Certifiably Optimal Mutual Localization with Anonymous Bearing Measurements

Yingjian Wang[1,2], Xiangyong Wen[1,2], Longji Yin[2], Chao Xu[1,2], Yanjun Cao[2], Fei Gao[1,2]

***Abstract*— Mutual localization is essential for coordination and cooperation in multi-robot systems. Previous works have tackled this problem by assuming available correspondences between measurements and received odometry estimations. However, the correspondence is difficult to acquire, especially for unified robotteams. In this paper, we present a certifiably optimal algorithm that uses only anonymous bearing measurements to formulate a novel mixed-integer quadratically constrained quadratic problem (MIQCQP). Then, we relax the original nonconvex problem into a semidefinite programming (SDP) problem and obtain a certifiably global optimum using off-the-shelf solvers. As a result, if we obtain sufficient independent bearing measurements, our method can determine bearing-pose correspondences and furthermore recover initial relative poses between robots with optimality guarantee. We compare our method with local optimization methods on extensive simulations under different noise levels to show our advantage in global optimality and robustness. Real-world experiments are conducted to show the practicality and robustness.**

I. Introduction

Recently, due to the inherent advantage, multi-robot systems have received increasing attention in many applications, such as formation [1], exploration [2] and package delivery [3]. To execute each subtask correctly and complete the full task collaboratively, robots in a team are expected to be located in a common reference frame. However, this requirement is not satisfied in wild environments like underground caves where global coordinate systems are not available. Launching robots in a predetermined relative pose is another solution. However, it is obviously time-consuming and prone to failure in large-scale environments.

To bridge this gap, self-localization using onboard sensors and relative pose recovery are irreplaceable in multi-robot systems. There are majorly two ways to estimate the initial relative transformations between robots in a team. They are map-based localization which relies on exchanging environment features, and mutual localization which depends on robot-to-robot measurements. Most research [4]–[6] focuses on the map-based relative pose recovery method, which can be easily adapted from loop-closing modules of existing simultaneous localization and mapping (SLAM) systems [7]. However, it requires robots to observe the same scene and send observed environment information to others, leading to degeneration in the environments with many similar or texture-less scenes.

[1]State Key Laboratory of Industrial Control Technology, Institute of Cyber-Systems and Control, Zhejiang University, Hangzhou, 310027, China.
[2]Huzhou Institute of Zhejiang University, Huzhou, 313000, China.
E-mails:{yj_wang, fgaoaa}@zju.edu.cn.

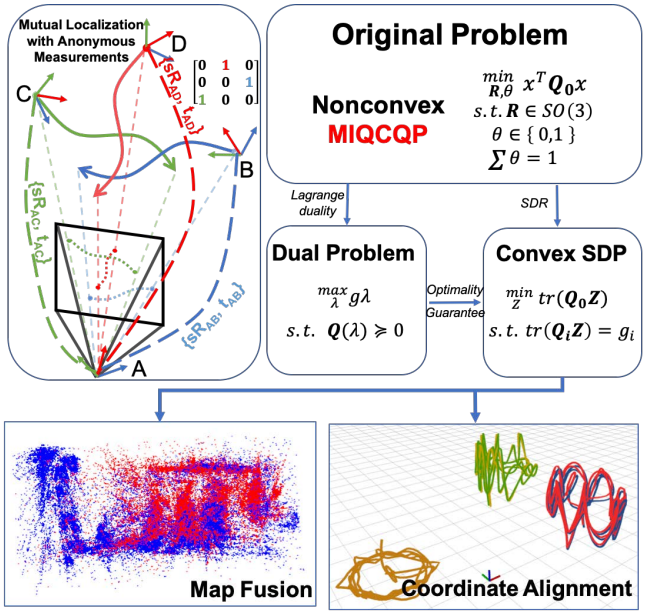

Fig. 1. Considering one observer robot A and multiple unified observed robots (B, C and D). Given robots' trajectories (three full lines) and anonymous bearing measurements (three dotted lines), our proposed method can obtain not only correct bearing-pose estimation (permutation matrix) but also accurate relative pose (sR and t). Our certifiably optimal solution can be used for map fusion and coordinate alignment in multi-robot tasks.

Our study focuses on mutual localization using bearing measurements, which only utilize detected robots' 2D coordinates in the observer's image and observed robots' estimated odometry. Compared to map-based localization, it is less influenced by environments and needs less bandwidth. Despite its appeals, as we do not rely on any specialized devices, like visual tags or external sensors, data association between the visual detection and robot identifications in a team of unified robots is challenging. Thus, in our paper, we aim to recover the data association and relative poses jointly. We firstly introduce binary variables representing the data association and mix them with multiple SO(3) variables representing the relative poses, formulating a mixed-integer problem. Furthermore, we rewrite it as a non-convex MIQCQP problem and employ tight convex relaxation to obtain a SDP problem. Then we provide a sufficient condition, under which we can obtain a *certifiably globally optimal solution* to avoid local minimum of the origin non-convex problem in noise-free cases. The optimized relative poses with optimality-guarantee can be used for map fusion and coordinate alignment, as shown in Fig.1. Extensive experiments on synthetic real-world datasets show the robustness of our method under different levels of noise.

Our contributions in this paper are:

1) We provide an innovative formulation which jointly solves data association and relative poses in a MIQCQP problem. To the best of our knowledge, there is no such work in mutual localization.
2) We propose an algorithm for the non-convex MIQCQP problem, which adopts semidefinite relaxation (SDR) to make it convex. Furthermore, we provide a condition to guarantee the tightness of the relaxation.
3) We conduct sufficient simulation and real-world experiments to validate the practicality and robustness of our proposed method.
4) We release the implementation of our method in MATLAB and C++ for the reference of our community [1].

II. Related Works

A. Relative Pose Estimation

There are mainly two ways to solve multi-robot relative pose estimation (RPE) problems: interloop detection based methods and mutual observation based methods. Most interloop detection based methods, including centralized [4] and decentralized architectures [5, 6], firstly determine whether the robots in a team visited the same places using loop detection technique, then conduct the relative pose recovery. However, interloop detection-based methods require significant computation and bandwidth and have poor performance in environments with many similar scenes.

Most mutual observation-based methods employ robot-to-robot range or bearing measurements to recover relative poses. Early work [8, 9] take extended Kalman filter (EKF) as nonlinear estimator using prior identified range measurements. Zhou [10] provides a set of 14 minimal analytical solutions that cover any combination of range and bearing measurements. However, their proposed algorithm has poor performance under noise because it only uses minimal measurements. Besides, all the above works assume that correspondence between measurement and estimated poses is known, which is not common in practical applications.

Cognetti [11] solves mutual localization problem with particle filters (PF) using anonymous measurements. Indelman [12] and Dong [13] formulate a multi-robot pose graph problem and utilize the expectation-maximization (EM) approach to estimate initial relative poses between robots. However, it is well known that PF and EM all require extensive computation. Nguyen [14] adapts the coupled probabilistic data association filter to estimate relative pose with vision sensor and IMU. In [15], Jang proposes an alternating minimization algorithm to optimize relative poses in multi-robot monocular SLAM. However, these local optimization methods are sensitive to initial values and cannot work with multiple bearing measurements in one image. Compared with the above work, our proposed method solves correspondence and relative poses together without extra sensor inputs.

