

Multi-Robot Relative Pose Estimation in SE(2) with Observability Analysis: A Comparison of Extended Kalman Filtering and Robust Pose Graph Optimization

Gihun Shin, Hyunjae Sim, Seungwon Nam, Yonghee Kim, Jae Heo and Kwang-Ki K. Kim*

Abstract—In this study, we address challenges in multi-robot localization, with a particular focus on cooperative localization and the observability analysis of relative pose estimation. Cooperative localization enhances each robot’s information accuracy through communication networks and message passing. When odometry data from a target robot can be transmitted to an ego robot, the observability of their relative pose estimation can be achieved using either range-only or bearing-only measurements, provided that both robots have non-zero linear velocities. However, if the target robot’s odometry data is not directly transmitted and must instead be estimated by the ego robot, both range and bearing measurements are necessary to ensure observability. This research validates the feasibility of relative pose estimation in ground-based multi-robot systems by exploring six different sensing and communication structures. The subsequent discussion compares and analyzes the estimation methods employed and proposes models for potential application. Additionally, the results of simulations and real-world experiments with robots are presented and discussed. In ROS/Gazebo simulations, we compare the estimation accuracy of extended Kalman filter (EKF) and pose graph optimization (PGO) methods, incorporating different robust loss functions through filtering and smoothing techniques, with varying sliding window batch sizes. For hardware experiments, two Turtlebot3 robots equipped with UWB modules were used to estimate inter-robot relative poses in real-world scenarios, employing both EKF and PGO-based methods. The real-world experiments demonstrate the practical applicability of the proposed decentralized relative pose estimation methods, which rely solely on onboard sensing without the need for inter-robot communication. The results indicate that, compared to scenarios where additional inter-robot communication data is utilized, estimation accuracy degrades by only 0 to 8%.

Code: https://github.com/iASL/MRS/MR_RPEstm

Video: <https://youtu.be/KG3AJOKxPka>

Index Terms—Autonomous mobile robots (AMRs), Range and bearing measurements, Cooperative localization, Relative pose estimation, Nonlinear observability, Observability rank condition, Extended Kalman filtering (EKF), Pose graph optimization (PGO), Angle-aided SLAM (RA-SLAM), Distributed data fusion (DDF), Multi-robot multi-target tracking, UWB.

The authors are with the department of electrical and computer engineering at Inha University, Incheon 22212, Korea.

*Corresponding author (Email: kwangki.kim@inha.ac.kr)

The first two authors, K. Shin and H. Sim, contributed equally.

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant NRF-2022R1F1A1076260 and in part by the BK21 Four Program funded by the Ministry of Education(MOE, Korea) and National Research Foundation of Korea (NRF).

I. INTRODUCTION

The field of multi-robot systems is rapidly growing, and a critical problem in the operation and control of these systems is the accurate estimation of poses, i.e., orientations and positions, of the robots. Pose estimation is fundamental for the coordination, navigation, and interaction between multiple autonomous robots, enabling them to work collaboratively or independently in complex environments.

In fully autonomous mobile robot systems (AMRS), each robot within a team must estimate both its absolute pose (position and orientation) and the relative poses of its neighboring robots in the body-fixed reference frame of the corresponding ego robot [1]–[3]. This challenge is known as multi-robot localization [4], a critical capability for AMRS in applications such as search and rescue missions, warehouse logistics, formation control, and sensor coverage [5]–[7]. As illustrated in Fig. 1, this problem involves determining both the absolute and relative poses of the robots. Of particular interest is distributed multi-robot localization, where each robot determines its ego-centric state, absolute pose, and the relative poses of neighboring robots within the ego robot’s reference coordinates.

In multi-robot state estimation, cooperative localization—encompassing both absolute and relative pose estimation through the exchange of messages, including state, odometry, and raw data—has been extensively studied [8]–[10]. Centralized cooperative localization in autonomous multi-robot systems (AMRS) estimates the absolute poses of all robots by utilizing both relative measurements between robots and absolute measurements from landmarks. In contrast, distributed cooperative localization allows each ego robot in an AMRS to determine both the absolute and relative poses of its neighboring robots.

Real-time distributed relative pose estimation is essential for the cooperative control of multi-robot systems [3], [11], [12]. Various sensors, such as cameras, LiDAR, and UWB, along with estimation methods like maximum likelihood estimation (MLE), Kalman filtering (KF), and factor graph optimization (FGO), have been employed for relative pose estimation in these systems. Notable examples include vision-based relative pose estimation using manifold optimization [13], [14] and LiDAR-based relative pose estimation utilizing point-cloud registration with an adaptive cubature split-covariance inter-

section filter [15].

Observability is a fundamental property in robotic perception, crucial for state estimation in dynamic systems. For state estimation in mobile multi-robot systems, the theory of nonlinear observability [16]–[18] is applied to analyze observability in mobile robot localization [19]–[21] and multi-robot localization using various types of sensor measurements, including (i) range and bearing measurements [22]–[25], (ii) range measurements [26], and (iii) bearing measurements [27]. Observability analysis for visual-inertial navigation systems (VINS) in mobile robots has also been extensively studied [28]–[30], [30]–[34], with observability-constrained navigation [35]–[37] presenting a key challenge due to the limited field of view in visual perception. Beyond addressing the observability of multi-robot systems with known kinematic or dynamic inputs, extending the observability rank condition to nonlinear systems driven by unknown inputs remains a complex task [38], [39].

The contributions of this study are outlined as follows:

- *Theoretical Insights on Observability:* This paper provides a comprehensive overview of nonlinear observability analysis for wheeled mobile robots operating in SE(2). For simplicity, we focus on two-robot systems. Building on the work of Martinelli [22], we establish that when both robots have non-zero linear velocities, either range-only or bearing-only measurements can ensure observability in relative pose estimation, provided that each robot's odometry data is available to the other. We extend this analysis to scenarios where the odometry data of the neighboring robot is not available to the ego robot.
- *Communication-Independent Relative Pose Estimation:* While previous work [22]–[27] addresses cooperative robot-to-robot relative pose estimation under the assumption of a reliable communication network, this assumption proves impractical in many real-world applications. In dynamic or harsh environments, communication networks often suffer from instability, leading to breakdowns in coordination or inefficient and suboptimal pose estimation. Such environments, where communication is unreliable or intermittent, require robust methods capable of functioning independently of a communication network. Our work tackles this limitation by proposing a method that can be applied to both cooperative and non-cooperative robot-to-robot relative pose estimation, offering enhanced flexibility and resilience. By eliminating the dependence on constant communication, our approach is designed to maintain effectiveness in a broader range of operational environments, whether communication is present or absent, thus improving the autonomy and robustness of multi-robot systems in complex scenarios.
- *Performance Validation:* The observability analysis presented in this paper is validated through ROS/Gazebo simulation environments. We apply and compare two widely-used state estimation methods—Extended Kalman Filter (EKF) and Pose Graph Optimization (PGO)—across six distinct information structures.
- *Real-World Hardware Experiments:* We further test the observability analysis and estimation methods using real-

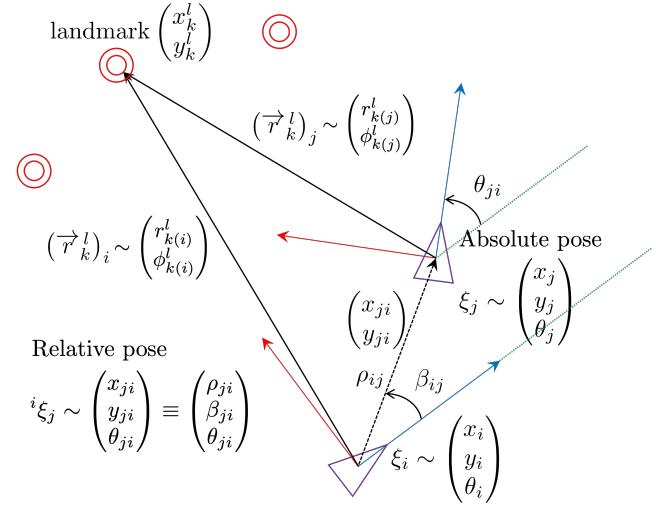


Fig. 1. Schematic for multi-robot localization: (a) (Centralized) multi-robot localization is to estimate the set of *absolute* poses $\{\xi_i\}_{i \in \mathcal{N}}$ for multiple robots in the set \mathcal{N} ; (b) Distributed multi-robot localization is to estimate the state $\{\xi_i, \{^i\xi_j\}_{j \in \mathcal{N}_i}\}$ concatenating the absolute pose of the ego-robot and the *relative* poses of the neighborhood-robots for each Robot $R_i, i \in \mathcal{N}$ where the sets of robots \mathcal{N} and \mathcal{N}_i (for $i \in \mathcal{N}$) could be time-varying; and (c) Cooperative localization is to solve (a) or (b) by communicating extra information such as odometry data and state estimate.

world hardware, specifically two Turtlebot3 robots. These experiments demonstrate the feasibility of EKF and PGO-based robot-to-robot relative pose estimation in the absence of inter-robot data communication. The results indicate accuracy differences ranging from 0 to 8%.

The remainder of this paper is organized as follows: Section II reviews key methods in multi-robot state estimation, focusing on localization and mapping techniques using distributed Extended Kalman Filter (EKF) and Pose Graph Optimization (PGO) methods. Section III provides an in-depth observability analysis for inter-robot relative pose estimation under varying measurement and information structures. In Section IV, we demonstrate and compare EKF and PGO-based estimation methods across six distinct information structures, including scenarios with outliers, using simulations in Gazebo. Section V presents real-world experimental results, evaluating the performance of EKF and PGO-based methods in hardware tests. Section VI explores future research directions and potential applications of our work in multi-robot systems. Finally, Section VII concludes the paper.

II. MULTI-ROBOT LOCALIZATION

We examine two mobile ground robots, denoted as R_i and R_j , situated in a 2D workspace. The configuration state of this multiple-robot system can be represented as the vector $X = [x_i, y_i, \theta_i, x_j, y_j, \theta_j]^T \in \text{SE}(2) \times \text{SE}(2)$ that includes the Cartesian absolute coordinates and absolute orientation of the two robots. For Robot R_i , the state of interest is a combination of its absolute pose $X_i^a = [x_i, y_i, \theta_i] \in \text{SE}(2)$ and the relative pose of the Robot R_j with respect to the ego-Robot R_i that can be represented by either $X_{ji}^c = [x_{ji}, y_{ji}, \theta_{ji}]^T$ in the cartesian coordinates with the atan2 orientation or $X_{ji}^p = [\rho_{ji}, \beta_{ji}, \theta_{ji}]^T$ in the polar coordinates with the atan2

orientation. The augmented state for the robot R_i is defined as the concatenation of the absolute and relative poses, that is, $X_i = [(X_i^a)^\top, X_{ji}^\top]^\top \in \text{SE}(2) \times \text{SE}(2)$.

The two direction-related words, bearing and orientation, are similar but used for different meanings. In this paper, as shown in Fig. 1, bearing angle denoted by β refers to the positional angle of an object in the two-dimensional polar coordinates, whereas an orientation angle denoted by θ is defined by the directional cosine of two different coordinate frames in 2D space.

A. Motion model

The dynamics of the augmented state X_i can be described by the following dynamic system equation:

$$\dot{X}_i = f_i(X_i, U_i) \quad (1)$$

where $U_i = [v_i, \omega_i, v_j, \omega_j]^\top$ denotes the control input vector. For a unicycle vehicle model, the analytical expression for the vector field $f_i : \text{SE}(2)^2 \times \mathbb{R}^4$ is as follows:

$$f_i(X_i^a, X_{ji}, U_i) = \begin{bmatrix} f_i^a(X_i^a, U_i) \\ f_{ji}^r(X_{ji}, U_i) \end{bmatrix} \quad (2)$$

where the absolute pose kinematics are given by

$$\dot{X}_i^a = f_i^a(X_i^a, U_i) = \begin{bmatrix} v_i \cos \theta_i \\ v_i \sin \theta_i \\ \omega_i \end{bmatrix} \quad (3)$$

and the relative pose kinematics are expressed as

$$\dot{X}_{ji}^c = f_{ji}^r(X_{ji}^c, U_i) = \begin{bmatrix} v_j \cos \theta_{ji} \\ v_j \sin \theta_{ji} \\ \omega_j \end{bmatrix} + \begin{bmatrix} y_{ji} \omega_i - v_i \\ -x_{ji} \omega_i \\ -\omega_i \end{bmatrix} \quad (4)$$

for $X_{ji} = X_{ji}^c = [x_{ji}, y_{ji}, \theta_{ji}]^\top$ or

$$\dot{X}_{ji}^p = f_{ji}^r(X_{ji}^p, U_i) = \begin{bmatrix} v_j \cos(\theta_{ji} - \beta_{ji}) - v_i \cos \beta_{ji} \\ v_j (\sin(\theta_{ji} - \beta_{ji}) - \sin \beta_{ji}) / \rho_{ji} \\ \omega_j - \omega_i \end{bmatrix} \quad (5)$$

for $X_{ji} = X_{ji}^p = [\rho_{ji}, \beta_{ji}, \theta_{ji}]^\top$.

In this study, we consider both the Cartesian (4) and polar (5) coordinate systems to define the configuration state vector for (cooperative) localization. For notational convenience, we use $c(\cdot)$ for $\cos(\cdot)$ and $s(\cdot)$ for $\sin(\cdot)$.

[Case M-I] With communicating odometry data Define the state vector $X_i = [x_i, y_i, \theta_i, x_{ji}, y_{ji}, \theta_{ji}]^\top$ or $X_i = [x_i, y_i, \theta_i, \rho_{ji}, \beta_{ji}, \theta_{ji}]^\top$ with the odometry vector $U_i = [u_i^\top, u_j^\top]^\top = [v_i, \omega_i, v_j, \omega_j]^\top$.

$$\dot{X}_i = \sum_{k=1}^4 g_k(X_i) U_{ik} + W_i \quad (6)$$

where $g_k(\cdot)$ is either $g_k^r(\cdot)$ defined as

$$g_1^r = \begin{bmatrix} c(\theta_i) \\ s(\theta_i) \\ 0 \\ -1 \\ 0 \\ 0 \end{bmatrix}, g_2^r = \begin{bmatrix} 0 \\ 0 \\ 1 \\ y_{ji} \\ -x_{ji} \\ -1 \end{bmatrix}, g_3^r = \begin{bmatrix} 0 \\ 0 \\ 0 \\ c(\theta_{ji}) \\ s(\theta_{ji}) \\ 0 \end{bmatrix}, g_4^r = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (7)$$

for $X_i = [x_i, y_i, \theta_i, x_{ji}, y_{ji}, \theta_{ji}]^\top$, that is, the relative pose of the rectangular coordinates. or $g_k^p(\cdot)$ defined as

$$g_1^p = \begin{bmatrix} c(\theta_i) \\ s(\theta_i) \\ 0 \\ -c(\beta_{ji}) \\ -\frac{s(\beta_{ji})}{\rho_{ji}} \\ 0 \end{bmatrix}, g_2^p = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}, g_3^p = \begin{bmatrix} 0 \\ 0 \\ 0 \\ c(\psi_{ji}) \\ \frac{s(\psi_{ji})}{\rho_{ji}} \\ 0 \end{bmatrix}, g_4^p = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (8)$$

where $X_i = [x_i, y_i, \theta_i, \rho_{ji}, \beta_{ji}, \theta_{ji}]^\top$. Here, $\psi_{ji} = \theta_{ji} - \beta_{ji}$, that is, the relative pose of the polar coordinates. The process noise W_i is assumed to be a Gaussian random process for Kalman filtering-based state estimation.

[Case M-II] Without communicating odometry data Define the state vector $X_i = [x_i, y_i, \theta_i, x_{ji}, y_{ji}, \theta_{ji}, v_j, \omega_j]^\top$ or $X_i = [x_i, y_i, \theta_i, \rho_{ji}, \beta_{ji}, \theta_{ji}, v_j, \omega_j]^\top$ with the odometry vector $u_i = [v_i, \omega_i]^\top$.

$$\dot{X}_i = \sum_{k=1}^2 a_k(X_i) + \sum_{k=1}^2 b_k(X_i) u_{ik} + W_i \quad (9)$$

where $a_k(\cdot)$ and $b_k(\cdot)$ are either $a_k^r(\cdot)$ or $b_k^r(\cdot)$ defined as follows:

$$a_1^r = \begin{bmatrix} 0 \\ 0 \\ 0 \\ v_j c(\theta_{ji}) \\ v_j s(\theta_{ji}) \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, a_2^r = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \omega_j \\ 0 \\ 0 \\ 0 \end{bmatrix}, b_1^r = \begin{bmatrix} c(\theta_i) \\ s(\theta_i) \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, b_2^r = \begin{bmatrix} 0 \\ 0 \\ 1 \\ y_{ji} \\ -x_{ji} \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (10)$$

for $X_i = [x_i, y_i, \theta_i, x_{ji}, y_{ji}, \theta_{ji}, v_j, \omega_j]^\top$, that is, the relative pose of the rectangular coordinates. or $a_k^p(\cdot)$ and $b_k^p(\cdot)$ defined as

$$a_1^p = \begin{bmatrix} 0 \\ 0 \\ 0 \\ v_j c(\psi_{ji}) \\ v_j \frac{s(\psi_{ji})}{\rho_{ji}} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, a_2^p = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \omega_j \\ 0 \\ 0 \\ 0 \end{bmatrix}, b_1^p = \begin{bmatrix} c(\theta_i) \\ s(\theta_i) \\ 0 \\ -c(\beta_{ji}) \\ -\frac{s(\beta_{ji})}{\rho_{ji}} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, b_2^p = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (11)$$

for $X_i = [x_i, y_i, \theta_i, \rho_{ji}, \beta_{ji}, \theta_{ji}, v_j, \omega_j]^\top$, that is, the relative pose of the polar coordinates. The process noise W_i is assumed to follow a Gaussian random process, a common assumption in Kalman filtering-based state estimation. Furthermore, we posit that the kinematics of a neighboring robot are unknown to the ego robot in the context of relative pose estimation. Brownian motion is employed as the motion model, wherein the time derivatives of the linear and angular velocities are considered as white Gaussian noise.

Remark 1: More rigorously, the odometry noise and process noise (or disturbance) can be separately modeled as the following generalization of (6) and (9):

$$\dot{X}_i = a(X_i) + b(X_i)(U_i + W_i^u) + W_i^d \quad (12)$$

Here, W_i^u denotes the odometry noise and W_i^d corresponds to the combination of model uncertainty and external disturbances. In this study, the noise is modeled as lumped noise $W_i := b(X_i)W_i^u + W_i^d$. However, the separated noise model described above can also be applied for state estimation using Kalman filtering (KF) and pose-graph optimization (PGO).

B. Observation model

Three types of sensing and communication information are available for state estimation:

- proprioceptive sensors (encoders, etc.);
- exteroceptive sensors (LiDAR, Camera, UWB, etc.);
- communication network (V2X).

We consider the following measurement or sensor model:

$$Y_i = h_i(X_i) + V_i$$

where $h_i(\cdot)$ defines the relationship between the unknown state and the measurements, and V_i is the measurement noise.

