2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

Multi-Robot Pose Graph Localization and Data Association from Unknown Initial Relative Poses via Expectation Maximization

Vadim Indelman*, Erik Nelson†, Nathan Michael†, and Frank Dellaert*

***Abstract*— This paper presents a novel approach for multi-robot pose graph localization and data association without requiring prior knowledge about the initial relative poses of the robots. Without a common reference frame, the robots can only share observations of interesting parts of the environment, and trying to match between observations from different robots will result in many outlier correspondences. Our approach is based on the following key observation: while each multi-robot correspondence can be used in conjunction with the local robot estimated trajectories, to calculate the transformation between the robot reference frames, only the inlier correspondences will be similar to each other. Using this concept, we develop an expectation-maximization (EM) approach to efficiently infer the robot initial relative poses and solve the multi-robot data association problem. Once this transformation between the robot reference frames is estimated with sufficient measure of confidence, we show that a similar EM formulation can be used to solve also the full multi-robot pose graph problem with unknown multi-robot data association. We evaluate the performance of the developed approach both in a statistical synthetic-environment study and in a real-data experiment, demonstrating its robustness to high percentage of outliers.**

I. Introduction

A key capability in multi-robot autonomous systems is collaborative localization and mapping in challenging, partially unknown environments. By sharing information between the robots, the performance of individuals in the group can be significantly improved, allowing for cooperatively performing complicated tasks in different domains including surveillance, search and rescue, and object manipulation. The robotics community has been addressing this important line of research over the past decade, in an effort to extend simultaneous localization and mapping (SLAM) approaches from single robot scenarios to multi-robot scenarios.

Multi-robot autonomy introduces a number of significant challenges over those found in the single robot case. In this paper we address two of these challenges*: determining the robots initial relative poses* and *reliably establishing multi-robot data association*. These issues are coupled: multi-robot data association is essential to establish constraints between poses of different robot or between mutually observed 3D points. These constraints, together with efficient optimization techniques, are essential in enabling reliable robot team performance. A reliable data association between the robots is therefore critical, as introducing false constraints, i.e. outliers, may lead to a dramatic degradation in performance.

As in the single-robot case, multi-robot approaches can be roughly divided into two main categories: Full SLAM and PoseSLAM. In the former case, the robots explicitly estimate and share 3D points between themselves, or any other parametrization of the observed environment. Multi-robot SLAM attracted an extensive amount of research, including [9], [25], [13], [14], [1], [22], [5]. Research addressing multi-robot data association without assuming known initial relative poses between the robots include [5], [22] that perform robust data association using variations of the RANSAC algorithm [10], with an emphasis on distributed performance, and [2] where inconsistent data association is identified and removed in a decentralized framework.

In this paper we focus on the second category in multi-robot localization and mapping, multi-robot PoseSLAM, which has attracted only limited attention from the research community. The PoseSLAM framework avoids explicit estimation of landmarks, removing the necessity for finding a good initialization for these variables while being computationally efficient and robust [12], [20], [8], [18]. Different from Full SLAM approaches, robot state estimation is performed based on relative pose constraints that relate between different robot poses. Multi-robot PoseSLAM approaches differ in the way these relative pose constraints are generated.

Many methods assume the robots are capable of making *direct* observations of each other's pose, as well as identifying the robot the relative pose measurement refers to. The latter trivializes multi-robot data association, however requires sophisticated classification algorithms or tagging each robot with a unique mark, with drawbacks for each alternative. Methods in this category focus on inferring the robots poses without assuming prior knowledge on the initial relative pose between the robots [14], [25], [1], [3]. Recent work [11] relaxes the aforementioned assumption and considers direct relative pose observations that do not include the identity of the measured robots.

Another approach in multi-robot PoseSLAM relies on *indirect* relative pose constraints between the robots, that can be established whenever the same scene is observed by several robots. Different from direct relative pose constraints, indirect constraints do not require rendezvous or direct line of sight between the robots. Furthermore, these constraints can be established between poses of different robots from different time instances. Existing work typically assumes the relative poses of the robots are either known [9] or can be accurately inferred [13]. The use of multiple view geometry for multi-robot localization, tightly connected with PoseSLAM, was proposed in [16], [15]. Distributed multi-robot PoseSLAM was considered in [19], [17], [24]. In [19] the multi-robot PoseSLAM problem was formulated within a single centralized optimization framework where the

*Institute for Robotics and Intelligent Machines (IRIM), Georgia Institute of Technology, Atlanta, GA 30332, USA.
†Robotics Institute, Carnegie Mellon, Pittsburgh, PA 15213, USA.

978-1-4799-3684-7/14/$31.00 ©2014 IEEE

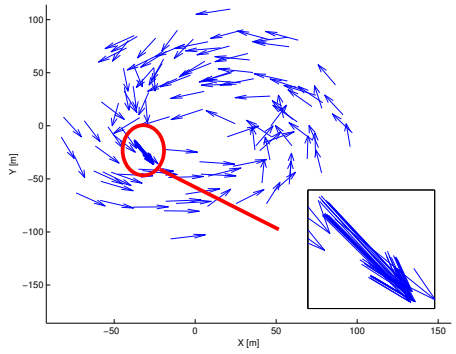


Fig. 1: Distribution of the transformation $T_{r_2}^{r_1}$ for all multi-robot candidate correspondences. Only inlier correspondences result in similar transformations.

unknown initial relative poses between the robots are treated as variables and inferred in conjunction with the robot poses. The common aspect in all these approach is the assumption of known multi-robot data association.

In this work we relax this restricting assumption and develop an approach for multi-robot PoseSLAM localization and data association from unknown initial robot relative poses. Our approach is based on the observation that by analyzing the distribution of multi-robot relative pose constraints (illustrated in Figure 1) it is possible to estimate the transformation between the robot reference frames and reject the outliers. Based on this insight we develop an expectation-maximization (EM) approach to efficiently perform this inference and show that once this transformation has been estimated with sufficient measure of confidence, it is possible to solve the multi-robot PoseSLAM problem with unknown multi-robot data association.

