IEEE Transactions on Robotics (T-RO) paper, presented at ICRA 2026, Vienna, Austria. Cite as T-RO paper.

Robust and Scalable Multi-Robot Localization Using Stereo UWB Arrays

Hanying Zhao, *Member, IEEE*, Lingwei Xu, *Student Member, IEEE*, Yi Li, *Student Member, IEEE*, Feiyang Wen, *Student Member, IEEE*, Haoran Gao, Changwu Liu, Jincheng Yu, *Member, IEEE*, Yu Wang, *Fellow, IEEE*, and Yuan Shen, *Senior Member, IEEE*

***Abstract*—In environments where robots operate with limited global navigation satellite system accessibility, ultra-wideband (UWB) localization technology is a popular auxiliary solution to assist visual–inertial odometry systems. However, current UWB approaches lack 3-D pairwise localization capability and suffer from rapidly declining localization update rates as the network scales, limiting their effectiveness for swarm robotic applications. This article presents a novel UWB sensor that enables 3-D pairwise localization and a localization scheme that can deliver robust, scalable, and accurate position awareness for multi-robot systems. Our approach begins with calibrating intrinsic UWB errors from hardware deviations and propagation effects, yielding high-accuracy distance and direction measurements. Using these measurements, we perform distributed relative localization through inter- and intra-node cooperation by integrating UWB and inertial measurement unit data. To enable swarm-scale operation, our platform implements the signal-multiplexing network ranging protocol to maximize update rates and network capacity. Experimental results show that our approach achieves centimeter-level localization accuracy at high update rates (100 Hz with UWB only), validating its robustness, scalability, and accuracy for robotic applications.**

***Index Terms*—3-D bearing, clock synchronization, multi-robot localization, ranging, relative localization, ultra-wideband (UWB).**

Received 3 November 2024; revised 20 May 2025; accepted 2 June 2025. Date of publication 10 July 2025; date of current version 1 October 2025. This work was supported in part by the National Natural Science Foundation of China under Grant 62401318, Grant 62271285, and Grant 62203257; in part by Tsinghua University-Meituan Joint Institute for Digital Life; and in part by Shanghai AI Laboratory. This article was recommended for publication by Associate Editor F. Morbidi and Editor M. Schwager upon evaluation of the reviewers' comments. *(Hanying Zhao, Lingwei Xu, and Yi Li contributed equally to this work.) (Corresponding author: Yuan Shen.)*

Hanying Zhao was with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China. She is now with the Division of Information Science and Engineering at KTH Royal Institute of Technology, SK-100 44 Stockholm, Sweden (e-mail: hanying@kth.se).

Lingwei Xu, Yi Li, Feiyang Wen, Haoran Gao, Changwu Liu, Jincheng Yu, and Yu Wang are with the Department of Electronic Engineering, Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China (e-mail: xlw22@mails.tsinghua.edu.cn; li-yi20@mails.tsinghua.edu.cn; wenfy24@mails.tsinghua.edu.cn; ghr22@mails.tsinghua.edu.cn; liucw_ee@tsinghua.edu.cn; yu-jc@tsinghua.edu.cn; yu-wang@tsinghua.edu.cn).

Yuan Shen is with the Department of Electronic Engineering, Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China, and also with the Shanghai AI Laboratory, Shanghai 201112, China (e-mail: shenyuan_ee@tsinghua.edu.cn).

The video is available at the following link: https://cloud.tsinghua.edu.cn/f/d9bae49585f54d6e8563/

Digital Object Identifier 10.1109/TRO.2025.3587854

I. Introduction

MULTI-ROBOT systems can substantially enhance capabilities over single-robot systems, particularly in perception and action [1], [2], [3]. Achieving these benefits relies on effective localization, where each robot determines its position and the relative locations of others to facilitate seamless environmental interaction [3], [4], [5]. To broaden the application fields of multi-robot systems, such as indoor services and outdoor low-altitude economy, precise, self-organized localization without reliance on predeployed infrastructure, such as global navigation satellite system (GNSS) is necessary.

Self-organized multi-robot localization is typically accomplished by integrating multimodal sensors [3]. Visual–inertial odometry, which combines visual and inertial measurements, is widely employed to deliver robust and accurate localization and mapping for mobile robotic systems [4], [5], [6], [7]. Several visual–inertial odometry (VIO) systems employ GNSS for absolute positioning to initialize state estimates and correct accumulated errors. However, GNSS reliability degrades in challenging environments across multiple robot application domains [3], [8], [9], [10]. Network localization is increasingly used to augment VIO systems [11], [12], [13], [14], [15], [16], [17]. This approach leverages wireless signals to measure interrobot ranges or relative angles, enabling position estimation through trilateration, multilateration, or triangulation. Key wireless technologies encompass ultra-wideband (UWB) [11], [12], [13], [14], Wi-Fi [15], laser [16], and radio frequency (RF) [17]. Among these, UWB is preferred for its high delay resolution and robustness in challenging propagation environments.

Despite their effectiveness, existing UWB localization methods for robotic applications face three critical challenges that complicate their implementation. The primary limitation is the inability of current UWB sensors to provide high-precision directional measurements. This deficiency forces each robot to maintain wireless connections with a minimum of three neighboring robots to enable localization, thus increasing communication overheads and reducing robustness in complex propagation environments where packet loss is unavoidable. Second, due to clock asynchronism among UWB devices, two-way ranging (TWR) is widely performed to achieve joint ranging and clock synchronization. However, this scheme quickly becomes infeasible as the measurement frequency drops sharply with the growth of the network scale; this problem is further exacerbated when robots move quickly, where the

©2026 IEEE

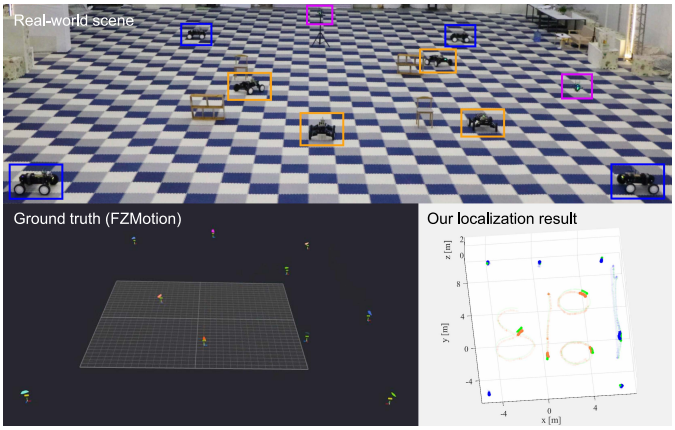


Fig. 1. Snapshot of the real-world experiment, where the ground truth in the bottom left is collected using a motion capture system, and the bottom right shows our localization results. Our method can achieve centimeter-level localization accuracy with an update rate of 100 Hz.

delayed measurements incur significantly accumulated errors. The scalability is constrained. Third, hardware limitations and complex propagation environments cause significant measurement errors that require resolution.

To address the above challenges, we propose a new UWB sensor and a relative localization scheme for multi-robot systems. The effectiveness of the proposed approach is validated through real-world experiments, as shown in Fig. 1. The architecture and features of our method are shown in Figs. 2 and 3, respectively. The main contributions of this article are as follows.

1) We design a *new localization sensor featuring UWB arrays*, as shown in Fig. 4. This stereo UWB array is lightweight and compatible with microrobots. Using this array, we can measure both 3-D directions and distances between pairs of nodes, achieving pairwise relative localization without relying on external infrastructures. The UWB array imposes minimal restrictions on wireless connectivity, thus enhancing the robustness of the localization system in complex propagation environments.
2) We propose *a new localization method that extracts high-accuracy position estimates from UWB signals.* Raw UWB measurements are susceptible to hardware deviations and complex propagation environments; thus, our approach first calibrates these intrinsic errors to achieve precise direction and distance estimations. We then present a distributed localization scheme that leverages UWB and inertial measurement unit data to determine the relative positions between robots. This scheme ensures real-time alignment of local reference frames with high robustness, efficiency, and accuracy.
3) We develop *a platform that leverages the proposed localization scheme on our UWB arrays*. We propose a new UWB transmission flow and new UWB data frames to perform signal-multiplexing network ranging (SM-NR), enabling network measurement and clock synchronization with minimal signal transmissions. An implementation attains centimeter-level localization accuracy at an update rate of 100 Hz (UWB only), demonstrating the scalability and timeliness for accommodating larger swarm sizes.

Our method is designed to serve as a foundational infrastructure for multi-robot systems, equipping them with reliable position awareness to navigate safely, collaborate effectively, and accomplish complex tasks to revolutionize traditional practices in diverse domains.

Notation: In this article, we denote variables, vectors, and matrices using italic letters x, bold italic letters $\boldsymbol{x}$, and bold capital italic letters $\boldsymbol{X}$, respectively; random variables, random vectors, and random matrices are represented by sans serif letter x, bold letters $\mathbf{x}$, and bold capital letters $\mathbf{X}$, respectively.

II. Related Work

A. Localization for Swarm Robotics

Current research in multi-robot localization focuses on using relative relationships among robots to offset the drift encountered in self-positioning odometry systems [13], [18], [19]. In addition, there is a growing body of research employing relative pose estimation for building and merging maps in multi-robot systems [20], [21], [22]. These studies implement a graph optimization approach to determine the state of each robot, where the relative pose measurements between robots are conceptualized as factors within the graph. Generally, there are three ways to generate relative factors between robots as follows.

1) Feature-Based: Some research employs neural networks to extract place descriptors. These descriptors are then matched between robots to identify key frames in different robots that have encountered the same location. Place descriptors can be extracted from lidar, as demonstrated in [20], and camera, as shown in [23]. Beyond using place descriptors to determine if robots are in the same location, key-point matching can also be employed to calculate the relative poses between robots. Previous works extract key points in key frames as landmarks and match key points across different keyframes from various robots [18], [22]. However, these methods require different robots to observe the same scene.

2) Visual Object Detection: If one robot is visible within the camera view of another robot, the relative pose between the detecting robot and the detected robot can be estimated based on the frames that capture the detected robot [13], [18], [24]. However, this method is contingent on the robots being within visual range of each other.

3) UWB Measurements: Both of the aforementioned methods depend on the visibility of the robot swarm. This dependence on the robots' relative positions diminishes the applicability of these techniques.

To address this issue, previous work has introduced UWB sensors in swarm localization [13], [18], [24]. The UWB technology contributes by adding distance measurements between robots as factors in the optimization graph. However, these UWB sensors are limited to providing 1-D distance measurements, which may result in undetermined problems in 3-D graph

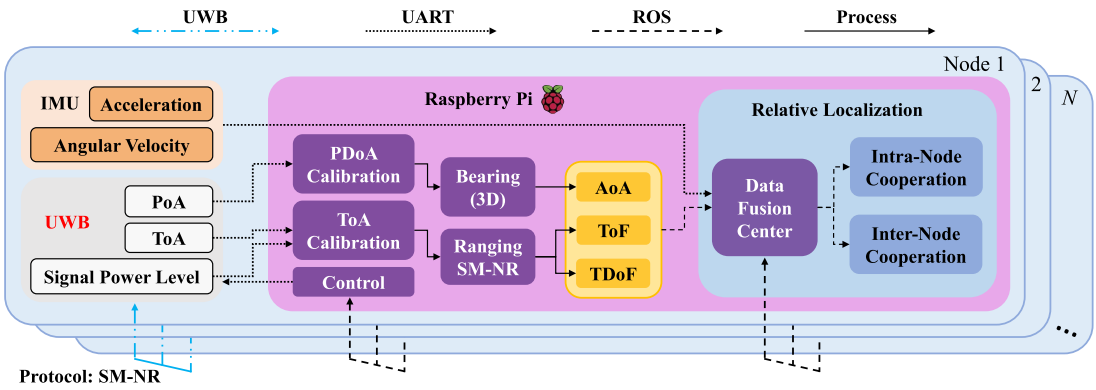


Fig. 2. Block diagram illustrating the full pipeline of the proposed localization scheme. This diagram includes the hardware realization and illustrates the data exchange between the various components involved.

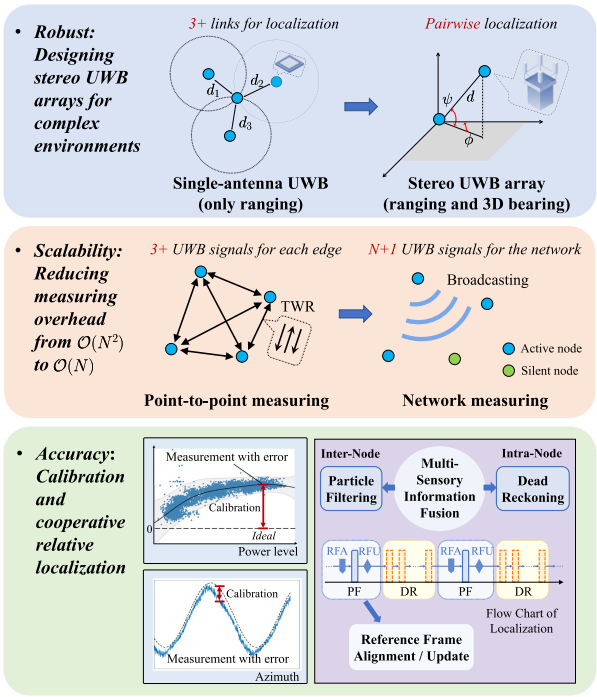


Fig. 3. Proposed localization method can provide robust, scalable, and accurate location-awareness for multi-robot systems: a novel UWB sensor that enables pairwise relative localization without external infrastructure, a wireless protocol that ensures scalability and timeliness for accommodating larger swarms, and advanced signal processing methods for high-accuracy position estimation.

optimization. Thus, for these methods to enable 3-D swarm localization, it is essential to either increase the number of groups within the cluster or optimize timing in tandem with ego-positioning odometry.

B. UWB Localization

This section provides an overview of UWB positioning in multi-robot systems.

1) 3-D Angle Estimation and Phase Calibration: Current lightweight UWB sensors face challenges in delivering reliable

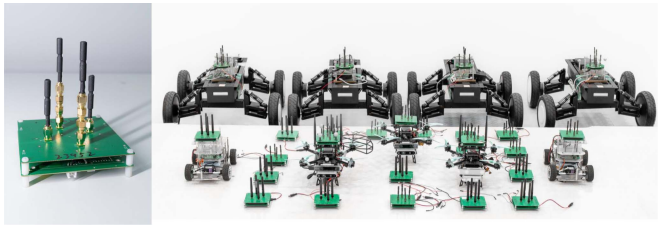


Fig. 4. Stereo UWB arrays.

3-D directional measurements, resulting in high communication overhead and reduced robustness in complex propagation environments. These limitations stem from significant errors in phase-of-arrival (PoA) measurements, which can arise due to phase center deviations, uncertainties in antenna positions, phase ambiguity, and measurement noise [25], [26], [27]. Although several phase calibration methods have achieved great success [28], [29], [30], [31], they focus on linear and planar arrays, with only a few studies addressing stereo arrays. In fact, stereo arrays are more suitable for 3-D scenarios, as they achieve higher resolution in elevation angle estimation. Extending methods for linear/planar arrays to stereo arrays comes with three primary limitations:

a) Works like [28] focus on arrays with two antennas, which restricts their applicability to arrays with more antennas.
b) Training deep neural networks for 3-D angle estimation (e.g., [29], [30], [31]) demands a large amount of accurately labeled data, whose acquisition poses significant practical and economic challenges.
c) High computational complexity hinders the application on end sides.

