

# Data driven MCMC for Appearance-based Topological Mapping

Ananth Ranganathan and Frank Dellaert  
College of Computing  
Georgia Institute of Technology

**Abstract**— Probabilistic techniques have become the mainstay of robotic mapping, particularly for generating metric maps. In previous work, we have presented a hitherto nonexistent general purpose probabilistic framework for dealing with topological mapping. This involves the creation of Probabilistic Topological Maps (PTMs), a sample-based representation that approximates the posterior distribution over topologies given available sensor measurements. The PTM is inferred using Markov Chain Monte Carlo (MCMC) that overcomes the combinatorial nature of the problem. In this paper, we address the problem of integrating appearance measurements into the PTM framework. Specifically, we consider appearance measurements in the form of panoramic images obtained from a camera rig mounted on a robot. We also propose improvements to the efficiency of the MCMC algorithm through the use of an intelligent data-driven proposal distribution. We present experiments that illustrate the robustness and wide applicability of our algorithm.

## I. INTRODUCTION

Mapping an unknown and uninstrumented environment is one of the foremost problems in robotics. For this purpose, both metric maps [3] [12] and topological maps [15] [1] have been explored in depth as viable representations of the environment. In both cases, probabilistic approaches have had great success in dealing with the inherent uncertainties associated with robot sensor-motor control and perception, that would otherwise make map-building a very brittle process.

This work deals with the problem of topological mapping. Topological maps attempt to capture the spatial connectivity of the environment by representing it as a graph with arcs connecting the nodes that designate significant places in the environment, such as corridor junctions and room entrances [10]. Arguably the hardest problem in topological mapping is the perceptual aliasing problem, which is an instance of the data association problem, also variously known as “closing the loop” [5] or “the revisiting problem” [18]. It is the problem of determining whether sensor measurements taken at different points in time correspond to the same physical location. When a robot receives a new measurement, it has to decide whether to assign this measurement to one of the locations it has visited previously, or to a completely new location. The aliasing problem is hard because the number of possible choices grows combinatorially with the number of measurements.

In previous work [14], we presented the concept of Probabilistic Topological Maps (PTMs) that deal with the perceptual aliasing problem in a systematic probabilistic manner. A PTM is a probability distribution over the discrete space of all possible topologies, and is obtained by computing the

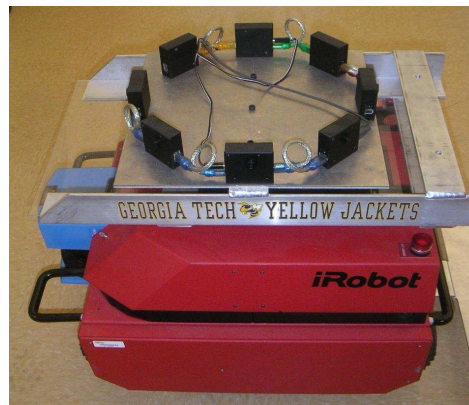


Fig. 1. Camera rig mounted on the robot to obtain panoramic images

posterior distribution over this space given the measurements. However, due to the combinatorial size of the state space, it is not possible to compute the PTM analytically, and hence, a sample-based approximation to the posterior distribution is used for this purpose. The sample-based approximation, in turn, is computed using the Markov Chain Monte Carlo (MCMC) sampling algorithm [4] that overcomes the combinatorial nature of the state space.

The intuitive reason for computing the posterior is to solve the aliasing problem for topologies in a systematic manner. The set of all possible correspondences between measurements and the physical locations from which the measurements are taken is exactly the set of all possible topologies. By inferring the posterior on this set, whereby each topology is assigned a probability, it is possible to locate the more probable topologies without committing to a specific correspondence at each step, as most current algorithms do. Thus, a general solution to the perceptual aliasing problem is obtained. Even in pathological environments, where almost all current algorithms fail, our technique provides a quantification of uncertainty by pegging a probability of correctness to each topology.

In [14], we considered the case where the measurements consist of odometry measurements alone. We demonstrated that even in this case, with infinite perceptual aliasing, PTMs perform well. However, the use of appearance information, if available, clearly provides an advantage which was not utilized in that work. Further, the proposal distribution used to mix the Markov chain in state space was constructed using a simple split-merge algorithm that does not take into account



Fig. 2. A panoramic image obtained from the robot camera rig

any domain knowledge. This leads to slow mixing and results in inefficiency in some cases.

In this paper, we address both the above shortcomings in our previous work. Hence, our contribution is two-fold. First, we present a general model for incorporating appearance measurements in the construction of PTMs. As a specific instance of this model, we propose the use of panoramic images obtained from a camera rig mounted on a robot (shown in Figure 1), to obtain appearance measurements. An example of such a panoramic image is given in Figure 2.

As a second major contribution, we provide a new data-driven proposal distribution for use in the MCMC sampler. The proposal uses domain knowledge in the form of expected landmark locations, and leads to faster mixing of the Markov chain, thus making the PTM algorithm more efficient. A more general aspect of this work is that it demonstrates a means to include pose information into any MCMC proposal that deals with the space of all possible clusterings. This is true since the space of topologies is exactly the same as that of all possible clusterings of available measurements.

We use the Fourier signature [6] of a panoramic image as the appearance measurements in our appearance model. Fourier signatures, which have previously been used in the context of localization using omni-directional vision [11], are a low-dimensional representation of images using Fourier coefficients. They allow inexpensive matching of images to determine correspondence to physical locations. Further, due to the periodicity of panoramic images, Fourier signatures are rotation-invariant. This property is of prime importance when determining correspondence since the robot may be moving in different directions when the images are obtained. We present a generative model for the appearance measurements that enables their use in the PTM algorithm. The advantage of using appearance measurements in addition to odometry is illustrated through experiments.

