

# Probabilistic Scan Matching: Bayesian Pose Estimation from Point Clouds

Rico Mendrzik\* and Florian Meyer†

\*Ibeo Automotive Systems GmbH, Hamburg, Germany

†University of California San Diego, San Diego, USA

Email: rico.mendrzik@ibeo-as.com, flmeyer@ucsd.edu

**Abstract**—Estimating position and orientation change of a mobile platform from two consecutive point clouds provided by a high-resolution sensor is a key problem in autonomous navigation. In particular, scan matching algorithms aim to find the translation and rotation of the platform such that the two point clouds coincide. The association of measurements in point cloud one with measurements in point cloud two is a problem inherent to scan matching. Existing methods perform non-probabilistic data association, i.e., they assume a single association hypothesis. This leads to overconfident pose estimates and reduced estimation accuracy in ambiguous environments. Our probabilistic scan matching approach addresses this issue by considering all association hypotheses with their respective likelihoods. We formulate a holistic Bayesian estimation problem for both data association and pose inference and present the corresponding joint factor graph. Near-optimum maximum a posteriori (MAP) estimates of the sensor pose are computed by performing iterative message passing on the factor graph. Our numerical study shows performance improvements compared to non-probabilistic scan matching methods that are based on the normal distributions transform (NDT) and implicit moving least squares (IMLS).

## I. INTRODUCTION

High-resolution sensors for situational awareness such as Light Detection and Ranging (LiDAR) and imaging sonar are paramount for autonomous navigation due to their unique sensing capabilities. LiDAR sensors can provide extremely high resolution accompanied by precise depth information and are thus ideal for automated driving and advanced driver assistance systems (ADAS) [1]. Imaging sonar generates accurate echo intensity profiles of the scanned area and can enable autonomous underwater vehicles (AUV) to navigate in an unknown and possibly unstructured environment [2]. Beyond object tracking and classification or environmental perception, high-resolution sensors can be cornerstones for ego-state estimation which can help navigating in challenging environments [3]–[10]. By matching two scans obtained at different time instances, information on the change of position and orientation (also known as *relative pose*) of the sensor can be extracted. A simplistic example from the LiDAR domain is depicted in Fig. 1.

State-of-the-art scan matching approaches [3]–[10] typically use heuristics that require a good initial guess on the relative pose of the sensor. Moreover, they rely on a key as-

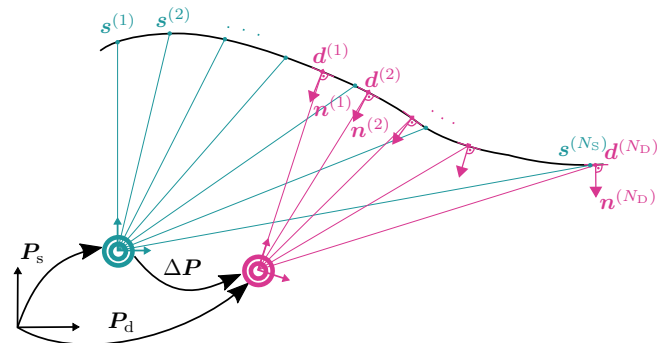


Fig. 1. Scenario: Our goal is to infer the relative pose,  $\Delta P$ , of LiDAR sensor (concentric circles) by matching the source point cloud  $s = [s^{(1)T} s^{(1)T} \dots s^{(N_s)T}]^T$  (green dots) and destination surface, consisting of scan points  $d = [d^{(1)T} d^{(2)T} \dots d^{(N_D)T}]^T$  (magenta diamonds) and normal vectors  $n = [n^{(1)T} n^{(2)T} \dots n^{(N_D)T}]^T$  (magenta arrows). Laser rays are indicated by thin green and magenta lines originating from the LiDAR sensor.

sumption: The association of scan data can be described by a single deterministic association hypothesis. This assumption can lead to poor and overconfident estimates of the relative pose if the scans contain ambiguous data or non-associable artifacts. State-of-the-art algorithms typically fail if the scan data association is erroneous, the generating environment does not allow for an unambiguous estimation of the relative pose, or scan points originate from detection errors belong to dynamic objects. Another frequently-occurring challenge of the scan matching paradigm, is partial observability of the environment, which describes the phenomenon that two scans recorded at different poses almost never depict the identical scene.

In contrast to state-of-the-art method, we formulate scan matching as a Bayesian inference problem. This approach aims to increase resilience to sensing artifacts and partial observability by performing scan associations probabilistically [11]–[14]. Our main contributions can be summarized as follows:

- We derive a stochastic model which entangles both scan association and pose information in a holistic fashion.
- We present a Bayesian inference algorithm for relative pose inference based on probabilistic scan matching.

*Notation:* Random variables are displayed in sans serif,

upright fonts; their realizations in serif, italic fonts. Vectors and matrices are denoted by bold lowercase and uppercase letters, respectively. For example, a random variable and its realization are denoted by  $\mathbf{x}$  and  $x$ ; a random vector and its realization are denoted by  $\mathbf{x}$  and  $\mathbf{x}$ . Sets are denoted by calligraphic font, e.g.  $\mathcal{X}$ . Furthermore,  $\mathbf{x}^T$  denotes the transpose of vector  $\mathbf{x}$ .  $f(\mathbf{x})$  denotes the probability density function (PDF) or probability mass function (PMF) of continuous or discrete random vector  $\mathbf{x}$ . Finally,  $\mathbf{0}_n$  is the row vector of size  $n$  containing all zeros.

## II. FUNDAMENTALS AND PROBLEM FORMULATION

We define the notion of poses, describe our sensor and measurement models, and present our problem formulation.

### A. Preliminaries

1) *Pose Representation*: We refer to the position and orientation of a rigid object as pose. The pose of the sensor in  $N_{\text{dim}}$ -dimensional space can be conveniently represented by the following matrix from the special Euclidean group  $SE(N_{\text{dim}})$  [15]

$$\mathbf{P} = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_{N_{\text{dim}}} & 1 \end{pmatrix}$$

where  $\mathbf{R}$  is a  $N_{\text{dim}} \times N_{\text{dim}}$  rotation matrix and  $\mathbf{t} \in \mathbb{R}^{N_{\text{dim}}}$  is a translation vector. From source pose  $\mathbf{P}_s$  and destination pose  $\mathbf{P}_d$ , we can determine the relative pose  $\Delta\mathbf{P} \in SE(N_{\text{dim}})$  as

$$\Delta\mathbf{P} = \mathbf{P}_d \mathbf{P}_s^{-1} = \begin{pmatrix} \Delta\mathbf{R} & \Delta\mathbf{t} \\ \mathbf{0}_{N_{\text{dim}}} & 1 \end{pmatrix}$$

where the  $N_{\text{dim}} \times N_{\text{dim}}$  rotation matrix  $\Delta\mathbf{R}$  and the translation vector  $\Delta\mathbf{t} \in \mathbb{R}^{N_{\text{dim}}}$  describe the change of orientation and displacement between  $\mathbf{P}_s$  and  $\mathbf{P}_d$ , respectively.

