Probabilistic Topological Mapping for Mobile Robots using Urn Models

Ananth Ranganathan **Frank Dellaert**
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332

Abstract

We present an application of Bayesian modeling and inference to topological mapping in robotics. This is a potentially difficult problem due to (a) the combinatorial nature of the state space, and (b) perceptual aliasing by which two different landmarks in the environment can appear similar to the robot's sensors. Hence, this presents a challenging approximate inference problem, complicated by the fact that the form of the prior on topologies is far from obvious. We deal with the latter problem by introducing the use of urn models, which very naturally encode prior assumptions in the domain of topological mapping. Secondly, we advance simulated tempering as the basis of two rapidly mixing approximate inference algorithms, based on Markov chain Monte Carlo (MCMC) and Sequential Importance Sampling (SIS), respectively. These algorithms converge quickly even though the posterior being estimated is highly peaked and multimodal. Experiments on real robots and in simulation demonstrate the efficiency and robustness of our technique.

1 Introduction

Bayesian modeling has proved successful in tackling a varied array of problems involving uncertainty. We are interested in robotic mapping, one of the foremost problems in robotics, and an area in which probabilistic solutions have had great success in recent years [22]. Maps created by robots are generally of two types, namely metric and topological maps. While metric maps have been more popular, topological maps provide a light weight, scalable representation of the environment and hence, are useful in many situations [23].

This paper deals with topological mapping using robots. Formally, a topological map is defined as a graph where the nodes correspond to landmarks in the environment and edges represent the connectivity between landmarks. In addition, edges may be annotated with metric or navigational information [11].

Topological mapping is a difficult problem due to, among other things, the occurrence of perceptual aliasing, which occurs when many places in the environment look alike to the robot sensors. This results in the robot being unable to distinguish between previously visited locations, and also being unable to decide whether a given location was previously visited or not. Moreover, since perceptual aliasing is an instance of the data association problem, the state space is combinatorial, making the problem even harder [16].

A Bayesian solution to topological mapping involves computing the posterior distribution over the space of topologies. This can be achieved by maintaining a discrete probability distribution over all the hypotheses for each measurement that is received. The joint distribution of all the measurements is precisely the distribution over the space of topologies. However, computing this is not straightforward due to the associated challenges in modeling and inference.

An important modeling challenge faced by a Bayesian implementation is the definition of a prior in the space of topologies. If the assumptions made about the environment are properly encoded in the prior, the efficiency of the posterior computation can be dramatically improved. On the other hand, inference of the posterior cannot be done analytically due to the combinatorial state space. Instead, recourse has to be taken to sample based approximations. However, the use of basic sampling techniques leads to slow mixing and convergence since the posterior is often highly peaked and multi-modal.

A major contribution in this paper is the use of urn models [9] as priors over topologies. More specifically, we introduce a new urn model that encodes the incremental probability of the robot observing a particular landmark during its run. We model the measurements arising from the same landmark as belonging to the same class. Thus, in our urn

Figure 1: A panoramic image obtained from the robot camera rig

models, the urns correspond to landmarks and the balls correspond to measurements.

As a second contribution, we provide two fast-mixing simulated tempering [7] algorithms, based on Markov chain Monte Carlo (MCMC) and Sequential Importance Sampling (SIS) respectively, for computing the posterior over topologies. Moreover, the MCMC algorithm incorporates a data-driven sampler to improve convergence. We demonstrate that our algorithms converge quickly even though the posterior is multi-modal and highly peaked. This is in stark contrast to standard algorithms that take an inordinately large amount of time.

To prove the validity of our technique, we present experiments in simulation and on a robot operating in an indoor environment. The experiments use odometry and appearance measurements to compute the posterior over topologies. We use Fourier signatures [8] of panoramic images as appearance measurements in our appearance model. The panoramic images (example in Figure 1) are obtained through automatic mosaicking of images from an eight camera rig mounted on the robot. Fourier signatures are a low-dimensional representation of images using Fourier coefficients and hence, allow inexpensive image matching [13]. We use a generative model for the appearance measurements to evaluate the appearance likelihood.

2 Probabilistic Topological Maps

We begin by giving a brief overview of Probabilistic Topological Maps (PTMs) that were introduced in [20]. We start by assuming that the robot is equipped with a "landmark detector" that detects a nearby landmark. We define a PTM as a histogrammed sample-based representation that approximates the posterior distribution $P(T|Z)$ over topologies T given measurements Z. The problem then is to compute the discrete posterior probability distribution $P(T|Z)$ over the space of topologies.

