Landmark-Based Localization Using Range Measurements: A Stochastic Geometry Perspective

Haozhou Hu, *Student Member, IEEE*, Harpreet S. Dhillon, *Fellow, IEEE*, R. Michael Buehrer, *Fellow, IEEE*

***Abstract*—Many modern wireless devices with accurate positioning needs have access to many vision sensors, such as a camera, radar, and Light Detection and Ranging (LiDAR). In numerous scenarios where wireless-based positioning is either inaccurate or unavailable, using information from vision sensors becomes highly desirable for determining the precise location of the wireless device. While localization utilizing vision information has been explored from different algorithmic perspectives, the underlying mathematical underpinnings of this problem space remain largely unexplored. Inspired by this, we develop a new analytical framework for vision-based localization in which error-free distance measurements in vision data are utilized to accurately determine the position of the target. Compared to wireless-based positioning, a notable differentiation of this approach is the inclusion of *non-unique* landmarks, such as lampposts, which may lack distinguishable features in the vision data. For instance, when the target is located close to a lamppost, it becomes challenging to precisely identify the specific lamppost (among several in the region) that is near the target. By assuming that the landmarks of various types follow a marked Poisson point process (PPP), we establish that three range measurements are sufficient for determining the correct combination of landmarks in a two-dimensional plane. When the number of measurements is less than three, there exists a potential for making errors in associating these range measurements with the corresponding landmark combination. We provide a mathematical characterization of this *probability of error*, which involves a novel joint distribution of key random variables.**

***Index Terms*—Landmark-based localization, localizability, stochastic geometry, Poisson point process.**

I. Introduction

Wireless-based positioning systems, such as the global navigation satellite system (GNSS), WiFi, and cellular networks, have revolutionized how we navigate and localize ourselves in the modern world. In order to obtain a precise position fix, these systems require measurements from a certain minimum number of *anchor* nodes. In the case of GNSS, this may not be possible in certain situations, such as in urban alleys and in the tunnels. On the other hand, since wireless networks (such as cellular networks) are deployed to optimize communication performance, getting line of sight measurements from a sufficient number of anchors is often challenging. Moreover, all these systems are susceptible to infrastructure failures and jamming attacks. Given these limitations, it is essential to investigate complementary methods for self-localization that rely on environmental information. This naturally leads to vision-based positioning, where the target estimates its location using data from the imaging or vision sensors, such as cameras, radars, and LiDAR systems.

The classical vision-based positioning is done by matching the features of its surroundings with location-labeled images, referred to as the *map*. Since it is natural to apply computer vision, and more recently deep learning, approaches to this setting, the problems often reduce to selecting and extracting features that express constraints among observations. These features are often hand-crafted depending upon the type of content and application, such as object-level content [1], geometric shapes [2], and visual vocabulary [3]. From the deep learning perspective, the problem is usually handled using a variety of convolutional neural networks, such as PoseNet [4], MapNet [5], and CamNet [6], that extract useful features to directly provide the target location. Two specific examples of works along these lines are [7], [8], where [7] focuses on reducing the candidate locations for the image and localizing the image of mountains by matching skylines computed from an elevation map, and [8] focuses on urban localization with distinctive landmarks using Scale Invariant Feature Transform.

While vision-based localization has seen major advances from the algorithmic perspective, the mathematical underpinnings of this problem have not been explored with the same depth as wireless-based positioning. In [9] and this paper, we initiate this study by formulating a rigorous mathematical problem of estimating the position of a target using information about geographical landmarks. These landmarks are different from *anchors* used in wireless-based positioning in the sense that they may not necessarily be unique. For example, without additional information, we may not necessarily know which *exact* lamppost the target is close to, as many may look similar. This differs significantly from wireless-based localization where anchors are unique and easily identifiable. Intuitively, being near a unique combination of landmarks, like a lamppost, bus stop, and bank, may suggest a location not replicated elsewhere in the city.

We assume that we have *range (or depth) measurements* from a set of landmarks, which can be obtained from the depth analysis of the images [10]. We focus on utilizing pairwise geometric constraints to localize the target and derive the corresponding error probability. We formulate this problem mathematically by using ideas from stochastic geometry. We first develop a statistical model in which landmarks of different types are modeled as a marked PPP. Assuming three perfect (error-free) ranges to landmarks, we rigorously establish that it is possible to almost surely identify the correct combination of landmarks that corresponds to these measurements. Further,

The authors are with Wireless@VT, Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, 24061, USA. Email: {haozhouhu, hdhillon, rbuehrer}@vt.edu. The support of the US NSF (Grants CNS-2107276 and CNS-2225511) is gratefully acknowledged.

if we have two range measurements, it is not possible to identify the correct combination of landmarks every time. For this scenario, we derive an analytical expression for the error probability in identifying the correct landmark combination. In addition to characterizing the final result, this analysis also exposes the underlying structure of the problem which is expected to inspire further investigation of this model.

II. System Model

In this section, we describe the system model, including different attributes of the landmarks, the map, and the problem statement. As the statistical approach developed in this paper represents a fundamentally new direction, this system model also stands as a notable contribution of our work.

A. Landmark Types

We categorize landmarks into different types, such that landmarks within the same type share similar appearances and are indistinguishable from one another. For example, using vision information, we may be able to determine that the target is near a lamppost, but pinpointing which exact lamppost (out of potentially many in a given region) may not be feasible. This is due to two interconnected factors: (a) the limited resolution of vision sensors, which might not allow us to distinguish between two landmarks that are similar in appearance, and (b) the prohibitive amount of prior information and computational resources are required to match the exact landmark seen in the visual information, which is not practically achievable for mobile devices.

B. Visibility Regions of Landmarks

The concept of a visibility region captures the fact that some types of landmarks, like lampposts, are visible only from short distances, while others, such as telephone towers on hilltops, can be seen from much farther away. We assume that each of the landmarks has an associated *visibility region*, denoted as $\mathbf{b}(\mathbf{x}, d_{\mathrm{v}})$, which is a ball with radius d_{v} centered at $\mathbf{x}$, where d_{v} is the *maximum visibility distance* and $\mathbf{x}$ is the landmark location. This notation assumes that the landmarks are *isotropic*, meaning they are visible from all directions. This simple step-function assumption for landmark visibility (landmark visible until a certain distance and not visible beyond that) is a reasonable first step to understanding the structure of the problem. We will explore more general visibility functions in the future work.

