Ultra-Wideband and Odometry-Based Cooperative Relative Localization With Application to Multi-UAV Formation Control

Kexin Guo, Xiuxian Li, *Member, IEEE*, and Lihua Xie, *Fellow, IEEE*

***Abstract*—This paper puts forth an infrastructure-free cooperative relative localization (RL) for unmanned aerial vehicles (UAVs) in global positioning system (GPS)-denied environments. Instead of estimating relative coordinates with vision-based methods, an onboard ultra-wideband (UWB) ranging and communication (RCM) network is adopted to both sense the inter-UAV distance and exchange information for RL estimation in 2-D spaces. Without any external infrastructures prepositioned, each agent cooperatively performs a consensus-based fusion, which fuses the obtained direct and indirect RL estimates, to generate the relative positions to its neighbors in real time despite the fact that some UAVs may not have direct range measurements to their neighbors. The proposed RL estimation is then applied to formation control. Extensive simulations and real-world flight tests corroborate the merits of the developed RL algorithm.**

***Index Terms*—Cooperative relative localization (RL), distributed formation control, GPS-denied environments, ultra-wideband ranging and communication (RCM) network, unmanned aerial vehicles (UAVs).**

I. Introduction

FORMATION control of multi-UAV systems has been a popular research topic in recent years [1] due to its extensive applications, such as surveillance, forest monitoring, and assistance in disaster areas. However, to realize real-life flights, relative positions between unmanned aerial vehicles (UAVs) are a prerequisite for most formation control methods [2]–[7]. Moreover, UAVs for formation applications should be designed to be as efficient as possible. This means that payload, sensing, and computational capabilities are limited. Therefore, how to simultaneously achieve relative position estimates with fewer resources and perform flight formation is a worthy research topic.

One simple method to obtain relative positions is to construct a communication network in which each UAV knows its own global position in a shared frame and communicates it with others. The UAVs can infer a relative estimate based on the shared data [3], [8]–[11]. Formation experimental results using global positions, particularly those involving quadcopter UAVs, already exist in both simulation and actual flight, but agents here must depend on some external infrastructure for positioning. Examples include global positioning system (GPS) [8], [9]; motion tracking systems [3], [10]; and radio-based positioning with anchors such as ultra-wideband (UWB) networks [11]. However, infrastructure dependence is disadvantageous. GPS is unreliable in cluttered urban environments, and the aforementioned alternatives require carefully setting up the operation area and have limited range. In addition, the accuracy of the infrastructures (e.g., positions of anchors and camera calibration) seriously affects the positioning accuracy. Especially, in some unknown environments such as forests, it is difficult to set up such infrastructures. Therefore, these infrastructure-dependent positioning systems are unsuitable for exploration tasks in an unknown environment. Even in [8], where the formation is distance-based and, thus, intended for use with interagent range measurements without infrastructure, the experiment reported was still GPS dependent.

On the other hand, infrastructure-independent examples also exist, but are deficient in some ways. In [12], a vision-based method was presented using onboard markers for relative localization (RL). Although promising, it is inherently restricted by limited view angles, occlusion and lighting, and heavy computations. The simultaneous localization and mapping (SLAM) method is able to provide position information by generating an environmental map onboard during flight, and obtain relative positions by sharing the position estimates through communication networks [13]. However, it still cannot be applied in some real-life environments, such as open areas, featureless hallways, an indoor environment with slanted walls, and forests with leaves and branches of tress. If the mission does not need to build a map, then mapping is a resource-intensive process to be discouraged. In [14], a Bluetooth-based RL method was proposed and intended for collision avoidance, but experimental results show its restrictions in flight duration and test area. A convex optimization-based RL was formulated in [15]. However, the method has high computational complexity and requires centralized implementation. Source localization with sampled range measurements was proposed in [16], where all mobile agents cooperatively estimate their

Manuscript received November 10, 2018; accepted February 21, 2019. Date of publication April 2, 2019; date of current version May 7, 2020. This work was supported in part by the Projects of Major International (Regional) Joint Research Program NSFC under Grant 61720106011. This paper was recommended by Associate Editor Y. Shi. *(Corresponding author: Lihua Xie.)*

K. Guo is with the School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China, and also with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: guok0005@e.ntu.edu.sg).

X. Li and L. Xie are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: elhxie@ntu.edu.sg).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCYB.2019.2905570

relative positions to a static source with an unknown position. However, it is still an infrastructure-dependent method and no practical tests were conducted to validate their approach. An integrated RL and leader–follower formation control in continuous time was proposed in [17]. Numerical simulations were carried out but there is no discussion on the influence of measurement noises. In [18], simultaneous velocity and relative position estimation with application to formation control was studied, but it requires that each agent must conduct a combined linear and circular motion during the entire formation period. Accurate angular velocity required in the circular motion makes it impractical for applications in a real platform. All of these observations motivate us to investigate an integrated system of RL and formation control, which should be practical and efficient for real-world applications.

In this paper, we present an infrastructure-free cooperative RL algorithm and demonstrate its application in formation control. The algorithm is validated by both simulations and experiments on a group of three quadcopters. The RL method uses only onboard sensors and interagent range measurements leveraging the same onboard UWB technology as reported in some previous works in [19] and [20], which can achieve submeter accuracy. The centimeter level of ranging accuracy within the radius of 150 m, and light weight, small size, and low power consumption make UWB a reliable and feasible ranging technique that can be applied for UAV. Unlike [20], which only discussed the problem of direct RL to a static UAV, this paper aims to localize each UAV relative to its neighbors. A consensus-based fusion scheme is proposed to fuse direct and indirect estimates which provides a more robust RL estimate in the presence of measurement noises and packet dropouts. The proposed RL estimation algorithm is then applied to formation control. Unlike [8], the system here is complete, integrating RL estimates with a suitable formation controller. The contributions of this paper are summarized as follows.

1) A cooperative RL strategy that is independent of infrastructures using only onboard sensors and interagent ranging: The RL estimation error is shown to be bounded under bounded measurement noises. Simulations and real-world tests are in accordance with the theoretical finding.
2) A consensus-based RL fusion to enhance robustness to measurement dropouts: The boundedness of the fused RL estimation error is analyzed, and experimental results showcase the efficacy of the developed approach.
3) Distributed formation flights using RL estimates: Numerical simulations and real-world formation flights of quadcopter UAVs demonstrate the effectiveness of the complete system, which integrates RL into formation control, by utilizing a customized UWB ranging and communication (RCM) network.

The remainder of this paper is structured as follows. Section II introduces the problem formulation and graph theory. Section III proposes a cooperative RL method and consensus-based RL fusion estimation, followed by a UWB network-based RL initialization strategy. Section IV discusses a discrete-time formation control algorithm by using the RL estimates. Simulations and flight experiments are conducted in Sections V and VI, respectively. Finally, concluding remarks are drawn in Section VII.

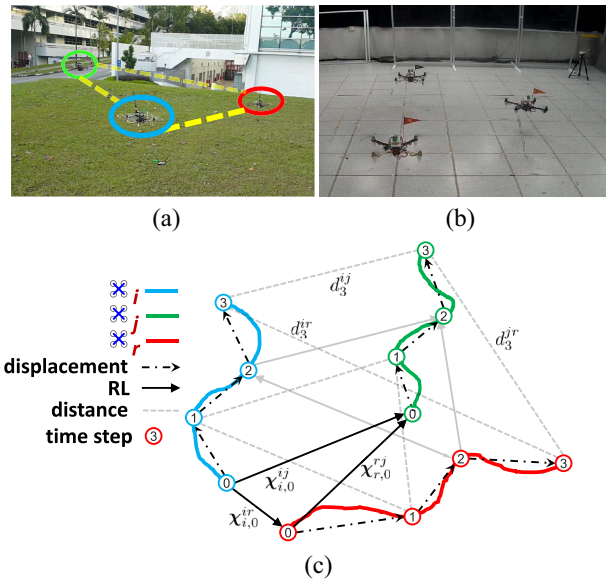


Fig. 1. Simultaneous infrastructure-free cooperative RL and distributed formation control strategy is proposed to overcome the shortcomings of widely used global positioning and/or visual detecting-based UAV swarm applications. (a) and (b) Distributed formation flights in different scenes. (c) Distances, displacements, and relative positions of three UAVs during their simultaneous movements. For clarity, all variables shown in (c) are depicted at different instants.

Notation: Throughout this paper, $\mathbb{N}$, $\mathbb{R}$, $\mathbb{R}^+$, $\mathbb{C}$, and $\mathbb{Z}^+$ represent the set of natural, real, positive real, complex, and positive integer numbers, respectively. For a matrix $\boldsymbol{A}$, notations $\|\boldsymbol{A}\|$ and $\boldsymbol{A}^\top$ denote its Euclidean norm and transpose.

II. Problem Formulation and Background

A. Problem of Interest

Consider a team of N UAVs (labeled consecutively from 1 to N) working collaboratively to localize themselves with respect to their neighbors. Meanwhile, each UAV employs these relative estimates to jointly perform formation.

Let $\mathcal{N}_i$ be the UAVi's neighbor set, in which UAVi is able to obtain distance d^{ij} from UAVj if $j \in \mathcal{N}_i$. That is, when performing a formation, neighboring UAVj is actually within an area with a perceiving radius $\bar{d}$, that is, $d^{ij} \leq \bar{d}$. Leveraging the obtained sensing information, UAVi aims to estimate the relative position of UAVj, $\boldsymbol{\chi}^{ij}_{i,k}$, at $t = t_{i,k}, k \in \mathbb{N}$ in UAVi's own frame $\mathcal{F}^M_i$ as depicted in Fig. 1(c). Then, based on the relative position estimates of its neighbors, each UAV generates a corresponding formation control command. In this paper, two scenarios are considered: 1) UAVi is able to sense UAVj and conduct direct RL estimation, that is, UAVj is a primary node for UAVi and 2) UAVi is only able to range and communicate with UAVr which is a neighbor of UAVj, but fails to sense UAVj directly, and UAVj herein serves as a secondary node.

Suppose each UAVi has an inertial frame $\mathcal{F}^I_i$ and a moving frame $\mathcal{F}^M_i$, for $i = 1, \ldots, N$. The moving frame $\mathcal{F}^M_i$, whose orientation is the same as that of $\mathcal{F}^I_i$, is attached to its body.

This can be satisfied if the UAVs carry compasses. In $\mathcal{F}_i^I$, UAVi is able to access its own velocity $\boldsymbol{v}_i$. Moreover, we assume that each UAVi is equipped with an onboard UWB module, which can measure the range d_k^{ij} to UAVj and obtain its corresponding change rate $\dot{d}_k^{ij}$ at $t = t_k, k \in \mathbb{N}$ in $\mathcal{F}_i^M$. In this paper, the 2-D case is considered. Consequently, it follows that $d_k^{ij} = d_k^{ji} = \|\boldsymbol{\chi}_k^{ij}\|$ and $\dot{\boldsymbol{\chi}}_k^{ij} = \boldsymbol{v}_k^{ij}$, where $\boldsymbol{\chi}_k^{ij} = (x_{j,k} - x_{i,k}, y_{j,k} - y_{i,k})^\top$ and $\boldsymbol{v}_k^{ij} = \boldsymbol{v}_{j,k} - \boldsymbol{v}_{i,k}$ are the relative coordinates and relative velocity of UAVj in UAVi's local frame, respectively, and (x_i, y_i) denotes the global coordinates[1] of UAVi. In addition, the velocity $\boldsymbol{v}_i$ of each UAVi, $i = 1, \ldots, N$, is continuously differentiable and satisfies that $\|\boldsymbol{v}_{i,k}\| \leq \bar{v}$, where $\bar{v}$ is a positive constant. That is, for any $i, j \in \mathcal{N}_i$, d^{ij} is differentiable and $\|\boldsymbol{v}_{i,k}^{ij}\| \leq 2\bar{v}$. The first objective is to develop an RL estimator such that each UAVi can estimate $\boldsymbol{\chi}_{i,k}^{ij}$, $j \in \mathcal{N}_i$. With these RL estimates and inter-UAV distance measurements, the next objective is to design a distributed distance-based formation control for real-world UAV swarm applications. To proceed, graph theory is introduced first.

B. Graph Theory

If each UAV is regarded as a node, then its coupling topology can be described by a graph. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denote a directed graph called a *sensing graph* which consists of a node set $\mathcal{V} = \{1, 2, \ldots, N\}$ and arc set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. A directed edge $(j, i) \in \mathcal{E}$ means that UAVi can measure its distance and the derivative of its distance to UAVj. To study the RL problem (e.g., estimate the relative positions to UAVj), we are also concerned with another *weighted digraph* graph $\mathcal{G}^j = (\mathcal{V}^j, \mathcal{E}^j, \mathcal{A}^j)$, which is composed of a set of nodes $\mathcal{V}^j = \mathcal{V}\backslash\{j\}$, a directed edge set $\mathcal{E}^j \subseteq \mathcal{V}^j \times \mathcal{V}^j$, and a weighted adjacency matrix $\mathcal{A}^j = [\kappa_{ir}^I] \in \mathbb{R}^{(N-1)\times(N-1)}$. $\mathcal{V}^j$ is the set of $N-1$ nodes excluding j, and if $r \in \mathcal{N}_i\backslash\{j\}$, there is an edge $(i, r) \in \mathcal{E}^j$. Moreover, $\forall i, r \in \mathcal{V}^j$, $\kappa_{ii}^I = 0$, $\kappa_{ir}^I > 0$ if $(i, r) \in \mathcal{E}^j$, and $\kappa_{ir}^I = 0$ otherwise. Furthermore, we also associate the N-UAV system with an undirected graph $\bar{\mathcal{G}} = \{\mathcal{V}, \bar{\mathcal{E}}\}$ called a *communication graph*, which is the corresponding undirected graph of $\mathcal{G}$. For an undirected edge $(j, i) \in \bar{\mathcal{E}}$, UAV$i$ and UAVj can exchange velocity and data packets mutually.