In subsequent sections, after providing related work, we give a brief overview of PTMs and the means for estimating the posterior over the space of topologies through MCMC sampling. Subsequently, we describe our appearance model for use in the PTM algorithm, and this is followed by an explanation of the data-driven proposal distribution. In Section VI, we give details of the experiments we conducted and the results obtained, following which we conclude.

## II. RELATED WORK

Maintaining the posterior distribution over the space of topologies results in a systematic and robust solution to the aliasing problem that plagues topological mapping. Though

probabilistic methods have been used in conjunction with topological maps before, none exist that are capable of dealing with the inference of the posterior distribution over the space of topologies. A recent approach by Remolina and Kuipers [15], improved upon by Savelli and Kuipers [16], gives an algorithm to build a tree of all possible topological maps that conform to the measurements, but in a non-probabilistic manner. Most instances of previous work extant in the literature that incorporate uncertainty in topological map representations do not deal with general topological maps, but with the use of Markov decision processes to learn a policy that the robot follows to navigate the environment.

Shatkay and Kaelbling [17] use the Baum-Welch algorithm, a variant of the EM algorithm used in the context of HMMs, to solve the aliasing problem for topological mapping. However, this approach is well-known to be prone to local minima in the solution space. The use of a limited, multiple-hypothesis space over correspondences through the use of POMDPs is prevalent also in the literature [20].

Others use a non-probabilistic approach to the perceptual aliasing problem by applying a clustering algorithm to the measurements to identify distinctive places, an instance being [9]. Numerous approaches also exist for the use of local appearance in place recognition, for example [19] [2]. However, all these methods are inherently brittle in the sense that they are prone to failing silently in environments with severe perceptual aliasing.

Data-driven proposals have previously been used various fields - for example in Computer Vision for image segmentation [21], and in Statistics to analyze mixture models [7]. In general, data-driven proposals cause a significant speed-up in the sampling algorithm in cases where the state space being considered is enormous. In such cases, a normal proposal would provide a number of samples that are from regions of low probability and hence get rejected, wasting the computation involved in their generation. A proposal that utilizes the data, on the other hand, directs the proposed samples towards regions of higher probability, thus increasing the MCMC acceptance ratio and reducing the number of cases where the proposed sample is rejected.

## III. PROBABILISTIC TOPOLOGICAL MAPS

We begin by giving a brief overview of Probabilistic Topological Maps (PTMs). A PTM is a sample-based representation that approximates the posterior distribution  $P(T|Z)$  over topologies  $T$  given measurements  $Z$ . While the space of possible maps is combinatorial, a probability density over this space can be approximated by drawing a sample of possible

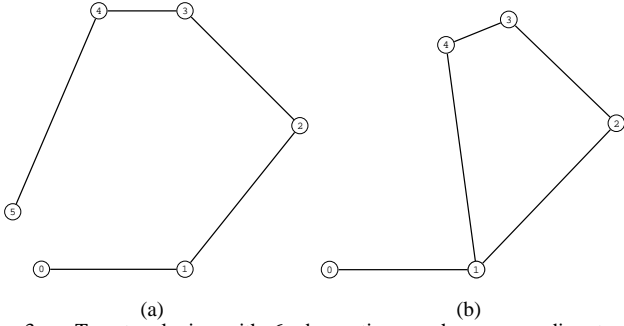


Fig. 3. Two topologies with 6 observations each corresponding to set partitions (a) with six landmarks  $\{\{0\}, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}\}$  and (b) with five landmarks  $\{\{0\}, \{1, 5\}, \{2\}, \{3\}, \{4\}\}$  where the second and sixth measurement are from the same landmark.

maps from the distribution. Using the samples, it is possible to construct a histogram on the support of this sample set.

For the purpose of this work, we assume the availability of a “landmark detector” that detects a landmark when it is near. While the problem of landmark detection is an important one in itself, we do not consider it in this paper. No knowledge of the correspondence between landmark measurements and the actual landmarks is given to the robot - indeed, that is exactly the topology that we seek. The problem then is to compute the discrete posterior probability distribution  $P(T|Z)$  over the space of topologies.

#### A. Topologies as set partitions

To infer the PTM from the measurements, we exploit the equivalence between topologies of an environment and set partitions of landmark measurements, which group the measurements into a set of equivalence classes. When all the measurements of the same landmark are clustered together, this naturally defines a partition on the set of measurements. Let the set of measurements be denoted as  $Z = \{z_i | i \in [1, N]\}$ , where  $N$  is the number of measurements (the number of landmarks seen by the robot). If the number of distinct landmarks in the environment is  $M$  ( $M \leq N$ ), then a topology  $T$  can be represented as the set partition of the set  $Z$ ,  $T = \{S_j | j \in [1, M]\}$ , where each  $S_j$  is a set of measurements such that  $S_{j1} \cap S_{j2} = \phi \forall j1, j2 \in [1, M], j1 \neq j2$  and  $\bigcup_{j=1}^M S_j = Z$ . The set  $S_j$  contains the measurements corresponding to the  $j$ th distinct landmark in the environment. As an aside, we note that we only deal with planar topologies in this work.

It can be seen that a topology is nothing but an assignment of measurements to sets in the partition. This results in the above mentioned isomorphism between topologies and set partitions. An example of the encoding of topologies as set partitions is shown in Figure 3. The number of possible topologies is thus equal to the number of set partitions of the set of measurements (and hence, also to the set of all possible clusterings of the set of measurements). This number is called the *Bell number* [13], and grows hyper-exponentially with the number of measurements.

