2024 IEEE International Conference on Robotics and Automation (ICRA 2024)
May 13-17, 2024. Yokohama, Japan

Tight Fusion of Odometry and Kinematic Constraints for Multiple Aerial Vehicles in Physical Interconnection

Yingjun Fan, Chuanbeibei Shi, Ganghua Lai, Ruiheng Zhang, Yushu Yu, Fuchun Sun, Yiqun Dong

Abstract— **Integrated aerial Platforms (IAPs), comprising multiple aircrafts, are typically fully actuated and hold significant potential for aerial manipulation tasks. Differing from a multiple aerial swarm, the aircrafts within the IAP are interconnected, presenting promising opportunities for enhancing localization. Incorporating the physical constraints of these multiple aircrafts to improve the accuracy and reliability of integrated aircraft positioning and navigation systems is a challenging yet highly significant problem. In this paper, we introduce a distributed multi-aircraft visual-inertial-range odometry system that analyzes the position, velocity, and attitude constraints within the IAP. Leveraging constraint relationships in the IAP, we propose corresponding methods that tightly fuse visual-inertial-range odometry and kinematic constraints to optimize odometry accuracy. Our system's performance is validated using a collected dataset, resulting in a notable 28.7% reduction in drift compared to the baseline.**

I. Introduction

In recent years, the aircraft has attracted the attention of numerous researchers because of its merits in maneuverability [1] - [2], [21] - [25]. It is also noticed that the under-actuation property in the traditional aircraft has prevented it from becoming effective manipulation platforms. Yet, the integrated aerial platform (IAP) composed of multiple aerial platforms is usually fully-actuated, While the IAP system has the potential to advance the payload of by increasing the number of aircrafts involved [3], Because of the above advantage, the IAP is a ideal platform for manipulation and observation [26] - [33].

To achieve the functionality of multiple aircrafts, obtaining precise poses for these vehicles is of paramount importance. The necessity for streamlined and precise positioning and navigation of multiple aircrafts holds significant relevance for control and planning algorithms.

There has been substantial research on the localization problem of multiple aircrafts, e.g., fusing the odometry, LiDAR, and UWB, and so on [4] - [8]. These methods aim to optimize the multi-aircraft SLAM technology by leveraging mutual observations between the aircraft and integrating Ultra-Wideband (UWB) distance measurements as additional constraints. By incorporating these constraints, it was seen that the positioning accuracy of the multiple aircraft can be improved.

However, differing from a multiple aerial swarm [9], [10], the multiple aircrafts within the IAP are interconnected. These interconnections hold the potential for utilization in the context of localization. The pose of one vehicle is interrelated with the pose of its connected counterpart. This information can be leveraged to amalgamate the odometry data from multiple aircrafts within the IAP, potentially enhancing localization accuracy. This becomes particularly crucial for the multiple interconnected aircraft in the IAP to collaborate effectively in unfamiliar environments. The incorporation of physical constraints into the onboard positioning and navigation systems of the IAP presents a novel challenge that has not been extensively explored in previous research. Hence, it represents a challenging yet highly meaningful problem to address.

This paper introduces a novel approach to multi-aircraft odometry technology through the application of physical constraints. Specifically, a odometry fusion method is presented for IAP, which capitalizes on prior knowledge of distance and attitude velocity constraints. By analyzing the kinematic constraints within the IAP, the paper establishes a distributed optimization problem that tightly integrates odometry with these kinematic constraints. Since the IAP's kinematic constraints solely comprise relative pose information, to further enhance localization accuracy, the distributed fusion is combined with a UWB anchor range data fusion module. The ranging data from anchors can provide absolute information to the odometry, suppressing odometry drift. To experimental validate the proposed approach, dataset on two types of IAPs are constructed. On the self-constructed dataset, it is demonstrated that this combination significantly improves the localization accuracy.

The main contributions of this article are,

- We propose a theoretical framework to analyze the position, velocity, and attitude constraints induced by the multiple aircrafts in the IAP, and propose the corresponding method to optimize the odometry accuracy. Our proposed approach improves the localization quality of the IAP, and we refer to this framework as Visual-Inertial-Range-Physical Odometry(VIRPO).
- The effectiveness of the proposed fusion algorithm has been verified on our self-collected dataset.

*This work was supported by the National Natural Science Foundation of China under Grant 62173037, National Key R. D. Program of China, and State Key Laboratory of Robotics and Systems (HIT). (Corresponding author: Yushu Yu)

Yingjun Fan, Chuanbeibei Shi, Ganghua Lai, Ruiheng Zhang and Yushu Yu are with the School of Mechatronical Engineering, Beijing Institute of Technology, Beijing 100081, China yushu.yu@bit.edu.cn

Fuchun Sun is with the Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China fcsun@mail.tsinghua.edu.cn

Yiqun Dong is with the Department of Aeronautics and Astronautics, Fudan University, Shanghai 200433, China yiqundong@fudan.edu.cn

II. System Overview

A. Variable Symbol

N represents the total count of aircrafts in the Integrated Aerial Platform (IAP), denoted as the set $\{ID_n\}$, where $n \in \{0, 1, 2, \ldots, N\}$. To ensure clarity, we define $\{ID_0\}$ as the central platform of the IAP.

$\{L_n\}$ represents the local odometry frame of $\{ID_n\}$, with its z-axis aligned to the direction of gravity.

$\{B_n\}$ and $\{C_n\}$ denote the IMU body frame and camera frame of $\{ID_n\}$, respectively.

$\{W\}$ represents the world coordinate system. The computation of the transformation matrix between $\{W\}$ and $\{L\}$ relies on actual data from the motion capture system. ${}^{W}_{L_i}\mathbf{T}_{rot}$ and ${}^{W}_{L_i}\mathbf{T}_{tran}$ refer to the rotation and translation components of the transformation matrix, mapping the local system $\{L_i\}$ of the i-th sub-aircraft to the world system $\{W\}$.

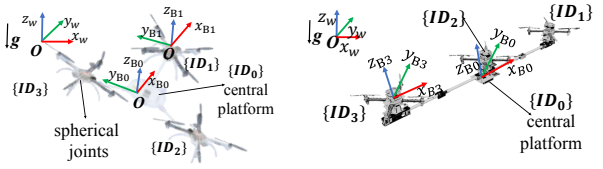
$(\cdot)_{xyz}$ represents the xyz component of the vector $(\cdot)$.

