Cooperative Transportation of Cable Suspended Payloads With MAVs Using Monocular Vision and Inertial Sensing

Guanrui Li, Rundong Ge, and Giuseppe Loianno

***Abstract*—Micro Aerial Vehicles (MAVs) have the great potential to be deployed in commercial or health care services such as e-commerce package delivery, transportation of medicines, same-day food delivery, and other time-sensitive transportation tasks. A team of MAVs can cooperatively transport objects to overcome the physical limitations of a single vehicle, while concurrently increasing the system's resilience to vehicles' failures. In this letter, we address the state estimation, control and trajectory tracking problems of cooperative transportation of cable suspended rigid body payloads with MAVs using monocular vision and inertial sensing. The key contributions are (a) a distributed vision-based coordinated control of the cable-suspended rigid body payload on SE(3), (b) a distributed estimation approach that allows each agent to estimate its cable direction and velocity independently, and (c) a new cooperative estimation scheme that can infer the payload's full 6-DoF states. This is obtained by sharing the robots' local position estimates and their relative position with respect to their corresponding attachment points on the payload. It allows to infer the payload's 6-DoF state when it is not directly measurable by each vehicle. The proposed solution runs in real-time on board and is validated through experimental results with multiple quadrotors transporting a rigid body payload via cables.**

***Index Terms*—Aerial systems, aerial systems, applications, perception and autonomy.**

I. Introduction

MICRO Aerial Vehicles (MAVs), equipped with on board sensors have great potential to help or substitute humans in complex or dangerous tasks such as exploration [1], inspection [2], mapping [3], interaction with the environment [4], search and rescue [5], and transportation [6], [7]. Particularly, aerial transportation offers a faster and more versatile solution compared to ground transportation in a congested urban area. In addition, in post-disaster scenarios, a team of aerial robots can deliver supplies and establish communication in areas where GPS signal is intermittent or unavailable. To increase the payload capacity, one can either use a large aerial vehicle or utilize a team of MAVs that cooperatively transport a payload. Although the complexity of the system increases with the number of robots, a team of MAVs can potentially provide additional resilience to the task compared to a single MAV in case of vehicles' failures. Furthermore, [8] suggests that small scale drone-based delivery systems can potentially reduce greenhouse gas emissions and energy use in the freight sector. In this work, we investigate the control, state estimation, and trajectory tracking problems of distributed cooperative transportation of cable-suspended rigid body payloads with MAVs using a monocular camera and inertial sensor as depicted in Fig. 1.

Fig. 1. Three quadrotors during a transportation task carrying a payload via cables using monocular vision and IMU.

Several approaches have been proposed to achieve aerial transportation with multiple MAVs. In [9] aerial robots equipped with active grippers can grasp and rigidly attach to the payload. However, active grasping mechanisms are more difficult to design and more energy-consuming than passive ones. Another approach is to use passive rigid mechanisms like magnets [10]. However, despite rigid connection generates a compact system, it increases the system inertia making it sluggish and incapable of quickly responding to perturbations.

Another passive connection is to suspend the payload from the MAVs through cables. This connection is low cost and requires minimal or no mechanical design for the vehicles and loads. Moreover, the additional mobility introduced between the vehicles and the load can introduce faster system response than the system with rigid connections. In some research literature,

Manuscript received October 15, 2020; accepted February 16, 2021. Date of publication March 11, 2021; date of current version May 4, 2021. This letter was recommended for publication by Associate Editor M. Garratt and Editor P. Pounds upon evaluation of the reviewers' comments. This work was supported in part by Qualcomm Research, the Technology Innovation Institute, Nokia, NYU Wireless, and in part by Young Researchers Program "Rita Levi di Montalcini" 2017 Grant PGR17W9W4N. *(Corresponding author: Guanrui Li.)*

The authors are with the Tandon School of Engineering, New York University, New York, NY 11201 USA (e-mail: lguanrui@nyu.edu; rundong.ge@nyu.edu; loiannog@nyu.edu).

This article has supplementary downloadable material available at https://doi.org/10.1109/LRA.2021.3065286, provided by the authors.

