Online Probabilistic Change Detection in Feature-Based Maps

Fernando Nobre, Christoffer Heckman, Paul Ozog, Ryan W. Wolcott and Jeffrey M. Walls

Abstract— Sparse feature-based maps provide a compact representation of the environment that admit efficient algorithms, for example simultaneous localization and mapping. These representations typically assume a static world and therefore contain static map features. However, since the world contains dynamic elements, determining when map features no longer correspond to the environment is essential for long-term utility. This work develops a feature-based model of the environment which evolves over time through feature persistence. Moreover, we augment the state-of-the-art sparse mapping model with a correlative structure that captures spatio-temporal properties, e.g. that nearby features frequently have similar persistence. We show that such relationships, typically addressed through an ad hoc formalism focusing only on feature repeatability, are crucial to evaluate through a probabilistically principled approach. The joint posterior over feature persistence can be computed efficiently and used to improve online data association decisions for localization. The proposed algorithms are validated in numerical simulation and using publicly available data sets.
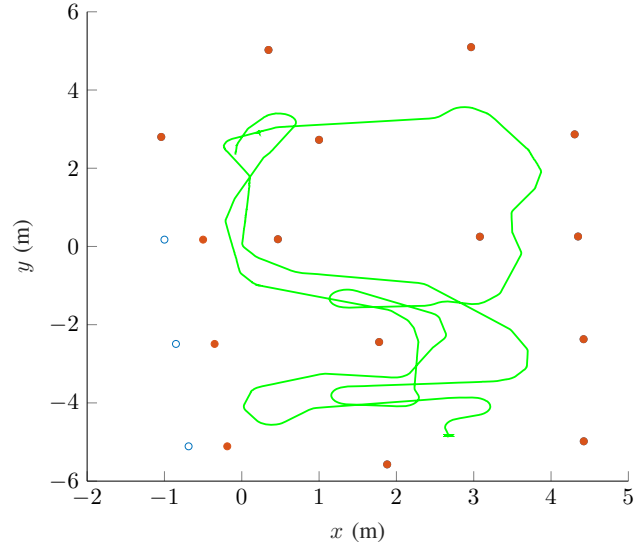
Fig. 1: Localizing against a stale map: the green trajectory shows the estimated poses from range/bearing and odometry measurements, given a prior map which is no longer current (circles represent incorrect map elements).

I. INTRODUCTION

In the context of mobile robots and autonomous driving, accurate high resolution *spatial awareness* is necessary for successfully navigating an environment. This can be cast as a *data fusion* problem in which noisy measurements from sensors undergoing uncertain dynamic motions must be combined into a single underlying state estimate. This problem is typically addressed through a procedure known as Simultaneous Localization and Mapping (SLAM). Some applications, such as autonomous driving, require high fidelity state estimates which need to be robust to sensor and environmental changes. Current SLAM algorithms can be fragile in two aspects: algorithmic foundations and hardware robustness. The former includes failure modes induced by limitations in current SLAM algorithms (i.e. difficulty handling dynamic environments); the latter includes failures due to sensor degradation. We focus on one of these algorithmic limitations: the reliance on static maps. In general, one cannot make simplifying assumptions such as the existence of a static world on arbitrary real-world environments; stop signs are removed, buildings change appearance, and road construction is pervasive. Explicitly addressing this failure mode is critical for safe long-term operation.

It is impossible to talk about algorithmic failures without mentioning data association. Data association matches each measurement to the portion of the state the measurement

refers to (e.g. associating a visual feature to a specific landmark in visual SLAM). Incorrect associations can quickly cause the SLAM estimate to diverge [1]. This is especially critical in feature-based maps which lack a notion of appearance (i.e visual or structural descriptor).

In the static-world case, perceptual aliasing makes data association a challenging problem; this problem is worsened by the presence of unmodeled dynamics in the environment, which include both short-term and seasonal changes. It is fairly common for current SLAM approaches to make a *static world assumption*, which holds true for independent, short mapping runs in small-scale environments. However, when mapping in large environments over long periods of time, change is inevitable.

Change in the environment can be especially difficult to detect when dealing with an increased amount of clutter or when the change is subtle. For example, lane markings may be re-painted a few centimeters from the original position due to construction, or traffic signs may be slightly re-located. These scenarios are especially susceptible to incorrect data associations resulting in localization error. We present a robust solution to detecting changes in feature-based maps which leverages information about both single features and neighboring features to produce a globally consistent belief over individual and joint feature persistence. Figure 1 shows an example of robust localization against a stale map: the
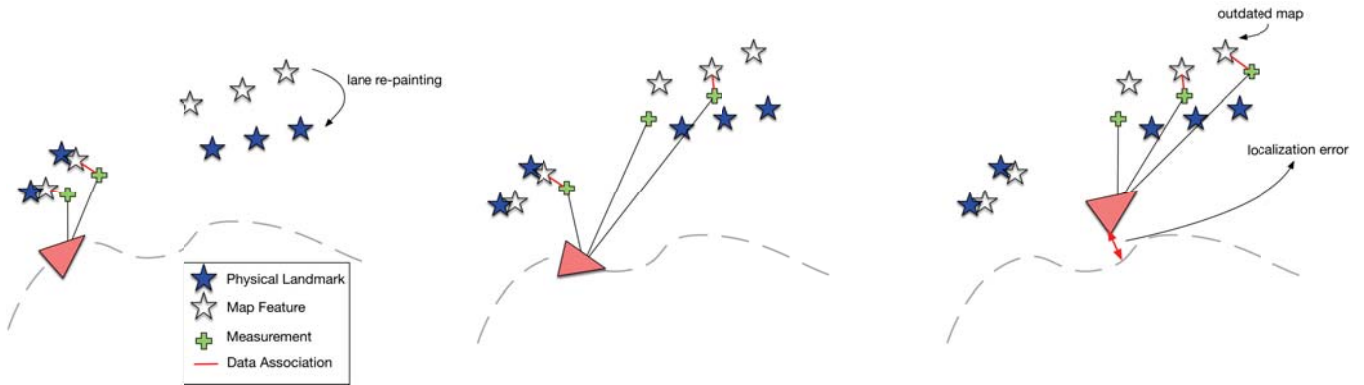
Fig. 2: Graphical example of the problem: in the left figure, the measurements are correctly associated to map features; in the middle figure, some landmarks have moved, for example lane markings that were re-painted (blue stars), and the map is outdated (white stars). Here, one measurement (green cross) is correctly identified as a new measurement, while the other is incorrectly associated to the outdated map feature due to landmark clutter, pose and measurement uncertainties. In the right figure, the incorrect data association causes the least-squares solution to converge to an incorrect estimate. Our proposed solution leverages the correlation between the three landmarks which moved to estimate the joint feature persistence and improve data associations.