2.1 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, $T = \{S_j \mid j \in [1,M]\}$, 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 jth distinct landmark in the environment.

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. The number of possible topologies is thus equal to the number of set partitions of the set of measurements. This number is called the *Bell number* [17], and grows hyper-exponentially with the number of measurements.

The posterior $P(T|Z)$ is computed through the use of Bayes law. We assume the availability of odometry and appearance measurements, so that $Z = \{O,A\}$, where O is the set of odometry measurements and A is the set of appearance measurements. Clearly now, the appearance and odometry measurements are conditionally independent given the topology. Hence, the posterior can be written as

$$\begin{aligned} P(T \mid O,A) &\propto P(O,A|T)P(T) \\ &= P(O|T)P(A|T)P(T) \end{aligned} \tag{1}$$

The above equation consists of the odometry and appearance likelihoods and the prior on the topologies. The odometry and appearance likelihoods are discussed below while the prior is presented in detail in Section 3.

2.2 The odometry likelihood

To evaluate the odometry likelihood $P(O|T)$ in (1), we marginalize over the set of landmark locations X and calculate the marginal distribution :

$$P(O|T) = \int_X P(O|X,T)P(X|T) \tag{2}$$

where $P(O|X,T)$ is the measurement model, a probability density on O given X and T, and $P(X|T) = P(X)$ is a prior over landmark locations.

For the prior over landmark locations, we assume a cubic penalty function that states that two distinct landmarks are

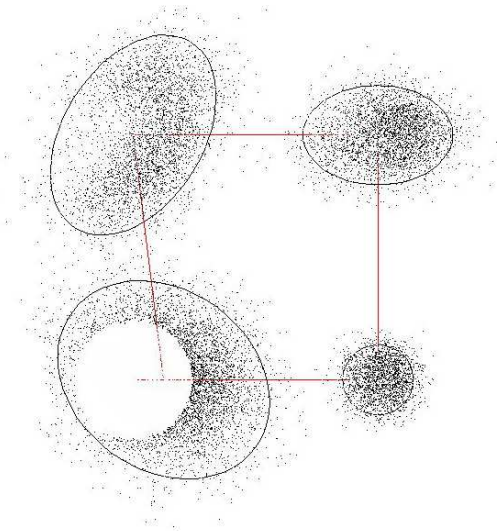


Figure 2: Example square topology with weighted samples used to illustrate the use of importance sampling for evaluating the odometry likelihood. Note the "hole" around the bottom left landmark due to the landmark prior which says that two distinct landmarks cannot be close together. This makes the above topology improbable.

not found close together in the environment. This cubic function is taken to be the negative log likelihood of the prior. Note that since the measurement model in (2) is not known and is non-linear in any case, analytical evaluation of the integral is not possible.

We use importance sampling to evaluate the integral in (2). The proposal distribution for importance sampling is obtained by using Laplace's approximation to the measurement model. This is computed as a Gaussian $Q(X|O,T)$ centered on the maximum likelihood robot path (without taking into consideration the landmark prior) obtained by optimizing the error caused by the topological constraints on the odometry. Subsequently, the likelihood is evaluated as

$$\int_X P(O|X,T)P(X|T) \approx \frac{1}{N}\sum_{i=1}^{N}\frac{P(O|X^{(i)},T)P(X^{(i)})}{Q(X^{(i)}|O,T)} \quad (3)$$

where $X^{(i)}$ are samples from the space of landmark locations obtained from the Gaussian proposal distribution $Q(X|O,T)$ as explained above, and N is the number of samples. An illustration of the evaluation for an example topology is shown in Figure 2. More details are in [20].

2.3 The appearance likelihood

To compute the appearance likelihood $P(A|T)$, we begin by denoting the set of appearance measurements as $A = \{a_i | 1 \leq i \leq N\}$. Evaluation of the appearance likelihood $P(A|T)$ is performed by clustering the appearance measurements according to the sets in the partition corresponding to the topology T.