C. The Map

Using some side information, such as the IDs of the serving base stations, the target can be placed in a specific region, which we term the *Area of Interest (AOI)*. Note that this AOI could span tens of kilometers in some situations (such as when the target is connected to a Low Earth Orbit (LEO) satellite). For analytical purposes, we consider the AOI to be a circular area with radius d_{bs}, centered at $\mathbf{x}_0$, denoted as $S = \mathbf{b}(\mathbf{x}_0, d_{\mathrm{bs}})$. The map of AOI includes labels for all landmarks, each represented as a tuple $(\mathbf{x}_i, m_i)$, where

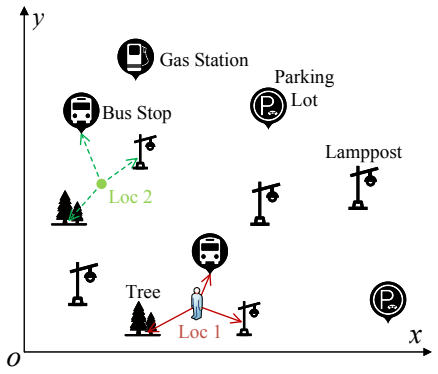


Fig. 1. An illustration of the map. There are five types of landmarks: lamppost, bus stop, parking lot, gas station, and tree. Three landmarks can be seen from Loc 1. However, we cannot uniquely localize the target since the exact same combination is visible from somewhere else in the map (Loc 2).

$\mathbf{x}_i \in \mathbb{R}^2$ denotes the location of the landmark and m_i is the landmark type. We consider that the maps of different places form realizations of the random set $\Phi = \{(\mathbf{x}_i, m_i)\}$, which can be naturally modeled as a marked point process. For the sake of analysis, we assume that Φ is a marked PPP. We restrict our attention to $\mathbb{R}^2$ because of its practical relevance, but our analysis can also be extended to $\mathbb{R}^n$.

D. Range Measurements

By utilizing data from imaging or vision sensors, the target can, in principle, determine the *depth* of each landmark detected. In the context of localization applications, we will refer to this depth as the *range measurement* of the landmark from the target, or simply, the *range*. To understand the structure of the problem, we assume that the *exact* ranges (i.e., free of noise or error) to visible landmarks are available at the target. Since landmarks of the same type are indistinguishable, the range measurement cannot be associated with a specific landmark, but simply be labeled by the corresponding landmark type. Therefore, the ranges and their corresponding types can be represented as a set of two-tuples $\mathcal{I} = \{(m_i, r_i)\}$, where r_i is the range measurement from one of the landmarks of type m_i. The noise and error in the measurements will be included in the journal extension of this work.

E. Problem Formulation

Assuming that the target is located within the AOI, our goal is to localize it by determining the correct set of landmarks that are seen in its vision data. We assume that the target measures ranges to a subset of visible landmarks, which are denoted as $\mathcal{I}$. Since the ranges are only associated with landmark types, the problem is equivalent to identifying the combination of correct observed landmarks c^* using type-labeled ranges $\mathcal{I}$ and a specific map φ, where φ represents a realization of the random set Φ. Since the actual geometric configuration of landmarks and the target is unknown, our interest lies in evaluating the performance across the whole map. Specifically, we aim to determine the probability of correctly identifying the

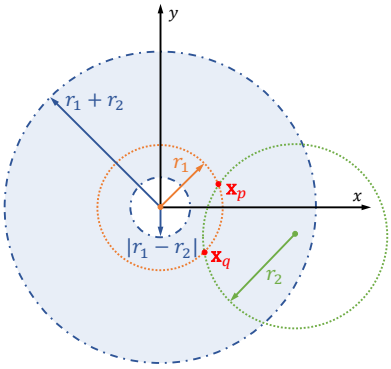


Fig. 2. An illustration of possible locations for the second landmark. The orange point represents the location of the first landmark, while the red points indicate the potential locations of the target. The blue annulus shows possible locations for the second landmark, such that the orange and green circles intersect.

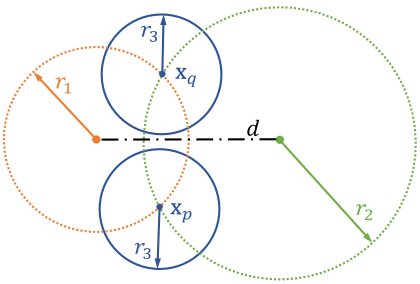


Fig. 3. An illustration of possible locations for the third landmark.

combination of landmarks using $\mathcal{I}$. This probability is defined as $\mathbb{P}(\hat{c} = c^* \mid N)$, where $N = |\mathcal{I}|$ is the number of ranges, c^* is the correct landmark combination, and $\hat{c}$ is the estimate of the landmark combination obtained from $\mathcal{I}$. We term this the *localizability probability* or simply localizability. In order to understand the technical challenges as well as the fundamental differences of this problem from more familiar wireless-based localization, please refer to Fig. 1, where the target is located at Loc 1 and can see three distinct landmarks. It turns out that the same combination of landmarks is also visible from a different location (Loc 2) on the map. Therefore, there is no way of uniquely determining the location of the target in this case. However, as is perhaps intuitive, as we detect more landmarks or the detected landmarks (or the combination of landmarks) are rare, our chances of uniquely determining the target location improve significantly.

From a mathematical perspective, our objective in this paper is to determine the exact combination of landmarks that generated the measurements in $\mathcal{I}$. In other words, our goal is to label each element of $\mathcal{I}$ with the correct landmark identification. Once accomplished, we can then unambiguously determine the location of the target on the map.

III. Identifying Correct Landmarks

To understand *localizability*, let us begin with the simplest case of a single measurement $\mathcal{I} = \{(m_1, r_1)\}$. This represents that a landmark of type m_1 is visible from the target location at a distance of r_1. As a result, the target is located on a circle with a radius of r_1 from one of the landmarks of type m_1. By definition, it is impossible to uniquely identify the specific landmark of type m_1 that is visible at the target. Therefore, it is not possible to localize the target in this case.