---

#### Algorithm 1 The Metropolis-Hastings algorithm

---

- 1) Start with a valid initial topology  $T_t$ , then iterate once for each desired sample
- 2) Propose a new topology  $T'_t$  using the *proposal distribution*  $Q(T_t \rightarrow T'_t)$
- 3) Calculate the *acceptance ratio*

$$a = \frac{P(T'_t|Z^t) Q(T'_t \rightarrow T_t)}{P(T_t|Z^t) Q(T_t \rightarrow T'_t)} \quad (1)$$

where  $Z^t$  is the set of measurements observed up to and including time  $t$ .

- 4) With probability  $p = \min(1, a)$ , accept  $T'_t$  and set  $T_t \leftarrow T'_t$ . If rejected we keep the state unchanged (i.e. return  $T_t$  as a sample).
- 

#### B. Inferring PTMs using Markov chain Monte Carlo

The aim of inference in the space of topologies is to obtain the posterior probability distribution on topologies  $P(T|Z)$ . We use Markov chain Monte Carlo (MCMC) sampling to perform inference in the combinatorial state space of topologies,

All MCMC methods work by running a Markov chain over the state space with the property that the chain ultimately converges to the target distribution of interest, in this case the posterior over topologies. Once the chain has converged, subsequent states visited by the chain are considered to be samples from the target distribution. The Markov chain itself is generated using a proposal distribution that is used to propose the next state in the chain, a move in state space, possibly by conditioning on the current state. The Metropolis-Hastings algorithm, a general MCMC method, provides a technique whereby the Markov chain can converge to the target distribution using any arbitrary proposal distribution, the only important restriction being that the chain be capable of reaching all the states in the state space.

The pseudo-code to generate a sequence of samples from the posterior distribution  $P(T|Z)$  over topologies  $T$  using the Metropolis-Hastings algorithm is shown in Algorithm 1 (adapted from [4]). Intuitively, the algorithm samples from the desired probability distribution  $P(T|Z)$  by rejecting a fraction of the moves generated by a proposal distribution  $Q(T'_t; T_t)$ , where  $T_t$  is the current state and  $T'_t$  is the proposed state. The fraction of moves rejected is governed by the acceptance ratio  $a$  given by (1), the computation of which requires the design of a proposal density and evaluation of the target density.

The target distribution  $P(T|Z)$  is computed through the use of Bayes Law

$$P(T|Z) \propto P(Z|T)P(T) \quad (2)$$

where  $P(T)$  is a prior on topologies and  $P(Z|T)$  is the observation likelihood. In [14], we considered the case where  $Z$  is just the set of odometry measurements. The odometry likelihood was computed by Rao-Blackwellization of landmark locations using a prior distribution on the landmark locations. We used a simple split-merge proposal distribution to move the Markov chain through the state space. For more details, see [14].

#### IV. INCORPORATING APPEARANCE MODELS IN PTMS

If, in addition to odometry, appearance measurements are also taken into consideration, the set of measurements  $Z$  consists of odometry measurements  $O$  and appearance measurements  $A$ , so that  $Z = \{A, O\}$ . Also, note that the odometry and appearance measurements are conditionally independent given the topology  $T$ , since the topology determines the correspondence between measurements and physical landmarks. Using this independence in (2), we get

$$\begin{aligned} P(T | O, A) &= kP(O, A|T)P(T) \\ &= kP(O|T)P(A|T)P(T) \end{aligned} \quad (3)$$

where  $k$  is the normalization constant. The evaluation of the odometry likelihood  $P(O|T)$  is discussed in [14] and is not considered here. We deal with modeling appearance to evaluate the appearance likelihood  $P(A|T)$  in this section.

Fourier signatures, which we use as appearance measurements, are computed by calculating the 1-D Fourier transform of each row of the panoramic image and storing only the first few coefficients corresponding to the lower spatial frequencies [11]. While more popular dimensionality reduction techniques such as PCA [8] exist, the drawback of such systems is the need to further preprocess the measurement images in order to obtain rotational invariance. In contrast, the magnitudes of Fourier coefficients in a Fourier signature are rotation-invariant since panoramic images are periodic. Hence, a Fourier signature yields a low-dimensional, rotation-invariant representation of the image. We use images obtained from an eight-camera rig mounted on a robot to produce panoramic images as shown in Figures 1 and 2. The eight images thus obtained are mosaicked automatically to form a  $360^\circ$  view of the environment.

In our case, Fourier signatures are calculated using a modification of the procedure given in [11]. First, a single row image obtained by averaging the rows of the input image is calculated and subsequently, the one-dimensional Fourier transform of this image is performed. This gives us the Fourier signature of the image. It is to be noted that Fourier signatures do not comprise an error-free source of measurements. If that were the case, then the need for a probabilistic treatment would not arise. Most of the errors in the measurements take the form of false positives, in the sense that images from distinct physical locations often yield similar Fourier signatures. This is due to perceptual aliasing and the extreme compression of the image data into a Fourier signature. However, when used in the PTM algorithm in conjunction with odometry, they still produce good results as we demonstrate in Section VI.