2) *Sensor Characteristics*: We consider a high-resolution sensors that illuminates the environment and uses reflected and scattered signal components to infer spatial information on the observed scene. Typically, sets of scan points (so-called point clouds) are the results of sensing. Here, we consider two distinct sets: *source* and *destination* point clouds. We define the destination point cloud as

$$\mathbf{d} \triangleq [\mathbf{d}^{(1)T} \mathbf{d}^{(2)T} \dots \mathbf{d}^{(N_D)T}]^T,$$

where  $\mathbf{d}^{(l)} \in \mathbb{R}^{N_{\text{dim}}}$  denotes the  $l^{\text{th}}$  noisy scan point in Cartesian coordinates. Similarly, we introduce the source point cloud  $\mathbf{s} \triangleq [\mathbf{s}^{(1)T} \mathbf{s}^{(2)T} \dots \mathbf{s}^{(N_S)T}]^T$ .

3) *Surface Information*: Surface normal vectors are a common method to preserve information on the shapes of the underlying surfaces without the need of having explicit representations of them [16]–[18]. We build upon this idea and augment every destination scan point  $\mathbf{d}^{(i)}$  by a normal vector  $\mathbf{n}^{(i)} \in \mathbb{R}^{N_{\text{dim}}}$ ,  $i = 1, 2, \dots, N_D$ . The normal vector  $\mathbf{n}^{(i)}$  of destination scan point  $\mathbf{d}^{(i)}$  is determined based on the set of neighboring destination scan points  $\mathcal{N}_D^{(i)} = \{\mathbf{d}^{(i')}\mid \|\mathbf{d}^{(i')} - \mathbf{d}^{(i)}\| \leq d_{\text{TH}}\}$  where the threshold  $d_{\text{TH}}$  controls the size of the neighborhood. The literature on the computation of surface normal vectors is rich and

the interested reader can find an overview in [18]. In our work, we determine  $i^{\text{th}}$  surface normal vector,  $\mathbf{n}^{(i)}$ , as the eigenvector which corresponds to the smallest eigenvalue of the sample covariance matrix of  $\mathcal{N}^{(i)}$ . Hereafter, we will refer to destination scan points and their associated normal vector as destination surface points.

### B. Measurement Model

Before introducing our measurement model, it is worth noting that all scan points are considered in homogeneous coordinates to cope with our pose representation, i.e.,  $\mathbf{d}_H^{(i)} \triangleq [\mathbf{d}^{(i)T} 1]^T$ ,  $\mathbf{n}_H^{(i)} \triangleq [\mathbf{n}^{(i)T} 0]^T$ ,  $i = 1, 2, \dots, N_D$ , and  $\mathbf{s}_H^{(j)} \triangleq [\mathbf{s}^{(j)T} 1]^T$ ,  $j = 1, 2, \dots, N_S$ . For mathematical convenience, we introduce stacked vectors of source points, destination points and their normals as follows:  $\mathbf{s}_H = [\mathbf{s}_H^{(1)T} \mathbf{s}_H^{(2)T} \dots \mathbf{s}_H^{(N_S)T}]^T$ ,  $\mathbf{d}_H = [\mathbf{d}_H^{(1)T} \mathbf{d}_H^{(2)T} \dots \mathbf{d}_H^{(N_D)T}]^T$ , and  $\mathbf{n}_H = [\mathbf{n}_H^{(1)T} \mathbf{n}_H^{(2)T} \dots \mathbf{n}_H^{(N_D)T}]^T$ .

1) *Point-to-plane Error Metric*: We can link the relative pose  $\Delta\mathbf{P}$  to source scan points,  $\mathbf{s}^{(j)}$ ,  $\forall j$ , and destination surface points,  $\{\mathbf{d}^{(i)}, \mathbf{n}^{(i)}\}$ ,  $\forall i$ , via a point-to-plane metric [19]. In particular, our measurement model is given by

$$\mathbf{n}_H^{(i)} \left( \mathbf{d}_H^{(i)} - \Delta\mathbf{P} \mathbf{s}_H^{(j)} \right) = e^{(j)} \quad j = 1, 2, \dots, N_S \quad (1)$$

in which  $e^{(j)}$  is an error term with known<sup>1</sup> PDF  $f(e^{(j)})$ . The left-hand side of (1) is typically referred to as pseudo measurements [20]. For future reference, we also introduce the joint measurement vectors  $\mathbf{z}^{(i,j)} \triangleq [\mathbf{d}_H^{(i)T} \mathbf{n}_H^{(i)T} \mathbf{s}_H^{(j)T}]^T$ ,  $\mathbf{z}^{(i)} = [\mathbf{d}_H^{(i)T} \mathbf{n}_H^{(i)T} \mathbf{s}_H^T]^T$ , and  $\mathbf{z} = [\mathbf{d}_H^T \mathbf{n}_H^T \mathbf{s}_H^T]^T$ . Note that from (1) and the error statistics  $f(e^{(j)})$ , we can directly obtain conditional PDFs  $f(\mathbf{z}^{(i,j)}|\Delta\mathbf{P})$  for each pair  $(i, j)$ .

The measurement model (1) is briefly discussed in what follows. Suppose we knew the relative pose  $\Delta\mathbf{P}$  and the association  $(i, j)$  of source point  $\mathbf{s}_H^{(j)}$  and destination surface point  $\{\mathbf{d}_H^{(i)}, \mathbf{n}_H^{(i)}\}$ . Then, the expression within the parenthesis on the left-hand side of (1) implies that we can transform the source scan point  $\mathbf{s}_H^{(j)}$  via relative pose  $\Delta\mathbf{P}$  and we will retrieve the destination scan point  $\mathbf{d}_H^{(i)}$ . Note that the sensor does not measure the exact same point on a surface from different poses due to its finite resolution (see Fig. 1). We account for this by embedding the local information of the surface. This is achieved by multiplying with the surface normal  $\mathbf{n}_H^{(i)}$ . Thus, the deviation of the transformed source point from destination point is measured only in the direction of the surface normal. The remaining error is given by  $e^{(j)}$ .

