



## **5G SLAM Using the Clustering and Assignment Approach with Diffuse Multipath**

Downloaded from: <https://research.chalmers.se>, 2024-04-23 06:25 UTC

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

Ge, Y., Wen, F., Kim, H. et al (2020). 5G SLAM Using the Clustering and Assignment Approach with Diffuse Multipath. *Sensors*, 20(16): 1-31. <http://dx.doi.org/10.3390/s20164656>

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

## Article

# 5G SLAM Using the Clustering and Assignment Approach with Diffuse Multipath

Yu Ge <sup>1,\*</sup> , Fuxi Wen <sup>1</sup> , Hyowon Kim <sup>2</sup> , Meifang Zhu <sup>3</sup>, Fan Jiang <sup>1</sup> , Sunwoo Kim <sup>2</sup> ,  
Lennart Svensson <sup>1</sup> and Henk Wymeersch <sup>1</sup> 

<sup>1</sup> Department of Electrical Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden; fuxi@chalmers.se (F.W.); fanji@chalmers.se (F.J.); lennart.svensson@chalmers.se (L.S.); henkw@chalmers.se (H.W.)

<sup>2</sup> Department of Electronic Engineering, Hanyang University, Seoul 04763, Korea; khw870511@hanyang.ac.kr (H.K.); remero@hanyang.ac.kr (S.K.)

<sup>3</sup> Department of Electrical and Information Technology, Lund University, 221 00 Lund, Sweden; meifang.zhu@eit.lth.se

\* Correspondence: yuge@chalmers.se

Received: 16 July 2020; Accepted: 14 August 2020; Published: 18 August 2020



**Abstract:** 5G communication systems operating above 24 GHz have promising properties for user localization and environment mapping. Existing studies have either relied on simplified abstract models of the signal propagation and the measurements, or are based on direct positioning approaches, which directly map the received waveform to a position. In this study, we consider an intermediate approach, which consists of four phases—downlink data transmission, multi-dimensional channel estimation, channel parameter clustering, and simultaneous localization and mapping (SLAM) based on a novel likelihood function. This approach can decompose the problem into simpler steps, thus leading to lower complexity. At the same time, by considering an end-to-end processing chain, we are accounting for a wide variety of practical impairments. Simulation results demonstrate the efficacy of the proposed approach.

**Keywords:** mmWave single base station SLAM; diffuse multipath; clustering

## 1. Introduction

Progressive generations of cellular communication systems have a common property—each generation relies on a larger bandwidth and higher carrier frequency than the previous one. This is particularly noticeable in 5G, where in Frequency Range 2 (FR2), carriers above 24 GHz are utilized in combination with contiguous bandwidths of up to 400 MHz. Early commercial deployments in the US have demonstrated data rates beyond 1 Gbps [1]. The same increases in bandwidth and carrier frequency also have direct benefits for positioning [2]. This has led to intense research activities into 5G positioning as well as a dedicated study item in 3GPP. The final report of this study in 3GPP [3] revealed that by using both delay and angle measurements, it is possible to satisfy commercial positioning performance requirements. Beyond positioning, the sparse nature of the channel at higher carriers allows the receiver to resolve the multipath components, which transforms them from foe to friend [4,5]. In particular, the geometric nature of the channel at mmWave frequencies leads to a relation between the multipath parameters and the physical environment, as the parameters are related to the location of the user (UE) with respect to the base station (BS) and the propagation environment. The simultaneous localization and mapping problem is then invoked to invert multipath parameters into geometric information that determines the user's position and the locations of objects, based on the signal from a single base station. Several papers have addressed this problem, exploiting both

line-of-sight (LOS) and non-LOS (NLOS) paths for position estimation, synchronization, and mapping in mmWave multiple-input multiple-output (MIMO) [6–8] and mmWave multiple-input single-output (MISO) [9,10] contexts. However, this inversion is challenging due to a number of reasons: (i) estimation of the channel parameters involves solving a high-dimensional problem; (ii) the association between the objects in the environment and the measurements is not available; (iii) an object may give rise to multiple measurements, due to diffuse multipath; (iv) the measurements per object are not clustered. We will now treat each of these four challenges in turn.

Channel estimation for mmWave is a rich research area, which we cannot cover in detail here. Rather, we categorize these methods as search-based and search-free. Search-based methods, such as those using maximum likelihood (ML) [11] and compressed sensing (CS) techniques [12,13], require an exhaustive search in the high-dimensional space of channel parameters, which entails high complexity. On the other hand, search-free methods [14], such as matrix or tensor decomposition based subspace methods, directly provide estimates of the channel parameters [15] or rely on low-dimensional search [16], thus avoiding the need for high-dimensional optimization. An important challenge of channel estimation for SLAM is that the different dimensions (angles of arrival and departure, delays and gains) should be correctly matched.

The unknown association between measurements and objects is a common problem in SLAM, and powerful methods to address it can be found in the literature [17,18]. SLAM with radio-based measurements has been considered in the context of ultra-wideband (UWB) communication [19,20] using only distance measurements, which is referred to as multipath-assisted SLAM, or channel SLAM. Focusing on the application of SLAM in a 5G mmWave context (which we designate ‘5G SLAM’), message passing-based estimators were introduced in References [7,21], based on the concept of non-parametric belief propagation, without the data association (DA). Extension of such methods to include the hidden DAs is possible, following the approaches from Reference [22]. In Reference [23], the probability hypothesis density (PHD) filter, which is a random-finite-set filter, was used to solve the 5G SLAM problem, considering only one measurement per object. In Reference [24], a more powerful random-finite-set filter, Poisson multi-Bernoulli mixture (PMBM) filter, was used, which enumerates all possible DAs.

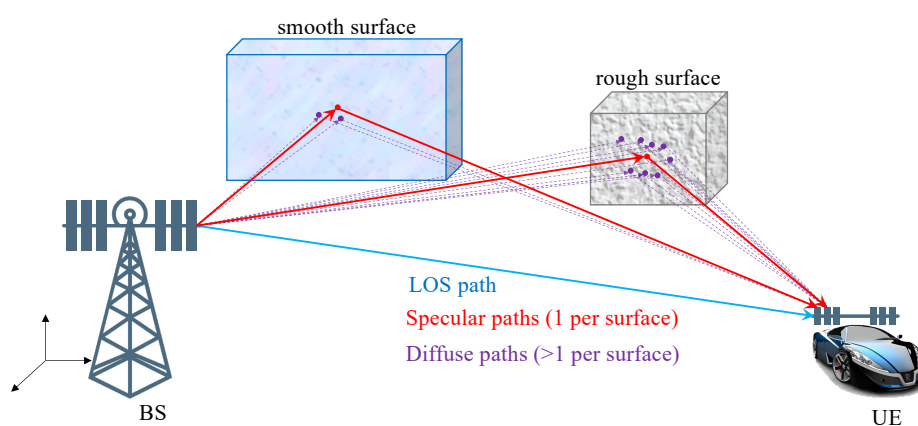
Multiple measurements per object, while common in the SLAM and extended object tracking literature [25], have been generally ignored in the above 5G SLAM works. The multiple measurements per object are caused by diffuse multipath due to object roughness with respect to the wavelength, depicted in Figure 1. In Reference [5], diffuse multipath is seen as a perturbation, leading to false measurements. In Reference [26], exploitation of the diffuse multipath in radar is proposed by means of diffuse multipath statistics. In Reference [27], surface roughness was considered in radar applications, modeled as a number of sub-reflectors, in an environment with known wall geometry. A similar model with random sub-reflectors was evaluated in Reference [28], where the estimated diffuse paths were used for positioning and mapping, using a simple geometric approach.

Finally, when measurements from an object are not clustered, the unknown grouping can be considered within the SLAM filter [29–31], though this comes at a high computational cost. In Reference [28], a K-means clustering was utilized, but this requires a priori knowledge of the number of clusters. In Reference [32], K-power-means was proposed, as well as several criteria to decide the number of clusters. In Reference [24], the perfect clustering was assumed.

In this paper, we address the aforementioned challenges, building on the extensive literature in each of the above research areas in order to provide an end-to-end framework for SLAM harnessing diffuse multipath. The proposed end-to-end framework provides a general approach for user localization and environment mapping in 5G downlink transmissions from a single BS. Therefore, the purposed framework can be utilized in many application areas, including personal navigation [4], localization of cars and robots [33], smart homes [34], indoor location analysis [35], immersive customer experiences [36], location-aided communication [37], personal radar [38], to name but a few. Moreover, the proposed framework can form a foundation for future Beyond 5G and 6G

localization and sensing approaches [39]. Our framework is built on a layered approach, comprising three main parts (channel estimation, clustering, and SLAM), which are evaluated separately and end-to-end. The main contributions of this paper are as follows:

- The description of an end-to-end framework for SLAM harnessing diffuse multipath and its performance evaluation.
- The evaluation of clustering and assignment methods, which is suitable for estimated channel parameters under both specular and diffuse multipath, as well as a method to utilize the estimated channel gains for improving the clustering in the 5G SLAM problem.
- The extension of the 5G SLAM likelihood function, in order to harness both specular and diffuse multipath components and to classify different object types according to their roughness, while accounting for clustering errors.



**Figure 1.** In mmWave simultaneous localization and mapping (SLAM) applications, each object can give rise to a specular path as well as multiple diffuse paths. The number and spread of these diffuse paths depend on the roughness of the object. At the receiver side, diffuse paths from an object have similar delays and angles, so that they can only be resolved with sufficient bandwidth and number of antennas.

The novelty of the proposed approach compared to the existing random finite set (RFS) based 5G SLAM work [23,40] is three-fold—first of all, References [23,40] did not use a real channel estimator, which makes the problem easier. Secondly, they assumed at most one measurement from an object, which is not the real case. Finally, the PHD filter is not optimal, which does not contain the enumeration of the different data associations. In Reference [24], although the measurements are from the ESPRIT channel estimator and the diffuse multipath is considered, the channel gain is still ignored and the channel estimation results are assumed to be well grouped based on the source. In the current paper, we study the whole framework, from downlink signals to SLAM filter. We also fully use the information given by the channel estimator, including diffuse multipath and channel gain. The PMBM filter is used, which is optimal and enumerates all possible data associations.

The remainder of this paper is structured as follows. The system model is described in Section 2, including the signal model, environment model, sensor, and measurement model. The end-to-end framework is then presented in Section 3, specifying the components that will be detailed in the subsequent sections, starting with channel estimation in Section 4, clustering in Section 5, and the novel likelihood in Section 6. Simulation results are presented in Section 7, followed by our conclusions in Section 8. The paper also contains several appendices describing the geometric expressions of the channel parameters, as well as the details of the SLAM method.

## Notations

Scalars (e.g.,  $x$ ) are denoted in italic, vectors (e.g.,  $\mathbf{x}$ ) in bold, matrices (e.g.,  $\mathbf{X}$ ) in bold capital letters, sets (e.g.,  $\mathcal{X}$ ) in calligraphic, tensors (e.g.,  $\mathcal{X}$ ) in bold calligraphic. Transpose and Hermetian are denoted by  $\cdot^T$  and  $\cdot^H$ , and  $j = \sqrt{-1}$ . Furthermore,  $|\cdot|$  denotes the absolute value of a scalar, or the cardinality of a set;  $\|\cdot\|$  denotes the Euclidean norm of a vector. A Gaussian density with mean  $\mu$  and covariance  $\Sigma$ , evaluated in value  $x$  is denoted by  $\mathcal{N}(x; \mu, \Sigma)$ . Finally, the notations of important variables are summarized in Table 1.

**Table 1.** Notations of important variables.

Notation	Definition	Notation	Definition
$s$	state of the user	$\mathbf{x}_{\text{LM}}$	landmark location
$m$	landmark type	$\theta$	angle of arrival (AOA) pair
$\phi$	angle of departure (AOD) pair	$\tau$	time of arrival (TOA)
$g$	channel gain	$\mathcal{Z}$	measurement set
$z$	single measurement	$c$	speed of light
$p_D$	detection probability	$k$	time index
$i$	surface index	$l, p$	path index
$s$	subcarrier index	$r$	dimension index

## 2. System Model

In this section, we describe the basic models of the user, the environment, the propagation channel, and the observed waveform. We use the 5G positioning reference signals (PRS) [41] as pilot signals. These signals are broadcast by the BS, and used for positioning according to the 3GPP standards [3,42]. Since the signals are broadcast, there is no multi-user issue or interference, so that we only focus on a single UE, where the proposed algorithms can be assumed to be executed per UE.

### 2.1. User Model

In this paper, we only consider the single-user scenario, so the cooperation between different users is beyond the scope of this paper. The dynamic state of the user at time step  $k$  is denoted by  $\mathbf{s}_k$ , which contains the position of the user  $\mathbf{x}_{\text{UE},k} = [x_k, y_k, z_k]^T$ , heading  $\omega_k$ , translation speed  $\zeta_k$ , turn rate  $\varrho_k$  and clock bias  $B_k$ . The dynamic model of  $\mathbf{s}_k$  is

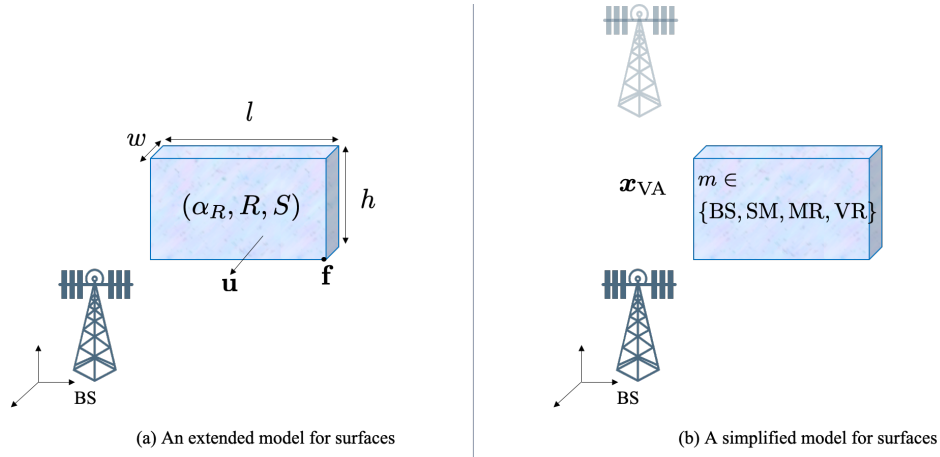
$$\mathbf{s}_k = \mathbf{v}(\mathbf{s}_{k-1}) + \mathbf{q}_k, \quad (1)$$

where  $\mathbf{v}(\cdot)$  denotes a known transition function;  $\mathbf{q}_k$  is the process noise, modeled as a zero-mean Gaussian with known covariance  $\mathbf{Q}_k$ . The above dynamic model can be expressed, equivalently, in terms of a transition density  $f(\mathbf{s}_k|\mathbf{s}_{k-1}) = \mathcal{N}(\mathbf{s}_k; \mathbf{v}(\mathbf{s}_{k-1}), \mathbf{Q}_k)$ .

### 2.2. Environment Model

We consider an environment with a few landmarks. There is a fixed and known BS, located at  $\mathbf{x}_{\text{BS}} \in \mathbb{R}^3$ . The other unknown landmarks are modeled as different types of surfaces (see Figure 2). A surface can be described with an extended state comprising:

- A point  $\mathbf{f}$  on a corner of the surface and a vector  $\mathbf{u}$  normal to the surface.
- The size of the surface in length  $l$  and height  $h$ . The width  $w$  is not relevant.
- The smoothness of the surface, denoted by  $\alpha_R \geq 0$ .
- The scattering attenuation  $1 \geq S \geq 0$  and reflection attenuation  $1 \geq R \geq 0$ , with  $R^2 + S^2 \leq 1$ , in which the remaining power is absorbed in the surface.



**Figure 2.** Two different surface models: (a) shows a high-dimensional model, while (b) is a compact model that expresses the location of the surface via a virtual anchor.

To avoid dealing with such a high-dimensional surface state description, we use a simplified state, containing a fixed virtual anchor (VA) with location  $\mathbf{x}_{VA} \in \mathbb{R}^3$ , which is the reflection of the BS with respect to the surface, that is,

$$\mathbf{x}_{VA} = (\mathbf{I} - 2\mathbf{u}\mathbf{u}^T)\mathbf{x}_{BS} + 2\mathbf{f}^T\mathbf{u}\mathbf{u}. \quad (2)$$

Note that the VA location is surface-specific. We only consider 3 different types of surfaces: smooth surface (SM, with  $S = 0, R = 0.8, \alpha_R = 100$ ), medium rough surface (MR, with  $S = 0.4, R = 0.6, \alpha_R = 4$ ) and very rough surface (VR, with  $S = 0.8, R = 0, \alpha_R = 0$ ). Signals from the base station can only be reflected via SM to the receiver; while they can be reflected and diffused via MR to the receiver; whereas VR can only diffuse signals. This allows us to have a compact state representation, with the landmark state  $\mathbf{x} = [\mathbf{x}_{LM}^T, m]^T$ , where  $m \in \{BS, SM, MR, VR\}$  and  $\mathbf{x}_{LM} = \mathbf{x}_{BS}$  for  $m = BS$ , while  $\mathbf{x}_{LM} = \mathbf{x}_{VA}$ , for  $m \neq BS$  [24].