We model the appearance measurements using a mixture of Gaussians with M components. All the components in the mixture are assumed to have the same variance σ^2 but distinct means given by the vector $Y = \{y_j | 1 \leq j \leq M\}$, where each y_j is the mean of a component in the mixture. The appearance likelihood can be written by marginalization as

$$P(A \mid T) = \int_{Y,\sigma^2} P(A \mid Y,\sigma^2,T)P(Y,\sigma^2|T) \quad (4)$$

where $P(A|Y,\sigma^2,T)$ is the measurement model and $P(Y,\sigma^2|T) = P(Y,\sigma^2)$ is the prior on the appearance. The mean appearance of a landmark is independent of all other landmarks, so that each y_j is independent of all other $y_{j'}$. Additionally, the appearance measurements can also be factored given Y, so that (4) can be written as

$$P(A \mid T) = \int_{y_j} \prod_{j=1}^{M} P(y_j,\sigma^2)^{|S_j|} \prod_{i=1}^{N} P(a_i \mid y_{j_i},\sigma^2,T) \quad (5)$$

where $\{S_j | 1 \leq j \leq M\}$ is the set partition corresponding to T, y_{j_i} is the mean of the set corresponding to a_i, and $P(y_j,\sigma^2)$ is a prior on appearance means and variance. The prior is modeled as a Gaussian-Inverse Gamma distribution so as to be conjugate to the likelihood [5]. This particular choice of conjugate priors allows the integration in (5), and hence the evaluation of the appearance likelihood, to be performed analytically.

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 [13].

3 Urn model priors over topologies

The prior plays an important role in our context since it provides a distribution on the number of distinct landmarks in a sequence of measurements obtained by the robot. We use urn models as priors over topologies. The generic urn model consists of one or more urns in which balls of different colors are added or removed according to a fixed set of stochastic rules. Assumptions regarding the problem at hand can be encoded in the urn model by appropriately defining these rules.

One class of urn models that fit our needs perfectly are Species Sampling Models (SSMs) [18]. An SSM specifies the probability distribution on the number of distinct species in a sequence of samples from some population. The total number of species in the population is initially unknown. Moreover, the probability of obtaining a new species with each successive sample in the sequence is never zero. Regarding the sequence of landmark observations as a sequence of samples from an SSM gives us a prior over topologies. A well-known example of an SSM is

the infinite Polya urn model, which can be shown to be an SSM obtained using the Ewens Sampling Formula [4].

We illustrate the computation of PTMs using an urn model that assumes constant probability for a new landmark at all times. Also, the probability of re-visiting a landmark is independent of the number of times it was visited prevously. This is in contrast to the popular Polya urn model. An urn model with these assumptions was recently proposed in [2] as the Yule-Zipf-Simon model and independently discovered by us. Let the topology T consist of the sequence of landmark observation $s_1, s_2, \ldots, s_n$. The probability of the nth landmark observation, conditioned on the previous observations, is given by the Yule-Zipf-Simon model [2] as

$$p(s_n = k | s_{1:n-1}) = \begin{cases} (1-u)\frac{kz_k(n)}{n} & k \neq 0, n > 1 \\ u & k = 0, n > 1 \\ 1 & n = 1 \end{cases} \quad (6)$$

where u is the constant probability of seeing new landmarks, $z_k(n)$ is the number of landmarks that have been visited k times so far, and $kz_k(n)$ is the total frequency of these landmarks.

The joint probability of the complete sequence of landmarks (i.e. the topology) cannot be evaluated in closed form using (6) [2], unlike, for example, the Polya urn. However, given a specific topology, its probability can be evaluated. Since the sampling algorithms that we use only require the evaluation of the probability of a given topology, the unavailability of an analytical expression for the joint prior on the topology is of no concern.

4 Sampling algorithms for constructing PTMs

We now describe the use of prior and the likelihoods in the sampling algorithms for the construction of PTMs.

4.1 An MCMC algorithm for constructing PTMs

We use the posterior distribution given by (1) as the target distribution in the Metropolis-Hastings (MH) algorithm [1] to compute PTMs. The acceptance ratio of the algorithm is given by

$$a(T', T) = \min\left(1, \frac{q(T' \to T)}{q(T \to T')} \frac{P(T')}{P(T)} \frac{P(O|T')P(A|T')}{P(O|T)P(A|T)}\right) \quad (7)$$

where T is the current state of the Markov chain and T' is the proposed topology. Evaluation of the acceptance ratio involves computing the prior ratio $\frac{P(T')}{P(T)}$ using (6), and the likelihoods ratios using (3) and (5) respectively. $r = \frac{q(T' \to T)}{q(T \to T')}$ is the proposal ratio and comprises the Hastings part of the MH algorithm.