Denote $\boldsymbol{\mathcal{L}}^j$ as the Laplacian matrix of $\mathcal{G}^j$, whose diagonal elements $l_{ii} = \sum_{r\in\mathcal{N}_i\backslash\{j\}} \kappa_{ir}^I$ and for $i \neq r$, there holds $l_{ir} = -\kappa_{ir}^I$. For $\bar{\mathcal{G}}$, if there is a path in $\mathcal{G}$ from every UAVi in $\bar{\mathcal{G}}$ to UAVj, and we say that UAVj is *globally reachable* in $\mathcal{G}$, which is much weaker than *strong connectedness* (every node is reachable from every other node). Define $\boldsymbol{\mathcal{B}}^j \in \mathbb{R}^{(N-1)\times(N-1)}$ to be a j adjacency matrix associated with $\mathcal{G}$ with diagonal elements κ_{ij}^D for $i \in \mathcal{V}^j$, where $\kappa_{ij}^D > 0$ if UAVj is a neighbor of UAVi, that is, UAVi has a direct distance measurement of UAVj and $\kappa_{ij}^D = 0$ otherwise.

III. Cooperative RL Estimation

In this section, a cooperative RL estimator based on the UWB RCM network is presented. To enhance the robustness, a consensus-based RL fusion estimation method is also introduced. Finally, an RL initialization strategy for real implementations is given.

A. Persistent Excitation-Based RL Estimation

1) Incorporating Local Measurements: Suppose each UAVi is able to measure the relative distance d_k^{ij} to its neighbor j at sampling time instant $t_{i,k}$. Meanwhile, UAVi will transmit certain information, including $\boldsymbol{v}_{i,k}$, and local RL estimates $\hat{\boldsymbol{\chi}}_{i,k}^{ij}$, etc., to its neighbors (e.g., UAVj) by encoding the data packet into one ranging request command. For the listener's side, although UAVj can acquire its own velocity $\boldsymbol{v}_{j,t}$ continuously, $\boldsymbol{v}_{j,k}$ is actually sampled for calculation at $t = t_{j,k}$ once UAVi's information is received. A time-synchronization signal has been entrained in our UWB RCM[2] protocol for each cycle; thus, without consideration of the transmission time, we assume $t_{i,k} = t_{j,k}$. Since $j \in \mathcal{N}_i$, UAVi can receive UAVj's information as well. Assume that all of the clocks have been synchronized, based on the obtained information on UAVi, we aim to design an estimator for each UAVi in a team to localize the relative position $\boldsymbol{\chi}_{i,k}^{ij}$ in its own frame $\mathcal{F}_i^M$ at $t = t_{i,k}, j \in \mathcal{N}_i, k \in \mathbb{N}$.

Since the orientations of reference frames $\mathcal{F}_i^M$ and $\mathcal{F}_j^M$ for $i \neq j$ are consistent, it follows that $\boldsymbol{\chi}_{k+1}^{ij} = \boldsymbol{\chi}_k^{ij} + T\boldsymbol{v}_k^{ij}$ and $\boldsymbol{\chi}_{i,k}^{ij} = -\boldsymbol{\chi}_{j,k}^{ji}$. Taking the derivative of both sides of $d_t^{ij^2} = ||\boldsymbol{\chi}_t^{ij}||^2$ with respect to time in the continuous-time version, we can get $d_t^{ij}\dot{d}_t^{ij} = \boldsymbol{v}_t^{ij^\top} \cdot \boldsymbol{\chi}_t^{ij}$. By discretizing it, one obtains $d_k^{ij}\dot{d}_k^{ij} = \boldsymbol{v}_k^{ij^\top} \cdot \boldsymbol{\chi}_k^{ij}$. By taking the sensor noises into consideration, the RL estimation on UAVi at $t = t_{i,k}$ is given by

$$\begin{aligned}\hat{\boldsymbol{\chi}}_{i,k+1}^{ij} = \hat{\boldsymbol{\chi}}_{i,k}^{ij} + T\left(\boldsymbol{v}_{i,k}^{ij} + \boldsymbol{\epsilon}_k\right) + \gamma T\left(\boldsymbol{v}_{i,k}^{ij} + \boldsymbol{\epsilon}_k\right) \\ \times \left[\left(d_k^{ij} + \epsilon_k^d\right)\left(\dot{d}_k^{ij} + \epsilon_k^{\dot{d}}\right) - \left(\boldsymbol{v}_{i,k}^{ij} + \boldsymbol{\epsilon}_k\right)^\top \hat{\boldsymbol{\chi}}_{i,k}^{ij}\right]\end{aligned} \tag{1}$$

where $\boldsymbol{\epsilon}_k$, ϵ_k^d, and $\epsilon_k^{\dot{d}}$ are the measurement noises of $\boldsymbol{v}_{i,k}^{ij}$, d_k^{ij}, and $\dot{d}_k^{ij}$, respectively, at time k; $\gamma \in \mathbb{R}^+$ is a tunable constant gain; and T is the sampling period. It will be shown later that the RL estimation error is bounded in the noise-contaminated case.

2) RL Error Analysis With Corrupted Noise: The distance measurement sensor adopted herein (i.e., UWB) is able to achieve centimeter-level measurement accuracy [19]. Compared with the larger errors (tens of centimeters per second) in velocity estimation (i.e., Px4Flow) [21], [22], we simply assume that ϵ_k^d and $\epsilon_k^{\dot{d}}$ are equal to 0 in (1). To proceed, suppose there exists a constant $\bar{\delta} > 0$ such that $\|\boldsymbol{\epsilon}_k\| \leq \bar{\delta}$.

For $i = 1, \ldots, N$, we let $\tilde{\boldsymbol{\chi}}_i^{ij} = \hat{\boldsymbol{\chi}}_i^{ij} - \boldsymbol{\chi}_i^{ij}$. Since $d_k^{ij}\dot{d}_k^{ij} = \boldsymbol{v}_k^{ij^\top} \cdot \boldsymbol{\chi}_k^{ij}$, the error dynamic can be obtained as follows:

$$\tilde{\boldsymbol{\chi}}_{i,k+1}^{ij} = \boldsymbol{A}_{i,k}\tilde{\boldsymbol{\chi}}_{i,k}^{ij} + \boldsymbol{\varepsilon}_{i,k} \tag{2}$$

where $\boldsymbol{A}_{i,k} := \boldsymbol{I} - \gamma T(\boldsymbol{v}_{i,k}^{ij} + \boldsymbol{\epsilon}_k)(\boldsymbol{v}_{i,k}^{ij} + \boldsymbol{\epsilon}_k)^\top$ and $\boldsymbol{\varepsilon}_{i,k} := T\boldsymbol{\epsilon}_k - \gamma T(\boldsymbol{v}_{i,k}^{ij} + \boldsymbol{\epsilon}_k)\boldsymbol{\epsilon}_k^\top \boldsymbol{\chi}_{i,k}^{ij}$.

[1]The RL estimation method can work in 3-D spaces as well if the distance measurement, velocity, and relative position are all calculated in 3-D.

[2]Each UAV has both transmitter and listener roles in our UWB RCM protocol. Note that the larger the data packet, the longer the time duration of one ranging slot and, thus, practically, a tradeoff between the ranging update rate and the quantity of data transmission should be considered for a UWB RCM protocol design.

Before moving on, it is necessary to provide the following assumption.

Assumption 1 (Persistent Excitation Condition): There exist $m \in \mathbb{Z}^+$ and $\mu > 0$ such that $\forall k \in \mathbb{Z}^+$, the relative velocity measurement $\underline{\boldsymbol{v}}_{i,k}^{ij} := \boldsymbol{v}_{i,k}^{ij} + \boldsymbol{\epsilon}_k$, satisfies

$$\sum_{l=k}^{k+m-1} \underline{\boldsymbol{v}}_{i,l}^{ij} \underline{\boldsymbol{v}}_{i,l}^{ij\top} \geq \mu \boldsymbol{I}. \tag{3}$$

Remark 1: The physical meaning of Assumption 1 is that the two components of $\underline{\boldsymbol{v}}_{i,l}^{ij}$, denoted by $v_{ix,l}^{ij}$ and $v_{iy,l}^{ij}$, are linearly independent[3] for $l \in \{k, \ldots, k+m-1\}$. Expanding the left side of (3), we can get

$$\sum_{l=k}^{k+m-1} \underline{\boldsymbol{v}}_{i,l}^{ij} \underline{\boldsymbol{v}}_{i,l}^{ij\top} = \begin{bmatrix} \sum_{l=k}^{k+m-1} \left(v_{ix,l}^{ij}\right)^2 & \sum_{l=k}^{k+m-1} v_{ix,l}^{ij} v_{iy,l}^{ij} \\ \sum_{l=k}^{k+m-1} v_{ix,l}^{ij} v_{iy,l}^{ij} & \sum_{l=k}^{k+m-1} \left(v_{iy,l}^{ij}\right)^2 \end{bmatrix}.$$

Abiding by the Cauchy–Schwarz inequality, we have

$$\left(\sum_{l=k}^{k+m-1} v_{ix,l}^{ij} v_{iy,l}^{ij}\right)^2 \leq \sum_{l=k}^{k+m-1} \left(v_{ix,l}^{ij}\right)^2 \sum_{l=k}^{k+m-1} \left(v_{iy,l}^{ij}\right)^2$$

for any $k \in \mathbb{Z}^+$. And the two sides of the above inequality are equal if and only if $v_{ix,l}^{ij}$ and $v_{iy,l}^{ij}$ are linearly dependent, $l \in \{k, \ldots, k+m-1\}$. Hence, Assumption 1 is equivalent to the linear independence of $v_{ix,l}^{ij}$ and $v_{iy,l}^{ij}$, $l \in \{k, \ldots, k+m-1\}$.

Now, we are in a position to give the following result.

Theorem 1: For UAVi, under Assumption 1, if the sampling period T satisfies

$$0 < T < \frac{1}{\gamma\left(2\bar{v} + \bar{\delta}\right)^2} \tag{4}$$

then the estimation error $\|\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\|$ is bounded by $c := [(\xi_1 + \sqrt{\xi_1^2 + 4\sigma\xi_2})/2\sigma]$, where

$$\begin{aligned} \sigma &:= \mu\alpha\left[1 - \gamma T\left(2\bar{v} + \bar{\delta}\right)^2\right]^{2(m-1)} \\ \xi_1 &:= 2m\xi + \alpha\xi m(m-1)\left(2\bar{v} + \bar{\delta}\right)^2 \\ \xi_2 &:= \left[m^2 + \frac{1}{6}\alpha m(m-1)(2m-1)\right]\xi^2 \\ \alpha &:= 2\gamma T - \gamma^2 T^2\left(2\bar{v} + \bar{\delta}\right)^2 \\ \xi &:= T\bar{\delta}\left[1 + \gamma\left(2\bar{v} + \bar{\delta}\right)\bar{d}\right]. \end{aligned}$$

The proof can be found in the Appendix. ∎

Remark 2: Note that if there exists no velocity noise, that is, $\bar{\delta} = 0$, then $c = 0$ and the RL estimation error will asymptotically converge to zero. The effect of $\boldsymbol{\epsilon}$ on the RL estimation error and its upper bound c, will be discussed in Section V-A.

The proposed infrastructure-free cooperative RL strategy does not rely on any anchors and no precalibration is required. Leveraging only onboard UWB RCM and the built-in sensors of UAVs, an RL estimate with bounded error can be achieved under bounded measurement noises. Besides, the direct RL estimation, an indirect RL estimate can also be obtained through UWB RCM. Next, we will introduce a consensus-based RL fusion method to fuse the direct and indirect RL estimates to improve the robustness of RL estimation in noise-contaminated scenarios.

[3]The functions $f_1(l), \ldots, f_s(l)$ are called linearly dependent in an interval $\mathcal{T}$ if there exist $a_1, \ldots, a_s$, not all zero, such that $\sum_{i=1}^{s} a_i f_i(l) = 0, \forall l \in \mathcal{T}$, and linearly independent otherwise.

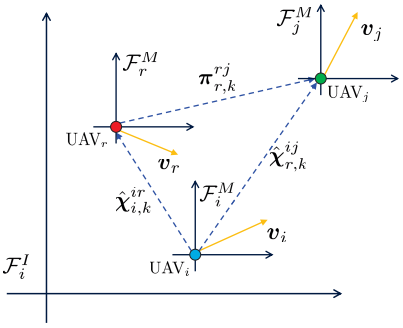


Fig. 2. Indirect RL estimate $\hat{\boldsymbol{\chi}}_{r,k}^{ij}$ from UAVi to UAVj through UAVr.

B. Consensus-Based RL Fusion Estimation

If UAVj is both a primary node and secondary node for UAVi, that is, UAVj is not only the neighbor of UAVi but also a neighbor of UAVr and $r \in \mathcal{N}_i$, then $\boldsymbol{\chi}^{ij}$ can be estimated indirectly. Let $\hat{\boldsymbol{\chi}}_{r,k}^{ij}$ be an *indirect* relative estimate of $\boldsymbol{\chi}^{ij}$ obtained by UAVi at $t = t_{i,k}$ through an intermediate UAVr, then $\hat{\boldsymbol{\chi}}_{r,k}^{ij} = \hat{\boldsymbol{\chi}}_{i,k}^{ir} + \boldsymbol{\pi}_{r,k}^{rj}$, as shown in Fig. 2. Moreover, based on the RL error analysis of (1) in Section III-A, we know that $\|\tilde{\boldsymbol{\chi}}_k^{ij}\|$ is bounded for each UAV.

