Cooperative Transport of a Suspended Payload via Two Aerial Robots with Inertial Sensing

XIE, Heng; DONG, Kaixu; CHIRARATTANANON, Pakpong

Published in:
IEEE Access

Published: 01/01/2022

Document Version:
Final Published version, also known as Publisher’s PDF, Publisher’s Final version or Version of Record

License:
CC BY

Publication record in CityU Scholars:
Go to record

Published version (DOI):
10.1109/ACCESS.2022.3194932

Publication details:

Citing this paper
Please note that where the full-text provided on CityU Scholars is the Post-print version (also known as Accepted Author Manuscript, Peer-reviewed or Author Final version), it may differ from the Final Published version. When citing, ensure that you check and use the publisher's definitive version for pagination and other details.

General rights
Copyright for the publications made accessible via the CityU Scholars portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights. Users may not further distribute the material or use it for any profit-making activity or commercial gain.

Publisher permission
Permission for previously published items are in accordance with publisher's copyright policies sourced from the SHERPA RoMEO database. Links to full text versions (either Published or Post-print) are only available if corresponding publishers allow open access.

Take down policy
Contact lbscholars@cityu.edu.hk if you believe that this document breaches copyright and provide us with details. We will remove access to the work immediately and investigate your claim.
Cooperative Transport of a Suspended Payload via Two Aerial Robots With Inertial Sensing

HENG XIE, KAIXU DONG, AND PAKPONG CHIRARATTANANON, (Member, IEEE)
Department of Biomedical Engineering, City University of Hong Kong, Hong Kong, SAR, China
Corresponding author: Pakpong Chirarattananon (pakpong.c@cityu.edu.hk)
This work was supported by the Research Grants Council of the Hong Kong Special Administrative Region of China under Grant CityU 11215117.

ABSTRACT In this article, we address the problem of aerial transport of a suspended payload by two vehicles with minimal sensing and computational requirements. Herein, the leader-follower paradigm is adopted, allowing the follower robot to estimate and stabilize the formation using only feedback from its onboard inertial sensor. This is achieved through an analysis of the underlying dynamics of the coupled system. The gained insights lead to the introduction of a two-stage Kalman-based estimation strategy and the linear control laws that are capable of robustly dealing with the underactuated leader-payload-follower system. The vision-less methods have been verified through a series of flight experiments. The efficiency and the robustness of the schemes enable the cooperative payload transport to be accomplished by two sub-100-g robots with highly limited computational power in both indoor and outdoor environments.

INDEX TERMS Autonomous vehicles, intelligent transportation systems, state estimation.

I. INTRODUCTION

Potential applications of Micro Aerial Vehicles (MAVs), including quadcopters, have accelerated the developments in several related topics, such as enhanced flight capabilities [1]–[3], locomotion [4]–[6], aerial manipulation [7], and swarm coordination [8]. A single flying robot carrying a payload has been successfully deployed for an inspection, surveillance, or monitoring task. Nevertheless, there remain limitations on the range, endurance, and payload capacity, which exacerbate as the platform size reduces [9]. Simultaneous uses of multiple MAVs offer a possible solution to enable a team of robots to collaboratively carry more, cover larger areas, or robustly deal with failures. Nevertheless, the cooperation usually demands coordinated localization, agent communication, and resource allocation [8].

When it comes to payload capacity, the thrust limit of a single robot restricts the weight of the payload it can accommodate. A larger platform, in addition to being less human friendly and more costly, may suffer from reduced agility. Subsequently, it has been proposed to appoint multiple MAVs to carry a single load [10]. For a single rotorcraft to transport an object, the payload is usually rigidly attached [2], [11] or hoisted [12]–[14]. To have a payload fixed to the robot, considerations must be given to the grasping mechanism [15], weight distribution [2], and the parcel size [11]. The suspension approach, while benefits from a simpler load attachment, introduces passive degrees of freedom to the MAV, necessitating the robot to dynamically regulate the slung payload [13], [16].

The cargo transport task with a group of robots can be tackled in a similar manner, by rigid attachments [2], [17], [18] or payload suspension [19], [20]. Both possess challenges and constraints akin to the single robot operation, but with the complexity of multi-robot interactions and coupled dynamics. Realization of such feats must consider issues related to (i) trajectory planning [10], [19]–[21], (ii) formation stabilization [22]–[25], and (iii) feedback and communication [18], [26], [27].

In the first step, the trajectory planning concerns the search for feasible trajectories provided the dynamics of aerial robots and the constraints between the payload and the robots. For rigidly attached payloads, this task is arguably simplified due to the reduced degrees of freedom. In the case of a suspended
payload, for instance, it was shown that the system is differentially flat when the cable tensions are strictly positive [20]. This means the control inputs for the robots can be produced from the desired payload trajectory. In more complicated situations, path planning is also required to identify collision-free path [21]. After the trajectory generation, controllers are devised to ensure that the agents are able to follow the planned trajectories and the formation is maintained and stable. Several control strategies have been proposed, including optimization-based [28], passivity-based [29], and energy-based [25] controllers. However, many of these methods assume the full knowledge of the system. That is, the locations and/or attitude of all agents and the payload must be available for each robot to compute the command. Therefore, the validation can only be achieved via numerical simulations or with the help of motion capture systems and centralized communication networks.

To overcome the remaining obstacle of achieving aerial cooperative transport in real-world settings, more recently, a few vision-based [26], [30], [31], and inertial-based [18] strategies have been introduced to workaround the absence of motion capture feedback. In [30], the authors employed two robots, each equipped with a downward camera, to autonomously carry a single payload by leveraging fixed attachments to form a single rigid structure. The fixture radically reduces the degrees of freedom. With visual odometry, each robot initially independently estimates its pose. With the known fixed structure between the two robots, the two pose estimates are combined and refined using the loop closure constraint to yield a single fuse estimate for the entire system, allowing stable flights to be realized.

Unlike [30], in [18], the authors address the problem of multiple robots cooperatively transporting a payload by grasping with a spherical joint, under the leader-follower paradigm. With the spherical joints, the entire system is no longer a single rigid structure. Still, without cables, the system contains fewer degrees of freedom compared with a system with a cable suspended payload. In [18], the decentralized approach is developed by means of admittance control, the follower estimates the external force acting on the vehicle using only the onboard Inertial Measurement Unit (IMU) and commands. The follower then regulates its dynamics to follow the reference trajectory and compensate for the external force.

To deal with a tethered load, the authors in [26] allow the payload assume different poses (i.e., the payload is not rigidly attached). The vision-based method in [26] eradicates the need for communication, also by using a leader-follower approach. Each robot visually tracks the load using tags and the follower tracks the leader through another tag. The control scheme makes an important simplifying assumption that the payload can be split symmetrically into two halves, each to be regarded as a simple pendulum by each vehicle. This neglects the full interactions between the two robots via the payload. The follower ensures a stable heading direction of the transported object. The framework enables the payload to be stably transported across the 3.3 m distance. Similarly, in [31], the aerial transport of a suspended rigid body payload was accomplished with three robots thanks to the used of monocular vision and inertial sensing. To do so, the use of visual perception is critical and communication between agents is used so that each agent is aware of their global and relative positions.

In this work, we take a vision-less approach to tackle the challenge of aerial transport of a suspended point-mass payload by two robots under the leader-follower format without the use of motion capture systems. Similar to recent state-of-the-art methods for the transport of a suspended payload [26], [31], the proposed framework uses only onboard sensors for estimation, allowing the method to be deployed in the real world. The vision-less implementation for a suspended payload, however, has not been accomplished in [26], [31]. The omission of visual feedback for estimation and control substantially reduces the computational complexity. Despite having to deal with the highly coupled dynamics stemming from the slung load, the lightweight inertia-based method highly benefits from computational efficiency and minimal hardware requirements. Without added components, the system does not suffer from the reduced payload capacity, rendering it feasible to be implemented on sub-100-g robots without a camera or a companion computing unit.