2) *Associability of Scan Points*: High-resolution sensing, with effects such as occlusion and finite sensing range, make scenes appear differently even if they are observed from similar poses. As a result, the number of source points  $N_S$  and destination points  $N_D$  are generally different and not all scan points are associable. For robust scan matching, we need to account for scan points which cannot be associated. In particular, we introduce the destination-oriented

<sup>1</sup>In partice, full knowledge of  $f(e^{(j)})$  is typically unavailable and the PDF needs to be approximated by its most significant moments.

data association vector  $\mathbf{a} = [a^{(1)} a^{(2)} \dots a^{(N_D)}]^T$  in which  $a^{(i)}, i \in \{1, 2, \dots, N_D\}$  refers to the association variable of the  $i^{\text{th}}$  destination surface point.  $a^{(i)}$  can take any value from the set  $\{0, 1, \dots, N_S\}$ , where  $a^{(i)} = j$  means that the  $i^{\text{th}}$  destination surface point relates to the  $j^{\text{th}}$  source point. Moreover,  $a^{(i)} = 0$  means that the destination surface point  $\{\mathbf{d}_H^{(i)}, \mathbf{n}_H^{(i)}\}$  is not associable to any source point  $\mathbf{s}_H^{(j)}, \forall j$ .

### C. Problem Formulation

Fig. 1 graphically depicts the problem we attempt to solve. Our goal is to infer the relative pose  $\Delta \mathbf{P}$  of the sensor based on the source and destination point clouds by transforming the source point cloud onto the destination point cloud. Note that the associations between source scan points and destination surface points  $\mathbf{a}$  are unknown and need to be inferred jointly with the relative pose  $\Delta \mathbf{P}$  which is a random matrix on the manifold  $SE(N_{\text{dim}})$ .

In the pursuit of solving this problem robustly, we resort to the probabilistic domain. Our goal is to obtain a Bayes-optimum estimate of the relative pose  $\Delta \mathbf{P}$  considering all valid data association hypotheses  $\mathbf{a}$ . First, we marginalize the joint posterior distribution  $f(\Delta \mathbf{P}, \mathbf{a} | \mathbf{z})$  to obtain  $f(\Delta \mathbf{P} | \mathbf{z})$  which factors in all valid data association hypotheses. Then, we derive an maximum a posteriori (MAP) estimate by maximizing the marginal posterior distribution, i.e.,

$$\Delta \hat{\mathbf{P}}_{\text{MAP}} = \arg \max_{\Delta \mathbf{P}} f(\Delta \mathbf{P} | \mathbf{z}). \quad (2)$$

Note that determining  $f(\Delta \mathbf{P} | \mathbf{z})$  by direct marginalization (i.e.,  $f(\Delta \mathbf{P} | \mathbf{z}) = \sum_{\mathbf{a}} f(\Delta \mathbf{P}, \mathbf{a} | \mathbf{z})$ ) is infeasible due to the high-dimensionality of the summation space as discussed in Section III. We consider feasible approximate calculation based on the notion of message passing [21]–[23].

## III. STOCHASTIC MODELING

Here, we present our novel stochastic model which couples data association, sensor pose, and point cloud measurements.

### A. Modeling of the Point-to-Surface-Point Associations

We attempt to associate  $N_S$  source points to  $N_D$  destination surface points. Recall, though, that not all points in the source and destination scan are associable. We assume the number of non-associable source points,  $N_{NA}$ , to be random and Poisson distributed with mean  $\lambda_{NA}$ . Hence the total number of source points,  $N_S$ , is also random and given by

$$N_S = N_{NA} + |\mathcal{D}_a|, \quad (3)$$

where  $\mathcal{D}_a \triangleq \{i \in \{1, 2, \dots, N_D\} : a^{(i)} \neq 0\}$ .

Valid associations are those in which each source point  $\mathbf{s}_H^{(j)}$  is associated to at most one destination surface point  $\{\mathbf{d}_H^{(i)}, \mathbf{n}_H^{(i)}\}$ , and each destination surface point  $\{\mathbf{d}_H^{(i)}, \mathbf{n}_H^{(i)}\}$  is associated to at most one source point  $\mathbf{s}_H^{(j)}$ . This assumptions imposes constraints on  $\mathbf{a}$ . In particular,  $\mathbf{a}$  is valid if and only if  $a^{(i)} = j \in \{1, \dots, N_S\}$ ,  $a^{(i')} \neq j, \forall i \neq i'$ . These constraints can be conveniently expressed by

$$\Psi(\mathbf{a}) = \begin{cases} 0, & a^{(i)} = a^{(i')} = j, i \neq i', j \neq 0 \\ 1, & \text{otherwise.} \end{cases} \quad (4)$$

Note that the cardinality of valid hypothesis is huge, rendering enumeration of all hypotheses intractable [12], [24]. To address this issue, we introduce a redundant formulation of  $\mathbf{a}$ : the source-oriented data association vector  $\mathbf{b} = [b^{(1)} b^{(2)} \dots b^{(N_S)}]^T$ , in which  $b^{(j)} \in \{1, \dots, N_D\}$  and  $b^{(j)} = i$  means that the  $j^{\text{th}}$  source point is associated to the  $i^{\text{th}}$  destination surface point. If  $b^{(j)} = 0$ , the  $j^{\text{th}}$  source point is non-associable. For source-oriented data association vectors  $\mathbf{b}$  we have  $b^{(j)} = i \in \{1, \dots, N_D\}$ ,  $b^{(j')} \neq i, \forall j' \neq j$ . Note that the set of all valid vectors  $\mathbf{b}$  contains exactly the same information as that of all valid vectors  $\mathbf{a}$ .

The reason for introducing this redundant data association representation is that validity of association hypotheses can now be described via pairs of  $a^{(i)}$  and  $b^{(j)}$  instead of the entire vector  $\mathbf{a}$ , making evaluation tractable. This is described in more detail in [12], [24]. It can be easily verified that valid data associations can be described by

$$\Psi(\mathbf{a}, \mathbf{b}) = \prod_{i=1}^{N_D} \prod_{j=1}^{N_S} \psi_{i,j}(a^{(i)}, b^{(j)}), \quad (5)$$

where

$$\psi_{i,j}(a^{(i)}, b^{(j)}) \triangleq \begin{cases} 0, & a^{(i)} = j, b^{(j)} \neq i \\ & b^{(j)} = i, a^{(i)} \neq j \\ 1, & \text{otherwise.} \end{cases} \quad (6)$$

Note that we treat both  $\mathbf{a}$  and  $\mathbf{b}$  as random vectors that are inferred jointly with the relative pose  $\Delta \mathbf{P}$ .

