Channel Parameter Estimation and TX Positioning with Multi-Beam Fusion in 5G mmWave Networks

Mike Koivisto, Jukka Talvitie *Member, IEEE*, Elizaveta Rastorgueva-Foi, Yi Lu *Student Member, IEEE*, and Mikko Valkama *Senior Member, IEEE*

Abstract—Since the beginning of the fifth generation (5G) standardization process, positioning has been considered as a key element in future cellular networks. In order to perform accurate positioning, solutions for estimating and processing location-related measurements such as direction of arrival (DoA) and time of arrival (ToA) for various use-cases need to be developed. In this paper, building on the existing 5G new radio (NR) specifications and millimeter wave frequencies, we propose a novel estimation and tracking solution of the DoA and ToA such that only analog/RF beamforming-based observations are utilized. In addition to the proposed extended Kalman filter (EKF)-based estimation and tracking approach, we derive Cramér-Rao lower bounds (CRLBs) for the considered RF multi-beam system, and propose an information-based criterion for selecting the necessary beams for the estimation process in order to provide highly accurate performance with feasible computational complexity. The performance of the proposed method is evaluated using extensive ray-tracing simulations and numerical evaluations, and the results are compared with other estimation and beam-selection approaches. Based on the obtained results, beam-selection at the receiver can have a significant impact on the DoA and ToA estimation performance as well as on the subsequent positioning accuracy. Finally, we demonstrate the highly accurate performance of the methods when extended to joint multi-receiver-based device positioning and clock synchronization.

Index Terms—5G, beamforming, Cramér-Rao lower bound, direction of arrival, extended Kalman filter, millimeter waves, positioning, time of arrival, tracking

I. INTRODUCTION

IN addition to the demanding communication requirements of the fifth generation (5G) mobile radio networks, it has been envisioned that accurate positioning will play a key role not only in personal navigation and various vertical applications but also in empowering network functionalities in terms of location-aware communications [1]–[4]. Especially the 5G millimeter wave (mmWave) networks will employ large bandwidths as well as antenna arrays with increasing number of antenna elements, thus creating a convenient environment for device positioning based on, e.g., direction of arrival (DoA) and time of arrival (ToA) measurements, which are already supported

in 5G new radio (NR) specifications [5]. Furthermore, the mmWave networks rely highly on line of sight (LoS) conditions which in turn implies even more densified deployments [6] as well as more accurate positioning performance if designed properly. However, when the number of antenna elements increases, thus theoretically improving the DoA estimation accuracy [7], [8], the overall positioning performance may suffer from signal-to-noise ratio (SNR) degradation due to possible beam-misalignments in a fixed-beam system [9]. In order to compensate for such occasional performance degradation, we propose a novel extended Kalman filter (EKF)-based DoA and ToA estimation solution that is able to fuse information from several available beams in a polarimetric 5G mmWave system with analog/RF beamforming. Importantly, facilitating efficient positioning and tracking through RF-only beamforming is even further emphasized when the networks evolve towards the sub-THz regime, paving the way for future 6G systems [10].

In the recent studies, channel variables such as DoA and ToA are typically estimated by utilizing compressed sensing (CS)-based estimation approaches [11]–[13], through subspace-based methods [14], or by employing sequential estimation solution building on, e.g., EKFs [15]–[18]. While the sequential estimation methods can be found more attractive compared to classical estimation approaches due to more elaborate fusion of the prior information of the system [19], [20], the authors in [15], [16] assume that each antenna element in the arrays is equipped with a radio frequency (RF) chain. This, however, may not be feasible in 5G mmWave networks with large antenna arrays due to the cost and complexity of the hardware [21]. Despite several works on actual mmWave positioning being available in the existing literature [4], [7], [8], [17], [22], [23], the methods either consider different location-related measurements with a known transmitter (TX)-side antenna and beamforming information [17], [22], [23], provide only theoretical performance bounds [4], [7], [8] or consider obtaining the location-related measurements from a pre-defined statistics [24], [25] without practical estimators.

Building on the existing 5G NR specifications and dual-polarized antenna arrays with inexpensive RF-beamforming capabilities, we therefore propose a novel EKF-based solution for estimating and tracking the azimuth and co-elevation DoA angles and ToA through sophisticated and adaptive beam-selection in 5G mmWave networks. Despite various beam-selection methods are available in the existing literature [26], [27], the proposed beam-selection method is analyzed through its theoretical and practical positioning performance instead of

The authors are with the Department of Electrical Engineering, Tampere University, FI-33720 Tampere, Finland (e-mail: mike.koivisto@tuni.fi; jukka.talvitie@tuni.fi; elizaveta.rastorgueva-foi@tuni.fi; yi.lu@tuni.fi; mikko.valkama@tuni.fi).

This manuscript contains multimedia material, available at https://research.tuni.fi/wireless/research/positioning/TWC2021/.

communication metrics [27]–[29] or only theoretical limits [30] typically considered in the available literature. In addition to the normal communication between the TX and receiver (RX), the proposed solution does not require any information about the TX-side antenna array or additional communication overhead in terms of, e.g., dedicated positioning reference signals or two-way communications, thus being fundamentally different to the methods, e.g., in [17], [18]. Despite the fact that theoretical bounds for DoA and ToA estimates and subsequent positioning performance are derived, e.g., in [7], [31]–[35], the Cramér-Rao lower bounds (CRLBs) for the considered RF multi-beam formulation are not readily available in the existing literature. Besides the proposed estimation method, we therefore derive such theoretical bounds for the DoA and ToA estimates based on multi-beam fusion and compare the results with those obtained using a digital beamformer at the RX-side. Finally, the performance of the proposed method and its extension to multi-RX and joint TX synchronization scenarios are evaluated using extensive ray-tracing simulations in realistic environments while considering the latest 5G NR numerology.

The contributions of this article are summarized as follows:

- We propose a novel and adaptive beam-selection method for analog beamforming systems where the TX-side antenna and beamforming information is assumed to be unknown.
- We derive CRLBs for the DoA and ToA estimates as well as for the subsequent positioning for the proposed beam-selection method.
- We propose a practical EKF-based solution for DoA and ToA estimation by employing multiple RX-beams while assuming that the TX-side information is unknown.
- We evaluate the performance of the proposed method building on the uplink (UL) sounding reference signal (SRS) transmissions by employing the latest 5G NR specifications and extensive ray-tracing simulations at 28 GHz in realistic urban canyon and open area scenarios.

The rest of the paper is organized as follows. In Section II, we provide a general concept and necessary assumptions considered in the proposed framework including the assumed signal and channel models. In Section III, we first derive the theoretical lower bounds for the considered DoA and ToA estimates and then provide the notion of position error bound (PEB) for the subsequent TX location estimate. In Section IV, we then derive the proposed estimation and tracking algorithm employing a beam-selection scheme that is stemming from the available information of the received beam-based observations. The ray-tracing environment as well as the considered scenarios are described in Section V, after which the results are provided and analyzed in Section VI, where the proposed method is also extended to cover multi-RX and joint TX synchronization scenarios. Finally, conclusions are drawn in Section VII.

II. SYSTEM MODEL AND GENERAL ASSUMPTIONS

In this section, we shortly describe the general system model and assumptions regarding the problem formulation in the context of the considered mmWave 5G NR networks.

Thereafter, we derive the signal and channel models with proper polarimetric antenna models, which are then revised for the assumed scenario by reformulating the models with respect to the desired DoA and ToA channel parameters. Finally, the proposed beam-selection method is described.

A. General System Model and Assumptions

Building on the premises of 5G mmWave networks, an active TX at an unknown location $\mathbf{p}_{\mathrm{TX}} = [x_{\mathrm{TX}}, y_{\mathrm{TX}}, z_{\mathrm{TX}}]^{\mathrm{T}}$ is assumed to transmit beamformed reference signals (RSs) in a periodic manner exploiting orthogonal frequency division multiplexing (OFDM) waveforms. In particular, we do not set any further constraints for the number of transmitted beams, thus implying that such RSs can be transmitted by the TX either through a single beam based on, e.g., beam-correspondence, or by using a set of fixed TX-beams $\mathcal{B}_{\mathrm{T}}$ in a beam-sweep process. After propagating through the LoS-dominant multi-path channel, which is typically considered at high mmWave frequency communications, the transmitted RSs are then received by the RX through a set of its own fixed beams $\mathcal{B}_{\mathrm{R}}$ during a beam-sweep process as illustrated in Fig. 1. In this work, we assume that the given RX knows its exact location $\mathbf{p}_{\mathrm{RX}} = [x_{\mathrm{RX}}, y_{\mathrm{RX}}, z_{\mathrm{RX}}]^{\mathrm{T}}$ as well as the dual-polarized antenna array structure that is exploited at the RX. For practicality and improved cost-efficiency, we also assume that the beamforming at both ends is carried out using inexpensive RF-beamforming by the means of phase-shifters instead of digital or hybrid beamformers.

In the proposed channel parameter estimation and subsequent positioning solution, the received beam-based RS observations are first deployed to extract the information of the most powerful TX-beam. Given the most feasible TX-beam index, the RX then determines the candidate RX-beams $\mathcal{M}_{\mathrm{B}} \subseteq \mathcal{B}_{\mathrm{R}}$ using the proposed beams-selection method that employs the available information of the received beam-based observations in an adaptive manner. After determining the set of candidate RX-beams $\mathcal{M}_{\mathrm{B}}$, the beam-based observations corresponding to the candidate RX-beams are facilitated in the proposed EKF-based solution that estimates and tracks the DoA angles and the ToA of the LoS-path (see Fig. 1). Such estimates can be then fused in the subsequent positioning stage into the sequential location estimates similar to, e.g., in [16], [18].

Finally, it is to note that the beamformer as well as the antenna model of the TX are both assumed to be unknown at the RX side in the proposed beam-selection and subsequent two-stage EKF-based estimation solution, which in turn results in the fact that we cannot estimate the direction of departure (DoD) angles from the received signals, for instance. Moreover, we do not essentially restrict the role of the TX and RX in the considered system, thus implying that the communication direction can be either UL or downlink (DL) assuming that the orientation of the RX is known or jointly estimated. However, in the numerical evaluations, we consider a network-centric approach where the network elements are receiving beamformed UL SRS having a form of an extended Zadoff-Chu sequence [36]. Such a choice is stemming from the fact that the UL SRSs are beamformed, they are considered as potential
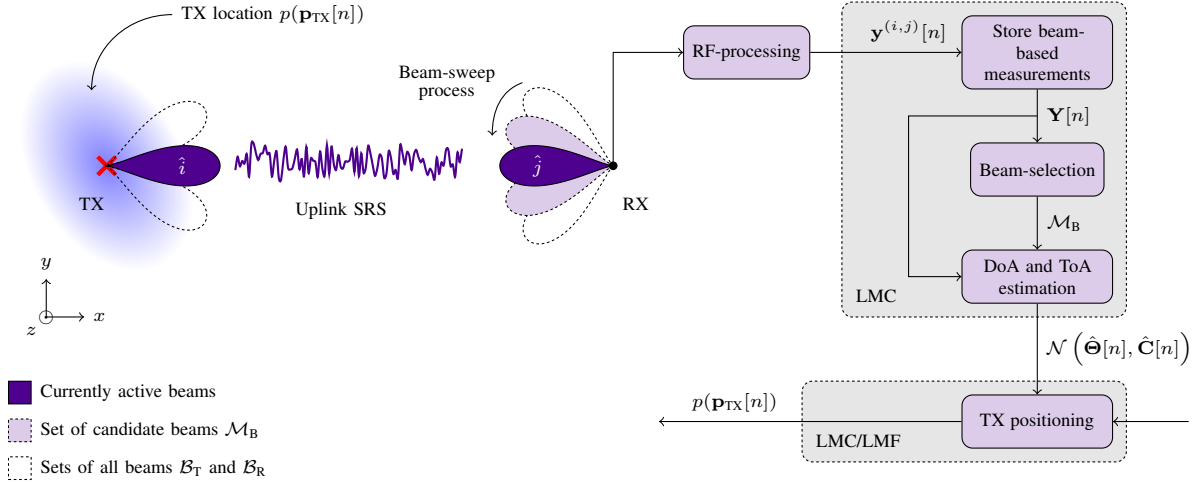
Fig. 1: Proposed two-fold estimation and positioning framework, with an adaptive beam-selection method. The proposed approach first determines the set of candidate RX-beams $\mathcal{M}_\text{B}$, which are then facilitated in joint DoA and ToA estimation and subsequent TX positioning. The processing can run in a local location management component (LMC), or alternatively at a centralized location management function (LMF) allowing to fuse the DoA and ToA estimates from several RX nodes

positioning-related RSs in 5G NR, and a specific SRS-for-positioning information element configuration is currently under consideration in the specification process [36]. In the network-centric approach, locations of the RXs are also readily available, whereas in the device-centric positioning, such information is not typically available. In order to reduce the positioning latency, it also expected that the logical positioning functionality can be brought closer to the edge by the means of LMC [37].

B. Signal and Channel Models

Let us assume that at a given time-instant n, a TX equipped with an antenna array containing $\mathcal{M}_\text{T}$ antenna elements, transmits beamformed OFDM symbols by using one of its TX-beams. In particular, the transmitted complex-valued symbol allocated at the mth subcarrier is denoted as $x^{(i)}[m,n] \in \mathbb{C}$, where the superscript $i \in \mathcal{B}_\text{T}$ indicates the index of the utilized TX-beam. As mentioned in Section II-A, we assume that the beamformers at both TX and RX side are designed purely using RF phase-shifters and such a TX-side RF-beamformer is denoted as $\mathbf{F}_\text{T}^{(i)}[n] \in \mathbb{C}^{\mathcal{M}_\text{T} \times 1}$ for which $\|\mathbf{F}_\text{T}^{(i)}[n]\|^2 = \mathcal{M}_\text{T}$. However, the design of the RF-beamformers is not restricted to the considered assumptions, and other similar antenna array and beamforming designs such as those in [38], [39] can be considered as well.