[1]https://github.com/ZJU-FAST-Lab/CertifiableMutualLocalization

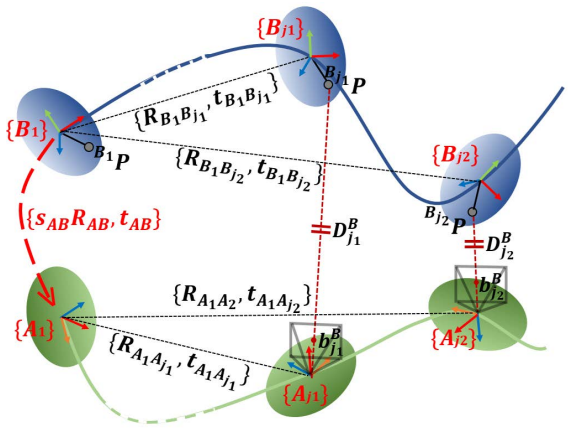

Fig. 2. Illusion of bearing-based mutual localization in two-robot case.

B. Certifiably Global Optimization

Recently, based on semidefinite relaxation and advanced optimization theory, the research community has developed certifiably optimal non-minimal solvers for many computer vision and robotics problems that are non-convex and NP-hard. In [16], Carlone uses Lagrangian duality to verify the optimality of candidate solution of pose graph optimization (PGO). Exploiting the strong duality of PGO, SE-Sync [17] obtain the optimal solution of PGO under acceptable noise. In [18], point registration with outliers is formulated as a QCQP by binary cloning, relaxed using SDR, and finally globally optimized by adding redundant constraints. Besides, SDR is also leveraged in 3d registration [19], camera pose estimation [20], extrinsic calibration [21] and so on. All of these problems involve optimization over SO(3) or SE(3) variables and add orthogonality constraints to make the convex relaxation tight. In this paper, our solution procedure is similar to [18]. Differently, we keep binary variables and introduce *binary constraint* and *correspondence constraint* to formulate a MIQCQP problem. As far as we know, our proposed algorithm is the first method that can obtain a globally optimal solution for the mutual localization problem using anonymous measurements.

III. Formulation of Relative Pose Estimation

In this section, we formulate the RPE problem with anonymous measurements as a QCQP problem. Firstly, we define a loop error for mutual localization of one observed robot case in Sec.III-A. Then in Sec.III-B, we extend the error to multiple observed robots case, introduce binary variables for data association, and formulate the optimization as a mixed-integer programming problem. Finally, we marginalize distance variables, define auxiliary variables, and derive a QCQP problemin Sec.III-C.

A. Loop Error for One Observed Robot

In this subsection, we consider two robots, observer robot A and observed robot B, moving along two 3D trajectories. As shown in Fig. 2, their camera coordinates frame at time j are denoted by $\{A_j\}$ and $\{B_j\}$, where $j \in J$, J is the timestamp collection. Robot A observes feature of robot B at

time j and gets the bearing measurement b_j^B in frame $\{A_j\}$. Assuming B be rigid body, the inner bias ${}^BP$ between the feature and camera on B are time-invariant, i.e., ${}^BP = {}^{B_j}P$. Then ${}^{A_j}P$, the feature coordinate in $\{A_j\}$ is given by

$$ {}^{A_j}P = D_j^B b_j^B = R_{A_jB_j}{}^{B_j}P + t_{A_jB_j} = R_{A_jB_j}{}^BP + t_{A_jB_j}, \tag{1} $$

where D_j^B is the distance between A's camera and the observed feature. $R_{A_jB_j}$ and $t_{A_jB_j}$ denote the relative rotation and translation between $\{A_j\}$ and $\{B_j\}$. For simplicity, we set $D_j = D_j^B$ and $b_j = b_j^B$ in two robots' case. And for each time j, we have

$$ R_{A_1A_j}{}^{A_j}P + t_{A_1A_j} = s_{AB}R_{AB}(R_{B_1B_j}{}^BP + t_{B_1B_j}) + t_{AB}, \tag{2} $$

where s_{AB} denotes the scale ratio between local maps of A and B, and $\{s_{AB}R_{AB}, t_{AB}\}$ is the corresponding relative pose. After subtraction between Equ.(2) of $j_1, j_2 \in J$, we eliminate variable t_{AB} and derive the *loop error*:

$$ \begin{aligned} e_{j_1j_2}^{AB} = {} & s_{AB}R_{AB}\widehat{t}_{B_{j_1}B_{j_2}} + R_{AB}\widehat{R}_{B_{j_1}B_{j_2}}{}^B\bar{P} - \\ & \widehat{t}_{A_{j_1}A_{j_2}} + R_{A_1A_{j_1}}b_{j_1}D_{j_1} - R_{A_1A_{j_2}}b_{j_2}D_{j_2}, \end{aligned} \tag{3} $$

where $\widehat{t}_{X_{j_1}X_{j_2}} = t_{X_1X_{j_2}} - t_{X_1X_{j_1}}, X \in \{A, B\}$, $\widehat{R}_{B_{j_1}B_{j_2}} = R_{B_1B_{j_2}} - R_{B_1B_{j_1}}$ and ${}^B\bar{P} = {}^BP/s_{AB}$. If $s_{AB}R_{AB}$ and ${}^B\bar{P}$ are recovered, ${}^BP$ can be determined solely. This expression is found in [15]. In this paper, we reformulate it in a linear expression which will be used to get a quadratic cost in Sec.III-B. Firstly we define the following variables:

$$ \begin{aligned} r_s &\doteq \text{vec}(s_{AB}R_{AB}) \in \mathbb{R}^{9\times 1}, \\ r_p &\doteq \text{vec}({}^B\bar{P}^T \otimes R_{AB}) \in \mathbb{R}^{27\times 1}, \end{aligned} \tag{4} $$

where $\otimes$ is the Kronecker product, $\text{vec}(M)$ is the vectorization (applied column-wise) of matrix M. Then we introduce an additional variable y and constraint $y^2 = 1$ to define

$$ x_{j_1j_2}^{AB} \doteq [\ r_s^T,\ r_p^T,\ y,\ D_{j_1},\ D_{j_2}\ ]^T \in \mathbb{R}^{(9+27+1+2)\times 1}. \tag{5} $$

Then the loop error of the edge $\{j_1, j_2\}$ is rewritten as

$$ \begin{aligned} e_{j_1j_2}^{AB} = {} & [\ \widehat{t}_{B_{j_1}B_{j_2}}^T \otimes I,\ \text{vec}(\widehat{R}_{B_{j_1}B_{j_2}})^T \otimes I, \\ & -\widehat{t}_{A_{j_1}A_{j_2}},\ R_{A_1A_{j_1}}b_{j_1},\ -R_{A_1A_{j_2}}b_{j_2}\ ]x_{j_1j_2}^{AB}, \end{aligned} \tag{6} $$

The derivation of Equ.(6) from Equ.(3) is given in supplementary material.