II. Problem Formulation

We consider a group of R robots deployed to collaboratively operate in some unknown environment and assume the robots start operating from different locations, *initially unaware* of each other. Each robot r is assumed to be capable of estimating its trajectory X^r based on observations Z^r from its onboard sensors. We represent this estimation problem in a pose graph probabilistic formulation

$$p\left(X^r|Z^r\right) \propto p\left(x_0^r\right) \prod_i p\left(u_{i-1,i}^r|x_{i-1}^r, x_i^r\right), \qquad (1)$$

where $x_i^r \in X^r$ is the robot's pose at time t_i, expressed relative to some reference frame, and $p\left(x_0^r\right)$ is a prior. Since we assume no a priori knowledge about the environment and the initial pose of the robots, the reference frame of each robot is arbitrarily set to coincide with the initial pose.

The measurement likelihood term $p\left(u_{i-1,i}^r|x_{i-1}^r, x_i^r\right)$ in Eq. (1) involves the relative pose measurement $u_{i-1,i}^r$ that can be either directly obtained from odometry measurements or calculated from vision or laser sensor observations at the two time instances t_{i-1} and t_i. We follow the standard assumption in SLAM community and model the measurement likelihood as a Gaussian:

$$p\left(u_{i-1,i}^r|x_{i-1}^r, x_i^r\right) \propto \exp\left(-\frac{1}{2}\left\|u_{i-1,i}^r \ominus h\left(x_{i-1}^r, x_i^r\right)\right\|_\Sigma^2\right), \qquad (2)$$

with Σ being the measurement noise covariance and h the measurement model that, in the case of relative pose observations and robot poses expressed in the *same* reference frame, is $h\left(x_{i-1}^r, x_i^r\right) \doteq x_{i-1}^r \ominus x_i^r$. We follow Lu and Milios [21] and use the notation $\ominus$ in $a \ominus b$ to express b locally in the frame of a for any two poses a, b.

The maximum a posteriori (MAP) estimate of the rth robot poses X^r using only local information is then given by

$$\hat{X}^r = \arg\max_{X^r} p\left(X^r|Z^r\right). \qquad (3)$$

In the multi-robot case, relative pose constraints between different robots can be established to substantially improve the estimate of each individual trajectory and allow for coordination between the robots.

We denote by $\mathcal{F}$ the set of multi-robot data association, with each individual data association $(r_1, r_2, k, l) \in \mathcal{F}$ representing a relative pose constraint $u_{k,l}^{r_1,r_2}$ relating between the pose of robot r_1 at time t_k and the pose of robot r_2 at time t_l. This constraint can represent both direct observation of one robot pose relative to another robot, and also the estimated relative pose based on observation of a common scene by two robots. In the latter case, it is computed from the measurements of the two robots $z_k^{r_1} \in Z^{r_1}$ and $z_l^{r_2} \in Z^{r_2}$, that can represent, for example, laser scans or image observations.

Assuming multi-robot data association $\mathcal{F}$ has been established and appropriate constraints $u_{k,l}^{r_1,r_2}$ have been calculated, we can write a probabilistic formulation for the multi-robot joint pdf for all the robots as follows:

$$p\left(X|Z\right) \propto \prod_r p\left(X^r|Z^r\right) \prod_{(r_1,r_2,k,l)\in\mathcal{F}} p\left(u_{k,l}^{r_1,r_2}|x_k^{r_1}, x_l^{r_2}\right), \qquad (4)$$

where X and Z represent, respectively, the trajectories and the measurements of all the robots in the group.

As the robots express their local trajectories with respect to *different* reference systems, the measurement likelihood term in Eq. (4) is

$$p\left(u_{k,l}^{r_1,r_2}|x_k^{r_1}, x_l^{r_2}\right) \propto \exp\left(-\frac{1}{2}\left\|err\left(u_{k,l}^{r_1,r_2}, x_k^{r_1}, x_l^{r_2}\right)\right\|_\Sigma^2\right), \qquad (5)$$

with

$$err\left(u_{k,l}^{r_1,r_2}, x_k^{r_1}, x_l^{r_2}\right) \doteq u_{k,l}^{r_1,r_2} \ominus h\left(x_k^{r_1}, x_l^{r_2}\right), \qquad (6)$$

and

$$h\left(x_k^{r_1}, x_l^{r_2}\right) \doteq x_k^{r_1} \ominus \left(T_{r_2}^{r_1} \oplus x_l^{r_2}\right). \qquad (7)$$

The notation $\oplus$ represents the compose operator [21], and $T_{r_2}^{r_1}$ is a transformation between the reference frames of robots r_1 and r_2. Since the robots start operating from

different unknown locations, this transformation is *initially unknown*.

While the formulation (4) assumes multi-robot data association $\mathcal{F}$ is given, in practice it is *unknown* ahead of time and should therefore be established. In this paper our goal is to reliably infer the multi-robot data association $\mathcal{F}$ in a multi-robot PoseSLAM framework, without assuming prior knowledge on initial relative poses between the robots, i.e., unknown $T_{r_2}^{r_1}$ for all pairs $r_1, r_2 \in [1, \ldots, R]$.

III. Approach

We assume each robot r shares carefully chosen N_r high-quality measurements $\{z_i^r\}_{i=1}^{N_r}$ with other robots. For any two robots r_1 and r_2, the data association problem can then be formulated as identifying the inliers among the constraints $u_{k,l}^{r_1,r_2}$ that are calculated for each $z_k^{r_1} \in \{z_i^{r_1}\}_{i=1}^{N_{r_1}}$ and $z_l^{r_2} \in \{z_i^{r_2}\}_{i=1}^{N_{r_2}}$. Instead of assuming data association to be given, we introduce a latent binary variable $j_{k,l}^{r_1,r_2}$ for each possible multi-robot data association $(r_1, r_2, k, l) \in \mathcal{F}$, and use the intuitive convention that the data association is an inlier if $j_{k,l}^{r_1,r_2} = inlier$ and accordingly outlier when $j_{k,l}^{r_1,r_2} = outlier$. We collect all such latent variables into the set $\mathcal{J}$ that becomes part of the inference.

