2013 IEEE International Conference on Robotics and Automation (ICRA)
Karlsruhe, Germany, May 6-10, 2013

Switchable Constraints vs. Max-Mixture Models vs. RRR – A Comparison of Three Approaches to Robust Pose Graph SLAM

Niko Sünderhauf and Peter Protzel

***Abstract*— SLAM algorithms that can infer a correct map despite the presence of outliers have recently attracted increasing attention. In the context of SLAM, outlier constraints are typically caused by a failed place recognition due to perceptional aliasing. If not handled correctly, they can have catastrophic effects on the inferred map. Since robust robotic mapping and SLAM are among the key requirements for autonomous long-term operation, inference methods that can cope with such data association failures are a hot topic in current research. Our paper compares three very recently published approaches to *robust* pose graph SLAM, namely switchable constraints, max-mixture models and the RRR algorithm. All three methods were developed as extensions to existing factor graph-based SLAM back-ends and aim at improving the overall system's robustness to false positive loop closure constraints. Due to the novelty of the three proposed algorithms, no direct comparison has been conducted so far.**

I. Introduction

Probabilistic inference in factor graphs has become the state of the art for solving large scale SLAM problems in robotics. Current approaches like g^2o [7] or iSAM2 [5] eventually express the SLAM problem as a nonlinear least squares optimization problem and solve it using iterative techniques like Levenberg-Marquardt, Powell's Dog-Leg or Gauss-Newton. The key to efficiency is to exploit the sparse structure of the underlying problem and the sparsity of the resulting algebraic structures.

It is however well known that least squares methods are vulnerable to outliers. In the context of pose graph SLAM, outlier constraints typically are false positive loop closure detections. That means, due to perceptional aliasing, i.e. self similarity of the environment, the place recognition module in the front-end fails and erroneously declares a loop closure between two places that do not correspond in reality. This typically results in heavily distorted and unusable maps. Coping with these false positive loop closure constraints is a non-trivial problem and standard techniques like the so called *robust cost functions* (e.g. the Huber [4] function) are not sufficient.

Instead, different specialized approaches have been proposed in the very recent literature that aim at considerably increasing the robustness of the overall SLAM system against such outliers. These are in particular the switchable constraints introduced in our earlier work [13], the max-mixture models by Olson and Agarwal [9], and the RRR algorithm [8] by Latif, Cadena, and Neira.

The authors are with the Department of Electrical Engineering and Information Technology, Chemnitz University of Technology, 09111 Chemnitz, Germany. Contact: niko@etit.tu-chemnitz.de Website: http://www.tu-chemnitz.de/etit/proaut

Given the novelty of these approaches, no direct comparison has been conducted so far. The contribution of our paper therefore is to evaluate all three methods on a set of synthetic and real-world datasets. We will compare their estimation accuracy in terms of trajectory errors and precision and recall. Before we explain the conducted experiments and the results in detail, we begin by shortly introducing each of the examined algorithms.

II. Approaches to Robust Pose Graph SLAM

A. Switchable Constraints

To gain robustness against false positive loop closures, we introduced the switchable constraints approach in our earlier work [13], [12]. The main idea behind the switchable constraints is that the topology of the factor graph that represents the pose graph SLAM problem, should be partially variable and subject to the optimization instead of being kept fixed. This way, edges representing outlier constraints can be removed from the graph during the optimization. This is achieved by augmenting the original problem and introducing an additional type of hidden variable: A *switch variable* is associated with each factor that could potentially represent an outlier. This additional variable acts as a multiplicative scaling factor on the information matrix associated with that constraint. Depending on the state of the switch variable (a value between 0 and 1), the resulting information matrix is either the original matrix (when the switch is equal to 1) or $\mathbf{0}$ (when the switch is 0) or something between both ends. Notice that if the switch variable is equal to 0, the associated constraint is completely removed and has no influence on the overall solution.

