

Orientation and Location Tracking of XR Devices: 5G Carrier Phase-Based Methods

Jukka Talvitie , Member, IEEE, Mikko Säily , Member, IEEE, and Mikko Valkama , Fellow, IEEE

Abstract—Accurate knowledge of the three-dimensional (3D) orientations and 3D locations of the user devices, such as wearable glasses, is of paramount importance in different extended reality (XR) use cases and applications. In this article, we address the corresponding six degrees-of-freedom (6DoF) tracking challenge of 5G-empowered XR devices. We describe a new uplink (UL) carrier phase measurements based estimation approach, allowing for low-latency 3D orientation and 3D location tracking directly at the 5G network base-stations or gNodeBs (gNBs). extended Kalman filter (EKF) based practical signal processing algorithms are described while also the applicable Cramér-Rao lower-bounds (CRLBs) are derived and presented. Also, the related aspect of over-the-air estimation of the XR headset antenna constellation or antenna geometry is addressed. Additionally, the important practical challenges related to user equipment (UE) clock drifting as well as integer ambiguities in carrier phase based methods are both considered. Finally, an extensive set of numerical results is provided in an example indoor factory like environment, covering both 3.5 GHz and 28 GHz network deployments. The obtained results demonstrate the feasibility of continuous 6DoF tracking through the proposed approach, with root mean squared error (RMSE) accuracies below one degree for the 3D orientation and below one centimeter for the 3D location, respectively. The results also demonstrate that UE clock drifting and carrier phase integer ambiguities can both be efficiently estimated and tracked, as part of the overall proposed concept and methods.

Index Terms—3D orientation estimation, 3D positioning, 5G NR, 6G, Bayesian filtering, carrier phase, clock drifting, extended reality, integer ambiguity, six degrees-of-freedom, tracking.

I. INTRODUCTION

MOBILE communication networks, most notably the fifth generation (5G) New Radio (NR) standardized by the 3rd Generation Partnership Project (3GPP), offer improved bit rates and largely reduced latencies compared to earlier cellular systems [1], [2]. XR is one of the timely new use cases of 5G and 5G-Advanced, paving the way eventually towards the future metaverse in the sixth generation (6G) era [3]. XR is generally an umbrella term, covering different variants of augmented reality

Manuscript received 14 February 2023; revised 19 June 2023 and 16 August 2023; accepted 16 August 2023. Date of publication 28 August 2023; date of current version 16 November 2023. This work was supported in part by the Academy of Finland under Grant 323244, Grant 338224, Grant 345654, Grant 352754, and Grant 357730 and in part by Nokia Corporation. The guest editor coordinating the review of this manuscript and approving it for publication was Dr. Yongpeng Wu. (Corresponding author: Mikko Valkama.)

Jukka Talvitie and Mikko Valkama are with the Tampere University, 33100 Tampere, Finland (e-mail: jukka.talvitie@tuni.fi; mikko.valkama@tuni.fi).

Mikko Säily is with the Nokia Standards, 02610 Espoo, Finland (e-mail: mikko.saily@nokia.com).

Digital Object Identifier 10.1109/JSTSP.2023.3309463

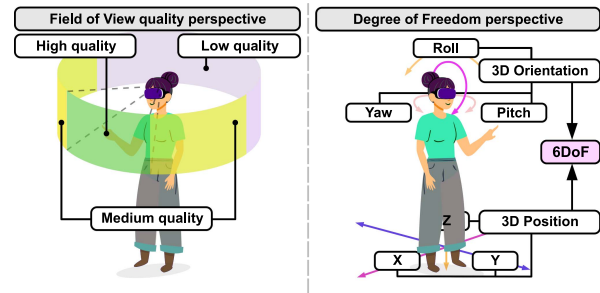


Fig. 1. Basic XR concept illustration highlighting the field-of-view and its potential impact on the content quality, together with the involved six degrees-of-freedom containing the XR device 3D position and 3D orientation.

(AR), virtual reality (VR), and mixed reality (MR), with wide variety of consumer-scale and industrial applications.

The requirements and technical aspects of XR service characteristics in 5G NR are generally well defined and summarized in [4]. One of the distinct features is the requirement for continuous and highly accurate XR device orientation and location estimation, commonly referred to as 6DoF tracking [5], [6]. Depending on the more specific XR application, such 6DoF pose information – illustrated conceptually in Fig. 1 – is needed at the network side even at millisecond-level update intervals while the accuracy requirements can be in the order of centimeters [4], [5], [7]. While user equipment (UE) positioning – and more broadly radio-based situational awareness [8], [9] – are generally inherent features of 5G and 5G-Advanced, the accuracy of the existing UL or downlink (DL) based methods is far below the XR requirements, especially when it comes to the 3D orientation estimation of the XR device [10], [11]. Terminology-wise, we use the terms UE and XR device inter-changeably in the rest of this article, as the XR device operates as a UE from the cellular network perspective.

UE orientation estimation is conventionally performed at the UE side by using antenna array measurements, camera-based methods [12], [13] or inertial sensors such as accelerometers, gyros and magnetometers [14], [15]. When using antenna arrays, the orientation can be estimated directly as part of the channel parameter estimation, where the orientation is specifically linked to the angle-of-arrival (AoA) observations or measurements. Such methods have been developed and described, e.g. in [16], [17], [18], [19], [20], [21], [22], [23]. However, in order to reach high orientation estimation accuracy, large array sizes are commonly required, which increases implementation cost and potentially introduces also additional reference signal overheads and increased latency. Furthermore, and very importantly, most of the existing angle-based methods build on UE side processing and DL measurements, hence, the orientation information has to

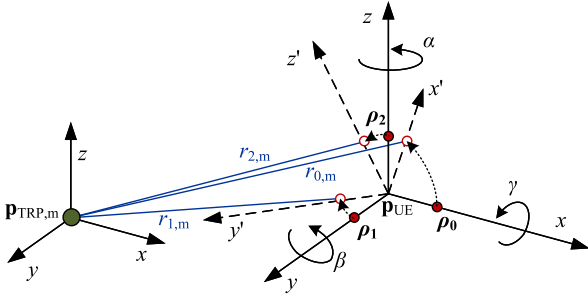


Fig. 2. Illustration of the fundamental problem geometry with a single TRP and a single UE. For presentation simplicity, the UE contains three antenna points illustrated in red. The figure also highlights the impact of 3D rotation at the UE side. The proposed concept builds on acquiring and processing UE antenna-level ranging estimates through 5G NR reference signals.

be separately communicated from the XR device to the network via applicable messages. Such signaling or message protocols are always subject to latency while also consuming the UL bandwidth.

In general, inertial measurement sensors are able to provide accurate UE orientation estimation locally, see for example [14], [24] and the references therein. This applies, however, only under appropriate calibration conditions. For example, accelerometers are able to measure orientation displacement over time, but when time evolves, also the errors accumulate larger and larger [14]. Furthermore, fusing camera-based or other optical data with inertial measurement unit (IMU) data can facilitate the orientation tracking but in case of, e.g., new environments, variable lighting condition and/or UE mobility, the camera- and IMU-based orientation tracking may cause loss of absolute orientation information, thus making the XR content rendering challenging. Use of local sensors is generally a valid approach, but in this article we deliberately pursue pure cellular radio based 6DoF tracking solutions.

To this end, in this article, we focus on the 6DoF XR device tracking through cellular carrier phase measurements. However, the proposed methods can be applied also in the context of other applications calling for 6DoF tracking. While carrier phase based methods are conceptually known and broadly utilized in global navigation satellite system (GNSS) system context, see for example [25] and the references therein, they are currently also considered in 3GPP standardization for enhanced positioning capabilities in 5G-Advanced evolution [26], [27], [28]. Compared to the existing state-of-the-art in cellular orientation estimation, we show that the proposed carrier phase based approach enables accurate orientation estimation with small numbers of antennas – as an extreme example, even a single antenna element can be exploited during orientation tracking. Additionally, and importantly, the UE orientation estimation can be performed based on UL signals, which enables real-time UE orientation estimation also at the gNB or the network transmission and reception point (TRP) side. Hence, the proposed approach enables an effective over-the-air (OTA) inertial sensor like functionality, with minimum latency, while avoiding the error accumulation problems of local IMUs and offering support for full 6DoF tracking.

For clarity, it is noted that the utilization of 5G/cellular carrier phase measurements has been addressed in the recent literature, e.g., in [29], [30], [31]. In [29], carrier phase measurements are utilized for line-of-sight (LoS)/non-line-of-sight

(NLoS) identification in complex propagation environments. Actual carrier phase based ranging and indoor positioning are addressed in [30] where the time-of-arrival (ToA) estimates obtained using already standardized DL reference signals are further enhanced through carrier phase based processing. The work in [31], in turn, focuses on improved UE clock synchronization through carrier phase measurements, and thereon for improved localization accuracy through time-based positioning methods. However, none of these works consider the important 6DoF challenge and the related 3D orientation estimation – the main technical problem addressed in this article.

In relation to the existing literature, the contributions and novelty of this article can be stated and summarized as follows:

- We describe a new carrier phase ranging based 6DoF tracking concept for continuous estimation of XR device 3D orientation and 3D location in 5G and beyond cellular systems;
- We provide applicable CRLBs, and assess how the carrier phase based measurement accuracy impacts the 6DoF estimation performance;
- We formulate EKF based state models and tracking algorithms for practical signal processing implementations;
- We propose and describe new concepts and algorithms for over-the-air estimation or calibration of the XR device antenna constellation;
- We provide further extended problem definition and EKF processing solutions to allow for estimating and tracking also the UE clock drifting and the carrier phase integer ambiguities;
- We provide extensive numerical performance results in the context of 5G NR with realistic deployment assumptions, demonstrating the feasibility of continuous 6DoF tracking using the proposed approach, even in the presence of UE clock drifting and carrier phase integer ambiguities.

The rest of this article is organized as follows: Section II describes the assumed system model in terms of the 6DoF geometry, XR device and network TRP antenna systems as well as the carrier phase measurement model. Section III provides the proposed 6DoF tracking methods while also addressing the performance bounds in the form of CRLBs. Also applicable UL-based and DL-based variants of the proposed scheme are discussed. In Section IV, important extensions to UE clock drifting and integer ambiguity tracking are provided. A large collection of numerical results is provided in Section V, while the conclusions are drawn in Section VI. Finally, selected mathematical details are provided in the Appendix.

Notations: Vectors are denoted by bold lowercase letters (i.e., \mathbf{a}), bold uppercase letters are used for matrices (i.e., \mathbf{A}) and scalars are denoted by normal font (i.e., a). The operators $(\cdot)^T$, $\mathbb{E}_x\{\cdot\}$, and $\|\cdot\|$ denote the transpose, expectation with respect to random variable x , and Euclidian norm, respectively.

II. SYSTEM MODEL

In this work, we consider standalone wearable XR devices with built-in 5G NR modem and computing capabilities, such as the device types XR5G-A4 or XR5G-A5 defined in [4]. Specifically, we consider the 6DoF scenario including the estimation of the 3D position and 3D orientation of the XR device, with the basic system geometry being illustrated in Fig. 2. Our emphasis is on the utilization of 5G NR carrier phase-based ranging measurements that can be obtained from physical signals as

described, e.g., in [30], [31]. Hence, our emphasis in this article is on the concepts and signal processing on top of the raw carrier phase measurements – instead of individual ranging algorithms. For clarity, it is noted that the basic system model considered in this section excludes UE clock drifting and carrier phase integer ambiguity challenges. However, these are addressed explicitly in Section IV.

To this end, the estimated unknown 6DoF state vector $\tilde{\mathbf{s}}$ is defined as

$$\tilde{\mathbf{s}} = [\boldsymbol{\theta}^T \quad \mathbf{p}_{\text{UE}}^T]^T, \quad (1)$$

where $\boldsymbol{\theta} = [\alpha, \beta, \gamma]^T$ describes the 3D UE orientation given in rotations α , β and γ , which refer to yaw, pitch (zero at horizon) and roll of the device, respectively. Furthermore, $\mathbf{p}_{\text{UE}} = [x_{\text{UE}}, y_{\text{UE}}, z_{\text{UE}}]^T$ is the UE position, defined as the center point of rotation, where $x_{\text{UE}}, y_{\text{UE}}, z_{\text{UE}}$ are the x-coordinate, y-coordinate and z-coordinate of the UE, respectively. Furthermore, for presentation simplicity, we consider a system with a single UE (i.e., the XR device) and M active TRPs located at $\mathbf{p}_{\text{TRP},m}$ for $m = 0, \dots, M-1$. The UE is equipped with N antennas with distinct transmit and receive RF chains. In this article, to support a straightforward generalization of the provided system model on scenarios with multiple gNBs and arbitrary antenna configurations, the TRP is defined as a single antenna element. Therefore, it is possible to consider arbitrary-shaped antenna arrays at the gNB side by simply deploying nearby TRPs according to the desired antenna array structure, and similarly, to define different gNB transceivers by deploying TRPs far away from each other. In practice, the TRP can also be implemented, e.g., as a phased antenna array as long as there is an unambiguous reference position or reference point for determining the range estimate.

By utilizing antenna-level carrier phase measurements, the range $r_{n,m}$ between the n th UE antenna and the m th TRP, can in general be measured either at the TRP based on UL signals, or at the UE based on DL signals. Considering M TRPs and N UE antennas, the measurement model for the carrier phase-based range measurements is defined as

$$\mathbf{y} = \mathbf{h}(\tilde{\mathbf{s}}) + \mathbf{n}, \quad (2)$$

where $\mathbf{h} : \mathbb{R}^6 \rightarrow \mathbb{R}^{MN}$ is the measurement function, $\mathbf{n} \in \mathbb{R}^{MN}$ is measurement noise, which determines the ranging error as $\mathbf{n} \sim \mathcal{N}(0, \boldsymbol{\Sigma})$, and $\boldsymbol{\Sigma} \in \mathbb{R}^{MN \times MN}$ is the measurement covariance matrix. Moreover, the measurement function can be further divided into TRP-wise models as

$$\mathbf{h}(\tilde{\mathbf{s}}) = [\mathbf{h}_0(\tilde{\mathbf{s}})^T \quad \mathbf{h}_1(\tilde{\mathbf{s}})^T \quad \dots \quad \mathbf{h}_{M-1}(\tilde{\mathbf{s}})^T]^T, \quad (3)$$

where the measurement function of the m th TRP $\mathbf{h}_m(\tilde{\mathbf{s}}) : \mathbb{R}^6 \rightarrow \mathbb{R}^N$ is given as

$$\mathbf{h}_m(\tilde{\mathbf{s}}) = [h_{0,m} \quad h_{1,m} \quad \dots \quad h_{N-1,m}]^T. \quad (4)$$

Finally, the measurement function, describing the range $r_{n,m}$ between the m th TRP and n th antenna element of the UE/XR device can be written as

$$h_{n,m}(\tilde{\mathbf{s}}) = r_{n,m} = \|\mathbf{p}_{\text{UE}} + \mathbf{R}(\boldsymbol{\theta})\boldsymbol{\rho}_n - \mathbf{p}_{\text{TRP},m}\|, \quad (5)$$

where $\mathbf{R}(\boldsymbol{\theta}) \in \mathbb{R}^{3 \times 3}$ is the 3D rotation matrix, which applies the rotation determined by the UE orientation $\boldsymbol{\theta}$. The rotation matrix is defined according to 5G NR specifications in [32], and is also provided in the Appendix for presentation completeness and readers' convenience. Furthermore, $\boldsymbol{\rho}_n \in \mathbb{R}^3$, with $n = 0, \dots, N-1$, is the position of the n th antenna element relative to \mathbf{p}_{UE} (UE local coordinates) without rotation (i.e.,

$\alpha = \beta = \gamma = 0$). An alternative method for representing orientations is to employ quaternions [33], which in specific use cases are able to introduce potential advantage, such as improved numerical stability. However, we follow the approach determined by the 3GPP [32], and therefore utilize the rotation matrix-based representation for the orientation in this work. Terminology-wise, we call the set $\boldsymbol{\rho}_n$, $n = 0, 1, \dots, N-1$, the XR device antenna constellation in the following.