Based on the RL error analysis of (1) in Section III-A, we know that $\|\tilde{\boldsymbol{\chi}}_k^{ij}\|$ is bounded for each UAV. Inspired by [16], we consider a noise-corrupted consensus-like RL fusion strategy for UAVi as follows:

$$\begin{aligned} \boldsymbol{\pi}_{i,k+1}^{ij} = \boldsymbol{\pi}_{i,k}^{ij} + T\left(\boldsymbol{v}_{i,k}^{ij} + \boldsymbol{\epsilon}_k\right) + \kappa_{ij}^D\left[\hat{\boldsymbol{\chi}}_{i,k}^{ij} - \boldsymbol{\pi}_{i,k}^{ij}\right] \\ + \sum_{r \in \mathcal{N}_i \setminus \{j\}} \kappa_{ir}^I\left[\hat{\boldsymbol{\chi}}_{r,k}^{ij} - \boldsymbol{\pi}_{i,k}^{ij}\right] \end{aligned} \tag{5}$$

where κ_{ij}^D and κ_{ir}^I are constants representing the *direct* estimation weight and *indirect* estimation weight, respectively. $\kappa_{ij}^D \neq 0$ if UAVi has a direct RL estimate to UAVj and $\kappa_{ij}^D = 0$ otherwise, which means that UAVs can obtain reliable and continual RL estimates in real time despite the fact that some UAVs may not have direct range measurements to their neighbors.

Theorem 2: If the conditions of Theorem 1 are satisfied and UAVj is globally reachable in $\mathcal{G}$ and $\kappa_{ij}^D + \sum_{r \in \mathcal{N}_i \setminus \{j\}} \kappa_{ir}^I < 1$, then the fused RL estimate $\boldsymbol{\pi}_i^{ij}$ of $\boldsymbol{\chi}_i^{ij}$ in (5) has a bounded estimation error.

This theorem is an extension of [16], but the difference is that here the fused RL estimation error is bounded under bounded direct and indirect RL estimates in the presence of measurement noises. Before proving the result, we present some preliminaries.

Definition 1 [23]: Consider a nonlinear dynamic system

$$\boldsymbol{x}(k+1) = f(\boldsymbol{x}(k), \boldsymbol{u}(k))$$

with the state $\boldsymbol{x}(k)$ and input $\boldsymbol{u}(k)$. It is said to be (globally) input-to-state stable (ISS) if for every initial state $\boldsymbol{\xi} = \boldsymbol{x}(0) \in \mathbb{R}^n$ and each $\boldsymbol{u}(k) \in \mathbb{R}^m$, there exists a class $\mathcal{KL}$

function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and a class $\mathcal{K}$ function η satisfying

$$|\boldsymbol{x}(k, \boldsymbol{\xi}, \boldsymbol{u})| \leq \beta(|\boldsymbol{\xi}|, k) + \eta(\|\boldsymbol{u}\|)$$

for each $k \in \mathbb{N}$.

Lemma 1 [23]: A linear discrete-time system

$$\boldsymbol{x}(k+1) = \boldsymbol{A}\boldsymbol{x}(k) + \boldsymbol{B}\boldsymbol{u}(k)$$

is ISS when $\boldsymbol{A}$ is a Schur matrix.

Lemma 2 [24]: The matrix $\boldsymbol{\mathcal{L}}^j + \boldsymbol{\mathcal{B}}^j$ is positive stable (its eigenvalues have positive real-parts) if and only if UAVj is globally reachable in $\mathcal{G}$.

Proof of Theorem 2: For $i = 1, \ldots, N$ and $i \neq j$, we let $\tilde{\boldsymbol{\pi}}_i^{ij} = \boldsymbol{\pi}_i^{ij} - \boldsymbol{\chi}_i^{ij}$. Then, it follows from (5) that:

$$\tilde{\boldsymbol{\pi}}_{i,k+1}^{ij} = \tilde{\boldsymbol{\pi}}_{i,k}^{ij} - \kappa_{ij}^D \tilde{\boldsymbol{\pi}}_{i,k}^{ij} - \sum_{r \in \mathcal{N}_i \backslash \{j\}} \kappa_{ir}^I \left[\tilde{\boldsymbol{\pi}}_{i,k}^{ij} - \tilde{\boldsymbol{\pi}}_{r,k}^{rj} \right] + \boldsymbol{u}_{i,k}^{ij} \tag{6}$$

where $\boldsymbol{u}_{i,k}^{ij} = \kappa_{ij}^D \tilde{\boldsymbol{\chi}}_{i,k}^{ij} + \sum_{r \in \mathcal{N}_i \backslash \{j\}} \kappa_{ir}^I \tilde{\boldsymbol{\chi}}_{i,k}^{ir} + T\boldsymbol{\epsilon}_k$ with $\|\boldsymbol{u}_{i,k}^{ij}\| \leq c(\kappa_{ij}^D + \sum_{r \in \mathcal{N}_i \backslash \{j\}} \kappa_{ir}^I) + T\bar{\delta}$.

Assume that at $t = t_{k+1}$, UAVj is globally reachable. We aggregate all of the systems (6) for $i \in \mathcal{V}^j$, and then we have

$$\tilde{\boldsymbol{\pi}}_{k+1}^{*j} = \left(\left(\boldsymbol{I} - \boldsymbol{\mathcal{L}}^j - \boldsymbol{\mathcal{B}}^j \right) \otimes \boldsymbol{I}_2 \right) \tilde{\boldsymbol{\pi}}_k^{*j} + \boldsymbol{u}_k^{*j} \tag{7}$$

where $\tilde{\boldsymbol{\pi}}_k^{*j} = [\tilde{\boldsymbol{\pi}}_{1,k}^{1j^\top} \ldots \tilde{\boldsymbol{\pi}}_{j-1,k}^{j-1j^\top}, \tilde{\boldsymbol{\pi}}_{j+1,k}^{j+1j^\top} \ldots \tilde{\boldsymbol{\pi}}_{N,k}^{Nj^\top}]$ and $\boldsymbol{u}_k^{*j} = [\boldsymbol{u}_{1,k}^{1j^\top} \ldots \boldsymbol{u}_{j-1,k}^{j-1j^\top}, \boldsymbol{u}_{j+1,k}^{j+1j^\top} \ldots \boldsymbol{u}_{N,k}^{Nj^\top}]$ with $* \in \mathcal{V}^j$. $\boldsymbol{\mathcal{L}}^j$ is the weighted Laplacian matrix of the graph $\mathcal{G}^j$ and $\boldsymbol{\mathcal{B}}^j \in \mathbb{R}^{(N-1)\times(N-1)}$ is a j adjacency matrix associated with $\mathcal{G}$.

Based on the Gersgorin disk theorem [25], all of the eigenvalues of $\boldsymbol{\mathcal{L}}^j + \boldsymbol{\mathcal{B}}^j$, denoted by λ, are located in the union of discs as follows:

$$\begin{aligned} &\mathrm{Ger}\left(\boldsymbol{\mathcal{L}}^j + \boldsymbol{\mathcal{B}}^j\right) \\ &:= \bigcup_{i=1}^{N-1} \left\{ \lambda_i \in \mathbb{C} : \sum_{r \in \mathcal{N}_i} \bigcup_{i=1}^{N} \left| \lambda_i - \left(\kappa_{ij}^D + \sum_{r \in \mathcal{N}_i \backslash \{j\}} \kappa_{ir}^I \right) \right| \right. \\ &\qquad \left. \leq \sum_{r \in \mathcal{N}_i \backslash \{j\}} \kappa_{ir}^I \right\}. \end{aligned} \tag{8}$$

By the condition in Theorem 2, we know that $\sum_{r \in \mathcal{N}_i \backslash \{j\}} \kappa_{ir}^I \leq 1$ and $\kappa_{ij}^D + \sum_{r \in \mathcal{N}_i \backslash \{j\}} \kappa_{ir}^I \leq 1$. Thus, we can infer that λ_i lies in a union of discs which are inside the unit circle centered at $(1, 0)$. Moreover, since UAVj is globally reachable in $\mathcal{G}$, by Lemma 2 that all the eigenvalues of $\boldsymbol{\mathcal{L}}^j + \boldsymbol{\mathcal{B}}^j$ are of positive real parts, it indicates that λ_i is strictly inside the unit circle centered at $(1, 0)$. Therefore, all eigenvalues of $\boldsymbol{\mathcal{L}}^j + \boldsymbol{\mathcal{B}}^j$ lie strictly inside the unit circle centered at the origin and accordingly, $\boldsymbol{\mathcal{L}}^j + \boldsymbol{\mathcal{B}}^j$ is a Schur matrix. Based on Lemma 1, (7) is ISS, which means for all $k \geq 0$, $\|\tilde{\boldsymbol{\pi}}_k^{*j}(\tilde{\boldsymbol{\pi}}_0^{*j}, \boldsymbol{u}^{*j}(k))\| \leq \beta(\|\tilde{\boldsymbol{\pi}}_0^{*j}\|, k) + \eta(\|\boldsymbol{u}_k^{*j}\|)$.

According to Theorem 1, for any positive constants κ_{ij}^D, κ_{ir}^I, $\|\boldsymbol{u}_{i,k}^{ij}\| \leq c(\kappa_{ij}^D + \sum_{r \in \mathcal{N}_i \backslash \{j\}} \kappa_{ir}^I) + T\bar{\delta}$. Since $\eta(\cdot)$ is a class $\mathcal{K}$ function, it follows that $\eta(\|\boldsymbol{u}_k^{*j}\|) \to \rho$ as $k \to \infty$, where ρ is a constant. Because $\|\tilde{\boldsymbol{\pi}}_0^{*j}\|$ is bounded and $\beta(\cdot, \cdot)$ is a class $\mathcal{KL}$ function, $\beta(\|\tilde{\boldsymbol{\pi}}_0^{*j}\|, k) \to 0$ as $k \to \infty$. Consequently, $\|\tilde{\boldsymbol{\pi}}_k^{*j}(k, \tilde{\boldsymbol{\pi}}_0^{*j}, \boldsymbol{u}^{*j}(k))\| \to \rho$ as $k \to \infty$. Thus, the proof is completed. ■

Remark 3: It should be noted that the graph $\mathcal{G}^j$ might be time varying due to possible failures of range measurements. If a digraph $\mathcal{G}$ is strongly connected at $t = t_{k+1}$ (i.e., there exists a path between any two distinct UAVs), besides the relative estimation to UAVj, for example, each UAV estimates the relative position to UAVq, $q \in \mathcal{V}\backslash\{j\}$, accordingly, $\tilde{\boldsymbol{\pi}}_k^{*q}$, $\boldsymbol{\mathcal{L}}^q$, and $\boldsymbol{\mathcal{B}}^q$ can be obtained.

Remark 4: Before the UAVs take off, a nonlinear regression (NLR)-based method [11] can be applied to estimate the coordinates of the static UAVs, and these estimates will serve as an initial guess of the RL estimators in Section III to improve RL accuracy. To apply this technique, a UWB RCM scheme is developed in Section VI-B.

Remark 5: The proposed RL estimation method can be seen as a weighted mean value of the direct and indirect estimates which can increase the convergence speed and provide a more robust RL estimate in noise-corrupted cases.

IV. Application to Formation Control

Combining interagent RL with distributed formation control is still challenging for practical applications especially in the presence of noise. Based on global position information, Deghat *et al.* [26] proposed a continuous-time algorithm for a double-integrator model. However, this continuous-time version cannot be implemented directly since range measurements between agents are discrete. Thus, in this section, a discrete-time formation control law involving RL estimates and measurement noises is proposed as follows by discretizing the aforementioned continuous-time algorithm:

$$\begin{aligned} \boldsymbol{p}_{i,k+1} &= \boldsymbol{p}_{i,k} + T\boldsymbol{v}_{i,k} \\ \boldsymbol{v}_{i,k+1} &= \boldsymbol{v}_{i,k} + \gamma_1 T \sum_{j \in \mathcal{N}_i} \underline{\boldsymbol{v}}_{i,k}^{ij} + \gamma_2 T \sum_{j \in \mathcal{N}_i} (d_k^{ij^2} - d_{ij}^{*2}) \boldsymbol{\pi}_{i,k}^{ij} \end{aligned} \tag{9}$$

where T is the sampling period, γ_1 and γ_2 are small positive constants, and $\underline{\boldsymbol{v}}_{i,k}^{ij}$ and $\boldsymbol{\pi}_{i,k}^{ij}$ are relative velocity measurements and fused RL estimates obtained from (5).

By combining the RL algorithm developed in the last section and the formation control algorithm (9), an integrated cooperative RL and formation control algorithm is proposed in Algorithm 1.

V. Simulations

In this section, we demonstrate the cooperative RL strategy with six mobile UAVs in consideration of the sudden change in trajectory, and then apply the estimates to multi-UAV formation control.