Since in pose graph SLAM, every loop closure factor could be an outlier, we associate each loop closure edge with one of the newly introduced switch variables. With the switchable constraints, the optimization therefore works on an augmented problem, searching for the joint optimal configuration of the original variables and the newly introduced switch variables, hence searching the optimal graph topology.

For the experiments described later on, we used an implementation of the latest version of the switchable constraints approach as described in [13]. This implementation is available in our software package Vertigo [1] and uses the SVN-version of g^2o [7] from www.openslam.org

B. Max-Mixture Model

The max-mixture model was introduced by Olson and Agarwal [9]. The authors first considered to use Gaussian mixtures to model the likelihood of the loop closure

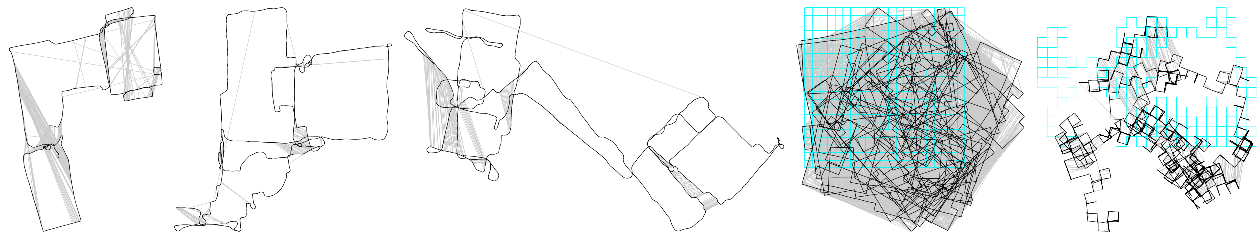

Fig. 1. Five of the datasets used in this comparison. From left to right: Bicocca, Bovisa04, Bovisa06, City10000, Manhattan. Ground truth for the synthetic datasets is overlaid in cyan.

constraints. In the conventional formulation of pose graph SLAM using factor graphs, it is assumed that all distributions are *unimodal* Gaussians. The reason why a false positive loop closure constraint is an outlier, is because this assumed Gaussian error model does not cover its true likelihood. Olson and Agarwal argue that an additional second mixture component, best modelled to be a uniform distribution over the whole world, would cover for the low, but nonzero probability that a loop closure constraint could be a false positive and therefore connect to literally anywhere. This uniform distribution can be approximated by a Gaussian with a very large covariance and the same mean as the first Gaussian component. The main insight of [9] is that a Gaussian sum mixture cannot be converted into the least squares formulation, because the sum cannot be pushed through the log that is taken to transform the Gaussian distribution into its negative log form. A max-mixture model however, can be transformed as needed and is a suitable approximation of the sum model. From the implementation point of view, a maximum likelihood selection process is applied each time a max-mixture factor is to be evaluated by the solver. The selection process picks the component that locally maximizes the likelihood of the factor's error function.

For our experiments we used the implementation from the original authors that is available online (`github.com/agpratik/max-mixture.git`, revision as of October 25th 2012). Since the SVN-version of g^{2}o does not work with the max-mixture factors, we used the GIT-version (Jan 11th, 2013) of g^{2}o that is available from `github.com/RainerKuemmerle/g2o.git`.

C. RRR

The *Realizing, Reversing, Recovering* (RRR) algorithm by Latif, Cadena, and Neira [8] aims at identifying false positive loop closures by checking the consistency between loop closures and the trajectory estimates with a number of χ^2 tests. This approach works by first clustering the loop closure constraints based on their timestamps. Afterwards the consistency of these single clusters is checked. This so called *intra cluster consistency* check determines if the whole cluster or single constraints of it shall be rejected as outliers. It is performed by optimizing a graph that contains all poses and all odometry constraints, but only the loop closures from the cluster currently under inspection. If the χ^2 error after the optimization exceeds a threshold, the whole cluster is rejected, i.e. all loop closure constraints in it are declared to be false positives. Otherwise, all of the constraints contained in the cluster are checked individually, again based on the χ^2 error. Single constraints whose residual error exceeds a threshold are rejected. This scheme of χ^2 tests is applied again in a second phase that checks the *inter cluster consistency*. In this part, the RRR algorithm tries to find sets of clusters that are mutually consistent. Again this is performed by repeatedly solving a graph that contains subsets of the loop closure links. See [8] for further details.