In Fig. 2, the considered system geometry is illustrated for a scenario with a single TRP and a single UE equipped with three antennas. The reference coordinate frame is defined by the TRP as (x, y, z) , which determines the positions of the UE antennas $\boldsymbol{\rho}_n$ with zero rotation. By applying specific rotation $\mathbf{R}(\boldsymbol{\theta})$, the UE orientation can be altered, resulting in a new UE coordinate frame of (x', y', z') . The UE orientation is defined based on yaw α , pitch β and roll γ , which describe rotations of z-axis, y-axis and x-axis in respective order. By measuring the ranges $r_{n,m}$, as shown conceptually in Fig. 2, the UE 3D orientation and 3D position can be estimated and tracked through the proposed methods, described next in Section III.

III. PROPOSED METHODS AND BOUNDS

In this section, we describe the actual 6DoF estimation and tracking methods, while also provide the corresponding performance bounds in the form of CRLBs. Also selected UL-based and DL-based variants of the proposed scheme are discussed.

A. 6DoF Performance Bounds

In order to analyze the performance of carrier phase-based 3D UE orientation and 3D position estimation at a theoretical level, we derive the CRLB, which provides a theoretical lower-bound for the covariance matrix of an unbiased estimator $\hat{\tilde{\mathbf{s}}}$, expressed formally as [34]

$$\text{CRLB}(\hat{\tilde{\mathbf{s}}}) = \mathbf{J}^{-1}(\hat{\tilde{\mathbf{s}}}) \leq \mathbb{E}\{(\hat{\tilde{\mathbf{s}}} - \tilde{\mathbf{s}})(\hat{\tilde{\mathbf{s}}} - \tilde{\mathbf{s}})^T\}, \quad (6)$$

where $\mathbf{J}(\tilde{\mathbf{s}}) \in \mathbb{R}^{6 \times 6}$ is the Fisher Information Matrix (FIM). The FIM element at the i th row and j th column is defined as

$$[\mathbf{J}(\tilde{\mathbf{s}})]_{i,j} = -\mathbb{E}_{\tilde{\mathbf{s}}}\left\{\frac{\partial^2 \ell(\mathbf{y}|\tilde{\mathbf{s}})}{\partial \tilde{s}_i \partial \tilde{s}_j}\right\}, \quad (7)$$

where $\ell(\mathbf{y}|\tilde{\mathbf{s}})$ is the log-likelihood function of the measurements given the estimated state vector $\tilde{\mathbf{s}}$, \tilde{s}_i denotes the i th element of the state vector $\tilde{\mathbf{s}}$, while the state vector $\tilde{\mathbf{s}}$ is defined in (1).

Considering the Gaussian measurement model defined in (2), the FIM can be conveniently obtained as [34]

$$\mathbf{J}(\tilde{\mathbf{s}}) = \tilde{\mathbf{H}}(\tilde{\mathbf{s}})^T \boldsymbol{\Sigma}^{-1} \tilde{\mathbf{H}}(\tilde{\mathbf{s}}), \quad (8)$$

where $\tilde{\mathbf{H}}(\tilde{\mathbf{s}}) \in \mathbb{R}^{MN \times 6}$ is the Jacobian matrix of the measurement function, defined with respect to the state vector $\tilde{\mathbf{s}}$. Furthermore, the Jacobian matrix is constructed from two sub-matrices as

$$\tilde{\mathbf{H}}(\tilde{\mathbf{s}}) = [\mathbf{H}_\theta(\boldsymbol{\theta}) \quad \mathbf{H}_p(\mathbf{p}_{\text{UE}})] \quad (9)$$

where $\mathbf{H}_\theta(\boldsymbol{\theta}) \in \mathbb{R}^{MN \times 3}$ relates to the UE orientation state and $\mathbf{H}_p(\mathbf{p}_{\text{UE}}) \in \mathbb{R}^{MN \times 3}$ relates to the UE position state. The

Jacobian matrix related to the UE orientation state is given as

$$\mathbf{H}_\theta = \begin{bmatrix} \frac{\partial h_{0,0}}{\partial \alpha} & \frac{\partial h_{0,0}}{\partial \beta} & \frac{\partial h_{0,0}}{\partial \gamma} \\ \vdots & \vdots & \vdots \\ \frac{\partial h_{N-1,0}}{\partial \alpha} & \frac{\partial h_{N-1,0}}{\partial \beta} & \frac{\partial h_{N-1,0}}{\partial \gamma} \\ \frac{\partial h_{0,1}}{\partial \alpha} & \frac{\partial h_{0,1}}{\partial \beta} & \frac{\partial h_{0,1}}{\partial \gamma} \\ \vdots & \vdots & \vdots \\ \frac{\partial h_{N-1,M-1}}{\partial \alpha} & \frac{\partial h_{N-1,M-1}}{\partial \beta} & \frac{\partial h_{N-1,M-1}}{\partial \gamma} \end{bmatrix}, \quad (10)$$

where the partial derivatives, derived based on the fundamental system geometry and measurement function in (5), read as

$$\begin{aligned} \frac{\partial h_{n,m}}{\partial \alpha} &= \frac{(\mathbf{p}_{\text{UE}} + \mathbf{R}(\boldsymbol{\theta})\boldsymbol{\rho}_n - \mathbf{p}_{\text{TRP},m})^T \mathcal{R}_\alpha(\boldsymbol{\theta})\boldsymbol{\rho}_n}{\|\mathbf{p}_{\text{UE}} + \mathbf{R}(\boldsymbol{\theta})\boldsymbol{\rho}_n - \mathbf{p}_{\text{TRP},m}\|} \\ \frac{\partial h_{n,m}}{\partial \beta} &= \frac{(\mathbf{p}_{\text{UE}} + \mathbf{R}(\boldsymbol{\theta})\boldsymbol{\rho}_n - \mathbf{p}_{\text{TRP},m})^T \mathcal{R}_\beta(\boldsymbol{\theta})\boldsymbol{\rho}_n}{\|\mathbf{p}_{\text{UE}} + \mathbf{R}(\boldsymbol{\theta})\boldsymbol{\rho}_n - \mathbf{p}_{\text{TRP},m}\|} \\ \frac{\partial h_{n,m}}{\partial \gamma} &= \frac{(\mathbf{p}_{\text{UE}} + \mathbf{R}(\boldsymbol{\theta})\boldsymbol{\rho}_n - \mathbf{p}_{\text{TRP},m})^T \mathcal{R}_\gamma(\boldsymbol{\theta})\boldsymbol{\rho}_n}{\|\mathbf{p}_{\text{UE}} + \mathbf{R}(\boldsymbol{\theta})\boldsymbol{\rho}_n - \mathbf{p}_{\text{TRP},m}\|}. \end{aligned} \quad (11)$$

The matrices $\mathcal{R}_\alpha(\boldsymbol{\theta})$, $\mathcal{R}_\beta(\boldsymbol{\theta})$, $\mathcal{R}_\gamma(\boldsymbol{\theta}) \in \mathbb{R}^{3 \times 3}$ represent the partial derivatives of the rotation matrix $\mathbf{R}(\boldsymbol{\theta})$ with respect to α , β and γ in respective order, expressed formally as

$$\mathcal{R}_\alpha(\boldsymbol{\theta}) = \frac{\partial \mathbf{R}(\boldsymbol{\theta})}{\partial \alpha}, \quad \mathcal{R}_\beta(\boldsymbol{\theta}) = \frac{\partial \mathbf{R}(\boldsymbol{\theta})}{\partial \beta}, \quad \mathcal{R}_\gamma(\boldsymbol{\theta}) = \frac{\partial \mathbf{R}(\boldsymbol{\theta})}{\partial \gamma}, \quad (12)$$

whose detailed descriptions are provided for readers' convenience in the Appendix. The corresponding Jacobian matrix related to the UE position state can, in turn, be given as

$$\mathbf{H}_p = \begin{bmatrix} \frac{\partial h_{0,0}}{\partial \mathbf{p}_{\text{UE}}} & \dots & \frac{\partial h_{N-1,0}}{\partial \mathbf{p}_{\text{UE}}} & \frac{\partial h_{0,1}}{\partial \mathbf{p}_{\text{UE}}} & \dots & \frac{\partial h_{N-1,M-1}}{\partial \mathbf{p}_{\text{UE}}} \end{bmatrix}^T. \quad (13)$$

Moreover, based on the considered measurement function in (5), the partial derivatives can be expressed as

$$\frac{\partial h_{n,m}}{\partial \mathbf{p}_{\text{UE}}} = \frac{\mathbf{p}_{\text{UE}} + \mathbf{R}(\boldsymbol{\theta})\boldsymbol{\rho}_n - \mathbf{p}_{\text{TRP},m}}{\|\mathbf{p}_{\text{UE}} + \mathbf{R}(\boldsymbol{\theta})\boldsymbol{\rho}_n - \mathbf{p}_{\text{TRP},m}\|}. \quad (14)$$

At last, utilizing (6) and (8), the CRLBs for the unknown 6DoF state parameters can be stated as follows:

$$\begin{aligned} \text{CRLB}(\hat{\alpha}) &= \mathbf{J}^{-1}(\tilde{\mathbf{s}})_{[1,1]}, & \text{CRLB}(\hat{x}_{\text{UE}}) &= \mathbf{J}^{-1}(\tilde{\mathbf{s}})_{[4,4]} \\ \text{CRLB}(\hat{\beta}) &= \mathbf{J}^{-1}(\tilde{\mathbf{s}})_{[2,2]}, & \text{CRLB}(\hat{y}_{\text{UE}}) &= \mathbf{J}^{-1}(\tilde{\mathbf{s}})_{[5,5]} \\ \text{CRLB}(\hat{\gamma}) &= \mathbf{J}^{-1}(\tilde{\mathbf{s}})_{[3,3]}, & \text{CRLB}(\hat{z}_{\text{UE}}) &= \mathbf{J}^{-1}(\tilde{\mathbf{s}})_{[6,6]}. \end{aligned} \quad (15)$$

The above CRLBs can be further exploited for obtaining the orientation error bound (OEB) and positioning error bound (PEB) [17] as

$$\text{OEB} = \sqrt{\mathbf{J}^{-1}(\tilde{\mathbf{s}})_{[1,1]} + \mathbf{J}^{-1}(\tilde{\mathbf{s}})_{[2,2]} + \mathbf{J}^{-1}(\tilde{\mathbf{s}})_{[3,3]}} \quad (16)$$

$$\text{PEB} = \sqrt{\mathbf{J}^{-1}(\tilde{\mathbf{s}})_{[4,4]} + \mathbf{J}^{-1}(\tilde{\mathbf{s}})_{[5,5]} + \mathbf{J}^{-1}(\tilde{\mathbf{s}})_{[6,6]}}, \quad (17)$$

where the OEB represents the lower bound for the RMSE of the 3D orientation estimation, and PEB represents the lower bound for the 3D positioning RMSE.

Furthermore, when there is a-priori information on the estimated parameters, or the parameters are estimated recursively in time so that the preceding estimates provide information on the current estimate, the FIM can be (re-)expressed with two additive components as [35]

$$\mathbf{J}_{\text{post}}(\mathbf{s}) = \mathbf{J}_{\text{prior}}(\mathbf{s}) + \mathbf{J}_{\text{meas}}(\mathbf{s}). \quad (18)$$

In (18), the state vector notation is deliberately changed to \mathbf{s} to allow for including also rate-of-change parameters in the state definition – as shown concretely in (22). Additionally, $\mathbf{J}_{\text{post}}(\mathbf{s})$ is the FIM of the a-posteriori estimate, $\mathbf{J}_{\text{prior}}(\mathbf{s})$ is the FIM of the a-priori estimate, and $\mathbf{J}_{\text{meas}}(\mathbf{s})$ is the FIM obtained from the measurements, calculated conceptually similar to (8). The prior information of the state is determined as

$$[\mathbf{J}_{\text{prior}}(\mathbf{s})]_{i,j} = -\mathbb{E}_{\mathbf{s}} \left\{ \frac{\partial^2 \log \mathcal{P}(\mathbf{s})}{\partial s_i \partial s_j} \right\}, \quad (19)$$

where $\mathcal{P}(\mathbf{s})$ denotes the prior probability distribution of state \mathbf{s} . Considering a system with dynamic parameters, which are estimated and tracked recursively over time steps k , the a-posteriori FIM can be re-expressed as

$$\mathbf{J}_{\text{post}}[k] = \mathbf{J}_{\text{prior}}[k] + \mathbf{J}_{\text{meas}}[k], \quad (20)$$

wherein for presentation convenience, we have omitted the state vector from the FIM notations while emphasizing time k . Furthermore, the a-priori FIM is defined as

$$\mathbf{J}_{\text{prior}}[k] = \mathbf{F}\mathbf{J}_{\text{post}}^{-1}[k-1]\mathbf{F}^T + \mathbf{Q}[k], \quad (21)$$

where \mathbf{F} describes the state-transition matrix, and $\mathbf{Q}[k]$ is the process covariance, which are both defined and further described along the upcoming EKF formulations in Section III-B, specifically (25) and (26). Furthermore, numerical results of the stated bounds are provided in Section V.

B. 6DoF Tracking Through Extended Kalman Filtering

In the following, we next formulate an efficient EKF-based approach for estimating and tracking the desired 6DoF parameters of an XR device in a time-varying mobile scenario. To this end, we first extend and redefine the static state vector, given in (1), to support the desired mobility, and define the time-varying state vector at the k th time step as

$$\mathbf{s}[k] = [\boldsymbol{\theta}[k]^T \quad \Delta\boldsymbol{\theta}[k]^T \quad \mathbf{p}_{\text{UE}}[k]^T \quad \mathbf{v}_{\text{UE}}[k]^T]^T \in \mathbb{R}^{12} \quad (22)$$

where $\boldsymbol{\theta}[k] = [\alpha[k], \beta[k], \gamma[k]]^T$ is the 3D UE orientation, and $\Delta\boldsymbol{\theta}[k] = [\Delta\alpha[k], \Delta\beta[k], \Delta\gamma[k]]^T$ is the rotational rate-of-change, including distinct rates-of-change of rotation angles α , β and γ , respectively. In addition, $\mathbf{p}_{\text{UE}}[k] = [x_{\text{UE}}[k], y_{\text{UE}}[k], z_{\text{UE}}[k]]^T$ denotes the 3D UE position, and $\mathbf{v}_{\text{UE}}[k] = [v_x[k], v_y[k], v_z[k]]^T$ is the 3D UE velocity, defined relative to the x-axis, y-axis and z-axis, respectively.