1) *Measurements of absolute pose using landmark observation model:* The measurement model relates the current (absolute) pose of the ego robot to its LiDAR range and bearing measurements $[r_{k(i)}^l, \phi_{k(i)}^l]^\top$ for $k \in \mathcal{N}_i^l$

$$\begin{aligned} Y_{ki}^l &= h_{ki}^l(X_i, X_k^l) + V_{ki}^l \\ &= \begin{bmatrix} \sqrt{(x_k^l - x_i - d_i c\theta_i)^2 + (y_k^l - y_i - d_i s\theta_i)^2} \\ \text{atan2}(y_k^l - y_i - d_i s\theta_i, x_k^l - x_i - d_i c\theta_i) - \theta_i \end{bmatrix} + V_{ki}^l \end{aligned} \quad (13)$$

where x_k^l and y_k^l are the ground truth coordinates of the landmark k that can be observed by Robot R_i , x_i and y_i and θ_i represent the current pose of Robot R_i , and d_i is the known distance between robot center and laser rangefinder (LiDAR). For brevity, we use $c\theta_i = \cos \theta_i$ and $s\theta_i = \sin \theta_i$. The landmark measurement noise V_{ki}^l is assumed to be a Gaussian random process for Kalman filtering-based state estimation.

2) *Measurements of relative pose using range, bearing, and orientation observation models of neighboring robots:* The measurement model relates the current relative pose of robot R_j with respect to robot R_i , which is denoted as ${}^i\xi_j$, to the range, bearing, and orientation measurements $[\rho_{ji}, \beta_{ji}, \theta_{ji}]^\top$:

$$Y_{ji}^{rbo} = h_{ji}^{rbo}(X_i) + V_{ji}^{rbo} = \begin{bmatrix} Y_{ji}^r \\ Y_{ji}^b \\ Y_{ji}^o \end{bmatrix} + V_{ji}^{rbo} \quad (14)$$

where $Y_{ji}^r = h_{ji}^r(X_i) = \rho_{ji}$, $Y_{ji}^b = h_{ji}^b(X_i) = \beta_{ji}$, $Y_{ji}^o = h_{ji}^o(X_i) = \theta_{ji}$. The measurement function can be rewritten as follows: $\rho_{ji} = \sqrt{x_{ji}^2 + y_{ji}^2} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}$ and $\beta_{ji} = \text{atan2}(y_{ji}, x_{ji}) = \text{atan2}(y_j - y_i, x_j - x_i) - \theta_i$ in Cartesian coordinates. The measurement noise V_{ji}^{rbo} is assumed to be a Gaussian random process for Kalman filtering-based state estimation. Note that the two direction-related measurements, bearing β_{ji} and orientation θ_{ji} , refer to the relative positional and rotational angles, respectively, of the target object indexed by j with respect to the ego-robot R_i 's reference direction (e.g., the ego-robot's heading direction). In contrast, the absolute rotation angle θ_i or θ_j of a robot indexed by i or j is defined

with respect to a specific reference coordinate system (e.g., the global coordinate system).

Remark 2: The measurement variables that relate the current relative pose of the Robot R_i with respect to the Robot R_j satisfy the relations $\rho_{ij} = \rho_{ji}$, $\beta_{ij} = \text{atan2}(\sin \psi_{ji}, \cos \psi_{ji})$, and $\theta_{ij} = -\theta_{ji}$, where $\psi_{ji} = \theta_{ji} - \beta_{ji}$.

3) *Communication of odometry data:* The wheel-odometry data $u_j = [v_j, \omega_j]^\top$ of a neighboring robot R_j are not directly measured but can be transmitted via a communication network.

$$Y_{ji}^{od} = h_{ji}^{od}(u_j) = \begin{bmatrix} v_j \\ \omega_j \end{bmatrix} + V_{ji}^{od} \quad (15)$$

where V_{ji}^{od} corresponds to a combination of communication and measurement noise, which is also assumed to be a Gaussian random process for Kalman filtering-based state estimation.

C. Message passing model

Let $h_{ji}^{mp}(t)$ denote the message passing from robot R_j to robot R_i at time t . Assume that this message (set) $m_{ji}(\cdot)$ belongs to the power set $2^{\mathcal{M}_{ji}}$ with the set $\mathcal{M}_{ji} = \{x_j, y_j, \theta_j, \rho_{ij}, \beta_{ij}, \theta_{ij}, v_j, \omega_j\} = X_j \cup U_j$. For the robot state estimation, this message provides additional information as a model.

$$Y_{ji}^{mp} = h_{ji}^{mp}(X_j, U_j) + V_{ji}^{mp} \quad (16)$$

where V_{ji}^{mp} is the communication noise or uncertainty, which can also be assumed to be a Gaussian random process.

D. Extended Kalman filtering for multi-robot localization

Research has explored extensions of EKF-based single-robot state estimation to multi-robot environments, including distributed EKF-based approaches and SLAM [40], [41]. Roumeliotis et al. [4] adapted the centralized EKF into a decentralized form, where each robot runs a local EKF independently. In this setup, the Kalman gain is determined based on the robot's own measurements and intermittent communication with neighboring robots, which includes their states and measurements.

In the context of multi-robot cooperative localization, the Distributed Kalman Filter (KF) enhances collaboration compared to centralized KF by using consensus algorithms to achieve a common state estimate from local observations. The Multi-State Constraint Kalman Filter (MSCKF) framework [42], originally developed for cooperative visual-inertial odometry (VIO), has recently been applied to filter-based cooperative multi-robot localization [43], [44]. Additionally, distributed MSCKF models that incorporate inter-robot ranging information have been explored [45], [46]. For foundational theoretical and algorithmic developments of distributed KF, readers are directed to [47]–[51].

For the robotic state estimation, we consider a discrete-time version of the motion model

$$\begin{aligned} X_i(t) &= f_i^{\text{dt}}(\hat{X}_i(t-1), U_i(t)) \\ &= a_i^{\text{dt}}(X_i(t-1)) + b_i^{\text{dt}}(X_i(t-1))U_i(t) + W_i(t) \end{aligned} \quad (17)$$

and the associated sampled observation model.

$$Y_i(t) = h_i(X_i(t), t) + V_i(t) \quad (18)$$

It is crucial to note that the measurements available at time t are usually time-varying, given that the set of landmarks $\mathcal{N}_i^l(t)$ and the set of neighboring robots $\mathcal{N}_i(t)$ are subject to change over time. This temporal variation implies that the observation models in the state estimation explicitly depend on time.

1) *Decentralized EKF*: Multi-robot localization using the decentralized EKF framework, as presented in [4], can be summarized as follows:

▷ Prediction

$$\begin{aligned}\check{X}_i(t) &= f_i^{\text{dt}}(\hat{X}_i(t-1), U_i(t)) \\ \check{P}_i(t) &= F_i(t-1)\hat{P}_i(t-1)F_i^{\top}(t-1) + \Sigma_{w,i}(t) \\ \check{Y}_i(t) &= h_i(\check{X}_i(t), t)\end{aligned}\quad (19)$$

where $F_i = \partial f_i^{\text{dt}} / \partial X_i$ denotes the Jacobian matrix of the vector field f_i^{dt} evaluated at $(\hat{X}_i(t-1), U_i(t))$. A positive definite matrix $\hat{P}_i(t-1)$ is the error covariance matrix of the posterior probability of the state at the previous time step and $\Sigma_{w,i}(t)$ is the second moment of the probabilistic disturbance $W_i(t)$. The observation model h_i relates both robot R_i's onboard sensor measurements and messages passed from neighbors $\mathcal{N}_i(t)$ to robot R_i's state vector.

▷ Computation of Kalman gains

$$K_i(t) = \check{P}_i(t)H_i^{\top}(t)(H_i(t)\check{P}_i(t)H_i^{\top}(t) + \Sigma_{v,i}(t))^{-1} \quad (20)$$

where $H_i = \partial h_i / \partial X_i$ is the Jacobian matrix of the measurement function h_i evaluated at $\check{X}_i(t)$ which is the predicted mean of the state, and $\Sigma_{v,i}(t)$ is the second moment of random process noise $V_i(t)$.

▷ Correction

$$\begin{aligned}\hat{X}_i(t) &= \check{X}_i(t) + K_i(t)(Y_i(t) - \check{Y}_i(t)) \\ \hat{P}_i(t) &= (I - K_i(t)H_i(t))\check{P}_i(t)\end{aligned}\quad (21)$$

where $Y_i(t)$ denotes a noisy measurement vector encompassing both the onboard sensing and communication information. The resulting probabilistic inference of the state estimate is assumed to follow a Gaussian random process.

$$X_i(t) \sim \mathcal{N}(\hat{X}_i(t), \hat{P}_i(t)). \quad (22)$$

2) *Distributed EKF*: Similar to but different from the decentralized EKF, the distributed EKF takes into account coupled constraints and strives for consensus among robots when there are overlapping state variables of estimation¹. The consensus Kalman filtering, as presented in [47], [48], can be summarized as follows:

▷ Prediction-consensus

$$\begin{aligned}\hat{X}_i(t) &= \check{X}_i(t) + K_i(t)(Y_i(t) - \check{Y}_i(t)) \\ &\quad + \sum_{j \in \mathcal{N}_i(t)} \hat{K}_{ji}(t)(C_{ji}\check{X}_i(t) - C_{ij}\check{X}_j(t))\end{aligned}\quad (23)$$

where $K_i(t)$ denotes the Kalman gain corresponding to the on-board sensing and $\hat{K}_{ji}(t)$ is the distributed Kalman gain

¹In this paper, the distributed EKF incorporates consensus constraints explicitly, as described by (23) and (24). In contrast, the decentralized EKF relies on information exchange only as specified in (19) and (21), and does not require consensus on state estimates among neighboring robots. For a detailed discussion on consensus Kalman filtering, see [47], [48].

Function	Loss $\ell(e)$	Weight $\gamma(e)$
L_2	$\frac{e^2}{2}$	1
Laplace	$t e $	$\frac{t}{ e }$
Huber	$\begin{cases} \frac{e^2}{2} & \text{for } e \leq t \\ t(e - t/2) & \text{o.w.} \end{cases}$	$\begin{cases} 1 & \text{for } e \leq t \\ \frac{1}{ e } & \text{o.w.} \end{cases}$
Cauchy	$\frac{t^2}{2} \ln(1 + \frac{e^2}{t^2})$	$\frac{t^2}{t^2 + e^2}$
Fair	$t^2 \left(\frac{t}{ e } - \ln(1 + \frac{t}{ e }) \right)$	$\frac{t}{t + e }$
Geman-McClure	$\frac{e^2}{2(t + e^2)}$	$\frac{t^2}{(t + e^2)^2}$
Welsch	$\frac{t^2}{2} (1 - \exp(-\frac{e^2}{t^2}))$	$\exp(-\frac{e^2}{t^2})$
Switchable-Constraint	$\begin{cases} \frac{e^2}{2} & \text{for } e^2 \leq t \\ \frac{2te^2}{t + e^2} - \frac{t}{2} & \text{o.w.} \end{cases}$	$\begin{cases} 1 & \text{for } e^2 \leq t \\ \frac{4t^2}{(t + e^2)^2} & \text{o.w.} \end{cases}$
Tukey	$\begin{cases} \frac{t^2(1 - (1 - \frac{e^2}{t^2})^3)}{2} & \text{for } e \leq t \\ \frac{t^2}{2} & \text{o.w.} \end{cases}$	$\begin{cases} (1 - \frac{e^2}{t^2})^2 & \text{for } e \leq t \\ 0 & \text{o.w.} \end{cases}$
Max.Dist.	$\begin{cases} \frac{e^2}{2} & \text{for } e \leq t \\ \frac{t^2}{2} & \text{o.w.} \end{cases}$	$\begin{cases} 1 & \text{for } e \leq t \\ 0 & \text{o.w.} \end{cases}$

TABLE I
A LIST OF ROBUST LOSS FUNCTIONS FOR M-ESTIMATION [53], [54].

corresponding to the consensus constraints of *predicted* relative pose, $\check{\rho}_{ij}(t) = \check{\rho}_{ji}(t)$, $\check{\beta}_{ij}(t) = \check{\theta}_{ji}(t) - \check{\beta}_{ji}(t)$, and $\check{\theta}_{ij}(t) = -\check{\theta}_{ji}(t)$, in the polar coordinates. Matrices C_{ji} and C_{ij} are defined by following the consensus constraints of the relative poses X_{ji} and X_{ij} .

▷ Correction-consensus

$$\begin{aligned}\hat{X}_i(t) &= \hat{X}_i(t-1) \\ &\quad + \sum_{j \in \mathcal{N}_i(t)} \hat{K}_{ji}(t)(C_{ji}\hat{X}_i(t-1) - C_{ij}\hat{X}_j(t-1))\end{aligned}\quad (24)$$

where $\hat{K}_{ji}(t)$ is the distributed Kalman gain corresponding to the consensus constraints of *corrected* relative pose, $\hat{\rho}_{ij}(t) = \hat{\rho}_{ji}(t)$, $\hat{\beta}_{ij}(t) = \hat{\theta}_{ji}(t) - \hat{\beta}_{ji}(t)$, and $\hat{\theta}_{ij}(t) = -\hat{\theta}_{ji}(t)$, in the polar coordinates. The update (24) is indeed iterative, whereas the update described in (23) does not require iterations. Consequently, it is important to analyze the convergence of the iterative process in (24). Since this process is linear, eigenvalue or spectral analysis can be employed to investigate its convergence [52].

E. Optimization-based state estimation

1) *Nonlinear least squares methods*: Consider online optimization methods to estimate the state variables of a nonlinear Markov process in the following form:

$$\begin{aligned}X_i(t) &= f_i(X(t-1)) + W_i(t), \\ Y_i(t) &= h_i(X(t)) + V_i(t)\end{aligned}\quad (25)$$

where $X_i \in \mathbb{R}^n$ refers to the state of the system $Y_i \in \mathbb{R}^m$ denotes measured output. $W_i \in \mathbb{R}^n$ denotes the disturbance and $V_i \in \mathbb{R}^m$ denotes measurement noise. Here, the odometry-dependence of f_i is hidden for simplicity. We further assume

that the state, disturbance, and noise belong to compact convex constraint sets: $\mathbb{X}_i(t) \subseteq \mathbb{R}^n$, $\mathbb{W}_i(t) \subseteq \mathbb{R}^n$, and $\mathbb{V}_i(t) \subseteq \mathbb{R}^m$, respectively, which can be represented as (linear or quadratic) inequalities.

The estimation problem with full information on T measurement sequences $\{Y_i(1), \dots, Y_i(T)\}$ is represented by the following nonlinear programming (NLP):

$$\begin{aligned} & \underset{\{X_i, W_i, V_i\}}{\text{minimize}} \quad q_0(X_i(0)) + \sum_{t=1}^T q(X_i(t-1), X_i(t)) \\ & \text{subject to } X_i(t) = f_i(X_i(t-1)) + W_i(t), \\ & \quad Y_i(t) = h_i(X_i(t)) + V_i(t), \\ & \quad X_i(t) \in \mathbb{X}_i(t), W_i(t) \in \mathbb{W}_i(t), V_i(t) \in \mathbb{V}_i(t) \end{aligned} \quad (26)$$

where T refers to the current time step and the stage loss function is defined as

$$\begin{aligned} & q(X_i(t-1), X_i(t)) \\ & = (X_i(t) - f_i(X_i(t-1)))^\top \Psi (X_i(t) - f_i(X_i(t-1))) \\ & \quad + (Y_i(t) - h_i(X_i(t)))^\top \Phi (Y_i(t) - h_i(X_i(t))) \\ & = W_i^\top(t) \Psi W_i(t) + V_i^\top(t) \Phi V_i(t) \\ & = q_w(W_i(t)) + q_v(V_i(t)) \end{aligned} \quad (27)$$

and $q_0(\cdot)$ denotes the arrival cost function of the initial state $X_i(0) \in \mathbb{R}^n$.

The arrival cost function $q_0 : \mathbb{X}_i(0) \rightarrow \mathbb{R}$ is used to summarize the prior information or statistics regarding the state at time $t = 0$, and it is assumed to satisfy $q_0(\bar{X}_{i,0}) = 0$ and $q_0(X_{i,0}) > 0$ for all $X_{i,0} \neq \bar{X}_{i,0}$ where $\bar{X}_{i,0}$ denotes the best state estimate at time $t = 0$. An example of the arrival cost $q_0(\cdot)$ is the quadratic form $q_0(X_{i,0}) = (X_{i,0} - \bar{X}_{i,0})^\top P_{i,0}^{-1} (X_{i,0} - \bar{X}_{i,0})$ where $P_{i,0}$ is a positive definite matrix corresponding to the error covariance of the initial state $X_i(0) \sim \mathcal{GP}(\bar{X}_{i,0}, P_{i,0})$.

2) Robust M-estimation methods: Instead of using quadratic forms for the loss functions $q_0(\cdot)$ and $q(\cdot, \cdot)$, robust loss functions for M estimation can also be used.

$$\begin{aligned} \tilde{q}_0(X_{i,0}) &= \ell\left(\sqrt{q_0(X_{i,0})}\right), \\ \tilde{q}(W_i(t), V_i(t)) &= \ell\left(\sqrt{q_w(W_i(t))}\right) + \ell\left(\sqrt{q_v(V_i(t))}\right), \end{aligned} \quad (28)$$

where the function $\ell : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is selected as the loss function as shown in Table I. Owing to the nonconvexity of the *robust* loss functions listed in Table I, the resulting nonlinear least-squares problem should be solved iteratively. The iteratively reweighted least-squares (IRLS) method [55], [56] can be used to approximate the loss function $\ell(e(X))$:

$$\ell(e(X)) \approx \gamma(e(\check{X})) e^2(X) \quad (29)$$

and iteratively updated as follows:

$$\hat{X} = \arg \min_{X \in \mathbb{X}} \sum_{t=1}^{T+1} \gamma(e(\check{X}(t-1))) e^2(X(t-1)) \quad (30)$$

and repeat (30) with an update $\check{X} = \hat{X}$, where $\check{X} = (\check{X}(t))_{t=0}^T$ and $\hat{X} = (\hat{X}(t))_{t=0}^T$ are the previous and current guesses of the estimate, respectively. The function $\gamma : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is the weight associated with the loss function. The relations between several loss and weight functions are also listed in Table I.

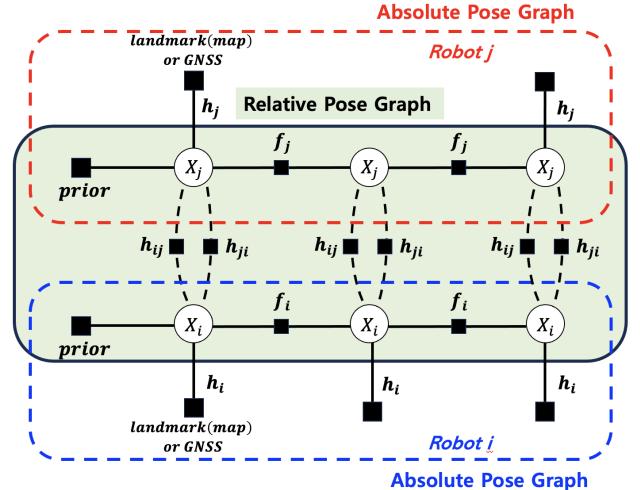


Fig. 2. Factor graph representation of multi-robot absolute and relative pose estimation.

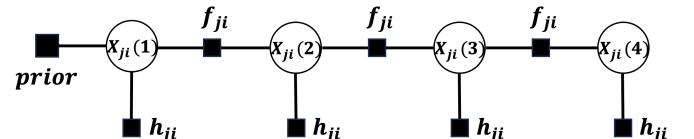


Fig. 3. Factor graph representation of inter-robot relative pose estimation: f_{ji} and h_{ji} refer to the kinematics and measurement functions associated with X_{ji} that is the relative pose of Robot R_j with respect to Robot R_i.