### 2.3. Channel Model

At time step  $k$ , the channel from the BS to the user in the frequency domain is given by [11]:

$$H_k(f, \boldsymbol{\theta}, \boldsymbol{\phi}) = \sum_{i=0}^{I_k-1} \sum_{l=0}^{L_k^i-1} g_k^{i,l} \delta(\boldsymbol{\theta} - \boldsymbol{\theta}_k^{i,l}) \delta(\boldsymbol{\phi} - \boldsymbol{\phi}_k^{i,l}) e^{-j2\pi f \tau_k^{i,l}}, \quad (3)$$

where  $\delta(\cdot)$  is the delta Dirac function,  $f$  is the frequency,  $\boldsymbol{\theta}$  is the angle of arrival (AOA),  $\boldsymbol{\phi}$  is the angle of departure (AOD),  $I_k$  is the number of landmarks in the environment ( $i = 0$  represents the BS);  $L_k^i$  is the number of paths from each landmark ( $l = 0$  represents the specular path). For each surface, there is at most one deterministic specular component among the paths, while all remaining paths are diffuse components and thus random [28]. We denote the total number of paths by  $P_k = \sum_{i=1}^{I_k} L_k^i$ . Each path  $i, l$  can be described by a complex gain  $g_k^{i,l}$ , a time of arrival (TOA)  $\tau_k^{i,l}$ , an AOA pair  $\boldsymbol{\theta}_k^{i,l}$  in both azimuth and elevation, and an AOD pair  $\boldsymbol{\phi}_k^{i,l}$  in both azimuth and elevation. The generative model for each of these parameters is described as

- *LOS path:* When  $i = 0, L_k^i = 1$ , the gain has uniform phase and has power

$$|g_k^{0,0}|^2 = \frac{\lambda^2}{(4\pi)^2 \|\mathbf{x}_{UE} - \mathbf{x}_{BS}\|^2}, \quad (4)$$

where  $\lambda$  is the wavelength, and the TOA, AOA, and AOD follow the geometric relations between the BS and the UE. They are given in Appendix A.

- *Specular path from surface  $i$ :* For  $i > 0, l = 0$ , the point of incidence on the  $i$ -th surface (denoted by  $\mathcal{L}_i$  with virtual anchor  $\mathbf{x}_{VA,i}$ ) is the intersection of the surface  $\mathcal{L}_i$  and the line between the  $i$ -th virtual anchor  $\mathbf{x}_{VA,i}$  and the UE  $\mathbf{x}_{UE}$ . The specular path gain has uniform phase and power

$$|g_k^{i>0,0}|^2 = \frac{\lambda^2 R^2}{(4\pi)^2 \|\mathbf{x}_{UE} - \mathbf{x}_{VA,i}\|^2}. \quad (5)$$

The TOA, AOA, and AOD follow the relative position of the UE, BS, and the incidence point on the surface. They are given in Appendix A.

- *Diffuse paths from surface  $i$ :* For  $i > 0, l > 0$ , the number of paths per surface and their spread in angle and delay, as well as the channel gains, depend on the roughness of that surface. These paths can be interpreted as coming from random points  $\mathbf{x}_k^{i,l}$  on the surface, with a spatial distribution  $p_i(\mathbf{x}_{sc}|\mathbf{x}_{UE}, \mathbf{x}_{BS})$  that depends on the roughness, where  $\mathbf{x}_{sc}$  is the random variable that describes the position of the diffuse point. The diffuse points are generated from the distribution  $\mathbf{x}_k^{i,l} \sim p_i(\mathbf{x}_{sc}|\mathbf{x}_{UE}, \mathbf{x}_{BS})$  ([43], Chapter 3)

$$p_i(\mathbf{x}_{sc}|\mathbf{x}_{UE}, \mathbf{x}_{BS}) \propto \begin{cases} \frac{(1+\cos \nu_1(\mathbf{x}_{sc}))^{\alpha_R}}{2} \frac{|\cos \nu_2(\mathbf{x}_{sc})| |\cos \nu_3(\mathbf{x}_{sc})|}{\|\mathbf{x}_{sc}-\mathbf{x}_{BS}\|^2 \|\mathbf{x}_{sc}-\mathbf{x}_{UE}\|^2} & \mathbf{x}_{sc} \in \mathcal{L}_i \\ 0 & \text{else,} \end{cases} \quad (6)$$

where  $\mathcal{L}_i \subset \mathbb{R}^3$  denotes the set of points that make the  $i$ -th surface,  $\nu_1(\mathbf{x}_{sc})$  is the deviation of the scattering angle with respect to the angle of the specular path (i.e.,  $\nu_1(\mathbf{x}_{sc}) = 0$  when  $\mathbf{x}_{sc}$  is the incidence point of the specular path),  $\nu_2(\mathbf{x}_{sc})$  is the angle between the impinging ray (i.e., from the transmitter to  $\mathbf{x}_{sc}$ ) and the surface normal, and  $\nu_3(\mathbf{x}_{sc})$  is the angle between the departing ray and the surface normal. The diffuse paths have uniform and independent phases and equal power

$$|g_k^{i>0,l>0}|^2 = \frac{\lambda^2 S^2 (\alpha_R + 1)}{(4\pi)^2 (L_k^i - 1) (1 - 2^{-(\alpha_R+1)}) (\alpha_R + 2) \|\mathbf{x}_{UE} - \mathbf{x}_{VA,i}\|^2}. \quad (7)$$

The locations of the diffuse points  $\mathbf{x}_k^{i,l}$  fully determine their corresponding TOA, AOA, and AOD, provided in Appendix A.

#### 2.4. Signal Model

We will assume a transmitter and receiver with uniform rectangular arrays (URAs) with  $M_T = M_1 \times M_2$  (i.e.,  $M_1$  columns and  $M_2$  rows) and  $M_R = M_3 \times M_4$  elements, respectively. The corresponding steering vectors are

$$\mathbf{a}_T(\boldsymbol{\phi}) = [\phi_{az}, \phi_{el}] = \mathbf{a}_c(\boldsymbol{\phi}) \otimes \mathbf{a}_r(\boldsymbol{\phi}) \quad (8)$$

$$[\mathbf{a}_c(\boldsymbol{\phi})]_{m_c} = \exp(j\pi m_c \sin(\phi_{el})), \quad m_c \in \{0, \dots, M_2 - 1\} \quad (9)$$

$$[\mathbf{a}_r(\boldsymbol{\phi})]_{m_r} = \exp(j\pi m_r \cos(\phi_{el}) \sin(\phi_{az})), \quad m_r \in \{0, \dots, M_1 - 1\} \quad (10)$$

and

$$\mathbf{a}_R(\boldsymbol{\theta}) = [\theta_{az}, \theta_{el}] = \mathbf{a}_c(\boldsymbol{\theta}) \otimes \mathbf{a}_r(\boldsymbol{\theta}) \quad (11)$$

$$[\mathbf{a}_c(\boldsymbol{\theta})]_{m_c} = \exp(j\pi m_c \sin(\theta_{el})), \quad m_c \in \{0, \dots, M_4 - 1\} \quad (12)$$

$$[\mathbf{a}_r(\boldsymbol{\theta})]_{m_r} = \exp(j\pi m_r \cos(\theta_{el}) \sin(\theta_{az})), \quad m_r \in \{0, \dots, M_3 - 1\}, \quad (13)$$

in which  $\otimes$  is the Kronecker product.



In addition, we consider OFDM transmission with  $M_S = M_5$  subcarriers and subcarrier spacing  $\Delta f$ . The received signal at subcarrier  $s$  at time step  $k$  is then [44]

$$\mathbf{Y}_{s,k} = \underbrace{\sum_{i=0}^{I_k-1} \sum_{l=0}^{L_k^i-1} g_k^{i,l} \mathbf{a}_R(\boldsymbol{\theta}_k^{i,l}) \mathbf{a}_T^H(\boldsymbol{\phi}_k^{i,l}) e^{-j2\pi s \Delta f \tau_k^{i,l}}}_{\mathbf{H}_s} \mathbf{S}_s + \mathbf{N}_{s,k}, \quad (14)$$

where  $\mathbf{S}_s \in \mathbb{C}^{M_T \times T}$  is the pilot signal over subcarrier  $s$  (spanning  $T$  OFDM symbols),  $\mathbf{N}_{s,k}$  is white Gaussian noise with  $\text{vec}(\mathbf{N}_{s,k}) \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{M_R T})$ , and  $\mathbf{H}_s$  is the channel frequency response,

### 3. Methodology and End-to-End Framework

Using the raw measurements  $\mathbf{Y}_s$  directly in SLAM is challenging, due to the high dimensionality of the measurement, the complex nonlinear relation to the user and landmark states, and the fact that not all paths may be resolvable, due to a limited number of transmitter and receiver antennas and bandwidth. While such direct localization has performance benefits [45–47], we instead consider a *layered approach*, visualized in Figure 3, comprising the following steps after downlink transmission of the signals  $\mathbf{Y}_s$ .

1. First of all, channel estimation is performed to recover the channel parameters (angles, delays, gains). Due to the finite resolution at the receiver side, not all paths are resolvable. Hence, the number of estimated paths (denoted by  $\hat{P}_k$ ) will be much smaller than  $I_k \times L_k^i$ . The channel estimator thus provides a set of channel parameter estimates  $\mathcal{Z}_k$  at time  $k$ ,  $\mathcal{Z}_k = \{\mathbf{z}_k^0, \mathbf{z}_k^1, \dots, \mathbf{z}_k^{\hat{P}_k-1}\}$ . Each element  $\mathbf{z}_k^p \in \mathcal{Z}_k$  is either a clutter, which is caused by noise peaks that are detected as paths during channel estimation, with clutter intensity  $c(\mathbf{z})$  or follows

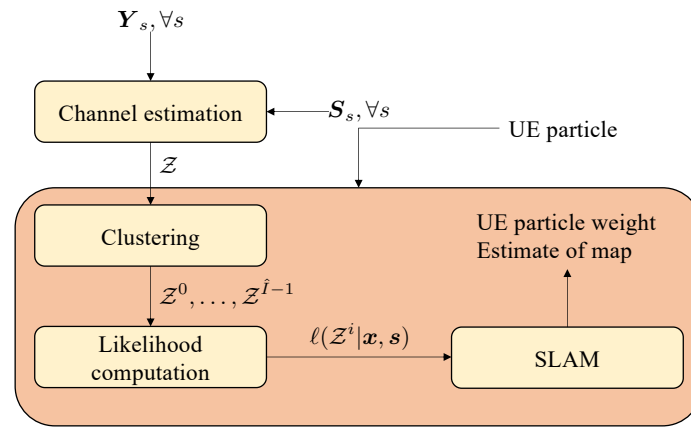
$$\mathbf{z}_k^p = [\tau_k^p, (\boldsymbol{\theta}_k^p)^\top, (\boldsymbol{\phi}_k^p)^\top, g_k^p]^\top + \mathbf{w}_k^p \quad (15)$$

$$= \mathbf{h}(\mathbf{x}_k^p, \mathbf{s}_k) + \mathbf{w}_k^p, \quad (16)$$

where  $\mathbf{w}_k^p$  denotes the measurement noise;  $\mathbf{x}_k^p$  is  $\mathbf{x}_{BS}$  for LOS, the incidence point of the deterministic specular components, or a (random) point on the surface for a NLOS component. We recall that the underlying geometric relation  $\mathbf{h}(\mathbf{x}_k^p, \mathbf{s}_k)$  can be found in Appendix A. We describe the channel estimator in Section 4.

2. After channel estimation, we group the unordered elements in  $\mathcal{Z}_k$  in clusters  $\mathcal{Z}_k^i$ , where each cluster should correspond to one landmark. This removes the need to consider all possible partitions of the measurements in the SLAM method, drastically reducing overall complexity. Clustering is challenging as measurement clusters may be non-convex. In addition, diffuse paths may be far away from the specular paths, leading to possible miss-classifications. The proposed clustering method is described in Section 5.
3. Finally, after clustering, the SLAM method requires a likelihood function that expresses the statistical relation between the state and the clustered measurements,  $\ell(\mathcal{Z}_k^i | \mathbf{x}, \mathbf{s}_k)$ . The SLAM method is deferred to Appendix B, while in the main text we focus on the proposed likelihood function in Section 6. The SLAM filter follows a Rao-Blackwellized approach, where we use a set of particles (indexed by  $n$ ) to represent the user state, and use PMBM densities conditioned on each particle to represent the map. Clustering and likelihood computation are conditioned in the user state and are thus performed per particle.





**Figure 3.** Proposed layered approach for SLAM from the observations  $\mathbf{Y}_s$ . The time index  $k$  is omitted.

#### 4. Channel Estimation

In this section, the ESPRIT channel estimator is introduced.

##### 4.1. Background

The standard formulation of the channel estimation problem is an ML problem, where

$$\hat{\Theta}_k = \arg \min_{\Theta} - \sum_s \log p(\mathbf{Y}_{s,k} | \Theta, \mathbf{S}_s), \quad (17)$$

in which  $\Theta$  contains the delays, gains, and angles of all the paths, as well as the number of paths. Since the number of paths is unknown and paths are not all resolvable, model order selection techniques can be applied, for example, by adding a regularizer to (17) ([11], Section 5.2), for example, following the minimum description length (MDL) or Akaike information criterion (AIC). The dense multipath can be treated as a stochastic process, by writing (14) as

$$\mathbf{Y}_{s,k} = \sum_{i=0}^{I_k-1} g_k^{i,0} \mathbf{a}_R(\theta_k^{i,0}) \mathbf{a}_T^H(\phi_k^{i,0}) e^{-j2\pi s \Delta f \tau_k^{i,0}} \mathbf{S}_s + \mathbf{W}_{s,k}^{\text{dm}} + \mathbf{N}_{s,k}, \quad (18)$$

where  $\mathbf{W}_{s,k}^{\text{dm}}$  is the dense multipath containing all the contributions from the diffuse paths. The dense multipath is then modeled as complex Gaussian process with zero mean and covariance  $\mathbf{R}_k^{\text{dm}}$ , possibly with unknown elements. Then the entire noise process has covariance  $\mathbf{R}_k^{\text{dm}} + \sigma^2 \mathbf{I}_{M_R T}$ . Alternatively, we can estimate the geometric parameters of the dominant diffuse multipath components from MIMO channel measurements, which leads to a multidimensional harmonic retrieval (HR) problem [48]. Numerous HR techniques have been developed, ranging from multidimensional searching or optimization based, such as multiple signal classification (MUSIC), ML and CS techniques [49], and polynomial-rooting or matrix-shifting based search-free methods, such as root-MUSIC [14] and estimation of parameters by rotational invariant techniques (ESPRIT) and their multidimensional extensions [50]. Due to its simplicity and high-resolution capability, ESPRIT-type algorithms have become one of the popular HR techniques [51].

##### 4.2. ESPRIT Channel Estimator

For notational convenience, we drop the time index  $k$ . From (14), the received signal on subcarrier  $s$  is of the form

$$\mathbf{Y}_s = \mathbf{H}_s \mathbf{S}_s + \mathbf{N}_s, \quad (19)$$

where  $\mathbf{H}_s$  is the channel frequency response,  $\mathbf{S}_s$  is a known pilot signal with orthogonality property ( $\mathbf{S}_s \mathbf{S}_s^H = \beta \mathbf{I}_{M_T \times M_T}$ ), and  $\mathbf{N}_s$  is i.i.d. Gaussian noise. Then we have

$$\mathbf{X}_s = \frac{1}{\beta} \mathbf{Y}_s \mathbf{S}_s^H = \mathbf{H}_s + \frac{1}{\beta} \mathbf{N}_s \mathbf{S}_s^H = \mathbf{H}_s + \mathbf{W}_s, \quad (20)$$

where  $\mathbf{W}_s$  is also i.i.d. Gaussian noise with a scaled covariance matrix.

#### 4.2.1. Observations in Tensor Form

We utilize a Tensor framework to exploit the  $R$ -D grid structure inherent in the data, as well as the Vandermonde structure in angle and delay domains. To this end, we map from geometric channel parameters to spatial frequencies by

$$\begin{aligned} \mu_{(1)}^p &= \pi \sin(\theta_{az}^p) \cos(\theta_{el}^p), & \mu_{(2)}^p &= \pi \sin(\theta_{el}^p), \\ \mu_{(3)}^p &= \pi \sin(\phi_{az}^p) \cos(\phi_{el}^p), & \mu_{(4)}^p &= \pi \sin(\phi_{el}^p), \\ \mu_{(5)}^p &= 2\pi \Delta_f \tau^p. \end{aligned} \quad (21)$$

For subcarrier  $i$ ,  $\mathbf{X}_i$  and  $\mathbf{W}_i$  are  $M_3 M_4 \times M_1 M_2$  matrices. We convert these  $M_5$  matrices (one per subcarrier) to a 5D tensor of suitable dimension,  $\mathcal{X}$ ,  $\mathcal{H}$  and  $\mathcal{W} \in \mathbb{C}^{M_1 \times M_2 \times M_3 \times M_4 \times M_5}$ . For the  $p$ -th path, the equivalent 5D array steering tensor can be written as

$$\mathcal{A}^p = \mathbf{a}(\mu_{(1)}^p) \circ \mathbf{a}(\mu_{(2)}^p) \circ \cdots \circ \mathbf{a}(\mu_{(5)}^p), \quad (22)$$

where  $\circ$  represents the outer product (Note that  $A = \mathbf{a}_1 \circ \mathbf{a}_2 = \mathbf{a}_1 \mathbf{a}_2^T$ , and  $\mathcal{A} = \mathbf{a}_1 \circ \mathbf{a}_2 \circ \mathbf{a}_3$  with  $A_{ijk} = a_{1i} a_{2j} a_{3k}$ ),  $\mu_{(r)}^p$  is the spatial frequency of the  $p$ -th source in the  $r$ -th dimension,  $\mathbf{a}(\mu_{(r)}^p) \in \mathbb{C}^{M_r \times 1}$  is equivalent to the uniform linear array steering vector composed of  $M_r$  sensors,  $p = 1, 2, \dots, P$ , where  $P$  is the total number of paths, and  $r = 1, 2, \dots, 5$ . This allows us to express the observation as

$$\mathcal{X} = \sum_{p=1}^P g^p \mathcal{A}^p + \mathcal{W} \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_5}, \quad (23)$$

where  $g^p$  denotes the complex path gain of the  $p$ -th path.