To accomplish this, this paper investigates the dynamics of the leader-payload-follower system from the follower’s perspective. The linear analysis reveals that, around the equilibrium state, the state decouples into two marginally stable subsystems. This allows the state variables to be separately and efficiently estimated using only IMU feedback and motor commands. The inertia-based estimation herein is distinct from [18] as the state, not external forces, are directly estimated. In other words, the localization and control of the follower and the payload is attained without extra parts. This has not been accomplished in [18], [26], [30], [31]. Following the estimation, the control laws are proposed based on the linearized dynamics of the entire coupled leader-follower system. The linear strategy proves to be highly efficient and robust due to the inherent marginal stability of the system about its equilibrium. That is, our approach exploits simplifications emerged from the insights gained from the linear analysis. The proposed control method, therefore, differs from those found in [18], [26], [30]. In [18], the admittance controller was used for the robot to compensate for the external wrench, without considering the coupled dynamics of the system in the control scheme. In [30], the authors also employ a simple linear controller for the follower. The controller was designed to control the external force applied to the robot, separating the dynamics of the leader and the payload from the follower. In [30], the entire system was treated as a rigid body. Hence, a nonlinear controller commonly used for quadcopters [32] can be readily employed.

All in all, the estimation and control methods proposed in this paper enables the follower robot to stabilized a suspended payload and the formation using only the existing onboard
IMU, not previously demonstrated by [18], [26], [30]. The preliminary outcome of this work was first presented in [33], but only the planar dynamics of the leader-payload-follower system was examined, resulting in a reduced system. The extension to three dimensions in this work and its associated linear analysis covers the pendulum-like motion of the payload that is under-actuated. This complicates the stabilization effort. Unlike in [33], the non-zero acceleration of the leader is also estimated in this article. The extension allows the follower to robustly follow the leader with more agile trajectories as proven in extensive flight experiments, both indoor and outdoor. Despite considering the entire system for estimation and control, the limitation remains the extension to the transport with more than two robots.

The paper is structured as follows. Section II provides the description of the system’s dynamics in the generalized coordinates, together with the equilibrium solution and the linearized equations of motion. Next, the Kalman-based estimation approach is detailed in Section III. This is followed by the stability analysis and the derivation of the control laws in Section IV. Section V assesses the performance of the proposed strategy as well as the robustness through expansive flight experiments. This was accomplished by the onboard implementation of the devised estimation and control algorithms on two vision-less sub-100-g robots. Lastly, conclusion is given in Section VI.

II. LEADER-FOLLOWER SYSTEM DYNAMICS

We consider the transport of a point mass payload suspended by two robots in the leader-follower manner. For the follower, the IMU is the only sensing unit, and there is no communication between the two agents. For simplicity, the cables are assumed to be inelastic, massless, and always taut.

A. GENERALIZED COORDINATES AND DYNAMIC MODEL

The leader-payload-follower system is shown in Fig 1 with the inertia frame \{x_l, y_l, z_l\}. The body-fixed frame of the follower is \{x_f, y_f, z_f\}. Defining the leader’s position in the global frame as \( p = [p_1, p_2, p_3]^T \), the position of the payload relative to the leader as seen in the inertial frame is vector \( l_p \). The location of the follower with respect to the payload as seen in the inertial frame is \( l_f \). Since the cables are assumed always taut, this brings two associated holonomic constraints

\[
I_p^T l_p - l_p^2 = 0 \quad \text{and} \quad I_f^T l_f - l_f^2 = 0, \tag{1}
\]

where \( l_p = |l_p| \) and \( l_f = |l_f| \) are the tethers’ lengths.

First, focusing on the relative positions of the payload and the follower to the leader. With two constraints above, this leaves four degrees of freedom (DoF) required to describe the two relative positions. We opt to employ the following generalized coordinates \( q \in \mathbb{R}^4 \)

\[
q = [q_1^T \quad q_2^T \quad q_3^T \quad q_4^T]^T, \quad q_1 = [\phi_p \quad \psi_p]^T, \quad q_2 = [\theta_p \quad \psi_f]^T. \tag{2}
\]

As illustrated in Fig 1, two auxiliary planes and one coordinate frame are constructed. The auxiliary frame \{x_a, y_a, z_a\} is located at the intersection of the vector \( l_f \) and a horizontal line that passes through the leader, with \( y_p \) pointing towards the leader and \( z_a \) pointing vertically up. The first plane (in pale yellow) is \( x_a-z_a \). The second plane (in light blue), normal to \( l_p \times l_f \), passes through the leader. The angle \( \theta_p \) is the angle between the vertical and the intersection between the two planes, measured about \( y_a \). Two in-plane angles, \( \phi_p \) and \( \phi_f \) (measured about \(-l_p \times l_f \)), such that \( \phi_p \) is negative and \( \phi_f \) is positive, are defined to partially describe \( l_p \) and \( l_f \) as detailed in Fig 1. Lastly, the angle \( \psi_f \) (measured about \( z_a \)) presents the location of the intersection of \( y_p \) and cable \( l_f \) with respect to the leader after projected onto the \( x_a-y_a \) plane. Together, we collectively refer to \( \phi_p \) and \( \phi_f \) as in-plane angles \( q_1 \) and \( \theta_p \) and \( \psi_f \) as out-of-plane angles \( q_4 \) for convenience.

Based on these definitions, the translational dynamics of the payload and the follower of masses \( m_p \) and \( m_f \) can be derived with the following Lagrangian

\[
L(q, \dot{q}) = T(q, \dot{q}) - U(q), \tag{3}
\]

in which

\[
T = \frac{1}{2} m_p (\dot{l}_p + \dot{\rho})^T (\dot{l}_p + \dot{\rho}) + \frac{1}{2} m_f (\dot{l}_f + \dot{\rho} + \dot{l}_f)^T (\dot{l}_f + \dot{\rho} + \dot{l}_f),
\]

\[
U = e_3^T (m_p g(p + l_p) + m_f g(p + l_p + l_f)), \tag{4}
\]

where

\[
l_p = \begin{bmatrix}
-S(\phi_p)c(\psi_f) - c(\phi_p)S(\phi_f)c(\psi_f) \\
c(\phi_p)c(\psi_f) - c(\phi_f)S(\phi_p)c(\psi_f) \\
-S(\phi_p)c(\psi_f) - c(\phi_f)S(\phi_p)c(\psi_f) \\
c(\phi_p)c(\psi_f) - c(\phi_f)S(\phi_p)c(\psi_f)
\end{bmatrix} l_p, \tag{5}
\]

\[
l_f = \begin{bmatrix}
-c(\phi_f)c(\psi_f) - S(\phi_f)c(\phi_p)c(\psi_f) - S(\phi_f)S(\phi_p)c(\psi_f) \\
c(\phi_p)c(\psi_f)S(\phi_f) + c(\phi_f)c(\phi_p)c(\psi_f) \\
-c(\phi_p)c(\psi_f) - S(\phi_p)c(\phi_f)c(\psi_f) \\
c(\phi_p)c(\psi_f)c(\phi_f) - c(\phi_p)c(\phi_f)c(\psi_f)
\end{bmatrix} l_f, \tag{6}
\]

and \( e_3 = [0, 0, 1]^T \). Solving the Euler-Lagrange equation yields the equations of motion

\[
H(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) = \tau_f + \tau_l, \tag{7}
\]
in which \( \tau_f = B_f(q)f_f \) and \( \tau_l = B_l(q)\ddot{q} \) are the generalized torque attributed to the collective thrust \( f_f = [f_{f1}, f_{f2}, f_{f3}]^T \) of the follower robot (see Fig. 1), and the acceleration of the leader \( \ddot{p} \) is evaluated from D’Alembert’s principle of virtual work. Complete expressions of \( H, C, G, B_f, \) and \( B_l \) are provided in the Supplemental Materials. Eqs. (S2)-(S11). Furthermore, the follower’s thrust \( f_f \) can be written as \( f_f = f_f R_f e_3 \), where \( R_f \) is the rotation matrix describing the attitude of the follower and \( f_f = |f_f|/l_f R_f \) can be represented with ZYX Euler angles as \( R_f = R_\psi(\psi)R_\theta(\theta)R_\theta(\phi) \). Since the robot’s yaw angle can be separately controlled, without loss of generality, we assume the yaw angle \( \psi = 0 \), or the \( x_f \) and \( x_a \) axes are parallel. Hence, \( R_f = R_\theta(\theta)R_\theta(\phi) \). In cases where \( \psi \neq 0 \), the follower’s frame can be re-aligned.