$[\mathbf{q}]_{xyz}$ represents the Euler angle vector from a quaternion $\mathbf{q}$.

$\mathbf{S}_{\mathbf{x}_A}, \mathbf{S}_{\mathbf{y}_A}, \mathbf{S}_{\mathbf{z}_A}$ signify the orthogonal basis of the coordinate system $\{A\}$ within the world system.

B. Mechanical Structure and Dataset Collection

We conducted data collection using two types of integrated aircraft platforms: a star-shaped Integrated Aircraft Platform (SIAP) and a line-integrated aircraft platform (LIAP). Both of them incorporate a central platform. IMU and PX4 flight control were deployed on both the central platform and each sub-aircraft, while ZED cameras were attached to each sub-aircraft. A spherical joint connects the sub-aircraft to the central platform, resulting in relative motion between the sub-aircraft and the IAP. The mechanical structure diagram of the IAP is depicted in Fig. 1a and Fig. 1b. Data acquisition took place within an Indoor Infrared Motion Capture System, providing centimeter-level positioning results for the IAP through surrounding infrared cameras.During the experiments, the virtual simulation will utilize ground-based UWB anchors, with their positions based on the ground truth values from the motion capture system. The algorithm will then utilize simulated distance measurements between the aircraft and these virtual anchors.



(a) (b)

Fig. 1: (a) and (b) represent the mechanical structures of the SIAP and LIAP.

III. Methodologies

A. Tightly Coupled Multi VIRPO

The VIRPO algorithm, as depicted in Fig. 2, operates on each sub-aircraft, utilizing onboard cameras, IMUs, UWB sensor data, and data from the IAP. It integrates UWB ranging data and incorporates the IAP's physical constraints into VIRPO to enhance localization accuracy. The improved results are shared among sub-aircraft. This algorithm is built upon the open-source VINS-Fusion code [11], incorporating UWB and physical structure constraints into the localization process. The exchanged data, which encompasses pose, velocity, and angular velocity, undergoes Bundle Adjustment (BA) optimization. This optimization encompasses position, velocity, angular velocity, and UWB residuals within the VIRPO framework, as illustrated in Figure 3.

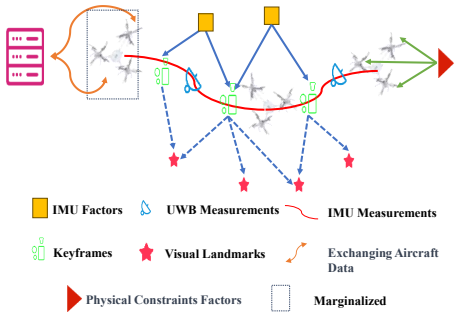


Fig. 3: In the tightly coupled formulation of VIRPO, the sliding window integrates information from IMUs, visual inputs, physical constraints, and UWB measurements.

The state vector of the i-th sub aircraft is defined as

$$\boldsymbol{\chi}^i = [\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2 ... \mathbf{x}_K, \mathbf{x}_c^b, \lambda_0, \lambda_1 .... \lambda_M]$$

$$\mathbf{x}_k = [\mathbf{p}_{b_k}, \mathbf{v}_{b_k}, \mathbf{q}_{b_k}, \mathbf{ba}, \mathbf{bg}]$$

where $\mathbf{x}_k^i$ is the IMU state at the time when the k-th image of i-th sub-aircraft is captured. It contains position, velocity, and orientation of the IMU in the i-th aircraft's local frame $\{L_i\}$, acceleration bias $\mathbf{b_a}$ and gyroscope bias $\mathbf{b_g}$ in the IMU body frame. K is the total number of keyframes. λ_m^i is the inverse distance of the m-th feature of i-th sub-aircraft from its first observation. A visual-inertial-UWB-physical constraint bundle adjustment is performed to estimate the states. The formulation of i-th sub-aircraft is

$$\begin{aligned}\min_{\boldsymbol{\chi}^i} \Big\{ & ||\mathbf{r}_p - \mathbf{H}_p \boldsymbol{\chi}^i||^2 + \sum_{k \in \mathfrak{B}} ||\mathbf{r}_{inertial}(\hat{\mathbf{z}}_{b_{k+1}}^{b_k}, \boldsymbol{\chi}^i)||^2_{\mathbf{P}_{b_{k+1}}^{b_k}} \\ & + \sum_{(m,k) \in \mathfrak{C}} ||\mathbf{r}_{visual}(\hat{\mathbf{z}}_m^{c_k}, \boldsymbol{\chi}^i)||^2_{\mathbf{P}_m^{c_k}} \\ & + ||\mathbf{r}_{kp}(\boldsymbol{\chi}^i, \mathcal{F})||^2_{\sigma_l^{-2}} + ||\mathbf{r}_{kv}(\boldsymbol{\chi}^i, \mathcal{F})||^2_{\sigma_\omega^{-2}} \\ & + ||\mathbf{r}_{kq}(\boldsymbol{\chi}^i, \mathcal{F})||^2_{\sigma_{att}^{-2}} + ||\mathbf{r}_{uwb}(\boldsymbol{\chi}^i, \mathcal{J})||^2_{\sigma_d^{-2}} \\ & + \sum_{o \neq i} ||\mathbf{r}_{drift}(\boldsymbol{\chi}^o)||_{\sigma_{drift}^{-2}} \Big\}\end{aligned} \tag{1}$$

where $\mathbf{r}_{inertial}(\hat{\mathbf{z}}_{b_{k+1}}^{b_k}, \boldsymbol{\chi}^i)$ and $\mathbf{r}_{visual}(\hat{\mathbf{z}}_m^{c_j}, \boldsymbol{\chi}^i)$ are residuals for IMU and visual measurements, respectively, $\mathfrak{B}$ is the set of all IMU measurements, $\mathfrak{C}$ is the set of features that have been observed at least twice in the current sliding