#### 4.2.2. Shift Invariance

We now introduce the multidimensional shift invariant structure in each dimension  $r$ , in order to apply Tensor-ESPRIT. We first introduce so-called *selection matrices*  $J_{1,(r)}, J_{2,(r)} \in \mathbb{R}^{M_r^{(\text{sel})} \times M_r}$ , which select  $M_r^{(\text{sel})}$  out of  $M_r$  elements belonging to the first and the second subarray in the  $r$ -th mode, respectively. For example, when  $M_r^{(\text{sel})} = M_r - 1$ , we have

$$J_{1,(r)} = [\mathbf{I}_{M_r-1} \quad \mathbf{0}_{(M_r-1) \times 1}] \text{ and } J_{2,(r)} = [\mathbf{0}_{(M_r-1) \times 1} \quad \mathbf{I}_{M_r-1}], \quad (24)$$

where  $\mathbf{I}$  and  $\mathbf{0}$  denote the identity matrix and zero vector, respectively, and the sizes are defined in subscripts. Secondly, we introduce  $\mathbf{A}_{(r)} = [\mathbf{a}(\mu_{(r)}^1) \quad \mathbf{a}(\mu_{(r)}^2) \quad \cdots \quad \mathbf{a}(\mu_{(r)}^P)]$ , which has the shift invariance property

$$J_{1,(r)} \mathbf{A}_{(r)} \Phi_{(r)} = J_{2,(r)} \mathbf{A}_{(r)}, \quad (25)$$

where  $\Phi_{(r)} = \text{diag}(e^{j\mu_{(r)}^1}, e^{j\mu_{(r)}^2}, \dots, e^{j\mu_{(r)}^L})$  is a diagonal matrix that contains the unknown parameters  $\mu_l^{(r)}, l = 1, 2, \dots, L$ , in dimension  $r$ .

### 4.2.3. Tensor-ESPRIT

In order to obtain the subspace spanned by  $\mathbf{A}_{(r)}$ , we take CANDECOMP/PARAFAC (CP) decomposition on  $\mathcal{X}$  [52]. Because the total number of paths  $P$  is unknown, model order selection techniques [53] can be utilized to estimate  $\hat{P}$ . In general, the estimated  $\hat{P} \ll P$  for rough surfaces with hundreds of closely located diffuse paths.

$$\mathcal{X} = \sum_{p=1}^{\hat{P}} \hat{g}^p \mathbf{u}_{(1)}^p \circ \mathbf{u}_{(2)}^p \circ \cdots \circ \mathbf{u}_{(5)}^p \equiv [\hat{\mathbf{g}}; \mathbf{U}_{(1)}, \mathbf{U}_{(2)}, \cdots, \mathbf{U}_{(5)}], \quad (26)$$

where  $\hat{\mathbf{g}} = [\hat{g}^1 \ \hat{g}^2 \ \cdots \ \hat{g}^{\hat{P}}]$  are the estimated path gains, and  $\mathbf{U}_{(r)} = [\mathbf{u}_{(r)}^1 \ \mathbf{u}_{(r)}^2 \ \cdots \ \mathbf{u}_{(r)}^{\hat{P}}]$  is the estimated signal subspace with normalized column vectors  $\mathbf{u}_{(r)}^p$ . Note that after taking CP decomposition, the path gains  $g^p$  and channel parameter associations are achieved synchronously. In other words,  $g^p$  corresponds to a unique  $\mathbf{u}_{(r)}^p$  in each dimension  $r$ , so that the output of (26) is  $\{\hat{g}^p, \mathbf{u}_{(1)}^p, \mathbf{u}_{(2)}^p, \cdots, \mathbf{u}_{(5)}^p\}, p = 1, 2, \cdots, \hat{P}$ .

Since  $\mathbf{A}_{(r)}$  and  $\mathbf{U}_{(r)}$  span the same subspace,  $\mathbf{A}_{(r)} = \mathbf{U}_{(r)} \mathbf{T}_{(r)}$ , where  $\mathbf{T}_{(r)} \in \mathbb{C}^{\hat{P} \times \hat{P}}$  is a non-singular transform matrix. Entering this relation in (25), we obtain

$$J_{1,(r)} \mathbf{U}_{(r)} \mathbf{\Psi}_{(r)} = J_{2,(r)} \mathbf{U}_{(r)}, \quad (27)$$

where  $\mathbf{\Psi}_{(r)} = (\mathbf{T}_{(r)})^{-1} \mathbf{\Phi}_{(r)} \mathbf{T}_{(r)}$ . In (27), only  $\mathbf{\Psi}_{(r)}$  is unknown, but can be estimated via the least squares method:

$$\hat{\mathbf{\Psi}}_{(r)} = (J_{1,(r)} \mathbf{U}_{(r)})^{\dagger} J_{2,(r)} \mathbf{U}_{(r)}, \quad (28)$$

where  $\cdot^{\dagger}$  denotes the pseudo-inverse. Since the matrix  $\mathbf{\Psi}_{(r)}$  is similar to the diagonal matrix  $\mathbf{\Phi}_{(r)}$ , they share the same eigenvalues. Hence, the spatial frequency can be recovered as  $\hat{\mu}_{(r)}^p = \arg(\lambda_{(r)}^p)$ , where  $\lambda_{(r)} = [\lambda_{(r)}^1, \lambda_{(r)}^2, \cdots, \lambda_{(r)}^{\hat{P}}]$  are the eigenvalues of  $\mathbf{\Psi}_{(r)}$ . The  $\{\hat{\mu}_{(r)}^p\}, r = 1, 2, \cdots, 5$ , map to delay  $\hat{\tau}^p$ , AOA pair  $\hat{\theta}^p = [\hat{\theta}_{az}^p, \hat{\theta}_{el}^p]^T$ , and AOD pair  $\hat{\phi}^p = [\hat{\phi}_{az}^p, \hat{\phi}_{el}^p]^T$  via (21).

Finally, each tuple  $[\hat{\tau}^p, \hat{\theta}_{az}^p, \hat{\theta}_{el}^p, \hat{\phi}_{az}^p, \hat{\phi}_{el}^p, \hat{g}^p]$ , for  $p \in \{1, \dots, \hat{P}\}$  is returned as the output of the channel estimator. We denote this combined output as  $\mathcal{Z}$ .

The most computationally demanding part of channel parameter estimation is the CP decomposition. In (Reference [54], Table 1), the complexities of major computations in popular CP decomposition algorithms is summarized. For example, the alternating least squares (ALS) algorithm with line search has a complexity of order  $\mathcal{O}(2^R \hat{P} J + R \hat{P}^3)$ , where  $J = \prod_{r=1}^R M_r$  and  $R = 5$  and  $\hat{P}$  denotes the total number of paths.

## 5. Channel Parameter Clustering

At time step  $k$ , the channel estimator provides a set of channel parameter estimates  $\mathcal{Z}_k$ ,  $\mathcal{Z}_k = \{z_k^0, z_k^1, \dots, z_k^{\hat{P}-1}\}$ , and we need to cluster  $\mathcal{Z}_k$  based on different landmarks. However, clustering is challenging, since  $\mathcal{Z}_k$  is in a 6D space, and the delay, angle, and gain are in different scales and spaces, so that they should be properly weighted. We also do not have any advance knowledge of the number of landmarks in the environment. In this section, we try to address these challenges. For brevity, we will omit the time index  $k$ .

### 5.1. Background

As a general term, a *cluster* is defined as a collection of objects that are similar to each other in some agreed-upon sense [55]. In radio channel analysis, a cluster is usually described as a group of multipath components (MPCs) with the same parameter distribution. The MPCs that have similar

values in delay and angular domain are jointly classified as a single cluster, which reflects the physical environment, as similar MPCs are usually spatially close to each other, for example, reach the receiver via the same landmark.

Clustering methodology in radio channel analysis has been expanded from visual clustering to automatic clustering, where the visual clustering is usually valid for data with limited dimensions [55,56], while the automatic clustering can handle data with more than three dimensions [57]. Automatic clustering algorithms, focusing on the parameter space of MPCs, such as Hierarchical, K-means, and Gaussian mixture, have been widely used in radio channel characterizations [32,58–61]. Hierarchical clustering algorithms cluster the data based on a binary cluster tree, which is limited by the data size, and therefore limited to small data sets. K-means [62] assumes a known number of clusters and iteratively assigns data to each cluster and computes cluster centroids until the cluster centroids are converged. While K-means is a widely used clustering method, it suffers from several drawbacks: (i) the number of clusters is predefined, which limits the capabilities to reflect the reality of the environment; (ii) it cannot cope with outliers, which implies all the observed data will eventually be part of some clusters, even observations that are far away in 6D space and should be considered as outliers; (iii) it usually gives spherical-like shape clusters. MPCs that reflect other physical properties can be assigned as the edge of a spherical cluster, which has notable effects on the cluster properties, that is, cluster centroids, and loses the ability to link to physical reality. The more sophisticated affinity propagation [63] partially avoids these problems, but still leads to circular clusters and has very high complexity. Gaussian mixture clustering [57] gives more variation to the shapes of the clusters extracted with K-means, and follows the similar structure of K-means. It can give clusters in different shape if the original data is not distributed circularly. However, the method is complex and converges slowly.

Density-based spatial clustering of applications with noise (DBSCAN) [64] is one of the most widely applied clustering methods for data sorting. Without a predefined number of clusters, DBSCAN utilizes two critical parameters—the minimum number of points clustered together for a region to be considered dense, and a distance measure to locate the points in the neighborhood of any point. The density-based clustering method has no limitations on the number of clusters, uses a dense distance, and gives more freedom regarding the shape of clusters. While we have evaluated different methods in Section 7, here we only present the best performing method, which is an extension of the DBSCAN algorithm.

## 5.2. Modified DBSCAN

In this section, we will present the DBSCAN clustering method with the obtained channel estimates. The method involves three phases—(i) the mapping of the geometric channel parameters (delay and angles) to a 3D point; (ii) clustering with DBSCAN; (iii) refinement by using the channel gain.

### 5.2.1. Phase 1: 5D to 3D Mapping

Distinct from existing work where the channel parameters are used as features for clustering [28,61], we transform the 5D channel parameter into a 3D position. In the absence of noise, this point should be on the surface of a landmark for NLOS paths, while in the presence of noise, the most likely estimation of the point can be founded close to the surface. The transformation reduces the dimensionality, and avoids different scales and spaces problems. Hence, the 5D parameters, that is, the delay  $\hat{\tau}^p$ , the AOA pair  $\hat{\theta}^p = [\hat{\theta}_{az}^p, \hat{\theta}_{el}^p]^T$  and the AOD pair  $\hat{\phi}^p = [\hat{\phi}_{az}^p, \hat{\phi}_{el}^p]^T$  can be converted to a point  $\hat{x}^p$ .

The general idea to transform the 5D parameters into a 3D point is illustrated in Figure 4. Specifically, for the  $p$ -th diffuse multipath, a virtual anchor position can be estimated as

$$\hat{x}_{LM}^p = x_{UE} + c(\hat{\tau}^p - \hat{B})f_R^p, \quad (29)$$

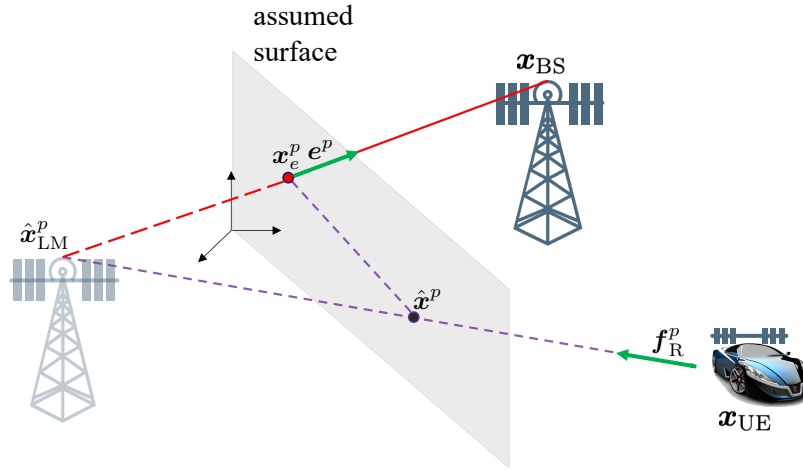
where  $c$  is the speed of light,  $\hat{B}$  is the estimated clock bias,  $\mathbf{f}_R^p$  is the unit vector pointing along with the direction of arrival, given by

$$\mathbf{f}_R^p = \begin{bmatrix} \cos \hat{\theta}_{az}^p \cos \hat{\theta}_{el}^p \\ \sin \hat{\theta}_{az}^p \cos \hat{\theta}_{el}^p \\ \sin \hat{\theta}_{el}^p \end{bmatrix}. \quad (30)$$

Given  $\hat{\mathbf{x}}_{LM}^p$  and  $\mathbf{x}_{BS}$ , a hypothesized surface can be defined via a point on the surface  $\mathbf{x}_e^p = (\mathbf{x}_{BS} + \hat{\mathbf{x}}_{LM}^p)/2$  and the normal vector to the surface  $\mathbf{e}^p = (\mathbf{x}_{BS} - \hat{\mathbf{x}}_{LM}^p)/\|\mathbf{x}_{BS} - \hat{\mathbf{x}}_{LM}^p\|$ . With knowledge of  $\hat{\mathbf{x}}_{LM}^p$  and the surface, the incidence point location can be computed from from Snell's law of reflection as

$$\hat{\mathbf{x}}^p = \hat{\mathbf{x}}_{LM}^p + \frac{(\mathbf{x}_e^p - \hat{\mathbf{x}}_{LM}^p)^T \mathbf{e}^p}{(\mathbf{x}_{UE} - \hat{\mathbf{x}}_{LM}^p)^T \mathbf{e}^p} (\mathbf{x}_{UE} - \hat{\mathbf{x}}_{LM}^p). \quad (31)$$

Note that this method also works for the LOS path, and it will return a point near  $\hat{\mathbf{x}}_{BS}$ , since  $\hat{\mathbf{x}}_{LM}^p$  and  $\mathbf{x}_e^p$  are both approximate to  $\hat{\mathbf{x}}_{BS}$ . From  $\mathcal{Z}$ , we obtain the corresponding  $\hat{P}$  3D points  $\mathcal{D} = \{\hat{\mathbf{x}}^0, \hat{\mathbf{x}}^1, \dots, \hat{\mathbf{x}}^{\hat{P}_k-1}\}$ . Next, we will perform the clustering methods with the 3D positions of points.



**Figure 4.** Illustration of the 3D transformation from the 5D parameters: from the channel estimates, a hypothesized landmark (a virtual anchor) location  $\hat{\mathbf{x}}_{LM}^p$  is determined. From the landmark and BS locations, the incidence point  $\hat{\mathbf{x}}^p$  is derived.

### 5.2.2. Phase 2: Clustering with DBSCAN

In DBSCAN [64], all points are classified into either clusters or identified as noise. In a specific cluster  $\mathcal{C}$ , the *core points* have an  $\epsilon$ -neighbourhood with  $N_{\min} \in \mathbb{N}$  points. We define a distance measure  $d: \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$  from which the  $\epsilon$ -neighbourhood of the  $p$ -th point is defined as

$$\mathcal{C}_p \triangleq \{\hat{\mathbf{x}}^q \in \mathcal{D} : d(\hat{\mathbf{x}}^p, \hat{\mathbf{x}}^q) < \epsilon\}. \quad (32)$$

In this paper, we use  $d(\hat{\mathbf{x}}^p, \hat{\mathbf{x}}^q) = \|\hat{\mathbf{x}}^p - \hat{\mathbf{x}}^q\|$ . The main idea of DBSCAN is that clusters are not characterized by their variance (as in K-means), but by *density-reachability*. In particular,  $\hat{\mathbf{x}}^q$  is directly density-reachable from  $\hat{\mathbf{x}}^p$ , if and only if  $|\mathcal{C}_p| \geq N_{\min}$  and  $\hat{\mathbf{x}}^q \in \mathcal{C}_p$ . In addition,  $\hat{\mathbf{x}}^p$  is density-reachable from  $\hat{\mathbf{x}}^q$ , if there exists an ordered sequence of points  $\hat{\mathbf{x}}^{p_0} = \hat{\mathbf{x}}^q, \hat{\mathbf{x}}^{p_1}, \dots, \hat{\mathbf{x}}^{p_n} = \hat{\mathbf{x}}^p$  in  $\mathcal{D}$ , where  $\hat{\mathbf{x}}^{p_{i+1}}$  is directly density-reachable from  $\hat{\mathbf{x}}^{p_i}$ . If the sphere of  $\hat{\mathbf{x}}^p$  contains at least  $N_{\min}$  points, that is,  $|\mathcal{C}_p| \geq N_{\min}$ , the point  $\hat{\mathbf{x}}^p$  will be a core point of a cluster. The border points can have a smaller size  $\epsilon$ -neighbourhood, but can be density-reachable from the core points.

