Circumventing the Feature Association Problem in SLAM

Martin Adams, John Mullane[†], Ba-Ngu Vo[‡]

Universidad de Chile, Santiago, Chile

[†]*Projective Space, Singapore*

[‡]*Curtin University, Australia*

martin@ing.uchile.cl, [†]*jmullane@projective-space.com,* [‡]*ba-ngu.vo@curtin.edu.au*

Abstract

In autonomous applications, a vehicle requires reliable estimates of its location and information about the world around it. To capture prior knowledge of the uncertainties in a vehicle's motion response to input commands and sensor measurements, this fundamental task has been cast as probabilistic Simultaneous Localization and Map building (SLAM). SLAM has been investigated as a stochastic filtering problem in which sensor data is compressed into features, which are consequently stacked in a vector, referred to as the map.

Inspired by developments in the tracking literature, recent research in SLAM has recast the map as a Random Finite Set (RFS) instead of a random vector, with huge mathematical consequences. With the application of recently formulated Finite Set Statistics (FISST), such a representation circumvents the need for fragile feature management and association routines, which are often the weakest component in vector based SLAM algorithms.

This tutorial demonstrates that true sensing uncertainty lies not only in the spatial estimates of a feature, but also in its existence. This gives rise to sensor probabilities of detection and false alarm, as well as spatial uncertainty values. By re-addressing the fundamentals of SLAM under an RFS framework, it will be shown that it is possible to estimate the map in terms of true feature number, as well as location. The concepts are demonstrated with short range radar, which detects multiple features, but yields many false measurements. Comparison of vector, and RFS SLAM algorithms shows the superior robustness of RFS based SLAM to such realistic sensing defects.

Index Terms

Autonomous Navigation, Random Finite Set (RFS), SLAM, robotic mapping, short range radar, Finite Set Statistics (FISST).

I. INTRODUCTION

Due to the imperfect nature of all real sensors, the lack of predictability in real environments and the necessary approximations to achieve computational decisions, robotics is a science which depends on probabilistic algorithms. Autonomous vehicles have now evolved beyond tasks within factory environments, and therefore require a robust computational representation of their complex surroundings and their related uncertainties. Such a representation is referred to as the map. Also of crucial importance to any autonomous task is the vehicle's positional estimate. In indoor and under water environments, GPS does not function, and in urban areas, satellite outages and multi-path reflections result in unacceptable GPS positional errors for autonomous vehicle applications. Therefore, in many applications, the fundamental necessity to autonomously acquire an environmental map, and infer vehicle location exists, which has been referred to as the "Holy grail" of autonomous robotics research - *Simultaneous Localization and Map Building* (SLAM) [1].

This tutorial demonstrates that the commonly used vector based methods for feature map representation in SLAM, summarized in Figure 1, suffer many fundamental disadvantages when applied to realistic situations. These correspond to situations in which
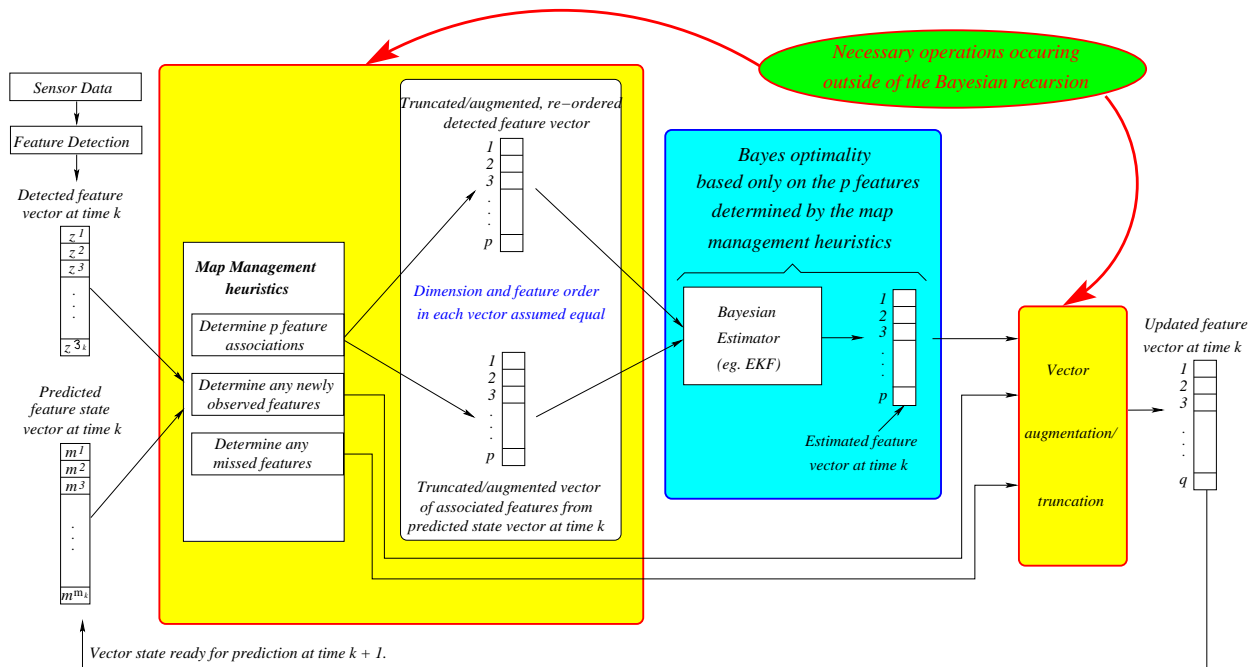

Fig. 1: *Vector based feature mapping and its mathematical restrictions, for a single update cycle, in SLAM.*

an a-priori unknown number of features are to be estimated, in the presence of realistic sensor and/or feature detection defects such as missed detections and false alarms, as well as spatial sensing errors. The tutorial therefore reviews the basic estimation principles of SLAM and, stemming from developments in the tracking community [2], summarizes recent SLAM investigations, which suggest that a landmark map is more appropriately represented as a *set* of landmarks, requiring the tools of Random Finite Set (RFS) theory, known as Finite Set Statistics (FISST) [3], [4]. An RFS is a random variable that take values as finite sets. It is defined by a discrete distribution that, importantly, characterizes the number of elements in the set, and a family of joint distributions which characterize the distribution of the element's values, conditioned on the number [5].

While defining a vector or set valued feature map representation may appear to be a trivial case of terminology, it will be demonstrated that the vector valued feature map has numerous mathematical consequences, namely an inherent rigid ordering of variables and a fixed length. This has caused robotics researchers to resort to SLAM state vector re-ordering and augmentation methods based on fragile feature management rules and association methods, as high-lighted in the yellow boxes (Figure 1), which necessarily occur before and after the SLAM Bayesian recursion. Bayes optimality can only take place with the p map states, which are considered to exist according to the external map management heuristics/filters. The RFS on the other hand, offers a powerful map representation, which will be shown to naturally encapsulate all possible permutations of a feature map and measurements thus circumventing the necessity of fragile feature management and association heuristics/methods. This results in more robust SLAM estimates in the presence of realistic sensor defects and clutter[1].

---

[1]Returned sensor readings incorrectly interpreted as useful features.

The SLAM experiments presented in this tutorial rely on short range, millimeter wave (MMW) radar as the exteroceptive sensor. Contrary to the spatial modelling school of thought primarily adopted by robotic engineers, to radar engineers, received power values at successive range, bearing, and possibly elevation, values are related to the possible *existence* of an object in space. Hence, uncertainty in the sensing process is primarily formulated in terms of the decision itself as to whether or not a landmark even exists at the range, bearing/elevation coordinates in question. Such an analysis usually gives rise to the well known concepts (in the radar domain) of probabilities of detection and false alarm, and significantly less emphasis, if any, is placed on the possible spatial uncertainty of a hypothesized landmark. Therefore, this tutorial will examine various measurement models, which are applied to range finding sensors. The fundamental difference between *detection* and *spatial* error modelling will also be shown, together with the unique ability of the RFS concept to capture both error models in a joint manner.

For comparison purposes both FISST based SLAM and state of the art vector based SLAM techniques are compared in an urban environment.

II. RELATED WORK & MOTIVATION

A. Prior Work in SLAM

There is a plethora of work in SLAM, with the origins of the feature based map representation being traced back to the seminal work of Smith *et. al.* [6]. This inspired a Gaussian approximation to Bayes theorem in the form of the Extended Kalman Filter (EKF) for the propagation of the SLAM state, and its assumed Gaussian, spatial uncertainty. An extension to this philosophy was advocated in [7], in which a Gaussian Mixture (GM) implementation of Bayes theorem was applied to SLAM. Non-parametric approximations to Bayes theorem followed with a Rao-Blackwellized (RB) particle filter yielding a Factored Solution To SLAM (FastSLAM) [8]. Further approximations to Bayesian SLAM have adopted Graph Based methods [9] and the Sparse Extended Information Filter (SEIF) [10].

An RFS formulation for SLAM was first proposed in [11] with preliminary studies using 'brute force' implementations also appearing in [12]. The approach modelled the joint vehicle trajectory and map as a single RFS, and recursively propagated its first order moment, known as the Probability Hypothesis Density (PHD). More refined approaches which adopted a factored concept, in the form of a particle representation of the vehicle state, each maintaining its own PHD filter for the estimation of the map state, are detailed in [3], [4]. Recent work by Lee *et al.* has modelled the SLAM problem with a single cluster PHD filter, which models the vehicle state as a parent process and the map state as a Poisson spatial point daughter process [13].

B. Motivation & Prior Work in Robotic Navigation with Radar

The prime motivation for radar in the experimental analysis of this tutorial, as opposed to laser, vision and ultrasonic based sensors, is its ability to operate in environments containing high levels of dust, rain, fog or snow. In contrast to the other sensors, radar has the ability to penetrate such atmospheric effects and also has the potential to provide more information, than for example laser range finders, as it provides multiple line-of-sight data. As an example, Figure 5 shows superimposed laser range data (black points) and received power data (plotted in color) from the car park environment of Figure 6 (left). Two of the lamp posts in the scene are missed by the laser range finders (labelled) as are the occluded trees, which are shielded by the shrubs. The radar data however yields significant received power values from these, as well as the occluding, objects, albeit at a reduced angular resolution. The specific details of how radar measurements can be used in SLAM are left until Section V.

Because of these advantages, radar sensors have been adopted by quite a number of research groups worldwide. [1] used a W-band radar sensor for feature based SLAM experimental analyses, while reflectivity patterns from leisure craft were examined in [14]. Further SLAM and mapping investigations using W-band radar were presented in [15], [16] examining the signal statistics and their influence on the resulting localization and map estimates. Further motivation for, and a review of, the use of radar in SLAM, with an introduction to related processing methods, is given in [17].

III. REVISITING SLAM FUNDAMENTALS

The aim of SLAM is to concurrently estimate the pose (position and orientation) of a vehicle X_k and the positions of its surrounding features, typically stacked in a map vector M_k at time k. Due to the uncertainty in vehicle motion and sensor measurements, SLAM estimates a joint probability density function (PDF) on these values $p_k(X_k, M_k)$. As a vehicle acquires more measurements, this distribution is updated according to Bayes Theorem:

$$p_{k|k}(X_k, M_k | Z^k, U^{k-1}, X_0) = c p_{k|k-1}(X_k, M_k | Z^{k-1}, U^{k-1}, X_0) g(Z_k | M_k, X_k) \qquad (1)$$

where $p_{k|k}(X_k, M_k | Z^k, U^{k-1}, X_0)$ is an estimate of the joint PDF on X_k and M_k, given all measurements Z^k up to and including time k, all input values U^{k-1} used to control the vehicle motion, up to and including time $k-1$ and the vehicle's initial pose X_0. c is a normalizing value to ensure that $p_{k|k}(X_k, M_k | Z^k, U^{k-1}, X_0)$ is a PDF and $g(Z_k | M_k, X_k)$ is the sensor measurement likelihood, given the map and vehicle location.

Primarily, this tutorial demonstrates the performance of recent experimental results which remodel the measurement *vector* Z_k as an RFS $\mathcal{Z}_k$ and the map vector M_k as an RFS $\mathcal{M}_k$. To motivate these changes, sensor measurement models are initially examined in Sections IV and V, followed by justifications for a set based model in Section VI.

IV. MEASUREMENTS AND THEIR TRUE UNCERTAINTIES

This section high lights the importance of a sensor's and/or feature detection algorithm's *detection* uncertainty as well as the usually considered *spatial* uncertainty. This plays a key role in the principled incorporation of measurements into stochastic SLAM algorithms.

