \forall t, 1 \leq i \leq K \), the message from variable node \( \mathbf{p}_U^{(t)} \) to factor node \( \varphi_i^{(t)} \) is given by

\[
\Delta \psi_i^{(t)} - \theta_i^{(t)} = \mathbb{E}(\psi_i^{(t)} | \theta_i^{(t)}) - \theta_i^{(t)} \tag{34}
\]

where

\[
\mathbb{E}(\psi_i^{(t)} | \theta_i^{(t)}) = \sum_{j=1}^{K} \Delta \psi_j^{(t)} - \theta_j^{(t)} \mathbb{E}(\psi_j^{(t)} | \theta_j^{(t)}) \tag{35}
\]

differs from (29) for only one term and can be approximated as a Gaussian distribution with mean vector \( \mathbf{m}_{\psi_i^{(t)}} \) and the covariance matrix \( \mathbf{C}_{\psi_i^{(t)}} \) similarly. Therefore, \( \Delta \psi_i^{(t)} - \theta_i^{(t)} \) is also a Gaussian message with its mean vector and covariance matrix calculated by

\[
\mathbf{m}_{\psi_i^{(t)}} - \varphi_i^{(t)} = \mathbf{C}_{\psi_i^{(t)}} - \varphi_i^{(t)} \left( \mathbf{C}^{-1} \mathbf{G}^{(t)} \mathbf{C}_{\psi_i^{(t)}} + \mathbf{C}^{-1} \mathbf{p}_U^{(t)} \mathbf{C}_{\psi_i^{(t)}} \mathbf{p}_U^{(t)} \right) \tag{36a}
\]

\[
\mathbf{C}^{-1} \mathbf{p}_U^{(t)} - \varphi_i^{(t)} = \mathbf{C}^{-1} \mathbf{G}^{(t)} + \mathbf{C}^{-1} \mathbf{p}_U^{(t)} \mathbf{p}_U^{(t)} \tag{36b}
\]

4) Messages from \( \varphi_i^{(t)} \) to \( \theta_1^{(t)} \): For \( \forall t, 1 \leq i \leq K \), the message from factor node \( \varphi_i^{(t)} \) to variable node \( \theta_1^{(t)} \) is given by

\[
\Delta \varphi_i^{(t)} - \theta_1^{(t)} = \int \mathcal{M}(\theta_1^{(t)} | \varphi_i^{(t)} - \theta_1^{(t)}) \mathbb{P}(\varphi_i^{(t)} | \theta_1^{(t)}) d\theta_1^{(t)} \tag{37}
\]

The integral in (37) does not have a closed-form expression. To facilitate the subsequent message passing, we approximate this message by a VM distribution based on the far-field assumption as

\[
\Delta \varphi_i^{(t)} - \theta_1^{(t)} = \int \mathcal{M}(\theta_1^{(t)} | \varphi_i^{(t)} - \theta_1^{(t)}) \mathbb{P}(\varphi_i^{(t)} | \theta_1^{(t)}) d\theta_1^{(t)} \tag{38}
\]

with

\[
\mu_{\varphi_i^{(t)} - \theta_1^{(t)}} = \pi \bar{\theta}_i^{(t)} \tag{39a}
\]

\[
k_{\varphi_i^{(t)} - \theta_1^{(t)}} = \frac{d_i^2}{\pi^2 \left(1 - (\bar{\theta}_i^{(t)} \cdot \mathbf{v}_i)^2\right) \mathbf{v}_i^T \mathbf{C}_{\varphi_i^{(t)} - \theta_1^{(t)}} \mathbf{v}_i} \tag{39b}
\]

In (39), \( \mathbf{v}_i \) is the unit directional vector perpendicular to \((\mathbf{p}_{R,i} - \mathbf{m}_{\psi_i^{(t)}})\) and is in the plane spanned by \( \mathbf{e}_U \) and \((\mathbf{p}_{R,i} - \mathbf{m}_{\psi_i^{(t)}})\); \( \bar{\theta}_i^{(t)} \) is the mean AoA; \( d_i \) is the Euclidean distance between \( \mathbf{p}_{R,i} \) and \( \mathbf{m}_{\psi_i^{(t)}} \) as shown in Fig. 3. \( \mathbf{v}_i, \bar{\theta}_i^{(t)} \) and \( d_i \) are given respectively by

\[
\mathbf{v}_i = \left( \frac{(\mathbf{m}_{\psi_i^{(t)}} - \mathbf{p}_{R,i}) \times \mathbf{e}_U}{\| (\mathbf{m}_{\psi_i^{(t)}} - \mathbf{p}_{R,i}) \times \mathbf{e}_U \|} \right) \left( \frac{(\mathbf{m}_{\psi_i^{(t)}} - \mathbf{p}_{R,i}) \times \mathbf{e}_U}{\| (\mathbf{m}_{\psi_i^{(t)}} - \mathbf{p}_{R,i}) \times \mathbf{e}_U \|} \right) \tag{40a}
\]

\[
\bar{\theta}_i^{(t)} = \left( \frac{\mathbf{p}_{R,i} - \mathbf{m}_{\psi_i^{(t)}}}{\| \mathbf{p}_{R,i} - \mathbf{m}_{\psi_i^{(t)}} \|} \right)^T \mathbf{e}_U \tag{40b}
\]

\[
d_i = \| \mathbf{p}_{R,i} - \mathbf{m}_{\psi_i^{(t)}} \| \tag{40c}
\]

where \( \times \) denotes the cross product. The derivation of (39) can be found in Appendix B.

D. Output Estimations at Time Slot \( t \)

The mean of the Gaussian message \( \Delta \varphi_i^{(t)} - \theta_1^{(t)} \) in (28), denoted by \( \mathbf{m}_{\psi_i^{(t)}} - \varphi_i^{(t)} \), is used as an output estimate of user position \( \mathbf{p}_U^{(t)} \). For the estimation of the user AoAs, we express the message at variable node \( \theta_i^{(t)} \) as

\[
\Delta \varphi_i^{(t)} - \theta_i^{(t)} \propto \Delta \varphi_i^{(t)} - \theta_i^{(t)} \mathbb{P}(\varphi_i^{(t)} | \theta_i^{(t)}) \mathbb{P}(\theta_i^{(t)} | \theta_1^{(t)}) \Delta \varphi_i^{(t)} - \theta_i^{(t)} \mathbb{P}(\theta_i^{(t)} | \theta_1^{(t)}) \tag{41}
\]

From (22)-(24), \( \Delta \varphi_i^{(t)} - \theta_i^{(t)} \) is approximated as a VM distribution. We obtain an output estimate of \( \theta_i^{(t)} \) as \( \mu_{\theta_i^{(t)}} / \pi \), where \( \mu_{\theta_i^{(t)}} \) is the mean direction of \( \theta_i^{(t)} \).