Digital Object Identifier 10.1109/LRA.2021.3065286

centralized control solutions are proposed for these types of system [11]–[14], but the control problem can also be approached in distributed fashion with the advantage of not requiring a centralized coordinator. Recently, a team formation-based method is introduced in [15]. However, the payload is not considered as an integral part of the system, but as a disturbance. It is stabilized by minimizing the cable swing which restricts the full–range motion of the system. The load is indirectly displaced as the agents follow a consensus trajectory. A leader-follower approach is proposed in [16], [17], however, such approaches are subject to the single point of failure since the system relies on the master agent as the central unit for navigation. Another type of distributed approach is proposed in [18] and [19] to control the 6 Degrees of Freedom (DoF) of cable suspended rigid body payloads considering the full nonlinear system dynamics. In both papers, the vehicles share a common goal and each agent calculates its control actions based on its attach point on the payload. However, both papers only provide a theoretical stability analysis and simulation results without dealing with the experimental design challenges on real robots equipped with embedded on board sensors and computationally limited units. In our recent work [20], we experimentally demonstrate a distributed cable-suspended payload control solution with multiple quadrotors, whereas in [21] a similar but semi-distributed solution installs a small computer on the payload acting as an extra agent in the system.

All the previous works did not address the perception, state estimation, and vision-based coordinated control research challenges for cooperative transportation of cable suspended payloads in 6–DoF. They circumvent this problem by either considering simplified dynamic models or using an external sensing system like GPS and motion capture system. In this paper, we address the challenges of bridging the gap between the closed-loop control method and the real world deployment by closing the loop with on-board real time state estimation without any external sensing system, such as GPS or motion capture system.

This work presents multiple contributions. First, we propose a vision-based position and attitude control of the cable-suspended payload with multiple quadrotors. The control scheme allows independent control of each vehicle. Second, each robot, leveraging its sensor information, estimates its cable direction and velocity in distributed fashion. Third, we propose a new cooperative state estimation scheme that can estimate the full 6-DoF of cable-suspended rigid-body payload. If the payload's states is not directly measurable by each agent, our method can estimate them by combing the shared informations among the team. Each agent will share its local position and its relative position with respect to the payload. Finally, we validate our approach with trajectory tracking experiments using three quadrotors with on board perception, state estimation, and control. We also discuss the scalability with respect to the number of vehicles at both control and perception levels. To the best of our knowledge, this is the first time that real-time perception, state estimation, and control are designed and integrated on board with MAVs for distributed cable-suspended rigid-body payload transportation. We believe that the proposed system can be deployed in real-world settings such as warehouses and GPS-denied environments.

The paper is organized as follows. In Section II, we review the system dynamics and the proposed control framework. Section III presents the state estimation strategy for autonomous flight. Section IV presents results on transportation tasks with three quadrotors and discusses the scalability. Section V concludes the work and proposes multiple future research directions.

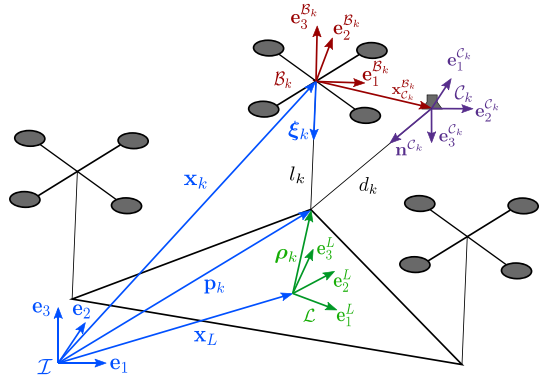

Fig. 2. System convention definition $\mathcal{I}, \mathcal{L}, \mathcal{B}_k, \mathcal{C}_k$ denote the inertial, payload, the k^{th} robot body and the k^{th} camera frames respectively.

TABLE I
NOTATION TABLE

$\mathcal{I}, \mathcal{L}, \mathcal{B}_k$	inertial frame, payload frame, k^{th} robot frame
$m_L, m_k \in \mathbb{R}$	mass of payload, k^{th} robot
$\mathbf{x}_L, \mathbf{x}_k \in \mathbb{R}^3$	position of payload, k^{th} robot in $\mathcal{I}$
$\dot{\mathbf{x}}_L, \ddot{\mathbf{x}}_L \in \mathbb{R}^3$	linear velocity, acceleration of payload in $\mathcal{I}$
$\dot{\mathbf{x}}_k, \ddot{\mathbf{x}}_k \in \mathbb{R}^3$	linear velocity, acceleration of k^{th} robot in $\mathcal{I}$
$\boldsymbol{R}_L \in SO(3)$	orientation of payload with respect to $\mathcal{I}$
$\boldsymbol{R}_k \in SO(3)$	orientation of k^{th} robot with respect to $\mathcal{I}$
$\boldsymbol{\Omega}_L, \dot{\boldsymbol{\Omega}}_L \in \mathbb{R}^3$	angular velocity, acceleration of payload in $\mathcal{L}$
$\boldsymbol{\Omega}_k \in \mathbb{R}^3$	angular velocity of k^{th} robot in $\mathcal{B}_k$
$f_k \in \mathbb{R}$, $\mathbf{M}_k \in \mathbb{R}^3$	total thrust, moment at k^{th} robot in $\mathcal{B}_k$.
$\mathbf{J}_L, \mathbf{J}_k \in \mathbb{R}^{3\times3}$	moment of inertia of payload, k^{th} robot
$\boldsymbol{\xi}_k \in S^2$	unit vector from k^{th} robot to attach point in $\mathcal{I}$
$\boldsymbol{\omega}_k \in \mathbb{R}^3$, $l_k \in \mathbb{R}$	angular velocity, length of k^{th} cable
$\boldsymbol{\rho}_k \in \mathbb{R}^3$	position of k^{th} attach point in $\mathcal{L}$

