

Successive Pose Estimation and Beam Tracking for mmWave Vehicular Communication Systems

Cen Liu*, Guangxu Zhu[†], Fan Liu[‡], Yuanwei Liu[§] and Kaibin Huang[¶]

*National University of Singapore, Singapore [†]Shenzhen Research Institute of Big Data, Shenzhen, China

[‡]Southern University of Science and Technology, Shenzhen, China [§]Queen Mary University of London, London, U.K.

[¶]The University of Hong Kong, Hong Kong

Email: liucen@u.nus.edu, gxzhu@sribd.cn, liuf6@sustech.edu.cn, yuanwei.liu@qmul.ac.uk, huangkb@eee.hku.hk

Abstract—The millimeter wave (mmWave) radar sensing-aided communications in vehicular mobile communication systems is investigated. To alleviate the beam training overhead under high mobility scenarios, a successive pose estimation and beam tracking (SPEBT) scheme is proposed to facilitate mmWave communications with the assistance of mmWave radar sensing. The proposed SPEBT scheme first resorts to a Fast Conservative Filtering for Efficient and Accurate Radar odometry (Fast-CFEAR) approach to estimate the vehicle pose consisting of 2-dimensional position and yaw from radar point clouds collected by mmWave radar sensor. Then, the pose estimation information is fed into an extend Kalman filter to perform beam tracking for the line-of-sight channel. Owing to the intrinsic robustness of mmWave radar sensing, the proposed SPEBT scheme is capable of operating reliably under extreme weather/illumination conditions and large-scale global navigation satellite systems (GNSS)-denied environments. The practical deployment of the SPEBT scheme is verified through rigorous testing on a real-world sensing dataset. Simulation results demonstrate that the proposed SPEBT scheme is capable of providing precise pose estimation information and accurate beam tracking output, while reducing the proportion of beam training overhead to less than 5% averagely.

Index Terms—Sensing-aided communications, millimeter wave, pose estimation, beam tracking, radar odometry, extended Kalman filter.

I. INTRODUCTION

In the past few years, millimeter wave (mmWave) technology and its promising applications have been getting growing interest from both wireless communications and robotics communities. On one hand, mmWave communications is known as a key technology in 5G New Radio (NR) and beyond 5G cellular network [1], [2]. The extremely high frequency (EHF) band of mmWave is able to provide enormous amount of spectrum resources to meet the needs of diverse user applications. On the other hand, mmWave radar sensing plays an indispensable role in modern autonomous robotic systems. Owing to its intrinsic robustness towards bad illumination and harsh weather conditions, mmWave radar sensing is capable of supplying more stable and reliable perception information than vision and LiDAR solutions under global navigation satellite systems (GNSS)-denied environments [3]–[5].

To counteract high attenuation of free space propagation as well as blockage penetration in EHF band, mmWave communications generally resorts to large-scale multi-input multi-output (MIMO) antennas in forming high directional

beams aligned to the wireless terminals via transmit precoding and receive beamforming [6]. Nevertheless, it tends to incur unbearable beam training overhead in vehicular mobile communication systems since the vehicle moves at a high speed and thus the position of mobile station (MS) changes intensively. To address this issue, beam tracking serves as a commonly-adopted technique to alleviate the burden incurred by frequent beam training [7].

Several works related to the beam tracking in mmWave mobile communication systems have emerged in recent years [8]–[13]. The authors of [8] proposed a dual-timescale model on characterizing abrupt channel changes as well as tracking slow variations of angles of departure (AoDs) and angles of arrival (AoAs) for multiple dominant paths by assuming the path gains are priori known. On the contrary, in [9], the AoD/AoA and the path gain of a single line-of-sight (LoS) path are tracked using extended Kalman filter (EKF) algorithm through training a single beam pair at each timeslot. Assuming parallelism between the antenna arrays of base station (BS) and MS, a joint minimum mean square error (MMSE) robust beamforming and EKF tracking scheme is proposed in [10] to minimize the mismatch on beamforming angles between BS and MS. In [11], the authors proposed a second-order AoD/AoA variation model by taking the slow variations of AoD/AoA angular velocity into account. [12] investigated the EKF-based beam tracking approach for vehicular communication systems by incorporating 1-dimensional (1D) position and velocity of the vehicle into state variables instead of tracking AoD/AoA directly. The recent work [13] utilized 2-dimensional (2D) position and angular velocity of MS to back up the EKF tracking approach under high mobility scenarios, and offered insights towards the formulation of state evolution models with the assistance of spatial information.

Inspired by the vision of integrated sensing and communications (ISAC) [14]–[16], a successive pose estimation and beam tracking (SPEBT) scheme is proposed in this paper to alleviate the beam training overhead for mmWave vehicular mobile communication systems. With the knowledge of pose information consisting of vehicle position and yaw offered by mmWave radar sensing, both BS and MS are capable of tracking the dynamically changing channel and forming directional beams to facilitate mmWave communications. To the best of our knowledge, this is the first work investigating

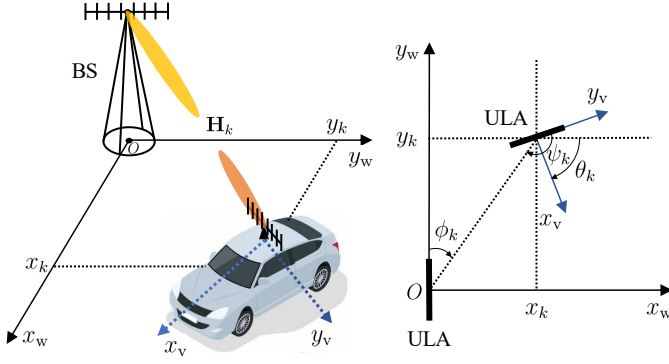


Fig. 1. mmWave vehicular communications system. Left: 3D view of the MIMO system. Right: 2D bird's eye view of the MIMO system.

the mmWave radar sensing-aided mmWave communications for vehicular mobile communication systems under large-scale GNSS-denied environments. The main contributions of this paper are summarized below:

- 1) A SPEBT scheme is proposed to reduce the beam training overhead by performing mmWave beam tracking in light of pose estimation offered by mmWave radar sensing in a vehicular mobile communication system, which can normally work under extreme conditions and GNSS-denied environments.
- 2) An open-source real-world mmWave sensing platform is used to verify the feasibility and effectiveness for the practical deployment of the proposed SPEBT scheme.

Notations: $(\cdot)^H$, $(\cdot)^T$ and $(\cdot)^*$ stand for the Hermitian, the transpose and the conjugate operations, respectively. $|\cdot|$ represents the modulus of a complex number or the cardinality of a set. $\|\cdot\|$ represents the ℓ_2 norm of a real vector. $\text{diag}(\cdot)$ denotes the diagonal matrix whose diagonals are the elements of the input vector. $\mathbb{SE}(2)$ and $\mathbb{SO}(2)$ stand for the 2D special Euclidean group and 2D special orthogonal group.