3) Factor graph optimization: Factor graph optimization is widely employed in optimization-based robot state estimation [57]–[59], particularly in Pose Graph Optimization (PGO) as a back-end for Simultaneous Localization and Mapping (SLAM) [60], [61].

- Fig. 2 shows a factor graph representation corresponding to multi-robot absolute and relative pose estimation, in which the absolute poses (X_i, X_j), and the relative pose between the robots (X_{ji}, X_{ij}) should be estimated by using the relations between temporal kinematics (f_i, f_j), ego-pose measurements (h_i, h_j), and inter-robot relative measurements (h_{ji}, h_{ij}))
- Inter-robot relative pose estimation (shaded in Fig. 2) can be equivalently represented in the factor graph of Fig. 3.

For graph optimization for distributed multi-robot localization, consider the three types of error functions: The function associated with the state-transition (forward-dynamics) factor is defined as

$$e_i^f(X_i, X'_i) = (X_i - f_i(X'_i, U_i))^\top \Sigma_{w,i}^{-1} (X_i - f_i(X'_i, U_i)), \quad (31)$$

the function associated with the map-based absolute position measurements is given by

$$\begin{aligned} e_{ki}^l(X_i, X_k^l) \\ = (Y_{ki}^l - h_{ki}^l(X_i, X_k^l))^\top \Sigma_{l,ki}^{-1} (Y_{ki}^l - h_{ki}^l(X_i, X_k^l)), \end{aligned} \quad (32)$$

and the function corresponding to the robot-to-robot measurements and message passing is defined as

$$\begin{aligned} e_{ji}^{r2r}(X_i, X_j) \\ = (Y_{ji}^{r2r} - h_{ji}^{r2r}(X_{ji}))^\top \Sigma_{r2r,ki}^{-1} (Y_{ji}^{r2r} - h_{ji}^{r2r}(X_{ji})) \end{aligned} \quad (33)$$

where the robot-to-robot measurement vector involves concatenating the range-bearing-orientation measurements and the odometry data for the target or neighboring Robot R_j , denoted as $Y_{ji}^{r2r} = (Y_{ji}^{rbo}, Y_{ji}^{od})^\top$, and the function $h_{ji}^{r2r}(\cdot)$ is defined accordingly.

The associated optimization problem of absolute pose estimation is defined as follows:

$$\begin{aligned} \min_{(X_i(t))_{t=1}^T} \quad & p(X_i(0)) + \sum_{t=1}^T e_i^f(X_i(t), X_i(t-1)) \\ & + \sum_{t=1}^T \sum_{k \in \mathcal{N}_i^l(t)} e_{ki}^l(X_i(t), X_k^l) \\ & + \sum_{t=1}^T \sum_{j \in \mathcal{N}_i(t)} e_{ji}^{r2r}(X_{ji}(t)). \end{aligned} \quad (34)$$

Similar to the absolute pose estimation (34), the relative pose estimation between the two robots R_i and R_j can be represented as the following optimization problem:

$$\begin{aligned} \min_{(X_{ji}(t))_{t=1}^T} \quad & p(X_{ji}(0)) + \sum_{t=1}^T e_i^f(X_{ji}(t), X_{ji}(t-1)) \\ & + \sum_{t=1}^T e_{ji}^{r2r}(X_{ji}(t)) \end{aligned} \quad (35)$$

where $j \in \mathcal{N}_i(t)$ is assumed for all $t = 1, \dots, T$.

III. OBSERVABILITY IN RELATIVE POSE ESTIMATION

This section summarizes existing observability analyses for inter-robot planar relative pose estimation within ground mobile robots. It also introduces a novel analysis for scenarios where an inter-robot communication network is unavailable. Since the motion models considered are nonlinear dynamic systems, we employ the Lie derivative-based observability rank condition. This approach includes explicit computation of state-dependent observability spaces for six different information structures, using both Polar and Cartesian coordinates to represent robot poses in $\text{SE}(2)$.

A. Observability with global odometry data

The observability rank condition, based on Lie derivatives, was initially introduced in [17]. Subsequently, it was applied to the observability analysis of multirobot relative pose estimation [22] with inter-robot range-only measurements [26] and bearing-only measurements [62]. Notably, the measurements of the angular speed odometry data (ω_i, ω_j) did not change the rank of the observation space. We have rewritten and reinterpreted the results previously presented in [22], [26].

1) *Range-only measurement*: We considered two different coordinates of the system representation for observability analysis: Polar and Cartesian coordinates.

a) *Observability analysis in the Polar coordinates*: For observability analysis of the relative pose estimation between the two mobile robots, we consider the following model equations described in polar coordinates:

- Kinematic motion model: $\dot{x} = \sum_{k=1}^4 g_k(x)u_k$ which is a control-linear system given in (6) and (8) and
- Measurement model: $y = h(x) = \rho_{ji} = x_1$ given in (14).

By applying an observability analysis method based on Lie derivatives [17], we obtain the following sequence of expanding observability subspaces (codistributions):

$$\begin{aligned} \mathcal{O}_0 &= \text{span}\{\mathcal{L}^0 h\} = \text{span}\{[1, 0, 0]^\top\} \\ \mathcal{O}_1 &= \mathcal{O}_0 + \sum_{k=1}^4 \text{span}\{\nabla \mathcal{L}_{g_k} h\} \\ &= \text{span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ s(x_2) \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ s(x_{32}) \\ -s(x_{32}) \end{bmatrix}\right\} \\ \mathcal{O}_\ell &= \mathcal{O}_1 \text{ for all } \ell \geq 2 \end{aligned} \quad (36)$$

where $x_{32} = x_3 - x_2$ is the angular difference between the orientation and bearing angles of the relative pose, represented by the eco-robot's local coordinates. The rank of observability (sub-)space is given by

$$\text{rank } \mathcal{O}_1 = 3 \quad (37)$$

unless $x_2 = n\pi$ or $x_2 - x_3 = \bar{n}\pi$ for $n, \bar{n} \in \mathbb{N}$. This implies that the corresponding system is *locally observable* at almost every location.

b) *Observability analysis in the Cartesian coordinates*::

Under the same information structure but with a different coordinate system of pose representations, we consider

- Kinematic motion model: $\dot{x} = \sum_{k=1}^4 g_k(x)u_k$ which is a control-linear system represented in (6) and (7) and
- Measurement model: The distance-related model formulation $y = h(x) = \frac{1}{2}(x_1^2 + x_2^2)$ is equivalent to the model represented in (14). The introduction of this measurement function serves primarily for mathematical convenience and was considered in [26] specifically for observability analysis purposes.

Similar to the case of polar coordinates, we apply the Lie derivative-based observability analysis and obtain the following sequence of observability codistributions:

$$\begin{aligned} \mathcal{O}_0 &= \text{span}\{\mathcal{L}^0 h\} = \text{span}\{[x_1, x_2, 0]^\top\} \\ \mathcal{O}_1 &= \mathcal{O}_0 + \sum_{k=1}^4 \text{span}\{\nabla \mathcal{L}_{g_k} h\} \\ &= \text{span}\left\{\begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} c(x_3) \\ s(x_3) \\ x_2 c(x_3) - x_1 s(x_3) \end{bmatrix}\right\} \\ \mathcal{O}_\ell &= \mathcal{O}_1 \text{ for all } \ell \geq 2. \end{aligned} \quad (38)$$

The rank condition is given by

$$\text{rank } \mathcal{O}_1 = 3 \quad (39)$$

unless $x_2 = 0$ or $\text{atan}(x_2/x_1) - x_3 = n\pi$ for $n \in \mathbb{N}$. This coincides with the result obtained from the Cartesian coordinate pose representation with the same information structure.

2) *Bearing-only measurement*: Similar to range-only measurements, we consider both Polar and Cartesian coordinates.

a) *Observability analysis in the Polar coordinates::*

First, we consider the following models of relative motion and bearing-only measurements in polar coordinates:

- Kinematic motion model: $\dot{x} = \sum_{k=1}^4 g_k(x)u_k$ which is a control-linear system represented in (6) and (8) and
- Measurement model: $y = h(x) = \beta_{ji} = x_2$ represented in (14).

By applying the method of observability analysis based on Lie derivatives, we obtain a sequence of codistributions

$$\begin{aligned}\mathcal{O}_0 &= \text{span}\{\mathcal{L}^0 h\} = \text{span}\{[0, 1, 0]^\top\} \\ \mathcal{O}_1 &= \mathcal{O}_0 + \sum_{k=1}^4 \text{span}\{\nabla \mathcal{L}_{g_k} h\} \\ &= \text{span}\left\{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} s(x_2)/x_1^2 \\ -c(x_2)/x_1^2 \\ 0 \end{bmatrix}, \begin{bmatrix} s(x_{23})/x_1^2 \\ -c(x_{23})/x_1 \\ c(x_{23})/x_1 \end{bmatrix}\right\} \quad (40) \\ \mathcal{O}_\ell &= \mathcal{O}_1 \text{ for all } \ell \geq 2\end{aligned}$$

where a nontrivial displacement is assumed; that is, $x_1 \neq 0$. The rank of observability (sub)space is given by:

$$\text{rank } \mathcal{O}_1 = 3 \quad (41)$$

almost every, unless $x_2 = n\pi$ or $c(x_2) = c(x_{23}) = 0$. Therefore, we conclude that the corresponding system with bearing-only measurements is *locally observable*.

b) *Observability analysis in the Cartesian coordinates::*

Consider the following models of relative motion and bearing-only measurements in Cartesian coordinates:

- Kinematic motion model: $\dot{x} = \sum_{k=1}^4 g_k(x)u_k$ which is a control-linear system represented in (6) and (7) and
- Measurement model: $y = h(x) = \tan^{-1}(x_1/x_2)$ represented in (14).

To apply the method of observability analysis based on Lie derivatives, we computed a sequence of observability subspaces:

$$\begin{aligned}\mathcal{O}_0 &= \text{span}\{\mathcal{L}^0 h\} = \text{span}\{[x_2/d^2, -x_1/d^2, 0]^\top\} \\ \mathcal{O}_1 &= \mathcal{O}_0 + \sum_{k=1}^4 \text{span}\{\nabla \mathcal{L}_{g_k} h\} \\ &= \text{span}\left\{\begin{bmatrix} x_2/d^2 \\ -x_1/d^2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2x_1x_2)/d^4 \\ -2x_2^2)/d^4 \\ 0 \end{bmatrix}, \begin{bmatrix} -2x_1x_2c(x_3))/d^4 - 2x_1^2s(x_3)/d^4 \\ -2x_2^2c(x_3))/d^4 + c(x_3)/d^2 - 2x_1x_2s(x_3))/d^4 \\ -x_2s(x_3)/d^2 + x_1c(x_3)/d^2 \end{bmatrix}\right\} \quad (42) \\ \mathcal{O}_\ell &= \mathcal{O}_1 \text{ for all } \ell \geq 2\end{aligned}$$

where $d = \sqrt{x_1^2 + x_2^2}$ is the distance in the 2D space between the two robots. The rank of observability (sub-)space is given as

$$\text{rank } \mathcal{O}_1 = 3 \quad (43)$$

unless $x_2 = 0$ or $x_3 = n\pi$ for $n \in \mathbb{N}$, implying that the corresponding system is *locally observable*.

3) *Orientation-only measurement:* For observability analysis of a system with orientation-only measurements, either Polar or Cartesian coordinate system representations can be considered.

- Kinematic motion model: $\dot{x} = \sum_{k=1}^4 g_k(x)u_k$ which is a control-linear system represented in (6) and (7) (or (8)), and
- Measurement model: $y = h(x) = x_3$ represented in (14).

Similar to previous cases, the observability subspaces can be computed as follows:

$$\begin{aligned}\mathcal{O}_0 &= \text{span}\{\mathcal{L}^0 h\} = \text{span}\{[0, 0, 1]^\top\} \\ \mathcal{O}_\ell &= \mathcal{O}_1 \text{ for all } \ell \geq 1\end{aligned} \quad (44)$$

and the rank of observability (sub)space is.

$$\text{rank } \mathcal{O}_1 = 1, \quad (45)$$

regardless of the state x and the input u . It can be concluded that the relative position between robots cannot be estimated using only relative orientation estimation, even though the velocities are known.

B. Observability without global odometry data

In Section III-A, we assume that wheel odometry or kinematic velocity data can be shared via a communication network. However, when odometry inputs from a neighboring robot are unavailable, range-only or bearing-only measurements alone do not ensure observability in inter-robot relative pose estimation. Our study shows that with both range and bearing measurements available, observability can be guaranteed without needing to communicate odometry data. This analysis is relevant to the literature on nonlinear unknown input observability [38], [39], [63] and disturbance observers [64]–[66]. We focus on the observability analysis method introduced in Section III-A, which considers an augmented state-space model where the velocities of neighboring robots are treated as additional unknown states.

For the nonlinear observability analysis, we assume that the odometry inputs of a neighboring robot follow Brownian motion, as described in (9), using polar coordinates. Similar to Section III-A, we apply a nonlinear observability method based on Lie derivatives. The following augmented state-space model and range-bearing measurement equations are considered:

- Kinematic motion model: $\dot{x} = \bar{g}_0(x) + \sum_{k=1}^2 \bar{g}_k(x)u_k$ which is a control-affine system represented in (9) and (11), where $x = [\rho_{ji}, \beta_{ji}, \theta_{ji}, v_j, \omega_j] \in \mathbb{R}^2 \times \mathbb{S} \times \mathbb{R}^2$ and

$$\bar{g}_0(x) = \begin{bmatrix} x_4c(x_{32}) \\ s(x_{32}) \\ x_4 \\ x_1 \\ x_5 \\ 0 \\ 0 \end{bmatrix}, \quad \bar{g}_1(x) = \begin{bmatrix} -c(x_2) \\ s(x_2) \\ x_1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \bar{g}_2(x) = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

- Measurement model: $y_1 = h_1(x) = \rho_{ji} = x_1$ and $y_2 = h_2(x) = \beta_{ji} = x_2$ represented in (14).

Using the Lie derivatives, we compute the vectors spanning the state-dependent observability space as

$$O_{01} = \nabla \mathcal{L}^0 h_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, O_{02} = \nabla \mathcal{L}^0 h_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$O_{110} = \nabla \mathcal{L}_{\bar{g}_0}^1 h_1 = \nabla(\nabla h_1 \cdot \bar{g}_0) = \begin{bmatrix} 0 \\ x_4 s(x_{32}) \\ -x_4 s(x_{32}) \\ c(x_{32}) \\ 0 \end{bmatrix},$$

$$O_{120} = \nabla \mathcal{L}_{\bar{g}_0}^1 h_2 = \nabla(\nabla h_2 \cdot \bar{g}_0) = \begin{bmatrix} -\frac{x_4 s(x_{32})}{x_1^2} \\ x_4 c(x_{32}) \\ \frac{x_1}{0} \\ \frac{s(x_{32})}{x_1} \\ 0 \end{bmatrix},$$

$$O_{111} = \nabla \mathcal{L}_{\bar{g}_1}^1 h_1 = \nabla(\nabla h_1 \cdot \bar{g}_1) = \begin{bmatrix} 0 \\ s(x_2) \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$O_{121} = \nabla \mathcal{L}_{\bar{g}_1}^1 h_2 = \nabla(\nabla h_2 \cdot \bar{g}_1) = \begin{bmatrix} \frac{s(x_2)}{x_1^2} \\ \frac{c(x_2)}{x_1} \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

In addition to the first-order Lie derivatives and associated observability subspaces, we are interested in a nontrivial higher-order Lie derivative as follows:

$$O_{2100} = \mathcal{L}_{\bar{g}_0 \bar{g}_1}^2 h_1 = \nabla(\nabla \mathcal{L}_{\bar{g}_0}^1 h_1 \cdot \bar{g}_1)$$

$$= \begin{bmatrix} \frac{x_4 s^2(x_{32})}{x_1^2} \\ -\frac{x_4^2 s(x_{32}) c(x_{32})}{x_1} + x_4 x_5 c(x_{32}) \\ \frac{-x_4^2 s(x_{32}) c(x_{32})}{x_1} - x_4 x_5 c(x_{32}) \\ \frac{2 x_4 s^2(x_{32})}{x_1} - x_5 s(x_{32}) \\ -x_4 s(x_{32}) \end{bmatrix}$$

Thus, we have the following sequence of observability subspaces

$$\mathcal{O}_0 = \text{span}\{O_{01}, O_{02}\},$$

$$\mathcal{O}_1 = \mathcal{O}_0 + \text{span}\{O_{110}, O_{120}, O_{1101}, O_{121}\},$$

$$\mathcal{O}_2 \supseteq \tilde{\mathcal{O}}_2 := \mathcal{O}_1 + \text{span}\{O_{2100}\}$$

and the rank condition

$$\text{rank } \tilde{\mathcal{O}}_2 = 5 \Rightarrow \text{rank } \mathcal{O}_2 = 5,$$

Measurements	Odometry inputs				Observable states
	v_i	ω_i	v_j	ω_j	
ρ_{ji}	○	△	○	△	$(\rho_{ji}, \beta_{ji}, \theta_{ji})$
	○	△	×	△	(ρ_{ji}, β_{ji})
	×	△	○	△	(ρ_{ji}, β_{ji})
	×	△	×	△	ρ_{ji}
β_{ji}	○	△	○	△	$(\rho_{ji}, \beta_{ji}, \theta_{ji})$
	○	△	×	△	(ρ_{ji}, β_{ji})
	×	△	○	△	(ρ_{ji}, β_{ji})
	×	△	×	△	β_{ji}
θ_{ji}	△	△	△	△	θ_{ji}
(ρ_{ji}, β_{ji})	○	△	△	△	$(\rho_{ji}, \beta_{ji}, \theta_{ji})$

TABLE II
OBSERVABILITY SUBSPACES ASSOCIATED WITH MEASUREMENTS AND ODOMETRY DATA. (FOR ODOMETRY INPUTS, ○ MEANS *non-zero*, × MEANS *zero*, AND △ MEANS EITHER ZERO OR NON-ZERO.)

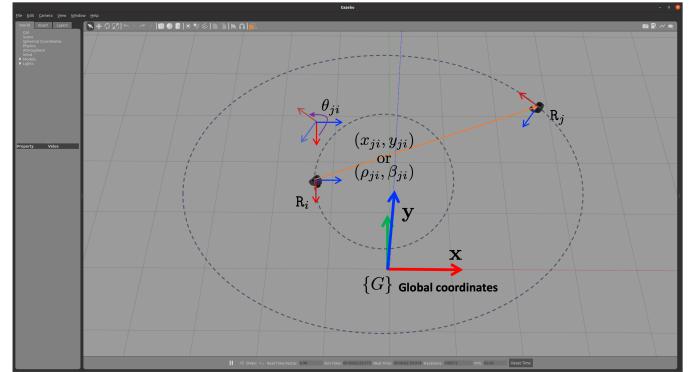


Fig. 4. ROS/Gazebo simulation environments for inter-robot relative pose estimation of two Turtlebot3 robots.

which implies that an inter-robot system with both range and bearing measurements is *locally observable* almost everywhere.

The observability analyses for various information structures are summarized in Table II. The new results demonstrate that the system of inter-robot relative pose dynamics remains observable when both range and bearing measurements are available, even without velocity information from the neighboring robot.

Remark 3 (Velocity tracking): In addition to indirect velocity estimation through position measurements (range and bearing), direct measurement and computation of velocity using the Doppler effect in radar [67], [68] and LiDAR [69] have been extensively explored. These methods are applied in areas such as target tracking [69]–[71] and ego-motion estimation [72], [73]. Intuitively, direct velocity tracking can ensure observability even with range-only or bearing-only measurements, thus eliminating the need for communicating wheel-odometry data.