localization estimate remains consistent even though some elements of the map have moved.

We extend the work in [2], which introduced the notion of a Bayesian filter to model feature persistence in a time-varying feature-based environmental model. We differ from [2] by proposing a general formulation for persistence which takes into account correlation between features. We focus on sparse feature based maps with no assumptions on sensor-specific feature descriptors (i.e visual descriptors) which is broadly applicable to any sparse feature-based representation. We are interested in estimating the existence of each feature in the map and bounding the localization error due to incorrect data associations, we show that by capturing the underlying structure of the environment we are able to make better informed decisions on data associations. In summary, we propose a novel joint probabilistic formulation over feature persistence. We show that a joint formulation over feature persistence can be made informative by imposing or learning the structure of the environment. We further show that the joint and marginal persistence estimates are amenable to constant-time operation. Finally, we demonstrate that by incorporating the joint belief over feature persistence in the data-association step, we are able to perform robust localization even in the presence of hard-to-detect changes (e.g. small changes). We demonstrate the benefits of estimating map persistence in a graph-SLAM [3], [4] implementation, potentially enabling long-term autonomous applications which are robust to arbitrarily small map changes.

II. PRIOR WORK

The challenge of dealing with dynamic and semi-static environments is a recurring problem in the robotics community and has been addressed from multiple fronts. The principal challenges with semi-static environments are the need to *detect* a change in the environment and *update* an existing map so that it reflects the most current state of the world. One way to tackle this problem is to use environmental representations that are suited for dynamic environments. One such representation is the seminal work

by Biber and Duckett [5], [6] who update a sparse map built from 2D laser scans by randomly selecting a fixed fraction of the scans every revisit to update the prior map. Morris et al. presents a multiple-map approach [7] where many map instances are stored and the one best fitting the current set of sensor measurements is used, an approach suitable for environments with a discrete set of possible configurations. Other approaches [8] model each "place" as a set of experiences which has proven to be robust to drastic seasonal changes. The main difference between these methods and what we propose is that while one targets *localization* we are interested in producing a geometrically and temporally consistent representation of the environment suitable for continued localization.

Other approaches [9], [10], [11] are capable of recovering a geometrically-consistent map robust to dynamic environments. However these solutions are tailored to specific sensor modalities, such as the work in [11] which models places as a collection of camera images, connected by 6-DOF transformations between camera poses. Another example is [9], which proposes the *Dynamic Pose Graph SLAM* but limits its use to 2D laser scanners and lacks an underlying probabilistic model for reasoning about change while abstracting the sensor modalities.

The Occupancy Grid is also a common choice for environmental representation, as proposed by Meyer-Delius et al. [12]. This work uses a *dynamic occupancy grid* which adapts the classic occupancy grids [13] to dynamic environments by modeling each cell as a stationary two-state Markov process. Other works such as [14] also propose a form of dynamic occupancy grids; [15] incorporated the occupancy grid model into a particle filtering framework for a Bayesian model-based mapping solution. These methods have in common the restriction of representing the environment by its volumetric geometry, which limits their accuracy in cases of interest such as when using visual appearance information for mapping.

Recently there has been some work on semi-static feature-abstracted environments such as Krajnk et al. [16] with

Fourier analysis for predicting future states but that is designed for *periodic* activity. The work most similar to this paper was presented by Rosen et al. [2] which proposes an information-theoretic formulation for feature persistence taking into account sensor errors. However, this approach ignores potential correlations between feature persistence by assuming each feature persistence is marginally independent. This is not validated with any real data; furthermore, the impact of the persistence model on improving the map and data associations is not addressed. In contrast to these prior works, we present a general unified formulation for feature persistence that captures potential correlation between features, while maintaining a tractable posterior for constant-time estimation and showing the necessity of modeling feature persistence jointly for robust localization.

III. METHODOLOGY

We cast feature persistence estimation within the context of probabilistic SLAM. The quantities of interest that are directly useful for higher-level tasks are the robot's pose in time and the positions of the map features; these quantities will be represented by the state vector $\mathcal{X} = [\mathbf{x_p}, \mathbf{x_{l_{1:M}}}]$ where $\mathbf{x_p}$ are all the poses and $\mathbf{x_l}$ are the M map features. Given a set of sensor measurements $\mathcal{Z}$, the Maximum-a-Posteriori (MAP) SLAM problem is to maximize the posterior $p(\mathcal{X}|\mathcal{Z})$. However this problem as stated is intractable since it requires summing over all possible data associations, an intractably large problem due to its combinatorial nature [17]. Letting J be the vector of all data association hypothesis one might then wish to estimate the optimal data association vector $\arg\max_J p(J|Z)$; however the difficulty in evaluating the likelihood of a spurious measurement (i.e. the likelihood of seeing a new feature) makes this approach undesirable. The usual solution is to solve for data associations with a search over possible associations, using techniques such as Joint Compatibility Branch and Bound (JCBB) [1] and then condition the state estimate on data associations: $p(\mathcal{X}|\mathcal{Z}, J)$ where J is the vector of data associations that assign a feature in the state vector to a measurement. If we drop the static-world assumption and allow features to have an associated "survival time," an additional set of discrete random variables $\Theta^t \in \{0, 1\}$ which represent if a feature exists at a specified time t. An appealing approach would be to jointly estimate data associations and feature persistence $p(J, \Theta|\mathcal{Z})$, however we run into the same problem of evaluating the likelihood of a spurious measurement. Therefore we take a similar approach to estimating the state vector $\mathcal{X}$ and condition the feature persistence on the output of the data association step $p(\Theta^t|J)$. This is done by first estimating J using a data association technique such as JCBB, which is then used to estimate Θ.