### B. Joint Posterior Distribution

In what follows, we derive the joint posterior distribution  $f(\Delta \mathbf{P}, \mathbf{a}, \mathbf{b} | \mathbf{z})$  and the corresponding factor graph. For  $\mathbf{z}$  observed and thus fixed, we obtain

$$f(\Delta \mathbf{P}, \mathbf{a}, \mathbf{b} | \mathbf{z}) \propto f(\mathbf{z}, \mathbf{a}, \mathbf{b}, N_S | \Delta \mathbf{P}) f(\Delta \mathbf{P}). \quad (7)$$

$f(\Delta \mathbf{P})$  is a random matrix pdf with support  $SE(N_{\text{dim}})$  that incorporates prior knowledge on the relative pose  $\Delta \mathbf{P}$ . The conditional PDF  $f(\mathbf{z}, \mathbf{a}, \mathbf{b}, N_S | \Delta \mathbf{P})$  is given by

$$f(\mathbf{z}, \mathbf{a}, \mathbf{b}, N_S | \Delta \mathbf{P}) \quad (8)$$

$$= \frac{e^{-\lambda_{NA}}}{N_{NA}!} \Psi(\mathbf{a}, \mathbf{b}) v(\Delta \mathbf{P}; \mathbf{s}_H) \prod_{i=1}^{N_D} q_i(\Delta \mathbf{P}, a^{(i)}; \mathbf{z}^{(i)})$$

where we introduced  $v(\Delta \mathbf{P}; \mathbf{s}_H) \triangleq \prod_{j=1}^{N_S} f_{NA}(\mathbf{s}_H^{(j)} | \Delta \mathbf{P})$  and

$$q_i(\Delta \mathbf{P}, a^{(i)}; \mathbf{z}^{(i)}) = \begin{cases} f_A(\Delta \mathbf{P}) \frac{f(\mathbf{z}^{(i)}, a^{(i)} | \Delta \mathbf{P})}{\lambda_{NA} f_{NA}(\mathbf{s}_H^{(a^{(i)})} | \Delta \mathbf{P})} & a^{(i)} \neq 0 \\ 1 - f_A(\Delta \mathbf{P}) & a^{(i)} = 0 \end{cases} \quad (10)$$

and used  $\Psi(\mathbf{a}, \mathbf{b})$  from (5). A detailed derivation of (8) together with the underlying assumptions is provided in Appendix A. By using (8) in (7), we obtain the joint posterior distribution  $f(\Delta \mathbf{P}, \mathbf{a}, \mathbf{b} | \mathbf{z})$  shown in (9).

Let us discuss the expression in (9) to gain further insights into our model. Suppose that we have a destination

$$f(\Delta P, \mathbf{a}, \mathbf{b} | \mathbf{z}) \propto f(\Delta P) v(\Delta P; \mathbf{s}_H) \prod_{i=1}^{N_D} q_i(\Delta P, a^{(i)}; \mathbf{z}^{(i)}) \prod_{j=1}^{N_S} \psi_{i,j}(a^{(i)}, b^{(j)}) \quad (9)$$

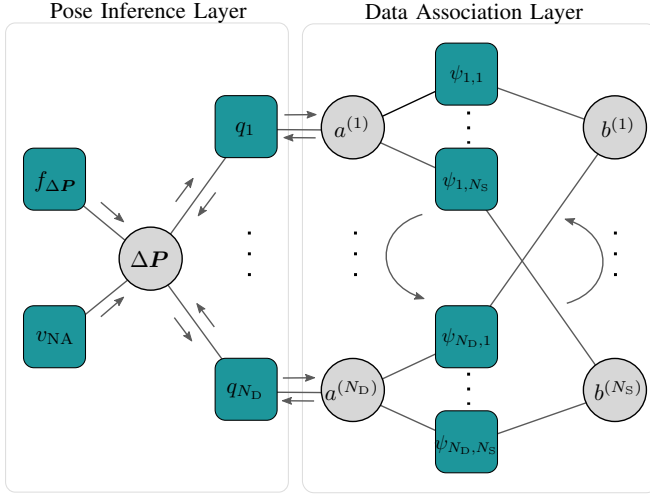


Fig. 2. Factor graph of the joint posterior distribution in (9). The graph can be sub-divided into a data association layer and a pose inference layer. In the latter, our goal is to infer the marginal posterior distribution  $f(\Delta P | \mathbf{z})$  via message passing (messages are indicated with grey arrows). This PDF will be maximized in a subsequent step to obtain a MAP estimate of the relative pose  $\Delta P$ . We use the following shorthand notations within the graph:  $f_{\Delta P} \triangleq f(\Delta P)$ ,  $q_i \triangleq q_i(\Delta P, a^{(i)}; \mathbf{z}^{(i)})$ ,  $\psi_{i,j} \triangleq \psi_{i,j}(a^{(i)}, b^{(j)})$ ,  $v_{NA} \triangleq v(\Delta P; \mathbf{s}_H)$ .

surface point  $i$ ,  $\{\mathbf{d}_H^{(i)}, \mathbf{n}_H^{(i)}\}$ , that is associated with source point  $\mathbf{s}_H^{(j)}$ , i.e.,  $a^{(i)} = j$ . Then, we have  $q_i(\Delta P, j, \mathbf{z}^{(i)}) = f_A(\Delta P) f(\mathbf{z}^{(i,j)} | \Delta P) / (\lambda_{NA} f_{NA}(\mathbf{s}_H^{(j)} | \Delta P))$ , where  $f_A(\cdot)$  accounts for the fact that destination surface point  $i$  is associable and contributes pseudo measurement  $\mathbf{z}^{(i,j)}$  with likelihood  $f(\mathbf{z}^{(i,j)} | \Delta P)$ . Lastly, the term in the denominator,  $f_{NA}(\mathbf{s}_H^{(j)} | \Delta P)$ , reflects the fact that source point  $j$ ,  $\mathbf{s}_H^{(j)}$ , is associable as well since it cancels the  $j^{\text{th}}$  term in  $v(\Delta P; \mathbf{s}_H)$ . On the contrary, if the  $i^{\text{th}}$  destination surface point is not associated with any source point, we have  $q_i(\Delta P, 0, \mathbf{z}^{(i)}) = 1 - f_A(\Delta P)$ . This term takes the likelihood of non-associability,  $1 - f_A(\Delta P)$ , of destination surface point  $i$  into account. Finally,  $\Psi(\mathbf{a}, \mathbf{b})$  enforces validity of associations.

