



Integrated Snapshot and Filtering-based Bistatic Radio SLAM in mmWave Networks

Downloaded from: <https://research.chalmers.se>, 2025-01-19 17:02 UTC

Citation for the original published paper (version of record):

Kaltiokallio, O., Talvitie, J., Rastorgueva-Foi, E. et al (2024). Integrated Snapshot and Filtering-based Bistatic Radio SLAM in mmWave Networks. IEEE Workshop on Signal Processing Advances in Wireless Communications, SPAWC: 301-305.
<http://dx.doi.org/10.1109/SPAWC60668.2024.10694114>

N.B. When citing this work, cite the original published paper.

© 2024 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, or reuse of any copyrighted component of this work in other works.

Integrated Snapshot and Filtering-based Bistatic Radio SLAM in mmWave Networks

Ossi Kaltiokallio[†], Jukka Talvitie[†], Elizaveta Rastorgueva-Foi[†], Henk Wymeersch^{*}, and Mikko Valkama[†]

[†]Unit of Electrical Engineering, Tampere University, Tampere, Finland

^{*}Department of Electrical Engineering, Chalmers University of Technology, Göteborg, Sweden

Abstract—The inherent geometric relations between the propagation environment and millimeter-wave (mmWave) signals can be leveraged for simultaneous localization and mapping (SLAM) in 5G and beyond networks. Conventional solutions either rely on nonlinear filtering techniques or only solve the SLAM problem for a single user equipment (UE) location. This paper presents an integrated snapshot and filtering-based bistatic radio SLAM approach which utilizes low complexity linear Bayesian filtering equations for processing the measurements sequentially over time. The proposed method is validated using 60 GHz experimental data, indicating that the proposed processing pipeline combines good accuracy with low computational overhead, outperforming state-of-the-art benchmark algorithms.

Index Terms—millimeter wave, 5G/6G, simultaneous localization and mapping, Bayesian filtering

I. INTRODUCTION

In addition to the improved data communication quality and other connectivity related features, the evolving 5G and forthcoming 6G mobile networks can enable precise localization and sensing functions [1]. To extend situational awareness beyond ordinary user equipment (UE) localization, one central ingredient is related to appropriate processing and treatment of the non-line-of-sight (NLoS) paths and the underlying environment incidence points (IPs) [2]. Jointly estimating the unknown location of the UE and those of the environment IPs, using uplink or downlink reference signals, leads to the notion of bistatic cellular simultaneous localization and mapping (SLAM) [3] – being also the main scope of this paper.

Developing new methods for 5G/6G cellular SLAM is of growing interest as it can enhance the UE positioning accuracy [2] while allowing to carry out positioning without line-of-sight (LoS) [4] – and potentially also with just a single base station (BS) entity [3]. In general, one established approach builds on *sequential SLAM processing or filtering* where the UE location and the map are jointly estimated via recursive processing of the measurements gathered over time. The other alternative is *snapshot SLAM* where the problem is solved independently at different measurement locations. One example filtering approach utilizes random finite set (RFS) theory [5], [6] and the attractiveness of the RFS approach lies in the fact that it enables a fully integrated Bayesian framework for SLAM. The snapshot SLAM problem on the other hand is typically solved using geometry-based methods [4], [7] or

This work has been supported by the Academy of Finland (grants #338224, #345654, #352754 and #359095), by Business Finland (6G-ISAC project), and by the Vinnova B5GPOS Project under Grant 2022-01640. Email: {ossi.kaltiokallio, jukka.talvitie, elizaveta.rastorgueva-foi, mikko.valkama}@tuni.fi, henkw@chalmers.se

maximum likelihood estimation [2], [8], and the approaches are fundamentally important as they serve as a baseline for what can be done with radio signals alone. One fundamental drawback of RFS approaches is that the nonlinear observation model is approximated using local linearization leading to approximations in the filtering recursion. The main drawback of snapshot SLAM methods is, since the problem is solved for a single UE location only, that the available information is not fully utilized if observations from multiple nearby UE locations exist. This paper addresses the aforementioned drawbacks.

In this paper, we present an integrated snapshot and filtering-based bistatic radio SLAM approach. While conventional methods utilize channel parameter estimates as input to SLAM algorithms, we propose to utilize the snapshot SLAM algorithm as an additional processing step. This can be seen as a transformation of the observations from the nonlinear angle/delay domain to linear Euclidean space. As a consequence, the resulting measurement models are linear which enables us to utilize low complexity linear Bayesian filtering equations for processing the measurements sequentially over time. The contributions of the paper can be summarized as follows:

- 1) We present a new linear bistatic radio SLAM approach which integrates snapshot and filtering-based techniques;
- 2) A multi-hypothesis filtering framework based on the nearest neighbour (NN) algorithm and the probability hypothesis density (PHD) filter is devised to sequentially solve the bistatic radio SLAM problem;
- 3) Using experimental data with off-the-shelf 60 GHz millimeter-wave (mmWave) MIMO radios, we demonstrate that the developed algorithm provides notable performance gains over the state-of-the-art benchmark algorithms.