We begin by denoting the set of appearance measurements as  $A = \{a_i | 1 \leq i \leq N\}$ , where  $N$  is the number of measurements (the number of landmarks observed by the robot during its run). Let the number of distinct landmarks in the environment be  $M$  ( $M \leq N$ ). Evaluation of the appearance likelihood is performed by introducing the hidden parameter  $Y = \{y_j | 1 \leq j \leq M\}$ , where each  $y_j$  corresponds to a distinct landmark. The hidden parameter denotes the “true appearance” of each landmark in the topology. As we do not

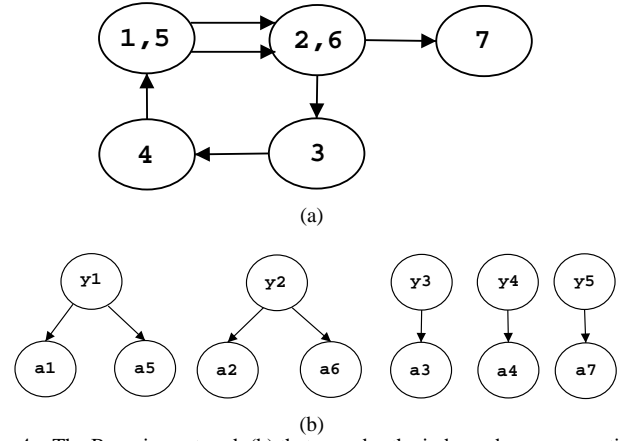


Fig. 4. The Bayesian network (b) that encodes the independence assumptions for the appearance measurements in the topology (a) given the true appearance  $Y = \{y_1, \dots, y_5\}$  at all the landmark locations.

need to compute  $Y$  when inferring topologies, we marginalize over it so that

$$P(A | T) = \int_Y P(A | Y, T)P(Y | T) \quad (4)$$

where  $P(A|Y, T)$  is the measurement model and  $P(Y | T)$  is the prior on the appearance. We assume that the appearance of a landmark is independent of all other landmarks, so that each  $y_j$  is independent of all other  $y_{j'}$ . The prior  $P(Y | T)$  can thus be factored into a product of priors on the individual  $y_j$ .

$$P(Y | T) = \prod_{j=1}^M P(y_j) \quad (5)$$

As seen in Section III-A, the topology  $T$  introduces a partition on the set of appearance measurements by determining which “true appearance”  $y_j$  each measurement  $a_i$  actually measures, i.e the partition encodes the correspondence between the set  $A$  and the set  $Y$ . Also, given  $Y$ , the likelihood of the appearance can be factored into a product of likelihoods of the individual appearance instances. This is illustrated using an example topology in Figure 4, where the Bayesian network encodes the independence assumptions in the appearance measurements. Hence, denoting the  $j$ th set in the partition as  $S_j$ , we rewrite  $P(A | Y, T)$  as -

$$P(A | Y, T) = \prod_{j=1}^M \prod_{a_i \in S_j} P(a_i | y_j) \quad (6)$$

where the dependence on  $T$  is subsumed in the partition. Combining Equations (4), (5) and (6), we get the expression for the appearance likelihood as

$$P(A | T) = \prod_{j=1}^M \int_{y_j} P(y_j) \prod_{a_i \in S_j} P(a_i | y_j) \quad (7)$$

In our case, each appearance measurement  $a_i$  is a Fourier signature vector given as  $a_i = \{a_{i1}, a_{i2}, \dots, a_{iK}\}$ , where  $a_{ik}$  is the  $k$ th Fourier component in the Fourier signature. Also,

we assume a similar vector form for the hidden appearance variables  $y_i$ , so that  $y_i = \{y_{i1}, y_{i2}, \dots, y_{iK}\}$ . Moreover, we assume that the frequency components of the Fourier signature given the corresponding appearance variable are independent, and hence, can be factored, as can be the prior over the hidden appearance variables. Consequently, we can rewrite (7) to get the expression for the appearance likelihood as

$$P(A | T) = \prod_{j=1}^M \prod_{k=1}^K \int_{y_{jk}} P(y_{jk}) \prod_{a_i \in S_j} P(a_{ik} | y_{jk}) \quad (8)$$

In the above equation,  $P(y_{jk})$  is a prior on appearance in the environment, and  $P(a_{ik} | y_{jk})$  is the appearance measurement model. Evaluation of the appearance likelihood requires the specification of these two quantities.

We assume the measurement noise in the Fourier signatures to be Gaussian distributed so that the model for appearance instance  $a_{ik}$ , belonging to the  $j$ th set  $S_j$ , is also a Gaussian centered around the ‘‘true appearance’’  $y_{jk}$  with variance  $\sigma_{jk}^2$ . Since we do not know either of these parameters, we further model them hierarchically in a proper Bayesian manner. Hierarchical conjugate priors are placed on  $\sigma_{jk}^2$  and  $y_{jk}$ : the prior on  $\sigma_{jk}^2$  being an inverse gamma distribution while the prior on  $y_{jk}$  is taken to be a Gaussian distribution with mean  $\mu$  and variance  $\frac{\sigma_{jk}^2}{\kappa}$ . This particular choice of conjugate priors allows the integration in (8) to be performed analytically. The appearance model can then be summarized as

$$\begin{aligned} a_{ik} &\sim \mathcal{N}(a_{ik}; y_{jk}, \sigma_{jk}^2) \quad \text{where } a_i \in S_j \\ y_{jk} &\sim \mathcal{N}(y_{jk}; \mu, \frac{\sigma_{jk}^2}{\kappa}) \\ \sigma_{jk}^2 &\sim IG(\sigma_{jk}^{-2}; \alpha_k, \beta_k) \end{aligned} \quad (9)$$

where  $IG$  denotes the inverse gamma distribution. Note that while the value of  $\kappa$  is generally chosen so that the prior on  $y_{jk}$  is vague, we usually have some extra ‘‘world knowledge’’ that can be used to set the values of the hyper-parameters  $\alpha_k$  and  $\beta_k$ . For example, if we expect the value of the Fourier signature to vary by only a small amount in the neighborhood of a given location, the prior on  $\sigma_{jk}^2$  should reflect this knowledge by being peaked about a specific value.