Given the parameters  $\epsilon$  and  $N_{\min}$ , we can discover a cluster in a two-step approach: (i) choose an arbitrary point from the database satisfying the core point conditions as a seed; (ii) find all points

that are density-reachable from the seeds [64]. We follow these two steps, that is, to identify a core point, to find a cluster with the core point, and to find all clusters. The algorithms are provided in Algorithm 1. Note that Algorithm 2 is used in Algorithm 1 to find all points in each cluster. The output of the clustering is a partitioning of  $\mathcal{Z}$  (or  $\mathcal{D}$ ) into clusters  $\mathcal{P}_l$  and a set of noise points  $\mathcal{P}_N$ .

---

**Algorithm 1** DBSCAN for Clustering
 

---

**Input:** Points  $\{\hat{x}^0, \hat{x}^1, \dots, \hat{x}^{P_k-1}\}$ , threshold  $\epsilon$ , and  $N_{\min}$ ;

**Output:** All clusters and the associated points.

```

1: Compute pair-wise Euclidean distance for all points, store them in D, and label all points unvisited;
2:  $l \leftarrow 0$  and  $p \leftarrow 0$ ; // cluster index  $l$  and point index  $p$ 
3: while  $p < P$  do
4:   if Point  $p$  is unvisited then
5:     Label point  $p$  as visited;
6:     Find all neighbours within  $\epsilon$  distance of  $p$ ;
7:     if  $\epsilon$ -neighbourhood is less than  $N_{\min}$  then
8:       Classify the point  $p$  as noise;
9:     else
10:       $l \leftarrow l + 1$ ; // new cluster
11:      Discover all points in cluster  $l$  by using Algorithm 2.
12:    end if
13:  end if
14: end while

```

---



---

**Algorithm 2** Find All Points in Cluster  $l$ 


---

**Input:** Cluster index  $l$ , point index  $p$  and its  $\epsilon$ -neighbourhood  $\mathcal{C}_p$ ;

**Output:** The associated points in cluster  $l$ .

```

1: Classify  $p$  into cluster  $l$ ;
2: Store all indexes of the points in  $\mathcal{C}_p$  in vector c;
3:  $k \leftarrow 1$ ;
4: while  $k \leq |\mathbf{c}|$  do
5:    $j \leftarrow \mathbf{c}_k$ ;
6:   if Point  $\hat{x}^j$  is unvisited then
7:     Label point  $\hat{x}^j$  as visited;
8:     Find all neighbours within  $\epsilon$  distance of  $\hat{x}^j$ , that is,  $\mathcal{C}_j$ ;
9:     if  $|\mathcal{C}_j| \geq N_{\min}$  then
10:      Add all indexes in  $\mathcal{C}_j$  to c;
11:    end if
12:  end if
13:  if  $\hat{x}^j$  is not classified to any cluster then
14:    Classify  $\hat{x}^j$  in cluster  $l$ ;
15:  end if
16:   $k \leftarrow k + 1$ ;
17: end while

```

---

### 5.2.3. Phase 3: Extract Isolated Specular Paths and Outliers Using Channel Gain

The DBSCAN clustering has two drawbacks in our application:

- The tensor ESPRIT channel estimator from Section 4 can generate estimates, whose 3D points (as obtained in Section 5.2.1) are still on or near the corresponding surfaces, but are far away from the cluster centers. Hence, they are informative for the SLAM algorithm, but are part of  $\mathcal{P}_N$ , so they are not clustered correctly. We have observed that the channel gains of these paths are very small.
- The LOS path and specular paths from smooth surfaces are not part of any cluster, as such landmarks have one or few associated paths. We have observed that the channel gains of these paths are very large (approximately following the path loss models from Section 2.3).

To solve these problems, we use the estimated channel gains. We partition  $\mathcal{P}_N$  into two sets according to the channel gains.

$$\mathcal{P}_{N,L} = \{q : c(\hat{\tau}^p - \hat{B})|\hat{g}^p| \frac{4\pi}{\lambda} \leq \delta\} \quad (33)$$

$$\mathcal{P}_{N,H} = \{q : c(\hat{\tau}^p - \hat{B})|\hat{g}^p| \frac{4\pi}{\lambda} > \delta\}, \quad (34)$$

where  $\delta$  is a predefined threshold,  $\hat{B}$  is the estimated clock bias, and  $c$  is the speed of light. The scaling with  $c(\hat{\tau}^p - \hat{B})\frac{4\pi}{\lambda}$  is added to compensate for the path loss. For each of the high-gain paths in  $\mathcal{P}_{N,H}$ , we create a new cluster. For each of the low-gain paths in  $p \in \mathcal{P}_{N,L}$ , we find the nearest cluster  $l^*$  as

$$l^* = \arg \min_l \min_{q \in \mathcal{P}_l} d(\hat{x}^p, \hat{x}^q). \quad (35)$$

and add  $p$  to cluster  $\mathcal{P}_{l^*}$ .

The overall computations in channel parameter clustering consist of three parts, that is, the computations in each phase. Given the number of paths  $\hat{P}$ , the transition from 5D to 3D mapping requires  $\mathcal{O}(\hat{P})$  computations. According to Reference [64], the complexity of the DBSCAN algorithm without the use of index structure for acceleration is  $\mathcal{O}(\hat{P}^2)$ . The last phase requires  $\mathcal{O}(\hat{P})$  computations. Therefore, the overall complexity is  $\mathcal{O}(\hat{P}^2)$ .

## 6. Likelihood Function for SLAM

After clustering, the measurements  $\mathcal{Z}_k$  are grouped into clusters  $\mathcal{Z}_k = \{\mathcal{Z}_k^0, \mathcal{Z}_k^1, \dots, \mathcal{Z}_k^{\hat{I}_k-1}\}$ , using Algorithm 1, where  $\hat{I}_k$  is the number of estimated clusters. The clustered measurements will be the input into the SLAM filter during the update process, via an appropriate likelihood function. For brevity, we omit the time index  $k$  again. We work with the compact representation of the landmark state from Section 2.2. Hence, our goal is to determine  $\ell(\mathcal{Z}^i | \mathbf{x}_{LM}, \mathbf{s}, m)$ , where  $\mathbf{x}_{LM}$  is the landmark position (i.e., the BS location or a virtual anchor location) and  $m \in \{\text{BS}, \text{SM}, \text{MR}, \text{VR}\}$  is the landmark state. This likelihood function will influence the probabilities of different associations between landmarks and clusters, and it will also influence the particle weight and the distribution of the landmark given a state  $\mathbf{s}$  and the association.

### 6.1. Background

In previous 5G SLAM works, Reference [21] assumed there is at most one measurement from each source; the channel gain is not used, and the inter-path interference is not considered. The measurements are only generated by known multivariate Gaussian distributions, then the likelihood function falls into the likelihood of the signal measurements, which follows the Gaussian format. According to the PMBM formalism, multiple measurements per source should follow a Poisson or Bernoulli distribution. Under the standard point model, the likelihood function corresponds to a Bernoulli distribution, with [65]

$$\ell(\mathcal{Z}^i | \mathbf{x}_{LM}, \mathbf{s}, m) = \begin{cases} 1 - p_D & \mathcal{Z}^i = \emptyset \\ p_D p(\mathbf{z}^{i,0} | \mathbf{x}_{LM}, \mathbf{s}, m) & \mathcal{Z}^i = \{\mathbf{z}^{i,0}\} \\ 0 & \text{else,} \end{cases} \quad (36)$$

where  $p_D$  is the detection probability. Under the extended target model, the likelihood function corresponds to a Poisson point process (PPP) ([25], Equation (5)):

$$\ell(\mathcal{Z}^i | \mathbf{x}_{LM}, \mathbf{s}, m) = e^{-\gamma_m} \gamma_m^{|\mathcal{Z}^i|} \prod_{l=0}^{|\mathcal{Z}^i|-1} p(\mathbf{z}^{i,l} | \mathbf{x}_{LM}, \mathbf{s}, m), \quad (37)$$



where  $\gamma_m \geq 0$  is the Poisson rate for surfaces of type  $m$ . For both (36) and (37), it has been proven that the PMBM density is a conjugate prior.

## 6.2. Likelihood Function

In this section, we describe the proposed likelihood function. We here assume that the measurements within each cluster  $\mathcal{Z}^i$  are independent, as diffuse points are generated independently. We also assume the number of paths  $|\mathcal{Z}^i|$  only depends on the source type  $m$ . It is important to note that for measurements from diffuse paths, a measurement is a function of a *random incidence point on the surface*. In order to express the measurements as a function of the compact state, we propose the following approach.

We first separate  $\mathcal{Z}^i$  into two parts, the path with the shortest delay  $\{z^{i,0}\}$ , which is the closest path to the specular component, and any remaining paths  $\tilde{\mathcal{Z}}^i = \{z^{i,1}, z^{i,2}, \dots, z^{i,|\mathcal{Z}^i|-1}\}$ , which we view as diffuse multipath. We associate the deterministic incidence point  $x^{i,0}$  with the assumed specular component, which can be derived by (31) using  $x_{BS}$  and  $x_{LM}$ . Therefore, we can write  $p(z^{i,0}|x_{LM}, s, m) = p(z^{i,0}|x^{i,0}, s, m)$ , for  $z^{i,0}$ , which is in the desired form (16). Hence

$$\ell(\mathcal{Z}^i|x_{LM}, s, m) = \begin{cases} 1 - p_D & \mathcal{Z}^i = \emptyset \\ p_D p(z^{i,0}|x_{LM}, s, m) \ell(\tilde{\mathcal{Z}}^i|x_{LM}, s, m) & \text{else.} \end{cases} \quad (38)$$

In other words, the object may be miss-detected with probability  $1 - p_D$ . If the object is detected, then the likelihood contains a contribution from the first arriving path (the specular path, which is directly related to the VA location or BS location) and from the set of all remaining paths.

Note that as a special case for  $m = BS$  there can be at most one associated measurement, so that

$$\ell(\mathcal{Z}^i|x_{LM}, s, m = BS) = \begin{cases} 1 - p_D & \mathcal{Z}^i = \emptyset \\ p_D p(z^{i,0}|x_{LM}, s, m) & \mathcal{Z}^i = \{z^{i,0}\} \\ 0 & \text{else.} \end{cases} \quad (39)$$

which is in the desired form. Hence, we can focus on the diffuse paths. As the measurements contain both a specular component and diffuse paths, the likelihood function does not follow the standard models (36)–(37). Nevertheless, we use the likelihood in the PMBM filter without any proof of conjugacy or optimality.

### 6.2.1. Likelihood for Diffuse Paths

We recall from Section 2.3 that diffuse multipath originates from random points  $x^{i,l}$  on the surface and from Section 4 that the estimates points will differ from those from Section 2.3, due to the finite resolution of the receiver. Hence, it is *fundamentally impossible* to estimate the original random incidence points  $x^{i,l}$  from the observations  $\tilde{\mathcal{Z}}^i$ . To avoid estimating  $x^{i,l}$ , we propose estimating an *artificial incidence point*  $\tilde{x}^{i,l}$  from  $z^{i,l>0}$  as the projection of  $\hat{x}^{i,l}$  (obtained using the method in Section 5.2.1) onto the surface. The projection point  $\tilde{x}^{i,l}$  can be derived by

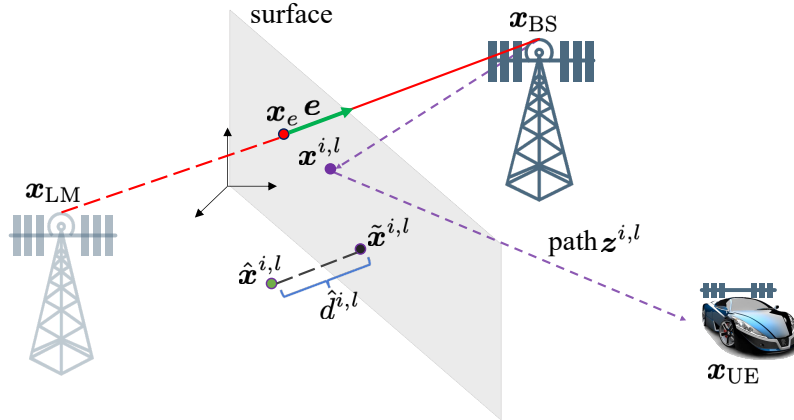
$$\tilde{x}^{i,l} = \hat{x}^{i,l} + (x_e - \hat{x}^{i,l})^\top e e, \quad (40)$$

where  $x_e = (x_{BS} + x_{LM})/2$  is a point on the surface, and  $e = (x_{BS} - x_{LM})/\|x_{BS} - x_{LM}\|$  is a normal to the surface. Then,  $p(z^{i,l}|x_{LM}, s, m)$  can be expressed in the desired form  $p(z^{i,l}|\tilde{x}^{i,l}, s, m)$ , where  $\tilde{x}^{i,l}$  is the artificial incidence point that gave rise to measurement  $z^{i,l}$ . The likelihood can be further simplified by noticing the error component of  $\tilde{x}^{i,l}$ , which is only due to the the projection distance of  $\hat{x}^{i,l}$  to the

surface. We therefore use this distance directly as a compressed measurement to replace the delay and angles

$$\hat{d}^{i,l}(z^{i,l}) = \mathbf{e}^T(\mathbf{x}_e - \hat{\mathbf{x}}^{i,l}). \quad (41)$$

Figure 5 shows the principle of calculating  $\hat{d}^{i,l}(z^{i,l})$ .



**Figure 5.** The principle of finding  $\hat{\mathbf{x}}^{i,l}$  and calculating  $\hat{d}^{i,l}$  using  $\mathbf{z}^{i,l}$ ,  $\mathbf{x}_{\text{LM}}$  and  $\mathbf{x}_{\text{BS}}$ .

We thus have a general model for the likelihood function associated with the diffuse paths, where the cardinality distribution is arbitrary and (cf. (Reference [18], Equation (1))

$$\ell(\tilde{\mathcal{Z}}^i | \mathbf{x}_{\text{LM}}, \mathbf{s}, m) = p(|\tilde{\mathcal{Z}}^i| | m) |\tilde{\mathcal{Z}}^i|! \prod_{l=1}^{|\tilde{\mathcal{Z}}^i|} p([\hat{d}^{i,l}(z^{i,l}), \hat{g}^{i,l}] | \mathbf{x}_{\text{LM}}, \mathbf{s}, m), \quad (42)$$

in which all constituent distributions can be obtained from the simulation of a channel estimator or provided directly in closed-form by a channel estimator. In (42), the first term  $p(|\tilde{\mathcal{Z}}^i| | m)$  represents the cardinality distribution,  $|\tilde{\mathcal{Z}}^i|!$  is a normalization constant to ensure that  $\ell(\tilde{\mathcal{Z}}^i | \mathbf{x}_{\text{LM}}, \mathbf{s}, m)$  integrates to one over  $\tilde{\mathcal{Z}}^i$ . The element in  $\tilde{\mathcal{Z}}^i$  are represented in compressed format through  $[\hat{d}^{i,l}(z^{i,l}), \hat{g}^{i,l}]$  and are modeled as independent and identically distributed.

### 6.2.2. Clustering Errors and Marginal Likelihood Function

Due to clustering errors and channel estimator errors, it is possible that the specular path is not associated with the set of the specular paths. Then the shortest path in  $\mathcal{Z}^i$  may be a diffuse path, leading to low likelihood  $p(z^{i,0} | \mathbf{x}_{\text{LM}}, \mathbf{s}, m)$ , so that the SLAM method is likely to view the source of this measurement cluster as a new landmark, which is not desirable. In order to solve this problem, we introduce  $p_m \in (0, 1]$ , which is the probability that the first path in a cluster is in fact the specular path. This probability can be determined by observation of the overall clustering performance. By marginalizing over the event of missing the specular path in a cluster, the likelihood function can be written as

$$\begin{aligned} p(z^{i,0} | \mathbf{x}_{\text{LM}}, \mathbf{s}, m) &= p_m p(z^{i,0} | \mathbf{x}_{\text{LM}}, \mathbf{s}, m) \\ &+ (1 - p_m) \frac{p(|\tilde{\mathcal{Z}}^i| + 1 | m)(|\tilde{\mathcal{Z}}^i| + 1)}{p(|\tilde{\mathcal{Z}}^i| | m)} p([\hat{d}^{i,0}(z^{i,0}), \hat{g}_k^{i,0}] | \mathbf{x}_{\text{LM}}, \mathbf{s}, m), \end{aligned} \quad (43)$$

where the factor  $p(|\tilde{\mathcal{Z}}^i| + 1 | m)(|\tilde{\mathcal{Z}}^i| + 1) / p(|\tilde{\mathcal{Z}}^i| | m)$  is due to there being an additional specular path, which was not present in the cluster. The complexity of the likelihood function is  $\mathcal{O}(|\tilde{\mathcal{Z}}^i|)$ , as we consider  $(|\tilde{\mathcal{Z}}^i| + 1)$  paths.