II. BACKGROUND

In the following, we consider a bistatic scenario in which the transmitter and receiver are at the BS and UE side, respectively. The focus is on describing the methods in the 2D/azimuth domain (position $\mathbf{p} = [x, y]^T$ and angle α) to simplify the notation but extension to 3D is possible (conceptually similar to, e.g., [4]). In the following, time index k is omitted whenever possible. The BS state is $\mathbf{x}_{\text{BS}} = [\mathbf{p}_{\text{BS}}^T, \alpha_{\text{BS}}]^T \in \mathbb{R}^3$ and the UE state is $\mathbf{x}_{\text{UE}} = [\mathbf{p}_{\text{UE}}^T, \alpha_{\text{UE}}, b_{\text{UE}}]^T \in \mathbb{R}^4$ in which b_{UE} denotes the error between the UE and BS clocks. The map of the environment with M landmarks is modeled as a RFS with random cardinality and random states, given by, $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_M\}$ in which state of the i th landmark $\mathbf{p}_i \in \mathbb{R}^2$ represents the IP of the i th single-bounce NLoS path. Furthermore, let $\mathbf{z}_i = [\text{ToA}_i, \text{AoD}_i, \text{AoA}_i]^T \in \mathbb{R}^3$

denote the time-of-arrival (ToA), angle-of-departure (AoD), and angle-of-arrival (AoA) estimates of the i th propagation path and $\mathcal{Z} = \{\mathbf{z}_1, \dots, \mathbf{z}_N\}$ the set of channel parameter estimates.

A. Robust Snapshot SLAM

The work in [7] introduced an algorithm that can solve the SLAM problem using \mathcal{Z} which contains outliers, that is, channel parameter estimates that originate from multi-bounce propagation paths and false detections. Let N_{\min} denote the minimum number of channel parameter estimates required to solve the snapshot SLAM problem ($N_{\min} = 4$ in our case), let $\mathcal{Z}_{\min} \in \mathbb{R}^{1 \times N_{\min}}$ denote a subset of \mathcal{Z} and let $\mathcal{Z}_{\text{all}} \in \mathbb{R}^{L \times N_{\min}}$ denote all L such possible combinations of \mathcal{Z} . In essence, the robust snapshot SLAM method first uses $\mathcal{Z}_{\min} \in \mathcal{Z}_{\text{all}}$ to compute an initial solution. Then, according to the principles of the random sample consensus (RANSAC) algorithm, the channel parameter estimates are partitioned into a set of inliers and a set of outliers based on the initial solution. Lastly, the problem is re-solved using the inlier set and a cost for the solution is computed, denoted by $C(\alpha_{\text{UE}}, \mathcal{Z}_{\min})$. The robust snapshot SLAM algorithm performs the above three steps for all possible combinations \mathcal{Z}_{all} and UE orientations $\mathcal{A} = (-\pi, \pi]$. Then, the UE state estimate $\hat{\mathbf{x}}_{\text{UE}}$ and estimated inlier set $\hat{\mathcal{Z}}$ are obtained by solving the following optimization problem

$$\underset{\alpha_{\text{UE}} \in \mathcal{A}, \mathcal{Z}_{\min} \in \mathcal{Z}_{\text{all}}}{\text{minimize}} C(\alpha_{\text{UE}}, \mathcal{Z}_{\min}), \quad (1)$$

in which the UE position and clock bias estimates are given in closed-form when solving (1). The problem is solved using a grid search and we consider the more challenging scenario by assuming that the LoS does not exist meaning that four or more NLoS paths are required to solve (1). After solving $\hat{\mathbf{x}}_{\text{UE}}$ and for every element of the inlier set $\hat{\mathcal{Z}}$, landmark $\hat{\mathbf{p}}_i$ can be estimated independently by solving a nonlinear optimization problem as presented in [7]. From now on, $\hat{\mathbf{x}}_{\text{UE}}$ and $\hat{\mathbf{p}}_i$ are referred to as measurements since they are used as measurement inputs to a tracking filter in Section III.

B. Measurement Covariance

To tackle the online SLAM problem appropriately in a Bayesian way, the measurement covariance matrix is required. For notational convenience, let us assume $M = N$ and that each channel parameter estimate is associated with a single-bounce propagation path so that we can express the joint UE and map state as $\mathbf{x} = [\mathbf{x}_{\text{UE}}^T, \mathbf{p}_1^T, \dots, \mathbf{p}_N^T]^T$. The Cramér-Rao lower bound (CRLB) provides a lower bound on the variance of an unbiased estimator, defined by the inverse of the Fisher information matrix (FIM). Therefore, the unbiased estimator $\hat{\mathbf{x}}$ satisfies $\text{Cov}(\hat{\mathbf{x}}) = \text{E}[(\hat{\mathbf{x}} - \mathbf{x})(\hat{\mathbf{x}} - \mathbf{x})^T] \geq \mathbf{J}(\mathbf{x})^{-1}$ [9], where $\mathbf{J}(\mathbf{x}) \in \mathbb{R}^{4+2N}$ denotes the FIM. Let us assume the likelihood function of the single-bounce NLoS paths is Gaussian, $\mathcal{N}(\mathbf{z}_i | \mathbf{g}(\mathbf{x}_{\text{UE}}, \mathbf{p}_i), \mathbf{W}_i)$, with covariance \mathbf{W}_i and mean given by