II. SYSTEM MODEL

A. MmWave Vehicular Communications Model

We consider a mmWave MIMO vehicular communications system, as illustrated in Fig. 1. The BS and vehicle are equipped with uniform linear arrays (ULAs) of N and M antennas, respectively. The mmWave wireless downlink can be described by a time-varying geometric channel model $\mathbf{G}_k \in \mathbb{C}^{M \times N}$ as follows [17]

$$\mathbf{G}_k = \sum_{l=1}^{L_k} \alpha_{l,k} \mathbf{a}_r(\psi_{l,k}) \mathbf{a}_t^H(\phi_{l,k}) \quad (1)$$

where $\alpha_{l,k}$, $\phi_{l,k}$ and $\psi_{l,k}$ are the complex path gain, AoD and AoA of the l -th path among total L_k paths at timeslot k .

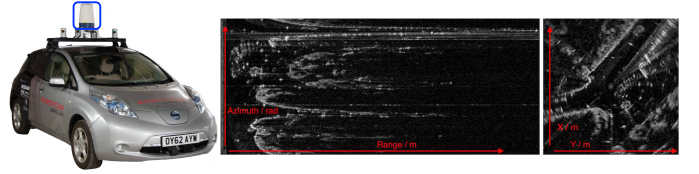


Fig. 2. mmWave vehicular sensing platform [19]. Left: a Navtech CTS350-X mmWave FMCW radar (marked with a blue box) mounted on a Nissan LEAF automobile. Middle: a sample of raw radar point cloud in polar form. Right: a sample of raw radar point cloud in Cartesian form.

$\mathbf{a}_t(\cdot) \in \mathbb{C}^N$ and $\mathbf{a}_r(\cdot) \in \mathbb{C}^M$ are the transmit and receive array response vectors given by

$$\mathbf{a}_t(\phi) = \frac{1}{\sqrt{N}} \left[1, e^{-j \frac{2\pi d}{\lambda} \cos \phi}, \dots, e^{-j(N-1) \frac{2\pi d}{\lambda} \cos \phi} \right]^T \quad (2)$$

$$\mathbf{a}_r(\psi) = \frac{1}{\sqrt{M}} \left[1, e^{-j \frac{2\pi d}{\lambda} \cos \psi}, \dots, e^{-j(M-1) \frac{2\pi d}{\lambda} \cos \psi} \right]^T \quad (3)$$

with the antenna spacing d and carrier wavelength λ .

Because of the sparsity on spatial scattering in angular domain and high free-space path loss [18], mmWave channel is dominated by LoS component which is deemed as the target of beam tracking [9]–[13]. The time-varying LoS channel model $\mathbf{H}_k \in \mathbb{C}^{M \times N}$ can be written as

$$\mathbf{H}_k \triangleq \mathbf{G}_k^{\text{LoS}} = \alpha_k \mathbf{a}_r(\psi_k) \mathbf{a}_t^H(\phi_k) \quad (4)$$

where the path index is omitted for simplicity. Both BS and the vehicle adopt analog beamforming to enhance directivity of transmit and receive beam patterns. The signal received by the vehicle at timeslot k can be written as

$$z_k = \mathbf{g}_k^H \mathbf{H}_k \mathbf{f}_k b_k + v_k \quad (5)$$

where b_k and v_k denote the transmitted symbol and additive white Gaussian noise (AWGN), respectively, with $\mathbb{E}[|b_k|^2] = 1$ and $v_k \sim \mathcal{CN}(0, \sigma_v^2)$. $\mathbf{f}_k \in \mathbb{C}^N$ denotes an analog precoder at BS, and $\mathbf{g}_k \in \mathbb{C}^M$ denotes an analog combiner at the vehicle.

B. MmWave Vehicular Sensing Model

The mmWave vehicular mobile system applies *Oxford Radar RobotCar* [19] as a sensing platform to perform vehicle pose estimation with the purpose of assisting beam tracking for mmWave communications. As demonstrated in Fig. 2, a mmWave frequency-modulated continuous-wave (FMCW) radar sensor mounted on a vehicle emits radio waves to the ambient environment and receives radio waves reflected from the surrounding objects. Through detecting the reflective radio power returns captured by radar point cloud, the radar sensor is capable of inferring relative location (i.e., range and azimuth) of the surrounding objects with respect to the sensor itself.

As depicted in Fig. 1, $x_w y_w$ -axis indicate the world coordinate frame \mathcal{F}_w , while $x_v y_v$ -axis indicate the vehicle coordinate frame \mathcal{F}_v . By processing raw radar point cloud collected at timeslot k , the relative position vector $\mathbf{r}_k = [x_k, y_k]^T$ and yaw θ_k of \mathcal{F}_v with respect to \mathcal{F}_w (or in other words, the absolute pose of the vehicle) can be estimated, which will be elaborated in Section III-A.

III. SUCCESSIVE POSE ESTIMATION AND BEAM TRACKING

In this section, we propose a SPEBT scheme for reducing the beam training overhead incurred in mmWave vehicular mobile communication systems. The proposed scheme involves two stages, namely 1) estimating the vehicle pose based on Fast-CFEAR approach and 2) carrying out beam tracking subsequently by employing EKF.

A. Fast-CFEAR Based Pose Estimation

Based on the state-of-the-art (SOTA) Conservative Filtering for Efficient and Accurate Radar odometry (CFEAR) framework [20], an approach named Fast-CFEAR is proposed to estimate the vehicle pose consisting of 2D position and yaw. The basic idea of Fast-CFEAR is to remove noise in raw radar point clouds resorting to conservative filtering, such that the underlying noise-less sparse representations can be recovered. As a result, the vehicle pose can be estimated via deploying point cloud registration between consecutive sparse representations. Fast-CFEAR is capable of offering accurate 2D position and yaw estimations with low computational complexity in real time.

The Fast-CFEAR approach can be decomposed into the following four stages:

1) *K-Strongest Conservative Filtering*: For each azimuth of a radar point cloud in polar form, only points with the K strongest pixel values exceeding a threshold κ_{\min} are preserved. The point cloud after conservative filtering is denoted by $\mathcal{P}_F \subset \mathbb{R}^2$. At the risk of possibly missing landmarks with low pixel values, the conservative filtering technique reduces the noise originated from false positive detections.

2) *Downsampling*: Grid cells with side length d_D/f_D are applied to downsample \mathcal{P}_F in Cartesian form. Only the centroid of each grid cell is kept in the downsampled point cloud $\mathcal{P}_D \subset \mathbb{R}^2$. d_D is a distance parameter controlling the number of points in estimating the surface normal and f_D is a resampling factor adjusting the density of \mathcal{P}_D .

