Analyzing Gaussian Proposal Distributions for Mapping with Rao-Blackwellized Particle Filters

Cyrill Stachniss* Giorgio Grisetti* Wolfram Burgard* Nicholas Roy†

Abstract— **Particle filters are a frequently used filtering technique in the robotics community. They have been successfully applied to problems such as localization, mapping, or tracking. The particle filter framework allows the designer to freely choose the proposal distribution which is used to obtain the next generation of particles in estimating dynamical processes. This choice greatly influences the performance of the filter. Many approaches have achieved good performance through informed proposals which explicitly take into account the current observation. A popular approach is to approximate the desired proposal distribution by a Gaussian. This paper presents a statistical analysis of the quality of such Gaussian approximations. We also propose a way to obtain the optimal proposal in a non-parametric way and then identify the error introduced by the Gaussian approximation. Furthermore, we present an alternative sampling strategy that better deals with situations in which the target distribution is multi-modal. Experimental results indicate that our alternative sampling strategy leads to accurate maps more frequently that the Gaussian approach while requiring only minimal additional computational overhead.**

I. Introduction

Particle filters are a frequently used technique in robotics for dynamical system estimation. They have been used to localize robots [4], to build both feature-maps [12], [13] and grid-maps [7], [8], [9], and to track objects based on vision data [10]. A particle filter approximates the posterior by a set of random samples and updates it in a recursive way. The particle filter framework specifies how to update the sample set but leaves open how to choose the so-called proposal distribution. The proposal is used to draw the next generation of samples at the subsequent time step in the dynamical process. For example, in the context of localizing a robot, the odometry motion model is a good choice for the proposal in that it can be easily sampled and then easily transformed into the target distribution by such techniques as weighted importance sampling. In practice, the design of the proposal has a major influence on the performance and robustness of the filtering process. On the one hand, the closer the proposal is to the target distribution, the better is the estimation performance of the filter. On the other hand, the computational complexity of the calculation of the proposal distribution should be small in order to run the filter online. For this reason, the majority of particle filter applications restrict the proposal distribution to a Gaussian since one can efficiently draw samples from such a distribution.

Murphy, Doucet, and colleagues [6], [14] introduced factored particle filters, known as "Rao-Blackwellization", as an effective means to solve the simultaneous localization and mapping (SLAM) problem. By applying this factorization, several efficient mapping algorithms have been presented [7], [8], [9], [12] and we can note that all of these algorithms have used Gaussians to obtain the next generation of particles.

In this paper, we analyze how well such Gaussian proposal distributions approximate the optimal proposal in the context of mapping. We apply well-founded statistical measures to carry out the comparisons. To the best of our knowledge, this question has not been addressed in the context of particle filter applications in robotics so far. It turns out that Gaussians are often an appropriate choice but there exist situations in which multi-modal distributions are needed to appropriately sample the next generation of particles. Based on this insight, we present an alternative sampling technique that has the same complexity as the Gaussian approximation but can appropriately capture distributions with multiple modes, resulting in more robust mapping systems.

This paper is organized as follows. After a discussion of related approaches, we briefly introduce in Section III the ideas of mapping with Rao-Blackwellized filters. In Section IV, we explain how to actually represent and sample from the optimal proposal. We then present an efficient variant that allows us to deal with multi-modal proposals in an efficient way. In Section VI, we introduce the statistical tests that are used in the experimental section for evaluation.

II. Related Work

Particle filters have been applied to various kinds of robotic state estimation problems such as localization [4], mapping [7], [8], [9], [12], visual tracking [10], or data association problems [20]. Murphy, Doucet, and colleagues were the first that presented an approach based on a Rao-Blackwellized particle filter that learns grid maps [6], [14]. The first efficient approach for mapping with Rao-Blackwellized particle filters was the FastSLAM algorithm by Montemerlo *et al.* [13]. It uses a set of Kalman filters to represent the map features conditioned on a sampled robot pose. A Gaussian process model is used to sample the odometry motion model and generate the proposal distribution on the next step. The grid-based variant presented by Haehnel *et al.* [9] performs scan-matching as a preprocessing step. In this way, they are able to draw samples from Gaussians with lower variances compared to proposals computed based on the odometry only. This reduces the number of required particles and allows a robot to maintain a map estimate online. In contrast to that, Eliazar *et al.* [7] focus on an efficient grid map representation which allows the particles