E. Overall Algorithm

Based on the discussions in the preceding subsections, our BULT algorithm is summarized in Algorithm 1. Algorithm 1 is an online algorithm that estimates the user position \( \mathbf{p}_U^{(t)} \) and AoAs \( \theta_i^{(t)} \) jointly in each time slot by iteratively passing messages between the position tracking module and the AoA estimation module. Specifically, lines 5-6 of Algorithm 1 correspond to the calculation of message passing from the position tracking module to the AoA estimation module. Lines 7-8 correspond to the calculation of message passing from the the AoA estimation module to the position tracking module. Line 2 and line 11 correspond to the calculation of message passing along the Markov chain. Note that the factor graph in Fig. 2 is loopy, so the convergence of the BULT algorithm cannot be guaranteed. As inspired by [35], we apply the damping technique on the messages \( \{\Delta \psi_i^{(t)} - \theta_1^{(t)}, \mu_{\psi_i^{(t)}} - \theta_1^{(t)}\} \) to improve the convergence.
Algorithm 1 BULT Algorithm

Input: Observed signal $y^{(1)}, \ldots, y^{(t)}$, initial user position estimation $\mathbf{r}_0^{(i)} \rightarrow \theta^{(i)}(0)$ and with its covariance matrix $\mathbf{C}_0^{(i)} \rightarrow \theta^{(i)}(0)$. 
Output: User position estimations $\{\mathbf{m}_i^{(t)}(t)\rightarrow \theta_i^{(t)}(t)\}$, user AoA estimations $\{\mu_{\theta_i^{(t)}}(t)/\pi\}$, and equivalent path gain estimations $\{\hat{\rho}_i^{(t)}\}$ for $1 \leq i \leq K$, $1 \leq t \leq t_0$. 
1: for $t = 1$ to $t_0$ do 
2: Calculate $\mathbf{m}_i^{(t)}(t) \rightarrow \theta_i^{(t)}(t)$ and $\mathbf{C}_i^{(t)} \rightarrow \theta_i^{(t)}(t)$ by (33).
3: Initialization: for all $i$, $\mathbf{m}_{\theta_i^{(t)}}(t) = \mathbf{m}_i^{(t)}(t) \rightarrow \theta_i^{(t)}(t)$, $\mathbf{C}_{\theta_i^{(t)}}(t) = \mathbf{C}_i^{(t)} \rightarrow \theta_i^{(t)}(t)$. 
4: repeat 
5: For all $i$, update the message $\Delta_{\mathbf{r}_i^{(t)}(t) \rightarrow \theta_i^{(t)}(t)}$ by (36). 
6: For all $i$, update the message $\Delta_{\mathbf{r}_i^{(t)}(t) \rightarrow \theta_i^{(t)}(t)}$ by (38). 
7: For all $i$, update the posterior pdf estimation $P(\theta_i^{(t+1)}(t)|Y^{(t)})$ and equivalent path gain estimation $\hat{\rho}_i^{(t)}$ by the VALSE algorithm. 
8: For all $i$, update the message $\Delta_{\mathbf{r}_i^{(t)}(t) \rightarrow \theta_i^{(t)}(t)}$ by (24). 
9: For all $i$, update $\mathbf{m}_{\theta_i^{(t)}}(t)$ and $\mathbf{C}_{\theta_i^{(t)}}(t)$ by (35). 
10: until stopping criterion 
11: Calculate $\mathbf{m}_i^{(t)}(t) \rightarrow \phi_i^{(t)}(t)$ and $\mathbf{C}_i^{(t)} \rightarrow \phi_i^{(t)}(t)$ by and (32). 
12: end for 
13: return $\{\mathbf{m}_i^{(t)}(t) \rightarrow \theta_i^{(t)}(t)\}$, $\{\mu_{\theta_i^{(t)}}(t)/\pi\}$, and $\{\hat{\rho}_i^{(t)}\}$ for $1 \leq t \leq t_0$, $1 \leq i \leq K$.

As for the initialization in the first time slot $t = 1$, we assume that a coarse user position information in the form of mean $\mathbf{m}_0^{(0)} \rightarrow \theta^{(0)}(0)$ and covariance matrix $\mathbf{C}_0^{(0)} \rightarrow \theta^{(0)}(0)$ is available, such information can be provided, e.g., by the Global Positioning System (GPS). Alternatively, we can run the BULT algorithm without any prior position information (e.g., by setting $\mathbf{m}_0^{(0)} \rightarrow \theta^{(0)}(0) = \mathbf{0}$ and $\mathbf{C}_0^{(0)} \rightarrow \theta^{(0)}(0) = \eta I$ where $\eta$ is a large positive number). In this case, the estimation of $\{\theta_i^{(1)}(1)\}_i=1^K$ given by the AoA estimation module are uncorrelated which causes the mismatch problem between the variable nodes $\{\theta_i^{(1)}(1)\}_i=1^K$ and the factor nodes $\{\phi_i^{(1)}(1)\}_i=1^K$. To obtain a correct match, we exhaustively search all possible matching schemes and select the one with the minimum $\text{Tr}(\mathbf{C}_i^{(1)} \rightarrow \theta_i^{(1)}(1))$.

F. Complexity of BULT

The complexity of the BULT algorithm mainly arises from the variational Bayesian method and the Gaussian message approximation. For any time slot, the complexity of the variational Bayesian method according to (29) is $O(n_1K^3N_1 + n_1K^2N_2)$, where $n_1$ is the number of iterations in the variational Bayesian method. The complexity of Gaussian message approximation is $O(n_2K)$, where $n_2$ is the iteration number of GDM in (34). Therefore, the overall complexity of BULT is given by $O(n_1K^3N_1 + n_3K^2N_2 + n_3K)$, where $n_3$ is the iteration number of message passing between the position tracking module and the AoA estimation module, which depends on the stopping criterion and is empirically less than 15.

