Inference In The Space Of Topological Maps: An MCMC-based Approach

Ananth Ranganathan and Frank Dellaert
{ananth,dellaert}@cc.gatech.edu
College of Computing
Georgia Institute of Technology, Atlanta, GA

***Abstract*— While probabilistic techniques have been considered extensively in the context of metric maps, no general purpose probabilistic methods exist for topological maps. We present the concept of Probabilistic Topological Maps (PTMs), a sample-based representation that approximates the posterior distribution over topologies given the available sensor measurements. The PTM is obtained through the use of MCMC-based Bayesian inference over the space of all possible topologies. It is shown that the space of all topologies is equivalent to the space of set partitions of all available measurements. While the space of possible topologies is intractably large, our use of Markov chain Monte Carlo sampling to infer the approximate histograms overcomes the combinatorial nature of this space and provides a general solution to the correspondence problem in the context of topological mapping. We present experimental results that validate our technique and generate good maps even when using *only* odometry as the sensor measurements.**

I. Introduction

One way for a robot to navigate successfully in an uninstrumented environment is for it to build a map. Both metric [1][2][3] and topological maps [4][5][6] have been explored in depth in the mobile robotics community. In both cases, probabilistic approaches have been very successful in dealing with the inherent uncertainties associated with robot perception, that would otherwise make map-building a very brittle process. However, no previous method has dealt with inference in the complete space of topological maps, which is perceived as intractably large.

In this paper we introduce a novel concept, *Probabilistic Topological Maps* (PTMs), a sample-based representation that captures the posterior distribution over all possible topological maps given the available sensor measurements. The key realization is that a distribution over this combinatorially large space can be succinctly approximated by a sample set drawn from this distribution.

The idea of defining a probability distribution over the space of topologies and using sampling to obtain this distribution is, to the best of our knowledge, a completely novel idea. As a second major contribution, we show how to perform inference in the space of topologies given uncertain sensor data from the robot, the outcome of which is exactly a PTM. Specifically, we use Markov chain Monte Carlo (MCMC) sampling [7] to extend the highly successful Bayesian probabilistic framework to the space of topological maps.

Sampling over topologies is accomplished by encoding a topology as a *set partition* over the set of landmark measurements. Each set in the partition corresponds to the measurements arising from a single physical landmark. We then sample over the space of set partitions, using as target distribution the posterior probability over topologies.

PTMs can also be seen as a principled, probabilistic way of dealing with the correspondence problem or "closing the loop" in the context of topological mapping. Previous solutions to the correspondence problem [8][9] commit to a specific correspondence at each step, so that once a wrong decision has been made, the algorithm has difficulty recovering. Computing the posterior distribution over topologies helps solve the correspondence problem in a robust manner. The key to making this work is assuming a simple but very effective prior on the density of landmarks in the environment. We demonstrate that given this prior the additional sensor information used can be very scant indeed. In fact, while our method is general and can deal with any type of sensor measurement, the results we present were obtained using *only* odometry measurements and yet yield nice maps of the environment.

II. Related Work

A major part of the extant work in probabilistic mapping applies to the creation of metric maps, especially as part of the Simultaneous Localization and Mapping (SLAM) problem [3][10]. A good survey of current techniques given in [11] includes various approaches such as Extended Kalman filters, the EM algorithm, particle filters and hybrid methods. Recently, genetic algorithms have been used to search over the space of metric maps [12], where each map is encoded as a chromosome string. The space of candidate solutions is then progressively refined to obtain the maximum-likelihood map. Metric maps suffer from the disadvantage that constructing geometrically accurate maps depends to a large extent on errors in sensors and actuators of the robot [13] and often results in brittle systems. In addition, it is well-known that not all the information in a metric map is required for navigation [14].