A. Cooperative RL Simulation Results

1) Configuration: Six mobile UAVs are labeled as $1, 2, \ldots, 6$, respectively. The sensing graph of the six UAVs is depicted in Fig. 3. It can be checked that in the sensing graph, UAV1 is *globally reachable*. UAV2, 3, 4 have direct RL estimates to UAV1 while UAV5 and UAV6 only have indirect RL estimates to UAV1. Besides, UAV3 is able to obtain two

Algorithm 1: Pseudocode for Simultaneous RL Estimation $\boldsymbol{\pi}_{i,k}^{ij}$ and Distributed Formation Control

Procedure: RL Initialization

Ranging and communication, then conduct NLR: **while** $\Delta\boldsymbol{\phi} > \kappa$ **do**
  for $i \leftarrow 1$ *to* N **do**
    if $i == N$, $j \leftarrow 0$; otherwise $j \leftarrow i+1$
    UAVi sends range request and data package to UAVj
    Calculate the relative distance $\underline{d}_k^{ij}$
  Conduct NLR based RL initialization [11]
Each UAV stores the initial RL estimates $\hat{\boldsymbol{\chi}}_0^{ij}$

Procedure: Cooperative RL and Formation Control

while *CooperativeRL* **do**
  for $i \leftarrow 1$ *to* N **do**
    if $i == N$, $j \leftarrow 0$; otherwise $j \leftarrow i+1$
    UAVi sends range request and data package to UAVj
    Calculate the relative distance $\underline{d}_k^{ij}$, distance rate $\dot{\underline{d}}_k^{ij}$, and indirect relative velocity $\underline{\boldsymbol{v}}_{i,k}^{ij}$
    $\hat{\boldsymbol{\chi}}_{i,k+1}^{ij} \leftarrow (1)$
    $\boldsymbol{\pi}_{i,k+1}^{ij} \leftarrow (5)$
    Calculate the control command $\leftarrow (9)$
    Send command to flight control board

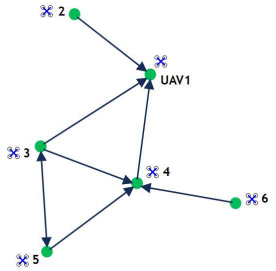


Fig. 3. Sensing graph.

indirect RL estimates to UAV1 through UAV4 and UAV5, and UAV5 has two indirect RL estimates to UAV1 through UAV3 and UAV4. The six UAVs' movements are as follows:

$$\boldsymbol{v}_{1,k} = \begin{bmatrix} \cos\left(\frac{1}{3}k\right) \\ -\frac{5}{3}\sin\left(\frac{1}{3}k\right) \end{bmatrix}, \boldsymbol{v}_{2,k} = \begin{bmatrix} -2\sin(k) \\ 2\sin(k) \end{bmatrix}$$

$$\boldsymbol{v}_{3,k} = \begin{bmatrix} \cos\left(\frac{1}{5}k\right) & -\sin\left(\frac{1}{5}k\right) \\ \sin\left(\frac{1}{5}k\right) & \cos\left(\frac{1}{5}k\right) \end{bmatrix} \begin{bmatrix} 1 \\ \cos(k) \end{bmatrix}, \boldsymbol{v}_{5,k} = \begin{bmatrix} \frac{1}{6} \\ 0 \end{bmatrix}$$

$$\boldsymbol{v}_{4,k} = \begin{bmatrix} -3\sin(k) \\ 3\cos(k) \end{bmatrix}, \boldsymbol{v}_{6,k} = \begin{bmatrix} -\frac{10}{3}\sin(\frac{1}{3}k) \\ \frac{5}{3}\cos(\frac{1}{3}k) \end{bmatrix}$$

and they are initially located at $(0, 0)$, $(2, -30)$, $(20, -15)$, $(-20, 8)$, $(-14, 8)$, and $(-10, -30)$, respectively.

2) Fused RL Estimation π: To demonstrate the success of the proposed estimation scheme, we set UAV1 as a relative target to estimate. In the first simulation, we let all of the UAVs take different continuous motions as seen in Fig. 4(a), which include both linear and nonlinear motions. In the second simulation, UAV4 suffers from a sudden deviation of velocity at $t = 100$ s, while others remain the same as shown in Fig. 4(d).

TABLE I
STATISTICS OF RL ERRORS WITH $\bar{\delta} = 0.5$ m/s, $\gamma = 0.5$ (UNIT: m)

Case		π_{21}	π_{31}	π_{41}	π_{51}	π_{61}
Fig. 4(b)	mean	0.1145	0.1359	0.2078	0.2123	0.2105
	variance	0.0048	0.0046	0.0086	0.0071	0.0122
Fig. 4(e)	mean	0.1231	0.1742	0.2305	0.2274	0.2167
	variance	0.0185	0.0123	0.0198	0.0133	0.0177

In this section, the positions of all UAVs relative to UAV1 are discussed and demonstrated since each UAV is able to *reach* UAV1 as shown in the sensing graph of Fig. 3. Velocity noise is bounded with $\bar{\delta} = 0.5$ m/s, and ϵ^d and $\epsilon^{\dot{d}}$ are considered as well with a bound of 0.05 m. Choose $T = 0.05$ s (20 Hz),[4] $\gamma = 0.5$, $\kappa_{ij}^D = [\alpha_{ij}/(|\mathcal{N}_i| + 1 + \alpha_{ij})]$, and $\kappa_{ij}^I = [1/(|\mathcal{N}_i| + 1 + \alpha_{ij})]$, where $|\mathcal{N}_i|$ is the cardinality of $\mathcal{N}_i$, and $\alpha_{ij} = 1$ if UAVi has a direct measurement of UAVj, and $\alpha_{ij} = 0$ otherwise. Accordingly, the conditions of Theorems 1 and 2 are satisfied. Since each UAV conducts the RL estimation from itself to UAV1 and performs the RL fusion in its own frame, for the sake of brevity, here, we denote $\boldsymbol{\pi}_i^{i1}$ by $\boldsymbol{\pi}_{i1}$ and its corresponding real RL is $\boldsymbol{\chi}_{i1}$. The evolution curves of the estimation errors $||\boldsymbol{\pi}_{i1} - \boldsymbol{\chi}_{i1}||$ for the first and second simulation are depicted in Fig. 4(b) and (e), respectively, and the RL estimation errors are bounded with 0.4 m in both cases. It is worth noting that in Fig. 4(e), the fused estimates $\boldsymbol{\pi}_{31}$, $\boldsymbol{\pi}_{51}$, and $\boldsymbol{\pi}_{61}$ fluctuate slightly in the presence of a significant variation of $\boldsymbol{\pi}_{41}$ at $t = 100$ s, and this demonstrates that our proposed cooperative RL method is robust to unexpected maneuverings. The statistics of the errors are shown in Table I which reveals that the mean estimation errors of these two cases fall into the same magnitude and the errors in Fig. 4(e) are slightly larger than those in Fig. 4(b) due to the jerk of UAV4. Although the variances of the errors in Fig. 4(e) are an order of magnitude larger than those in Fig. 4(b), it is still acceptable since they are in centimeter-level.

On the other hand, in the presence of noises and data loss, the fused estimator averages the impact of noises and data dropouts. To some extent, additional information in indirect estimation helps speed up convergence rate of the RL estimation. For example, see the evolutions of direct, indirect, and fused estimation error of $\boldsymbol{\chi}_{31}$ in the presence of noises and data dropouts in Fig. 4(f). Obviously, the simulation result with fusion converges faster than the direct estimation no matter in the beginning or when the data transmission recovers.

3) Direct RL Estimation $\hat{\boldsymbol{\chi}}$ and the Influence of ϵ and γ: To show the influence of different velocity noises on the accuracy of the RL estimation, $\boldsymbol{\epsilon}$ bounded by $\bar{\delta} = 0.1$–0.7 m/s are demonstrated. Other configurations are the same as the second simulation. For a given γ, without loss of generality, the evolutions of $\|\hat{\boldsymbol{\chi}}_{53} - \boldsymbol{\chi}_{53}\|$ under different conditions of $\bar{\delta}$ are shown in Fig. 5(a) and their statistical results of the estimation

[4]UWB module PulsON 440 can provide a ranging measurement interval of 25 ms (40 Hz) by default configuration, which is sufficient for the inner control loop of a UAV in most scenarios.

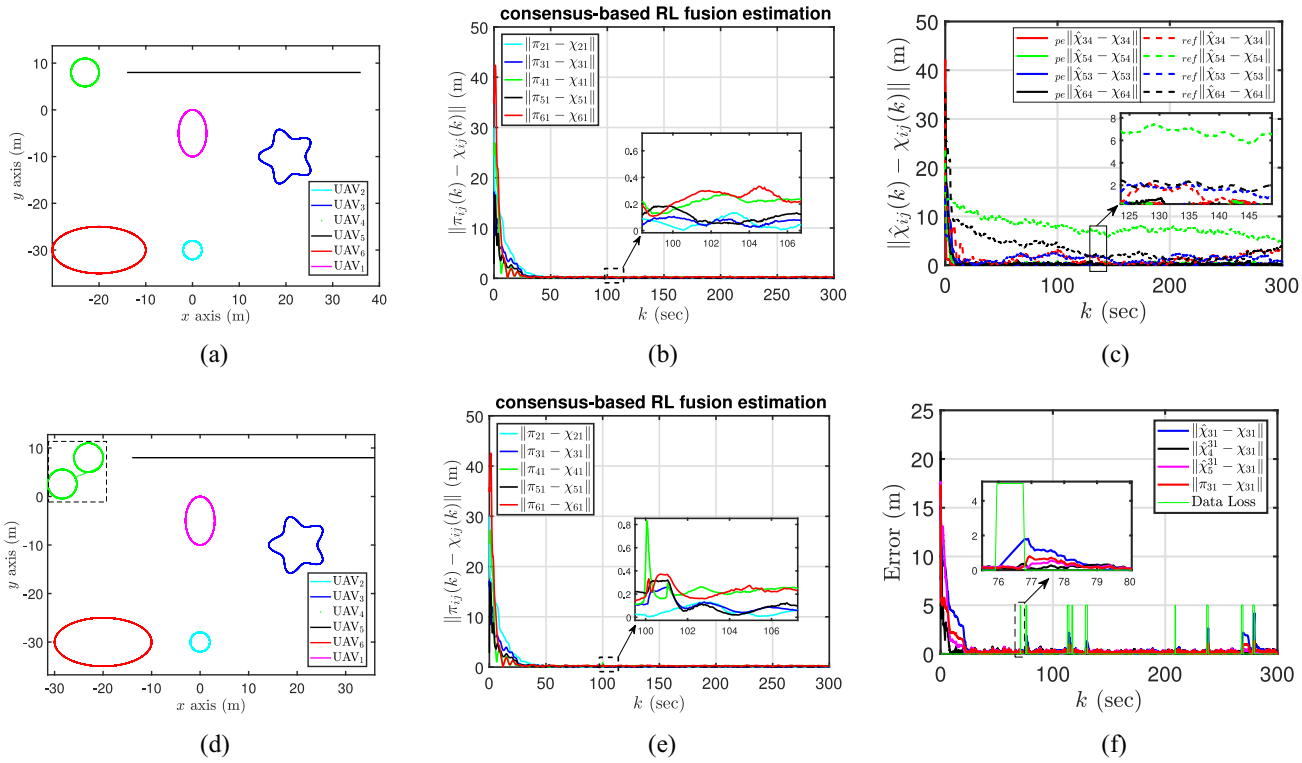


Fig. 4. RL estimation results in the cases of continuous and hopping trajectories. (a) Continuous trajectories of six UAVs. (b) Evolution of RL estimation error $\|\boldsymbol{\pi}_{i1} - \boldsymbol{\chi}_{i1}\|$ calculated on UAVi, $i = 2, \ldots, 6$. (c) Comparison of direct RL estimation errors between the proposed method, ${}_{pe}\|\cdot\|$, and the method in [20], ${}_{ref}\|\cdot\|$. (d) Trajectories of six UAVs with a sudden change of UAV4. (e) Evolution of the RL estimation error in the presence of jumping on UAV4 at 100 s. (f) Evolutions of direct, indirect, and fused estimation errors of $\boldsymbol{\chi}_{31}$.

TABLE II
STATISTICS OF RL ERRORS OF $\boldsymbol{\pi}_{53}$ UNDER DIFFERENT $\bar{\delta}$, γ (UNIT: m)

γ		$\bar{\delta}$			
		0.1m/s	0.3m/s	0.5m/s	0.7m/s
0.1	mean	0.0685	0.1397	0.1541	0.1765
	variance	0.0020	0.0084	0.0092	0.0148
0.5	mean	0.1034	0.2439	0.3530	0.3991
	variance	0.0032	0.0223	0.0677	0.0688

errors are presented in Table II. It can be seen that regardless of $\gamma = 0.1$ or 0.5, the mean and variance values of the errors become larger with the increase of $\bar{\delta}$, which is consistent with Theorem 1.

On the other hand, for given $\bar{\delta}$, γ should be chosen to satisfy (4). The trajectories of $\|\hat{\boldsymbol{\chi}}_{53} - \boldsymbol{\chi}_{53}\|$ and its upper bound with $\gamma = 0.5$ and 0.1 are exhibited in Fig. 5(b) and (c), respectively. The statistics of RL errors are shown in Table II. It can be seen that under the same $\bar{\delta}$, a smaller γ can help to reduce the estimation errors but decelerate convergence speed while a larger γ is reversed. Therefore, a dynamic γ can be adopted, which means at the beginning of RL, a larger γ can be chosen to speed up the convergence of RL, then reduce γ to narrow the RL error. A time window of 3 s is chosen for Assumption 1, that is, $m = 3/T$. If μ is obtained at the end of flight, then it should be fixed and c will be a constant. In this simulation, we adopt m latest measurements to dynamically calculate μ with the increase of time step k. Thus, the upper bound curve is wavy but the peak value can be chosen for reference such as 0.6 m in Fig. 5(c). To summarize, the relationship among $\bar{\delta}$, γ, and c can be seen in Fig. 6. Next, we will combine the proposed cooperative RL with UWB-based formation control.