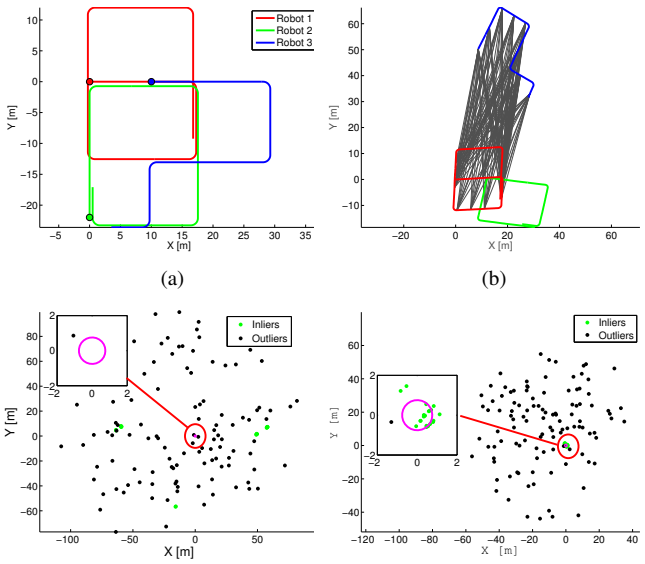
The probabilistic formulation (4) then turns into

$$p\left(X, \mathcal{J}|Z\right) \propto \prod_r p\left(X^r|Z^r\right) \prod_{(r_1,r_2,k,l)\in\mathcal{F}} p\left(j_{k,l}^{r_1,r_2}\right) p\left(u_{k,l}^{r_1,r_2}|x_k^{r_1}, x_l^{r_2}, j_{k,l}^{r_1,r_2}\right). \quad (8)$$

Since the robots are unaware of each other's locations, only a small fraction of the multi-robot data associations in $\mathcal{F}$ will be inliers. One may argue that outliers can be directly identified and rejected by matching algorithms, such as RANSAC-based fundamental matrix estimation in the case of image observations or ICP matching in the case of laser measurements. However, this argument is only partially true: while these algorithms are capable of accurate relative pose estimation given observations of a *common* scene, identifying the fact that two given observations were acquired from *different* parts of the environment is a much more challenging task. For example, ICP will often produce some relative pose (with a reasonable uncertainty covariance and number of matched points between the two scans) when fed with two laser scans from different yet similar in appearance parts of the environment (e.g. corridor, hallway). It is for this reason that the outliers ratio in the set $\mathcal{J}$ may be quite high.

What complicates this problem is the fact that the transformation $T_{r_2}^{r_1}$, for a given pair of robots $r_1, r_2 \in [1, \ldots, R]$, is unknown. Assuming some arbitrary value $T_{r_2}^{r_1}$, each candidate multi-robot data association (r_1, r_2, k, l) with a corresponding constraint $u_{k,l}^{r_1,r_2}$ will typically result in high discrepancy between $u_{k,l}^{r_1,r_2}$ and the prediction $h\left(x_k^{r_1}, x_l^{r_2}\right)$ from Eq. (7). These high errors (6) will be obtained *both for inlier and outlier correspondences*.

We illustrate this fact in a simple synthetic planar scenario of 3 robots shown in Figure 2. The ground truth robot trajectories are shown in Figure 2a, while Figure 2b shows the estimated robot trajectories, using odometry and loop closures, where the initial pose of each robot was set to some *arbitrary* value. Additionally, the figure illustrates the candidate multi-robot relative pose constraints (the set $\mathcal{F}$) between the red and the blue robot, 85% of which are outliers. The errors (6) for all these constraints, evaluated using arbitrary initial pose values and the correctly-estimated values are shown respectively in Figures 2c and 2d, along with the 1σ ellipse representing the measurement noise covariance Σ. One can observe that in the former case, the majority of the errors are high and *no inliers* can be found in the vicinity of the 1σ ellipse.



(a) (b) (c) before (d) after

Fig. 2: (a) Ground truth synthetic scenario; (b) Arbitrary robot initial relative poses and all candidate correspondences between the red and blue robots; (c)-(d) Constraints errors (6) evaluated using arbitrary and estimated initial robot poses. Also shown is 1σ uncertainty ellipse corresponding to Σ.

Therefore, attempting to directly identify inlier correspondences in the full problem (8), for example by using robust estimation techniques (e.g. [23]), will typically fail as further discussed in Section V. On the other hand, if the transformation $T_{r_2}^{r_1}$ was known, the distribution of the errors would be different, with *all the inliers* located in the vicinity of the 1σ ellipse, as illustrated in Figure 2d, making it feasible to correctly identify the corresponding data associations as inliers and rejecting the outliers.

Consequently, we propose first to infer the reference frame transformations $T_{r_2}^{r_1}$ and only when these transformations are known with significant levels of confidence we approach the full problem (8), then much easier. Therefore, in our approach the robots first establish a mutual reference frame and only then proceed to multi-robot information fusion.

IV. Inference over Robots Reference Frames Transformation

Our approach for inferring $T_{r_2}^{r_1}$ for any two robots r_1 and r_2 is based on the following **key observation**: *each* candidate multi-robot correspondence (r_1, r_2, k, l), regardless if it is inlier or outlier, suggests a solution for the transformation $T_{r_2}^{r_1}$. However, *only* the inlier correspondences will produce similar transformations, while those calculated from outlier correspondences will typically disagree amongst each other.

This concept is illustrated in Figure 1 for the synthetic scenario from Figure 2. The distribution of the calculated transformation $T_{r_2}^{r_1}$, between the green and blue robots, for each candidate multi-robot correspondence is shown (pose is represented by x-y coordinates and an angle). One can observe the *cluster* representing the correct transformation $T_{r_2}^{r_1}$, while the rest of the data points are *scattered*.