IV. ROS/GAZEBO SIMULATION RESULTS: TWO MOBILE ROBOT CASES

This section presents the simulation results for inter-robot relative pose estimation using EKF and PGO across var-

ious information structures. The simulations demonstrating observability analysis and estimation were conducted using the Robot Operating System (ROS) and Gazebo environments, as shown in Fig. 4. These environments facilitate a smooth transition from simulation to real hardware implementation. The estimation source code and simulations were run on a laptop equipped with an Intel i7-9700K CPU (3.6 GHz), 32 GB LPDDR5 RAM, a 512 GB PCIe NVMe x2 SSD, and an Ubuntu 20.04 LTS operating system.

For ROS-Gazebo simulations, we created two Turtlebot3 robot objects and performed relative pose estimations between them. Gazebo provides the ground-truth data of robot pose $X = [p_x, p_y, p_z, q_x, q_y, q_z, q_w]$ (3D position and 3D orientation using quaternions)² for each Turtlebot3 object in the simulation based on the origin of the map, that is, global coordinates. Using the absolute pose information of each robot, we generated observation data for range, bearing with random fluctuation induced by Gaussian white noise. The range measurements noise is modeled as $\mathcal{N}(0, 0.001)$, while the bearing measurements, which are typically less precise due to angular sensor limitations, are represented by $\mathcal{N}(0, 0.01)$. For the process noise, which accounts for uncertainties in the system's dynamics and the integration of externally provided information, the noise associated with the communication is modeled as $\mathcal{N}(0, 0.5^2)$ for the relative velocity, and $\mathcal{N}(0, 0.1^2)$ for the angular velocity and position, due to potential transmission delays, errors, and the inherent variability in these quantities when applied to the motion model. The linear and angular velocities $U = [v_i, w_i, v_j, w_j]$ which are the kinematic control inputs of the two robots R_i and R_j , are set as constants $U = [0.2, 0.1, 0.4, 0.09]$. Therefore, the two robots moved in concentric circles with radii of 2 m and 4 m. More details of simulation environments and source codes are available in the accompanying GitHub page.

We considered six different experimental scenarios, each utilizing distinct information structures for inter-robot state estimation:

- Case 1: Inter-robot communication network, onboard wheel-encoder sensing, and inter-robot ranging measurements.
- Case 2: Inter-robot communication network, onboard wheel-encoder sensing, and inter-robot bearing measurements.
- Case 3: Inter-robot communication network, onboard wheel-encoder sensing, with both inter-robot ranging and bearing measurements.
- Case 4: Onboard sensor measurements for inter-robot ranging and bearing without an inter-robot communication network.
- Case 5: Onboard sensor measurements for inter-robot ranging without an inter-robot communication network.
- Case 6: Onboard sensor measurements for inter-robot bearing without an inter-robot communication network.

²Since our work considers the planar motion and pose in SE(2), we only need to export data of 2D position (p_x, p_y) and 1D orientation, the yaw angle, from the quaternions as $\theta = \text{atan}2(2(q_y q_z + q_w q_x), 1 - 2(q_x^2 + q_y^2))$.

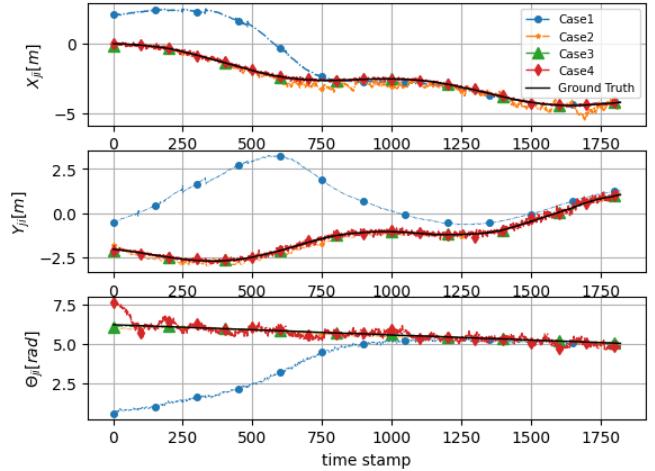


Fig. 5. EKF-based relative pose estimation results with different information structures (Cases 1, 2, 3, and 4).

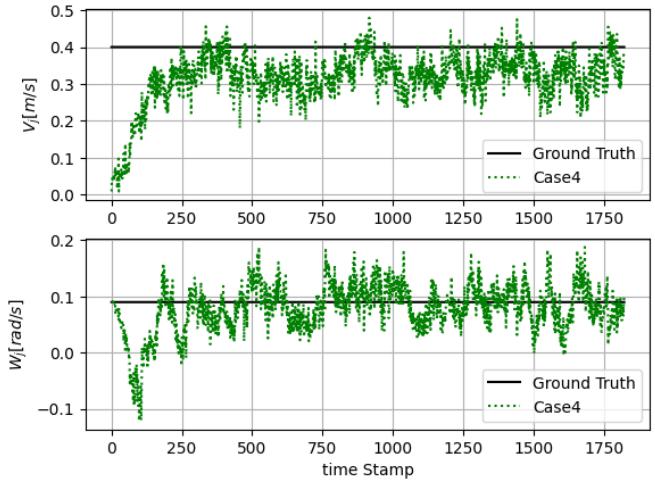


Fig. 6. EKF-based velocity (v_j, w_j) estimation results for Case 4.

A. EKF estimation

Fig. 5 presents the results of EKF-based estimation and ground truth for four different information structures: Cases 1, 2, 3, and 4, which are theoretically expected to ensure observability, based on ROS/Gazebo simulations. In Case 1, where only distance measurements were available, the initial estimation accuracy was low but improved over time. Notably, Case 2, which utilized only bearing measurements, achieved higher estimation accuracy. Results for Case 4, where wheel-odometry data or velocities (v_j, w_j) of a neighboring robot were not provided but were estimated alongside the relative pose, are shown in Fig. 6.

Table III compares the estimation accuracy using RMSE values for EKF-based state estimation across the four cases with different measurement types and information structures. Case 1, relying solely on range measurements, exhibited the highest RMSE. In Case 4, where no inter-robot communication network was present, the RMSE for orientation estimation θ_{ji} was high. However, the estimation of position (x_{ji}, y_{ji}) was relatively accurate.

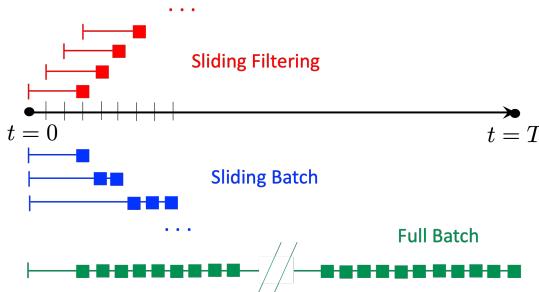


Fig. 7. Diagram of NLS-based PGO for state estimation with different data processing strategies: Sliding Filtering, Sliding Batch, and Full Batch. The colored solid lines refer to the horizon of measurement data considered for optimization-based estimation, while the colored square boxes correspond to the state estimates resulting from the applied methods.

B. Nonlinear least-square estimation

Similar to the EKF-based estimation, we performed simulation experiments for NLS estimation of inter-robot relative pose across each information structure scenario. The Ceres Solver [74] was used for numerical optimization. Three methods were implemented for data processing and optimization:

- **Sliding Filtering (SF) Method:** This method solves NLS problems using a fixed-size sliding window for semi-batch optimization. It updates only the most recent state estimation within the estimation horizon, resulting in partial smoothing.
- **Sliding Batch (SB) Method:** In this approach, the sliding window size is gradually increased, and the corresponding batch optimization problems are solved. The SB method updates state estimates within the expanded sliding window, providing full smoothing within the window while maintaining partial smoothing beyond it.
- **Full Batch (FB) Method:** This method considers all measurements simultaneously, estimating the entire state trajectory over the entire runtime.

Tables IV and V provide a comparative analysis of RMSE values obtained through the Kalman filter and those derived from the optimization method. The batch method exhibits the smallest estimation error, showing an improvement of approximately 94.04% compared to the EKF-based method. When comparing each case, the RMSE result of Case 4 indicates a performance degradation of about 3.24% compared to Case 3, suggesting that while the error is slightly larger, the method still maintains high accuracy. This tendency is also evident in Figs. 8 and 9, which depict the relative pose estimation results using the SF and SB methods, respectively. In the case of the SF method, while the performance improvement rate is

lower compared to EKF than the other numerical optimization methods, it exhibits a high level of computational efficiency.

Tables IV and V present a comparative analysis of RMSE values obtained from the Kalman filter versus those derived from the optimization methods. The Full Batch (FB) method demonstrates the smallest estimation error, with an improvement of approximately 94.04% over the EKF-based method. Comparing each case, the RMSE result for Case 4 shows a performance degradation of about 3.24% compared to Case 3. This indicates a slight increase in error, though the method still maintains high accuracy. This trend is also reflected in Figs. 8 and 9, which illustrate the relative pose estimation results using the Sliding Filtering (SF) and Sliding Batch (SB) methods, respectively. While the SF method shows a lower performance improvement compared to EKF and other numerical optimization methods, it is noted for its high computational efficiency.

TABLE IV
COMPARISON OF RMSE VALUES FOR STATE (RELATIVE POSE IN PLANAR MOTIONS) ESTIMATION USING EKF AND NLS METHODS ACROSS DIFFERENT INFORMATION STRUCTURE CASES

	State	EKF	SF	SB	FB
Case 1	x_{ji}	1.6827	1.0198	1.8363	0.0464
	y_{ji}	2.8620	1.9465	2.1228	0.0377
	θ_{ji}	2.6385	0.2030	0.1872	0.0122
	Total	7.1834	3.1694	4.1464	0.0963
Case 2	x_{ji}	0.3133	0.7459	0.0425	0.0279
	y_{ji}	0.1601	0.5113	0.0757	0.0246
	θ_{ji}	0.0906	0.0675	1.3821	0.0065
	Total	0.5641	1.5244	0.1856	0.0591
Case 3	x_{ji}	0.0213	0.0176	0.0138	0.0129
	y_{ji}	0.0331	0.0152	0.0122	0.0092
	θ_{ji}	0.0198	0.0090	0.0057	0.0049
	Total	0.0744	0.0419	0.0318	0.0270
Case 4	x_{ji}	0.0706	0.0211	0.0145	0.0131
	y_{ji}	0.0994	0.0207	0.0137	0.0096
	θ_{ji}	0.2971	0.0254	0.0117	0.0052
	v_j	0.1099	0.0069	0.0002	0.0021
	ω_j	0.0442	0.0071	0.0005	0.0005
	Total	0.4673	0.0672	0.0399	0.0279

TABLE V
COMPARISON OF PERFORMANCE (%) FOR EACH PARAMETER BASED ON THE NETWORK COMMUNICATION AVAILABILITY OR ESTIMATION METHOD

Reference	Improvement	x_{ji} [%]	y_{ji} [%]	θ_{ji} [%]	Total [%]
Network	Case 3 → Case 4	-0.99	-3.98	-7.79	-3.24
Method	EKF → SF	70.06	79.16	91.46	85.61
	EKF → SB	79.45	86.26	96.06	91.45
	EKF → FB	81.49	90.37	98.23	94.03

TABLE III
RMSE OF STATE (RELATIVE POSE IN PLANAR MOTIONS) ESTIMATION FOR EACH CASE OF INFORMATION FUSION USING EKF

	x_{ji}	y_{ji}	θ_{ji}
Case 1	1.6827	2.8620	2.6385
Case 2	0.3133	0.1601	0.0906
Case 3	0.0213	0.0331	0.0198
Case 4	0.0706	0.0994	0.2971

TABLE VI
COMPARISON OF ESTIMATION FAILURE RESULTS FOR RANGE-ONLY (CASE 5) AND BEARING-ONLY (CASE 6) MEASUREMENTS WITHOUT AN INTER-ROBOT COMMUNICATION NETWORK ACROSS DIFFERENT INFORMATION STRUCTURE CASES

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
Total RMSE	0.0963	0.0591	0.0270	0.0279	2.6922	2.7409

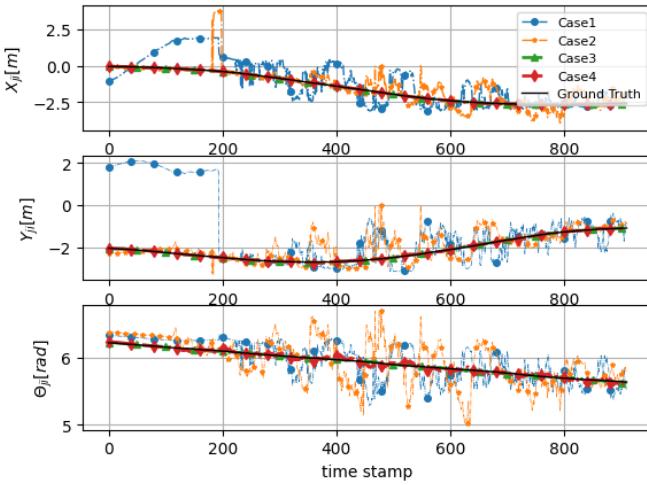


Fig. 8. Relative pose estimation result of each Cases 1, 2, 3, and 4 in SF

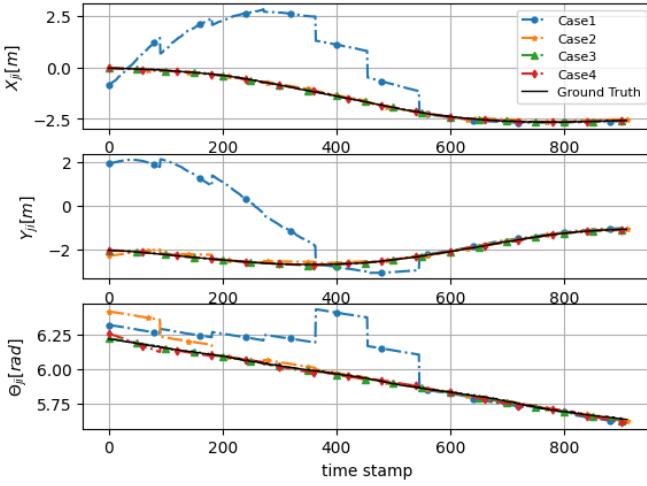


Fig. 9. Relative pose estimation result of each Cases 1, 2, 3, and 4 in SB

In addition, to enhance the clarity of the Case 4 study, we compared the estimation failure results for range-only (Case 5) and bearing-only (Case 6) measurements without an inter-robot communication network, as shown in Table VI. When comparing the total RMSE results of the NLS-based estimation methods, there is a significant performance degradation compared to other information structure cases and it can be seen that the estimation results do not converge, as shown in Figs. 10 and 11.

To further clarify the results for Case 4, we compared the estimation outcomes for range-only (Case 5) and bearing-only (Case 6) measurements without an inter-robot communication network, as detailed in Table VI. The comparison of total RMSE values for the NLS-based estimation methods reveals significant performance degradation in these cases compared to others. The results indicate that the estimation does not converge, as illustrated in Figs. 10 and 11.

C. Robust M-estimation

For the robust M-estimation experiment, outliers were introduced into the sensor data using the interquartile range (IQR) method [75]. The ratio of outliers was incrementally

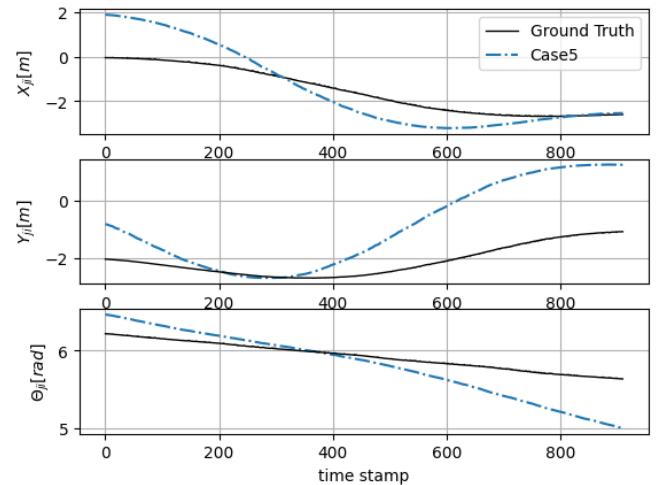


Fig. 10. Relative pose estimation failure results for Case 5 which assumes availability range-only sensing without an inter-robot communication network.

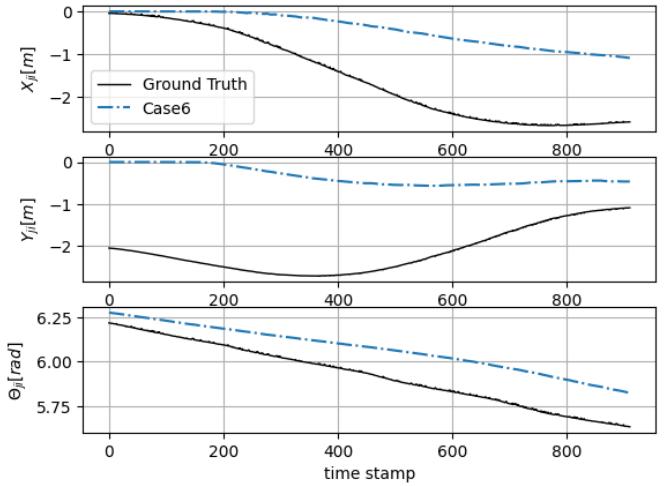


Fig. 11. Relative pose estimation failure results for Case 6 which assumes availability bearing-only sensing without an inter-robot communication network.

increased from 10% to 50% in 10% intervals, with various kernel functions (Huber, Cauchy, Tukey, Arctan) provided by the Ceres Solver being employed. The experiments focused exclusively on Case 4 (odometry + range and bearing sensors).

Fig. 12 displays the estimation results using each kernel function when the sensor data contained 30% outliers. The L2 norm represents the results obtained using the least-squares method without any kernel function. It is evident that the least-squares estimation deviates significantly from the true values due to the influence of outliers. In contrast, the estimates obtained using the kernel functions were more accurate, demonstrating the effectiveness of robust M-estimation in mitigating the impact of outliers.

Fig. 13 compares the Root Mean Square Error (RMSE) values obtained using various kernel functions as the outlier rate increases in Case 4. The threshold values for each kernel function were empirically determined and set as follows: Huber (0.5), Cauchy (0.5), Tukey (3.0), and Arctan (3.0). The tuning parameter values t are listed in Table I and were selected experimentally. The figure shows that the RMSE increases for

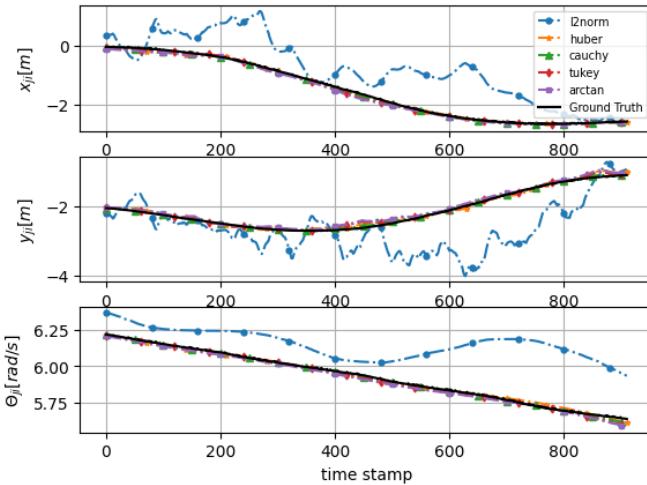


Fig. 12. Relative pose estimation results of M-estimation using different kernel functions are showcased for Case 4, wherein no communication between robots occurs, and only range and bearing measurements are utilized for estimation.