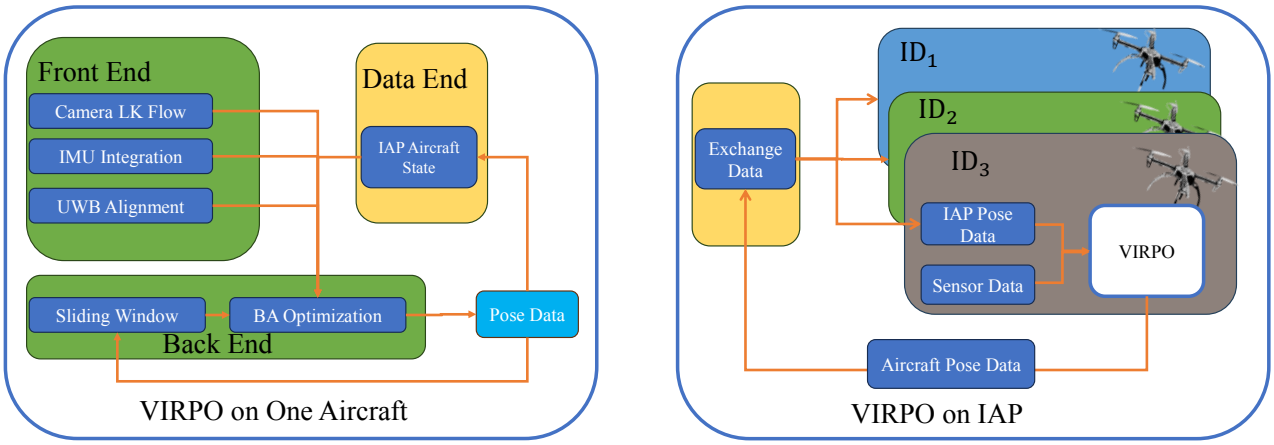


Fig. 2: In the IAP, each sub-aircraft utilizes data from cameras, IMUs, UWB sensors, and odometer data from the IAP in BA optimization to execute the VIRPO algorithm. Additionally, it shares its own odometer data with other sub-aircraft.

window, $\mathbf{r}_p\,\mathbf{H}_p$ are the prior information from marginalization, $\mathbf{r}_{kp}(\boldsymbol{\chi}^i)$, $\mathbf{r}_{kv}(\boldsymbol{\chi}^i)$, $\mathbf{r}_{kq}(\boldsymbol{\chi}^i)$ are residual for position, angular, velocity and attitude constraints between sub aircraft, $\mathbf{r}_{uwb}(\boldsymbol{\chi}^i)$ is residual for UWB measurements. $\mathcal{F}, \mathcal{J}$ represents the set of pose data obtained for the other sub-aircrafts in the IAP (including the central platform) and the corresponding set of UWB range data. σ_l, σ_v, σ_{att}, and σ_d represent the covariance of the position, velocity, attitude, and range constraints, respectively. In light of the observed accumulated drift errors in other aircraft within the IAP, we incorporate residual drift error terms $||\mathbf{r}_{drift}(\boldsymbol{\chi}^o)||$ from these aircraft into our tightly coupled optimization. This inclusion aims to enhance the robustness of our system and mitigate the influence of outliers. The detailed information of the residual terms will be presented in Sections III-C to III-E. The Ceres solver is used for solving this nonlinear problem in our implementation.

B. Visual and Inertial Residual

1) IMU Preintegration and IMU Measurement Residual: A IMU can measure the acceleration $\mathbf{a}$ and the angular velocity $\boldsymbol{\omega}$ of the body frame with respect to the inertial frame. However, the raw measurements $\hat{\boldsymbol{\omega}}$ and $\hat{\mathbf{a}}$ are affected by bias and white noise [12]. We adopt a continuous-time quaternion-based IMU preintegration method, which includes handling IMU biases and employs the kinematics described in [5]. This approach avoids the need for re-propagation when the body state changes, and the pre-integrated measurements serve as constraint factors between successive keyframes. When IMU bias changes with a small increments [13], instead of computing pre-integrated measurements iteratively, use a first-order approximation to update [14]. The residual of IMU preintegration measurement is referenced from [11].

2) Visual Measurement Residual: Referring to the VINS-Fusion approach, using the measurement error of the camera defined on the unit sphere [11], consider the m-th feature observed first in the k_1-th image, the residual for the feature observation in the k_2-th image is defined as

$$
\begin{aligned}
r_{visual}(\mathbf{z}_m^{c_{k_2}}, \boldsymbol{\chi}) &= \left[\mathbf{b}_1, \mathbf{b}_2\right]^T \left(\hat{\mathcal{P}}_m^{c_{k_2}} - \frac{\mathcal{P}_m^{c_{k_2}}}{||\mathcal{P}_m^{c_{k_2}}||}\right) \\
\hat{\mathcal{P}}_m^{c_{k_2}} &= \pi_c^{-1}\left(\begin{bmatrix}\hat{u}_m^{c_{k_2}} \\ \hat{v}_m^{c_{k_2}}\end{bmatrix}\right) \\
\mathcal{P}_m^{c_{k_2}} &= \mathbf{R}_B^C\Big(\mathbf{R}_L^{B^{k_2}}\Big(\mathbf{R}_{B^{k_1}}^L\Big(\mathbf{R}_C^B\frac{1}{\lambda_m}\pi_c^{-1}\Big(\begin{bmatrix}\hat{u}_m^{c_{k_1}} \\ \hat{v}_m^{c_{k_1}}\end{bmatrix}\Big) \\
&\quad + \mathbf{p}_C^B\Big) + \mathbf{p}_{B^{k_1}}^L - \mathbf{p}_{B^{k_2}}^L\Big) - \mathbf{p}_C^B\Big)
\end{aligned}
\tag{2}
$$

C. Drift Residual

Despite obtaining the transformation matrices for $\{L\}$ and $\{W\}$ through the motion capture system, minor errors may arise due to factors such as time delays and measurement precision. Additionally, the execution of VIRPO by the sub-aircraft introduces cumulative drift errors. Therefore, the optimization process integrates the drift biases from other aircraft's pose data. VIO involves four degrees of freedom, as outlined in [15], encompassing translations along the xyz-axis and yaw rotation. The drift can be appropriately described as a random walk of the local odometry frame, as discussed in [16]:

$$
\begin{aligned}
{}_L^W\mathbf{T}_{t_{k+1}} &= {}_L^W\mathbf{T}_{t_k} + \mathcal{N}(\mathbf{0}, \sigma_{drift}^2) \\
{}_L^W\hat{\mathbf{T}}_{t_k} &= \delta_L^W\mathbf{T}_{t_k}\,{}_L^W\mathbf{T}_k
\end{aligned}
\tag{3}
$$