3) *Estimating Oriented Surface Points*: For each point $\mathbf{p}_i = [p_i^{[1]}, p_i^{[2]}]^T \in \mathcal{P}_D$, its all neighborhoods $\mathbf{p}_j = [p_j^{[1]}, p_j^{[2]}]^T \in \mathcal{P}_F$ within a radius d_D centering at \mathbf{p}_i are utilized to estimate an oriented surface point $\{\boldsymbol{\mu}_i, \mathbf{n}_i\}$. The surface point $\boldsymbol{\mu}_i$ is computed by sample mean $\frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} \mathbf{p}_j$, where $\mathcal{N}_i = \{j \mid \|\mathbf{p}_i - \mathbf{p}_j\| \leq d_D, \forall \mathbf{p}_i \in \mathcal{P}_D, \forall \mathbf{p}_j \in \mathcal{P}_F\}$. The surface normal \mathbf{n}_i is defined as the smallest eigenvector of sample covariance matrix

$$\boldsymbol{\Sigma}_i = \frac{1}{|\mathcal{N}_i| - 1} \sum_{j \in \mathcal{N}_i} \begin{bmatrix} (p_j^{[1]} - \mu_i^{[1]})^2 & c_i \\ c_i & (p_j^{[2]} - \mu_i^{[2]})^2 \end{bmatrix} \quad (6)$$

where $c_i = (p_j^{[1]} - \mu_i^{[1]})(p_j^{[2]} - \mu_i^{[2]})$.

Through estimating the oriented surface points $\{\boldsymbol{\mu}_i, \mathbf{n}_i\}$ for all points in \mathcal{P}_D , a sparse representation $\mathcal{M} = \{\{\boldsymbol{\mu}_i, \mathbf{n}_i\} \mid \forall \mathbf{p}_i \in \mathcal{P}_D\}$ of a radar point cloud can be obtained. The readers are suggested to refer to a sample sparse representation visualized by Fig. 5 in [20] for understanding intuitively.

4) *Point Cloud Registration*: Rather than resorting to multiple keyframes registration adopted in [20] with high computational complexity, Fast-CFEAR registers the current representation \mathcal{M}_k generated at k -th timeslot to the previous one \mathcal{M}_{k-1} generated at $(k-1)$ -th timeslot directly. The vehicle pose with respect to $\vec{\mathcal{F}}_w$ at timeslot k can be represented by a homogeneous transformation matrix $\mathbf{T}_k \in \mathbb{SE}(2)$ [21]

$$\mathbf{T}_k = \begin{bmatrix} \mathbf{R}_k & \mathbf{r}_k \\ \mathbf{0}^T & 1 \end{bmatrix}, \quad \mathbf{R}_k = \begin{bmatrix} \cos \theta_k & -\sin \theta_k \\ \sin \theta_k & \cos \theta_k \end{bmatrix} \quad (7)$$

where $\mathbf{R}_k \in \mathbb{SO}(2)$ is a 2D rotation matrix.

The point cloud registration can be formulated as an unconstrained minimization problem on *point-to-line* (P2L) distance

$$\min_{\mathbf{T}_k} f_{\text{P2L}}(\mathbf{T}_k \mid \mathbf{T}_{k-1}, \mathcal{M}_{k-1}, \mathcal{M}_k) \quad (8)$$

$$\Leftrightarrow \min_{\mathbf{r}_k, \theta_k} \sum_{\forall i \in \mathcal{M}_k} \mathcal{L}_\delta(\mathbf{n}_i^T (\Delta \mathbf{R}_k \boldsymbol{\mu}_i + \Delta \mathbf{r}_k - \boldsymbol{\mu}_j)) \quad (9)$$

which means that we seek to estimate the vehicle pose $\{\hat{\mathbf{r}}_k, \hat{\theta}_k\}$ at timeslot k for best registering the consecutive sparse representations \mathcal{M}_{k-1} and \mathcal{M}_k given the vehicle pose $\{\mathbf{r}_{k-1}, \theta_{k-1}\}$ at timeslot $k-1$.

The objective function in (9) is formulated by taking a summation over P2L distances between the surface points $\boldsymbol{\mu}_i \in \mathcal{M}_k$ and their neighborhoods $\boldsymbol{\mu}_j \in \mathcal{M}_{k-1}$, where Huber loss $\mathcal{L}_\delta(\cdot)$ is used to reduce the objective's sensitivity against outliers. The relative vehicle pose from timeslot $k-1$ to timeslot k can be computed by

$$\mathbf{T}_{k-1}^{-1} \mathbf{T}_k = \begin{bmatrix} \Delta \mathbf{R}_k & \Delta \mathbf{r}_k \\ \mathbf{0}^T & 1 \end{bmatrix}, \quad \Delta \mathbf{R}_k = \begin{bmatrix} \cos \Delta \theta_k & -\sin \Delta \theta_k \\ \sin \Delta \theta_k & \cos \Delta \theta_k \end{bmatrix} \quad (10)$$

where $\Delta \mathbf{r}_k = \mathbf{r}_k - \mathbf{r}_{k-1}$ is the relative position vector and $\Delta \mathbf{R}_k \in \mathbb{SO}(2)$ is the relative rotation matrix with relative yaw $\Delta \theta_k = \theta_k - \theta_{k-1}$.

For each surface point $\boldsymbol{\mu}_i \in \mathcal{M}_k$, its neighborhoods $\boldsymbol{\mu}_j \in \mathcal{M}_{k-1}$ must satisfies two conditions: firstly, the angle between surface normals \mathbf{n}_i and \mathbf{n}_j is within a tolerance Θ ; secondly, $\boldsymbol{\mu}_j$ should be the closest point to $\boldsymbol{\mu}_i$ within the radius d_D centering at $\boldsymbol{\mu}_i$.

The proposed Fast-CFEAR approach applies a non-linear optimization library Ceres Solver [22] to numerically solve the unconstrained minimization problem (9) using BFGS line search method.

B. EKF Based Beam Tracking

The EKF based beam tracking algorithm is composed of state evolution model and measurement model, which correspond to channel prediction and update stages, respectively.