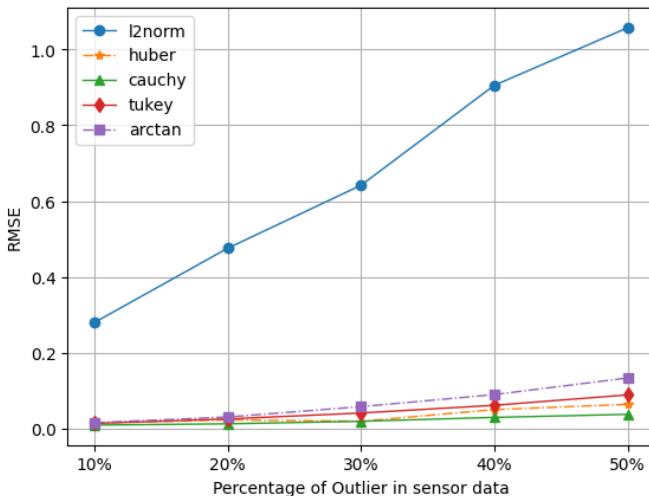


Fig. 13. RMSE of M-estimation results using different kernel functions with varying ratios of outliers for Case 4.

all kernel functions as the proportion of outliers rises. However, compared to the rapid increase in RMSE observed with the L2 norm, the RMSE increases more gradually when kernel functions are applied. This demonstrates the effectiveness of kernel functions in mitigating the impact of outliers.

V. HARDWARE EXPERIMENTS

A series of experiments was conducted to assess the accuracy of inter-robot relative pose estimation and its applicability to real-world hardware. We focused on Case 4, where only onboard sensor measurements of inter-robot ranging and bearing were used without any inter-robot communication network. The successful results for Case 4 inherently validate the efficacy and observability of other information structures as well.

Four distinct motion scenarios were designed for the robots, including a mirrored motion scenario, to reflect potential relative positional dynamics in inter-robot systems. Both Extended Kalman Filter (EKF) and Pose Graph Optimization (PGO)

methods were applied to each scenario, and their performance was compared. A video of the experiments is available in the *Supplementary Material*.

A. Experimental setup

Our experimental setup is designed to closely mimic practical scenarios by implementing and testing the proposed methods on real-world hardware. The experiments are conducted in indoor environments that reflect realistic conditions for multi-robot systems operating in controlled spaces, such as warehouses, research facilities, and indoor navigation settings. Although these testing environments are not yet certifiable to industry standards, they provide valuable insights into the practical performance of our approach. The proposed methods are particularly relevant for practical applications where robots must operate in environments with communication constraints or interruptions. Our current setup provides a strong foundation for demonstrating the feasibility of our methods.

Two Turtlebot3 robots were used in the experimental trials, designated as the *ego-robot* and the *neighbor-robot*. Each robot was equipped with a UWB module from NoopLoop, offering precise range and bearing measurements based on Time of Arrival (TOA) and Angle of Arrival (AOA) signals. Each robot had a tag and anchor pair, providing data at a frequency of 200 Hz, with an accuracy of 5 cm for range measurements and 5° for direction-angle determination. The operational range of the tag was within ±90° relative to the anchor coordinates, with a slight decrease in measurement accuracy as the tag approached these ±90° extremes.

For *EKF-based estimation*, the algorithm was executed on a LattePanda Alpha single-board computer (SBC) equipped with an Intel Core m3-7Y30 CPU (2.6 GHz) and 8 GB of Dual Channel RAM. *PGO-based estimation* computations were performed on a laptop with an Intel i7-9700K CPU (3.5 GHz), 32 GB LPDDR5 RAM, and a 512 GB PCIe NVMe x2 SSD. All software was developed in ROS and C++, running on Ubuntu 20.04 LTS.

In terms of runtime and computational cost, the EKF-based estimation achieved an average data processing rate of 12.46 Hz, consuming 18.87 MiB of memory during inference. In comparison, the offline NLS-based estimation averaged 8.56 Hz with a memory usage of 556.65 MiB. During the indoor experiments, the positions and orientations of the robots were tracked as ground truth using a motion capture (MoCap) system equipped with ten Qualisys ARQUS A5 cameras. This system operated at a frame rate of 700 to 1400 fps, with resolutions between 1 to 5 MP and a 3D tracking resolution of 0.06 mm, while maintaining a camera latency of 1.4 ms.

The experimental design included four distinct scenarios, each carefully crafted to analyze variations in performance based on the robots' relative positions, linear velocities, and angular velocities:

- Scenario 1: The two robots move in concentric circles at constant velocities, with radii of 1.5 m and 2.0 m, respectively.
- Scenario 2: The robots follow a motion pattern similar to Scenario 1 but with varying speeds over time, introducing changes in dynamic behavior.

- Scenario 3: The neighbor robot moves in a circle with a radius of 0.8 m, while the ego robot remains stationary at a fixed point outside the circle.
- Scenario 4: Similar to Scenario 3, the neighbor robot moves in a circular path, but the ego robot moves back and forth toward the center of the circle from a point outside it, introducing additional complexity in the interaction dynamics.

B. UWB sensor calibration

Furthermore, we conducted a meticulous sensor calibration between UWB and MoCap systems and the outcomes of which are illustrated in Fig. 14. Recognizing and adjusting for offsets, excluding noise, between MoCap-tracked data and UWB-AOA measurements are crucial for ensuring high fidelity in our estimation.

The least-squares method was employed as a robust technique for refining distance and orientation measurements, considering the following equations:

$$z_{ij}^{\rho} = \rho_{ji} + \chi^{\rho}(X_{ji}, Y_{ji}) + \nu_{ji}^{\rho} \quad (46)$$

where z_{ij}^{ρ} is the range measured by MoCap, ρ_{ji} is the range measured by UWB, and ν_{ji}^{ρ} is the noise or uncertainty of the range measurements. The regressor $\chi^{\rho}(X_{ji}, Y_{ji})$ is a range calibration factor defined as $\chi^{\rho}(X_{ji}, Y_{ji}) = a_0 + a_1 X_{ji} + a_2 Y_{ji}$ and the parameters are determined using the least-squares method with offline experiments of UWB sensing.

Similarly, the bearing measurements can be represented as

$$z_{ij}^{\beta} = \beta_{ji} + \chi^{\beta}(X_{ji}, Y_{ji}) + \nu_{ji}^{\beta} \quad (47)$$

where z_{ij}^{β} is the range measured by MoCap, β_{ji} is the bearing measured by UWB, ν_{ji}^{β} is a noise or uncertainty of bearing, and the regressor $\chi^{\beta}(X_{ji}, Y_{ji})$ is a bearing calibration factor defined as $\chi^{\beta}(X_{ji}, Y_{ji}) = b_0 + b_1 X_{ji} + b_2 Y_{ji}$. The parameters (b_0, b_1, b_2) are determined using the least-squares method with offline experiments of UWB sensing.

For hardware demonstrations, the actual measurements used for the state estimation are

$$\begin{aligned} \tilde{\rho}_{ji} &= \rho_{ji} + \chi^{\rho}(X_{ji}, Y_{ji}) + \nu_{ji}^{\rho}, \\ \tilde{\beta}_{ji} &= \beta_{ji} + \chi^{\beta}(X_{ji}, Y_{ji}) + \nu_{ji}^{\beta} \end{aligned} \quad (48)$$

where ρ_{ji} and β_{ji} are the measurements from the UWB anchor-tag module, and $\chi^{\rho}(\cdot)$ and $\chi^{\beta}(\cdot)$ are the pre-calibrated factors of range and bearing sensing, respectively. Fig. 14 shows a comparison of before and after calibration in UWB sensor measurements.

Fig. 15 depicts the MoCap experimental environment used for hardware implementation and validation. Ground-truth data were captured using an optoelectronic marker-based motion capture system, which features ten Qualisys ARQUUS A5 cameras. These cameras were synchronized in both space and time via the Qualisys Track Manager, providing a unified global frame as well as a local reference frame centered on each robot platform. Special care was taken in the placement of marker and sensor modules to adhere to the specifications for resolution, focus, and exposure time.

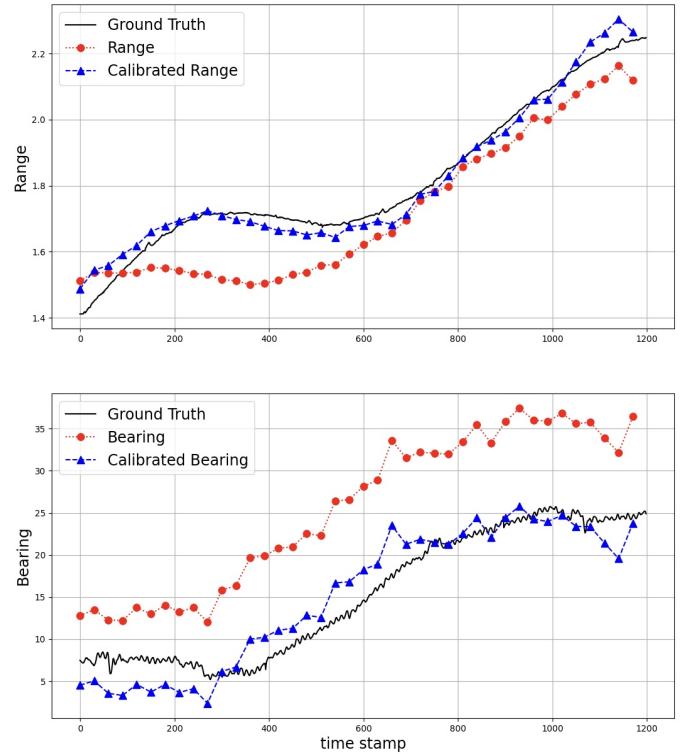


Fig. 14. Results of UWB-AOA calibration for range and bearing: the black solid line represents the ground truth, the red solid line signifies raw UWB measurements, and the blue solid line depicts calibrated measurements as per(48), which are utilized for estimation.

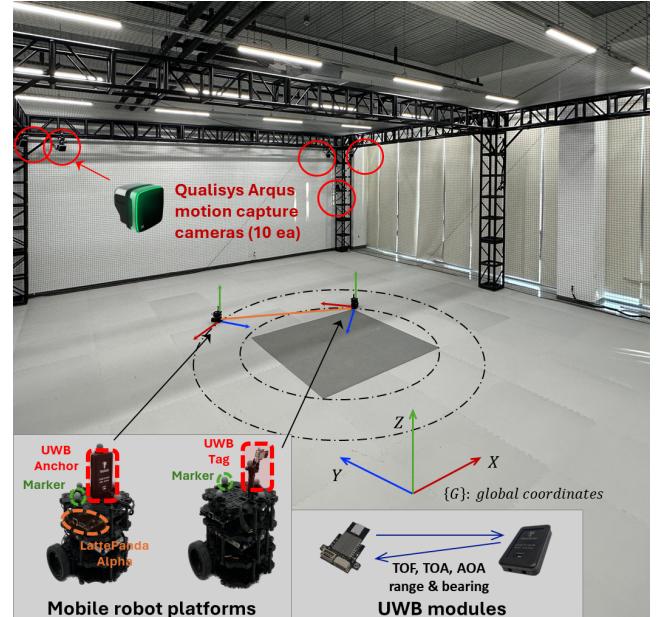


Fig. 15. Photo of the MoCap facility and mobile robot platforms used for hardware experiments, in which a LattePanda Alpha SBC is used for on-board computing.

C. EKF estimation

Hardware experiments were conducted to evaluate the performance of EKF-based estimation across various motion scenarios. Specifically, Scenario 1, which mirrors the simulated environment, provided a real-world test for inter-robot relative pose estimation. Figs. 16 and 17 display the EKF-based

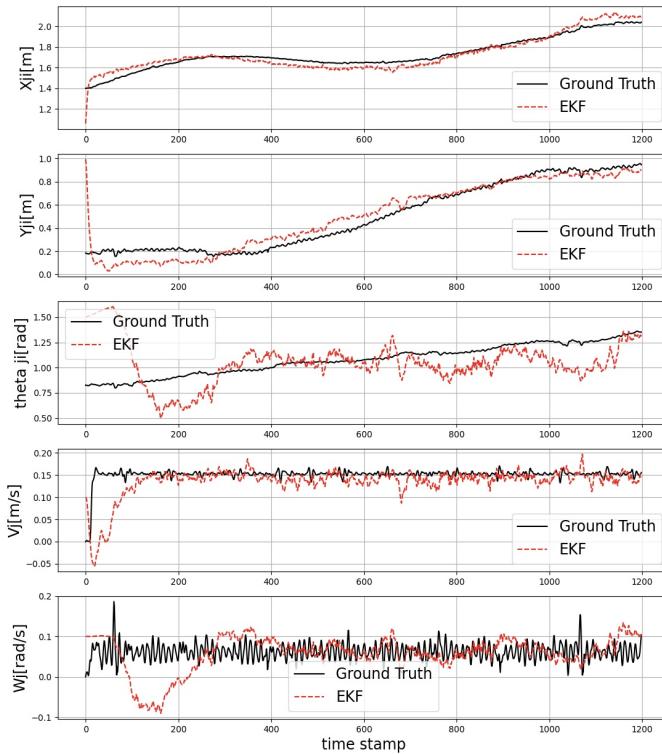


Fig. 16. EKF relative pose (x_{ji} , y_{ji} , θ_{ji}) and velocity (v_j , w_j) estimation results with Scenario 1.

estimation results alongside the ground truth for Scenarios 1 and 2. The experimental results demonstrate that the estimated poses in both scenarios increasingly align with the expected values, even in the absence of wheel-odometry data or velocity information from the neighboring robot. Although there is a noticeable reduction in estimation precision in Scenario 2—where the robots move at varying speeds—the estimated trajectory remains closely aligned with the ground truth, similar to Scenario 1.

The estimation algorithm exhibited robustness against irregular pulses or outliers, which may have been caused by uneven surfaces or vibrations in the experimental environment. Table VII provides a comparative analysis of the RMSE values for EKF-based state estimation across Scenarios 1, 2, and Case 4 in the simulation study. While the RMSE for orientation estimation θ_{ji} was relatively high, the position estimates x_{ji} and y_{ji} were generally accurate. In Scenarios 3 and 4, as shown in Figs. 18 and 19, the limited Field of View (FOV) of the UWB-AOA sensors, which provide range and bearing measurements, negatively impacted the accuracy of the estimations. We suggest that utilizing alternative sensors with higher precision could potentially improve the estimation accuracy.

D. Nonlinear least-square estimation

Similar to the EKF-based relative pose estimation experiments, we conducted a series of nonlinear least-squares (NLS) experiments for Scenarios 1 and 2, employing the Ceres Solver for numerical optimization. For the hardware implementation, we applied full-batch nonlinear least-squares (FB NLS) methods, utilizing both L2 norm and Cauchy loss kernel functions

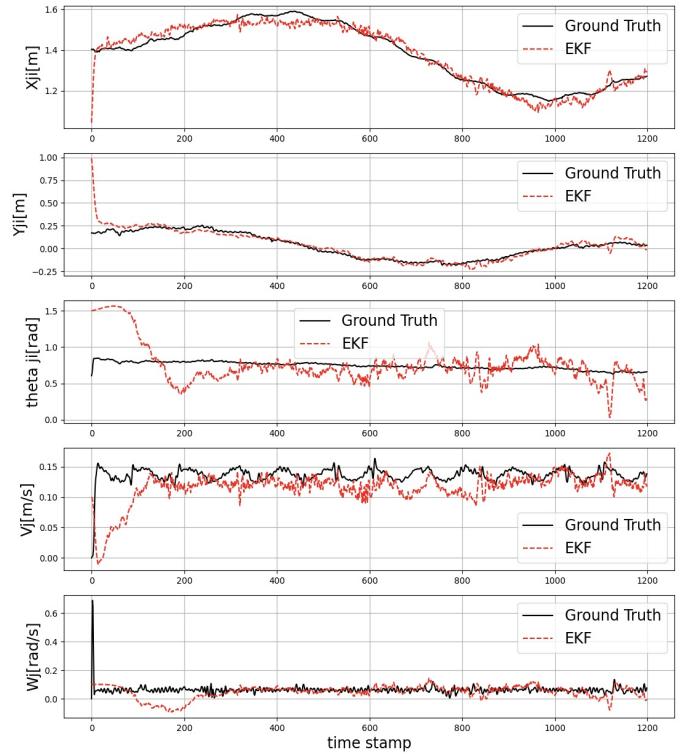


Fig. 17. EKF relative pose (x_{ji} , y_{ji} , θ_{ji}) and velocity (v_j , w_j) estimation results with Scenario 2.

TABLE VII
RMSE VALUES OF STATE (RELATIVE POSE IN PLANAR MOTIONS)
ESTIMATION FOR SIMULATION (CASE 4) AND EACH SCENARIO OF
HARDWARE IMPLEMENTATION USING EKF

	x_{ji}	y_{ji}	θ_{ji}
Simulation (Case 4)	0.0706	0.0994	0.2971
Hardware Experiment Scenario 1	0.0509	0.0889	0.2142
Hardware Experiment Scenario 2	0.0356	0.0631	0.2061

as robust M-estimation techniques to manage measurement noise.

Fig. 20 and Fig. 21 illustrate that the NLS-based estimation of position x_{ji} , y_{ji} , and orientation θ_{ji} shows progressively reduced error and smoother trajectories in Scenarios 1 and 2, respectively. While the precision of these estimations is more dependent on measurement accuracy and onboard computational capability, as determined by the hardware, the results demonstrate substantial effectiveness across all examined scenarios. Table VIII provides a comparison of the RMSE values for EKF and NLS optimization methods using L2 norm and Cauchy loss. Notably, in Scenario 2, both NLS-based methods prove effective in managing significant deviations from actual values.

VI. DISCUSSION AND FUTURE DIRECTIONS

A. Sensor-agnostic relative pose estimation

Although Ultra-Wideband (UWB) is often favored for relative positioning due to its high accuracy in short-range measurements, it is not essential for the implementation of our proposed method. Our approach is sensor-agnostic and can operate with any system that provides range and/or bearing

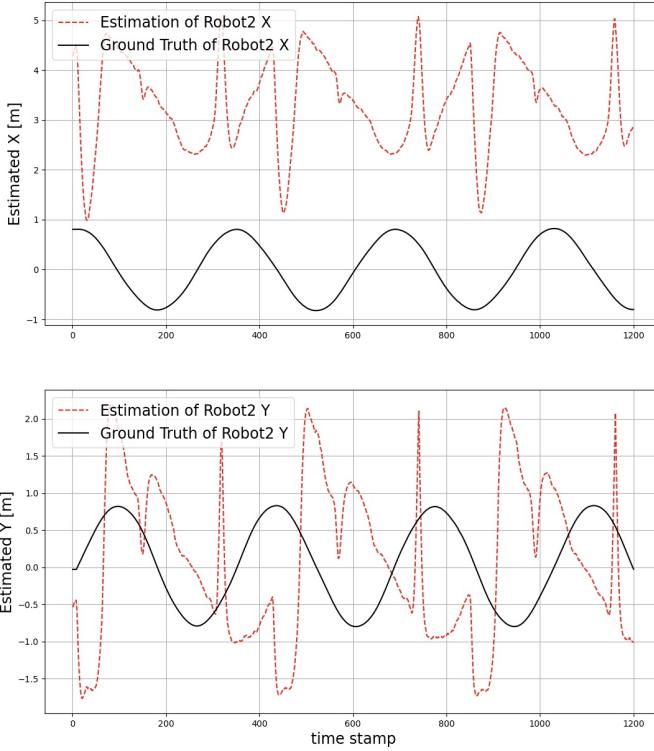


Fig. 18. EKF relative position (x_{j_i}, y_{j_i}) estimation results with Scenario 3.