Based on the models for IMU bias and the analysis of VIO drift modeling in reference [17], it is essential to maintain a consistent time interval $\delta_L^W\mathbf{T}$ between adjacent frames with a small time gap. Consequently, the residual can be expressed as:

$$
\mathbf{r}_{drift}(\boldsymbol{\chi}^o) = \sum_{k=0}^{K}\sum_{j\in\mathcal{F}_k^o}\left|\left|\delta_{L_o}^W\mathbf{T}_{t_{j+1}} - \delta_{L_o}^W\mathbf{T}_{t_j}\right|\right| \tag{4}
$$

D. Kinematic Constraints Residual

The VIRPO algorithm determines the aircraft's velocity, position, and attitude in the coordinate system $\{L\}$. It

carefully accounts for the interaction between the spherical joint connection and the aircraft's structural integrity to compute the velocity and position of the spherical joint within $\{L\}$. Subsequently, the transformation matrix linking frame $\{L\}$ with frame $\{W\}$ is utilized to determine the velocity and position of the spherical joint within frame $\{W\}$. The position of the spherical joint in frame $\{B\}$ of the sub-aircraft, with the carrier system situated in the right-front-upper, is denoted as ${}_{sp}\mathbf{p}(0,0,x)$. Here x signifies the precise Euclidean distance between the spherical joint and the IMU, a value that can be determined with a high degree of accuracy through measurement. Subsequently, the position and velocity of the spherical joint in the $\{W\}$ is derived as expressed in the following formula:

$$ {}^{W}_{sp}\mathbf{p}^{i}_{t_j} = {}^{W}_{L_i}\hat{\mathbf{T}}_{rot}({}^{L_i}\mathbf{p}^{i}_{t_j} + \mathbf{q}^{i}_{t_j}\,{}^{B_i}_{sp}\mathbf{p}) + {}^{W}_{L_i}\hat{\mathbf{T}}_{tran} \tag{5} $$

$$ {}^{B_i}\mathbf{v}_{ij} = (\mathbf{q}^{i}_{t_j})^{-1}\,{}^{L_i}\mathbf{v}^{i}_{t_j} \tag{6} $$

$$ {}^{B_i}_{sp}\mathbf{v}^{i}_{t_j} = {}^{B_i}\mathbf{v}^{i}_{t_j} + \boldsymbol{\omega}^{i}_{t_j} \times {}^{B_i}_{sp}\mathbf{p} \tag{7} $$

$$ {}^{L_i}_{sp}\mathbf{v}^{i}_{t_j} = \mathbf{q}^{i}_{t_j}\,{}^{B}_{sp}\mathbf{v}^{i}_{t_j} \quad {}^{W}_{sp}\mathbf{v}^{i}_{t_j} = {}^{W}_{L_i}\hat{\mathbf{T}}_{rot}\,{}^{B}_{sp}\mathbf{v}^{i}_{t_j} \tag{8} $$

where $\boldsymbol{\omega}^{i}_{t_j}$ represents the angular velocity of aircraft i at time t_j. It is obtained by subtracting the gyro zero drift $\mathbf{bg}$ from the IMU data $\hat{\boldsymbol{\omega}}$. Let $\mathcal{F}^{i}_{k}$ denote the set of physical constraint's measurements between the k-th keyframe and the $(k+1)$th keyframe of the $\{ID_i\}$. The pose data at t_j between the two keyframes is acquired through pre-integration.

1) Position Constraint Residual: The separation between spherical joints within the IAP remains constant and can be accurately measured. To illustrate, consider the spherical joint associated with the aircraft o_1 and o_2. We denote $l^{o_1\,o_2}$ as the Euclidean distance between the respective spherical joints of the aircraft o_1 and o_2. This measurement leads to the derivation of positional constraint residuals for the IAP:

$$ ||{}^{W}_{sp}\mathbf{p}^{o_1}_{t_j} - {}^{W}_{sp}\mathbf{p}^{o_2}_{t_j}|| = l^{o_1\,o_2} \tag{9} $$

$$ \rho_{\mathfrak{D}}({}^{W}_{sp}\mathbf{p}^{o_1}_{t_j}, {}^{W}_{sp}\mathbf{p}^{o_2}_{t_j}) = ||{}^{W}_{sp}\mathbf{p}^{o_1}_{t_j} - {}^{W}_{sp}\mathbf{p}^{o_2}_{t_j}|| - l^{o_1\,o_2} \tag{10} $$

$$ \mathbf{r}_{kp}(\chi^i) = \sum_{o\in\{ID\}}\sum_{o\neq i}\sum_{k=0}^{K}\sum_{j\in\mathcal{F}^{i}_{k}} ||\rho_{\mathfrak{D}}({}^{W}_{sp}\mathbf{p}^{i}_{t_j}, {}^{W}_{sp}\mathbf{p}^{o}_{t_j})|| \tag{11} $$

2) Angular Velocity Constraint Residual: The angular velocity of a rigid body can be computed based on the velocities and positions of three points that are not collinear on the rigid body. Utilizing formula (12) to compute the angular velocity of the central platform based on the velocity and position of the three spherical joints, the velocity residual of the physical constraint, along with the residuals of velocity and angular velocity, can be expressed as follows. Here, ${}^{B_0}_{Center}\boldsymbol{\omega}_{t_j}$ represents the angular velocity at time t_j of the central platform IMU:

$$ \begin{aligned} &\mathbf{r}^3 = \mathbf{p}^{o_1\,o_2} \times \mathbf{p}^{o_1\,o_3} = r^3\mathbf{e}^3 \quad r^{i\,j} = \mathbf{p}^{i\,o_1}\cdot\mathbf{p}^{j\,o_1} \\ &\boldsymbol{\lambda} = \frac{1}{r^3}(r^{o_2\,o_3}\mathbf{e}^3\cdot\mathbf{v}^{o_1\,o_2} - r^{o_2\,o_2}\mathbf{e}^3\cdot\mathbf{v}^{o_1\,o_3}) \\ &\boldsymbol{\mu} = \frac{1}{r^3}(r_{o_3\,o_3}\mathbf{e}^3\cdot\mathbf{v}^{o_1\,o_2} - r^{o_2\,o_3}\mathbf{e}^3\cdot\mathbf{v}^{o_1\,o_3}) \\ &{}_{est}\boldsymbol{\omega} = \frac{1}{r^{o_2\,o_2} + r^{o_3\,o_3}}\Big\{[\boldsymbol{\lambda}\mathbf{I} + \boldsymbol{\Omega}(\mathbf{v}^{o_1\,o_2})]\mathbf{p}^{o_1\,o_2} + \\ &[\boldsymbol{\mu}\mathbf{I} + \boldsymbol{\Omega}(\mathbf{v}^{o_1\,o_3})]\mathbf{p}^{o_1\,o_3}\Big\} \end{aligned} \tag{12} $$