After propagating through a LoS-dominant multiple-input multiple-output (MIMO) multi-path channel, the transmitted signal is then received at the RX through its current RF-beamformer with an index $j \in \mathcal{B}_\text{R}$. In a similar manner as with the TX, the RF-beamformer of the jth RX-beam is denoted as $\mathbf{F}_\text{R}^{(j)}[n] \in \mathbb{C}^{\mathcal{M}_\text{R} \times 1}$ for which we have $\|\mathbf{F}_\text{R}^{(j)}[n]\|^2 = \mathcal{M}_\text{R}$. The received complex-valued sample between the TX-RX beam-pair (i,j) after taking the fast Fourier transform (FFT) is denoted as $y^{(i,j)}[m,n] \in \mathbb{C}$ and it is given as

$$y^{(i,j)}[m,n] = \mathbf{F}_\text{R}^{(j)}[n]^\text{H} (\mathbf{H}[m]\mathbf{F}_\text{T}^{(i)}[n]x^{(i)}[m,n] + \mathbf{n}[m,n]), \quad (1)$$

where the noise term $\mathbf{n}[m,n] \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_{\mathcal{M}_\text{R}})$, representing the noise at RX antenna elements, is modeled as the complex-

Gaussian white-noise with a power spectral density of σ^2. Moreover, $\mathbf{H}[m] \in \mathbb{C}^{\mathcal{M}_\text{R} \times \mathcal{M}_\text{T}}$ is the double-directional polarimetric MIMO channel matrix, which according to [15], [40], can be written as a superposition of $P + 1$ propagation paths as

$$\mathbf{H}[m] = \sum_{p=0}^{P} \mathbf{A}_\text{R}(\varphi_{\text{R},p}, \vartheta_{\text{R},p}) \mathbf{\Gamma}_p \mathbf{A}_\text{T}^\text{H}(\varphi_{\text{T},p}, \vartheta_{\text{T},p}) e^{-j2\pi f_m \tau_p}, \quad (2)$$

where $p = 0$ denotes the LoS-path, and $e^{-j2\pi f_m \tau_p}$ is a phase-shift introduced by the propagation delay $\tau_p \in \mathbb{R}$ of the pth propagation path and the mth subcarrier with a baseband frequency of f_m. In addition, $\mathbf{\Gamma}_p \in \mathbb{C}^{2 \times 2}$ is a polarimetric complex-valued path-gain having the radio wave polarization coefficients of the considered p^th path as its elements [15], [40]. In this work, it is assumed that the whole beam-sweep process occurs within a coherence time of the channel, thus implying that the response $\mathbf{H}[m]$ remains almost constant during a single beam-sweep process, and therefore, we omit the time-index n from its definition.

Finally, the polarimetric antenna array responses of the TX and the RX in (2) are given as

$$\mathbf{A}_\text{T}(\varphi_{\text{T},p}, \vartheta_{\text{T},p}) = [\mathbf{A}_{\text{T},\text{V}}(\varphi_{\text{T},p}, \vartheta_{\text{T},p}), \mathbf{A}_{\text{T},\text{H}}(\varphi_{\text{T},p}, \vartheta_{\text{T},p})] \quad (3)$$

$$\mathbf{A}_\text{R}(\varphi_{\text{R},p}, \vartheta_{\text{R},p}) = [\mathbf{A}_{\text{R},\text{V}}(\varphi_{\text{R},p}, \vartheta_{\text{R},p}), \mathbf{A}_{\text{R},\text{H}}(\varphi_{\text{R},p}, \vartheta_{\text{R},p})], \quad (4)$$

where $\mathbf{A}_{\text{T},X}(\varphi_{\text{T},p}, \vartheta_{\text{T},p}) \in \mathbb{C}^{\mathcal{M}_\text{T} \times 1}$ and $\mathbf{A}_{\text{R},X}(\varphi_{\text{R},p}, \vartheta_{\text{R},p}) \in \mathbb{C}^{\mathcal{M}_\text{R} \times 1}$ denote the antenna array responses of the TX and the RX for the considered polarization component $X \in \{\text{H}, \text{V}\}$, respectively. Moreover, the angle-pair $(\varphi_{\text{T},p}, \vartheta_{\text{T},p})$ denotes the DoD angle-pair and $(\varphi_{\text{R},p}, \vartheta_{\text{R},p})$ is the DoA angle-pair of the pth propagation path. Here, the azimuth angles are denoted as $\varphi \in [0, 2\pi)$ and the co-elevation angles are denoted as $\vartheta \in [0, \pi]$ and both of them are defined in a global coordinate system.

C. Considered Antenna and Single-path Channel Models

Since the proposed estimation approach estimates the DoA angles and the ToA corresponding to the LoS-path, we first

re-formulate the signal model in (1) accordingly. Due to the considered LoS-path assumption, we ignore the path-index p from the channel variables throughout the rest of the paper. However, despite the simplified single-path channel model used in the theoretical derivations as well as in the algorithm derivations, all the essential multi-path components are modeled and included in the final numerical evaluations using extensive ray-tracing simulations.

Let us first collect all the received frequency-domain samples at the RX into a single multicarrier measurement vector denoted as $\mathbf{y}^{(i,j)}[n] \in \mathbb{C}^{\mathcal{M}_f \times 1}$, where $\mathcal{M}_f$ is the number of active subcarriers and (i,j) is the current beam-pair between the TX and the RX. Applying the results in [15], [40] into the signal model (1), the multicarrier observation at the RX between the considered beam-pair (i,j) can be written in a simplified form as

$$\mathbf{y}^{(i,j)}[n] = \mathbf{B}^{(i,j)}(\mathbf{\Theta}[n])\boldsymbol{\psi}^{(i)}[n] + \mathbf{n}^{(i,j)}[n], \quad (5)$$

where $\mathbf{\Theta}[n] = [\tau[n], \varphi_\text{R}[n], \vartheta_\text{R}[n]]^\text{T}$ denotes the channel parameters of the LoS-path. Moreover,

$$\boldsymbol{\psi}^{(i)}[n] = [\psi_\text{V}^{(i)}[n], \psi_\text{H}^{(i)}[n]]^\text{T} \quad (6)$$
$$= \mathbf{\Gamma}\mathbf{A}_\text{T}^\text{H}(\varphi_\text{T}[n], \vartheta_\text{T}[n])\mathbf{F}_\text{T}^{(i)}[n] \in \mathbb{C}^{2\times 1} \quad (7)$$

is a combination of the unknown path-gains and the TX side antenna response including the beamforming gain, whereas the matrix-valued function $\mathbf{B}^{(i,j)}(\mathbf{\Theta}[n]) \in \mathbb{C}^{\mathcal{M}_f \times 2}$ reads

$$\mathbf{B}^{(i,j)}(\mathbf{\Theta}[n]) = (\mathbf{F}_\text{R}^{(j)}[n]^\text{H}\mathbf{A}_\text{R}(\varphi_\text{R}[n], \vartheta_\text{R}[n])) \odot \mathbf{B}_f(\tau[n]), \quad (8)$$

where $\odot$ denotes the Khatri-Rao product, i.e., a column-wise Kronecker product, and $\mathbf{B}_f(\tau[n]) = \mathbf{X}^{(i)}[n]\mathbf{G}_f\mathbf{d}(\tau) \in \mathbb{C}^{\mathcal{M}_f \times 1}$ is a combination of the transmitted signal and the combined frequency response of the RF-chains, in which [40]

$$\mathbf{d}(\tau) = e^{-j2\pi f_0 \tau \left[-\frac{(\mathcal{M}_f-1)}{2}, ..., \frac{(\mathcal{M}_f-1)}{2}\right]}, \quad (9)$$

where f_0 is the subcarrier spacing. Furthermore, $\mathbf{X}^{(i)}[n] \in \mathbb{C}^{\mathcal{M}_f \times \mathcal{M}_f}$ is a matrix containing the transmitted multicarrier SRS reference symbols of the ith TX-beam on the diagonal with equally distributed energy over the subcarriers. Finally, the noise-term in (5) after the RX-side beamformer is now given as $\mathbf{n}^{(i,j)}[n] \sim \mathcal{CN}(0, \sigma^2 \mathcal{M}_\text{R}\mathbf{I}_{\mathcal{M}_f})$, and it is assumed to be independent of the RX beamformer. In particular, the structure of $\mathbf{B}^{(i,j)}(\mathbf{\Theta}[n])$ in (8) is assumed to be known at the RX whereas $\boldsymbol{\psi}^{(i)}[n]$ in (7) is unknown and therefore, it needs to be estimated jointly with the unknown model noise variance σ^2 and the desired channel parameters $\mathbf{\Theta}[n]$.

In the proposed estimation method as well as in the conducted theoretical derivations, the polarimetric antenna array responses are modeled by employing effective aperture distribution functions (EADFs) which can be considered as a numerically efficient solution in representing the antenna array manifold using Fourier series expansion [40], [41]. Hence, the antenna array responses in (3)-(4) can be given as

$$\mathbf{A}_{\text{T},X}(\varphi_{\text{T},p}, \vartheta_{\text{T},p}) = \mathbf{G}_{\text{T},X}\mathbf{d}(\varphi_{\text{T},p}, \vartheta_{\text{T},p}) \quad (10)$$
$$\mathbf{A}_{\text{R},X}(\varphi_{\text{R},p}, \vartheta_{\text{R},p}) = \mathbf{G}_{\text{R},X}\mathbf{d}(\varphi_{\text{R},p}, \vartheta_{\text{R},p}) \quad (11)$$

where $\mathbf{G}_{\text{T},X} \in \mathbb{C}^{\mathcal{M}_\text{T} \times \mathcal{M}_a \mathcal{M}_e}$ and $\mathbf{G}_{\text{R},X} \in \mathbb{C}^{\mathcal{M}_\text{R} \times \mathcal{M}_a \mathcal{M}_e}$ denote the EADFs of the TX and the RX to the given polarization direction, respectively. Moreover, $\mathbf{d}(\varphi_{\text{T},p}, \vartheta_{\text{T},p})$ and $\mathbf{d}(\varphi_{\text{R},p}, \vartheta_{\text{R},p})$ denote the antenna domain sampling vectors which are formed such that $\mathbf{d}(\varphi_{\cdot,p}, \vartheta_{\cdot,p}) = \mathbf{d}(\vartheta_{\cdot,p}) \otimes \mathbf{d}(\varphi_{\cdot,p})$, where $\otimes$ denotes the Kronecker product and [16], [40], [41]

$$\mathbf{d}(\vartheta_{\cdot,p}) = e^{j\vartheta_{\cdot,p}\left[-\frac{(\mathcal{M}_e-1)}{2}, ..., \frac{(\mathcal{M}_e-1)}{2}\right]} \in \mathbb{C}^{\mathcal{M}_e \times 1} \quad (12)$$
$$\mathbf{d}(\varphi_{\cdot,p}) = e^{j\varphi_{\cdot,p}\left[-\frac{(\mathcal{M}_a-1)}{2}, ..., \frac{(\mathcal{M}_a-1)}{2}\right]} \in \mathbb{C}^{\mathcal{M}_a \times 1} \quad (13)$$

denote the Vandermonde-structured vectors. Here, $\mathcal{M}_a$ and $\mathcal{M}_e$ are the number of modes, i.e., spatial harmonics of the antenna response in the azimuth and co-elevation directions, respectively [40], [42].

D. Proposed Information-based Beam-Selection Scheme

Since the information from several available RX-beams is fused in the proposed estimation process, we shortly present the notation for the combined beam-based observations. After the whole beam-sweep process and storing the beam-based multicarrier observations, the RX needs to then determine the set of beam-based observations that are the most significant in the actual estimation phase. Therefore, the RX first determines the best TX side beam-index $\hat{i} \in \mathcal{B}_\text{T}$ based on the received beam-based reference signal received power (RSRP) measurements [43] over all the RX-beams such that

$$\hat{i} = \underset{i}{\operatorname{argmax}} \sum_{j \in \mathcal{B}_\text{R}} \left(\frac{1}{\mathcal{M}_f} \sum_{m=1}^{\mathcal{M}_f} |y^{(i,j)}[m,n]|^2\right). \quad (14)$$

In particular, the TX-beam corresponding to the obtained index $\hat{i}$ is the most powerful beam from the RX perspective. If the TX transmits the signals using a single beam, e.g., through beam-correspondence, the best TX beam-index $\hat{i}$ is thus the index of such beam.