$$\mathbf{g}(\mathbf{x}_{\text{UE}}, \mathbf{p}_i) = \begin{bmatrix} \|\mathbf{p}_{\text{BS}} - \mathbf{p}_i\|/c + \|\mathbf{p}_i - \mathbf{p}_{\text{UE}}\|/c + b_{\text{UE}} \\ \text{atan2}(y_i - y_{\text{BS}}, x_i - x_{\text{BS}}) - \alpha_{\text{BS}} \\ \text{atan2}(y_i - y_{\text{UE}}, x_i - x_{\text{UE}}) - \alpha_{\text{UE}} \end{bmatrix}, \quad (2)$$

in which $\|\cdot\|$ denotes the Euclidean norm and c the speed of light. The FIM of the channel parameters, $\mathbf{g}(\mathbf{x}) = [\mathbf{g}(\mathbf{x}_{\text{UE}}, \mathbf{p}_1)^T, \dots, \mathbf{g}(\mathbf{x}_{\text{UE}}, \mathbf{p}_N)^T]^T$, is $\mathbf{J}(\mathbf{z}) = \text{blkdiag}(\mathbf{W}_1^{-1}, \dots, \mathbf{W}_N^{-1})$ [3], in which $\text{blkdiag}(\cdot)$ denotes a block diagonal matrix. Now, the FIM of state \mathbf{x} is

$$\mathbf{J}(\mathbf{x}) = [\nabla_{\mathbf{x}} \mathbf{g}(\mathbf{x})]^T \mathbf{J}(\mathbf{z}) [\nabla_{\mathbf{x}} \mathbf{g}(\mathbf{x})], \quad (3)$$

in which $\nabla_{\mathbf{x}} \mathbf{g}(\mathbf{x}) \in \mathbb{R}^{3N \times (4+2N)}$ denotes the Jacobian of $\mathbf{g}(\mathbf{x})$ evaluated with respect to \mathbf{x} with element in row i and column j given by $[\nabla_{\mathbf{x}} \mathbf{g}(\mathbf{x})]_{i,j} = \partial[\mathbf{g}(\mathbf{x})]_i / \partial[\mathbf{x}]_j$. Now, covariance of the UE and landmark measurements are given by

$$\mathbf{R}^{\text{UE}} = [\mathbf{J}(\mathbf{x})^{-1}]_{1:4,1:4}, \quad (4)$$

$$\mathbf{R}_i^{\text{MAP}} = [\mathbf{J}(\mathbf{x})^{-1}]_{2i+3:2i+4,2i+3:2i+4}. \quad (5)$$

III. PROPOSED MULTI-HYPOTHESIS LINEAR SLAM FILTER

Ideally, the cost in (1) should only have one minimum that locates near the ground truth UE orientation, but due to noise and outliers, the cost can have multiple local minima, as defined after (6), and the global minimum is not always the one that is closest to the ground truth as illustrated in Fig. 1. To overcome this challenge, we present a detector that finds every local minima and since the detector gives rise to multiple measurements, we formulate the UE tracking problem as single object tracking (SOT) in clutter [10] which is a special case of multiple object tracking (MOT) [11]. In addition, a map that is conditioned on the UE state is computed and the mapping problem is formulated as a MOT problem.

A. Multi-Hypothesis Detector

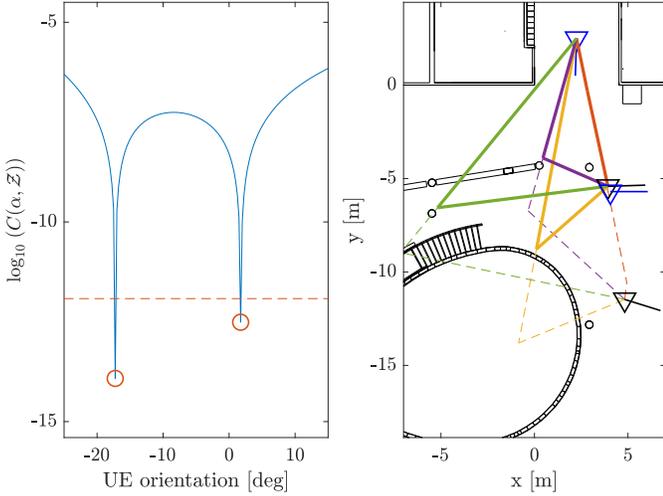
The measurements and covariances correspond to a specific time instant k , denoted using subscript k from now on. The measurements at time k are modeled using a RFS with random cardinality and random elements, given by

$$\mathcal{Y}_k^{\text{UE}} = \{\hat{\mathbf{x}}_k^{\text{UE}} \mid C(\alpha_{\text{UE}}, \mathcal{Z}_{\min}) \leq C(\alpha_{\text{UE}}^*, \mathcal{Z}_{\min}) \wedge C(\alpha_{\text{UE}}, \mathcal{Z}_{\min}) \leq \Gamma\} \quad (6)$$