In previous work [20], we used a simple split-merge proposal distribution in the MH algorithm. The proposal merges two sets inside a partition or splits an existing set to create a new partition that is proposed. This proposal has the major problem that many of the proposed topologies are rejected since the acceptance ratio is small. To overcome this problem, we use a data-driven proposal that uses the odometry to propose partitions that are more likely.

The basic computations for the split and merge probabilities are explained in [20] and are not provided here. The major difference in the data-driven proposal is that a merge move is chosen not at random but using the knowledge of the odometry measurements. Intuitively, measurements that are taken with the robot pose close together 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 an optimization similar to the one used to create the importance sampling distribution in (3). The probability of the merge step is then obtained, using the distance D between the landmarks to be merged, as $\exp\left(-\frac{D^2}{\sigma^2}\right)$, where σ^2 is a variance that encodes the odometry error and the scale of the environment.

The analogous calculation for the split step is simple since the probability of the split is just the inverse of total possible moves, merge and split, from T. The data-driven proposal distribution is given in Algorithm 1.

4.1.1 Simulated Tempering for fast mixing

The sampling algorithm using the above proposal still has the problem of slow mixing. This is due to the fact that the posterior distribution is multi-modal and highly peaked, so that the algorithm takes a long time to move from one mode to another. We solve this problem using the Simulated Tempering algorithm, of which we use the Metropolis-Coupled MCMC (MC-cubed) variant [6].

The MC-cubed algorithm works by running N coupled Markov chains in parallel with the first chain having the target distribution of interest $P(T|Z)$ and the other chains having "heated" target distributions $P(T|Z)^\beta$ where $\beta = \frac{1}{1+(i-1)t}$ for the ith chain and t is a constant temperature increment. The peaks in the heated target distributions get increasingly smoothed out so that these chains mix more rapidly. After every chain has been moved one step, an exchange of states takes place between the chains which also enables fast mixing of the original chain. Only samples from the original chain are considered for output[1]. The MC-cubed algorithm we use is given in Algorithm 2.

[1] Considering the time it would take for a single chain to converge, this is still advantageous for reasonable N.

Algorithm 1 The Proposal Distribution

1. Let N_M and N_S be the number of possible split and merge moves from the current topology T. T' is the new topology to be computed below. r is the proposal ratio.

2. Select a merge or a split with probability $\left\{\frac{N_M}{N_M+N_S}, \frac{N_S}{N_M+N_S}\right\}$

3. **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 $q(T \rightarrow T') = \frac{\exp\left(-\frac{D^2}{\sigma^2}\right)}{N_M+N_S}$
     (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'

4. **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 in 2(c)

Algorithm 2 The Metropolis Coupled MCMC Algorithm

1. Let T_i be the current state of the ith chain - total number of chains being N.

2. For all chains $i \in (1, 2, \ldots, n)$ do

   - Propose a new value for T_i using the proposal distribution and acceptance ratio

$$a_i = \min\left(1, \frac{q(T'_i \rightarrow T_i)}{q(T_i \rightarrow T'_i)}\left(\frac{P(T'_i|Z)}{P(T_i|Z)}\right)^{\beta_i}\right)$$
$$\beta_i = \frac{1}{1+(i-1)t}$$

3. After all chains have advanced one cycle, for each consecutive pair of chains i and $i-1$ (starting with n), swap the states of the chains with probability

$$r = \min\left(1, \frac{P(T_i|Z)^{\beta_{i-1}}}{P(T_{i-1}|Z)^{\beta_{i-1}}}\frac{P(T_{i-1}|Z)^{\beta_i}}{P(T_i|Z)^{\beta_i}}\right)$$

4. Goto step 2

Algorithm 3 The SIS algorithm for constructing PTMs

Repeat steps 1 and 2 to obtain N samples, then normalize the weights.

1. Initialize weight $w = 1$

2. For each new measurement i do
   (a) Generate s_i from the multinomial distribution (8)
   (b) Calculate $w_i = p(a_i|s_{1:i-1}, a_{1:i-1})$ from (9)
   (c) Calculate the importance sampling weight for the sample as $w = w \times w_i$.