In order to determine the suitable RX side beam-set $\mathcal{M}_\text{B} \subseteq \mathcal{B}_\text{R}$, different approaches based on, e.g., noise-threshold or some heuristics can be applied [17], [22]. In order to find the best possible set of RX-beams, solutions relying on, e.g., Bayesian information criterion (BIC) could be applied, but finding such a beam-set requires evaluating the BIC for all the possible RX beam-combinations, which is computationally extremely demanding. In this work, the set $\mathcal{M}_\text{B}$ is defined as a unique set of beam-indices over all the desired channel parameters for which the cumulative sum of the normalized Fisher information matrix (FIM) values exceeds a certain predefined threshold $\alpha_\text{thr} \in (0, 1]$. At each time-instant, the set of RX beam-indices can be thus determined using a mathematical set notation as

$$\mathcal{M}_\text{B} = \bigcup_{z \in \mathcal{I}_\Theta} \left\{ j : \sum_{j \in \mathcal{B}_{\text{R,sort}}} \left(\frac{\mathbf{J}_z^{(j)}(\hat{\mathbf{\Theta}}[n])}{\sum_{j \in \mathcal{B}_\text{R}} \mathbf{J}_z^{(j)}(\hat{\mathbf{\Theta}}[n])}\right) > \alpha_\text{thr} \right\}, \quad (15)$$

where $\mathbf{J}_z^{(j)}(\hat{\mathbf{\Theta}}[n])$ denotes the FIM of the state-variable $z \in \mathcal{I}_\Theta = \{\varphi_\text{R}, \vartheta_\text{R}, \tau\}$ which can be calculated as in (19) by considering only the jth beam-based observation $\mathbf{y}^{(i,j)}[n]$ instead of the whole concatenated set of beam-based observations. In addition, $\mathcal{B}_{\text{R,sort}}$ is a set of RX beam-indices after ordering them based on the normalized FIM values, thus ensuring

that the most informative beam-indices are taken into account in $\mathcal{M}_\text{B}$. In practice, α_thr can be interpreted as a percentage threshold describing the information portion (i.e., a cumulative sum of normalized beam-based information values) that needs to be exceeded for a given state-variable z. As a concrete example, if $\alpha_\text{thr} = 1$, it would imply that all the RX-beams are selected for the estimation process. As shortly discussed in Section IV, the computational complexity of the proposed method is proportional to the number of facilitated RX-beams while the performance of the method can be improved by increasing the value of α_thr. Hence, the proposed beam-selection method allows for adjusting the computational complexity and the estimation performance through tuning α_thr even between state-variables. In the numerical evaluations, we compare the proposed information-based beam-selection scheme with the one used in [18] by analyzing the estimation and subsequent positioning performance with both beam-selection approaches.

Finally and for the notational convenience, given the set of concatenated beam-based observations $\mathbf{Y}[n] \in \mathbb{C}^{|\mathcal{M}_\text{B}|\mathcal{M}_f \times 1}$, we define the combined signal model for the corresponding set of $\mathcal{M}_\text{B}$ candidate RX-beams simply as

$$\mathbf{Y}[n] = \mathbf{B}(\mathbf{\Theta}[n])\boldsymbol{\psi}[n] + \mathbf{n}[n], \qquad (16)$$

where the signal model function $\mathbf{B}(\mathbf{\Theta}[n]) \in \mathbb{C}^{\mathcal{M} \times 2}$, with $\mathcal{M} = |\mathcal{M}_\text{B}|\mathcal{M}_f$, reads

$$\mathbf{B}(\mathbf{\Theta}[n]) = (\mathbf{F}_\text{R}[n]^\text{H}\mathbf{A}_\text{R}(\varphi_\text{R}[n], \vartheta_\text{R}[n])) \odot \mathbf{B}_f(\tau[n]), \quad (17)$$

in which $\mathbf{F}_\text{R}[n] \in \mathbb{C}^{\mathcal{M}_\text{R} \times |\mathcal{M}_\text{B}|}$ is the RF-beamforming matrix containing all the selected RX-beams. Furthermore, $\mathbf{n}[n] \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_\mathcal{M})$ is the combined noise-term over the chosen RX-beams.

III. CRAMÉR-RAO LOWER BOUNDS

In this work, we seek to estimate the DoA angles and the ToA by fusing information from $\mathcal{M}_\text{B}$ available RX-beams given the best TX-beam with an index $\hat{i}$. In order to demonstrate the impact of such a measurement fusion on the accuracy of the desired channel parameter estimates and on the subsequent positioning performance, we first derive the CRLBs for the DoA and ToA under the considered system model. Thereafter, we provide PEBs for analyzing the instantaneous positioning performance at a theoretical level. In particular, the CRLB provides a theoretical lower-bound for the covariance matrix of an unbiased estimator $\hat{\mathbf{x}}$ such that [19]

$$\text{CRLB}(\mathbf{x}) = (\mathbf{J}(\mathbf{x}))^{-1} \leq \mathbb{E}\{(\hat{\mathbf{x}} - \mathbf{x})(\hat{\mathbf{x}} - \mathbf{x})^\text{T}\}, \qquad (18)$$

where $\mathbf{J}(\mathbf{x})$ is the FIM of the vector-valued variable $\mathbf{x}$. By considering the ith and j^th scalar-valued elements of $\mathbf{x}$, denoted as x_i and x_j, respectively, the FIM is obtained through [19], [40]

$$[\mathbf{J}(\mathbf{x})]_{i,j} = -\mathbb{E}\left\{ \frac{\partial^2 \ell(\mathbf{Y}|\mathbf{x})}{\partial x_i \partial x_j} \right\}, \qquad (19)$$

where $[\cdot]_{i,j}$ denotes the element of the matrix argument located at the index-pair (i, j), and $\ell(\mathbf{Y}|\mathbf{x})$ is the log-likelihood of the observations given the variable $\mathbf{x}$. Throughout this section, we omit the time-index n from the derivations for notational simplicity.

A. Theoretical Bounds of the Channel Variables

In general, the CRLBs of the channel variables can be derived similar to e.g., in [8], [32], [40], however, we derive the CRLBs such that the estimators as well as the CRLBs employ the same models where the impact of existing multi-path components is also incorporated. Due to the fact that we are not actually interested in estimating the unknown noise variance σ^2 or the combined channel and TX information $\boldsymbol{\psi}$ in the proposed framework, we employ a subspace fitting technique through variable projection [44] in a similar manner as, e.g., in [16], [18]. Building on such a variable projection approach, we thus derive the CRLBs for the DoA and the ToA estimates in order to emphasize the role of multi-beam fusion in the estimation process.

For the notational convenience, let us first define a vector $\mathbf{x}$ that contains all the unknown variables in the considered beam-based signal model (16), as

$$\mathbf{x} = [\mathbf{\Theta}^\text{T}, \boldsymbol{\psi}^\text{T}, \sigma^2]^\text{T} \in \mathbb{C}^{6 \times 1}, \qquad (20)$$

where $\mathbf{\Theta} = [\tau, \varphi_\text{R}, \vartheta_\text{R}]^\text{T}$ denotes the desired channel variables. Given the model for the complex-valued beam-based multi-carrier observations in (16), the likelihood function for the observations $\mathbf{Y} \in \mathbb{C}^{\mathcal{M} \times 1}$ can be expressed as [40]

$$p(\mathbf{Y}|\mathbf{x}) = p(\mathbf{Y}|\tau, \varphi_\text{R}, \vartheta_\text{R}, \boldsymbol{\psi}, \sigma^2) \qquad (21)$$

$$= \frac{1}{\pi^\mathcal{M} \det(\sigma^2 \mathbf{I}_\mathcal{M})} e^{-\frac{1}{\sigma^2}\|\mathbf{Y} - \mathbf{B}(\varphi_\text{R}, \vartheta_\text{R}, \tau)\boldsymbol{\psi}\|^2}, \qquad (22)$$

which can be transformed into the log-likelihood function in a straightforward manner such that

$$\ell(\mathbf{Y}|\mathbf{x}) = \log(p(\mathbf{Y}|\mathbf{x})) \qquad (23)$$

$$= -\mathcal{M}\log(\pi\sigma^2) - \frac{1}{\sigma^2}\|\mathbf{Y} - \mathbf{B}(\varphi_\text{R}, \vartheta_\text{R}, \tau)\boldsymbol{\psi}\|^2, \quad (24)$$

where $\log$ denotes the natural logarithm.

In particular, since (24) is separable with respect to the nuisance variables σ^2 and $\boldsymbol{\psi}$, and they can be solved in a closed-form, we first solve them from the obtained log-likelihood function. Based on (24), the maximum likelihood (ML) estimates of σ^2 and $\boldsymbol{\psi}$ are then given as

$$\hat{\sigma}^2 = \frac{\|\mathbf{P}_\perp(\mathbf{\Theta})\mathbf{Y}\|^2}{\mathcal{M}}, \quad \text{and} \quad \hat{\boldsymbol{\psi}} = (\mathbf{B}(\mathbf{\Theta}))^\dagger \mathbf{Y}, \qquad (25)$$

where $(\cdot)^\dagger$ denotes the Moore-Penrose pseudo-inverse, and $\mathbf{P}_\perp(\mathbf{\Theta}) = \mathbf{I}_\mathcal{M} - \mathbf{B}(\mathbf{\Theta})(\mathbf{B}(\mathbf{\Theta}))^\dagger$ denotes the orthogonal projection matrix onto the null-space of $\mathbf{B}(\mathbf{\Theta})$. Substituting the ML estimators (25) into the log-likelihood (24), we obtain the concentrated log-likelihood function

$$\hat{\ell}(\mathbf{Y}|\mathbf{\Theta}) = \hat{\ell}(\mathbf{Y}|\varphi_\text{R}, \vartheta_\text{R}, \tau) \propto -\mathcal{M}\log\left(\|\mathbf{P}_\perp(\mathbf{\Theta})\mathbf{Y}\|^2\right), \quad (26)$$

where the proportionality takes into account only the parts that play a role in the following FIM determination. Given the definition (19) and after straightforward manipulation, the observed FIM for the proposed model can be calculated according to

$$\mathbf{J}(\mathbf{\Theta}) = \frac{2}{\hat{\sigma}^2} \Re\left\{ \left(\frac{\partial \mathbf{P}_\perp(\mathbf{\Theta})\mathbf{Y}}{\partial \mathbf{\Theta}}\right)^\text{H} \frac{\partial \mathbf{P}_\perp(\mathbf{\Theta})\mathbf{Y}}{\partial \mathbf{\Theta}} \right\} + c(\mathbf{\Theta}), \quad (27)$$

where $\hat{\sigma}^2$ has the form as given in (25), and where the term $c(\boldsymbol{\Theta})$ vanishes after taking the expectation value (see, e.g., [40]). Utilizing the differentiation results for projection matrices from [44], [45], the partial derivatives in (27) with respect to each channel variable have the form

$$\frac{\partial \mathbf{P}_{\perp}(\boldsymbol{\Theta})\mathbf{Y}}{\partial \tau} = -\left(\mathbf{g}_{\tau}(\boldsymbol{\Theta}) + \mathbf{g}_{\tau}(\boldsymbol{\Theta})^{\mathrm{H}}\right)\mathbf{Y}, \quad (28)$$

where the auxiliary variable is given as

$$\mathbf{g}_{\tau}(\boldsymbol{\Theta}) = \mathbf{P}_{\perp}(\boldsymbol{\Theta})\frac{\partial \mathbf{B}(\boldsymbol{\Theta})}{\partial \tau}\mathbf{B}(\boldsymbol{\Theta})^{\dagger}. \quad (29)$$

The corresponding formulation can be conducted in a similar manner for the remaining partial derivatives as well.

Since we model the antenna array responses by the means of the EADFs, the partial derivatives of $\mathbf{B}(\boldsymbol{\Theta})$ in (28) with respect to each desired channel variable can be given as

$$\frac{\partial \mathbf{B}(\boldsymbol{\Theta})}{\partial \tau} = (\mathbf{F}_{\mathrm{R}}\mathbf{A}_{\mathrm{R}}(\varphi_{\mathrm{R}}, \vartheta_{\mathrm{R}}))\odot(\boldsymbol{\Xi}_{\tau}\mathbf{B}_f(\tau)) \quad (30)$$

$$\frac{\partial \mathbf{B}(\boldsymbol{\Theta})}{\partial \varphi_{\mathrm{R}}} = (\mathbf{F}_{\mathrm{R}}[\mathbf{G}_{\mathrm{R,V}}, \mathbf{G}_{\mathrm{R,H}}]\odot(\boldsymbol{\Xi}_{\varphi}\mathbf{d}(\varphi_{\mathrm{R}}, \vartheta_{\mathrm{R}})))\odot\mathbf{B}_f(\tau) \quad (31)$$

$$\frac{\partial \mathbf{B}(\boldsymbol{\Theta})}{\partial \vartheta_{\mathrm{R}}} = (\mathbf{F}_{\mathrm{R}}[\mathbf{G}_{\mathrm{R,V}}, \mathbf{G}_{\mathrm{R,H}}]\odot(\boldsymbol{\Xi}_{\vartheta}\mathbf{d}(\varphi_{\mathrm{R}}, \vartheta_{\mathrm{R}})))\odot\mathbf{B}_f(\tau), \quad (32)$$

where the auxiliary matrices are

$$\boldsymbol{\Xi}_{\tau} = \mathrm{diag}\left(-\mathrm{j}2\pi f_0\left[-\frac{\mathcal{M}_f - 1}{2}, \ldots, \frac{\mathcal{M}_f - 1}{2}\right]\right) \quad (33)$$

$$\boldsymbol{\Xi}_{\varphi} = \mathrm{diag}\left(\mathrm{j}\left[-\frac{\mathcal{M}_a - 1}{2}, \ldots, \frac{\mathcal{M}_a - 1}{2}\right]\right) \otimes \mathbf{I}_{\mathcal{M}_e} \quad (34)$$

$$\boldsymbol{\Xi}_{\vartheta} = \mathbf{I}_{\mathcal{M}_a} \otimes \mathrm{diag}\left(\mathrm{j}\left[-\frac{\mathcal{M}_e - 1}{2}, \ldots, \frac{\mathcal{M}_e - 1}{2}\right]\right), \quad (35)$$

where $\mathcal{M}_f$ denotes the number of active subcarriers, and $\mathcal{M}_a$ and $\mathcal{M}_e$ are the spatial modes of the EADFs, respectively.