where $\exists \delta > 0$ s.t. $C(\alpha_{\text{UE}}, \mathcal{Z}_{\min}) \leq C(\alpha_{\text{UE}}^*, \mathcal{Z}_{\min}) \forall \alpha_{\text{UE}}^* \in (\alpha_{\text{UE}} - \delta, \alpha_{\text{UE}} + \delta)$ is the condition defining the local minima in which δ is the grid spacing of \mathcal{A} . The second condition $C(\alpha_{\text{UE}}, \mathcal{Z}_{\min}) \leq \Gamma$ is used to only account for local minima for which the cost is below threshold Γ . The threshold is defined as $\Gamma = (1 + \beta) \times C_{\min}$ in which $C_{\min} = \min C(\alpha_{\text{UE}}, \mathcal{Z}_{\min})$ is the global minimum and β is a tuning parameter between two extremes. If $\beta = 0$ only the global minimum is accounted for and if $\beta = \infty$ all local minima are taken into account. The total number of measurements at time k is $m_k = |\mathcal{Y}_k^{\text{UE}}|$, in which $|\mathcal{Y}_k^{\text{UE}}|$ denotes cardinality of the set. For every measurement $i \in (1, \dots, m_k)$, a set of landmark measurements $\mathcal{Y}_{k,i}^{\text{MAP}} = \{\hat{\mathbf{p}}_{k,1}, \dots, \hat{\mathbf{p}}_{k,|\hat{\mathcal{Z}}|}\}$ and the covariances defined in Section II-B are associated.

B. Nearest Neighbour Filter for UE Tracking

The filtering recursion of SOT in clutter follows the conventional prediction and update steps of Bayesian filtering but since the number of hypotheses grows exponentially, approximations



(a) Cost function vs. UE orientation (b) Two candidate solutions

Fig. 1. In (a), cost function for a given set \mathcal{Z} with two local minima and in (b), the corresponding snapshot SLAM solutions are illustrated. In (b), solution corresponding to the global minima is illustrated using dashed lines and the other solution shown using solid lines. The ground truth UE position is illustrated using (∇) . The detection threshold is shown in (a) using a dashed red line and the detection threshold value in the example is $(1 + \beta) = 100$.

are required to obtain a computationally feasible algorithm [10]. In the following, we devise a NN filter which only keeps one hypothesis after each update step and the rest are pruned.

1) *Models*: Posterior, transition density and likelihood function of measurement $\mathbf{y}_{k,j}^{\text{UE}} \triangleq \{\mathcal{Y}_k^{\text{UE}}\}_j$ are all represented using Gaussians:

$$p(\mathbf{x}_k^{\text{UE}} | \mathcal{Y}_{1:k}^{\text{UE}}) \approx \mathcal{N}(\mathbf{x}_k^{\text{UE}}; \mathbf{m}_k^{\text{UE}}, \mathbf{P}_k), \quad (7)$$

$$p(\mathbf{x}_k^{\text{UE}} | \mathbf{x}_{k-1}^{\text{UE}}) = \mathcal{N}(\mathbf{x}_k^{\text{UE}}; \mathbf{F}\mathbf{x}_{k-1}^{\text{UE}}, \mathbf{Q}), \quad (8)$$

$$p(\mathbf{y}_{k,j}^{\text{UE}} | \mathbf{x}_k^{\text{UE}}) = \mathcal{N}(\mathbf{y}_{k,j}^{\text{UE}}; \mathbf{H}\mathbf{x}_k^{\text{UE}}, \mathbf{R}_k^{\text{UE}}). \quad (9)$$

The posterior in (7) is parameterized by the mean \mathbf{m}_k^{UE} and covariance \mathbf{P}_k . The transition density in (8) describes the state evolution in which \mathbf{F} is the state transition matrix and \mathbf{Q} covariance of the process noise, both assumed time-invariant. The connection between the UE measurements and the UE state is represented using the likelihood function in (9) in which \mathbf{H} is a linear measurement matrix and \mathbf{R}_k^{UE} is the measurement noise covariance given in (4) at time k .

2) *Filter Recursion*: Assume that at time $k-1$ the posterior is Gaussian and given by $\mathcal{N}(\mathbf{x}_{k-1}^{\text{UE}}; \mathbf{m}_{k-1}^{\text{UE}}, \mathbf{P}_{k-1})$. In SOT, the predicted density is given by the conventional Bayesian filtering prediction step [10, Eq. 4.11]

$$p(\mathbf{x}_k^{\text{UE}} | \mathcal{Y}_{1:k-1}^{\text{UE}}) = \mathcal{N}(\mathbf{x}_{k-1}^{\text{UE}}; \mathbf{m}_{k|k-1}^{\text{UE}}, \mathbf{P}_{k|k-1}), \quad (10)$$

where $\mathbf{m}_{k|k-1}^{\text{UE}} = \mathbf{F}\mathbf{m}_{k-1}^{\text{UE}}$ and $\mathbf{P}_{k|k-1} = \mathbf{F}\mathbf{P}_{k-1}\mathbf{F}^T + \mathbf{Q}$. The posterior at time k is:

$$p(\mathbf{x}_k^{\text{UE}} | \mathcal{Y}_{1:k}^{\text{UE}}) = \sum_{\theta_k=0}^{m_k} w_k^{\theta_k} \mathcal{N}(\mathbf{x}_k^{\text{UE}}; \mathbf{m}_k^{\theta_k}, \mathbf{P}_k^{\theta_k}), \quad (11)$$

where θ_k is a data association (DA) variable, $w_k^{\theta_k}$ represent the DA probabilities and $\sum_{\theta_k=0}^{m_k} w_k^{\theta_k} = 1$. Let P^{D} denote the