*University of Freiburg, Department of Computer Science, D-79110 Freiburg. †MIT, 77 Massachusetts Ave., Cambridge, MA 02139-4307

to share a map. Subsequently, Montemerlo *et al.* published FastSLAM2 [12] that uses an informed proposal based on the most recent sensor observation to restrict the space for sampling. Again, to efficiently draw the next generation of particles, the distribution is assumed to be Gaussian. Grisetti *et al.* [8] extended FastSLAM2 to deal with large-scale occupancy grid maps. This technique combines scan-matching on a per particle basis with informed Gaussian proposal distributions.

To the best of our knowledge, there exists no evaluation of how well the Gaussian proposal distributions approximate the optimal proposal which in general is non-Gaussian in the context of mapping. There exist approaches that show that the uncertainty of certain SLAM techniques monotonically decreases over time. For example, Newman proved this property for the relative map filter and also showed that "in the limit, as the number of observations increases, the relative map becomes perfectly known" [15]. In the context of particle filters for SLAM, Montemerlo *et al.* [12] showed that FastSLAM2 "converges [...] for a restricted class of linear Gaussian problems". It, however, makes no statement about the validity of Gaussian approximations in real world settings.

III. Learning Maps with Rao-Blackwellized Particle Filters

A particle filter requires three sequential steps to update its estimate. Firstly, one draws the next generation of samples from the so-called proposal distribution π. Secondly, one assigns a weight to each sample. The weights account for the fact that the proposal distribution is in general not equal to the target distribution. The third step is the resampling step in which the target distribution is obtained from the weighted proposal by drawing particles according to their weight.

In the context of the SLAM problem, one aims to estimate the trajectory of the robot as well as a map of the environment. The key idea of a Rao-Blackwellized particle filter for SLAM is to separate the estimate of the trajectory $x_{1:t}$ of the robot from the map m of the environment. This is done by the following factorization

$$\begin{aligned} &p(x_{1:t}, m \mid z_{1:t}, u_{1:t-1}) = \\ &\quad p(m \mid x_{1:t}, z_{1:t}) \cdot p(x_{1:t} \mid z_{1:t}, u_{1:t-1}), \end{aligned} \tag{1}$$

where $z_{1:t}$ is the observation sequence and $u_{1:t-1}$ the odometry information. In practice, the first term of Eq. (1) is estimated using a particle filter and the second term turns into "mapping with known poses".

One of the main challenges in particle filtering is to choose an appropriate proposal distribution. The closer the proposal is to the true target distribution, the more precise is the estimate represented by the sample set. Typically, one requires the proposal π to fulfill the assumption

$$\begin{aligned} \pi(x_{1:t} \mid z_{1:t}, u_{1:t-1}) \;=\; &\pi(x_t \mid x_{1:t-1}, z_{1:t}, u_{1:t-1}) \\ &\cdot \pi(x_{1:t-1} \mid z_{1:t-1}, u_{1:t-2}). \end{aligned} \tag{2}$$

According to Doucet [5], the distribution

$$\begin{aligned} &p(x_t \mid m_{t-1}^{(i)}, x_{t-1}^{(i)}, z_t, u_{t-1}) = \\ &\qquad \frac{p(z_t \mid m_{t-1}^{(i)}, x_t) p(x_t \mid x_{t-1}^{(i)}, u_{t-1})}{p(z_t \mid m_{t-1}^{(i)}, x_{t-1}^{(i)}, u_{t-1})} \end{aligned} \tag{3}$$

is the optimal proposal for particle i with respect to the *variance of the particle weights* that satisfies Eq. (2). This proposal minimizes the degeneracy of the algorithm (Proposition 4 in [5]). As a result, the computation of the weights turn into

$$w_t^{(i)} = w_{t-1}^{(i)} \frac{\eta p(z_t \mid m_{t-1}^{(i)}, x_t^{(i)}) p(x_t^{(i)} \mid x_{t-1}^{(i)}, u_{t-1})}{p(x_t \mid m_{t-1}^{(i)}, x_{t-1}^{(i)}, z_t, u_{t-1})} \tag{4}$$