Finally, the overall FIM can be expressed as a real-valued matrix

$$\mathbf{J}(\boldsymbol{\Theta}) = \begin{bmatrix} \mathbf{J}_{\tau,\tau}(\boldsymbol{\Theta}) & \mathbf{J}_{\tau,\varphi_{\mathrm{R}}}(\boldsymbol{\Theta}) & \mathbf{J}_{\tau,\vartheta_{\mathrm{R}}}(\boldsymbol{\Theta}) \\ \mathbf{J}_{\varphi_{\mathrm{R}},\tau}(\boldsymbol{\Theta}) & \mathbf{J}_{\varphi_{\mathrm{R}},\varphi_{\mathrm{R}}}(\boldsymbol{\Theta}) & \mathbf{J}_{\varphi_{\mathrm{R}},\vartheta_{\mathrm{R}}}(\boldsymbol{\Theta}) \\ \mathbf{J}_{\vartheta_{\mathrm{R}},\tau}(\boldsymbol{\Theta}) & \mathbf{J}_{\vartheta_{\mathrm{R}},\varphi_{\mathrm{R}}}(\boldsymbol{\Theta}) & \mathbf{J}_{\vartheta_{\mathrm{R}},\vartheta_{\mathrm{R}}}(\boldsymbol{\Theta}) \end{bmatrix} \in \mathbb{R}^{3\times3}, \quad (36)$$

while the root mean squared errors (RMSEs) of the estimates can be obtained as

$$\mathrm{RMSE}\ \hat{\tau} = \sqrt{\mathrm{CRLB}_{\tau}(\boldsymbol{\Theta})} = \sqrt{\mathbf{J}_{\tau,\tau}^{-1}(\boldsymbol{\Theta})} \quad (37)$$

$$\mathrm{RMSE}\ \hat{\varphi}_{\mathrm{R}} = \sqrt{\mathrm{CRLB}_{\varphi_{\mathrm{R}}}(\boldsymbol{\Theta})} = \sqrt{\mathbf{J}_{\varphi_{\mathrm{R}},\varphi_{\mathrm{R}}}^{-1}(\boldsymbol{\Theta})} \quad (38)$$

$$\mathrm{RMSE}\ \hat{\vartheta}_{\mathrm{R}} = \sqrt{\mathrm{CRLB}_{\vartheta_{\mathrm{R}}}(\boldsymbol{\Theta})} = \sqrt{\mathbf{J}_{\vartheta_{\mathrm{R}},\vartheta_{\mathrm{R}}}^{-1}(\boldsymbol{\Theta})}. \quad (39)$$

These RMSE values are used to analyze the theoretical estimation performance of the individual channel variables in the conducted simulations and numerical evaluations. We also note that the derived FIM in (27) contains the stochastic signal properties due to the variable projection and hence, the final FIM utilized in the theoretical results is averaged over several Monte Carlo realizations. Finally, in order to allow for proper comparisons, we compare the theoretical lower bounds of the proposed approach to those obtained using a digital beamformer at the RX similar to the available existing literature [7], [31]–[35].

B. Positioning Error Bounds

Since the DoA angles and the ToA can be related to the target position using simple geometric relations, the obtained FIMs can be translated to the PEBs [19], [32]. In particular, the PEB gives the theoretical lower bound for the instantaneous position estimate through position FIM. In this work, the PEB is analyzed in order to provide better insights into how the CRLBs of the DoA and ToA jointly affect the theoretical positioning performance. Now, by denoting the unknown location of the TX as $\mathbf{p}_{\mathrm{TX}} = [x_{\mathrm{TX}}, y_{\mathrm{TX}}, z_{\mathrm{TX}}]^{\mathrm{T}}$ and the corresponding known RX location as $\mathbf{p}_{\mathrm{RX}} = [x_{\mathrm{RX}}, y_{\mathrm{RX}}, z_{\mathrm{RX}}]^{\mathrm{T}}$, the relation between the device location $\mathbf{p}_{\mathrm{TX}}$ and the DoA angles and the ToA is given with a differentiable function $[\tau, \varphi_{\mathrm{R}}, \vartheta_{\mathrm{R}}]^{\mathrm{T}} = \mathbf{h}(\mathbf{p}_{\mathrm{TX}}) \in \mathbb{R}^{m\times1}$, where [32]

$$\mathbf{h}(\mathbf{p}_{\mathrm{TX}}) = \begin{bmatrix} \dfrac{\|\mathbf{p}_{\mathrm{RX}} - \mathbf{p}_{\mathrm{TX}}\|}{c} \\ \arctan\left(\dfrac{y_{\mathrm{TX}} - y_{\mathrm{RX}}}{x_{\mathrm{TX}} - x_{\mathrm{RX}}}\right) \\ \arccos\left(\dfrac{z_{\mathrm{TX}} - z_{\mathrm{RX}}}{\|\mathbf{p}_{\mathrm{RX}} - \mathbf{p}_{\mathrm{TX}}\|}\right) \end{bmatrix}, \quad (40)$$

where c is the speed of light.

Considering then the mapping in (40), the FIMs of the azimuth and the co-elevation DoA angles and the ToA can be transformed into the 3D PEB according to

$$\mathrm{PEB} = \sqrt{\mathrm{trace}(\mathbf{J}_{\mathbf{p}_{\mathrm{TX}}}^{-1}(\mathbf{p}_{\mathrm{TX}}))} \quad (41)$$

$$= \sqrt{\mathrm{trace}((\mathbf{H}(\mathbf{p}_{\mathrm{TX}})^{\mathrm{T}}\mathbf{J}(\boldsymbol{\Theta})\mathbf{H}(\mathbf{p}_{\mathrm{TX}}))^{-1})}, \quad (42)$$

where $\mathbf{H}(\mathbf{p}_{\mathrm{TX}}) \in \mathbb{R}^{3\times3}$ denotes the Jacobian matrix of the measurement model function $\mathbf{h}(\mathbf{p}_{\mathrm{TX}})$ with respect to the TX location $\mathbf{p}_{\mathrm{TX}}$. While mutually synchronized TX and RX as well as known RX orientation are assumed in the theoretical evaluations in this work, we note that such effects can also be taken into account in the considered CRLB and PEB derivations in a similar manner as, e.g., in [8], [32]. Furthermore, in the conducted simulations, we extend and demonstrate the performance of the proposed method also in joint TX positioning and synchronization scenarios.

IV. PROPOSED EKF-BASED ESTIMATION AND TRACKING SOLUTION

The proposed sequential estimation framework consists of two stages, as illustrated at a general level in Fig. 1. The fundamental idea behind the two-fold approach is that the output of the first-stage estimation process allows for efficient positioning through Gaussian approximations without the need of communicating the raw channel measurements between the network elements, e.g., when several network nodes are participating in device positioning. In this section, we derive the proposed first-stage EKF where the beam-based multicarrier observations are first fused into the DoA angle and the ToA estimates. Thereafter, we shortly describe the second-stage EKF where the obtained DoA and ToA estimates are facilitated in the final device positioning, thus providing essential information on the performance of the proposed method through one and easily understandable performance metric. It is to note that the second-stage positioning EKF is not considered as a novelty
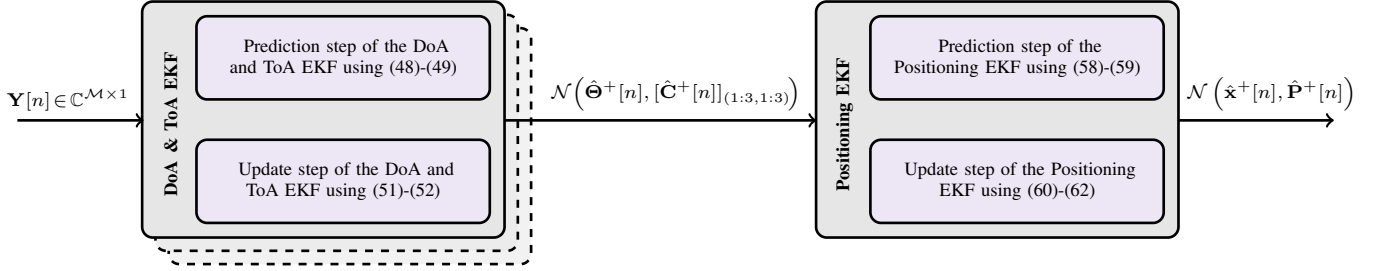
Fig. 2: Schematic of the proposed estimation and tracking framework, where the beam-based measurements are fused into the TX location estimate and the corresponding location uncertainty after employing the information-based beam-selection method. The dashed boxes indicate the possibility to fuse measurements from several RX nodes, as shortly illustrated in Section VI-D.

of the work as such, since similar works on fusing DoA and ToA measurements in positioning are readily available in the existing literature. The more detailed schematic of the proposed two-fold estimation framework is depicted in Fig. 2.

A. Joint DoA and ToA Estimation and Tracking EKF

After obtaining the most significant beams according to (15), the DoA angles and the ToA are then jointly estimated by fusing the beam-based measurements in an EKF-based solution. In general, an EKF is a sequential Bayesian estimation approach for the systems with non-linear state and measurement models where the noise is assumed to be Gaussian distributed [20], [46]–[48]. In contrast to classical estimation methods, an EKF is able to incorporate prior information of the state to the estimation process which in turn can improve the overall estimation performance [19], [20], while being computationally more attractive compared to other Bayesian estimation methods, e.g., particle filters [46].

In this work, a constant white-noise acceleration (CWNA) model is considered as a state evolution model for the system, implying that the considered state variables, that is, the DoA angles and the ToA, are assumed to evolve with almost a constant rate over time. At a time-instant n, the state of the system can be written as

$$\mathbf{s}[n] = \left[\mathbf{\Theta}^{\mathrm{T}}[n], \Delta\mathbf{\Theta}^{\mathrm{T}}[n] \right]^{\mathrm{T}} \tag{43}$$

$$= [\varphi_{\mathrm{R}}[n], \vartheta_{\mathrm{R}}[n], \tau[n], \Delta\varphi_{\mathrm{R}}[n], \Delta\vartheta_{\mathrm{R}}[n], \Delta\tau[n]]^{\mathrm{T}}, \tag{44}$$

where $\Delta\mathbf{\Theta}[n] = [\Delta\varphi_{\mathrm{R}}[n], \Delta\vartheta_{\mathrm{R}}[n], \Delta\tau[n]]^{\mathrm{T}}$ denote the rate-of-change in the azimuth and co-elevation DoA angles, and in the ToA, respectively. Given the CWNA state-transition model and the considered multi-beam observation model, the state-space model can be thus written as

$$\mathbf{s}[n] = \mathbf{F}\mathbf{s}[n-1] + \mathbf{w}[n] \tag{45}$$

$$\mathbf{Y}[n] = \mathbf{B}(\mathbf{s}[n])\boldsymbol{\psi}[n] + \mathbf{n}[n], \tag{46}$$

where the measurement model is as shown in (17), while the state-transition matrix $\mathbf{F} \in \mathbb{R}^{6\times6}$ with Δt denoting the time between two consecutive estimation time-instances reads

$$\mathbf{F} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \otimes \mathbf{I}_3. \tag{47}$$

Furthermore, the state-process noise is given as $\mathbf{w}[n] \sim \mathcal{N}(0, \mathbf{Q}[n])$, where the state-model driving noise covariance

matrix can be obtained through discretization of the continuous time model [46]. Using a similar notation as in [20], the a priori estimates of the mean $\hat{\mathbf{s}}^{-}[n]$ and the covariance matrix $\hat{\mathbf{C}}^{-}[n]$ can be then obtained such that

$$\hat{\mathbf{s}}^{-}[n] = \mathbf{F}\hat{\mathbf{s}}^{+}[n-1] \tag{48}$$

$$\hat{\mathbf{C}}^{-}[n] = \mathbf{F}\hat{\mathbf{C}}^{+}[n-1]\mathbf{F}^{\mathrm{T}} + \mathbf{Q}[n], \tag{49}$$

where $\hat{\mathbf{s}}^{+}[n-1]$ and $\hat{\mathbf{C}}^{+}[n-1]$ denote the previous a posteriori estimates of the mean and the covariance matrix, respectively. After the prediction step, the measurements can be then incorporated to the estimation process in the update step of the EKF. Instead of a more common Kalman-gain form of the EKF, the a posteriori estimates are updated in this work in the information form[1] by employing the gradient and observed FIM of the measurement model function [15]. Such a decision is stemming from the fact that the information form of the update step can be found computationally more attractive than the Kalman-gain form especially in a case of having large number of complex-valued measurements [16], [49].

Next, in order to derive the gradient and the observed FIM for the proposed EKF, we assume that the TX beam-index $\hat{i}$ as well as the set of RX-beams $\mathcal{M}_{\mathrm{B}}$ have been determined and available at the RX. As described in Section III, we are not interested in tracking the noise variance σ^2 or the combined channel and TX information $\boldsymbol{\psi}[n]$ in the proposed estimation framework. Therefore, we employ a subspace fitting technique through variable projection [44] and concentrate the log-likelihood function with respect to the nuisance variables in a similar manner as in the CRLB derivations in Section III. In particular, we deploy the observed FIM defined in (27), and thus we only need to determine the gradient of the log-likelihood function for the proposed EKF. Given the concentrated log-likelihood function in (26), we can thus calculate the gradient, i.e., the score-function, by taking the partial derivative of the concentrated log-likelihood with respect to the related channel parameters $\mathbf{\Theta}[n]$ according to [19], [40]

$$\mathbf{q}(\mathbf{\Theta}[n]) = \frac{\partial}{\partial\mathbf{\Theta}[n]}\hat{\ell}(\mathbf{Y}[n]|\mathbf{\Theta}[n]) \tag{50}$$

$$= -\frac{2}{\hat{\sigma}^2}\Re\left\{ \left(\frac{\partial\mathbf{P}_{\perp}(\mathbf{\Theta}[n])\mathbf{Y}[n]}{\partial\mathbf{\Theta}[n]} \right)^{\mathrm{H}} \mathbf{P}_{\perp}(\mathbf{\Theta}[n])\mathbf{Y}[n] \right\},$$

---

[1]In contrast to the information filter in [20], the exploited information form of the update step follows the notion used, e.g., in [15] and it is essentially the same as a single update iteration of the iterated Kalman filter [49].