B. LINEARIZED DYNAMICS

To gain some insights into the system’s dynamics for the estimation task, the dynamics of the payload and the follower near the nominal equilibrium state (denoted by the superscript \( \bullet \)), defined as \( \dot{q} = 0 \) (or \( q = q^* \)) and \( \ddot{p} = 0 \) are considered.

1) EQUILIBRIUM SOLUTION

In equilibrium, the equations of motion (Eq. (7)) reduces to
\[
G^* q^* = \dot{q}^* q^* - B_f^* f_f^*.
\]
(8)

It can be shown that the solution is obtained when \( \dot{q}^* = 0 \). Under such condition, one may choose any \( \phi^*_p \in (-\frac{\pi}{2}, 0) \), \( \phi^*_q \in (0, \frac{\pi}{2}) \), and \( \psi^*_f \in (-\frac{\pi}{2}, \frac{\pi}{2}) \) as the nominal state. It is then straightforward to solve Eq. (8) for \( f_f^* \) as provided by Eqs. (S12)-(S14) in the Supplemental Materials.

2) LINEARIZED EQUATIONS OF MOTION

Based on the equilibrium state, the dynamics of the system at small deviation \( \delta q \) from \( q^* \) can be obtained by linearizing Eq. (7) as
\[
H^* \delta q + D^* \delta q = B_f^* \dot{f}_f + B_l^* \ddot{p},
\]
(9)

where \( H^* = H(q^*), D^* = \frac{\partial H(q) q^*}{\partial q} - \frac{\partial \tau_f(q)}{\partial q} q^* \), \( B_f^* = B_f(q^*) \), and \( B_l^* = B_l(q^*) \) as \( C(q^*, \dot{q}^*) = 0 \). It turns out that \( H^* \) and \( G^* \) are block diagonal matrices. In other words, they take the following forms:
\[
H^* = \begin{bmatrix}
    H_f^* & 0_{2 \times 2} \\
    0_{2 \times 2} & H_O^*
\end{bmatrix}, \quad D^* = \begin{bmatrix}
    D_f^* & 0_{2 \times 2} \\
    0_{2 \times 2} & D_O^*
\end{bmatrix},
\]
(10)

where the definitions of \( (\cdot)^* \)’s and \( (\cdot)_O \)’s are provided in the Supplemental Materials (Eqs. (S15)-(S17)). The outcome of the linearization process implies that the in-plane and out-of-plane dynamics are only loosely coupled and can be approximately separated for the estimation and control purposes to reduce the complexity of the analysis.

III. EKF-BASED ESTIMATION VIA INERTIAL SENSING

In order to for the follower robot to stabilize its position with respect to the leader, it first must estimate its relative location, or \( q \). In this work, we emphasize on a lightweight and efficient solution. Therefore, this is achieved with only the IMU measurements from the follower itself. The efficiency of the estimation strategy is boosted through two avenues. First, the estimation loops for in-plane and out-of-plane dynamics are separated to reduce the dimension of the state vectors. This is motivated by the loose coupling between \( q_f \) and \( q_o \) as suggested by the linearized dynamics above. Next, a variant of a Kalman filter akin to an Extended Kalman Filter (EKF) is proposed in order to yield the estimation accuracy parallel to that of an Unscented Kalman Filter (UKF) at the computational demand of an EKF.

A. MEASUREMENT MODEL

Prior to presenting the estimation framework, we define the measurements. Using only the IMU feedback, the accelerometer’s reading \( a = [a_1, a_2, a_3]^T \) is taken as the output as \( y = a \). Since \( a \) is the specific acceleration of the follower in its body frame (see Fig. 1), it is
\[
a = -(f_f/m_f)e_3 - (1/m_f)R_f^Tf_f,
\]
(11)

where the vector \( f_f = [f_{f1}, f_{f2}, f_{f3}]^T \) is the tension of cable \( l_f \). With the pre-determined mapping of the input commands and collective thrust, \( f_f = [f_f] \) is assumed known. Moreover, a conventional sensor fusion algorithm is capable of robustly estimating \( R_f \). This leaves \( f_f \) that relates \( a \) to the state of the robot \( q \).

To express \( f_f \) in terms of the state \( q \) and \( \dot{q} \), a similar approach to the derivation of the system’s equations of motion is applied. In short, we consider the leader-payload sub-system, of which the external non-conservative force is \( f_f \). The derivation and full expression of \( f_f \) is provided in the Supplemental Materials as Eqs. (S20) and (S24).

B. DECOUPLED STATES

Conventionally, a Kalman-based estimator designed for estimation of \( q \in \mathbb{R}^4 \) and \( \dot{p} \in \mathbb{R}^3 \) would require a state vector \( x = [q^*, \dot{q}^*, \ddot{p}]^T \in \mathbb{R}^{11} \). The real-time evaluation of the state prediction step \( (\hat{x} = f(x,u)) \) and the Jacobian \( (\partial f/\partial x) \in \mathbb{R}^{11 \times 11} \) turns out to be a challenge for small MAVs without a companion computer. This is particularly true for a UKF, as in each iteration step, the dynamics \( f(x,u) \) must be evaluated multiple times depending on the number sigma points considered. To alleviate the computational burden, we leverage the fact that the in-plane and out-of-plane dynamics are loosely coupled as found earlier to separate the state estimation to two staged tasks.

To facilitate the decoupling of the states \( q_l \) and \( q_o \) for the estimation purpose, we employ an additional simplifying assumption that \( \psi_f \) is nominally slowly time-varying. That is the desired trajectory of \( \psi_f(t) \) does not change rapidly. This permits us to presume \( \psi_f^* = 0 \) (or \( \psi_f \approx 0 \)) by re-defining the direction of the inertia frame \( (x_w,y_w,z_w) \) such that the vector \( l_f + l_f \) is nominally in the \( y_w-z_w \) plane (see Fig. 1). This process reduces the measurement model and aids the estimation of \( \ddot{p} \) as described below.
1) IN-PLANE STATE

For the estimation of $q_1$, it is assumed that the out-of-plane dynamics is in equilibrium, or $\theta_p, \dot{\theta}_p, \psi_f, \dot{\psi}_f = 0$. As a result, we may deduce the reduced equations of motion from Eq. (7) as

$$\vec{H}_1 \ddot{q}_1 + \vec{C}_1 \dot{q}_1 + \vec{G}_1 = \vec{B}_{f,1} \begin{bmatrix} f_{f,2}, f_{f,3} \end{bmatrix}^T + \vec{B}_{l,1} \begin{bmatrix} \ddot{p}_2, \ddot{p}_3 \end{bmatrix}^T,$$

(12)

where $\vec{H}_1(q_1)$ and $\vec{C}_1(q_1, \dot{q}_1)$ are the first 2 \times 2 elements of $\vec{H}$ and $\vec{C}$ evaluated at $q_o, \dot{q}_o = 0$. Similarly, $\vec{G}_1(q_1), \vec{B}_{f,1}(q_1) \in \mathbb{R}^{2 \times 2}$, and $\vec{B}_{l,1}(q_1) \in \mathbb{R}^{2 \times 2}$ are taken from the top two rows of $\vec{G}$, $\vec{B}_f$, and $\vec{B}_l$ evaluated under the same condition (see Eqs. (S25)-(S27)). Note that approximation $\psi_f \approx 0$ renders $f_{f,1}$ and $\ddot{p}_1$ irrelevant to the in-plane dynamics.