B. Mutual Localization with Anonymous Measurements

In this section, we extend the loop error to multi-observed-robot cases. Considering te scenario that there is one observer robot and $N \geq 2$ observed robots, the correspondence of bearing measurement sequence $b^X = \{b_j^X\}_{j\in J}$ and estimated pose trajectory $T_Y = \{R_{Y_j}, t_{Y_j}\}_{j\in J}$ is unknown. Here X, Y are indexes of measurement sequence and estimated trajectory respectively. Recovering the correspondence is called anonymity recovery problem. To solve it, we introduce binary variables $\Theta = \{\theta_{XY}\}_{X,Y\in[1,N]}$, in which the *binary constraint* ($\theta_{XY} = \{0, 1\}$) indicates whether the X^{th} bearing measurement corresponds to the Y^{th} trajectory

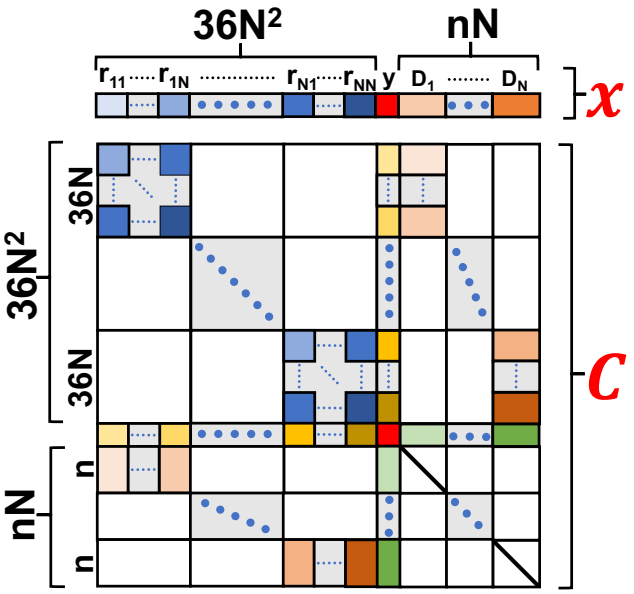

Fig. 3. Structure of decision variable x and cost matrix C in Problem 3.1.

($\theta_{XY} = 1$) or not ($\theta_{XY} = 0$). And we introduce the *correspondece constraints* :

$$ \sum_X \theta_{XY} = 1, \sum_Y \theta_{XY} = 1, \forall X, Y \in [1, N]. \tag{7} $$

The above constraints are to guarantee that the measurement sequences and the estimated trajectories have one-to-one correspondence.

We then utilize the unknown variables $\{\theta_{XY}\}$ to rewrite the loop error Equ.(3) for the X^{th} measurement as follow

$$ \begin{aligned} e_{j_1j_2}^X = {} & R_{A_1A_{j_2}}b_{j_2}^X D_{j_2}^X - R_{A_1A_{j_1}}b_{j_1}^X D_{j_1}^X + \widehat{t}_{A_{j_1}A_{j_2}} - \\ & \sum_{Y=1}^N \theta_{XY}(R_{AY}\widehat{R}_{Y_{j_1}Y_{j_2}}{}^Y\bar{P} + s_{AY}R_{AY}\widehat{t}_{Y_{j_1}Y_{j_2}}). \end{aligned} \tag{8} $$

Now we convert the mixed-integer expression to a linear form. Firstly, we denote the parameters that need to be estimated for robot Y as ${}^Y\mathbb{P} \doteq [s_{AY}, {}^Y\bar{P}^T]^T$. Then we define extra variables ${}^Y\mathbb{P}_X \doteq \theta_{XY}{}^Y\mathbb{P}$. Furthermore, we define the following unknown decision variables,

$$ r_{XY} \doteq \text{vec}({}^Y\mathbb{P}_X^T \otimes R_{AY}) \in \mathbb{R}^{36\times 1}, \tag{9} $$

$$ r_X \doteq \text{vstack}(\{r_{XY}\}_{Y=1}^N) \in \mathbb{R}^{36N\times 1}, \tag{10} $$

$$ D_X \doteq \text{vstack}(\{D_j^X\}_{j\in J}) \in \mathbb{R}^{n\times 1}, \tag{11} $$

$$ r \doteq \text{vstack}(\{r_X\}_{X=1}^N) \in \mathbb{R}^{36N^2\times 1}, \tag{12} $$

$$ D \doteq \text{vstack}(\{D_X\}_{X=1}^N) \in \mathbb{R}^{nN\times 1}, \tag{13} $$

$$ x \doteq [\ r^T,\ y,\ D^T\ ]^T \in \mathbb{R}^{(36N^2+1+nN)\times 1}. \tag{14} $$

where the notation $\text{vstack}(G)$ stacks all variable in G vertically and n is the number of measurements. r_{XY} can be seen as augmented version of r_s and r_p in multi-robot cases. We then rewrite Equ. (8) in linear form as $e_{j_1j_2}^X = c_{j_1j_2}^X x, X \in [1, N]$, in which x is unknown decision variable and $c_{j_1j_2}^X$ is derived using collected data. We refer readers to supplementary material for detailed derivation. Then the

error of each measurement sequence is used to formulate a nonconvex least-square problem

Problem 3.1 (Original Problem):

$$\begin{aligned} x^* =& \arg\min_{x} \sum_{\substack{X\in[1,N]\\ \{j_1,j_2\}\in J}} (e^X_{j_1j_2})^T w^X_{j_1j_2} e^X_{j_1j_2} \\ =& \arg\min_{x} \sum_{\substack{X\in[1,N]\\ \{j_1,j_2\}\in J}} x^T (c^X_{j_1j_2})^T w^X_{j_1j_2} c^X_{j_1j_2} x \\ =& \arg\min_{x} \; x^T \underbrace{\Big(\sum_{\substack{X\in[1,N]\\ \{j_1,j_2\}\in J}} (c^X_{j_1j_2})^T w^X_{j_1j_2} c^X_{j_1j_2} \Big)}_{:=C} x \\ s.t. & D^X_j > 0,\; s_{AY} > 0,\; R_{AY} \in SO(3), \\ & \sum_X \theta_{XY} = 1, \sum_Y \theta_{XY} = 1, \theta_{XY} \in \{0,1\}, \end{aligned} \tag{15}$$

where $w^X_{j_1j_2}$ is the measurement confidence parameter. The structure of x and C is shown in Fig.3. Note that C is a Gram matrix, so it is positive semidefinite and symmetric.