Now, let us suppose that we have two range measurements $\mathcal{I} = \{(m_1, r_1), (m_2, r_2)\}$. Our objective is to determine whether it is possible to uniquely map these measurements to the correct pair of landmarks c^*, which is visible from the target's location. The answer to this question is no, which can be understood through a simple geometric argument illustrated in Fig. 2. Let us consider the landmark that could potentially correspond to (m_1, r_1), and place it at the origin. Using the information in $\mathcal{I}$, we can constrain the location of the second landmark to an annulus centered at the origin, with an outer radius of $r_1 + r_2$ and an inner radius of $|r_1 - r_2|$. This comes from the triangle inequality constraint, as discussed in Section III-A. If we draw two circles, each centered at one of the two landmarks and with radii r_1 and r_2 respectively, the possible locations of the target are the two intersection points, $\mathbf{x}_p$ and $\mathbf{x}_q$. Since the measure of the possible locations for the second landmark is nonzero, there is a non-zero probability that any pair of landmark combinations can generate the given range measurements r_1 and r_2. Therefore, it is impossible to correctly identify the landmark combination with certainty using only two measurements. In Section IV, we explore the probability of incorrectly identifying the landmark pair.

However, if we assume that we have three error-free measurements, it is possible to uniquely identify the correct set of three landmarks corresponding to these measurements. In the following lemma, we prove that there is almost certainly no other combination of three landmarks (other than the correct one) capable of yielding the given three measurements.

Lemma 1. *Assuming that the landmarks form a marked PPP on $\mathbb{R}^2$, there is almost surely no combination of landmarks other than the current combination c^* that will lead to three error-free range measurements at any target location.*

Proof: Consider a pair of landmarks at a distance d that satisfies $|r_1 - r_2| \leq d \leq r_1 + r_2$, where r_1 and r_2 are two measurements corresponding to the first and second landmarks, respectively. As illustrated in Fig. 3, the target could be located at one of the intersection points, either $\mathbf{x}_p$ or $\mathbf{x}_q$. Now, if we were to position the third landmark at a distance of r_3 from the target, the potential locations for this landmark would be represented by the two blue circles in Fig. 3. Since the landmark locations are modeled as a PPP, the probability that the third landmark will land exactly on one of the two blue circles in Fig. 3 is zero. Therefore, in this scenario, the probability that combinations of three landmarks, other than the correct combination, could yield the specific measurements r_1, r_2, and r_3 is also zero. This concludes the proof. ■

Lemma 1 demonstrates that it is possible to identify the correct combination of visible landmarks using three error-free measurements. This result largely stems from two factors: the assumption of error-free measurements, and the positioning of landmarks as a PPP in continuous space.

From the algorithmic perspective, we can determine the correct combination by utilizing the geometric constraints between landmarks, which is the information readily available on the map. Such constraints can be the distance between different landmarks or their relative placement on the plane. In the rest of this section, we will delve into a specific pairwise constraint, the triangle inequality that appeared above.

A. The Triangle Inequality

When we view the target and two landmarks as vertices of a triangle, the lengths of the triangle's edges must satisfy the triangle inequality, given as

$$|r_i - r_j| \leq d_{ij} \leq r_i + r_j, \tag{1}$$

where $d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ is the distance between two landmarks, r_i and r_j are two range measurements. As depicted in Fig. 2, this constraint can be interpreted as follows: when we draw circles centered at two landmarks (x_i, y_i) and (x_j, y_j) with radii r_i and r_j, respectively, the two possible target locations are the intersection points of these circles. The triangle inequality needs to be satisfied for these circles to intersect. If we have more than two measurements in $\mathcal{I}$, this constraint must hold pairwise for any two landmarks selected from the correct landmark combination.

IV. Localizability Analysis

In the previous section, we thoroughly investigated the problem of identifying the correct set of landmarks visible from the target using three or more range measurements. However, there are numerous situations where we might only have access to two measurements. For instance, landmarks may be sparse in certain regions (such as rural areas), or it might simply be too time-consuming or resource-intensive to take additional measurements, even if more landmarks are visible. While Section III has already established that we cannot always identify the correct set of visible landmarks using just two range measurements, this section aims to characterize how often we can correctly identify these landmarks, with a specific focus on the probability of incorrectly identifying the landmark pair. Our approach relies on the triangle inequality to constrain the set of possible landmark combinations. When there are multiple candidate combinations, we select one uniformly at random as the solution.

Given two labeled range measurements, denoted as $\mathcal{I} = \{(m_i, r_i), (m_j, r_j)\}$, we next define two key entities: the *combination set* and the *solution set*.

Definition 1 (Combination Set)**.** *The combination set, denoted as $\mathcal{C}$, is defined as the set of all possible pairs of landmarks of corresponding types p and q, within the AOI. It can be mathematically written as follows:*

$$\mathcal{C} = \left\{ \{\mathbf{x}_i, \mathbf{x}_j\} \mid \forall \mathbf{x}_i \in \mathcal{B}_p, \mathbf{x}_j \in \mathcal{B}_q \right\}, \tag{2}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are the locations of two potential landmarks that correspond to the measured ranges, and $\mathcal{B}_p = \Phi_p \cap S$ and $\mathcal{B}_q = \Phi_q \cap S$ are the random sets of the landmark locations of type p and q, respectively, within the AOI S.

For each element in the combination set, there exists a corresponding distance $d_{ij} = |\mathbf{x}_i - \mathbf{x}_j|$ between two landmarks. We derive the distribution of these distances next.

Lemma 2. *For an arbitrary landmark combination $c \in \mathcal{C}$, the distribution of the distance between landmarks d_{ij} is*

$$f_D(s) = \frac{2s}{d_{\text{bs}}^2} - \frac{s^2\sqrt{4d_{\text{bs}}^2 - s^2}}{\pi d_{\text{bs}}^4} - \frac{4s \sin^{-1}\left(\frac{s}{2d_{\text{bs}}}\right)}{\pi d_{\text{bs}}^2}. \tag{3}$$

Proof: Since landmarks are modeled as a PPP, conditioned on the number of landmarks in AOI, they are distributed uniformly at random independently from each other in $S = \mathbf{b}(\mathbf{x}_0, d_{\text{bs}})$ [11]. Therefore, the calculation essentially reduces to determining the distribution of the distance between two points that are uniformly distributed at random within S. This is a well-established result, as referenced in [12]. ■

Remark 1. *The distribution of the distance $f_D(s)$ is independent of the size of the combination set. For any given pair of landmarks $c_{ij} \in \mathcal{C}$, the corresponding distance d_{ij} follows the same distribution $f_D(s)$.*

Since mapping from two range measurements to the correct combination of landmarks is not always possible, we first consider mapping to the set of all *eligible combinations* that satisfy the triangle inequality. Following this, we construct the *solution set*, as defined below.