For our experiments, we used the original implementation of the RRR algorithm that is available from `github.com/ylatif/rrr.git`, (revision of August 22nd 2012).

TABLE I
THE DATASETS USED DURING THE EVALUATION.

Dataset	synthetic / real	Source	Poses	Loop Closures
Manhattan	synthetic	iSAM	3500	2099
City10000	synthetic	iSAM	10000	10688
Ring	synthetic	novel	434	26
RingCity	synthetic	novel	2361	901
Bicocca	real	RRR	8358	23 ...446
Bovisa-04	real	RRR	11393	197
Bovisa-06	real	RRR	10744	219

III. THE EXPERIMENTS

In order to compare all three approaches, we conducted a set of experiments on the synthetic and real-world datasets illustrated in Fig. 1 and Fig. 6. Table I summarizes their important properties. All experiments were conducted on an Intel Core i7 CPU.

We benchmark the three approaches in terms of RMSE and precision-recall, as well as the required runtimes. The RMSE (root mean squared error) is the mean deviation of the poses in the resulting map from their respective ground truth positions. Notice that only the errors in the xy-plane are considered. Although the RPE (relative pose error) would be a better error metric [2], it is unavailable since we do not have exact transformation information for the edges in the real-world datasets.

Table II summarizes the parameters used for the different algorithms. Since fine-tuning the parameters for different

TABLE II

PARAMETERS USED DURING THE EVALUATION.

Method	Parameter	Value
SC	Ξ	1
MM	mixture weight factor	0.01
MM	mixture scale factor	10e-12
RRR	odometry rate for Manhattan, Ring, RingCity	0
RRR	odometry rate otherwise	5
RRR	place recognition rate	1

datasets is not practical, it was not conducted in the following. We consider it desirable that the free parameters of any algorithm can either be referred e.g. from the known properties of an environment or dataset, or otherwise there should be default values that can be used regardless of the dataset at hand.

For the swichable constraints (SC) approach, the parameter Ξ (which is the covariance of the switch variable prior factor), was set to 1 as discussed in our earlier work e.g. [13]. For the max-mixtures (MM) approach, there are two parameters that need to be set. From the analysis presented in [10] we selected the given values for the mixture and covariance weights since they appeared to give the best results in general. For RRR the parameters for odometry and place recognition rate were set (after referring with the authors of [8]) to either the correct values for the respective dataset or (for the synthetic datasets) to those values that are supposed to give the best results.

A. The synthetic Datasets Manhattan and City10000

Two of the synthetic 2D datasets we used in our evaluation are common benchmark datasets in the community. The Manhattan dataset has been first published by [11]. The City10000 dataset shipped with the open source release of iSAM [6].

These datasets originally do not contain any false positive loop closure constraints. In order to test the robustness of all evaluated algorithms, we therefore spoiled the datasets with additional false positive loop closure edges. We followed the same procedure as described in [13], i.e. we added an increasing number of up to 1000 outliers (0, 50, 100, 250, 500, 1000) and applied four different policies of adding these outliers. These policies comprise random outliers, local outliers, consistent groups (with 10 constraints per group), and consistent local groups. See [13] for a more detailed explanation. The outliers were added using the script `generateDataset.py` that ships with our software Vertigo [1]. We applied the option `-p` that produces perfectly matching loop closure constraints, i.e. the assumed translation and rotation between the two connected poses was 0. Each pairing of policy vs. number of outliers was created and solved 30 times, resulting in a total of 720 trials per dataset for each algorithm.

B. The synthetic Datasets Ring and RingCity

We created these two new datasets in order to demonstrate a principal weakness of the max-mixture approach. Both datasets are illustrated in Fig. 6.