For the proposed EKF formulation, we assume that the UE velocity, as well as the rate-of-change of the UE rotation, are almost constant over a small time period. Consequently, we utilize a linear constant white-noise acceleration (CWNA) state-transition model [36], and accordingly define the state-space model, consisting of the state-transition model and measurement model as

$$\mathbf{s}[k] = \mathbf{F}\mathbf{s}[k-1] + \mathbf{w}[k] \quad (23)$$

$$\mathbf{y}[k] = \mathbf{h}(\tilde{\mathbf{s}}[k]) + \mathbf{n}[k], \quad (24)$$

where $\mathbf{F} \in \mathbb{R}^{12 \times 12}$ is the state-transition matrix, and $\mathbf{w}[k] \sim \mathcal{N}(0, \mathbf{Q}[k]) \in \mathbb{R}^{12}$ is the driving noise of the state-transition model. Furthermore, the state-transition matrix is expressed as

$$\mathbf{F} = \begin{bmatrix} \bar{\mathbf{F}} \otimes \mathbf{I}_3 & \mathbf{0}_{6 \times 6} \\ \mathbf{0}_{6 \times 6} & \bar{\mathbf{F}} \otimes \mathbf{I}_3 \end{bmatrix}, \quad \text{where } \bar{\mathbf{F}} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}, \quad (25)$$

and Δt is the time duration between two consecutive time steps. According to CWNA state-transition model [36], the process

covariance matrix can be given as

$$\mathbf{Q}[k] = \begin{bmatrix} \sigma_{\Delta\theta}^2 \bar{\mathbf{Q}}[k] \otimes \mathbf{I}_3 & \mathbf{0}_{6 \times 6} \\ \mathbf{0}_{6 \times 6} & \sigma_v^2 \bar{\mathbf{Q}}[k] \otimes \mathbf{I}_3 \end{bmatrix} \in \mathbb{R}^{12 \times 12}, \text{ where } (26)$$

$$\bar{\mathbf{Q}}[k] = \begin{bmatrix} \frac{\Delta t^3}{3} & \frac{\Delta t^2}{2} \\ \frac{\Delta t^2}{2} & \Delta t \end{bmatrix}. \quad (27)$$

The magnitude of process covariance can be scaled with $\sigma_{\Delta\theta}^2$ and σ_v^2 , which describe the power spectral density of angular acceleration of each rotation angle (α , β and γ), and the power spectral density of UE acceleration with respect to each component axis (x_{UE} , y_{UE} , z_{UE}), respectively. The measurement model in (24) follows the formulation provided in (2). Accordingly, $\mathbf{y}[k] \in \mathbb{R}^{MN}$ denotes the measurement vector, which is build on the measurement function $\mathbf{h} : \mathbb{R}^6 \rightarrow \mathbb{R}^{MN}$, defined in (3), and on the measurement noise $\mathbf{n}[k] \sim \mathcal{N}(0, \Sigma[k])$. For the measurement function input, parameters of velocity and rotational rate-of-change are omitted from the EKF state vector so that $\hat{\mathbf{s}}[k] = [\boldsymbol{\theta}[k]^T, \mathbf{p}_{\text{UE}}[k]^T]^T \in \mathbb{R}^6$.

Given the state-space model in (23) and (24), the EKF-based estimation and tracking proceeds through iterations of a prediction step and an update step. First, at time index k , the prediction step of the EKF is expressed as

$$\begin{aligned} \hat{\mathbf{s}}^-[k] &= \mathbf{F}\hat{\mathbf{s}}^+[k-1] \\ \hat{\mathbf{P}}^-[k] &= \mathbf{F}\hat{\mathbf{P}}^+[k-1]\mathbf{F}^T + \mathbf{Q}[k] \end{aligned} \quad (28)$$

where $\hat{\mathbf{s}}^-[k] \in \mathbb{R}^{12}$ and $\hat{\mathbf{P}}^-[k] \in \mathbb{R}^{12 \times 12}$ are the a-priori estimates of the mean and covariance of the state $\mathbf{s}[k]$, respectively. After this, the a-posteriori estimates of the mean and the covariance, denoted as $\hat{\mathbf{s}}^+[k] \in \mathbb{R}^{12}$ and $\hat{\mathbf{P}}^+[k] \in \mathbb{R}^{12 \times 12}$, respectively, can be obtained by the update step of the EKF as

$$\begin{aligned} \hat{\mathbf{s}}^+[k] &= \hat{\mathbf{s}}^-[k] + \mathbf{K}[k](\hat{\mathbf{y}}[k] - \mathbf{h}(\hat{\mathbf{s}}^-[k])) \\ \hat{\mathbf{P}}^+[k] &= (\mathbf{I}_{12} - \mathbf{K}[k]\mathbf{H}[k])\hat{\mathbf{P}}^-[k], \end{aligned} \quad (29)$$

where the Kalman gain $\mathbf{K}[k] \in \mathbb{R}^{12 \times MN}$ is defined as

$$\mathbf{K}[k] = \hat{\mathbf{P}}^-[k]\mathbf{H}[k]^T(\mathbf{H}[k]\hat{\mathbf{P}}^-[k]\mathbf{H}[k]^T + \Sigma[k])^{-1}. \quad (30)$$

In addition, $\hat{\mathbf{y}}[k]$ denotes the obtained measurements at time index k , and $\hat{\mathbf{s}}^-[k]$ is the estimated a-priori state vector without the velocity and rotational rate-of-change parameters, as given in (24). Furthermore, $\mathbf{H}[k] \in \mathbb{R}^{MN \times 12}$ is the Jacobian matrix for attaining the linear approximation of the non-linear measurement function. For the considered EKF formulation, the Jacobian matrix can be written as

$$\mathbf{H}[k] = [\mathbf{H}_\theta[k] \quad \mathbf{0}_{NM \times 3} \quad \mathbf{H}_p[k] \quad \mathbf{0}_{NM \times 3}], \quad (31)$$

where $\mathbf{H}_\theta[k]$ and $\mathbf{H}_p[k]$ are evaluated at the a-priori mean estimate $\hat{\mathbf{s}}^-[k]$ as derived and expressed in (10) and (13), respectively.

Finally, it is noted and highlighted that the considered nonlinear filtering problem could be solved through other alternative methods, such as the unscented Kalman filter (UKF) or the cubature Kalman filter (CKF), that also build on Gaussian density approximations. However, since the underlying models are differentiable and because the EKF is shown through the numerical results to essentially reach the respective posterior bounds, EKF is a well-argued nonlinear filtering approach and thus adopted in this article. In general, the main advantage of extended Kalman filter over alternative nonlinear filtering solutions, such as the UKF or the particle filter, is its relatively low computational complexity compared to its filtering performance [37]. We have

also verified through concrete numerical evaluations that UKF and CKF yield essentially the same performance as the EKF-based solutions.

C. Uplink vs. Downlink Aspects

In general, the proposed methods can be utilized for either DL-based or UL-based measurements, depending primarily on which reference signals are available and utilized for carrier phase ranging. Furthermore, depending on the more specific approach, there is need to report certain parameters or measurements between the TRPs and the UE. This directly affects the estimation latency and the amount of the side-information signaling, and the used approach for a given XR application should be eventually selected based on the underlying performance requirements defined in [4].

In the UL-based approach, a feasible physical estimation resource is the uplink sounding reference signal (SRS) [1], [10] and hence the TRPs can obtain the ranging estimates through the corresponding carrier phase measurements. If the actual 6DoF UE tracking is then also pursued at the network, information about the UE antenna constellation needs to be communicated as side-information – or alternatively, estimated over-the-air as described in Section III-D. As a further alternative, the ranging measurements can be communicated from the network towards the UE, together with the TRP locations, after which UE can execute the 6DoF tracking.

In the DL-based approach, in turn, the most feasible estimation resource is the downlink positioning reference signal (PRS) [1], [10] with which the UE can calculate the antenna-level range estimates through the carrier phase measurements. If the 6DoF tracking is then also pursued at the UE, side-information about the TRP locations is needed through DL signaling. As an alternative, the DL PRS based ranging measurements can be communicated towards the network, together with the UE antenna constellation information, which then allow the 6DoF tracking to run at the network.

Overall, given the strict latency requirements for photorealistic rendering in different XR use cases [4], [7], the UL-based approach with UE antenna configuration or constellation signaling allows for lowest latency. Such an approach can be seen as an effective over-the-air inertial sensor like functionality with support for full 6DoF tracking without error accumulation problems commonly observed with, e.g., IMUs [14]. Finally, it is noted that the provided EKF formulation provides also information about the prevailing rate-of-change of the XR device orientation, due to the adopted state variable formulation in (22). When such information is available in the network, with minimum latency through the UL-based approach, viewpoint prediction, proactive rendering and other similar schemes discussed, e.g., in [38], become basically technically feasible.

While the baseline approach for acquiring the headset antenna configuration or constellation at the network is through uplink feedback signaling, we next describe a novel OTA measurement and estimation approach for obtaining such knowledge directly from the uplink radio signals.

D. Extension to Over-the-Air Estimation or Calibration of Headset Antenna Constellation

The 6DoF tracking through EKF, presented in Section III-B, assumes a known headset antenna constellation. That is, the position of each antenna element, $\boldsymbol{\rho}_n$, with respect to the UE

position with zero-rotation were assumed to be known during the tracking process. However, in certain practical scenarios, such information on the antenna constellation might be unavailable, or the information is not desired to be shared with the entity performing the 6DoF tracking. In such cases, it is possible to extend the proposed EKF to enable over-the-air estimation of the antenna constellation.

Considering that the headset antenna constellation, described by the relative antenna element positions ρ_n , is not known, it is unfeasible to jointly estimate and track the antenna constellation, the device position \mathbf{p}_{UE} and device orientation θ – all simultaneously. As defined in the measurement function in (5), the element positions ρ_n are determined with respect to the device position (center point-of-rotation) and zero-orientation, which results in implicit relation between the unknown antenna constellation and the unknown device position and orientation. Nonetheless, although the absolute joint estimates of the above parameters are out of reach, it is possible to estimate and track the constellation so that the antenna elements are positioned with respect to each other, revealing the fundamental geometric structure of the antenna constellation. The estimated antenna constellation is then connected to the true constellation through specific rotation and position offset.

For the proposed over-the-air calibration of the XR headset antenna constellation, we consider four separate approaches:

- 1) *Unknown Movement and Rotation*: In the first approach, the XR device is moving and rotating in the area without any prior knowledge of the device positions or orientations. Consequently, the EKF state vector is given as $[\theta[k]^T, \Delta\theta[k]^T, \check{\mathbf{p}}_{\text{UE}}[k]^T, \mathbf{v}_{\text{UE}}[k]^T, \check{\rho}_0[k]^T, \dots, \check{\rho}_{N-1}[k]^T]^T \in \mathbb{R}^{12+3N}$, where the $\theta[k]$, $\check{\mathbf{p}}_{\text{UE}}[k]$ and $\check{\rho}_n[k]$ are the position and orientation of the UE/XR device, and the relative position of the n th antenna element, respectively. These state values are connected to the corresponding true values through a specific rotation and position offset.
- 2) *Unknown Rotation at a Fixed Known Position*: In the second considered approach, the XR device is rotating at a fixed and known position without knowing the orientations. Since the position of the UE/XR device is known, the EKF state vector is reduced to $[\theta[k]^T, \Delta\theta[k]^T, \check{\rho}_0[k]^T, \dots, \check{\rho}_{N-1}[k]^T]^T \in \mathbb{R}^{6+3N}$, thus including only tracking of the orientation and antenna constellation.
- 3) *Unknown Movement and Fixed Known Rotation*: In the third approach, the XR device is moving in the area with a fixed and known orientation. In this case, the EKF state vector is given as $[\check{\mathbf{p}}_{\text{UE}}[k]^T, \mathbf{v}_{\text{UE}}[k]^T, \check{\rho}_0[k]^T, \dots, \check{\rho}_{N-1}[k]^T]^T \in \mathbb{R}^{6+3N}$, and therefore considering only tracking of the position and antenna constellation.
- 4) *Known Movement and Rotation*: In the fourth approach, the XR device is moving and rotating in the area by assuming perfect knowledge of the device positions and orientations at each time instant. As both the position and orientation of the UE/XR device are known, the EKF state vector includes only the unknown antenna constellation given as $[\check{\rho}_0[k]^T, \dots, \check{\rho}_{N-1}[k]^T]^T \in \mathbb{R}^{3N}$.

Out of these, the approach or scenario 4) serves as the performance benchmark, while items 1), 2) and 3) are more realistic in practical applications.

For all the considered approaches defined above, the estimation and tracking of the headset antenna constellation can be carried out by using the EKF formulations, presented in

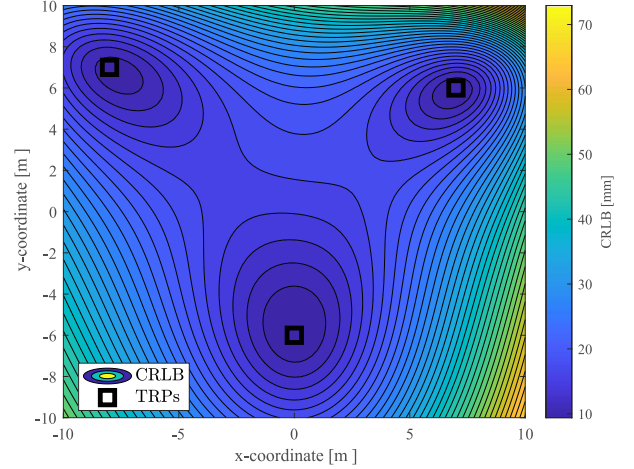


Fig. 3. Illustration of the CRLB for estimating the antenna constellation ρ_n , $n = 0, \dots, N-1$, at different UE positions with three TRPs. The CRLB refers to the average of the single-shot measurement bounds of individual antennas.

Section III-B. However, the Jacobian matrix, shown in (31), must be extended according to the unknown antenna constellation parameters. To this end, considering measurements from the m th TRP, the Jacobian matrix $\mathbf{H}_{\rho,m} \in \mathbb{R}^{N \times 3N}$ for the antenna constellation ρ_n with $n = 0, \dots, N-1$ can be written as

$$\mathbf{H}_{\rho,m} = \begin{bmatrix} \frac{\partial h_{0,m}}{\partial \rho_0} & \mathbf{0}_{1 \times 3} & \cdots & \mathbf{0}_{1 \times 3} \\ \mathbf{0}_{1 \times 3} & \frac{\partial h_{1,m}}{\partial \rho_1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{1 \times 3} \\ \mathbf{0}_{1 \times 3} & \cdots & \mathbf{0}_{1 \times 3} & \frac{\partial h_{N-1,m}}{\partial \rho_{N-1}} \end{bmatrix}, \text{ where } (32)$$

$$\frac{\partial h_{n,m}}{\partial \rho_n} = \mathbf{R}(\theta)^T \frac{\mathbf{p}_{\text{UE}} + \mathbf{R}(\theta)\rho_n - \mathbf{p}_{\text{TRP},m}}{\|\mathbf{p}_{\text{UE}} + \mathbf{R}(\theta)\rho_n - \mathbf{p}_{\text{TRP},m}\|}. \quad (33)$$

For example, regarding the first approach with unknown movement and rotation, the Jacobian matrix can be constructed as $[\mathbf{H}_\theta, \mathbf{0}_{NM \times 3}, \mathbf{H}_p, \mathbf{0}_{NM \times 3}, [\mathbf{H}_{\rho,0}^T, \dots, \mathbf{H}_{\rho,M-1}^T]^T]$, which is an extension of the original Jacobian matrix shown in (31).