The work in [32] was based on stereo arrays but exhibits high computational complexity due to the grid search on the spherical surface. To our knowledge, an effective and computationally efficient method for 3-D angle estimation using UWB signals requires further exploration.

2) Joint Ranging and Clock Synchronization: UWB signals can achieve centimeter-level ranging accuracy but require precise clock synchronization between the nodes. The TWR

paradigm is the most popular method for joint ranging and clock synchronization, which measures time of flights (ToFs) from round-trip time (RTT) in a point-to-point manner [33], [34], [35]. However, this paradigm is not scalable because the signal overhead increases quadratically as the number of nodes grows. To address this challenge, several passive ranging methods are proposed [36], [37], [38], which estimates the difference between distances from multiple transmitters by calculating the time difference of flights (TDoFs). In this way, the receiver positions can be determined without signal transmission. However, the measurement frequency remains inadequate. To illustrate, flying drones typically require localization update rates exceeding 100 Hz for safe operation. Although existing approaches support pairwise ranging at tens of hertz, their performance will collapse to sub-1-Hz frequencies when scaled to full network size in large swarms. This scalability limitation will cause low update rates and introduce substantial accumulating errors. Moreover, while existing methods [33], [34], [35], [36], [37], [38] demonstrate effectiveness in static environments, they suffer performance degradation in dynamic scenarios due to unmodeled motion dynamics.

3) Ranging Error Calibration: Hardware-inherent biases, and channel-induced factors, including antenna delays [39], non-line-of-sight (NLOS) conditions [40], and multipath effects, will significantly affect ranging accuracy. Various statistical models have been introduced to relate those random influential factors to the ranging bias. Preter et al. [41] utilized a linear model to estimate the ranging error based on the received signal strength. Alternatively, in [42], the range measurement was modeled as a Gaussian process with respect to the poses of the antennas. The authors in [39] and [43] used deep learning models to regress the ranging errors from the received channel impulse response (CIR), while Wang et al. [40] proposed to model the ranging bias using a support vector machine, where several waveform statistics extracted from the CIR are chosen as the input features. Although prior methods have effectively improved the accuracy of UWB measurements, they will incur a large training overhead as the network scales up, since their calibration complexity is $\mathcal{O}(N^2)$ for networks with N nodes.

4) Relative Localization: Instead of concentrating on absolute positions, relative localization focuses on the relative configuration that describes the "shape" of the network, which is pervasive in multi-robot applications, such as autonomous driving and formation control [44]. Theoretical frameworks have been established by introducing the relative error for state estimates based on the concept of the equivalent state class [45]. For the algorithm designs of relative localization, several spatial cooperation algorithms are developed for static networks, such as multidimensional scaling (MDS), Chan algorithm, least squares (LS), and semidefinite programming algorithms [46], [47], [48], [49], [50], [51]. In dynamic networks, temporal cooperation has also gained increasing attention with emerging multisource information fusion schemes [52], [53], [54], [55], [56], [57]. Filtering and graphical models are adopted for spatiotemporal cooperation [52], [53], [54], while the characteristics of relative localization should be further explored. In [55], an infrastructure-free relative localization system was realized by fusing the range and inertial measurements with map information, while environmental knowledge is essential to imposing positional constraints. In [56], a dead reckoning (DR) framework with particle-based intermittent belief propagation was proposed; however, it assumes that the initial positions are known, which may not satisfy in practice. Moreover, the characteristics of relative localization are not fully exploited. In [57], a backbone-listener scheme was developed with DR and error-state Kalman filtering for information fusion, while the implementations are constrained in 2-D space.

C. Difference With Previous Work

Key distinctions from our prior work include: for bearing estimation, Li et al. [57] developed a planar UWB array for azimuth angle estimation but was limited by unaddressed phase errors and lacked elevation bearing capability. This article advances this by presenting a stereo UWB array enabling high-accuracy elevation angle estimation and a joint azimuth-elevation estimation method that incorporates phase calibration and ambiguity resolution to improve accuracy and reliability. Regarding ranging, while the principle of SM-NR was presented in [58] and [59], these works did not mitigate timestamp measurement errors and were only validated by simulations. In this article, we develop a hardware version of SM-NR, perform it on designed UWB arrays, and further validate its performance through real-world experiments. The added features in the hardware version include network bearing, timestamp calibration, optimized UWB signal transmission flow, and new UWB data frames. For localization, we proposed a relative localization scheme for 2-D scenarios in [57] and [59]. This article extends this approach to distributed 3-D scenarios using ranging and 3-D bearing measurements from SM-NR. Moreover, we present hardware implementation details and extensive real-world field experiments to show the robustness, scalability, and accuracy of our localization scheme.

III. Overview

Herein, we present a new UWB sensor that enables 3-D pairwise localization and a localization scheme that can achieve robust, scalable, and accurate multi-robot localization without relying on external infrastructures. Given the asynchronous nature of devices, we begin with performing network measurement via multiplexing signals (see Section IV), which achieves network ranging and clock synchronization with minimal signal transmissions. Second, we address phase errors and present an online 3-D angle estimation method (see Section V). Next, we eliminate clock errors and coordinate timestamp errors for high-accuracy ranging (Section VI). With distance and direction information, relative positioning is achieved by integrating UWB measurements and inertial measurement unit (IMU) data with reference frame alignment (RFA) and update (See Section VII). Details on hardware implementation are presented in Section VIII.

Our UWB-based localization system improves upon current UWB approaches by enabling 3-D bearing for pairwise localization, optimizing wireless protocols for

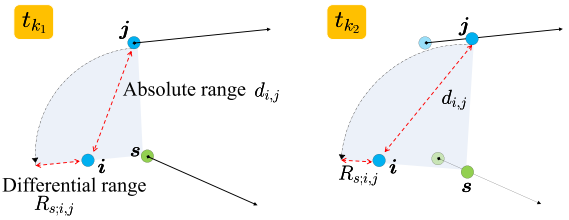


Fig. 5. Geometrical relationship between the absolute range and the differential range in dynamic scenarios.

scalability, correcting timestamp and phase errors, and facilitating distributed relative localization. These features collectively enhance the accuracy, scalability, and robustness of the proposed localization system in complex propagation environments.

IV. Preliminaries

A. Active-Silent Localization

Consider a 3-D multi-robot system consisting of N nodes, where each node represents either a robot or any other device that can assist in the localization process.[1] Denote the node index set as $\mathcal{N} \triangleq \{1, 2, \ldots, N\}$. Each node is equipped with a UWB array for both ranging and bearing measurements for pairwise localization.

We present an active-silent localization paradigm in which several nodes actively transmit UWB signals while others work only as recipients. For brevity, we refer to the transmitting nodes as active nodes and the nodes that receive only as silent nodes. Suppose that there are N_a active nodes and N_s silent nodes, where $\mathcal{N}_\mathrm{a} \triangleq \{1, 2, \ldots, N_\mathrm{a}\}$ and $\mathcal{N}_\mathrm{s} \triangleq \{N_\mathrm{a}+1, N_\mathrm{a}+2, \ldots, N_\mathrm{a}+N_\mathrm{s}\}$ are the index sets of active and silent nodes, respectively, and $\mathcal{N} = \mathcal{N}_\mathrm{a} \cup \mathcal{N}_\mathrm{s}$. The number of active nodes and the selection of nodes operating in active mode are predetermined.[2] Our method allows silent nodes to achieve comparable localization performance with active ones.

When a UWB signal is transmitted, the UWB chips on the transmitter side provide the time of departures (ToDs), while on the receiver side, they deliver the time of arrivals (ToAs), PoAs, and received signal strength. The PoAs are used for bearing estimation, while the ToAs and ToDs are used for ranging. For bearing, all nodes estimate azimuth and elevation angle of arrivals (AoAs). For ranging, active and silent nodes estimate the ToFs and TDoFs to calculate absolute ranges and differential ranges, respectively. The geometrical relationship between these two ranges is shown in Fig. 5. Let $d_{i,j}$ and $T_\mathrm{oF}(i,j)$, respectively, denote the absolute range and the ToF between nodes i and j, which satisfies

$$d_{i,j} = cT_\mathrm{oF}(i,j) \tag{1}$$

where c is the signal transmission speed. For node pairs (i,s) and (j,s), the differential range between them is defined as $R_{s;i,j} = d_{i,s} - d_{j,s}$ and satisfies

$$R_{s;i,j} = cT_\mathrm{d}(s;i,j) \tag{2}$$

where TDoF $T_\mathrm{d}(s;i,j) = T_\mathrm{oF}(i,s) - T_\mathrm{oF}(j,s)$. Active nodes estimate their absolute ranges with other active ones, while silent nodes only estimate the differential ranges between silent-active node pairs. This is because silent nodes do not engage in round-trip measuring processes, resulting in their absolute ranges to other nodes not being measured.

B. Protocol of Network Measurement

We perform network measurement using the SM-NR proposed in our previous work [58], [59], which achieves clock synchronization with a minimum of signal overhead. The meaning of "signal-multiplexing" is that it executes multiple round-trip measuring processes simultaneously via multiplexing signals through broadcast communication.

We explain the protocol of SM-NR. For simplicity, assume that the order of signal transmission aligns with the index numbers assigned to each active node. The active node 1 initiates the measurement process by transmitting the first measuring signal. After a predefined delay, the active node 2 transmits the second measurement signal. Each active node transmits signals in sequence, each with a predefined delay. When active node N_a has transmitted its signal, one complete run finishes, and a new run commences with node 1 transmitting. Silent nodes only receive signals from the active ones.

Since the network may not be fully connected in practice, denote the neighbor sets within the communication range of node n as $\mathcal{N}^{(n)}$. Then, at time instant k when active node i ($i \in \mathcal{N}_\mathrm{a}$) broadcasts a UWB signal, the following parameters are updated by SM-NR:

$$\mathbf{d}_k = \{\hat{\mathsf{d}}_{i,j} | j \in \mathcal{N}^{(i)} \cap \mathcal{N}_\mathrm{a}\} \tag{3a}$$

$$\mathbf{R}_k = \{\widehat{\mathsf{R}}_{s;i,j} | s \in \mathcal{N}^{(i)} \cap \mathcal{N}^{(j)} \cap \mathcal{N}_\mathrm{s},\ j \in \mathcal{N}^{(i)} \cap \mathcal{N}_\mathrm{a}\} \tag{3b}$$

$$\boldsymbol{\Phi}_k = \{\hat{\phi}_{i,j} \,|\, j \in \mathcal{N}^{(i)}\} \tag{3c}$$

$$\boldsymbol{\Psi}_k = \{\hat{\psi}_{i,j} \,|\, j \in \mathcal{N}^{(i)}\} \tag{3d}$$

which represent the absolute range, differential range, azimuth, and elevation angle estimations, respectively.

Definition 1 (Network Measurement Update Rate): The network measurement update rate is defined as the distances and directions that can be updated per measuring signal.

Our method (SM-NR) can significantly enhance the scalability, timeliness, and robustness of relative localization. For illustration, consider a fully connected network. Every wireless signal in SM-NR can update $N_\mathrm{a}-1$ absolute ranges, $N_\mathrm{s}(N_\mathrm{a}-1)$ differential ranges, and $2(N_\mathrm{a}+N_\mathrm{s}-1)$ angles. In contrast, TWR-based methods [33], [34], [35], [60] update $\leq 1/3$ absolute ranges and $2(N_\mathrm{a}+N_\mathrm{s}-1)$ angles. High update rates can ensure timely measurement data and minimize errors that accumulate from using outdated measurements. Moreover, high update rates enhance robustness in complex propagation environments, particularly when line-of-sight (LOS) signals are blocked or affected by multipath propagation environments.

[1] Our method can also be applied to localizing single robots when nodes equipped with stereo UWB arrays are available. By measuring distances and directions, the robot's position can be determined through triangulation.

[2] Determining the optimal N_a and dynamically selecting which nodes should operate in active mode will be addressed in our future work.

Regarding scalability, SM-NR requires only $N_\mathrm{a}+1$ signals to determine the network topology. In contrast, existing methods typically require at least $3\binom{N_\mathrm{a}}{2}$ signals to achieve the same goal. This reduction from $\mathcal{O}(N_\mathrm{a}^2)$ to $\mathcal{O}(N_\mathrm{a})$ is advantageous when dealing with large swarm sizes. In fact, $N_\mathrm{a}+1$ is the minimum number necessary [58].

C. Spatial-Temporal Cooperative Localization

To enhance robustness and lower computation burden, distributed localization is preferred for multi-robot systems [9], where each node aims to estimate the relative position of the network consisting of its neighbors within the communication range. The relative position describes the "shape" of the network and the relative error is defined as

$$\varepsilon_\mathrm{r}(\hat{\mathbf{p}}, \boldsymbol{p}) = \min_{\boldsymbol{\Gamma}} \|\boldsymbol{p} - S_{\boldsymbol{\Gamma}}(\hat{\mathbf{p}})\| \tag{4}$$

where $S_{\boldsymbol{\Gamma}}(\hat{\mathbf{p}}) = (\boldsymbol{I}_N \otimes \boldsymbol{R})\hat{\mathbf{p}} + \mathbf{1}_N \otimes \boldsymbol{t}_\mathrm{p}$ is the equivalent geometry of a network position estimation $\hat{\mathbf{p}}$ to eliminate the effects of translation, rotation, and reflection, and $\boldsymbol{\Gamma} = \{\boldsymbol{R}, \boldsymbol{t}_\mathrm{p}\}$ is the transformation parameter with $\boldsymbol{R} \in \mathcal{O}(3), \boldsymbol{t}_\mathrm{p} \in \mathbb{R}^3$. The goal of relative localization is to minimize the relative error.

For node n, the goal of distributed relative localization is to estimate the poses of each node n along with its neighbors in $\mathcal{N}^{(n)}$ in the local reference frame of n. Since no absolute position reference is available, it is important to establish a common consensus on the coordinate datum in real time. To this end, nodes will share wireless measurements and sensing data with one another through wireless communication technologies. Denote $\mathbf{z}_{\mathrm{IMU},k}^{(n)} = [\mathbf{a}_k^{(n)\mathrm{T}}, \boldsymbol{\omega}_k^{(n)\mathrm{T}}]^\mathrm{T}$ as the IMU measurements with acceleration and angular velocity. Then, the complete measurements for the active and silent nodes are

$$\begin{aligned}
\mathbf{z}_k^{(n)} &= \left[\mathbf{d}_k^{(n)\mathrm{T}}, \boldsymbol{\Phi}_k^{(n)\mathrm{T}}, \boldsymbol{\psi}_k^{(n)\mathrm{T}}, \mathbf{z}_{\mathrm{IMU},k}^{(n)\mathrm{T}}\right]^\mathrm{T}, \ n \in \mathcal{N}_\mathrm{a} \\
\mathbf{z}_k^{(m)} &= \left[\mathbf{R}_k^{(m)\mathrm{T}}, \boldsymbol{\Phi}_k^{(m)\mathrm{T}}, \boldsymbol{\psi}_k^{(m)\mathrm{T}}, \mathbf{z}_{\mathrm{IMU},k}^{(m)\mathrm{T}}\right]^\mathrm{T}, \ m \in \mathcal{N}_\mathrm{s}
\end{aligned} \tag{5}$$

where $\mathbf{d}_k^{(n)}$, $\mathbf{R}_k^{(m)}$, $\boldsymbol{\Phi}_k^{(n)}$, and $\boldsymbol{\psi}_k^{(n)}$ can be derived from (3), as measurements related to node n and its neighbors.