1) *State Evolution Model*: The channel state vector can be defined as

$$\mathbf{s}_k = [|\alpha_k|, \arg \alpha_k, \phi_k, \psi_k]^T \quad (12)$$

where $|\alpha_k|$ and $\arg \alpha_k$ denote the magnitude and argument of the path gain. The channel state evolution model can be represented by

$$\mathbf{s}_k = \mathbf{f}(\mathbf{s}_{k-1}) + \mathbf{u}_k + \mathbf{w}_k \quad (13)$$

$$\mathbf{w}_k = \left[\frac{\gamma}{2} \hat{\rho}_k |\alpha_{k-1}| \left(\frac{\hat{\mathbf{r}}_k^T \boldsymbol{\epsilon}_k^r}{\|\hat{\mathbf{r}}_k\|^2} - \frac{\hat{\mathbf{r}}_{k-1}^T \boldsymbol{\epsilon}_{k-1}^r}{\|\hat{\mathbf{r}}_{k-1}\|^2} \right), \frac{2\pi}{\lambda} \left(\frac{\hat{\mathbf{r}}_{k-1}^T \boldsymbol{\epsilon}_{k-1}^r}{\|\hat{\mathbf{r}}_{k-1}\|^2} - \frac{\hat{\mathbf{r}}_k^T \boldsymbol{\epsilon}_k^r}{\|\hat{\mathbf{r}}_k\|^2} \right), \frac{\hat{\mathbf{r}}_{k-1}^T \boldsymbol{\epsilon}_{k-1}^r}{\|\hat{\mathbf{r}}_{k-1}\|^2} - \frac{\hat{\mathbf{r}}_k^T \boldsymbol{\epsilon}_k^r}{\|\hat{\mathbf{r}}_k\|^2}, \frac{\hat{\mathbf{r}}_{k-1}^T \boldsymbol{\epsilon}_{k-1}^r}{\|\hat{\mathbf{r}}_{k-1}\|^2} - \frac{\hat{\mathbf{r}}_k^T \boldsymbol{\epsilon}_k^r}{\|\hat{\mathbf{r}}_k\|^2} + \Delta \epsilon_\theta^r \right]^T \quad (11)$$

where $\mathbf{f}(\cdot)$ is the evolution function. \mathbf{u}_k and \mathbf{w}_k denote the control input and process noise, respectively.

Proposition 1 (Ground Truth Pose Based Channel State Evolution Model): Based on ground truth of the vehicle pose, the channel state evolution model can be established as

$$\mathbf{s}_k = \mathbf{F}_k \mathbf{s}_{k-1} + \mathbf{u}_k \quad (14)$$

where $\mathbf{F}_k = \text{diag}([\rho_k, 1, 1, 1])$ is an evolution matrix with $\rho_k = (\|\mathbf{r}_{k-1}\|/\|\mathbf{r}_k\|)^{\frac{\gamma}{2}}$ standing for correlation coefficient on the magnitude of path gain. The path loss exponent is given by γ . $\mathbf{u}_k = [0, \frac{2\pi}{\lambda} (\|\mathbf{r}_k\| - \|\mathbf{r}_{k-1}\|), u_k^{[3]}, u_k^{[3]} - \Delta\theta_k]^T$ is the deterministic control input, where $u_k^{[3]} = \arctan(\frac{x_k}{y_k}) - \arctan(\frac{x_{k-1}}{y_{k-1}})$.

Proof: See Appendix A. ■

Nevertheless, the ground truth of vehicle pose is not available in practice. Instead, the vehicle pose estimation is accessible as we elaborated in Section III-A. Taking pose estimation error into consideration, another channel state evolution model can be derived by the following proposition.

Proposition 2 (Pose Estimation Based Channel State Evolution Model): Based on the vehicle pose estimation, the channel state evolution model can be established as

$$\mathbf{s}_k \approx \hat{\mathbf{F}}_k \mathbf{s}_{k-1} + \hat{\mathbf{u}}_k + \mathbf{w}_k \quad (15)$$

where $\hat{\mathbf{F}}_k$ and $\hat{\mathbf{u}}_k$ are estimated evolution matrix and control input, which have the same expressions as \mathbf{F}_k and \mathbf{u}_k but are calculated with the pose estimation instead of ground truth. $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$ is the process noise given in (11) which is placed on the top of this page. The covariance \mathbf{Q}_k of the process noise is given by

$$\mathbf{Q}_k = \begin{bmatrix} \mathbf{A}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_k \end{bmatrix}, \quad \mathbf{B}_k = \beta_k \sigma_r^2 \begin{bmatrix} 1 & 1 \\ 1 & 1 + \frac{2\sigma_\theta^2}{\beta_k \sigma_r^2} \end{bmatrix} \quad (16)$$

$$\mathbf{A}_k = \sigma_r^2 \begin{bmatrix} \left(\frac{\gamma}{2} \hat{\rho}_k |\alpha_{k-1}| \right)^2 \beta_k & -\frac{\pi}{\lambda} \gamma \hat{\rho}_k |\alpha_{k-1}| \zeta_k \\ -\frac{\pi}{\lambda} \gamma \hat{\rho}_k |\alpha_{k-1}| \zeta_k & \left(\frac{2\pi}{\lambda} \right)^2 \end{bmatrix} \quad (17)$$

where $\beta_k = \frac{1}{\|\hat{\mathbf{r}}_{k-1}\|^2} + \frac{1}{\|\hat{\mathbf{r}}_k\|^2}$ and $\zeta_k = \frac{1}{\|\hat{\mathbf{r}}_{k-1}\|} + \frac{1}{\|\hat{\mathbf{r}}_k\|}$.

Proof: See Appendix B. ■

Remark 1: For mmWave channel, the argument of path gain $\arg \alpha_k$ cannot be accurately tracked even based on mm-level localization techniques.

Proof: The standard deviation on tracking $\arg \alpha_k$ is given by $(\mathbf{Q}_k^{[2,2]})^{\frac{1}{2}} = \frac{2\pi}{\lambda} \sigma_r$ in Proposition 2, which ranges approximately $600\sigma_r \sim 6000\sigma_r$ rads for mmWave channel. Even with localization techniques of mm-level accuracy (i.e., $\sigma_r \approx 10^{-3}\text{m}$), the standard deviation on tracking $\arg \alpha_k$ still ranges about $0.6 \sim 6.0$ rads (i.e., $34^\circ \sim 340^\circ$). Such a large standard deviation indicates that the estimation on $\arg \alpha_k$ is always inaccurate. ■