How to automatically estimate the transformation $T_{r_2}^{r_1}$ while being robust to high percentage of outliers? One alternative is to apply clustering over the distribution of the transformations. As the number of clusters is unknown ahead of time, a good candidate is the meanshift algorithm that infers the number of clusters instead of pre-determining it. Following this approach, one can identify dominant clusters and then determine the transformation from correspondences that belong to the most dominant cluster. The advantage is that a global search is performed, treating each data point, in our case the transformations for each correspondence, as a potential different cluster. A downside is the computational complexity which, for the basic version of mean shift, is $O\left(\tau n^2\right)$ with τ number of iterations and n being the number of data points, i.e. the number of candidate multi-robot correspondences.

A. Expectation-Maximization Formulation

Instead, we develop an alternative approach that is based on the expectation-maximization (EM) algorithm [7]. We assume the trajectories X^r estimated by each robot based on local observations according to Eq. (3) are of reasonable accuracy and denote all these trajectories by

$$\hat{X}^{SR} \doteq \left\{\hat{X}^r\right\}_{r=1}^{R},$$

where the superscript "SR" stands for "single robot". Considering these trajectories as given the MAP estimate of the transformation $T_{r_2}^{r_1}$ can be written as

$$\hat{T}_{r_2}^{r_1} = \arg\max_{T_{r_2}^{r_1}} p\left(T_{r_2}^{r_1}|\hat{X}^{SR}, Z\right) = \arg\max_{T_{r_2}^{r_1}} \sum_{\mathcal{J}} p\left(T_{r_2}^{r_1}, \mathcal{J}|\hat{X}^{SR}, Z\right), \quad (9)$$

where the summation refers to all the possible values for *each* of the latent binary variables $j_{k,l}^{r_1,r_2}$ in $\mathcal{J}$.

Following the EM approach, we re-write the above as

$$\hat{T}_{r_2}^{r_1} = \arg\max_{T_{r_2}^{r_1}} p\left(\mathcal{J}|\hat{T}_{r_2}^{r_1}, \hat{X}^{SR}, Z\right) \log p\left(T_{r_2}^{r_1}, \mathcal{J}|\hat{X}^{SR}, Z\right), \quad (10)$$

alternating between the E step, that infers the data association given current estimate $\hat{T}_{r_2}^{r_1}$ (and also given $\hat{X}^{SR}, Z$ that remain fixed), and the M steps where the transformation $T_{r_2}^{r_1}$ is re-estimated using updated multi-robot data association. Eq. (10) represents a single EM iteration (iteration number is not shown explicitly).

Recalling Eq. (8) and applying Bayes rule we can write:

$$\log p\left(T_{r_2}^{r_1}, \mathcal{J}|\hat{X}^{SR}, Z\right) = \log p\left(T_{r_2}^{r_1}|\hat{X}^{SR}\right) + \sum_{(r_1,r_2,k,l)} \log p\left(j_{k,l}^{r_1,r_2}\right) p\left(u_{k,l}^{r_1,r_2}|\hat{x}_k^{r_1}, \hat{x}_l^{r_2}, j_{k,l}^{r_1,r_2}, T_{r_2}^{r_1}\right).$$

Since $p\left(J|\hat{T}_{r_2}^{r_1}, \hat{X}^{SR}, Z\right) p\left(u_{k,l}^{r_1,r_2}, j_{k,l}^{r_1,r_2}|\hat{X}^{SR}\right) \equiv p\left(j_{k,l}^{r_1,r_2}|\hat{T}_{r_2}^{r_1}, \hat{X}^{SR}, Z\right) p\left(u_{k,l}^{r_1,r_2}, j_{k,l}^{r_1,r_2}|\hat{X}^{SR}\right)$, and assuming an uninformative prior $p\left(T_{r_2}^{r_1}|\hat{X}^{SR}\right)$, we can re-write Equation (10) as

$$\hat{T}_{r_2}^{r_1} = \arg\max_{T_{r_2}^{r_1}} \sum_{(r_1,r_2,k,l)} p\left(j_{k,l}^{r_1,r_2}|\hat{T}_{r_2}^{r_1}, \hat{X}^{SR}, Z\right) \log p\left(j_{k,l}^{r_1,r_2}\right) p\left(u_{k,l}^{r_1,r_2}|\hat{x}_k^{r_1}, \hat{x}_l^{r_2}, j_{k,l}^{r_1,r_2}, T_{r_2}^{r_1}\right).$$

Further, as the latent variable $j_{k,l}^{r_1,r_2}$ is binary, there are only two cases to consider (inlier or outlier). Defining the set M of these possible cases as $M \doteq \{inlier, outlier\}$ we get:

$$\hat{T}_{r_2}^{r_1} = \arg\max_{T_{r_2}^{r_1}} \sum_{(r_1,r_2,k,l)} \sum_{m \in M} p\left(j_{k,l}^{r_1,r_2} = m|\hat{T}_{r_2}^{r_1}, \hat{X}^{SR}, Z\right) \log p\left(j_{k,l}^{r_1,r_2} = m\right) p\left(u_{k,l}^{r_1,r_2}|\hat{x}_k^{r_1}, \hat{x}_l^{r_2}, j_{k,l}^{r_1,r_2} = m, T_{r_2}^{r_1}\right). \quad (11)$$

Calculation of the weight $p\left(j_{k,l}^{r_1,r_2} = m|\hat{T}_{r_2}^{r_1}, \hat{X}^{SR}, Z\right)$ proceeds by applying the Bayes rule, followed by normalization of the weights for the latent variable $j_{k,l}^{r_1,r_2}$ to sum to 1.