IV. BAYESIAN CRAMÉR-RAO BOUND

In this section, we derive the BCRB for the estimation of the user position and AoAs over time. The BCRB acts as a benchmark of our tracking algorithm. Considering the geometric constraint in (7), we take a parameter set in the time slot $t$ as $\gamma^{(t)} = [(\mathbf{p}_U^{(t)})^T, (\varphi_i^{(t)})^T, (\rho_i^{(t)})^T]^T \in \mathbb{R}^{2K+1}$, where the $i$-th term of $\varphi_i^{(t)}$ and $\rho_i^{(t)}$ are $\varphi_i^{(t)}$ and $\rho_i^{(t)}$, respectively represent the angle and amplitude of the equivalent complex path gain $\rho_i^{(t)}$. In the sequel, we first derive the Fisher information matrix (FIM) in a single time slot by neglecting the probability model and regarding $\gamma^{(t)}$ as a deterministic parameter set. We then compute the BCRB based on the derived FIM.

A. FIM Calculation for a Single Time Slot

Consider the problem of estimating a deterministic parameter set $\gamma^{(t)}$ in time slot $t$ without any prior information from time slot $t-1$. The MSE of the unbiased estimator is lower bounded by the Cramér Rao bound (CRB) which is the diagonal of the inverse of FIM $\mathbf{J}^{(t)} \in \mathbb{R}^{(2K+1) \times (2K+1)}$. For the signal model in (10b), the $(i, j)$-th entry of $\mathbf{J}^{(t)}$ can be calculated according to the following lemma (36):

**Lemma 1.** For an $N$-dimensional complex Gaussian noise disturbed observation signal $\mathbf{y}^{(t)} \sim \mathcal{CN}(\mathbf{y}^{(t)}; \mu^{(t)}, \Sigma^{(t)})$, the $(i, j)$-th entry of the FIM is given by

$$
\mathbf{J}^{(t)}_{i,j} = 2R \left[ \frac{\partial (\mu_i^{(t)})^H}{\partial y_i^{(t)}} \Sigma_i^{(t)^{-1}} \frac{\partial (\mu_j^{(t)})}{\partial y_j^{(t)}} + \text{Tr} \left( \Sigma_i^{(t)^{-1}} \frac{\partial (\Sigma_i^{(t)})}{\partial y_j^{(t)}} \Sigma_j^{(t)^{-1}} \frac{\partial (\Sigma_j^{(t)})}{\partial y_i^{(t)}} \right) \right].
$$

(42)

In our scenario, we have $\mu_i^{(t)} = \sum_{i=1}^{K} \mu_i^{(t)}(\theta_i^{(t)}), \mu_j^{(t)} = \sum_{i=1}^{K} \mu_j^{(t)}(\theta_j^{(t)})$. The FIM of $\gamma^{(t)}$ is given as

$$
\mathbf{J}^{(t)} = 2 \sum_{i=1}^{K} \mathbb{E} \left[ \frac{\partial (\mu_i^{(t)})}{\partial y_i^{(t)}} \frac{\partial (\mu_j^{(t)})}{\partial y_j^{(t)}} + \frac{\partial I(\mu_i^{(t)})}{\partial y_i^{(t)}} \frac{\partial I(\mu_j^{(t)})}{\partial y_j^{(t)}} \right],
$$

(43)

where $\mu_i^{(t)}$ is the $n$-th entry of $\mu^{(t)}$ given by

$$
\mu_i^{(t)} = \sum_{i=1}^{K} \left| \rho_i^{(t)} \right| e^{j(\pi(n-1)\theta_i^{(t)} + \gamma_i^{(t)})}.
$$

(44)

For $1 \leq n \leq N_U$ and $1 \leq i \leq K$, we have

$$
\frac{\partial \mathbb{E} \left[ \mu_i^{(t)} \right]}{\partial \mathbf{p}_U^{(t)}} = \sum_{i=1}^{K} \left| \rho_i^{(t)} \right| (\varphi_i^{(t)} - \varphi_i^{(t,t+1)}) \frac{\partial \theta_i^{(t)}(t)}{\partial \mathbf{p}_U^{(t)}}.
$$

(45a)

$$
\frac{\partial \mathbb{E} \left[ \mu_i^{(t)} \right]}{\partial \mathbf{p}_U^{(t)}} = \sum_{i=1}^{K} \left| \rho_i^{(t)} \right| (\varphi_i^{(t)} - \varphi_i^{(t,t+1)}) \frac{\partial \rho_i^{(t)}(t)}{\partial \mathbf{p}_U^{(t)}}.
$$

(45b)
where \( \frac{\partial \theta_{t,i}^{(t)}}{\partial p_{U,i}^{(t)}} \) is calculated according to (44) given by
\[
\frac{\partial \theta_{t,i}^{(t)}}{\partial p_{U,i}^{(t)}} = \frac{-e_{U} + (e_{i}^{(t)})^{T}e_{U,i}^{(t)}}{\left\| p_{R,i} - p_{U,i}^{(t)} \right\|_2},
\]
with \( e_{i}^{(t)} = \frac{(p_{R,i} - p_{U,i}^{(t)})}{\left\| p_{R,i} - p_{U,i}^{(t)} \right\|_2} \). Similarly, we have
\[
\frac{\partial \theta_{t,i}^{(t)}}{\partial \rho_{i}^{(t)}} = -|\rho_{i}^{(t)}| \sin \left( \pi (n - 1) \theta_{U,i}^{(t)} + \rho_{i}^{(t)} \right),
\]
\[
\frac{\partial \theta_{t,i}^{(t)}}{\partial \mu_{i}^{(t)}} = |\rho_{i}^{(t)}| \cos \left( \pi (n - 1) \theta_{U,i}^{(t)} + \rho_{i}^{(t)} \right),
\]
\[
\frac{\partial \theta_{t,i}^{(t)}}{\partial |\rho_{i}^{(t)}|} = \cos \left( \pi (n - 1) \theta_{U,i}^{(t)} + \rho_{i}^{(t)} \right),
\]
\[
\frac{\partial \theta_{t,i}^{(t)}}{\partial \mu_{i}^{(t)}} = \sin \left( \pi (n - 1) \theta_{U,i}^{(t)} + \rho_{i}^{(t)} \right).
\]

By substituting (45)-(47) into (43), we obtain the FIM in the current time slot without considering the prior information from the preceding time slot.