After the estimation and tracking procedure, or at any time instant during the procedure, the estimated antenna constellation can be transformed into the desired frame of reference as

$$\hat{\rho}_n = \mathbf{R}(\theta_{\text{REF}})^T (\mathbf{R}(\hat{\theta})\check{\rho}_n + \check{\mathbf{p}}_{\text{UE}} - \mathbf{p}_{\text{UE,REF}}) \quad (34)$$

where $\hat{\rho}_n$ is the estimated relative position of the m th antenna while θ_{REF} and $\mathbf{p}_{\text{UE,REF}}$ are the orientation and position of the XR device, all in the desired frame of reference. By defining the frame of reference according to the original definition of ρ_n (i.e., antenna elements defined relative to the device position with zero rotation), it is possible to evaluate the estimation accuracy for the antenna constellation. Furthermore, since the described procedure involves setting a specific reference orientation and position, it can also be considered as one kind of a calibration process, where the estimated frame of reference is associated with the used global frame of reference.

The theoretical estimation accuracy for the antenna constellation can be derived by applying (32) to the derivations presented in Section III-A. In Fig. 3, the behavior of the corresponding CRLB for estimating the antenna constellation is illustrated using single-shot measurements from three TRPs, whose exact locations are provided in Section V. The figure presents the average CRLB over all antennas at different XR device locations, and

shows that the estimation accuracy is the highest when the device is located close to a TRP. Although not visible in the figure, at a fixed location, the device orientation does not significantly affect the calibration bound. However, minor variations due to orientation can be observed, as different orientations alter the distance between the antenna elements and TRPs. The figure also clearly illustrates that the calibration performance is favorable when the device is located inside the convex hull composed by the three TRPs. This is due to the more favorable geometry in the corresponding area, compared to being located outside. Additional concrete numerical examples will be provided in Section V.

IV. ADDRESSING INTEGER AMBIGUITY AND CLOCK CHALLENGES

In this section, we extend the problem definition to cover carrier phase measurements with integer ambiguity as well as unknown UE clock drifting. Furthermore, we formulate a joint EKF-based tracking solution for UE position, UE orientation, integer ambiguities, and UE clock. After this, we describe the corresponding tracking method with differential measurements.

A. Carrier Phase Measurements With Unknown Integer Ambiguity and UE Clock Drifting

The carrier phase measurement for the n th antenna element, obtained at the m th TRP, can be written as [31]

$$\phi_{n,m} = r_{n,m} + c\delta - N_{n,m}^{\text{int}}\lambda + \nu_{n,m}, \quad (35)$$

where $r_{n,m}$ is the range between the n th antenna element and the m th TRP, as described in (5). In addition, c is the speed of light, δ is the clock error between the UE and TRPs, λ is the carrier wavelength, and $\nu_{n,m} \sim \mathcal{N}(0, \sigma_\nu^2)$ is measurement noise. Moreover, $N_{n,m}^{\text{int}}$ denotes the integer ambiguity for the pseudorange between n th antenna element and m th TRP, where the pseudorange refers to the range with a clock error as $r_{n,m} + c\delta$. The unknown UE clock δ is considered time-varying and can be modeled at time instant k as [39], [40]

$$\begin{aligned} \delta[k] &= \delta[k-1] + \Delta\delta[k]\Delta t \\ \Delta\delta[k] &= \Delta\delta[k-1] + \xi[k], \end{aligned} \quad (36)$$

where $\Delta\delta[k]$ denotes a time-varying clock skew, and $\xi[k] \sim \mathcal{N}(0, \sigma_\xi^2)$.

B. 6DoF Tracking Under Integer Ambiguities and UE Clock Errors

For joint tracking of position, orientation, integer ambiguities, and clock error, the state vector shown in (22) is extended with new state variables incorporating the integer ambiguities and clock error as

$$\begin{aligned} \mathbf{s}_{\text{int\&clk}}[k] &= [\mathbf{s}[k]^T \quad \bar{\mathbf{s}}_{\text{int\&clk}}[k]^T]^T \in \mathbb{R}^{14+2NM}, \text{ where} \\ \bar{\mathbf{s}}_{\text{int\&clk}}[k] &= [N_{0,0}^{\text{int}}[k] \quad \dots \quad N_{N,M}^{\text{int}}[k] \quad \Delta N_{0,0}^{\text{int}}[k] \quad \dots \\ &\quad \Delta N_{N,M}^{\text{int}}[k] \quad \delta[k] \quad \Delta\delta[k]]^T \in \mathbb{R}^{2+2NM}. \end{aligned} \quad (37)$$

In (37), $\delta[k]$, $\Delta\delta[k]$ and $N_{n,m}^{\text{int}}[k]$ are as defined above, while $\Delta N_{n,m}^{\text{int}}[k]$ denotes the integer ambiguity rate-of-change for the n th antenna element and m th TRP. Compared to conventional approaches in the literature such as [41], [42], where integer ambiguity resolution is performed outside of the Kalman filter,

we estimate all parameters within the EKF. By this way, we pursue towards more computationally attractive solution for practical signal processing implementations – especially with multiple UE antennas and TRPs.

For tracking purposes, the state-transition models of integer ambiguity and clock are assumed to follow a CWNA model similar to the device position and orientation. Consequently, to obtain the a-priori estimates of the mean and covariance, denoted as $\hat{\mathbf{s}}_{\text{int\&clk}}^-[k]$ and $\hat{\mathbf{P}}_{\text{int\&clk}}^-[k]$, respectively, the EKF prediction step considering integer ambiguities and clock bias can be performed by following the same approach as shown in (28). However, the state transition matrix and process covariance matrix are extended to block-diagonal matrices $\mathbf{F}_{\text{int\&clk}} = \text{diag}(\mathbf{F}, \bar{\mathbf{F}} \otimes \mathbf{I}_{NM}, \bar{\mathbf{F}}) \in \mathbb{R}^{14+2NM}$ and $\mathbf{Q}_{\text{int\&clk}}[k] = \text{diag}(\mathbf{Q}[k], \sigma_{\Delta N^{\text{int}}}^2 \mathbf{Q}[k] \otimes \mathbf{I}_{NM}, \sigma_{\Delta\delta}^2 \bar{\mathbf{Q}}[k]) \in \mathbb{R}^{14+2NM}$, respectively. In the covariance matrix $\mathbf{Q}_{\text{int\&clk}}[k]$, the parameters $\sigma_{\Delta N^{\text{int}}}^2$ and $\sigma_{\Delta\delta}^2$ denote the power spectrum densities of process acceleration for integer ambiguity and clock in respective order. Furthermore, the a-posteriori estimates of the mean and the covariance, denoted as $\hat{\mathbf{s}}_{\text{int\&clk}}^+[k]$ and $\hat{\mathbf{P}}_{\text{int\&clk}}^+[k]$, can be obtained according to (29) by considering modified state variables and measurement model. To this end, by considering the measurement model with integer ambiguity and clock bias in (35), the Jacobian matrix for the EKF update phase is extended with partial derivatives with respect to the tracked integer ambiguities and clock as $\partial\phi_{n,m}/\partial N_{n,m}^{\text{int}} = -\lambda$ and $\partial\phi_{n,m}/\partial\delta = c$ in respective order. Since the partial derivatives with respect to the orientation $\boldsymbol{\theta}$ and position \mathbf{p}_{UE} are identical to the ones presented in Section III-A, the Jacobian matrix for tracking with integer ambiguities and clock is given as $\mathbf{H}_{\text{int\&clk}}[k] = [\mathbf{H}[k], -\lambda \mathbf{I}_{NM}, \mathbf{0}_{NM \times NM}, c \mathbf{1}_{NM}, \mathbf{0}_{NM \times 1}]$, where $\mathbf{1}_{NM}$ is a vector of all-ones with NM elements, and $\mathbf{H}[k]$ is as described in (31).

As defined in (37), the proposed EKF formulation with unknown integer ambiguity and UE clock considers all state variables as real numbers, which does not reflect the true behaviour of integer ambiguities. Therefore, we propose using an additional EKF update phase, where the real-valued integer ambiguity estimates from $\hat{\mathbf{s}}_{\text{int\&clk}}^+[k]$ are rounded and fed back to the EKF in form of *virtual integer ambiguity measurements*. This approach enables estimation of rounded (integer precision) integer ambiguities inside the same EKF without introducing supplementary algorithms for integer ambiguity resolution. Accordingly, the additional EKF update step can be defined as

$$\begin{aligned} \hat{\mathbf{s}}_{\text{int\&clk}}^{++}[k] &= \hat{\mathbf{s}}_{\text{int\&clk}}^+[k] + \mathbf{K}_{\text{int}}[k](\hat{\mathbf{y}}^{\text{int}}[k] - \mathbf{H}_{\text{N}}^{\text{int}} \hat{\mathbf{s}}_{\text{int\&clk}}^+[k]) \\ \hat{\mathbf{P}}_{\text{int\&clk}}^{++}[k] &= (\mathbf{I}_{14+2NM} - \mathbf{K}_{\text{int}}[k] \mathbf{H}_{\text{N}}^{\text{int}}) \hat{\mathbf{P}}_{\text{int\&clk}}^+[k], \end{aligned} \quad (38)$$

where the Kalman gain $\mathbf{K}_{\text{int}}[k] \in \mathbb{R}^{(14+2NM) \times MN}$ reads as

$$\mathbf{K}_{\text{int}}[k] = \hat{\mathbf{P}}_{\text{int\&clk}}^+[k] \mathbf{H}_{\text{N}}^{\text{int}T} (\mathbf{H}_{\text{N}}^{\text{int}} \hat{\mathbf{P}}_{\text{int\&clk}}^+[k] \mathbf{H}_{\text{N}}^{\text{int}T} + \boldsymbol{\Sigma}_{\text{N}}^{\text{int}}[k])^{-1},$$

and $\hat{\mathbf{y}}^{\text{int}}[k] = [\hat{y}_{0,0}^{\text{int}}[k], \dots, \hat{y}_{N,M}^{\text{int}}[k]]^T \in \mathbb{R}^{MN}$ is the measurement vector including the rounded integer estimates. Furthermore, $\mathbf{H}_{\text{N}}^{\text{int}} = [\mathbf{0}_{NM \times 12}, \mathbf{I}_{NM}, \mathbf{0}_{NM \times (NM+2)}]$ and $\boldsymbol{\Sigma}_{\text{N}}^{\text{int}}[k] \in \mathbb{R}^{NM \times NM}$ are the constant measurement matrix and the covariance matrix for the virtual integer ambiguity measurements, respectively. To obtain the rounded integer ambiguity measurements for the additional update step, we consider a simple rounding function $\hat{y}_{n,m}^{\text{int}}[k] = \text{round}(\hat{N}_{n,m}^{\text{int}}[k])$, where $\hat{N}_{n,m}^{\text{int}}[k]$ is the real-valued integer ambiguity estimate from $\hat{\mathbf{s}}_{\text{int\&clk}}^+[k]$.

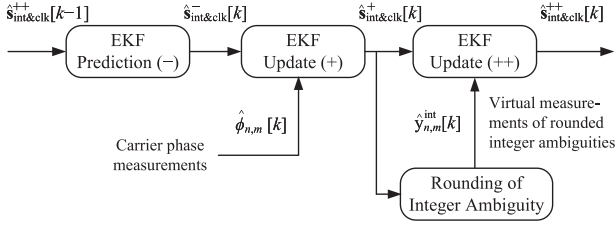


Fig. 4. EKF processing steps for tracking with integer ambiguities and unknown clock bias. Conceptually similar processing steps are also utilized with differential carrier phase measurements. Instead of rounding, also alternative integer ambiguity estimation methods can be deployed.

Alternatively, one could also consider the least-squares ambiguity decorrelation adjustment (LAMBDA) method [43], where with the estimated a-posteriori covariance $\hat{\mathbf{P}}_{\text{int\&clk}}^{+}[k]$, all integer ambiguities can be jointly estimated based on the least squares principle. Compared to the simple rounding scheme, the LAMBDA method can provide increased estimation robustness in challenging operation environments at the cost of higher computational complexity. This is because the LAMBDA method includes possibility to scale the integer ambiguity search space, and thus improves integer ambiguity resolution under challenging cycle slips. However, since the search space grows exponentially when increasing the number of UE antennas and TRPs, the resulting computational complexity must be carefully considered in practical scenarios with real-time processing. An illustration of the proposed EKF processing steps is shown in Fig. 4.

C. 6DoF Tracking Through Differential Carrier Phase Measurements

Considering the carrier phase measurement of the n_{ref} th antenna element and m_{ref} th TRP as a reference, a *differential phase measurement* is described for $(n, m) \neq (n_{\text{ref}}, m_{\text{ref}})$ as

$$\begin{aligned} \phi_{n,m}^{\text{diff}} &= \phi_{n,m} - \phi_{n_{\text{ref}},m_{\text{ref}}} \\ &= r_{n,m} - r_{n_{\text{ref}},m_{\text{ref}}} - N_{n,m}^{\text{diff}} \lambda + \Delta \nu_{n,m}, \end{aligned} \quad (39)$$

where $N_{n,m}^{\text{diff}} = N_{n,m}^{\text{int}} - N_{n_{\text{ref}},m_{\text{ref}}}^{\text{int}}$ denotes the integer ambiguity difference between a specific measurement and the corresponding reference. Moreover, $\Delta \nu_{n,m} = \nu_{n,m} - \nu_{n_{\text{ref}},m_{\text{ref}}}$ is the effective measurement noise of the differential measurement. Compared to the formulations and approach in Section IV-B, the differential measurement approach allows to suppress the UE clock offset impact without explicit clock tracking.

The tracked state vector with differential phase measurements can now be written as

$$\mathbf{s}_{\text{diff}}[k] = [\mathbf{s}[k]^T \quad \bar{\mathbf{s}}_{\text{diff}}[k]^T]^T \in \mathbb{R}^{12+2(NM-1)}, \quad (40)$$

where $\bar{\mathbf{s}}_{\text{diff}}[k] \in \mathbb{R}^{2(NM-1)}$ for $(n, m) \neq (n_{\text{ref}}, m_{\text{ref}})$ reads

$$\bar{\mathbf{s}}_{\text{diff}}[k] = [N_{0,0}^{\text{diff}}[k] \cdots N_{N,M}^{\text{diff}}[k] \Delta N_{0,0}^{\text{diff}}[k] \cdots \Delta N_{N,M}^{\text{diff}}[k]]^T,$$

and $N_{n,m}^{\text{diff}}[k]$ and $\Delta N_{n,m}^{\text{diff}}[k]$ are the integer ambiguity difference and its rate-of-change. Again, assuming the CWNA-based state-transition model, the EKF prediction step to obtain the a-priori estimates of the mean $\hat{\mathbf{s}}_{\text{diff}}^{-}[k]$ and covariance $\hat{\mathbf{P}}_{\text{diff}}^{-}[k]$ can be performed based on (28) by considering the following modifications. Firstly, the state-transition matrix and the process covariance matrix for tracking with differential carrier phase measurements are

described as block-diagonal matrices $\mathbf{F}_{\text{diff}} = \text{diag}(\mathbf{F}, \bar{\mathbf{F}} \otimes \mathbf{I}_{NM-1})$ and $\mathbf{Q}_{\text{diff}}[k] = \text{diag}(\mathbf{Q}[k], \sigma_{\Delta N_{\text{diff}}}^2 \bar{\mathbf{Q}}[k] \otimes \mathbf{I}_{NM-1})$, respectively. Moreover, $\sigma_{\Delta N_{\text{diff}}}^2$ denotes the process acceleration for integer ambiguity difference.