The trajectory of the *Ring* dataset is a simple octagonal loop, measuring 150 m in diameter. The robot drives around this octagon and closes the loop as it returns to its starting position and re-visits the first 26 poses. However, the accumulated odometry error around the loop is fairly large (almost 27 meter). As we are going to see, although all 26 detected loop closures in the dataset are correct, the max-mixtures approach is going to reject them all.

The *RingCity* dataset is built on the Ring dataset: After closing the large outer loop, the robot continues its trajectory inside the octagon, closing a number of further smaller loops as it goes along.

Both datasets are available online as part of Vertigo [1].

C. The Real-World Datasets

All three real-world datasets are available as part of the RRR release (`github.com/ylatif/rrr.git`) and have been originally recorded during the Rawseeds project [3].

The datasets Bovisa-04 and Bovisa-06 are mixed indoor/outdoor while Bicocca has been recorded indoors. All datasets provide odometry constraints and loop closure information that was created using a place recognition system based on bag of words. By changing a parameter of this BoW system, [8] created a total of 41 datasets for the Bicocca scenario. The difference between them regards only the loop closure constraints. Different settings of the BoW algorithm were used to create between 23 and 446 loop closure edges. The odometry constraints however remain unchanged. For the two Bovisa scenarios, only one dataset is available, respectively.

Although ground truth information is provided, the datasets and their respective ground truth are not properly aligned. Therefore, [8] suggests to align the estimated trajectory and the ground truth by finding the optimal transformation that minimizes the misalignment. This is implemented as part of the ATE (absolute trajectory error) error metric provided by the Rawseeds toolkit. We adopt this method in our evaluation to compute the error between the estimated and the ground truth trajectories. Notice that we call the resulting error RMSE (root mean squared error), but for the real-world datasets this is essentially the ATE calculated by the Rawseeds toolkit.

We did not add additional outliers to these real-world datasets, but rather processed them as they were. From visual inspection we can see that there are several false positives present, but it is unknown, *which* constraints are outliers. We therefore cannot conduct a presicion-recall analysis for the Bicocca and Bovisa experiments.

IV. RESULTS

The results of our experiments with the different datasets and algorithms are summarized in Table III that lists the trajectory errors (RMSE) and required runtimes. In the last column we indicate which algorithm performed best on the respective dataset, where the “best” algorithm is the one

TABLE III

EVALUATION RESULTS.

	Switchable Constraints				Max-Mixtures				RRR				Best
	RMSE [m]			Time [s]	RMSE [m]			Time [s]	RMSE [m]			Time [s]	
Dataset	median	mean	max	mean	median	mean	max	mean	median	mean	max	mean	
Manhattan	1.16	1.36	26.42	9.7	1.18	1.49	38.28	13.9	7.38	11.64	37.40	9.8	SC
City	0.063	0.063	0.063	38.8	0.058	0.251	64.18	47.7	0.94	1.60	5.11	523.3	SC
Ring	-	4.39	-	0.07	-	15.06	-	0.12	-	5.21	-	0.19	SC
RingCity	-	1.82	-	0.41	-	41.13	-	2.0	-	4.18	-	54.0	SC
Bovisa-04	-	2.39	-	1.1	-	11.81	-	1.4	-	3.01	-	5.9	SC
Bovisa-06	-	9.38	-	1.1	-	7.67	-	1.4	-	3.95	-	2.9	RRR
Bicocca	2.73	2.67	2.98	0.8	3.93	3.92	5.59	1.1	1.10	1.64	2.96	2.29	RRR