4) Comparison With the Method in [20]: Note that the method introduced in [20] only considered the direct RL without fusion. The 2-norm of the direct RL estimation errors of $\hat{\boldsymbol{\chi}}_{34}$, $\hat{\boldsymbol{\chi}}_{54}$, $\hat{\boldsymbol{\chi}}_{53}$, and $\hat{\boldsymbol{\chi}}_{64}$ are depicted in Fig. 4(c). The sensing graph and configurations are the same as Fig. 4(a). In these two cases, $\hat{\boldsymbol{\chi}}_{34}$, $\hat{\boldsymbol{\chi}}_{54}$, $\hat{\boldsymbol{\chi}}_{53}$, and $\hat{\boldsymbol{\chi}}_{64}$ share the same initial positions of $[-4, 1]^\top$, $[17, -5]^\top$, $[23, -11]^\top$, and $[-4, 3]^\top$. It can be seen that based on the proposed method, all direct RL estimation errors are bounded by 0.5 m at 20 s. However, the errors using the method in [20] can be up to 8 m, which are much larger than the results of the proposed method.

B. Combining Cooperative RL and Distributed Formation Control

Consider a three-UAV team where each UAV is able to measure the distance to the other two UAVs. The control goal is to form two triangular formations with velocity consensus. In the simulation, we set $\gamma_1 = 1$ and $\gamma_2 = 1.4 \times 10^{-2}$. Other parameters are the same as those in Section V-A. Two equilateral triangular formations are preconfigured with $d_{ij}^* = 15$ and 30 m sequentially.

Fig. 7(a) and (b) shows that two equilateral triangular formations are formed successively and sustained by true relative positions and RL estimates, respectively. Compared with the

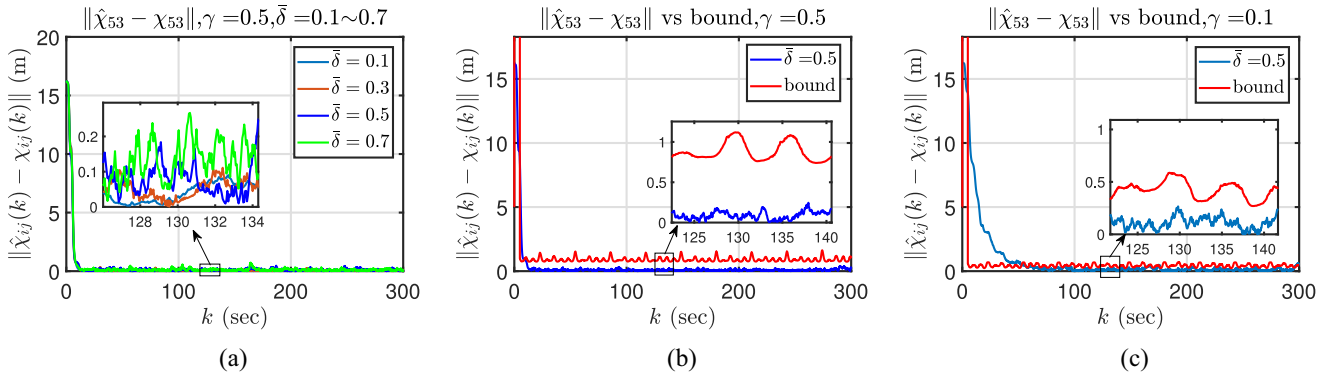


Fig. 5. Absolute values of RL estimation error, $\|\hat{\chi}_{53}-\chi_{53}\|$, and its upper bound c. (a) $\|\hat{\chi}_{53}-\chi_{53}\|$ with respect to $\bar{\delta}$ = 0.1–0.7 m/s. (b) $\|\hat{\chi}_{53}-\chi_{53}\|$ and its upper bound c with $\gamma = 0.5$ and $\bar{\delta} = 0.5$ m/s. (c) $\|\hat{\chi}_{53}-\chi_{53}\|$ and its upper bound c with $\gamma = 0.1$ and $\bar{\delta} = 0.5$ m/s.

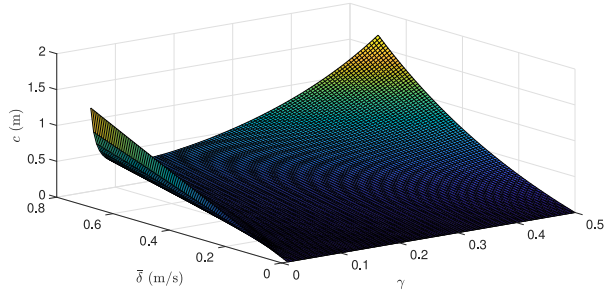


Fig. 6. Relationship among $\bar{\delta}$, γ, and c.

noiseless case in Fig. 7(a) (without measurement noises and using true relative positions), the significant difference in Fig. 7(b) is the flight trajectories, which are affected by the measurement noises. The desired formations are achieved at 11 and 15 s, respectively, and the inter-UAV distances are maintained at 30 m as seen in Fig. 7(c). Both of these two cases achieve velocity consensus as shown in Fig. 7(d), and the fluctuating relative velocity between 0 and 20 s indicates the change of formation. The performances demonstrate that the proposed RL algorithm is able to support multiple UAVs to achieve formation in the presence of measurement noises.

VI. Flight Tests

In this section, we present extensive flight experimental evaluations to demonstrate our proposed simultaneous cooperative RL and distributed formation control system. At first the performance of cooperative RL is evaluated in terms of average absolute error in comparison with the ground truth provided by VICON[5] (a motion capture system with milimeter-level positioning accuracy). Then, this cooperative RL combined with the proposed distributed formation control is demonstrated by three-quadcopter flights.

A. System Configuration

Fig. 8 illustrates the hardware configuration and signal flow of the UWB-based RL system. UWB modules PulsON 440[6] are installed on three quadcopters. Due to the large bandwidth (from 3.1 to 5.3 GHz), UWB is robust to multipath and non-line-of-sight effects, and provides a reliable long distance ranging with an accuracy of 10 cm. Algorithms, sensing, and communication are executed on a credit-card-sized mobile-level Raspberry Pi2. The quadcopter is equipped with Pixhawk and Px4Flow[7] for flight control and onboard velocity estimation. Note that other equivalent velocity or displacement sensing systems can be adopted as well. Moreover, high-precision velocity measurements can help to improve the RL accuracy.

The UWB module on each quadcopter actively sends ranging requests to neighboring UAVs for distance measuring and communication based on the two-way time-of-flight (TWToF) ranging method. A customized UWB RCM network protocol will be introduced in Section VI-B. In this network, an instant distance estimate obtained by UAVi will be calibrated first through linear regression [19]. Then, the corrected range measurement together with the received information go through Algorithm 1. The output of (9) is an acceleration command which will be sent to the low-level control board, Pixhawk. In the control board, this acceleration command will be transformed into attitude command and further into pulse-width modulating (PWM) signals to control motors for UAV flight as depicted in Fig. 9. All the processes above are controlled by Raspberry Pi2 and XBee only receives each UAV's global position and velocity from VICON as ground truth.

B. Ultra-Wideband Ranging and Communication Network Configuration

A contention-free time division multiple access (TDMA) time-slotted network is used to coordinate communication and range measurements. Each node transmits its previous range and velocity of range measurement provided by UWB firmware, together with its current velocity, RL and height estimate as data packet embedded in the range request packet. All nodes can overhear this data and provide it to the processor. An example of slot map with three nodes identified as 200–202 (i.e., UAV0-2) is given in Table III. Each slot supports

[5]https://www.vicon.com/

[6]https://timedomain.com/products/pulson-440/

[7]https://pixhawk.org/modules/px4flow

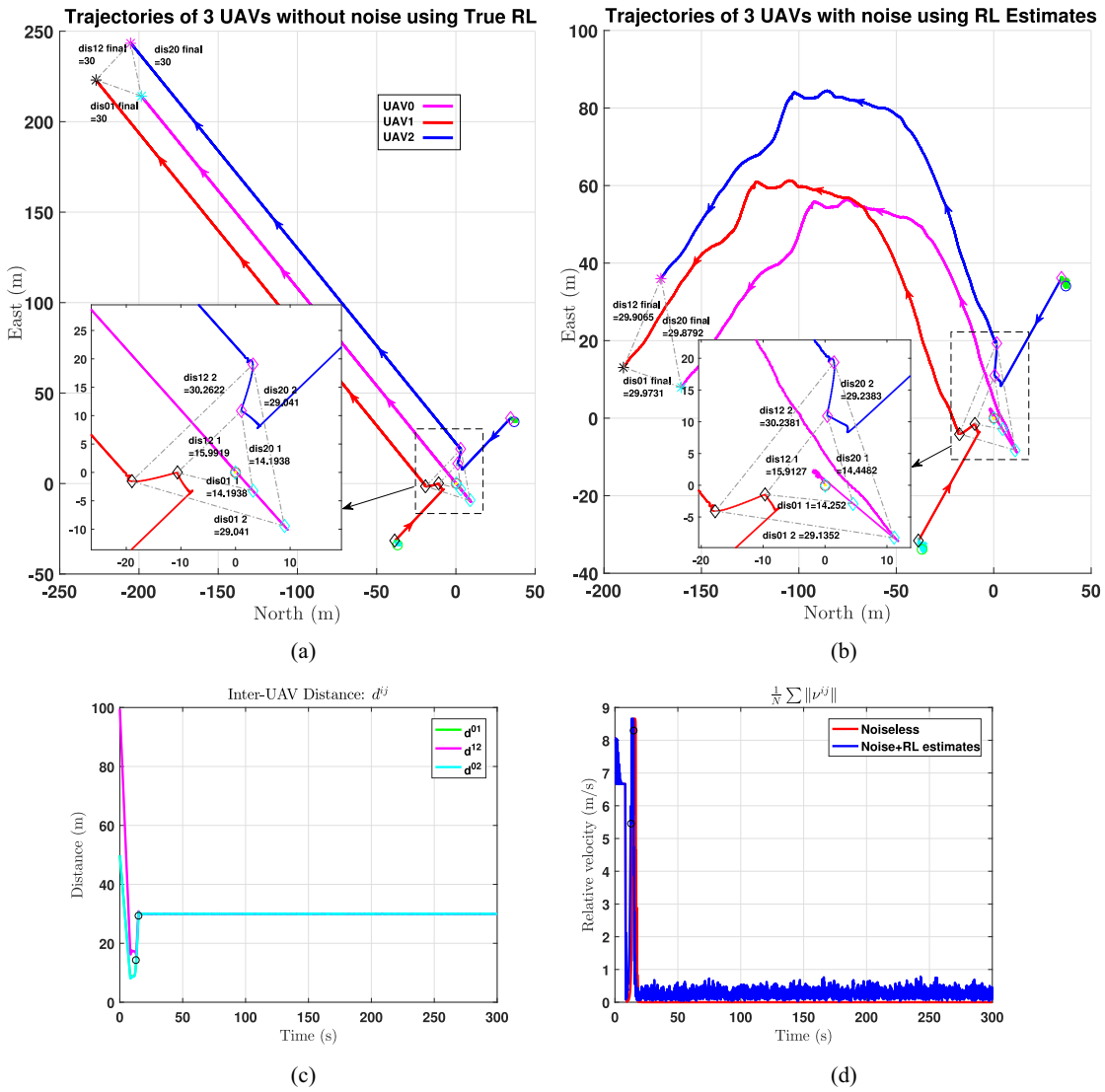


Fig. 7. Simulation results of formation control using RL estimates in noise-corrupted case and the system dynamics are governed by (9). (a) and (b) Flight trajectories of three UAVs using true RL in noise-free conditions, and using RL estimates in noise-corrupted case, respectively. (c) Evolutions of inter-UAV distances in noise-corrupted case. (d) Average of the 2-norm of the relative velocities in two cases: noiseless measurements + true RL (red) case and noisy measurements + RL estimates (blue) case.

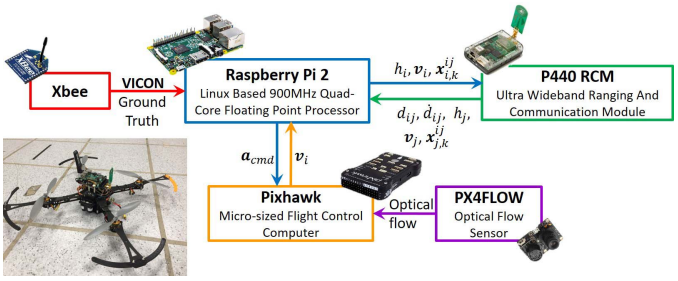


Fig. 8. Flight platform and diagram of system workflow.

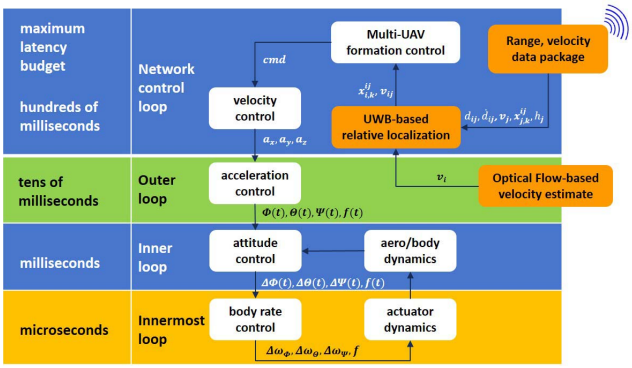


Fig. 9. Diagram of flight control workflow.

a TWToF range conversation with associated data package. On completion of slot 3, the process will repeat from slot 0.