For the EKF update step with differential carrier phase measurements, the a-posteriori estimates of the mean and covariance, denoted as $\hat{\mathbf{s}}_{\text{diff}}^{+}[k]$ and $\hat{\mathbf{P}}_{\text{diff}}^{+}[k]$, can be obtained based on (29) by appropriately modifying the state variables and measurement model. Accordingly, given the measurement model in (39), the Jacobian matrix for differential phase measurements includes additional terms of integer ambiguity differences, given as $\partial \phi_{n,m}^{\text{diff}} / \partial N_{n,m}^{\text{diff}} = -\lambda$, and can be further described as $\mathbf{H}_{\text{diff}}[k] = [\bar{\mathbf{H}}_{\text{diff}}[k], -\lambda \mathbf{I}_{NM-1}, \mathbf{0}_{(NM-1) \times (NM-1)}]$. Moreover, $\bar{\mathbf{H}}_{\text{diff}}[k] \in \mathbb{R}^{(NM-1) \times 12}$ is defined according to (31) by replacing the partial derivatives with respect to the position \mathbf{p}_{UE} and rotation α , and similarly for β and γ , as $\partial \phi_{n,m}^{\text{diff}} / \partial \mathbf{p}_{\text{UE}} = \partial h_{n,m} / \partial \mathbf{p}_{\text{UE}} - \partial h_{n_{\text{ref}},m_{\text{ref}}} / \partial \mathbf{p}_{\text{UE}}$, and $\partial \phi_{n,m}^{\text{diff}} / \partial \alpha = \partial h_{n,m} / \partial \alpha - \partial h_{n_{\text{ref}},m_{\text{ref}}} / \partial \alpha$, respectively, when $(n, m) \neq (n_{\text{ref}}, m_{\text{ref}})$. The descriptions of the partial derivatives are presented in Section III-A.

Since also the integer ambiguity differences $\Delta N_{n,m}^{\text{diff}}[k]$ are integer values, the additional EKF update step, described in (38), is also performed for the differential measurements to obtain the a-posteriori estimates of the mean and covariance as $\hat{\mathbf{s}}_{\text{diff}}^{++}[k]$ and $\hat{\mathbf{P}}_{\text{diff}}^{++}[k]$, respectively. Regarding the additional EKF update step with differential measurements, the constant measurement model and the measurement covariance matrix are given as $\mathbf{H}_{\text{N}}^{\text{diff}} = [\mathbf{0}_{(NM-1) \times 12}, \mathbf{I}_{NM-1}, \mathbf{0}_{(NM-1) \times (NM-1)}]$ and $\Sigma_{\text{N}}^{\text{diff}}[k] \in \mathbb{R}^{(NM-1) \times (NM-1)}$ in respective order. Moreover, $\hat{\mathbf{y}}_{\text{N}}^{\text{diff}}[k] = [\hat{y}_{0,0}^{\text{diff}}[k], \dots, \hat{y}_{N,M}^{\text{diff}}[k]]^T \in \mathbb{R}^{MN}$ is the measurement vector including the estimates of rounded integer ambiguity differences, where $\hat{y}_{n,m}^{\text{diff}}[k] = \text{round}(\hat{N}_{n,m}^{\text{diff}}[k])$ for $(n, m) \neq (n_{\text{ref}}, m_{\text{ref}})$. The EKF processing steps illustrated in Fig. 4 are conceptually also valid for the differential measurements.

D. Further Discussion

The key driver towards the proposed EKF processing with integer ambiguity estimation is feasible computational complexity to enable practical real-time processing which can scale up to several user device antennas. The utilized rounding approach to handle the integer ambiguities is simple, but effective when utilized with the EKF through virtual integer ambiguity measurements, as shown in Fig. 4. However, the proposed overall concept with antenna-level measurements allows, as such, estimating integer ambiguities also with more advanced methods, such as the LAMBDA method, with further details available e.g. in [31], [41], [42], [44]. In [41] and [42], a Kalman filter is processed independently using real (float) numbers, and a subsequent estimation method is used separately to resolve the integer ambiguity. In [42], a separate batch weighted nonlinear least-squares method is used to jointly estimate the UE position and real valued integer ambiguities, but eventually both [41] and [42] utilize the LAMBDA method for integer ambiguity solution. Moreover, a mixture Kalman filter is proposed in [44] for elegant, but computationally fairly evolved solution involving particle sampling.

In practical setups with measurement outliers, or with longer measurement intervals, the mixture Kalman filter can provide robustness against cycle slips. Furthermore, by dynamically

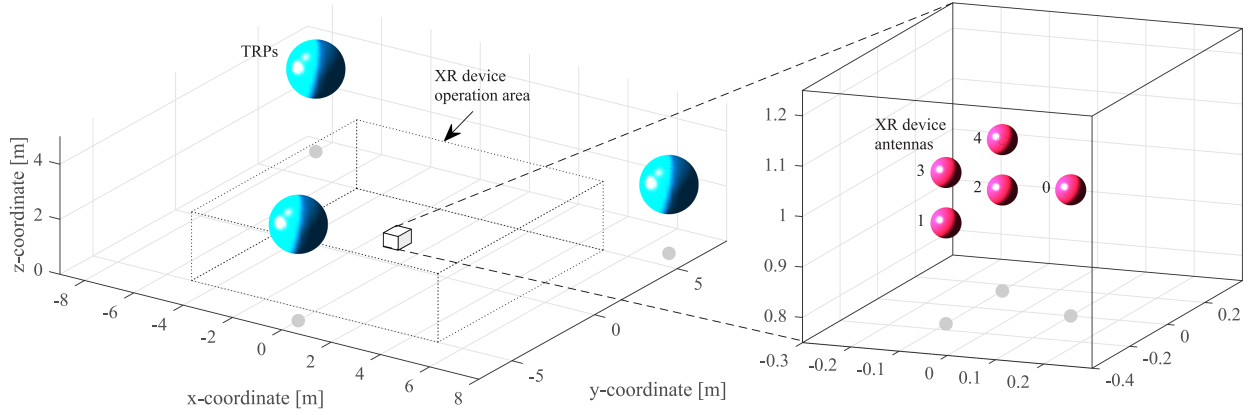


Fig. 5. Illustration of the overall scenario, indicating the 3 TRP positions as well as the mobile XR device operation area (on the left), and the configuration of the antenna elements at the XR device (on the right).

scaling the computational complexity according to underlying estimation uncertainty, for example, by varying the number of used particles, the mixture Kalman filter could provide a favorable trade-off between computational complexity and robustness for the considered XR scenario.

V. NUMERICAL RESULTS

A. Scenario and Assumptions

The performance of the proposed concept and methods is next numerically assessed in an indoor factory like setting, following the 3GPP evaluation guidelines and assumptions stated in [26], [32]. The size of the factory area within which a person wearing the XR headset is moving is limited to $10\text{ m} \times 10\text{ m}$ while both 3.5 GHz and 28 GHz network deployments with three TRPs are considered. In the baseline evaluations, the XR headset is assumed to contain five antenna points, while the number is also later varied. The considered gNB or TRP deployment locations are illustrated in Fig. 5, where the exact TRP positions are $\mathbf{p}_{\text{TRP},0} = [0, -6, 3.5]^T$, $\mathbf{p}_{\text{TRP},1} = [-8, 7, 3]^T$ and $\mathbf{p}_{\text{TRP},2} = [7, 6, 2.5]^T$. In addition to TRP locations, the antenna configuration of the XR device is illustrated for a device position at $\mathbf{p}_{\text{UE}} = [0, 0, 1]^T$. Accordingly, the antenna positions relative to the device position are $\boldsymbol{\rho}_0 = [0.1, 0, 0]^T$, $\boldsymbol{\rho}_1 = [-0.1, -0.15, -0.05]^T$, $\boldsymbol{\rho}_2 = [-0.1, 0.15, -0.05]^T$, $\boldsymbol{\rho}_3 = [-0.1, -0.15, 0.05]^T$ and $\boldsymbol{\rho}_4 = [-0.1, 0.15, 0.05]^T$, comprising the so-called antenna constellation.

To account for different practical aspects and limitations, such as oscillator impairments, antenna reference point errors and antenna phase center offsets, we utilize the extensive carrier phase estimation accuracy results, collected and reported in [26], while also vary the corresponding ranging estimation error variance to understand how the antenna-level ranging performance impacts the corresponding performance of the proposed 6DoF estimation and tracking concept. Furthermore, all numerical results are evaluated based on 100 separate UE test tracks with varying UE position and orientation. Within any individual UE test track, carrier phase measurements are obtained at intervals of $\Delta t = 10\text{ ms}$, which is the duration of one radio frame in 5G NR. In practice, the required update interval depends on the expected maximum velocities (or rate-of-change) of tracked parameters and varies according to different system parameters,

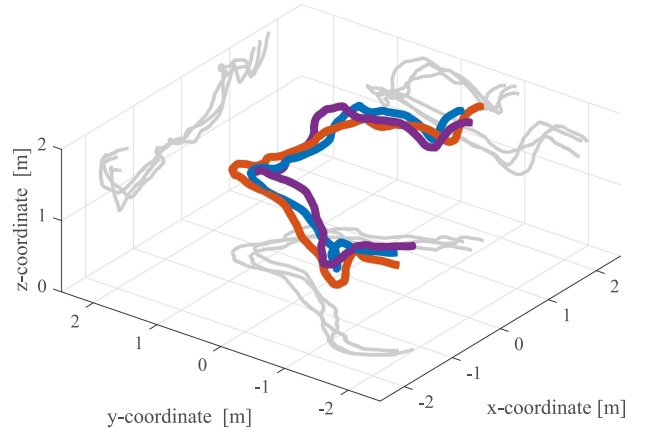


Fig. 6. Example illustration of one simulation track showing the 3D positions of individual UE antenna elements as functions of time. For improved clarity, only the positions of 3 antenna elements out of 5 are shown. The grey curves show the projection of the antenna positions to the XY-, XZ- and YZ-planes.

such as measurement noise. With certain considered simulation setups, larger intervals could be used, especially in ones with low measurement noise. Nonetheless, in 5G NR, it is possible to transmit PRSs or SRSs even with less than 1 ms intervals, and thereon impact the update interval. For the test tracks, the true UE movement and the true evolution of orientation angles are modeled using a non-linear kinematic model, which builds on the 3D-extension of the model presented in [45]. Each user track is 30 s long, and bounded within the $10\text{ m} \times 10\text{ m} \times 2.5\text{ m}$ operation area as illustrated in Fig. 5. Moreover, the values of orientation angles are limited to $-180\text{ deg} \leq \alpha < 180\text{ deg}$ and $-60\text{ deg} \leq \beta, \gamma \leq 60\text{ deg}$ to reflect physically feasible XR headset rotation. Inside a track, the UE velocity and the rate-of-change per orientation angle vary up to 4 km/h and 90 deg/s, respectively. Such a velocity corresponds essentially to human walking speed while the orientation rate of change is well inline with the human head moving rates reported, e.g., in [46]. An example of a considered 3D UE track with varying 3D UE orientation is illustrated in Fig. 6 by showing separate antenna positions as functions of time. For visual clarity, only the positions of 3 antenna elements are shown, instead of all 5. In addition, to assist with perception of 3D space, projections of

the antenna positions are shown in the XY-plane, XZ-plane and YZ-plane.

Assuming uncorrelated measurements, the covariance matrix of the measurement noise is defined as $\Sigma = \sigma_r^2 \mathbf{I}_{MN}$, where σ_r is the standard deviation of the ranging error. By following the reported carrier phase-based positioning and ranging results in [26], we evaluate results for the theoretical 6DoF bounds in Section V-B by varying the ranging error from $\sigma_r = 0.1$ mm to $\sigma_r = 50$ mm. Such fairly large range covers well the individual accuracy results reported and discussed in [26], reflecting different practical assumptions. Moreover, for the 6DoF tracking results in Section V-C, we consider two fixed ranging error deviations, given as $\sigma_r = 50$ mm and $\sigma_r = 5$ mm, which reflect to 3.5 GHz and 28 GHz network deployments, respectively. These stem from the comprehensive ranging performance results provided in [26], serving as realistic practical example scenarios. These ranging standard deviation values are in the order of half the wavelength of the considered frequency bands, hence, one can argue that the ranging accuracy assumptions are even somewhat pessimistic – or at least not overly optimistic and thus safe.

In order to provide statistical variation for the 6DoF tracking results in Section V-C, each track is further run over 100 trials with different measurement noise realizations. In the beginning of each track realization, the used EKF is initialized without assuming any prior information on the UE position, UE orientation, or related motion. The initial UE position and UE orientation as well as their covariance, are obtained based on Gauss-Newton algorithm [34] by exploiting the derived Jacobian matrix in (9). Moreover, the initial UE velocity and the rotational rate-of-change are set to zero with very large covariance. Regarding the state-transition model, the power spectral densities of angular acceleration and UE acceleration of the process covariance $\mathbf{Q}[k]$ are defined as $\sigma_{\Delta\theta}^2 = 9 \cdot 10^{-4}$ and $\sigma_v^2 = 4 \cdot 10^{-4}$, respectively. Throughout the evaluations we assume a known measurement covariance matrix $\Sigma[k]$. However, in practice the measurement covariance can be estimated on the run, or learned in advance based on training data. Finally, the baseline results build on the assumption of known XR headset antenna geometry, however, results related to over-the-air estimation of the antenna constellation are also provided towards the end. Similarly, the extended problem scope with unknown integer ambiguities and drifting UE clock is considered towards the end of the numerical results.

B. 6DoF Bounds

In the following, we first study the measurement-related (no prior) bounds and their behavior as functions of the standard-deviation of the carrier phase -based ranging error, σ_r . The behaviors of the corresponding posterior bounds are illustrated and studied along the actual EKF-based tracking results in the next subsection.

First, CRLBs are evaluated and shown separately for the orientation parameters α , β and γ in Fig. 7. The RMSE curves are obtained based on the derived expressions in (15) as $\sqrt{\text{CRLB}(\hat{\alpha})}$, $\sqrt{\text{CRLB}(\hat{\beta})}$ and $\sqrt{\text{CRLB}(\hat{\gamma})}$, respectively. As can be observed, highly accurate estimation of the orientation parameters is theoretically feasible, with 1° RMSE of an individual parameter calling for approximately 3...5 mm ranging accuracy in the considered system geometry. Based on [26], such ranging

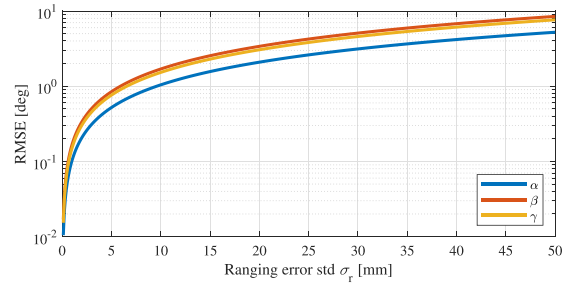


Fig. 7. CRLBs for the estimation of orientation parameters α , β and γ as functions of ranging error standard deviation σ_r .