Active sensors, such as laser range finders, sonars and radars transmit various forms of energy into their environments. After transmission, the receiver section of the sensor then converts any power incident on its aperture into a received signal. In some sensors, this received signal, as a function of range from the sensor, is available for processing. For example, the recent Sick LMS 5xx laser range finders, yield the largest 5 received power spikes at different ranges. Underwater sonars also provide such received power arrays. In the case of radar, the received power, as a function of distance from the radar, is called an "A-Scope". Many active sensors however keep this signal "hidden" from the user, and apply their own form of detection theory to provide a simple range decision - i.e. a range value, which is typically interpreted as corresponding to the detection of a valid feature. Importantly, in any sensor, the output is directly related to the time varying power measured at its receiver and, whether hidden from the user or not, vital information about an environment depends on how this signal is processed.

A. *Interpreting Received Power versus Range Signals*

As a case study, an A-Scope is examined, which resulted when a radar pointed in the direction θ of a landmark i. Note that this signal could be the (usually internal) received power signal, as a function of range in any active range finder. The A-Scope is illustrated in Figure 2, Graph (a), which plots an array of linearized power values $\mathbf{S}_k^{\mathrm{lin}}(\theta) = [S_k^{\mathrm{lin}}(1)\ S_k^{\mathrm{lin}}(2)\ \ldots\ S_k^{\mathrm{lin}}(q)\ \ldots\ S_k^{\mathrm{lin}}(Q)],$

with elements $S_k^{\text{lin}}(q)$ containing the received power corresponding to a particular range bin[2] q, recorded at discrete time k. Graph
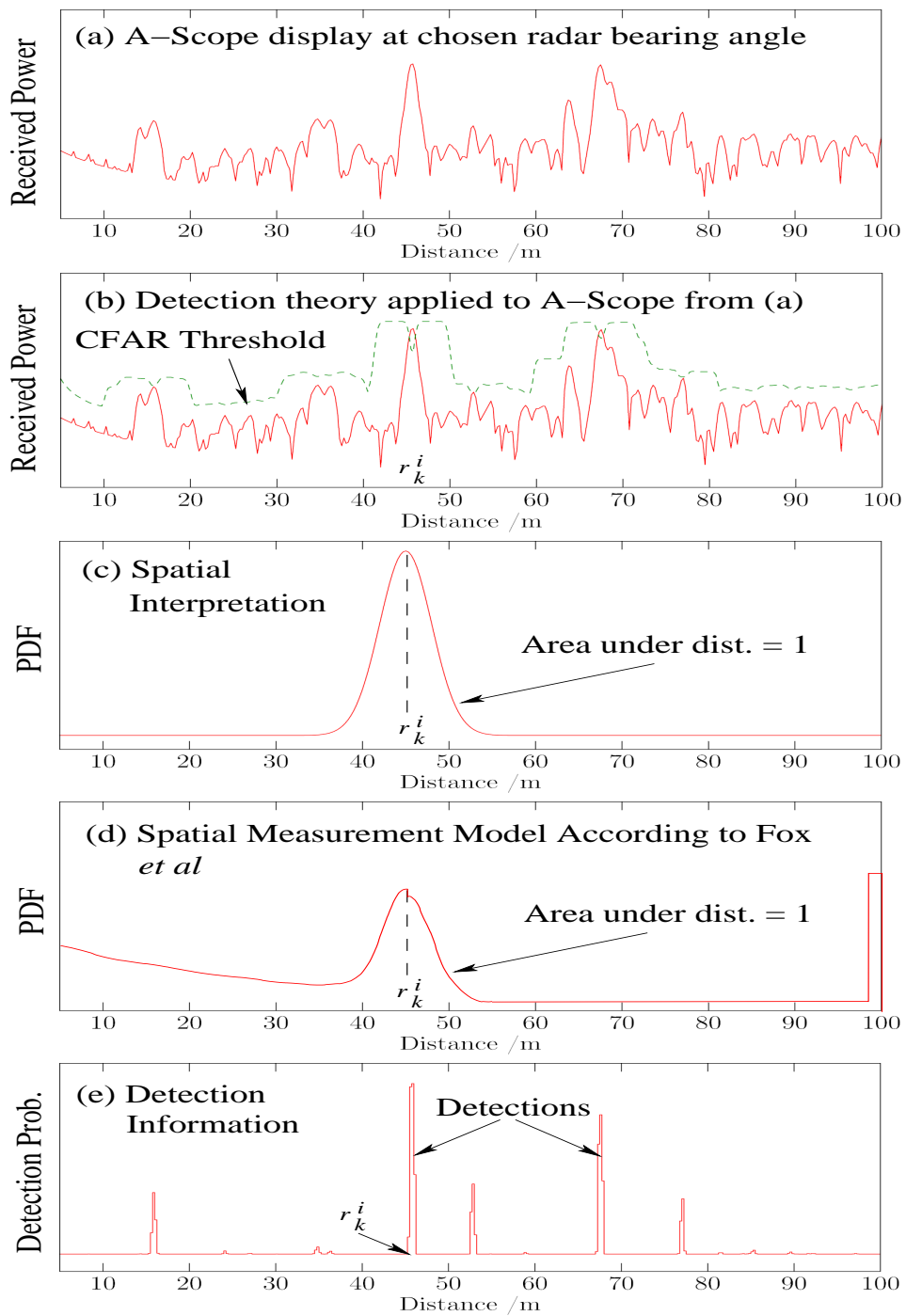

Fig. 2: *The interpretation of measurement detection in robotics and radar/tracking research.*

(b) illustrates the concept of detection theory, applied to the A-Scope in Graph (a). Most detection theoretic methods aim to identify power signals emanating from true targets, with a constant (typically very small) probability of false alarm (incorrectly

---

[2]A range bin is a discretized segment of the range space, in which a single received power value is received. It corresponds to the range resolution of the sensor.

declaring a target) and a quantifiable probability of detection (probability of correctly declaring a target). Such methods are referred to as *Constant False Alarm Rate* (CFAR) processors. Note that in the example of Figure 2(b), the received power from the true target at range r_k^i exceeds the detection threshold, as does another power value further down range. Graph (c) illustrates the way in which most of the spatial modelling approaches in the robotics literature have interpreted this information. Its interpretation exists only in the spatial domain. A probability distribution is created in the range space with its mean located at the range bin of the closest range detection, and spatial variance based on prior information about the sensor. Note that by definition, the area under the distribution must be unity, implying that *one* landmark is assumed to exist with unity probability somewhere in the range space - i.e. a unity *Probability of Detection* ($P_D = 1$) is assumed.

Work by Fox *et. al.* proposed a sensor likelihood to represent the spatial information, from sensors yielding single range values, based on error densities associated with object detections as well as false, missing and unexpected data as illustrated in Graph (d) [18]. The distribution is the weighted sum of that in Graph (c) corresponding to a correct detection corrupted with Gaussian noise, as well as an exponential distribution to model unexpected objects, a uniform distribution over the sensor's operating range to model false alarms and finally a Dirac distribution at maximum range to model missed detections. While this model offers a significant advance over that of Graph (c), it still assumes the existence of a *single* object over the sensor's range space, and is only valid for sensors which yield maximum range when a missed detection occurs.

Graph (e) illustrates the different way in which most of the detection modelling approaches in the radar and tracking literature have interpreted this form of measurement information. Based on a pre-chosen *Probability of False Alarm*, P_{fa}, from which the detection threshold (dashed line in Figure 2(b)) is derived, every range bin q is accompanied by a calculated P_D and a hypothesis value $\mathcal{H}_0$ or $\mathcal{H}_1$ corresponding to "No landmark detected" or "landmark detected" respectively. Hence, in the interpretation in Graph (e), two landmarks would be hypothesized to exist. Once it is acknowledged that detections may not correspond to true objects, the concept of multiple line of sight detections makes sense[3]

The question naturally arises as to which interpretation is correct? Since, in Graph (e), more than one range bin can be associated with hypothesis $\mathcal{H}_1$, contrary to the spatial interpretation (Graphs (c) or (d)), no assumption on the number of possible landmarks is made. However under this interpretation it is not clear how any prior knowledge of spatial uncertainty could be incorporated into the measurement information. Therefore all interpretations fail to represent *all* of the available information. The fundamental difficulty in applying standard, vector based, estimation theory to this problem is that the information to be estimated has different dimensions. It is necessary to estimate the *number* of features due to detection uncertainty, which is a dimensionless quantity, and the *location* of the features themselves due to spatial uncertainty, which have spatial dimensions.

An RFS framework can readily overcome these issues. For example, the PHD approximation of an RFS does not maintain an existence estimate on each feature, but instead propagates a density which represents the mean number of features in the map as well as their spatial densities. Fundamentally, the concept of a *set based* measurement and state representation will be introduced, which allows *all* of the measurement information, spatial and detection, to be incorporated into joint Bayesian navigation and mapping frameworks.

Before introducing an RFS based solution to SLAM, the following sections clarify the differences between the measurement models under each of the above interpretations and introduce a set based measurement model.

---

[3]It should also be noted that radar has the ability to penetrate certain materials, and has a wider beam width than laser range finders, thus providing the potential for received power spikes from multiple objects down range.

A. Relating Spatial Measurements to the Map/SLAM State

Spatial modelling is the standard method of incorporating measurements into the SLAM estimation problem in autonomous robotics research. For the simple planar motion and sensing case, a range/bearing measurement from feature i is shown diagrammatically in Figure 3. Such a measurement corresponds to an assumed detection ($P_D = 1$) at range value r_k^i and
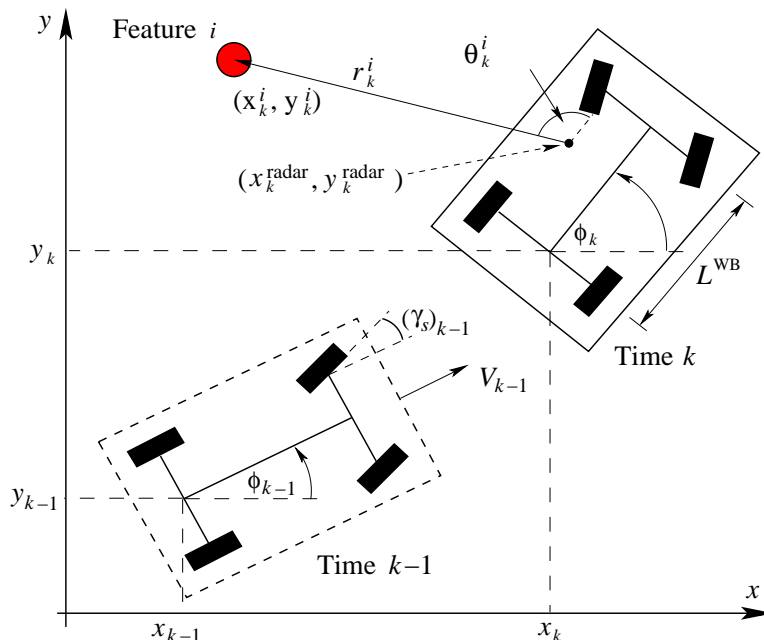

Fig. 3: *Relating target location to vehicle motion.*

bearing angle θ_k^i, and in robotics research, is often interpreted in a stochastic manner in which the sensed object is "smeared" in the range (as in Graph (c) in Figure 2) and bearing space, according to a multi-variate Gaussian distribution. From Figure 3 it can be seen that it is possible to relate the spatial measurement vector $[r_k^i\ \theta_k^i]^T$ to the vehicle pose $X_k = [x_k\ y_k\ \phi_k]^T$ and the ith environmental feature $m_k^i = [x_k^i\ y_k^i]^T$ states

$$r_k^i = \sqrt{(x_k^i - x_k^{\text{radar}})^2 + (y_k^i - y_k^{\text{radar}})^2} + (w_k^r)^i \tag{2}$$

$$\theta_k^i = \arctan\left[\frac{y_k^i - y_k^{\text{radar}}}{x_k^i - x_k^{\text{radar}}}\right] - \phi_k + (w_k^\theta)^i \tag{3}$$

where $\left(x_k^{\text{radar}}, y_k^{\text{radar}}\right)$ are the coordinates of the sensor and $(w_k^r)^i$ and $(w_k^\theta)^i$ are the, assumed additive, spatial measurement random noise terms in range and bearing respectively.