The generative model for Fourier signature measurements specified by (9) is now used to compute the appearance likelihood given by (8). In addition to integrating over  $y_{jk}$ , we also integrate over the variance  $\sigma_{jk}^2$  as we are not interested in its value. It follows that

$$P(A | T) = \prod_{j=1}^M \prod_{k=1}^K \int_{\sigma_{jk}^2} IG(\alpha_k, \beta_k) \times \int_{y_{jk}} \mathcal{N}(\mu, \frac{\sigma_{jk}^2}{\kappa}) (\mathcal{N}(y_{jk}, \sigma_{jk}^2))^{|S_j|} \quad (10)$$

Due to the use of conjugate priors, this computation can be performed analytically.

Now that the appearance likelihood can be evaluated from (10), this can be used in (3) to compute the target distribution.

---

## Algorithm 2 The Proposal Distribution

---

- 1) Select a merge or a split with probability  $\left\{ \frac{N_M}{N_M+N_S}, \frac{N_S}{N_M+N_S} \right\}$
  - 2) **Merge move:**
    - if  $T$  contains only one set, re-propose  $T' = T$ , hence  $r = 1$
    - otherwise select two sets at random, say  $R$  and  $S$ 
      - a) Let  $D$  be the distance between the locations corresponding to  $R$  and  $S$  obtained by optimizing the odometry wrt  $T$
      - b)  $T' = (T - \{R\} - \{S\}) \cup \{R \cup S\}$  and  $Q(T \rightarrow T') = \frac{\exp\left(-\frac{D^2}{\sigma^2}\right)}{\frac{N_M+N_S}{N_M+N_S}}$
      - c)  $Q(T' \rightarrow T)$  is obtained from the reverse case 3(c), hence  $r = \frac{N_M+N_S}{N'_M+N'_S} \exp\left(\frac{D^2}{\sigma^2}\right)$ , where  $N'_M$  and  $N'_S$  are the number of merge and split moves possible from  $T'$
  - 3) **Split move:**
    - if  $T$  contains only singleton sets, re-propose  $T' = T$ , hence  $r = 1$
    - otherwise select a non-singleton set  $U$  at random from  $T$  and split it into two sets  $R$  and  $S$ .
      - a)  $T' = (T - \{U\}) \cup \{R, S\}$
      - b) Let  $D$  be the distance between the locations corresponding to  $R$  and  $S$  obtained by optimizing the odometry wrt  $T'$
      - c)  $Q(T \rightarrow T') = \frac{1}{N_M+N_S}$
      - d)  $Q(T' \rightarrow T)$  is obtained from the reverse case 2(b), hence  $r = \frac{N_M+N_S}{N'_M+N'_S} \exp\left(-\frac{D^2}{\sigma^2}\right)$ , where  $N'_M$  and  $N'_S$  are as defined in 2(c)
- 

The target distribution, in turn, is needed to compute the acceptance ratio in the Metropolis-Hastings algorithm, as has been mentioned before.

The appearance model presented above is not specific to Fourier signatures. Indeed, it is a general purpose clustering model that assumes that the data are distributed as a mixture of Gaussians. A topology labels each data instance as arising out of one of the mixture components, where the number of mixture components is equal to the number of sets in the set partition corresponding to the topology.

## V. DATA-DRIVEN PROPOSAL DISTRIBUTION

Consider a topology  $T = \{S_j | j \in [1, M]\}$ , where the  $S_j$  are sets in a set partition of the measurements as before. If the Markov chain is currently in the state  $T$ , the task of the proposal distribution is to propose a new topology  $T'$  from  $T$ . With reference to the calculation of the Metropolis-Hastings acceptance ratio in (1), the probability of the step from  $T$  to  $T'$  as well as the reverse step has to be computed. We now present a data-driven proposal distribution that accomplishes this task in an efficient manner so that the Markov chain mixes rapidly across the state space.

The basic steps of the proposal consist of the merge and the split moves as in [14]. A *merge* move occurs when two of the sets in the set partition corresponding to  $T$  are merged to yield  $T'$ . Analogously, a *split* move occurs when a set in  $T$  is split into two. The number of ways in which a merge move



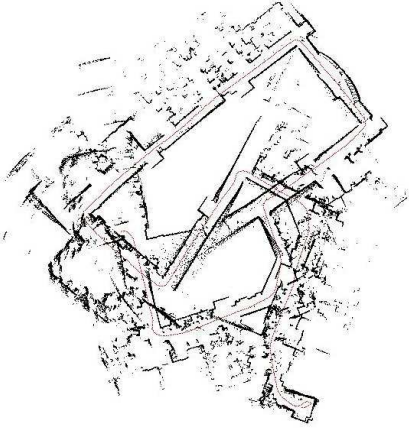


Fig. 5. Odometry of the robot plotted with the laser measurements for the first experiment.

can occur is given as  $N_M = \binom{M}{2}$ ,  $M > 1$ . To calculate the number of possible ways a split move can occur, let  $N_{St}$  be the number of non-singleton sets in the partition. Clearly,  $N_{St}$  is the number of sets in the partition that can be split. Out of these  $N_{St}$  sets, we pick a random set  $R$  to split. The number of possible ways to split  $R$  into two subsets is given by  $\left\{ \begin{smallmatrix} |R| \\ 2 \end{smallmatrix} \right\}$ . Here  $\left\{ \begin{smallmatrix} n \\ m \end{smallmatrix} \right\}$  denotes the Stirling number of the second kind that gives the number of possible ways to split a set of size  $n$  into  $m$  subsets [13]. Hence, the total number of ways a split move can occur on  $T$  is given by  $N_S = N_{St} \left\{ \begin{smallmatrix} |R| \\ 2 \end{smallmatrix} \right\}$ .