Topological maps overcome these drawbacks of metric maps through use of a more qualitative spatial representation [15][4]. Topological maps, as used in this work, are typically graphs where the vertices denote rooms or

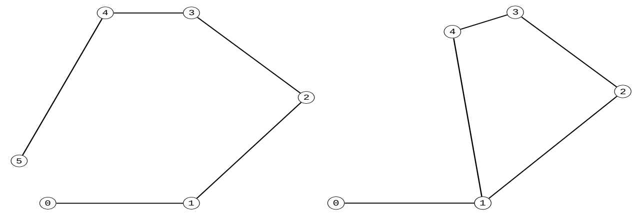

Fig. 1. Two topologies with 6 observations each corresponding to set partitions (a) $(\{0\},\{1\},\{2\},\{3\},\{4\},\{5\})$ and (b) $(\{0\},\{1,5\},\{2\},\{3\},\{4\})$. In (b), the second and sixth measurement are from the same landmark.

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* $Q(T_t'; T_t)$
3) Calculate the *acceptance ratio*

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

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).

other recognizable places, and the edges denote traversals between these places. Such maps are quite useful for planning and symbolic manipulation and, unlike metric maps, do not require precise knowledge of the robot's environment. Unfortunately, they are difficult to build in large scale environments in the presence of ambiguous sensing, for example if two or more recognizable places look very similar [16].

Though probabilistic methods have been used in conjunction with topological maps before, none exist that are capable of dealing with a general, multi-hypothesis, topological space. Most instances of previous work do not deal with general topological maps, but with the use of decision theory to learn a policy to navigate in the environment [17][6][18]. Probabilistic methods for metric SLAM have also been applied to generating topological maps with some success [19].

An approach that is closer to the one presented here, but applicable only to indoor environments, is given by Tomatis and Nourbakhsh [20]. However, while they do maintain a multi-hypothesis space, it is used only to detect points where the probability mass splits in two parts. Finally, the work by Kuipers and Beeson [21] focuses on the identification of distinctive places, but is not concerned with inference about the topologies themselves.

III. Inference in the Space of Topologies

We begin our consideration by assuming that the robot observes N "special places" or landmarks during a run. We assume that the robot is equipped with a "landmark detector" that simply recognizes a landmark when it is near (or on) a landmark, i.e. it is a binary measurement that tells us when landmarks were encountered. We denote by $\{Z_i | 1 \leq i \leq N\}$ the set of sensor measurements recorded by the robot. The results we present in this paper use only odometry measurements though, in general, Z can also include appearance measurements obtained from the landmark locations.

No knowledge of the correspondence between landmark observations and the actual landmarks is given to the robot: indeed, that is exactly the topology that we seek. Given the framework described above, the problem is to compute the discrete posterior probability distribution $P(T|Z)$ over the space of topologies T given the measurements Z.

In this paper, we use the equivalence between the topology of an environment and a set partition of landmark measurements, which groups all measurements into a set of equivalence classes. When all the measurements of the same landmark are grouped together, this naturally defines a partition on the set of measurements. It can be seen that a topology is nothing but the assignment of measurements to sets in the partition, resulting in the above mentioned isomorphism between topologies and set partitions. Figure 1 shows an example encoding of topologies as set partitions.

Formally, for the N element measurement set Z, a partition P can be represented as $P = \{S_i \mid i \in [1, M]\}$, where the S_i are disjoint sets of measurements and $M \leq N$ is the number of sets in the partition (and also the number of distinct landmarks in the environment). In the context of topological mapping, all members of the set S_i represent landmark observations of the ith landmark. The cardinality of the space of topologies over a set of N landmark observations is identical to the number of disjoint set partitions of the N-set. This number is called the Bell number b_N [22], defined as $b_N = \frac{1}{e} \sum_{k=0}^{\infty} \frac{k^N}{k!}$, and grows hyper-exponentially with N, for example $b_1 = 1$, $b_3 = 5$ but $b_{15} = 190,899,522$. The combinatorial nature of this space makes exhaustive evaluation impossible for all but trivial environments.