Algorithm 1 The Proposed SPEBT Scheme

- 1: **input:** $\mathbf{r}_0, \theta_0, \mathcal{M}_0$
 - 2: **output:** $\bar{\mathbf{s}}_{k|k} = [|\alpha_k|, \phi_k, \psi_k]^T$
 - 3: **initialization:** Set timeslot $k = 0$. Initialize channel state vector $\bar{\mathbf{s}}_1$ via mmWave initial access techniques. Set $\bar{\mathbf{s}}_{0|0} = \bar{\mathbf{s}}_1$ and $\mathbf{P}_{0|0} = \bar{\mathbf{Q}}_0$. Set thresholds $\tilde{\alpha}, \tilde{\phi}, \tilde{\psi} > 0$.
 - 4: **repeat** $k \leftarrow k + 1$
 - // **Fast-CFEAR Based Pose Estimation**
 - 5: The latest radar point cloud is available at timeslot k .
 - 6: Obtain a sparse representation \mathcal{M}_k by performing the first three stages of Fast-CFEAR approach.
 - 7: Estimate the vehicle pose $(\hat{\mathbf{r}}_k, \hat{\theta}_k)$ via point cloud registration.
 - // **EKF Based Beam Tracking**
 - 8: Beam tracking via configuring the analog precoder $\mathbf{f}_k = \mathbf{a}_t(\phi_{k-1})$ and the analog combiner $\mathbf{g}_k = \mathbf{a}_r(\psi_{k-1})$.
 - **Channel Prediction Stage:**
 - 9: A Priori Channel State Vector:
 $\bar{\mathbf{s}}_{k|k-1} = \bar{\mathbf{F}}_k \bar{\mathbf{s}}_{k-1|k-1} + \bar{\mathbf{u}}_k$
 - 10: A Priori Channel Covariance Matrix:
 $\mathbf{P}_{k|k-1} = \bar{\mathbf{F}}_k \mathbf{P}_{k-1|k-1} \bar{\mathbf{F}}_k^T + \bar{\mathbf{Q}}_k$
 - **Channel Update Stage:**
 - 11: Kalman Gain:
 $\mathbf{K}_k = \frac{\mathbf{P}_{k|k-1} \nabla \bar{h}(\bar{\mathbf{s}}_{k|k-1})}{\nabla^T \bar{h}(\bar{\mathbf{s}}_{k|k-1}) \mathbf{P}_{k|k-1} \nabla \bar{h}(\bar{\mathbf{s}}_{k|k-1}) + \sigma_{v_k}^2}$
 - 12: A Posteriori Channel State Vector:
 $\bar{\mathbf{s}}_{k|k} = \bar{\mathbf{s}}_{k|k-1} + \mathbf{K}_k (\bar{z}_k - \nabla^T \bar{h}(\bar{\mathbf{s}}_{k|k-1}) \bar{\mathbf{s}}_{k|k-1})$
 - 13: A Posteriori Channel Covariance Matrix:
 $\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \nabla^T \bar{h}(\bar{\mathbf{s}}_{k|k-1})) \mathbf{P}_{k|k-1}$
 - **Deviation Checking for Re-initialization:**
 - 14: **if** $||\alpha_k| - |\alpha_1|| > \tilde{\alpha}$ or $|\phi_k - \phi_1| > \tilde{\phi}$ or $|\psi_k - \psi_1| > \tilde{\psi}$
 - 15: Re-initialize $\bar{\mathbf{s}}_1$ and set $\bar{\mathbf{s}}_{k|k} = \bar{\mathbf{s}}_1$.
 - 16: **end**
 - 17: **until** The vehicle arrives at destination.
-

Leaving the tracking on $\arg \alpha_k$ out of account, the channel state vector can be re-defined as

$$\bar{\mathbf{s}}_k = [|\alpha_k|, \phi_k, \psi_k]^T \quad (18)$$

and the pose estimation based channel state evolution model can be modified accordingly

$$\bar{\mathbf{s}}_k \approx \bar{\mathbf{F}}_k \bar{\mathbf{s}}_{k-1} + \bar{\mathbf{u}}_k + \bar{\mathbf{w}}_k \quad (19)$$

where $\bar{\mathbf{F}}_k$, $\bar{\mathbf{u}}_k$ and $\bar{\mathbf{w}}_k$ are calculated by $\hat{\mathbf{F}}_k$, $\hat{\mathbf{u}}_k$ and \mathbf{w}_k excluding the 2nd row (and the 2nd column) relevant to $\arg \alpha_k$.

2) *Measurement Model:* At each timeslot k , a pilot symbol b_k undergoing the downlink channel \mathbf{H}_k is transmitted from BS to the vehicle. As given in (5), a complex-valued measurement model can be established as

$$z_k = h(\mathbf{s}_k) + v_k \quad (20)$$

TABLE I
SIMULATION PARAMETERS

Pose Estimation Subsystem	Parameter Value
Selected sensing range	5m ~ 100m
K -strongest filtering	$K = 3$
Pixel value threshold (integer: 0 ~ 255)	$\kappa_{\min} = 55$
Side length of downsampling grid cell	$d_D = 3.5\text{m}$
Resampling factor	$f_D = 1.0$
Angle tolerance	$\Theta = 30^\circ$
Huber loss \mathcal{L}_δ	$\delta = 0.1$
Beam Tracking Subsystem	Parameter Value
Carrier frequency and wavelength	$f = 50\text{GHz}, \lambda = 6\text{mm}$
Number of transmit/receive antennas	$N = M = 4$
Antenna spacing	$d = \lambda/2$
Path loss exponent	$\gamma = 2.2$
Magnitude of path gain at $d_0 = 1\text{m}$	$ \alpha_{\text{ref}} = 5 \times 10^{-4}$
Noise power	$\sigma_v^2 = -90\text{dBm}$
Standard deviation of estimation error	$\sigma_r = 1.0\text{m}, \sigma_\theta = 3^\circ$
Channel re-initialization thresholds	$\bar{\alpha} = 5 \times 10^{-7}$ $\bar{\phi} = \bar{\psi} = 7.5^\circ$
Coordinate of vehicle initial position	(0m, 0m)
Coordinate of BS position	(-30m, -125m)

where $h(\mathbf{s}_k) = \alpha_k \mathbf{g}_k^H \mathbf{a}_r(\psi_k) \mathbf{a}_t^H(\phi_k) \mathbf{f}_k b_k$ is the complex-valued measurement function. Nonetheless, a real-valued measurement model is still in need for the convenience of deploying EKF algorithm that generally works in real field. Since the phase information of measurement z_k is inessential as the tracking on $\arg \alpha_k$ is not involved, a real-valued measurement model can be obtained by taking the magnitude of z_k as follows

$$\bar{z}_k = \bar{h}(\bar{\mathbf{s}}_k) + \tilde{v}_k \approx \bar{h}(\bar{\mathbf{s}}_k) + \bar{v}_k \quad (21)$$

where the real-valued measurement and measurement function are given by $\bar{z}_k = |z_k|^2 - \sigma_v^2$ and $\bar{h}(\bar{\mathbf{s}}_k) \triangleq |h(\mathbf{s}_k)|^2 = |b_k|^2 |\alpha_k|^2 |\mathbf{g}_k^H \mathbf{a}_r(\psi_k)|^2 |\mathbf{f}_k^H \mathbf{a}_t(\phi_k)|^2$, respectively. The measurement noise is given by $\tilde{v}_k = 2\Re[h(\mathbf{s}_k)v_k^*] + |v_k|^2 - \sigma_v^2$ with zero mean and variance $\sigma_{\tilde{v}_k}^2 = (\sigma_v^2 + 2\bar{h}(\bar{\mathbf{s}}_k))\sigma_v^2$.

Last but not the least, it can be observed that \tilde{v}_k is no longer a Gaussian random variable. Instead, \tilde{v}_k is a summation of a Gaussian random variable and a chi-squared random variable with 2 degrees of freedom. For the sake of guaranteeing the assumption of Gaussian measurement noise in EKF algorithm, a Gaussian noise $\bar{v}_k \sim \mathcal{N}(0, \sigma_{\bar{v}_k}^2)$ with the same variance as \tilde{v}_k is adopted to substitute the non-Gaussian \tilde{v}_k in the final measurement model (21).