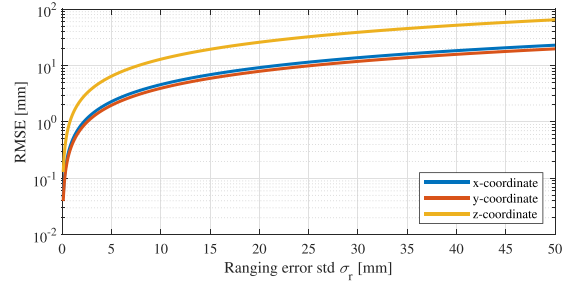


Fig. 8. CRLBs for the estimation of x-coordinate, y-coordinate and z-coordinate as functions of ranging error standard deviation σ_r .

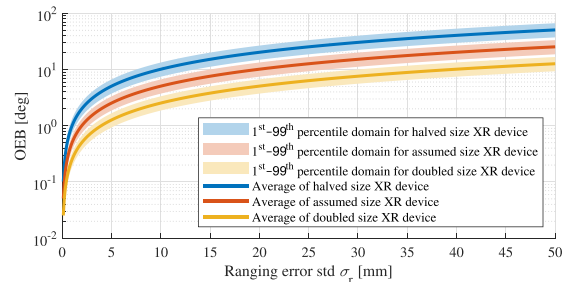


Fig. 9. OEBs as functions of ranging error standard deviation σ_r for three different XR device sizes. Also the corresponding 1st to 99th percentile distribution ranges are illustrated through shaded regions.

accuracies are feasible – especially at the millimeter-wave bands. We can also observe that the estimation of α is somewhat more accurate, theoretically, stemming from the favourable TRP geometry in horizontal plane, in which the α directly operates.

The corresponding no-prior bounds for the x-coordinate, y-coordinate and z-coordinate are illustrated in Fig. 8. The RMSE curves are evaluated based on the derived expressions in (15), as $\sqrt{\text{CRLB}(\hat{x}_{\text{UE}})}$, $\sqrt{\text{CRLB}(\hat{y}_{\text{UE}})}$ and $\sqrt{\text{CRLB}(\hat{z}_{\text{UE}})}$, respectively. Based on the bounds and their numerical behavior, below centimeter level positioning accuracy is theoretically feasible with the 10 mm ranging performance. It can also be observed from the figure that the estimation of the z-coordinate is less accurate, at least to some extent, stemming from the fact that all the network TRPs are deployed at mutually similar antenna heights as described in the previous subsection. Hence, taking this aspect into account in the TRP or gNB deployment is essential to maximize the 6DoF estimation performance.

Next, the average OEB behavior, as defined and expressed in (16), is evaluated and shown in Fig. 9 as a function of the ranging error standard deviation σ_r . Besides the assumed XR

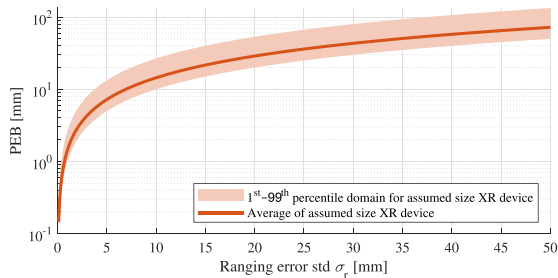


Fig. 10. PEB as a function of ranging error standard deviation σ_r with the default XR device size. Also the corresponding 1st to 99th percentile distribution range is illustrated through shaded region.

device antenna configuration, results include also two additional configurations, where the dimensions of the device and thereon the antenna geometry are either halved or doubled. This provides further insight to the error and performance dependency on the assumed antenna separations at the XR device. The shaded areas in the figure show the OEB domains between the 1st and 99th percentiles over all time steps and test tracks. It can clearly be observed that for a given antenna-level ranging standard deviation, the performance is very stable across the different paths and points along the tracks. We can also observe that in order to reach a true one degree 3D orientation RMSE, the ranging accuracy should be just a few millimeters. The figure also clearly shows that increasing the device size has a clear positive impact on the 3D orientation estimation accuracy, due to the improved antenna geometry.

Finally, the average PEB behavior, as defined and expressed in (17), is evaluated and shown in Fig. 10 while again varying the ranging error standard deviation σ_r . Regarding the PEB, the RMSE values with half-sized array or double-sized array are practically identical for the considered system configuration, and are thus not shown in the figure. The results in Fig. 10 again highlight the consistency of the estimation performance along the tracks and points of the individual tracks, assuming a given ranging standard deviation. Concretely, in order for the 3D location RMSE to be below 1 cm with high probability, antenna ranging accuracies in the order of 5 mm are required in the considered scenario.

C. 6DoF Tracking Results

Next, we study and present the performance of the proposed EKF and the corresponding posterior bounds derived and expressed in (20). We consider two alternative ranging error standard deviations, $\sigma_r = 5$ mm and $\sigma_r = 50$ mm = 5 cm, chosen based on the extensive carrier phase estimation studies in [26], and corresponding to 28 GHz and 3.5 GHz network deployments, respectively. As noted earlier, the results build on 100 independent UE tracks with time-varying position and orientation, while each track is also sampled 100 times with different measurement noise realizations for statistical purposes.

First, the average orientation estimation RMSEs for the proposed EKF as functions of time are shown in Fig. 11 with the two considered ranging error standard deviations. The figure also shows the corresponding posterior OEBs for reference. It can be clearly observed that after the initial transient, the proposed EKF yields performance very close to the respective posterior bound, hence implying an efficient estimator. Application-wise, it can

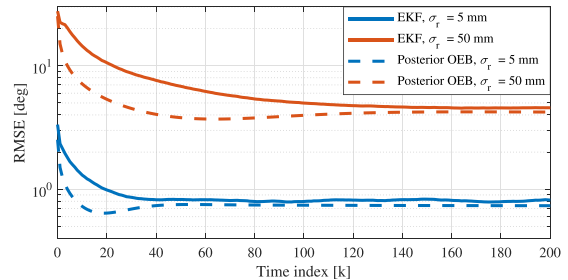


Fig. 11. Average 3D orientation estimation RMSE for the proposed EKF and the corresponding posterior OEB with the two considered ranging error standard deviations.

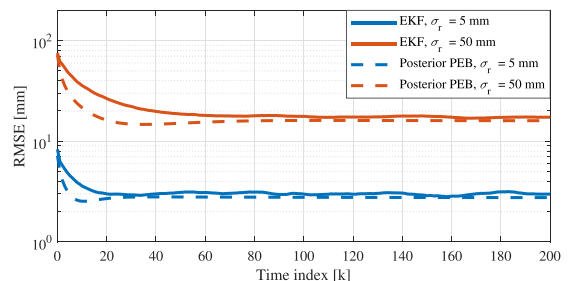


Fig. 12. Average 3D positioning RMSE for the proposed EKF and the corresponding posterior PEB with the two considered ranging error standard deviations.

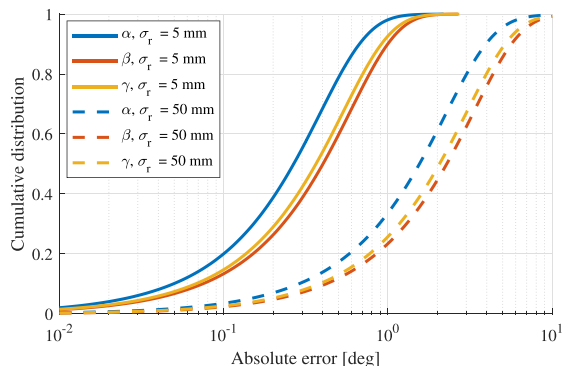


Fig. 13. Cumulative distributions of the absolute estimation error for orientation angles α , β and γ with the two considered ranging error standard deviations.

also be observed that below one degree 3D orientation tracking is feasible, especially at the millimeter-wave frequencies.

Similarly, the average positioning RMSEs for the proposed EKF as functions of time are shown in Fig. 12, together with the applicable posterior PEBs, for the two considered ranging error standard deviations. The results highlight again the efficiency of the EKF solution, while also evidence close to one centimeter 3D location tracking RMSE at the millimeter-waves.

The actual cumulative distributions of the absolute estimation errors with the proposed EKF for orientation angles α , β and γ are shown in Fig. 13, again considering the two antenna-level ranging error standard deviations of 5 mm and 50 mm. Similarly, the corresponding cumulative distributions of the absolute estimation errors for x-coordinate, y-coordinate and z-coordinate are shown in Fig. 14. The cumulative distributions demonstrate, overall, highly consistent tracking performance across the tracks and noise realizations. Additionally, similar to

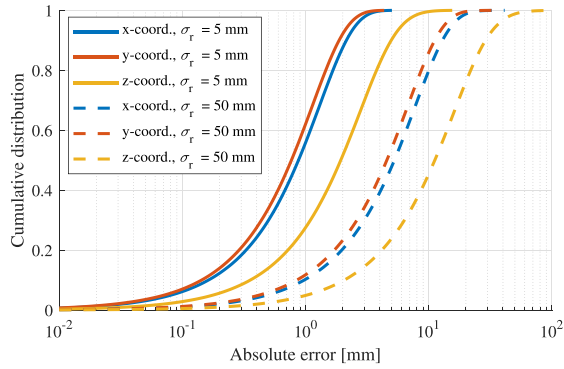


Fig. 14. Cumulative distributions of the absolute estimation error for x-coordinate, y-coordinate and z-coordinate with the two considered ranging error standard deviations.

TABLE I

AVERAGE RMSEs FOR 3D POSITION AND 3D ORIENTATION ESTIMATION, TOGETHER WITH THE CORRESPONDING BOUNDS AT FR1 (3.5 GHz, 50 MM RANGING STD) AND FR2 (28 GHz, 5 MM RANGING STD)

RMSE	FR1		FR2	
	Static	EKF	Static	EKF
Orientation estimate [deg]	27.7	4.79	4.80	0.83
OEB [deg]	25.4	4.21	4.30	0.75
Position estimate [mm]	75.2	18.1	18.1	3.05
PEB [mm]	72.1	16.3	16.4	2.83

the no-prior performance findings in the previous subsection, the EKF performance is relatively better for α , compared to other orientation parameters, due to the more favourable horizontal geometry of the TRPs. The distributions also show that tracking the device z-coordinate is more challenging, as noted earlier, due to the fact that the antennas of all three TRPs are located essentially at the same height.

Finally, we summarize selected key performance measures in Table I, covering both the frequency range 1 (FR1, 3.5 GHz) and FR2 (28 GHz) deployments. The table shows the 3D orientation and 3D position RMSEs with the proposed EKF formulation together with the RMSEs of the corresponding static estimator which refers to the Gauss-Newton based estimation utilized as the EKF initialization approach. Additionally, the respective bounds are shown for comparison purposes. The results in the table clearly high-light the added value of a true tracking solution that incorporates information also from the preceding estimates. Specifically, as shown by the results in Table I, essentially one order of magnitude RMSE improvements can be obtained through EKF-based tracking compared to static estimation. Overall, the results demonstrate the feasibility of extremely accurate 6DoF XR device tracking, with below one degree and below one centimeter orientation and position RMSEs available at the millimeter-wave bands.

Besides the proposed EKF-based approach, corresponding CKF and UKF based solutions were also implemented and tested. However, the obtained results with CKF and UKF are practically identical with the EKF, and thus they are omitted from the numerical results shown in this article. It should be emphasized that under the condition of similar performance between EKF, CKF and UKF, the EKF is preferred from the computational complexity perspective.

TABLE II

AVERAGE EKF TRACKING RMSEs WITH REDUCED HEADSET ANTENNA COUNTS FOR 3D POSITION AND 3D ORIENTATION ESTIMATION, TOGETHER WITH CORRESPONDING BOUNDS AT FR1 (3.5 GHz, 50 MM RANGING STD) AND FR2 (28 GHz, 5 MM RANGING STD)

RMSE	FR1			FR2		
	3	4	5	3	4	5
Orientation estimate [deg]	6.54	5.83	4.79	1.14	1.06	0.83
OEB [deg]	5.89	5.11	4.21	1.03	0.95	0.75
Position estimate [mm]	22.2	19.7	18.1	3.76	3.59	3.05
PEB [mm]	20.1	17.9	16.3	3.47	3.29	2.83

D. Results With Reduced Headset Antenna Counts

Next, we shortly study how reducing the antenna count in the XR headset, from 5 down to 4 and 3, impacts the 6DoF tracking performance. For the considered 4 antenna scenario, we employ the antenna elements with indices 0, 1, 2 and 3 from the original antenna constellation which was earlier illustrated in Fig. 5. Furthermore, for the 3 antenna scenario we remove yet an additional element, and thus employ only the antenna elements with indices 0, 1 and 2.

The average EKF tracking RMSEs of 3D positioning and 3D orientation estimation with reduced antenna counts are shown in Table II for both the FR1 (3.5 GHz) and FR2 (28 GHz) deployments. In addition to the EKF tracking results, the corresponding performance bounds (PEB, OEB) are again provided. All results are evaluated based on the same test tracks and parameters as with the original 5 antenna setup. According to the shown RMSEs, it is visible that reducing the antenna count at the FR2 deployment has a rather small effect. Specifically, when reducing the antenna count from 5 to 3, the EKF RMSE for the 3D orientation increases from 0.83 deg to 1.14 deg, and correspondingly for the 3D position from 3.05 mm to 3.76 mm. However, the performance drop is more notable at FR1, where the RMSE of 3D orientation estimation is roughly doubled, and the RMSE of 3D positioning almost tripled, when reducing the antenna count from 5 to 3. The reason for this is that in the FR1 deployment, the assumed 50 mm ranging error standard deviation inflicts antenna-level measurement errors that are already at the level of antenna separation distances, which in turn introduces challenges to the estimator when the antenna count is reduced towards to the critical number of 3 antennas. Nevertheless, the proposed EKF is still able to achieve RMSEs close to the theoretical bounds, and thus proving its feasibility to the considered tracking of XR device orientation and location even with only 3 headset antenna points. These findings also clearly motivate for clever XR headset design in terms of optimizing the antenna positions.