IV. Inferring PTMs using MCMC

We define a *Probabilistic Topological Map* to be a data structure that approximates the posterior distribution $P(T|Z)$ over topologies T, given measurements Z. While the space of possible topological maps is combinatorially large, a PDF over this space can be approximated by drawing samples from the distribution over possible maps. The samples are obtained using Markov chain Monte Carlo sampling. The PTM is then a histogram constructed on the support of this sample set.

We use the Metropolis-Hastings (MH) algorithm [7], a very general MCMC method, for performing inference.

Algorithm 2 The Proposal Distribution

1) select a merge or a split with probability 0.5
2) if a merge is selected go to 3, else go to 4
3) **merge move**:
   - if T contains only one set, re-propose $T' = T$, hence $r = 1$
   - otherwise select two sets at random P and Q, and
     a) $T' = (T - \{P\} - \{Q\}) \cup \{P \cup Q\}$ and $q(T'|T) = \frac{1}{N_M}$
     b) $q(T|T')$ is obtained from the reverse case 4(b), hence $r = N_M^{-1} N_S \left\{ {|P \bigcup Q| \atop 2} \right\}$, where N_S is the number of possible splits in T'
4) **split move**:
   - if T contains only one set, re-propose $T' = T$, hence $r = 1$
   - otherwise select a non-singleton set R at random from T, split it into two sets P and Q, and
     a) $T' = (T - \{R\}) \cup \{P, Q\}$ and $q(T'|T) = \left(N_S \left\{ {|R| \atop 2} \right\} \right)^{-1}$
     b) $q(T|T')$ is obtained from the reverse case 3(b), hence $r = N_M \left(N_S \left\{ {|R| \atop 2} \right\} \right)^{-1}$, where N_M is the number of possible merges in T'

All MCMC methods work by generating a sequence of *states* from a Markov chain, with the property that the generated states are samples from the target distribution. In our case the state space that is sampled over is the space of set partitions, where each partition represents a different topology of the environment. The pseudo-code to generate a sequence of samples from the target distribution using the MH algorithm is shown in Algorithm 1 (adapted from [7]). The MH algorithm uses a *proposal distribution* $Q(T_t; T_t')$ to propose moves, the tentative next state in the Markov chain at each step, in the space of topologies. Intuitively, the algorithm samples from the desired probability distribution $P(T|Z)$ by rejecting a fraction of the moves generated by the proposal distribution. The fraction of moves rejected is governed by the acceptance ratio a, where most of the computation takes place.

The two hurdles to sampling using an MCMC sampler are the design of the proposal density and evaluation of the target density. The details of these are discussed below.

A. The Proposal Distribution

The proposal distribution operates by proposing one of two moves, a *split* or a *merge*, with equal probability at each step. Given that the current sample topology has M distinct landmarks, the next sample is obtained by splitting a set or merging two sets in the partition T and may have M, $M + 1$, or $M - 1$ distinct landmarks.

The **merge move** merges two randomly selected sets in the partition to produce a new partition with one less set than before. The probability of a merge is simply $1/N_M$, N_M being the number of possible merges given by $\binom{M}{2}$.

The **split move** splits a randomly selected set in the partition to produce a new partition with one more set than before. To calculate the probability of a split move, let N_S be the number of sets in the partition with more than one element. Clearly, N_S is the number of sets in the partition that can be split. Out of these N_S sets, we pick a random set R to split. Then, the number of possible ways to split R into two subsets is given by the Stirling number of second kind for R, $\left\{ {|R| \atop 2} \right\}$, where the Stirling number itself is defined recursively as $\left\{ {n \atop m} \right\} \triangleq \left\{ {n-1 \atop m-1} \right\} + m \left\{ {n-1 \atop m} \right\}$ [22]. Hence, the probability of the split is $\left(N_S \left\{ {|R| \atop 2} \right\} \right)^{-1}$. A random split of R can be generated efficiently by using the recursive algorithms described in [22].