$$\propto w_{t-1}^{(i)} \frac{p(z_t \mid m_{t-1}^{(i)}, x_t^{(i)}) p(x_t^{(i)} \mid x_{t-1}^{(i)}, u_{t-1})}{\frac{p(z_t \mid m_{t-1}^{(i)}, x_t) p(x_t \mid x_{t-1}^{(i)}, u_{t-1})}{p(z_t \mid m_{t-1}^{(i)}, x_{t-1}^{(i)}, u_{t-1})}} \tag{5}$$

$$= w_{t-1}^{(i)} \cdot p(z_t \mid m_{t-1}^{(i)}, x_{t-1}^{(i)}, u_{t-1}) \tag{6}$$

$$= w_{t-1}^{(i)} \cdot \int p(z_t \mid x') p(x' \mid x_{t-1}^{(i)}, u_{t-1}) \, dx'. \tag{7}$$

Unfortunately, the optimal proposal distribution is in general not available in closed form or in a suitable form for efficient sampling. As a result, most efficient mapping techniques use a Gaussian approximation of the optimal proposal. This approximation is easy to compute and allows the robot to sample efficiently. As we will show in this paper, the Gaussian assumption is not always justified. To provide examples for this statement, we first compute the optimal proposal explicitly and then compare it to the Gaussian approximation. Using the optimal proposal in a mapping system leads to computationally expensive operations which are explained in the next section in more detail.

IV. Computing and Sampling from the Optimal Proposal

This section explains how to compute the optimal proposal and how to sample from that distribution. In mapping as well as in many other problems, there is no closed form solution available but we can arrive at a high-fidelity numerical solution for the likelihood function. In our case, the numerator of Eq. (3) is the product of the observation likelihood and the odometry motion model. When using laser range finders, the dominating factor is the observation likelihood. To point-wise evaluate the observation likelihood, we use the so called "beam endpoint model" [19]. In this model, the individual beams within a scan are considered to be independent. Furthermore, the likelihood of a beam is computed based on the distance between the endpoint of the beam and the closest obstacle from that point. Using this point-wise evaluation of the observation likelihood, we can compute a three-dimensional histogram providing the observation likelihood for the different poses.

The second term in Eq. (3) is the robot motion model. In this paper, we consider the "banana-shaped" distribution known from most approaches to Monte-Carlo localization [4]. The likelihood for the individual poses is computed

point-wise and is stored in a histogram. This histogram describes the likelihood function in a non-parametric form. Histograms, however, are affected by discretization errors. To smooth this effect, we furthermore apply the Parzen window/kernel estimator [1] based on the evaluated data points. Let x^j be the evaluated poses, then this estimator is defined as

$$\hat{p}(x) \quad = \quad \frac{p(x^j)}{h} \sum_{j=1}^{n} K\left(\frac{x - x^j}{h}\right) \tag{8}$$

where h is called Parzen window. We chose the kernel $K(u)$ as

$$K(u) \quad = \quad \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right). \tag{9}$$

This technique allows us to smooth the histogram data and in this way avoid the discontinuities which are inherent in the histogram representation itself. Furthermore, we can make the likelihood of the smoothed histogram arbitrarily close to the optimal distribution of Eq. (3) by increasing the resolution of the local grid map and reducing the size of the histogram bins.

Given this non-parametric estimator, we can perform rejection sampling to draw the next generation of particles. Obviously, this results in a highly inefficient mapping system with respect to the computation time. However, it allows us to sample from an arbitrarily close approximation to the optimal proposal distribution and to compare it to its Gaussian approximation.

As we will illustrate in the experiments, in most cases the proposal can be safely approximated by a Gaussian. This explains why existing methods based on this particular approximation have been so successful. In certain situations, however, the distribution is highly non-Gaussian and often multi-modal so that the Gaussian does not properly approximate the true distribution which in turn can lead to the divergence of the filter. To overcome this problem, we present an alternative sampling method in the following section. This sampling strategy is able to handle multiple modes in the likelihood functions used as the proposal distribution. Note that our approach does not require any significant computational overhead compared to existing mapping systems that apply scan-matching in combination with a Gaussian proposal [8].