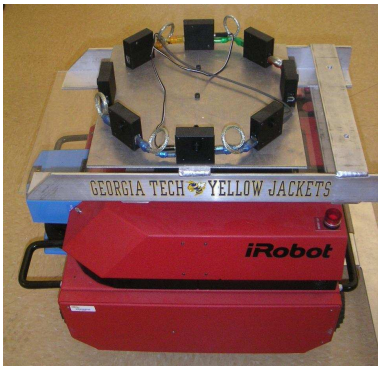


Figure 3: The camera rig mounted on the robot used to obtain panoramic images

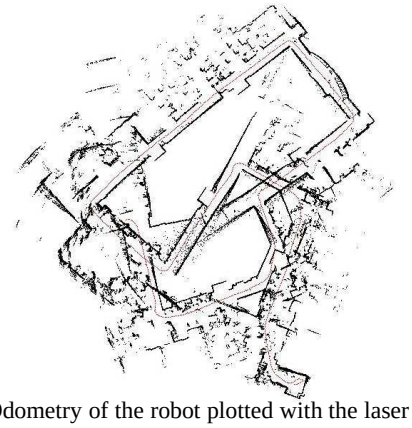


Figure 4: Odometry of the robot plotted with the laser measurements for the first experiment.

4.2 The SIS algorithm for constructing PTMs

The Sequential Importance Sampling (SIS) algorithm [3] for PTMs is based on the sequential nature of the urn model prior, i.e the urn model specifies a discrete distribution on the occurrence of the next landmark given all previous landmarks. We take advantage of this property by combining the appearance likelihood with the urn model prior. This results in an infinite mixture model for the appearance.

The SIS algorithm for PTMs is a slight modification of the Rao-Blackwellized SIS algorithm given in [12]. Upon receiving a landmark measurement, the set of particles is updated by sampling from a predictive distribution involving the odometry and appearance and subsequently weighting the particles. The predictive distribution (or the motion model) for the ith landmark is given by

$$p(s_i = j|P_i, a_i, o_{1:i}) \propto p(o_{1:i}|s_i = j, T_{i-1}) \; p(a_i|s_i, P_i)p(s_i|T_{i-1}) \quad (8)$$

where $P_i = \{s_{1:i-1}, a_{1:i-1}\}$ and $T_{i-1} = s_{1:i-1}$ is the topology with the first $i-1$ measurements.

The prediction step (8) is implemented by sampling from the prior (6) and weighting the samples using the appearance and odometry likelihoods. Note that the odome-

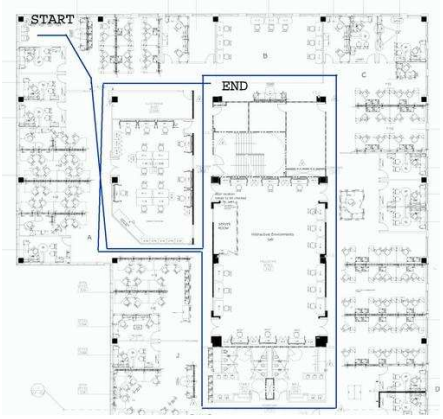


Figure 5: Schematic of robot path overlaid on a floorplan of the environment for the first experiment.

try likelihood cannot be evaluated incrementally (i.e. as $P(o_i|o_{1:i-1}, s_{i:i})$) since the distance between landmarks is not known.

It is now straightforward, using a sequential imputation procedure [10], to derive the expression for the weights at the ith measurement as

$$p(a_i|a_{1:i-1}, s_{1:i-1}) \quad = \quad \sum_{j=1}^{i-1} p(a_i|s_i = j, s_{1:i-1}, a_{1:i-1}) \quad (9)$$

More details can be found in [19]. A summary of the complete SIS algorithm is given in Algorithm 3.

Note that the SIS algorithm does not involve any explicit use of simulated tempering as in the case of the MC-cubed algorithm. However, it has been shown [15] that the Rao-Blackwellized SIS sampler using sequential imputation is also equivalent to simulated tempering. The Rao-Blackwellized posterior obtained upon processing the ith measurement acts in a similar manner to the "heated" target distribution in the MC-cubed algorithm for the $(i+1)$th step. Hence, this results in rapid convergence.

5 Experiments

The experiments were performed on an ATRV-Mini robot mounted with an eight camera rig shown in Figure 3.