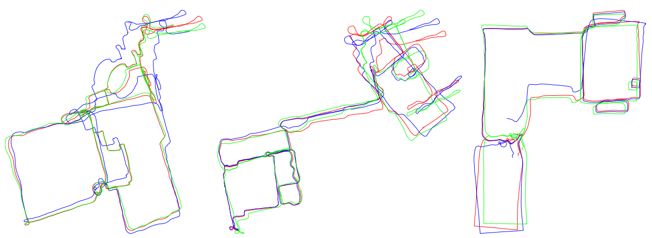

Fig. 2. Estimated trajectories for the Bovisa-04 (left), Bovisa-06 (middle), and one of the Bicocca (right) datasets. Colors indicate the used method: switchable constraints (red), max-mixtures (blue), RRR (green). The trajectories are plotted in their optimal alignment with the ground truth (not shown) according to the `get_ATE()` error function from the Rawseeds toolkit.

with the lowest mean trajectory error. We are now going to shortly discuss every dataset and comment on the algorithms' individual outcomes.

A. The Bovisa Datasets

Fig. 2 compares the estimated trajectories of both Bovisa datasets for all three algorithms. The numerical results are summarized in Table III. From there we can see that the two Bovisa datasets exhibited a different behaviour. For Bovisa-04, the switchable constraint (SC) approach produced more accurate results than RRR, while for Bovisa-06, RRR was clearly superior. The max-mixture approach (MM) was not able to produce accurate results as it disabled many of the correct loop closures.

B. The Bicocca Datasets

The Bicocca scenario consisted of 41 single datasets which differed in the number of loop closure constraints. An exemplary estimation result of all three algorithms is depicted in Fig. 2. The RMSE values over all datasets are compared in the boxplots of Fig. 3 and summarized in Table III. It is apparent that RRR has the lowest median and mean RMSE values, although the spread of the results is rather large. The switchable constraints perform worse than RRR in most cases, but still better than MM.

The two free parameters of RRR were set equally for the Bicooca and Bovisa datasets: The odometry rate was set to 5 and the place recognition rate to 1. According to the authors of [8], these are the correct values for the recorded data.

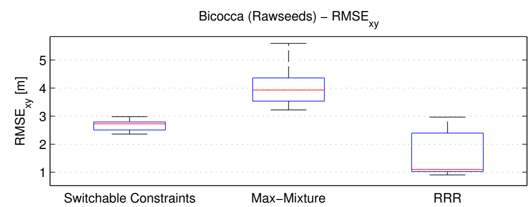

Fig. 3. Distribution of the trajectory errors (RMSE) for all 41 Bicocca datasets. RRR clearly produced the most accurate results.

C. The Manhattan Dataset

In contrast to the real-world datasets Bovisa and Bicocca, the synthetic Manhattan dataset was spoiled by additional false positive loop closure links as explained above. The detailed results for the RMSE are shown in Fig. 4.

From these plots it is apparent that the max-mixtures and the switchable constraints perform comparably well. Both achieve low RMSE errors in general. However, a few outliers of the RMSE distribution can be spotted in the boxplots for SC and MM in Fig. 4. This behaviour was already reported in [13] for the switchable constraints: A small percentage of datasets could not be solved successfully, i.e. single false positive loop closure constraints could not be removed. Here we see the same behaviour for the max-mixture approach. The total number of incorrectly handled datasets is slightly larger for MM (there are more outliers in the error boxplots), which can also be seen from the lower recall rate in Table IV. Notice that a recall of 100% means that all outliers were correctly rejected, while precision of 100% expresses that no correct loop closures were erroneously disabled.

We reported before in [13] that combining the switchable constraints with the Huber robust error function increases its robustness further. We could confirm these findings here, since mean RMSE dropped slightly from 1.36 to 1.21 for SC, however even with Huber, some outlier loop closures could not be rejected and the overall recall did not improve significantly. The same effect could be observed for MM. A combination of Huber however helped improving the precision value for both approaches, at the cost of slower convergence.