E. Results With OTA Calibration of Headset Antennas

We next demonstrate and evaluate the proposed over-the-air estimation or calibration of the XR headset antenna constellation. The evaluations are performed using the original antenna constellation with 5 antenna elements according to Fig. 5, including the four different OTA estimation approaches, described in Section III-D. The RMSEs of the antenna constellation estimates as functions of time for the four considered approaches are shown in Fig. 15. For all the approaches, we assume the FR2 deployment with 5 mm ranging standard deviation, and

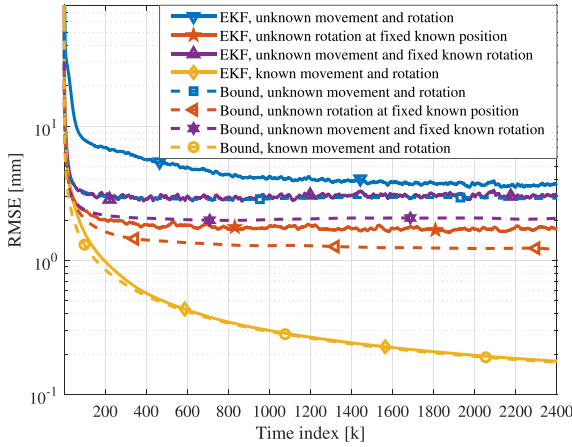


Fig. 15. Average 3D positioning RMSE and the corresponding posterior bounds for the online OTA calibration of the headset antenna constellation with four different approaches, while assuming 5 mm ranging standard deviation.

evaluate the performances by exploiting the same test tracks and orientations as described in Section V-A. Besides the EKF tracking results, the corresponding theoretical posterior bounds are derived and shown to illustrate the EKF performance with respect to an estimator with optimum performance.

In the first approach, namely the *Unknown Movement and Rotation* case, the XR device is moving and rotating in the operation area, while the EKF tracks jointly the device position and device orientation as well as the antenna constellation. It should be emphasized that during the tracking process the estimated parameters do not reflect the true parameters due to the implicit relation between the position, orientation and relative antenna coordinates. Thus, the accuracy of the antenna constellation estimation can be evaluated only after the estimated antenna constellation is transformed to the original frame of reference, as discussed and defined in (34). As seen in Fig. 15, from the four considered approaches, the approach with unknown movement and rotation is the most challenging, reaching to antenna constellation estimation RMSE of below 4 mm. However, it can also be seen that the proposed EKF gets close to the posterior performance bound found to be approximately at 3 mm at the end of the tracking process.

In the second approach of estimating the antenna constellation, namely the *Unknown Rotation at a Fixed Known Position* scenario, the XR device is rotating at a fixed known position, while the EKF tracks jointly the device orientation and the antenna constellation. For these evaluations, the XR device is located at a fixed position in the center of the area ($\mathbf{p}_{\text{UE}} = [0, 0, 1.5]^T$), while the orientation changes according to the orientation of the earlier considered test tracks. As seen in Fig. 15, compared to the first approach, the estimation RMSE is improved due to the reduced process covariance through known device position, reaching antenna constellation accuracy of below 2 mm already. In the third approach, namely the *Unknown movement and fixed known rotation*, the orientation is known instead and fixed at $\boldsymbol{\theta} = [0, 0, 0]^T$, while the device is moving with unknown position. Due to an interchangeable relationship between the antenna constellation and the device position, this approach results in lower estimation accuracy compared to the one with known position, eventually reaching around 3 mm accuracy.

Whereas the first three approaches can be considered more essential and relevant for practical scenarios and applications, the fourth approach, namely the *Known Movement and Rotation* reference case, works as a performance benchmark for estimating the antenna constellation. In this approach, similar to the first and third ones, the XR device is moving and rotating in the operation area, but assuming now that the device position and orientation are perfectly known at each time instant. Thus, the EKF tracks now only the unknown antenna constellation. As seen in Fig. 15, the RMSE of antenna constellation estimation reaches accuracies that are only fractions of a millimeter. Moreover, unlike with the other considered OTA calibration approaches, the estimation accuracy is now able to continuously improve in time due to the absence of process noise related to the device position and/or the orientation.

F. Results With Integer Ambiguity and UE Clock Drifting

For the results with integer ambiguity and clock bias, as well as with differential carrier phase measurements, we consider the same general parameters described in Section V-A, except that the filter update interval is reduced to 1 ms to support tracking of integer ambiguities. The standard deviation of the clock skew noise $\xi[k]$ is defined as $\sigma_{\xi} = 10^{-10}$, which can result in clock variations corresponding to tens of meters of ranging error during one track. Moreover, the considered true integer ambiguities stem directly from the time-varying position, orientation and clock of the UE. For the standard deviation of carrier phase measurement error we vary three different values, including $\sigma_{\nu} = 0.01\lambda$ m, $\sigma_{\nu} = 0.05\lambda$ m, and $\sigma_{\nu} = 0.1\lambda$ m, where λ is the carrier wavelength. These correspond to standard deviations of carrier phase errors of 3.6 deg, 18 deg, and 36 deg in respective order.

In the EKF process model, the additional parameters related to clock drift and integer ambiguities have been chosen based on brief testing over different parameter combinations. Regarding the approach with tracking the clock and integer ambiguities, we define the power spectrum density of clock acceleration as $\sigma_{\Delta\delta}^2 = 10^{-20}$. Furthermore, the power spectrum density of process acceleration for integer ambiguity is defined $\sigma_{\Delta N^{\text{int}}}^2 = 10^8$ for FR1, and $\sigma_{\Delta N^{\text{int}}}^2 = 10^{10}$ for FR2. The relatively large process noise is adopted in order to allow fast integer updates with the considered simple rounding scheme. However, the process noise magnitude can be considerably decreased when using more advanced integer resolution methods, such as the LAMBDA method. Then, regarding the approach with differential carrier phase measurements, the power spectrum density of process acceleration for integer ambiguity difference is defined $\sigma_{\Delta N^{\text{diff}}}^2 = 10^{10}$ for FR1, and $\sigma_{\Delta N^{\text{diff}}}^2 = 10^{12}$ for FR2. The covariance matrix for the virtual integer ambiguity measurements in the additional EKF update step is given as $\Sigma_N^{\text{int}}[k] = 10^{-8} \cdot \mathbf{I}_{NM}$ for tracking with clock, and $\Sigma_N^{\text{diff}}[k] = 10^{-8} \cdot \mathbf{I}_{NM-1}$ for differential measurements.

For all results with integer ambiguity, we initialize the EKF using differential measurements at FR1 (defined in (39)) and a non-linear least squares estimation based on the Gauss-Newton method with multiple initial starting points for the iterations. With the differential measurements, the initial clock error can be removed, allowing accurate UE position estimation. The initialization of the EKF is accomplished using 100 samples before proceeding to the proposed EKF processing steps, shown in Fig. 4. In practice, the computational burden and latency

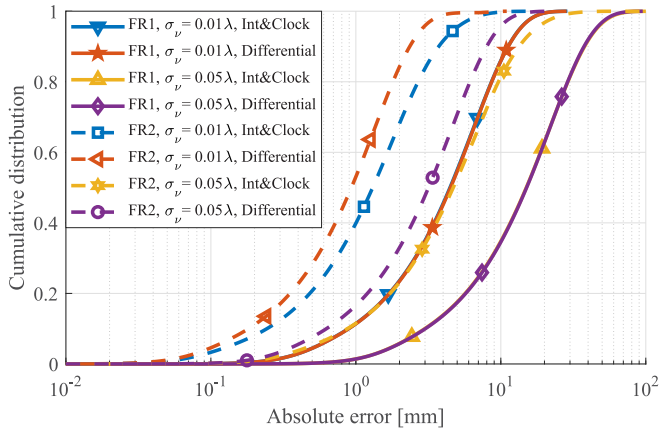


Fig. 16. Cumulative distributions of the absolute 3D positioning errors for tracking with integer ambiguities and unknown clock, as well as for tracking with differential measurements.

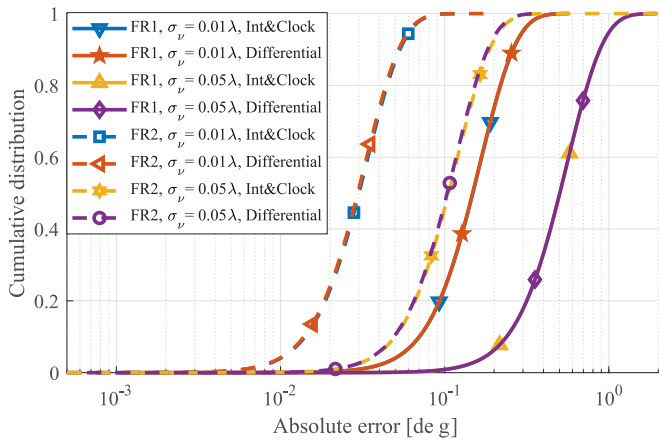


Fig. 17. Cumulative distributions of the absolute 3D orientation estimation errors for tracking with integer ambiguities and unknown clock, as well as for tracking with differential measurements.

of such brute force approach could be mitigated by exploiting additional initial positioning methods supported in 5G NR, such as multi-round-trip-time (Multi-RTT), time-difference-of-arrival (TDoA), or AoA.

In Fig. 16, cumulative distributions of absolute 3D positioning errors are presented for both FR1 and FR2, considering tracking with integer ambiguities and unknown clock as well as two separate standard deviations of carrier phase measurement error. When comparing the positioning accuracy between FR1 and FR2, at FR2 higher accuracy is achieved due to the shorter wavelength. At FR1, EKF-based tracking with integers and clock (“Int&Clock”), presented in Section IV-B, performs similarly compared to the corresponding tracking with differential measurements (“Differential”), presented in Section IV-C. However, at FR2, tracking with integers and clock becomes more challenging due to more frequent integer ambiguity cycle slips. The 90% error percentiles with measurement standard deviation of $\sigma_\nu = 0.01\lambda$ are below 12 mm at FR1 and below 4 mm at FR2, while with $\sigma_\nu = 0.05\lambda$, below 40 mm at FR1 and below 14 mm at FR2. The corresponding cumulative distributions for absolute 3D orientation errors are shown in Fig. 17. On contrary to the positioning results, there are no visible differences between

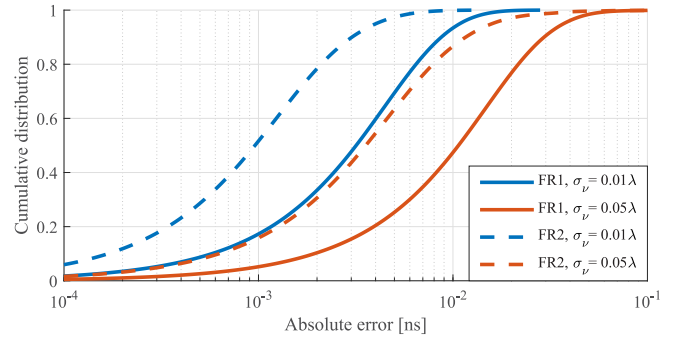


Fig. 18. Cumulative distributions of the absolute clock estimation error for tracking with integer ambiguities and unknown clock.

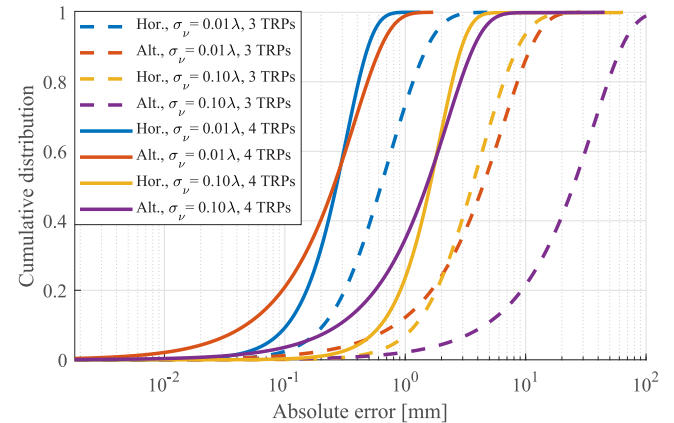


Fig. 19. Cumulative distributions of the absolute 2D horizontal errors and altitude errors for 3 and 4 TRPs, while assuming tracking with integer ambiguities and unknown clock at FR1.

tracking with integers and clock vs. tracking with differential measurements. This is due to fact that orientation estimation is based on relative measurements between different antennas, and is thus insensitive to common clock error between the antennas. With measurement standard deviation of $\sigma_\nu = 0.01\lambda$, the 90% error percentiles are below 0.27 deg at FR1 and below 0.06 deg at FR2, whereas with $\sigma_\nu = 0.05\lambda$, the percentiles are below 0.89 deg at FR1 and below 0.19 deg at FR2.

Cumulative distributions of the clock estimation error are shown in Fig. 18 for the tracking with integers and clock. Again, the estimation accuracy at FR2 is higher compared to FR1. Considering the measurement standard deviation of $\sigma_\nu = 0.05\lambda$, the 90% error percentiles for the clock estimation are below 0.03 ns for FR1 and below 0.02 ns at FR2.

As shown earlier in Fig. 14, the positioning performance is limited due to poor TRP geometry from the z-coordinate (altitude) point of view. Thus, for final evaluation, we introduce an additional TRP to the position $\mathbf{p}_{\text{TRP},3} = [0, 0, 4]^T$, which can be considered to be placed on the ceiling in the middle of the operation area. In Fig. 19, cumulative distributions of positioning errors are compared between 3 TRPs and 4 TRPs for tracking with integers and clock (“Int&Clock”), while assuming FR1 and two measurement standard deviations of $\sigma_\nu = 0.01\lambda$ and 0.10λ . Moreover, the cumulative distributions are shown separately for horizontal 2D error defined in xy-plane (“Hor.”), and altitude error defined in direction of z-axis (“Alt.”). From the figure, it can be seen that with 3 TRPs the altitude error is

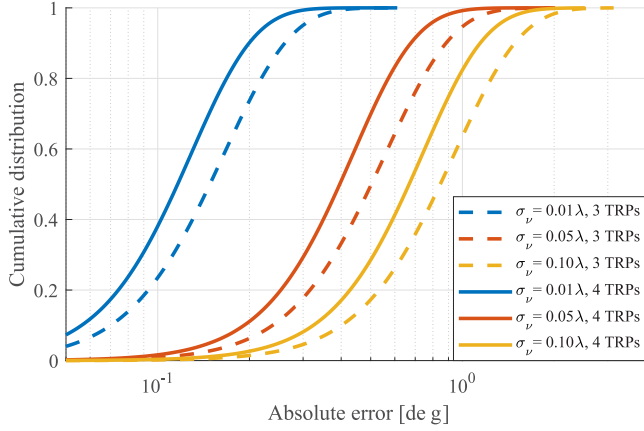


Fig. 20. Cumulative distributions of the absolute 3D orientation estimation errors for 3 and 4 TRPs, while assuming tracking with integer ambiguities and unknown clock at FR1.

considerably larger compared to the horizontal error, especially with $\sigma_\nu = 0.10\lambda$. However, with 4 TRPs the horizontal and altitude errors are at the same range, and most importantly, there is a significant improvement in the altitude estimation accuracy compared to the original 3 TRP case. In addition, with 4 TRPs the 90% error percentiles for both the horizontal and altitude errors reach below 4 mm with $\sigma_\nu = 0.10\lambda$. Finally, in Fig. 20, the corresponding cumulative distributions for 3D orientation error are shown for the 3 TRP and 4 TRP scenarios. It can be seen that introducing the fourth TRP does not provide similar increase in accuracy with orientation estimation as with positioning. Thus, orientation estimation accuracy is not as sensitive to the TRP geometry as positioning accuracy. Nonetheless, overall, when applying 6DoF tracking in the considered 5G NR scenario, it is essential to ensure an appropriate TRP geometry.