Starting from an initial value for the transformation $T_{r_2}^{r_1}$, the nonlinear optimization (11) is guaranteed to converge to a *local* maximum of (9). However, choosing an initial guess far away from the true solution will lead to a local minimum, especially in the presence of many outliers. In the next section we discuss a simple method for addressing this problem, and suggest a measure to quantify the confidence in the estimated transformation $T_{r_2}^{r_1}$. The latter can be used to decide whether to accept the estimate $\hat{T}_{r_2}^{r_1}$ and proceed to full multi-robot localization we discuss in Section V.

B. Initial Guess and Measure of Confidence

We propose a simple approach for identifying several promising candidates for good initial guesses of the transformation $T_{r_2}^{r_1}$. Recalling the key observation from Section IV, we look at the distribution of the transformations, calculated for each candidate multi-robot data association (r_1, r_2, k, l), and identify dominant values for each element in $T_{r_2}^{r_1}$ separately (i.e., x axis, y-axis etc.).

This basic clustering results in a small set of initial values for the transformation $T_{r_2}^{r_1}$. We then perform the optimization

(11) for each such initial guess of $T_{r_2}^{r_1}$, typically leading to different estimations of $T_{r_2}^{r_1}$, one for each initial guess. We merge nearby initial guesses, therefore guaranteeing all initial guesses substantially differ from each other.

Now the question is - which estimate of $T_{r_2}^{r_1}$ to choose? Selecting the solution that minimizes the cost in Eq. (11) is not a good approach as the cost is expected to be lower for solutions with only a few (or none at all) identified inliers. Instead, we examine for each solution, how many inliers were identified, i.e. how many multi-robot correspondences (r_1, r_2, k, l) are with high inlier probability $p\left(j_{k,l}^{r_1,r_2} = inlier\right)$, and choose the solution with the largest number of identified inliers. The latter can be used as measure of confidence, considering the transformation between the robots as established once the number of identified inliers is above a threshold.

C. Scalability to More than Two Robots

The proposed approach can be trivially generalized to any number of robots. To this end, the robot poses are expressed in an arbitrary common reference frame. The objective then becomes inferring the transformations between the local frame of each robot r and that reference frame.

Without loss of generality, if we set the reference frame to the origin of (some) robot r_0, then the transformations to be estimated $\left\{T_{r_i}^{r_0}\right\} \doteq \left\{T_{r_2}^{r_0}, \ldots, T_{r_R}^{r_0}\right\}$ are given by a slight modification of the EM formulation (11):

$$\left\{\hat{T}_{r_i}^{r_0}\right\} = \arg\max_{T_{r_i}^{r_0}} \sum_{(r_1,r_2,k,l)} p\left(j_{k,l}^{r_1,r_2} | T_{r_1}^{r_0}, T_{r_2}^{r_0}, \hat{X}^{SR}, Z\right) \log p\left(j_{k,l}^{r_1,r_2}\right) p\left(u_{k,l}^{r_1,r_2} | \hat{x}_k^{r_1}, \hat{x}_l^{r_2}, j_{k,l}^{r_1,r_2}, T_{r_1}^{r_0}, T_{r_2}^{r_0}\right). \quad (12)$$

The measurement likelihood is accordingly changed into

$$p\left(u_{k,l}^{r_1,r_2} | \hat{x}_k^{r_1}, \hat{x}_l^{r_2}, j_{k,l}^{r_1,r_2}, T_{r_1}^{r_0}, T_{r_2}^{r_0}\right) \propto \exp\left(-\frac{1}{2}\left\|u_{k,l}^{r_1,r_2} \ominus \left(T_{r_1}^{r_0} x_k^{r_1} \ominus T_{r_2}^{r_0} x_l^{r_2}\right)\right\|_{\Sigma}^2\right). \quad (13)$$

As observed from Eqs. (12)-(13), the number of variables in the optimization (12) scales *linearly* with the number of robots in the group, while all multi-robot correspondences in $\mathcal{F}$ can be accommodated.

V. Multi-Robot Localization via Expectation-Maximization

After establishing the transformation between the robot's reference frames, it is possible to perform multi-robot localization, expressing the robots trajectories in the *same* reference frame. Since the approach discussed in Section IV depends on the quality of the local trajectories X^r, the estimated transformations are not exact. Moreover, although one could use the identified multi-robot inlier correspondences, multi-robot data association is still required for any new incoming observations from different robots.

We therefore continue considering a probabilistic formulation in which the data association is represented by latent variables, as given by Eq. (8). At this point, since the transformation between the robots is approximately known, the errors for inlier multi-robot correspondences in $\mathcal{F}$ will typically be small while the outlier correspondences will produce high errors (see, e.g., Figure 2d). One alternative therefore would be to use robust graph optimization approaches (e.g. [23]).

However, instead we propose an EM framework to efficiently infer the robots trajectories X as $\hat{X} = \arg\max_X p\left(\mathcal{J}|\hat{X}, Z\right) \log p\left(X, \mathcal{J}|Z\right)$. Performing a derivation similar to the one presented in Section IV-A, we get

$$\hat{X} = \arg\max_X \sum_r \log p\left(X^r|Z^r\right) - \sum_{(r_1,r_2,k,l)} \sum_{m \in M} p\left(j_{k,l}^{r_1,r_2} = m | \hat{X}, Z\right) \cdot \log p\left(j_{k,l}^{r_1,r_2} = m\right) p\left(u_{k,l}^{r_1,r_2} | x_k^{r_1}, x_l^{r_2}, j_{k,l}^{r_1,r_2} = m\right), \quad (14)$$

where, as earlier, $M = \{inlier, outlier\}$, and the measurement likelihood is given by Eqs. (5) and (7). If desired, the multi-robot correspondences identified as inliers in Section IV can be initialized with a high prior $p\left(j_{k,l}^{r_1,r_2} = inlier\right)$.