C. Experimental Evaluations of RL Estimation

To verify our proposed RL system, onboard tests were carried out on three simultaneous moving UAV platforms at Internet of Things laboratory as seen in Fig. 1(b), where a VICON system is installed for the ground truth. Each UAV concurrently conducts ranging measurement and communication.

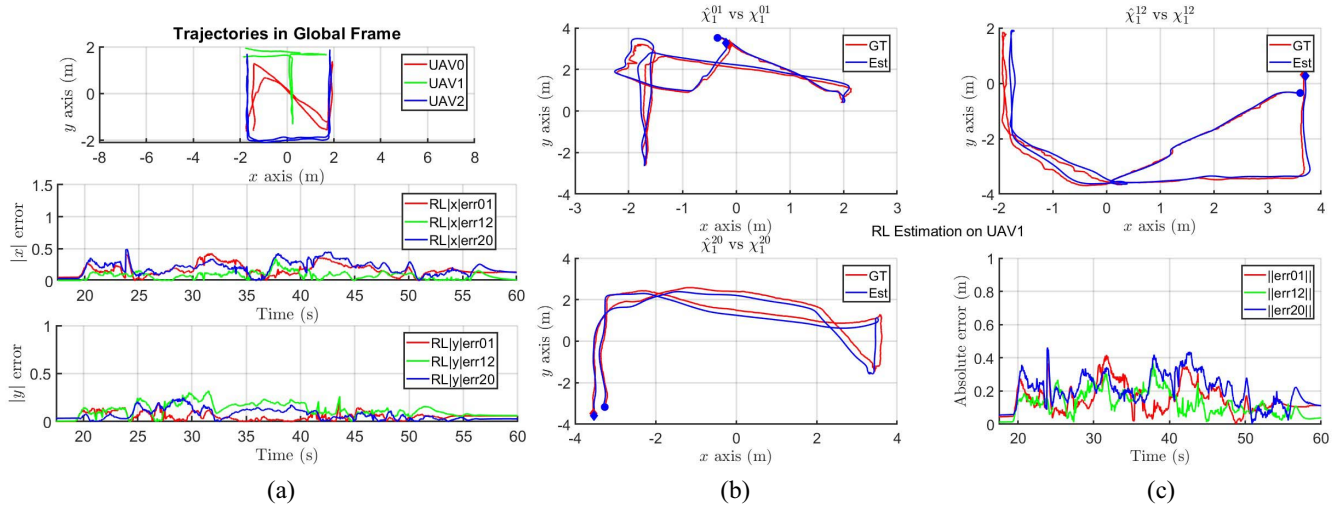


Fig. 10. Distributed RL estimates of Test 13 on each UAV platform. (a) Global trajectories and absolute RL estimation errors in x and y directions. (b) Comparison of RL trajectories calculated on UAV1 and its corresponding 2-norm errors. **GT** represents the RL ground truth obtained from VICON system and **Est** is the estimation curve. The solid red and blue circles are the starting points of ground truth and RL estimates, and the diamonds represent the corresponding ending points.

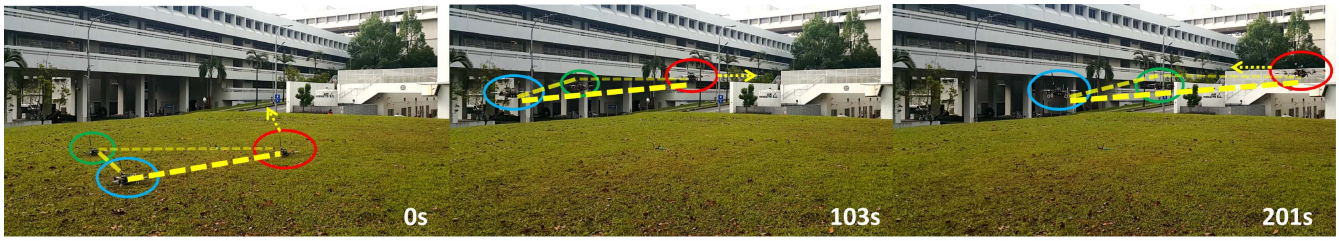


Fig. 11. Video snapshot of three UAVs simultaneously estimating relative positions and flying as a formation. The red circled UAV is a designated leader labeled as 1 and it can be controlled by a pilot to prevent the UAVs from flying toward the crowds. The blue and green circles, labeled as 0 and 2, respectively, are two followers to achieve and maintain a desired formation with the movements of UAV1 (yellow arrow indicates the flight direction of the leader at certain time step). Forward, backward, right, and left movements of the formation are demonstrated. In fact, these three UAVs run the same code and no leader is assigned beforehand. Here, the designated leader is merely for safety considerations, otherwise the formation centroid may be random due to the perturbations and control accuracy.

TABLE III
TDMA SLOTMAP CONFIGURATION FOR THREE UAVS

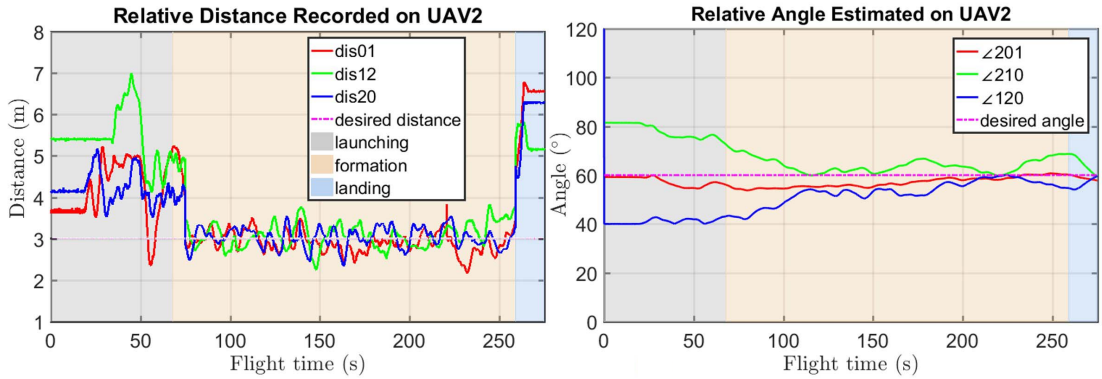
Slot	Requester ID	Responder ID	Data package (16 bytes)
00	200	201-202	time-synchronization message
01	200	201	yes
02	201	202	yes
03	202	200	yes

In order to cover as distinct situations as possible, 13 tests were conducted with different speeds and trajectories, including **S**tatic, **C**ircle, **T**riangle, **L**ine, random **W**alk, and **NTU** letters. Since the RL performances of these three UAVs are similar, without losing generality, here we show the average absolute RL estimation errors of UAV2 as an example in Table IV, where 2**S**1**C**_2 means that two UAVs are static and one UAV moves in a circle in the second time. The statistical results illustrate that the mean of our proposed RL estimation error along one axis is around 0.2 m in the presence of measurement noises.

Fig. 10(a) illustrates the plot of overhead 2-D trajectory recorded on UAV1 from Test no.13. It can be seen that the average absolute RL estimation errors in x and y directions are both less than 0.4 m. Since the RL estimation is conducted distributively, the RL estimation errors generated on each UAV have the same magnitude with slight differences. Without loss of generality, the evolution curves of RL trajectories and estimation errors on UAV1 are shown in Fig. 10(b).

TABLE IV
AVERAGE ABSOLUTE RL ESTIMATION ERRORS $(1/n)\sum_{i=1}^{n}|\hat{x}_i - x_i^{GT}|$ CALCULATED ON UAV2 (UNIT: m)

Test	$\hat{\chi}_2^{01}-\chi_{GT}^{01}$		$\hat{\chi}_2^{12}-\chi_{GT}^{12}$		$\hat{\chi}_2^{20}-\chi_{GT}^{20}$	
	$\lvert\delta x\rvert$	$\lvert\delta y\rvert$	$\lvert\delta x\rvert$	$\lvert\delta y\rvert$	$\lvert\delta x\rvert$	$\lvert\delta y\rvert$
01_2**S**1**C**_1	0.132	0.110	0.141	0.098	0.121	0.111
02_2**S**1**C**_2	0.150	0.101	0.136	0.122	0.141	0.118
03_2**S**1**T**_1	0.245	0.117	0.230	0.155	0.237	0.122
04_2**S**1**T**_2	0.258	0.133	0.249	0.148	0.234	0.161
05_1**S**2**C**_1	0.267	0.120	0.233	0.123	0.244	0.131
06_1**S**2**C**_2	0.296	0.094	0.261	0.103	0.266	0.139
07_1**S**2**T**_1	0.202	0.167	0.215	0.187	0.216	0.151
08_1**S**2**T**_2	0.227	0.115	0.212	0.176	0.249	0.133
09_1**L**2**C**	0.190	0.186	0.197	0.189	0.201	0.182
10_3**W**	0.203	0.223	0.210	0.216	0.219	0.206
11_3**T**	0.217	0.204	0.215	0.216	0.205	0.221
12_3**C**	0.247	0.220	0.218	0.237	0.232	0.226
13_**NTU**	0.222	0.190	0.211	0.208	0.238	0.198
mean	**0.219**	**0.152**	**0.210**	**0.167**	**0.215**	**0.161**



(a) (b)

Fig. 12. (a) Inter-UAV distance measurements evolution on each UAV during flight. (b) Bearings of each UAV relative to its neighbors deriving from RL estimates.

D. Quadcopters Formation Flight

In this section, we implemented the proposed distributed formation control utilizing cooperative RL on quadcopters with the same configuration as described in Section VI-A, and conducted a series of real-world flight tests in a grass field at Nanyang Technological University and an abandoned basketball court near Science Center Singapore. These outdoor experiments also demonstrate the robustness of the systems against wind disturbances. In both these fields, an equilateral triangle formation of three UAVs was achieved successfully.[8]

There are three UAVs involved in the experiments, which comprise of one leader (UAV1) and two followers (UAV0 and 2). The control of these UAVs are set in a global coordinate frame with x-axis align with magnetic East and y-axis with magnetic North. Initially, all three UAVs are placed on the ground, roughly following a predefined configuration for consistency purpose. In particular, we assign UAV0 as a launching origin and place UAV1 along the East direction with respect to UAV0 in order to keep the flight control coordinate and the RL coordinate system consistent. Meanwhile, UAV2 can be put arbitrarily in North zone. Upon power-startup, the UWB modules will measure the distances and start communication, and the Raspberry Pi2 will then derive initialization configuration and start estimating relative positions. With estimated relative positions, control outputs can be calculated accordingly. Second, when onboard control outputs are available for all UAVs, the pilots would then manually takeoff and ascend the UAVs to 1.5 m altitude, with respect to the terrain. This altitude is chosen to minimize ground aerodynamic disturance while maintaining good visual contact for the optical flow velocity estimator. When the target altitude is reached, two pilots would then switch their UAVs from manual flight mode to formation flight mode, where the UAV would automatically control its velocity to satisfy formation requirements. One pilot, controlling the leader UAV1, would then manually fly the UAV in predefined trajectories to test the behavior of follower UAVs. If the follower UAVs could automatically maintain the predefined formation against unknown movements from leader UAV, the experiment is considered successful. A flight video snapshot is shown in Fig. 11, where three UAVs reach the desired formation of an equilateral triangle with side length of 3 m and maintain the formation with the movements of the leader UAV. Online inter-UAV distance measurements and the relative bearing evolutions (generated by RL estimates) in the period of formation forming, are depicted in Fig. 12(a) and (b), where three different background colors, respectively, show the launching, formation keeping, and landing phase. After a period of adjustment, the desired formation is achieved and kept from 56 to 257 s with the inter-UAV distance of 3 m and relative azimuth angle of 60°.

[8]Flight video can be found in: https://youtu.be/E713k93_Vws.

VII. Conclusion

This paper proposed a combined distributed cooperative RL and distance-based formation control scheme for UAV swarms without infrastructures, global positions, and computing for pattern detecting. Based on the capability of the designed UWB RCM network, each UAV is able to estimate the relative positions of its neighbors by utilizing our proposed consensus-based RL fusion scheme. These RL estimates combined with the inter-UAV distances and relative velocity estimates are directly fed into a discrete-time distributed formation control law to achieve formation flights. Extensive flight experiments validated the effectiveness of our proposed system. The cooperative RL and formation control with applications to mobile vehicles in some switching topology case can be investigated in the future.