probability of detection and $\lambda_c(\mathbf{y}_{k,\theta_k}^{\text{UE}})$ the clutter intensity. Parameters of the posterior for misdetection ($\theta_k = 0$) are:

$$w_k^{\theta_k=0} = 1 - P^{\text{D}}, \quad \mathbf{m}_k^{\theta_k=0} = \mathbf{m}_{k|k-1}^{\text{UE}} \quad \text{and} \quad \mathbf{P}_k^{\theta_k=0} = \mathbf{P}_{k|k-1}, \quad (12)$$

and for detection ($\theta_k \in \{1, \dots, m_k\}$) they are:

$$w_k^{\theta_k} = \frac{P^{\text{D}}}{\lambda_c(\mathbf{y}_{k,\theta_k}^{\text{UE}})} \mathcal{N}(\mathbf{y}_{k,\theta_k}^{\text{UE}}; \mathbf{H}\mathbf{m}_{k|k-1}^{\text{UE}}, \mathbf{S}_k), \quad (13)$$

$$\mathbf{m}_k^{\theta_k} = \mathbf{m}_{k|k-1}^{\text{UE}} + \mathbf{K}_k(\mathbf{y}_{k,\theta_k}^{\text{UE}} - \mathbf{H}\mathbf{m}_{k|k-1}^{\text{UE}}), \quad (14)$$

$$\mathbf{P}_k^{\theta_k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{S}_k\mathbf{K}_k^T, \quad (15)$$

where $\mathbf{y}_{k,\theta_k}^{\text{UE}}$ denotes element θ_k of $\mathcal{Y}_k^{\text{UE}}$. The predicted innovation covariance \mathbf{S}_k and Kalman gain \mathbf{K}_k can be computed using the conventional Kalman filter update step [12, Eq. 4.21]. The NN filter only keeps the most probable hypothesis and prunes the rest, such that the posterior at time step k is approximated as [10, Eq. 4.12]

$$p(\mathbf{x}_k^{\text{UE}} | \mathcal{Y}_{1:k}^{\text{UE}}) \approx \mathcal{N}(\mathbf{x}_k^{\text{UE}}; \mathbf{m}_k^{\theta_k^*}, \mathbf{P}_k^{\theta_k^*}), \quad \text{where} \quad (16)$$

$$\theta_k^* = \arg \max_{\theta_k} w_k^{\theta_k}. \quad (17)$$

At the first time step $k=1$, the NN filter is initialized using the measurement with lowest cost.

C. PHD Filter for Mapping

The PHD filter can account for the RFS nature of the landmarks and is therefore utilized for mapping [6]. However, a significant difference with respect to prior works is that the observation model of the proposed PHD filter is linear, whereas in prior works the observation model is nonlinear [5], [6], [13]. The PHD filter recursively estimates the first-order statistical moment of the posterior density $p(\mathcal{P}_k | \mathcal{Y}_{1:k}^{\text{MAP}}, \mathbf{x}_k^{\text{UE}})$ and in this paper, the filter utilizes the measurement associated with the most probable hypothesis θ_k^* given in (17) such that the measurement for time step k is $\mathcal{Y}_k^{\text{MAP}} \triangleq \mathcal{Y}_{k,\theta_k^*}^{\text{MAP}}$.

1) *Models*: The PHD is represented using a Gaussian mixture, $v_k(\mathbf{p}) = \sum_{j=1}^{J_k} \eta_k^j \mathcal{N}(\mathbf{p}; \boldsymbol{\mu}_k^j, \boldsymbol{\Sigma}_k^j)$, where η_k^j , $\boldsymbol{\mu}_k^j$ and $\boldsymbol{\Sigma}_k^j$ denote the weight, mean and covariance of mixture component j , and J_k is the number of components. The transition density and likelihood function are represented using Gaussians:

$$p(\mathbf{p}_k | \mathbf{p}_{k-1}) = \mathcal{N}(\mathbf{p}_k; \mathbf{F}\mathbf{p}_{k-1}, \mathbf{Q}), \quad (18)$$

$$p(\mathbf{y}_{k,j}^{\text{MAP}} | \mathbf{p}_k) = \mathcal{N}(\mathbf{y}_{k,j}^{\text{MAP}}; \mathbf{H}\mathbf{p}_k, \mathbf{R}_k^{\text{MAP}}), \quad (19)$$

in which \mathbf{F} is a linear transition matrix, \mathbf{Q} the process noise covariance, \mathbf{H} the linear measurement matrix and $\mathbf{R}_k^{\text{MAP}}$ the measurement covariance matrix. The measurement covariance is $\mathbf{R}_k^{\text{MAP}} = \frac{1}{|\mathcal{Z}|} \sum_{i=1}^{|\mathcal{Z}|} \mathbf{R}_{k,i}^{\text{MAP}}$, in which $\mathbf{R}_{k,i}^{\text{MAP}}$ is given in (5). The covariance is averaged over all measurements since the DA between measurements and existing landmarks is unknown.