The proposal distribution is summarized in pseudo-code format in Algorithm 2, where q is the proposal distribution and $r = \frac{q(T'|T)}{q(T|T')}$ is the proposal ratio. It is to be noted that this proposal distribution does not incorporate any domain knowledge but uses only the combinatorial properties of set partitions to propose moves.

B. Evaluating the Target Distribution

In addition to proposing new moves in the space of topologies, we also need to evaluate the posterior probability $P(T|Z)$ for each proposed topology change. Using Bayes Law, we obtain

$$P(T \mid Z) \propto P(Z|T)P(T) \tag{2}$$

where $P(T)$ is a prior and $P(Z|T)$ is the observation likelihood. In this work, we assume a non-informative uniform prior over all topologies, but it is also possible to use a Poisson distribution on the number of landmarks in the environment if some evidence for this exists.

It is not possible to evaluate the likelihood $P(Z|T)$ without knowledge of the landmark locations. Hence, in a process called Rao-Blackwellization [24], we integrate over the set of landmark locations X to calculate the marginal distribution $P(Z|T)$ from the joint distribution $P(Z, X|T)$. The likelihood $P(Z|T)$ is then given as

$$P(Z|T) \propto \int_X P(Z|X, T) P(X|T) \tag{3}$$

where $P(Z|X, T)$ is the measurement model, an arbitrary density on Z given X and T, and $P(X|T)$ is the prior on landmark locations. As an example, in a 2D environment, commonly assumed in the robotics literature, we have $X = \{X_t = (x_t, y_t, \theta_t) | 1 \le t \le N\}$.

A prior on the distribution of the landmark locations X given the topology T, $P(X|T)$, is required to evaluate (3). In our case, the prior is used to encode the assumption that distinct landmarks do not lie close together in the environment. For this purpose, a penalty function is used to penalize topologies containing distinct landmark measurements that are spatially close. Specifically, the penalty function used is a cubic function as in Figure 2. The prior

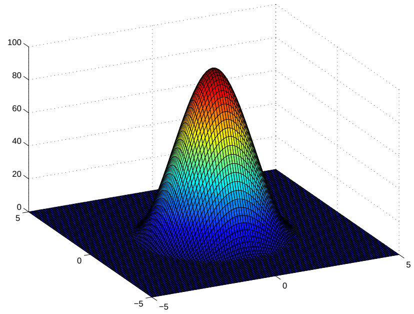

Fig. 2. Cubic penalty function used as a prior over landmark density

on landmark locations, $P(X|T)$, is then

$$P(X|T) \quad = \quad e^{-\sum_{1\le i<j\le N} f(X_i,X_j)} \qquad (4)$$

where f is the penalty function, and X_i and X_j do not belong to the same set.

Assuming the availability of only odometry measurements, we can write the negative log-likelihood function corresponding to $P(Z|T)$ in (3) as

$$L(X) = \left(\frac{X - X_O}{\sigma_O}\right)^2 + \sum_{S\in T}\sum_{i,j\in S}\left(\frac{X_i - X_j}{\sigma_T}\right)^2 + \sum_{1\le i<j\le N} f(X_i, X_j) \qquad (5)$$

where S is a set in the partition corresponding to T, σ_O and σ_T are standard deviations explained below, and X_o is the set of landmark locations obtained from the odometry measurements. The intuition here is that the topology T constrains some measurements as being from the same location even though the odometry may put these locations far apart. The log-likelihood function accounts for the error from distorting the odometry, the first term in (6), and the error for not conforming to the topology T, the second term in (6). The error for not conforming to a topology is expressed through a set of "soft constraints". These constraints try to place two observations that are ascribed to the same landmark by the topology at the same physical location. The standard deviations for the odometry and soft constraints, σ_O and σ_T respectively, encode the amount of error that we are willing to tolerate in each of these quantities. The final term in (5), where the sum is over all X_i and X_j that do not belong to the same set, is simply the negative log-likelihood of the prior in (4).