#### IV. GRAPH-BASED INFERENCE

The factor graph of the joint posterior distribution in (9) is depicted in Fig. 2. Our goal is to infer the marginal posterior distribution  $f(\Delta P | \mathbf{z}) = \sum_{\mathbf{a}} \sum_{\mathbf{b}} f(\Delta P, \mathbf{a}, \mathbf{b} | \mathbf{z})$  needed for the calculation of an MAP estimate  $\hat{\Delta P}_{\text{MAP}}$ . We employ the notation of message passing according to the sum-product rules [21]–[23] to calculate an approximation of the marginal posterior distribution  $f(\Delta P | \mathbf{z})$ . The messages passed along the edges of the factor graph in Fig. 2. Messages that involve integration  $\int d\Delta P$  over the sensor pose  $\Delta P$  are approximated by sampling-based integration [25]. We present the steps of our inference method in Algorithm 1.

#### Algorithm 1 Obtaining the Marginal Posterior $f(\Delta P | \mathbf{z})$

- 1: Initialize prior information
 
$$\mu_{f_{\Delta P} \rightarrow \Delta P}(\Delta P) = f(\Delta P)$$

$$\mu_{v_{NA} \rightarrow \Delta P}(\Delta P) = v(\Delta P; \mathbf{s}_H)$$
- 2: Forward prior information
 **For**  $i = 1, 2, \dots, N_D$ 

$$\mu_{\Delta P \rightarrow q_i}(\Delta P) = \mu_{f_{\Delta P} \rightarrow \Delta P}(\Delta P) \mu_{v_{NA} \rightarrow \Delta P}(\Delta P)$$
**End For**
- 3: Evaluate associations
 **For**  $i = 1, 2, \dots, N_D$ 

$$\mu_{q_i \rightarrow a^{(i)}}(a^{(i)}) = \int q_i(\Delta P, a^{(i)}; \mathbf{z}^{(i)}) \mu_{\Delta P \rightarrow q_i}(\Delta P) d\Delta P$$

$$\approx \sum_{p=1}^{N_P} q_i(\Delta P^{(p)}, a^{(i)}; \mathbf{z}^{(i)}) v(\Delta P^{(p)}; \mathbf{s}_H)$$
- where  $\Delta P^{(p)}$ ,  $p = 1, 2, \dots, N_P$  are random samples drawn from  $\mu_{f_{\Delta P} \rightarrow \Delta P}(\Delta P) = f(\Delta P)$ .
 **End For**
- 4: Perform  $N_{DA}$  data association iterations
 **For**  $l = 1, 2, \dots, N_{DA}$ 
 Message exchange between  $a^{(i)}$  and  $b^{(j)}$  see [12], [24]
 
$$(\mu_{a^{(i)} \rightarrow \psi_{i,j}}^{(l)}, \mu_{\psi_{i,j} \rightarrow b^{(j)}}^{(l)}, \mu_{b^{(j)} \rightarrow \psi_{i,j}}^{(l)}, \mu_{\psi_{i,j} \rightarrow a^{(i)}}^{(l)})$$
**End For**
- 5: Forward data association result
 **For**  $i = 1, 2, \dots, N_D$ 

$$\mu_{a^{(i)} \rightarrow q_i}(a^{(i)}) = \prod_{j=1}^{N_S} \mu_{\psi_{i,j} \rightarrow a^{(i)}}^{(N_{DA})}(a^{(i)})$$
**End For**
- 6: Marginalize associations
 **For**  $i = 1, 2, \dots, N_D$ 

$$\mu_{q_i \rightarrow \Delta P}(\Delta P) = \sum_{a^{(i)}=0}^{N_S} q_i(\Delta P, a^{(i)}; \mathbf{z}^{(i)}) \mu_{a^{(i)} \rightarrow q_i}(a^{(i)})$$
**End For**
- 7: Determine marginalized posterior  $f(\Delta P | \mathbf{z})$ 

$$f(\Delta P | \mathbf{z}) \approx f(\Delta P) v(\Delta P; \mathbf{s}_H) \prod_{i=1}^{N_D} \mu_{q_i \rightarrow \Delta P}(\Delta P) \quad (11)$$

First, we perform Algorithm 1 to compute an accurate approximation of the marginal posterior  $f(\Delta P | \mathbf{z})$ . Then, our goal is to derive the MAP estimate,  $\hat{\Delta P}_{\text{MAP}} = \arg \max_{\Delta P} f(\Delta P | \mathbf{z})$ , by maximizing (11) with  $N_{IT}$ . To that end, we employ iterative numerical solvers that exploit the structure of  $SE(N_{\text{dim}})$  [26]. As initial guess, we use the pose sample that maximizes the marginal posterior, i.e.,

$$\Delta P_{\text{init}} = \arg \max_{\Delta P^{(p)}} f(\Delta P^{(p)} | \mathbf{z}), \quad p \in \{1, 2, \dots, N_P\} \quad (12)$$

Relative Trans. Error	Proposed	NDT [3]	IMLS [6]
$e_{\text{trans},50}$	1.87%	1.800%	6.60%
$e_{\text{trans},95}$	7.723%	94.65%	28.06%
Relative Rotation Error	Proposed	NDT [3]	IMLS [6]
$e_{\text{rot},50}$	0.0349°/m	0.0203°/m	0.2487°/m
$e_{\text{rot},95}$	0.1123°/m	0.0759 °/m	2.2016°/m

TABLE I  
ERROR QUANTILES OF SIMULATED POSE ESTIMATION METHODS.

where  $\Delta \mathbf{P}^{(p)}$ ,  $p = 1, 2, \dots, N_P$  are random samples drawn from  $f(\Delta \mathbf{P})$ .

We emphasize that our approach determines the full marginal posterior distribution  $f(\Delta \mathbf{P}|z)$ . Obtaining an estimate by finding its mode is by no means the only information that we can extract from it. We could derive, e.g., uncertainty information on our estimate that we can use to reason about the quality of matching. Such knowledge is typically required for closing loops in simultaneous localization and mapping (SLAM) applications. Beyond scan-to-scan matching, our proposed stochastic model can be embedded into online or offline factor graph-based trajectory estimation.

The computational complexities in the pose inference and data association layers (see Fig. 2) as well as for optimization are given by  $\mathcal{O}(N_D N_S N_*)$ ,  $*$   $\in \{P, DA, IT\}$ . To avoid high complexity, segmentation of point clouds [27] can be used to generate semantic clusters (e.g., road surface, guardrails, poles, etc.) where each cluster contains significantly fewer points than the original point cloud. Running multiple instances (one for each cluster) of our proposed algorithm in parallel reduces the complexity as  $N_{D,i} N_{S,i} \ll N_D N_S$ , where  $N_{D,i}$  and  $N_{S,i}$  denote the number of destination and source points in cluster  $i$ , respectively.

## V. NUMERICAL EVALUATION

In this section, we evaluate the our scan matching approach based on simulated data and present the results.