Definition 2 (Solution Set)**.** *The solution set is a subset of the combination set, where all pairs of landmarks satisfy the triangle inequality, written as*

$$\mathcal{S} = \left\{ c_{ij} \mid c_{ij} \in \mathcal{C}, |r_i - r_j| \leq d_{ij} \leq r_i + r_j \right\}, \tag{4}$$

where $d_{ij} = |\mathbf{x}_i - \mathbf{x}_j|$ is the distance of two landmarks.

Remark 2. *In the noise-free case, the correct combination, denoted by $c^* = \{\mathbf{x}_i^*, \mathbf{x}_j^*\}$, will always be contained in the solution set. Therefore, the cardinality of $\mathcal{S}$ is lower bounded by one, i.e., $|\mathcal{S}| \geq 1$.*

When the size of the solution set $|\mathcal{S}| = 1$, we can directly determine the correct solution. However, when $|\mathcal{S}| \geq 2$, a unique solution is not guaranteed. In these situations, we select one solution uniformly at random from $\mathcal{S}$, denoted as $\hat{c} \in \mathcal{S}$. The probability of selecting an incorrect solution is

$$\mathbb{P}(\hat{c} \neq c^* \mid |\mathcal{S}| = k) = \frac{k-1}{k},\ k \in \mathbb{N}^+, \tag{5}$$

which is a direct result of selecting one solution uniformly and randomly.

Next, we consider the conditional probability related to the size of the solution set, as given in the following Lemma.

Lemma 3. *The size of the solution set depends on the range measurements $\mathbf{r}$, their corresponding types $\mathbf{m}$, and the size of*

the combination set $|\mathcal{C}|$. The associated conditional probability mass function (pmf) is as follows:

$$\mathbb{P}(|\mathcal{S}| = k \mid \mathbf{r}, \mathbf{m}, |\mathcal{C}| = n) = \binom{n-1}{k-1} p_c^{k-1} (1-p_c)^{n-k}, \quad (6)$$

where

$$p_c = \int_{|r_i - r_j|}^{r_i + r_j} f_D(s) \mathrm{d}s, \quad (7)$$

is the probability that any pair of landmarks satisfies the triangle inequality.

Proof: The size of the solution set is the number of landmark pairs that satisfy the triangle inequality. When given two range measurements $\mathbf{r} = [r_i, r_j]$, the probability that an arbitrary landmark pair $c_{ij} \in \mathcal{C} \backslash \{c^*\}$ satisfies the triangle inequality is

$$\mathbb{P}(I_{ij} \mid \mathbf{r}, \mathbf{m}) = \mathbb{E}_{D_{ij}}\{\mathbb{1}(|r_i - r_j| \leq D_{ij} \leq r_i + r_j)\} \quad (8)$$

$$= \int_{|r_i - r_j|}^{r_i + r_j} f_D(s) \, \mathrm{d}s, \quad (9)$$

where $I_{ij} = \mathbb{1}(|r_i - r_j| \leq D_{ij} \leq r_i + r_j)$ is a binary random variable that serves as an indicator of the event where the triangle inequality is satisfied by c_{ij}. We assume that for all landmark pairs $c_{ij} \in \mathcal{C}$, the corresponding indicator I_{ij} is independent and identically distributed. Under aforementioned conditioning, the size of $|\mathcal{S}|$ is the sum of the binary random variables I_{ij}, hence a binomial random variable with conditional pmf

$$\mathbb{P}(|\mathcal{S}| = k \mid \mathbf{r}, \mathbf{m}, |\mathcal{C}| = n)) \quad (10)$$

$$= \mathbb{P}\left(\sum_{c_{ij} \in \mathcal{C}} I_{ij} = k \mid \mathbf{r}, \mathbf{m}, |\mathcal{C}| = n \right) \quad (11)$$

$$= \mathbb{P}(I^* = 1 \mid \mathbf{r}, \mathbf{m}, |\mathcal{C}| = n) \times \mathbb{P}\left(\sum_{c_{ij} \in \mathcal{C} \backslash \{c^*\}} I_{ij} = k - 1 \mid \mathbf{r}, \mathbf{m}, |\mathcal{C}| = n \right) \quad (12)$$

$$= \binom{n-1}{k-1} \mathbb{P}(I = 1 \mid \mathbf{r}, \mathbf{m})^{k-1} \, \mathbb{P}(I \neq 1 \mid \mathbf{r}, \mathbf{m})^{n-k}, \quad (13)$$

where $I^* = 1$ is the indicator that the correct pair of landmarks satisfies the triangle inequality, while I is the indicator that any arbitrary pair of landmarks, excluding the correct one, satisfies the triangle inequality. This concludes the proof. ∎

Remark 3. *As noted above, $|\mathcal{S}| = 1$ will ensure unique identification of the correct landmark combination. The probability of this event can be obtained from Lemma 3 as*

$$\mathbb{P}(|\mathcal{S}| = 1 \mid \mathbf{r}, \mathbf{m}, |\mathcal{C}| = n) = (1 - p_c)^{n-1}. \quad (14)$$

As expected, we can infer from the above equation that when the size of combinations set n increases, the probability of obtaining the correct solution decreases.

We now mathematically characterize the error probability for the case with two observations, defined as

$$\mathbb{P}(\hat{c} \neq c^* \mid N = 2) = \mathbb{E}_{\mathbf{R}, \mathbf{M}, |\mathcal{C}|}\{\mathbb{P}(\hat{c} \neq c^* \mid \mathbf{r}, \mathbf{m}, |\mathcal{C}| = n)\}, \quad (15)$$

where N is the number of range measurements. Using the previous Lemmas, we derive this in the next Theorem.

Theorem 1. *The probability of making an error in identifying the correct pair of landmarks using two observations is*

$$\mathbb{P}(\hat{c} \neq c^* \mid N = 2) = \mathbb{E}_{\mathbf{R}, \mathbf{M}, |\mathcal{C}|}\left\{ 1 - \frac{1 - (1 - p_c)^{|\mathcal{C}|}}{|\mathcal{C}| \cdot p_c} \right\}. \quad (16)$$

Proof: For the setup described above, the probability of making an error depends upon the size of the solution set. Therefore, we can write this probability as follows:

$$\mathbb{P}(\hat{c} \neq c^* \mid \mathbf{r}, \mathbf{m}, |\mathcal{C}| = n)$$

$$\overset{(a)}{=} \sum_{k=1}^{n} \mathbb{P}(\hat{c} \neq c^* \mid |\mathcal{S}| = k) \, \mathbb{P}(|\mathcal{S}| = k \mid \mathbf{r}, \mathbf{m}, |\mathcal{C}| = n) \quad (17)$$

$$= 1 - \sum_{k=1}^{n} \frac{1}{k} \cdot \mathbb{P}(|\mathcal{S}| = k \mid \mathbf{r}, \mathbf{m}, |\mathcal{C}| = n), \quad (18)$$

where (a) follows from the law of total probability and the fact that the error probability is only dependent on the size of the solution set.