Eq. (12) captures the simplified nonlinear dynamics of $q_1$ and can be employed in the state prediction step of a Kalman-based filter. For the prediction and update of the covariances, one needs to acquire the linearized version of Eq (12) (about the current $q_1$ state). This is partially obtained via Eq. (9) as

$$\vec{H}_1 \delta q_1 + \vec{D}_1 \delta q_1 = \vec{B}_{f,1} \begin{bmatrix} \delta f_{f,2}, \delta f_{f,3} \end{bmatrix}^T + \vec{B}_{l,1} \begin{bmatrix} \ddot{p}_2, \ddot{p}_3 \end{bmatrix}^T,$$

(13)

where $\vec{D}_1(q_1) = \vec{D}_{f,1}^T \begin{bmatrix} q_1, q_1 \end{bmatrix} = 0$. As a consequence, we may obtain the state-space representation of the linearized in-plane motion from Eq. (13) as

$$x_1 = \begin{bmatrix} \delta q_1^T, \ddot{p}_2, \ddot{p}_3 \end{bmatrix}^T, \quad y_1 = \begin{bmatrix} a_2 & a_3 \end{bmatrix} \begin{bmatrix} x_1 \end{bmatrix}, \quad \dot{x}_1 = \vec{A}_1 x_1 + \vec{B}_1 u_1, \quad u_1 = \begin{bmatrix} \delta f_{f,2}, \delta f_{f,3} \end{bmatrix}^T,$$

(14)

where the state vector includes the unknown leader’s acceleration, $\ddot{p}_2, \ddot{p}_3$. Remark that $\ddot{p}_2, \ddot{p}_3$ are treated as constants ($\ddot{p}_3/dt, \ddot{p}_3/dt = 0$) to be estimated. The full expressions of $\vec{A}_1$ and $\vec{B}_1$ are provided by Eqs. (S29)-(S30). The Jacobian $\vec{C}_1$ is determined from Eq. (11). Again, when assuming $\theta_p, \psi_f = 0$, the measurement model from Eq. (11) indicates that $a_1$ is not related to $q_1$ and, therefore, omitted from the output vector $y_1$. This is reasonable as in this case, the direction of $a_1$ coincides with the out-of-plane direction.

2) OUT-OF-PLANE STATE

For the out-of-plane analysis, the static assumption on the in-plane state is used: $\phi_p, \phi_f = 0$. The reduced nonlinear dynamics taken from Eq. (7) becomes

$$\vec{H}_0 \ddot{q}_o + \vec{C}_o \dot{q}_o + \vec{G}_o = \vec{B}_{o,f} f_f + \vec{B}_{o,l} \dot{p},$$

(15)

where $\vec{H}_o(q_1, q_o)$ and $\vec{C}_o(q_1, q_o, \dot{q}_o)$ are taken from the last 2 \times 2 elements of $\vec{H}(q)$ and $\vec{C}(q, \dot{q})|_{q_1=0}$. Meanwhile, $\vec{G}_o(q_1), \vec{B}_{o,f}(q_1) \in \mathbb{R}^{2 \times 3}$, and $\vec{B}_{o,l}(q_1) \in \mathbb{R}^{2 \times 3}$ are from the bottom two rows of $\vec{G}$, $\vec{B}_f$, and $\vec{B}_l$ (see Eqs. (S33) and (S34)). Unlike Eq. (12), the simplified out-of-plane dynamics is influenced by all elements of $f_f$ and $\dot{p}$. This resultant equation is to be used for the state prediction of the proposed estimator.

The linearization of Eq. (15) is evaluated, about the state $q_1, q_o$ with $\dot{q}_1 = 0$, partially using the result from Eq. (9) as

$$\vec{H}_o \delta q_o + \vec{D}_o \delta q_o = \vec{B}_{o,f} \begin{bmatrix} \delta f_f \end{bmatrix} + \vec{B}_{o,l} \begin{bmatrix} \ddot{p}_1, \ddot{p}_2, \ddot{p}_3 \end{bmatrix}^T,$$

(16)

where $\vec{D}_o(q_1, q_o) = \vec{D}_{o,1}^T [q_1, q_o]$. From here, Eq. (16) can be re-arranged into the state space form:

$$x_o = \begin{bmatrix} \delta q_o^T, \delta \dot{q}_o^T, \ddot{p}_1, \ddot{p}_2, \ddot{p}_3 \end{bmatrix}^T, \quad y_o = \begin{bmatrix} a_1 \end{bmatrix} = \vec{C}_o x_o.$$

$$\dot{x}_o = \vec{A}_o x_o + \vec{B}_o u_o. \quad u_o = \begin{bmatrix} \delta f_f, \ddot{p}_2, \ddot{p}_3 \end{bmatrix}^T,$$

(17)

where it can be seen that only $\ddot{p}_1$ is included in $x_o$ as other elements of $\dot{p}$ are treated as the known system’s inputs. In this case, only $a_1$ is taken as the measurement as it is anticipated to reflect most prominently to the change in $\theta_p$ and $\psi_f$. See Eqs. (S36) and (S37) for the breakdown of $\vec{A}_o$ and $\vec{B}_o$. The corresponding Jacobian $\vec{C}_o$ can be computed from the measurement model or Eq. (11).

The concept of decoupled dynamics and their associated state-space equations derived here form the basis of the efficient two-stage EKF-based estimation strategy outlined below.

C. OBSERVABILITY

Prior to deriving the observers, we investigate the observability of the in-plane and out-of-plane subsystems based on the linearized dynamics captured by Eqs. (14) and (17).

To numerically evaluate the observability of the in-plane state $x_1$, we construct the observability matrix from the pair of $(\vec{A}_1, \vec{C}_1)$ and numerically determine the rank for all pairs of $\theta_p$ and $\psi_f$ with $\phi_p$ ranging from $-90^\circ$ to $0^\circ$ and $\phi_f$ ranging from $0^\circ$ to $90^\circ$ in increments of $1^\circ$. The results reveal that the observability matrix is full rank except when $\phi_p = -90^\circ, 0^\circ$ or $\phi_f = 0^\circ, 90^\circ$. This suggests the in-plane state $x_1$ is observable for those physically feasible configurations.

Similarly, for estimating the out-of-plane state, the observability is determined by the pair of $(\vec{A}_o, \vec{C}_o)$. We tested the rank condition for all the possibilities of $\phi_p, \phi_f, \theta_p$ and $\psi_f$ ranging from $-90^\circ$ to $0^\circ$ for $\phi_p, 0^\circ$ to $90^\circ$ for $\phi_f, 0^\circ$ to $90^\circ$ for $\theta_p$, and $0^\circ$ to $90^\circ$ for $\psi_f$ (in increments of $1^\circ$). Similar to the in-plane dynamics, the rank condition is satisfied except at the extreme angles, suggesting that the out-of-plane state is also observable for realistic configurations.

D. EKF-BASED ESTIMATORS

To estimate the state vectors $x_1$ and $x_o$, which include the acceleration of the leader, from the IMU reading, an EKF-like filter is proposed. The preliminary study reveals that an EKF is unable to robustly estimate the states and often results in divergence. This is likely caused by the nonlinearity of the dynamics. On the other hand, UKF was found to be reliable and accurate in an offline implementation, but its relative complexity impedes a successful onboard deployment.