II. MODELING AND CONTROL

In this section, we introduce the dynamic model and controller for the system, where a rigid-body load is suspended by cables from n quadrotors. The controller is formulated based on the system dynamics presented in Section II-A. We employ a hierarchical formulation starting from a payload controller described in Section II-B generating forces and moments to control the position and orientation of the rigid–body load. Subsequently, these quantities are then distributed to desired thrust and moment on each robot based on connecting cables as presented in Section II-C. Our solution enables explicitly closed-loop control of both position and orientation of the payload to track a desired trajectory in 3D space in distributed fashion. These represent key advantages of this method compared to existing state of the art solutions.

A. Dynamic Model

We consider n quadrotors that cooperatively transport a rigid body payload through massless cables with respect to the world frame $\mathcal{I}$ as depicted in Fig. 2. Without loss of generality, we define that $\mathcal{I}$ is fixed at the ground and coincident with the initial payload frame $\mathcal{L}$ that is located at the center of mass of the payload. The relevant variables in our problem are stated in Table. I. Assuming the cables are taut, the geometric relationship between the center of mass of the payload and the k^{th} quadrotor

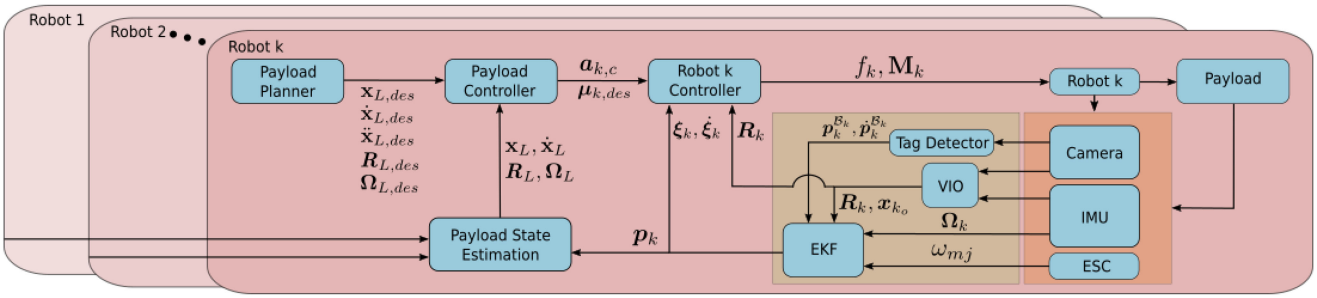

Fig. 3. System architecture of the cooperative transportation system.

is

$$\mathbf{x}_k = \mathbf{x}_L + \boldsymbol{R}_L \boldsymbol{\rho}_k - l_k \boldsymbol{\xi}_k.$$

The kinetic energy $\mathcal{T}$ and potential energy $\mathcal{U}$ of the system are respectively

$$\mathcal{T} = \frac{1}{2} m_L \left\| \dot{\mathbf{x}}_L \right\|_2^2 + \frac{1}{2} \boldsymbol{\Omega}_L \cdot \mathbf{J}_L \boldsymbol{\Omega}_L + \sum_{k=1}^{n} \frac{1}{2} m_k \left\| \dot{\mathbf{x}}_k \right\|_2^2 + \frac{1}{2} \boldsymbol{\Omega}_k \cdot \mathbf{J}_k \boldsymbol{\Omega}_k,$$

$$\mathcal{U} = m_L \mathbf{g} \cdot \mathbf{x}_L + \sum_{k=1}^{n} m_k \mathbf{g} \cdot \mathbf{x}_k, \tag{1}$$

where $\mathbf{g} = g\mathbf{e}_3$, $g = 9.81$ m/s^2 and $\mathbf{e}_3 = [0 \quad 0 \quad 1]^\top$.