Using the result from Lemma 3, we get

$$\mathbb{P}(\hat{c} \neq c^* \mid \mathbf{r}, \mathbf{m}, |\mathcal{C}| = n) = 1 - \frac{1 - (1 - p_c)^n}{n \cdot p_c}. \quad (19)$$

Now, using the definition in (15) completes the proof. ∎

Equation (16) provides an analytical expression for the probability of making an error. This probability is an expectation over range measurements $\mathbf{r}$, each with their corresponding types $\mathbf{m}$, and the size of the combination set $|\mathcal{C}|$. To obtain the average probability of error across the entire region, or equivalently, over the whole point process of landmarks, we need the joint distribution of these three random variables, which we will derive in the next section.

V. The Joint Distribution

In Section IV, we considered two landmarks and utilized the pairwise constraint to find the target location. The error probability is expressed in the form of an expectation over three key parameters: the range measurement $\mathbf{R}$ with its corresponding types $\mathbf{M}$ and the size of the combination set $|\mathcal{C}|$. In order to determine the joint distribution of these parameters, we need to specify how the two landmarks of interest were chosen from amongst the visible landmarks. For this analysis, we assume that the two landmarks (and consequently the two ranges) are selected independently and uniformly at random from the visibility region. Further, we consider m types of landmarks with different visibility distances $\{d_{\mathrm{v}_1}, d_{\mathrm{v}_2}, \ldots, d_{\mathrm{v}_m}\}$. To simplify the notation, we denote the PPP of the p^{th} type of landmarks as Φ_p, with a density of λ_p. Under this setup, the joint distribution can be expressed as

$$f_{\mathbf{R}, \mathbf{M}, |\mathcal{C}|}(\mathbf{r}, \mathbf{m}, n) = \mathbb{P}(\mathbf{M} = \mathbf{m}) \cdot f_{\mathbf{R}|\mathbf{M}}(\mathbf{r} \mid \mathbf{m}) \cdot \mathbb{P}(|\mathcal{C}| = n \mid \mathbf{m}), \quad (20)$$

which follows from the fact that the two landmarks are selected uniformly at random from the visibility region. Therefore, the

joint distribution of range measurements is independent of the size of $\mathcal{C}$ when conditioned on the landmark type. In the rest of this section, we derive expressions for each of the three terms in the expression above.

Lemma 4. *The marginal distribution of landmark types that correspond to range measurements is*

$$\mathbb{P}(\mathbf{M}=\mathbf{m}) = \frac{\Lambda_p\big(S_{\mathrm{v}_p}\big)\cdot\Lambda_q\big(S_{\mathrm{v}_q}\big)}{\big(\sum_{i=1}^{m}\Lambda_i(S_{\mathrm{v}_i})\big)^2}, \tag{21}$$

where $\mathbf{m}=[p,q]$ *are two different landmark types, and* $\Lambda_p\big(S_{\mathrm{v}_p}\big)=\lambda_p\pi d_{\mathrm{v}_p}^2$.

Proof: Because the ranges are obtained from two landmarks selected uniformly at random from the visibility region, the two observations (and hence their types) are independent,

$$\mathbb{P}(\mathbf{M}=\mathbf{m}) = \mathbb{P}(M_1=p)\,\mathbb{P}(M_2=q)\,. \tag{22}$$

Therefore, to obtain the joint probability in (21), it is sufficient to characterize the marginal distribution of one of the landmark types, for example, $\mathbb{P}(M_1=p)$. Since the visibility distances of landmarks vary among different types, we scale the distance to $\mathbf{x}_0$ for all point processes $\Phi_1,\ldots,\Phi_m$, such that visible landmarks of different types are mapped onto a unit circle, denoted as $\mathbf{b}(\mathbf{x}_0,1)$. The scaled (or transformed) PPP of the p^{th} type of landmarks $\tilde{\Phi}_p$, has density $\lambda_p'=\lambda_p d_{\mathrm{v}_p}^2$. Then, equivalently, we can select marked points uniformly at random from $\tilde{\Phi}=\bigcup_{p=1}^{m}\tilde{\Phi}_p$. The probability of selecting a point of a particular type, which is independent of their locations, is given by

$$\mathbb{P}(M_1=p) = \frac{\Lambda_p\big(S_{\mathrm{v}_p}\big)}{\sum_{i=1}^{m}\Lambda_i(S_{\mathrm{v}_i})}. \tag{23}$$

We can derive $\mathbb{P}(M_2=q)$ in the same way and this completes the proof. ∎

Lemma 5. *The conditional joint distribution of distance measurements is* $f_{\mathbf{R}|\mathbf{M}}(\mathbf{r}\mid\mathbf{m})=$

$$\frac{2r_1\mathbb{1}_{r_1}((0,d_{\mathrm{v}_p}])}{d_{\mathrm{v}_p}^2}\cdot\frac{2r_2\mathbb{1}_{r_2}((0,d_{\mathrm{v}_q}])}{d_{\mathrm{v}_q}^2}. \tag{24}$$

Proof: Because of the independence of the two observations, it is again sufficient to consider the marginal distribution of range values to obtain (24).