2) *Filter Recursion*: If the PHD at time $k-1$ is a Gaussian mixture and since the models are linear Gaussian, it follows that the predicted PHD is a Gaussian mixture [11, Eq. 24]

$$v_{k|k-1}(\mathbf{p}) = v_{k|k-1}^{\text{S}}(\mathbf{p}) + v_k^{\text{B}}(\mathbf{p}) \quad (20)$$

and also the updated PHD is a Gaussian mixture [11, Eq. 32]

$$v_k(\mathbf{p}) = (1 - P^{\text{D}})v_{k|k-1}(\mathbf{p}) + \sum_{\mathbf{y}_k^{\text{MAP}} \in \mathcal{Y}_k^{\text{MAP}}} v_k^{\text{D}}(\mathbf{p}; \mathbf{y}_k^{\text{MAP}}). \quad (21)$$

The recursion to compute the means and covariances in (20) and (21) can be efficiently computed using the Kalman filter prediction and update steps [11, Table I]. Utilizing the Gaussian component reduction algorithm [11, Table II], the PHD filter complexity can be further reduced. Similar to [14], we utilize an informative prior for the birth intensity $v_k^B(\mathbf{p})$ as follows. In the PHD update step, we utilize gating to eliminate unlikely observations and if an observation is not used to update any landmark, we use the observation to create a new landmark. For example, if element i of $\mathcal{Y}_k^{\text{MAP}}$ is not associated to any landmark, the mean of the birth intensity is $\boldsymbol{\mu}_{k,i}^B = [\mathcal{Y}_k^{\text{MAP}}]_i$ and covariance $\boldsymbol{\Sigma}_{k,i}^B = \mathbf{R}_{k,i}^{\text{MAP}}$ is computed from the FIM using (5).

IV. EXPERIMENTAL RESULTS

A. Experimental Setup

The performance of the proposed and benchmark SLAM algorithms is evaluated using real-world 60 GHz measurement data, obtained indoors at Tampere University campus, with floorplan as illustrated in Fig. 3. Altogether 45 UE locations were measured. Beamformed measurements were obtained using 400 MHz transmission bandwidth utilizing 5G NR-specified downlink positioning reference signals. Details of the experiment, hardware and channel estimator are provided in [8].

In the following, the proposed algorithm is quantitatively evaluated in two phases. In the first phase, the evaluation solely focuses on the multi-hypothesis snapshot SLAM detector whereas in the second phase, the entire processing pipeline is considered. The proposed method is benchmarked against a robust snapshot SLAM algorithm [7] referred to as BM1 and a conventional filtering-based method that uses the channel parameter estimates as input to a Rao-Blackwellized PHD-SLAM filter [5] referred to as BM2. Since the ground truth landmark locations are unknown, which is a common problem in SLAM when using real-world data, the quantitative evaluation focuses on the UE state. Accuracy is assessed using the position, heading and clock bias root mean squared errors (RMSEs).

The filters are evaluated using two different dynamic models. The first is a random walk (RW) model for which $\mathbf{F} = \mathbf{I}_4$ and $\mathbf{Q} = \text{diag}([10 \text{ m}^2, 10 \text{ m}^2, 0.01 \text{ rad}^2, 33 \text{ ns}^2])$, where \mathbf{I}_4 denotes a 4×4 identity matrix and $\text{diag}(\cdot)$ a diagonal matrix. The second is a RW model with a deterministic control input (RW+C) for which $\mathbf{F} = \mathbf{I}_4$, $\mathbf{Q} = \text{diag}([0.1 \text{ m}^2, 0.1 \text{ m}^2, 0.01 \text{ rad}^2, 33 \text{ ns}^2])$ and the ground truth 2D translation is added to the predicted UE state. The PHD also utilizes a RW model with $\mathbf{F} = \mathbf{I}_2$ and $\mathbf{Q} = \text{diag}([0.1 \text{ m}^2, 0.1 \text{ m}^2])$. The measurement matrices are $\mathbf{H} = \mathbf{I}_4$ and $\mathbf{H} = \mathbf{I}_2$. The NN and PHD filters utilize the same values for $P^D = 0.9$ and $\lambda_c(\cdot) = 10^{-6}$. The covariance of the channel parameters is $\mathbf{W} = \text{diag}([1 \text{ ns}, \frac{3\pi}{180} \text{ rad}, \frac{3\pi}{180} \text{ rad}]^2)$. The snapshot SLAM algorithm and BM1 use the same parameters that were used in [7]. BM2 uses 1000 particles and the other parameters are tuned to maximize the filter performance.

B. Quantitative Results

1) *Multi-Hypothesis Snapshot SLAM Accuracy*: We evaluate properties of the multi-hypothesis snapshot SLAM detector by computing the position RMSE in three different ways. At

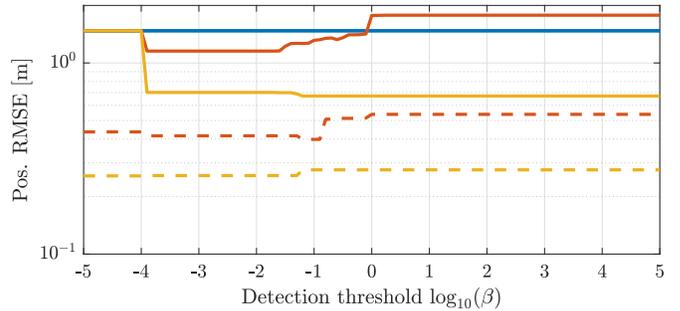


Fig. 2. Position RMSE as a function of detection threshold β . The RMSE is computed using *global minima* (—), *oracle* (—) and *weighted average* (—) methods. The dashed lines (RW (---), RW+C (---)) illustrate performance of the proposed filter with different detection thresholds.