where $\mathbf{I}$ is a identity matrix and $\boldsymbol{\Omega}$ is skew-symmetric matrix of a vector. The residual equation for velocity constraints is derived as follows:

$$ \mathbf{r}_{kv}(\chi^i) = \sum_{\substack{o_1,o_2\in\{ID\}\\ o_1\neq o_2\neq i\,o_1\neq i}}\sum_{k=0}^{K}\sum_{j\in\mathcal{F}^{i}_{k}} \Big|\Big|\,||{}^{W}_{est}\boldsymbol{\omega}_{t_j}|| - ||{}^{B_0}_{Center}\boldsymbol{\omega}_{t_j}||\,\Big|\Big| \tag{13} $$

3) Attitude Constraint Residual: In the case of an IAP comprised of multiple sub-aircraft, the combined positions of these sub-aircraft can be utilized to infer a partial orientation for the central rigid body. The central rigid body's flight control module integrates a 9-axis IMU sensor, which further enhances the accuracy of attitude estimation for the central rigid body. The PX4 flight control module employs a standard Extended Kalman Filter (EKF) for estimating the orientation of $\{B_0\}$ with respect to $\{L_0\}$, denoted as ${}^{L_0}_{B_0}\mathbf{R}$. It is essential to emphasize that the $\{L_0\}$ serves as a fixed reference frame. This reference frame is frequently referred to as the ENU (East-North-Up) frame within the PX4 flight control module.

In the scenario of a non-collinear deployment of SIAP, the unit normal vector of the plane formed by the three spherical joints can be defined as follows:

$$ {}^{W}\mathbf{n} = \frac{{}^{W}\mathbf{p}^{o_1\,o_2} \times {}^{W}\mathbf{p}^{o_1\,o_3}}{||{}^{W}\mathbf{p}^{o_1\,o_2} \times {}^{W}\mathbf{p}^{o_1\,o_3}||} \tag{14} $$

where $1 \le o_1 < o_2 < o_3 \le N$, $\mathbf{p}^{o_1\,o_2}$ represents the position vectors of sub-aircraft o_1 and o_2 relative to the spherical joints. Taking into account that both $\mathbf{z}_{L_0}$ and $\mathbf{z}_W$ are aligned in the opposite direction to gravity, $(\mathbf{n})_z$ represents the cosine of the angle α between $\mathbf{n}$ and the z-axis of the $\{L_0\}$, which results in the calculation of the subsequent attitude residual:

$$ \mathbf{r}_{kq}(\chi^i) = \sum_{\substack{o_1,o_2\in\{ID\}\\ o_1\neq o_2\neq i}}\sum_{k=0}^{K}\sum_{j\in\mathcal{F}^{i}_{k}} ||\cos\hat{\alpha} - ({}^{W}\mathbf{n})_z|| \tag{15} $$

In contrast to the SIAP, the LIAP features a co-linear deployment and does not require the computation of plane normal vectors. Nevertheless, it is still possible to determine the pitch angle of the central rigid body based on the positions of spherical joints. This information enables the calculation of attitude residuals:

$$ \mathbf{r}_{kq}(\chi^i) = \sum_{\substack{o\in\{ID\}\\ o\neq i}}\sum_{k=0}^{K}\sum_{j\in\mathcal{F}^{i}_{k}} ||\cos\hat{\beta} - ({}^{W}\mathbf{e}^{\mathbf{io}})_z|| \tag{16} $$

where $\mathbf{e}^{i\,o}$ is unit vector of $\mathbf{p}^{i\,o}$, $\hat{\beta}$ is the angle between the $\mathbf{x}_{B_0}$ and $\mathbf{z}_{L_0}$. When using the direction cosine matrix to represent ${}^{L_0}_{B_0}\mathbf{R}$, $\cos\hat{\alpha}$ and $\cos\hat{\beta}$ correspond to the elements in the third column of ${}^{L_0}_{B_0}\mathbf{R}$, specifically, the third and first elements, respectively.

E. UWB Measurements Residual

The UWB module provides distance measurements between the UWB node on the aircraft and stationary ground UWB anchors, yielding UWB distance measurements represented as z_t. ${}^{W}_{anchor}\boldsymbol{p}$ is the location of the anchor point in the $\{W\}$ system. These measurements are further characterized by a Gaussian noise model:

$$z_t = ||{}^{W}\mathbf{p}_t - {}^{W}_{anchor}\,\boldsymbol{p}|| + \mathcal{N}(0, \sigma_d^2) \tag{17}$$

A range-focused methodology in [18] has been implemented to optimize the utilization of UWB ranging data between two keyframes. The set of UWB measurements between the k-th and $(k+1)$-th keyframes of $\{ID_i\}$ is denoted as $\mathcal{J}_i^k$:

$$\mathcal{J}_k^i = \{(z_{t_j}, \Delta\hat{\mathbf{p}}_{t_k\,t_j})\}_{t_k \le t_j < t_{k+1}} \tag{18}$$

where $\Delta\hat{\mathbf{p}}_{t_k\,t_j}$ denotes the positional change from time t_k to time t_j, calculated via pre-integration. The UWB range residuals are expressed as [19]:

$$\begin{aligned}\mathbf{r}_{uwb}(\chi^i) &= \sum_{k=0}^{K}\sum_{j\in\mathcal{J}_k^i}\rho_{uwb}({}^{W}\mathbf{p}_{t_k}^i, z_{t_j}, \Delta\hat{\mathbf{p}}_{t_k\,t_j}) \\ \rho_{uwb}({}^{W}\mathbf{p}_{t_k}^i, z_{t_j}, \Delta\hat{\mathbf{p}}_{t_k\,t_j}) &= \Big|\Big|\,||{}^{W}\mathbf{p}_{t_k}^i + \Delta\hat{\mathbf{p}}_{t_k\,t_j} - {}^{W}_{anchor}\boldsymbol{p}|| \\ &\quad - z_{t_j}\Big|\Big|\end{aligned} \tag{19}$$

Incorporating residual absolute information obtained from distance measurements between VINS and stationary anchors can effectively reduce VINS drift. This ensures that the odometry error of each aircraft remains within a relatively low range, thereby maintaining the accuracy of pose data for other sub-aircraft within the kinematic constraint residuals.