The range/bearing Equations 2 and 3 can be written as a single vector equation. When augmented for all landmarks detected at different sensor bearing angles within a scan, it takes the form of a stacked vector of *spatial* values, which relates to the map and vehicle states via

$$Z_k = \begin{bmatrix} r_k^i \\ \theta_k^i \end{bmatrix} = h^{\text{spatial}}(M_k, X_k) + w_k^{\text{spatial}} \tag{4}$$

where $h^{\text{spatial}}()$ is generally a non-linear function, mapping the landmark and vehicle locations into the relative range and bearing measurements with w_k^{spatial} being the additive augmented vector of range and bearing measurement noises over all detected

features within a scan. To demonstrate such spatial measurements, Figure 4 shows a vehicle[4] carrying a 360° scanning, radar[5], capable of range measurements of up to 200m with a range bin resolution of 25cm. The vehicle also has two 180° scanning
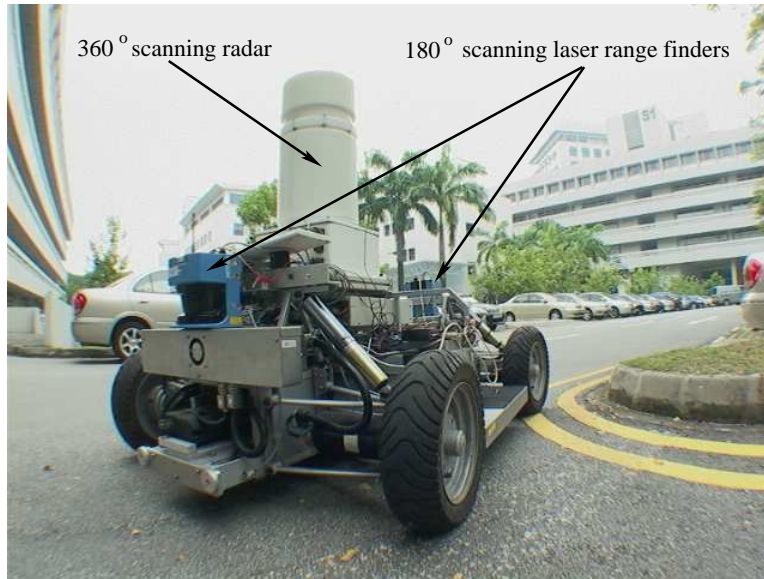

Fig. 4: *A vehicle carrying 2 laser range finders and a 360° scanning MMW radar.*

laser range finders[6] (labelled).

Experiments were conducted in a car park, causing radar returns from trees, lamp posts, buildings as well as clutter from ground measurements, due to the rolling of the vehicle during maneuvers. The results of a single, 360° laser scan of the environment are shown in Figure 5. Laser range decisions are shown as black dots, relative to the robot location at the origin.

B. Relating Power Measurements to the Map/SLAM State

Sensing methods exist, which yield measurements that are not spatial in nature. As examples, cameras produce pixel intensity values corresponding to objects within a scene and radar primarily produces received power values from an object, even though range can be inferred. In the case of radar, linearized received power measurements $\mathbf{S}^{\text{lin}}$ can be used as the measurements at specific range bins q, which are also related to the position of the vehicle X_k and a map state M_k. For example, if the map state M_k were defined to incorporate point landmark *Radar Cross Sections* (RCSs), as well as spatial locations, then the measurements obtained from a radar scan of an environment could take the form

$$Z_k^{\text{power}} = \mathbf{S}_k^{\text{lin}}(\theta : 0 \to 2\pi) = h^{\text{power}}(M_k, X_k) + w_k^{\text{power}} \tag{5}$$

where Z_k^{power} now corresponds to a stacked vector of A-Scope power values $\mathbf{S}_k^{\text{lin}}(\theta : 0 \to 2\pi)$ at time k, taken at each discrete bearing angle θ. Based on a state M_k containing landmark RCSs and spatial coordinates, $h^{\text{power}}()$ could be a vector of functions

---

[4]The vehicle is called the "RobuCAR" from Robosoft (http://www.robosoft.com).

[5]The radar is from Navtech Radar Ltd, and is a 77GHz, Frequency Modulated Continuous Wave (FMCW) device.

[6]The laser range finders are the Sick LMS 200 devices, capable of reported range measurements of up to 80m, with a distance resolution of 10mm.
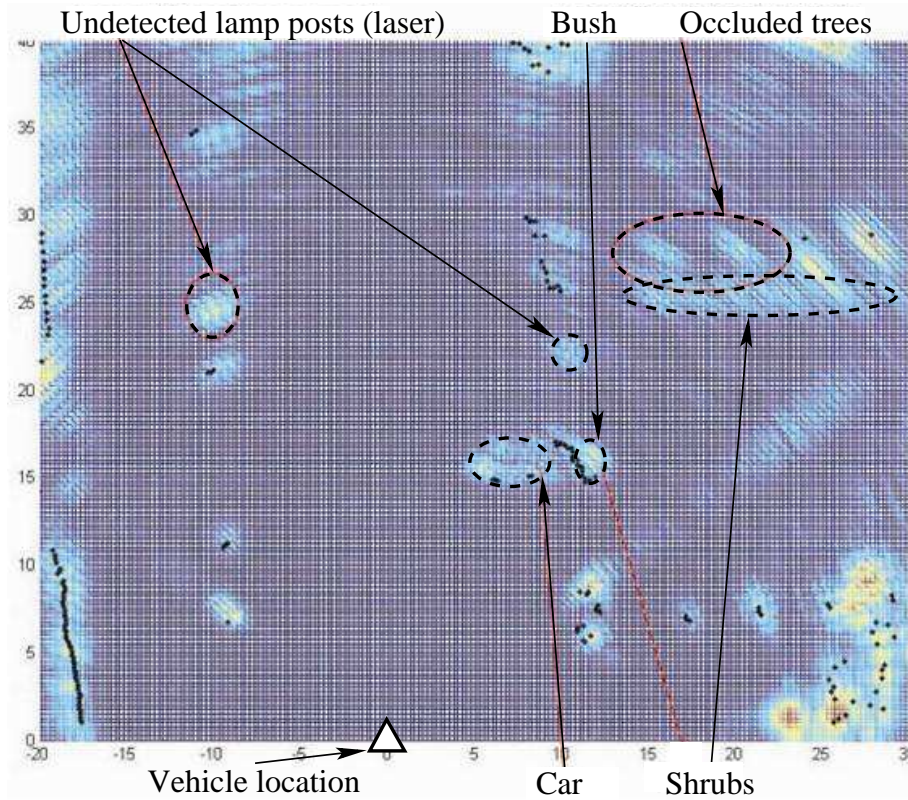
Fig. 5: *Received laser range and radar power data superimposed on the same scan. Note the missed detections by the laser range finder.*

based on the radar equation, which relates received power to target RCS and range. w_k^{power} could then be the received power noise associated with each power measurement.

To demonstrate received *power* type measurements, a single, 360° scan from the radar on board the vehicle of Figure 4 is superimposed onto the laser data of Figure 5. The received power values from the radar are shown as color values, with dark blue indicating minimum, increasing to light blue and yellow to red indicating maximum received power. Such a display can be envisaged as the top view of several A-Scopes (Figure 2(a)) swept about the origin, where color represents received power. In radar terminology, this is known as a sector *Plan Position Indicator* (PPI) plot.

C. Relating Detection Measurements to the Map/SLAM State

In a radar based application, it can be more convenient to consider the measurement Z_k to be the simple binary detection hypotheses $\mathcal{H}_{0/1}$ corresponding to detection/no detection. The measurements in this case can be defined as

$$Z_k^{\text{det}} = \mathbf{H}_{0/1} = h^{\text{detection}}(\mathbf{S}_k^{\text{lin}}(\theta : 0 \rightarrow 2\pi)), \tag{6}$$

where Z_k^{det} now corresponds to a vector of binary detection hypotheses $\mathbf{H}_{0/1}$ with elements $\mathcal{H}_{0/1}^\theta(q)$ corresponding to each range bin q at each bearing angle θ. In this case, $h^{\text{detection}}()$ would correspond to a vector of functions derived from a chosen detector. Note that $\mathbf{S}_k^{\text{lin}}(\theta : 0 \rightarrow 2\pi)$ is then related to the map M_k and vehicle X_k states through Equation 5. Each detection measurement is accompanied by unique P_D and P_{fa} values. For those measurements corresponding to an assumed detection, they can be located spatially according to Equations 2 and 3.

To illustrate this form of measurement, Figure 6 (right) shows *detections* produced by a standard radar detection method[10], known as the *Ordered Statistics* (OS)-CFAR processor, superimposed after three consecutive loops within a car park environment. At this stage, the detections are superimposed onto the known ground truth trajectory of the vehicle, to demonstrate the quality
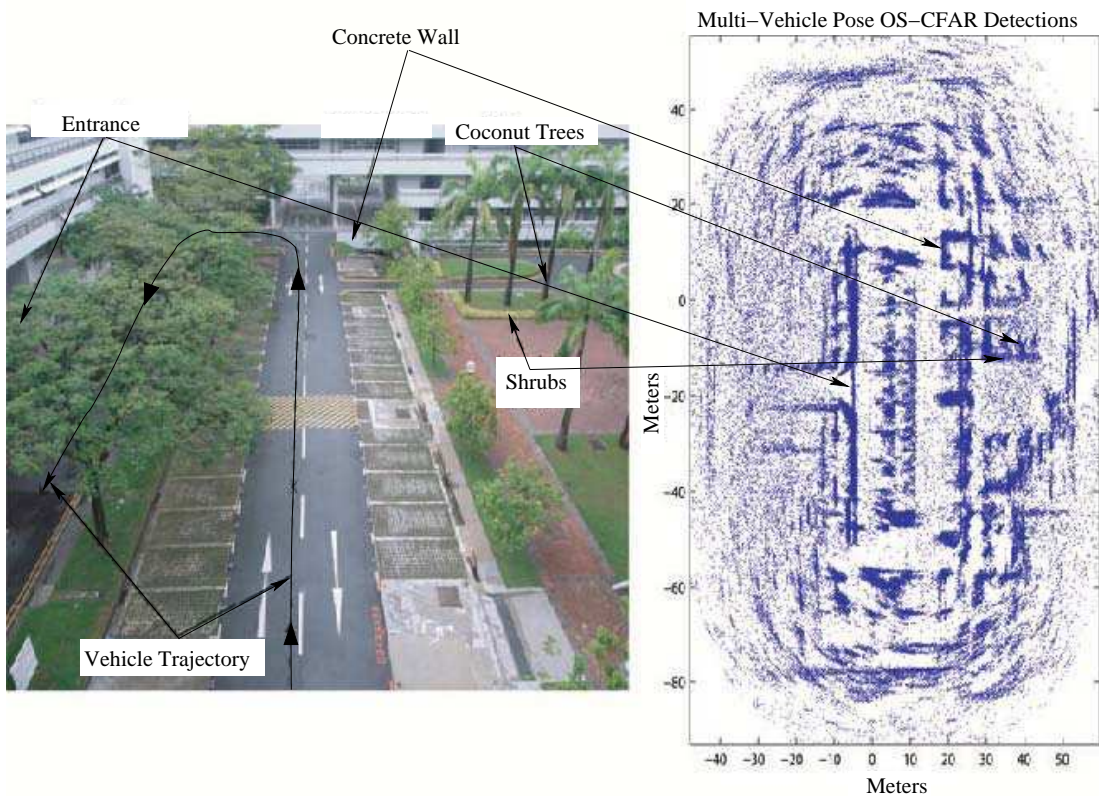

Fig. 6: *(Left) Photograph of a car park scene, with the vehicle's trajectory and objects labelled. (Right) OS-CFAR radar detections, superimposed onto the vehicle's multi-loop trajectory, within the car park environment.*

of the radar detections alone, without including the vehicle localization errors. Many of the true objects in the photo (left) can be identified among the radar detections (right) and some are labelled. Also clearly evident however, are many *false alarms* in the radar data, which can result from incorrect detector behavior, noise and pitching of the vehicle, causing random ground returns.

D. Set Based Measurements

It is interesting to note that for SLAM, measurement Equations 4, 5 and 6 fail to completely encapsulate all of the information provided in a general feature measurement[7]. Equation 4 fails to incorporate detection and false alarm probabilities, while Equations 5 and 6 fail to explicitly make use of any knowledge of a sensor's spatial uncertainty.