There exist two major differences between EKFs and UKFs, both underpinning the propagation of the covariance. While the EKF relies on the linearized dynamics for the propagation, the unscented transformation in the UKF employs full nonlinear models for the sampling of sigma points around the mean [34]. While resulting in more accurate estimates of the covariance, the use of nonlinear model could be costly for systems with large dimensions as more sample
points are required, especially when it involves the evaluation of trigonometric or other complex functions. Moreover, the UKF incorporates tuning factors to control the spread of the sigma points [34], effectively re-scaling the covariance in each propagation. These two steps underline the advantage of UKF.

The proposed estimator is designed to alleviate the inaccuracy from the linearization in the covariance propagation of the EKF. To do so, we integrate a tuning parameter that controls the spread of the covariance, equivalent to the tuning parameter for the sigma points in the UKF [34], while retaining the linearization process in order to minimize the computational demand. The scaling factor likely mitigates the impact of the scented covariance propagation, improving the accuracy.

1) FORMULATION
The process of estimating $x_I$ and $x_O$ follows closely that of a standard EKF. The dynamics of $x_I$ and $x_O$ are discretized with the Euler method. Let the subscript $k$ denote the time index, the subscript $j = \{I, O\}$ is used to collectively refer to in-plane or out-of-plane variables. The models for the estimation are

$$
\begin{align*}
x_{j,k} &= x_{j,k-1} + f_j(x_{j,k-1}, u_{j,k-1})T + w_{j,k}, \\
y_{j,k} &= h_j(x_{j,k}, u_{j,k}) + v_{j,k},
\end{align*}
$$

where $T$ is a sample time, $w_{j,k}$ and $v_{j,k}$ are zero-mean Gaussian white noises with the associated covariance matrices $Q_{j,k}$ and $R_{j,k}$. The nonlinear models $f_j(\cdot)$ and $h_j(\cdot)$ are derived from the simplified nonlinear equations of motion obtained during the introduction of the decoupled dynamics earlier. For $f_j$, Eq. (12) is applied (assuming $\theta_p, \psi_f = 0$, see also Eq. (S38)-(S39)). For $f_O$, it has been assumed that $\phi_p, \phi_f = 0$ in the derivation of Eq. (15) (see Eq. (S40)-(S41)).

The measurement models $h_j$ are taken from Eq. (11).

In the following prediction and update steps, the hat symbol is used to denote a priori $\hat{x}_{j,k}$ and a posteriori estimates $\hat{x}_{j,k}$.

2) PREDICTION
Letting $P_{x_{j,k}}$ be the estimate of the covariance, the prediction step follows

$$
\begin{align*}
\hat{x}_{j,k}^+ &= \hat{x}_{j,k-1}^+ + f_j(\hat{x}_{j,k-1}^+, u_{j,k-1})T, \\
P_{x_{j,k}}^- &= \lambda \mathcal{A}_{j,k}\hat{x}_{j,k-1}^+ \mathcal{A}_{j,k}^T + Q_{j,k},
\end{align*}
$$

where $\lambda$ is the proposed scalar tuning parameter, taking after the scaling factor in UKF [34]. The state Jacobian $\mathcal{A}_{j,k}$ is either taken from Eq. (14) or (17).

3) UPDATE
Following the UKF framework, the update of the innovation covariance $P_{y_{j,k}}$ and Kalman gain $K_{j,k}$ must also include $\lambda$ as

$$
\begin{align*}
P_{y_{j,k}}^- &= \lambda \mathcal{C}_{j,k} P_{x_{j,k}}^- \mathcal{C}_{j,k}^T + R_{j,k}, \\
K_{j,k} &= \lambda P_{x_{j,k}}^- \mathcal{C}_{j,k}^T P_{y_{j,k}}^- 1,
\end{align*}
$$

where $\mathcal{C}_{j,k}$ is the innovation Jacobian taken from Eq. (14) or (17). As a consequence, the a posteriori state and covariance estimates are

$$
\begin{align*}
\hat{x}_{j,k}^- &= \hat{x}_{j,k}^+ + K_{j,k} (y_{j,k} - h_j(x_{j,k}, u_{j,k})), \\
P_{x_{j,k}}^- &= P_{x_{j,k}}^- - \lambda K_{j,k}\mathcal{C}_{j,k} P_{x_{j,k}}^- 1.
\end{align*}
$$

When $\lambda = 1$, the estimator becomes an EKF. $\lambda > 1$ or $\lambda < 1$, therefore, enlarges or compresses the updated covariances as they are propagated through the linearization process.

4) TWO-STAGE IMPLEMENTATION
In the formulation of the EKF-based estimator for in-plane and out-of-plane states provided by Eq. (18), $f_j$ is obtained by imposing $\theta_p, \psi_f = 0$. This implies the in-plane state is completely independent of $x_O$. Nevertheless, the evaluation of $f_O$ (as well as $A_O, B_O$, and $C_O$) presumes $\phi_p, \phi_f = 0$, but still depends on actual $\phi_p$ and $\phi_f$ as well as $\dot{q}_2$ and $\dot{p}_3$. Thus, $\dot{x}_I$ is used in the estimation of $\dot{x}_O$ as schematically shown in Fig. 2(a). Due to these reasons, two estimation routines are executed in a cascaded fashion, rather than in parallel.

To sum up, as captured by Eqs. (12)-(14), the first routine corresponding to the in-plane dynamics provides the estimates of $q_1, \dot{q}_1, \dot{p}_2, \dot{p}_3$ while assuming $\theta_p, \psi_f = 0$. Then, the second estimator, for the out-of-plane motion, leverages $\ddot{q}_1$ and $\ddot{p}_2, \ddot{p}_3$ in order to produce $\ddot{q}_O, \ddot{q}_O, \ddot{p}_1$. This two-stage strategy turns out to be highly robust and efficient in the flight experiments presented in Section V.

IV. CONTROL AND STABILITY
In this section, the objective is to determine the system’s input, in the form of the force generated by the follower robot $f_j$, that stabilizes the system to the nominal equilibrium state $q^*$ based on the estimates $\{\ddot{q}, \ddot{q}, \ddot{p}\}$ from the preceding section. This implies that the relative position between the two robots is controlled regardless of the motion of the leader.

The challenges in controlling the leader-payload-follower system stem from the nonlinear and underactuated dynamics.

![FIGURE 2. (a) Block diagrams detailing the two-stage EKF-based estimation strategy. (b) The architecture of the implemented estimation and control framework onboard the follower robot.](image)
Herein, we propose a strategy based on the outlined decoupled linear dynamics. As shown below, neglecting aerodynamic drag, the linearized dynamics are marginally stable in open-loop. This suggests a simple linear controller, previously proven effective for some cooperative transport methods [18], [26], has potential to stabilize the system adequately. The linear approach benefits from the efficiency and proven local stability. The low computational cost is advantageous for small robots with limited power. Despite the linearization, the method yields robust results as demonstrated in Section V.

A. LINEAR STABILITY ANALYSIS

To begin, we consider the decoupled dynamic equations from Section II-B2 to inspect their the open-loop stability before presenting the proposed control laws.

1) OPEN-LOOP STABILITY

Recall the dynamics of the two subsystems from Eqs. (13) and (16), when they are linearized about the nominal equilibrium point \( \mathbf{q}^* \) with \( \psi^* = 0 \) and subject to no additional inputs \( \delta f_j, \delta p \)

\[
\begin{align*}
\tilde{H}_I^e \delta \mathbf{q}_I + \tilde{D}_I^e \delta \mathbf{q}_I &= 0, \\
\tilde{H}_O^e \delta \mathbf{q}_O + \tilde{D}_O^e \delta \mathbf{q}_O &= 0,
\end{align*}
\]

where the superscript \( * \) indicates the fixed linearization point at \( \mathbf{q}^* \). It is known that both \( \tilde{H}_I^e \) and \( \tilde{H}_O^e \) are positive definite as they represent the systems’ inertia.