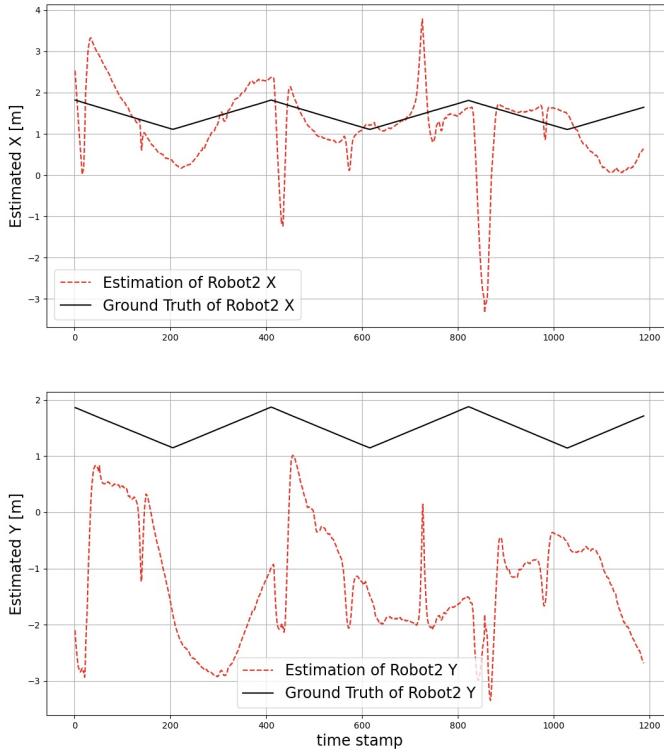


Fig. 19. EKF relative position (x_{j_i}, y_{j_i}) estimation results with Scenario 4.

measurements, making it adaptable to a variety of sensing technologies. As long as the measurement models are available and the sensor specifications are well-defined, our method can be applied using any of these technologies, offering versatility and adaptability in real-world applications. This flexibility significantly enhances the applicability of our method across

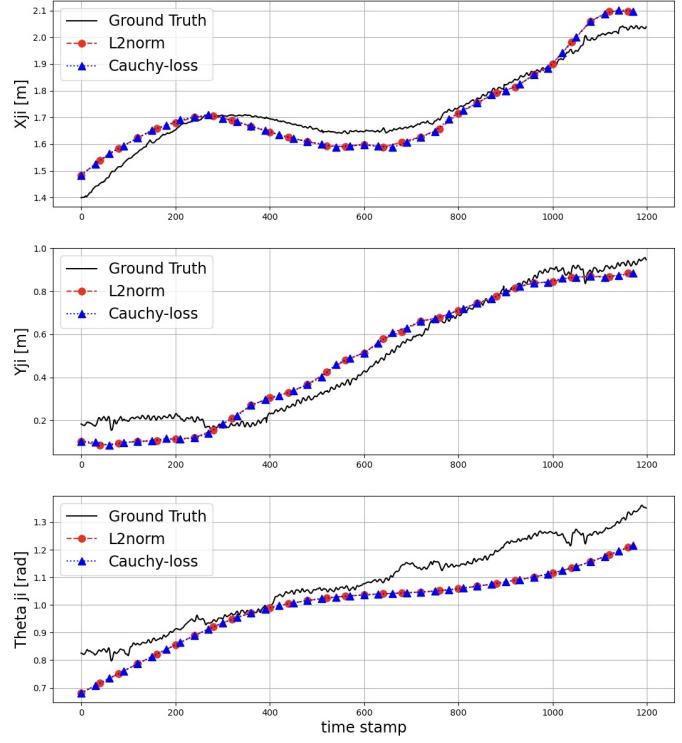


Fig. 20. Relative pose estimation result of Scenario 1 for FB NLS with two different loss functions.

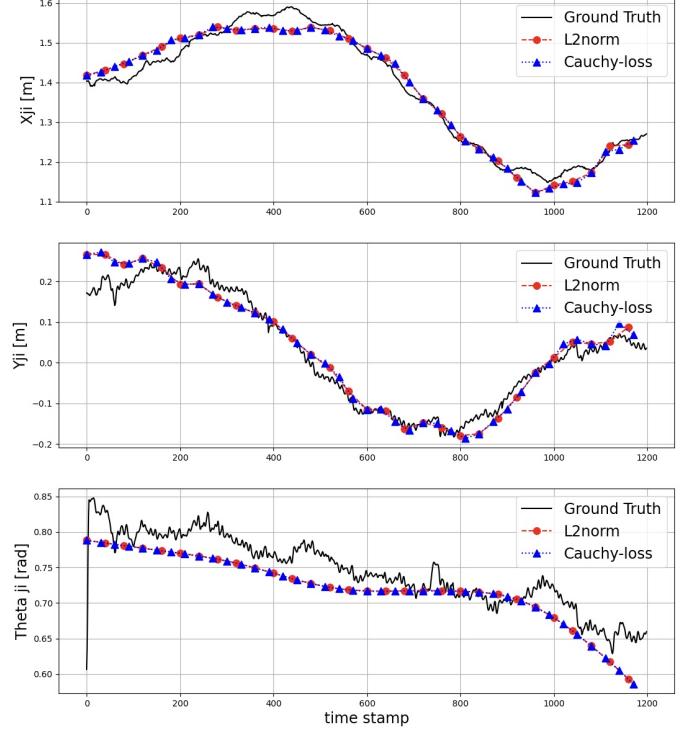


Fig. 21. Relative pose estimation result of Scenario 2 for FB NLS with two different loss functions.

diverse environments and scenarios.

Several alternative sensors that can be used for our method of inter-robot relative pose estimation include:

- Depth Cameras: Stereo cameras or monocular cameras paired with depth sensors offer both range and bearing measurements. These systems are particularly well-suited

TABLE VIII
COMPARISON OF RMSE VALUES FOR STATE (RELATIVE POSE IN PLANAR MOTIONS) ESTIMATION USING EKF AND NLS METHODS ACROSS DIFFERENT MOTION SCENARIOS

	State	EKF	FB with L2norm	FB with Cauchy
Hardware Experiment Scenario 1	x_{ji}	0.05093	0.04860	0.04860
	y_{ji}	0.08895	0.07019	0.07020
	θ_{ji}	0.21172	0.08667	0.08656
	Total	0.35406	0.20536	0.20537
Hardware Experiment Scenario 2	x_{ji}	0.03559	0.02748	0.02747
	y_{ji}	0.06306	0.03152	0.03153
	θ_{ji}	0.20613	0.03347	0.03350
	Total	0.30479	0.09248	0.09251

for visual-based localization in both indoor and outdoor environments, providing rich spatial information for relative pose estimation.

- LiDAR: Light Detection and Ranging (LiDAR) sensors excel at producing accurate range data, especially in structured environments. LiDAR is widely used in Simultaneous Localization and Mapping (SLAM) applications and can be seamlessly integrated into our method for precise inter-robot positioning.
- Radar: Radar systems are a robust alternative, especially in outdoor environments where optical sensors may struggle due to adverse weather conditions, such as fog or rain. Radar provides reliable range and bearing information, often over longer distances, making it suitable for outdoor multi-robot systems.

B. Uncertainty propagation in distributed data fusion

An important application of inter-robot relative pose estimation is distributed data fusion (DDF), which can be utilized for:

- *Collaborative multi-robot target tracking* [76], where the goal is to estimate the state of one or multiple targets by integrating data from a team of robots.
- *Collaborative multi-robot SLAM* [77], [78], which aims to jointly perform localization and mapping of a mission space through the collective efforts of multiple robots.
- *Collaborative multi-robot localization* [79]–[81], which includes intra-robot (time-to-time or event-to-event) and inter-robot (robot-to-robot) relative pose estimation.
- *Multi-robot SLAM with inter-robot loop closure* [82], [83], where relative pose (loop closure) between two robots is used to minimize errors in UWB ranging, allowing robots to localize and map collaboratively even in the absence of visual loop closures.

Mapping involves constructing estimates of object locations or poses relative to a world or local frame. For clarity, we consider a collaborative two-robot single-static-target tracking problem. In this scenario, each robot's pose is represented probabilistically. Let Robot R_i have an estimated absolute pose denoted as $T_i = (x_i, y_i, \theta_i)$, following a normal distribution $\mathcal{N}(\hat{T}_i, \Sigma_i)$. The target's location or pose, $\zeta^i = (\zeta_x^i, \zeta_y^i, \zeta_\theta^i)$, is also represented probabilistically in Robot R_i 's local coordinate frame as $\mathcal{N}(\hat{\zeta}^i, \Phi^i)$.

$$\begin{aligned} P(\zeta^i(t)|z_i(t)) &\approx \mathcal{N}(\hat{\zeta}^i(t), \Phi^i(t)) \\ P(T_i(t)|T_i(0), y_i(0:t)) &\approx \mathcal{N}(\hat{T}_i(t), \Sigma_i(t)) \\ P(T_{ji}(t)|T_{ji}(0), y_{ji}(0:t)) &\approx \mathcal{N}(\hat{T}_{ji}(t), \Sigma_{ji}(t)) \end{aligned}$$

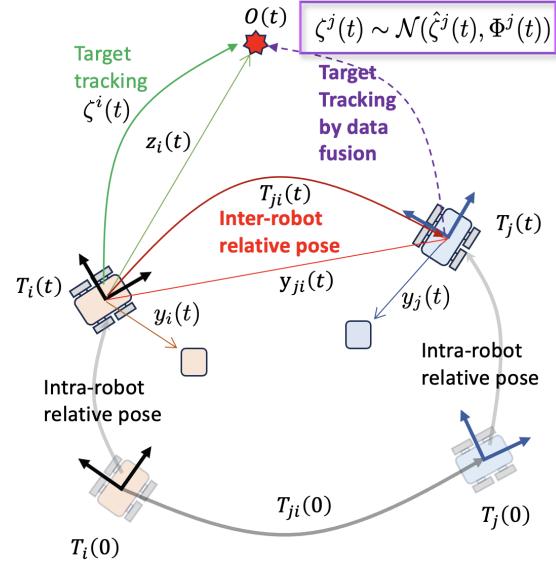


Fig. 22. The representation and propagation of spatial uncertainty in a two-robot distributed data fusion for single-target tracking are illustrated in which Robot R_j doesn't need to have a direct observation of the target O but uses the transformations of spatial information given in (49). The uncertainties of (a) target estimation, (b) intra-robot pose estimation, and (c) inter-robot relative pose estimation propagate through the process of distributed data fusion for target tracking. The propagated accumulative uncertainty in data fusion is quantified and approximated as the covariance $\Phi^j(t) = C(\zeta^j(t))$ at each time-step t .

Probabilistic estimates are not limited to Gaussian distributions, but in this context, we focus on uncertainty quantification and propagation by approximating the first and second moments of random variables, such as positions or poses. Specifically, we assume that the relative pose of Robot R_j with respect to Robot R_i is represented probabilistically. This relative pose is denoted as $T_{ji} = (x_{ji}, y_{ji}, \theta_{ji})$ and follows a normal distribution $\mathcal{N}(\hat{T}_{ji}, \Sigma_{ji})$.

Fig. 22 illustrates the process of Distributed Data Fusion (DDF) for tracking a single target using a team of two mobile robots. In this scenario, Robot R_j does not have a direct measurement or estimation of the target's position. Instead, Robot R_i provides the spatial information of the target O to Robot R_j . Robot R_j then transforms this spatial information into its own local coordinate frame by considering the inter-robot relative pose T_{ji} (or equivalently, T_{ij}). In DDF-based target tracking, the central challenge is to compute or estimate the probability distribution of the target pose ζ^j in Robot R_j 's local coordinates. Specifically, the goal of DDF is to compute the transformed spatial information using the formula:

$$\zeta^j = \ominus T_{ji} \oplus (T_i \oplus \zeta^i) \quad (49)$$

where $T_i \oplus \zeta^i$ represents the pose composition of T_i and ζ^i in the head-to-tail configuration, $\ominus T_{ji} = T_{ij}$ denotes the inverse transformation of T_{ji} , and $\ominus T_{ji} \oplus \zeta^i$ is the tail-to-tail pose composition, with $\zeta^i := T_i \oplus \zeta^i$. These operations define

the spatial relationships between the information and variables involved.

As shown in Fig. 22, given the first and second moments of the uncertain spatial variables (T_{ji}, T_i, ζ^i) , approximating the first and second moments of the transformed data ζ^j for which the mean and covariance are estimated is possible as follows [84], [85]:

$$\begin{aligned}\hat{\zeta}^j &\approx \ominus \hat{T}_{ji} \oplus (\hat{T}_i \oplus \hat{\zeta}^i) \\ C(\zeta^j) &\approx J_{\oplus} \begin{bmatrix} J_{\ominus} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \Sigma_{ji} & 0 \\ 0 & C(\zeta^i) \end{bmatrix} \begin{bmatrix} J_{\ominus} & 0 \\ 0 & I \end{bmatrix}^{\top} J_{\oplus}^{\top}\end{aligned}\quad (50)$$

where $\hat{\zeta}$ and $C(\zeta)$ denote the mean and covariance of a random variable ζ , respectively. The compound relation and associated uncertainty (or error) propagation can be computed as

$$\begin{aligned}\tilde{\zeta}^i &= \begin{bmatrix} \tilde{\zeta}_x^i \\ \tilde{\zeta}_y^i \\ \tilde{\zeta}_\theta^i \end{bmatrix} = T_i \oplus \zeta^i = \begin{bmatrix} \zeta_x^i \cos \theta_i - \zeta_y^i \sin \theta_i + x_i \\ \zeta_x^i \sin \theta_i + \zeta_y^i \sin \theta_i + y_i \\ \theta_i + \zeta_\theta^i \end{bmatrix}, \\ C(\tilde{\zeta}^i) &= C(T_i \oplus \zeta^i) \approx \tilde{J}_{\oplus} \begin{bmatrix} \Sigma_i & 0 \\ 0 & \Phi^i \end{bmatrix} \tilde{J}_{\oplus}^{\top}\end{aligned}$$

where the Jacobians are given as

$$\begin{aligned}\tilde{J}_{\oplus} &= \frac{\partial \tilde{\zeta}^i}{\partial (T_i, \zeta^i)} = \begin{bmatrix} 1 & 0 & -(\zeta_x^i - y_i) \cos \theta_i - \sin \theta_i & 0 \\ 0 & 1 & (\zeta_x^i - x_i) & \sin \theta_i & \cos \theta_i & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}, \\ J_{\oplus} &= \frac{\partial T_{ij}}{\partial T_{ji}} = \begin{bmatrix} -\cos \theta_{ji} & -\sin \theta_{ji} & y_{ji} \\ \sin \theta_{ji} & -\cos \theta_{ji} & -x_{ji} \\ 0 & 0 & -1 \end{bmatrix}, \\ J_{\oplus} &= \frac{\partial \zeta^j}{\partial (T_{ji}, \tilde{\zeta}^j)} = \begin{bmatrix} 1 & 0 & -(\tilde{\zeta}_y^i - y_{ji}) \cos \theta_{ji} & -\sin \theta_{ji} & 0 \\ 0 & 1 & (\tilde{\zeta}_x^i - x_{ji}) & \sin \theta_{ji} & \cos \theta_{ji} & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}\end{aligned}$$

that are evaluated at the mean values of the spatial representations $(\hat{T}_{ji}, \hat{T}_i, \hat{\zeta}^i)$.

Remark 4 (Uncertainty propagation): From the previous computations that are crucial for a robust DDF, one can see the importance of the inter-robot relative pose estimation $T_{ji} = (x_{ji}, y_{ji}, \theta_{ji})$ as well as the intra-robot relative pose estimation $T_i = (x_i, y_i, \theta_i)$. In addition, spatial uncertainty of a transformed information ζ^j quantified as its covariance $C(\zeta^j)$ can be approximated as a function of the mean $(\hat{T}_{ji}, \hat{T}_i, \hat{\zeta}^i)$ and the covariance $(\Sigma_{ji}, \Sigma_i, \Phi^i)$, as shown in (50).

Remark 5 (Symmetries and Perturbation Map): In the realm of distributed data fusion (DDF), the exploration extends to a probabilistic representation encompassing both mobile robot localization and map features with interdependent relationships, known as the symmetry and perturbation map (SPmap) [86] in multi-sensor SLAM literature. This approach addresses the inherent complexities in distributed data fusion (DDF) where both localization and mapping are involved, and where sensors and robots may be interconnected in a networked system.

Remark 6 (KF-based probabilistic inference vs. NLP-based PGO): Essentially, robot state estimation considerably benefits from methods that explicitly and accurately consider uncertainty. Techniques based on Kalman filter (KF) recursive

probabilistic inference, such as Error-State EKF [87] and Invariant EKF [88], Unscented Kalman Filter (UKF) [89], [90], and Particle Filter (PF) [91], [92] on manifolds, inherently compute estimation uncertainty as the covariance matrix. In contrast, Nonlinear Programming (NLP)-based Pose Graph Optimization (PGO) methods, like those used in our approach, do not explicitly compute uncertainty quantification. Nonetheless, they exhibit robustness against outliers, as discussed in Section VI-C and demonstrated in simulation and hardware experiments (Sections IV and V). Future work should explore an iterative EKF that incorporates robust kernel functions in the correction step [93], extending it to robot state estimation on manifolds. This could offer a promising alternative solution, combining the advantages of KF-based recursive probabilistic inference and NLP-based (semi-)batch PGO, particularly concerning uncertainty quantification and propagation in robot state estimation and spatial information.

C. Outlier robust relative pose estimation

The presence of outliers in sensor measurements can significantly degrade estimation performance. As demonstrated in Section IV, nonlinear least-squares methods and standard EKF are particularly vulnerable to outliers. In contrast, Pose Graph Optimization (PGO) utilizing M-estimation with robust cost functions [53], [54] offers high-fidelity and reliable state estimation. However, the effectiveness of M-estimators is closely tied to the tuning of parameters within the robust cost functions. This highlights the need for adaptive tuning methods to optimize Iteratively Reweighted Least Squares (IRLS) approaches. Several strategies for adaptive parameter tuning in M-estimation have been proposed [94]–[97].

In addition to selecting the appropriate M-estimator and tuning its parameters, the time-horizon of measurement processing plays a crucial role in determining both the accuracy and computational efficiency of Pose Graph Optimization (PGO)-based estimation due to its inherent smoothing nature. To balance accuracy with computational time, robust PGO-based smoothing can either use sliding-window measurements (partial information) or full-batch measurements (complete information). However, finding the optimal sliding window size can be challenging and often involves a trial-and-error process. Alternatively, employing dynamic programming-based Moving Horizon Estimation (MHE) [98]–[100] can enhance performance by utilizing a shorter sliding window horizon, thereby improving both robustness and efficiency in PGO-based smoothing for robot state estimation.

Alongside integrating M-estimators with PGO, Random Sample Consensus (RANSAC) is a prominent iterative method for estimating parameters and system states from data contaminated with outliers. Previous studies [101], [102] have examined the combination of RANSAC with Kalman Filters (KF) to bolster robust state estimation for dynamic systems. More recently, an iterative Extended Kalman Filter (EKF) incorporating robust loss functions within the M-estimation framework has been proposed [93], [103], [104]. This approach utilizes Iteratively Reweighted Least Squares (IRLS) optimization in the correction step to enhance robustness against outliers.