A. Feature Persistence Model

We follow the survivability formulation introduced in [2], which we will briefly describe here. Each feature i in the map has a latent "survival-time" $T_i \in [0, \infty)$ which represents the

time when feature i ceases to exist, as well as a persistence variable Θ_i^t:

$$T_i \sim p_{T_i}(\cdot)$$
$$\Theta_i^t | T_i = \begin{cases} 1, & t \leq T_i \\ 0, & t > T_i \end{cases}, \tag{1}$$

where Θ_i^t is a boolean random variable representing whether feature i exists at time t, and $p_{T_i}(\cdot)$ encodes some prior distribution over the survival time T_i. We are interested in estimating for each feature i, its marginal persistence probability $p(\Theta_i^t = 1|J^{1:N})$, where $J^{1:N}$ are all the feature detections collected until time t_N. Note that for a map with M features $J^{1:N} = \{J_1^{1:N}, ..., J_M^{1:N}\}$ with $J_k^{1:N} \triangleq \{j_k^{t_i}\}_{i=1}^N$, $k \in [1, M]$, $j \in [0, 1]$. The feature detections are the output of the data-association step, indicating if feature k was detected at time t_i.

B. Estimating Feature Persistence

We are interested in estimating the full joint feature persistence posterior at a certain time t, given all data association decisions from time t_1 to t_N:

$$p(\Theta^t = 1|J^{1:N}), \tag{2}$$

where Θ^t is the joint persistence over all M map features $\Theta^t \triangleq \{\Theta_1^t, ..., \Theta_M^t\}$ at time $t \in [t_N, \infty)$. This implies that we are interested in estimating the joint posterior probability over feature existence in the present and future, given the sequence of detections for all features. It is important to note that we only estimate the persistence probability for times equal to or greater than the last received measurement t_N. Noting that $p(\Theta^t = 1|J^{1:N}) = p(T \geq t|J^{1:N})$, with $T \triangleq \{T_i\}_{i=1}^M$ the vector of survival times for all M map features and using Bayes' Rule to compute the posterior probability in (2)

$$p(\Theta^t = 1|J^{1:N}) = \frac{p(J^{1:N}|T \geq t)p(T \geq t)}{p(J^{1:N})}. \tag{3}$$

We will now derive a closed-form expression for evaluating each of the joint posterior terms, starting with the joint detection likelihood $P(J^{1:N}|T \geq t)$. We make the assumption that a sequence of detections $J^{1:N}$ of feature i depend only on the feature i itself; that is, $p(J_i^{1:N}|T \geq t) = p(J_i^{1:N}|T_i \geq t)$

$$p(J^{1:N}|T \geq t) = \prod_{i=1}^M p(J_i^{1:N}|T_i \geq t)$$
$$= \prod_{i=1}^m \prod_{k=1}^n p(j_i^{t_k}|T_i \geq t), \tag{4}$$

where we are also making the assumption that the sequence of detections $\{j_i^{t_1}, j_i^{t_2}, \ldots, j_i^{t_N}\}$ for feature i is conditionally independent from other detections, given the persistence Θ_i^t. The intuition behind this is that given existence of a feature, its sequence of detections should not depend on other features. We are still left with evaluating the individual measurement likelihood $P(j_i^t|T_i)$ which is the
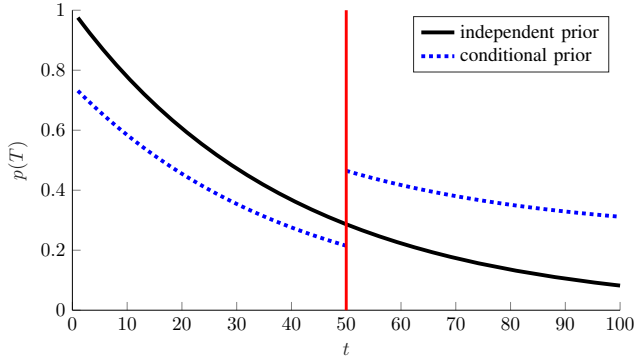
Fig. 3: $p(T_i)$ (solid black line) is the independent prior on the survival time T_i modeled as an exponential decaying function. Given some correlated feature k with $T_k \geq 50$ we model $p(T|T_k \geq 50)$ (dotted blue line) with weights $\pi_i = 0.75$ $\pi_k = 0.25$, which re-enforces our prior belief over T_i if $T_i \geq 50$

probability of detecting a feature given its survival time T_i. If data-associations were always perfect this would simply be 1 if $T_i \geq t$ and 0 if $T_i < t$. Since that is not the case, we follow the formulation in [2] and define a *probability of missed detections* P_M and *probability of false alarm* P_F. Using the model defined in (1)

$$p(j_i^t|T_i) = \begin{cases} P_M^{(1-j_i^t)}(1-P_M)^{j_i^t}, & T_i \geq t \\ P_F^{j_i^t}(1-P_F)^{(1-j_i^t)}, & T_i < t \end{cases}, \quad (5)$$

where P_M models the probability that the feature exists but was not detected, and P_F the probability that the feature no longer exists but was detected, which may happen due to incorrect data associations or spurious measurements. These quantities are dependent on a series of factors such as the amount of clutter in the environment, the data-association process, and therefore the state uncertainty. We present P_M and P_F as constants, but they could be modified per observation, e.g., to include occlusions.