The RRR algorithm does not perform well on the Man-

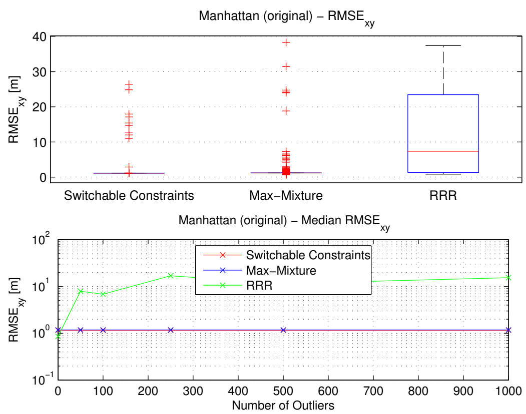

Fig. 4. Top: Boxplots showing the RMSE metric for all 720 trials of the Manhattan dataset for all three evaluated algorithms. Bottom: Comparison of the median RMSE error metrics plotted against the increasing number of false positive loop closure constraints. Notice the logarithmic scale in the bottom plot and how the quality of RRR's results quickly degrades with increasing number of outliers.

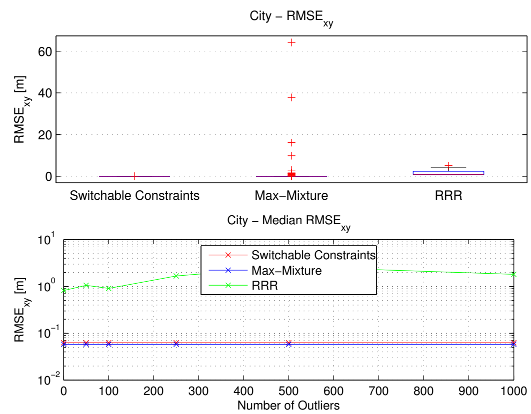

Fig. 5. Comparison of the RMSE metric for all trials of the City dataset. MM and SC have similar low median errors, but MM produces more outliers, i.e. unresolved trajectories where false positives are still present. RRR's results are an order of magnitude worse.

TABLE IV
COMPARISON OF MEAN PRECISION / RECALL [%]

Dataset	SC	MM	RRR
Manhattan	99.69 / 99.54	98.98 / 98.71	29.17 / 96.57
City10000	100 / 99.99	99.19 / 99.94	8.54 / 97.15

hattan dataset. We can see in Fig. 4 that the errors are much higher, roughly by one order of magnitude. After referring with the authors of [8], we used an odometry rate of 0 and place recognition rate 1, which was supposed to give the best results. With these settings, RRR considers each loop closure to be its own individual cluster.

D. The City1000 dataset

The outcomes for the City dataset are similar to those of the Manhattan dataset. Again we observe max-mixtures and switchable constraints to perform comparably, while RRR is not able to handle this particular dataset well. The median RMSE is compared in Fig. 5. Although the median RMSE is smallest for the max-mixture approach, its lower recall rate produced bad trajectory results for some of the 720 trials. Therefore, the smallest mean RMSE over all trials is achieved by the switchable constraints methods, which was able to successfully resolve all trials and reached a almost perfect 99.99% recall rate. Notice that it was not possible to use the supposedly best parameters for RRR (which would be place recognition rate 1, odometry rate 0), since with these settings, even after a whole week of constant processing, RRR still was not finished with the experiment. To bring the required runtime into a bearable region, we set the place recognition rate to 1 and odometry rate to 5 and only ran each pairing of outliers and policy once, which resulted in 24 trials for RRR.

E. The Ring and RingCity Datasets

These two datasets reveal a weakness in the max-mixture approach: It is prone of erroneously disabling correct loop closure constraints if the initial errors for these constraints are high. Fig. 6 illustrates this behaviour. Even for the very simple Ring dataset, MM was not able to close the loop and rejected all 26 loop closure edges. SC and RRR handled the dataset well, although SC gained lower RMSE values.

The same results can in principle be found for the more complex RingCity scenario, where MM again rejected all 901 loop closures as outliers. RRR erroneously rejected 154 of them, which lead to worse RMSE results than SC. Notice that since there are no outliers in these datasets, precision-recall statistics cannot be calculated.