UE location k , let $\hat{\mathbf{P}}_{\text{UE},k} \in \mathbb{R}^{2 \times m_k}$ denote the m_k position measurements of the detector and the corresponding costs as $\hat{C}_k \in \mathbb{R}^{1 \times m_k}$, such that we can define the squared Euclidean norm for the i th measurement as

$$e_{k,i} = (\mathbf{p}_{\text{UE},k} - [\hat{\mathbf{P}}_{\text{UE},k}]_{\cdot,i})^\top (\mathbf{p}_{\text{UE},k} - [\hat{\mathbf{P}}_{\text{UE},k}]_{\cdot,i}). \quad (22)$$

Now, the RMSE is computed as follows: (i) *Global minima* – RMSE is computed as $(\frac{1}{K} \sum_{k=1}^K e_{k,i})^{1/2}$ in which $i = \arg \min_{m_k} \hat{C}_k$ is index of the measurement with lowest cost; (ii) *Oracle* – RMSE is computed as $(\frac{1}{K} \sum_{k=1}^K e_{k,j})^{1/2}$ in which $j = \arg \min_{m_k} e_k$ is index of the measurement that is closest to the ground truth; (iii) *Weighted average* – RMSE is computed as $(\frac{1}{K} \sum_{k=1}^K \sum_{i=1}^{m_k} w_{k,i} e_{k,i})^{1/2}$ in which $w_{k,i} = \hat{C}_{k,i}^{-1} / \sum_{i=1}^{m_k} \hat{C}_{k,i}^{-1}$. The position RMSE using the three different methods as a function of detection threshold β is illustrated in Fig. 2 and it is to be noted that all methods yield comparative performance when $\beta < 10^{-4}$ since there is only one measurement. The global minima method is equivalent to solving (1) and β does not impact the RMSE since the global minima is used as the measurement. On the other hand, RMSE computed using the oracle method monotonically decreases as β increases, which implies that the optimal solution given by (1) is not always the one that is closest to the ground truth and a multi-hypothesis detector should be considered instead. Interestingly, the RMSE saturates to a nearly fixed value already at $\beta \approx 10^{-1}$ indicating that the accuracy cannot be improved indefinitely just by increasing β , which in general increases the number of measurements. The weighted average method reflects the benefit and disadvantage of the multi-hypotheses detector. As illustrated, the weighted average method outperforms the global minima method as long as the detection threshold is set properly $10^{-4} < \beta < 10^0$ indicating that it is beneficial to consider multiple snapshot SLAM solutions. On the other hand, the RMSE increases when $\beta \geq 10^0$ since the multi-hypothesis detector produces very inaccurate measurements and such clutter measurements have a negative impact on SLAM.

2) *Online SLAM Performance*: The positioning RMSE of the proposed filter as a function of β is illustrated in Fig. 2 and in this section, we fix the threshold to $\beta = 10^{-1}$. The performance of BM1 and the proposed approach is illustrated in Fig. 3. BM1 yields good SLAM performance in majority of the UE locations, but as illustrated in Fig. 3a, in few locations

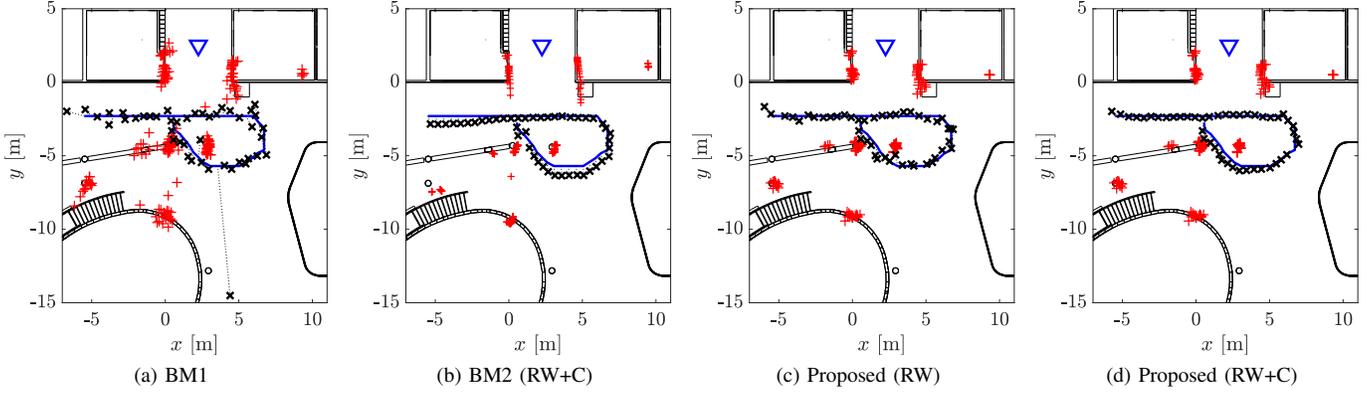


Fig. 3. Example SLAM performance using the benchmark algorithms and the proposed filtering approach with two different dynamic models. In the figures $\hat{\mathbf{p}}_{\text{UE}}$ and $\hat{\mathbf{p}}_i$'s shown for all 45 UE locations and the legend of the figures is: \mathbf{p}_{BS} (∇), \mathbf{p}_{UE} (—), $\hat{\mathbf{p}}_{\text{UE}}$ (\times) and $\hat{\mathbf{p}}_i$'s ($+$).