By applying the Lagrange d'Alembert principle, similarly to [18], after integrating by parts, the system dynamics are

$$\sum_{k=1}^{n} \bar{\mathbf{u}}_k^{\parallel} = m_L \left(\ddot{\mathbf{x}}_L + \mathbf{g} \right), \tag{2}$$

$$\sum_{k=1}^{n} \hat{\boldsymbol{\rho}}_k \boldsymbol{R}_L^\top \bar{\mathbf{u}}_k^{\parallel} = \mathbf{J}_L \dot{\boldsymbol{\Omega}}_L + \hat{\boldsymbol{\Omega}}_L \mathbf{J}_L \boldsymbol{\Omega}_L, \tag{3}$$

$$\hat{\boldsymbol{\xi}}_k^2 \left(\mathbf{u}_k - m_k \boldsymbol{a}_k \right) = m_k l_k \left(\ddot{\boldsymbol{\xi}}_k + \left\| \dot{\boldsymbol{\xi}}_k \right\|_2^2 \boldsymbol{\xi}_k \right), \tag{4}$$

$$\mathbf{M}_k = \mathbf{J}_k \dot{\boldsymbol{\Omega}}_k + \boldsymbol{\Omega}_k \times \mathbf{J}_k \boldsymbol{\Omega}_k, \tag{5}$$

where $\boldsymbol{a}_k$ is the acceleration of the k^{th} attachment point,

$$\boldsymbol{a}_k = \ddot{\mathbf{x}}_L + \mathbf{g} - \boldsymbol{R}_L \hat{\boldsymbol{\rho}}_k \dot{\boldsymbol{\Omega}}_L + \boldsymbol{R}_L \hat{\boldsymbol{\Omega}}_L^2 \boldsymbol{\rho}_k,$$

$$\bar{\mathbf{u}}_k^{\parallel} = \mathbf{u}_k^{\parallel} - m_k l_k \left\| \boldsymbol{\omega}_k \right\|_2^2 \boldsymbol{\xi}_k - m_k \boldsymbol{\xi}_k \boldsymbol{\xi}_k^\top \boldsymbol{a}_k,$$

$$\mathbf{u}_k^{\parallel} = \boldsymbol{\xi}_k \boldsymbol{\xi}_k^\top \mathbf{u}_k, \quad \mathbf{u}_k^{\perp} = \left(\mathbf{I}_{3\times 3} - \boldsymbol{\xi}_k \boldsymbol{\xi}_k^\top \right) \mathbf{u}_k,$$

$\mathbf{I}_{3\times 3}$ is the identity matrix, and the hat map $\hat{\cdot} : \mathbb{R}^3 \to SO(3)$ is defined such that $\hat{\boldsymbol{a}}\mathbf{b} = \boldsymbol{a} \times \mathbf{b}, \forall \boldsymbol{a}, \mathbf{b} \in \mathbb{R}^3$.

B. Payload Control

The payload controller enables the load to follow a desired trajectory in a closed loop. The subscript $*_{des}$ denotes the desired value given by the trajectory planner. The desired forces and moments acting on the payload are designed as

$$\mathbf{F}_{des} = m_L \boldsymbol{a}_{L,c}, \tag{6}$$

$$\boldsymbol{a}_{L,c} = \mathbf{K}_p \mathbf{e}_{\mathbf{x}_L} + \mathbf{K}_d \mathbf{e}_{\dot{\mathbf{x}}_L} + \mathbf{K}_i \int_0^t \mathbf{e}_{\mathbf{x}_L} d\tau + \ddot{\mathbf{x}}_{L,des} + \mathbf{g}$$

$$\mathbf{M}_{des} = \mathbf{K}_{R_L} \mathbf{e}_{R_L} + \mathbf{K}_{\boldsymbol{\Omega}_L} \mathbf{e}_{\boldsymbol{\Omega}_L} + \mathbf{J}_L \boldsymbol{R}_L^\top \boldsymbol{R}_{L,des} \dot{\boldsymbol{\Omega}}_{L,des} + \left(\boldsymbol{R}_L^\top \boldsymbol{R}_{L,des} \boldsymbol{\Omega}_{L,des} \right)^\wedge \mathbf{J}_L \boldsymbol{R}_L^\top \boldsymbol{R}_{L,des} \boldsymbol{\Omega}_{L,des}, \tag{7}$$