TABLE I: APPROXIMATED NUMBER OF FLOPs REQUIRED IN THE DOA AND TOA ESTIMATION AND TRACKING EKF ($p \ll \mathcal{M}$).

Prediction step	$4p^3 + 3p^2$
Update step	$(18 + 82p)\mathcal{M}^2 + 196p\mathcal{M} + 4|\mathcal{M}_\mathrm{B}|\mathcal{M}_\mathrm{R}^2 + 8\mathcal{M}_\mathrm{R}\mathcal{M}_f^2$

TABLE II: APPROXIMATED NUMBER OF FLOPs REQUIRED IN THE POSITIONING EKF.

Prediction step	$4\ell^3 + 3\ell^2$
Update step	$6m\ell^2 + 4\ell m^2 + m^3 + \ell^3 + 2\ell^2 + 2m\ell$

where $\hat{\sigma}^2$ is given in (25), and the partial derivatives can be obtained as described in Section III.

Finally, given the gradient (50) and the observed FIM (27), the a posteriori covariance matrix and mean estimates denoted as $\hat{\mathbf{C}}^+[n]$ and $\hat{\mathbf{s}}^+[n]$, respectively, can be obtained using the information form of the update step [15] as

$$\hat{\mathbf{C}}^+[n] = ((\hat{\mathbf{C}}^-[n])^{-1} + \mathbf{J}(\hat{\mathbf{s}}^-[n]))^{-1} \tag{51}$$

$$\hat{\mathbf{s}}^+[n] = \hat{\mathbf{s}}^-[n] + \hat{\mathbf{C}}^+[n]\mathbf{q}(\hat{\mathbf{s}}^-[n]), \tag{52}$$

where $\mathbf{q}(\hat{\mathbf{s}}^-[n])$ and $\mathbf{J}(\hat{\mathbf{s}}^-[n])$ are the gradient and the observed FIM, both evaluated at a priori mean estimate $\mathbf{s}^-[n]$. Since the log-likelihood function does not depend on the state-variables $\Delta\mathbf{\Theta}[n]$, the corresponding values in the gradient and in the observed FIM are zero. The computational complexity of the proposed DoA and ToA EKF is in the order of $\mathcal{O}(\mathcal{M}^2)$ as presented in Table I by the means of approximate floating point operations (FLOPs), thus implying that the complexity is dominated by the utilized bandwidth and the number of facilitated beams.

In this work, the initialization of the DoA and ToA estimation EKF is assisted by first selecting the most powerful beam-indices in a similar manner as in [18] and then calculating the ML estimates for the DoA angles and ToA based on the selected beams. Such a choice is stemming from the fact that the prior information of the TX location is not assumed to be available in the initialization and thus the information-based beam-selection approach cannot be facilitated as such. However, other initialization methods can be applied to the considered estimation problem as well. Specifically, we initialize the DoA and ToA estimates according to

$$\hat{\mathbf{\Theta}}^+[0] = \underset{\mathbf{\Theta}}{\arg\max} \; \hat{\ell}(\mathbf{Y}[0]|\mathbf{\Theta}) \tag{53}$$

$$[\hat{\mathbf{C}}^+[0]]_{1:3,1:3} = (\mathbf{J}(\hat{\mathbf{\Theta}}^+[0]))^{-1}, \tag{54}$$

where $[\cdot]_{(k:l,k:l)}$ denotes the submatrix between the indices k and l. Moreover, $\mathbf{J}(\hat{\mathbf{\Theta}}^+[0])$ is the observed FIM as given in (27) and it is evaluated at the obtained ML estimate. Without the loss of generality, the velocity components of the state-vector can be initialized after two consecutive state-estimates.

B. Positioning EKF

We next express the corresponding positioning EKF utilizing the estimated DoA angles and ToA values as measurements. In the considered positioning EKF, we assume that the velocity of the target is almost constant and therefore, the linear CWNA state-transition model is employed. Given the considered state-model, we can write the state-vector of the positioning EKF as

$$\mathbf{x}[n] = [x[n], y[n], z[n], \Delta x[n], \Delta y[n], \Delta z[n]]^\mathrm{T} \in \mathbb{R}^{6\times 1}, \tag{55}$$

where $\mathbf{p}_{\mathrm{TX}}[n] = [x[n], y[n], z[n]]^\mathrm{T}$ denotes the 3D Cartesian coordinates of the target, and $[\Delta x[n], \Delta y[n], \Delta z[n]]^\mathrm{T}$ denotes the corresponding velocity components. Due to the considered non-linear measurement model, through which the estimated DoA angles and the ToA are related to the state-vector, the state-space model of the positioning EKF can be thus written as

$$\mathbf{x}[n] = \mathbf{F}\mathbf{x}[n-1] + \mathbf{u}[n] \tag{56}$$

$$\hat{\mathbf{\Theta}}^+[n] = \mathbf{h}(\mathbf{x}[n]) + \mathbf{v}[n], \tag{57}$$

where the measurement $\hat{\mathbf{\Theta}}^+[n] = [\tau^+[n], \varphi^+[n], \vartheta^+[n]]^\mathrm{T}$ is the a posteriori estimate of the joint DoA and ToA estimation and tracking EKF, and the state-transition model matrix $\mathbf{F}$ is identical to that in (47). For the sake of simplicity, we here assume that the orientation of the RX is known and the RXs and the TX are synchronized, thus allowing us to analyze the pure positioning performance of the proposed method. However, such assumptions can be relaxed through joint estimation of the involved variables [8], [13], [17], [18], by employing a time difference of arrival (TDoA)-based approach, or by using a two-way communications [50]. Hence, the considered measurement model function $\mathbf{h}(\mathbf{x}[n])$ relates the obtained DoA and ToA estimates to the current state of the positioning EKF as given in (40). Moreover, the state-model driving noise $\mathbf{u}[n] \sim \mathcal{N}(0, \mathbf{U}[n])$, and the measurement noise $\mathbf{v}[n] \sim \mathcal{N}(0, [\hat{\mathbf{C}}^+[n]]_{(1:3,1:3)})$, where $[\hat{\mathbf{C}}^+[n]]_{(1:3,1:3)}$ is the covariance matrix of the DoA and ToA estimates obtained using the first-stage EKF.

For a given state-space model, the prediction step of the positioning EKF can be expressed as

$$\hat{\mathbf{x}}^-[n] = \mathbf{F}\hat{\mathbf{x}}^+[n-1] \tag{58}$$

$$\hat{\mathbf{P}}^-[n] = \mathbf{F}\hat{\mathbf{P}}^+[n-1]\mathbf{F}^\mathrm{T} + \mathbf{U}[n] \tag{59}$$

after which the DoA and ToA estimates can be fused in the update step of the proposed positioning EKF by employing the Kalman-gain form of the EKF. This is expressed as

$$\mathbf{K}[n] = \hat{\mathbf{P}}^-[n]\mathbf{H}[n]^\mathrm{T}(\mathbf{H}[n]\hat{\mathbf{P}}^-[n]\mathbf{H}[n]^\mathrm{T} + \hat{\mathbf{C}}^+[n])^{-1} \tag{60}$$

$$\mathbf{P}^+[n] = (\mathbf{I} - \mathbf{K}[n]\mathbf{H}[n])\hat{\mathbf{P}}^-[n] \tag{61}$$

$$\mathbf{x}^+[n] = \mathbf{x}^-[n] + \mathbf{K}[n](\hat{\mathbf{\Theta}}^+[n] - \mathbf{h}(\mathbf{x}^-[n])), \tag{62}$$

where $\mathbf{H}[n]$ is the Jacobian matrix of the measurement model function (40) evaluated at the a priori mean estimate $\mathbf{x}^-[n]$. The computational complexity of the positioning EKF is in the order of $\mathcal{O}(\ell^3)$ as given in Table II, where ℓ and m denote the dimensions of the state and observation vectors, respectively.

In this work, the positioning EKF is initialized in a similar manner as the proposed DoA and ToA EKF by adopting applicable ML estimates. After the two consecutive time-instants, the corresponding velocity components of the state can be initialized based on the difference between the first TX

(a) Urban canyon scenario
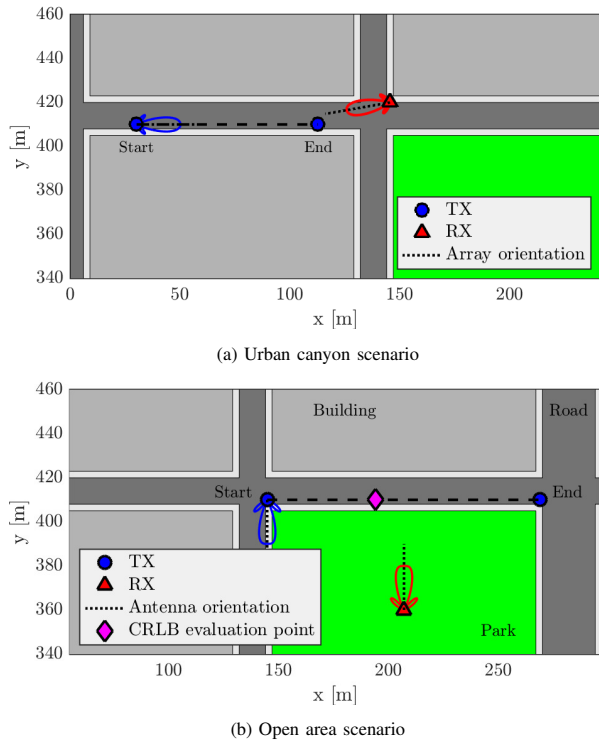

(b) Open area scenario

Fig. 3: Considered a) urban canyon and b) open area scenarios on top of the METIS Madrid grid environment. In both scenarios, the TX is moving from left to right with $30\,\mathrm{km/h}$, thus attempting to imitate a realistic vehicle movement in an urban environment. Furthermore, the TX location utilized in CRLB evaluation results in Fig. 5 is indicated in b).

TABLE III: BASELINE EVALUATION NUMEROLOGY AND ASSUMPTIONS

Parameter	Value
Carrier frequency	28 GHz
Subcarrier spacing	60 kHz
Bandwidth	25 MHz
Antenna elements	3GPP patch-elements [51]
TX array model	ULA (4 dual-polarized elements)
RX array model	URA (4×4 dual-polarized elements)
TX velocity	30 km/h
RX height	7.5 m
TX height	1.5 m
TX power	+10 dBm
SRS periodicity	periodic, every 100 ms
SRS resource allocation	one OFDM symbol, comb2 [36]

location estimates. In addition to the state-vector, the covariance estimate corresponding to the velocity components is initialized as a relatively large diagonal matrix.

V. EVALUATION SCENARIOS AND ASSUMPTIONS

In order to evaluate the performance of the proposed solution in terms of the channel parameter estimation and the subsequent positioning accuracy, simulations employing a full ray-tracing channel model [51], [52] are carried out on top of the METIS Madrid grid [53] with two different deployment scenarios, namely the street canyon and the open area scenarios depicted in Fig. 3a and Fig. 3b, respectively. The system-level numerology for the baseline evaluation scenarios is given in details in Table III. In particular, we assume that the TX is equipped with a uniform linear array (ULA) whereas the RX is equipped with a uniform rectangular array (URA), and the array elements at both link ends are modeled as dual-polarized patch-elements as described in [51], [54]. Moreover, the number of analog RF beams is assumed to be limited to the number of antenna array elements while it is noted that a denser grid of beams could also be considered in the estimation process with a cost of an increasing communication overhead.

In general, the network is considered to operate at $28\,\mathrm{GHz}$ carrier frequency with a subcarrier spacing of $60\,\mathrm{kHz}$. In the evaluations, the TX moving with $30\,\mathrm{km/h}$ along the predefined trajectory transmits SRS employing a single OFDM symbol in time-domain at every $100\,\mathrm{ms}$ (every 400th slot). In this work, we implement the comb2-configuration indicating that the SRS

occupies every second subcarrier in the frequency domain, whereas various bandwidth allocations can be considered depending on the selected resource block (RB) configuration [36], [55]. The user equipment (UE) transmit power is set to $+10\,\mathrm{dBm}$, and the noise spectral density and the base-station RX noise figure are defined to be $-174\,\mathrm{dBm/Hz}$ and $5\,\mathrm{dB}$, respectively. Such choices result in SNRs between $15\text{-}35\,\mathrm{dB}$ depending on the TX location as well as the chosen bandwidth and antenna array configurations. In the conducted simulations, the beam-selection threshold of the proposed beam-selection method is selected from the set $\alpha_{\mathrm{thr}} \in \{0.3, 0.5, 0.9\}$.

In evaluating the theoretical lower bounds and PEB of the considered multi-beam fusion, we utilize various configurations of the TX and RX antenna arrays while considering only the open area scenario in order to allow for more intuitive analysis by the means of active beam-directions and observed SNRs. Thereafter, the proposed EKF-based solution facilitating the proposed beam-selection method is compared with the similar EKF-based solution employing an BRSRP-based beam-selection method presented in [18] in the urban area scenario with more severe multi-path characteristics. The DoA and ToA estimation performance of the EKF-based methods are compared with the similar ML-based approaches which can be obtained in a straightforward manner by facilitating the provided gradient and Hessian matrix. In order to avoid divergence when employing consecutive estimates in ML algorithm, an initial guess in ML is drawn from the Gaussian distribution using the reference channel parameters as the mean at each time-instant. Besides the DoA and ToA estimation performance, the EKF-based solutions are finally compared by the means of 3D positioning performance in order to analyze the joint performance of the DoA and ToA estimates with a more convenient metric.