TABLE I
PERFORMANCE SUMMARY OF THE ALGORITHMS

| Method | Model | Pos. [m] | Head. [deg] | Clk [ns] | Time [ms] |
|----------|-------|----------|-------------|----------|-----------|
| BM1 | - | 1.47 | 4.62 | 4.95 | 6.00 |
| BM2 | RW | 0.48 | 2.45 | 1.30 | 165.13 |
| BM2 | RW+C | 0.46 | 2.23 | 1.50 | 152.02 |
| Proposed | RW | 0.40 | 2.28 | 1.34 | 6.44 |
| Proposed | RW+C | 0.28 | 2.09 | 0.94 | 6.40 |
| BM2* | RW | 0.35 | 2.43 | 0.76 | 166.33 |
| BM2* | RW+C | 0.20 | 2.00 | 0.34 | 156.56 |

*Initialized using the ground truth UE state.

$\hat{\mathbf{p}}_{\text{UE}}$ is very inaccurate, for example at $\mathbf{p}_{\text{UE}} = [3.56, -5.70]^T$. At these locations, the cost function is typically multimodal and BM1 selects the mode with minimum cost as the estimate which is not the one closest to the ground truth. The proposed method generates a measurement for every local minima that has low enough cost which are then used as input to the NN filter. Since the minima closest to the ground truth is typically also the one closest to the predicted state estimate, the NN filter is able to accurately estimate the UE state at every UE location as illustrated in Fig. 3. Moreover, comparing Figs. 3c and 3d, it is evident that a better dynamic model enhances the UE tracking performance, but it is important to note that the mapping performance is equivalent with both dynamic models since the UE and landmark measurements are decoupled using the proposed approach.

The performance of the algorithms are summarized in Table I. With respect to BM1, the proposed method improves the accuracy notably with only a slight increase in computational cost as tabulated in the last column of Table I. The performance with respect to BM2 depends on how the filter is initialized. If BM2 is initialized utilizing the snapshot SLAM algorithm, the proposed algorithm clearly outperforms BM2. The main drawback of BM2 when initialized using the measurements is that typically the initial uncertainty of the UE state is high and occasionally the filter converges to the wrong mode. Several different countermeasures were tried, such as utilizing more particles and using a burn-in period after initialization, but they were not effective in practice. On the other hand, if BM2 is initialized using the ground truth UE state, BM2 outperforms the proposed method but this initialization procedure is very unrealistic. BM2 also has a notably higher computational cost as tabulated in Table I.

V. CONCLUSIONS

In this paper, we presented an online bistatic radio SLAM method that combines properties of conventional snapshot and filtering-based solutions. The main advantage of the proposed method is that the underlying measurement models are linear so that the filtering recursion can be implemented using low complexity linear Bayesian filtering equations. Experiments conducted using real-world 60 GHz measurements demonstrated that the proposed method combines good accuracy with low computational overhead, outperforming state-of-the-art snapshot and filtering-based bistatic radio SLAM methods.

REFERENCES

- [1] 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, 2021.
- [2] A. Shahmansoori *et al.*, "Position and orientation estimation through millimeter-wave MIMO in 5G systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1822–1835, 2018.
- [3] H. Wymeersch *et al.*, "5G mmWave downlink vehicular positioning," in *IEEE GLOBECOM*, 2018, pp. 206–212.
- [4] F. Wen *et al.*, "5G synchronization, positioning, and mapping from diffuse multipath," *IEEE Wireless Commun. Lett.*, vol. 10, no. 1, pp. 43–47, 2021.
- [5] O. Kaltiokallio *et al.*, "Towards real-time radio-SLAM via optimal importance sampling," in *IEEE SPAWC*, 2022, pp. 1–5.
- [6] H. Kim *et al.*, "5G mmWave cooperative positioning and mapping using multi-model PHD filter and map fusion," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 3782–3795, Jun. 2020.
- [7] O. Kaltiokallio *et al.*, "Robust snapshot radio SLAM," 2024. [Online]. Available: <https://arxiv.org/abs/2404.10291>
- [8] E. Rastorgueva-Foi *et al.*, "Millimeter-wave radio SLAM: End-to-end processing methods and experimental validation," *IEEE J. Select. Areas Commun.*, pp. 1–1, 2024.
- [9] P. Tichavsky *et al.*, "Posterior Cramer-Rao bounds for discrete-time nonlinear filtering," *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1386–1396, 1998.
- [10] S. Challa *et al.*, *Fundamentals of Object Tracking*. Cambridge University Press, 2011.
- [11] B.-N. Vo *et al.*, "The Gaussian mixture probability hypothesis density filter," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4091–4104, 2006.
- [12] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge University Press, 2013.
- [13] O. Kaltiokallio *et al.*, "mmWave simultaneous localization and mapping using a computationally efficient EK-PHD filter," in *IEEE FUSION 2021*, 2021, pp. 1–8.
- [14] J. S. Mullane *et al.*, *Random Finite Sets for Robot Mapping & SLAM: New Concepts in Autonomous Robotic Map Representations*. Springer Publishing Company, Incorporated, 2013.