C. Marginalization and Auxiliary Variables

In this subsection, we following the procedures in [21] to marginalize the distance variables using Schur Complement. We write cost matrix C as

$$C = \begin{bmatrix} C_{\bar{\mathcal{D}},\bar{\mathcal{D}}} & C_{\bar{\mathcal{D}},\mathcal{D}} \\ C_{\mathcal{D},\bar{\mathcal{D}}} & C_{\mathcal{D},\mathcal{D}} \end{bmatrix}, \tag{16}$$

where the subindex $\mathcal{D}$ stands for the set of indexes corresponding to the distance variables or not (subindex $\bar{\mathcal{D}}$). Then we eliminate distance variables D and obtain

Problem 3.2 (Marginalized Problem):

$$\begin{aligned} z^* &= \arg\min_{z} \; z^T \bar{C} z \\ s.t. \quad & s_{AX} > 0, R_{AY} \in SO(3), \\ & \sum_X \theta_{XY} = 1, \sum_Y \theta_{XY} = 1, \theta_{XY} \in \{0,1\}, \end{aligned} \tag{17}$$

where $z = [r^T,\; y]^T$ and $\bar{C} = C/C_{\mathcal{D},\mathcal{D}} = C_{\bar{\mathcal{D}},\bar{\mathcal{D}}} - C_{\bar{\mathcal{D}},\mathcal{D}} C^{-1}_{\mathcal{D},\mathcal{D}} C_{\mathcal{D},\bar{\mathcal{D}}}$. Note that after marginalization, the number of involved variables is solely related to N. In contrast, exiting local optimization methods' computation is not only related to N but also the number of measurements.

Then we formulate all involved constraints by introducing auxiliary variables. To represent the involved $SO(3)$ constraint in $r_{XY} = \text{vec}(\theta_{XY}\, {}^Y\mathbb{P} \otimes R_{AY})$, we introduce $\mu = \theta_{XY} h_Y$ and generalize the constraint as follows

$$(\mu R_{AY})^T (\mu R_{AY}) = \mu^2 I, \tag{18}$$

$$(\mu R_{AY})(\mu R_{AY})^T = \mu^2 I, \tag{19}$$

$$\begin{aligned} (\mu R_{AY})^{(i)} \times (\mu R_{AY})^{(j)} = \mu (\mu R_{AY})^{(k)}, \\ \forall (i,j,k) = \{(1,2,3),(2,3,1),(3,1,2)\}. \end{aligned} \tag{20}$$

where h_Y could be the term s_{AY}, ${}^Y\bar{P}^{(1)}$, ${}^Y\bar{P}^{(2)}$ or ${}^Y\bar{P}^{(3)}$. Variable μ is unknown. We introduce this variable only for modeling constraints. Note that constraints (19) and (20) are actually redundant [18], and we will study how they make convex relaxation tighter in Sec.V. Then we process the binary constraint and correspondence constraints introduced in Sec.III-B. After rewriting the binary constrain as $\theta^2_{XY} - \theta_{XY} = 0$, all these constraints can be written as quadratic in term of θ_{XY}. However, until now our decision variables only include $\text{vec}(\theta_{XY} h_Y R_{AY})$ and auxiliary $\mu = \theta_{XY} h_Y$ but not θ_{XY}. Thus, we need to apply the two type of constraints to decision variables by introducing other auxiliary unknown variable h_Y and $\text{vec}(h_Y R_{AY})$. They link original variables with θ_{XY} like a "bridge" by following equality constraints

$$\underline{\theta_{XY} h_Y} = \underline{\theta_{XY}}\; \underline{h_Y}, \tag{21}$$

$$\underline{\text{vec}(\theta_{XY} h_Y R_{AY})} = \underline{\theta_{XY}}\; \underline{\text{vec}(h_Y R_{AY})}. \tag{22}$$

where the underlines denote independent variables.

Until now, we have introduced all necessary constraints, including $SO(3)$, binary, correspondence and equality constraints. All of them can be formulated as quadratic.

Summarize all necessary auxiliary variables as follows

1) **Lifted Rotation Variable:** ℓ

$$\ell_Y \doteq vec({}^Y\mathbb{P}^T \otimes R_{AY}) \in \mathbb{R}^{36\times 1}, \tag{23}$$

$$\ell \doteq \text{vstack}(\{\ell_Y\}_{Y=1}^N) \in \mathbb{R}^{36N\times 1}. \tag{24}$$

2) **Binary Variable:** φ_θ

$$\varphi^X_\theta \doteq \text{vstack}(\{\theta_{XY}\}_{Y=1}^N) \in \mathbb{R}^{N\times 1}, \tag{25}$$

$$\varphi_\theta \doteq \text{vstack}(\{\varphi^X_\theta\}_{X=1}^N) \in \mathbb{R}^{N^2\times 1}. \tag{26}$$

3) **Scale Ratio and Inner Bias Variable:** φ_h

$$\varphi_h \doteq \text{vstack}(\{{}^Y\mathbb{P}\}_{X=1}^N) \in \mathbb{R}^{4N\times 1}. \tag{27}$$

4) **Lifted Scale Ratio and Inner Bias Variable:** φ_μ

$$\varphi^X_\mu \doteq \text{vstack}(\{{}^Y\mathbb{P}_X\}_{Y=1}^N) \in \mathbb{R}^{4N\times 1}, \tag{28}$$

$$\varphi_\mu \doteq \text{vstack}(\{\varphi^X_\mu\}_{X=1}^N) \in \mathbb{R}^{4N^2\times 1}. \tag{29}$$

Now we define the final decision variable

$$f \doteq [z^T, \ell^T, \varphi^T_\theta, \varphi^T_h, \varphi^T_\mu]^T,$$

and use it to formulate all constraints in quadratic terms $f^T Q_i f = g_i, i \in [1,m]$, where m is the number of constraints. For detailed derivation of Q_i and g_i, we refer readers to supplementary material. Now we obtain

Problem 3.3 (QCQP Probem):

$$\begin{aligned} f*_{\text{QCQP}} &= \min_{f} \; f^T Q_0 f \\ s.t. \quad & f^T Q_i f = g_i, i = 1,..,m, \end{aligned} \tag{30}$$

where $Q_0 = \begin{bmatrix} \bar{C} & 0_{d_z\times d_a} \\ 0_{d_a\times d_z} & 0_{d_a\times d_a} \end{bmatrix}$. d_z and d_a are dimensions of z and auxiliary variables respectively.

However, Problem 3.3 is non-convex and nontrivial to solve. In next section, we provide a complete algorithm using SDR to get the global optimal solution of Problem 3.3.

IV. Certifiably Global Optimization by Semidefinite Relaxation

In this section, we will firstly apply semidefinite relaxation to Problem 3.3 in Sec.IV-A. Then we recover data correspondence and relative poses from the solution of the SDP problem in Sec.IV-B. Lastly, we provide a condition under which the zero-duality-gap and one-rank-solution can be guaranteed in noise-free cases in Sec.IV-C.

A. Semidefinite Relaxation and Dual Problem

As stated above, Problem 3.3 is non-convex. Fortunately, it can be relaxed to a convex SDP, known as Shor's relaxation. By introducing matrix variable $F \doteq ff^T$, we have

$$f^T Q_i f = \text{tr}\,(f^T Q_i f) = \text{tr}\,(Q_i f f^T) = \text{tr}\,(Q_i F), \quad (31)$$

where $\text{tr}(M)$ is the trace of matrix M. Together with Equ.(31) and dropping the constraint of $\text{rank}\,(F) = 1$, we obtain the following problem.