The first experiment was conducted over an entire floor of a building and consisted of a robot run containing two loops. Twelve landmarks were observed by the robot during the run, shown overlaid on a floorplan of the experimental area in Figure 5. The odometry of the robot with the laser plotted on top is shown in Figure 4. The five most probable topologies in the PTM[2] are given in Figure 6. The ability of our inference to cope with perceptual aliasing in the environment is demonstrated by the large probability mass

[2]Note that the SIS and MCMC algorithms produce the same PTM since they are computing the same posterior.

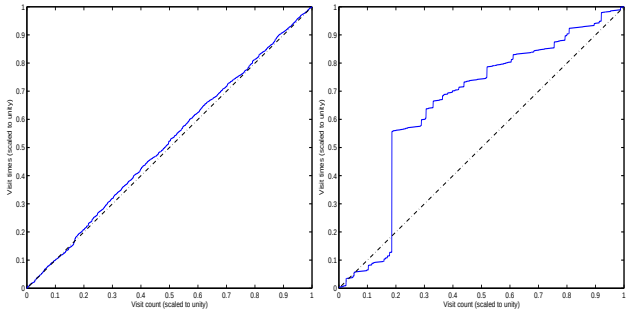


Figure 7: SRQ plots for (a) MC-cubed algorithm (b) single chain MCMC obtained using 15000 samples. The chain produces stable estimates if there are no significant deviations from unit slope.

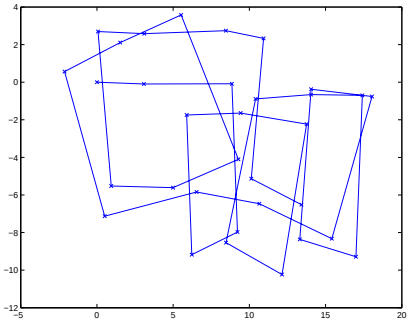
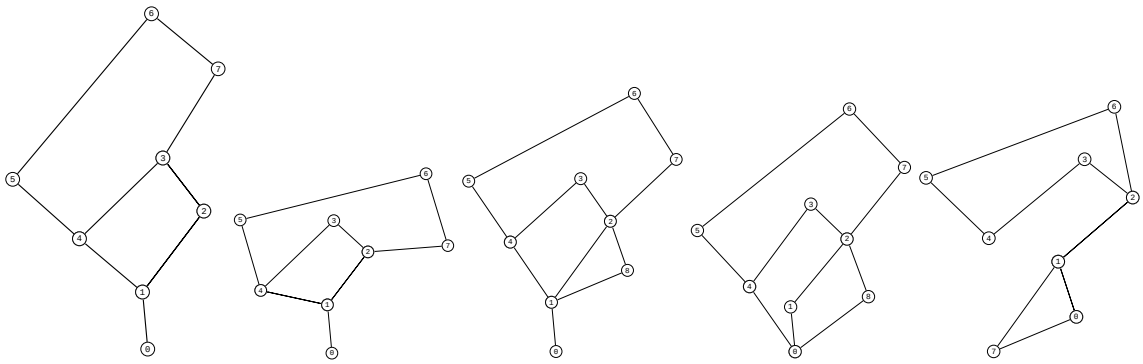


Figure 8: Landmark locations obtained from simulated odometry for the second experiment.

concentrated on the ground truth topology. Computing the PTM using the MC-cubed algorithm takes approximately 7 minutes while the SIS algorithm takes only about half this time on average. In both cases, a value of 0.2 was used for the urn parameter u. Three chains with heating temperatures $t = 10, 20, 30$ were used in the MC-cubed algorithm.

We tested convergence of the MC-cubed algorithm using Scaled Regeneration Quantile (SRQ) plots [14]. An SRQ plot displays visit times to a particular state plotted against the visit count, with both axes scaled to unity. If the chain has converged, the plot should be close to a straight line through the origin with unit slope. Significant changes from unit slope, especially horizontal or vertical segments, signal poor convergence. Figure 7 shows the SRQ plots for the first experiment. It can be seen that the MC-cubed algorithm converges closely to the posterior as opposed to a simple MCMC scheme. The SIS algorithm yields the same posterior as the MC-cubed algorithm and hence, also converges quickly.