VI. CONCLUSION

In this article, the so-called six degrees-of-freedom tracking challenge of XR headset was addressed, in the context of 5G and beyond cellular networks. The proposed approach builds on antenna-level carrier phase measurements, obtained through

either uplink and downlink reference signals, combined with Bayesian filtering in the form of EKF to continuously track the XR headset 3D orientation and 3D location. The EKF formulation also allows for tracking the device velocity and orientation rate-of-change that can provide added value in mobile XR applications, e.g., for viewpoint and/or field-of-view prediction and proactive rendering with minimum latency when operating in the uplink-based mode. Also applicable no-prior and posterior Cramér-Rao lower-bounds were stated and derived to assess the 6DoF performance limits. Furthermore, over-the-air estimation concepts and algorithms were proposed for acquiring the XR headset antenna geometry. Additionally, extended problem definitions and corresponding EKF solutions were provided to account for the important integer ambiguity and UE clock drifting challenges. An extensive set of numerical results were provided, evaluating and assessing the performance of the proposed methods and the corresponding performance bounds, covering results for both 3.5 GHz and 28 GHz network deployments. To obtained results clearly demonstrate that especially in millimeter-wave networks, the proposed concept can facilitate highly accurate 3D orientation and 3D location tracking – with the best numerical examples showing RMSE accuracies below one degree and below one centimeter, respectively. Also, estimating the XR headset antenna constellation at RMSEs closing towards one millimeter as well as accurate tracking of the carrier phase integer ambiguities and UE clock drifting were shown to be technically feasible.

APPENDIX

DEFINITIONS OF 3D ROTATION MATRICES

We consider the 3D rotation matrix $\mathbf{R}(\boldsymbol{\theta}) \in \mathbb{R}^{3 \times 3}$ defined in [32], and reproduced in (41), shown at the bottom of this page applying rotation in z-, y- and x-axis in respective order according to the rotation angles $\boldsymbol{\theta} = [\alpha, \beta, \gamma]^T$, where α, β and γ are the yaw, pitch and roll. The partial derivatives of the rotation matrix $\mathbf{R}(\boldsymbol{\theta})$ with respect to distinct rotation angles α, β and γ , given as $\mathcal{R}_\alpha(\boldsymbol{\theta}) = \partial \mathbf{R}(\boldsymbol{\theta}) / \partial \alpha$, $\mathcal{R}_\beta(\boldsymbol{\theta}) = \partial \mathbf{R}(\boldsymbol{\theta}) / \partial \beta$ and $\mathcal{R}_\gamma(\boldsymbol{\theta}) = \partial \mathbf{R}(\boldsymbol{\theta}) / \partial \gamma$, are stated in (42) shown at the bottom of this page, (43) and (44) shown at the bottom of this page, in respective order.

$$\mathbf{R}(\boldsymbol{\theta}) = \begin{bmatrix} \cos(\alpha) \cos(\beta) & \cos(\alpha) \sin(\beta) \sin(\gamma) - \sin(\alpha) \cos(\gamma) & \cos(\alpha) \sin(\beta) \cos(\gamma) + \sin(\alpha) \sin(\gamma) \\ \sin(\alpha) \cos(\beta) & \sin(\alpha) \sin(\beta) \sin(\gamma) + \cos(\alpha) \cos(\gamma) & \sin(\alpha) \sin(\beta) \cos(\gamma) - \cos(\alpha) \sin(\gamma) \\ -\sin(\beta) & \cos(\beta) \sin(\gamma) & \cos(\beta) \cos(\gamma) \end{bmatrix} \quad (41)$$

$$\mathcal{R}_\alpha(\boldsymbol{\theta}) = \begin{bmatrix} -\sin(\alpha) \cos(\beta) & -\sin(\alpha) \sin(\beta) \sin(\gamma) - \cos(\alpha) \cos(\gamma) & -\sin(\alpha) \sin(\beta) \cos(\gamma) + \cos(\alpha) \sin(\gamma) \\ \cos(\alpha) \cos(\beta) & \cos(\alpha) \sin(\beta) \sin(\gamma) - \sin(\alpha) \cos(\gamma) & \cos(\alpha) \sin(\beta) \cos(\gamma) + \sin(\alpha) \sin(\gamma) \\ 0 & 0 & 0 \end{bmatrix} \quad (42)$$

$$\mathcal{R}_\beta(\boldsymbol{\theta}) = \begin{bmatrix} -\cos(\alpha) \sin(\beta) & \cos(\alpha) \cos(\beta) \sin(\gamma) & \cos(\alpha) \cos(\beta) \cos(\gamma) \\ -\sin(\alpha) \sin(\beta) & \sin(\alpha) \cos(\beta) \sin(\gamma) & \sin(\alpha) \cos(\beta) \cos(\gamma) \\ -\cos(\beta) & -\sin(\beta) \sin(\gamma) & -\sin(\beta) \cos(\gamma) \end{bmatrix} \quad (43)$$

$$\mathcal{R}_\gamma(\boldsymbol{\theta}) = \begin{bmatrix} 0 & \cos(\alpha) \sin(\beta) \cos(\gamma) + \sin(\alpha) \sin(\gamma) & \sin(\alpha) \cos(\gamma) - \cos(\alpha) \sin(\beta) \sin(\gamma) \\ 0 & \sin(\alpha) \sin(\beta) \cos(\gamma) - \cos(\alpha) \sin(\gamma) & -\sin(\alpha) \sin(\beta) \sin(\gamma) - \cos(\alpha) \cos(\gamma) \\ 0 & \cos(\beta) \cos(\gamma) & -\cos(\beta) \sin(\gamma) \end{bmatrix} \quad (44)$$

REFERENCES

- [1] H. Holma, A. Toskala, and T. Nakamura, *5G Technology: 3GPP New Radio*. Hoboken, NJ, USA: Wiley, 2019.
- [2] E. Dahlman, S. Parkvall, and J. Skold, *5G NR: The Next Generation Wireless Access Technology*, 2nd ed. Amsterdam, The Netherlands: Elsevier, 2020.
- [3] F. Tang, X. Chen, M. Zhao, and N. Kato, "The roadmap of communication and networking in 6G for the metaverse," *IEEE Wireless Commun.*, early access, Jun. 24, 2022, doi: [10.1109/MWC.019.2100721](https://doi.org/10.1109/MWC.019.2100721).
- [4] 3rd Generation Partnership Project (3GPP), "Extended reality (XR) in 5G," 3GPP, Sophia Antipolis, France, Tech. Rep. 26.928 V18.0.0, Mar. 2023. [Online]. Available: <http://www.3gpp.org>
- [5] S. Gül et al., "Kalman filter-based head motion prediction for cloud-based mixed reality," in *Proc. ACM Int. Conf. Multimedia*, 2020, pp. 3632–3641.
- [6] A. Clemm, M. T. Vega, H. K. Ravuri, T. Wauters, and F. D. Turck, "Toward truly immersive holographic-type communication: Challenges and solutions," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 93–99, Jan. 2020.
- [7] R. Ju et al., "Ultra wide view based panoramic VR streaming," in *Proc. ACM Workshop Virtual Reality Augmented Reality Netw.*, 2017, pp. 19–23.
- [8] R. Mendrzik, F. Meyer, G. Bauch, and M. Z. Win, "Enabling situational awareness in millimeter wave massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 1196–1211, Sep. 2019.
- [9] N. Garcia, H. Wymeersch, E. G. Larsson, A. M. Haimovich, and M. Coulon, "Direct localization for massive MIMO," *IEEE Trans. Signal Process.*, vol. 65, no. 10, pp. 2475–2487, May 2017.
- [10] 3rd Generation Partnership Project (3GPP), "Study on NR positioning support (Release 16)," 3GPP, Sophia Antipolis, France, Tech. Rep. 38.855 V16.0.0, Mar. 2019. [Online]. Available: <http://www.3gpp.org>
- [11] 3rd Generation Partnership Project (3GPP), "Service requirements for the 5G system (Release 19)," 3GPP, Sophia Antipolis, France, Tech. Specification 22.261 V19.1.0, Dec. 2022. [Online]. Available: <http://www.3gpp.org>
- [12] J. Jang, Y. Jo, M. Shin, and J. Paik, "Camera orientation estimation using motion-based vanishing point detection for advanced driver-assistance systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 10, pp. 6286–6296, Oct. 2021.
- [13] R. Frohlich, L. Tamas, and Z. Kato, "Absolute pose estimation of central cameras using planar regions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 2, pp. 377–391, Feb. 2021.
- [14] M. Nazarahari and H. Rouhani, "40 years of sensor fusion for orientation tracking via magnetic and inertial measurement units: Methods, lessons learned, and future challenges," *Inf. Fusion*, vol. 68, pp. 67–84, 2021.
- [15] J. Wu, Z. Zhou, H. Fourati, and Y. Cheng, "A super fast attitude determination algorithm for consumer-level accelerometer and magnetometer," *IEEE Trans. Consum. Electron.*, vol. 64, no. 3, pp. 375–381, Aug. 2018.
- [16] W. Zheng and N. González-Prelcic, "Joint position, orientation and channel estimation in hybrid mmWave MIMO systems," in *Proc. IEEE 53rd Asilomar Conf. Signals, Syst., Comput.*, 2019, pp. 1453–1458.
- [17] 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.
- [18] A. Shahmansoori, B. Uguen, G. Destino, G. Seco-Granados, and H. Wymeersch, "Tracking position and orientation through millimeter wave lens MIMO in 5G systems," *IEEE Signal Process. Lett.*, vol. 26, no. 8, pp. 1222–1226, Aug. 2019.
- [19] 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.
- [20] Z. Abu-Shaban, H. Wymeersch, T. Abhayapala, and G. Seco-Granados, "Single-anchor two-way localization bounds for 5G mmWave systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 6, pp. 6388–6400, Jun. 2020.
- [21] J. Talvitie, M. Koivisto, T. Levanen, M. Valkama, G. Destino, and H. Wymeersch, "High-accuracy joint position and orientation estimation in sparse 5G mmWave channel," in *Proc. IEEE Int. Conf. Commun.*, 2019, pp. 1–7.
- [22] J. Talvitie, M. Valkama, G. Destino, and H. Wymeersch, "Novel algorithms for high-accuracy joint position and orientation estimation in 5G mmWave systems," in *Proc. IEEE Globecom Workshops*, 2017, pp. 1–7.
- [23] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch, "Position and orientation estimation through millimeter-wave MIMO in 5G systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1822–1835, Mar. 2018.
- [24] M. Kok and T. B. Schön, "A fast and robust algorithm for orientation estimation using inertial sensors," *IEEE Signal Process. Lett.*, vol. 26, no. 11, pp. 1673–1677, Nov. 2019.
- [25] W. Ding, W. Sun, Y. Gao, and J. Wu, "Carrier phase-based precise heading and pitch estimation using a low-cost GNSS receiver," *Remote Sens.*, vol. 13, no. 18, 2021, Art. no. 3642.
- [26] 3rd Generation Partnership Project (3GPP), TR 38.859 V18.0.0, "Study on expanded and improved NR positioning (Release 18)," Dec. 2022. [Online]. Available: <http://www.3gpp.org>
- [27] M. Säily, O. N. C. Yilmaz, D. S. Michalopoulos, E. Pérez, R. Keating, and J. Schaepperle, "Positioning technology trends and solutions toward 6G," in *Proc. IEEE 32nd Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun.*, 2021, pp. 1–7.
- [28] 3rd Generation Partnership Project (3GPP), "Expanded and improved NR positioning work item description (Release 18)," 3GPP, Sophia Antipolis, France, Tech. Rep. RP-223549, Dec. 2022. [Online]. Available: <http://www.3gpp.org>
- [29] O. L. A. López, D. Kumar, and A. Tölli, "Polarization diversity-enabled LOS/NLOS identification via carrier phase measurements," *IEEE Trans. Commun.*, vol. 71, no. 3, pp. 1678–1690, Mar. 2023.
- [30] L. Chen, X. Zhou, F. Chen, L.-L. Yang, and R. Chen, "Carrier phase ranging for indoor positioning with 5G NR signals," *IEEE Internet Things J.*, vol. 9, pp. 10908–10919, Jul. 2022.
- [31] S. Fan, W. Ni, H. Tian, Z. Huang, and R. Zeng, "Carrier phase-based synchronization and high-accuracy positioning in 5G new radio cellular networks," *IEEE Trans. Commun.*, vol. 70, no. 1, pp. 564–577, Jan. 2022.
- [32] 3rd Generation Partnership Project (3GPP), "Study on channel model for frequency spectrum from 0.5 to 100 GHz (Release 17)," 3GPP, Sophia Antipolis, France, Tech. Rep. 38.901 V17.0.0, Mar. 2022. [Online]. Available: <http://www.3gpp.org>
- [33] A. Kim and M. F. Golnaraghi, "A quaternion-based orientation estimation algorithm using an inertial measurement unit," in *Proc. IEEE Position Location Navigation Symp.*, 2004, pp. 268–272.
- [34] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, 1st ed., vol. 1. Englewood Cliffs, NJ, USA: Prentice Hall, Apr. 1993.
- [35] 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.
- [36] T. K. Y. Bar-Shalom and X. Rong Li, *Estimation With Applications to Tracking and Navigation: Theory, Algorithms and Software*. Hoboken, NJ, USA: Wiley, 2001.
- [37] S. Särkkä, *Bayesian Filtering and Smoothing, Ser. Institute of Mathematical Statistics Textbooks*. Cambridge, U.K.: Cambridge Univ. Press, 2013.
- [38] E. Bastug, M. Bennis, M. Medard, and M. Debbah, "Toward interconnected virtual reality: Opportunities, challenges, and enablers," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 110–117, Jun. 2017.
- [39] M. Koivisto et al., "Joint device positioning and clock synchronization in 5G ultra-dense networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2866–2881, May 2017.
- [40] H. Kim, X. Ma, and B. R. Hamilton, "Tracking low-precision clocks with time-varying drifts using Kalman filtering," *IEEE/ACM Trans. Netw.*, vol. 20, no. 1, pp. 257–270, Feb. 2012.
- [41] D. B. Cox Jr. and J. D. W. Brading, "Integration of LAMBDA ambiguity resolution with Kalman filter for relative navigation of spacecraft," *Navigation*, vol. 47, no. 3, pp. 205–210, 2000.
- [42] J. Khalife and Z. M. Kassas, "Differential framework for submeter-accurate vehicular navigation with cellular signals," *IEEE Trans. Intell. Veh.*, vol. 8, no. 1, pp. 732–744, Jan. 2023.
- [43] P. Teunissen, "The least-squares ambiguity decorrelation adjustment: A method for fast GPS integer ambiguity estimation," *J. Geodesy*, vol. 70, pp. 65–82, Nov. 1995.
- [44] K. Berntorp, A. Weiss, and S. D. Cairano, "Integer ambiguity resolution by mixture Kalman filter for improved GNSS precision," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 4, pp. 3170–3181, Aug. 2020.
- [45] T. Bailey, J. Nieto, J. Guivant, M. Stevens, and E. Nebot, "Consistency of the EKF-SLAM algorithm," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2006, pp. 3562–3568.
- [46] Z. Hu, A. Bulling, S. Li, and G. Wang, "EHTask: Recognizing user tasks from eye and head movements in immersive virtual reality," *IEEE Trans. Vis. Comput. Graph.*, vol. 29, no. 4, pp. 1992–2004, Apr. 2023.