Appendix
Proof of Theorem 1

A Lyapunov function $J(\tilde{\boldsymbol{\chi}}_{i,k}^{ij}) = \tilde{\boldsymbol{\chi}}_{i,k}^{ij^\top}\tilde{\boldsymbol{\chi}}_{i,k}^{ij}$ is defined, and the difference of this Lyapunov function in m time steps can be written as

$$\begin{aligned}\Delta J_m^k &= J\left(\tilde{\boldsymbol{\chi}}_{i,k+m}^{ij}\right) - J\left(\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\right)\\ &= \tilde{\boldsymbol{\chi}}_{i,k+m}^{ij^\top}\tilde{\boldsymbol{\chi}}_{i,k+m}^{ij} - \tilde{\boldsymbol{\chi}}_{i,k}^{ij^\top}\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\\ &= \sum_{l=k}^{k+m-1}\left(\tilde{\boldsymbol{\chi}}_{i,l+1}^{ij^\top}\tilde{\boldsymbol{\chi}}_{i,l+1}^{ij} - \tilde{\boldsymbol{\chi}}_{i,l}^{ij^\top}\tilde{\boldsymbol{\chi}}_{i,l}^{ij}\right)\end{aligned}$$

$$
\begin{aligned}
&= \sum_{l=k}^{k+m-1} \left(\gamma^2 T^2 \|\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\|^2 - 2\gamma T\right) \\
&\quad \times \tilde{\boldsymbol{\chi}}_{i,l}^{ij^\top} \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)\left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)^\top \tilde{\boldsymbol{\chi}}_{i,l}^{ij} \\
&\quad + 2 \sum_{l=k}^{k+m-1} \boldsymbol{\varepsilon}_{i,l}^\top \boldsymbol{A}_{i,l} \tilde{\boldsymbol{\chi}}_{i,l}^{ij} + \sum_{l=k}^{k+m-1} \|\boldsymbol{\varepsilon}_{i,l}\|^2 \\
&\le \left[\gamma^2 T^2 \left(2\bar{v} + \bar{\delta}\right)^2 - 2\gamma T\right] \\
&\quad \times \left[\sum_{l=k}^{k+m-1} \tilde{\boldsymbol{\chi}}_{i,l}^{ij^\top} \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)\left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)^\top \tilde{\boldsymbol{\chi}}_{i,l}^{ij}\right] \\
&\quad + 2 \sum_{l=k}^{k+m-1} \boldsymbol{\varepsilon}_{i,l}^\top \boldsymbol{A}_{i,l} \tilde{\boldsymbol{\chi}}_{i,l}^{ij} + \sum_{l=k}^{k+m-1} \|\boldsymbol{\varepsilon}_{i,l}\|^2. \qquad (10)
\end{aligned}
$$

In addition, by using induction method on the system (2), RL estimation error at l step is obtained as

$$
\tilde{\boldsymbol{\chi}}_{i,l}^{ij} = \underbrace{\prod_{t=k}^{l-1} \boldsymbol{A}_{i,t}}_{:=\boldsymbol{A}_{i,k:l-1}} \tilde{\boldsymbol{\chi}}_{i,k}^{ij} + \sum_{t=0}^{l-k-1} \left(\prod_{s=k+t+1}^{l-1} \boldsymbol{A}_{i,s}\right) \boldsymbol{\varepsilon}_{i,k+t} \qquad (11)
$$

where $\boldsymbol{A}_{i,k:l-1} \coloneqq \boldsymbol{I}$ when $l = k$. Due to (4), $\|\boldsymbol{A}_{i,l}\| \le 1, \forall l \in \mathbb{Z}^+$. Since $\|\boldsymbol{\varepsilon}_{i,l}\| \le T\bar{\delta}[1 + \gamma(2\bar{v} + \bar{\delta})\bar{d}] = \xi$, then

$$
\left\|\tilde{\boldsymbol{\chi}}_{i,l}^{ij}\right\| \le \left\|\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\right\| + \sum_{t=0}^{l-k-1} \left\|\boldsymbol{\varepsilon}_{i,k+t}\right\| \le \left\|\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\right\| + (l-k)\xi.
$$

Thus,

$$
\begin{aligned}
2 \sum_{l=k}^{k+m-1} \boldsymbol{\varepsilon}_{i,l}^\top \boldsymbol{A}_{i,l} \tilde{\boldsymbol{\chi}}_{i,l}^{ij} &\le 2\xi \sum_{l=k}^{k+m-1} \left[\|\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\| + (l-k)\xi\right] \\
&= 2m\xi \|\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\| + m(m-1)\xi^2.
\end{aligned}
$$

Now, (10) can be written as

$$
\begin{aligned}
\Delta J_m^k \le &-\alpha \sum_{l=k}^{k+m-1} \tilde{\boldsymbol{\chi}}_{i,l}^{ij^\top} \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)\left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)^\top \tilde{\boldsymbol{\chi}}_{i,l}^{ij} \qquad (12a) \\
&+ 2m\xi \|\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\| + m^2\xi^2. \qquad (12b)
\end{aligned}
$$

To separate $\tilde{\boldsymbol{\chi}}_{i,k}^{ij}$ from $\tilde{\boldsymbol{\chi}}_{i,l}^{ij}$, next we expand (12a) individually. According to (11), (12) can be rewritten as

$$
\begin{aligned}
\Delta J_m^k \le &-\alpha \sum_{l=k}^{k+m-1} \tilde{\boldsymbol{\chi}}_{i,k}^{ij^\top} \boldsymbol{A}_{i,k:l-1}^\top \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right) \\
&\times \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)^\top \boldsymbol{A}_{i,k:l-1} \tilde{\boldsymbol{\chi}}_{i,k}^{ij} \\
&- \alpha \sum_{l=k}^{k+m-1} \tilde{\boldsymbol{\chi}}_{i,k}^{ij^\top} \boldsymbol{A}_{i,k:l-1}^\top \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)\left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)^\top \\
&\times \left(\sum_{t=0}^{l-k-1} \boldsymbol{A}_{i,k+t+1:l-1} \boldsymbol{\varepsilon}_{i,k+t}\right) \\
&- \alpha \sum_{l=k}^{k+m-1} \left(\sum_{t=0}^{l-k-1} \boldsymbol{A}_{i,k+t+1:l-1} \boldsymbol{\varepsilon}_{i,k+t}\right)^\top \\
&\times \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)\left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)^\top \boldsymbol{A}_{i,k:l-1} \tilde{\boldsymbol{\chi}}_{i,k}^{ij} \\
&- \alpha \sum_{l=k}^{k+m-1} \left(\sum_{t=0}^{l-k-1} \boldsymbol{A}_{i,k+t+1:l-1} \boldsymbol{\varepsilon}_{i,k+t}\right)^\top \\
&\times \left(\sum_{t=0}^{l-k-1} \boldsymbol{A}_{i,k+t+1:l-1} \boldsymbol{\varepsilon}_{i,k+t}\right) \\
&+ 2m\xi \|\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\| + m^2\xi^2 \\
\le &-\alpha \sum_{l=k}^{k+m-1} \tilde{\boldsymbol{\chi}}_{i,k}^{ij^\top} \boldsymbol{A}_{i,k:l-1}^\top \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right) \\
&\times \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)^\top \boldsymbol{A}_{i,k:l-1} \tilde{\boldsymbol{\chi}}_{i,k}^{ij} \\
&+ 2\alpha \sum_{l=k}^{k+m-1} \left[(l-k)\xi(2\bar{v} + \bar{\delta})^2 \|\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\|\right] \\
&+ \alpha \sum_{l=k}^{k+m-1} \left[(l-k)^2\xi^2\right] + 2m\xi \|\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\| + m^2\xi^2 \\
\le &-\alpha \sum_{l=k}^{k+m-1} \tilde{\boldsymbol{\chi}}_{i,k}^{ij^\top} \boldsymbol{A}_{i,k:l-1}^\top \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right) \\
&\times \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)^\top \boldsymbol{A}_{i,k:l-1} \tilde{\boldsymbol{\chi}}_{i,k}^{ij} \qquad (13a) \\
&+ \left[2m\xi + \alpha\xi m(m-1)(2\bar{v} + \bar{\delta})^2\right] \|\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\| \qquad (13b) \\
&+ \left[m^2 + \frac{1}{6}\alpha m(m-1)(2m-1)\right]\xi^2. \qquad (13c)
\end{aligned}
$$

To proceed, $\boldsymbol{A}_{i,k:l-1}$ should be extracted from the summation in (13a), and thus we need to prove the following inequality first:

$$
\begin{aligned}
&\boldsymbol{A}_{i,k:l-1}^\top \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)\left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)^\top \boldsymbol{A}_{i,k:l-1} \\
&\quad \le \boldsymbol{A}_{i,k:l-2}^\top \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)\left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)^\top \boldsymbol{A}_{i,k:l-2} \qquad (14)
\end{aligned}
$$

$\forall l = k, k+1, \ldots, k+m-1$, where $\boldsymbol{A}_{i,k:s} \coloneqq \boldsymbol{I}$ if $s < k$.

The following is the proof of (14). $\forall \boldsymbol{x} \in \mathbb{R}^2$:

$$
\begin{aligned}
&\boldsymbol{x}^\top \boldsymbol{A}_{i,k:l-1}^\top \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)\left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)^\top \boldsymbol{A}_{i,k:l-1} \boldsymbol{x} \\
&= \underbrace{\boldsymbol{x}^\top \boldsymbol{A}_{i,k:l-2}^\top}_{:=\boldsymbol{Z}^\top} \boldsymbol{A}_{i,l-1}^\top \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)\left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)^\top \boldsymbol{A}_{i,l-1} \underbrace{\boldsymbol{A}_{i,k:l-2} \boldsymbol{x}}_{:=\boldsymbol{Z}} \\
&= \boldsymbol{Z}^\top \boldsymbol{A}_{i,l-1}^\top \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)\left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)^\top \boldsymbol{A}_{i,l-1} \boldsymbol{Z} \\
&= \boldsymbol{Z}^\top \left[\boldsymbol{I} - \gamma T (\boldsymbol{v}_{i,l-1}^{ij} + \boldsymbol{\epsilon}_{l-1})(\boldsymbol{v}_{i,l-1}^{ij} + \boldsymbol{\epsilon}_{l-1})^\top\right]^\top \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right) \\
&\quad \times \left(\boldsymbol{v}_{i,l}^{ij} + \boldsymbol{\epsilon}_l\right)^\top \left[\boldsymbol{I} - \gamma T \underbrace{(\boldsymbol{v}_{i,l-1}^{ij} + \boldsymbol{\epsilon}_{l-1})(\boldsymbol{v}_{i,l-1}^{ij} + \boldsymbol{\epsilon}_{l-1})^\top}_{:=\boldsymbol{B}_{i,l-1}}\right] \boldsymbol{Z}. \qquad (15)
\end{aligned}
$$

Before proceeding, we denote $\|\boldsymbol{v}_{i,l-1}^{ij} + \boldsymbol{\epsilon}_{l-1}\|^2$ by $\lambda_{i,l-1}$, the *zerospace* of $\boldsymbol{B}_{i,l-1}$ by

$$
\mathrm{Null}\left(\boldsymbol{B}_{i,l-1}\right) = \left\{\boldsymbol{v} \in \mathbb{R}^2 | \boldsymbol{B}_{i,l-1} \boldsymbol{v} = \boldsymbol{0}\right\}
$$

and the *eigenspace* of $\lambda_{i,l-1}$ by

$$
E_{\lambda_{i,l-1}}^{(\boldsymbol{B}_{i,l-1})} = \left\{\boldsymbol{u} \in \mathbb{R}^2 | \boldsymbol{B}_{i,l-1} \boldsymbol{u} = \lambda_{i,l-1} \boldsymbol{u}\right\}.
$$

It is easy to see that $||\boldsymbol{v}_{i,l-1}^{ij}+\boldsymbol{\epsilon}_{l-1}||^2$ is the only nonzero eigenvalue of $\boldsymbol{B}_{i,l-1}$ with eigenvector $\boldsymbol{v}_{i,l-1}^{ij}+\boldsymbol{\epsilon}_{l-1}$, which is simple. Therefore, $\lambda_{i,l-1}=||\boldsymbol{v}_{i,l-1}^{ij}+\boldsymbol{\epsilon}_{l-1}||^2$.

If $\boldsymbol{Z}\in \text{Null}(\boldsymbol{B}_{i,l-1})$

$$\begin{aligned}&\boldsymbol{x}^\top \boldsymbol{A}_{i,k:l-1}^\top\Big(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l\Big)\Big(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l\Big)^\top \boldsymbol{A}_{i,k:l-1}\boldsymbol{x}\\&\quad=\boldsymbol{Z}^\top\Big(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l\Big)\Big(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l\Big)^\top\boldsymbol{Z}.\end{aligned}\tag{16}$$

If $\boldsymbol{Z}\in E_{\lambda_{i,l-1}}^{(\boldsymbol{B}_{i,l-1})}$, without loss of generality, we choose $\boldsymbol{Z}=\iota(\boldsymbol{v}_{i,l-1}^{ij}+\boldsymbol{\epsilon}_{l-1})$, $\iota\in\mathbb{R}$, then

$$\begin{aligned}&\boldsymbol{x}^\top \boldsymbol{A}_{i,k:l-1}^\top\Big(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l\Big)\Big(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l\Big)^\top \boldsymbol{A}_{i,k:l-1}\boldsymbol{x}\\&\quad=\Big(1-\gamma T\|\boldsymbol{v}_{i,l-1}^{ij}+\boldsymbol{\epsilon}_{l-1}\|^2\Big)^2\boldsymbol{Z}^\top\Big(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l\Big)\Big(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l\Big)^\top\boldsymbol{Z}\\&\quad<\boldsymbol{Z}^\top\Big(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l\Big)\Big(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l\Big)^\top\boldsymbol{Z}.\end{aligned}\tag{17}$$

It is known that for any $\boldsymbol{Z}\in\mathbb{R}^2$, it has $\boldsymbol{Z}=\boldsymbol{Z}_1+\boldsymbol{Z}_2$ for some $\boldsymbol{Z}_1\in\text{Null}(\boldsymbol{B}_{i,l-1})$, $\boldsymbol{Z}_2\in E_{\lambda_{i,l-1}}^{(\boldsymbol{B}_{i,l-1})}$, together with (16) and (17), which implies (14).