IV. Experimental Results

A. Performance Metrics and Experiment Environment

We employ Root Mean Square Error (RMSE) as the evaluation metric for Absolute Trajectory Error (ATE). It is important to note that the entire trajectory is considered in the trajectory alignment process. The distance between spherical joints was precisely measured with a σ_l of 0.04. Angular velocity and attitude constraints rely on IMU angular velocity and attitude, with σ_{att} and σ_ω both set to 0.01 due to their high precision, as indicated in [20]. Distance measurements are subject to UWB accuracy. Despite the absence of UWB sensors in our data, we simulated real values by introducing zero-mean Gaussian noise with a variance of 0.05 [18], resulting in σ_d equal to 0.05. The motion trajectory of the central platform in the data collection experiment is shown in the Figure 4.

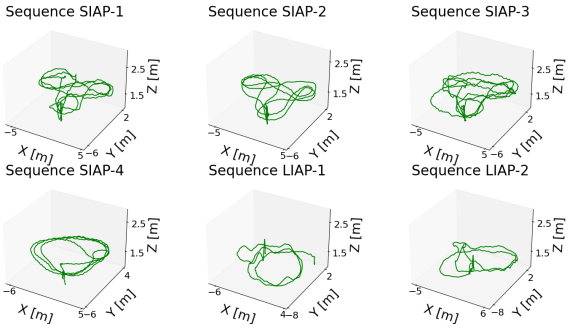


Fig. 4: Central platform experimental trajectory display

B. VIRPO on Collective Dataset

TABLE I: RMSE Values of ATE (m) Results in Collective Dataset

Sequence ID		ATE(m)			
		VINS	VINS+UWB	VINS+UWB+ self-Phy	VINS+UWB+ full-Phy
SIAP-1	ID_1	0.533	0.272	0.196	**0.180**
	ID_2	0.343	0.262	0.162	**0.126**
	ID_3	0.455	0.204	0.159	**0.140**
SIAP-2	ID_1	0.549	0.323	0.201	**0.185**
	ID_2	0.407	0.175	0.144	**0.120**
	ID_3	0.443	0.158	0.132	**0.096**
SIAP-3	ID_1	0.278	0.249	0.199	**0.170**
	ID_2	0.489	0.196	0.144	**0.152**
	ID_3	0.280	0.202	0.183	**0.153**
SIAP-4	ID_1	0.599	0.172	0.168	**0.165**
	ID_2	0.680	0.218	0.251	**0.208**
	ID_3	0.410	0.363	0.228	**0.190**
LIAP-1	ID_1	0.294	0.243	0.188	**0.182**
	ID_2	0.601	0.195	0.157	**0.133**
	ID_3	0.292	0.139	0.142	**0.120**
LIAP-2	ID_1	Failed	0.191	0.165	**0.158**
	ID_2	0.399	0.178	0.177	**0.167**

Table I presents verified results from a dataset where we used different positioning methods: VINS, VINS+UWB, VINS+UWB+full-Phy, and VINS+UWB+self-Phy. The baseline method was VINS-Fusion, and the VINS+UWB method incorporated UWB ranging data at a 40Hz frequency. VINS+UWB+full-Phy added physical constraint terms to all sub-aircraft in tightly coupled optimization, while VINS+UWB+self-Phy added them to one sub-aircraft while using UWB as the foundation. "Self-Phy" refers to the sub-aircraft labeled in this row utilizing physical constraint residuals, while the remaining sub-aircraft solely employ the VINS+UWB method. We measured the ATE for each aircraft ($\{ID_1\}$, $\{ID_2\}$ and $\{ID_3\}$) in each sequence, showing accuracy improvements compared to the baseline and VINS+UWB methods. Our full-Phy and self-Phy methods improve global localization accuracy by leveraging physical constraints and sub-aircraft states to enhance estimation. Experimental sequences included SIAP, LIAP, and two-agent LIAP combinations, validating our approach's effectiveness. This underscores the generalizability of our method across multiple configurations, indicating that similar approaches can be adopted for IAP in other configurations in the future. Compared to the VINS+UWB method, the full-Phy approach exhibited a maximum decrease of 0.173 m in the ATE metric, with the maximum decrease percentage reaching

52.0%. The average percentage decrease was 28.7%. The quartile diagrams in Fig. 5a and Fig. 5b illustrate quartiles for three algorithms in experimental sequences $SIAP-3$ and $LIAP-1$. This demonstrates that the mean, maximum, and minimum errors can all be reduced using our approach. For exceptional outlier estimation results, it demonstrates a notable suppressive effect.

C. Physical Constraints on Collective Dataset

TABLE II: Assessment of ATE Results for Various Kinematic Constraints on the Dataset. **pos** - Position constraint, **velo** - Angular velocity constraint, **att** - Attitude constraint, **drift** - Accumulated Drift Estimation. **All – Kin** - Using all types of kinematic constraints

Sequence ID	Aircraft ID	pos+drift	pos+velo+drift	pos+att+drift	All-Kin	All-kin+drift
SIAP-1	ID_1	0.237	0.225	0.189	0.188	**0.180**
SIAP-1	ID_2	0.215	0.192	0.131	0.143	**0.126**
SIAP-1	ID_3	0.183	0.175	0.152	0.146	**0.140**
SIAP-2	ID_1	0.259	0.241	0.192	0.189	**0.185**
SIAP-2	ID_2	0.187	0.173	0.135	0.137	**0.120**
SIAP-2	ID_3	0.162	0.155	0.102	0.113	**0.096**
SIAP-3	ID_1	0.227	0.202	0.179	0.177	**0.170**
SIAP-3	ID_2	0.204	0.193	0.166	0.157	**0.152**
SIAP-3	ID_3	0.183	0.170	0.156	0.158	**0.153**
SIAP-4	ID_1	0.21	0.181	0.169	0.172	**0.165**
SIAP-4	ID_2	0.256	0.237	0.211	0.308	**0.208**
SIAP-4	ID_3	0.285	0.202	0.193	**0.184**	0.190
LIAP-1	ID_1	0.183	/	/	0.182	**0.182**
LIAP-1	ID_2	0.163	/	/	0.138	**0.133**
LIAP-1	ID_3	0.143	/	/	0.128	**0.120**
LIAP-2	ID_1	0.168	/	/	0.165	**0.158**
LIAP-2	ID_2	0.182	/	/	0.174	**0.167**