C. Numerical Evaluation of the Target Distribution

Though in some cases integral (3) may be evaluated analytically using the functional form of the log-likelihood given in (5), this is not possible in general. Instead, we use a Monte Carlo approximation to evaluate the integral, using importance sampling [25] to approximate the integrand $P(Z|X,T)P(X|T)$. Given a target distribution to be sampled, importance sampling requires a proposal distribution

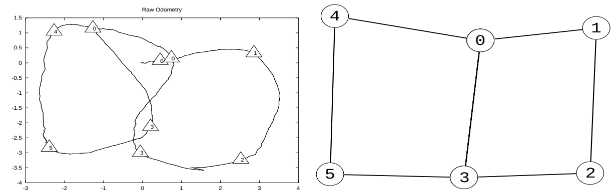

Fig. 3. Raw odometry (left) and Ground truth topology (right) from the first experiment involving 9 observations

from which samples are actually obtained. Subsequently, these samples are weighted by their "importance", i.e. the ratio of the target distribution to the proposal distribution at the sample point. The weighted samples can then be used in Monte Carlo integration.

In our case, the importance sampling proposal distribution is an approximation of the log-likelihood in (5). Firstly, ignoring the final term corresponding to the prior in (5), we obtain the function

$$\psi(X) \quad = \quad \left(\frac{X - X_O}{\sigma_O}\right)^2 + \sum_{S\in T}\sum_{i,j\in S}\left(\frac{X_i - X_j}{\sigma_T}\right)^2 \qquad (6)$$

Subsequently, Laplace's method is employed to obtain a multivariate Gaussian distribution from $\psi(X)$, which is used as the proposal distribution. This is achieved by computing the maximum likelihood path $X^\star$ through a non-linear optimization of $\psi(X)$, and creating a local Gaussian approximation $Q(X|Z,T)$ around $X^\star$

$$X^\star = \underset{X}{\text{argmax}} \quad \psi(X)$$

$$Q\left(X \mid Z, T\right) = \frac{1}{\sqrt{|2\pi\Sigma|}} e^{-\frac{1}{2}(X - X^\star)^T \Sigma^{-1} (X - X^\star)}$$

where Σ is the covariance matrix relating to the curvature of $\psi(X)$ around X^*. The distribution $Q(X|Z,T)$ is then used as the proposal distribution for the importance sampler.

The posterior given by (3) is now evaluated using the Monte Carlo approximation

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

where $X^{(i)}$ denote the N samples obtained from the Gaussian proposal distribution $Q(X|Z,T)$.

V. Results

We performed two experiments consisting of runs with nine observations each, a short run of about 15 meters and a longer one covering a complete floor of a building. The platform used for the experiments was an I-Robot ATRV-Mini with a frontal SICK laser range finder. In all cases, the sampler was initialized with a topology that assigned distinct landmarks to each observation.

In the first experiment we explored the influence of the penalty-term in a small, lab-like environment, the results of