V. 3-D Bearing Using UWB Arrays

This section shows our estimation of elevation and azimuth angles in conjunction with the calibration of phase biases and phase ambiguities. As the bearing estimation step is conducted at the node end, the superscript $(\cdot)^{(n)}$, which indicates the observation made by node n, is omitted for brevity.

A. UWB Measurements

Taking antenna 1 as the reference point, the phase difference of arrival (PDoA) between antennas $m \in \mathcal{M} := \{1, 2, \ldots, M\}$ and 1 can be described by[3]

$$\Delta\varphi_m = \mathrm{Wrap}\left(\frac{2\pi}{\lambda}(\boldsymbol{q}_m - \boldsymbol{q}_1)^\mathrm{T}\boldsymbol{u}(\phi,\psi) + \mathsf{b}_m + \mathsf{n}_m\right) \tag{6}$$

[3]This far-field model holds when the distance is larger than 0.5m for UWB signals, which is commonly satisfied in practice.

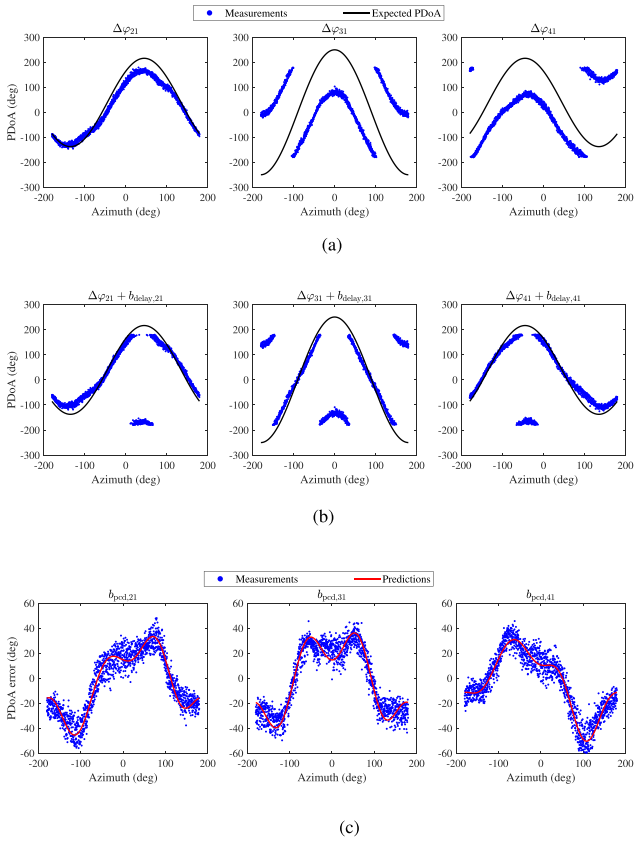


Fig. 6. Received PoAs and PoAs after calibration. (a) Expected PDoA (black solid lines) and raw PDoA measurements (blue dots) corrupted by phase biases and phase ambiguities. (b) Expected PDoA (black solid lines) and PDoA measurements after calibrating the transmission delay b_delay (blue dots). (c) Phase measurements bias induced by the variant phase center deviation b_pcd.

where the operator $\mathrm{Wrap}(\cdot)$ wraps the PDoA to $[-\pi, \pi]$, as phases fall within this interval, $\boldsymbol{q}_m \in \mathbb{R}^3$ denotes the position of antenna m, and

$$\boldsymbol{u}(\phi,\psi) = [\cos(\phi)\cos(\psi)\ \sin(\phi)\cos(\psi)\ \sin(\psi)]^\mathrm{T} \tag{7}$$

with ϕ and ψ being the azimuth and elevation angles, respectively. The notions b_m and n_m represent the difference in phase bias and observation noise, respectively. We can estimate the AoAs based on (6). However, raw PoA data are significantly affected by phase biases and observation noise. Fig. 6(a) shows the raw PDoA data provided by the UWB hardware and the expected PDoA data. Here, the expected PDoA is calculated by $\Delta\varphi_m = \frac{2\pi}{\lambda}(\boldsymbol{q}_m - \boldsymbol{q}_1)^\mathrm{T}\boldsymbol{u}(\phi,\psi)$, which excludes the effects of phase bias, ambiguity, and noises. The significant discrepancies between the raw and expected PDoA data highlight the necessity of phase error calibration and phase ambiguity resolution during direction estimation.

There are no widely accepted models for phase biases and phase observation noises. Phase biases can result from delays between the phase center and the processing chip, phase center variations, mutual coupling effects, and other hardware imperfections. Phase observation noise can arise from complex propagation environments [25], [26], [27]. To enhance spatial resolution and reduce antenna coupling effects, our UWB arrays

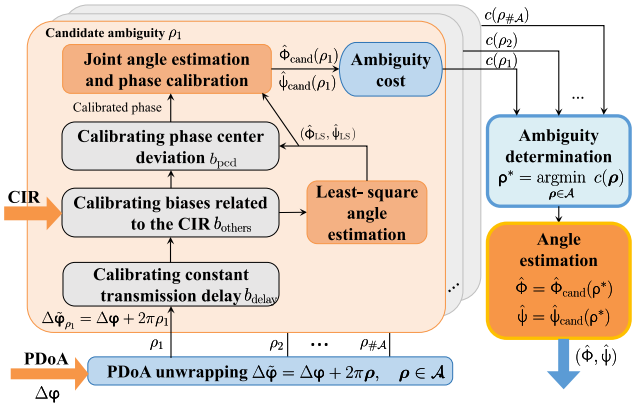


Fig. 7. Implementation of a parallel computing mechanism in our method for simultaneous angle estimation and phase calibration. The cardinality of a set $\mathcal{A}$ is denoted by $\#\mathcal{A}$.

are designed with antenna spacing larger than half the signal wavelength. However, this design choice further exacerbates the phase ambiguity problem, causing the PDoA between two antennas to exceed 2π.

To extract high-accuracy angle estimates from raw UWB PoA data, we implement a parallel computing framework, as shown in Fig. 7. This process begins with unwrapping the raw PoA data across all potential ambiguities. For each candidate ambiguity, we perform joint angle estimation and phase calibration in parallel. The AoA estimate is derived by minimizing the ambiguity cost function across all candidate solutions.

B. Phase Unwrapping

For received PDoA data $\Delta\boldsymbol{\varphi}$ in (6), the possible unwrapped PDoA can be described as

$$\Delta\tilde{\boldsymbol{\varphi}}(\boldsymbol{\rho}) = \Delta\boldsymbol{\varphi} + 2\pi\boldsymbol{\rho}, \quad \boldsymbol{\rho} \in \mathcal{A} \tag{8}$$

where $\mathcal{A}$ is the ambiguity set, defined by

$$\mathcal{A} = \prod_{m\in\mathcal{M}\setminus\{1\}} \mathcal{A}_m = \prod_{m\in\mathcal{M}\setminus\{1\}} \{0, \pm 1, \ldots\}$$

in which $\mathcal{A}_m$ is determined by the array configuration, and the operator $\Pi(\cdot)$ denotes the Cartesian product of input sets. For each candidate ambiguity $\boldsymbol{\rho}$, a procedure of joint angle estimation and phase calibration is conducted.

C. Phase Bias Model

To model the biased PoA data, a mass of phase data from various arrays is collected, encompassing different azimuths and elevations. Based on extensive experimental data, we develop a phase bias model, which can be described by

$$\mathsf{b}_m = b_{\text{delay},m} + b_{\text{pcd},m}(\phi, \psi) + \mathsf{b}_{\text{others},m}(\text{CIR}_k). \tag{9}$$

The term b_m characterizes the phase bias between antennas m and 1. The first term $b_{\text{delay},m}$ denotes a constant phase bias that is induced by transmission delays. The second term $b_{\text{pcd},m}(\phi, \psi)$ characterizes the phase biases induced by phase center deviations. Given the presence of several unclear factors whose effects can be inferred from CIR data—such as antenna coupling and wireless propagation—we encapsulate them into $\mathsf{b}_{\text{others}}(\text{CIR}_k)$, where k denotes the kth time instant.

Given an array, we collect a training dataset $\{(\phi_i, \psi_i, \Delta\boldsymbol{\varphi}_i, \text{CIR}_i),\ i = 1, 2, \ldots, D\}$ to estimate phase error b_m in (9). To reduce computational complexity, we use the CIR samples at the peak points (one sample per antenna), rather than the entire CIR samples of the received UWB signal. First, the transmission delays that incur the phase bias $b_{\text{delay},m}$ are estimated by calculating the average difference between PDoA measurements and the expected PDoAs. After estimating and calibrating these delays, the PDoA data, presented in Fig. 6(b), is significantly improved. This result indicates that $b_{\text{delay},m}$ is a principal factor influencing the phase bias.

Second, upon analyzing the experimental results in Fig. 6(c), we note that the phase center deviation is affected by the AoAs. Thus, the phase bias induced by the phase center deviations can be expressed as

$$b_{\text{pcd},m}(\phi, \psi) = \frac{2\pi}{\lambda}(\delta\boldsymbol{q}_m - \delta\boldsymbol{q}_1)^{\mathrm{T}}\boldsymbol{u}(\phi, \psi) \tag{10}$$

where $(\delta\boldsymbol{q}_m - \delta\boldsymbol{q}_1)$ represents the unknown phase center deviation between antennas m and 1. Based on the results in Fig. 6(c), we approximate this effect using spherical harmonics

$$(\delta\boldsymbol{q}_m - \delta\boldsymbol{q}_1)(\phi, \psi) = \sum_{l=1}^{L}\sum_{j=-l}^{l} \boldsymbol{a}_{m,l,j} Y_l^j(\phi, \psi)$$

where $Y_l^j(\phi, \psi)$ are the spherical harmonics, and $\boldsymbol{a}_{m,l,j}$ are coefficients to be determined. The coefficients of spherical harmonics are estimated using an LS method. Fig. 6(c) demonstrates that the accuracy of the phase bias calibration is further improved after this step. Finally, the third component, $\mathsf{b}_{\text{others}}(\text{CIR}_k)$, in (9) accounts for unclear factors, such as mutual coupling and antenna scattering effects. We approximate it using a Gaussian process, adhering to the variational Gaussian process regression (GPR) methodology, as outlined in [61].

After determining the parameters of the phase bias model in (9) during the training stage, we then employ it for high-accuracy AoA estimation. Next, we present a method for online angle estimation along with phase calibration.

D. Phase Calibration and Angle Estimation

As indicated in (10), the biases induced by phase center deviations, $b_{\text{pcd},m}(\phi, \psi)$, exhibit strong directional dependence, necessitating joint estimation of phase biases and AoAs to achieve high accuracy in practical deployments. For a candidate ambiguity $\boldsymbol{\rho}$, we jointly estimate the phase biases and the AoAs by solving the following LS problem:

$$\underset{\phi,\psi}{\text{minimize}} \sum_{m\in\mathcal{M}\setminus\{1\}} \left\| \frac{2\pi}{\lambda}(\boldsymbol{q}_m - \boldsymbol{q}_1)^{\mathrm{T}}\boldsymbol{u}(\phi, \psi) + \mathsf{b}_m(\phi, \psi) - \Delta\tilde{\varphi}_m(\boldsymbol{\rho}) \right\|^2. \tag{11}$$

Objective function (11) is nonconvex, making a precise initial point crucial for ensuring the accuracy of the optimization. Recall (9), since the first and third terms are irrelevant to the AoAs,

we can derive a solution $(\hat{\phi}_{\text{LS}}, \hat{\psi}_{\text{LS}})$ that only calibrates these two terms. As shown in Fig. 6(a) and (b), the first term $b_{\text{delay},m}$ is the predominant factor in phase bias. Therefore, $(\hat{\phi}_{\text{LS}}, \hat{\psi}_{\text{LS}})$ serves as an effective initialization.

With $(\hat{\phi}_{\text{LS}}, \hat{\psi}_{\text{LS}})$, the Newton method is applied to optimize (11). The resulting solution is our angle estimation, which corresponds to the candidate ambiguity $\boldsymbol{\rho}$, denoted as $(\hat{\phi}_{\text{cand}}(\boldsymbol{\rho}), \hat{\psi}_{\text{cand}}(\boldsymbol{\rho}))$. Its ambiguity cost is defined by

$$c(\boldsymbol{\rho}) = \sum_{m \in \mathcal{M}\setminus\{1\}} \|\Delta\varphi_m - \hat{\varphi}_m(\boldsymbol{\rho})\|^2 \tag{12}$$

where

$$\hat{\varphi}_m(\boldsymbol{\rho}) = \text{Wrap}\left((\boldsymbol{q}_m - \boldsymbol{q}_1)^{\text{T}} \frac{2\pi}{\lambda} \boldsymbol{u}(\hat{\phi}_{\text{cand}}(\boldsymbol{\rho}), \hat{\psi}_{\text{cand}}(\boldsymbol{\rho})) + \mathsf{b}_m\left(\hat{\phi}_{\text{cand}}(\boldsymbol{\rho}), \hat{\psi}_{\text{cand}}(\boldsymbol{\rho})\right)\right). \tag{13}$$

Finally, considering the ambiguities, we complete the estimation process and provide an AoA estimation obeying

$$\hat{\phi}, \hat{\psi} = (\hat{\phi}_{\text{cand}}, \hat{\psi}_{\text{cand}})(\arg\min_{\boldsymbol{\rho}\in\mathcal{A}} c(\boldsymbol{\rho})).$$

VI. Network Ranging

In this section, we estimate the distances between robots using UWB timestamp measurements.

A. UWB Measurements

Inter-robot distances are computed based on (1) and (2) by estimating the time of flight and time difference of flight from timestamp measurements. In ideal conditions, ToF is the difference between ToA at the receiver and ToD at the transmitter. However, due to the asynchronous nature of the devices and timestamp measurement errors, this calculation method can result in ranging errors of hundreds of meters. According to the IEEE 802.15.4a standard [62] and related works [33], [34], [35], [36], [37], [38], [60], the ToA/ToD timestamp observed by node n at time instant k can be modeled as

$$\mathsf{t}_k^{(n)} = (1 + \mathsf{e}_k^{(n)}) t_k^{(n)} + \theta^{(n)} + \mathsf{w}_k^{(n)} \tag{14}$$

where $t_k^{(n)}$ represents the actual time; and $\mathsf{e}_k^{(n)}$, $\theta^{(n)}$, and $\mathsf{w}_k^{(n)}$ denote the clock frequency deviation, the time offset associated with the clock boot time, and the timestamp measurement error, respectively. Following the Decawave DW1000 user manual [63], time offset $\theta^{(n)}$ can be viewed as time-invariant, while clock drift $\mathsf{e}_k^{(n)}$ is a random variable with a range of up to ± 20 parts per million (ppm). At the millisecond-level, this clock drift parameter can be viewed as time-invariant [36], [37], [38]. Timestamp measurement errors $\mathsf{w}_k^{(n)}$ arise from antenna delays [39], NLOS conditions [40], and multipath effects, for which there is currently no well-recognized model.

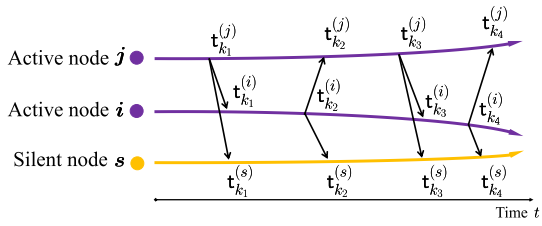


Fig. 8. Timestamp measurements of SM-NR.