It also explores the connection between Kalman gain and Gauss-Newton iterations, presenting a promising alternative for robust state estimation in the presence of outliers.

D. Range-aided multi-robot state estimation

Recently, there has been significant focus on range-aided (RA) multi-robot state estimation, where inter-robot range measurements are crucial for cooperative localization, mapping, and Simultaneous Localization and Mapping (SLAM). Particle Filter (PF)-based methods for multi-robot SLAM, which use inter-robot range and bearing measurements communicated over ad hoc networks, have been explored in [105]–[107]. Research into RA cooperative localization with Ultra-Wideband (UWB) range measurements using a multistate constrained Kalman filter (MSCKF) is presented in [45], [46]. Pose Graph Optimization (PGO)-based Range-Aided SLAM (RA-SLAM) [108]–[112] has also gained prominence in multi-robot SLAM.

In RA-SLAM, the nonconvex nonlinear optimization problem in the back-end can be relaxed to convex optimization problems such as second-order conic programming (SOCP) or semidefinite programming (SDP). Recent advancements include certifiable RA-SLAM methods, which provide a lower bound and a suboptimal solution for the original RA-SLAM problem [113]–[117]. This suboptimal solution can serve as an initial estimate for iterative methods that solve the original back-end nonlinear least-squares optimization. The use of UWB devices and signals for inter-robot range measurements in RA-SLAM is reviewed in [118]–[121]. Additionally, sensor fusion involving Camera-IMU-UWB is considered for indoor localization [122]–[127] and navigation [128]–[131].

For cooperative multi-robot range-aided localization using the PGO framework, the associated *centralized* manifold optimization is defined as follows:

$$\begin{aligned} \min_{(i_t, j_s) \in \mathcal{E}} & w_{j_s i_t}^{\text{rot}} \|R_{j_s} - R_{i_t} \hat{R}_{j_s i_t}\|_{\text{F}}^2 \\ & + w_{j_s i_t}^{\text{tran}} \|\tau_{j_s} - \tau_{i_t} - R_{i_t} \hat{\tau}_{j_s i_t}\|_2^2 \\ & + w_{j_s i_t}^{\text{rang}} (\|\tau_{j_s} - \tau_{i_t}\|_2 - \tilde{\rho}_{j_s i_t})^2 \end{aligned} \quad (51)$$

$$\text{s.t. } R_{i_t} \in \text{SO}(d), \tau_{i_t} \in \mathbb{R}^d, \forall i, \forall t$$

where both inter-robot spatial relations are encoded in a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Each variable node in \mathcal{V} corresponds to a single pose $T_{i_t} = (R_{i_t}, \tau_{i_t}) \in \text{SE}(d)$ owned by a robot R_i at time t and an edge $(i_t, j_s) \in \mathcal{E}$ is formed, if there is relative spatial information between robots such as relative pose estimates $\hat{T}_{j_s i_t} = (\hat{R}_{j_s i_t}, \hat{\tau}_{j_s i_t})$ or range measurements $\tilde{\rho}_{j_s i_t}$ from T_{i_t} to T_{j_s} .

Central manifold optimization (51) can be decomposed into the following distributed incremental optimization for each robot R_i :

$$\begin{aligned} \min_{j_s: (i_t, j_s) \in \mathcal{E}} & w_{j_s i_t}^{\text{rot}} \|R_{j_s}^+ - R_{i_t} \hat{R}_{j_s i_t}\|_{\text{F}}^2 \\ & + w_{j_s i_t}^{\text{tran}} \|\tau_{j_s}^+ - \tau_{i_t} - R_{i_t} \hat{\tau}_{j_s i_t}\|_2^2 \\ & + w_{j_s i_t}^{\text{rang}} (\|\tau_{j_s}^+ - \tau_{i_t}\|_2 - \tilde{\rho}_{j_s i_t})^2 \end{aligned} \quad (52)$$

$$\begin{aligned} \text{s.t. } R_{i_t} &= R_{i_t}^- \boxplus \delta R_{i_t} \in \text{SO}(d), \tau_{i_t} = \tau_{i_t}^- \boxplus \delta \tau_{i_t} \in \mathbb{R}^d, \forall t \\ \|\delta R_{i_t}\|_{\text{F}} &\leq \epsilon_R, \|\delta \tau_{i_t}\| \leq \epsilon_\tau \end{aligned}$$

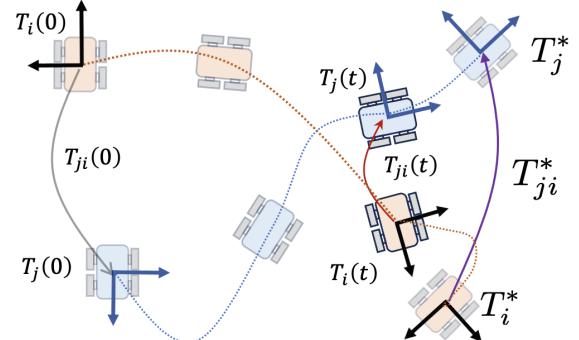


Fig. 23. Formation control of two-robot to achieve the desired absolute pose $T_i^* (T_j^*)$ and the desired relative pose $T_{ji}^* (T_{ij}^*)$ for Robot R_i (R_j).

where $d \in \{2, 3\}$ refers to the dimension and $\boxplus : \mathcal{M}(d) \times \mathbb{R}^d \rightarrow \mathcal{M}(d)$ defines an infinitesimal addition preserving the corresponding manifold structure, the threshold parameters $(\epsilon_R, \epsilon_\tau) > 0$ characterize the trust region optimization (TRO) for updates of orientation and translation variables, respectively. The superscripts $(\cdot)^-$ and $(\cdot)^+$ denote the prior and post-updates of the variables, respectively.

It is assumed that a neighboring robot R_j with $j \in \mathcal{N}_i$ updated his or her poses over time and shared the current guess of these pose estimates $T_{j_s}^+ = (R_{j_s}^+, \tau_{j_s}^+)$ for all time s with Robot R_i . This is only for the clarity of presentation, and the updates can be either synchronous or asynchronous. Once the optimization (52) is solved with a solution $(\delta R_i^*, \delta \tau_i^*)$, then it is updated as $R_i^+ = R_i^- \boxplus \delta R_i^*$ and $\tau_i^+ = \tau_i^- \boxplus \delta \tau_i^*$. Note that the distributed TRO on manifolds can be equivalently rewritten as a Euclidean optimization with a proper parameterization of the perturbed rotation matrix $\delta R_i \in \mathbb{R}^d$.

E. Formation control using the inter-robot relative pose estimation

In the field of cooperative coordination for mobile multi-robot systems, substantial efforts from both industry and academia have been dedicated to improving efficiency, robustness, scalability, and reliability across a variety of applications. These applications include sensor network localization, search and rescue operations, object detection and tracking, environment mapping, and surveillance.

Cooperative coordination problems can be classified based on the desired configuration of a multi-robot system. These classifications include: (a) position-based, (b) relative position-based, (c) range-based, (d) bearing-based, and (e) orientation-based formation control. This paper specifically addresses the general coordination problem of absolute and relative pose-based formation control, which encompasses the more specific formation problems listed from (a) to (e). For a thorough review of multi-robot formation control, readers are encouraged to consult recent monographs [3] and comprehensive reviews [132].

1) Estimation-based output feedback formation control: Using the sensing measurement $y_{ji}(t) = h_{ji}(T_{ji}(t))$ related with the inter-robot relative pose, the formation control input

of the robot R_i can be represented as a function of the relative measurement

$$u_{ji}(t) = \pi_{ji}(y_{ji}(t) - y_{ji}^*, t)$$

or the relative pose estimate $\hat{T}_{ji}(t)$

$$u_{ji}(t) = \pi_{ji}(\hat{T}_{ji}(t) - T_{ji}^*, t)$$

where y_{ji}^* and T_{ji}^* define the desired configurations of robots R_i and R_j in the R_i 's local frame that shown in Fig. 23. From a practical perspective, the actual relative measurement is different from that extracted from the relative pose estimation, that is, $y_{ji}(t) \neq \hat{y}_{ji}(t) := h_{ji}(\hat{T}_{ji}(t))$, owing to measurement noise. In practice, it is better to use the relative pose estimate $\hat{T}_{ji}(t)$ or the associated relative measurement estimate $\hat{y}_{ji}(t) = h_{ji}(\hat{T}_{ji}(t))$ for feedback control of cooperative formation, even when the desired formation is defined in terms of the relative measurement, y_{ji}^* .

Similarly, the absolute pose-based control of robot R_i can be represented as

$$u_{ii}(t) = \pi_{ii}(\hat{T}_i(t) - T_i^*, t)$$

where T_i^* denotes the desired pose described in the local frame of robot R_i , which is generally defined as its initial pose, that is, $T_i(0) = 0$ without loss of generality. As stated in [133], it is not appropriate for mobile robots to use a single global frame; however, a set of local frames is more suitable for mobile robot estimation and control problems.

The output feedback formation control is represented as a weighted sum of the absolute and relative pose-based control inputs.

$$u_i(t) = \lambda_{ii}(t)u_{ii}(t) + \sum_{j \in \mathcal{N}_i} \lambda_{ji}(t)u_{ji}(t) \quad (53)$$

where the time-varying weights satisfy the conditions of unit sum and non-negativity, $\sum_{k \in \mathcal{N}_i \cup \{i\}} \lambda_{ki}(t) = 1$ and $\lambda_{ki}(t) \in [0, 1]$ for all (i, k) and t .

VII. CONCLUSIONS

In this paper, we present a comprehensive overview of observability analysis and estimation methods for wheeled-mobile multi-robot localization. Our novel observability analysis of inter-robot relative pose estimation, performed without information exchange, reveals that wheel-odometry information exchange is unnecessary if inter-robot range and bearing measurements are available. We apply and compare both the Extended Kalman Filter (EKF) and optimization-based M-estimation techniques in ROS/Gazebo simulations for two-robot relative pose and velocity estimation. Our findings indicate that robust Pose Graph Optimization (PGO)-based estimation provides greater reliability than EKF-based estimation, particularly in the presence of outliers. We also validate these estimation methods through hardware experiments using two Turtlebot3 robots.

REPRODUCIBILITY

For further information and access to project resources, please visit:

https://github.com/i-ASL/MRS/MR_RPEstm

REFERENCES

- [1] R. Siegwart, I. R. Nourbakhsh, and D. Scaramuzza, *Introduction to Autonomous Mobile Robots*. Cambridge, Massachusetts: MIT Press, 2011.
- [2] P. Yang, R. A. Freeman, and K. M. Lynch, "Multi-agent coordination by decentralized estimation and control," *IEEE Transactions on Automatic Control*, vol. 53, no. 11, pp. 2480–2496, 2008.
- [3] K. Sakurama, T. Sugie *et al.*, "Generalized coordination of multi-robot systems," *Foundations and Trends® in Systems and Control*, vol. 9, no. 1, pp. 1–170, 2021.
- [4] S. Roumeliotis and G. Bekey, "Distributed multirobot localization," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 5, pp. 781–795, 2002.
- [5] A. Howard, "Multi-robot simultaneous localization and mapping using particle filters," *The International Journal of Robotics Research*, vol. 25, no. 12, pp. 1243–1256, 2006.
- [6] S. Thrun and Y. Liu, "Multi-robot slam with sparse extended information filers," in *Robotics Research. The Eleventh International Symposium*, 2005, pp. 254–266.
- [7] J. P. Queraltà, J. Taipalmaa, B. Can Pullinen, V. K. Sarker, T. Nguyen Gia, H. Tenhunen, M. Gabbouj, J. Raitoharju, and T. Westerlund, "Collaborative multi-robot search and rescue: Planning, coordination, perception, and active vision," *IEEE Access*, vol. 8, pp. 191617–191643, 2020.
- [8] A. Martinelli, F. Pont, and R. Siegwart, "Multi-robot localization using relative observations," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2005, pp. 2797–2802.
- [9] A. Miller, K. Rim, P. Chopra, P. Kelkar, and M. Likhachev, "Cooperative perception and localization for cooperative driving," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 1256–1262.
- [10] S. Wang, Y. Wang, D. Li, and Q. Zhao, "Distributed relative localization algorithms for multi-robot networks: A survey," *Sensors*, vol. 23, no. 5, p. 2399, 2023.
- [11] J. Desai, J. Ostrowski, and V. Kumar, "Modeling and control of formations of nonholonomic mobile robots," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 6, pp. 905–908, 2001.
- [12] A. K. Das, R. Fierro, V. Kumar, J. P. Ostrowski, J. Spletzer, and C. J. Taylor, "A vision-based formation control framework," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 5, pp. 813–825, 2002.
- [13] M. Garcia-Salguero, J. Briales, and J. Gonzalez-Jimenez, "Certifiable relative pose estimation," *Image and Vision Computing*, vol. 109, p. 104142, 2021.
- [14] J.-H. Kim and K.-K. K. Kim, "Vision-based markerless robot-to-robot relative pose estimation using RGB-D data," *Journal of Institute of Control, Robotics and Systems*, vol. 29, no. 6, pp. 495–0503, 2023.
- [15] S. Fang, H. Li, and M. Yang, "Adaptive cubature split covariance intersection filter for multi-vehicle cooperative localization," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1158–1165, 2022.
- [16] G. Haynes and H. Hermes, "Nonlinear controllability via Lie theory," *SIAM Journal on Control*, vol. 8, no. 4, pp. 450–460, 1970.
- [17] R. Hermann and A. Krener, "Nonlinear controllability and observability," *IEEE Transactions on automatic control*, vol. 22, no. 5, pp. 728–740, 1977.
- [18] A. van der Schaft, "Controllability and observability for affine nonlinear Hamiltonian systems," *IEEE Transactions on Automatic Control*, vol. 27, no. 2, pp. 490–492, 1982.
- [19] A. Bicchi, D. Prattichizzo, A. Marigo, and A. Balestrino, "On the observability of mobile vehicle localization," in *Theory and Practice of Control and Systems*. World Scientific, 1998, pp. 142–147.
- [20] F. Conticelli, A. Bicchi, and A. Balestrino, "Observability and nonlinear observers for mobile robot localization," *IFAC Proceedings Volumes*, vol. 33, no. 27, pp. 663–668, 2000.
- [21] F. Lorussi, A. Marigo, and A. Bicchi, "Optimal exploratory paths for a mobile rover," in *IEEE International Conference on Robotics and Automation (ICRA)*, vol. 2, 2001, pp. 2078–2083.
- [22] A. Martinelli and R. Siegwart, "Observability analysis for mobile robot localization," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2005, pp. 1471–1476.
- [23] G. P. Huang, A. I. Mourikis, and S. I. Roumeliotis, "Observability-based rules for designing consistent EKF SLAM estimators," *The International Journal of Robotics Research*, vol. 29, no. 5, pp. 502–528, 2010.