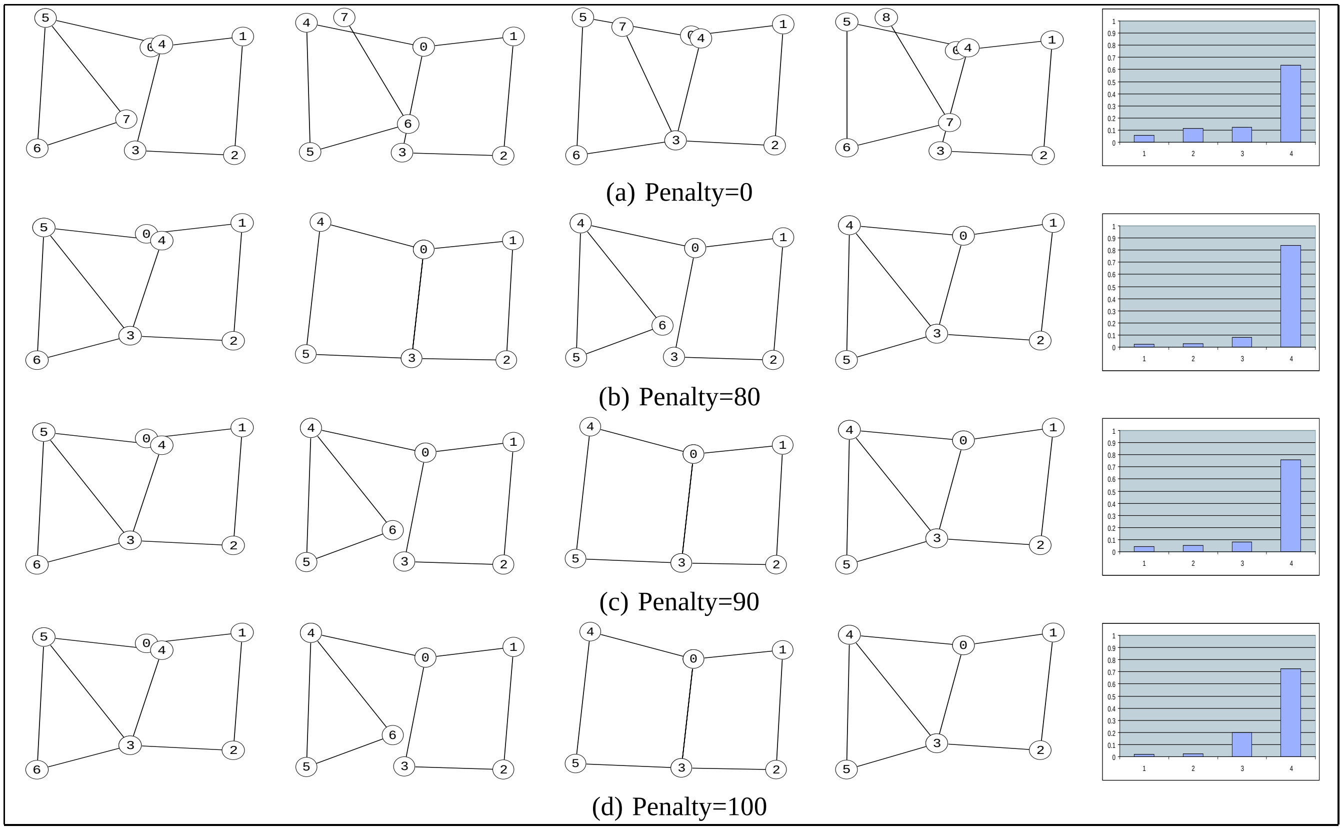

TABLE I

CHANGE IN PROBABILITY MASS WITH MAXIMUM PENALTY OF THE FOUR MOST PROBABLE TOPOLOGIES IN THE HISTOGRAMMED POSTERIOR. THE HISTOGRAM AT THE END OF EACH ROW GIVES THE PROBABILITY VALUES FOR EACH TOPOLOGY IN THE ROW.

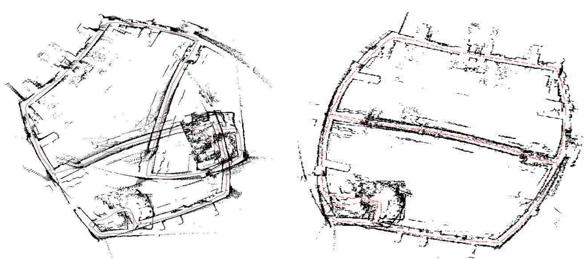

Fig. 4. Map obtained by plotting laser on raw odometry (left) and the laser plot corresponding to odometry from the correct topology (right)