The data-driven proposal distribution, which computes the proposal ratio  $r$  used in the calculation of the acceptance ratio in (1), is given in Algorithm 2. The proposal begins by picking a split or merge move according to the proportion of the number of ways these moves are possible.

Subsequently, if a merge move is chosen, two random sets from the partition are chosen to be merged. Now, however, the knowledge of the odometry measurements is used to calculate the probability of the proposal. Intuitively, measurements that are taken when the robot pose is almost the same have a higher probability of being from the same landmark, and should have a higher probability of being merged.

The landmark locations corresponding to the sets to be merged are obtained from the optimal robot trajectory, which in turn is obtained by optimizing the odometry under the constraints required by topology  $T$  [14]. The topology  $T$  requires certain landmark measurements to correspond to the same physical landmark, i.e. to occur at the same physical location. However, enforcing this constraint causes the trajectory of the robot to diverge from the odometry measurements. The optimal trajectory minimizes the total error due to divergence from the odometry measurements and not enforcing the constraints dictated by the topology  $T$ . The probability of the merge step is then obtained using the distance  $D$  between the landmarks that we are proposing to merge as  $\exp\left(-\frac{D^2}{\sigma^2}\right)$ , where  $\sigma^2$  is a variance that encodes our belief in the distance between landmarks, or equivalently, the scale of the environment.

The analogous calculation for the split step is now simple

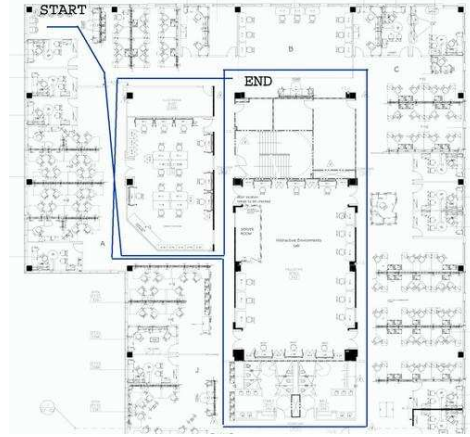


Fig. 6. Schematic of robot path overlaid on a floorplan of the environment for the first experiment.

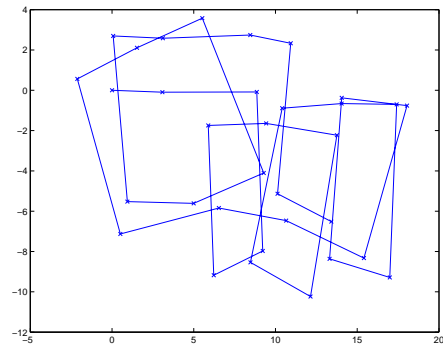


Fig. 7. Landmark locations obtained from simulated odometry for the second experiment.

	<i>Data-driven proposal</i>	<i>General proposal</i>
1 <sup>st</sup> experiment	9 minutes	46 minutes
2 <sup>nd</sup> experiment	51 minutes	> 6 hours

Fig. 8. Running times for computing the PTM using the two proposals in both the experiments. The data-driven proposal speeds up the algorithm by at least a factor of five.

since the probability of the split itself is just the inverse of total possible moves from  $T$ . Note that the merge moves proposed in the above scheme will often have a low probability since the majority of landmarks chosen for merging will not be close together. To prevent this, a gating scheme is used that selects the landmarks to be merged preferentially based on their being closer together than a pre-defined threshold.

## VI. EXPERIMENTS

We now present experiments to validate our algorithm. The experiments were performed on an ATRV-Mini robot mounted with an eight camera rig and the landmarks in the runs were selected manually.

The first experiment was conducted over an entire floor of a building and consisted of a robot run containing two loops, a bigger loop enclosing a smaller loop. Twelve landmarks were observed by the robot during the run, shown overlaid

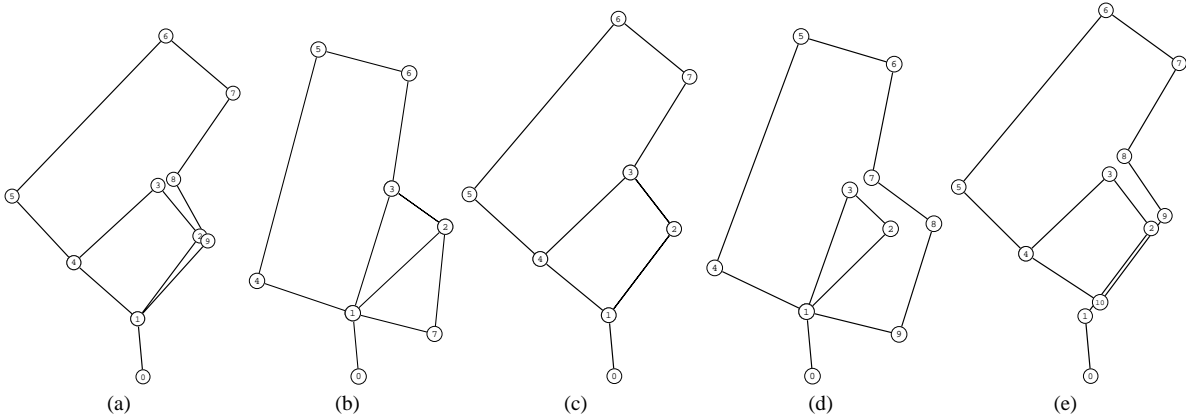


Fig. 9. Topologies with highest posterior probability mass for the first experiment using only odometry. (a) receives 43% of the probability mass while (b), (c), (d) and (e) receive 14%, **7.3%**, 3.9% and 2.8% of the probability mass respectively. The ground truth topology is (c).