### B. BCRB for Successive Time Slots

The BCRB is derived based on the Bayesian Fisher information matrix (BFIM). For an arbitrary time slot \( t \), the BFIM of \( \gamma^{(t)} \) denoted by \( J_{B}^{(t)} \) is recursively calculated by
\[
J_{B}^{(t)} = J^{(t)} + G_{22}^{(t)} - G_{21}^{(t)}J_{B}^{(t-1)} + G_{11}^{(t-1)}G_{11}^{(t)},
\]
where \( J^{(t)} \) is given by (43) and the else items in (48) are defined as
\[
G_{11}^{(t)} = \mathbb{E} \left[ -\frac{\partial^2 \log p(\gamma^{(t)}|\gamma^{(t-1)})}{\partial \gamma^{(t-1)} \partial \gamma^{(t-1)}^{T}} \right],
\]
\[
G_{12}^{(t)} = \mathbb{E} \left[ -\frac{\partial^2 \log p(\gamma^{(t)}|\gamma^{(t-1)})}{\partial \gamma^{(t-1)} \partial \gamma^{(t-1)}^{T}} \right],
\]
\[
G_{21}^{(t)} = \mathbb{E} \left[ -\frac{\partial^2 \log p(\gamma^{(t)}|\gamma^{(t-1)})}{\partial \gamma^{(t-1)} \partial \gamma^{(t-1)}^{T}} \right],
\]
\[
G_{22}^{(t)} = \mathbb{E} \left[ -\frac{\partial^2 \log p(\gamma^{(t)}|\gamma^{(t-1)})}{\partial \gamma^{(t-1)} \partial \gamma^{(t-1)}^{T}} \right].
\]

Due to the independence of \( \rho^{(t)} \) in different time slots, we obtain \( p(\gamma^{(t)}|\gamma^{(t-1)}) = p_{U}^{(t)}|p_{U}^{(t-1)} \). Further calculation yields \( G_{22}^{(t)} = G_{11}^{(t)} = -G_{12}^{(t)} = G_{21}^{(t)} \in \mathbb{R}^{(2K+3)(2K+3)} \), where
\[
G_{11}^{(t)} = \begin{bmatrix}
-\frac{\partial^2 \log p(\gamma^{(t)}|\gamma^{(t-1)})}{\partial \gamma^{(t-1)} \partial \gamma^{(t-1)}^{T}} & 0 \\
0 & 0
\end{bmatrix},
\]
with \( \mathbb{E} \left[ -\frac{\partial^2 \log p(\gamma^{(t)}}{\partial \gamma^{(t-1)} \partial \gamma^{(t-1)}^{T}} \right] = C_{q}^{-1} \). By initializing \( J_{B}^{(0)} = \mathbb{E} \left[ -\frac{\partial^2 \log p(\gamma^{(0))}}{\partial \gamma^{(0)} \partial \gamma^{(0)}^{T}} \right] \), the BCRB of the estimation of \( \hat{\gamma}^{(t)} \) in the \( t \)-th time slot is given by
\[
\mathbb{E} \left[ (\hat{\gamma}^{(t)} - \gamma^{(t)})(\hat{\gamma}^{(t)} - \gamma^{(t)})^{H} \right] \geq [(J_{B}^{(t)})^{-1}]_{1,i}, \quad \text{for} \quad \forall 1 \leq i \leq 2K+3.
\]

As for the BCRB of AoAs estimation \( \theta^{(t)} \), we exploit the theory of parameter transformation for CRLB [38]. Assuming a function mapping \( \theta^{(t)} = g_{t}(\gamma^{(t)}) \) associates \( \gamma^{(t)} \) and \( \theta^{(t)} \), the FIM of \( \theta^{(t)} \) denoted by \( J_{\theta}^{(t)} \in \mathbb{R}^{K \times K} \) is obtained by
\[
J_{\theta}^{(t)} = T_{\theta}^{(t)}J_{B}^{(t)}(T_{\theta}^{(t)})^{T},
\]
where \( T_{\theta}^{(t)} \in \mathbb{R}^{K \times (2K+3)} \) is the Jacobian matrix of \( g_{t}(\gamma^{(t)}) \) whose entries are obtained by
\[
T_{\theta}^{(t)} = \begin{bmatrix}
\frac{\partial \theta^{(t)}}{\partial (\gamma^{(t)})^{T}} & \frac{\partial \theta^{(t)}}{\partial (\gamma^{(t)})^{T}} & \frac{\partial \theta^{(t)}}{\partial (\gamma^{(t)})^{T}}
\end{bmatrix},
\]
where \( \frac{\partial \theta^{(t)}}{\partial \gamma^{(t)}} \) is given according to (46) and \( \frac{\partial \theta^{(t)}}{\partial \gamma^{(t)}} = 0 \). Similarly, for any \( 1 \leq i \leq K \), the BCRB for estimation of \( \hat{\theta}^{(t)} \) in the \( t \)-th time slot is given by
\[
\mathbb{E} \left[ (\hat{\theta}^{(t)} - \theta^{(t)})(\hat{\theta}^{(t)} - \theta^{(t)})^{H} \right] \geq [(J_{\theta}^{(t)})^{-1}]_{1,i}, \quad \text{for} \quad \forall 1 \leq i \leq K.
\]
The above BCRB is useful in the beamforming design discussed in the next section.

### V. BEAMFORMING DESIGN FOR BS AND RISs

The beamforming (BF) of the BS and the passive beamforming (PBF) of the RISs need to be appropriately designed to assist the user tracking. In this section, we first introduce the beamforming design based on the minimization of the BCRB. We then provide a position-based directional beamforming design as a low-complexity solution.