which are shown in Table I. The experiment was performed on a short run approximately 15 meters long during which the robot observed nine landmarks. The raw odometry from the run and the corresponding ground truth topology are shown in Figure 3. The table shows the evolution of the Markov Chain sampler for different values of the maximum penalty. In our algorithm, it is the penalty term that facilitates merging of nodes in the map that are the same. Without the penalty, the system has no incentive to move toward a topology with lesser number of nodes as this increases the odometry error. Table I(a) illustrates this case. It can be seen that the topology that is closest to the odometry data and also having the maximum possible nodes gets the maximum probability mass. For the next two cases with maximum penalties equal to 80 and 90 respectively, the most likely solution is a compromise between the ground truth solution and the odometry. Also, it is to be noted that the large error in odometry makes the ground truth topology less likely compared to topologies such as the most likely one in Table I(b). In spite of this, as the penalty is increased the effect of the odometry is diminished and the ground truth topology gains probability mass. However, a very large penalty swamps the odometry data and makes absurd topologies more likely.

The second experiment demonstrates that PTMs have the power to close the loop even in large environments. This experiment involved a complete floor of the building containing our lab during which nine landmark observations were recorded. The raw odometry with laser readings plotted over it is shown in Fig 4. Also shown is the map obtained by plotting the maximum likelihood path with laser readings on top. It can be seen from Figure 5, which gives the most probable topologies in the posterior, that the correct topology receives the largest probability mass.

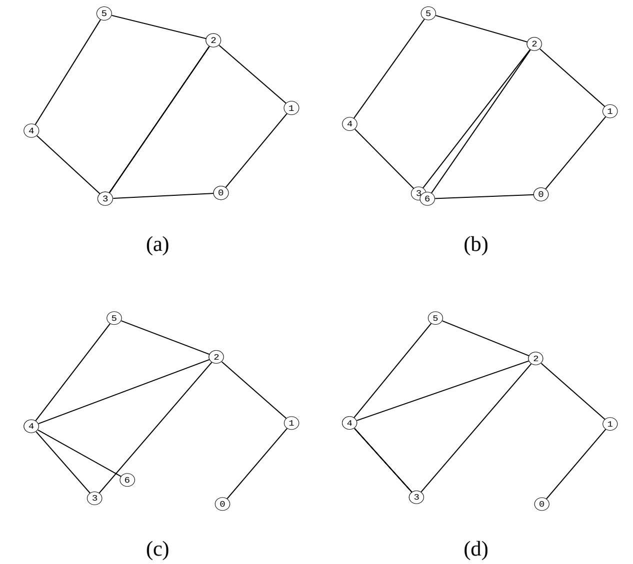

Fig. 5. Topologies with highest probability mass for the second experiment (a) The correct topology receives 97% of the mass (b), (c) and (d) receive 2%, 0.5% and 0.5% of the mass respectively

VI. Discussion

We presented the novel idea of computing discrete probability densities over the space of all possible topological maps. Probabilistic Topological Maps are computed using Markov chain Monte Carlo sampling over set partitions that are used to encode the topologies. We use a simple spatial prior in the form of a cubic penalty function that disallows proximity among landmarks. Experimental results on environments with varied sizes hold promise for the applicability and further improvements of PTMs.

One advantage of our approach is that an estimate of topology is possible even if only a meager amount of information is available. It is not the purpose of this work to find the best topological map but to compute the posterior distribution over topological space as per the Bayesian approach. We have shown this capability in experiments that use only odometry to create distributions that can either correspond to the odometry or the prior (in this case the spatial penalty function) as parameters are varied.

The next step is to include range sensors and appearance models in our technique. It is also future work to induct domain-specific knowledge into the proposal distribution of the MCMC sampler and include a more informative prior. Finally, the present algorithm is sensitive to parameter settings of the penalty function, which needs to be addressed.

Acknowledgments

We would like to thank Emanuele Menegatti for valuable comments, S. Charles Brubaker for his earlier work on this project, and Ruth Conroy-Dalton for many helpful discussions on spatial priors. We are also grateful to Michael Kaess for providing the data used in the experiments.