C. The Proposed SPEBT Scheme

The SPEBT scheme is summarized in Algorithm 1. At the start of the scheme, the initial channel state vector $\bar{\mathbf{s}}_1$ should be established via mmWave initial access techniques [23], which is out of the scope of this paper. The vehicle performs beam tracking in light of the pose estimation whenever the latest point cloud collected by the radar sensor is available. To guarantee the accuracy and timeliness of the tracked channel, the algorithm should re-initialize the channel state vector once the deviation between the tracked channel and the initial channel are greater than the given thresholds. At

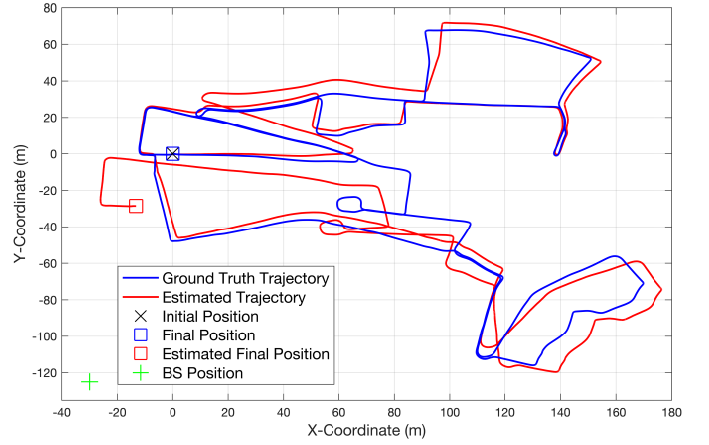


Fig. 3. The estimated vehicle moving trajectory compared to ground truth. Both trajectories are aligned by the initial position marked with \times , and the final positions are marked with \square . The position of BS is marked with $+$.

the channel update stage, $\nabla \bar{h}(\cdot)$ represents the gradient of the measurement function with respect to the channel state vector, whose expression is provided in Appendix C.

IV. SIMULATION RESULTS

In this section, experimental simulations are carried out to validate the performance of the proposed SPEBT scheme. For the simplicity of implementation, the simulation follows a divide and conquer strategy using two subsystems: a pose estimation subsystem and a pose estimation based beam tracking subsystem. On the one hand, as depicted in Fig. 2, a sequence named 2019-01-15-14-24-38 in *Oxford Radar RobotCar* dataset is selected to perform vehicle pose estimation. The sequence is consisted of 6388 consecutive frames collected by a Navtech CTS350-X mmWave FMCW radar in an urban environment. This subsystem is implemented by C++ programming language in Visual Studio 2022 for the sake of real-time processing. On the other hand, as illustrated by the 2D bird's eye view in Fig. 1, the beam tracking is performed in a 2D mmWave MIMO mobile wireless system, and this subsystem is implemented in MATLAB R2023a regarding the pose estimation information as input. The simulation parameters of both subsystems are summarized in Table I.

In Fig. 3, we demonstrate the estimated vehicle moving trajectory and ground truth trajectory, respectively. In addition, Fig. 4 compares the estimated pose with ground truth pose on the coordinate and yaw against the timeslot. It can be observed from both figures that the estimations are in high accuracy at the start of the trajectory, and they gradually deviate from the ground truth as the vehicle moving due to the accumulation of estimation errors. As a consequence, an obvious disparity is produced at vehicle final position. This observation brings an insight that it is crucial to take the pose estimation error into account in the beam tracking modeling as we elaborated in Proposition 2. That is also the reason why we cannot apply the simpler model offered by Proposition 1 directly. Table II provides the vehicle pose estimation evaluated by root-

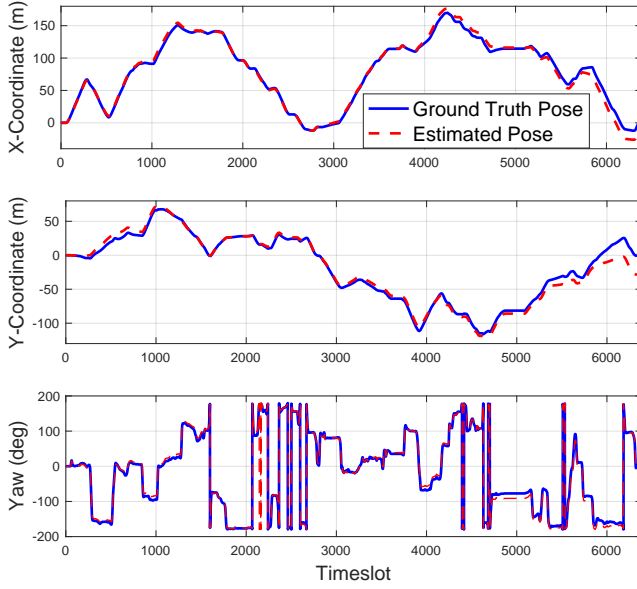


Fig. 4. The estimated pose compared to ground truth pose on coordinate (x_k, y_k) and yaw θ_k against timeslot k .

TABLE II
VEHICLE POSE ESTIMATION ERROR

Position Estimation Error		Yaw Estimation Error	
RMSE (m)	KITTI (%)	RMSE (deg)	KITTI (deg/m)
(2.772, 4.083)	6.562	3.550	0.0125

mean-squared error (RMSE) metric and KITTI benchmark [24] respectively, which are generated by an open-source toolbox [25] that is commonly adopted in pose evaluation. The statistics confirm that our pose estimation results are comparable to those of the SOTA approaches.

In Fig. 5, we compare a beam tracking realization generated by SPEBT scheme with ground truth channel on the magnitude of path gain, AoD and AoA against the timeslots, respectively. Similarly, the tracking on magnitude is almost in accordance with ground truth in the first half 3000 timeslots, but it fluctuates intensively around ground truth in the second half duration, which coincides with the gradual reduction in confidence of pose estimations. The beam tracking performance evaluated by RMSE is provided in Table III, which is the average result of 10^5 i.i.d. Monte Carlo simulations. From both the qualitative and quantitative results in Fig. 5 and Table III, it is interesting to find that the tracking on AoD performs much better than the tracking on AoA. It is because that AoD is only related to the vehicle position owing to the stationary BS antenna. However, AoA is additionally relevant to the vehicle yaw since the orientation of vehicular antenna keeps changing.

Last but not the least, it is satisfying to point out that the average times on channel re-initialization is less than 300 out of the total 6388 timeslots, which means that the proposed SPEBT scheme can greatly reduce the proportion of beam training overhead to less than 5% for mmWave vehicular mobile communication systems.