When conditioned on the type of landmarks, the corresponding visibility distance is known. The location of the selected landmark in type p is uniformly distributed in the circle with radius d_{v_p}, which gives

$$f_{R_1|M_1}(r_1\mid p) = \frac{2r_1\mathbb{1}_{r_1}((0,d_{\mathrm{v}_p}])}{d_{\mathrm{v}_p}^2}. \tag{25}$$

This is a straightforward consequence of the PPP assumption of the landmark locations. ∎

Lemma 6. *The conditional probability of the size of the combination set is* $\mathbb{P}(|\mathcal{C}|=n\mid\mathbf{M}=\mathbf{m})=$

$$\begin{aligned}\sum_{n_1n_2=n}\Bigg\{&\sum_{i_1=1}^{n_1}\frac{c_p^{n_1-i_1}b_p^{i_1}(b_p-a)^{-i_1}}{(n_1-i_1)!}\cdot\\&\frac{\Gamma(i_1)-\Gamma(i_1,b_p-a)}{\Gamma(i_1)}\exp(-a-c_p)\Bigg\}\cdot\\\Bigg\{&\sum_{i_2=1}^{n_2}\frac{c_q^{n_2-i_2}b_p^{i_2}(b_q-a)^{-i_2}}{(n_2-i_2)!}\cdot\\&\frac{\Gamma(i_2)-\Gamma(i_2,b_p-a)}{\Gamma(i_2)}\exp(-a-c_q)\Bigg\},\end{aligned} \tag{26}$$

where p *and* q *are different landmark types,* $a=\sum_{i=1}^{m}\Lambda_i(S_{\mathrm{v}_i})$, $b_p=\Lambda_p\big(S_{\mathrm{v}_p}\big)$ *and* $c_p=\Lambda_p\big(S\backslash S_{\mathrm{v}_p}\big)=\lambda_p\pi\Big(d_{\mathrm{bs}}^2-d_{\mathrm{v}_p}^2\Big)$.

Proof: When two distance measurements are of different types, the size of the combination set is

$$|\mathcal{C}| = N(\mathcal{B}_p)\cdot N(\mathcal{B}_q)\,, \tag{27}$$

where $N(\mathcal{B}_p)$ and $N(\mathcal{B}_q)$ are the number of landmarks of type p and q in the AOI S, respectively. Since distance measurements are independent, the numbers of the landmarks in type p and q, respectively, are independent, which gives

$$\begin{aligned}&\mathbb{P}(N(\mathcal{B}_p)=n_1,N(\mathcal{B}_q)=n_2\mid\mathbf{M}=\mathbf{m})=\\&\mathbb{P}(N(\mathcal{B}_p)=n_1\mid M_1=p)\,\mathbb{P}(N(\mathcal{B}_q)=n_2\mid M_2=q)\,.\end{aligned} \tag{28}$$

Now, using the total probability law, the conditional distribution of $N(\mathcal{B}_p)$ is

$$\begin{aligned}&\mathbb{P}(N(\mathcal{B}_p)=n_1\mid M_1=p)\\&=\sum_{i_1=1}^{n_1}\Big\{\mathbb{P}\big(N(\mathcal{B}_p)=n_1\mid M_1=p,N\big(\mathcal{B}_p^{\mathrm{V}}\big)=i_1\big)\\&\qquad\cdot\mathbb{P}\big(N\big(\mathcal{B}_p^{\mathrm{V}}\big)=i_1\mid M_1=p\big)\Big\},\end{aligned} \tag{29}$$

where $\mathcal{B}_p^{\mathrm{V}}=\Phi_p\cap\mathbf{b}\big(\mathbf{x},d_{\mathrm{v}_p}\big)$. The first term denotes the conditional probability of the presence of landmarks of type p within the AOI, given that $N\big(\mathcal{B}_p^{\mathrm{V}}\big)=i_1$ landmarks of type p already exist within the visibility region. The second term represents the probability associated with the number of landmarks of type p present within the corresponding visibility region $\mathbf{b}(\mathbf{x}_0,d_{\mathrm{v}_p})$. The first term in (29) can be further expressed as

$$\begin{aligned}&\mathbb{P}\big(N(\mathcal{B}_p)=n_1\mid M_1=p,N\big(\mathcal{B}_p^{\mathrm{V}}\big)=i_1\big)\\&=\mathbb{P}\big(N\big(\mathcal{B}_p\backslash\mathcal{B}_p^{\mathrm{V}}\big)=n_1-i_1\mid M_1=p\big)\qquad(30)\\&=\frac{c_p^{n_1-i_1}}{(n_1-i_1)!}\exp(-c_p)\,,\qquad(31)\end{aligned}$$

where $c_p=\Lambda\big(S\backslash S_{\mathrm{v}_p}\big)=\lambda_p\pi\Big(d_{\mathrm{bs}}^2-d_{\mathrm{v}_p}^2\Big)$. This is a direct consequence of the PPP assumption [11].

Now, to derive the second term in (29), we consider the joint probability that the type of the selected landmark is p, the number of landmarks of type p in the visibility region is

$N(\mathcal{B}_p^{\mathrm{V}})$, and the total number of landmarks in the visibility region is $N(\mathcal{B}^{\mathrm{V}})$. First, $N(\mathcal{B}^{\mathrm{V}})$ is written as

$$N(\mathcal{B}^{\mathrm{V}}) = \sum_{i=1}^{m} N(\Phi_i \cap \mathbf{b}(\mathbf{x}_0, d_{\mathrm{v}_i})) \tag{32}$$

$$= \sum_{i=1}^{m} N\Big(\tilde{\Phi}_i \cap \mathbf{b}(\mathbf{x}_0, 1)\Big) \tag{33}$$

$$= N\Big(\tilde{\Phi} \cap \mathbf{b}(\mathbf{x}_0, 1)\Big), \tag{34}$$

where $\tilde{\Phi}_i$ and $\tilde{\Phi}$ are the transformed PPPs defined in the proof of Lemma 4. Then, we can write the joint probability as

$$\begin{aligned} &\mathbb{P}\big(M_1 = p, N(\mathcal{B}_p^{\mathrm{V}}) = i_1, N(\mathcal{B}^{\mathrm{V}}) = j\big) \\ &= \mathbb{P}\big(M_1 = p \mid N(\mathcal{B}_p^{\mathrm{V}}) = i_1, N(\mathcal{B}^{\mathrm{V}}) = j\big) \\ &\quad \cdot \mathbb{P}\big(N(\mathcal{B}_p^{\mathrm{V}}) = i_1, N(\mathcal{B}^{\mathrm{V}}) = j\big), \end{aligned} \tag{35}$$

where the first component of (35) is the probability that the selected landmark is of type p, given that we select (uniformly at random) one landmark out of $N(\mathcal{B}_p^{\mathrm{V}})$ visible landmarks, which consist of $N(\mathcal{B}^{\mathrm{V}})$ landmarks in type p, written as

$$\mathbb{P}\big(M_1 = p \mid N(\mathcal{B}_p^{\mathrm{V}}) = i_1, N(\mathcal{B}^{\mathrm{V}}) = j\big) = \frac{i_1}{j}. \tag{36}$$