## 7. Results

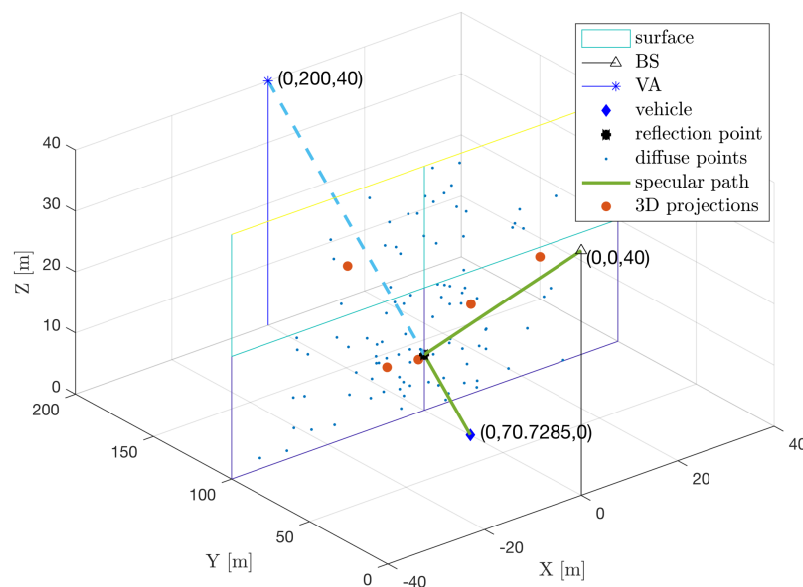
In the implementation, all the codes are written in MATLAB, and the simulations and experiments are run on a MacBook Pro (15-inch, 2019) with a 2.6 GHz 6-Core Intel Core i7 processor and 16 Gb memory.

### 7.1. Simulation Parameters

We consider a scenario with a single vehicle with the initial state  $[70.7285\text{ m}, 0\text{ m}, 0\text{ m}, \pi/2\text{ rad}, 22.22\text{ m/s}, \pi/10\text{ rad/s}, 1\text{ }\mu\text{s}]^T$ , a BS located at  $[0\text{ m}, 0\text{ m}, 40\text{ m}]^T$ , a SM, with the VA at  $[200\text{ m}, 0\text{ m}, 40\text{ m}]^T$ , two MRs, with VAs at  $[0\text{ m}, \pm 200\text{ m}, 40\text{ m}]^T$  and a VR, with the VA at  $[0\text{ m}, -200\text{ m}, 40\text{ m}]^T$ . The BS and the receiver are both equipped with a URA with  $8 \times 8$  antennas. The vehicle moves around the BS with a known constant turn rate. Every time step, the BS sends  $10 \times 64$  OFDM symbols to the vehicle with 200 subcarriers using the transmit power of 5.05 W at carrier frequency of 28 GHz. The subcarrier spacing is 0.5 MHz. The transmitted signals are disturbed by the noise with power spectral density of  $4.0049 \times 10^{-9}$  mW/Hz. The transition function of the movement, the initial vehicle state, the landmark locations, the process noise, the initial prior, the survival probability, the detection probability, the birth rate, the clutter intensity, and pruning thresholds are the same as in Reference [23].

### 7.2. Channel Estimation Results

In this section, we show the results of the tensor ESPRIT channel estimator for a MR surface, which generated one specular path and 100 diffuse paths. To visualize the channel estimator output, we transform the estimated TOAs, AOAs, AODs into 3D points. The result is shown in Figure 6, where the original diffuse points are marked with blue dots and the estimated diffuse points with red discs. We observe that the channel estimator returns only 5 paths, much less than the original 101 paths, due to the finite resolution of the receiver. All projected points are close to the surface and there is a projected point very close to the deterministic reflection point. This is because the specular path of the MR has larger power than the other diffuse paths, so it is less affected by inter-path interference.



**Figure 6.** The 3D projection of channel estimation results of 100 diffuse path and a specular path from an MR.

### 7.3. Clustering Performance Evaluation

To evaluate the clustering performance, we generate 40 snapshots of the channel with known landmarks. We then obtain the channel estimates with 5D geometric parameters, and transform them into 3D point positions. The DBSCAN clustering algorithm is applied to each snapshot, with  $\epsilon = 30$  m and  $N_{\min} = 2$ . Then we use the estimated channel gain to modify the results, with  $\delta = 0.25$ . As a comparison, we also compare with other clustering methods, including the well-known K-means, gap statistics ('GS') [66], and affinity propagation ('AP') [63] on the same data set. For the K-means, we assume that the number of clusters,  $K$ , is known, while for the remaining methods this priori knowledge is not needed.

To evaluate the clustering performance, we consider the following performance metrics, which use as ground-truth hand-labeled data: the clustering accuracy (CA), which is the percentage of correctly clustered points, and the impurity (IMP), which is the percentage of points clustered into a wrong cluster

$$CA(\mathcal{G}, \mathcal{C}) = \frac{\max_{\gamma \in \Gamma} (\hat{P} + \sum_{(i,j) \in \gamma} (|\mathcal{G}_i \cap \mathcal{C}_j| - |\mathcal{C}_j|))}{\hat{P}} \quad (44)$$

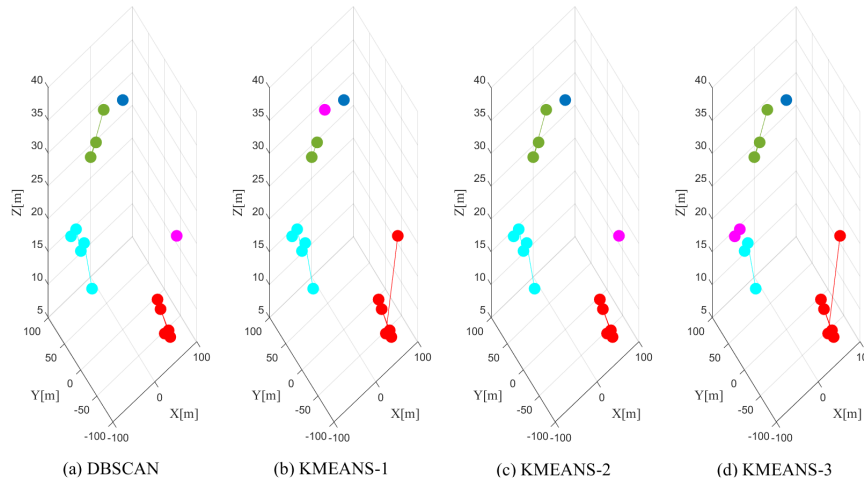
$$IMP(\mathcal{G}, \mathcal{C}) = \frac{\sum_{k \in \{1, \dots, N\}} (|\mathcal{C}_k| - \max_{m \in \{1, \dots, N\}} |\mathcal{G}_m \cap \mathcal{C}_k|)}{\hat{P}}, \quad (45)$$

where  $\Gamma$  is the set of all possible assignments;  $\mathcal{G} = \{\mathcal{G}_1, \dots, \mathcal{G}_N\}$  is the ground truth clusters;  $\mathcal{C} = \{\mathcal{C}_1, \dots, \mathcal{C}_{\hat{N}}\}$  is the cluster results provided by the clustering algorithm;  $\hat{P}$  is the number of all target points. The performance is presented in Table 2. The proposed clustering method based on DBSCAN provides the best clustering performance, with an average clustering accuracy over 99%. The modified DBSCAN goes through all noise points using channel gain and refines the results, so it performs slightly better. K-means achieves an average accuracy of 94.63%. GS-based clustering method has an average accuracy of only 68.64% since the estimated value of the number of clusters can be erroneous. The AP clustering method has an average accuracy of 84.35% since it still based on space partition and the estimated value of the number of clusters can be erroneous either. In terms of impurity, the last three algorithms cluster measurements from different sources together when errors occur, but for DBSCANs this never happens, so that all points in each cluster correspond to one landmark. By observing the clustering results of modified DBSCAN, losing the shortest path from the same landmark happens rarely, and we have estimated  $p_m \approx 1/200$ .

Though the K-means algorithm shows good average clustering performance, it suffers from a few limitations as indicated in Section 5.1. This can be demonstrated in Figure 7, where we present the clustering results with DBSCAN and K-means with the data from the 33rd snapshot. We use different colors to indicate the different clusters obtained from the corresponding clustering methods. The proposed DBSCAN can correctly cluster the measurements into the right clusters shown in Figure 7a; however, K-means shows unstable clustering performance indicated in Figure 7b–d. This is because the clustering performance of K-means is highly dependent on the initialization of the partitions. In addition, we observe from Figure 7b,d that in K-means, the data points from the same cluster can be divided into separated clusters, and the points from different clusters can be clustered in a single cluster. All these mismatches will cause performance degradation in SLAM, which will be further demonstrated in Section 7.5.

**Table 2.** Performance comparison of different clustering methods.

Clustering Method	Clustering Accuracy	Impurity
Modified DBSCAN	99.61%	0
DBSCAN	99.07%	0
K-means	94.63%	5.37%
Gap Statistics (GS)	68.64%	27.19%
Affinity Propagation (AP)	84.35%	12.54%

**Figure 7.** The clustering performance of DBSCAN and K-means.

#### 7.4. Estimated Likelihoods

Given the channel estimation and the clustering performance, we now describe the obtained likelihood functions. For simplicity, we consider all dimensions independent and report the observed distributions of the number of paths and error of the specular path gain, delay and 4 angles. For diffuse paths, we report the distribution of the distances  $\hat{d}^{i,l}$  and the channel gains. All these distributions are based on gathered channel estimation and clustering results in various vehicle locations. We use the setting in Section 7.1, and run the channel estimator many times, collect all channel estimation results and study statistics of collected data. Since the channel gains are strongly affected by path loss, (4)–(7) reduce the impact of the distance dependence by considering

$$\check{g}^{i,l} [\text{dB}] = 20 \log_{10} \left( |\hat{g}^{i,l}| \|\mathbf{x}_{\text{UE}} - \mathbf{x}_{\text{LM}}\| \frac{4\pi}{\lambda} \right). \quad (46)$$

We first focus on the case  $m = \text{MR}$  as an example, and study  $p(|\tilde{Z}^i| | \text{MR})$ ,  $p(c\tau^{i,0} | \mathbf{x}_{\text{LM}}, \mathbf{s}, \text{MR})$ ,  $p(\phi_{\text{el}}^{i,0} | \mathbf{x}_{\text{LM}}, \mathbf{s}, \text{MR})$ ,  $p(\check{g}^{i,0} | \mathbf{x}_{\text{LM}}, \mathbf{s}, \text{MR})$ ,  $p(\hat{d}^{i,l} | \mathbf{x}_{\text{LM}}, \mathbf{s}, \text{MR})$  and  $p(\check{g}^{i,l>0} | \mathbf{x}_{\text{LM}}, \mathbf{s}, \text{MR})$  in Figure 8. Figure 8a shows the histogram of the number of paths, as well as a Poisson approximation. We observe that there are always a (assumed) specular path and 1 to 5 diffuse paths, and the existing of more than one diffuse path follows a Poisson distribution. Figure 8b shows the histogram of the delay error of the first estimated path  $\tau^{i,0}$  (i.e., subtracted with the delay of the specular path and multiple by the speed of light  $c$ ) as well as a Gaussian fit. We observe that there is a shift of delay of 0.17 m, which is caused by inter-path interference. Figure 8c shows the histogram of the elevation AOD error of first estimated path  $\phi_{\text{el}}^{i,0}$  (i.e., subtracted with the angle of the specular path) as well as a Gaussian fit. We observe that inter-path interference does not influence the angle so much. Figure 8d shows the histogram and Gaussian fit of the distances  $\hat{d}^{i,l}$ ,  $l > 0$ . As the inter-path interference leads to a positive delay shift, the estimated diffuse points are more likely to be behind the surface. Finally, Figure 8e,f shows the histogram of  $\check{g}^{i,0}$  and  $\check{g}^{i,l>0}$  as well as Gaussian fits. We observe that the specular path is stronger than

diffuse paths, as the mean of  $\tilde{g}^{i,0}$  is larger, and diffuse paths are influenced by inter-path interference more seriously, as the distribution of  $\tilde{g}^{i,l>0}$  is more spread.

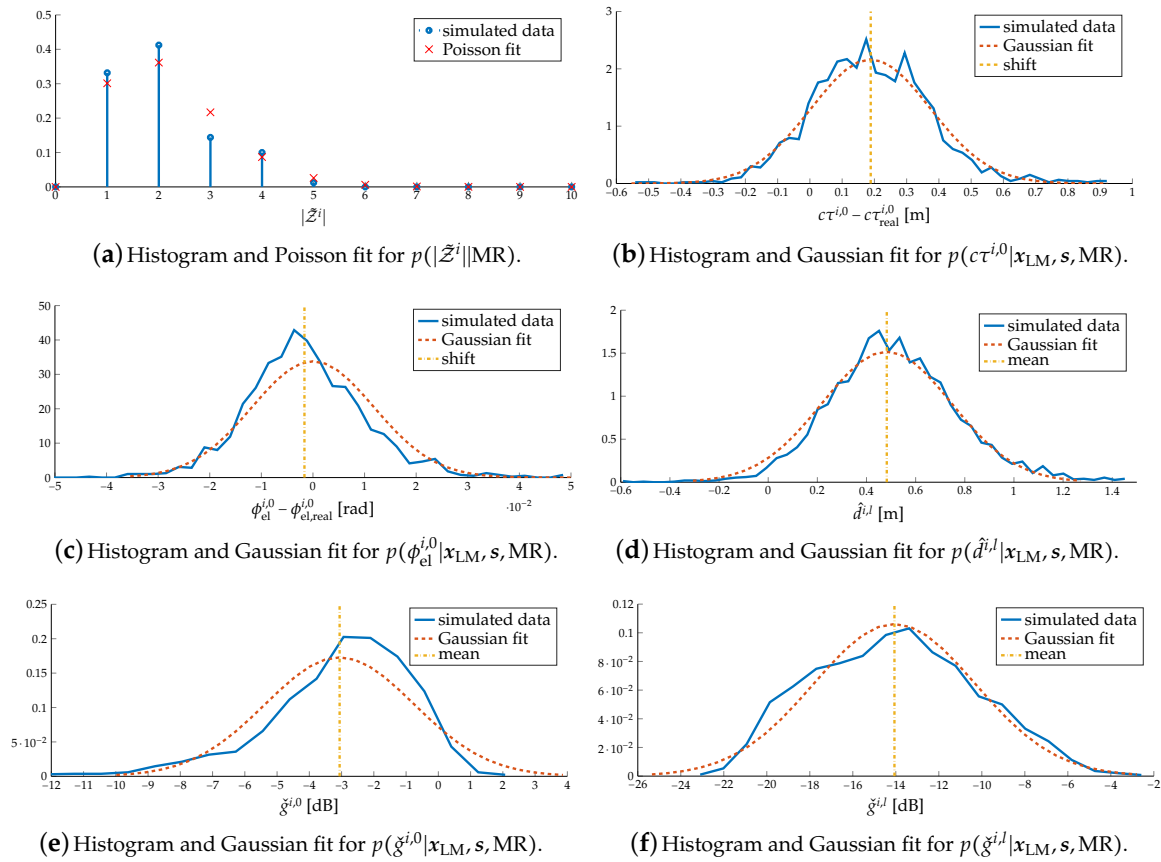


Figure 8. Some components of the likelihood for MR.

Table 3 provides a complete overview of the likelihood function for all landmark types. Based on all components of the likelihood functions for all landmarks, we have the following observations: for both cases  $m = BS$  and  $m = SM$ , there is always one path presents and all the angles, delays and  $\tilde{g}^{i,0}$  follow Gaussian distributions. For case  $m = MR$ , apart from the assumed specular path, there are 3 to 12 more paths present, and angles, delays,  $\tilde{g}^{i,l}$  and  $\hat{d}^{i,l}$  also follow Gaussian distributions. We also find that the smoother surface has a higher gain for the specular path and lower gain for the diffuse path.

However, since the clustering algorithm is not perfect, it is possible that the cluster associated with a landmark may have fewer or more paths than the ground truth cluster. For example, the cluster associated with the MR may only have a specular path, due to the clustering error. We therefore set those possibility none zero in  $p(|Z^i||m)$  manually, in order to increase the tolerance of the wrong clustering, as shown in Table 3.