According to (14), together with (13), we can get

$$\begin{aligned}\Delta J_m^k&\le-\alpha\sum_{l=k}^{k+m-1}\tilde{\boldsymbol{\chi}}_{i,k}^{ij\top}\boldsymbol{A}_{i,k:k+m-2}^\top\Big(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l\Big)\Big(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l\Big)^\top\\&\quad\times\boldsymbol{A}_{i,k:k+m-2}\tilde{\boldsymbol{\chi}}_{i,k}^{ij}+\xi_1\|\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\|+\xi_2\\&=-\alpha\tilde{\boldsymbol{\chi}}_{i,k}^{ij\top}\boldsymbol{A}_{i,k:k+m-2}^\top\left[\sum_{l=k}^{k+m-1}\Big(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l\Big)\Big(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l\Big)^\top\right]\\&\quad\times\boldsymbol{A}_{i,k:k+m-2}\tilde{\boldsymbol{\chi}}_{i,k}^{ij}+\xi_1\|\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\|+\xi_2\\&\le-\mu\alpha\tilde{\boldsymbol{\chi}}_{i,k}^{ij\top}\boldsymbol{A}_{i,k:k+m-2}^\top\boldsymbol{A}_{i,k:k+m-2}\tilde{\boldsymbol{\chi}}_{i,k}^{ij}+\xi_1\|\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\|+\xi_2.\end{aligned}\tag{18}$$

Now, we need to prove

$$\boldsymbol{A}_{i,k:k+m-2}^\top\boldsymbol{A}_{i,k:k+m-2}\ge\Big[1-\gamma T\big(2\bar{v}+\bar{\delta}\big)^2\Big]^{2(m-1)}\boldsymbol{I}.\tag{19}$$

The following is the proof of (19).

For $\boldsymbol{A}_{i,l}^\top\boldsymbol{A}_{i,l}=[\boldsymbol{I}-\gamma T(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l)(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l)^\top]^\top[\boldsymbol{I}-\gamma T(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l)(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l)^\top]$, denote $\boldsymbol{B}_{i,l}:=(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l)(\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l)^\top$ and $\lambda_{i,l}:=\|\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l\|^2$. Similar with the proof of (14), $\forall\boldsymbol{x}\in\mathbb{R}^2$

$$\begin{aligned}&\boldsymbol{x}^\top\boldsymbol{A}_{i,l}^\top\boldsymbol{A}_{i,l}\boldsymbol{x}\\&\quad=\begin{cases}\boldsymbol{x}^\top\boldsymbol{x} & \text{if } \boldsymbol{x}\in\text{Null}\big(\boldsymbol{B}_{i,l}\big)\\\Big(1-\gamma T\|\boldsymbol{v}_{i,l}^{ij}+\boldsymbol{\epsilon}_l\|^2\Big)^2\boldsymbol{x}^\top\boldsymbol{x} & \text{if } \boldsymbol{x}\in E_{\lambda_{i,l}}^{(\boldsymbol{B}_{i,l})}\end{cases}\end{aligned}\tag{20}$$

thus, $\boldsymbol{A}_{i,l}^\top\boldsymbol{A}_{i,l}\ge[1-\gamma T(2\bar{v}+\bar{\delta})^2]^2\boldsymbol{I}$. According to (4), (19) holds.

So far the RL estimation error bound can be presented as follows. Combining (18) with (19), we have

$$\begin{aligned}\Delta J_m^k&\le\mu\alpha\Big[\gamma T\big(2\bar{v}+\bar{\delta}\big)^2-1\Big]^{2(m-1)}\|\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\|^2+\xi_1\|\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\|+\xi_2\\&=-\sigma J\Big(\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\Big)+\xi_1\sqrt{J\Big(\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\Big)}+\xi_2\end{aligned}$$

where $\sigma>0$, $\xi_1>0$, and $\xi_2>0$, therefore, $\sqrt{J(\tilde{\boldsymbol{\chi}}_{i,k}^{ij})}$ is bounded by $[(\xi_1+\sqrt{\xi_1^2+4\sigma\xi_2})/2\sigma]$ and $[(\xi_1-\sqrt{\xi_1^2+4\sigma\xi_2})/2\sigma]$ in the right and left half plane, respectively. Since $|[(\xi_1+\sqrt{\xi_1^2+4\sigma\xi_2})/2\sigma]|>|[(\xi_1-\sqrt{\xi_1^2+4\sigma\xi_2})/2\sigma]|$ and $[(\xi_1+\sqrt{\xi_1^2+4\sigma\xi_2})/2\sigma]>0$, then we can obtain that $\|\tilde{\boldsymbol{\chi}}_{i,k}^{ij}\|\le c$, where $c=[(\xi_1+\sqrt{\xi_1^2+4\sigma\xi_2})/2\sigma]$, and the proof is completed.

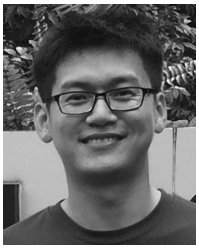
ACKNOWLEDGMENT

The authors would like to thank M.-C. Hoang, T. M. Nguyen, R. X. Foo, M. Cao, T. Ji, and H. Wang for their assistance in numerous experiments. They would also like to thank S. Liu, D. Han, and L. Xu for helpful discussions.

REFERENCES

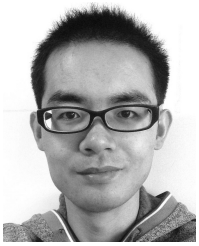
[1] K.-K. Oh, M.-C. Park, and H.-S. Ahn, "A survey of multi-agent formation control," *Automatica*, vol. 53, pp. 424–440, Mar. 2015.
[2] Z. Han, L. Wang, Z. Lin, and R. Zheng, "Formation control with size scaling via a complex Laplacian-based approach," *IEEE Trans. Cybern.*, vol. 46, no. 10, pp. 2348–2359, Oct. 2016.
[3] J. A. Preiss, W. Honig, G. S. Sukhatme, and N. Ayanian, "Crazyswarm: A large nano-quadcopter swarm," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2017, pp. 3299–3304.
[4] H. Du, W. Zhu, G. Wen, Z. Duan, and J. Lü, "Distributed formation control of multiple quadrotor aircraft based on nonsmooth consensus algorithms," *IEEE Trans. Cybern.*, vol. 49, no. 1, pp. 342–353, Jan. 2019. doi: 10.1109/TCYB.2017.2777463.
[5] D. Li, W. Zhang, W. He, C. Li, and S. S. Ge, "Two-layer distributed formation-containment control of multiple Euler–Lagrange systems by output feedback," *IEEE Trans. Cybern.*, vol. 49, no. 2, pp. 675–687, Feb. 2019. doi: 10.1109/TCYB.2017.2786318.
[6] T. Han, Z. Lin, R. Zheng, and M. Fu, "A barycentric coordinate-based approach to formation control under directed and switching sensing graphs," *IEEE Trans. Cybern.*, vol. 48, no. 4, pp. 1202–1215, Apr. 2018.
[7] X. Li and L. Xie, "Dynamic formation control over directed networks using graphical Laplacian approach," *IEEE Trans. Autom. Control*, vol. 63, no. 11, pp. 3761–3774, Nov. 2018. doi: 10.1109/TAC.2018.2798808.
[8] S.-M. Kang, M.-C. Park, and H.-S. Ahn, "Distance-based cycle-free persistent formation: Global convergence and experimental test with a group of quadcopters," *IEEE Trans. Ind. Electron.*, vol. 64, no. 1, pp. 380–389, Jan. 2017.
[9] X. Dong, Y. Zhou, Z. Ren, and Y. Zhong, "Time-varying formation tracking for second-order multi-agent systems subjected to switching topologies with application to quadrotor formation flying," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 5014–5024, Jun. 2017.
[10] W. Jasim and D. Gu, "Robust team formation control for quadrotors," *IEEE Trans. Control Syst. Technol.*, vol. 26, no, 4, pp. 1516–1523, Jul. 2018.
[11] T. M. Nguyen, A. H. Zaini, K. Guo, and L. Xie, "An ultra-wideband-based multi-UAV localization system in GPS-denied environments," in *Proc. Int. Micro Air Veh. Conf. Competitions (IMAV)*, Beijing, China, Oct. 2016, pp. 56–61.
[12] M. Saska *et al.*, "System for deployment of groups of unmanned micro aerial vehicles in GPS-denied environments using onboard visual relative localization," *Auton. Robots*, vol. 41, no. 4, pp. 919–944, 2017.
[13] D. Scaramuzza *et al.*, "Vision-controlled micro flying robots: From system design to autonomous navigation and mapping in GPS-denied environments," *IEEE Robot. Autom. Mag.*, vol. 21, no. 3, pp. 26–40, Sep. 2014.
[14] M. Coppola, K. N. McGuire, K. Y. Scheper, and G. C. de Croon, "On-board communication-based relative localization for collision avoidance in micro air vehicle teams," *Auton. Robots*, vol. 42, no. 8, pp. 1787–1805, 2018
[15] S. Wang, D. Gu, L. Chen, and H. Hu, "Single beacon-based localization with constraints and unknown initial poses," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 2229–2241, Apr. 2016.

[16] G. Chai, C. Lin, Z. Lin, and W. Zhang, "Consensus-based cooperative source localization of multi-agent systems with sampled range measurements," *Unmanned Syst.*, vol. 2, no. 3, pp. 231–241, 2014.
[17] Z. Han, K. Guo, L. Xie, and Z. Lin, "Integrated relative localization and leader–follower formation control," *IEEE Trans. Autom. Control*, vol. 64, no. 1, pp. 20–34, Jan. 2019. doi: 10.1109/TAC.2018.2800790.
[18] B. Jiang, M. Deghat, and B. D. O. Anderson, "Simultaneous velocity and position estimation via distance-only measurements with application to multi-agent system control," *IEEE Trans. Autom. Control*, vol. 62, no. 2, pp. 869–875, Feb. 2017.
[19] K. Guo *et al.*, "Ultra-wideband-based localization for quadcopter navigation," *Unmanned Syst.*, vol. 4, no. 1, pp. 23–34, 2016.
[20] K. Guo, Z. Qiu, W. Meng, L. Xie, and R. Teo, "Ultra-wideband based cooperative relative localization algorithm and experiments for multiple UAVs in GPS-denied environments," *Int. J. Micro Air Veh.*, vol. 9, no. 3, pp. 169–186, 2017.
[21] K. McGuire, G. de Croon, C. De Wagter, K. Tuyls, and H. Kappen, "Efficient optical flow and stereo vision for velocity estimation and obstacle avoidance on an autonomous pocket drone," *IEEE Robot. Autom. Lett.*, vol. 2, no. 2, pp. 1070–1076, Apr. 2017.
[22] Y. Zhang, T. Wang, Z. Cai, Y. Wang, and Z. You, "The use of optical flow for UAV motion estimation in indoor environment," in *Proc. IEEE Chin. Guid. Navig. Control Conf. (CGNCC)*, Aug. 2016, pp. 785–790.
[23] Z.-P. Jiang and Y. Wang, "Input-to-state stability for discrete-time nonlinear systems," *Automatica*, vol. 37, no. 6, pp. 857–869, 2001.
[24] J. Hu and Y. Hong, "Leader-following coordination of multi-agent systems with coupling time delays," *Physica A Stat. Mech. Appl.*, vol. 374, no. 2, pp. 853–863, 2007.
[25] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
[26] M. Deghat, B. D. O. Anderson, and Z. Lin, "Combined flocking and distance-based shape control of multi-agent formations," *IEEE Trans. Autom. Control*, vol. 61, no. 7, pp. 1824–1837, Jul. 2016.



Kexin Guo received the B.S. degree in control theory and control engineering from the Beijing Institute of Technology, Beijing, China, in 2011, the M.S. degree in instrumentation science and technology from Beihang University, Beijing, in 2014, and the Ph.D. degree from the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, in 2018.

Since 2019, he has been a Post-Doctoral Fellow of the "Zhuoyue" program with the School of Automation Science and Electrical Engineering, Beihang University. His current research interests include global positioning system (GPS)-denied localization, formation control, and unmanned systems.



Xiuxian Li (M'16) received the B.S. degree in mathematics and applied mathematics and the M.S. degree in pure mathematics from Shandong University, Jinan, China, in 2009 and 2012, respectively, and the Ph.D. degree in mechanical engineering from the University of Hong Kong, Hong Kong, in 2016.

Since 2016, he has been a Research Fellow with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, and a Senior Research Associate with the Department of Biomedical Engineering, City University of Hong Kong, Hong Kong. His current research interests include distributed optimization, distributed control, formation control, and multiagent networks.

Lihua Xie (F'07) received the B.E. and M.E. degrees in electrical engineering from the Nanjing University of Science and Technology, Nanjing, China, in 1983 and 1986, respectively, and the Ph.D. degree in electrical engineering from the University of Newcastle, Callaghan, NSW, Australia, in 1992.

Since 1992, he has been with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, where he is currently a Professor and the Director of the Delta-NTU Corporate Laboratory for Cyber-Physical Systems. He served as the Head of the Division of Control and Instrumentation from 2011 to 2014. He held teaching appointments with the Department of Automatic Control, Nanjing University of Science and Technology, from 1986 to 1989. His current research interests include robust control and estimation, networked control systems, multiagent networks, localization, and unmanned systems.

Dr. Xie is the Editor-in-Chief of *Unmanned Systems* and an Associate Editor of the IEEE TRANSACTIONS ON NETWORK CONTROL SYSTEMS. He has served as an Editor for IET Book Series in Control and an Associate Editor for a number of journals, including the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, *Automatica*, the IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY, and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—PART II: EXPRESS BRIEFS. He is an Elected Member of the Board of Governors, IEEE Control System Society, from 2016 to 2018. He is a fellow of IFAC.