VI. NUMERICAL EVALUATIONS AND RESULTS

A. CRLBs Along Open Area Trajectory

We first analyze the impact of the proposed beam-selection method on the theoretical ToA and DoA angle estimation performance and the subsequent PEB along the open area trajectory as depicted in Fig. 3b. Besides illustrating the performance of the proposed method, we compare the theoretical bounds to those obtained using a digital beamformer at the RX side. Evaluating and analyzing the theoretical bounds only with the presented open area scenario is stemming mainly from the
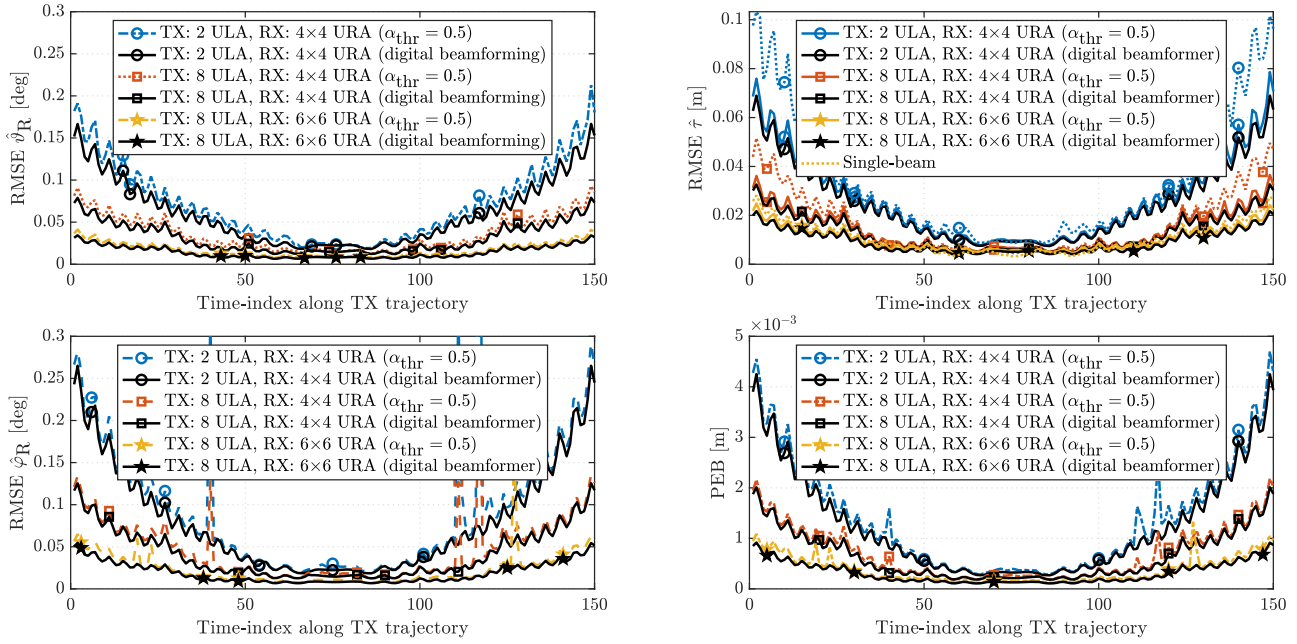
Fig. 4: Theoretical lower bounds for the DoA elevation and azimuth estimates (top row), ToA estimates (bottom row on the left), and positioning performance (bottom row on the right) along the trajectory in the considered open area scenario. Digital RX beamformer based bounds are also shown for reference.

fact that it allows for easier performance evaluation due to the more intuitive and symmetric SNR behavior along the trajectory. However, the performance of the proposed EKF-based solutions are still evaluated and compared to the performance of the reference ML estimator in the urban canyon scenario with more severe multipath propagation characteristics.

Since the theoretical ToA and DoA estimation performance is in general highly dependent on the observed SNR, which in turn reflects the current TX location as well as the deployed antenna configurations with the assumed bandwidth, we first provide the theoretical lower bounds along the TX trajectory with couple of antenna configurations and with $25\,\text{MHz}$ RS bandwidth. The obtained theoretical ToA estimation performance of the proposed multi-beam fusion is illustrated in Fig. 4 together with an estimator that considers only a single RX-beam in the ToA estimation. Despite the fact that the performance of the single-beam estimator is close to those of the proposed multi-beam and the digital beamformer approaches, e.g., in the RX boresight direction, a clear difference in the performance can be observed when the beam-directions between the TX and RX are not perfectly aligned. Such occasions can be observed in the case of TX having two dual-polarized elements (blue line), e.g., before and after the boresight direction of the RX, and in the beginning and in the end of the trajectory. Based on a rather straightforward derivation from (30), such an improvement in the ToA estimation performance of the proposed multi-beam method is proportional to the sum of beam-based SNRs over the selected RX-beams. This in turn implies that when the TX and RX beams are perfectly aligned and almost all the received signal power is in the single beam-pair, there is no larger difference in the ToA estimation between the single-beam and the proposed multi-beam approach. In general, we can observe that the performance of the proposed multi-beam

fusion is extremely close to that obtained by employing a digital beamformer at the RX. The small difference in the results is stemming from the limited spatial sampling of the channel in the assumed analog/RF beamforming scheme compared to that available with the digital beamformer.

Moreover, a similar SNR-dependency can be observed also in the theoretical DoA estimation performance as depicted in Fig. 4, where the same beam-selection threshold of $\alpha_{\text{thr}} = 0.5$ is employed. Occasional peaks in the theoretical estimation performance are stemming from bad realizations and the utilized projection approach, which in turn can be filtered out with proper sequential estimation solution, for instance. Based on the detailed analysis of the obtained results, at least two RX-beams are required in both azimuth and co-elevation direction in order to ensure theoretically feasible estimation accuracy without divergence in the variance of the corresponding estimators. Therefore, we are not able to provide the theoretical DoA estimation performance for a single-beam approach. In contrast to the ToA estimation performance that is related to the beams with high beam reference signal received powers (BRSRPs), the DoA estimation accuracy is also significantly affected by the set of chosen RX-beams as expected. For instance, the CRLB of the azimuth DoA becomes smaller when the RX beam-set contains various beams in the azimuth plane, whereas improvement in co-elevation estimation accuracy is depending on the number of beams in the elevation domain. In order to analyze the impact of theoretical ToA and DoA estimation errors on the corresponding positioning performance, we translate the angular and temporal bounds to the PEB, and the obtained results are depicted in Fig. 4. Based on the PEB results, the proposed two-stage estimation approach can in theory facilitate extremely accurate positioning performance especially with high SNRs, while being able to

provide almost as good performance as the reference system employing a digital beamformer at the RX-side.

Based on a more detailed analysis regarding the number of selected RX-beams, we observed that the number of the beams is related to the observed SNR in two ways. First, in a case of poor SNR due to a possible beam-mismatch, the proposed solution combines more RX-beams in order to improve the estimation performance as expected. Second, for almost perfect TX-RX beam-alignment, the number of RX-beams in the proposed beam-selection method increases in order to gather enough information to keep the DoA estimation performance at a reasonable level. This observation can be explained with the chosen orthogonal set of RX-beams which in turn means that the gain of secondary RX-beams is small in such beam-alignment occasions.

B. CRLBs with Different Beam-Selection Thresholds and TX-RX Antenna Configurations

We next provide theoretical lower bounds for the considered channel parameters as a function of various TX and RX configurations given a single TX location on the open area trajectory. In Fig. 5a, the bounds for the ToA and DoA estimates are shown with respect to the number of TX antenna elements. As expected, the theoretical estimation performance increases when the number of antenna elements at the TX is increased. This is mainly due to the improved TX-gain which in turn improves the received SNR. Based on a more detailed analysis of the obtained results, the number of selected RX-beams is essentially independent of the number of elements in the TX ULA. Such an observation is relatively intuitive, since only the best TX-beam is selected in the proposed beam-selection process for the subsequent estimation. With the considered 4×4 URA at the RX, the number of selected RX-beams in the proposed beam-selection at such TX location is either 4, 6, or 12 depending only on $\alpha_{\text{thr}} \in \{0.3, 0.5, 0.9\}$, correspondingly. As expected, the performance of the proposed method increases when the number of RX-beams increases, that is, when the value of α_{thr} is increased. However, the computational complexity of the algorithm is also proportional to the number of RX-beams as indicated in Section IV, which implies that the higher performance is achieved at the cost of added computational complexity. Here, it is to note that the provided results in Fig. 5a-5b are obtained in a single TX-location and instantaneous SNRs may have an impact on the observed results. For instance, the slightly degraded performance in Fig. 5a and 5b in the case of 8 TX elements compared to those of 5-7 TX-side elements is due to the bad geometry of the TX-beams, i.e., the TX with 4 and 8 elements have the same notch in the considered TX location due to beam-mismatch, whereas the cases with 5-7 TX elements are more favorable in this given TX location.

Finally, the performance bounds with varying number of antenna elements at the RX-side are depicted in Fig. 5b. Because of the better resolution in angular domain and the improved RX beamforming gain, the estimation performance of both ToA and DoA increases when the number of antenna elements at the RX increases. However, the proposed beam-selection method adapts to the increasing number of antenna elements, and therefore, the number of facilitated beams at the RX also increases as expected. Based on the overall results shown in Fig. 5a-5b, the estimation performance of the co-elevation angle is slightly better than that of the azimuth angle. This is mainly due to the fact that the considered antenna system at the TX results in wider beams in the elevation domain and the height of the TX is resulting in a favorable elevation-angle that is not typically aligned with the employed RX-beams, thus distributing the received signal power to some extent across several beams in elevation domain.

C. Estimation Performance of Proposed Two-Stage EKF

In addition to assessing the theoretical performance, we also evaluate and provide the actual estimation performance results with the proposed two-stage EKF employing the proposed beam-selection method in the more complex urban canyon scenario with more severe multi-path propagation compared to the previous open area evaluations. The proposed method is compared with the same EKF-based solution that utilizes the beam-selection method where the candidate RX-beams are determined based on a fixed number of most powerful BRSRP measurements as used, e.g., in [18]. The number of RX-beams in the compared beam-selection method is selected based on α_{thr} such that the number of beams is on average at the same level or higher compared to the proposed beam-selection method. Moreover, the obtained EKF-based results are compared with the similar ML-based solutions in order to see the impact of sequential filtering on the achieved estimation performance. In the conducted simulations, we also compared the proposed method against two adaptive beam-selection methods proposed in [17], [22]. However, the number of facilitated RX-beams in [22], relying on a likelihood-ratio test, was all the time significantly higher than in the proposed and the non-adaptive [18] methods, thus making the fair comparison unfeasible. In addition, the method in [17] resulted in a divergence of the first-stage EKF due to extremely small covariance estimates and the fact that the RX-side beamformer is not designed as a function of angles. Hence, the adaptive beam-selection methods in [17], [22] are not comparable with the proposed method. The DoA and ToA estimation results for all the considered simulation scenarios are depicted in Fig. 6 and Fig. 7, respectively.

Based on the obtained DoA estimation results depicted in Fig. 6, the proposed EKF-based method is outperforming the ML-based approach not only in terms the median error (black horizontal line), but also in terms of the 50% and 99% confidence intervals (illustrated with boxes and whiskers). Such observations are caused by better noise filtering of the EKFs which take into account also the available prior information in the estimation process. In all the considered simulation cases, the average number of selected RX-beams in the case of the proposed beam-selection method (denoted as Info. in the corresponding figures) is lower compared to the BRSRP-based beam-selection method, while the estimation performance of the proposed method is still better than that of the BRSRP-based approach. This in turn highlights the importance of considering the available observed information
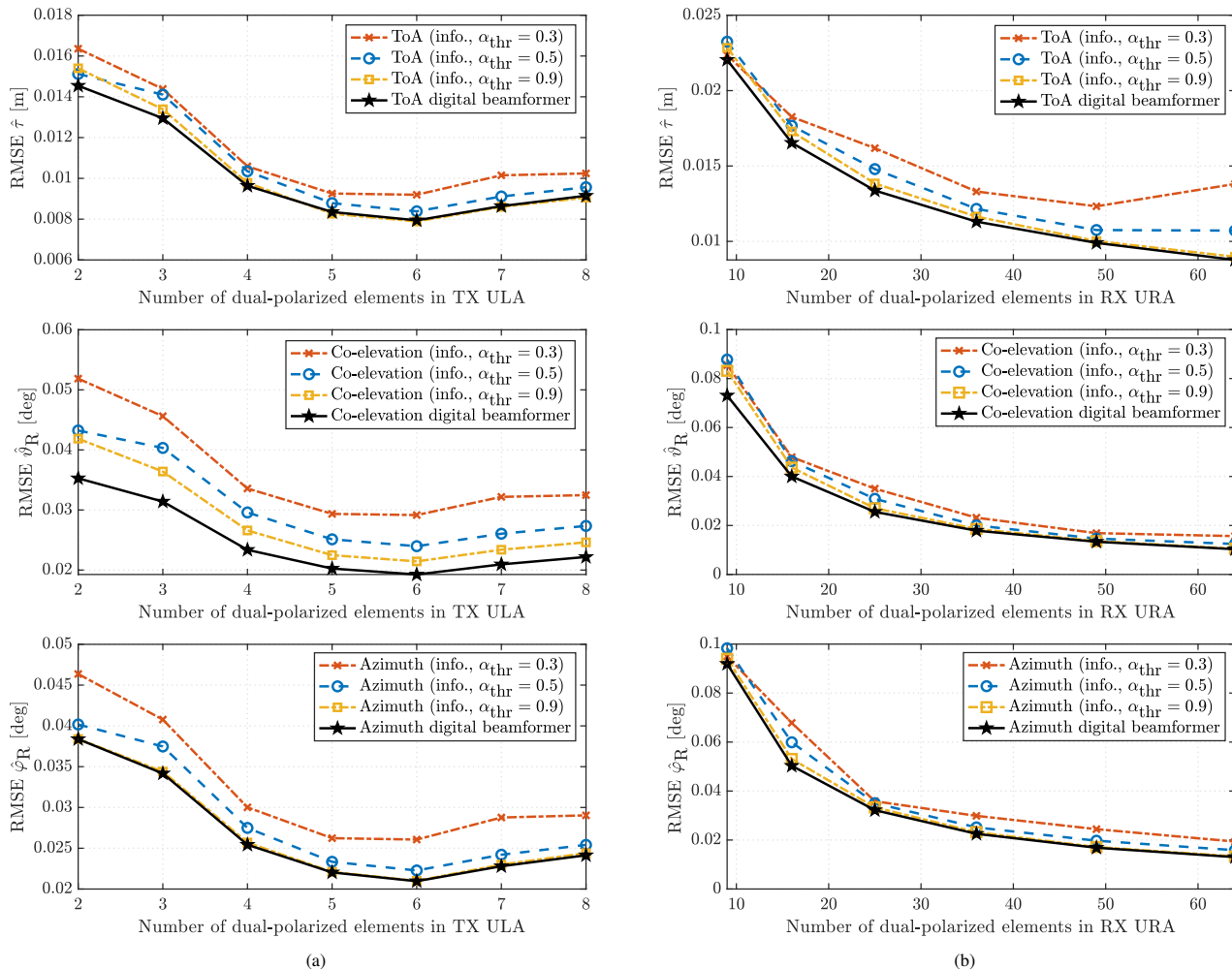
Fig. 5: Theoretical lower bounds for DoA angles and ToA as a function of the number of dual-polarized antenna elements a) at the TX-side and b) at the RX-side. Again, digital RX beamformer based bounds are shown for reference.