**Table 3.** Likelihood function for 5G SLAM for different types of landmarks.

Type $m$	$p( \tilde{\mathbf{z}}^i  m)$	$p([\hat{d}^{i,l}; \hat{g}^{i,l}] x_{LM}, s, m)$
BS	$( \tilde{\mathbf{z}}^i  - 1)^* \sim 0.9\Delta_0 + 0.1 \text{ Geo}(0.9)$	N/A
SM	$ \tilde{\mathbf{z}}^i  \sim 0.9\Delta_0 + 0.1 \text{ Geo}(0.9)$	N/A
MR	$( \tilde{\mathbf{z}}^i  - 1) \sim 0.1\Delta_{-1} + 0.9 \text{ Poi}(1.5)$	$\mathcal{N}([\hat{d}^{i,l}; \hat{g}^{i,l}]; [0.48; -14.1], \text{diag}([0.35, 3.8]^2))$
VR	$( \tilde{\mathbf{z}}^i  - 3) \sim 0.1 \sum_{n=1}^3 0.5^{n-1} \Delta_{-n} + 0.825 \text{ Geo}(0.27)$	$\mathcal{N}([\hat{d}^{i,l}; \hat{g}^{i,l}]; [0.48; -9.4], \text{diag}([0.35, 3.8]^2))$
	$p(\tilde{\mathbf{z}}^{i,0} x_{LM}, s, m)$	
BS	$\mathcal{N}(\tilde{\mathbf{z}}^{i,0}; [\mathbf{h}_{VA}; 0], \text{diag}([0.003, 0.0001 \times \mathbf{1}_4, 0.005]^2))$	
SM	$\mathcal{N}(\tilde{\mathbf{z}}^{i,0}; [\mathbf{h}_{VA}; -1.93], \text{diag}([0.01, 0.002 \times \mathbf{1}_4, 0.02]^2))$	
MR	$\mathcal{N}(\tilde{\mathbf{z}}^{i,0}; [\mathbf{h}_{VA}; -3.1] + [0.17, \mathbf{0}_{1 \times 5}]^T, \text{diag}([0.19, 0.012 \times \mathbf{1}_4, 2.3]^2))$	
VR	$\mathcal{N}(\tilde{\mathbf{z}}^{i,0}; [\mathbf{h}_{VA}; -6.3] + [0.8, \mathbf{0}_{1 \times 5}]^T, \text{diag}([0.5, 0.05 \times \mathbf{1}_4, 3.8]^2))$	

\*  $|Z| - M \sim p\Delta_n + (1 - p) \text{ Geo}(\gamma)$  means that  $|Z|$  is equal to  $M$  plus a random variable drawn from a discrete mixture distribution with two components: one with mass  $p$  at value  $n$  and one that is a scaled Geometric distribution (with scaling  $1 - p$ ).  $\text{Poi}(\gamma)$  represents a Poisson distribution.  $\tilde{\mathbf{z}}^{i,0}$  denotes  $[c\tau^{i,0}; \theta^{i,0}; \phi^{i,0}; \hat{g}^{i,0}[\text{dB}]]$ .  $\mathbf{h}_{BS}$ ,  $\mathbf{h}_{VA}$  are the geometric relations  $\mathbf{h}(x_{BS}, s)$  and  $\mathbf{h}(x_{VA}, s)$ , which can be found in Appendix A.

## 7.5. SLAM Performance Evaluation

### 7.5.1. Mapping Performance

Firstly, we study the performance of the proposed 5G SLAM scheme in mapping, conditioned on the true vehicle state. We use the generalized optimal subpattern assignment (GOSPA) distance [67] as the metric to evaluate the mapping result. The GOSPA metric considers the estimation error, which is the error between the estimated positions with the real positions of landmarks, the miss-detection error, which is the error of miss-detecting the existing landmarks, and the false alarm error, which is the error of detecting non-existing landmarks. The GOSPA between the estimate landmark position set  $\hat{\mathcal{X}} = \{\hat{\mathbf{x}}_i\}_{i=1}^{\hat{N}}$  and the real landmark position set  $\mathcal{X} = \{\mathbf{x}_j\}_{j=1}^N$  is defined as

$$d_{\text{GOSPA}}(\hat{\mathcal{X}}, \mathcal{X}) = \min_{\gamma \in \Gamma} \left( \sum_{(i,j) \in \gamma} d^{q_c}(\hat{\mathbf{x}}_i, \mathbf{x}_j) + \frac{q_c^{q_p}}{q_a} (N_{\text{miss}}^\gamma + N_{\text{false}}^\gamma) \right)^{\frac{1}{q_p}} \quad (47)$$

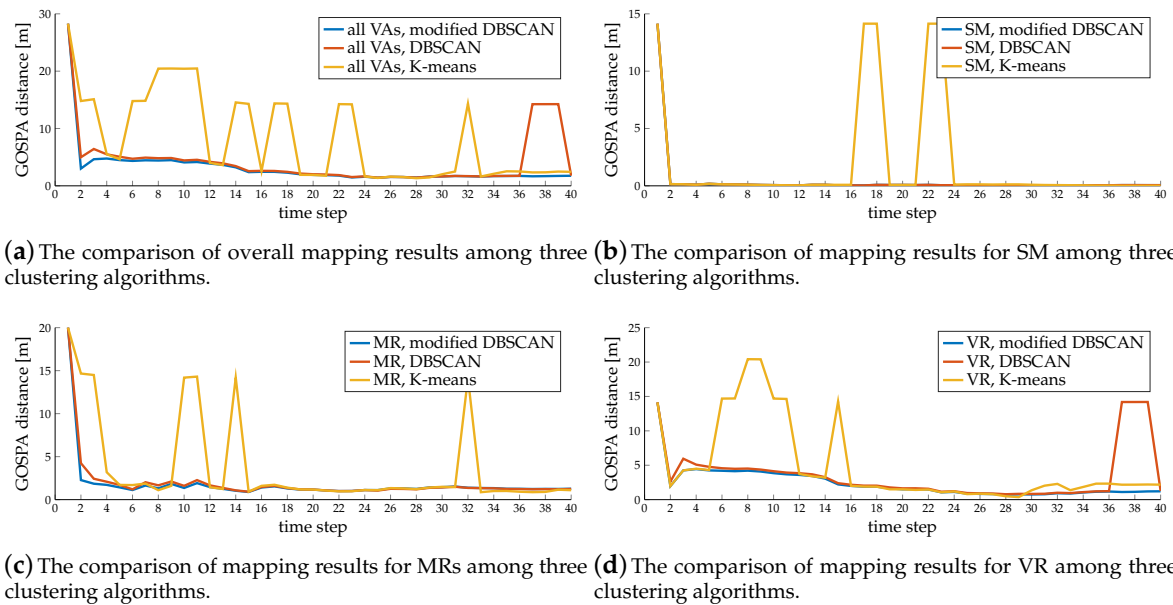
where  $\Gamma$  is the set of all possible assignments;  $N_{\text{miss}}^\gamma$  is the number of miss detection of assignment  $\gamma$ ;  $N_{\text{false}}^\gamma$  is the number of false alarm of assignment  $\gamma$ . The cardinality penalty factor  $q_a$ , the cut-off distance  $q_c$ , and the exponent factor  $q_p$  are set as 20, 2, and 2, respectively.

We make two comparisons: (i) we compare different clustering methods (SLAM filter using modified DBSCAN which is the proposed method, DBSCAN without channel gains, and K-means); (ii) we compare SLAM filters with different numbers of paths (using all paths in every signal cluster based on the proposed likelihood function, which is the proposed method; using all paths but without using channel gains [24]; using the single (specular) path in every cluster based on the proposed likelihood function; using the single (specular) path in every cluster but without using channel gains, which is the PMBM extension of Reference [23]).

Figure 9 shows the outcome of the first comparison. We observe that the SLAM filter with K-means clustering performs poorly, even though the number of clusters is already known. Although the K-means performance in Table 2 seems good, grouping measurements from different sources leads to significant errors. The reason is that K-means only divides the space linearly and cannot handle complicated shapes. It often groups projected points from different sources into the same clusters; for example, at time step 17, it groups the measurements from BS and SM together. Modified DBSCAN and simple DBSCAN perform much better, as the blue and red lines are much lower, since both of them never group measurements from different sources together. The tensor ESPRIT channel estimator may



create measurements that are far away from the specular path, and have low gains but still contain location information. Treating those outliers individually as single-element clusters in DBSCAN loses some information, and also forces the SLAM filter to create a new landmark, when similar outliers appear in successive time steps. This is apparent in the peak at time step 37–39 in Figure 9a, since there are some clutters pointing to a similar region between time step 36–39. With the help of the channel gains, the modified DBSCAN groups those measurements into the right clusters, which helps in mapping. That is the reason why the blue line is lower than the red line, for example at time step 2 in Figure 9c and time step 3 in Figure 9d.



**Figure 9.** Comparison of GOSPA results among three clustering algorithms.

Figure 10 shows the comparison of the GOSPA results among 4 different settings (with or without using diffuse multipath, with or without using channel gain). We observe that SLAM filter has better performance when using diffuse multipath, as the solid lines are lower than the dashed lines, and it can also help the system to identify the landmark types, as there are no sharp peaks in solid lines, which is because more paths provide more information. We also observe the gain can improve the performance, especially in helping the SLAM filter to recognize the surface type, as the blue lines are lower than the red lines. Without information from channel gains, the SLAM filter has trouble to distinguish SM and MR when only using the specular path, as indicated by the sharp peaks of red dashed lines in Figure 10b,c and the red solid line dropping more slowly than the blue solid line in Figure 10b. This is because the distributions of delays and angles of specular paths from SM and MR are very close to each other, and some disturbance may cause a wrong identification. The information from channel gains will alleviate this problem. However, the channel gain is usually disturbed by inter-path interference, which increases with the roughness of the surface. Therefore, the gain is more informative for paths from SM and MR, but less informative for paths from VR, so the blue dashed lines are much lower than the red dashed lines in Figure 10b,c, but only slightly lower than the red dashed line in Figure 10d.

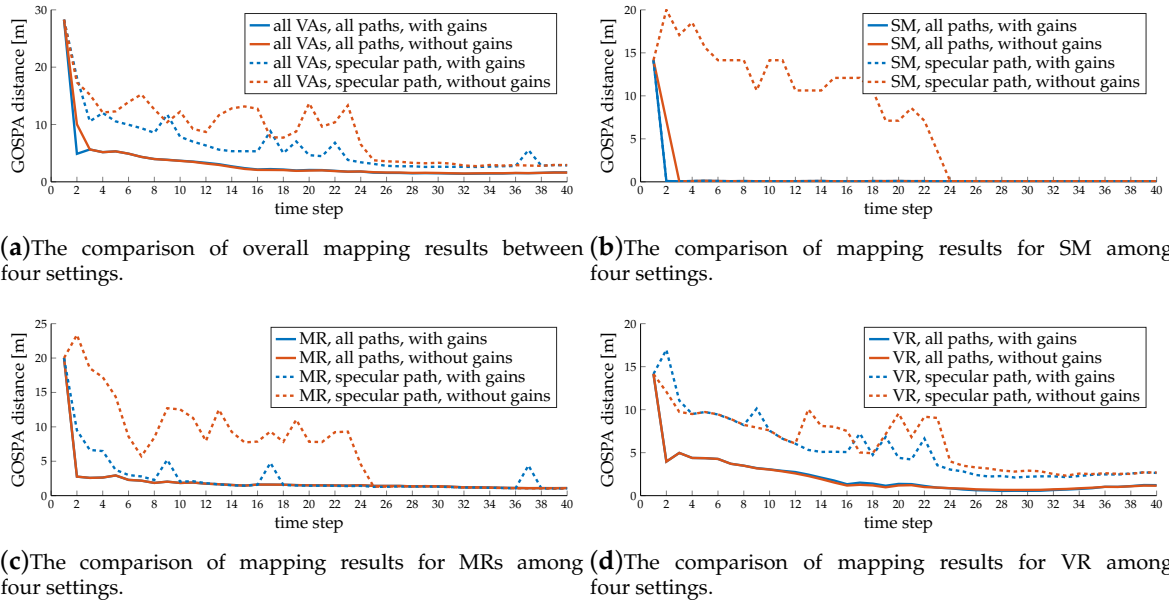


Figure 10. Comparison of GOSPA results among four different settings.

### 7.5.2. Localization Performance

Next, we study the performance of the proposed 5G SLAM scheme in vehicle state estimation and compare the estimation results among the SLAM filter with different settings (with or without using diffuse multipath, with or without using channel gain). The landmarks are unknown and are mapped with tracking the vehicle state. We add  $[0.9 \text{ m}, 0.9 \text{ m}, 0 \text{ m}, 4.5 \text{ deg}, 0 \text{ m/s}, 0 \text{ rad/s}, 3 \text{ ns}]^T$  bias to the initial state, and the initial covariance is set as  $\text{diag}([0.3 \text{ m}, 0.3 \text{ m}, 0 \text{ m}, 0.1 \text{ rad}, 0 \text{ m/s}, 0 \text{ rad/s}, 1 \text{ ns}]^2)$ . We use 2000 particles to represent the vehicle state, and obtain the mean absolute error (MAE) between the real vehicle state and the estimated vehicle state after time step 5 until the simulation ends at time step 40

$$\text{MAE} = \frac{\sum_{k=6}^{40} |s_k - \hat{s}_k|}{35}, \quad (48)$$

as shown in Figure 11. The clock bias is multiplied with the speed of light  $c$  for evaluation. Overall, when using diffuse multipath, the SLAM filter has better performance in positioning, as MAEs are lower. The reason is that diffuse multipath provides more information than the single path. Using channel gain also improves the positioning performance, especially very helpful to estimate the clock bias, since gain contains information about propagation distance directly without being affected by the clock bias.

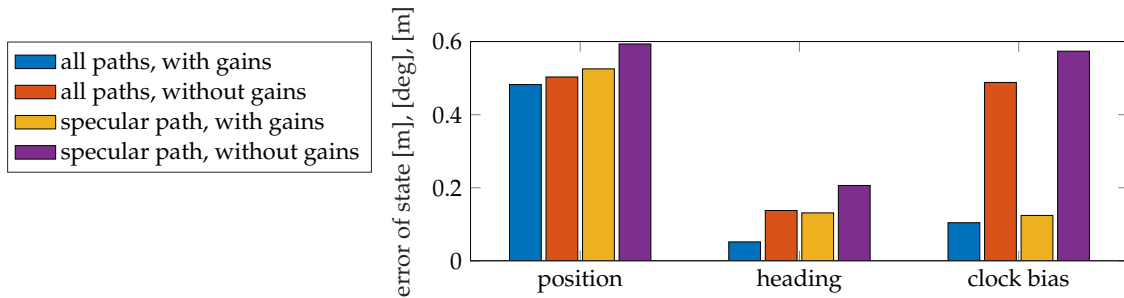


Figure 11. Comparison of vehicle state estimation performance, considering the utilization of channel gains and different sets of estimated propagation paths.

### 7.5.3. Complexity Evaluation

We also evaluate the complexity of the proposed framework by measuring the execution time for each phase—channel estimation, clustering, mapping, SLAM (per particle), as shown in Figure 12. We measured a runtime per time step of 8.12 s for the channel estimation. The reason for taking a long time is that the ESPRIT channel estimator is based on a high dimensional tensor decomposition with high complexity. We measured the modified DBSCAN takes 0.21 ms per snapshot. For mapping, it takes 110 ms per time step. Although the proposed SLAM filter considers all possible data associations, as there are not too many landmarks under the simulation scenario, the number of all possible data associations is not large. However, the runtime will increase with more data associations. For SLAM, the proposed SLAM filter takes 149 ms per each particle (298.52 s in total) per time step. As 2000 particles are used and we need to fuse maps of all particles, the complexity is much higher than mapping, where the true vehicle state is given.

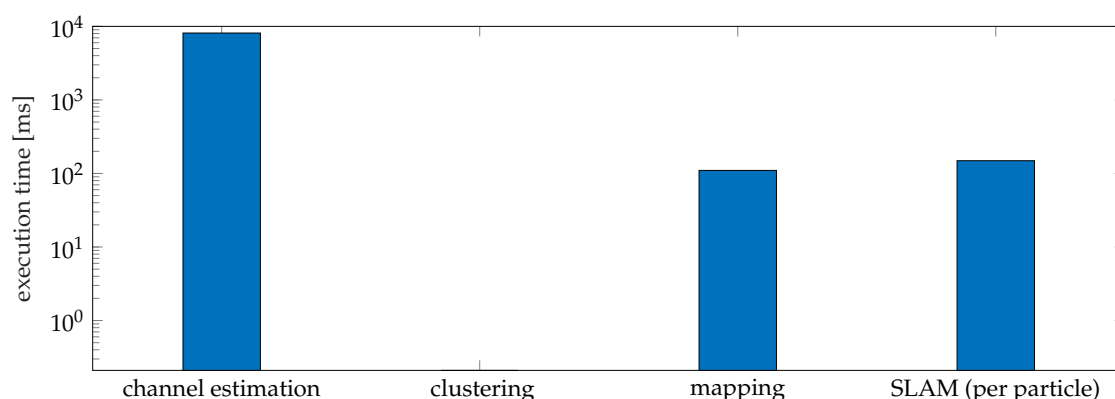


Figure 12. The execution times for each phase of the proposed framework.

## 8. Conclusions

In this paper, we have treated the 5G SLAM problem from an end-to-end perspective, including downlink data transmission, channel estimation, clustering, and the SLAM filter. In the 5G SLAM problem, we aim to localize and synchronize a user while mapping the propagation environment, with the help of downlink signals from a single base station. We have proposed a novel method to cluster the MPCs by projecting the high-dimensional data into 3D points and then cluster the points based on the DBSCAN algorithm, which we augmented to account for the channel gains. We have also proposed a novel likelihood function in the 5G SLAM filter, which accounts for both the specular path as well as the diffuse multipath components.

Our results show that the ESPRIT channel estimator can estimate the channel parameters of both specular and diffuse multipath, and that the proposed system can directly use the raw un-clustered channel estimation results by applying the proposed clustering algorithms. With the help of the novel likelihood function, the proposed scheme can accurately estimate the number of landmarks, their types (i.e., roughness), and positions, and the channel gain is helpful in clustering and mapping and positing. The results also confirm that the proposed method can handle mapping and vehicle state estimation simultaneously, and highlight the benefit of considering both specular and diffuse multipath. In addition, the channel gains turn out to be highly informative for synchronizing the user to the base station.