Problem 4.1 (Primal SDP):

$$f^*_{\text{primal}} = \min_F \ \text{tr}(Q_0 F)$$
$$s.t. F \succeq 0, \text{tr}(Q_i F) = g_i, i = 1, ..., m, \quad (32)$$

which is convex and can be solved by off-shelf solvers using primal-dual interior point method. Its dual problem is

Problem 4.2 (Dual SDP):

$$f^*_{\text{dual}} = \max_\lambda \ g^T \lambda$$
$$s.t. Q(\lambda) = Q_0 - \sum_i \lambda_i Q_i \succeq 0, i = 1, ..., m, \quad (33)$$

where $g = [g_1, ..., g_m]^T$, $\lambda = [\lambda_1, ..., \lambda_m]^T$.

Once F^*, the solution of Problem 4.1, is obtained, we denote the part of F^* that corresponds to variable z as $Z^* \doteq F^*_{[1:36N^2, 1:36N^2]}$. Moreover, if zero-duality-gap ($f^*_{\text{primal}} = f^*_{\text{dual}}$) and one-rank-solution ($\text{rank}\,(Z^*) = 1$) hold, we can obtain the global optimal solution z^* of Problem 3.2 as described in Sec. IV-B. Actually, both the above conditions are satisfied in noise-free cases, which is proved in Sec. IV-C.

B. Recovery from the tight SDP solution

Given Z^*, we need to recover the optimal correspondences and relative poses. According to the one-rank-solution ($\text{rank}(Z^*) = 1$), we firstly deploy a rank-one decomposition to obtain $z^* \in \mathbb{R}^{36N^2 \times 1}$. Denoting $r^*_{XY} \in \mathbb{R}^{36\times 1}$ as slices of z^* corresponding to variable r_{XY}, we define $M^X_{AY} \doteq {\theta^*_{XY}}^Y \mathbb{P}^{*T} \otimes R^*_{AY} = \text{mat}(r^*_{XY}, [12, 3])$, where $\text{mat}(v, [r, c])$ means reshape the vector v to one $r \times c$ matrix by col-first order. Note that M^X_{AY} is either zero matrix or non-zero matrix due to the binary variable θ_{XY}. So we set $\epsilon = 10^{-5}$ and take $||M^X_{AY}||_2 > \epsilon$ to indicate that the X^{th} measurement corresponds to the Y^{th} estimated trajectory.

Then for each M^X_{AY} whose corresponding $\theta_{XY} > \epsilon$, we recover the scale ratio and relative rotation

$$\mathbb{S}^*_{AY} := s^*_{AY} R^*_{AY} = M^X_{AY[1:3,1:3]}, \quad (34)$$

$$s^*_{AY} = \sqrt[3]{\det(\mathbb{S}^*_{AY})}, R^*_{AY} = \mathbb{S}^*_{AY}/s^*_{AY}, \quad (35)$$

and inner bias ${}^Y P^*$ similarly.

Recall that we have marginalized the distance variable D in Sec.III-C, we now recover the optimal D^* as

$$D^*(r^*) = -C^{-1}_{\mathcal{D},\mathcal{D}} C_{\mathcal{D},\bar{\mathcal{D}}} r^*. \quad (36)$$

Furthermore, the optimal relative translation t^*_{AY} is recovered using D^* as follow

$$t^*_{AY} = \sum_{j\in J} (t_{A_1 A_j} + R_{A_1 A_j}(D^{Y*}_j \ b^Y_j) - s^*_{AY} R^*_{AY}({R_{Y_1 Y_j}}^Y P^* + t_{Y_1 Y_j})). \quad (37)$$

C. Tightness of Semidefinite Relaxation

In this subsection, we aim to prove that there are zero-duality-gap and one-rank-solution in noise-free cases. Firstly, we introduce a lemma and a corank-one condition.

Lemma 4.1: If $C \in \mathbb{R}^{n\times n}$ be positive semidefinite and $x^T C x = 0$ for a vector x, then $Cx = 0$.

Definition 1: For Problem 3.1, the *corank-one condition* holds if the number of *independent* measurements n and the number of observed robots N, satisfy that $n \geq 18N + 2$, where *independent* measurements mean that $\{c^X_{j_1 j_2}\}$ are linearly independent vectors.

Based in this condition, we have

Lemma 4.2: Assume that the corank-one condition holds, then the cost matrix C is semidefinite and has corank one in noise-free cases. Furthermore, after Schur Compliment, $\bar{C}$ is also semidefinite and has corank one.

The detailed proof of above two lemmas can be seen in supplementary material. Then we apply Lemma 2.1 in [22] to our problem and introduce the following proposition.

Proposition 1: If bearing measurements are noise-free, their is zero-duality-gap between Problem 3.3 and Problem 4.2. Furthermore, once the corank-one condition is satisfied and given the solution Z^* of Problem 4.1, we have $\text{rank}(Z^*) = 1$, and its rank-one decomposition z^* is the global optimal minimum of Problem 3.3.

Proof: Let $\tilde{f} = [\tilde{z}^T, \tilde{\ell}^T, {\tilde{\varphi}_\theta}^T, {\tilde{\varphi}_p}^T, {\tilde{\varphi}_\mu}^T]^T$ be a feasible point in Problem 3.3 where $\tilde{z}, \tilde{\ell}, \tilde{\varphi}_\theta, \tilde{\varphi}_p, \tilde{\varphi}_\mu$ are all ground truth. Let $\tilde{\lambda} = 0$ be a feasible point in Problem 4.2. Then the zero-duality-gap is guaranteed since the below three conditions needed in Lemma 2.1 in [22] are satisfied: (i) Primal feasibility. In noise-free cases, the ground truth certainly satisfy constraints in Problem 3.3. (ii) Dual feasibility. $Q(\tilde{\lambda}) = Q_0 - \sum_{i=1}^m \tilde{\lambda}_i Q_i = Q_0 = \begin{bmatrix} \bar{C} & 0_{d_z \times d_a} \\ 0_{d_a \times d_z} & 0_{d_a \times d_a} \end{bmatrix} \succeq 0$. (iii) Lagrangian multiplier. Since $\tilde{f}$ is ground truth, $\tilde{f}^T Q_0 \tilde{f}$ equals to the optimal cost in Problem 3.1, which equals to 0. Recall that Q_0 is semidefinite according to Lemma. 4.2, so $Q(\tilde{\lambda})\tilde{f} = 0$ is obtained based on Lemma. 4.1.