V. Efficient Mapping with Multi-Modal Proposal Distributions

In this section, we present our alternative sampling strategy that can handle multiple modes in the distributions while at the same time keeping the efficiency of a Gaussian proposal distribution. Our approach is equivalent to computing a sum of weighted Gaussians to model the proposal but does not require the explicit computation of a sum of Gaussians. Note that an open source implementation of our mapping system using this technique is available online [18].

Our previous method [8] first applies scan-matching on a per-particle basis. It then computes a Gaussian proposal *for each sample* by evaluating poses around the pose reported by the scan-matcher. This technique yields accurate results in case of a uni-modal distribution, but encounters problems in that it focuses only on the dominant mode to which the scan-matching process converges. The left image in Figure 1 illustrates an example in which the scan-matching process converges to the dominant peak denoted as "mode 1". As a result, the Gaussian proposal samples only from this mode and at most a few particles cover "mode 2" (and only if the modes are spatially close). Even if such situations are rarely encountered in practice, we found in our experiments that they are one of the major reasons for filter divergence.

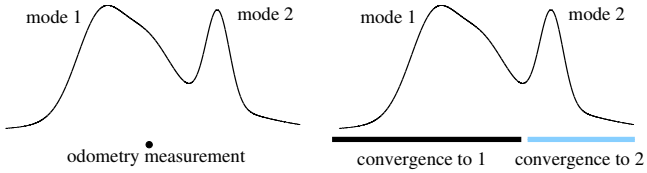

Fig. 1. The left image illustrates a 1D likelihood function and an odometry measurement. Conventional informed sampling first performs scan-matching starting from the odometry measurement. In this situation, the scan-matcher will find a local peak in the likelihood function (most likely mode 1) and the future sample will be drawn from a Gaussian centered at this single mode. The right image illustrates the new approach. It draws the sample first from the odometry model and applies scan-matching afterwards. When a drawn sample falls into the area colored black, the scan-matcher will converge to mode 1, otherwise, it will converge to mode 2. By sampling first from the odometry, then applying scan-matching, and finally computing local Gaussian approximations, multiple modes in the likelihood function are likely to be covered by the overall sample set.

One of the key ideas of our approach is to adapt the scan-matching/sampling procedure to better deal with multiple modes. It consists of a two step sampling. First, only the odometry motion model is used to propagate the samples. This technique is known from standard Monte-Carlo localization approaches (c.f. [4]) and allows the particles to cover possible movements of the robot. In a second step, gradient descent scan-matching is applied based on the observation likelihood and the denominator of Eq. (3). As a result, each sample converges to the mode in the likelihood function that is closest to its own starting position. Since the individual particles start from different locations, they are likely to cover the different modes in their corresponding likelihood functions as illustrated in the right image of Figure 1. Our approach leads to sample sets distributed according to a Gaussian *around the modes* in the observation likelihood functions. As we will demonstrate in the experimental results, this technique leads to proposal distributions which are closer to the optimal proposal given in Eq. (3) than the Gaussian approximations; when the distribution has only a single mode, the solution is equivalent to previous approaches [8].

VI. Statistical Tests

To analyze how close the Gaussian proposal as well as our new proposal are to the optimal proposal distribution, we make use of three statistical measures. First, we apply the Anderson-Darling test on normality [2]. This test is reported to be one of the most powerful tests in statistics for detecting most departures from normality. This test is

superior to the Kolmogorov-Smirnov test and has a similar performance than the Shapiro-Wilk test [16]. Second, we use the Kullback-Leibler divergence [11] to measure the distance between distributions. Third, we make use of a measure taken from the Cramér-von-Mises test [3], [21] to identify differences between distribution.

Given a set of n samples $\{y^1 < \ldots < y^n\}$ in ascending order of magnitude, the Anderson-Darling (AD) test computes the A statistic as

$$A = -n - \sum_{k=1}^{n} \frac{2k-1}{n} \left[\ln F(y^k) + \ln(1 - F(y^{n+1-k})) \right], \quad (10)$$

where F is the cumulated density function (CDF) of the distribution that is assumed to have generated the samples. In our case, F is the CDF of the normal distribution.

To determine if the samples are generated by a Gaussian or not, one needs to test if

$$A \cdot \left(1 + \frac{0.75}{n} + \frac{2.25}{n^2} \right) \quad \leq \quad c, \quad (11)$$

where c is the Anderson-Darling test value for normal distributions corresponding to a desired level of significance. For example, for a 95% confidence test of normality, the corresponding c is 0.752.