### A. Simulation Scenario

We consider a vehicle that drives in the urban environment depicted in Fig. 3. The vehicle moves with a constant velocity of 10 m/s along the indicated trajectory (red line). It is equipped with a 360°-LiDAR sensor that generates scans every  $\Delta T = 80$  ms. We treat the current scan as the source scan and the previous scan as the destination scan to obtain pose estimates that are inline with the motion of the vehicle.

The sensor has an angular resolution of 1 deg and a maximum sensing range of 100 m. The scanner obtains range and bearing estimates that are corrupted by zero-mean Gaussian noise with standard deviation  $\sigma_{\text{range}} = 0.05$  m and  $\sigma_{\text{bearing}} = 0.5$  deg, respectively. Non-associable source points are assumed to be uniformly distributed in range and bearing, being independent of  $\Delta \mathbf{P}$ . The number of non-associable source points is assumed Poisson distributed with mean  $\lambda_{NA} = 1$ . For surface normal computation, we use  $d_{TH} = 2$  m. An exemplary scan is depicted in Fig. 3. We

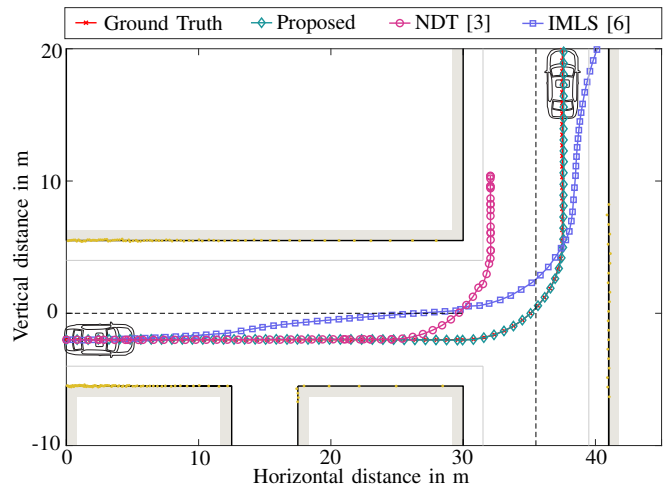


Fig. 3. Simulation scenario and results: The ground truth trajectory is shown against the accumulated pose change estimates of the our approach and its benchmarks. Scan points obtained at the first time instance are depicted by yellow dots to visualize the characteristics of the LiDAR sensor.

use the following parameters for our proposed algorithm. The probability of associability is constant  $f_A(\Delta \mathbf{P}) = 0.8$ . For each  $j \in \{1, 2, \dots, N_S\}$ , the noise statistics of  $f(e^{(j)})$  are assumed zero-mean Gaussian distributed with standard deviation  $\sigma_{e^{(j)}} = 0.03$  m. Translation and rotation prior are considered uniform within the intervals  $[-10, 10]$  m and  $[-90, 90]$  °. We performed  $N_{MC} = 100$  Monte Carlo trials.

### B. Performance Comparison

We compare our probabilistic scan matching approach to normal distributions transform (NDT) scan matching [3] and implicit moving least squares (IMLS) scan matching [6]. In our implementation of IMLS, we set the neighborhood parameter to  $h = 2$  and obtained estimates by minimizing [6, Eq. (1)]. We employ the following error metrics to evaluate the translation and rotation drift:  $e_{\text{trans}} = \|\Delta \hat{\mathbf{t}}_{\text{MAP}} - \Delta \mathbf{t}\| / \|\Delta \mathbf{t}\|$  and  $e_{\text{rot}} = \arccos((\text{Tr}(\Delta \hat{\mathbf{R}}_{\text{MAP}}^{-1} \Delta \mathbf{R}) - 1)/2) / \|\Delta \mathbf{t}\|$ , where  $\Delta \hat{\mathbf{t}}_{\text{MAP}}$  and  $\Delta \hat{\mathbf{R}}_{\text{MAP}}$  are the estimated translation and rotation extracted from  $\Delta \hat{\mathbf{P}}_{\text{MAP}}$ ,  $\Delta \mathbf{t}$  and  $\Delta \mathbf{R}$  are the ground truth translation and rotation, and  $\text{Tr}(\cdot)$  is the trace operator.

Results are depicted in Tab. I. Here,  $e_{\text{trans},q}$  refers to an error threshold such that  $q\%$  of all relative observed translation errors  $e_{\text{trans}}$  are below this threshold. The same logic applies to  $e_{\text{rot},q}$ . For example,  $e_{\text{trans},50}$  is the median relative translation error. From Tab. I we observe that median relative translation and rotation errors of NDT and our proposed approach both show low drifts. However, considering  $e_{\text{trans},q}$ , NDT shows a significantly larger drift compared to our proposed probabilistic approach. IMLS scan matching shows large drifts, presumably due to the absence of a good initial guess provided to the optimizers. Exemplary estimated trajectories obtained via a concatenation of relative pose estimates are also shown in Fig. 3 where we can see how closely our approach attains ground truth.

## VI. CONCLUSIONS

We derived a probabilistic scan matching model which jointly describes the pose change of a high-resolution sensor along with the unknown data associations of observed scan points. Moreover, we presented an inference algorithm for that model which determines the posterior distribution of pose change given scan data. Compared to state-of-the-art scan matching algorithms, our approach considers all data association hypotheses and allows for uncertainty quantification, as the full posterior distribution is available. Numerical studies show precise pose change estimates characterized by low drift in terms of both translation and rotation.

### APPENDIX A

#### DERIVATION OF $f(\mathbf{z}, \mathbf{a}, \mathbf{b}, N_S | \Delta \mathbf{P})$

We make the following assumptions:

- A1) Source points,  $\mathbf{s}_H^{(j)}, \forall j$ , are either associable or non-associable. Every associable source point can be associated with at most one destination surface point and every destination surface point can be associated to at most one associable source point.
- A2) Valid source-to-destination associations according to A1) are uniformly distributed a priori.
- A3) Associability of destination points,  $\boldsymbol{\theta}$ , and the number of non-associable source points,  $N_{NA}$ , are statistically independent given the relative pose,  $\Delta \mathbf{P}$ .
- A4) The associability of destination surface points  $i$  and  $i', \forall i \neq i', \theta^{(i)}$  and  $\theta^{(i')}$ , are independent Bernoulli-distributed random variables given the relative pose  $\Delta \mathbf{P}$ . The probability of a destination surface point  $i$  to be associable ( $\theta^{(i)} = 1$ ) is  $0 \leq f_A(\Delta \mathbf{P}) < 1, \forall i$ .
- A5) The number of non-associable source points,  $N_{NA}$ , is Poisson-distributed with intensity  $\lambda_{NA}$ .
- A6) A non-associable source point,  $\mathbf{s}_H^{(j)}$ , is statistically independent of an associable source points,  $\mathbf{s}_H^{(j')}$ . This holds for all associable and non-associable points.
- A7) Given the relative pose  $\Delta \mathbf{P}$ , pseudo measurements,  $\mathbf{z}^{(i,j')}, \forall i, j'$ , corresponding to associable source points,  $\mathbf{s}_H^{(j')}, \forall j'$ , are independent of each other and distributed according to  $f(\mathbf{z}^{(i,j')} | \Delta \mathbf{P})$ .
- A8) Non-associable source points,  $\mathbf{s}_H^{(j)}, \forall j$ , are independent and identically distributed (iid) with PDF  $f_{NA}(\mathbf{s}_H^{(j)} | \Delta \mathbf{P}) \neq 0, \Delta \mathbf{P} \in SE(N_{\dim})$ .

As a preliminary step, we derive the conditional PDF  $f(\mathbf{z}, \mathbf{a}, N_S | \Delta \mathbf{P})$ . This is closely related to the derivation of the joint probabilistic data association (JPDA) filter [11]. We start by decomposing  $f(\mathbf{z}, \mathbf{a}, N_S | \Delta \mathbf{P})$  as

$$f(\mathbf{z}, \mathbf{a}, N_S | \Delta \mathbf{P}) = f(\mathbf{z} | \mathbf{a}, N_S, \Delta \mathbf{P}) f(\mathbf{a}, N_S | \Delta \mathbf{P}). \quad (13)$$

Let us next define the vector  $\boldsymbol{\theta} \triangleq [\theta^{(1)} \ \theta^{(2)} \ \dots \ \theta^{(N_D)}]^T$  that indicates which destination surface point  $i \in \{1, 2, \dots, N_D\}$  is associable ( $\theta^{(i)} = 1$ ) or non-associable ( $\theta^{(i)} = 0$ ). Note that  $\theta^{(i)} = 1$  implies  $\mathbf{a}^{(i)} \neq 0$  and  $\theta^{(i)} = 0$  implies  $\mathbf{a}^{(i)} = 0$ . Due to (3), the mapping between  $(\mathbf{a}, \boldsymbol{\theta}, N_{NA})$  and  $(\mathbf{a}, N_S)$  is thus deterministic as well as bijective and

the factor  $f(\mathbf{a}, N_S | \Delta \mathbf{P})$  in (13) can also be written as  $f(\mathbf{a}, \boldsymbol{\theta}, N_{NA} | \Delta \mathbf{P}) = f(\mathbf{a}, N_S | \Delta \mathbf{P})$ .

As a result of the chain rule and by using A3), we obtain

$$\begin{aligned} f(\mathbf{a}, N_S | \Delta \mathbf{P}) &= f(\mathbf{a}, \boldsymbol{\theta}, N_{NA} | \Delta \mathbf{P}) \\ &= f(\mathbf{a} | \boldsymbol{\theta}, N_{NA}, \Delta \mathbf{P}) f(\boldsymbol{\theta} | \Delta \mathbf{P}, N_{NA}) f(N_{NA} | \Delta \mathbf{P}) \\ &= f(\mathbf{a} | \boldsymbol{\theta}, N_{NA}, \Delta \mathbf{P}) f(\boldsymbol{\theta} | \Delta \mathbf{P}) f(N_{NA} | \Delta \mathbf{P}). \end{aligned} \quad (14)$$

Note that A4) implies

$$f(\boldsymbol{\theta} | \Delta \mathbf{P}) = \left( \prod_{i \in \mathcal{D}_a} f_A(\Delta \mathbf{P}) \right) \prod_{i' \notin \mathcal{D}_a} (1 - f_A(\Delta \mathbf{P})) \quad (15)$$

and from A5), we obtain

$$f(N_{NA} | \Delta \mathbf{P}) = \frac{\lambda_{NA}^{N_{NA}}}{N_{NA}!} e^{-\lambda_{NA}}. \quad (16)$$

Let us denote by  $\mathcal{A}$  the set of all vectors  $\mathbf{a}$  that do not violate assumption A1). From basic results of combinatorics and A2), it is straightforward to show that  $f(\mathbf{a} | \boldsymbol{\theta}, N_{NA}, \Delta \mathbf{P}) = N_{NA}! / N_S!$  for  $\mathbf{a} \in \mathcal{A}$  and thus  $f(\mathbf{a} | \boldsymbol{\theta}, N_{NA}, \Delta \mathbf{P}) = \Psi(\mathbf{a}) N_{NA}! / N_S!$  for arbitrary  $\mathbf{a} \in \{0, 1, \dots, N_S\}^{N_D}$ . Using this result as well as (15) and (16) in (14), we obtain

$$\begin{aligned} f(\mathbf{a}, N_S | \Delta \mathbf{P}) &= \Psi(\mathbf{a}) \frac{\lambda_{NA}^{N_S - |\mathcal{D}_a|}}{N_S!} e^{-\lambda_{NA}} \left( \prod_{i \in \mathcal{D}_a} f_A(\Delta \mathbf{P}) \right) \\ &\quad \times \prod_{j \notin \mathcal{D}_a} (1 - f_A(\Delta \mathbf{P})). \end{aligned} \quad (17)$$

Furthermore, using A6)-A8)  $f(\mathbf{z} | \mathbf{a}, N_S, \Delta \mathbf{P})$  in (13), can be expressed as

$$\begin{aligned} f(\mathbf{z} | \mathbf{a}, N_S, \Delta \mathbf{P}) &= \left( \prod_{j=1}^{N_S} f_{NA}(\mathbf{s}_H^{(j)} | \Delta \mathbf{P}) \right) \prod_{i \in \mathcal{D}_a} \frac{f(\mathbf{z}^{(i, \mathbf{a}^{(i)})} | \Delta \mathbf{P})}{f_{NA}(\mathbf{s}_H^{(\mathbf{a}^{(i)})} | \Delta \mathbf{P})}. \end{aligned} \quad (18)$$