The core contribution of this work is in the modeling and evaluation of $p(\Theta^t)$, the joint prior distribution over feature persistence (at time t). In the trivial case each feature is independent and $p(\Theta^t) = \prod_{i=1}^{M} p(\Theta_i^t)$ however we cannot realistically make that independence assumption. In many environments the existence of one feature is clearly correlated to the existence of other features and exploiting that correlation is a crucial aspect estimating jointly consistent persistence. However, tracking the full joint distribution $P(\Theta^t)$ over all features is intractable as it grows 2^M with M map features. However, if we impose some structure to the environment (e.g. a feature drawn from a curb in the road is not affected if a sign-post is removed, however it is strongly correlated to other features in the same curb) the complexity of computing the joint prior is bounded by the maximum number of correlated features. Such structure is justified based on the intuition that the existence of a feature is only strongly correlated to a subset of the map. The formulation proposed in [2] makes the assumption that the persistence for each feature is marginally independent, which is a specific instance which falls out of the general

formulation in (3). We propose exploiting the underlying structure of the environment to leverage the correlation between features while still maintaining a tractable posterior. Imposing some structure on the feature map such that there exists a set of $L \leq M$ cliques $\{\tau_i\}_{i=1}^L$, with $\tau_i^t \subset \Theta^t$ and features $\mathfrak{z} \in \tau$, the joint posterior factors into

$$p(\Theta^t) \approx \prod_{i=1}^{L} p(\tau_i^t). \quad (6)$$

Given the assumption that features associated in a clique τ_i have strongly correlated persistence, the joint prior $p(\tau_i^t)$ can be approximately decomposed into:

$$p(\tau_i^t) \approx \prod_{\mathfrak{z} \in \tau} p(\theta_\mathfrak{z}^t|\theta_k^t) = \prod_{\mathfrak{z} \in \tau} p(T_\mathfrak{z}|T_k \geq t) \ \forall k \in \tau, \quad (7)$$

which states that for any feature k in a clique τ_i, the joint distribution $p(\tau_i)$ can be approximately factored into a product of conditional distributions on feature $\mathfrak{z}_k \in \tau_i$. The intuition behind this approximation is that for a set of correlated features (e.g a set of features all drawn from the same rigid body), conditioning on a single feature from that rigid body adds approximately the same amount of information as conditioning on all the features. The conditional prior $p(T_\mathfrak{z}|T_k \geq t)$ is defined as

$$p(T_\mathfrak{z}|T_k \geq t) \triangleq \begin{cases} \pi_\mathfrak{z} p(T_\mathfrak{z} < t_i) & t_i < t \\ \pi_\mathfrak{z} p(T_\mathfrak{z} \geq t_i) + \pi_k & t_i \geq t \end{cases}, \quad (8)$$

with $\sum_j \pi_j \triangleq 1$ pairwise weights associated to each feature pair in τ_i.

Having computed the likelihood and the prior from (3), we are left with computing the marginal measurement probability or evidence $P(J^{1:N})$. We leverage the fact that the detection likelihood is constant in the intervals between detections and the factorization of the full joint posterior into L cliques. Combined with a way to evaluate the cumulative distribution function of the survival time prior $p(T \leq t) \triangleq F_T(t)$, where $F_T(t)$ is the c.d.f. of the survival time prior as described in [2] to derive a closed-form expression for the evidence; defining $t_0 \triangleq 0$ and $t_{N+1} \triangleq \infty$

$$\begin{aligned} p(J^{1:N}) &= \prod_{i=1}^{L} p(J_{\tau_i}^{1:N}) \\ &= \prod_{i=1}^{L} \left(\prod_{\mathfrak{z} \in \tau_i} \left(\sum_{\mathfrak{u}=0}^{N} p(J_\mathfrak{z}^{1:N}|t_\mathfrak{u}) \int_{t_\mathfrak{u}}^{t_{\mathfrak{u}+1}} p(T_{\tau_i}) \right) \right), \end{aligned}$$
$$(9)$$

where we first decomposed the full joint evidence into the product of the cliques τ_i, then further decomposed each clique into the product of its individual terms, which are tied together by the joint prior $P(T_{\tau_i})$. We make use of (7) to write out the integral over joint prior survival times as a product of conditional distributions on one element of the clique $p(T_{\tau_i}) = \prod_{\mathfrak{z} \in \tau} p(T_\mathfrak{z}|T_k)$ where each conditional prior can be evaluated according to (8).
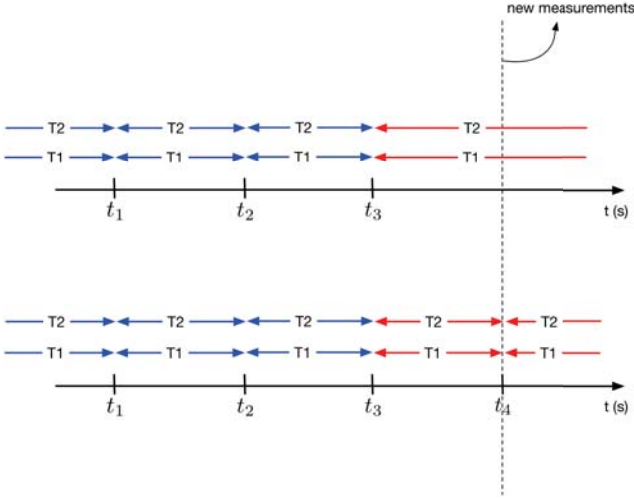
Fig. 4: Terms which need to be computed for the marginal evidence with two features $[T_1, T_2]$. The segments represented by arrows are the the possibility that each feature's survival time T is within that range. Each t_i represents the time in which a new detection was received. Computing the joint evidence requires summing over the detection likelihood for all measurements given every configuration of T_1, T_2. When a new measurement is received (dotted vertical line) at time t_4 we only need to recompute terms associated with the segments highlighted in red.