F. A Discussion of Precision and Recall Statistics

There is a major conceptual difference between the switchable constraints approach and MM and RRR: While the latter two make a *binary* decision on whether a loop closure should be accepted or rejected, SC assigns a *continuous* switch value between 0 and 1. This means that in terms of precision-recall statistics, MM and RRR produce a single point for each trial. SC however produces a curve that depends on the chosen threshold for the switch values to force a binary decision.

This inherently different behaviour complicates the analysis of the algorithms in terms of precision and recall. The values given in Table IV are the mean precision and recall over all 720 trials of one dataset for MM and RRR. For SC, we first determined the point with the maximum f-score on the precision recall curve for each trial and than averaged over these 720 points.

V. DISCUSSION AND CONCLUSIONS

To conclude this paper, we would have preferred to report that one of the three evaluated algorithms performed clearly superior on all of the examined datasets. However, this is not the case.

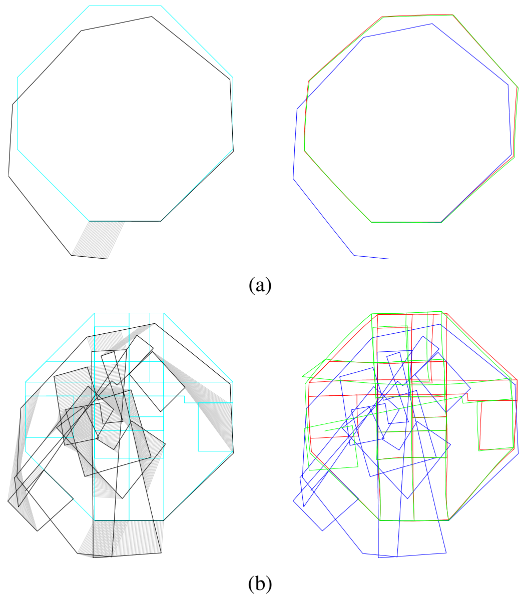

Fig. 6. Two simple new datasets, coined Ring (a) and RingCity (b), illustrate the limits of the Max-Mixture approach [9]. The left sides of (a) and (b) show the respective ground truth trajectories (cyan) for both datasets and the initial guesses from odometry (black) along with the loop closures (gray). Notice that all loop closures are correct, i.e. there are no false positives. On the right are the resulting maps produced by SC (red), MM (blue) and RRR (green). Even for the very simple Ring dataset, MM fails to close the loop. Due to the large accumulated odometry error along the loop, MM always selects the outlier component of the mixture model, despite all 26 loop closure constraints supporting each other. Only if the mixture scale factor is set to $\geq 10^{-7}$ (the default value is 10^{-12}), can MM close the loop. However, with this value it cannot cope with outliers in the Manhattan dataset anymore. RRR can process the simple Ring dataset as well as SC, but discards correct loop closures for the slightly more complex RingCity dataset. Only the switchable constraints approach (red) can cope well with both datasets.

RRR performed best on the real-world datasets Bovisa-06 and Bicocca, while SC and MM followed far behind. On the other hand, on the synthetic datasets Manhattan and City10000, SC and MM easily outperformed RRR, which was on the edge of being infeasible, especially for the large City10000 dataset Both the required runtime and RRR's ability to cope with false positive loop closure constraints were clearly inferior for these two synthetic datasets.

Compared to SC, MM has the advantage of being simpler to implement. Since there are no additional variables and factors that have to be evaluated, we can expect MM to converge quicker than SC does. This advantage is not captured or revealed by our experiments reported here, since both SC and MM always run for 50 iteration of Gauss-Newton, regardless of the residual error or convergence behaviour. This limitation is because the Gauss-Newton solver of g^2o currently does not detect convergence. However, our experiments with the new Ring and RingCity datasets revealed that the maximum likelihood selection scheme underlying MM is prone of picking the wrong mixture component when the error of the initial guess (i.e. according to odometry) is very high. We think this also explains MM's worse performance on the real-world datasets, where it declares most of the true positives to be outliers.