Furthermore, suppose Z^* is the optimal solution of Problem 4.1. Then $Z^* \neq 0$ since at least one $g_i \neq 0$. By complementary slackness, $\text{tr}(Q(\tilde{\lambda})Z^*) = \text{tr}(\begin{bmatrix} \bar{C}Z^* & 0_{d_z \times d_a} \\ 0_{d_a \times d_z} & 0_{d_a \times d_a} \end{bmatrix}) = 0$, so $\text{tr}(\bar{C}Z^*) = 0$. And since $\bar{C}$ and Z^* are both positive semidefinite, $\text{rank}(\bar{C}) + \text{rank}(Z^*) <= N$. So, if $\text{corank}(\bar{C}) = 1$, $\text{rank}(Z^*) = 1$. Moreover, its rank-one decomposition $\tilde{z}$ is the unique optimum of Problem 3.3. ■

V. Experiments

In this section, we firstly confirm the optimality and efficiency of our method by comparing it against the alternating minimization (AM) [15] and the Levenberg-Marquardt (LM) methods. Next, to present the robustness of our method, we compare its performances under different levels of noise. Then, we show the results of our method with different robot number and noise. Finally, we apply our algorithm in real-world, using estimated odometry from different sources. Our algorithm is implemented in MATLAB using cvx [23] and

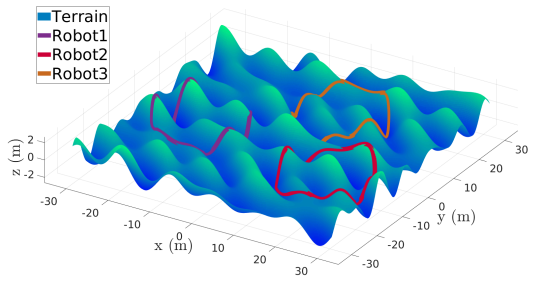

Fig. 4. Three random trajectories in a simulated environment.

in C++ using MOSEK [24] respectively. We run simulated experiments in PC (Intel i7-10900 CPU) and real-world experiments in NUC (i5-1135G7 CPU) with single core.

A. Experiments on Synthetic Data

To simulate bearing measurement, we generate random trajectories for multiple robots. An example simulated environment is shown in Fig. 4. Robots trace circular routes around different centers over a common landscape consisting of multiple random sinusoidal functions. All trajectories have the same length. Then, for observer robot A and observed robot Y, we use their global poses to generate noisy bearing measurement as follow

$$b_j^Y = R_{A_j}^{-1}(t_{Y_j} + \mathcal{N}(1,\sigma)R_{Y_j}{}^Y P - t_{A_j}). \tag{38}$$

where $\mathcal{N}(1,\sigma)$ is Gaussian distribution with standard deviation σ (unit:m) . We take the first pose of each trajectory as the local world frame and obtain each robot's local poses, which will be shared with other robots.

1) Optimality and Runtime: We firstly present how AM and LM is sensitive to initial values. As shown in Fig.5, if the Euclidean distance of initial value and the ground truth is greater than a threshold, both AM and LM will fall into local minimum. Fig.5 clearly states that the optimality of these methods can't be guaranteed.

Then we compare our method with these two methods in optimality and efficiency for four problems: RPE without scale ratio and inner bias (RPE-only), RPE with scale ratio (RPE-S), RPE with inner bias (RPE-B), and RPE with scale ratio and inner bias (RPE-SB). For each problem, we conduct 1000 experiments using different measurements. The left figure of Fig. 6 shows that for all problems, our method can always obtain the optimal solution, while both AM and LM are trapped in local minimums with random initial values. For efficiency, since our formulation fixes the number of variables by marginalizing the distance variables D (see Sec.III-C), its computing time is only related to the number of observed robots. In contrast, the number of variables in local optimization methods AM and LM increase with measurement number. The right figure of Fig. 6, which presents the mean runtime with 200 bearing measurements, show that our method solves all problems faster.

2) Robustness: To evaluate the robustness of our method and the effectiveness of the redundant constraints, we add different levels of noise into simulated measurements. We compare two versions of our method, the default version

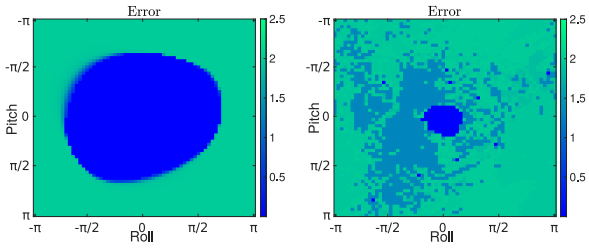

Fig. 5. Error heatmaps for two methods over a uniform sampled initial value R_{init}. We generate R_{init} by rotating ground truth with roll (x-axis) and pitch (y-axis) from $-\pi$ and π. Blue region (zero-error) denotes range of R_{init} that leads to global optimality, while green region (non-zero-error) denotes range of R_{init} that leads to local minimums.

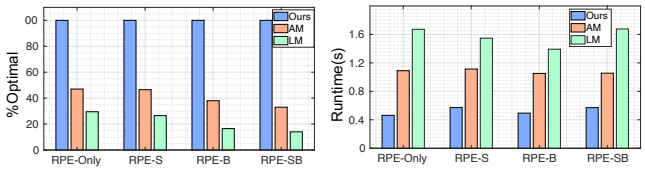

Fig. 6. Benchmark results between our method and local optimization algorithms for different problems.

(D) and the augmented version which is added redundant rotation constraint (D+R). As the left plot of Fig. 7 shows, for each noise level, the augmented version (D+R) recover an exact minimizer of the primal problem. However, the default version (D) does not obtain the one-rank solution under extreme noise ($\sigma \geq 0.4$).

Fig. 9 presents the performance of our method (D and D+R) and several local optimization methods. AM and LM utilize random rotation as the initial value, and AM (GT) and LM (GT) use the ground truth instead. Each figure represents 100 random trials on simulated data with different noise levels σ. As Fig. 9 shows, our method is consistently more accurate compared to AM and LM and has comparable performance with AM (GT). Besides, we observe that under extreme noise (σ = 0.5), our method still performs accurately.

Furthermore, we conduct experiments with multiple observed robots under noise. The right figure of Fig.7 presents the trend of rank(Z) when the noise level increases. Although as the number of robots increases, the zero-duality-gap is easily influenced by noise, our method always obtains one-rank solution with a common noise level ($\sigma < 0.4$). Moreover, we present the error distribution of obtained solutions in Fig. 10. The figure shows that as the measurement increase with the

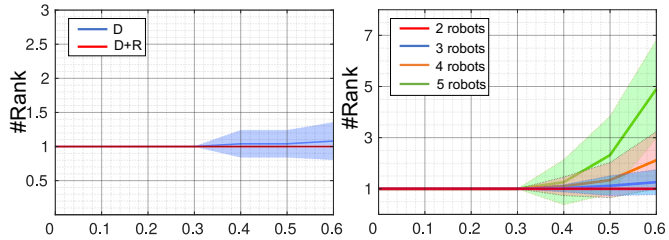

Fig. 7. (Left) Comparison between method w/ and w/o redundant constraints (Right) Comparison with different number of robots. (solid line: mean; shaded area: 1-sigma standard deviation).