1) Marginal Formulation: Suppose we wish to estimate the marginal persistence for a feature a which is correlated to another feature b, given a sequence of N detections $J_a^{1:N}$, $J_b^{1:N}$ from time $t \in [t_1, t_N]$. We may make the reduction:

$$
\begin{aligned}
p(\Theta_a^t = 1 | J_a^{1:N}, J_b^{1:N}) = \\
& p(T_a \geq t | J_a^{1:N}, J_b^{1:N}) \\
&= \int_0^\infty p(T_a \geq t, T_b | J_a^{1:N}, J_b^{1:N})\, dT_b \\
&= \int_0^\infty \frac{p(J_a^{1:N}, J_b^{1:N} | T_a \geq t, T_b) \cdot p(T_a \geq t, T_b)}{p(J_a^{1:N}, J_b^{1:N})}\, dT_b \\
&= \frac{p(J_a^{1:N} | t_N) p(T_a \geq t)}{p(J_a^{1:N}, J_b^{1:N})} \times \\
& \int_0^\infty p(J_b^{1:N} | T_b) p(T_b | T_a \geq t)\, dT_b,
\end{aligned}
$$
(10)

where we use the fact that $p(J^{1:N}|T)$ is constant in the intervals $[t_i, t_{i+1}]$ to define the integral in (10) as

$$
\begin{aligned}
\int_0^\infty p(J_b^{1:N} | T_b) p(T_b | T_a \geq t)\, dT_b \\
= \sum_{i=0}^N p(J_b^{1:N} | t_i) \int_{t_i}^{t_{i+1}} p(T_b | T_a \geq t)\, dT_b.
\end{aligned}
$$
(11)

Note that in the case where features a and b are independent, $p(T_b | T_a \geq t) = p(T_b)$ and (11) simply becomes the marginal $p(J_b^{1:N})$ which cancels part of the evidence in (10); this results in exactly the posterior defined in [2].

C. Feature Correlation Design

In this section we describe how to design the weights π in (8). Since there is no inherent structure to the sparse feature-based environmental representation, we design a prior structure that aims to capture the underlying structure of the environment. It is possible to learn feature correlations from the sensor data used to create the feature map (i.e image frames or point clouds) using a object detector to semantically segment the environment. However the original sensor data used to create the map is not always readily available. We deal with that scenario, where the only input to designing feature correlation is the sparse feature map itself.

We model features which where observed at a similar point in time, and are physically close to have correlated persistence. The intuition is that if a set of features is co-observed and geometrically close, the likelihood that they belong to the same semantic object (e.g. lane markings, sign post) is high. We define the set $\mathfrak{N}$ of all features that were observed within Δs, and within that set we employ a Euclidean nearest neighbors metric to group features in cliques of up to $n = 5$ features, which are within a maximum distance d_{max} to the center of the clique. When applied to sparse feature-based maps in which clutter is reduced this method is a general way of capturing the underlying scene structure. The weights π between features are then computed as the inverse Euclidean distance

$$
\pi_{ij} = \frac{1}{\|\hat{\mathcal{X}}_i - \hat{\mathcal{X}}_j\|},
$$
(12)

where π_{ij} is the normalized weight such that $\sum_j \pi_{ij} = 1$. Computing the set of cliques and their corresponding weights can be performed offline. Figure 3 demonstrates the effect the weights π have on the prior distribution $p(T_i)$; The conditional distribution $p(T_i | T_j > t)$ with given weights π_{ij} is a mixture model with reduced probability mass before t. In the case of independent features $\pi_{ii} = 1$ and Figure 3 becomes the solid line.

D. Recursive Estimation

Computing the full posterior in (3) every time a new observation is included is computationally expensive, in this section we show how to compute both the joint distribution $P(\Theta^t | J^{1:N})$ and the marginal $P(\Theta_i^t | J^{1:N})$ as described in Section III-B.1 in a recursive manner by re-using previously computed terms. This allows for a constant-time update when a new detection is received. When a new observation $j^{t_{N+1}}$ is appended to the observation vector $J^{1:N}$, $t_{N+1} > t_N$ and given the independence assumption in (4), the updated joint likelihood is

$$
p(J^{1:N+1} | T \geq t) = \prod_{i=1}^M \prod_{k=1}^{N+1} p(j_i^{t_k} | T_i \geq t).
$$
(13)

So the updated joint likelihood at time t_{N+1} can be written in terms of the previous at time t_N

$$
p(J^{1:N+1} | T \geq t) = p(J^{1:N} | T \geq t) \prod_{i=1}^M p(j_i^{N+1} | T_i \geq t).
$$
(14)