- [24] G. P. Huang, N. Trawny, A. I. Mourikis, and S. I. Roumeliotis, "Observability-based consistent EKF estimators for multi-robot cooperative localization," *Autonomous Robots*, vol. 30, no. 1, pp. 99–122, 2011.
- [25] A. Martinelli, "State estimation based on the concept of continuous symmetry and observability analysis: The case of calibration," *IEEE Transactions on Robotics*, vol. 27, no. 2, pp. 239–255, 2011.
- [26] B. Araki, I. Gilitschenski, T. Ogata, A. Wallar, W. Schwarting, Z. Choudhury, S. Karaman, and D. Rus, "Range-based cooperative localization with nonlinear observability analysis," in *IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019, pp. 1864–1870.
- [27] R. Sharma, R. W. Beard, C. N. Taylor, and S. Quebe, "Graph-based observability analysis of bearing-only cooperative localization," *IEEE Transactions on Robotics*, vol. 28, no. 2, pp. 522–529, 2012.
- [28] J. A. Hesch, D. G. Kottas, S. L. Bowman, and S. I. Roumeliotis, "Camera-IMU-based localization: Observability analysis and consistency improvement," *The International Journal of Robotics Research*, vol. 33, no. 1, pp. 182–201, 2014.
- [29] G. Panahandeh, C. X. Guo, M. Jansson, and S. I. Roumeliotis, "Observability analysis of a vision-aided inertial navigation system using planar features on the ground," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2013, pp. 4187–4194.
- [30] Y. Yang and G. Huang, "Aided inertial navigation: Unified feature representations and observability analysis," in *International Conference on Robotics and Automation (ICRA)*, 2019, pp. 3528–3534.
- [31] G. Panahandeh, S. Hutchinson, P. Händel, and M. Jansson, "Planar-based visual inertial navigation: Observability analysis and motion estimation," *Journal of Intelligent & Robotic Systems*, vol. 82, pp. 277–299, 2016.
- [32] M. K. Paul and S. I. Roumeliotis, "Alternating-stereo VINS: Observability analysis and performance evaluation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4729–4737.
- [33] Y. Yang and G. Huang, "Observability analysis of aided ins with heterogeneous features of points, lines, and planes," *IEEE Transactions on Robotics*, vol. 35, no. 6, pp. 1399–1418, 2019.
- [34] J. Huai, Y. Lin, Y. Zhuang, C. K. Toth, and D. Chen, "Observability analysis and keyframe-based filtering for visual inertial odometry with full self-calibration," *IEEE Transactions on Robotics*, vol. 38, no. 5, pp. 3219–3237, 2022.
- [35] J. A. Hesch, D. G. Kottas, S. L. Bowman, and S. I. Roumeliotis, "Observability-constrained vision-aided inertial navigation," *University of Minnesota, Dept. of Comp. Sci. & Eng., MARS Lab, Tech. Rep*, vol. 1, p. 6, 2012.
- [36] M. A. Gomaa, O. De Silva, G. K. Mann, and R. G. Gosine, "Observability-constrained vins for mavs using interacting multiple model algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 3, pp. 1423–1442, 2020.
- [37] C. Liu, C. Jiang, and H. Wang, "Variable observability constrained visual-inertial-GNSS EKF-based navigation," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 6677–6684, 2022.
- [38] A. Martinelli, "Nonlinear unknown input observability: Extension of the observability rank condition," *IEEE Transactions on Automatic Control*, vol. 64, no. 1, pp. 222–237, 2018.
- [39] ——, "Nonlinear unknown input observability and unknown input reconstruction: The general analytical solution," *Information Fusion*, vol. 85, pp. 23–51, 2022.
- [40] T. Sasaoka, I. Kimoto, Y. Kishimoto, K. Takaba, and H. Nakashima, "Multi-robot SLAM via information fusion extended Kalman filters," *IFAC-PapersOnLine*, vol. 49, no. 22, pp. 303–308, 2016.
- [41] S. Shreedharan, "DKF-SLAM: Distributed Kalman filtering for multi-robot SLAM." Ph.D. dissertation, UC San Diego, 2023.
- [42] A. I. Mourikis and S. I. Roumeliotis, "A multi-state constraint Kalman filter for vision-aided inertial navigation," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2007, pp. 3565–3572.
- [43] I. V. Melnyk, J. A. Hesch, and S. I. Roumeliotis, "Cooperative vision-aided inertial navigation using overlapping views," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2012, pp. 936–943.
- [44] P. Zhu, Y. Yang, W. Ren, and G. Huang, "Cooperative visual-inertial odometry," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 13 135–13 141.
- [45] B. Chenchana, O. Labbani-Igbida, S. Renault, and S. Boria, "Range-based collaborative MSCKF localization," in *25th International Conference on Mechatronics and Machine Vision in Practice (M2VIP)*, 2018, pp. 1–6.
- [46] S. Jia, Y. Jiao, Z. Zhang, R. Xiong, and Y. Wang, "FEJ-VIRO: A consistent first-estimate Jacobian visual-inertial-ranging odometry," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022, pp. 1336–1343.
- [47] R. Olfati-Saber, "Distributed Kalman filter with embedded consensus filters," in *Proceedings of the 44th IEEE Conference on Decision and Control (CDC)*, 2005, pp. 8179–8184.
- [48] R. Olfati-Saber and J. S. Shamma, "Consensus filters for sensor networks and distributed sensor fusion," in *Proceedings of the 44th IEEE Conference on Decision and Control (CDC)*, 2005, pp. 6698–6703.
- [49] U. A. Khan and J. M. Moura, "Distributing the Kalman filter for large-scale systems," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4919–4935, 2008.
- [50] R. Olfati-Saber, "Kalman-consensus filter: Optimality, stability, and performance," in *Proceedings of the 48h IEEE Conference on Decision and Control (CDC)*, 2009, pp. 7036–7042.
- [51] F. S. Cattivelli and A. H. Sayed, "Diffusion strategies for distributed Kalman filtering and smoothing," *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2069–2084, 2010.
- [52] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2508–2530, 2006.
- [53] M. Bosse, G. Agamennoni, and I. Gilitschenski, "Robust estimation and applications in robotics," *Foundations and Trends® in Robotics*, vol. 4, no. 4, pp. 225–269, 2016.
- [54] J. T. Barron, "A general and adaptive robust loss function," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4331–4339.
- [55] P. Bergström and O. Edlund, "Robust registration of point sets using iteratively reweighted least squares," *Computational optimization and applications*, vol. 58, no. 3, pp. 543–561, 2014.
- [56] Y. Tao and S. S.-T. Yau, "Outlier-robust iterative extended Kalman filtering," *IEEE Signal Processing Letters*, vol. 30, pp. 743–747, 2023.
- [57] G. Grisetti, R. Kümmerle, C. Stachniss, and W. Burgard, "A tutorial on graph-based SLAM," *IEEE Intelligent Transportation Systems Magazine*, vol. 2, no. 4, pp. 31–43, 2010.
- [58] F. Dellaert and M. Kaess, "Factor graphs for robot perception," *Foundations and Trends® in Robotics*, vol. 6, no. 1-2, pp. 1–139, 2017.
- [59] F. Dellaert, "Factor graphs: Exploiting structure in robotics," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 4, pp. 141–166, 2021.
- [60] Y. Latif, C. Cadena, and J. Neira, "Robust graph SLAM back-ends: A comparative analysis," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2014, pp. 2683–2690.
- [61] D. McGann, J. R. III, and M. Kaess, "Robust incremental smoothing and mapping (riSAM)," in *IEEE International Conference on Robotics and Automation (ICRA)*, London, GB, 2023, pp. 4157–4163.
- [62] R. Sharma, R. W. Beard, C. N. Taylor, and S. Quebe, "Graph-based observability analysis of bearing-only cooperative localization," *IEEE Transactions on Robotics*, vol. 28, no. 2, pp. 522–529, 2011.
- [63] K. Maes, M. Chatzis, and G. Lombaert, "Observability of nonlinear systems with unmeasured inputs," *Mechanical Systems and Signal Processing*, vol. 130, pp. 378–394, 2019.
- [64] W.-H. Chen, J. Yang, L. Guo, and S. Li, "Disturbance-observer-based control and related methods—An overview," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 2, pp. 1083–1095, 2015.
- [65] K. C. Veluvolu and Y. C. Soh, "High-gain observers with sliding mode for state and unknown input estimations," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 9, pp. 3386–3393, 2009.
- [66] A. Radke and Z. Gao, "A survey of state and disturbance observers for practitioners," in *American Control Conference (ACC)*, 2006, pp. 5183–5188.
- [67] V. C. Chen, F. Li, S.-S. Ho, and H. Wechsler, "Micro-Doppler effect in radar: Phenomenon, model, and simulation study," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 1, pp. 2–21, 2006.
- [68] V. C. Chen, *The micro-Doppler effect in radar*. Artech House, 2019.
- [69] Y. Ma, J. Anderson, S. Crouch, and J. Shan, "Moving object detection and tracking with Doppler LiDAR," *Remote Sensing*, vol. 11, no. 10, p. 1154, 2019.
- [70] W. Sun, Z. Pang, W. Huang, Y. Ji, and Y. Dai, "Vessel velocity estimation and tracking from Doppler echoes of T/RR composite compact HFSWR," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 4427–4440, 2021.
- [71] D. Kellner, M. Barjenbruch, J. Klappstein, J. Dickmann, and K. Dietmayer, "Instantaneous full-motion estimation of arbitrary objects

- using dual Doppler radar," in *IEEE Intelligent Vehicles Symposium Proceedings*, 2014, pp. 324–329.
- [72] D. Kellner, M. Barjenbruch, K. Dietmayer, J. Klappstein, and J. Dickmann, "Instantaneous lateral velocity estimation of a vehicle using Doppler radar," in *Proceedings of the 16th International Conference on Information Fusion*, 2013, pp. 877–884.
- [73] D. Kellner, M. Barjenbruch, J. Klappstein, J. Dickmann, and K. Dietmayer, "Instantaneous ego-motion estimation using Doppler radar," in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, 2013, pp. 869–874.
- [74] S. Agarwal, K. Mierle, and The Ceres Solver Team, "Ceres Solver," 10 2023. [Online]. Available: <https://github.com/ceres-solver/ceres-solver>
- [75] P. J. Rousseeuw and M. Hubert, "Robust statistics for outlier detection," *Wiley interdisciplinary reviews: Data mining and knowledge discovery*, vol. 1, no. 1, pp. 73–79, 2011.
- [76] M. B. Peterson, P. C. Lusk, and J. P. How, "MOTLEE: Distributed mobile multi-object tracking with localization error elimination," *arXiv preprint arXiv:2304.12175*, 2023.
- [77] A. Cunningham, M. Paluri, and F. Dellaert, "DDF-SAM: Fully distributed SLAM using constrained factor graphs," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2010, pp. 3025–3030.
- [78] A. Cunningham, V. Indelman, and F. Dellaert, "DDF-SAM 2.0: Consistent distributed smoothing and mapping," in *2013 IEEE international conference on robotics and automation*. IEEE, 2013, pp. 5220–5227.
- [79] D. Fox, W. Burgard, H. Kruppa, and S. Thrun, "A probabilistic approach to collaborative multi-robot localization," *Autonomous robots*, vol. 8, pp. 325–344, 2000.
- [80] S. I. Roumeliotis and I. M. Rekleitis, "Propagation of uncertainty in cooperative multirobot localization: Analysis and experimental results," *Autonomous Robots*, vol. 17, no. 1, pp. 41–54, 2004.
- [81] L. C. Carrillo-Arce, E. D. Nerurkar, J. L. Gordillo, and S. I. Roumeliotis, "Decentralized multi-robot cooperative localization using covariance intersection," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2013, pp. 1412–1417.
- [82] E. R. Boroson, R. Hewitt, N. Ayanian, and J.-P. de la Croix, "Inter-robot range measurements in pose graph optimization," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 4806–4813.
- [83] R. Liu, Z. Deng, Z. Cao, M. Shalihan, B. P. L. Lau, K. Chen, K. Bhowmik, C. Yuen, and U.-X. Tan, "Distributed ranging SLAM for multiple robots with Ultra-WideBand and odometry measurements," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022, pp. 13 684–13 691.
- [84] R. C. Smith and P. Cheeseman, "On the representation and estimation of spatial uncertainty," *The international journal of Robotics Research*, vol. 5, no. 4, pp. 56–68, 1986.
- [85] R. Smith, M. Self, and P. Cheeseman, *Estimating Uncertain Spatial Relationships in Robotics*. Berlin, Heidelberg: Springer-Verlag, 1990, p. 167–193.
- [86] J. A. Castellanos and J. D. Tardos, *Mobile robot localization and map building: A multisensor fusion approach*. Springer Science & Business Media, 2012.
- [87] J. Sola, "Quaternion kinematics for the error-state Kalman filter," *arXiv preprint arXiv:1711.02508*, 2017.
- [88] A. Barrau and S. Bonnabel, "Invariant Kalman filtering," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, pp. 237–257, 2018.
- [89] S. Hauberg, F. Lauze, and K. S. Pedersen, "Unscented Kalman filtering on Riemannian manifolds," *Journal of mathematical imaging and vision*, vol. 46, pp. 103–120, 2013.
- [90] M. Brossard, A. Barrau, and S. Bonnabel, "A code for unscented Kalman filtering on manifolds (UKF-M)," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 5701–5708.
- [91] C. Zhang, A. Taghvaei, and P. G. Mehta, "Feedback particle filter on Riemannian manifolds and matrix Lie groups," *IEEE Transactions on Automatic Control*, vol. 63, no. 8, pp. 2465–2480, 2017.
- [92] K. Li, F. Pfaff, and U. D. Hanebeck, "Unscented dual quaternion particle filter for SE(3) estimation," *IEEE Control Systems Letters*, vol. 5, no. 2, pp. 647–652, 2020.
- [93] Y. Tao and S. S.-T. Yau, "Outlier-robust iterative extended Kalman filtering," *IEEE Signal Processing Letters*, vol. 30, pp. 743–747, 2023.
- [94] G. Agamennoni, P. Furgale, and R. Siegwart, "Self-tuning M-estimators," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 4628–4635.
- [95] T. Pfeifer and P. Protzel, "Expectation-maximization for adaptive mixture models in graph optimization," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2019, pp. 3151–3157.
- [96] T. Pfeifer, S. Lange, and P. Protzel, "Advancing mixture models for least squares optimization," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3941–3948, 2021.
- [97] M. Ramezani, M. Mattamala, and M. Fallon, "AEROS: AdaptivE RObust least-squares for graph-based SLAM," *Frontiers in Robotics and AI*, vol. 9, p. 789444, 2022.
- [98] D. A. Allan and J. B. Rawlings, *Moving Horizon Estimation*. Cham: Springer International Publishing, 2019, pp. 99–124.
- [99] C. V. Rao, J. B. Rawlings, and D. Q. Mayne, "Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations," *IEEE transactions on automatic control*, vol. 48, no. 2, pp. 246–258, 2003.
- [100] J. B. Rawlings and L. Ji, "Optimization-based state estimation: Current status and some new results," *Journal of Process Control*, vol. 22, no. 8, pp. 1439–1444, 2012.
- [101] J. Civera, O. G. Grasa, A. J. Davison, and J. M. Montiel, "1-point RANSAC for extended Kalman filtering: Application to real-time structure from motion and visual odometry," *Journal of field robotics*, vol. 27, no. 5, pp. 609–631, 2010.
- [102] M. B. Alatise and G. P. Hancke, "Pose estimation of a mobile robot based on fusion of IMU data and vision data using an extended Kalman filter," *Sensors*, vol. 17, no. 10, 2017.
- [103] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. Mineola, NY: Dover Publications, Inc., 2007.
- [104] M. A. Skoglund, G. Hendeby, and D. Axehill, "Extended Kalman filter modifications based on an optimization view point," in *18th International Conference on Information Fusion (Fusion)*, 2015, pp. 1856–1861.
- [105] L. Carlone, M. K. Ng, J. Du, B. Bona, and M. Indri, "Rao-Blackwellized particle filters multi robot SLAM with unknown initial correspondences and limited communication," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2010, pp. 243–249.
- [106] L. Carlone, M. Kaouk Ng, J. Du, B. Bona, and M. Indri, "Simultaneous localization and mapping using rao-blackwellized particle filters in multi robot systems," *Journal of Intelligent & Robotic Systems*, vol. 63, pp. 283–307, 2011.
- [107] S. Saeedi, M. Trentini, M. Seto, and H. Li, "Multiple-robot simultaneous localization and mapping: A review," *Journal of Field Robotics*, vol. 33, no. 1, pp. 3–46, 2016.
- [108] A. Torres-González, J. R. Martinez-de Dios, and A. Ollero, "Range-only SLAM for robot-sensor network cooperation," *Autonomous Robots*, vol. 42, no. 3, pp. 649–663, 2018.
- [109] N. Funabiki, B. Morrell, J. Nash, and A.-a. Agha-mohammadi, "Range-aided pose-graph-based SLAM: Applications of deployable ranging beacons for unknown environment exploration," *IEEE Robotics and Automation Letters*, vol. 6, no. 1, pp. 48–55, 2020.
- [110] 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.
- [111] E. R. Boroson, R. Hewitt, N. Ayanian, and J.-P. de la Croix, "Inter-robot range measurements in pose graph optimization," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 4806–4813.
- [112] R. T. Rodrigues, N. Tsigiakas, A. Pascoal, and A. P. Aguiar, "Online range-based SLAM using B-spline surfaces," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1958–1965, 2021.
- [113] A. Papalia, A. Fishberg, B. W. O'Neill, J. P. How, D. M. Rosen, and J. J. Leonard, "Certifiably correct range-aided SLAM," *arXiv preprint arXiv:2302.11614*, 2023.
- [114] A. Papalia, J. Morales, K. J. Doherty, D. M. Rosen, and J. J. Leonard, "Score: A second-order conic initialization for range-aided slam," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 10 637–10 644.
- [115] F. Dümbgen, C. Holmes, B. Agro, and T. D. Barfoot, "Toward globally optimal state estimation using automatically tightened semidefinite relaxations," *arXiv preprint arXiv:2308.05783*, 2023.
- [116] C. Holmes, F. Dümbgen, and T. D. Barfoot, "On semidefinite relaxations for matrix-weighted state-estimation problems in robotics," *arXiv preprint arXiv:2308.07275*, 2023.
- [117] A. Goudar, F. Dümbgen, T. D. Barfoot, and A. P. Schoellig, "Optimal initialization strategies for range-only trajectory estimation," *arXiv preprint arXiv:2309.09011*, 2023.

- [118] C. Wang, H. Zhang, T.-M. Nguyen, and L. Xie, "Ultra-wideband aided fast localization and mapping system," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 1602–1609.
- [119] Y. Song, M. Guan, W. P. Tay, C. L. Law, and C. Wen, "UWB/LiDAR fusion for cooperative range-only SLAM," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2019, pp. 6568–6574.
- [120] T.-M. Nguyen, S. Yuan, M. Cao, T. H. Nguyen, and L. Xie, "VIRAL SLAM: Tightly coupled Camera-IMU-UWB-Lidar SLAM," *arXiv preprint arXiv:2105.03296*, 2021.
- [121] R. Liu, Z. Deng, Z. Cao, M. Shalihan, B. P. L. Lau, K. Chen, K. Bhowmik, C. Yuen, and U.-X. Tan, "Distributed ranging slam for multiple robots with ultra-wideband and odometry measurements," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022, pp. 13 684–13 691.
- [122] T. H. Nguyen, T.-M. Nguyen, and L. Xie, "Range-focused fusion of camera-IMU-UWB for accurate and drift-reduced localization," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1678–1685, 2021.
- [123] ——, "Flexible and resource-efficient multi-robot collaborative visual-inertial-range localization," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 928–935, 2021.
- [124] C. Zhang, X. Ma, and P. Qin, "LiDAR-IMU-UWB-based collaborative localization," *World Electric Vehicle Journal*, vol. 13, no. 2, p. 32, 2022.
- [125] T.-M. Nguyen, S. Yuan, M. Cao, Y. Lyu, T. H. Nguyen, and L. Xie, "NTU VIRAL: A visual-inertial-ranging-lidar dataset, from an aerial vehicle viewpoint," *The International Journal of Robotics Research*, vol. 41, no. 3, pp. 270–280, 2022.
- [126] G. Delama, F. Shamsfakhr, S. Weiss, D. Fontanelli, and A. Fomasier, "UVIO: An UWB-aided visual-inertial odometry framework with bias-compensated anchors initialization," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2023, pp. 7111–7118.
- [127] J. Wang, P. Gu, L. Wang, and Z. Meng, "RVIO: An effective localization algorithm for range-aided visual-inertial odometry system," *IEEE Transactions on Intelligent Transportation Systems*, 2023.
- [128] J. Hu, Y. Li, Y. Lei, Z. Xu, M. Lv, and J. Han, "Robust and adaptive calibration of UWB-aided vision navigation system for UAVs," *IEEE Robotics and Automation Letters*, 2023.
- [129] C. Hu, P. Huang, and W. Wang, "Tightly coupled visual-inertial-UWB indoor localization system with multiple position-unknown anchors," *IEEE Robotics and Automation Letters*, 2023.
- [130] H.-Y. Lin and J.-R. Zhan, "GNSS-denied UAV indoor navigation with UWB incorporated visual inertial odometry," *Measurement*, vol. 206, p. 112256, 2023.
- [131] A. Goudar, W. Zhao, and A. P. Schoellig, "Range-visual-inertial sensor fusion for micro aerial vehicle localization and navigation," *IEEE Robotics and Automation Letters*, vol. 9, no. 1, pp. 683–690, 2023.
- [132] J. K. Verma and V. Ranga, "Multi-robot coordination analysis, taxonomy, challenges and future scope," *Journal of intelligent & robotic systems*, vol. 102, pp. 1–36, 2021.
- [133] R. Brooks, "Visual map making for a mobile robot," in *IEEE International Conference on Robotics and Automation (ICRA)*, vol. 2, 1985, pp. 824–829.



Hyunjae Sim HYUNJAE SIM received his B.S. degree in Electrical Engineering from Inha University in Incheon, South Korea, in 2022. He is currently pursuing a Ph.D. in Electrical and Computer Engineering at Inha University. His research interests include optimization, optimal control, and multi-agent distributed control, with applications to robotics and autonomous vehicles.



Seungwon Nam SEUNGWON NAM received a B.S. degree in Mechanical Engineering from Inha University in Incheon, South Korea, in 2023. He is currently pursuing an M.S. in Electrical and Computer Engineering at Inha University. His research interests include multi-robot systems, autonomous systems, and control theory.



Yonghee Kim YONGHEE KIM received a B.S. degree in Electrical Engineering from Inha University in Incheon, South Korea, in 2024. He is currently pursuing an M.S. in Electrical and Computer Engineering at Inha University. His research interests include multi-robot systems, autonomous systems, and visual perception.



Jae Heo JAE HEO received a B.S. degree in Control and Instrumentation Engineering from Kangwon National University in Kangwon, South Korea, in 2024. He is currently pursuing an M.S. in Electrical and Computer Engineering at Inha University. His research interests include vision-aware AI-based robot control, combining model predictive control and deep learning techniques.



Kwang-Ki K. Kim KWANG-KI K. KIM received his M.S. and Ph.D. degrees in Aerospace Engineering from the University of Illinois at Urbana-Champaign, Urbana-Champaign, IL, USA, in 2009 and 2013, respectively. From 2013 to 2016, he was a postdoctoral fellow at the School of Electrical and Computer Engineering, Georgia Institute of Technology, GA, USA. Since 2017, he has been an associate professor in the Department of Electrical Engineering at Inha University. His research interests include control, learning, and optimization for autonomous



Gihun Shin GIHUN SHIN received his B.S. degree in Aerospace Engineering and M.S. degree in Electrical and Computer Engineering from Inha University in Incheon, South Korea, in 2022 and 2024, respectively. He is currently employed as a research engineer at Korea Aerospace Industries (KAI) in Sacheon, South Korea. His research interests focus on sensor fusion for the estimation of UAVs and multi-robot systems.

systems and robotics.