The proposed framework has two computational bottlenecks—the ESPRIT channel estimator and the particle filter used in the PMBM SLAM. In order to enable real-time execution, there is a need for faster solutions for both the channel estimation and the SLAM filter. The solutions could be either in the form of new algorithms, or by offloading the computation to more powerful edge computing systems, where edge servers can provide high-performance computing capability closer to end users [68–70].

**Author Contributions:** Y.G. has led the paper, drafted the main parts, derived the likelihood functions, and performed the simulations. F.W. has led the work related to channel estimation and relation between channel parameters and the location parameters. H.K. derived the SLAM methods. M.Z. provided expertise on clustering and channel estimation. F.J. has supported the work on clustering and proof-reading of the paper. S.K., L.S., and H.W. were supervisors of the research and acquired the funding. They have also proof-read and edited the paper. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by Knut and Alice Wallenberg Foundation, the Vinnova 5GPOS project under grant 2019-03085, by the Swedish Research Council under grant 2018-0370, by the MSIP, Korea, under the ITRC support program IITP-2020-2017-0-01637, and by the IITP grant funded by the Korea government No. 2019-0-01325.

**Acknowledgments:** The authors are grateful to Kamran Keykhosravi, Zohair Abu Shaban and M. Furkan Keskin for their comments on this manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Geometric Relations

The geometric relations are as follows. For LOS,  $h(\mathbf{x}_{BS}, \mathbf{s})$  returns

- TOA:  $\tau = \|\mathbf{x}_{UE} - \mathbf{x}_{BS}\| / c + B$ .
- AOD pair:  $\phi_{az} = \arctan(y_{UE} - y_{BS}, x_{UE} - x_{BS})$ ,  $\phi_{el} = \arcsin((z_{UE} - z_{BS}) / \|\mathbf{x}_{UE} - \mathbf{x}_{BS}\|)$ .
- AOA pair:  $\theta_{az} = \arctan(y_{BS} - y_{UE}, x_{BS} - x_{UE}) + \pi - \omega$ ,  $\theta_{el} = \arcsin((z_{BS} - z_{UE}) / \|\mathbf{x}_{UE} - \mathbf{x}_{BS}\|)$ .

For the case of a specular component,  $h(\mathbf{x}_{VA}, \mathbf{s})$  returns

- TOA:  $\tau = \|\mathbf{x}_{VA} - \mathbf{x}_{UE}\| / c + B$ .
- AOD pair:  $\phi_{az} = \arctan(y_{sp} - y_{BS}, x_{sp} - x_{BS})$ ,  $\phi_{el} = \arcsin((z_{sp} - z_{BS}) / \|\mathbf{x}_{sp} - \mathbf{x}_{BS}\|)$ .
- AOA pair:  $\theta_{az} = \arctan(y_{VA} - y_{UE}, x_{VA} - x_{UE}) - \omega$ ,  $\theta_{el} = \arcsin((z_{VA} - z_{UE}) / \|\mathbf{x}_{VA} - \mathbf{x}_{BS}\|)$ .

where  $\mathbf{x}_{sp}$  is the incident point, which can be calculated by (31). For the case of a diffuse component with diffuse point  $\mathbf{x}_{sc}$ ,  $h(\mathbf{x}_{sc}, \mathbf{s})$  returns

- TOA:  $\tau = (\|\mathbf{x}_{BS} - \mathbf{x}_{sc}\| + \|\mathbf{x}_{sc} - \mathbf{x}_{UE}\|) / c + B$ .
- AOD pair:  $\phi_{az} = \arctan(y_{sc} - y_{BS}, x_{sc} - x_{BS})$ ,  $\phi_{el} = \arcsin((z_{sc} - z_{BS}) / \|\mathbf{x}_{sc} - \mathbf{x}_{BS}\|)$ .
- AOA pair:  $\theta_{az} = \arctan(y_{sc} - y_{UE}, x_{sc} - x_{UE}) - \omega$ ,  $\theta_{el} = \arcsin((z_{sc} - z_{UE}) / \|\mathbf{x}_{sc} - \mathbf{x}_{BS}\|)$ .

## Appendix B. PMBM SLAM Filter

### Appendix B.1. Representation of PMBM Density

We adopt the finite set statistic (FISST) for the multi-target tracking [71], and we introduce RFS  $\mathcal{X}$  for a multi-target state, which consists of the undetected targets  $\mathcal{X}_U$  that have never been detected before and detected targets  $\mathcal{X}_D$  that have been detected at least once. The PMBM density is denoted by [65]

$$f(\mathcal{X}) = \sum_{\mathcal{X}_U \uplus \mathcal{X}_D = \mathcal{X}} f_P(\mathcal{X}_U) f_{MBM}(\mathcal{X}_D), \quad (A1)$$

where  $\uplus$  denotes the union of mutually disjoint sets,  $f_P(\cdot)$  denotes a Poisson point process (PPP) density, and  $f_{MBM}(\cdot)$  denotes an multi-Bernoulli mixture (MBM) density. The PPP density  $f_P(\mathcal{X}_U)$  is defined as

$$f_P(\mathcal{X}_U) = e^{-\int \lambda(x) dx} \prod_{j=1}^n \lambda(\mathbf{x}^j), \quad (A2)$$

where  $\lambda(\cdot)$  is the intensity function called as probability hypotheses density [72],  $n$  is the cardinality of  $\mathcal{X}_U$ . The MBM density  $f_{MBM}(\mathcal{X}_D)$  is represented as

$$f_{MBM}(\mathcal{X}_D) \propto \sum_h \sum_{\mathcal{X}^1 \uplus \dots \uplus \mathcal{X}^n = \mathcal{X}_D} \prod_{j=1}^n l^{h,j} f_B^{h,j}(\mathcal{X}^j), \quad (A3)$$

where  $h$  indicates the hypotheses index [18],  $n$  is the number of potentially detected objects, and  $f_B^{h,j}(\cdot)$  denotes the Bernoulli density of the object  $j$  under the global hypothesis  $h$ , and  $l^{h,j}$  is its weight. Each Bernoulli density follows

$$f_B^{h,j}(\mathcal{X}^j) = \begin{cases} 1 - r^{h,j} & \mathcal{X}^j = \emptyset \\ r^{h,j} f^{h,j}(\mathbf{x}) & \mathcal{X}^j = \{\mathbf{x}\} \\ 0 & \text{otherwise} \end{cases} \quad (\text{A4})$$

where  $r^{h,j}$  is the existence probability and  $f^{h,j}(\cdot)$  is the state density. Then, (A2) can be parameterized by  $\lambda(\mathbf{x})$ , and (A3) can be parameterized by  $\{l^{h,j}, r^{h,j}, f^{h,j}(\mathbf{x})\}_{j \in \mathbb{I}^h, h \in \mathbb{I}}$ , where  $\mathbb{I}$  is the index set.

### Appendix B.2. Implementation of PMBM SLAM Filter

The PMBM SLAM filter goes through prediction and update steps of the Bayesian filtering recursion as in [23]. For implementation, we adopt the Rao-Blackwellized approach; particles represent the user state, and PMBM densities conditioned on each particle indicate the landmarks. At the beginning of time  $k$ ,  $\{s_{0:k-1}^n, \omega_{k-1|k-1}^n\}_{n=1}^N$  is given, where  $\omega_{k-1|k-1}^n \geq 0$  and  $\sum_n \omega_{k-1|k-1}^n = 1$ , and we have a PMBM density with PPP parameter  $\lambda_{k-1|k-1}^n(\mathbf{x})$  and MBM parameters  $\{l_{k-1|k-1}^{n,h,j}, r_{k-1|k-1}^{n,h,j}, f_{k-1|k-1}^{n,h,j}(\mathbf{x})\}_{j \in \mathbb{I}^{n,h}, h \in \mathbb{I}^n}$  for all particle  $n$ . For notation simplicity, we omit the particle index  $n$  in the PPP and MBM parameters.

1. *User Prediction:* Using (1), the user particle  $s_{k-1|k-1}^n$  is predicted as  $s_{k|k-1}^n = v(s_{k-1|k-1}^n) + q_k^n$ , where  $q_k^n \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$  and  $\omega_{k|k-1}^n = \omega_{k-1|k-1}^n$ .
2. *Map Prediction:* Since the targets are static, the PPP parameter is predicted as  $\lambda_{k|k-1}(\mathbf{x}) = p_S \lambda_{k-1|k-1}(\mathbf{x}) + \lambda_{B,k}(\mathbf{x})$ , where  $p_S$  is the survival probability,  $\lambda_{B,k}(\mathbf{x})$  is the birth intensity. For the MBM components,  $l_{k|k-1}^{h,j} = l_{k-1|k-1}^{h,j}$ ,  $r_{k|k-1}^{h,j} = p_S r_{k-1|k-1}^{h,j}$ , and  $f_{k|k-1}^{h,j}(\mathbf{x}) = f_{k-1|k-1}^{h,j}(\mathbf{x})$ .
3. *Map Update:* The map update is divided into the following four cases [65]
  - (a) *Missed detections for undetected objects:* The undetected objects remain as the undetected objects, and thus is given by

$$\lambda_{k|k}(\mathbf{x}) = (1 - p_D) \lambda_{k|k-1}(\mathbf{x}), \quad (\text{A5})$$

where  $p_D$  indicates the detection probability.

- (b) *Detections for the first time:* Using the grouped measurement  $\mathcal{Z}_k^i$  and the PPP parameter  $\lambda_{k|k-1}(\mathbf{x})$ , we newly generate the MBM parameters as

$$r_{U,k|k}^i = \rho_{U,k|k-1}^i(\mathcal{Z}_k^i) / (c(\mathcal{Z}_k^i) + \rho_{U,k|k-1}^i(\mathcal{Z}_k^i)), \quad (\text{A6})$$

$$f_{U,k|k}^i(\mathbf{x}) = \ell(\mathcal{Z}_k^i | \mathbf{x}, \mathbf{s}_{k|k-1}) \lambda_{k|k-1}(\mathbf{x}) / \rho_{U,k|k-1}^i(\mathcal{Z}_k^i), \quad (\text{A7})$$

$$l_{U,k|k}^i = c(\mathcal{Z}_k^i) + \rho_{U,k|k-1}^i(\mathcal{Z}_k^i), \quad (\text{A8})$$

$$\rho_{U,k|k-1}^i(\mathcal{Z}_k^i) = \int \ell(\mathcal{Z}_k^i | \mathbf{x}, \mathbf{s}_{k|k-1}) \lambda_{k|k-1}(\mathbf{x}) d\mathbf{x}, \quad (\text{A9})$$

where  $c(\mathcal{Z}_k^i) = \delta(|\mathcal{Z}_k^i| = 1) c(\mathbf{z}_k^{i,0})$ .

- (c) *Missed detections for the previously detected objects:* The detected objects also remain as the detected objects, and then the MBM parameters have no measurement update

$$r_{k|k}^{h,j,0} = (1 - p_D) r_{k|k-1}^{h,j} / (1 - p_D r_{k|k-1}^{h,j}), \quad (\text{A10})$$

$$f_{k|k}^{h,j,0}(\mathbf{x}) = f_{k|k-1}^{h,j}(\mathbf{x}), \quad (\text{A11})$$

$$l_{k|k}^{h,j,0} = l_{k|k-1}^{h,j} (1 - r_{k|k-1}^{h,j} p_D). \quad (\text{A12})$$

- (d) *Detections for the previously detected objects:* Using  $\mathcal{Z}_k^i$ , the MBM parameters are computed as

$$r_{k|k}^{h,j,i} = 1, \quad (\text{A13})$$

$$f_{k|k}^{h,j,i}(\mathbf{x}) = \ell(\mathcal{Z}_k^i | \mathbf{x}, \mathbf{s}_{k|k-1}) f_{k|k-1}^{h,j}(\mathbf{x}) / \rho_{k|k-1}^{h,j,i}(\mathcal{Z}_k^i), \quad (\text{A14})$$

$$l_{k|k}^{h,j,i} = l_{k|k-1}^{h,j} r_{k|k-1}^{h,j} \rho_{k|k-1}^{h,j,i}(\mathcal{Z}_k^i), \quad (\text{A15})$$

$$\rho_{k|k-1}^{h,j,i}(\mathcal{Z}_k^i) = \int \ell(\mathcal{Z}_k^i | \mathbf{x}, \mathbf{s}_{k|k-1}) f_{k|k-1}^{h,j}(\mathbf{x}) d\mathbf{x}. \quad (\text{A16})$$

To reduce the exponentially increasing data association, we adopt the pruning and merging process [65] with the Murty's algorithm [73], where the cost matrix is generated by  $l_{U,k|k}^i$ ,  $l_{k|k}^{h,j,0}$  and  $l_{k|k}^{h,j,i}$  calculated in (A8), (A12) and (A15).

4. *User Update:* Each particle weight is updated as

$$\omega_{k|k}^n \propto \omega_{k|k}^n \sum_h l_{k|k}^{n,h}, \quad (\text{A17})$$

where  $l_{k|k}^{n,h}$  is the weight for the updated global hypothesis  $h$  and particle  $n$ . Then, the user state is estimated by  $\hat{\mathbf{s}}_{k|k} = \sum_n \omega_{k|k}^n \mathbf{s}_{k|k}^n$ .

Given the number of particles  $N$ , the user state prediction requires  $\mathcal{O}(N)$  computations. For each particle, the map prediction requires  $\mathcal{O}(N_B)$  computations (if there are  $N_B$  unique Bernoulli components for the given particle), and the map update process requires  $\mathcal{O}(N_B \hat{I})$  computations for updating Bernoulli components and  $\mathcal{O}(N_B N_G \sum_{t=0}^{\hat{I}} C_t^t A_{N_B}^{\hat{I}-t})$  for updating data associations and their weights (if there are  $\hat{I}$  measurement clusters and  $N_G$  global hypotheses in total for the given particle), where  $A$  represents the permutation operation and  $C$  represents the combination operation. The complexity of the particle weight updating is  $\mathcal{O}(N_G)$  for each particle, and the complexity of state estimation is  $\mathcal{O}(N)$ . Note: each particle may have different  $N_G$  and  $N_B$ .

### Appendix B.3. Map Fusion

We need to create a map using these particles. Since there is a different MBM for each particle, and an MB represents a possible map, a best map needs to be computed. Firstly, we need to compute an MB for each particle. The MBM with parameters  $\{l_{k|k}^{n,h,j}, r_{k|k}^{n,h,j}, f_{k|k}^{n,h,j}(\mathbf{x})\}_{j \in \mathbb{I}_k^{n,h}} \}_{h \in \mathbb{I}_k^n}$  can be written as  $\{q_{k|k}^{n,h}, \{r_{k|k}^{n,h,j}, f_{k|k}^{n,h,j}(\mathbf{x})\}_{j \in \mathbb{I}_k^{n,h}}\}_{h \in \mathbb{I}_k^n}$ , where  $q_{k|k}^{n,h}$  is the probability of the global hypothesis  $h$ , which follows

$$q_{k|k}^{n,h} \propto \prod_{j \in \mathbb{I}_k^{n,h}} l_{k|k}^{n,h,j}. \quad (\text{A18})$$

The existence probability of each new MB component will be [18]

$$\bar{r}_{k|k}^{n,j} = \sum_{h \in \mathbb{I}_k^n} q_{k|k}^{n,h} r_{k|k}^{n,h,j}, \quad (\text{A19})$$

and the density becomes

$$\tilde{f}_{k|k}^{n,j}(\mathbf{x}) = \frac{\sum_{h \in \mathbb{I}_k^n} q_{k|k}^{n,h} \tilde{f}_{k|k}^{n,h,j}(\mathbf{x})}{\tilde{r}_{k|k}^{n,j}}. \quad (\text{A20})$$

Then, we have an MB with parameters  $\{\tilde{r}_{k|k}^{n,j}, \tilde{f}_{k|k}^{n,j}(\mathbf{x})\}_{j \in \mathbb{I}_k^n}$  for particle  $n$ . When considering all particles, we would create a larger MB by concatenating all MBs of different particles together, with parameters  $\{\{\tilde{r}_{k|k}^{n,j}, \omega_{k|k}^n, \tilde{f}_{k|k}^{n,j}(\mathbf{x})\}_{j \in \mathbb{I}_k^n}\}_{n \in \mathbb{I}_k}$ . Finally, we merge components with Euclidean distances smaller than the given threshold in the new MB, and the final MB is the best map. The map fusion requires  $\mathcal{O}(NN_{GNB} + (NN_B)^2)$  computations, when considering the merging process.