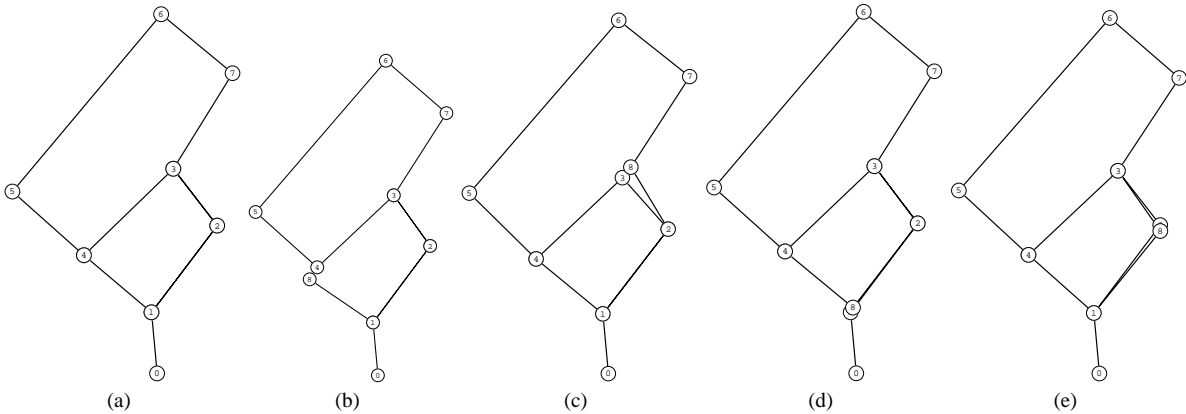


Fig. 10. Topologies constituting the PTM for the first experiment using both odometry and appearance. (a) receives **99.5%** of the probability mass while (b), (c), (d) and (e) receive 0.25%, 0.14%, 0.12% and 0.01% of the probability mass respectively. The ground truth topology is (a).

on a floorplan of the experimental area in Figure 6. The odometry of the robot with the laser plotted on top is shown in Figure 5. Using only the odometry measurements, the ground truth topology received a low probability mass due to noisy odometry. The five most probable topologies in the PTM obtained are given in Figure 9.

We now repeat the experiment, but this time also using the appearance measurements in addition to the odometry. The first five frequencies of the Fourier signatures were used for this purpose. The values of the variance hyper-parameters in the appearance model were set so that the prior over the variance is centered at 500 with a variance of 50. When appearance is also included, the results (shown in Figure 10) shift dramatically since there is little perceptual aliasing in this environment. Although five topologies appear in the PTM, the ground truth topology receives almost all of the probability mass. This experiment illustrates the fact that when reliable measurements are available, our approach produces a PTM that is sharply peaked in the space of topological maps. More specifically, the appearance measurements help disambiguate noisy odometry data in this case. The improvement in running time using the data-driven proposal is given in Table 8. To demonstrate the scalability of our algorithm, we conducted a second experiment in simulation in an environment where

the robot was made to traverse a number of loops. A total of 33 landmarks were observed by the robot in the run. The landmark locations obtained from odometry generated during the simulated run are shown in Figure 7. Using the data-driven proposal speeds up the algorithm by a factor of six (Table 8) as compared to the general split-merge proposal [14]. The four most likely topologies in the PTM are shown in Figure 11.

In the above experiments, we test the convergence of the MCMC procedure by initializing the chain randomly and observing if the chains converge to the same distribution. While this provides a rough estimate of convergence, it is not a robust theoretical measure. Implementing a theoretically-grounded convergence criterion is future work.

## VII. CONCLUSION

We presented a generative model for appearance and used it in addition to odometry to generate Probabilistic Topological Maps. It is seen from the experiments that addition of appearance information improves the results significantly by disambiguating noisy odometry. This is in spite of the fact that the appearance measurements are themselves noisy. In addition, we presented a data-driven proposal that significantly speeds up the algorithm by enabling rapid mixing of the Markov chain. This speed-up is at least by a factor of five

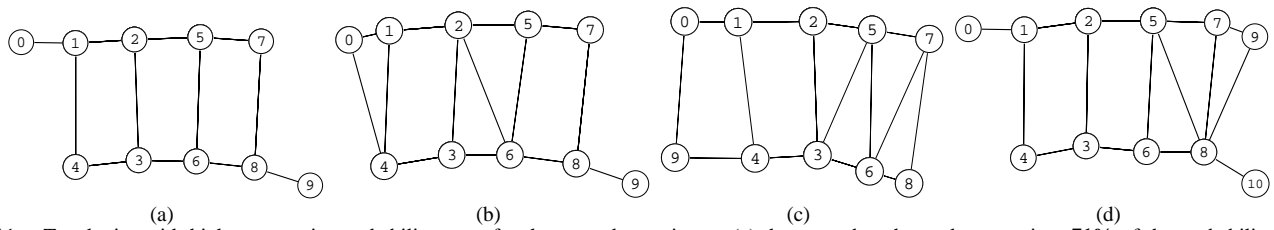


Fig. 11. Topologies with highest posterior probability mass for the second experiment. (a) the ground truth topology receives 71% of the probability mass while (b), (c), and (d) receive 9.1%, 8.2%, and 6% of the probability mass respectively. The ground truth topology is (a).

as seen from the results we have presented.