References

[1] A. Elfes, "Occupancy grids: A probabilistic framework for robot perception and navigation," *Journal of Robotics and Automation*, vol. RA-3, no. 3, pp. 249–265, June 1987.

[2] H. Moravec, "Sensor fusion in certainty grids for mobile robots," *AI Magazine*, vol. 9, pp. 61–74, 1988.

[3] 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.

[4] H. Choset and K. Nagatani, "Topological simultaneous localization and mapping (SLAM): toward exact localization without explicit localization," *IEEE Trans. on Robotics and Automation*, vol. 17, no. 2, April 2001.

[5] E. Remolina and B. Kuipers, "Towards a general theory of topological maps," *Artificial Intelligence*, vol. 152, no. 1, pp. 47–104, 2004.

[6] H. Shatkay and L. Kaelbling, "Learning topological maps with weak local odometric information," in *Proceedings of IJCAI-97*, 1997.

[7] W. Gilks, S. Richardson, and D. Spiegelhalter, Eds., *Markov chain Monte Carlo in practice*. Chapman and Hall, 1996.

[8] S. Thrun, D. Fox, and W. Burgard, "A probabilistic approach to concurrent mapping and localization for mobile robots," *Machine learning*, vol. 31, pp. 29–53, 1998.

[9] K. Konolige and J.-S. Gutmann, "Incremental mapping of large cyclic environments," in *International Symposium on Computational Intelligence in Robotics and Automation (CIRA'99)*, 1999.

[10] H. Durrant-Whyte, S. Majunder, S. Thrun, M. de Battista, and S. Scheding, "A Bayesian algorithm for simultaneous localization and map building," 2001, submitted for publication.

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

[12] T. Duckett, "A genetic algorithm for simultaneous localization and mapping," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2003, pp. 434–439.

[13] R. Brooks, "Aspects of mobile robot visual map making," in *Second. Int. Symp. Robotics Research*. MIT press, 1984.

[14] B. Kuipers, "The cognitive map: Could it have been any other way?" in *Spatial Orientation: Theory, Research, and Application*, H. L. P. Jr. and L. P. Acredolo, Eds. New York: Plenum Press, 1983.

[15] 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*, vol. 8, pp. 47–63, 1991.

[16] G. Dudek, M. Jenkin, E. Milios, and D. Wilkes, "Robotic exploration as graph construction," *IEEE Transactions on Robotics and Automation*, vol. 7, no. 6, pp. 859–865, 1991.

[17] R. Simmons and S. Koenig, "Probabilistic robot navigation in partially observable environments," in *Proc. International Joint Conference on Artificial Intelligence*, 1995.

[18] L. Kaelbling, A. Cassandra, and J. Kurien, "Acting under uncertainty: Discrete Bayesian models for mobile-robot navigation," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1996.

[19] S. Thrun, S. Gutmann, D. Fox, W. Burgard, and B. Kuipers, "Integrating topological and metric maps for mobile robot navigation: A statistical approach," in *AAAI*, 1998, pp. 989–995.

[20] N. Tomatis, I. Nourbakhsh, and R. Siegwart, "Hybrid simultaneous localization and map building: Closing the loop with multi-hypotheses tracking," in *Proc. of the IEEE Intl. Conf. on Robotics and Automation*, 2002.

[21] B. Kuipers and P. Beeson, "Bootstrap learning for place recognition," in *AAAI Nat. Conf. on Artificial Intelligence*, 2002.

[22] A. Nijenhuis and H. Wilf, *Combinatorial Algorithms*, 2nd ed. Academic Press, 1978.

[23] W. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, pp. 97–109, 1970.

[24] G. Casella and C. Robert, "Rao-Blackwellisation of sampling schemes," *Biometrika*, vol. 83, no. 1, pp. 81–94, March 1996.

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