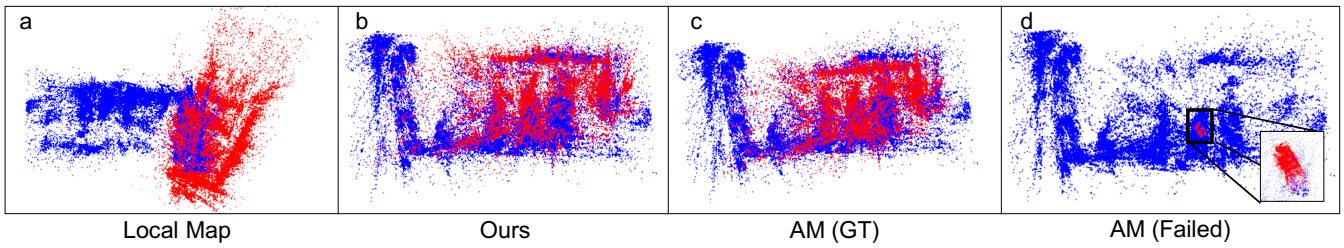

Fig. 8. Map fusion results using feature maps from two robots, which launch at different place and observe each other when they rendezvous.

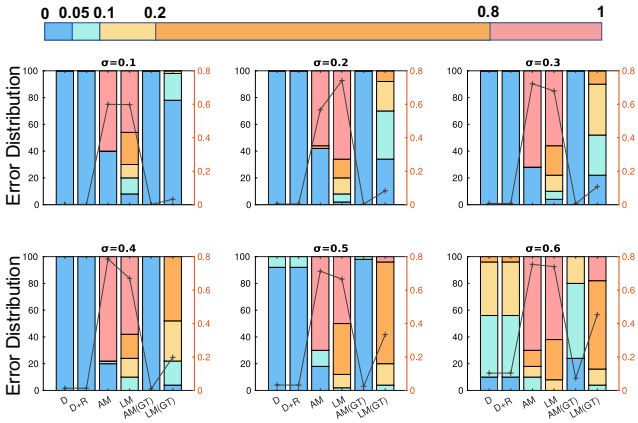

Fig. 9. Comparison of error distribution between different methods. The top colorbar presents colors corresponding to different error range. In each subfigure, each bar denotes the percentage of error range and the black line represents the mean error.

number of robots, the accuracy does not be affected majorly. And although there is no one-rank-solution under extreme noise, the result of one-rank decomposition has comparable accuracy with AM (GT).

3) Scalability: In our method, the number of variables is related to the squared number of observed robots. Fig. 11 presents runtime of our method with different numbers of robots. It shows that our method has acceptable runtime in real multi-robot applications when robot number is limited.

B. Real-world Experiments

We use UAVs as our experimental platform. Firstly, we carry on experiments using motion capture or VIO for odometry estimations and AprilTag for bearing measurements. Table V-B summarizes the results of comparison between our method and others. In these experiments, we walked randomly with drones to ensure that the observer robot could see other robots continuously. AM (C+GT) optimizes with groundtruth correspondence in experiments with three robots, while AM (w/o C+GT) does not. As Table V-B shows, AM and AM (w/o C+GT) will converge to small cost but obtain egregiously large error, which indicates that they are all trapped in a local minimum. Compared with them, our globally optimal approach obtains the minimum cost in all experiments with the most or second most small error. Note that, due to noise, obtaining the minimum cost does not mean obtaining the most accurate estimation.

Then, we carry on experiments with different detection equipment, including a fish-eye camera and unified LED marker. We let UAVs move along 3D random trajectories autonomously. Although the observer's obtained measurements are intermittent due to the affected line of sight, our proposed method also recovered accurate correspondence and relative poses. The results are as shown in Fig.11 and Fig.12. More details are available in the attached video.

Finally, we apply our algorithm in multi-robot map fusion, as Fig.7 shows. In this experiment, two robots obtain their local map by ORB-SLAM2 [7] and bearing measurements using ApirlTag. Then local maps from different robots are fused with optimized relative poses. Compared with AM, which traps into local minimum and fails to fuse maps, our proposed method fuses robots' maps correctly without any initialization, while AM (GT) needs.

We consume 250MB memory in two-robots experiments and 300 MB in three-robots experiments. Since only trajectory information is passed between robots, we consume bandwidth at 0.5KB per message. In all experiments, robots send odometry at 100hz, resulting in 50KB/s per robot.

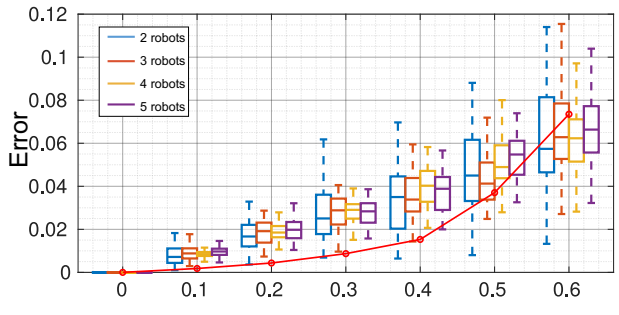

Fig. 10. Comparison of error distribution of our method with different number of robots under different levels of noise. The red line denotes the average error of AM (GT) with known correspondence.

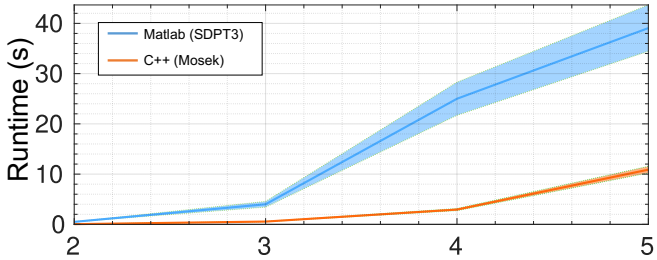

Fig. 11. Runtime comparison using Matlab / C++ with different number of robots. (solid line: mean; shaded area: 1-sigma standard deviation)

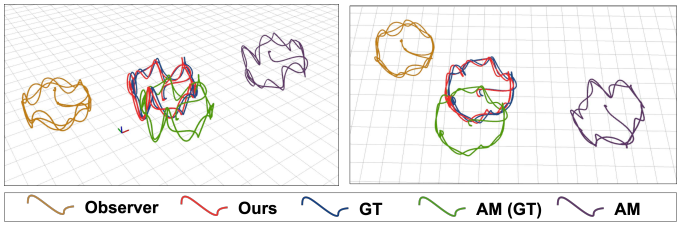

Fig. 12. Result of two-robots experiment. Two robots move along 1m-radius 3D curves at 1.5 m/s. Our proposed method recover the most accurate trajectory while other methods fall into local minimum.

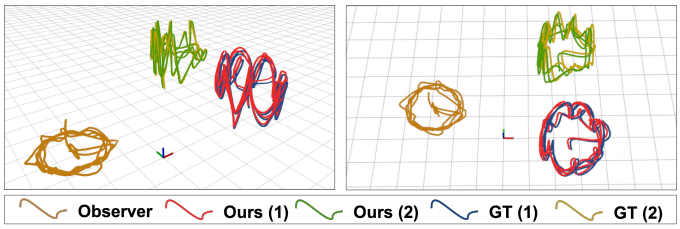

Fig. 13. Result of three-robots experiment. Three robots move along 0.8m-radius 3D curves at 1.5m/s. Our proposed method recover bearing-trajectory correspondence and relative poses jointly. Numbers are indexes of robots.