This test allows us to check if the optimal proposal is in fact a Gaussian distribution. An interesting property of the AD test is that it also provides a confidence level for its result. To apply this test, we only need to draw a sample set from the optimal proposal and compute Eq. (10) and Eq. (11). Performing this test for all proposals generated during a mapping experiment provides a measure of how often a sample set is generated from a wrong distribution.

Besides the Anderson-Darling test, we apply the Kullback-Leibler divergence (KLD) which is a frequently used technique to measure the distance between two arbitrary distributions. This allows us to also compare our proposal given in the previous section to the optimal proposal distribution. A KLD value of zero indicates that the distributions are equal and the higher the KLD, the bigger is the difference between them. The KLD between p and f is defined as

$$KLD(p, f) \quad = \quad \int p(x) \cdot \log \left(\frac{p(x)}{f(x)} \right) dx. \quad (12)$$

The KLD takes into account a quotient between two distributions. This can give a high weight to differences in the tails of the distributions (see Eq. (12), where $f(x)$ is small).

An alternative measure for comparison is used in the Cramér-von-Mises test [3], [21]. It measures the disparity of two distributions by taking into account their cumulative density functions (CDF). Since it does not use a quotient as the KLD does, it gives less weight to the tails of the distribution. It computes the integral over the squared distances between the CDFs. Let p and f be the distributions to compare and P and F the corresponding CDFs. Then,

$$d(p, f) \quad = \quad \int [P(x) - F(x)]^2 \, dP(x) \quad (13)$$

TABLE I

PROPOSAL DISTRIBUTIONS WHICH ARE REGARDED AS GAUSSIANS ACCORDING TO THE ANDERSON-DARLING TEST (95% CONFIDENCE).

Dataset	Gaussian proposal	Non-Gauss (unimodal)	Multi-modal proposal
Intel Research Lab	89.2%	7.2%	3.6%
FHW Museum	84.5%	10.4%	5.1%
Belgioioso	84.0%	10.4%	5.6%
MIT CSAIL	78.1%	15.9%	6.0%
MIT Killian Court	75.1%	19.1%	5.8%
Freiburg Bldg. 79	74.0%	19.4%	6.6%

provides a measure about the similarity of both distributions which is zero if both are equal.

The three techniques presented here are used in our experiments to identify the differences between the individual proposals and to illustrate potential weaknesses of the Gaussian proposals.

VII. EXPERIMENTS

The experiments presented in this paper are all based on real world data. We furthermore used freely available datasets to perform our analysis. The learned maps and the datasets used here are available online [17].

A. Quality of Gaussian Proposals

In the first experiment, we carried out the Anderson-Darling (AD) test with a confidence of 95% to determine if the optimal proposal can be considered as Gaussian. The results of the test are described in Table I. As can be seen, depending on the dataset, in the optimal proposal was non-Gaussian in 10% to 26% of all cases.

By visually inspecting the datasets and resulting maps, we observed two different scenarios in which non-Gaussian situations occurred. Firstly, we often observed non-Gaussian observation likelihood functions in highly cluttered environments where small changes in the position led to substantial changes of the likelihood. Multi-modal distributions are likely to occur and Gaussians are not well suited to serve as a proposal in these cases. Secondly, non-Gaussian proposals occurred when the robot was moving in environments with long corridors, a fact that surprised us. At first sight, this may appear counterintuitive since corridors are well-structured environments. However, in positions where the robot cannot observe the end of the corridor with its sensor, the likelihood along the main axis of the corridor is almost constant which is highly non-Gaussian and can lead to a negative result of the AD test. One example is MIT Killian Court, consisting mainly of long corridors. Note that even if the AD test fails in such situations, Gaussians can be still good proposals.