A closer investigation reveals that both \( \tilde{D}_I^e \) and \( \tilde{D}_O^e \) are also guaranteed to be positive definite for the nominal conditions \( \phi^*_p \in (-\frac{\pi}{2}, 0), \phi^*_O \in (0, \frac{\pi}{2}) \), and \( \mathbf{q}^*_0 = \mathbf{0} \) (see Eqs. (S42)-(S46) in the Supplemental Materials). This implies that both in-plane and out-of-plane motions, when subject to the nominal input \( \mathbf{f}^* \), are marginally stable (see Eqs. (S47)-(S50) for a complete proof). That is, their behaviors resemble those of an undamped mass-spring system. This is important for control and stability, as it implies that, for instance, only an introduction of a damping force would adequately stabilize the system, or a simple and efficient control strategy would be sufficient to robustly bring the system to the equilibrium and as the underlying dynamics are not inherently unstable. As a consequence, the linear control laws based on linearized dynamics are proposed as described below.

B. CLOSED-LOOP CONTROL

1) IN-PLANE MOTION

As described by Eq. (13), the in-plane motion can be approximately regarded as a two-dimensional second order system with two inputs. This suggests that the in-plane dynamics is fully actuated in state \( \mathbf{x} = [\mathbf{q}_I^T, \dot{\mathbf{q}}_I]^T \). Since \( \mathbf{B}_{I,f} \) (and \( \mathbf{B}_{I,f}(\mathbf{q}^*) \)) is invertible for \( \sin(\phi_p - \phi_f) \neq 0 \) (which is always true for the operating range of \( \phi_p \) and \( \phi_f \)), the following control law is employed to compute \( [\delta f_{j,2}, \delta f_{j,3}]^T \):

\[
\mathbf{B}_{I,f}^e \begin{bmatrix} \delta f_{j,2} \\ \delta f_{j,3} \end{bmatrix}^T = -\tilde{H}_I^e \begin{bmatrix} \mathbf{K}_{I,p} \delta \mathbf{q}_I + \mathbf{K}_{I,i} \int \delta \mathbf{q}_I \, dt \\ \mathbf{K}_{I,d} \delta \mathbf{q}_I \end{bmatrix} - \mathbf{B}_{I,f} \begin{bmatrix} \tilde{p}_2 \\ \tilde{p}_3 \end{bmatrix}^T,
\]

where \( \mathbf{K}_{I,*} \)'s are diagonal positive gain matrices. Neglecting the difference between the actual and estimated quantities, the closed-loop dynamics linearized about \( \mathbf{q}^* \) becomes

\[
\begin{align*}
\delta \dot{\mathbf{q}}_I + \mathbf{K}_{I,d} \delta \mathbf{q}_I + \left( \mathbf{K}_{I,p} + \tilde{H}_I^e \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix} \right) \delta \mathbf{q}_I + \mathbf{K}_{I,i} \int \delta \mathbf{q}_I \, dt \\
&= 0,
\end{align*}
\]

and the stability is guaranteed as shown by Eqs. (S51)-(S53). The inclusion of the integral terms alleviates the errors caused by inaccurate parameters.

2) OUT-OF-PLANE MOTION

Unlike the in-plane dynamics, as captured by Eq. (15), the out-of-plane motion is subject to only a single input \( f_{j,1} \) such that \( \text{rank}(\mathbf{B}_{O,f}) = 1 \). This renders the out-of-plane dynamics underactuated. The stability of the out-of-plane dynamics cannot be achieved in the same manner to that of the in-plane dynamics.

To obtain some insights into the closed-loop stability of the system, we re-write Eq. (16) in the state-space form as

\[
\begin{align*}
\mathbf{V} \begin{bmatrix} \delta \mathbf{q}_O \\ \delta \mathbf{q}_O \end{bmatrix} &= \begin{bmatrix} 0_{2 \times 2} & \mathbf{I}_{2 \times 2} \\ -\tilde{D}_O & 0_{2 \times 2} \end{bmatrix} \begin{bmatrix} \delta \mathbf{q}_O \\ \delta \mathbf{q}_O \end{bmatrix} + \mathbf{B}_{O,f} \delta f_{j,1},
\end{align*}
\]

where we have neglected the term with \( \tilde{p}_1 \) as disturbances,

\[
\mathbf{V} = \begin{bmatrix} \mathbf{I}_{2 \times 2} & 0_{2 \times 2} \\ 0_{2 \times 2} & \tilde{H}_O^e \end{bmatrix},
\]

is nonsingular, and

\[
\delta f_{j,1} = -\mathbf{B}_{O,f}^T \left( k_{O,p} \mathbf{q}_O + k_{O,d} \dot{\mathbf{q}}_O \right)
\]

is the proposed control law for positive scalar gains \( k_{O,p} \) and \( k_{O,d} \). Defining a projected state variable \( \tilde{x} = \mathbf{V}^{-1} [\delta \mathbf{q}_O^T, \dot{\delta \mathbf{q}}_O^T]^T \), the closed-loop expression of Eq. (25) reduces to

\[
\dot{\tilde{x}} = \begin{bmatrix} -\tilde{D}_O - \mathbf{B}_{O,f} \mathbf{B}_{O,f}^T k_{O,p} & -\mathbf{B}_{O,f} \mathbf{B}_{O,f}^T k_{O,d} \end{bmatrix} \mathbf{V}^{-1} \tilde{x} = \tilde{\mathbf{A}} \mathbf{V}^{-1} \tilde{x}.
\]

Since \( \tilde{H}_O^e \) and therefore, \( \mathbf{V}^{-1} \), is positive definite, the stability of the closed-loop dynamics is determined by the negative definiteness of \( \tilde{\mathbf{A}} \). In Section IV-B of the Supplemental Materials, we prove that this is readily attained. In practice, the integral term could be incorporated into Eq. (27) to further attenuate steady-state errors.

Since the control strategy leverages the fact that the in-plane and out-of-plane dynamics are decoupled for small deviations from the nominal state \( \mathbf{q}^* \), the method yields robust results as demonstrated in Section V.

V. EXPERIMENTAL EVALUATIONS

To assess the effectiveness of proposed methods, flight experiments were first conducted with only the follower robot. By replacing the leader robot with a stationary point, the impact of the leader’s acceleration is eliminated, allowing
the inherent stability and dynamics of the system to be thoroughly assessed. Then, both indoor and outdoor flights of the complete leader-payload-follower system were demonstrated. The list of experiments conducted are presented in Table 1.

<table>
<thead>
<tr>
<th>Experiments</th>
<th>Objectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>§V-C1: Open-loop experiments</td>
<td>To show the marginal stability of the open-loop system and validate the estimator</td>
</tr>
<tr>
<td>§V-C2: Flights with MOCAP feedback</td>
<td>To validate the controller</td>
</tr>
<tr>
<td>§V-C3: Flight with onboard estimation and control</td>
<td>To validate the control with the onboard feedback</td>
</tr>
<tr>
<td>§V-C4: Leader-payload-follower flights</td>
<td>To validate the estimation and control method with two robots via trajectory tracking flights</td>
</tr>
<tr>
<td>§V-C5: Flights with varying setpoints angles</td>
<td>To test the method with different setpoint angles and payload masses</td>
</tr>
<tr>
<td>§V-D: Disturbance rejection experiments</td>
<td>To test the robustness of the method against disturbances</td>
</tr>
<tr>
<td>§V-E: Outdoor flight demonstration</td>
<td>To demonstrate the aerial cooperative transport outdoor</td>
</tr>
</tbody>
</table>

A. EXPERIMENTAL SETUP

Indoor experiments were carried out in the arena with motion capture (MOCAP) cameras (NaturalPoint, OptiTrack Prime 13w) for ground truths. Two Mambo Minidrones (Parrot SA) were used as leader and follower robots. The inertia-based framework does not require vision or robot-to-robot communication, allowing small platforms to be used. The robot’s mass, including MOCAP markers, is 73 g. A dummy payload with markers was suspended by inelastic fishing lines (0.35 mm in diameter). The mass of the payload used, except for the flights in Section V-C5, was 30 g, significant in relation to the robot’s weight and thrust limit. The cable lengths (l₀ and l₁) were 0.73 m for all experiments. The estimation and control algorithms were implemented onboard with Simulink Support Package for Parrot Minidrones (MathWorks) and executed at 200 Hz. The support package enables low-level access to IMU data and control of the motor commands as needed by the estimation and control algorithms. Flight data were logged.