To encapsulate detection uncertainty, as well as spatial measurement noise, the detected features from a vehicle with pose X_k, at time k, can be mathematically modelled by an RFS $\mathcal{Z}_k$. This is formed by the union of a set of measurements expected to

---

[7]This actually applies to any sensor when landmarks are extracted using realistic landmark extraction algorithms. Such algorithms are always accompanied by non-unity probabilities of landmark detection and non-zero probabilities of false alarm, which may be quantifiable, and should then be used.

be generated under the current map estimate and a set of clutter (false alarm) measurements. Importantly, each set encapsulates[8] the aforementioned detection and spatial uncertainties and hence

$$\underbrace{\mathcal{Z}_k}_{\substack{\text{All}\\\text{Measurements}}} = \bigcup_{m \in \mathcal{M}_k} \underbrace{\mathcal{D}_k(m, X_k)}_{\substack{\text{Feature}\\\text{Measurements}}} \cup \underbrace{\mathcal{C}_k(X_k)}_{\text{Clutter}} \tag{7}$$

where $\mathcal{D}_k(m, X_k)$ is the RFS of measurements generated by a feature at location m and $\mathcal{C}_k(X_k)$ is the RFS of the spurious measurements at time k, which may depend on the vehicle pose X_k. $\mathcal{Z}_k = \{z_k^1, z_k^2, \ldots, z_k^{\mathfrak{z}_k}\}$ consists of a random number, $\mathfrak{z}_k$, of spatial measurements z_k^i, whose order of appearance has no physical significance with respect to the estimated map of features.

For each feature, $m \in \mathcal{M}_k$, and $z_k^i \in \mathcal{Z}_k$,

$$\mathcal{D}_k(m, X_k) = \{z_k^i\} \tag{8}$$

with probability density $P_D(m, X_k) g(z_k^i | m, X_k)$ and $\mathcal{D}_k(X_k, m) = \emptyset$ with probability $1 - P_D(m, X_k)$, where $P_D(m, X_k)$ is the probability of the sensor detecting feature m from pose X_k and $g(z_k^i | m, X_k)$ represents the sensor's likelihood of detecting z_k^i, given m and X_k. Note the principled association of both detection ($P_D(m, X_k)$) and spatial ($g(z_k^i | m, X_k)$) uncertainties with the RFS $\mathcal{D}_k(m, X_k)$. Spatial uncertainty term $g(z_k^i | m, X_k)$ would reflect the statistics of the spatial noise processes $(w_k^r)^i$ and $(w_k^\theta)^i$ in Equations 2 and 3. Similarly, the spurious measurement rate statistics corresponding to $\mathcal{C}_k(X_k)$ are typically a priori assigned based on an expected P_{fa}, available from the chosen detection equations.

1) Case Study: Landmark Extraction with MMW Radar: Prior to the selection of radar based landmarks, received power values at specific range bins q are classified into detection, $\mathcal{H}_1$ hypothesis, and the no detection or clutter $\mathcal{H}_0$ hypothesis, via an OS-CFAR detector (Figure 2(b)). Many power spikes in the radar A-Scopes correspond to clutter. If S^{lin} is the linearized received radar signal amplitude[8], then the empirical clutter amplitude, $p(S^{\text{lin}} | \mathcal{H}_0)$ can be obtained by Monte Carlo (MC) analysis over a large number of sample scans, using manually selected received power values containing only radar returns from areas free of landmarks. The OS-CFAR detector locally estimates the exponential distribution moment, in each range bin, q, and derives an adaptive threshold value, $S^{\text{OS-CFAR}}(q)$ (Figure 2(b)). $S^{\text{OS-CFAR}}(q)$ can be derived as a function of the desired false alarm rate $P_{fa}^{\text{OS-CFAR}}$. Typically $P_{fa}^{\text{OS-CFAR}}$ is chosen to be low, and in the experiments here is set to 10^{-5}. A landmark is considered detected in range bin q if the received power $S^{\text{lin}}(q) \geq S^{\text{OS-CFAR}}(q)$. Based on this detection criterion, each received power value $S^{\text{lin}}(q)$ is accompanied by a probability of detection $P_D^{\text{OS-CFAR}}(q)$. Importantly, both $P_{fa}^{\text{OS-CFAR}}$ and $P_D^{\text{OS-CFAR}}(q)$ will be utilized in the RFS based SLAM concept, rather than assuming $P_{fa}^{\text{OS-CFAR}} = 0$ and $P_D^{\text{OS-CFAR}}(q) = 1$ as is necessary in vector based SLAM.

OS-CFAR detections, based upon multiple radar scans as a vehicle traversed three loops of a car park environment, were already shown in Figure 6 (right). It shows that, while much useful information is present in OS-CFAR detections, it is still corrupted by many false alarms or clutter. Many more false detections result than those predicted by detection theory, and the real challenge in SLAM is to provide algorithms which are robust in the presence of such clutter measurements.

VI. SET BASED REPRESENTATION OF THE MAP STATE

In Section V-D the concept of an RFS measurement was justified. This section advocates that the map state itself is also more appropriately represented as a set. Consider a simplistic, hypothetical scenario in which a mobile robot traverses three different

---

[8]Note that for processing purposes, the linearized received power is used, however, due to the large dynamic range of the received power, a logarithmic scale is used for graphical purposes.
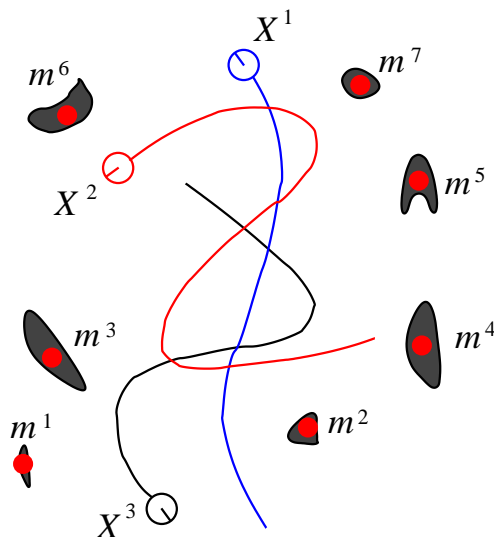
Fig. 7: *A hypothetical scenario in which a mobile robot executes three different trajectories X^1, X^2, X^3, amidst static objects (features) m^1 to m^7.*

trajectories, among static objects, as shown in Figure 7. If the trajectory taken by the robot were X^1 (blue), then it would seem logical that an on board sensor, with a limited range capability, may sense feature m^1 followed by m^2 followed by m^3 etc. Hence after completing trajectory X^1, if a vector M is used to represent the map, then the map could be represented as the first vector in Equations 9. Alternatively, had the robot pursued trajectory X^2 (red) instead, the order in which the features would be sensed would likely be very different, and the resulting estimated map could be the second vector in Equations 9, and had the robot pursued trajectory X^3 (black), the third map vector in Equations 9 could result. Since the order of the elements within a vector is of importance (a change in the order yields a different vector), three different map vectors result. However, since the map features themselves were assumed static, it seems odd that this estimated vector is actually dependent on the vehicle's trajectory. In a strict mathematical sense, the order of the features within the map estimate should not be significant, as any permutation of the vectors results in a valid representation of the map. By definition, the representation which captures all permutations of the elements within the vector, and therefore the features in the map, is a *finite set* $\mathcal{M}$, as shown in representation 9.

$$M = [m^1 \ m^2 \ m^3 \ m^4 \ m^5 \ m^6 \ m^7]^T$$
$$M = [m^4 \ m^2 \ m^3 \ m^1 \ m^5 \ m^7 \ m^6]^T$$
$$\vdots$$
$$\underbrace{M = [m^6 \ m^7 \ m^5 \ m^4 \ m^3 \ m^2 \ m^1]^T}$$

$$\mathcal{M} = \{m^1 \ m^2 \ m^3 \ m^4 \ m^5 \ m^6 \ m^7\} \tag{9}$$

A. Circumventing the Feature Association Problem

In vector based SLAM, the order in which features are acquired is generally unrelated to the order of the feature elements in the current map state estimate. This is illustrated in Figure 8 in which a measurement-to-state assignment problem is evident. A re-ordering of the observed feature vector Z is necessary, since in general, observed feature z^i will not correspond to the current estimate m^i (e.g. measurement z^5 resulted from feature m^4 etc.). The proposed RFS approach represents both features
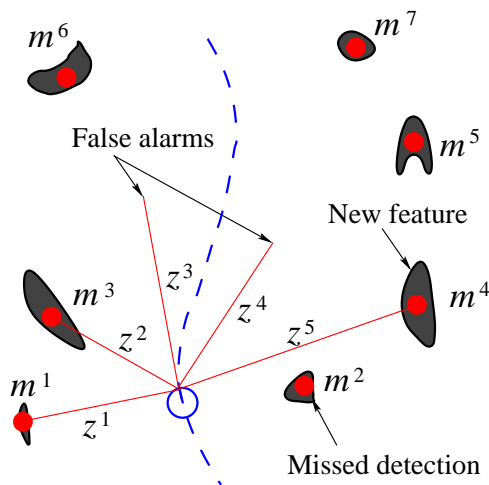
Fig. 8: *As the robot moves, the order in which features are detected will, in general, not correspond to the order of stored features in the map vector. Also, new features will enter the FoV of the sensor(s). In general, some features may be undetected and some false alarms may be declared.*

and measurements as finite valued sets $\mathcal{M}$ and $\mathcal{Z}$ respectively, which assume no distinct ordering of the features, as shown in representations 9 and 10.

$$Z = [z^1 \ z^2 \ z^3 \ z^4 \ z^5]^T$$
$$Z = [z^4 \ z^2 \ z^3 \ z^1 \ z^5]^T$$
$$\vdots$$
$$\underbrace{Z = [z^5 \ z^4 \ z^3 \ z^2 \ z^1]^T}$$
$$\mathcal{Z} = \{z^1 \ z^2 \ z^3 \ z^4 \ z^5\} \tag{10}$$

Since the finite set representations 9 and 10 naturally encapsulate all possible permutations of the feature map and measurement, a FISST representation of the set will circumvent the need for feature association.

B. Circumventing the Map Management Problem

The feature measurement to state assignment problem is further exacerbated since in reality, the number of measurements, $\mathfrak{z}_k$, at any given time is not fixed due to detection uncertainty, spurious measurements and unknown number of true features. For example, in Figure 8 it can be seen that 5 features have been detected, although there are seven features in the environment. Objects m^5, m^6 and m^7 lie out of range of the sensor, in the robot's current position. Due to sensor and/or feature detection algorithm imperfections, two false alarms z^3 and z^4 have occurred. These can originate from clutter, sensor noise or incorrect feature detection algorithm performance. Notice that although object m^2 lies within the FoV of the sensor, it has not been detected, and constitutes a *missed detection*.

Suppose that features m^1, m^2 and m^3 already exist at time $k-1$ in a vector based map representation, and that feature m^4 now falls into the robot's sensor(s) FoV. Feature m^4 should be initialized and included in the state estimate at time k. To date, many SLAM techniques use vector augmentation methods based on feature initiation rules. However if the map is defined as a

set, then a set based map transition function can be mathematically defined as

$$\mathcal{M}_{k-1} = \{m^1 \ m^2 \ m^3\}$$

$$\mathcal{M}_k = \{m^1 \ m^2 \ m^3\} \cup \{m^4\} \tag{11}$$

Contrary to vector based SLAM, a principled mathematical relationship between a set based measurement $\mathcal{Z}_k$ and the map $\mathcal{M}_k$, which takes into account the different dimensions of $\mathcal{Z}_k$ and $\mathcal{M}_k$ as well as spurious measurements (such as z^3 and z^4 in Figure 8), can be defined and was given in RFS Equation 7.