VI. Conclusions and Future Work

In this paper, we propose a certifiably globally optimal algorithm for mutual localization with anonymous bearing measurements. We can recover bearing-pose correspondences and relative poses by convex relaxation and SDP. We provide a necessary condition under which relaxation is tight in noise-less cases. Extensive experiments show that our method can be used without initial value but with optimality guarantee, compared with state-of-the-art. Moreover, constant runtime reveals that our method has considerable scalability for measurements and promising application in bootstrapping other algorithms that need relative pose for initialization. While extreme noise will affect the optimality, we will explore an exact noise tolerance threshold to provide a more powerful guarantee for application in the future.

TABLE I

Real-world Experiments Results

Scene	#Robots	Method	Cost	L^2 Error		Runtime (ms)
				Trans. (m)	Rot.	
VICON+ AprilTag	2	Ours (D+R)	**0.0018**	**0.24**	**0.063**	**343.3**
		AM	0.198	2.52	2.83	1391.5
		AM (GT)	0.132	0.323	0.087	660.5
	3	Ours (D+R)	**0.0006**	0.092	0.0688	**603.3**
		AM (C)	0.727	2.506	2.809	1289.1
		AM (C+GT)	0.082	**0.0305**	**0.0305**	632.7
		AM (w/o C+GT)	0.569	1.53	0.290	901.5
VINS+ AprilTag	2	Ours (D+R)	**0.553**	**0.429**	**0.0716**	180.1
		AM	0.667	2.597	2.823	312.3
		AM (GT)	0.650	0.601	0.159	**120.2**
	3	Ours (D+R)	**0.101**	0.943	0.205	603.3
		AM (C)	0.423	1.62	2.82	512.3
		AM (C+GT)	0.337	**0.773**	**0.118**	**131.1**
		AM (w/o C+GT)	1.308	7.24	2.82	305.6

References

[1] L. Quan, L. Yin, C. Xu, and F. Gao, "Distributed swarm trajectory optimization for formation flight in dense environments," *arXiv preprint arXiv:2109.07682*, 2021.

[2] Y. Gao, Y. Wang, X. Zhong, T. Yang, M. Wang, Z. Xu, Y. Wang, C. Xu, and F. Gao, "Meeting-merging-mission: A multi-robot coordinate framework for large-scale communication-limited exploration," *arXiv preprint arXiv:2109.07764*, 2021.

[3] K. Dorling, J. Heinrichs, G. G. Messier, and S. Magierowski, "Vehicle routing problems for drone delivery," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 1, pp. 70–85, 2016.

[4] P. Schmuck and M. Chli, "Ccm-slam: Robust and efficient centralized collaborative monocular simultaneous localization and mapping for robotic teams," *Journal of Field Robotics*, vol. 36, no. 4, pp. 763–781, 2019.

[5] T. Cieslewski, S. Choudhary, and D. Scaramuzza, "Data-efficient decentralized visual slam," in *2018 IEEE international conference on robotics and automation (ICRA)*.

[6] P.-Y. Lajoie, B. Ramtoula, Y. Chang, L. Carlone, and G. Beltrame, "Door-slam: Distributed, online, and outlier resilient slam for robotic teams," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1656–1663, 2020.

[7] R. Mur-Artal and J. D. Tardós, "Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras," *IEEE transactions on robotics*, vol. 33, no. 5, pp. 1255–1262, 2017.

[8] X. S. Zhou and S. I. Roumeliotis, "Multi-robot slam with unknown initial correspondence: The robot rendezvous case," in *2006 IEEE/RSJ international conference on intelligent robots and systems*.

[9] C.-H. Chang, S.-C. Wang, and C.-C. Wang, "Vision-based cooperative simultaneous localization and tracking," in *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, pp. 5191–5197.

[10] X. S. Zhou and S. I. Roumeliotis, "Determining 3-d relative transformations for any combination of range and bearing measurements," *IEEE Transactions on Robotics*, vol. 29, no. 2, pp. 458–474, 2012.

[11] M. Cognetti, P. Stegagno, A. Franchi, G. Oriolo, and H. H. Bülthoff, "3-d mutual localization with anonymous bearing measurements," in *2012 IEEE International Conference on Robotics and Automation*.

[12] V. Indelman, E. Nelson, N. Michael, and F. Dellaert, "Multi-robot pose graph localization and data association from unknown initial relative poses via expectation maximization," in *2014 IEEE International Conference on Robotics and Automation*.

[13] J. Dong, E. Nelson, V. Indelman, N. Michael, and F. Dellaert, "Distributed real-time cooperative localization and mapping using an uncertainty-aware expectation maximization approach," in *2015 IEEE International Conference on Robotics and Automation*.

[14] T. Nguyen, K. Mohta, C. J. Taylor, and V. Kumar, "Vision-based multi-mav localization with anonymous relative measurements using coupled probabilistic data association filter," in *2020 IEEE International Conference on Robotics and Automation*.

[15] Y. Jang, C. Oh, Y. Lee, and H. J. Kim, "Multirobot collaborative monocular slam utilizing rendezvous," *IEEE Transactions on Robotics*, vol. 37, no. 5, pp. 1469–1486, 2021.

[16] L. Carlone, D. M. Rosen, G. Calafiore, J. J. Leonard, and F. Dellaert, "Lagrangian duality in 3d slam: Verification techniques and optimal solutions," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems*.

[17] D. M. Rosen, L. Carlone, A. S. Bandeira, and J. J. Leonard, "Se-sync: A certifiably correct algorithm for synchronization over the special euclidean group," *The International Journal of Robotics Research*, vol. 38, no. 2-3, pp. 95–125, 2019.

[18] H. Yang and L. Carlone, "A quaternion-based certifiably optimal solution to the wahba problem with outliers," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 1665–1674.

[19] J. Briales and J. Gonzalez-Jimenez, "Convex global 3d registration with lagrangian duality," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4960–4969.

[20] J. Zhao, "An efficient solution to non-minimal case essential matrix estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.

[21] M. Giamou, Z. Ma, V. Peretroukhin, and J. Kelly, "Certifiably globally optimal extrinsic calibration from per-sensor egomotion," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 367–374, 2019.

[22] D. Cifuentes, S. Agarwal, P. A. Parrilo, and R. R. Thomas, "On the local stability of semidefinite relaxations," *Mathematical Programming*, pp. 1–35, 2021.

[23] M. Grant and S. Boyd, "Cvx: Matlab software for disciplined convex programming, version 2.1," 2014.

[24] M. ApS, *MOSEK Fusion API for C++. Ver. 9.3.*, 2019. [Online]. Available: https://docs.mosek.com/latest/cxxfusion/index.html