B. Network Ranging and Clock Synchronization

The goal of ranging is to obtain high-precision distance estimates from timestamp measurements, while clock errors, timestamp measurement errors, and node velocities are unknown. To mitigate clock errors, RTTs are measured, as shown in Fig. 8, the relationships between the RTTs and ToFs are described in (15),

$$\begin{aligned}
&t_{k_4}^{(j)} - t_{k_1}^{(j)} - (t_{k_4}^{(i)} - t_{k_1}^{(i)}) = T_{\text{oF}_{k_1}}(i,j) + T_{\text{oF}_{k_4}}(i,j)\\
&t_{k_3}^{(i)} - t_{k_2}^{(i)} - (t_{k_3}^{(j)} - t_{k_2}^{(j)}) = T_{\text{oF}_{k_2}}(i,j) + T_{\text{oF}_{k_3}}(i,j)\\
&t_{k_4}^{(s)} - t_{k_3}^{(s)} + T_{\text{oF}_{k_3}}(j,s) = T_{\text{oF}_{k_3}}(i,j) + T_{\text{oF}_{k_4}}(i,s) + t_{k_4}^{(i)} - t_{k_3}^{(i)}\\
&t_{k_4}^{(s)} - t_{k_1}^{(s)} + T_{\text{oF}_{k_1}}(j,s) = \sum_{l=1}^{3} T_{\text{oF}_{k_l}}(i,j) + T_{\text{oF}_{k_4}}(i,s)\\
&\qquad + t_{k_2}^{(i)} - t_{k_1}^{(i)} + t_{k_3}^{(j)} - t_{k_2}^{(j)} + t_{k_4}^{(i)} - t_{k_3}^{(i)}
\end{aligned} \tag{15}$$

where k_l $(l = 1, \ldots, 4)$ represents time instants, and $T_{\text{oF}_k}(i,j)$ denotes the ToF between nodes i and j at time instant k. The actual timestamp $t_k^{(n)}$ on the left-hand side of (15) is unknown in practice, and only the measurement $\mathsf{t}_k^{(n)}$ in (14) is available.

Our network ranging scheme consists of four steps: First, we calibrate timestamp measurement errors. Second, in the absence of a time server, we establish a *synchronization time*, which maintains a global average time as the global time. The node clocks are then synchronized with this reference time to eliminate clock errors. After the first two steps, clock and timestamp measurement errors are mitigated. Third, we derive idealized ToF and TDoF estimations from (15) under perfect clock synchronization and timestamp measurement conditions. Finally, we substitute the calibrated timestamps into the idealized range estimations derived in the third step.

The first step, which addresses timestamp measurement errors, will be presented in Section VI-C. The last three steps pertain to joint ranging and clock synchronization. We propose two methods: one for static systems [58]; and the other designed for dynamic systems that further removes velocity dependency [59]. The static method achieves network-wide clock synchronization with higher update frequencies but shows sensitivity to sudden velocity changes. In contrast, the dynamic method provides only pairwise clock synchronization at lower update rates, but offers superior resilience to rapid node movements. For technical details, see [58] and [59].

C. Calibration of Timestamp Measurement Errors

It is shown that our ToF estimation, given by [58, eq. (18)] or [59, eq. (21)], yields

$$\widehat{\mathsf{T}}^{\star}_{\mathrm{oF}_k}(i,j) \approx T_{\mathrm{oF}_k}(i,j) + \frac{1}{2}\left[\left(\mathsf{w}^{(j)}_{k_4} - \mathsf{w}^{(j)}_{k_3}\right) - \left(\mathsf{w}^{(i)}_{k_4} - \mathsf{w}^{(i)}_{k_3}\right)\right] \tag{16}$$

under the assumption that $t^{(n)}_{N_\mathrm{a}+1} - t^{(n)}_1 \gg \mathsf{w}^{(n)}_k$[4] and $\mathsf{w}^{(n)}_k \approx \mathsf{w}^{(n)}_{k-2}$.[5] This indicates that the clock-induced error is significantly mitigated by our ranging method, while the timestamp measurement error needs further careful calibration. The same result is consistent for TDoF. Thus, next, we address timestamp measurement error $\mathsf{w}^{(n)}_k$.

There is no well-recognized model on the timestamp measurement error $\mathsf{w}^{(n)}_k$ in (14). Here, we model it based on experimental results and prior research [39], [40], which can be described by

$$\mathsf{w}^{(n)}_k = \begin{cases} \zeta\left(\mathrm{CIR}^{(n)}_k\right) + o^{(n)}, & \text{ToA timestamp} \\ 0, & \text{ToD timestamp} \end{cases} \tag{17}$$

where $\zeta(\mathrm{CIR})$ and $o^{(n)}$ denote channel-induced and hardware-inherent errors, respectively. The first term, $\zeta(\mathrm{CIR})$, exists because the ToA timestamps are estimated using a leading-edge detection algorithm on CIR, which makes the channel-induced error a function dependent on CIR. The second term $o^{(n)}$ arises from hardware-related factors, including delays introduced by antennas and circuits, which are specific to individual nodes. Hardware-inherent errors are present in both ToA and ToD timestamps. Our model incorporates hardware-inherent errors into ToA timestamps while setting $\mathsf{w}^{(n)}_k = 0$ for ToD timestamps, which is equivalent to modeling them separately.

As the timestamp calibration step is conducted at the node end at each time instant, we omit the superscript $(\cdot)^{(n)}$ and the subscript $(\cdot)_k$, which indicate node n and time instant k, for brevity. Based on the sources of errors, we model the channel-induced error as a random variable and the hardware-inherent error as an unknown deterministic parameter. We describe the channel-induced error in (17) by

$$\zeta(\boldsymbol{P}) = \mathsf{b}(\boldsymbol{P}) + \mathsf{v} \tag{18}$$

where v denotes the observation noise, and $\mathsf{b}(\boldsymbol{P})$ represents a systematic bias that depends on the channel state, characterized by the vector $\boldsymbol{P} = [P_{\mathrm{R},1},\ P_{\mathrm{F},1}, \ldots,\ P_{\mathrm{R,M}},\ P_{\mathrm{F,M}},\ \boldsymbol{u}(\hat{\phi},\hat{\psi})^{\mathrm{T}},\ \mathsf{d}]^{\mathrm{T}}$. We adopt two channel parameters extracted from the CIR samples—the received power level P_R and the first path signal-to-noise ratio P_F instead of the CIR data because directly processing the high-dimensional CIR will lead to significant computational complexity. The notations $P_{\mathrm{R,m}}$ and $P_{\mathrm{F,m}}$ refer to the parameters measured by the mth antenna, while

[4]Note that $t^{(i)}_{N_\mathrm{a}+1} - t^{(i)}_1$ is at second-level, while $\mathsf{w}^{(n)}_k$ is at nanosecond-level.

[5]Our experimental results show that the root mean square value of $\widehat{\mathsf{T}}^{\star}_{\mathrm{oF}_k}(i,j) - \widehat{\mathsf{T}}^{\star}_{\mathrm{oF}_{k-2}}(i,j)$ is approximately 5 cm / c, while the root mean square error of $\widehat{\mathsf{T}}^{\star}_{\mathrm{oF}_k}(i,j)$ is more than 30 cm / c, indicating a strong temporal correlation of $\mathsf{w}^{(n)}_k$. Thus, it is feasible to assume that $\mathsf{w}^{(n)}_k \approx \mathsf{w}^{(n)}_{k-2}$.

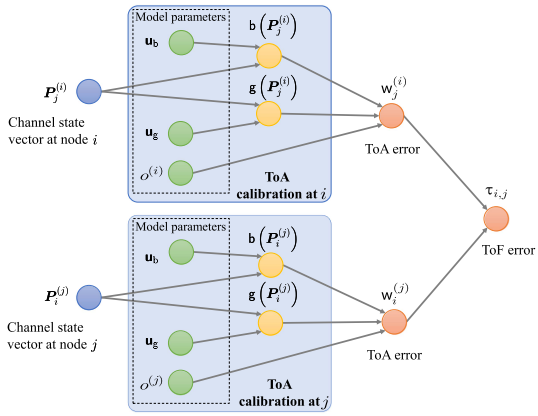


Fig. 9. Probabilistic graphical model for the calibration of timestamp measurement errors.

d is the range estimation using raw timestamp measurements. Essentially, $\boldsymbol{u}(\hat{\phi},\hat{\psi})$ in (7) and d are obtained from CIR and can provide additional information about the channel state apart from $P_\mathrm{R}, P_\mathrm{F}$. As for v, we model it as $p(\mathsf{v}) = \mathcal{N}(\mathsf{v}|0, \sigma^2_\mathrm{b}(\boldsymbol{P}))$, where $\sigma^2_\mathrm{b}(\boldsymbol{P})$ corresponds to the channel-related noise variance. This is based on observations concluded during experiments, which demonstrate that the variance of the observation noise is correlated with channel parameters and nodes, but these factors are independent of each other. For brevity, define $\mathsf{g}(\boldsymbol{P})$ such that $\sigma^2_\mathrm{b}(\boldsymbol{P}) = e^{\mathsf{g}(\boldsymbol{P})}$.

Two independent scalable variational Gaussian process (SVGP) models [64] are employed to estimate the probability distribution of $\mathsf{b}(\boldsymbol{P})$ and $\sigma^2_\mathrm{b}(\boldsymbol{P})$, respectively. First, assume the Gaussian process priors over $\mathsf{b}(\boldsymbol{P})$ and $\sigma^2_\mathrm{b}(\boldsymbol{P})$ as

$$p(\mathbf{b}) = \mathcal{N}(\mathbf{b}|\,\mathbf{0}, \boldsymbol{K}_\mathrm{b}) \quad \text{and} \quad p(\mathbf{g}) = \mathcal{N}(\mathbf{g}|\,\mathbf{0}, \boldsymbol{K}_\mathrm{g}) \tag{19}$$

where $\mathbf{b} = [\mathsf{b}_1\ \mathsf{b}_2\ \cdots\ \mathsf{b}_D]^{\mathrm{T}}$ is a vector containing D training data and $\boldsymbol{K}_\mathrm{b}$ is the covariance matrix derived from the kernel function with parameter $\boldsymbol{\theta}_\mathrm{b} \in \mathbb{R}^{2M+4}$, taking the form of

$$k\left(\boldsymbol{P}_{d_1}, \boldsymbol{P}_{d_2}; \boldsymbol{\theta}_\mathrm{b}\right) = \exp\left(-\sum_{l=1}^{2M+4} \frac{\left([\boldsymbol{P}_{d_1}]_l - [\boldsymbol{P}_{d_2}]_l\right)^2}{2[\boldsymbol{\theta}]^2_{\mathrm{b},l}}\right). \tag{20}$$

The vector $\mathbf{g}$ and the covariance matrix $\boldsymbol{K}_\mathrm{g}$ are defined in the same way. To reduce the computational complexity, we adopt the variational sparse heteroscedastic Gaussian process framework, where inducing points $\boldsymbol{Z}_\mathrm{b} = [\boldsymbol{Z}_{\mathrm{b},1}\, \boldsymbol{Z}_{\mathrm{b},2} \cdots \boldsymbol{Z}_{\mathrm{b},M}]^{\mathrm{T}}$ and $\boldsymbol{Z}_\mathrm{g} = [\boldsymbol{Z}_{\mathrm{g},1}\, \boldsymbol{Z}_{\mathrm{g},2} \cdots \boldsymbol{Z}_{\mathrm{g},M}]^{\mathrm{T}}$ and inducing variables $\mathbf{u}_\mathrm{b}$ and $\mathbf{u}_\mathrm{g}$, with variational distributions $q_\mathrm{b}(\mathbf{u}_\mathrm{b}) = \mathcal{N}(\mathbf{u}_\mathrm{b}|\boldsymbol{m}_\mathrm{b}, \boldsymbol{S}_\mathrm{b})$ and $q_\mathrm{g}(\mathbf{u}_\mathrm{g}) = \mathcal{N}(\mathbf{u}_\mathrm{g}|\boldsymbol{m}_\mathrm{g}, \boldsymbol{S}_\mathrm{g})$, are introduced for $\mathbf{b}$ and $\mathbf{g}$, respectively. The relationship between these parameters is shown in Fig. 9.

In the training stage, model parameters $\boldsymbol{\theta}_\mathrm{b}$, $\boldsymbol{\theta}_\mathrm{g}$, $\boldsymbol{m}_\mathrm{b}$, $\boldsymbol{S}_\mathrm{b}$, $\boldsymbol{m}_\mathrm{g}$, and $\boldsymbol{S}_\mathrm{g}$ are what we need to determine. Typically, these key parameters are trained by maximizing the variational lower bound of the log-likelihood of ToA estimation error $\mathsf{w}^{(n)}_k$, as discussed in [65]. However, it is difficult to obtain accurate ToA timestamps for training. To address this issue, we turn to

training the SVGP models for ToF errors, which are defined by $\tau_{i,j} = \hat{\mathsf{T}}_{\mathrm{oF}}(i,j) - T_{\mathrm{oF}}(i,j)$. The accurate values of these errors can be deduced from reliable position tracking systems, such as Optitrack. The variational lower bound for the marginal likelihood of the ToF measurement error $\mathcal{L}_\tau$ is derived, which can be described as

$$\begin{aligned}\mathcal{L}_\tau = &\sum_{i=1}^{N}\sum_{j=i+1}^{N} \log \mathcal{N}\left(\tau_{i,j}|\frac{\mu_j^{(i)}+\mu_i^{(j)}}{2}, \frac{r_{\mathrm{g},j}^{(i)2}+r_{\mathrm{g},i}^{(j)2}}{4}\right)\\ &-\sum_{i=1}^{N}\sum_{j=i+1}^{N}\left[\frac{1}{4}\left(\sigma_{\mathrm{g},j}^{(i)2}+\sigma_{\mathrm{g},i}^{(j)2}\right)+\frac{1}{2}\left(\frac{\sigma_{\mathrm{b},j}^{(i)2}}{r_{\mathrm{g},j}^{(i)2}}+\frac{\sigma_{\mathrm{b},i}^{(j)2}}{r_{\mathrm{g},i}^{(j)2}}\right)\right]\\ &-\mathrm{KL}\left[q_{\mathrm{b}}(\mathbf{u}_{\mathrm{b}})||p(\mathbf{u}_{\mathrm{b}})\right]-\mathrm{KL}\left[q_{\mathrm{g}}(\mathbf{u}_{\mathrm{g}})||p(\mathbf{u}_{\mathrm{g}})\right] \end{aligned} \tag{21}$$

where $\mu_j^{(i)} = \mu_{\mathrm{b},j}^{(i)} + o^{(i)}$, $r_{\mathrm{g},j}^{(i)2} = e^{\mu_{\mathrm{g},j}^{(i)} - \sigma_{\mathrm{g},j}^{(i)2}/2}$, in which $\mu_{\mathrm{b},j}^{(i)}$, $\mu_{\mathrm{g},j}^{(i)}$, $\sigma_{\mathrm{b},j}^{(i)2}$, $\sigma_{\mathrm{g},j}^{(i)2}$ are the predictive means and variances of $\mathbf{b}$ and $\mathbf{g}$, respectively, and $\mathrm{KL}[\cdot||\cdot]$ denotes Kullback-Leibler divergence.