By redefining the concept of Bayes optimality for estimation with unknown feature number, by formulating it as an RFS estimation problem, the formulation unifies the independent filters adopted by previous, vector based solutions, through the recursive propagation of a distribution of an RFS of features. This allows for the joint propagation of the Feature Based (FB) map density and leads to Bayes optimal map estimates in the presence of unknown map size, spurious measurements, feature detection and data association uncertainty. The RFS framework, summarized in Figure 9, further allows for the joint treatment of error in feature number and location estimates as it jointly propagates both the estimate of the number of landmarks and their corresponding states. Feature association is automatically dealt with under the multi-object statistical framework, and consequently *the need for feature association and management heuristics are circumvented.*
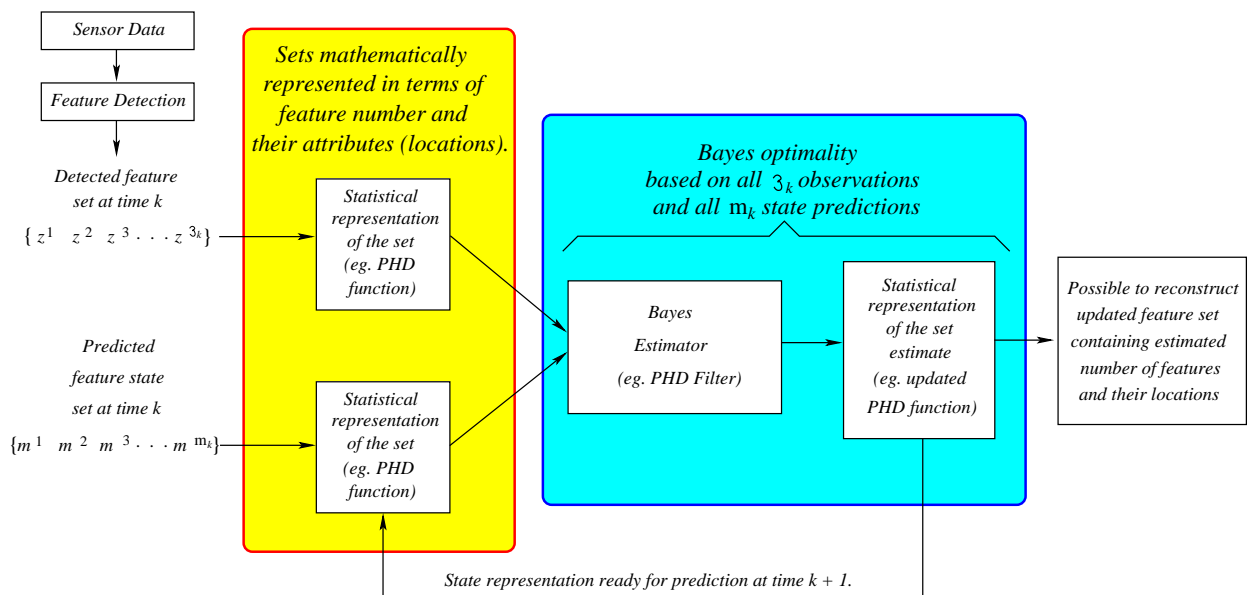

Fig. 9: *RFS based feature mapping, for a single update cycle, in SLAM. Notice that no map management or data association filters/heuristics are necessary.*

VII. SLAM WITH RANDOM FINITE SETS

A. Vector Representation of the Vehicle State

For the experiments adopted in this tutorial, a simple planar motion vehicle model is based on the Ackerman model [19]. With reference to Figure 3, the discrete time varying control inputs to the vehicle at time $k-1$ are forward speed V_{k-1} and

steering angle $(\gamma_s)_{k-1}$. The local origin on the vehicle, is defined as the center of the rear axle. In discrete time form, with the inclusion of noise terms in the control inputs, the vehicle state can be predicted from the following kinematic equations [1].

$$
\begin{bmatrix} x_k \\ y_k \\ \phi_k \end{bmatrix} = \begin{bmatrix} x_{k-1} \\ y_{k-1} \\ \phi_{k-1} \end{bmatrix} + \begin{bmatrix} \Delta T(V_{k-1} + v_{k-1}^V)\cos[\phi_{k-1}] \\ \Delta T(V_{k-1} + v_{k-1}^V)\sin[\phi_{k-1}] \\ \frac{\Delta T(V_{k-1}+v_{k-1}^V)\tan[(\gamma_s)_{k-1}+v_{k-1}^{\gamma_s}]}{L^{\mathrm{WB}}} \end{bmatrix}
\tag{12}
$$

where x_k represents the x coordinate of the vehicle at time k; x_{k-1} represents the previous x coordinate of the vehicle at time $k-1$ (and similarly for y and heading ϕ); ΔT is the time between updates $k-1$ and k and L^{WB} is the wheel base length defined in Figure 3. $v_{k-1}^V \backsim \mathcal{N}(0, (\sigma_V^2)_{k-1})$ and $v_{k-1}^{\gamma_s} \backsim \mathcal{N}(0, (\sigma_{\gamma_s}^2)_{k-1})$ represent the input noise terms in the model. $\mathcal{N}(\mu, \sigma^2)$ represents a Gaussian distribution with mean μ and variance σ^2.

Since the dimensions of this state are fixed as time progresses, and the order of the variables in the vector remain the same, the robot vehicle state is adequately modelled as a *random vector*. Note that, due to the random noise terms, Equation 12 is typically modelled as a PDF

$$
f^{\mathrm{veh}}(X_k | X_{k-1}, U_{k-1})
\tag{13}
$$

where U_{k-1} represents a vector comprising the input velocity signal and steering angle - i.e. $U_{k-1} = [V_{k-1} \ (\gamma_s)_{k-1}]^T$, which is derived from either odometric or desired input information. This general vehicle process model will be used in all of the SLAM algorithms, developed for comparison purposes, in this tutorial. In particular, the vehicle *trajectory* $X^k = [X_0 \ X_1 \ X_2 \ \cdots \ X_k]^T$ is stochastically represented by a set of N particles (samples) each of which is extended at each time interval from its component location X_{k-1} to X_k through Equation 12.

B. RFS Representation of the Map State

Let $\mathcal{M}$ be the RFS representing the entire unknown and unexplored static map. The explored map $\mathcal{M}_{k-1}$ then evolves in time according to,

$$
\underbrace{\mathcal{M}_k}_{\substack{\text{Current} \\ \text{Map Set}}} = \underbrace{\mathcal{M}_{k-1}}_{\substack{\text{Previous} \\ \text{Map Set}}} \cup \left(\underbrace{FoV(X_k)}_{\substack{\text{Sensor(s)} \\ \text{FoV}}} \cap \underbrace{\bar{\mathcal{M}}_{k-1}}_{\substack{\text{Complement of} \\ \text{Previous Map Set}}} \right)
\tag{14}
$$

where $\bar{\mathcal{M}}_{k-1} = \mathcal{M} - \mathcal{M}_{k-1}$, i.e the set of features not in $\mathcal{M}_{k-1}$. Let the new features which have entered the FoV, i.e. the second term of Equation 14, be modelled by the independent birth RFS, $\mathcal{B}_k(X_k)$. In this case, the RFS map transition density is given by,

$$
f^{\mathcal{M}}(\mathcal{M}_k | \mathcal{M}_{k-1}, X_k) = \sum_{\mathcal{W} \subseteq \mathcal{M}_k} f^{\mathcal{M}}(\mathcal{W} | \mathcal{M}_{k-1}) f_k^{\mathcal{B}}(\mathcal{M}_k - \mathcal{W} | X_k)
\tag{15}
$$

where $f^{\mathcal{M}}(\mathcal{W} | \mathcal{M}_{k-1})$ is the transition density of the set of features that are in the $FoV(X^{k-1})$ at time $k-1$ to time k, and $f_k^{\mathcal{B}}(\mathcal{M}_k - \mathcal{W} | X_k)$ is the birth density of the RFS, $\mathcal{B}(X_k)$, of the new features that pass within the FoV at time k. Implementation details related to the birth density will be given in Section VII-D.

Equation 15 is important since it defines the map transition function for set valued maps. Including the vector based vehicle state transition density $f^{\mathrm{veh}}(X_k | X_{k-1}, U_{k-1})$, the joint transition density of the map and the vehicle pose can be written as,

$$
f^{\mathrm{SLAM}}(\mathcal{M}_k, X_k | \mathcal{M}_{k-1}, X_{k-1}, U_{k-1}) = f^{\mathcal{M}}(\mathcal{M}_k | \mathcal{M}_{k-1}, X_k) f^{\mathrm{veh}}(X_k | X_{k-1}, U_{k-1}).
\tag{16}
$$

As in vector based SLAM, unfortunately, the general RFS recursion equivalent of Equation 1, is also mathematically intractable. A simple approximation to set-based estimation which is tractable, is to exploit the physical intuition of the first moment of an RFS, known as its PHD, v_k. The PHD at a point, gives the *density* of the expected number of features occurring at that point. A PHD function is any function which obeys the following two properties:

1) The mass (integral of the density over the volume) of the PHD within a given spatial region S, gives the expected number of features in S.

2) As a consequence, the peaks of the PHD indicate locations with high probability of feature existence.

To illustrate the application of the PHD estimator to 2D robotic mapping, graphical depictions of posterior PHDs after two consecutive measurements are shown in Figures 10 and 11. These illustrate a simple time varying PHD function, v_k, which is
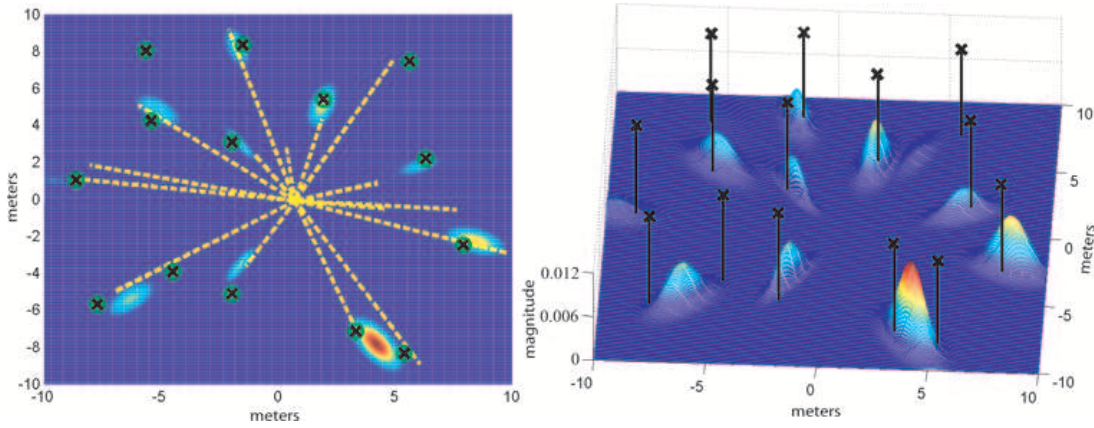

Fig. 10: *A sample map PHD at time $k-1$, with the true map represented by black crosses. The left hand map PHD is the plan view of the right hand 3D view. The measurement at time $k-1$ is represented by the dashed lines. The peaks of the PHD represent locations with the highest concentration of expected number of features. The local PHD mass in the region of most features is 1, indicating the presence of 1 feature. The local mass close to some unresolved features (for instance at (5,-8)) is closer to 2, demonstrating the unique ability of the PHD function to jointly capture the number of features.*

a Gaussian Mixture (GM) containing J_k Gaussian components, i.e.

$$v_k = \{\mu_k^{(j)}, P_k^{(j)}, w_k^{(j)}\}_{j=1}^{J_k}, \tag{17}$$

where $\mu_k^{(j)}$ is the mean spatial location of Gaussian component j at time k, with spatial covariance $P_k^{(j)}$ and weight $w_k^{(j)}$ corresponding to the number of features the jth component represents. In each figure, the PHD is plotted as a function of the spatial coordinates. Since the integral of the PHD is, by definition, the estimated number of features in the map, the mass (or integral) of each Gaussian can be interpreted as the number of features it represents. In the case of closely lying features (and large spatial measurement noise), the PHD approach may not be able to resolve the features, as demonstrated in Figure 10 at approximate coordinates (5, -8). However the PHD will represent the spatial density of L features by a singular Gaussian with a corresponding mass of L, which may improve the feature number estimate. This is only theoretically possible using the RFS framework. A graphical example for $L = 2$ is illustrated in Figure 10, which is then resolved through measurement updates into
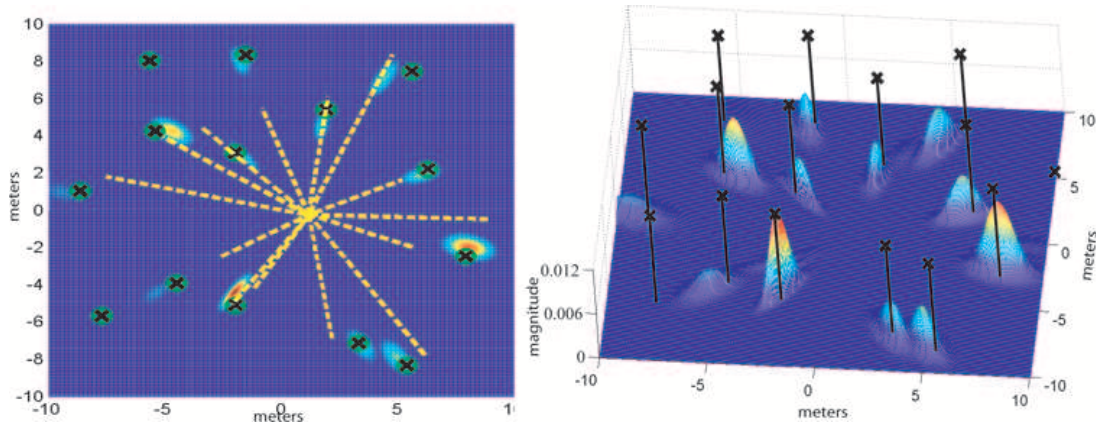
Fig. 11: *A sample map PHD and measurement at time k. Note that the features at (5,-8) are resolved due to well separated measurements, while at (-3,-4), a lone false alarm close to the feature measurement contributes to the local PHD mass. At (-5,-4) a small likelihood over all measurements, coupled with a moderate clutter PHD results in a reduced local mass.*

individual Gaussian components for each feature of mass $L \approx 1$, as shown in Figure 11 (the two peaks near coordinates (5, -8)). The PHD estimator has been proven to be powerful and effective in multi-target tracking [5].