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 8. The MC-cubed algorithm converges to the PTM, shown in



(a) (b) (c) (d) (e)

Figure 6: Topologies constituting the PTM for the first experiment using both odometry and appearance. (a) receives **72%** of the probability mass while (b), (c), (d) and (e) receive 12%, 7%, 3% and 2% of the probability mass respectively. The ground truth topology is (a).

Figure 10, in less than 15000 samples while a single chain MCMC algorithm takes over 50000 samples to converge. The MC-cubed and SIS algorithms take approximately 9 and 11 minutes respectively to compute the PTM. It can be seen that the noise in the simulated appearance measurements leads to a smaller proportion of the probability mass on the most probable topology.

The improvement in efficiency when using simulated tempering is quantified in Figure 9. The computation times shown are for plain MCMC, MCMC with the data-driven proposal, and the MC-cubed algorithm with the data-driven proposal. Topologies with 4 nodes (a simple square topology), 9 nodes (from a robot experiment not presented here), 12 and 33 nodes (experiments above) were considered for the comparison. As can be seen, the MC-cubed algorithm greatly improves the scalability of our technique.

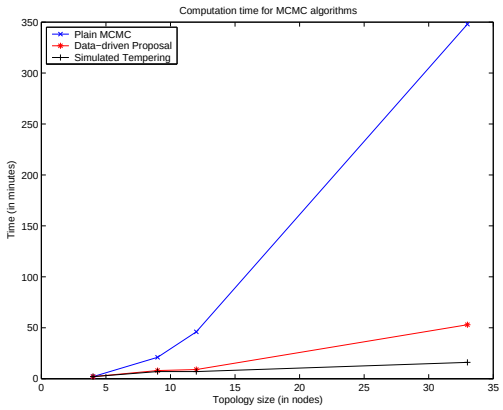


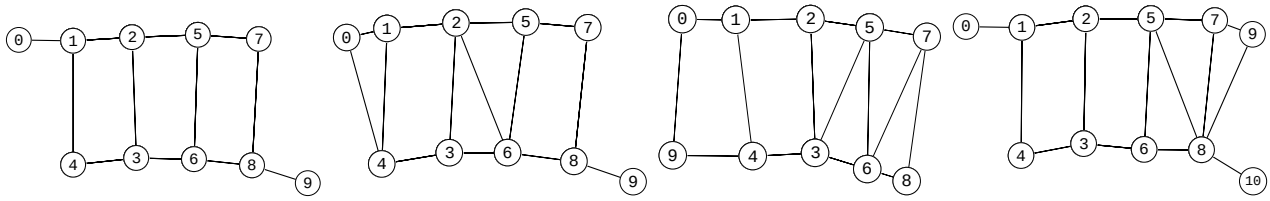
Figure 9: Computation times (rounded to the nearest minute) for the various MCMC algorithms for computing PTMs.

6 Conclusion and future work

We presented an application of Bayesian inference to robot topological mapping that provides a systematic solution to the perceptual aliasing problem. We demonstrated the use of urn models as priors over topologies and gave a characterization of the models that can be used with our technique as belonging to the class of Species Sampling Models (SSMs). Specifically, we introduced a new urn model that corresponds to common assumptions in topological mapping. We showed that the posterior is highly peaked and multi-modal so that ordinary sampling algorithms face the risk of high inefficiency. Instead, we use simulated tempering and a data-driven proposal in an MCMC framework, and a Rao-Blackwellized algorithm in an SIS framework to enable fast mixing and convergence.

In this work, we have illustrated our algorithms using the Yule-Zipf-Simon urn model. However, the algorithms are not restricted to this particular urn model. In particular, we have used the Polya urn model [19] and the Classical Occupancy Distribution [9] to test this claim. Used as priors, these models enable the inclusion of varied prior assumptions in different environments. However, it is to be noted that the Polya urn is a poor model for topological mapping in general since it links the future probability of visiting a landmark with the number of past visits to it. This is certainly not the case with landmarks visited by a robot.

A comparison of the MCMC and SIS algorithms shows that the SIS algorithm is faster for problems with a relatively small number of landmarks. However, as the number of landmarks increases, the SIS algorithm suffers from a problem of dimensionality due to importance sampling. Hence, the algorithm requires a larger number of samples in such cases. The MCMC algorithm, on the other hand, can be significantly slower in many moderately sized environments due to the multiple chains being run simultaneously. It is also seen that MCMC is less sensitive to param-



(a) (b) (c) (d)

Figure 10: Topologies with highest posterior probability mass for the second experiment. (a) the ground truth topology receives **59%** of the probability mass while (b), (c), and (d) receive 9.1%, 8.2%, and 6% of the probability mass respectively.

eter values than the SIS algorithm. This trade-off between the algorithms is an important one to be considered when deploying on a robot.