The second component of (35) is the joint probability of numbers of the visible landmarks, which is

$$\begin{aligned} &\mathbb{P}\big(N(\mathcal{B}_p^{\mathrm{V}}) = i_1, N(\mathcal{B}^{\mathrm{V}}) = j\big) \\ &\overset{(a)}{=} \mathbb{P}\big(N(\mathcal{B}^{\mathrm{V}} \backslash \mathcal{B}_p^{\mathrm{V}}) = j - i_1 \mid N(\mathcal{B}_p^{\mathrm{V}}) = i_1\big)\, \mathbb{P}(N(\mathcal{B}_p^{\mathrm{V}}) = i_1) \\ &= \frac{(a - b_p)^{j-i_1}}{(j-i_1)!} \exp(-a + b_p) \cdot \frac{b_p^{i_1}}{i_1!} \exp(-b_p) \cdot, \end{aligned} \tag{37}$$

where $a = \sum_{i=1}^{m} \Lambda_i(S_{\mathrm{v}_i})$ and $b_p = \Lambda_p(S_{\mathrm{v}_p})$.

Now, we have the joint distribution of M_1, $N(\mathcal{B}_p^{\mathrm{V}})$ and $N(\mathcal{B}^{\mathrm{V}})$. We can derive the expression of the second component of (29), written as

$$\begin{aligned} &\mathbb{P}\big(N(\mathcal{B}^{\mathrm{V}}) = i_1 \mid M_1 = p\big) \\ &= \sum_{j=i_1+1}^{\infty} \frac{\mathbb{P}\big(M_1 = p, N(\mathcal{B}_p^{\mathrm{V}}) = i_1, N(\mathcal{B}^{\mathrm{V}}) = j\big)}{\mathbb{P}(M_1 = p)} \end{aligned} \tag{38}$$

$$= \frac{a b_p^{i_1} \left(\Gamma(i_1) - \Gamma(i_1, b_p - a)\right)}{b_p (b_p - a)^{i_1} \Gamma(i_1)} \exp(-a). \tag{39}$$

Then, equation (29) can be summarized as

$$\begin{aligned} &\mathbb{P}(N(\mathcal{B}_p) = n_1 \mid M_1 = p) \\ &= \sum_{i_1=1}^{n_1} \Bigg\{ \frac{c_p^{n_1 - i_1} b^{i_1}}{(b_p - a)^{i_1} (n_1 - i_1)!} \\ &\qquad \cdot \frac{\Gamma(i_1) - \Gamma(i_1, b_p - a)}{\Gamma(i_1)} \exp(-a - c_p) \Bigg\}. \end{aligned} \tag{40}$$

We can derive $\mathbb{P}(N(\mathcal{B}_q) = n_2 \mid M_2 = q)$ using the same method. Because of the independence of $N(\mathcal{B}_p)$ and $N(\mathcal{B}_q)$, we can construct the distribution of $|\mathcal{C}|$ by easily taking the sum. This completes the proof. ■

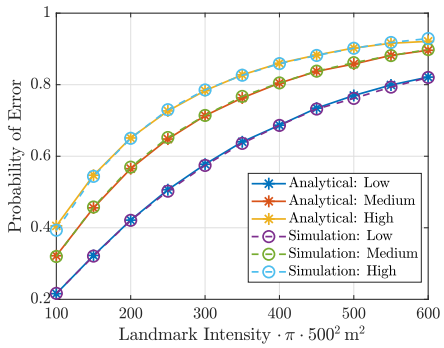


Fig. 4. The probability of making an error vs landmark densities.

Using the result from Lemmas 4, 5, 6, we obtain the joint distribution of range measurements $\mathbf{R}$, corresponding types $\mathbf{M}$ and the size of combination set $|\mathcal{C}|$, given next.

Theorem 2. *The joint distribution of* $\mathbf{R}$, $\mathbf{M}$, $|\mathcal{C}|$ *is given as*

$$\begin{aligned} f_{\mathbf{R},\mathbf{M},|\mathcal{C}|}(\mathbf{r}, \mathbf{m}, n) &= \frac{4 r_1 r_2 \lambda_p \lambda_q \mathbb{1}_{r_1}((0, d_{\mathrm{v}_p}]) \mathbb{1}_{r_2}((0, d_{\mathrm{v}_p}])}{a^2} \\ &\cdot \sum_{n_1 n_2 = n} \Bigg\{ \Bigg\{ \sum_{i_1=1}^{n_1} \frac{c_p^{n_1 - i_1} b_p^{i_1}}{(b_p - a)^{i_1} (n_1 - i_1)!} \\ &\qquad \cdot \frac{\Gamma(i_1) - \Gamma(i_1, b_p - a)}{\Gamma(i_1)} \exp(-a - c_p) \Bigg\} \\ &\cdot \Bigg\{ \sum_{i_2=1}^{n_2} \frac{c_q^{n_2 - i_2} b_q^{i_2}}{(b_q - a)^{i_2} (n_2 - i_2)!} \\ &\qquad \cdot \frac{\Gamma(i_2) - \Gamma(i_2, b_q - a)}{\Gamma(i_2)} \exp(-a - c_q) \Bigg\} \Bigg\}. \end{aligned} \tag{41}$$

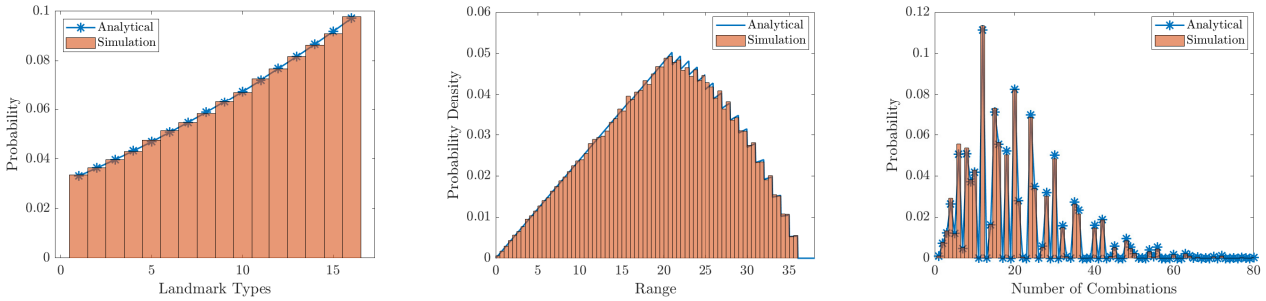
This theorem along with the result in Theorem 1 completely characterizes the probability of making an error in identifying the correct landmark pairs, which is one of the key technical contributions of this paper.