After the model training is completed, we calibrate ToA timestamp measurement error as follows. For ToA timestamp received at node n with signal power level $\boldsymbol{P}$, timestamp measurement error is predicted by

$$p(\mathsf{w}^{(n)}|\mathsf{b}(\boldsymbol{P}),\mathsf{g}(\boldsymbol{P})) = \mathcal{N}\left(\mathsf{w}^{(n)}|o^{(n)}+\mathsf{b}(\boldsymbol{P}), e^{\mathsf{g}(\boldsymbol{P})}\right)$$

where

$$\begin{aligned} p\left(\mathsf{b}(\boldsymbol{P})\right) &= \mathcal{N}\left(\mathsf{b}(\boldsymbol{P})|\mu_{\mathrm{b}}(\boldsymbol{P}),\sigma_{\mathrm{b}}^2(\boldsymbol{P})\right)\\ p\left(\mathsf{g}(\boldsymbol{P})\right) &= \mathcal{N}\left(\mathsf{g}(\boldsymbol{P})|\mu_{\mathrm{g}}(\boldsymbol{P}),\sigma_{\mathrm{g}}^2(\boldsymbol{P})\right) \end{aligned} \tag{22}$$

in which

$$\begin{aligned}\mu_{\mathrm{b}}(\boldsymbol{P}) &= \boldsymbol{k}_{\mathrm{b},M}\boldsymbol{K}_{\mathrm{b},M}^{-1}\boldsymbol{m}_{\mathrm{b}}\\ \mu_{\mathrm{g}}(\boldsymbol{P}) &= \boldsymbol{k}_{\mathrm{g},M}\boldsymbol{K}_{\mathrm{g},M}^{-1}\boldsymbol{m}_{\mathrm{g}}\\ \sigma_{\mathrm{b}}^2(\boldsymbol{P}) &= k(\boldsymbol{P},\boldsymbol{P};\boldsymbol{\theta}_{\mathrm{b}})\\ &\quad + \boldsymbol{k}_{\mathrm{b},M}\boldsymbol{K}_{\mathrm{b},M}^{-1}(\boldsymbol{S}_{\mathrm{b}}-\boldsymbol{K}_{\mathrm{b},M})(\boldsymbol{k}_{\mathrm{b},M}\boldsymbol{K}_{\mathrm{b},M}^{-1})^{\mathrm{T}}\\ \sigma_{\mathrm{g}}^2(\boldsymbol{P}) &= k(\boldsymbol{P},\boldsymbol{P};\boldsymbol{\theta}_{\mathrm{g}})\\ &\quad + \boldsymbol{k}_{\mathrm{g},M}\boldsymbol{K}_{\mathrm{g},M}^{-1}(\boldsymbol{S}_{\mathrm{g}}-\boldsymbol{K}_{\mathrm{g},M})(\boldsymbol{k}_{\mathrm{g},M}\boldsymbol{K}_{\mathrm{g},M}^{-1})^{\mathrm{T}}\end{aligned}$$

with kernel function defined in (20) and

$$\begin{aligned}[\boldsymbol{k}_{\mathrm{b},M}]_l &= k(\boldsymbol{P},[\boldsymbol{z}_{\mathrm{b}}]_l;\boldsymbol{\theta}_{\mathrm{b}})\\ [\boldsymbol{k}_{\mathrm{g},M}]_l &= k(\boldsymbol{P},[\boldsymbol{z}_{\mathrm{g}}]_l;\boldsymbol{\theta}_{\mathrm{g}})\\ [\boldsymbol{K}_{\mathrm{b},M}]_{l_1,l_2} &= k([\boldsymbol{z}_{\mathrm{b}}]_{l_1},[\boldsymbol{z}_{\mathrm{b}}]_{l_2};\boldsymbol{\theta}_{\mathrm{b}})\\ [\boldsymbol{K}_{\mathrm{g},M}]_{l_1,l_2} &= k([\boldsymbol{z}_{\mathrm{g}}]_{l_1},[\boldsymbol{z}_{\mathrm{g}}]_{l_2};\boldsymbol{\theta}_{\mathrm{g}}).\end{aligned}$$

The calibrated timestamps are then used for clock synchronization and ranging in Section VI-B.

VII. Relative Localization

In this section, we develop a multisensory information fusion scheme for distributed relative localization in dynamic scenarios. For the initialization of network relative position estimation, the MDS-based method and Chan algorithm are applied for active and silent nodes, respectively. The main focus in this article is to realize high-accuracy dynamic localization, for which we propose a scheme combining UWB and IMU measurements through particle filtering (PF) and DR that achieves inter- and intra-node cooperation.

A. Multisensory Information Fusion

Since the updating frequency of the IMU measurements is usually higher than the UWB signals, we perform PF when the UWB measurements are updated, and DR when only the IMU measurements are obtained, as shown in Fig. 3. The difference between active and silent nodes mainly stems from the measurement model, while the localization framework is rather similar. Therefore, in the following part, we would not distinguish between the mode of nodes.

For node n, denote its state vector at time instant k as

$$\boldsymbol{x}_k^{(n)} = \left[\boldsymbol{p}_k^{(n)\,\mathrm{T}}, \boldsymbol{v}_k^{(n)\,\mathrm{T}}, \boldsymbol{a}_k^{(n)\,\mathrm{T}}, \boldsymbol{o}_k^{(n)\,\mathrm{T}}, \boldsymbol{\omega}_k^{(n)\,\mathrm{T}}\right]^{\mathrm{T}} \tag{23}$$

in which the elements represent the position, velocity, acceleration, orientation, and angular velocity vector, respectively. The PF part concerns N_{p} particles with states $\{\mathbf{x}_k^{(n,l)}\}_{l=1}^{N_{\mathrm{p}}}$ and weights $\{\mathsf{w}_k^{(n,l)}\}_{l=1}^{N_{\mathrm{p}}}$, where l denotes the index of the particles. Define the time-variant state transition function and observation function as $\boldsymbol{f}_k(\cdot)$ and $\boldsymbol{h}_k(\cdot)$ that come from the physical models. The corresponding state transition and observation models, respectively, are

$$\mathbf{x}_k^{(n,l)} = \boldsymbol{f}_k\left(\mathbf{x}_{k-1}^{(n,l)}\right) + \mathbf{n}_{\mathrm{s},k}^{(n,l)} \tag{24}$$

$$\mathbf{z}_k^{(n)} = \boldsymbol{h}_k\left(\mathbf{x}_k^{(n,l)}, \mathbf{s}_{\mathcal{N},k-1}^{(n)}\right) + \mathbf{n}_{\mathrm{o},k}^{(n,l)} \tag{25}$$

where $\mathbf{n}_{\mathrm{s},k}^{(n,l)} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{Q}_k^{(n)})$, $\mathbf{n}_{\mathrm{o},k}^{(n,l)} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{R}_k^{(n)})$ are the state transition and observation noises. The details of PF are regular procedures [66], which are omitted due to the limited space, and the PF estimation at time k is

$$\widehat{\mathbf{x}}_{k,\mathrm{PF}}^{(n)} = \sum_{l=1}^{N_{\mathrm{p}}} \mathsf{w}_k^{(n,l)} \mathbf{x}_k^{(n,l)}. \tag{26}$$

As for dead reckoning, the IMU measurements are utilized as control vectors for the state update of the particles as well as the estimation result

$$\mathbf{x}_k^{(n,l)} = \boldsymbol{f}_k\left(\mathbf{x}_{k-1}^{(n,l)}, \mathbf{z}_{\mathrm{IMU},k-1}^{(n)}\right) + \mathbf{n}_{\mathrm{s},k}^{(n,l)} \tag{27}$$

$$\widehat{\mathbf{x}}_{k,\mathrm{DR}}^{(n)} = \boldsymbol{f}_k\left(\widehat{\mathbf{x}}_{k-1}^{(n)}, \mathbf{z}_{\mathrm{IMU},k-1}^{(n)}\right) \tag{28}$$

while the particle weights remain unchanged.

Remark 1: In our distributed relative localization scheme, each node estimates the relative positions within its neighborhood. In networks that are not fully connected, global relative positions are determined through geometry merging operations, as presented in our previous work [67], [68].

Extrinsic parameters including rotation and translation between IMU and UWB sensors are calculated in a stand-alone

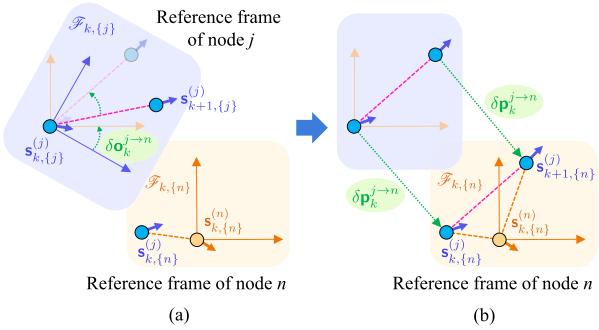


Fig. 10. Illustration of the process of RFA. (a) Rotation for the alignment of the orientations of the reference frames. (b) Translation for the alignment of the origins of the reference frames.

phase and set constant when PF is in operation. To be specific, an IMU-UWB sensor suite is intentionally fully excited in the calibration phase to conduct 6-DoF motion in space, and meanwhile onboard IMU and UWB data are collected. UWB measurements consist of range and bearing observations to four stationary stations with known positions. An optimization-based estimator is utilized to simultaneously reconstruct the above 6-DoF trajectory along with the IMU-UWB extrinsics. Preintegrated IMU factors with IMU biases and UWB measurement factors with our targeted extrinsics are building blocks of the factor graph structure of the estimator. Since no marginalization is conducted, the dimension of the estimated state is extremely large and the calibration process is done purely offline with hand measured extrinsics as initial values. We repeatedly implement such calibration of IMU-UWB alignment for each node.

B. Reference Frame Alignment and Update

Note that in relative localization, no absolute reference for any dimensions of the particle state $\mathbf{x}$ exists. Since each node performs the localization procedure in its own coordinate reference frame, their coordinate datum may differ from each other, and the cumulative drift over time may exist. This indicates the necessity of real-time alignment for the coordinate reference frames of different nodes in distributed relative localization.

Denote the local reference frame of node n at time k as $\mathscr{F}_{k,\{n\}}$, which is assumed to be an inertial coordinate system, where the subscript $\{\cdot\}$ denotes the index of the reference frame. Then, the transition from $\mathscr{F}_{k,\{j\}}$ to $\mathscr{F}_{k,\{n\}}$ $(j \in \mathcal{N}^{(n)})$ involves translation and rotation for the alignment of origins and axes, and the alignment of their translational velocities.

We first consider the dynamic part by introducing the indirect measurement of radial relative velocity in PF for the velocity alignment of reference frames. Note that during the update of particle weights, the ToF, TDoF, and AoA measurements can directly reflect the estimation confidence of position and orientation, and the IMU measurements reflect that of the acceleration and angular velocity, while the velocity is left out. Therefore, we estimate the radial relative velocity of each node pair by estimate the rate of change of ToF, which can be achieved by LS estimation. Then, we add it to the observation model, and exploit its reflection of the belief of relative velocity estimation

Algorithm 1: Distributed Relative Localization Algorithm.

Input: The measurement vectors $\mathbf{z}_k^{(n)}, n \in \mathcal{N}, k = 0, 1, \ldots$
Output: Position estimations $\widehat{\mathbf{p}}_k^{(n)}, n \in \mathcal{N}, k = 0, 1, \ldots$
Initialize the relative position estimation $\widehat{\mathbf{p}}_0^{(n)}$ and the calibration parameter $\mathbf{\Theta}_0^{(n)}$ $(n \in \mathcal{N})$;
Initialize particle parameters $\{\mathbf{x}_0^{(n,l)}, \mathsf{w}_0^{(n,l)}\}_{l=1}^{N_\mathrm{P}}$ $(n \in \mathcal{N})$;
for $k = 1, 2, \ldots$, nodes $n \in \mathcal{N}$ in parallel **do**
  if UWB measurements $\mathbf{z}_{\mathrm{UWB},k}^{(n)}$ updated **then**
    Perform RFA with $\mathbf{\Theta}_{k-1}^{(n)}$ to obtain $\mathbf{s}_{\mathcal{N},k}^{(n)}$;
    Update the particle states according to (24);
    Perform PF to obtain $\widehat{\mathbf{p}}_{k,\mathrm{PF}}^{(n)}$ by (26);
    Perform RFU by (30) and update $\mathbf{\Theta}_k^{(n)}$ accordingly;
  else if IMU measurements $\mathbf{z}_{\mathrm{IMU},k'}^{(n)}$ updated **then**
    Update the particle states according to (27);
    Perform dead reckoning to obtain $\widehat{\mathbf{p}}_{k',\mathrm{DR}}^{(n)}$ by (28);
  end if
end for
return Position estimations $\widehat{\mathbf{p}}_k^{(n)}, n \in \mathcal{N}, k = 0, 1, \ldots$

during weight update. On the macro perspective, this helps to align the translational velocities of the local reference frames of the neighbors in a natural manner.

Then, we tackle the geometrical alignment of different reference frames, as shown in Fig. 10. Let $\boldsymbol{s}_k^{(n)} = [\boldsymbol{p}_k^{(n)\mathrm{T}}, \boldsymbol{o}_k^{(n)\mathrm{T}}, \boldsymbol{v}_k^{(n)\mathrm{T}}]^\mathrm{T}$ be the pose vector. Denote the calibrated pose estimation of node j in $\mathscr{F}_{k,\{n\}}$ as $\widehat{\mathbf{s}}_{k,\{n\}}^{(j)}$. Let each node n maintain a calibration parameter table $\mathbf{\Theta}_k^{(n)} = [\cdots \boldsymbol{\vartheta}_k^{j\to n} \cdots]_{j\in\mathcal{N}^{(n)}}$, where the calibration parameters

$$\boldsymbol{\vartheta}_k^{j\to n} = \left[\delta\mathbf{p}_k^{j\to n\,\mathrm{T}}, \delta\mathbf{o}_k^{j\to n\,\mathrm{T}}\right]^\mathrm{T} \tag{29}$$

transform states from $\mathscr{F}_{k,\{j\}}$ to $\mathscr{F}_{k,\{n\}}$ with rotation $\delta\mathbf{o}_k^{j\to n}$ around $\mathbf{s}_{k,\{j\}}^{(j)}$ and translation $\delta\mathbf{p}_k^{j\to n}$. Note that the alignment of the states at time $(k+1)$ is performed in the reference frame determined at time k since the knowledge of the reference frame at time $(k+1)$ remains to be updated according to the new estimation results. With this operation of RFA, we obtain the neighbor information to perform the PF procedures precisely in the aligned frame.

The initial value of $\mathbf{\Theta}_0^{(n)}$ can be determined according to the initialization of network relative position estimation. Then, consider reference frame update (RFU) which updates the calibration parameters in real time. Denote the neighborhood pose vector of node n as $\mathbf{s}_{\mathcal{N},k}^{(n)} = [\cdots \mathbf{s}_k^{(j)\,\mathrm{T}} \cdots]_{j\in\mathcal{N}^{(n)}}^\mathrm{T}$ with $\mathbf{s}_k^{(n)}$ being the pose estimation, which is collected from broadcasting. The general idea is to reestimate the neighbor poses $\mathbf{s}_{\mathcal{N},k}^{(n)}$ while regarding the PF estimation result $\widehat{\mathbf{x}}_k^{(n,\mathrm{PF})}$ as known parameters, which is an optimization problem to minimize the mean squared error with solutions of the form

$$\widehat{\mathbf{s}}_{\mathcal{N},k}^{(n)} = \mathbf{s}_{\mathcal{N},k}^{(n)} + \boldsymbol{K}_k^{(n)}\left(\mathbf{z}_k^{(n)} - \boldsymbol{H}_k^{(n)}\,\mathbf{s}_{\mathcal{N},k}^{(n)}\right) \tag{30}$$

IEEE Transactions on Robotics (T-RO) paper, presented at ICRA 2026, Vienna, Austria. Cite as T-RO paper.