Table II presents verified results for four combinations of kinematic constraints in our dataset. The representation of position, velocity, and attitude constraints is denoted by **pos**, **velo**, and **att**, respectively. Additionally, the inclusion of accumulated drift estimation for other sub-aircraft is indicated by the variable **drift**. To assess the performance of these configurations, UWB ranging constraints have been incorporated into each of these combinations. The experiments show that incorporating **pos**, **velo**, and **att** constraints enhances IAP localization accuracy. Notably, position and attitude constraints have a more significant impact compared to velocity constraints. This is because the velocity constraint inherently contains position information, and VIO already provides relatively precise velocity measurements. As a result, introducing velocity constraints has a more modest impact. Additionally, including the **drift** residual term leads to a slight improvement, mitigating drift errors from other aerial vehicles.

It is important to mention that with a limited number of constraint types, the effect on localization accuracy can vary across different aerial vehicles in individual test sequences. This occurs because the tightly coupled method in this paper processes data within a sliding window, while the motion of the IAP covers a limited range compared to the other aerial swarms, resulting in an inadequate level of system observability.

D. Size of Physical Constraint

The dimension of $\mathcal{F}_k$ exerts an influence on the quantity of physical constraint residuals involved in BA optimization, subsequently impacting the results of odometry-based localization. To gauge this impact, constraints on the size

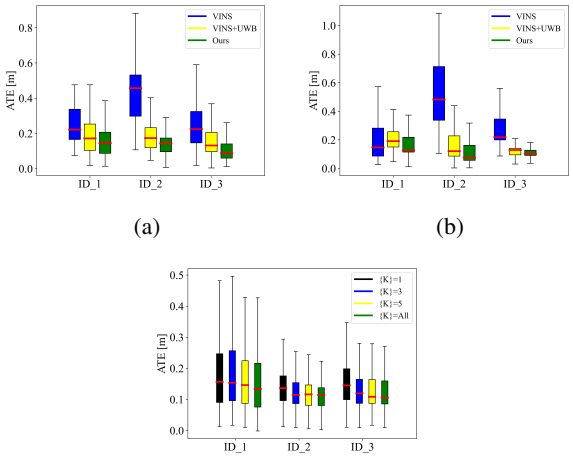


Fig. 5: (a): Quartile plot of errors for three aircrafts in the SIAP experiment sequence 3. (b): Quartile plot of errors for three aircrafts in the LIAP experiment sequence 1. (c): Quartile plots of localization errors for different $\mathcal{F}_k$ sizes: 1, 3, 5, and All on SIAP-1.

of $\mathcal{F}_k$ were intentionally imposed in the SIAP-1 sequence. Fig. 5c displays quartile plots of localization errors for different $\mathcal{F}_k$ sizes: 1, 3, 5, and All (using all available physical constraint data). The figure shows error reductions as $\mathcal{F}_k$ size increases, supporting the idea that more physical constraint residuals enhance localization accuracy. However, $\mathcal{F}_k$ contains redundant data from a single keyframe, and the short motion duration limits observability for optimization variables, posing a clear constraint on improving localization accuracy.

V. Conclusion

In this paper, a distributed multi-agent VIRPO system is introduced for the analysis of position, velocity, and attitude constraints within IAP. By leveraging the inherent constraint relationships within the IAP, corresponding methods are proposed to tightly integrate odometry and kinematic constraints, thereby enhancing odometry accuracy. The system's performance is rigorously validated using a collected dataset, resulting in a substantial 28.7% reduction in drift compared to the baseline. Additionally, experimental results confirm the effectiveness of various physical constraints. This work addresses a significant challenge in the realm of IAP, offering promising avenues for improving the accuracy and reliability of IAP positioning and navigation systems. Future research may explore more complex IAP configurations and additional types of constraints to further enhance system performance. This study holds significant importance in advancing the field of IAP and will provide valuable insights and inspiration for future related research.

REFERENCES

[1] Y. Yu and X. Ding, "Safe landing analysis of a quadrotor aircraft with two legs," *Journal of Intelligent & Robotic Systems*, vol. 76, pp. 527–537, 2014.

[2] Y. Yu, C. Shi, D. Shan, V. Lippiello, and Y. Yang, "A hierarchical control scheme for multiple aerial vehicle transportation systems with uncertainties and state/input constraints,"*Applied Mathematical Modelling*, vol. 109, pp. 651–678, 2022.

[3] Y. Yu, P. Li, and P. Gong, "Finite-time geometric control for underactuated aerial manipulators with unknown disturbances," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 13, pp. 5040–5061, 2020.

[4] X. Yu, P. T. Morrn, S. Salimpour, J. P. Queralta, and T. Westerlund, "Loosely coupled odometry, uwb ranging, and cooperative spatial detection for relative monte-carlo multi-robot localization," *arXiv preprint arXiv:2304.06264*, 2023.

[5] M. Kok, J. D. Hol, and T. B. Schön, "Using inertial sensors for position and orientation estimation," *arXiv preprint arXiv:1704.06053*, 2017.

[6] J. P. Queralta, Q. Li, F. Schiano, and T. Westerlund, "Vio-uwb-based collaborative localization and dense scene reconstruction within heterogeneous multi-robot systems," in *2022 International Conference on Advanced Robotics and Mechatronics (ICARM)*. IEEE, 2022, pp. 87–94.

[7] R. Liu, Z. Deng, Z. Cao, M. Shalihan, B. P. L. Lau, K. Chen, K. Bhowmik, C. Yuen, and U.-X. Tan, "Distributed ranging slam for multiple robots with ultra-wideband and odometry measurements," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 13 684–13 691.