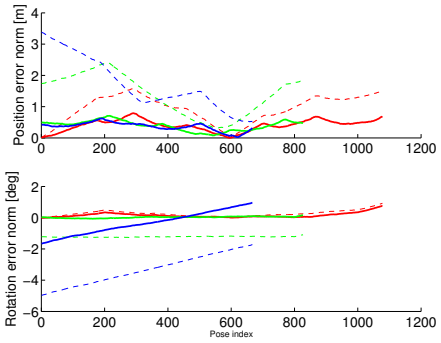
Remark: One could be tempted to directly apply the EM optimization (14), even before establishing the transformations between the robot reference frames. However, since this transformation is unknown and is set to an arbitrary value, the errors for any candidate multi-robot constraint will be high, regardless if it is inlier or outlier (see Section III). Since the inlier distributions is (by definition) narrower than the outlier distribution, i.e. $\Sigma_{inlier} \ll \Sigma_{outlier}$, the outlier distribution will *always* get the higher probability. As a consequence all the multi-robot constraints will be considered as outliers and therefore will be rejected.

VI. Results

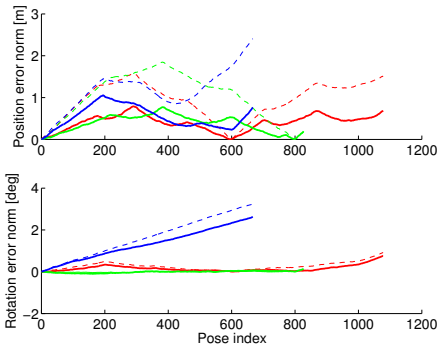
The developed approach was implemented within the GTSAM optimization library [6], and evaluated both in simulated environment and in an experiment with real sensors.

A. Simulation

We evaluate the robustness of the approach to outliers in a statistical study of 3 robots starting out with unknown relative position. The ground truth robot trajectories are given in Figure 2a, with circles denoting the true starting locations of the robots. The prior on the initial position was set to random values in each run, drawn from a zero-mean Gaussian distribution with 100 meters standard deviation (std). The robot local trajectories were calculated by integrating simulated odometry observations with 1 cm std on position and 0.01 degrees on rotation. Multi-robot relative pose constraints were calculated between the robot poses every 5 meters. A typical scenario with the calculated multi-robot constraints and some position priors is shown in Figure 2b. Some of the constraints were intentionally set to be outliers, according to the examined outlier ratio (10%, 40% and 90%).



(a)

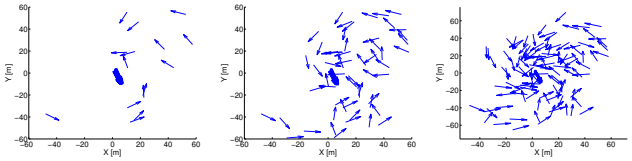


(b)

Fig. 3: (a) Global and (b) relative position and orientation estimation errors. The three robots are shown in different colors (red, green, blue) with dashed lines denoting performance after establishing transformation between reference frames but before the multi-robot optimization. Solid lines represent performance after multi-robot optimization.

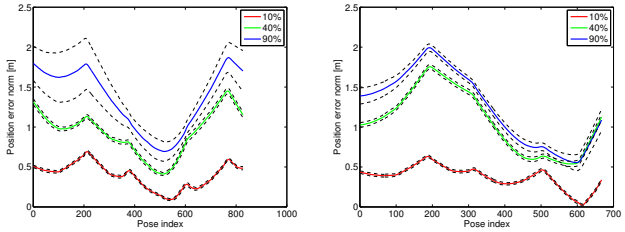
Figure 3 shows the position and orientation errors in a typical run (using 10% outliers). For each robot (red, blue, green), errors are evaluated using estimates after the reference frame transformation $T_{r_2}^{r_1}$ calculation (11) (dashed line), and after the complete multi-robot optimization (14). Both global and relative errors are shown, with the relative pose error calculated relative to the previous pose and compared to the appropriate ground truth relative pose. One can observe the global errors that correspond to the calculated transformation $T_{r_2}^{r_1}$ accuracy (dashed line) are within 3-4 meters, although the latter was set to initial values of 100m (1σ std). These level of errors are further reduced by the overall EM optimization (14), shown by solid line, both in the global and relative frame. The latter corresponds to improvement of robot's local trajectories via multi-robot localization.

First, observe that as the number of outlier increases (from 10% to 90%), the distribution of the transformation $T_{r_2}^{r_1}$, calculated for each multi-robot candidate correspondence, becomes more sporadic. This is shown in Figure 4. The inliers are still clustered together, however, as expected, it becomes increasingly difficult to identify this cluster as the outlier ration increases. Nevertheless, in our experiments we



(a) 10% outliers (b) 40% outliers (c) 90% outliers

Fig. 4: Distribution of the reference frame transformation $T_{r_2}^{r_1}$, calculated using each multi-robot candidate correspondence, for different outlier ratios (see Section IV).



(a) Robot 2 (green color in Figure 2a).(b) Robot 3 (blue color in Figure 2a).

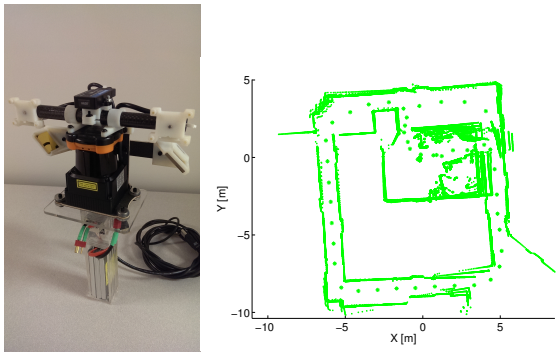
Fig. 5: Position estimation errors (norm) in statistical study of 50 runs. Mean errors are shown in bold, standard deviation about the mean in dashed line.

have observed the method to be resilient to this effect and were able to correctly identify the inliers even with high outlier ratio. As summarized in Table I below, reference frame transformation $T_{r_2}^{r_1}$ using the EM optimization (11) results in all the inliers correctly identified for 10% outlier ratio, a number that drops to 70% for 90% outlier ratio. The final optimization, initialized by the estimated transformation (Section V), further improves these numbers. Importantly, while not always all inliers were identified as such, the algorithm does not produce any false positives, i.e. outliers are *never* identified as inliers.