Our work can be viewed as an instance of the "revisiting problem" that involves the computation of probabilities of places outside the current knowledge of the robot. This problem was addressed in [21] by learning structural models of the environment using a Dirichlet prior while we address it using urn model priors. We believe such an analysis to be more advantageous than the model with a finite number of structure components considered in [21] because the ability of urn models to cope with an infinite number of distinct structures in the environment imparts more generality to the inference.

From our results, it can be seen that the PTM computation is currently not real-time. It is future work to address this issue by replacing the importance sampling in the computation of the odometry likelihood with an analytically computable form. This could also be done through the use of Reversible Jump MCMC. Variational approximations to the posterior are also likely to be a solution.

References

[1] S. Chib and E. Greenberg. Understanding the Metropolis-Hastings algorithm. *American Statistician*, 49(4):327–335, 1995.

[2] D. Costantini, S. Donadio, U. Garibaldi, and P. Viarengo. Herding and clustering in economics: the Yule-Zipf-Simon model. Working Draft, 2004.

[3] A. Doucet, N. de Freitas, and N. Gordon, editors. *Sequential Monte Carlo Methods In Practice*. Springer-Verlag, New York, 2001.

[4] W. J. Ewens. The sampling theory of selectively neutral alleles. *Theoretical Population Biology*, 3:87–112, 1973.

[5] A. Gelman, J. Carlin, H. Stern, and D. Rubin. *Bayesian Data Analysis*. Chapman and Hall, 1995.

[6] C. J. Geyer and E. A. Thompson. Annealing Markov chain Monte Carlo with applications to ancestral inference. *Journal of the American Statistical Association*, 90:909–920, 1995.

[7] Y. Iba. Extended ensemble Monte Carlo. *International Journal of Modern Physics*, C12:623–656, 2001.

[8] H. Ishiguro and S. Tsuji. Image-based memory of environment. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 1996.

[9] N. L. Johnson and S. Kotz. *Urn Models and their Applications*. John Wiley and Sons, 1977.

[10] A. Kong, J. S. Liu, and W. H. Wong. Sequential imputations and Bayesian missing data problems. *Journal of the American Statistical Association*, 89(425):278–288, 1994.

[11] B. 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.

[12] S. N. MacEachern, M. Clyde, and J.S.Liu. Sequential importance sampling for nonparametric Bayes models:the next generation. *Canadian Journal of Statistics*, 27:251–267, 1999.

[13] 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.

[14] P. Mykland, L. Tierney, and B. Yu. Regeneration in Markov chain samplers. *Journal of the American Statistical Association*, 90:233–241, 1995.

[15] R. M. Neal. Sampling from multimodal distributions using tempered transitions. *Statistics and Computing*, 6:353–366, 1996.

[16] J. Neira and J. Tardós. Data association in stochastic mapping using the joint compatibility test. *IEEE Trans. Robot. Automat.*, 17(6):890–897, December 2001.

[17] A. Nijenhuis and H. Wilf. *Combinatorial Algorithms*. Academic Press, 2 edition, 1978.

[18] J. Pitman. Some developments of the Blackwell-Macqueen urn scheme. In T. S. Ferguson, L. S. Shapley, and J. B. MacQueen, editors, *Statistics, Probability and Game Theory; Papers in honor of David Blackwell*, volume 30 of *Lecture Notes-Monograph Series*, pages 245–267.

[19] A. Ranganathan and F. Dellaert. Dirichlet process based Bayesian partition models for robot topological mapping. Technical Report GIT-GVU-04-21, GVU, College of Computing, 2004.

[20] 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.

[21] 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.

[22] S. Thrun. Robotic mapping: a survey. In *Exploring artificial intelligence in the new millennium*, pages 1–35. Morgan Kaufmann, Inc., 2003.

[23] S. Thrun and A. Bücken. Integrating grid-based and topological maps for mobile robot navigation. In *Proceedings of the AAAI Thirteenth National Conference on Artificial Intelligence*, 1996.