Prior to the flights, a benchtop experiment was performed to identify the mapping between the commands and the collective thrust generated by the robots. This is to obtain the value of \( f_f \) used in both the estimators and controllers using a loadcell (Nano17, ATI). The details of the procedure can be found in [5]. The collective thrust is modeled as a linear function of the motor commands and the battery voltage. The model coefficients were identified using N4SID [35].

B. ONBOARD IMPLEMENTATION

The architecture of the estimator and controller onboard of the follower robot is shown in Fig. 2(b). The state feedback is provided by the estimator. The flight controller computes the desired thrust \( f_f^* \) and desired roll and pitch angles. The low-level controller is responsible for regulating the thrust and attitude of the follower to realize the desired \( f_f \).

During flights, the in-plane setpoints were mostly chosen as \( \phi_p^* = -40^\circ \) and \( \phi_f^* = 40^\circ \) to strike the balance between efficiency (favoring small angles for reduced horizontal forces) and robustness (preferring large angles to avoid collisions). As outlined in Section II, the setpoint for out-of-plane angles are the linearized point: \( \theta_p^* = \psi_f^* = 0 \). The implemented controller did not compensate for the estimated \( \mathbf{p} \) as it was preliminary found to be insignificant. Controller gains were experimentally tuned. Each flight contained a total of 65 s of flying time, including 5 s for taking off and landing at the start and the end. Example flight videos are available as Supplemental Materials.

C. INDOOR EXPERIMENTS

We carried out several sets of indoor flights with MOCAP to separately assess the estimator and the controller. The flights were initially performed without a leader to eliminate the impact of \( \mathbf{p} \). This is achieved by substituting the leader robot with a fixed structure. Later, another Mambo minidrone was employed as the leader to showcase the performance of the estimator and controller together in a trajectory following flight. This mimics a real-world use where both robots travel over some distance.

1) OPEN-LOOP EXPERIMENTS WITH THE ESTIMATOR

At the beginning, open-loop flights were executed. This refers to the situation where the follower robot is commanded to produce the nominal thrust vector \( f_f^* \) (determined from Eq. (S14)), or all the controller gains \( K_{..} \)’s in Eqs. (23) and (27) are zero. This is to verify that the leader(payload)-follower system (in the absence of unmodeled dampings) is not stable even in the absence of the leader’s acceleration as suggested by Eq. (22). In the meantime, however, the EKF-based estimator can still be employed in such setting, despite a possible large deviation of the states from the nominal equilibrium point. Three flights were performed in open-loop without the leader. To emphasize the impact of initial conditions, we used a launch platform for the payload to produce a large initial \( \theta_p \) angle (see the supplementary video). An example of the recorded in-plane and out-of-plane angles, and their estimates, are shown in Fig. 3(a), S1, and S2. It can be seen that, without control, both in-plane angles oscillated around their nominal values with the amplitudes of \( \approx 20^\circ \). The pendulum-like oscillation, described by \( \theta_p \), slowly attenuated at the timescale of \( \approx 10 \) s, before reducing to \( \approx 10^\circ \). Meanwhile, \( \psi_f \) fluctuated with the amplitude of \( \approx 15^\circ \). Overall, oscillations were present in...
all modes, attesting the predicted marginal stability. Such behavior could destabilize the system when the leader robot is put in place. Statistically, the angular errors from all flights are summarized in Fig. 5(b) and S16(b). The trend is uniform among three flights.

Despite considerable deviations of the states from the equilibrium, the estimator was able to robustly provide accurate estimates as manifested by Fig. 3(a). The estimates of the angles and the rates follow the groundtruths closely. As seen in Fig. 5(a) and S16(a), the errors are generally below 5°, despite the large deviation of the states from the nominal conditions.

2) FLIGHTS WITH MOCAP FEEDBACK

In order to verify the performance of the proposed controller, in this set of experiments, three flights were conducted with closed-loop control without the leader, using the angular measurements from the MOCAP as feedback. This isolates the effects of the estimation from the system’s control and stability. The onboard estimator remained active to provide the estimated states for reference. The control laws, or Eqs (23) and (27), were used to compute $\delta f_f$ and translated into the setpoint attitude $R_f$ and thrust of the follower robot $f_f$.

The result of one representative flight is illustrated in Fig. 3(b), S3, and S4. It is evident that all angles quickly stabilized in the first 10 s, in sharp contrast with the open-loop flight (Fig. 3(a)). Fig. 5(b) and S16(b) confirm that angular and rate errors are markedly reduced from the open-loop flights. In this case, the estimation errors, as plotted in Fig. 5(a) and S16(a) are somewhat smaller when compared with the open-loop experiments, likely due to the controller that minimizes the deviation of the angles from the setpoints, mitigating the impact of the linearization used in the EKF-based observer. This suggests that the linearization of the dynamics does not substantially impact the performance of the estimator when it is concurrently deployed with the devised controller.

3) FLIGHTS WITH ONBOARD ESTIMATION AND CONTROL

Next, we executed three flights to test both the EKF-based estimator and controller together while the leader robot was still fixed.

An example trajectory and the estimates are shown in Fig. 3(c), S5, and S6. The flight results follow a similar trend observed in the MOCAP flights. The estimation and control errors were swiftly minimized within 10 s. As detailed in Fig. 5 and S16, the statistics of estimation errors are highly similar to those from flights controlled with MOCAP feedback. The state errors are slightly larger as the control here was based directly on the estimates instead of the MOCAP feedback. The results suggest that the complete onboard implementation does not significantly adversely affect the performance. The small amplification of the state errors is attributed to the estimation and control errors combined. Moreover, the small angle assumption made for the
the estimation errors of \( \psi_f \) are usually the largest among four angles owing to the poor onboard estimation of the yaw angle of the robot, which leads to the increased error in \( \psi_f \). However, the decrease in the tracking performance of \( \psi_f \) is not critical to the flight stability as the follower could achieve equilibrium at any value of \( \psi_f \).

5) FLIGHTS WITH VARYING SETPOINT ANGLES AND DIFFERENT PAYLOAD MASSES

To provide examples of flight with different configurations, two conditions were adjusted from the previous experiments. First, we changed the payload mass from 30 g to 25 g and 35 g (the model parameters for both the estimator and controller were adjusted accordingly). Second, we designed the reference \( \phi^*_p \) and \( \phi^*_f \) to alternate between \(-30^\circ\) and \(-50^\circ\) and \(30^\circ\) and \(50^\circ\), respectively. The trajectory of the leader was similar to the previous two-robot flights. The controller gains were kept unchanged from the previous tests.

The tracking results of two different payload weights are shown in Fig. 7, S9, and S10 for the 25-g payload, and Fig. 8, S11, and S12 for the 35-g payload. It can be seen that follower tracked the reference angles closely, with no significantly different behaviors compared with the previous sets of experiments.

D. DISTURBANCE REJECTION EXPERIMENT

To demonstrate robustness of the developed methods, we simulated the situations when an obstacle hits the follower robot or a cable. The two-robot hovering flight was conducted with a 30-gram payload. Similar to previous cooperative flights, MOCAP was only used for controlling the leader. Disturbances were injected at several occasions. At \( t \approx 9 \) s, the payload was pushed in the out-of-plane direction. After the stabilizing attempt, at \( t \approx 23 \) s, the payload was again pushed, but in the in-plane direction. This was followed by three pushes on the cable \( I_f \) to disturb the angle \( \phi_f \) (Fig. 6(b)).