Outlier ratio	10 %	40 %	90%
Correct inliers est. in EM (11)	100%	75%	70%
False positives in EM (11)	0%	0%	0%
Correct inliers in final EM (14)	100%	100%	80%
False positives in final EM (14)	0%	0%	0%

TABLE I: Percentage of correctly identified inliers in estimation of reference frame transformation $T_{r_2}^{r_1}$ and in the final multi-robot optimization (14).

Statistics for position estimation errors in the 50-run Monte-Carlo study for each outlier ratio level is shown in Figure 5 for two of the three robots. As seen, performance deteriorates with higher number of outliers. However, given the above explanation this deterioration is not due to the outliers but instead because of the smaller number of the actual multi-robot inlier constraints, which determines to what extend the robots local trajectories can be improved.



(a) (b)

Fig. 6: (a) CMU sensor suite; (b) Trajectory and map generated by one of the three robots.

B. Real-Data Experiments

We have evaluated the method in an experiment using laser data collected by the CMU sensor suite (Figure 6a), which also includes additional sensors. Three recording sessions were carried out, corresponding to three robots performing different trajectories. Figure 6b shows the trajectory and the map generated by one of the robots.

While the robots did start operating from the same location, this knowledge was assumed to be absent and the robot initial positions were initialized to arbitrary values, as shown in Figure 7. We generated the set of multi-robot correspondences $\mathcal{F}$, also shown in Figure 7, by performing ICP match between laser scans of different robots.

A basic methodology was used to choose the laser scans participating in this process: these were chosen by skipping 50 sequential scans and maintaining at least 1.5 meters with the preceding chosen scan. More sophisticated approaches can be used here, for example selecting only the informative scans. Performing ICP matching of two laser scans from different robots, we required a reasonable covariance and number of matched points above a threshold, to consider the match as a multi-robot candidate constraint and only then included it in $\mathcal{F}$. The set of all multi-robot candidate constraints, the majority of which are outliers, is shown in Figure 7 for the three robots. We use Censi's implementation [4] for both ICP matching and covariance calculation.

Figure 8a shows the robot trajectories after estimating the reference frame transformations $T_{r_2}^{r_1}$ and using these values to express the local robot trajectories within a common reference frame. The local robot trajectories were obtained by performing ICP matching between consecutive laser scans and manually-identified rare loop closures to maintain reasonable quality of the trajectories. From arbitrary initial relative poses (Figure 7), including a large rotation for the blue robot, and unknown multi-robot data association, this optimization step recovered the reference frame transformations with good accuracy. Observe the initial robot positions are close to each other (in reality the robots started from the same position).

The final EM optimization (14) yields the result shown in Figure 8b, where the multi-robot constraints identified as inliers are shown in black, and the rest of the multi-robot constraints shown in dashed cyan. Note that all the identified multi-robot constraints relate between adjacent poses of different robots. Also note that no inlier constraints between the green robot and the other robots were identified. The reason for this is that green robot was *traveling in an opposite direction* to the other two robots and therefore the laser scans represent different parts of environment although the robot trajectories are similar in practice.

VII. Conclusions

In this paper we addressed the problem of multi-robot localization and mapping when operating in unknown or uncertain environments, considering both the robot initial relative poses and multi-robot data association to be unknown. Our approach was based on the key observation that while each such multi-robot constraint can be used in conjunction with local robot estimated trajectories to calculate the transformation between the robot reference frames, only the inlier correspondences will yield a similar transformation. We developed an expectation-maximization (EM) approach to infer the robot initial relative poses *without* assuming multi-robot data association is given. Having estimated these transformations, we approached the full multi-robot localization problem with unknown data association that uses a similar EM formulation. We presented an evaluation of this approach using a multi-robot statistical study in a simulated environment, and in a real-data experiment. The results of this study demonstrated the method is resilient to high percentage of multi-robot outliers, correctly inferring most of the inlier correspondences. Both the experiment and statistical-study results showed the method correctly infers the robot initial relative poses even despite setting the starting positions of the robots to arbitrary values. Future work will focus on extending the developed approach to distributed and incremental framework and on further evaluation in more complicated scenarios involving larger robot groups.

References

[1] L. Andersson and J. Nygards. C-SAM : Multi-robot SLAM using square root information smoothing. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2008.

[2] R. Aragues, E. Montijano, and C. Sagues. Consistent data association in multi-robot systems with limited communications. In *Robotics: Science and Systems (RSS)*, Zaragoza, Spain, June 2010.

[3] L. Carlone, M. Kaouk Ng, J. Du, B. Bona, and M. Indri. Rao-Blackwellized particle filters multi robot SLAM with unknown initial correspondences and limited communication. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 243–249, 2010.

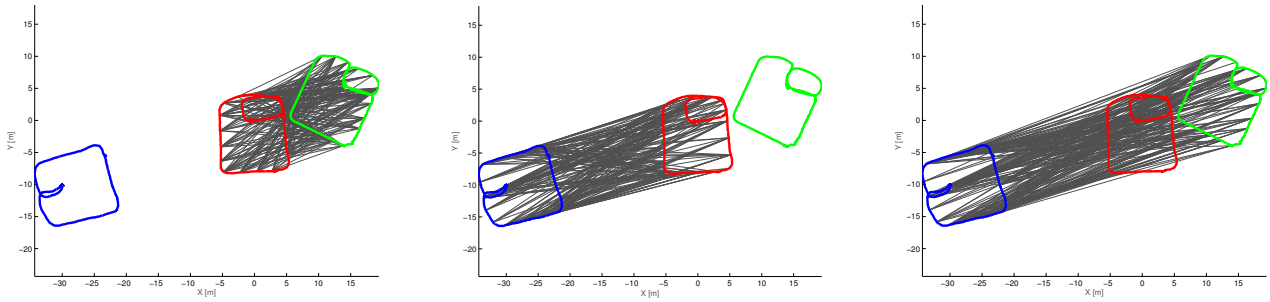
[4] Andrea Censi. An accurate closed-form estimate of icp's covariance. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 3167–3172. IEEE, 2007.