#### A. BCRB Based Beamforming Design

In Section [IV], we have derived the BCRB of the parameter estimation problem, which provides a metric for evaluating the tracking performance. By optimizing the BF and PBF jointly, we minimize the BCRB of the user position in the considered tracking problem corresponding to the first three diagonal elements of \( (J_{B}^{(t)})^{-1} \). Since \( \gamma^{(t)} \) is unknown, the user position \( p_{U} \) and equivalent complex path gain \( \rho^{(t)} \) are approximated respectively by the estimated user position \( m_{p_{c}}^{(t-1)} \rightarrow \tilde{\theta}_{i}^{(t)} \) and the vector valued function \( \tilde{g}^{(t-1)}(\omega^{(t)}, f^{(t)}) \), where \( \omega^{(t)} \in \mathbb{R}^{K \times N_{k}} \) collects the PBF vector of RISs as \( \omega^{(t)} = [\omega_{1}^{(t)}, \ldots, \omega_{K}^{(t)}]^{T} \). Referring to (11), the \( i \)-th term of function \( \tilde{g}^{(t-1)}(\omega^{(t)}, f^{(t)}) \) is expressed as
\[
\tilde{g}_{i}^{(t-1)}(\omega_{i}^{(t)}, f^{(t)}) = f_{i}^{(t-1)}a_{R}^{H}(\hat{\theta}_{R,i}^{(t)})^{H} \left( a_{R}(\hat{\theta}_{R,i}^{(t)})a_{R}^{H}(\hat{\theta}_{R,i}^{(t)}) \right)^{T},
\]
where \( \tilde{\theta}_{i}^{(t-1)} \) is the average AoD for the \( i \)-th RIS at the \( (t-1) \)-th time slot, i.e.,
\[
\tilde{\theta}_{R,i}^{(t-1)} = \begin{bmatrix}
(m_{p_{c}}^{(t-1)} \rightarrow \tilde{\theta}_{i}^{(t)} - p_{R,i})^{T} e_{R,i}
\end{bmatrix},
\]
with $\mathbf{e}_{R,i}$ being the known direction vector of the $i$-th RIS, and

$$
\hat{\rho}^{(t-1)}_{UB,i} = \hat{\rho}^{(t-1)}_i \mathbf{a}_R^H(\hat{\theta}_{R,i}) \mathbf{a}_H^R(\hat{\phi}_{B,i}) \mathbf{a}_B(\hat{\theta}_{B,i}),
$$

(57)

where $\hat{\rho}^{(t-1)}_i$ is the estimated equivalent complex path gain, $\hat{\theta}_{R,i}$ and $\hat{\phi}_{B,i}$ are the beamforming vectors in time slot $t-1$. Therefore, the optimization problem in time slot $t$ is formulated as

$$(P1) : \min_{\omega_i^{(t)},f^{(t)}} \sum_{j=1}^{3} \left( J_B \left( \mathbf{p}_{U}(\omega_i^{(t)}), \mathbf{f}^{(t)}(\omega_i^{(t)}), f^{(t)} \right) \right)_{j,j},$$

\text{s.t.} \quad ||\omega_i^{(t)}|| = 1, \text{ for } 1 \leq i \leq K, 1 \leq j \leq N_U, \quad ||f^{(t)}|| = 1.

The optimal solution of (P1) is difficult to be found due to the non-convexity of the objective function and the constraints. We propose to use the alternating gradient descent method (AGDM) which alternately optimizes the PBF of the RISs and the BF of the BS to obtain a locally optimal solution.

### B. Directional Beamforming Design

The complexity of the beamforming design in (57) is high due to the high-dimension of the optimization variables and the iterative nature of the design algorithm. We next give the position based directional BF and PBF design as a low-complexity solution.

For the passive beamforming design, the PBF for the $i$-th RIS in time slot $t$ is given by

$$\omega_i^{(t)} = \mathbf{a}_R(\hat{\theta}_{R,i}) \odot \mathbf{a}_H^R(\hat{\phi}_{B,i}),$$

(58)

which aligns the reflect beam of the $i$-th RIS to the estimation of user position $\mathbf{m}_{R,i+1}$. Thus we obtain the closed-form solution of PBF. Similar RIS PBF designs can be found in [39], [40].

For the beamforming design of the BS, as the positions of the RISs are fixed and known, a concise and effective BS beamforming design is given as

$$f^{(t)} = \sum_{i=1}^{K} w_i^{(t)} \mathbf{a}_B(\hat{\theta}_{B,i}),$$

(59)

where $w_i^{(t)}$ is the weight of the $i$-th directional beam obtained by solving the following optimization problem:

$$(P2) : \min_{\mathbf{w}^{(t)}} \sum_{j=1}^{3} \left( J_B \left( \mathbf{p}_U^{(t)}, \mathbf{f}^{(t)}(\omega_i^{(t)}), f^{(t)} \right) \right)_{j,j},$$

\text{s.t.} \quad \left\| \sum_{i=1}^{K} w_i^{(t)} \mathbf{a}_B(\hat{\theta}_{B,i}) \right\|_2 = 1,

where $\mathbf{w}^{(t)} = [w_1^{(t)}, \ldots, w_K^{(t)}]^T$. Here we also use the GDM to obtain a locally optimal solution of problem (P2). In the angle domain, the beamforming design in (59) provides multiple beam lobes with each aligned to a different RIS.

## VI. Numerical Experiments

In this section, we conduct numerical experiments to demonstrate the performance of the BULT algorithm and the beamforming design in different scenarios.

### A. System Parameters and Performance Metric

We deploy the RISs and the BS appropriately to ensure the far-field assumption used in (10a) holds. The parameter $\rho^{(t)}_{UB,i}$ in the equivalent complex path gains is generated by following [26], [41] as

$$P^{(t)}_{UB,i} = -\frac{\lambda}{4\pi} e^{-\frac{\alpha R}{2}} \left( d_{B,R,i} + d_{R,U,i} \right),$$

(60)

where $d_{B,R,i}$ and $d_{R,U,i}$ are the distances from the BS to the $i$-th RIS and the $i$-th RIS to the user respectively. We use the directional beamforming design for $\omega^{(t)}$ and $f^{(t)}$, unless otherwise specified. The system parameter settings are listed in Table I.