D. The PHD SLAM Filter

Similarly to the FastSLAM concept [20], the PHD-SLAM joint posterior can be factorized when the map is represented as a conditional PDF, conditioned on an entire vehicle trajectory X^k - i.e.

$$\underbrace{p_k(X^k, \mathcal{M}_k | \mathcal{Z}^k, U^{k-1}, X_0)}_{\text{Full SLAM posterior}} = \underbrace{p_k(X^k | \mathcal{Z}^k, U^{k-1}, X_0)}_{\substack{\text{Vector based vehicle} \\ \text{trajectory posterior}}} \times \underbrace{p_k(\mathcal{M}_k | \mathcal{Z}^k, X^k)}_{\substack{\text{Set based map} \\ \text{conditioned on trajectory}}}. \tag{18}$$

Under an RB implementation, the vehicle trajectory X^k is represented as vector particles $(X^k)^{(i)}$, each of which maintain their own set based map estimate $p_{k|k}\left(\mathcal{M}_k | \mathcal{Z}^k, (X^k)^{(i)}\right)$. The recursive Bayesian estimate of the map, per trajectory particle $(X^k)^{(i)}$, is then

$$p_{k|k}\left(\mathcal{M}_k | \mathcal{Z}^k, (X^k)^{(i)}\right) = \frac{p_{k|k-1}\left(\mathcal{M}_k | \mathcal{Z}^{k-1}, (X^k)^{(i)}\right) g\left(\mathcal{Z}_k | \mathcal{M}_k, (X_k)^{(i)}\right)}{g\left(\mathcal{Z}_k | \mathcal{Z}^{k-1}, (X^k)^{(i)}\right)} \tag{19}$$

The set based distribution $p_{k|k}\left(\mathcal{M}_k | \mathcal{Z}^k, (X^k)^{(i)}\right)$ is modelled with a GM implementation of the PHD $v_{k|k}(m| (X^k)^{(i)})$.

The PHD predictor equation is

$$v_{k|k-1}\left(m| (X^k)^{(i)}\right) = v_{k-1|k-1}\left(m| (X^{k-1})^{(i)}\right) + b\left(m|X_k^{(i)}\right) \tag{20}$$

where $v_{k-1|k-1}\left(m| (X^{k-1})^{(i)}\right)$ is previous GM estimate of the PHD, $v_{k|k-1}\left(m| (X^k)^{(i)}\right)$ is its prediction at time k and $b\left(m|X_k^{(i)}\right)$ is the GM PHD of the birth RFS, $\mathcal{B}\left(X_k^{(i)}\right)$. $b\left(m|X_k^{(i)}\right)$ is similar to the proposal function used in particle filters, and is used to give some *a priori* information to the filter about where features are likely to appear in the map. In SLAM, with no *a priori* information, $b\left(m|X_k^{(i)}\right)$, may be uniformly distributed in a non-informative manner about the space of features. However, in this work, the feature birth proposal at time k is chosen to be a GM containing $J_{b,k}$ Gaussian components, representing the set of measurements at time $k-1$, $\mathcal{Z}_{k-1}$ [4].

The PHD corrector is,

$$v_{k|k}\left(m|\left(X^k\right)^{(i)}\right) = v_{k|k-1}\left(m|\left(X^k\right)^{(i)}\right)\left[1-P_D^{\text{OS-CFAR}}\left(m|X_k^{(i)}\right)+\sum_{z\in\mathcal{Z}_k}\frac{\Lambda\left(m|X_k^{(i)}\right)}{c_k\left(z|X_k^{(i)}\right)+\int_{\mathcal{M}_k}\Lambda\left(\xi|X_k^{(i)}\right)v_{k|k-1}\left(\xi|\left(X^k\right)^{(i)}\right)d\xi}\right]$$

(21)

where $v_{k|k}\left(m|\left(X^k\right)^{(i)}\right)$ is the new GM estimate of the PHD at time k, $\Lambda\left(m|X_k^{(i)}\right) = P_D^{\text{OS-CFAR}}\left(m|X_k^{(i)}\right)g_k\left(z|m,X_k^{(i)}\right)$ and,

$$P_D^{\text{OS-CFAR}}\left(m|X_k^{(i)}\right) \quad = \quad \text{the probability of detecting a land-}$$
$$\text{mark at } m, \text{ from vehicle pose } X_k^{(i)}.$$

$$c_k\left(z|X_k^{(i)}\right) \quad = \quad \text{PHD of the clutter RFS } \mathcal{C}_k \text{ in}$$
$$\text{Equation 7 at time } k.$$

The integral $\int_{\mathcal{M}_k}\Lambda\left(\xi|X_k^{(i)}\right)v_{k|k-1}\left(\xi|\left(X^k\right)^{(i)}\right)d\xi$ in the denominator of corrector Equation 21 is calculated as the sum of the J_{k-1} predicted and $J_{b,k}$ birth Gaussians present in $v_{k|k-1}\left(m|\left(X^k\right)^{(i)}\right)$, each weighted by $\Lambda\left(m|X_k^{(i)}\right)$, over all N trajectory particles.

The clutter RFS, $\mathcal{C}_k$, represents the prior knowledge of the probability of false alarm P_{fa}. In a scan, in which N_d feature detection hypotheses are made (whether determined to be detections or not), an average of $\mathfrak{n}_c = P_{fa}N_d$ false measurements will result. Therefore, the clutter PHD is assumed Poisson distributed in number and uniformly spaced over the mapping region. The clutter intensity is therefore given by

$$c_k\left(z|X_k^{(i)}\right) = \mathfrak{n}_cVU(z)$$

(22)

where $\mathfrak{n}_c$ is the clutter rate per scan, V is the volume (or area in the 2D experiments presented here) corresponding to the FoV of the sensor(s) of the surveillance region and and $U(z)$ denotes a uniform distribution on the measurement space.

E. Particle Trajectory Weights

A fundamental difference between RB-PHD-SLAM and FastSLAM should be noted. In FastSLAM, each pose particle is used to generate a predicted measurement vector. The actual measurement vector, recorded from the unknown, true vehicle location, is then superimposed on to each particle. The likelihood of that measurement vector corresponding to that particle's predicted measurement vector is calculated to form a particle weight, through the measurement likelihood $g\left(Z_k|Z^{k-1},\left(X^k\right)^{(i)}\right)$. This requires the usual, fragile predicted and observed feature management and association routines, for which there is no concept within the RFS framework.

In RB-PHD-SLAM, the measurement likelihood is $g\left(\mathcal{Z}_k|\mathcal{Z}^{k-1},\left(X^k\right)^{(i)}\right)$, which is defined on the space of finite sets, unlike its FastSLAM counterpart, which is defined on a Euclidean space. Therefore, alternative methods are necessary to evaluate $g\left(\mathcal{Z}_k|\mathcal{Z}^{k-1},\left(X^k\right)^{(i)}\right)$, and hence the trajectory particle's new weight.

Fortunately, by rearranging Equation 19, it can be seen that $g\left(\mathcal{Z}_k|\mathcal{Z}^{k-1},\left(X^k\right)^{(i)}\right)$ is merely the normalizing constant,

$$g\left(\mathcal{Z}_k|\mathcal{Z}^{k-1},\left(X^k\right)^{(i)}\right) = \frac{g\left(\mathcal{Z}_k|\mathcal{M}_k,X_k^{(i)}\right)p_{k|k-1}\left(\mathcal{M}_k|\mathcal{Z}^{k-1},\left(X^k\right)^{(i)}\right)}{p_{k|k}\left(\mathcal{M}_k|\mathcal{Z}^k,\left(X^k\right)^{(i)}\right)}.$$

(23)

Note in the above, that the LHS does not contain the variable $\mathcal{M}_k$, while the RHS has $\mathcal{M}_k$ in both the denominator and numerator. In essence, $\mathcal{M}_k$ in Equation 23 is a dummy variable, and thus an exact implementation of Equation 23 would hold for *any*

arbitrary choice of $\mathcal{M}_k$. This theoretically allows the substitution of any choice of $\mathcal{M}_k$ to evaluate $g\left(\mathcal{Z}_k|\mathcal{Z}^{k-1},\left(X^k\right)^{(i)}\right)$. This[19] is an important result, which allows the likelihood of the measurement conditioned on the trajectory (but not the map), to be calculated in closed-form. The following subsection considers the simplest implementation of Equation 23 [3], [4].

It should be noted that an alternative approach has also been proposed in [13] and the resulting propagation equations have a similar form to [4], [21]. The approach in [13] does not require the calculation of the normalizing constant of Equation 23 and instead calculates weights for each Dirac based PHD parent process, representing the vehicle's pose, based on Poisson cluster assumptions.

1) The Empty Map Strategy: If the RFS $\mathcal{M}_k$ is approximated as a multi-target Poisson density (Poisson distributed in its number, and the points within $\mathcal{M}_k$ are IID distributed) then the probability density of $\mathcal{M}_k$ can be recovered exactly from the PHD intensity function. Similarly the predicted and posterior RFS maps can be approximated by Poisson RFSs with PHDs $v_{k|k-1}\left(m|\left(X^k\right)^{(i)}\right)$ and $v_{k|k}\left(m|\left(X^k\right)^{(i)}\right)$ respectively, giving

$$p_{k|k-1}\left(\mathcal{M}_k|\mathcal{Z}^{k-1},\left(X^k\right)^{(i)}\right) \approx \frac{\prod\limits_{m\in\mathcal{M}_k} v_{k|k-1}\left(m|\left(X^k\right)^{(i)}\right)}{\exp\left(\int v_{k|k-1}\left(m|\left(X^k\right)^{(i)}\right)dm\right)}, \tag{24}$$

$$p_{k|k}\left(\mathcal{M}_k|\mathcal{Z}^k,\left(X^k\right)^{(i)}\right) \approx \frac{\prod\limits_{m\in\mathcal{M}_k} v_{k|k}\left(m|\left(X^k\right)^{(i)}\right)}{\exp\left(\int v_{k|k}\left(m|\left(X^k\right)^{(i)}\right)dm\right)}. \tag{25}$$

Substituting Equations 24 and 25 into Equation 23, assigning $\mathcal{M}_k=\emptyset$, and rearranging Equation 23 gives,

$$g\left(\mathcal{Z}_k|\mathcal{Z}^{k-1},\left(X^k\right)^{(i)}\right) = g\left(\mathcal{Z}_k|\emptyset,X_k^{(i)}\right) \times \frac{\prod\limits_{m\in\mathcal{M}_k} v_{k|k-1}\left(m|\left(X^k\right)^{(i)}\right)}{\prod\limits_{m\in\mathcal{M}_k} v_{k|k}\left(m|\left(X^k\right)^{(i)}\right)} \times \frac{\exp\left(\int v_{k|k}\left(m|\left(X^k\right)^{(i)}\right)dm\right)}{\exp\left(\int v_{k|k-1}\left(m|\left(X^k\right)^{(i)}\right)dm\right)} \tag{26}$$