It is to be noted that the improvement in the results upon addition of appearance information is solely due to the fact that we have a more confident inference of the PTM. The PTM itself is simply the posterior over the space of topologies, and hence, the PTM obtained using only odometry is not incorrect in any sense. Addition of more information simply makes the posterior more peaked in the space of topologies. This improvement in the posterior upon use of more information is a feature of Bayesian inference in general.

Currently, the appearance model requires three parameters to be chosen by the user. These are the  $\alpha$  and  $\beta$  variance hyper-parameters and the number of frequency components to be considered in the Fourier signature. The variance hyper-priors encode the variation in appearance values from the same location in the environment. Changes in lighting, camera distortion and other measurement noise may make this variation large. Hence, the values of the hyper-parameters need to be empirically determined for each environment. It is our experience that there is rarely need to use more than the first five frequency components in the appearance model. This is because the higher frequency components mainly contain noise, which we do not seek to model. It is also to be noted that while we use Fourier signatures in this work, any other rotation-invariant dimensionality reduction technique can be used instead.

While we only use models for odometry and appearance, a simple extension to laser data is also possible.  $360^0$  laser scans at the landmark locations can be used to compute the likelihood between scans after an optimal alignment. This likelihood, extended to multiple scan comparison, can be used to sample over partitions. Similarly, we have only considered the use of odometry measurements in the data-driven proposal. Clearly, appearance information can also be used in a similar manner in conjunction to odometry. However, it is our experience that the use of odometry alone is sufficient to provide good proposals and hence, in the interests of space, we have not provided details of the analogous use of appearance measurements here.

#### ACKNOWLEDGEMENTS

The work on Fourier signatures was performed with Emanuele Menegatti and the code for computing them was also provided by him. We are grateful to Michael Kaess for gathering the robot datasets used in our experiments.

#### REFERENCES

- [1] H. Choset and K. Nagatani. Topological simultaneous localization and mapping (SLAM): toward exact localization without explicit localization. *IEEE Trans. Robot. Automat.*, 17(2):125 – 137, April 2001.
- [2] G. Dudek and D. Jugessur. Robust place recognition using local appearance based methods. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 1030–1035, 2000.
- [3] A. Elfes. Occupancy grids: A probabilistic framework for robot perception and navigation. *Journal of Robotics and Automation*, RA-3(3):249–265, June 1987.
- [4] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter, editors. *Markov chain Monte Carlo in practice*. Chapman and Hall, 1996.
- [5] D. Haehnel, W. Burgard, D. Fox, and S. Thrun. A highly efficient FastSLAM algorithm for generating cyclic maps of large-scale environments from raw laser range measurements. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2003.
- [6] H. Ishiguro, K. C. Ng, R. Capella, and M. M. Trivedi. Omnidirectional image-based modeling: three approaches to approximated plenoptic representations. *Machine Vision and Applications*, 14(2):94–102, 2003.
- [7] S. Jain and R. Neal. A split-merge Markov chain Monte Carlo procedure for the dirichlet process mixture model. *Journal of Computational and Graphical Statistics*, 13(1):158–182, March 2004.
- [8] I. T. Jolliffe. *Principal Component Analysis*. Springer, 1986.
- [9] Benjamin Kuipers and Patrick Beeson. Bootstrap learning for place recognition. In *AAAI Nat. Conf. on Artificial Intelligence*, pages 174–180, 2002.
- [10] B.J. Kuipers and Y.-T. Byun. A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations. *Journal of Robotics and Autonomous Systems*, 8:47–63, 1991.
- [11] E. Menegatti, M. Zoccarato, E. Pagello, and H. Ishiguro. Image-based Monte Carlo localisation with omnidirectional images. *Robotics and Autonomous Systems*, 48(1):17–30, 2004.
- [12] M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit. FastSLAM: A factored solution to the simultaneous localization and mapping problem. In *AAAI Nat. Conf. on Artificial Intelligence*, 2002.
- [13] A. Nijenhuis and H. Wilf. *Combinatorial Algorithms*. Academic Press, 2 edition, 1978.
- [14] A. Ranganathan and F. Dellaert. Inference in the space of topological maps: An MCMC-based approach. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2004.
- [15] E. Remolina and B. Kuipers. Towards a general theory of topological maps. *Artificial Intelligence*, 152(1):47–104, 2004.
- [16] F. Savelli and B. Kuipers. Loop-closing and planarity in topological map-building. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2004.
- [17] H. Shatkay and L. Kaelbling. Learning topological maps with weak local odometric information. In *Proceedings of IJCAI-97*, pages 920–929, 1997.
- [18] B. Stewart, J. Ko, D. Fox, and K. Konolige. The revisiting problem in mobile robot map building: A hierarchical Bayesian approach. In *Conf. on Uncertainty in Artificial Intelligence*, pages 551–558, 2003.
- [19] K. Sugihara. Some location problems for robot navigation using a single camera. *CVGIP:Image Understanding*, 42(1):112–129, April 1988.
- [20] N. Tomatis, I. Nourbakhsh, and R. Siegwart. Hybrid simultaneous localization and map building: Closing the loop with multi-hypotheses tracking. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 2749–2754, 2002.
- [21] Z.W. Tu and S.C. Zhu. Image segmentation by data-driven Markov chain Monte Carlo. *IEEE Trans. Pattern Anal. Machine Intell.*, 24(5):657–673, 2002.