An advantage of SC over the other two algorithms is that there is only one free parameter Ξ and that the results of the algorithm are rather insensitive to the exact value chosen. We kept this parameter fixed to 1 for all experiments we reported in this paper and our earlier work. RRR and MM have at least two free parameters that seem to be more delicate to tune. As a side remark, it is of course always possible to reach better results than reported here if one would fine-tune the parameters of the individual algorithms for each dataset. E.g. setting $\Xi = 0.1$ results in a RMSE of 3.09 for Bovisa-06, thus SC suddenly outperforms RRR. However, we think that the most appealing algorithm is the one with the least number of free parameters or the one which is most insensitive to the exact value of the parameters

Since none of the algorithms worked perfectly for all datasets, we have to conclude that more research on robust inference in factor graphs is necessary. A deeper understanding of the parameters of SC and MM, how they influence the estimation result, and if their optimal values could be referred directly from the characteristics of a dataset (i.e. not by an exhaustive search) is a worthwhile direction of future research.

Acknowledgements

The authors want to thank Cesar Dario Cadena Lerma and Yasir Latif for their valuable help with RRR and the included Rawseeds datasets. We furthermore thank Pratik Agarwal for helpful comments with respect to the max-mixtures approach.

References

[1] Vertigo – Versatile Extensions for Robust Inference using Graph Optimization. http://www.openslam.org/vertigo.

[2] W. Burgard, C. Stachniss, G. Grisetti, B. Steder, R. Kümmerle, C. Dornhege, M. Ruhnke, A. Kleiner, and Juan D. Tardós. A comparison of SLAM algorithms based on a graph of relations. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, St. Louis, MO, USA, 2009.

[3] S. Ceriani, G. Fontana, A. Giusti, D. Marzorati, M. Matteucci, D. Migliore, D. Rizzi, D. G. Sorrenti, and P. Taddei. Rawseeds ground truth collection systems for indoor self-localization and mapping. *Autonomous Robots*, 27(4):353–371, 2009.

[4] Peter J. Huber. Robust regression: Asymptotics, conjectures and monte carlo. *The Annals of Statistics*, 1(5):799–821, 1973.

[5] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. Leonard, and F. Dellaert. iSAM2: Incremental Smoothing and Mapping Using the Bayes Tree. *Intl. Journal of Robotics Research*, 2012.

[6] M. Kaess, A. Ranganathan, and F. Dellaert. iSAM: Incremental Smoothing and Mapping. *IEEE Transactions on Robotics*, 24(6), 2008.

[7] R. Kümmerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard. g2o: A General Framework for Graph Optimization. In *Proc. of Intl. Conf. on Robotics and Automation (ICRA)*, 2011.

[8] Yasir Latif, Cesar Cadena, and José Neira. Robust loop closing over time. In *Proceedings of Robotics: Science and Systems (RSS)*, Sydney, Australia, July 2012.

[9] Edwin Olson and Pratik Agarwal. Inference on Networks of Mixtures for Robust Robot Mapping. In *Proceedings of Robotics: Science and Systems (RSS)*, Sydney, Australia, July 2012.

[10] Edwin Olson and Pratik Agarwal. Inference on Networks of Mixtures for Robust Robot Mapping. *Int. J. of Robotics Research (IJRR)*, RSS Special Issue (to appear), 2013.

[11] Edwin Olson, John Leonard, and Seth Teller. Fast iterative optimization of pose graphs with poor initial estimates. In *Inl. Conf. on Robotics and Automation, ICRA*, 2006.

[12] Niko Sünderhauf. *Robust Optimization for Simultaneous Localization and Mapping*. PhD thesis, Chemnitz University of Technology, 2012.

[13] Niko Sünderhauf and Peter Protzel. Switchable Constraints for Robust Pose Graph SLAM. In *Proc. of IEEE International Conference on Intelligent Robots and Systems (IROS)*, Vilamoura, Portugal, 2012.