We assume that the tracking interval is 20 ms, which makes it possible to track the user in a high-speed mobile scenario. The user is bounded in a cuboid area, and the length, width and height of which are 30 meters, 30 meters and 6 meters, respectively. We set the initial user position as (-10, 0, 0). The user’s trajectory is generated based on the conditional probability $p(p_U^{(t)} | p_U^{(t-1)}) = N(p_U^{(t)}, \sigma^2_U)$, where $p_U$ is set as diag([0.03, 0.03, 0.01]T). Therefore, the average movement distance of user between adjacent time slots is given by $E \left[ \left\| p_U^{(t)} - p_U^{(t-1)} \right\|^2 \right] = \sqrt{0.07}$ m, which corresponds to a velocity of 48 Km/h considering the 20ms time interval. The total number of tracking time slots is set as $T = 300$.

To evaluate the performance of the proposed BULT algorithm, we define the root mean square error (RMSE) of the estimated user position and AoAs respectively as

$$\text{RMSE} \left( \mathbf{p}_U \right) = \sqrt{\frac{\sum_{t=1}^{T} \left\| \mathbf{p}_U^{(t)} - \mathbf{p}_U^{(t-1)} \right\|^2}{T}}, \quad (61a)$$

$$\text{RMSE} \left( \theta_{U,i} \right) = \sqrt{\frac{\sum_{t=1}^{T} \sum_{j=1}^{K} \left\| \theta_{U,i}^{(t)} - \theta_{U,i}^{(t-1)} \right\|^2}{TK}}. \quad (61b)$$

### TABLE I

| Parameter | Value |
|-----------|-------|
| $f_c$     | 28 GHz |
| $\sigma^2_R$ | -84 dBm |
| $N_B$     | 32    |
| $N_R$     | 32, 64, 96 |
| $N_U$     | 17    |
| $p_B$     | (20 m, 0 m, 0 m) |
| $p_{R,i}$ | (-35 m, 5 m, -10 m), (-30 m, 20 m, 10 m), (-20 m, 25 m, 20 m), (-10 m, 40 m, 10 m), (0 m, 20 m, 10 m), (10 m, 15 m, 20 m), (30 m, 20 m, 5 m) |
| $\mathbf{e}_{R,i}$ | (0, 1, 0), (1, 0, 0), (1, 0, 0), (1, 0, 0), (1, 0, 0), (0, 1, 0) |
| $\mathbf{e}_U$ | (1, 0, 0) |
Fig. 4. Evaluation of the VM approximation in (38), where the left peak and the right peak in the distribution correspond to scenario 1 and scenario 2 respectively.

Fig. 5. The user tracking and AoA estimation performance v.s. the transmission power with a varying number of RIS elements.

The expectations in (61) are numerically approximated by averaging from 30 independently generated trajectories.

**B. Message Approximation in (38)**

We conduct numerical experiments to compare the distribution given by (37) and its VM approximation given by (38). The covariance matrix of Gaussian message $p(\phi_{U,i} P_U)\Delta_{P_U}^{-\psi_{i}(t)}(P_U(t))$ is set by $C_{\psi_{i}(t)}^{-\psi_{i}(t)} = \text{diag}([2, 2, 2])$. We consider two scenarios: (1) the distance $d = 40$ m and $\tilde{d}_{U,i} = \frac{\sqrt{3}}{\lambda}$; (2) the distance $d = 20$ m and $\tilde{d}_{U,i} = \frac{\sqrt{3}}{\lambda}$. As the results show in Fig. 4, the approximation performs well when the far-field assumption is satisfied.

**C. Tracking and Estimation Performance**

In this subsection, we first evaluate the performance of the proposed BULT algorithm in different system parameter settings. We then compare the BULT algorithm with a straightforward tracking method that acts as a baseline scheme. Finally, we perform experiments to compare the tracking performance under different beamforming designs given in Section V.

1) **Tracking and AoAs Estimation**: We study the impact of the number of RIS elements and the number of RISs on the user tracking performance in (61) with varying transmission power. The results are shown in Fig. 5 and Fig. 6. For both the performance of user tracking and AoAs estimation, the increase of the RIS elements leads to performance improvement since the passive beamforming gain of the RISs is proportional to the number of the RIS elements. To gain insights into the relationship between the RIS number and the estimation error, we remove the RIS located in (-35, 5, -10) from the system with $K = 6$ and remove the RISs located in (-35, 5, -10) and (10, 15, 20) for $K = 5$. As given in Fig. 6, the additional RISs in system brings better performance in user tracking while the AoAs estimation performances are almost the same. The additional RISs provide more information for user position which improves the user tracking accuracy. However, it deteriorates the AoAs estimation performance slightly since more angle parameters are involved and need to be estimated. In Fig. 5 and Fig. 6, it is observed that both the user tracking errors and the AoAs estimation errors are very close to the BCRB at relatively high SNR, which validates the effectiveness of the proposed BULT algorithm.

2) **Comparison with Benchmark**: For the benchmark scheme, we consider the atomic norm based line spectrum inference method to provide the estimation of AoAs. Then the user position is estimated successively based on the estimated AoAs by using the method in [43]. The results are given in Fig. 7, where the proposed BULT algorithm significantly outperforms the atomic norm based user tracking. Our proposed algorithm considers the temporal correlation of the user position by building a probability module over time. The iterative message passing between the AoA estimation module and the position tracking module utilizes the geometric constraint which makes the estimation performance close to the BCRB.

3) **Comparison of Beamforming Methods**: As shown in Fig. 8, we compare the tracking performance of the directional beamforming design with that of the BCRB based beamforming...
VII. CONCLUSION

In this paper, we studied the user localization and tracking problem in the RIS aided MIMO system. We utilized the geometric relationship between the user AoAs and the user position to establish the state transition model for the mobile user. Then, we proposed a novel message-passing based algorithm for online user tracking and AoA estimation. Furthermore, we developed the BCRB to analyse the performance limits of the tracking problem which further guides the beamforming design of the BS and the RISs. Numerical simulation results show that the performance of the proposed BULT algorithm is close to the BCRB in user localization and tracking.