Since, $\mathcal{M}_k=\emptyset$, the empty set measurement likelihood is that of the clutter RFS (Poisson),

$$g\left(\mathcal{Z}_k|\emptyset,X_k^{(i)}\right) = \frac{\prod\limits_{z\in\mathcal{Z}_k} c_k\left(z|X_k^{(i)}\right)}{\exp\left(\int c_k\left(z|X_k^{(i)}\right)dz\right)}. \tag{27}$$

The PHDs $v_{k|k-1}$ and $v_{k|k}$ are empty, implying their product is 1, $\hat{\mathrm{m}}_{k|k-1} = \int v_{k|k-1}\left(m|\left(X^k\right)^{(i)}\right)dm$ and $\hat{\mathrm{m}}_{k|k} = \int v_{k|k}\left(m|\left(X^k\right)^{(i)}\right)dm$, giving,

$$g\left(\mathcal{Z}_k|\mathcal{Z}^{k-1},\left(X^k\right)^{(i)}\right) = \prod\limits_{z\in\mathcal{Z}_k} c_k\left(z|X_k^{(i)}\right)\exp\left(\hat{\mathrm{m}}_{k|k} - \hat{\mathrm{m}}_{k|k-1} - \int c_k\left(z|X_k^{(i)}\right)dz\right). \tag{28}$$

Note that $\int c_k\left(z|X_k^{(i)}\right)dz$ by definition gives the number of clutter measurements per scan $\mathrm{n}_c V$, and in the particular case that the clutter intensity is assumed equal for each robot pose $X_k^{(i)}$

$$g\left(\mathcal{Z}_k|\mathcal{Z}^{k-1},\left(X^k\right)^{(i)}\right) \propto \exp\left(\hat{\mathrm{m}}_{k|k} - \hat{\mathrm{m}}_{k|k-1}\right) \tag{29}$$

so that each particle's relative weight can be updated by the factor $\exp\left(\hat{\mathrm{m}}_{k|k} - \hat{\mathrm{m}}_{k|k-1}\right)$.

Note that while for the empty map choice, the likelihood $g\left(\mathcal{Z}_k|\mathcal{Z}^{k-1},\left(X^k\right)^{(i)}\right)$ does not contain a measurement likelihood term $g\left(\mathcal{Z}_k|\mathcal{M}_k,X_k^{(i)}\right)$, the history of measurements and trajectories are incorporated into the predicted and updated intensity terms, whose integrals appear as the terms $\hat{\mathrm{m}}_{k|k-1}$ and $\hat{\mathrm{m}}_{k|k}$ respectively. An implementation of the particle weight that considers a single feature map approximation to the Poisson distributions of Equations 24 and 25 is given in [3], [4], [17].

F. Extracting the Final SLAM Estimate

In MH-FastSLAM, the final trajectory is selected as the particle $\left(X^k\right)^{(i)}$ with the highest weight, and the final map estimate is its corresponding map. This is known as the maximum a-posteriori (MAP) map estimate. This strategy is also used in the comparative study here for RB-PHD-SLAM, meaning that its map PHD, $v_{k|k}\left(m|\left(X^k\right)^{(i)}\right)$, is selected as the final map. If $\hat{\mathrm{m}}_k = \int v_{k|k}\left(m|\left(X^k\right)^{(i)}\right)dm$, is the mass of the posterior map PHD, the expected map estimate can then be extracted by choosing the $\hat{\mathrm{m}}_k$ highest local maxima.

It should be noted that, in contrast to vector based SLAM algorithms, the PHD map representation yields a natural ability to average feature maps, giving an expected a-posteriori (EAP) map, even with map estimates of different size and without having to resolve the intra-map feature associations. Details of such an implementation, as well as pseudo-code implementations of both the MAP and EAP methods, are given in [3], [4], [17].

VIII. Benchmark Vector Based SLAM Implementations for Comparison

To demonstrate the effectiveness of an RFS representation of the map, state of the art vector based SLAM implementations were also carried out, based on the same OS-CFAR based landmarks extracted in Figure 6.

A. Nearest-Neighbor, Extended Kalman Filter SLAM

To date, arguably one of the most popular SLAM implementations has been based on approximating the joint vehicle (X_k) and map (M_k) SLAM state as a multi-variate Gaussian. Equation 1 then yields the EKF. For each OS-CFAR landmark detected, the EKF SLAM algorithm has to decide which landmark in the existing map it corresponds to, or whether a new landmark was observed. *Maximum likelihood* (ML) estimation calculates and compares a statistical distance (the *Mahalanobis distances*) between the detected feature locations and their predicted locations. This metric is then used to determine association between detections and stored features, referred to as the *Nearest Neighbor Standard Filter* (NNSF). In 2001, Neira *et. al.* highlighted limitations of the NN approach, since it ignores the fact that measurement prediction errors are correlated [22]. For this reason, the *Joint Compatibility Branch and Bound* (JCBB) data association method was proposed, which determines matches between *sets* of detected and predicted feature pairs, taking into account the measurement prediction correlations. This is referred to as JCBB-EKF-SLAM, and will be used as a vector based comparison with the RB-PHD-SLAM approach in Section IX.

B. Multiple Hypothesis FastSLAM

FastSLAM estimates the map on a per-particle basis, meaning that different particles can be associated with different features [20]. This means that the FastSLAM filter has the possibility to maintain different tracks for each possible hypothesis of each detected feature, known as *Multiple Hypothesis Tracking* (MHT) [23]. A new particle is created for each new hypothesis of each measurement, meaning that each particle is split into $n+2$ new particles, one for each of the n possible associations, one particle for the non-association hypothesis and the other particle for a new feature hypothesis. Particles with incorrect data association are, on average, more likely to be circumvented than those which were based on correct associations. This step reduces the number of particles back to its original number. This technique, referred to as MH-FastSLAM, will also be used in the results as a non-parametric, vector based SLAM approach for comparison. The MH-FastSLAM implementation was based exactly on the principles given in [20] with details of its associated MH data association possibility given in [24].

A. Performance Comparisons

If, instead of superimposing the OS-CFAR detections onto the vehicle ground truth positions, as in Figure 6 (right), they are superimposed onto the vehicle's odometry estimated positions (shown as the solid trajectory), Figure 12 (left) results. The
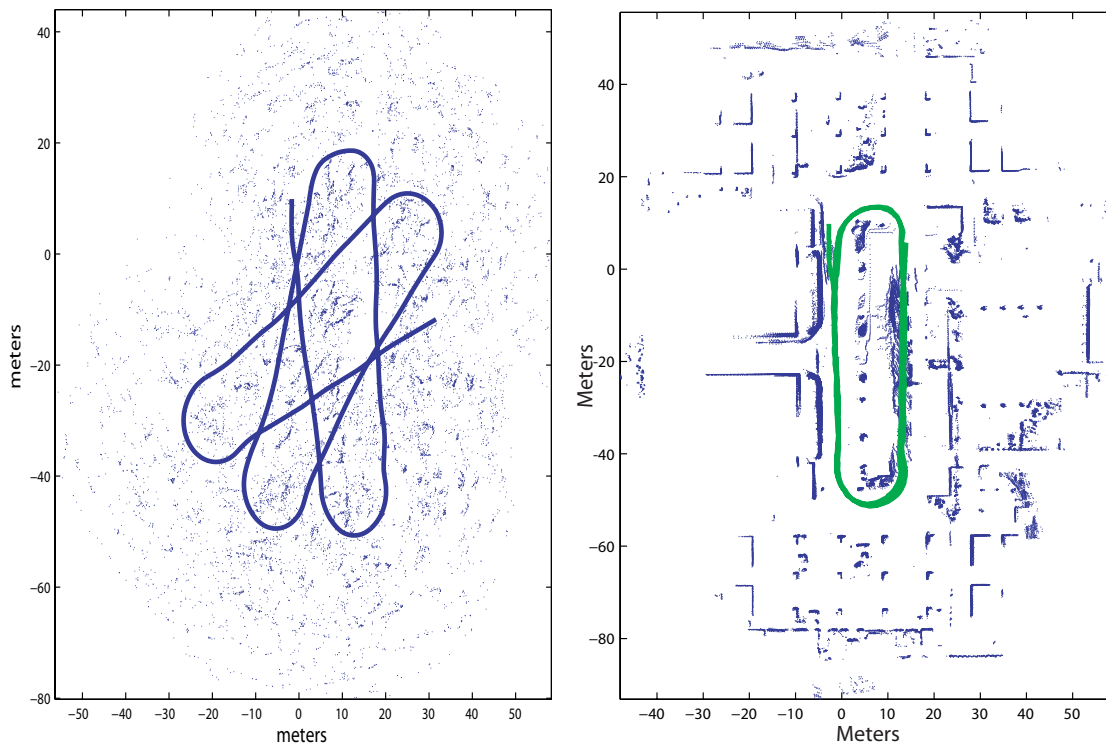

Fig. 12: *Left: Odometry and extracted clusters from the radar data, representing the raw inputs to the SLAM algorithms. Right: The ground truth trajectory (green) obtained by matching laser data due to a lack of GPS data.*

detections are based on 700 scans recorded during a 3 loop trajectory in the car park (Figure 6). The information displayed in this figure can be thought of as the information input to the SLAM algorithms. Given the tree coverage and surrounding buildings in the area, GPS was generally not available. Ground truth was thus obtained by manually matching successive scans from the laser range finders (labelled in Figure 4), with graphical verification also provided in Figure 12 (right). The vehicle was driven at approximately 1.5m/s with a control input frequency of 10Hz and a radar scan frequency of 2.5Hz.

In these experiments, JCBB-EKF-SLAM [22] and MH-FastSLAM [24], both using a mutual exclusion gate and a 95% χ^2 confidence gate, were used as the benchmark vector based comparisons. For each SLAM filter, identical sequences of control inputs and measurements were provided. The RB-PHD-SLAM filter used 50 trajectory particles, while for MH-FastSLAM a maximum limit of 2000 particles (number of hypotheses considered prior to re-sampling) was used.

Given the small-sized loop (with perimeter approximately 170m), the maximum range of the radar was set to 30m and the JCBB-EKF-SLAM, MH-FastSLAM and RB-PHD-SLAM algorithms were executed on the dataset. Figure 13 depicts the posterior estimated trajectory and map using the JCBB-EKF-SLAM algorithm (left) and that from MH-FastSLAM (middle). Finally, Figure 13 (right) shows the SLAM estimates based on the RB-PHD-SLAM algorithm.
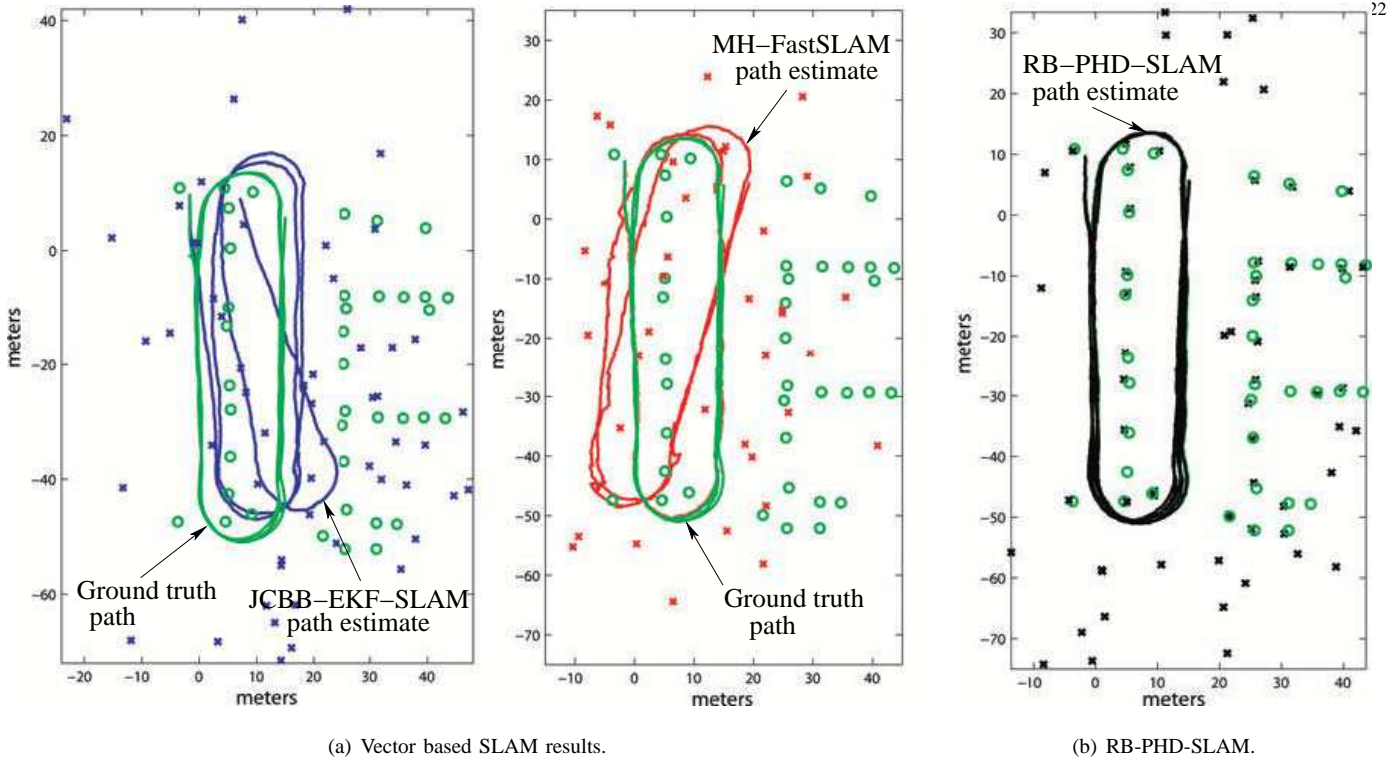
(a) Vector based SLAM results.