[8] F. Zhu, Y. Ren, F. Kong, H. Wu, S. Liang, N. Chen, W. Xu, and F. Zhang, "Swarm-lio: Decentralized swarm lidar-inertial odometry," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 3254–3260.

[9] H. Xu, L. Wang, Y. Zhang, K. Qiu, and S. Shen, "Decentralized visual-inertial-uwb fusion for relative state estimation of aerial swarm," in *2020 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2020, pp. 8776–8782.

[10] H. Xu, Y. Zhang, B. Zhou, L. Wang, X. Yao, G. Meng, and S. Shen, "Omni-swarm: A decentralized omnidirectional visual–inertial–uwb state estimation system for aerial swarms," *IEEE Transactions on Robotics*, vol. 38, no. 6, pp. 3374–3394, 2022.

[11] T. Qin, P. Li, and S. Shen, "Vins-mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004–1020, 2018.

[12] K. Nirmal, A. Sreejith, J. Mathew, M. Sarpotdar, A. Suresh, A. Prakash, M. Safonova, and J. Murthy, "Noise modeling and analysis of an imu-based attitude sensor: improvement of performance by filtering and sensor fusion," in *Advances in optical and mechanical technologies for telescopes and instrumentation II*, vol. 9912. SPIE, 2016, pp. 2138–2147.

[13] R. Buchanan, V. Agrawal, M. Camurri, F. Dellaert, and M. Fallon, "Deep imu bias inference for robust visual-inertial odometry with factor graphs," *IEEE Robotics and Automation Letters*, vol. 8, no. 1, pp. 41–48, 2022.

[14] M. George and S. Sukkarieh, "Tightly coupled ins/gps with bias estimation for uav applications," in *Proceedings of Australiasian Conference on Robotics and Automation (ACRA)*, 2005, pp. 1–7.

[15] G. Huang, "Visual-inertial navigation: A concise review," in *2019 international conference on robotics and automation (ICRA)*. IEEE, 2019, pp. 9572–9582.

[16] T. Ziegler, M. Karrer, P. Schmuck, and M. Chli, "Distributed formation estimation via pairwise distance measurements," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3017–3024, 2021.

[17] T. H. Nguyen, T.-M. Nguyen, and L. Xie, "Flexible and resource-efficient multi-robot collaborative visual-inertial-range localization," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 928–935, 2021.

[18] X. L. Nguyen T H, Nguyen T M, "Range-focused fusion of camera-imu-uwb for accurate and drift-reduced localization," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1678–1685, 2021.

[19] S. Jia, Y. Jiao, Z. Zhang, R. Xiong, and Y. Wang, "Fej-viro: A consistent first-estimate jacobian visual-inertial-ranging odometry," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 1336–1343.

[20] D. Lee, S. Lee, S. Park, and S. Ko, "Test and error parameter estimation for mems-based low cost imu calibration," *International Journal of Precision Engineering and Manufacturing*, vol. 12, pp. 597–603, 2011.

[21] Y. Yu and X. Ding, "A global tracking controller for underactuated aerial vehicles: design, analysis, and experimental tests on quadrotor," *IEEE/ASME Transactions on Mechatronics*, vol. 21, no. 5, pp. 2499–2511, 2016.

[22] Y. Yu and Y. Dong, "Global fault-tolerant control of underactuated aerial vehicles with redundant actuators," *International Journal of Aerospace Engineering*, vol. 2019, 2019.

[23] Y. Yu and V. Lippiello, "6d pose task trajectory tracking for a class of 3d aerial manipulator from differential flatness," *IEEE Access*, vol. 7, pp. 52 257–52 265, 2019.

[24] Y. Yu, K. Wang, R. Guo, V. Lippiello, and X. Yi, "A framework to design interaction control of aerial slung load systems: transfer from existing flight control of under-actuated aerial vehicles," *International Journal of Systems Science*, pp. 1–13, 2021.

[25] Y. Yu, R. Guo, K. Wang, and V. Lippiello, "Transfer from existing flight control of under-actuated aerial vehicles to interaction control of aerial slung load systems," in *2021 European Control Conference (ECC)*. IEEE, 2021, pp. 2045–2051.

[26] C. Shi and Y. Yu, "Design and implementation of a fully-actuated integrated aerial platform based on geometric model predictive control," *Micromachines*, vol. 13, no. 11, p. 1822, 2022.

[27] C. Shi, Y. Yu, Y. Ma, and D. E. Chang, "Constrained control for systems on matrix lie groups with uncertainties," *International Journal of Robust and Nonlinear Control*, 2022.

[28] J. Sun, Y. Yu, and B. Xu, "Towards flying carpet: Dynamics modeling, and differential-flatness-based control and planning," in *Cognitive Systems and Information Processing: 7th International Conference, ICCSIP 2022, Fuzhou, China, December 17-18, 2022, Revised Selected Papers*. Springer, 2023, pp. 351–370.

[29] J. Du, Y. Fan, K. Wang, Y. Feng, and Y. Yu, "Aerobotsim: A high-photo-fidelity simulator for heterogeneous aerial systems under physical interaction," in *Cognitive Systems and Information Processing: 7th International Conference, ICCSIP 2022, Fuzhou, China, December 17-18, 2022, Revised Selected Papers*. Springer, 2023, pp. 274–287.

[30] Y. Feng, C. Shi, J. Du, Y. Yu, F. Sun, and Y. Song, "Variable admittance interaction control of uavs via deep reinforcement learning," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 1291–1297.

[31] C. Shi, G. Lai, Y. Yu, M. Bellone, and V. Lippiello, "Real-time multi-modal active vision for object detection on uavs equipped with limited field of view lidar and camera," *IEEE Robotics and Automation Letters*, 2023.

[32] M. Zhang, M. Li, K. Wang, T. Yang, Y. Feng, and Y. Yu, "Zero-shot sim-to-real transfer of robust and generic quadrotor controller by deep reinforcement learning," in *International Conference on Cognitive Systems and Signal Processing*. Springer, 2023, pp. 27–43.

[33] M. Jiang, M. Zuo, X. Yu, R. Guo, R. Wang, and Y. Yu, "Perception-aware motion control of multiple aerial vehicle transportation systems," in *International Conference on Cognitive Systems and Signal Processing*. Springer, 2022, pp. 474–488.