TABLE I
HARDWARE CONFIGURATION AND ARRAY PERFORMANCE

Properties	Values	Properties	Values
Chip	4× DW1000	Data rate	6.8 Mbps
Center frequency	3993.6 MHz	UWB Channel	2
Weight	104 g	Antenna spacing	4.67 cm
Size	10×10×2.5 cm	Power	2.5 W
Range	1 km	Update frequency	⩾100 Hz

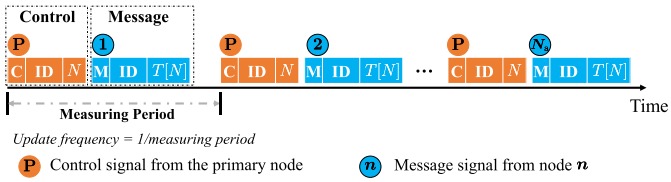


Fig. 11. Data frame of UWB signals.

in which the trivial details are omitted due to the limited space. Then, the calibration parameters $\vartheta_k^{j\to n}$ can be updated by the difference between $\widehat{\mathbf{s}}_{k,\{n\}}^{(j)}$ and $\mathbf{s}_{k,\{j\}}^{(j)}$ which achieves the real-time alignment of different reference frames. The complete proposed distributed relative localization scheme is summarized in Algorithm 1.

VIII. HARDWARE REALIZATION

This section presents the hardware implementation details of the proposed multi-robot localization method.

A. Hardware Configurations

The hardware realization is shown in Fig. 2. Each node is equipped with a stereo UWB array, which transmits and receives UWB signals, and a Raspberry Pi, which estimates the signal parameters and performs distributed positioning.

Each UWB array consists of an STM32 development board, which mounts four antennas, four Decawave DW1000 chips, and one Si5351B-B clock chip. The STM32 board retrieves raw measurements from registers and transmits data to the Raspberry Pi for postprocessing. Each antenna connects to a separate DW1000 UWB chip, with all UWB chips synchronized through a clock reference established by the Si5351B-B clock chip.[6] Despite careful design, asynchronism persists among the four UWB chips due to hardware imperfections, including differences in circuit lengths from the clock source to the UWB chips and differences in feed wires to individual antennas [32]. These imperfections introduce errors in PoA, ToA, and ToD measurements. We calibrate PoA measurement errors for bearing in Section V-C and address ToA/ToD measurement errors for ranging in Section VI-C.

Our stereo UWB array weighs 104 g with four antennas arranged in a regular tetrahedron configuration (4.67 cm spacing). It supports a measurement range of up to 1 km, enabling long-distance localization. Complete hardware specifications are provided in Table I. Overall, the designed UWB array is compact and light, making it suitable for diverse robots in their respective use cases.

[6]The Decawave DW3000 chip can directly provide AoA estimates by supporting two RF antenna ports. Compared with it, our bearing method achieves higher accuracy, since our array features four RF antenna ports by using four Decawave DW1000 chips and our method effectively addresses the critical challenges of the phase bias and phase ambiguity in bearing estimation.

B. Implementing SM-NR

Using the designed UWB arrays, we realize multi-robot localization as follows.

1) The Raspberry Pi releases the SM-NR measurement schedule to the corresponding UWB array through the serial port.
2) Each active node transmits UWB signals according to the SM-NR protocol and records the corresponding ToA, ToD, PoA, and signal strength measurements. These raw measurements are transmitted to the Raspberry Pi via a serial port for further processing.
3) The Raspberry Pi computes distances and directions, then shares these estimations with other nodes using ROS communication.
4) By aggregating information from other nodes, the Raspberry Pi at each node determines and outputs the localization results distributively.

To implement the above procedures on the hardware, we propose a new UWB transmission flow and new data frames, as illustrated in Fig. 11. Two types of UWB signals are defined: 1) control; and 2) message. A primary controller node is arbitrarily designated to broadcast the control UWB signal at each time instant, dictating the transmission order of active nodes. Upon receiving this control signal, the designated active node transmits a Message UWB signal for measurement. The receiver nodes process this signal to extract ToA, ToD, PoA, and signal strength data for ranging and bearing. Using these new UWB measurements, nodes update their position estimates. After a predefined interval, the primary controller sends the next control signal, and the measurement cycle repeats. In our scheme, the interval between two control UWB signals is 10 ms and the update frequencies for measurement and localization are 100 Hz. The shortest interval we have tested is 6 ms.

Our designed signal frame for SM-NR implementation contains four elements: 1) preamble; 2) start of frame delimiter; 3) physical layer header; and 4) data [63]. The preamble sequence is used to detect signals and estimate signal metrics, while the data frame is utilized for information exchange. Depicted in Fig. 11, both the control and message signal data frames comprise three segments. The first segment indicates the signal type (CMD or MSG); the second segment of the control signal specifies the ID of the next message transmitter, while the second segment of the message signal conveys the transmitter's ID; the third segment of the control signal specifies the number of active nodes within the network, while the message signal includes a timestamp array that records all sending and receiving

IEEE Transactions on Robotics (T-RO) paper, presented at ICRA 2026, Vienna, Austria. Cite as T-RO paper.

timestamps between the current transmission and the previous one.

IX. Experimental Analysis

In this section, we evaluate the proposed multi-robot localization method through real-world experiments. Section IX-A demonstrates its ability to achieve centimeter-level localization accuracy, and Section IX-B confirms its robustness and scalability.

For the experiments, aerial robots and vehicles are equipped with our UWB arrays. The dimensions of the test sites are 20× 25×4 m, and the precise positions of the nodes are obtained using the FZMotion motion capture system [69], which serve as the ground truth for performance evaluation.

A. Accuracy

The following are three typical application scenarios for multi-robot systems that are tested, as shown in Fig. 12.

1) *Road Overtaking:* Three vehicles travel normally on a two-lane road, while the fourth vehicle performs an overtaking maneuver, moving from the rear to the front. These four nodes are all active nodes.
2) *Formation:* A drone leads at the front, while four vehicles form a trapezoidal formation to navigate a U-shaped route. All five nodes are active nodes.
3) *Search and Rescue:* Positioned at strategic locations, three vehicles, two drones, and ten static UWB devices are actively broadcasting signals. The same set of robots—including three vehicles, two drones, and ten static UWB devices—are engaged in search operations in silent mode to reduce signal overhead. The complexity of this mission escalates due to high latency in the large-scale network and challenging propagation conditions, including NLOS and multipath effects caused by dense obstacles.

We evaluate the performance of UWB ranging, bearing, and localization using root mean square errors (RMSEs) as the performance metric, and the results are shown in Figs. 13–15, respectively.

1) Ranging Accuracy: The ranging performance is shown in Fig. 13. First, the RMSEs of raw active ranging across the three scenarios are 32.06, 26.90, and 34.19 cm,[7] respectively. After calibration, the RMSEs improve to 4.26, 6.22, and 6.48 cm. An approximate 80% improvement highlights the necessity of our calibration efforts. Second, the RMSE of silent ranging in the third scenario is 28.64 cm, which is inferior to that of active ranging due to the absence of timestamp measurement error calibration—an issue to be addressed in future work. Third, Fig. 16 visualizes the full date distribution before and after calibration, demonstrating the overall improvement in ranging performance.

[7]Note that the 30 cm-level errors before calibration arise from NLOS-induced timestamp inaccuracies and extended inter-node distances. By contrast, in controlled tests with nodes at 1–6 m separation (LOS conditions), the RMSE is 10.77 cm before calibration, which is consistent with the DW1000's specified accuracy.

TABLE II
Performance Evaluation With RMSEs

Module	Scenario 1	Scenario 2	Scenario 3	
	Active	Active	Active	Silent
Ranging (raw)	32.06 cm	26.90 cm	34.19 cm	28.64 cm
Ranging (calibrated)	4.26 cm	6.22 cm	6.48 cm	–
Elevation angle	2.5 deg	2.5 deg	2.8 deg	2.9 deg
Azimuth angle	4.4 deg	4.8 deg	4.8 deg	3.7 deg
Localization	6.23 cm	16.84 cm	11.26 cm	24.76 cm

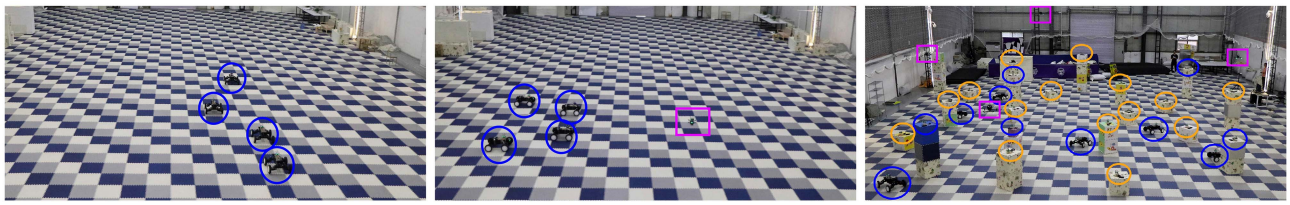
2) Bearing Accuracy: The AoA estimation performance under three dynamic scenarios is shown in Fig. 14. The baseline used for comparison calibrates only the phase biases caused by transmission delays, consistent with current UWB phase calibration practices. The RMSEs for elevation estimation across the three test scenarios are 2.5°, 2.5°, and 2.9°, respectively, while azimuth estimation RMSEs are 4.4°, 4.8°, and 4.3°, respectively. The proposed method significantly improves the elevation estimation accuracy, though performs similarly to the baseline on azimuth estimation. This improvement is attributed to our more comprehensive phase modeling and meticulous phase error calibration.

To further validate the effectiveness of the proposed AoA estimation method, we evaluate bearing measurement accuracy at three representative elevation angles (−0.5°, 14.4°, and 24.5°), selected randomly to test performance across different geometric configurations. For each elevation angle, PoA data across various azimuths are collected by rotating the receiver array around the z-axis. As shown in Fig. 17, our method achieves superior elevation estimation accuracy compared to the baseline, a critical enhancement for robust 3-D localization.

3) Localization Accuracy: The relative localization algorithm with the multisensory information fusion scheme proposed in Section VII is evaluated under the aforementioned three dynamic scenarios, as shown in Fig. 15. The groundtruth obtained by FZMotion is marked in green, while the estimated trajectories of active and silent nodes are marked in blue and orange, respectively. The relative RMSEs are listed in Table II.

The proposed algorithm performs effectively when the nodes move at moderate speeds. In Scenario 1 and most periods in Scenarios 2 and 3, the relative localization error is kept between 6 and 10 cm. When the drone accelerates violently, the measurements and the state updates suffer from the drastic changes of relative node poses, and the localization performance also deteriorates to a certain extent, which is a focus of our future work. The average localization performances in Scenarios 2 and 3 for the active nodes hold an accuracy of around 11–17 cm. For the silent nodes in Scenario 3, the silent nodes exhibit inferior performance compared to the active nodes, achieving an accuracy of about 25 cm. This is due to the lack of node cooperation and the larger biases caused by uncalibrated differential ranging. Calibrating differential ranges effectively will be addressed in our future work. These results demonstrate that our distributed localization system can attain centimeter-level accuracy.

IEEE Transactions on Robotics (T-RO) paper, presented at ICRA 2026, Vienna, Austria. Cite as T-RO paper.



(a) (b) (c)

Fig. 12. Three dynamic scenarios of real-world experiments: road overtaking, formation, and search and rescue. The blue/orange circles and magenta rectangles mark the vehicles as active/silent nodes, and the drones as active nodes, respectively. (a) Scenario 1: Road overtaking. (b) Scenario 2: Formation. (c) Scenario 3: Search and rescue.

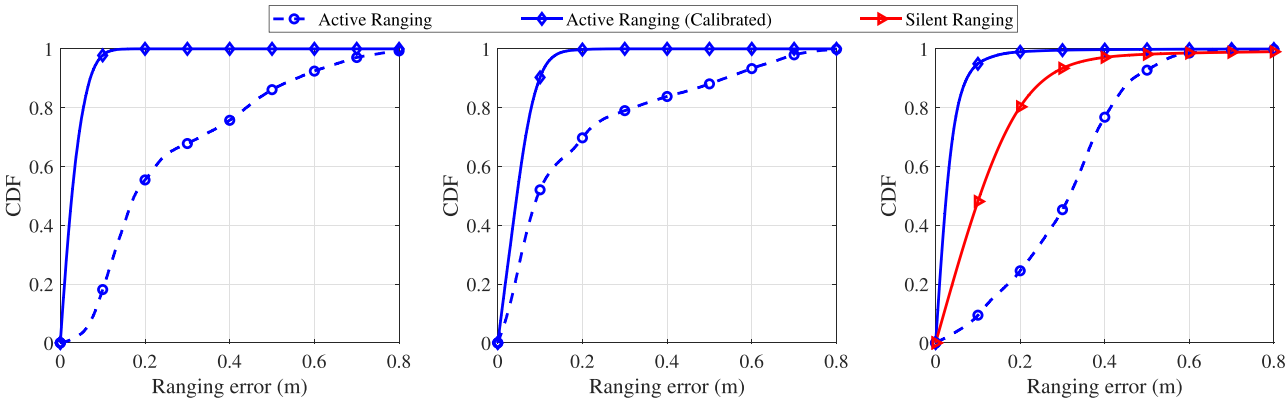


Fig. 13. Ranging performance under three dynamic scenarios.

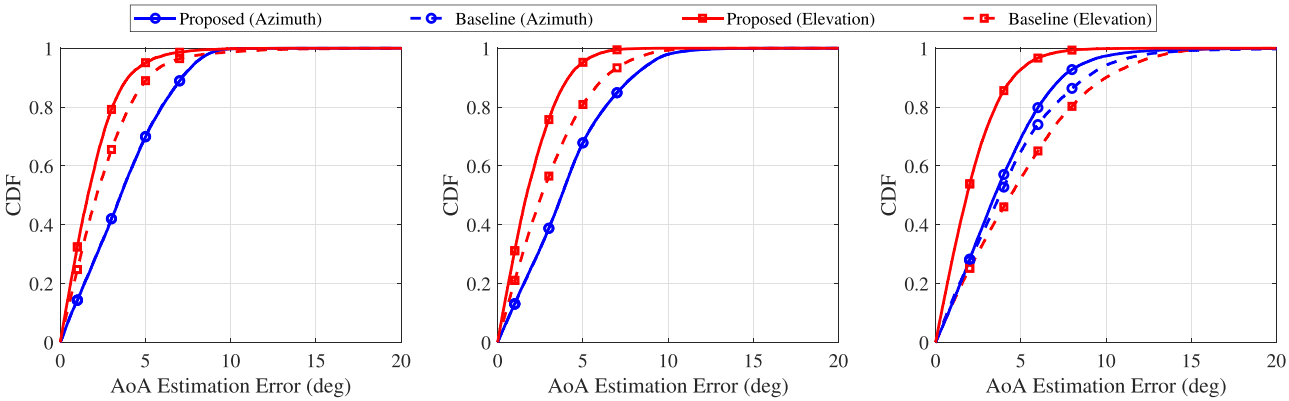


Fig. 14. AoA estimation performance under three dynamic scenarios.