[5] A. Cunningham, K. Wurm, W. Burgard, and F. Dellaert. Fully distributed scalable smoothing and mapping with robust multi-robot data association. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, St. Paul, MN, 2012.

[6] Frank Dellaert. Factor graphs and GTSAM: A hands-on introduction. Technical Report GT-RIM-CP&R-2012-002, Georgia Institute of Technology, 2012.

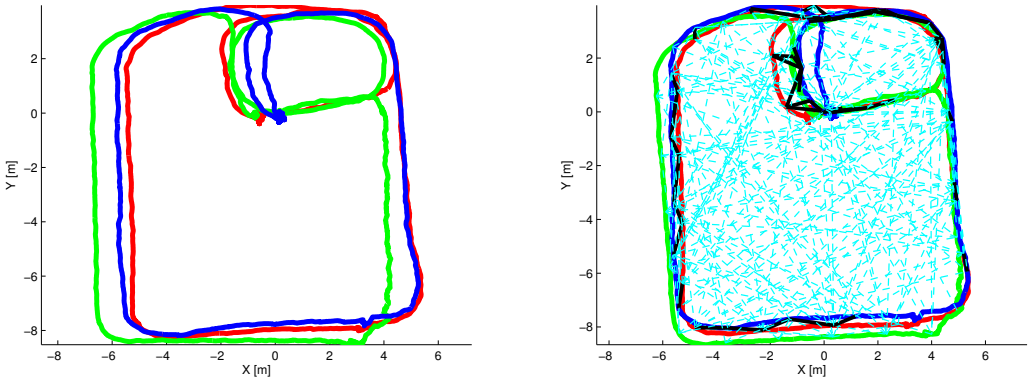
[7] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39(1):1–38, 1977.

[8] R.M. Eustice, H. Singh, and J.J. Leonard. Exactly sparse delayed-state filters for view-based SLAM. *IEEE Trans. Robotics*, 22(6):1100–1114, Dec 2006.



(a) (b) (c)

Fig. 7: Individual robot trajectories in the experiment optimized based on consecutive ICP scan matching and manual loop closure constraints. Arbitrary global position of the trajectories. Also shown are all the candidate multi-robot factors, vast majority of which are outliers: (a) constraints between the red and green, (b) red and blue, and (c) green and blue robots.



(a) (b)

Fig. 8: Robot trajectories (a) expressing local trajectories in a common reference frame given by the estimated transformation $T_{r_2}^{r_1}$; and (b) after multi-robot optimization (14), with the multi-robot constraints identified as inliers shown in black, and the rest represented by cyan dashed lines.

[9] J. Fenwick, P. Newman, and J. Leonard. Cooperative concurrent mapping and localization. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, volume 2, pages 1810–1817, 2002.

[10] M. Fischler and R. Bolles. Random sample consensus: a paradigm for model fitting with application to image analysis and automated cartography. *Commun. ACM*, 24:381–395, 1981.

[11] A. Franchi, G. Oriolo, and P. Stegagno. Mutual localization in multi-robot systems using anonymous relative measurements. *Intl. J. of Robotics Research*, 32:81–95, 2013.

[12] J.-S. Gutmann and K. Konolige. Incremental mapping of large cyclic environments. In *IEEE Intl. Symp. on Computational Intelligence in Robotics and Automation (CIRA)*, pages 318–325, 1999.

[13] A. Howard. Multi-robot simultaneous localization and mapping using particle filters. *Intl. J. of Robotics Research*, 25(12):1243–1256, 2006.

[14] A. Howard, G. S. Sukhatme, and M. J. Matarić. Multi-robot mapping using manifold representations. *Proceedings of the IEEE - Special Issue on Multi-robot Systems*, 94(9):1360 – 1369, Jul 2006.

[15] V. Indelman. *Navigation Performance Enhancement Using Online Mosaicking*. PhD thesis, Technion - Israel Institute of Technology, 2011.

[16] V. Indelman, P. Gurfil, E. Rivlin, and H. Rotstein. Distributed vision-aided cooperative localization and navigation based on three-view geometry. *Robotics and Autonomous Systems*, 60(6):822–840, June 2012.

[17] V. Indelman, P. Gurfil, E. Rivlin, and H. Rotstein. Graph-based distributed cooperative navigation for a general multi-robot measurement model. *Intl. J. of Robotics Research*, 31(9), August 2012.

[18] M. Kaess, A. Ranganathan, and F. Dellaert. iSAM: Incremental smoothing and mapping. *IEEE Trans. Robotics*, 24(6):1365–1378, Dec 2008.

[19] B. Kim, M. Kaess, L. Fletcher, J. Leonard, A. Bachrach, N. Roy, and S. Teller. Multiple relative pose graphs for robust cooperative mapping. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 3185–3192, Anchorage, Alaska, May 2010.

[20] K. Konolige. Large-scale map-making. In *Proc.* 21^{th} *AAAI National Conference on AI*, San Jose, CA, 2004.

[21] F. Lu and E. Milios. Globally consistent range scan alignment for environment mapping. *Autonomous Robots*, pages 333–349, Apr 1997.

[22] Eduardo Montijano, Sonia Martinez, and Carlos Sagues. Distributed robust data fusion based on dynamic voting. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 5893–5898. IEEE, 2011.

[23] N. Sunderhauf and P. Protzel. Towards a robust back-end for pose graph slam. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 1254–1261. IEEE, 2012.

[24] J. M. Walls and R. M. Eustice. An exact decentralized cooperative navigation algorithm for acoustically networked underwater vehicles with robustness to faulty communication: Theory and experiment. 2013.

[25] X. Zhou and S.I. Roumeliotis. Multi-robot SLAM with unknown initial correspondence: The robot rendezvous case. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, pages 1785–1792. IEEE, 2006.