Updating the joint evidence is less straightforward due to having to integrate over all possible survival times for all features. However using the decomposition of the joint prior distribution in (7) where we have L cliques τ_i, when a new measurement at t_{N+1} is observed, we can compute the updated joint evidence as

$$p(J^{1:N+1}) = \prod_{i=1}^{L} p(J_{\tau_i}^{1:N+1}). \tag{15}$$

If a clique τ_k has τ_M features the clique evidence $p(J_{\tau_k}^{1:N})$ can be computed as

$$p(J_{\tau_k}^{1:N}) = \sum_{i=0}^{N-1} p(J_{\tau_k}^{1:N}|t_i)p(t_{i-1} \le T_{\tau_k} < t_i) \\ + p(J_{\tau_i}^{1:N}|T_{\tau_i} \ge t_N), \tag{16}$$

which implies that the updated *partial evidence*, which excludes the last term in the sum (where the survival time is T_{N+1} after incorporating the measurement at time t_{N+1} is

$$p_L(J_{\tau_k}^{1:N+1}) \triangleq \left[\left(\sum_{i=0}^{N-1} p(J_{\tau_k}^{1:N}|t_i)p(t_{i-1} \le T_{\tau_k} < t_i) \right) \\ + \left(p(J_{\tau_k}^{1:N}|t_N)p(t_N \le T_{\tau_k} < t_{N+1}) \right) \right] \\ \times \prod_{i=1}^{\tau_M} p(j_i^{N+1}|T_i < t_{N+1}). \tag{17}$$

The full updated evidence can be obtained by combining the partial evidence from (17) with the measurement probability given feature survival times at time $T = t_{N+1}$

$$p_(J_{\tau_k}^{1:N+1}) = p_L(J_{\tau_k}^{1:N+1}) + \prod_{i=1}^{\tau_M} p(j_i^{N+1}|T_i \ge t_{N+1}) \tag{18}$$

The updated joint distribution can thus be computed for each clique τ_i using Eqs. (13), (15) and computing the prior using (7). Figure 4 shows the terms that need to be updated every time a new measurement is incorporated. In summary, we use the fact that the measurement likelihoods are constant between observations (e.g. between t_i and t_{i+1}) to discretize the integral over survival times T, and that when a new measurement is incorporated at time t_{N+1} the terms corresponding to survival times before t_N can simply be updated my multiplying by the probability of seeing measurement j^{N+1} given that the survival time for that feature is before t_{N+1} (segments in blue in Figure 4). The marginal persistence probability can be updated in an analogous manner and is omitted due to space constraints.

IV. ROBUST DATA ASSOCIATIONS AND LOCALIZATION

In this section we briefly present a common data association technique and show how our work estimating feature persistence can disambiguate the data association problem in the presence of semi-static scene elements. The *maximum likelihood* (ML) or *individual compatibility* solution to data association is based on probabilistic methods and can be summarized as taking into account the uncertainties between the robot's location and the landmark position in the map. It can be interpreted as a simple nearest neighbors data association, but with Euclidean distance replaced by the Mahalanobis distance. The standard approach is to model each measurement $z \sim \mathcal{N}(h(\mathcal{X}), \Gamma)$ as a Gaussian distribution with given covariance Γ. Taking the negative logarithm we obtain the maximum likelihood cost function

$$D^2 := ||h(\hat{\mathcal{X}}) - z||_{\Lambda}^2, \tag{19}$$

where D is the Mahalanobis distance and the covariance Λ is defined as

$$\Lambda := \frac{\partial h}{\partial \mathcal{X}} \Sigma \frac{\partial h}{\partial \mathcal{X}}^T + \Gamma, \tag{20}$$

with Σ the current state uncertainty. The hypothesis that a given measurement z was caused by the jth landmark can be evaluated based on a chi-square acceptance decision $D < \chi_{d,\alpha}^2$ where α is the desired confidence level and d the dimension of the measurement. Instead of considering every map feature as a candidate for data association, we use the marginal persistence estimate for each feature to weight the data associations. We are specifically interested in avoiding incorrect data associations when there are multiple hypothesis with similar cost; in the scenario where we consider feature persistence independently and the pose uncertainty is non-negligible, it is impossible to distinguish between two hypothesis: associate the measurement to a map feature or consider it a new feature, given that the static world assumption no longer holds. However given some environmental structure, the joint persistence estimate can capture the difference, as shown in Figure 5. The only assumption is that at some point at least one feature in a clique of correlated features was independently determined to have low persistence. For example, if a sequence of consecutive and co-linear lane markings are represented by point features and have correlated joint prior distribution over persistence, and those lane markings have shifted slightly the assumption is that at some point the pose uncertainty over the vehicle location is small enough that we are able to determine that at least a single feature in that group no longer exists; then even if the pose uncertainty grows and we are not able to individually determine that every feature in the group no longer exists, the joint persistence model allows us to infer the non-existence of the other lane markings, an example of this is shown in Figure 2.

V. EXPERIMENTS AND RESULTS

Our evaluation of the proposed joint persistence formulation has two principal objectives: establish the *improved persistence estimate by modeling correlation between features* and how can we use the persistence for each feature to perform robust data associations and successfully navigate challenging dynamic environments. To that extent we use both simulated and real datasets with ground truth information to assess persistence and localization error. Due to the lack of a dataset with naturally occurring changes of a semi-static nature, we impose changes on the environment to
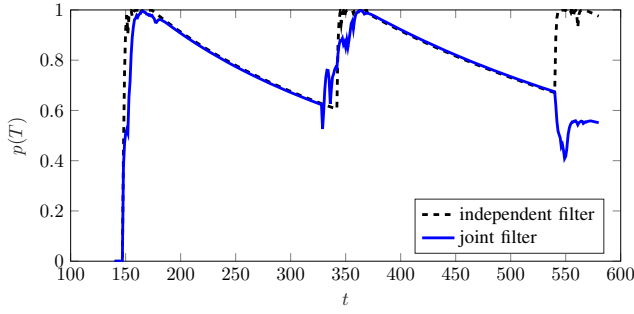
Fig. 5: Comparison of independent filter and joint filter for a single point feature. With no changes to the environment ($t \in [0, 540]$) the filters are equivalent. At time 540 the feature is moved by 1m in the x direction, due to increased pose uncertainty the measurements from the moved feature are associated to the old map point, and the independent filter (dashed line) renews its belief over the feature's existence. The joint filter indicates a drop in belief due to the correlation to neighboring features which have been identified as removed.

varying degrees. In order to assess the usefulness of estimating feature persistence in a factor-graph SLAM setting, we implement a 2D SLAM system in which odometry and range and bearing measurements are added to a keyframe-based fixed-lag smoothing [18] back-end which uses ceres-solver [19] to solve the non-linear least squares problem.