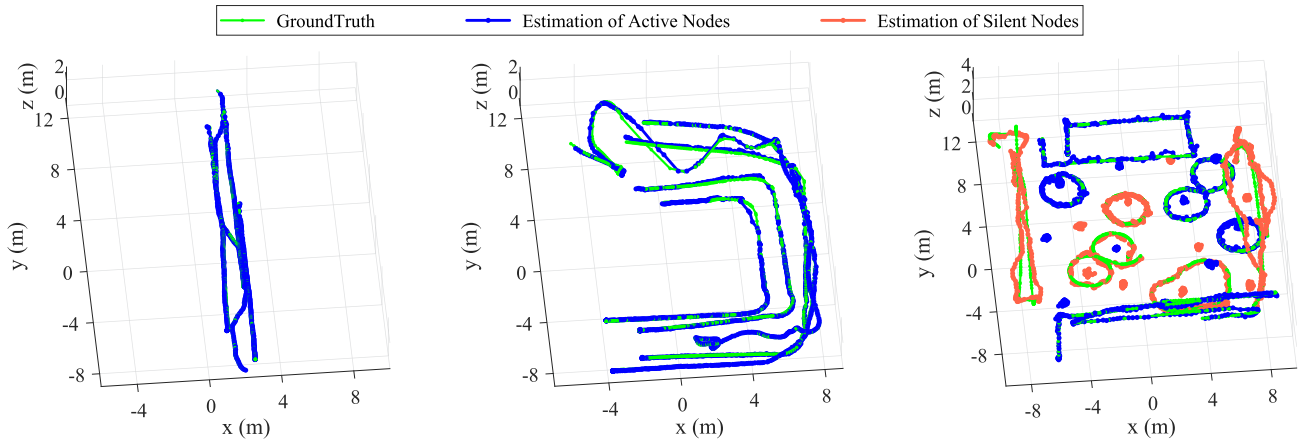


Fig. 15. Relative localization performance under three dynamic scenarios.

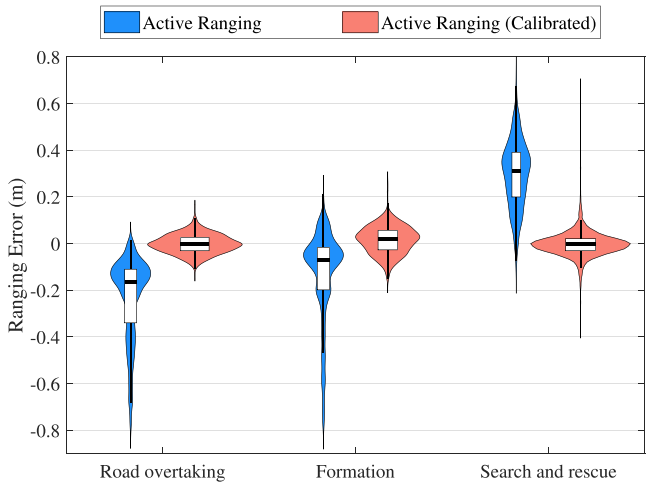


Fig. 16. Violin plots of the ranging performance under three dynamic scenarios.

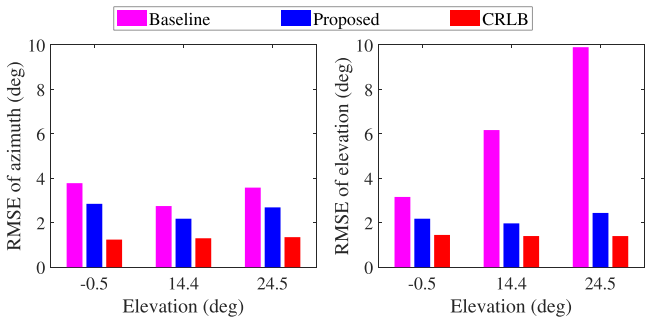


Fig. 17. Performance of azimuth/elevation estimation across various elevation angles.

Beyond improving accuracy, we also validate the efficiency and timeliness of the proposed network measurement method by comparing it with the widely used TWR in [33]. In this experiment, all nodes are active, serving as both transmitters and receivers. Fig. 18 shows the time-varying ranging errors of a single connection, and Fig. 19 shows the cumulative distribution function (CDF) of ranging results for the entire network. First, Figs. 18 and 19 show that our method exhibits greater robustness as the number of nodes increases, demonstrating its scalability for large swarms. Second, we note that UWB measurement errors consist of two components: 1) estimation errors; and 2) lag errors due to delayed measurements. Fig. 18 implies that lag errors are the dominant factor, especially in large-scale networks. Third, our ranging method preserves centimeter-level accuracy as the network scales, whereas the accuracy of TWR deteriorates to the meter level due to delayed measurements. This is because, in a network with N_{a} active nodes, our method updates $N_{\mathrm{a}}-1$ absolute range estimates per UWB signal, whereas TWR updates only $1/4$. Thus, the performance gain from using our method is proportional to N_{a}.

Furthermore, we obtain the following insights for localization system design. First, it is crucial to have a high rate of measurement updates, particularly for dynamic large-swarm systems. Second, while increasing the number of active nodes improves

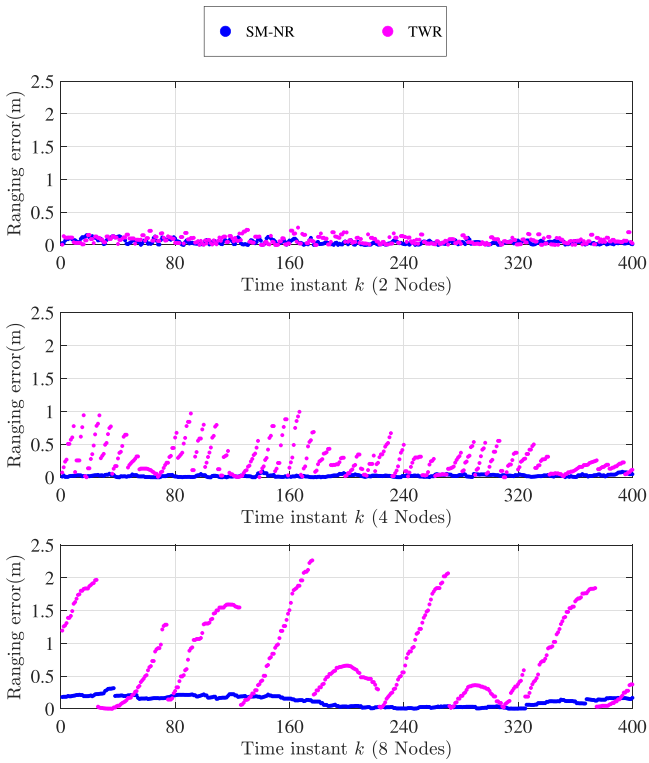


Fig. 18. Ranging errors of a single connection vary with time instant. The subfigures correspond to multi-robot systems with 2, 4, and 8 active nodes, respectively (w/o calibrating timestamp measurement errors).

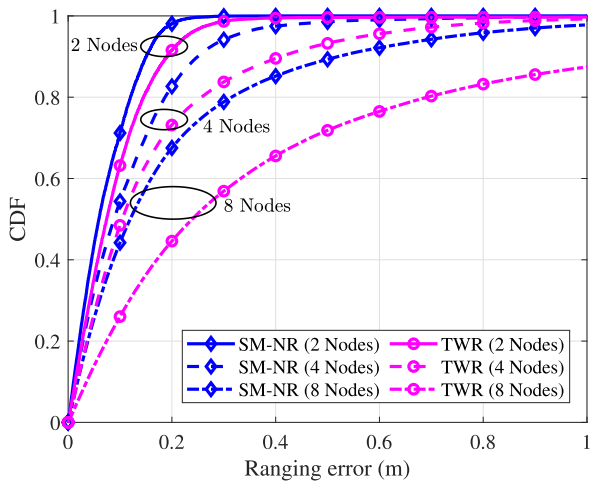


Fig. 19. CDF of the ranging errors when the number of active nodes is 2, 4, and 8 (w/o calibrating timestamp measurement errors).

ToF/TDoF estimation accuracy in static scenarios [58], it may introduce larger errors in dynamic scenarios due to delayed measurements, as the update rate decreases with more active nodes. Fig. 19 depicts the relationship between ranging accuracy and the number of active nodes. Hence, assigning more nodes to silent mode is beneficial for reducing resource consumption and essential for improving accuracy. Determining the optimal number of active nodes and dynamically selecting which nodes should operate in active mode need further investigation.

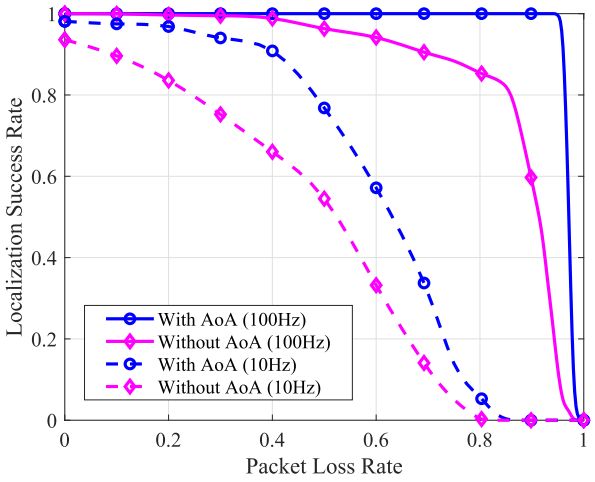


Fig. 20. Localization success rate with respect to different packet loss rates.

B. Robustness

The robustness of our localization method is validated under different packet loss rates. Packet loss refers to missing UWB measurements when UWB data fails to reach its intended destination, with higher loss rates indicating harsher propagation conditions. Using data collected from the 30-node experiment in Scenario 3, we conduct offline simulations with varying packet loss rates to model different levels of complex propagation environments. We block out a certain percentage of the UWB measurements according to the packet loss rate and perform localization using the incomplete data. We claim the localization in one trial is successful if the RMSE is less than 1 m. For each packet loss rate, we run simulations 200 times to derive the localization success rate curve as functions of packet loss rate varying from 0 to 100%. Note that a high packet loss rate has limited influence on initialization, since all nodes are stationary at the start of each test, enabling the collection of sufficient UWB measurements to establish reliable initial position estimates.

The performances with and without AoA measurements at the update rate of 100 and 10 Hz are shown in Fig. 20. The proposed localization method maintains a localization success rate of over 99% even when the packet loss rate reaches 93%. This performance is achieved by, on the one hand, enabling pairwise relative localization by incorporating AoA measurements and, on the other hand, leveraging SM-NR for high update frequency measurements. In comparison, when AoA measurements are unavailable, localization performance is significantly affected by the increase in the packet loss rate above 40%. This occurs because, without directional information, each node needs to establish wireless links with at least four neighboring nodes to achieve localization, which imposes restrictions on the connectivity of networks. As connectivity weakens, the localization failure rate rises. On the other hand, when the measurement update rate decreases to 10 Hz, the performance deteriorates significantly, implying the advantages of high update frequency for high-accuracy localization. The above results validate the robustness of our system in challenging propagation environments. They also highlight the importance of high-accuracy 3-D bearing capabilities and high update frequency for robust localization.

X. Conclusion

This article presents a robust and scalable multi-robot localization system using stereo UWB arrays. The designed array enables infrastructure-free 3-D pairwise localization, lowering network connectivity requirements and enhancing localization reliability in challenging environments. Beyond hardware design, we propose novel signal processing and localization methods to extract high-accuracy estimates from raw UWB measurements. The key components include UWB error calibration, robust direction and distance estimation, and distributed relative localization through spatial-temporal cooperation. Real-world experiments demonstrate that our localization method achieves centimeter-level accuracy at a 100-Hz update rate (UWB only), showing its robustness, scalability, and accuracy for multi-robot systems.

References

[1] E. Garcia, M. A. Jimenez, P. G. De Santos, and M. Armada, "The evolution of robotics research," *IEEE Robot. Autom. Mag.*, vol. 14, no. 1, pp. 90–103, Mar. 2007.

[2] M. Bloesch, S. Omari, M. Hutter, and R. Siegwart, "Robust visual inertial odometry using a direct EKF-based approach," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, Hamburg, Germany, 2015, pp. 298–304.

[3] R. Siegwart, I. R. Nourbakhsh, and D. Scaramuzza, *Introduction to Autonomous Mobile Robots*. Cambridge, MA, USA: MIT Press, 2011.

[4] S. Leutenegger et al., "Keyframe-based visual–inertial odometry using nonlinear optimization," *Int. J. Robot. Res.*, vol. 34, no. 3, pp. 314–334, Dec. 2015.

[5] Z. Huai and G. Huang, "Robocentric visual–inertial odometry," *Int. J. Robot. Res.*, vol. 41, no. 7, pp. 667–689, Jun. 2022.

[6] G. Huang, "Visual-inertial navigation: A concise review," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2019, pp. 9572–9582.

[7] M. Li and A. I. Mourikis, "High-precision, consistent EKF-based visual-inertial odometry," *Int. J. Robot. Res.*, vol. 32, no. 6, pp. 690–711, May 2013.

[8] W. Lee, K. Eckenhoff, P. Geneva, and G. Huang, "Intermittent GPS-aided VIO: Online initialization and calibration," in *Proc. IEEE Int. Conf. Robot. Autom.*, Paris, France, 2020, pp. 5724–5731.

[9] M. Z. Win, Y. Shen, and W. Dai, "A theoretical foundation of network localization and navigation," *Proc. IEEE*, vol. 106, no. 7, pp. 1136–1165, Jul. 2018.

[10] H. Zhao, N. Zhang, and Y. Shen, "Beamspace direct localization for large-scale antenna array systems," *IEEE Trans. Signal Process.*, vol. 68, pp. 3529–3544, May 2020.

[11] Y. Cao, C. Yang, R. Li, A. Knoll, and G. Beltrame, "Accurate position tracking with a single UWB anchor," in *Proc. IEEE Int. Conf. Robot. Autom.*, Paris, France, 2020, pp. 2344–2350.

[12] A. Ledergerber, M. Hamer, and R. D'Andrea, "A robot self-localization system using one-way ultra-wideband communication," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, Hamburg, Germany, 2015, pp. 3131–3137.

[13] X. Zhou et al., "Swarm of micro flying robots in the wild," *Sci. Robot.*, vol. 7, no. 66, May 2022, Art. no. eabm5954.

[14] J. González et al., "Mobile robot localization based on ultra-wide-band ranging: A particle filter approach," *Robot. Auton. Syst.*, vol. 57, no. 5, pp. 496–507, Nov. 2009.

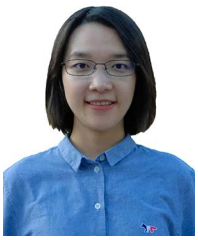
[15] W. Cui et al., "A robust mobile robot indoor positioning system based on Wi-Fi," *Int. J. Adv. Robot. Syst.*, vol. 17, no. 1, pp. 1–10, Jan. 2020.

[16] J. Fenwick, P. Newman, and J. Leonard, "Cooperative concurrent mapping and localization," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2002, pp. 1810–1817.

[17] E. Menegatti, A. Zanella, S. Zilli, F. Zorzi, and E. Pagello, "Range-only SLAM with a mobile robot and a wireless sensor networks," in *Proc. IEEE Int. Conf. Robot. Autom.*, Kobe, Japan, 2009, pp. 8–14.