## References

1. Narayanan, A.; Ramadan, E.; Carpenter, J.; Liu, Q.; Liu, Y.; Qian, F.; Zhang, Z.L. A first look at commercial 5G performance on smartphones. In Proceedings of the Web Conference, Taipei, Taiwan, 20–24 April 2020; pp. 894–905.
2. del Peral-Rosado, J.A.; Raulefs, R.; López-Salcedo, J.A.; Seco-Granados, G. Survey of cellular mobile radio localization methods: From 1G to 5G. *IEEE Commun. Surv. Tutorials* **2017**, *20*, 1124–1148. [CrossRef]
3. 3GPP TR 38.855 V16.0.0; Study on NR Positioning Support; Technical report; Available online: <https://www.3gpp.org> (accessed on 18 August 2020).
4. del Peral-Rosado, J.; Granados, G.; Raulefs, R.; Leitinger, E.; Grebien, S.; Wilding, T.; Dardari, D.; Lohan, E.; Wymeersch, H.; Floch, J.; et al. White Paper on New Localization Methods for 5G Wireless Systems and the Internet-of-Things. In COST Action CA15104. Available online: <http://www.iracon.org/wp-content/uploads/2018/03/IRACON-WP2.pdf> (accessed on 18 August 2020).
5. Witrisal, K.; Meissner, P.; Leitinger, E.; Shen, Y.; Gustafson, C.; Tufvesson, F.; Haneda, K.; Dardari, D.; Molisch, A.F.; Conti, A.; et al. High-accuracy localization for assisted living: 5G systems will turn multipath channels from foe to friend. *IEEE Signal Process. Mag.* **2016**, *33*, 59–70. [CrossRef]
6. Talvitie, J.; Valkama, M.; Destino, G.; Wymeersch, H. Novel algorithms for high-accuracy joint position and orientation estimation in 5G mmWave systems. In Proceedings of the IEEE Globecom Workshops (GC Wkshps), Singapore, 4–8 December 2017.
7. Mendrzik, R.; Wymeersch, H.; Bauch, G. Joint localization and mapping through millimeter wave MIMO in 5G systems. In Proceedings of the IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, UAE, 9–13 December 2018.
8. Li, X.; Leitinger, E.; Oskarsson, M.; Åström, K.; Tufvesson, F. Massive MIMO-based localization and mapping exploiting phase information of multipath components. *IEEE Trans. Wirel. Commun.* **2019**, *18*, 4254–4267. [CrossRef]
9. Fascista, A.; Coluccia, A.; Wymeersch, H.; Seco-Granados, G. Millimeter-wave downlink positioning with a single-antenna receiver. *IEEE Trans. Wirel. Commun.* **2019**, *18*, 4479–4490. [CrossRef]
10. Fascista, A.; Coluccia, A.; Wymeersch, H.; Seco-Granados, G. Downlink single-snapshot localization and mapping with a single-antenna receiver. *arXiv* **2020**, arXiv:2007.14679.
11. Richter, A. Estimation of Radio Channel Parameters: Models and Algorithms. Ph.D. Thesis, Ilmenau University of Technology, Ilmenau, Germany, 2005.
12. Alkhateeb, A.; El Ayach, O.; Leus, G.; Heath, R.W. Channel estimation and hybrid precoding for millimeter wave cellular systems. *IEEE J. Sel. Top. Signal Process.* **2014**, *8*, 831–846. [CrossRef]
13. Venugopal, K.; Alkhateeb, A.; Pircic, N.G.; Heath, R.W. Channel estimation for hybrid architecture-based wideband millimeter wave systems. *IEEE J. Sel. Areas Commun.* **2017**, *35*, 1996–2009. [CrossRef]
14. Gershman, A.B.; Rübsamen, M.; Pesavento, M. One- and two-dimensional direction-of-arrival estimation: An overview of search-free techniques. *Signal Process.* **2010**, *90*, 1338–1349. [CrossRef]
15. Wen, F.; Garcia, N.; Kulmer, J.; Witrisal, K.; Wymeersch, H. Tensor decomposition based beamspace ESPRIT for millimeter wave MIMO channel estimation. In Proceedings of the IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, UAE, 9–13 December 2018.



16. Zhou, Z.; Fang, J.; Yang, L.; Li, H.; Chen, Z.; Blum, R.S. Low-rank tensor decomposition-aided channel estimation for millimeter wave MIMO-OFDM systems. *IEEE J. Sel. Areas Commun.* **2017**, *35*, 1524–1538. [\[CrossRef\]](#)
17. Mullane, J.; Vo, B.N.; Adams, M.D.; Vo, B.T. A random-finite-set approach to Bayesian SLAM. *IEEE Trans. Robot.* **2011**, *27*, 268–282. [\[CrossRef\]](#)
18. Williams, J.L. Marginal multi-Bernoulli filters: RFS derivation of MHT, JIPDA, and association-based MeMBeR. *IEEE Trans. Aerosp. Electron. Syst.* **2015**, *51*, 1664–1687. [\[CrossRef\]](#)
19. Leitinger, E.; Meyer, F.; Hlawatsch, F.; Witrissal, K.; Tufvesson, F.; Win, M.Z. A belief propagation algorithm for multipath-based SLAM. *IEEE Trans. Wirel. Commun.* **2019**, *18*, 5613–5629. [\[CrossRef\]](#)
20. Gentner, C.; Jost, T.; Wang, W.; Zhang, S.; Dammann, A.; Fiebig, U.C. Multipath assisted positioning with simultaneous localization and mapping. *IEEE Trans. Wirel. Commun.* **2016**, *15*, 6104–6117. [\[CrossRef\]](#)
21. Kim, H.; Wymeersch, H.; Garcia, N.; Seco-Granados, G.; Kim, S. 5G mmWave vehicular tracking. In Proceedings of the 52nd IEEE Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 28–31 October 2018; pp. 541–547.
22. Meyer, F.; Braca, P.; Willett, P.; Hlawatsch, F. A scalable algorithm for tracking an unknown number of targets using multiple sensors. *IEEE Trans. Signal Process.* **2017**, *65*, 3478–3493. [\[CrossRef\]](#)
23. Kim, H.; Granström, K.; Gao, L.; Battistelli, G.; Kim, S.; Wymeersch, H. 5G mmWave cooperative positioning and mapping using multi-model PHD filter and map fusion. *IEEE Trans. Wirel. Commun.* **2020**. [\[CrossRef\]](#)
24. Ge, Y.; Kim, H.; Wen, F.; Svensson, L.; Kim, S.; Wymeersch, H. Exploiting diffuse multipath in 5G SLAM. *arXiv* **2020**, arXiv:2006.15603.
25. Granström, K.; Baum, M.; Reuter, S. Extended object tracking: Introduction, overview and applications. *arXiv* **2016**, arXiv:1604.00970.
26. Aubry, A.; De Maio, A.; Foglia, G.; Orlando, D. Diffuse multipath exploitation for adaptive radar detection. *IEEE Trans. Signal Process.* **2015**, *63*, 1268–1281. [\[CrossRef\]](#)
27. Setlur, P.; Negishi, T.; Devroye, N.; Erricolo, D. Multipath exploitation in non-LOS urban synthetic aperture radar. *IEEE J. Sel. Top. Signal Process.* **2013**, *8*, 137–152. [\[CrossRef\]](#)
28. Wen, F.; Kulmer, J.; Witrissal, K.; Wymeersch, H. 5G positioning and mapping with diffuse multipath. *arXiv* **2019**, arXiv:1912.08697.
29. Fatemi, M.; Granström, K.; Svensson, L.; Ruiz, F.J.; Hammarstrand, L. Poisson multi-Bernoulli mapping using Gibbs sampling. *IEEE Trans. Signal Process.* **2017**, *65*, 2814–2827. [\[CrossRef\]](#)
30. Granström, K.; Svensson, L.; Reuter, S.; Xia, Y.; Fatemi, M. Likelihood-based data association for extended object tracking using sampling methods. *IEEE Trans. Intell. Veh.* **2017**, *3*, 30–45. [\[CrossRef\]](#)
31. Granström, K.; Lundquist, C.; Orguner, O. Extended target tracking using a Gaussian-mixture PHD filter. *IEEE Trans. Aerosp. Electron. Syst.* **2012**, *48*, 3268–3286. [\[CrossRef\]](#)
32. Czink, N.; Cera, P.; Salo, J.; Bonek, E.; Nuutinen, J.P.; Ylitalo, J. A framework for automatic clustering of parametric MIMO channel data including path powers. In Proceedings of the IEEE Vehicular Technology Conference, Melbourne, Australia, 7–10 May 2006.
33. 5GPPP Association. 5G Automotive Vision, 5GPPP; White Paper. Available online: <https://5g-ppp.eu> (accessed on 18 August 2020).
34. Ahvar, E.; Daneshgar-Moghaddam, N.; Ortiz, A.M.; Lee, G.M.; Crespi, N. On analyzing user location discovery methods in smart homes: A taxonomy and survey. *J. Netw. Comput. Appl.* **2016**, *76*, 75–86. [\[CrossRef\]](#)
35. Yaeli, A.; Bak, P.; Feigenblat, G.; Nadler, S.; Roitman, H.; Saadoun, G.; Ship, H.J.; Cohen, D.; Fuchs, O.; Ofek-Koifman, S.; et al. Understanding customer behavior using indoor location analysis and visualization. *IBM J. Res. Dev.* **2014**, *58*, 3–1. [\[CrossRef\]](#)
36. Hwangbo, H.; Kim, Y.S.; Cha, K.J. Use of the smart store for persuasive marketing and immersive customer experiences: A case study of Korean apparel enterprise. *Mob. Inform. Syst.* **2017**, *2017*, 4738340.
37. Taranto, R.D.; Muppirisetty, S.; Raulefs, R.; Slock, D.; Svensson, T.; Wymeersch, H. Location-aware communications for 5G networks. *IEEE Signal Process. Mag.* **2014**, *31*, 102–112. [\[CrossRef\]](#)
38. Guidi, F.; Guerra, A.; Dardari, D. Personal mobile radars with millimeter-wave massive arrays for indoor mapping. *IEEE Trans. Mob. Comput.* **2015**, *15*, 1471–1484. [\[CrossRef\]](#)

39. Bourdoux, A.; Barreto, A.N.; van Liempd, B.; de Lima, C.; Dardari, D.; Belot, D.; Lohan, E.S.; Seco-Granados, G.; Srieddeen, H.; Wymeersch, H.; et al. 6G white paper on localization and sensing. *arXiv* **2020**, arXiv:2006.01779.
40. Kim, H.; Granström, K.; Kim, S.; Wymeersch, H. Low-complexity 5G SLAM with CKF-PHD filter. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*; IEEE: Piscataway, NJ, USA, 2020; pp. 5220–5224.
41. Ferre, R.M.; Seco-Granados, G.; Lohan, E.S. *Positioning Reference Signal Design for Positioning Via 5G*; National Committee for Radiology in Finland: Tampere, Finland, 2019.
42. 3GPP TS 38.211 NR V15.8.0; Physical Channels and Modulation; Technical Report. Available online: <https://www.3gpp.org> (accessed on 18 August 2020).
43. Kulmer, J. High-Accuracy Positioning Exploiting Multipath for Reducing the Infrastructure. Ph.D. Thesis, Graz University of Technology, Graz, Austria, 2019.
44. Heath, R.W.; Gonzalez-Prelcic, N.; Rangan, S.; Roh, W.; Sayeed, A.M. An overview of signal processing techniques for millimeter wave MIMO systems. *IEEE J. Sel. Top. Signal Process.* **2016**, *10*, 436–453. [[CrossRef](#)]
45. Vukmirović, N.; Erić, M.; Janjić, M.; Djurić, P.M. Direct wideband coherent localization by distributed antenna arrays. *Sensors* **2019**, *19*, 4582. [[CrossRef](#)] [[PubMed](#)]
46. Wang, P.; Pajovic, M.; Koike-Akino, T.; Sun, H.; Orlik, P.V. Fingerprinting-based indoor localization with commercial mmWave WiFi-part II: Spatial beam SNRs. In Proceedings of the IEEE Global Communications Conference (GLOBECOM), Waikoloa, HI, USA, 9–13 December 2019.
47. Guerra, A.; Guidi, F.; Dardari, D. Single-anchor localization and orientation performance limits using massive arrays: MIMO vs. beamforming. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 5241–5255. [[CrossRef](#)]
48. Liu, X.; Sidiropoulos, N.D.; Jiang, T.; Gershman, A. Multidimensional harmonic retrieval with applications in MIMO wireless channel sounding. In *Space-Time Processing for MIMO Communications*; Wiley Online Library: Hoboken, NJ, USA, 2005; pp. 41–75.
49. Yang, Z.; Li, J.; Stoica, P.; Xie, L. Sparse methods for direction-of-arrival estimation. In *Academic Press Library in Signal Processing*; Elsevier: Amsterdam, The Netherlands, 2018; Volume 7, pp. 509–581.
50. Haardt, M.; Pesavento, M.; Roemer, F.; El Korso, M.N. Subspace methods and exploitation of special array structures. In *Academic Press Library in Signal Processing*; Elsevier: Amsterdam, The Netherlands, 2014; Volume 3, pp. 651–717.
51. Haardt, M.; Roemer, F.; Del Galdo, G. Higher-order SVD-based subspace estimation to improve the parameter estimation accuracy in multidimensional harmonic retrieval problems. *IEEE Trans. Signal Process.* **2008**, *56*, 3198–3213. [[CrossRef](#)]
52. Kolda, T.G.; Bader, B.W. Tensor decompositions and applications. *SIAM Rev.* **2009**, *51*, 455–500. [[CrossRef](#)]
53. Liu, K.; da Costa, J.P.C.; So, H.C.; Huang, L.; Ye, J. Detection of number of components in CANDECOMP/PARAFAC models via minimum description length. *Digit. Signal Process.* **2016**, *51*, 110–123. [[CrossRef](#)]
54. Phan, A.; Tichavský, P.; Cichocki, A. CANDECOMP/PARAFAC decomposition of high-order tensors through tensor reshaping. *IEEE Trans. Signal Process.* **2013**, *61*, 4847–4860. [[CrossRef](#)]
55. Han, J.; Kamber, M.; Pei, J. *Data Mining, Concepts and Techniques*; Morgan Kaufmann Publishers: San Mateo, CA, USA, 2011.
56. Foerster, J.R.; Pendergrass, M.; Molisch, A.F. A UWB channel model for ultra wideband indoor communications. In Proceedings of the International Symposium on Wireless Personal Multimedia Communications, Yokosuka, Japan, 19–22 October 2003.
57. Czink, N. The Random-Cluster Model. Ph.D. Thesis, Technische Universität Wien, Vienna, Austria, 2007.
58. Wyne, S.; Czink, N.; Karedal, J.; Almers, P.; Tufvesson, F.; Molisch, A.F. A cluster-based analysis of outdoor-to-indoor office MIMO measurements at 5.2 GHz. In Proceedings of the IEEE Vehicular Technology Conference, Melbourne, Australia, 7–10 May 2006.
59. Salo, J.; Salmi, J.; Czink, N.; Vainikainen, P. Automatic clustering of nonstationary MIMO channel parameter estimates. In Proceedings of the Information and Communications Technologies, Cape Town, South Africa, 3–6 May 2005.
60. Czink, N.; Yin, X.; Ozelik, H.; Herdin, M.; Bonek, E.; Fleury, B.H. Cluster characteristics in a MIMO indoor propagation environment. *IEEE Trans. Wirel. Commun.* **2007**, *6*, 1465–1475. [[CrossRef](#)]

61. He, R.; Ai, B.; Molisch, A.F.; Stuber, G.L.; Li, Q.; Zhong, Z.; Yu, J. Clustering enabled wireless channel modeling using big data algorithms. *IEEE Commun. Mag.* **2018**, *56*, 177–183. [[CrossRef](#)]
62. Jain, A.K. Data clustering: 50 years beyond K-means. *Pattern Recognit. Lett.* **2010**, *31*, 651–666. [[CrossRef](#)]
63. Frey, B.J.; Dueck, D. Clustering by passing messages between data points. *Science* **2007**, *315*, 972–976. [[CrossRef](#)]
64. Ester, M.; Kriegel, H.P.; Sander, J.; Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. In Proceedings of the International Conference on Knowledge Discovery and Data Mining, Portland, OR, USA, 2–4 August 1996; pp. 226–231.
65. García-Fernández, Á.F.; Williams, J.L.; Granström, K.; Svensson, L. Poisson multi-Bernoulli mixture filter: Direct derivation and implementation. *IEEE Trans. Aerosp. Electron. Syst.* **2018**, *54*, 1883–1901. [[CrossRef](#)]
66. Tibshirani, R.; Walther, G.; Hastie, T. Estimating the number of clusters in a data set via the gap statistic. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2001**, *63*, 411–423. [[CrossRef](#)]
67. Rahmathullah, A.S.; García-Fernández, Á.F.; Svensson, L. Generalized optimal sub-pattern assignment metric. In Proceedings of the 20th IEEE International Conference on Information Fusion (Fusion), Xi'an, China, 10–13 July 2017.
68. Mitsis, G.; Tsiropoulou, E.E.; Papavassiliou, S. Data offloading in UAV-assisted multi-access edge computing systems: A resource-based pricing and user risk-awareness approach. *Sensors* **2020**, *20*, 2434. [[CrossRef](#)] [[PubMed](#)]
69. Mitsis, G.; Apostolopoulos, P.A.; Tsiropoulou, E.E.; Papavassiliou, S. Intelligent dynamic data offloading in a competitive mobile edge computing market. *Future Internet* **2019**, *11*, 118. [[CrossRef](#)]
70. Ferdowsi, A.; Challita, U.; Saad, W. Deep learning for reliable mobile edge analytics in intelligent transportation systems: An overview. *IEEE Veh. Technol. Mag.* **2019**, *14*, 62–70. [[CrossRef](#)]
71. Mahler, R.P. *Advances in Statistical Multisource-Multitarget Information Fusion*; Artech House: Norwood, MA, USA, 2014.
72. Mahler, R.P. Multitarget Bayes filtering via first-order multitarget moments. *IEEE Trans. Aerosp. Electron. Syst.* **2003**, *39*, 1152–1178. [[CrossRef](#)]
73. Murty, K.G. An algorithm for ranking all the assignments in order of increasing cost. *Oper. Res.* **1968**, *16*, 682–687. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).