In addition to testing acceptance as a Gaussian distribution, we analyzed the distance between the optimal proposal and its Gaussian approximation based on the KLD and the measure from the Cramér-von-Mises test (which is referred to as CvM in the remainder of this paper). Figure 2 plots the frequencies of the individual KLD and CvM values for the Intel and FHW datasets. As can be seen, the approximation error was small (values close to zero) in 94% to 97% of

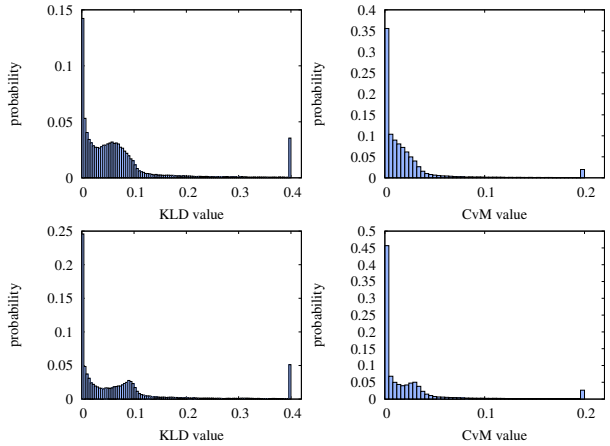

Fig. 2. Difference between the optimal proposal and the Gaussian approximation based on the Intel Research Lab (first row) and the FHW dataset (second row). The images on the left depict the frequencies of the individual Kullback-Leibler divergence values and the images on the right show the frequencies of the distance measure based on the Cramér-von-Mises test (see Eq. (13)). The right-most bin contains also all values larger or equal 0.4 (KLD) and 0.2 (CvM).

all cases. In all other cases, however, the distributions were substantially different. This fact is represented by the peak in the right-most bin of the histograms which contains all values larger or equal than 0.4 (KLD) and 0.2 (CvM). This peak corresponds to situations with multi-modal distributions which can only be badly approximated by a Gaussian. Note that similar results were obtained for the other datasets (see first row of Figure 3).

B. Multi-Modal Proposal Distribution

In the next experiment, we evaluated the alternative sampling strategy proposed in this paper. We used the KLD to compare our new proposal to the optimal proposal distribution. To actually perform the comparison, we computed all modes of the distribution explicitly, which is not required in the mapping system itself as described in Section V. To do so, we drew a set of samples and performed a gradient ascent in the likelihood function to find the individual modes. The modes were then approximated by Gaussians according to the sampled points.

The results of the comparison are shown in Figure 3 for different datasets. The plots in the first row show the KLD distance between the optimal proposal and its Gaussian approximation. The plots in the second row depict the corresponding comparison of our new proposal to the optimal one.

As can be seen, we obtained distributions that no longer approximated a significant fraction of the proposal distributions with large error (i.e., the right-most bin of the distance histograms). In contrast to this, the Gaussian approach approximates the optimal proposal inappropriately in 3% to 6% of all cases. The comparisons using the CvM value showed similar results and are omitted due to reasons of space.

Approaches using the Gaussian proposal have shown to build highly accurate maps of most datasets (compare the experiments in [8]) but there exist situations in which such

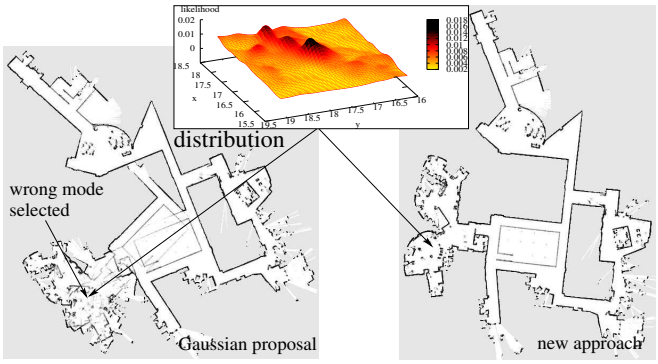

Fig. 4. Resulting map of the MIT CSAIL dataset using a Gaussian proposal (left) and our new approach (right). The Gaussian approach fails due to highly non-Gaussian likelihood functions in the cluttered room (illustrated for a given orientation θ in the top image). Trajectory length: 385m, recording time: 7 min, average speed: 0.9m/s.

TABLE II
EXECUTION TIME ON A 2.8 GHZ PC WITH A P4 SINGLE CORE CPU.