Initially the map is created with known data associations, since landmark initialization is not the focus of this work. Once the map of 2D point features has been created, Mahalanobis-distance data association as described in Section IV is used to chose a pairing between each range and bearing measurement and point feature in the map, using the covariance from the latest state estimate $\hat{\mathcal{X}}$ which can be efficiently recovered from the square-root information form using [20] and the covariance from the estimated map feature. This setup allows for realistic modeling of the feature detector (i.e data associations) in terms of the pose uncertainty, clutter in the environment and the type of change the landmark underwent (removed vs. moved).

A. Simulation

We generate a 2D map with N features drawn uniformly from a predefined set of environments (road, random landmarks, hallway) and simulate a robot traveling along a pre-determined path. Wheel odometry measurements are generated at 50Hz, composed of forward velocity (m/s) and steering angle (rad). The measurements are corrupted by zero-mean Gaussian noise with standard deviation of 0.1m/s and 0.09 rad ($\sim 10°$). Range and bearing measurements are generated for every landmark in the robot's field of view ($90°$, 10m) and corrupted by zero-mean Gaussian noise of 0.05m and 0.08 rad respectively. Visible landmark measurements are removed with probability P_M sampled uniformly $P_M \sim \mathcal{U}([0.01, 0.3])$ and spurious range and bearing measurements are sampled uniformly from the robot's field of view with probability $P_F \sim \mathcal{U}([0.01, 0.15])$. For each τ_i clique of features in the map we sample a survival time $T \sim [0, 1000]$, since we make the assumption that features in the same clique have correlated persistence. If we define a

(a) Removal precision vs. Δ_i
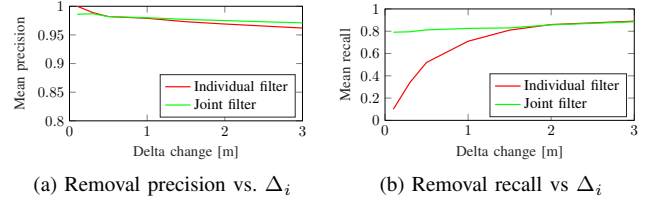

(b) Removal recall vs Δ_i

Fig. 6: Precision and recall for feature removal with removal threshold set to $P_d = 0.5$ over 200 observations of a single feature which was moved by Δ_i as described in Section V-A for a total duration $T = 1000$. The feature is moved by $\Delta_i \in \{0.1, 0.3, 0.5, 1, 1.5, 2, 3\}$ meters. The persistence used for determining if a feature should be removed or not is computed using the marginal persistence $p(T_i > t | J^{1:N})$ where $J^{1:N}$ are all the detections for every feature in the map.

persistence threshold $P_d = 0.5$ in which features are removed from the map the performance of the removal classifier in terms of the amount of change Δ_i each feature undergoes can be assessed in terms of feature removal precision and recall. In a map with 50 features in which a single feature is changed by varying amounts we average 200 observation sequences for that feature. Figure 6 shows the average removal precision and recall for both the individual persistence filter as in [2] and this work. The recall for the individual filter is considerably worse than the joint filter for small changes in the feature, as expected since given the uncertainties in the measurements the data association step will associate the measurement from the moved landmark to the stale map feature, thus renewing its belief and flagging the feature for removal.

B. Real Data

In order to validate the proposed method on real data we use the UTIAS [21] dataset which has ground truth poses sampled at 100Hz with accuracy in the order of $1e^{-3}$m. This dataset is composed of a 2D sparse point-feature based map with ground-truth data associations provided. Since there are no semi-static landmarks in the dataset, we impose change in the landmarks in a similar manner as described in Section V-A. We wish to assess the robustness of the proposed method to various environmental configurations, in terms of the pose RMSE. Figure 7 shows three datasets in which feature cliques were created based on co-linearity of features. The cliques were then moved by increments of 0.01m and a batch estimate was computed for each increment. A feature removal threshold of $P_d = 0.6$ was used for data associations. Figure 7d shows the evolution of the RMSE for each dataset vs. the clique translation. The RMSE increases with very small ($[0, 0.15]$m) translations since all features in the clique are still detected and the incorrect map features are still included in the optimization. The error reducing when at least one feature in the clique is determined to be below the removal threshold P_d. It is interesting to note that the detection, and subsequent update to persistence for each feature is entirely dependent on the data association, which in turn is a function of the structure of the environment (i.e. clutter, number of features) and the pose

(a) Dataset 1 mean trajectory   (b) Dataset 2 mean trajectory

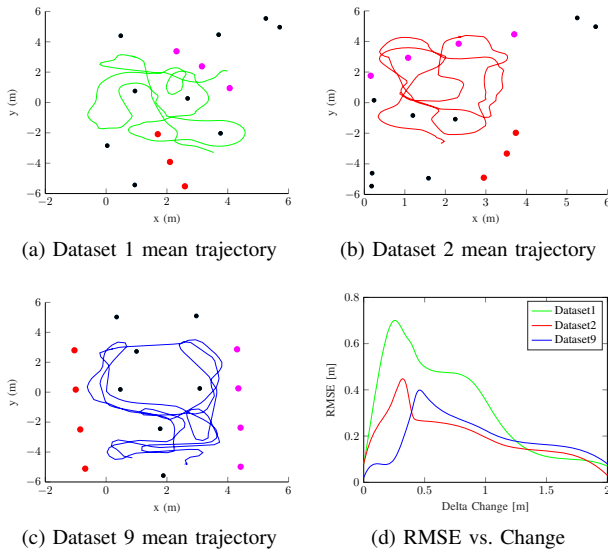(c) Dataset 9 mean trajectory   (d) RMSE vs. Change