APPENDIX A

GAUSSIAN MESSAGE APPROXIMATION IN (50)

The expression of $\prod_{j=1}^{K} \Delta_{\theta_{ij}} \rightarrow \mu_{ij}^{(t)}(p_{ij}^{(t)})$ is given by

$$\prod_{j=1}^{K} \Delta_{\theta_{ij}} \rightarrow \mu_{ij}^{(t)}(p_{ij}^{(t)}) \propto \exp \left( \sum_{j=1}^{K} \kappa_{\theta_{ij}}^{(t)} \cos(\pi(e^{(t)}_{j})^T e_{ij} - \mu_{\theta_{ij}}^{(t)} - \varphi_{ij}^{(t)}) \right),$$

(62)

where $e_{j}^{(t)}$ is defined as $e_{j}^{(t)} = \frac{p_{ij}^{(t)} - p_{ij}}{||p_{ij}^{(t)} - p_{ij}||_2}$. We resort to the gradient descent method to find the local maximum of (62) which is used as the mean vector $m_{G^{(t)}}$ of the approximated Gaussian message in (50). The covariance matrix $C_{G^{(t)}}$ of the approximated Gaussian message is given by the Hessian matrix at $p_{ij}^{(t)} = m_{G^{(t)}}$. We denote the exponential term of (62) as $f_{e}(p_{ij}^{(t)})$. Thus, the gradient of $f_{e}(p_{ij}^{(t)})$ is derived in

$$\nabla f_{e}(p_{ij}^{(t)}) = \frac{\partial f_{e}(p_{ij}^{(t)})}{\partial p_{ij}^{(t)}} = \sum_{j=1}^{K} \frac{\pi \kappa_{\theta_{ij}}^{(t)} \sin \left( \pi(e^{(t)}_{j})^T e_{ij} - \mu_{\theta_{ij}}^{(t)} - \varphi_{ij}^{(t)} \right) u_{ij}^{(t)}}{\| p_{ij}^{(t)} - p_{ij}^{(t)} \|_2},$$

(63)

where $\lambda$ is a predetermined step size, and $i$ is the iteration index. The iteration stops when it satisfies the stopping criterion. $m_{G^{(t)}}$ is set by the obtained local optimum. For the calculation of $C_{G^{(t)}}$, the Hessian matrix of $f_{e}(p_{ij}^{(t)})$ is expressed as

$$\frac{\partial^2 f_{e}(p_{ij}^{(t)})}{\partial p_{ij}^{(t)} \partial p_{ij}^{(t)}^T} = \sum_{j=1}^{K} \left( \frac{\pi \kappa_{\theta_{ij}}^{(t)} \cos \left( \pi(e^{(t)}_{j})^T e_{ij} - \mu_{\theta_{ij}}^{(t)} - \varphi_{ij}^{(t)} \right) u_{ij}^{(t)} \right)^T$$

$$\left( \frac{\pi \kappa_{\theta_{ij}}^{(t)} \sin \left( \pi(e^{(t)}_{j})^T e_{ij} - \mu_{\theta_{ij}}^{(t)} - \varphi_{ij}^{(t)} \right) u_{ij}^{(t)} \right)^T \frac{3e_{j}^{(t)} e_{j}^{(t)} - e_{j}^{(t)} e_{j}^{(t)} T - e_{j}^{(t)} e_{j}^{(t)} T}{\| p_{ij}^{(t)} - p_{ij}^{(t)} \|_2}.$$}

Then, we approximate $C_{G^{(t)}}$ as

$$C_{G^{(t)}} = \left( - \frac{\partial^2 f_{e}(p_{ij}^{(t)})}{\partial p_{ij}^{(t)} \partial p_{ij}^{(t)}^T} \right)^{-1}.$$
by projecting the user position error \((\mathbf{p}_{U}^{(t)} - \mathbf{p}_{U}^{(t)})\) on the directional vector \(v_i\). We represent the projection of \((\mathbf{p}_{U}^{(t)} - \mathbf{p}_{U}^{(t)})\) onto \(v_i\) by a random variable \(x_i\), following the distribution of \(N(x_i; 0, \mathbf{v}_i^T \mathbf{C}_{U,U} v_i)\). Recall from Fig. 3 that \(v_i\) is perpendicular to \(\mathbf{p}_{U,i} - \mathbf{m}_{U,i}^{(t)} v_i^{(t)}\). Thus, for a sufficiently large \(d_i\), the geometric constraint in (7) reduces to

\[
\tan(\arccos(\theta_{U,i}) - \arccos(\hat{\theta}_{U,i})) = \frac{x_i}{d_i}, \tag{67}
\]

Hence, the message in (37) is approximated as

\[
\Delta_{\psi_i}^{(t)} - \theta_{U,i}^{(t)}(\theta_{U,i}^{(t)}) \approx \int_{x_i} \frac{d}{dx_i} \tan(\arccos(\hat{\theta}_{U,i}^{(t)} - \arccos(\hat{\theta}_{U,i}^{(t)})) \exp \left( -\frac{x_i^2}{2 \mathbf{v}_i^T \mathbf{C}_{U,U} v_i} \right) \right) \exp \left( -\frac{d_i^2 \tan^2(\arccos(\hat{\theta}_{U,i}^{(t)} - \arccos(\theta_{U,i}^{(t)})))}{2 \mathbf{v}_i^T \mathbf{C}_{U,U} v_i} \right) \tag{68a}
\]

\[
\alpha \exp \left( -\frac{d_i^2 \tan^2(\arccos(\hat{\theta}_{U,i}^{(t)} - \arccos(\theta_{U,i}^{(t)})))}{2 \mathbf{v}_i^T \mathbf{C}_{U,U} v_i} \right), \tag{68b}
\]