Dataset	N	Execution time		
		optimal	[8]	new method
MIT Killian Court	80	155 h	112 min	113 min
Freiburg Bldg. 79	30	84 h	62 min	62 min
Intel Research Lab	30	40 h	29 min	29 min
FHW Museum	30	38 h	27 min	27 min
Belgioioso	30	18 h	13 min	13 min
MIT CSAIL	30	10 h	7 min	7 min

techniques are likely to fail. This is especially the case if the dominant mode in the likelihood function is not the correct one. Such a situation occurs, for example, in the CSAIL dataset [17] recorded at MIT. Our expectation is that modeling multiple modes in the proposal distribution leads to more robust filters. We carried out 10 experiments with different random seeds and evaluated the success rate of the approach using the Gaussian proposal and our new method. Using the Gaussian approximation for the proposal distribution, the final map had the correct topology (all loops closed, etc.) in only 20% of trials whereas our new approach generated a correct map every time. Figure 4 shows example maps using the Gaussian proposal (left) and our new approach (right).

C. Runtime

In principle, it is possible to avoid Gaussian approximations in the proposal distribution. The main disadvantage when sampling from the optimal proposal is the high computational overhead. To illustrate this overhead, Table II shows the execution time for the individual approaches as well as the number of samples used (N). As can be seen, sampling from the optimal proposal is not suitable for practical applications since it took up to one week to correct a single dataset. In contrast to this, the computational overhead of our new approach is negligible. It allows a robot to learn an accurate map online while moving through the environment.

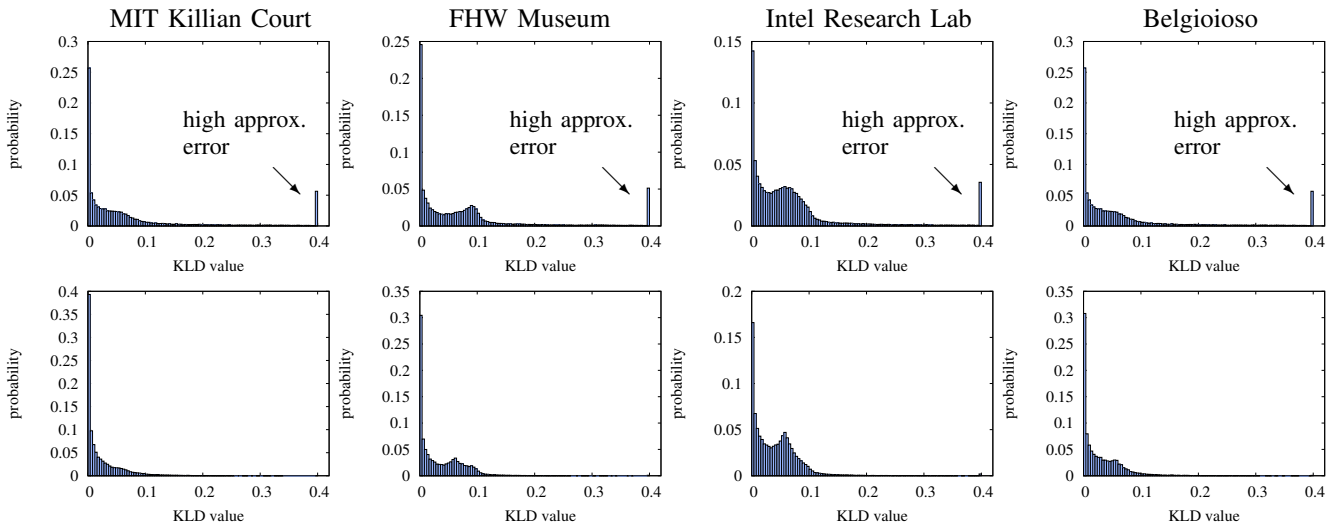

Fig. 3. The plots in the first row show the KLD between optimal proposal and its Gaussian approximation for different datasets. The plots in the second row depict the corresponding KLD between the optimal proposal and the proposal proposed in this paper. The right-most bin contains also all values larger or equal to 0.4. The right-most bin illustrates the mayor drawback of the Gaussian approximation since it described the situations in which the optimal proposal is highly non-Gaussian (e.g., multi-modal). Our new approach, however, can better deal with such situations.

VIII. Conclusion

In this paper, we analyzed how well Gaussian proposal distributions approximate the optimal proposal in the context of the application of Rao-Blackwellized particle filters to the simultaneous localization and mapping problem. We demonstrated that in around 5% of all cases, the Gaussian approximation is not sufficient to model the likelihood function. As such situations are one of the sources for the divergence of the filter, we presented an alternative sampling technique that is able to deal with multi-modal distributions while maintaining the same efficiency as the Gaussian proposal. This resulted in a more robust approach to mapping with Rao-Blackwellized particle filters. In experiments carried out with real data, we showed the efficiency and robustness of our approach.