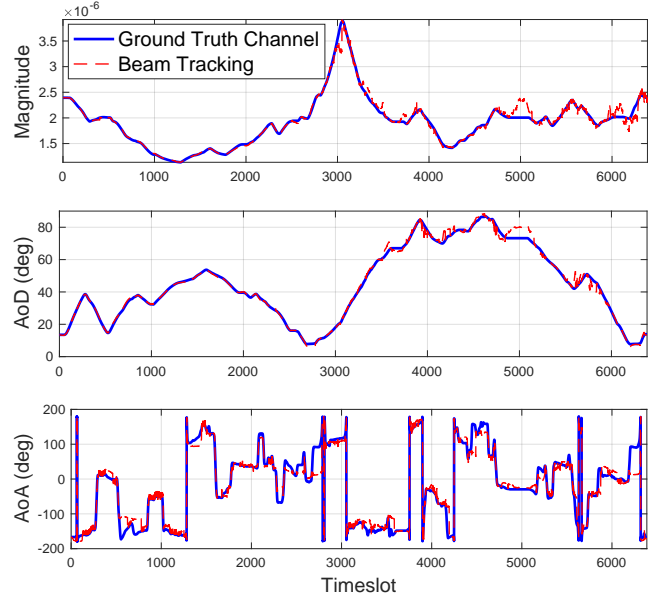


Fig. 5. A beam tracking realization generated by SPEBT scheme compared to ground truth channel on $|\alpha_k|$, ϕ_k and ψ_k against timeslot k .

TABLE III
BEAM TRACKING RMSE

Channel	Path Gain Mag. (%)	AoD (deg)	AoA (deg)
RMSE	1.580	0.220	4.616

V. CONCLUSION

We studied the mmWave radar sensing-aided mmWave communications aiming at reducing beam training overhead in high mobility vehicular mobile communication systems. A SPEBT scheme is proposed to achieve this goal, and simulations on an open-source real-world sensing dataset are conducted to verify its feasibility and effectiveness. A potential future direction is to extend the proposed scheme on 3-dimensional (3D) sensing dataset so as to generate precise 3D pose estimation results and accurate 3D beam tracking output.

APPENDIX A PROOF OF PROPOSITION 1

Based on simplified path loss model [26], the complex path gain of LoS channel is formulated as

$$\alpha = \alpha_{\text{ref}} (d_0 / \|\mathbf{r}\|)^{\frac{\gamma}{2}} e^{j \frac{2\pi}{\lambda} (\|\mathbf{r}\| - d_0)} \quad (22)$$

where α_{ref} is the complex path gain at reference distance d_0 . Therefore, the magnitude and argument of path gain are $|\alpha| = |\alpha_{\text{ref}}| (d_0 / \|\mathbf{r}\|)^{\frac{\gamma}{2}}$ and $\arg \alpha = \arg \alpha_{\text{ref}} + \frac{2\pi}{\lambda} (\|\mathbf{r}\| - d_0)$, respectively, which lead to the evolution equations $|\alpha_k| = \rho_k |\alpha_{k-1}|$ and $\arg \alpha_k = \arg \alpha_{k-1} + \frac{2\pi}{\lambda} (\|\mathbf{r}_k\| - \|\mathbf{r}_{k-1}\|)$. The derivation on the evolution equation of AoD is straightforward based on the 2D geometry illustrated in Fig. 1. Besides, the relation between yaw, AoD and AoA is described by $\theta - \phi + \psi = \pi$, which leads to the evolution equation of AoA accordingly.

APPENDIX B

PROOF OF PROPOSITION 2

The pose estimations are generally modeled as

$$\hat{\mathbf{r}}_k = \mathbf{r}_k + \epsilon_k^r, \quad \hat{\theta}_k = \theta_k + \epsilon_k^\theta \quad (23)$$

where $\hat{\mathbf{r}}_k$ and $\hat{\theta}_k$ are estimations on the position vector \mathbf{r}_k and yaw θ_k , and $\epsilon_k^r \sim \mathcal{N}(\mathbf{0}, \sigma_r^2 \mathbf{I}_2)$ and $\epsilon_k^\theta \sim \mathcal{N}(0, \sigma_\theta^2)$ are the corresponding estimation errors. Since the distance $\|\mathbf{r}_k\|$ between the two transceivers and the absolute value of yaw $|\theta_k|$ are usually much larger than the estimation errors $\|\epsilon_k^r\|$ and $|\epsilon_k^\theta|$, respectively, the evolution equation of AoA can be approximated by the first-order Taylor series as follows

$$\psi_k = f^{[4]}(\mathbf{r}_{k-1}, \mathbf{r}_k, \theta_{k-1}, \theta_k, \psi_{k-1}) \quad (24)$$

$$\approx f^{[4]}(\hat{\mathbf{r}}_{k-1}, \hat{\mathbf{r}}_k, \hat{\theta}_{k-1}, \hat{\theta}_k, \psi_{k-1}) + w_k^{[4]} \quad (25)$$

where $f^{[4]}(\cdot)$ is the evolution equation of AoA given in Proposition 1. The process noise $w_k^{[4]} = -(\frac{\partial f^{[4]}}{\partial \mathbf{r}_{k-1}} | \hat{\mathbf{r}}_{k-1})^\top \epsilon_{k-1}^r - (\frac{\partial f^{[4]}}{\partial \mathbf{r}_k} | \hat{\mathbf{r}}_k)^\top \epsilon_k^r - (\frac{\partial f^{[4]}}{\partial \theta_{k-1}} | \hat{\theta}_{k-1})^\top \epsilon_{k-1}^\theta - (\frac{\partial f^{[4]}}{\partial \theta_k} | \hat{\theta}_k)^\top \epsilon_k^\theta$ is consisted of the first-order Taylor polynomials and its final expression is provided by the fourth entry of (11), where $\tilde{\mathbf{r}}_k = [\hat{y}_k, -\hat{x}_k]^\top$ and $\Delta \epsilon_k^\theta = \epsilon_k^\theta - \epsilon_{k-1}^\theta$. In addition, $w_k^{[4]}$ is a zero mean Gaussian random variable with covariance $\mathbf{Q}_k^{[4,4]}$ (i.e., $w_k^{[4]} \sim \mathcal{N}(0, \beta_k \sigma_r^2 + 2\sigma_\theta^2)$) since it is a linear combination of zero mean Gaussian random variables. Besides, the first-order Taylor approximations on the evolution equations of other channel states can be derived in the same way, which are omitted here for avoidance of duplication.