The message in (68a) achieves its maximum at \(\theta_{U,i}^{(t)} = \hat{\theta}_{U,i}^{(t)}\). Since the VM distribution in (38) is maximized when \(\pi \theta_{U,i}^{(t)} = \mu_{\psi_i} - \theta_{U,i}^{(t)}\), we set \(\mu_{\psi_i} - \theta_{U,i}^{(t)} = \pi \hat{\theta}_{U,i}^{(t)}\). As for \(\kappa_{\psi_i} - \theta_{U,i}^{(t)}\), we take the Taylor series expansion of (38) and (68b) at the maximum with its second-order derivative given by

\[
d^2 \log \left( \frac{\Delta_{\psi_i}^{(t)} - \theta_{U,i}^{(t)}(\theta_{U,i}^{(t)})}{d\theta_{U,i}^{(t)}} \right) = -\frac{d_i^2}{\mathbf{v}_i^T \mathbf{C}_{U,U} v_i} \sin^2(\xi_i) \cos^4(\xi_i) \cos^4(\xi_i) \sin^2(\xi_i) \tag{69}
\]

\[
d^2 \log \left( \frac{d\theta_{U,i}^{(t)}}{d\theta_{U,i}^{(t)}} \right) = -\pi \kappa_{\psi_i} - \theta_{U,i}^{(t)} \cos(\pi \hat{\theta}_{U,i}^{(t)} - \mu_{\psi_i} - \theta_{U,i}^{(t)}), \tag{70}
\]

where \(\xi_i = \arccos(\theta_{U,i}^{(t)})\) and \(\hat{\xi}_i = \arccos(\hat{\theta}_{U,i}^{(t)})\). By letting

\[
d^2 \log \left( \frac{\Delta_{\psi_i}^{(t)} - \theta_{U,i}^{(t)}(\theta_{U,i}^{(t)})}{d\theta_{U,i}^{(t)}} \right) \left|_{\theta_{U,i}^{(t)} = \hat{\theta}_{U,i}^{(t)}} \right. = \frac{d_i^2}{\pi \sin^2(\xi_i) \mathbf{v}_i^T \mathbf{C}_{U,U} v_i} \tag{71}
\]

we obtain \(\kappa_{\psi_i} - \theta_{U,i}^{(t)} = \frac{d_i^2}{\pi \sin^2(\xi_i) \mathbf{v}_i^T \mathbf{C}_{U,U} v_i}\).
[22] S. E. Zegrar, L. Afeef, and H. Arslan, “A general framework for RIS-aided mmwave communication networks: Channel estimation and mobile user tracking,” arXiv preprint arXiv:2009.01180, 2020.

[23] Y. Liu, E. Liu, R. Wang, and Y. Geng, “Reconfigurable intelligent surface aided wireless localization,” in Proc. IEEE ICC, Aug. 2021, pp. 1–6.

[24] H. Zhang, H. Zhang, B. Di, K. Biao, Z. Han, and L. Song, “Metalo-localization: Reconfigurable intelligent surface aided multi-user wireless indoor localization,” IEEE Trans. Wireless Commun., Jun. 2021, early access.

[25] Y. Lin, S. Jin, M. Matthaiou, and X. You, “Channel estimation and user localization for IRS-assisted MIMO-OFDM systems,” IEEE Trans. Wireless Commun., Sep. 2021, early access.

[26] 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, Apr. 2021.

[27] J. He, H. Wymeersch, T. Sanguanpuak, O. Silvén, and M. Junnti, “Adaptive beamforming design for mmwave RIS-aided joint localization and communication,” in Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW), Apr. 2020, pp. 1–6.

[28] W. Tang, M. Z. Chen, X. Chen, J. Y. Dai, Y. Han, M. Di Renzo, Y. Zeng, S. Jin, Q. Cheng, and T. J. Cui, “Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement,” IEEE Trans. Wireless Commun., vol. 20, no. 1, pp. 421–439, Sep. 2020.

[29] M.-A. Badiu, T. L. Hansen, and B. H. Fleury, “Variational Bayesian inference of line spectra,” IEEE Trans. Signal Process., vol. 65, no. 9, pp. 2247–2261, May 2017.

[30] V. Va, H. Vikalo, and R. W. Heath, “Beam tracking for mobile millimeter wave communication systems,” in Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP), Dec. 2016, pp. 743–747.

[31] C. Zhang, D. Guo, and P. Fan, “Tracking angles of departure and arrival in a mobile millimeter wave channel,” in Proc. IEEE ICC, May 2016, pp. 1–6.

[32] K. V. Mardia and P. E. Jupp, Directional Statistics. John Wiley & Sons, 2009, vol. 494.

[33] Z. Xue, X. Yuan, and Y. Yang, “Denoising-based turbo message passing for compressed video background subtraction,” IEEE Trans. Image Process., vol. 30, pp. 2682–2696, Feb. 2021.

[34] D. L. Donoho, A. Maleki, and A. Montanari, “Message-passing algorithms for compressed sensing,” Proc. Nat. Acad. Sci., vol. 106, no. 45, pp. 18 914–18 919, Nov. 2009.

[35] J. T. Parker, P. Schniter, and V. Cevher, “Bilinear generalized approximate message passing—Part I: Derivation,” IEEE Trans. Signal Process., vol. 62, no. 22, pp. 5839–5853, Sep. 2014.

[36] S. L. Collier, “Fisher information for a complex gaussian random variable: Beamforming applications for wave propagation in a random medium,” IEEE Trans. Signal Processing, vol. 53, no. 11, pp. 4236–4248, Oct. 2005.

[37] P. Tichavsky, C. H. Muravchik, and A. Nehorai, “Posterior Cramér-Rao bounds for discrete-time nonlinear filtering,” IEEE Trans. Signal Process., vol. 46, no. 5, pp. 1386–1396, May 1998.

[38] S. M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory. Prentice-Hall, Inc., 1993.

[39] X. Hu, C. Zhong, Y. Zhang, X. Chen, and Z. Zhang, “Location information aided multiple intelligent reflecting surface systems,” IEEE Trans. Commun., vol. 68, no. 12, pp. 7948–7962, Aug. 2020.

[40] C. Cai, X. Yuan, W. Yan, Z. Huang, Y.-C. Liang, and W. Zhang, “Hierarchical passive beamforming for reconfigurable intelligent surface aided communications,” IEEE Wireless Commun. Lett., vol. 10, no. 9, pp. 1909–1913, Jun. 2021.

[41] 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, Aug. 2019.

[42] W. H. Foy, “Position-location solutions by Taylor-series estimation,” IEEE Transactions on Aerospace and Electronic Systems, vol. AES-12, no. 2, pp. 187–194, 1976.

[43] B. N. Bhaskar, G. Tang, and B. Recht, “Atomic norm denoising with applications to line spectral estimation,” IEEE Trans. Signal Process., vol. 61, no. 23, pp. 5987–5999, Jul. 2013.