Acknowledgment

This work has partly been supported by the DFG under contract number SFB/TR-8, by the EC under contract number FP6-IST-34120-muFly (action line: 2.5.2.: micro/nano based subsystems) and FP6-2005-IST-6-RAWSEEDS, and by the NSF under CAREER grant 0546467. Thanks to Dirk Hähnel for providing the Intel and the Belgioioso dataset as well as to Mike Bosse for the Killian Court dataset.

References

[1] E. Alaydin. *Introduction to Machine Learning*, chapter Nonparametric Density Estimation, pages 157–161. MIT Press, 2004.

[2] T. W. Anderson and D. A. Darling. Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes. *Annals of Mathematical Statistics*, 23:193–212, 1952.

[3] H. Cramér. On the composition of elementary errors. ii: statistical applications. *Skandinavisk Aktuarietidskrift*, 11:141–180, 1928.

[4] F. Dellaert, D. Fox, W. Burgard, and S. Thrun. Monte carlo localization for mobile robots. In *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, Leuven, Belgium, 1998.

[5] A. Doucet. On sequential simulation-based methods for bayesian filtering. Technical report, Signal Processung Group, Dept. of Engeneering, University of Cambridge, 1998.

[6] A. Doucet, J.F.G. de Freitas, K. Murphy, and S. Russel. Rao-Blackwellized partcile filtering for dynamic bayesian networks. In *Proc. of the Conf. on Uncertainty in Artificial Intelligence (UAI)*, 2000.

[7] A. Eliazar and R. Parr. DP-SLAM: Fast, robust simultainous localization and mapping without predetermined landmarks. In *Proc. of the Int. Conf. on Artificial Intelligence (IJCAI)*, pages 1135–1142, 2003.

[8] G. Grisetti, C. Stachniss, and W. Burgard. Improved techniques for grid mapping with rao-blackwellized particle filters. *IEEE Transactions on Robotics*, 23(1):34–46, 2007.

[9] D. Hähnel, W. Burgard, D. Fox, and S. Thrun. An efficient FastSLAM algorithm for generating maps of large-scale cyclic environments from raw laser range measurements. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pages 206–211, 2003.

[10] M. Isard and A. Blake. Contour tracking by stochastic propagation of conditional density. In *Proc. of the European Conference on Computer Vision*, pages 343–356, 1996.

[11] S. Kullback and R. A Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22:79–86, 1951.

[12] M. Montemerlo, S. Thrun D. Koller, and B. Wegbreit. FastSLAM 2.0: An improved particle filtering algorithm for simultaneous localization and mapping that provably converges. In *Proc. of the Int. Conf. on Artificial Intelligence (IJCAI)*, pages 1151–1156, 2003.

[13] M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit. FastSLAM: A factored solution to simultaneous localization and mapping. In *Proc. of the National Conference on Artificial Intelligence (AAAI)*, 2002.

[14] K. Murphy. Bayesian map learning in dynamic environments. In *Proc. of the Conf. on Neural Information Processing Systems (NIPS)*, pages 1015–1021, Denver, CO, USA, 1999.

[15] P.M. Newman. *On the structure and solution of the simultaneous localization and mapping problem*. PhD thesis, University of Sydney, Australia, 1999.

[16] S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52:591–611, 1965.

[17] C. Stachniss. Robotic datasets. http://www.informatik.uni-freiburg.de/∼stachnis/datasets, 2007.

[18] C. Stachniss and G. Grisetti. GMapping project at OpenSLAM.org. http://openslam.org, 2007.

[19] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*, chapter Robot Perception, pages 171–172. MIT Press, 2005.

[20] G.D. Tipaldi, A. Farinelli, L. Iocchi, and D. Nardi. Heterogeneous feature state estimation with rao-blackwellized particle filters. In *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, 2007.

[21] R. von Mises. *Wahrscheinlichkeitsrechnung und Ihre Anwendung in der Statistik und Theoretischen Physik*. Deuticke, Leipzig, Germany, 1931. In German.