[18] H. Xu et al., "Omni-swarm: A decentralized omnidirectional visual–inertial–UWB state estimation system for aerial swarms," *IEEE Trans. Robot.*, vol. 38, no. 6, pp. 3374–3394, Dec. 2022.
[19] T. Ziegler, M. Karrer, P. Schmuck, and M. Chli, "Distributed formation estimation via pairwise distance measurements," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 3017–3024, Apr. 2021.
[20] J. Yu et al., "SMMR-explore: Submap-based multi-robot exploration system with multi-robot multi-target potential field exploration method," in *Proc. IEEE Int. Conf. Robot. Autom.*, Xi'an, China, 2021, pp. 8779–8785.
[21] C. Cao et al., "Representation granularity enables time-efficient autonomous exploration in large, complex worlds," *Sci. Robot.*, vol. 8, no. 80, Jul. 2023, Art. no. eadf0970.
[22] P.-Y. Lajoie, B. Ramtoula, Y. Chang, L. Carlone, and G. Beltrame, "DOOR-SLAM: Distributed, online, and outlier resilient SLAM for robotic teams," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 1656–1663, Apr. 2020.
[23] Z. Zhang, J. Yu, J. Tang, Y. Xu, and Y. Wang, "MR-TopoMap: Multi-robot exploration based on topological map in communication restricted environment," *IEEE Robot. Autom. Lett.*, vol. 7, no. 4, pp. 10794–10801, Oct. 2022.
[24] H. Xu, L. Wang, Y. Zhang, K. Qiu, and S. Shen, "Decentralized visual-inertial-UWB fusion for relative state estimation of aerial swarm," in *Proc. IEEE Int. Conf. Robot. Autom.*, Paris, France, 2020, pp. 8776–8782.
[25] B. Friedlander and A. J. Weiss, "Direction finding in the presence of mutual coupling," *IEEE Trans. Antennas Propag.*, vol. 39, no. 3, pp. 273–284, Mar. 1991.
[26] T. Svantesson, "Modeling and estimation of mutual coupling in a uniform linear array of dipoles," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process.*, 1999, pp. 2961–2964.
[27] T. Liu, B. Li, and L. Yang, "Phase center offset calibration and multipoint time latency determination for UWB location," *IEEE Internet Things J.*, vol. 9, no. 18, pp. 17536–17550, Sep. 2022.
[28] I. Dotlic et al., "Angle of arrival estimation using DecaWave DW1000 integrated circuits," in *Proc. Workshop Positioning, Navigation Commun.*, Bremen, Germany, 2017, pp. 1–6.
[29] Z.-M. Liu, C. Zhang, and P. S. Yu, "Direction-of-arrival estimation based on deep neural networks with robustness to array imperfections," *IEEE Trans. Antennas Propag.*, vol. 66, no. 12, pp. 7315–7327, Dec. 2018.
[30] M. Naseri et al., "Machine learning-based angle of arrival estimation for ultra-wide band radios," *IEEE Commun. Lett.*, vol. 26, no. 6, pp. 1273–1277, Jun. 2022.
[31] T. Margiani, S. Cortesi, M. Keller, C. Vogt, T. Polonelli, and M. Magno, "Angle of arrival and centimeter distance estimation on a smart UWB sensor node," *IEEE Trans. Instrum. Meas.*, vol. 72, 2023, Art. no. 9508110.
[32] F. Ge and Y. Shen, "Single-anchor ultra-wideband localization system using wrapped PDoA," *IEEE Trans. Mobile Comput.*, vol. 21, no. 12, pp. 4609–4623, Dec. 2022.
[33] H. Kim, "Double-sided two-way ranging algorithm to reduce ranging time," *IEEE Commun. Lett.*, vol. 13, no. 7, pp. 486–488, Jul. 2009.
[34] Y. Jiang and V. C. M. Leung, "An asymmetric double sided two-way ranging for crystal offset," in *Proc. Int. Symp. Signals Syst. Electron.*, 2007, pp. 525–528.
[35] D. Neirynck, E. Luk, and M. McLaughlin, "An alternative double-sided two-way ranging method," in *Proc. Workshop Positioning, Navigation Commun.*, 2016, pp. 1–4.
[36] M. Pelka and H. Hellbrück, "S-TDoA–Sequential time difference of arrival—A scalable and synchronization free approach forl positioning," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Doha, Qatar, 2016, pp. 1–6.
[37] K. A. Horváth, G. Ill, and Á. Milánkovich, "Passive extended double-sided two-way ranging algorithm for UWB positioning," in *Proc. Int. Conf. Ubiquitous Future Netw.*, Milan, Italy, 2017, pp. 482–487.
[38] K. A. Horváth, G. Ill, and Á. Milánkovich, "Passive extended double-sided two-way ranging with alternative calculation," in *Proc. IEEE Int. Conf. Ubiquitous Wireless Broadband*, Salamanca, Italy, 2017, pp. 1–5.
[39] J. Tiemann, J. Pillmann, and C. Wietfeld, "Ultra-wideband antenna-induced error prediction using deep learning on channel response data," in *Proc. IEEE Semiannual Veh. Technol. Conf.*, Sydney, NSW, Australia, 2017, pp. 1–5.
[40] T. Wang, K. Hu, Z. Li, K. Lin, J. Wang, and Y. Shen, "A semi-supervised learning approach for UWB ranging error mitigation," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 688–691, Mar. 2021.
[41] A. D. Preter et al., "Range bias modeling and autocalibration of an UWB positioning system," in *Proc. Int. Conf. Indoor Positioning Indoor Navigation*, Pisa, Italy, 2019, pp. 1–8.
[42] A. Ledergerber and R. D'andrea, "Calibrating away inaccuracies in ultra wideband range measurements: A maximum likelihood approach," *IEEE Access*, vol. 6, pp. 78719–78730, 2018.
[43] Y. Li, S. Mazuelas, and Y. Shen, "A semi-supervised learning approach for ranging error mitigation based on UWB waveform," in *Proc. Mil. Commun. Conf.*, 2021, pp. 1–6.
[44] J. N. Ash and R. L. Moses, "On the relative and absolute positioning errors in self-localization systems," *IEEE Trans. Signal Process.*, vol. 56, no. 11, pp. 5668–5679, Nov. 2008.
[45] X. Shen, L. Xu, Y. Liu, and Y. Shen, "A theoretical framework for relative localization," *IEEE Trans. Inf. Theory*, vol. 70, no. 1, pp. 735–762, Jan. 2024.
[46] Y. Shang, W. Ruml, Y. Zhang, and M. Fromherz, "Localization from connectivity in sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 15, no. 11, pp. 961–974, Nov. 2004.
[47] M. Wei, A. Rosario, S. Carlos, and G. C. Calafiore, "Noisy range network localization based on distributed multidimensional scaling," *IEEE Sensors J.*, vol. 15, no. 3, pp. 1872–1883, Mar. 2015.
[48] Y. T. Chan and K. C. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 1905–1915, Aug. 1994.
[49] K. W. Cheung, H. C. So, W.-K. Ma, and Y. T. Chan, "Least squares algorithms for time-of-arrival-based mobile location," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 1121–1130, Apr. 2004.
[50] L. Xu, X. Shen, T. X. Han, R. Du, and Y. Shen, "An efficient relative localization method via geometry-based coordinate system selection," in *Proc. IEEE Int. Conf. Commun.*, Seoul, South Korea, 2022, pp. 4522–4527.
[51] Y. Liu, Y. Wang, J. Wang, and Y. Shen, "Distributed 3D relative localization of UAVs," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11756–11770, Oct. 2020.
[52] M. Z. Win, F. Meyer, Z. Liu, W. Dai, S. Bartoletti, and A. Conti, "Efficient multisensor localization for the Internet of Things: Exploring a new class of scalable localization algorithms," *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 153–167, Sep. 2018.
[53] Z. M. Kassas, M. Maaref, J. J. Morales, J. J. Khalife, and K. Shamei, "Robust vehicular localization and map matching in urban environments through IMU, GNSS, and cellular signals," *IEEE Intell. Transp. Syst.*, vol. 12, no. 3, pp. 36–52, Fall 2020.
[54] Z. Xiao, H. Wen, A. Markham, and N. Trigoni, "Indoor tracking using undirected graphical models," *IEEE Trans. Mobile Comput.*, vol. 14, no. 11, pp. 2286–2301, Nov. 2015.
[55] Z. Liu, W. Dai, and M. Z. Win, "Mercury: An infrastructure-free system for network localization and navigation," *IEEE Trans. Mobile Comput.*, vol. 17, no. 5, pp. 1119–1133, May 2018.
[56] Y. Li, Y. Wang, W. Yu, and X. Guan, "Multiple autonomous underwater vehicle cooperative localization in anchor-free environments," *IEEE J. Ocean. Eng.*, vol. 44, no. 4, pp. 895–911, Oct. 2019.
[57] X. Li et al., "A cooperative relative localization system for distributed multi-agent networks," *IEEE Trans. Veh. Technol.*, vol. 72, no. 11, pp. 14828–14843, Nov. 2023.
[58] Z. Zhang, H. Zhao, J. Wang, and Y. Shen, "Signal-multiplexing ranging for network localization," *IEEE Trans. Wireless Commun.*, vol. 21, no. 3, pp. 1694–1709, Mar. 2022.
[59] H. Zhao, Z. Zhang, L. Xu, Y. Wang, and Y. Shen, "Enhancing timeliness in asynchronous vehicle localization: A signal-multiplexing network measuring approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 10, pp. 13729–13743, Oct. 2024.
[60] R. T. Rajan and A.-J. van der Veen, "Joint ranging and synchronization for an anchorless network of mobile nodes," *IEEE Trans. Signal Process.*, vol. 63, no. 8, pp. 1925–1940, Apr. 2015.
[61] M. Titsias, "Variational learning of inducing variables in sparse Gaussian processes," in *Proc. 12th Int. Conf. Artif. Intell. Statist.*, D. van Dyk and M. Welling, Eds., Clearwater Beach, FL, USA, PMLR, 2009, pp. 567–574. [Online]. Available: https://proceedings.mlr.press/v5/titsias09a.html
[62] *IEEE Standard for Local and Metropolitan Area Networks–Part 15.4 Low-Rate Wireless Personal Area Networks (LR-WPANs)*, IEEE Standard 802.15.4-2011 (Revision of IEEE Std 802.15.4-2006), Sep. 2011.
[63] *Decawave DW1000 User Manual*, 2nd ed., Decawave Ltd, Dublin, Ireland, 2016.
[64] J. Hensman, A. Matthews, and Z. Ghahramani, "Scalable variational Gaussian process classification," in *Proc. Int. Conf. Artif. Intell. Statist.*, San Diego, CA, USA, 2015, pp. 351–360.
[65] H. Liu, Y.-S. Ong, and J. Cai, "Large-scale heteroscedastic regression via Gaussian process," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 2, pp. 708–721, Feb. 2021.

[66] F. Gustafsson et al., "Particle filters for positioning, navigation, and tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 425–437, Feb. 2002.
[67] Y. Liu, Y. Wang, X. Shen, J. Wang, and Y. Shen, "UAV-aided relative localization of terminals," *IEEE Internet Things J.*, vol. 8, no. 16, pp. 12999–13013, Aug. 2021.
[68] L. Xu, X. Shen, L. Wang, and Y. Shen, "A distributed relative localization scheme based on geometry merging priority," in *Proc. IEEE Glob. Telecommun. Conf.*, Rio de Janeiro, Brazil, 2022, pp. 1–6.
[69] LUSTER LightTech Co., Ltd., Hong Kong, "FZMotion mocap system," [Online]. Available: https://www.luster3ds.com



Hanying Zhao (Member, IEEE) received the B.E. degree in information engineering from Xidian University, Xi'an, China, in 2016, and the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 2021.

She was a Postdoctoral Researcher with Tsinghua University. Her research interests include statistical inference, multi-robot systems, and network localization.

Lingwei Xu (Student Member, IEEE) received the B.E. degree in electronic engineering, in 2022, from Tsinghua University, Beijing, China, where he is currently working toward the Ph.D. degree in electronic engineering.

His research interests include relative localization and cooperative networks.

Yi Li (Student Member, IEEE) received the B.E. degree in electronic engineering, in 2020, from Tsinghua University, Beijing, China, where he is currently working toward the Ph.D. degree in electronic engineering.

His research interests include array signal processing, cooperative networks, and UWB-based localization systems.

Feiyang Wen (Student Member, IEEE) received the B.S. degree in mechanics, in 2024, from Tsinghua University, Beijing, China, where he is currently working toward the M.E. degree in electronic and information engineering.

His research interests include cooperative networks and UWB-based localization systems.

Haoran Gao received the B.E. degree in electronic engineering, in 2022, from Tsinghua University, Beijing, China, where he is currently working toward the Ph.D. degree in electronic engineering.

His research interests include secure ranging and cooperative networks.

Changwu Liu received the B.E. degree in mechanics and the Ph.D. degree in aerospace engineering from Tsinghua University, Beijing, China, in 2018 and 2024, respectively.

He is currently a Postdoctoral Researcher with the Department of Electronic Engineering, Tsinghua University. His research interests include state estimation, nonlinear observers, geometric control methods and their various applications in robotics.

Jincheng Yu (Member, IEEE) received the B.S. and Ph.D. degrees in electronic engineering from Tsinghua University, Beijing, China, in 2016 and 2021, respectively.

He is currently a Research Assistant Professor with EE department, Tsinghua University. His research interests include DSA accelerator and AI robot application.

Dr. Yu is the Reviewer for ICRA, RAL and IROS, and also serves as an Assistant Editor for IROS 2025.

Yu Wang (Fellow, IEEE) received the B.S. and Ph.D. degrees (Hons.) in electronic engineering from Tsinghua University, Beijing, China, in 2002 and 2007, respectively.

He is currently a Full Professor with the Department of Electronic Engineering, Tsinghua University. His research interests include brain-inspired computing, parallel circuit analysis, application-specific acceleration and power/reliability aware circuit, and system design methodology.

Dr. Wang was the recipient of the Best Paper Award from ACM/SIGDA International Symposium on Field-Programmable Gate Arrays (FPGA) 2025, Asia and South Pacific Design Automation Conference (ASPDAC) 2019, FPGA 2017, IEEE Non-Volatile Memory Systems and Applications Symposium 2017, and IEEE Computer Society Annual Symposium on VLSI (ISVLSI) 2012, the Best Paper Honorable Mention Award from International Symposium on High-Performance Computer Architecture 2025, and the Best Poster Award from HEART 2012. He was also the recipient of the Alexander von Humboldt Fellowship in 2019, the DAC Under-40 Innovators Award in 2018, and the IBM X10 Faculty Award in 2010. He served as the TPC Chair for ASPDAC 2025, International Conference on Field-Programmable Technology (ICFPT) 2019 and 2011, ISVLSI 2018, the Finance Chair for International Symposium on Low Power Electronics and Design 2012–2016, the Track Chair for DATE 2017–2019 and Great Lakes Symposium on VLSI 2018. He is the Associate Editor for the IEEE TRANSACTIONS ON COMPUTER-AIDED DESIGN OF INTEGRATED CIRCUITS AND SYSTEM, *ACM Transactions on Embedded Computing Systems*, *ACM Transactions on Design Automation of Electronic Systems*, and IEEE EMBEDDED SYSTEM LETTER.

Yuan Shen (Senior Member, IEEE) received the B.E. degree in electronic engineering from Tsinghua University, Beijing, China, in 2005, and the S.M. and Ph.D. degrees in electrical engineering and computer science from Massachusetts Institute of Technology, Cambridge, MA, USA, in 2008 and 2014, respectively.

He is currently a Full Professor with the Department of Electronic Engineering, Tsinghua University. His research interests include network localization and navigation, integrated sensing and control, and multiagent systems.

Dr. Shen was the recipient of the IEEE ComSoc Fred W. Ellersick Prize and several best paper awards from IEEE conferences. He is currently an Editor of IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING, and *China Communications*. He has served as the TPC Symposium Co-Chair for IEEE ICC and IEEE Globecom for several times. He was the Elected Chair of the IEEE ComSoc Radio Communications Committee from 2019 to 2020.