(b) RB-PHD-SLAM.

Fig. 13: *The posterior SLAM estimates from JCBB-EKF-SLAM (left) and MH-FastSLAM (middle) and RB-PHD-SLAM (right) using the same dataset. The integrated Bayesian framework for SLAM, incorporating data association and feature management enhances the robustness of the SLAM algorithm given noisy measurements.*

For both the MH-FastSLAM and RB-PHD-SLAM algorithms, the map of the highest weighted trajectory estimate is used as the map estimate. In all the figures, the ground truth and estimated trajectories are labelled and the circles represent the ground truth feature locations. The crosses represent the estimated map. In the case of FastSLAM, this is derived with respect to the maximum a-posteriori (MAP) FastSLAM trajectory estimate (the particle (trajectory) with the final maximum weight). In the case of MH-FastSLAM, the trajectory "MH-FastSLAM path estimate" indicates the MAP trajectory estimate. Note that the RB-PHD-SLAM map yields feature and vehicle trajectory estimates which are much closer to the ground truth map than either JCBB-EKF-SLAM or MH-FastSLAM. This is due to its unique ability to jointly estimate both feature number as well as location, and its immunity to feature management and association errors.

B. Computational Cost Comparisons

The computational complexity of RB-PHD-SLAM is, $\mathcal{O}(\mathfrak{m}_k \mathfrak{z}_k N)$ i.e. linear in the number of features (in the FoV), linear in the number of measurements and linear in the number of trajectory particles. For a single thread implementation, Figure 14 shows that the computational time of RB-PHD-SLAM at each time index k is comparable with that of the MH-FastSLAM algorithm, both of which are less expensive than JCBB-EKF SLAM as its hypothesis tree grows in the presence of high clutter [4]. Note that due to the Rao-Blackwellized structure of RB-PHD-SLAM, binary tree based enhancements, such as those applied to traditional FastSLAM [20], can be readily developed to further reduce the complexity to $\mathcal{O}(\mathfrak{z}_k N \log(\mathfrak{m}_k))$.
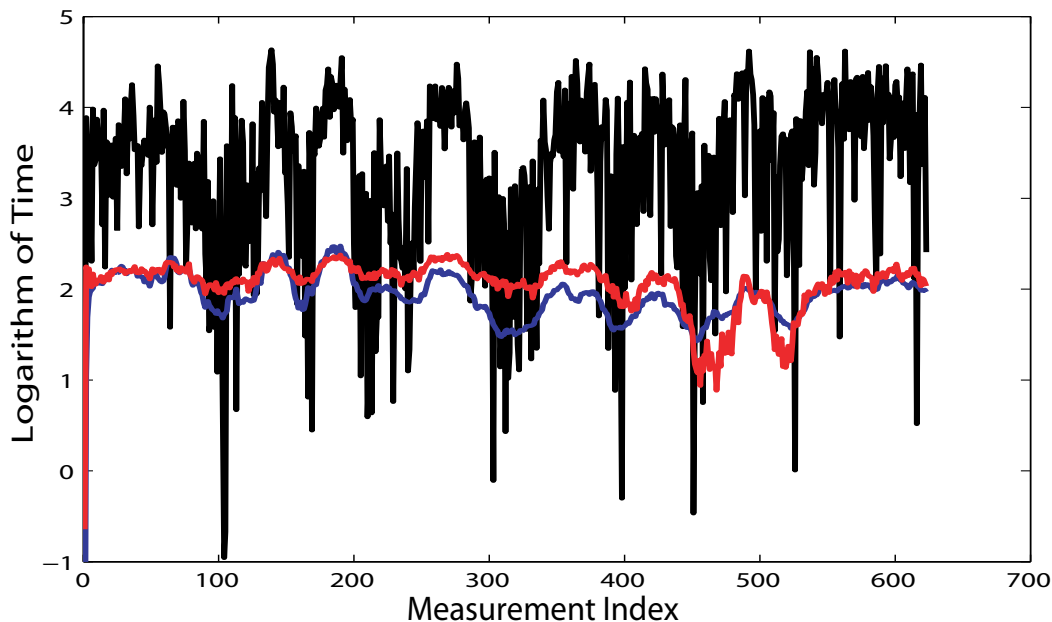
Fig. 14: *A comparison of the computation time per measurement update for RB-PHD-SLAM (blue), MH-FastSLAM (red) and JCBB-EKFSLAM (black).*

Furthermore, in contrast to data association based methods, the proposed approach admits numerous other computational enhancements, since the map PHD update can be segmented, executed in parallel and subsequently fused for state estimation.

X. Summary

In probabilistic SLAM algorithms, the concept of feature existence uncertainty is at least as important as a feature's spatial uncertainty. This tutorial high-lighted the importance of sensor models which are capable of incorporating both uncertainties, and demonstrated that such a model should be based on an RFS, rather than the commonly used vector representation. Further, to correctly take into account the facts that sensor based feature detections are typically not detected in the order in which their corresponding elements are stored in a map vector, and that they may differ in number, the map itself is also more appropriately modelled as an RFS. The recent development of FISST-based SLAM techniques, then allows mathematical representations of the map RFS to be propagated through Bayes theorem, allowing the estimation of feature locations and number in a joint manner. By taking into account prior knowledge of sensor based probabilities of detection, false alarm and spatial uncertainty parameters, the RB-PHD-SLAM approximation to Bayes theorem was demonstrated, circumventing the need for the fragile map management and feature association rules necessary in vector based SLAM. Its superiority over state of the art vector based methods, JCBB-EKF-SLAM and MH-FastSLAM, was demonstrated with short range radar and the presence of many clutter measurements. Since standard vector based methods inherently assume unity probabilities of detection and zero probability of false alarm, they are dependent on heuristics and routines, which occur outside of the Bayesian vector state update, to force the map and measurement state vectors to be of compatible dimensions before state update is possible. This restriction is removed when FISST representations of RFSs are adopted, as was demonstrated with the PHD filter.

The RFS concepts offer many avenues for future research, including the implementation of higher order FISST-based algorithms, such as the Cardinalized-PHD and Multi-target multi-Bernoulli Filters. Extended landmarks, rather than the point

features considered here, should also be examined together with the possibilities of joint mapping and landmark tracking from an autonomous vehicle.

REFERENCES

[1] M.W.M.G. Dissanayake, P. Newman, S. Clark, H.F. Durrant-Whyte, and M. Csorba. A solution to the simultaneous localization and map building (SLAM) problem. *IEEE Trans. Robotics and Automation*, 17(3):229–241, June 2001.

[2] B.N. Vo and W.K. Ma. The Gaussian mixture probability hypothesis density filter. *IEEE Transactions on Signal Processing*, 54(11):4091–4104, November 2006.

[3] J. Mullane, B.N. Vo, M.D. Adams, and B.T. Vo. A random-finite-set approach to Bayesian SLAM. *IEEE Transactions on Robotics*, 27(2):268–282, April 2011.

[4] J. Mullane, B.-N. Vo, M. Adams, and B.-T. Vo. *Random Finite Sets for Robot Mapping and SLAM*. Springer Tracts in Advanced Robotics 72. Springer, Berlin Heidelberg, 2011.

[5] R. Mahler. *Statistical Multisource Multitarget Information Fusion*. Artech House, 2007.

[6] R. Smith, M. Self, and P. Cheeseman. A Stochastic Map for Uncertain Spatial Relationships. In *The Fourth International Symposium of Robotics Research*, pages 467–474, 1987.

[7] H.F. Durrant-Whyte, S. Majumder, M. de Battista, and S. Scheding. A Bayesian algorithm for simultaneous localisation and map building. In *The Tenth International Symposium of Robotics Research (ISRR)*, Victoria, Australia, 2001.

[8] M. Montemerlo and S. Thrun. FastSLAM: A factored solution to the simultaneous localization and mapping problem. In *Eighteenth national conference on Artificial intelligence*, pages 593–598, Menlo Park, CA, USA, 2002. American Association for Artificial Intelligence.

[9] F. Lu and E. Milios. Robot Pose Estimation in Unknown Environments by Matching 2D Range Scans. *Journal of Intelligent and Robotic Systems*, 18(3):249–275, 1997.

[10] S. Thrun. Simultaneous localization and mapping with sparse extended information filters. *The International Journal of Robotics Research*, 23, 2004.

[11] J. Mullane, B.N. Vo, M. Adams, and W.S. Wijesoma. A random set formulation for Bayesian SLAM. In *proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, France, September 2008.

[12] J. Mullane, B.N. Vo, M. Adams, and W.S. Wijesoma. A random set approach to slam. In *proceedings of the IEEE International Conference on Robotics and Automation (ICRA) workshop on Visual Mapping and Navigation in Outdoor Environments*, Japan, May 2009.

[13] C.S. Lee, D.E. Clark, and J. Salvi. SLAM with single cluster PHD filters. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 2096–2101, St. Paul, Minnesota, May 2012.

[14] G. Brooker, C. Lobsey, and R. Hennessy. Low cost measurement of small boat rcs at 94GHz. In *Proceedings of IEEE International Conference on Control, Automation and Robotics*, 2006.

[15] J. Mullane, M. Adams, and W.S. Wijesoma. Robotic mapping using measurement likelihood filtering. *International Journal of Robotics Research*, 2(28):172–190, 2009.

[16] A. Foessel, J. Bares, and W.R.L. Whittaker. Three-Dimensional Map Building with MMW RADAR. In *Proceedings of the 3rd International Conference on Field and Service Robotics*, Helsinki, Finland, June 2001.

[17] M. Adams, J. Mullane, E. Jose, and B.N. Vo. *Robotic Navigation and Mapping with Radar*. Artech House, Boston, USA, 2012.

[18] D. Fox, W. Burgard, and S. Thrun. Markov localization for mobile robots in dynamic environments. *Journal of Artificial Intelligence Research*, 11:391–427, 1999.

[19] T.D. Gillespie. *Fundamentals of Vehicle Dynamics*. Society of Automotive Engineers, Inc., Warrendale. PA, USA, 1992.

[20] M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit. Fastslam: A factored solution to the simultaneous localization and mapping problem. In *Proc. AAAI National Conference on Artificial Intelligence*, pages 593–598, 2004.

[21] J. Mullane, B.N. Vo, and M. Adams. Rao-blackwellised PHD SLAM. In *proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Alaska, USA, May 2010.

[22] J. Niera and J.D. Tardos. Data association in stochastic mapping using the joint compatibility test. *IEEE Transactions on Robotics and Automation*, 17:890–897, December 2001.

[23] D. B. Reid. An algorithm for tracking multiple targets. *IEEE Transactions on Automatic Control*, 24:843–854, 1979.

[24] J. Nieto, J. Guivant, E. Nebot, and S. Thrun. Real time data association for fastSLAM. In *IEEE Internationlal Conference on Robotics and Automation*, volume 1, pages 412–418, September 2003.