VI. Simulation Results

Since the analytical results of Section III ensure unique identification of the visible landmarks in the presence of three or more observations, we will focus on the case of two observations in this Section. Our simulation results will provide sanity checks for the analytical results of Sections IV and V as well as additional system design insights.

The simulation scenario is inspired by landmarks found in the city. We consider the AOI with a radius of 500 m centered at the origin, where there are 16 different types of landmarks, each conforming to a PPP, with the same density λ. We consider different visibility distances as per the following three scenarios: *low* (ranging from 10 to 25 m), *medium* (from 20 to 35 m), and *high* (from 30 to 45 m). The target location $\mathbf{x}$ is placed within AOI, where at least two landmarks are visible.

First, for $\lambda = \frac{350}{\pi \cdot 500^2}$ and medium visibility distances, we plot the marginal distributions of $\mathbf{R}$, $\mathbf{M}$ and $|\mathcal{C}|$ in Figs. 5(a), 5(b), and 5(c), respectively. Since plotting their joint



(a) $\mathbb{P}(M = p)$ (b) $f_R(r)$ (c) $\mathbb{P}(|\mathcal{C}| = n)$

Fig. 5. The marginal distribution of $\mathbf{R}$, $\mathbf{M}$ and $|\mathcal{C}|$ corresponds to visibility distances of 16 types ranging from 21 to 36. The density λ is set at 350.

distribution is challenging, the comparison of these marginal distributions with their analytical counterparts from Section V serves as a sanity check for our analysis. As expected, Fig. 5(a) demonstrates that landmarks with greater visibility distances are more likely to be seen by the target. Furthermore, the probability density of the range measurements exhibits an intriguing *stepped* pattern due to the different visibility distances considered for different types. The probability mass function of the number of landmark combinations in Fig. 5(c) fluctuates significantly, while still showing a close match between the analytical and simulation results. This is because the number of combinations is determined by the multiplication of $N(\mathcal{B}_p)$ and $N(\mathcal{B}_q)$, both of which are integers.

Next, we vary the landmark density λ from $\frac{100}{\pi \cdot 500^2}$ to $\frac{600}{\pi \cdot 500^2}$ and simulate the probability of error. As Fig. 4 shows, the simulation matches with the analytical result derived in Theorem 1 with which we used the joint probability result from Theorem 2. This serves as a vital validation for our analyses in both Sections IV and V. It is worth noting that as the visibility distances decrease, the probability of error also reduces. This is to be expected, as a landmark with a smaller visibility region will restrict the target's location more tightly, hence providing more precise information than a landmark with a larger visibility region, assuming their densities are the same. Interestingly, we observe that lower landmark densities result in notably smaller error probabilities. This suggests the feasibility of localization using landmarks even in environments where we are restricted to only a few number of measurements.

VII. Conclusions

This paper developed a fundamentally new framework for landmark-based localization, leveraging stochastic geometry. The framework aims to determine the target's position using error-free depth/distance measurements to landmarks. These measurements are obtained from vision sensors such as cameras, radars, and LiDAR. Given that the landmarks detected in vision data may not necessarily be unique, this approach necessitates a distinct mathematical treatment compared to traditional wireless-based localization. Under the assumption that the landmarks form a marked PPP, we demonstrated that three error-free measurements are sufficient to accurately identify the set of visible landmarks corresponding to those measurements. For instances with two observations, we derived the probability of error when identifying the correct landmark combination. This new perspective presented in this paper opens up several potential avenues for further research. Two specific extensions related to this work include: (i) integrating a suitable noise (or error) model for the measurements in the analysis, and (ii) incorporating additional methods to select a limited set of observations from a larger set of visible landmarks. Overall, this paper connects stochastic geometry, localization, and computer vision, which could inspire a completely new direction of investigation.

References

[1] J. Civera, D. Gálvez-López, L. Riazuelo, J. D. Tardós, and J. M. M. Montiel, "Towards semantic SLAM using a monocular camera," in *2011 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2011.

[2] J. Engel, T. Schöps, and D. Cremers, "LSD-SLAM: Large-scale direct monocular slam," in *European Conf. on Computer Vision*, 2014, pp. 834–849.

[3] T. Sattler, B. Leibe, and L. Kobbelt, "Fast image-based localization using direct 2d-to-3d matching," in *Int. Conf. on Computer Vision*, 2011, pp. 667–674.

[4] A. Kendall, M. Grimes, and R. Cipolla, "Posenet: A convolutional network for real-time 6-dof camera relocalization," in *IEEE Int. Conf. on Computer Vision*, 2015, pp. 2938–2946.

[5] S. Brahmbhatt, J. Gu, K. Kim, J. Hays, and J. Kautz, "Geometry-aware learning of maps for camera localization," in *IEEE Conf. on Computer Vision and Pattern Recognition*, 2018, pp. 2616–2625.

[6] M. Ding, Z. Wang, J. Sun, J. Shi, and P. Luo, "Camnet: Coarse-to-fine retrieval for camera re-localization," in *IEEE Int. Conf. on Computer Vision*, 2019, pp. 2871–2880.

[7] G. Baatz, O. Saurer, K. Köser, and M. Pollefeys, "Large scale visual geo-localization of images in mountainous terrain," in *European Conf. on Computer Vision*, 2012, pp. 517–530.

[8] G. Schindler, M. Brown, and R. Szeliski, "City-scale location recognition," in *IEEE Conf. on Computer Vision and Pattern Recognition*, 2007.

[9] H. Hu, H. S. Dhillon, and R. M. Buehrer, "Stochastic geometry analysis of localizability in vision-based geolocation systems," in *Proc., IEEE Asilomar*, Oct. 2023.

[10] M. Berger, A. Tagliasacchi, L. M. Seversky, P. Alliez, G. Guennebaud, J. A. Levine, A. Sharf, and C. T. Silva, "A survey of surface reconstruction from point clouds," in *Computer Graphics Forum*, vol. 36, no. 1. Wiley Online Library, 2017, pp. 301–329.

[11] M. Haenggi, *Stochastic geometry for wireless networks*. Cambridge University Press, 2012.

[12] S. Lellouche and M. Souris, "Distribution of distances between elements in a compact set," *Stats*, vol. 3, no. 1, pp. 1–15, 2019.