Fig. 7: Evaluation of robust localization in the presence of small changes to landmarks in different configurations. Three scenarios from the UTIAS [21] dataset were used to assess the robustness of the proposed method to small changes. Individual features (black) do not have correlated persistence to any other features. Each scenario has two cliques (red and magenta) which have correlated persistence. In each of the three environments the features in a clique were moved in the x direction between $[0, 2]$m in increments of 0.01m. (a) - (c) represent the average estimated trajectory position over all change configurations, (d) shows the RMSE for each scenario vs. feature translation.

and feature uncertainty at the moment of data association, regardless of the data association scheme used (e.g. JCBB, IC). The three datasets in Figure 7 show a drop in RMSE at a similar change point ($\sim 0.4m$) indicating a realistic minimum feature change for any kind of persistence aided localization to be effective.

VI. CONCLUSION

We present an general formulation for feature persistence which makes use of correlation between feature persistence imposed by a joint prior distribution which may be learned or engineered. The key insight of this work is proposing a joint distribution over feature persistence which is computationally tractable in constant time. Our proposed formulation improves upon prior work on modeling individual feature persistence by demonstrating the necessity of a joint model when the environment is subject to change. Our approach allows for the use of the survival time prior distributions discussed in [2] while incorporating information about environmental structure. We show approximated constant-time online inference over feature persistence in a graph-slam environment in both simulated and real scenarios, subject to landmark change. The use of persistence for informed data associations allows navigating through challenging dynamic environments where considering the joint feature persistence is essential.

REFERENCES

[1] J. Neira and J. D. Tardós, "Data association in stochastic mapping using the joint compatibility test," *IEEE Transactions on robotics and automation*, vol. 17, no. 6, pp. 890–897, 2001.

[2] D. M. Rosen, J. Mason, and J. J. Leonard, "Towards lifelong feature-based mapping in semi-static environments," in *Robotics and Automation (ICRA), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1063–1070.

[3] F. Dellaert, "Factor graphs and GTSAM: A hands-on introduction," Georgia Institute of Technology, Tech. Rep., 2012.

[4] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. J. Leonard, and F. Dellaert, "isam2: Incremental smoothing and mapping using the Bayes tree," *The International Journal of Robotics Research*, vol. 31, no. 2, pp. 216–235, 2012.

[5] P. Biber, T. Duckett, *et al.*, "Dynamic maps for long-term operation of mobile service robots," in *Robotics: science and systems*, 2005, pp. 17–24.

[6] P. Biber and T. Duckett, "Experimental analysis of sample-based maps for long-term slam," *The International Journal of Robotics Research*, vol. 28, no. 1, pp. 20–33, 2009.

[7] T. Morris, F. Dayoub, P. Corke, G. Wyeth, and B. Upcroft, "Multiple map hypotheses for planning and navigating in non-stationary environments," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 2765–2770.

[8] W. Churchill and P. Newman, "Practice makes perfect? managing and leveraging visual experiences for lifelong navigation," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. IEEE, 2012, pp. 4525–4532.

[9] A. Walcott-Bryant, M. Kaess, H. Johannsson, and J. J. Leonard, "Dynamic pose graph SLAM: Long-term mapping in low dynamic environments," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*. IEEE, 2012, pp. 1871–1878.

[10] K. Konolige and J. Bowman, "Towards lifelong visual maps," in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*. IEEE, 2009, pp. 1156–1163.

[11] K. Konolige, J. Bowman, J. Chen, P. Mihelich, M. Calonder, V. Lepetit, and P. Fua, "View-based maps," *The International Journal of Robotics Research*, vol. 29, no. 8, pp. 941–957, 2010.

[12] D. Meyer-Delius, M. Beinhofer, and W. Burgard, "Occupancy grid models for robot mapping in changing environments." in *AAAI*, 2012.

[13] H. Moravec and A. Elfes, "High resolution maps from wide angle sonar," in *Robotics and Automation. Proceedings. 1985 IEEE International Conference on*, vol. 2. IEEE, 1985, pp. 116–121.

[14] J. Saarinen, H. Andreasson, and A. J. Lilienthal, "Independent markov chain occupancy grid maps for representation of dynamic environment," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*. IEEE, 2012, pp. 3489–3495.

[15] G. D. Tipaldi, D. Meyer-Delius, and W. Burgard, "Lifelong localization in changing environments," *The International Journal of Robotics Research*, vol. 32, no. 14, pp. 1662–1678, 2013.

[16] T. Krajnik, J. P. Fentanes, G. Cielniak, C. Dondrup, and T. Duckett, "Spectral analysis for long-term robotic mapping," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 3706–3711.

[17] F. Dellaert, "Monte-Carlo EM for data-association and its applications in computer vision," Ph.D. dissertation, Carnegie Mellon University Pittsburgh, PA, 2001.

[18] G. Sibley, L. Matthies, and G. Sukhatme, "Sliding window filter with application to planetary landing," *Journal of Field Robotics*, vol. 27, no. 5, pp. 587–608, 2010.

[19] S. Agarwal, K. Mierle, *et al.*, "Ceres solver," 2012.

[20] M. Kaess and F. Dellaert, "Covariance recovery from a square root information matrix for data association," *Robotics and autonomous systems*, vol. 57, no. 12, pp. 1198–1210, 2009.

[21] K. Y. Leung, Y. Halpern, T. D. Barfoot, and H. H. Liu, "The utias multi-robot cooperative localization and mapping dataset," *The International Journal of Robotics Research*, vol. 30, no. 8, pp. 969–974, 2011.