The flight results plotted in Fig. 9 and S13 indicate that the system remained stable during the 60-second flight. The follower spent less than \( \approx 10 \) s to settle to the setpoints after each disturbance was injected. The angular errors were mostly bounded within \( 10^\circ \) of the setpoint after the recovery. The outcomes verify the robustness of the proposed method.

E. OUTDOOR FLIGHT DEMONSTRATION

To show the feasibility for real-world applications, we carried out outdoor tests. In such conditions, the robots are subject to wind disturbances. The effects are significant for smaller robots such as the 73-gram Minidrones used in our demonstration.

Without ground truth measurements from the MOCAP, an optical flow sensor (Bitcraze, Flow deck v2) with a breakout board (Bitcraze, Crazyflie 2.0 board) and a supporting 1s battery were affixed underneath the leader to measure the velocity of the robot to provide some data on the flight trajectory. This added 17 g to the leader. The leader was remotely operated by a human pilot. During the experiment,
FIGURE 6. Photographs showing the payload being transported by the leader and the follower. (a) In the indoor flight, the leader’s position was controlled via the motion capture feedback. (b) Snapshot from the disturbance rejection flight showing a human pulling the cable between the follower and the payload. The white lines highlight the disturbed cable. (c) Outdoor, the leader was remotely controlled by a human pilot. The follower stabilizes the cables’ angles through its onboard estimation and control algorithms relying only on its IMU.

FIGURE 7. Leader-payload-follower flights with varying setpoint angles and 25-g payload. (a) Positions and velocities of two robots. Solid lines represent the leader and dashed lines represent the follower. (b) Trajectory of the follower robot in the generalized coordinates. The mean errors and standard deviations of $\phi_p$, $\phi_f$, $\theta_p$, and $\psi_f$ calculated from $t = 20 - 60$ s are $8.4^\circ \pm 4.3^\circ$, $5.8^\circ \pm 4.0^\circ$, $0.5^\circ \pm 2.0^\circ$, and $3.6^\circ \pm 5.9^\circ$

wind gusts were strong enough to visibly affect the flight of a single untethered robot.

A photo of an outdoor flight is shown in Fig. 6(c), with the optical flow data and estimated angles in Fig. 10 and Fig. S15. The position plot obtained from numerically integrating the optical flow suggests that the robots traversed over 20 m in 30 s, reaching the maximum speed of over 1 ms$^{-1}$. As visible in Fig. 10(b), during the first five seconds, the controller visibly brought down the initial errors in $\phi_p$ and $\phi_f$. The state errors were maintained to be under $10^\circ$ for the rest of the flight, with the exception of $\psi_f$, which deviated to just over $15^\circ$. The proposed methods, nonetheless, provided the system with the stability despite significant wind disturbances.

F. EKF-BASED ESTIMATOR VERSUS UKF
To highlight the advantage of the proposed estimator over a conventional UKF, we compare the performance and computational demand of the EKF-based observer and UKF.

The IMU data logged from the presented flight experiments and MOCAP data were used. We selected one flight from each set of indoor tests. These flights cover a range of contrasting conditions as some feature state angles that varied significantly from the nominal setpoint $q^*$ whereas the state angles were well controlled in others. Since the demand of the UKF exceeds the computational limit of Minidrones, we tested both estimators offline on a laptop computer (AMD
The flight experiments demonstrated that the proposed EKF-based estimator rivals the performance of the UKF in this setting, with the reduction of the computational cost by one order of magnitude. However, since it relies only on the robot’s existing IMU, it can either be applied to small quadrotors that are unable to carry a companion computer or implemented in conjunction with other vision-based approaches in larger robots to improve the flight performance or as a back-up mechanism in case of poor visibility or hardware failure. For future works, it remains to be seen whether the observability or controllability of the inertia-based method remains valid when more than two followers are involved. Nevertheless, we have shown that the performance of the vision-less method in the transport of a suspended payload compares favorably to the previous vision-based approaches in [26], [31]. For instance, in the leader-payload-follower flight in Section V-C4, the system tracked a inclined circular trajectory over 15.9 m in 60 seconds, while [26] transported a payload on a straight line across a distance of 3.3 m. The RMSE of the position of the payload is [6.4, 5.7, 4.0] cm, which slightly lower than [11.2, 10.8, 4.9] cm of the circular trajectory tracking flight in [31]. However, the direct comparison must be taken with care as the position of the payload in this work is defined with respect to the environment. The maximum transport speed demonstrated in this work (1.4 ms⁻¹) is also similar to those achieved through the vision-based strategies [26], [31] (≈1 ms⁻¹).

This work is the first step toward the applications of cooperative transport with minimal hardware requirements as navigation or a human pilot is only needed for the leader robot. All in all, this minimal solution, when compared with the vision-based methods, has its disadvantages and advantages. Despite the benefit of minimal computation and hardware requirements, the inertia-based method for aerial transport of a suspended payload presented in this work is not yet compatible with a rigid body payload as found in [26], [31]. This is because the limited sensing ability makes it difficult to infer the full pose of the suspended object. Moreover, it has been demonstrated in [31] that a vision-based approach supports the cooperation of more than two agents, whereas it remains unclear whether the observability and controllability of the vision-less approach remain valid when more than two robots are involved.

VI. CONCLUSION

In this paper, we have devised a vision-less method to enable two robots to transport a suspended payload in the leader-follower format. The advantage lies in the efficiency and minimal sensing requirement. Thanks to the comprehension of the underpinning dynamics of the system, the cooperative transport has been realized with relatively straightforward estimation and control schemes. The flight experiments validated the proposed strategy in both indoor and outdoor settings, using two sub-100-g MAVs with severely limited computational power.

Despite the benefit of minimal computation and hardware requirements, the inertia-based method for aerial transport of a suspended payload presented in this work is not yet compatible with a rigid body payload as found in [26], [31]. This is because the limited sensing ability makes it difficult to infer the full pose of the suspended object. Moreover, it has been demonstrated in [31] that a vision-based approach supports the cooperation of more than two agents, whereas it remains unclear whether the observability and controllability of the vision-less approach remain valid when more than two robots are involved. Nevertheless, we have shown that the performance of the vision-less method in the transport of a suspended payload compares favorably to the previous vision-based approaches in [26], [31]. For instance, in the leader-payload-follower flight in Section V-C4, the system tracked a inclined circular trajectory over 15.9 m in 60 seconds, while [26] transported a payload on a straight line across a distance of 3.3 m. The RMSE of the position of the payload is [6.4, 5.7, 4.0] cm, which slightly lower than [11.2, 10.8, 4.9] cm of the circular trajectory tracking flight in [31]. However, the direct comparison must be taken with care as the position of the payload in this work is defined with respect to the leader (as the leader is separately controlled), whereas in [31], all robots stabilized the payload with respect to the environment. The maximum transport speed demonstrated in this work (1.4 ms⁻¹) is also similar to those achieved through the vision-based strategies [26], [31] (≈1 ms⁻¹).

This work is the first step toward the applications of cooperative transport with minimal hardware requirements as navigation or a human pilot is only needed for the leader robot. All in all, this minimal solution, when compared with the vision-based methods, has its disadvantages and advantages. However, since it relies only on the robot’s existing IMU, it can either be applied to small quadrotors that are unable to carry a companion computer or implemented in conjunction with other vision-based approaches in larger robots to improve the flight performance or as a back-up mechanism in case of poor visibility or hardware failure. For future works, it remains to be seen whether the observability or controllability of the inertia-based method remains valid when more than two followers are involved. Similarly, the manipulation and transport of a rigid body payload is another possible research avenue.

REFERENCES