Now, we use the expression for  $f(\mathbf{a}, N_S | \Delta \mathbf{P})$  in (17) and the expression for  $f(\mathbf{z} | \mathbf{a}, N_S, \Delta \mathbf{P})$  in (18) into (14) and obtain

$$\begin{aligned} f(\mathbf{z}, \mathbf{a}, N_S | \Delta \mathbf{P}) &= \Psi(\mathbf{a}) \frac{\lambda_{NA}^{N_S - |\mathcal{D}_a|}}{N_S!} e^{-\lambda_{NA}} \left( \prod_{j=1}^{N_S} f_{NA}(\mathbf{s}_H^{(j)} | \Delta \mathbf{P}) \right) \\ &\quad \times \left( \prod_{i \in \mathcal{D}_a} \frac{f_A(\Delta \mathbf{P}) f(\mathbf{z}^{(i, \mathbf{a}^{(i)})} | \Delta \mathbf{P})}{f_{NA}(\mathbf{s}_H^{(\mathbf{a}^{(i)})} | \Delta \mathbf{P})} \right) \prod_{j \notin \mathcal{D}_a} (1 - f_A(\Delta \mathbf{P})). \end{aligned} \quad (19)$$

Finally, the final expression for  $f(\mathbf{z}, \mathbf{a}, \mathbf{b}, N_S | \Delta \mathbf{P})$  in (8) is obtained from the expression for  $f(\mathbf{z}, \mathbf{a}, N_S | \Delta \mathbf{P})$  in (19) by “opening” the factor  $\Psi(\mathbf{a})$ . In particular, we introduce  $\mathbf{b}$  as described in [24, Section V.B] and replace  $\Psi(\mathbf{a})$  by  $\Psi(\mathbf{a}, \mathbf{b})$  in (5).

## REFERENCES

- [1] A. Ziebinski, R. Cupek, H. Erdogan, and S. Waechter, "A survey of ADAS technologies for the future perspective of sensor fusion," in *Computational Collective Intelligence*. Springer International Publishing, 2016, pp. 135–146.
- [2] A. Mallios, P. Ridao, D. Ribas, F. Maurelli, and Y. Petillot, "EKF-SLAM for AUV navigation under probabilistic sonar scan-matching," in *Proc. IEEE IROS-10*, Taipei, Taiwan, 2010, pp. 4404–4411.
- [3] P. Biber and W. Strasser, "The normal distributions transform: a new approach to laser scan matching," in *Proc. IEEE IROS-03*, Las Vegas, NV, USA, Oct. 2003.
- [4] J. Zhang and S. Singh, "Low-drift and real-time lidar odometry and mapping," *Auton. Robots*, vol. 41, no. 2, pp. 401–416, Feb. 2016.
- [5] R. A. Srivatsan, M. Xu, N. Zevallos, and H. Choset, "Probabilistic pose estimation using a Bingham distribution-based linear filter," *Int J. Robot Res.*, vol. 37, no. 13–14, pp. 1610–1631, Jun. 2018.
- [6] J.-E. Deschaud, "IMLS-SLAM: Scan-to-model matching based on 3D data," in *Proc. IEEE ICRA-18*, Brisbane, Australia, May 2018.
- [7] T. Shan and B. Englot, "LeGO-LOAM: Lightweight and ground-optimized lidar odometry and mapping on variable terrain," in *Proc. IEEE IROS-18*, Madrid, Spain, Oct. 2018.
- [8] J. Tang, Y. Chen, X. Niu, L. Wang, L. Chen, J. Liu, C. Shi, and J. Hyyp, "LiDAR scan matching aided inertial navigation system in GNSS-denied environments," *Sensors*, vol. 15, no. 7, pp. 16710–16728, Jul. 2015.
- [9] E. Olson, "Real-time correlative scan matching," in *2009 IEEE International Conference on Robotics and Automation*. IEEE, May 2009.
- [10] A. Mallios, P. Ridao, D. Ribas, and E. Hernández, "Scan matching SLAM in underwater environments," *Autonomous Robots*, vol. 36, no. 3, pp. 181–198, Jun. 2013.
- [11] Y. Bar-Shalom, P. K. Willett, and X. Tian, *Tracking and Data Fusion: A Handbook of Algorithms*. Storrs, CT, USA: Yaakov Bar-Shalom, 2011.
- [12] J. L. Williams and R. Lau, "Approximate evaluation of marginal association probabilities with belief propagation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 4, pp. 2942–2959, Oct. 2014.
- [13] F. Meyer and M. Z. Win, "Scalable data association for extended object tracking," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 6, pp. 491–507, 2020.
- [25] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*. New York, NY, USA: Springer, 2001.
- [14] F. Meyer and J. L. Williams, "Scalable detection and tracking of extended objects," in *Proc. IEEE ICASSP 2020*, May 2020, pp. 8916–8920.
- [15] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. MIT Press Ltd, 2005.
- [16] J. C. Carr, R. K. Beatson, J. B. Cherrie, T. J. Mitchell, W. R. Fright, B. C. McCallum, and T. R. Evans, "Reconstruction and representation of 3D objects with radial basis functions," in *Proc. ACM SIGGRAPH-01*, Los Angeles, CA, USA, Aug. 2001.
- [17] J. Demantké, C. Mallet, N. David, and B. Vallet, "Dimensionality based scale selection in 3D lidar point clouds," *ISPRS Archives*, vol. XXXVIII-5/W12, pp. 97–102, Sep. 2012.
- [18] K. Klasing, D. Althoff, D. Wollherr, and M. Buss, "Comparison of surface normal estimation methods for range sensing applications," in *Proc. IEEE ICRA-09*, Kobe, Japan, May 2009.
- [19] K.-L. Low, "Linear least-squares optimization for point-to-plane ICP surface registration," Department of Computer Science, University of North Carolina at Chapel Hill, Tech. Rep., 2004.
- [20] P. Richards, "Constrained Kalman filtering using pseudo-measurements," in *IEE CATT-95*, London, UK, May 1995.
- [21] H.-A. Loeliger, J. Dauwels, J. Hu, S. Korl, L. Ping, and F. R. Kschischang, "The factor graph approach to model-based signal processing," *Proc. IEEE*, vol. 95, no. 6, pp. 1295–1322, Jun. 2007.
- [22] H. Loeliger, "An introduction to factor graphs," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 28–41, Jan. 2004.
- [23] F. Kschischang, B. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, 2001.
- [24] F. Meyer, T. Kropfreiter, J. L. Williams, R. Lau, F. Hlawatsch, P. Braca, and M. Z. Win, "Message passing algorithms for scalable multitarget tracking," *Proc. IEEE*, vol. 106, no. 2, pp. 221–259, Feb. 2018.
- [26] N. Boumal, "An introduction to optimization on smooth manifolds," Published online, Aug 2020. [Online]. Available: <http://www.nicolasboumal.net/book>
- [27] L. Tchammi, C. Choy, I. Armeni, J. Gwak, and S. Savarese, "SEG-Cloud: Semantic segmentation of 3d point clouds," in *2017 International Conference on 3D Vision (3DV)*. IEEE, Oct. 2017.