in selecting the candidate RX-beams in an adaptive manner for DoA estimation and subsequent positioning. However, the performance differences between the considered beam-selection methods are not significant with higher number of selected beams (case with 50% of beams or $\alpha_{\mathrm{thr}} = 0.5$) due to the fact that the most informative beams are most likely included in the candidate beam-sets of both beam-selection methods.

Furthermore, similar characteristics can be also observed from the obtained ToA estimation results, depicted in Fig. 7. At some time-instances with $100\,\mathrm{MHz}$ bandwidth, the BRSRP-based ML-estimator is not able to fully converge during the allowed Gauss-Newton iterations, thus resulting in the maximum ToA estimation errors around $4\,\mathrm{m}$. Overall, the observed ToA and DoA estimation results are extremely accurate. In general, increasing the bandwidth and the number of antenna elements at the RX side improves the estimation performance as expected. However, increasing the number of antenna elements at the RX also increases the number of available beams and therefore, also the number of selected candidate beams intuitively increases. Based on a more detailed analysis, we observed that at some time-instances the existing multi-path components cause a minor degradation in the

estimation performance. This in turn highlights the role of proper data-association, where beams corresponding to different propagation paths could be distinguished even better. Such a data-association and properly distinguished non-LoS (NLoS) paths could potentially improve the estimation performance of the considered system [56].

Finally, we evaluate the 3D positioning performance using the proposed two-stage EKF with both presented beam-selection methods, namely the proposed information-based method and the BRSRP-based method that is used in [18]. For the sake of simplicity, the performance of the previously presented ML estimators are ignored in the following positioning evaluations due to generally lower DoA and ToA estimation performance compared to the EKF-based solutions as seen in Figures 6-7. Based on the obtained 3D positioning results in Figure 8, the proposed beam-selection method is able to provide more accurate and consistent positioning performance than the reference method with the BRSRP-based beam-selection approach. In particular, significant positioning performance differences can be observed between the beam-selection methods, when the number of selected RX-beams is relatively low. This performance improvement is stemming
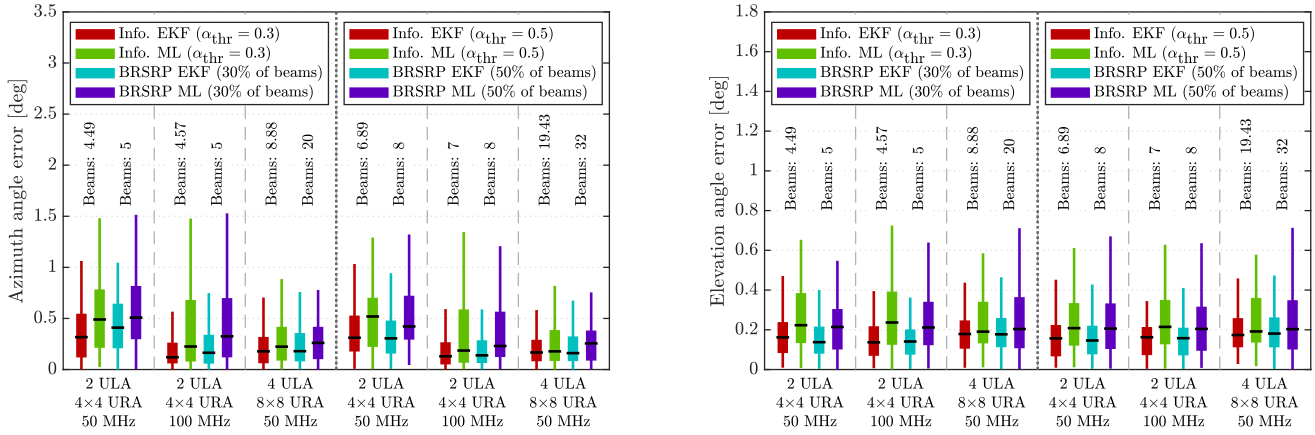
Fig. 6: Azimuth and elevation angle estimation performance of the proposed first-stage EKF-based solution with the proposed beam-selection method (denoted as Info.) in comparison with the non-adaptive beam-selection method (denoted as BRSRP) with the same EKF-based solution, and similar ML-based solutions. Three different simulation numerologies are considered in the urban canyon scenario while showing also two different beam-selection thresholds of $\alpha_{\text{thr}} = 0.3$ and $\alpha_{\text{thr}} = 0.5$. The average number of facilitated RX-beams with the considered beam-selection method are illustrated on top of the corresponding box-plots. The black horizontal line in each box-plot indicates the median value, whereas the wider box and the whiskers represent approximately the 50% and 99% confidence intervals, respectively.
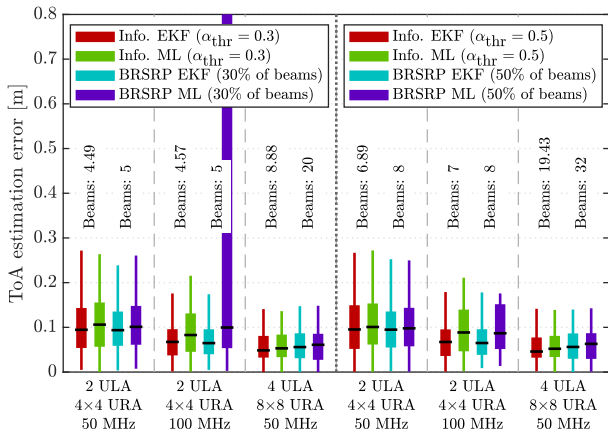

Fig. 7: ToA estimation performance of the proposed first-stage EKF-based solution with the proposed beam-selection method compared to another beam-selection method and to the similar ML-based solutions.
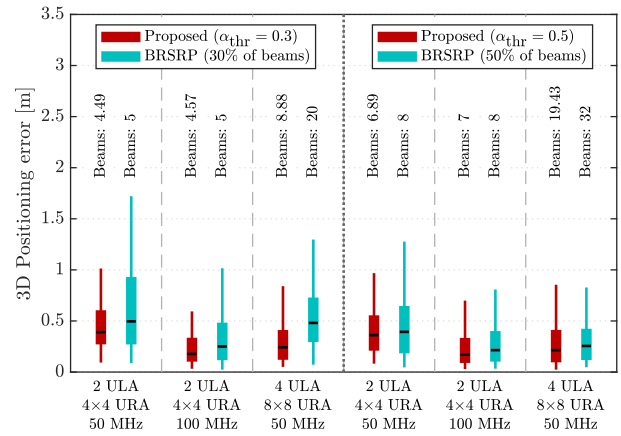
Fig. 8: 3D positioning performance results with the proposed beam-selection method and the proposed two-stage EKF compared with the same EKF-based solution under a different beam-selection method.

from fact that at some time-instances, the proposed beam-selection method can facilitate one or two beams more than the non-adaptive beam-selection method in order to improve the estimation performance at those unfavorable situations. However, the average number of utilized beams in the proposed beam-selection method is still lower than in the BRSRP-based method, which implies that the proposed beam-selection method is able to select more suitable RX-beams even when the number of selected beams is low.

D. Extension to Multi-RX Fusion and Joint TX Synchronization

Due to the cascaded structure of the proposed framework, the method can be extended in a straightforward manner to cover situations where several network nodes are in a favorable LoS condition towards a given TX. Such an extension enables not only improving the positioning performance through information fusion but also allows for estimating other closely associated parameters and variables including TX orientation and clock offsets [16]–[18] or employing TDoA-based positioning approach to compensate for the mutually asynchronous TX

and RXs clocks. In order to enable positioning in a practical and asynchronous network, we shortly present an extension of the proposed method to a joint TX positioning and synchronization where not only the position but also the clock offset and skew of the TX/UE are estimated by extending the state-vector (55) and by including the TX clock offset term in (57) in a similar manner as, e.g., in [16]. Furthermore, we assume that the RX nodes are mutually synchronized, as e.g. in time-sensitive networks [57]. Since extending the proposed solution to the multi-RX and joint TX synchronization scenarios is conceptually straightforward, we omit the further mathematical derivation for such cases in this paper. In short, the channel parameters are first estimated at individual LoS-RX nodes after which these estimates are then fused into the TX position and clock offset estimates in a sequential manner, conceptually similar to [16], [18].

In the considered multi-RX fusion scenario, a TX is assumed to be attached on top of a vehicle such that the orientation of the TX's antenna panel (8 ULA with 3GPP patch-elements) is parallel to the vehicle's heading direction. In particular, the
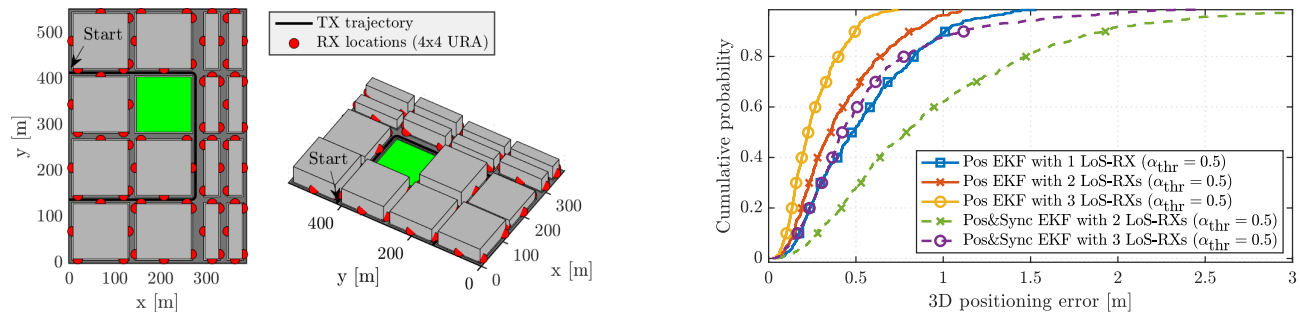
Fig. 9: Illustration of a multi-RX positioning scenario, with a TX attached to a vehicle moving according to a realistic acceleration profile (maximum speed of $40\,\mathrm{km/h}$) along a predefined trajectory at the METIS Madrid Map environment (left). The obtained positioning results (right) are illustrated for the positioning-only EKF (Pos EKF) and for the joint positioning and synchronization EKF (Pos&Sync EKF) in terms of cumulative distribution functions (CDFs).

TX transmits 8 SRS-beams using a $50\,\mathrm{MHz}$ bandwidth every $100\,\mathrm{ms}$ which are then received at LoS-RXs and processed using the proposed EKF-based scheme. In addition, the clock of the TX is modelled using the first-order auto-regressive model in a similar manner as described in [16], [58]. In order to model the propagation environment as realistically as possible, the channel is modelled using extensive ray-tracing simulations on top of the METIS Madrid Map environment depicted in Fig. 9.

As expected, based on the results shown in Fig. 9, the 3D positioning performance significantly increases when the DoA and ToA measurements from multiple available RX nodes are fused for TX positioning. Even with the joint positioning and synchronization case (denoted as Pos&Sync EKF in Fig. 9), better than $2\,\mathrm{m}$ positioning accuracy can be reached in more than 90% of the cases, while being able to obtain as accurate as $3.6\,\mathrm{ns}$ and $2.0\,\mathrm{ns}$ RMSEs for the TX clock offsets in the two and three LoS-RXs scenarios, respectively. For the sake of better visualization, we provide the same simulation scenario as a complementary online multimedia material at https://research.tuni.fi/wireless/research/positioning/TWC2021/.