APPENDIX C

GRADIENT OF MEASUREMENT FUNCTION

The gradient $\nabla \bar{h}(\cdot)$ of the measurement function $\bar{h}(\cdot)$ with respect to the channel state vector $\bar{\mathbf{s}}_k$ can be written as

$$\nabla \bar{h}(\bar{\mathbf{s}}_k) = \begin{bmatrix} 2|b_k|^2 |\alpha_k| |\mathbf{g}_k^H \mathbf{a}_r(\psi_k)|^2 |\mathbf{f}_k^H \mathbf{a}_t(\phi_k)|^2 \\ \frac{2\pi d \sin \phi_k}{N\lambda} |b_k|^2 |\alpha_k| |\mathbf{g}_k^H \mathbf{a}_r(\psi_k)|^2 |\mathbf{f}_k^H \Phi_k \mathbf{f}_k| \\ \frac{2\pi d \sin \psi_k}{M\lambda} |b_k|^2 |\alpha_k| |\mathbf{f}_k^H \mathbf{a}_t(\phi_k)|^2 \mathbf{g}_k^H \Psi_k \mathbf{g}_k \end{bmatrix} \quad (26)$$

where $\Phi_k \in \mathbb{C}^{N \times N}$ has the following expression

$$\Phi_k = \begin{bmatrix} \eta_0 & \eta_1^* & \eta_2^* & \cdots & \eta_{N-2}^* & \eta_{N-1}^* \\ \eta_1 & \eta_0 & \eta_1^* & \cdots & \eta_{N-3}^* & \eta_{N-2}^* \\ \eta_2 & \eta_1 & \eta_0 & \cdots & \eta_{N-4}^* & \eta_{N-3}^* \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \eta_{N-2} & \eta_{N-3} & \eta_{N-4} & \cdots & \eta_0 & \eta_1^* \\ \eta_{N-1} & \eta_{N-2} & \eta_{N-3} & \cdots & \eta_1 & \eta_0 \end{bmatrix} \quad (27)$$

with each entry specified by $\eta_n = n e^{-j(n \frac{2\pi d}{\lambda} \cos \phi_k - \frac{\pi}{2})}$, $n = 0, 1, \dots, N-1$. $\Psi_k \in \mathbb{C}^{M \times M}$ is of the same form with η_n replaced by $\xi_m = m e^{-j(m \frac{2\pi d}{\lambda} \cos \psi_k - \frac{\pi}{2})}$, $m = 0, 1, \dots, M-1$.

REFERENCES

- [1] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Feb. 2014.
- [2] L. Wei, R. Q. Hu, Y. Qian, and G. Wu, "Key elements to enable millimeter wave communications for 5G wireless systems," *IEEE Wireless Commun.*, vol. 21, no. 6, pp. 136–143, Dec. 2014.
- [3] S. H. Cen and P. Newman, "Precise ego-motion estimation with millimeter-wave radar under diverse and challenging conditions," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2018, pp. 6045–6052.
- [4] Z. Hong, Y. Petillot, and S. Wang, "RadarSLAM: Radar based large-scale SLAM in all weathers," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2020, pp. 5164–5170.
- [5] Y. Z. Ng, B. Choi, R. Tan, and L. Heng, "Continuous-time radar-inertial odometry for automotive radars," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2021, pp. 323–330.
- [6] A. Alkhateeb, J. Mo, N. Gonzalez-Prelcic, and R. W. Heath, "MIMO precoding and combining solutions for millimeter-wave systems," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 122–131, Dec. 2014.
- [7] Y. Wang, Z. Wei, and Z. Feng, "Beam training and tracking in mmWave communication: A survey," May 2022. [Online]. Available: <https://arxiv.org/abs/2205.10169>
- [8] C. Zhang, D. Guo, and P. Fan, "Tracking angles of departure and arrival in a mobile millimeter wave channel," in *Proc. IEEE Int. Conf. Commun.*, May 2016, pp. 1–6.
- [9] V. Va, H. Vikalo, and R. W. Heath, "Beam tracking for mobile millimeter wave communication systems," in *Proc. IEEE Global Conf. Signal Inf. Process.*, Dec. 2016, pp. 743–747.
- [10] S. Jayaprakasam, X. Ma, J. W. Choi, and S. Kim, "Robust beam-tracking for mmWave mobile communications," *IEEE Commun. Lett.*, vol. 21, no. 12, pp. 2654–2657, Sep. 2017.
- [11] F. Liu, P. Zhao, and Z. Wang, "EKF-based beam tracking for mmWave MIMO systems," *IEEE Commun. Lett.*, vol. 23, no. 12, pp. 2390–2393, Sep. 2019.
- [12] S. Shaham, M. Kokshoorn, M. Ding, Z. Lin, and M. Shirvanimoghadam, "Extended Kalman filter beam tracking for millimeter wave vehicular communications," in *Proc. IEEE Int. Conf. Commun. Workshops*, Jun. 2020, pp. 1–6.
- [13] L. Chen, S. Zhou, and W. Wang, "MmWave beam tracking with spatial information based on extended Kalman filter," *IEEE Wireless Commun. Lett.*, vol. 12, no. 4, pp. 615–619, Jan. 2023.
- [14] X. Li, F. Liu, Z. Zhou, G. Zhu, S. Wang, K. Huang, and Y. Gong, "Integrated sensing, communication, and computation over-the-air: MIMO beamforming design," *IEEE Trans. Wireless Commun.*, pp. 1–1, Jan. 2023.
- [15] Z. Wang, X. Mu, and Y. Liu, "STARS enabled integrated sensing and communications," *IEEE Trans. Wireless Commun.*, pp. 1–1, Feb. 2023.
- [16] G. Zhu, Z. Lyu, X. Jiao, P. Liu, M. Chen, J. Xu, S. Cui, and P. Zhang, "Pushing AI to wireless network edge: An overview on integrated sensing, communication, and computation towards 6G," *Sci. China Inf. Sci.*, vol. 66, no. 3, p. 130301, Feb. 2023.
- [17] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.
- [18] 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, Jan. 2014.
- [19] D. Barnes, M. Gadd, P. Murcutt, P. Newman, and I. Posner, "The Oxford Radar RobotCar dataset: A radar extension to the Oxford RobotCar dataset," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2020, pp. 6433–6438.
- [20] D. Adolfsson, M. Magnusson, A. Alhashimi, A. J. Lilienthal, and H. Andreasson, "Lidar-level localization with radar? The CFAR approach to accurate, fast, and robust large-scale radar odometry in diverse environments," *IEEE Trans. Robot.*, vol. 39, no. 2, pp. 1476–1495, 2023.
- [21] Y. Ma, S. Soatto, and J. Košecká, *An Invitation to 3-D Vision: From Images to Geometric Models*. Springer, 2004, vol. 26.
- [22] S. Agarwal, K. Mierle, and The Ceres Solver Team, "Ceres solver," Mar. 2022. [Online]. Available: <https://github.com/ceres-solver/ceres-solver>
- [23] M. Giordani, M. Mezzavilla, and M. Zorzi, "Initial access in 5G mmWave cellular networks," *IEEE Commun. Mag.*, vol. 54, no. 11, pp. 40–47, Nov. 2016.
- [24] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 3354–3361.
- [25] Z. Zhang and D. Scaramuzza, "A tutorial on quantitative trajectory evaluation for visual-(inertial) odometry," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2018, pp. 7244–7251.
- [26] A. Goldsmith, *Wireless Communications*. Cambridge university press, 2005.