VII. Conclusions

Building on the existing 5G NR specifications, we proposed a novel estimation and tracking solution for the DoA and ToA using an EKF-based solution for an mmWave system with emphasis on analog/RF beamforming capabilities. The proposed approach consists of a novel beam-selection method, which is able to determine a set of candidate RX-beams based on the available observed information in an adaptive manner. In order to evaluate the theoretical estimation performance of such multi-beam approaches, we derived and evaluated the CRLBs for the considered estimation problem, while also comparing to the corresponding bounds of digital RX beamforming-based system. Based on the extensive ray-tracing based evaluations covering both open area and street canyon type propagation scenarios on the METIS Madrid grid at 28 GHz, the proposed solution was shown to outperform the reference methods almost in all of the cases, while the most significant benefit was observed when the number of available RX-beams is low. In general, the ability to perform efficient positioning and tracking through RF-only beamforming is emphasized further when the

network center-frequencies are increasing towards the sub-THz regime, paving the way for future 6G systems.

References

[1] R. Di Taranto, S. Muppirisetty, R. Raulefs, D. Slock, T. Svensson, and H. Wymeersch, "Location-aware communications for 5G networks: How location information can improve scalability, latency, and robustness of 5G," *IEEE Signal Process. Mag.*, no. 6, pp. 102–112, Nov. 2014.

[2] M. Koivisto, A. Hakkarainen, M. Costa, P. Kela, K. Leppänen, and M. Valkama, "High-Efficiency Device Positioning and Location-Aware Communications in Dense 5G Networks," *IEEE Communications Magazine*, vol. 55, no. 8, pp. 188–195, Jul. 2017.

[3] 3GPP TR 22.872, "Study on positioning use cases," 2018. [Online]. Available: http://www.3gpp.org/dynareport/22872.htm

[4] Y. Lu, M. Koivisto, J. Talvitie, M. Valkama, and E. S. Lohan, "Positioning-Aided 3D Beamforming for Enhanced Communications in mmWave Mobile Networks," *IEEE Access*, vol. 8, Mar. 2020.

[5] 3GPP TR 38.855, "Study on NR positioning support," 2019. [Online]. Available: http://www.3gpp.org/dynareport/38855.htm

[6] N. Bhushan *et al.*, "Network densification: The dominant theme for wireless evolution into 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 82–89, Feb. 2014.

[7] A. Guerra, F. Guidi, and D. Dardari, "Single-Anchor Localization and Orientation Performance Limits Using Massive Arrays: MIMO vs. Beamforming," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5241–5255, Aug. 2018.

[8] 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, Mar. 2018.

[9] H. Wymeersch, G. Seco-Granados, G. Destino, D. Dardari, and F. Tufvesson, "5G mmWave Positioning for Vehicular Networks," *IEEE Wireless Communications*, vol. 24, no. 6, pp. 80–86, Dec. 2017.

[10] C. De Lima *et al.*, "Convergent communication, sensing and localization in 6G systems: An overview of technologies, opportunities and challenges," *IEEE Access*, vol. 9, pp. 26 902–26 925, Jan. 2021.

[11] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An Overview of Signal Processing Techniques for Millimeter Wave MIMO Systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[12] Z. Marzi, D. Ramasamy, and U. Madhow, "Compressive Channel Estimation and Tracking for Large Arrays in mm-Wave Picocells," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 514–527, Apr. 2016.

[13] J. Talvitie, M. Valkama, G. Destino, and H. Wymeersch, "Novel Algorithms for High-Accuracy Joint Position and Orientation Estimation in 5G mmWave Systems," in *2017 IEEE Globecom Workshops (GC Wkshps)*, Dec. 2017, pp. 1–7.

[14] F. Roemer, H. Becker, and M. Haardt, "Analytical performance assessment for multi-dimensional Tensor-ESPRIT-type parameter estimation algorithms," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Mar. 2010, pp. 2598–2601.

[15] J. Salmi, A. Richter, and V. Koivunen, "Detection and Tracking of MIMO Propagation Path Parameters Using State-Space Approach," *IEEE Trans. on Signal Process.*, vol. 57, no. 4, pp. 1538–1550, Apr. 2009.

[16] 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.

[17] 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.

[18] E. Rastorgueva-Foi, M. Costa, M. Koivisto, K. Leppänen, and M. Valkama, "User Positioning in mmW 5G Networks Using Beam-RSRP Measurements and Kalman Filtering," in *21st International Conference on Information Fusion, (FUSION)*, Sep. 2018, pp. 1150–1156.

[19] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, 1st ed. Prentice Hall, Apr. 1993.

[20] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*, 1st ed. Hoboken, N.J: Wiley-Interscience, Jun. 2006.

[21] H. Holma, A. Toskala, and T. Nakamura, *5G Technology: 3GPP New Radio*. Wiley, 2020.

[22] E. Rastorgueva-Foi, M. Costa, M. Koivisto, K. Leppänen, and M. Valkama, "Dynamic Beam Selection for Beam-RSRP Based Direction Finding in mmW 5G Networks," in *2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Sep. 2018, pp. 1–6, iSSN: 2162-7347.

[23] K. Witrisal *et al.*, "High-Accuracy Localization for Assisted Living: 5G systems will turn multipath channels from foe to friend," *IEEE Signal Processing Magazine*, vol. 33, no. 2, pp. 59–70, Mar. 2016.

[24] Z. Qi and W. Liu, "Three-dimensional millimetre-wave beam tracking based on smart phone sensor measurements and direction of arrival/time of arrival estimation for 5G networks," *IET Microwaves, Antennas & Propagation*, vol. 12, no. 3, pp. 271–279, Sep. 2018.

[25] M. Ruble and I. Güvenc, "Wireless localization for mmWave networks in urban environments," *EURASIP Journal on Advances in Signal Processing*, vol. 2018, no. 1, Jun. 2018.

[26] N. Garcia, H. Wymeersch, E. G. Ström, and D. Slock, "Location-aided mm-wave channel estimation for vehicular communication," in *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Aug. 2016, pp. 1–5.

[27] S. Shaham, M. Kokshoorn, M. Ding, Z. Lin, and M. Shirvanimoghaddam, "Extended kalman filter beam tracking for millimeter wave vehicular communications," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, Jul. 2020, pp. 1–6.

[28] A. Ali, N. González-Prelcic, and R. W. Heath, "Millimeter Wave Beam-Selection Using Out-of-Band Spatial Information," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1038–1052, Nov. 2018.

[29] F. Maschietti, D. Gesbert, P. de Kerret, and H. Wymeersch, "Robust Location-Aided Beam Alignment in Millimeter Wave Massive MIMO," in *IEEE Global Communications Conference (GLOBECOM)*, Jan. 2017, pp. 1–6.

[30] R. Koirala, B. Denis, B. Uguen, D. Dardari, and H. Wymeersch, "Localization Optimal Multi-user Beamforming with multi-carrier mmWave MIMO," in *IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Dec. 2018, pp. 1–7.

[31] 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.

[32] 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.

[33] Y. Shen and M. Z. Win, "Fundamental Limits of Wideband Localization - Part I: A General Framework," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4956–4980, Oct. 2010.

[34] M. Rath, J. Kulmer, E. Leitinger, and K. Witrisal, "Single-Anchor Positioning: Multipath Processing With Non-Coherent Directional Measurements," *IEEE Access*, May 2020.

[35] T. Wilding, S. Grebien, U. Mühlmann, and K. Witrisal, "Accuracy bounds for array-based positioning in dense multipath channels," *Sensors*, vol. 18, no. 12, p. 4249, Dec 2018. [Online]. Available: http://dx.doi.org/10.3390/s18124249

[36] 3GPP TS 38.211, "NR; Physical channels and modulation," 2019. [Online]. Available: http://www.3gpp.org/dynareport/38211.htm

[37] A. Ghosh, A. Maeder, M. Baker, and D. Chandramouli, "5G Evolution: A View on 5G Cellular Technology Beyond 3GPP Release 15," *IEEE Access*, vol. 7, pp. 127 639–127 651, Sep. 2019.

[38] A. A. Khalek, R. W. Heath, S. Rajagopal, S. Abu-Surra, and J. C. Zhang, "Cross-polarization RF precoding to mitigate mobile misorientation and polarization leakage," in *2014 IEEE 11th Consumer Communications and Networking Conference (CCNC)*, Jan. 2014, pp. 581–586, iSSN: 2331-9860.

[39] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially Sparse Precoding in Millimeter Wave MIMO Systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[40] A. Richter, "Estimation of radio channel parameters: Models and algorithms," Ph.D. dissertation, Ilmenau University of Technology, http://www.db-thueringen.de/servlets/DerivateServlet/Derivate-7407/ilm1-2005000111.pdf, 2005.

[41] M. Landmann and G. D. Galdo, "Efficient antenna description for MIMO channel modelling and estimation," in *7th European Conference on Wireless Technology, 2004*, Oct. 2004, pp. 217–220.

[42] M. Costa, A. Richter, and V. Koivunen, "DoA and Polarization Estimation for Arbitrary Array Configurations," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2330–2343, May 2012.

[43] 3GPP TS 38.215, "NR; Physical layer measurements," 2020. [Online]. Available: http://www.3gpp.org/dynareport/38215.htm

[44] M. Viberg, B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Trans. Signal Process.*, vol. 39, no. 11, pp. 2436–2449, Nov. 1991.

[45] G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," *SIAM Journal on Numerical Analysis*, vol. 10, no. 2, pp. 413–432, Apr. 1973.

[46] S. Särkkä, *Bayesian filtering and smoothing*. Cambridge University Press, 2013.

[47] R. Zekavat and R. M. Buehrer, *Handbook of Position Location: Theory, Practice and Advances*, 2nd ed. Wiley-IEEE Press, 2019.

[48] S. Sand, A. Dammann, and C. Mensing, *Positioning in Wireless Communications Systems*, 1st ed. Wiley, 2014.

[49] B. M. Bell and F. W. Cathey, "The iterated Kalman filter update as a Gauss-Newton method," *IEEE Transactions on Automatic Control*, vol. 38, no. 2, pp. 294–297, Feb. 1993.

[50] T. E. Abrudan, A. Haghparast, and V. Koivunen, "Time synchronization and ranging in ofdm systems using time-reversal," *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 12, pp. 3276–3290, 2013.

[51] 3GPP TR 38.901, "Study on channel model for frequencies from 0.5 to 100 GHz," 2020. [Online]. Available: http://www.3gpp.org/dynareport/38901.htm

[52] METIS, "D6.1 Simulation guidelines," Oct. 2013. [Online]. Available: https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D6.1_v1.pdf

[53] METIS, "D1.4 Channel models," Feb. 2015. [Online]. Available: https://www.metis2020.com/wp-content/uploads/METIS_D1.4_v3.pdf

[54] M. Rebato, L. Resteghini, C. Mazzucco, and M. Zorzi, "Study of Realistic Antenna Patterns in 5G mmWave Cellular Scenarios," in *2018 IEEE International Conference on Communications (ICC)*, May 2018, pp. 1–6.

[55] 3GPP TR 38.101-2, "NR; User Equipment (UE) radio transmission and reception; Part 2: Range 2 Standalone," 2020. [Online]. Available: http://www.3gpp.org/dynareport/38101-2.htm

[56] 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.

[57] S. Ruffini, M. Johansson, B. Pohlman, and M. Sandgren, "Ericsson: 5G Synchronization Requirements and Solutions (white paper)," Jan. 2021.

[58] H. Kim, X. Ma, and B. 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.

Mike Koivisto was born in Rauma, Finland, in 1989. He received the M.Sc. degree in mathematics from the Tampere University of Technology (TUT), Finland, in 2015, where he is currently pursuing the Ph.D. degree. From 2013 to 2016, he was a Research assistant with TUT. He is currently a Researcher with the Electrical Engineering unit at Tampere University. His research interests include positioning with an emphasis on network-based positioning and the utilization of location information in future mobile networks.

Jukka Talvitie was born in HyvinkÃ¤Ã¤, Finland, in 1981. He received the M.Sc. degree in automation engineering and the Ph.D. degree in computing and electrical engineering from the Tampere University of Technology (TUT), Finland, in 2008 and 2016, respectively. He is currently a University Lecturer at Tampere University, Finland. In addition to academic research, he has involved several years in industry-based research and development projects on wide variety of research topics, including radio signal waveform design, network based positioning and next generation WLAN, and cellular system design. His main research interests include signal processing for communications, wireless locations techniques, radio signal waveform design, and radio network system level development.

Elizaveta Rastorgueva-Foi received M.Sc. in electrical engineering from Tampere University (TAU), Finland, in 2019, where she is currently working towards Ph.D. degree. Her research interests include statistical signal processing, positioning and location-aware communications in mobile networks with an emphasis on 5G and beyond.

Yi Lu received the M.Sc degree in mobile communications (with distinction) from Heriot-Watt University, UK, in 2012. From 2014 to 2017, he worked as a Research Assistant at Universitat Autonoma de Barcelona (UAB), Spain. Since March 2018, he has been employed as a Doctoral Researcher at the Electrical Engineering unit of Tampere University, Finland. In addition, he serves as a TPC member of IEEE CIC ICCC since 2019. His research interests include the development of network-centric positioning system and location-aware communications scheme for the Industrial IoT system.

Mikko Valkama was born in Pirkkala, Finland, in 1975. He received the M.Sc. (Tech.) and D.Sc. (Tech.) degrees (Hons.) in electrical engineering from Tampere University of Technology (TUT), Finland, in 2000 and 2001, respectively. In 2002, he received the Best Ph.D. Thesis Award from the Finnish Academy of Science and Letters for his dissertation entitled Advanced I/Q Signal Processing for Wideband Receivers: Models and Algorithms. In 2003, he was a Visiting Post-Doc Researcher with the Communications Systems and Signal Processing Institute, SDSU, San Diego, CA, USA. He is currently a Full Professor and the Head of the Electrical Engineering Unit at the newly established Tampere University (TAU), Finland. His general research interests include radio communications, radio localization, and radio-based sensing, with emphasis on 5G and beyond mobile radio networks.