where $\mathbf{K}_p, \mathbf{K}_d, \mathbf{K}_i, \mathbf{K}_{R_L}, \mathbf{K}_{\boldsymbol{\Omega}_L}$ are diagonal positive constant matrices, and

$$\mathbf{e}_{\mathbf{x}_L} = \mathbf{x}_{L,des} - \mathbf{x}_L, \ \mathbf{e}_{\dot{\mathbf{x}}_L} = \dot{\mathbf{x}}_{L,des} - \dot{\mathbf{x}}_L,$$

$$\mathbf{e}_{R_L} = \frac{1}{2} \left(\boldsymbol{R}_L^\top \boldsymbol{R}_{L,des} - \boldsymbol{R}_{L,des}^\top \boldsymbol{R}_L \right)^\vee, \tag{8}$$

$$\mathbf{e}_{\boldsymbol{\Omega}_L} = \boldsymbol{R}_L^\top \boldsymbol{R}_{L,des} \boldsymbol{\Omega}_{L,des} - \boldsymbol{\Omega}_L.$$

The desired forces and moments are distributed to the desired tensions $\boldsymbol{\mu}_{k,des}$ along each cable as

$$\begin{bmatrix} \boldsymbol{\mu}_{1,des} \\ \vdots \\ \boldsymbol{\mu}_{n,des} \end{bmatrix} = \operatorname{diag}\left(\boldsymbol{R}_L, \ldots, \boldsymbol{R}_L \right) \mathbf{P}^\dagger \begin{bmatrix} \boldsymbol{R}_L^\top \mathbf{F}_{des} \\ \mathbf{M}_{des} \end{bmatrix}, \tag{9}$$

where $\mathbf{P}^\dagger$ denotes the pseudoinverse of $\mathbf{P}$ and

$$\mathbf{P} = \begin{bmatrix} \mathbf{I}_{3\times 3} & \mathbf{I}_{3\times 3} & \cdots & \mathbf{I}_{3\times 3} \\ \hat{\boldsymbol{\rho}}_1 & \hat{\boldsymbol{\rho}}_2 & \cdots & \hat{\boldsymbol{\rho}}_n \end{bmatrix}. \tag{10}$$

We arrange the cable attach points such that there are at least 3 non-collinear attach points on the payload. This guarantees $rank(\mathbf{P}) \geq 6$ and there are enough control inputs to move the payload in 6–DoF.

C. Robot Controller

The tension input of each cable of the system is selected as the projection of desired tension on the corresponding cable

$$\boldsymbol{\mu}_k = \boldsymbol{\xi}_k \boldsymbol{\xi}_k^\top \boldsymbol{\mu}_{k,des}, \tag{11}$$

The desired direction $\boldsymbol{\xi}_{k,des}$ and the desired angular velocity $\boldsymbol{\omega}_{k,des}$ of the k^{th} cable link can be obtained as

$$\boldsymbol{\xi}_{k,des} = -\frac{\boldsymbol{\mu}_{k,des}}{\left\| \boldsymbol{\mu}_{k,des} \right\|}, \ \boldsymbol{\omega}_{k,des} = \boldsymbol{\xi}_{k,des} \times \dot{\boldsymbol{\xi}}_{k,des},$$

where $\dot{\boldsymbol{\xi}}_{k,des}$ is the derivative of $\boldsymbol{\xi}_{k,des}$. The thrust f_k and moments $\mathbf{M}_k$ acting at the k^{th} quadrotor are

$$f_k = \mathbf{u}_k \cdot \boldsymbol{R}_k \mathbf{e}_3 = \left(\mathbf{u}_k^{\parallel} + \mathbf{u}_k^{\perp} \right) \cdot \boldsymbol{R}_k \mathbf{e}_3, \tag{12}$$

$$\mathbf{M}_k = \mathbf{K}_R \mathbf{e}_{R_k} + \mathbf{K}_{\boldsymbol{\Omega}} \mathbf{e}_{\boldsymbol{\Omega}_k} + \boldsymbol{\Omega}_k \times \mathbf{J}_k \boldsymbol{\Omega}_k - \mathbf{J}_k \left(\hat{\boldsymbol{\Omega}}_k \boldsymbol{R}_k^\top \boldsymbol{R}_{k,des} \boldsymbol{\Omega}_{k,des} - \boldsymbol{R}_k^\top \boldsymbol{R}_{k,des} \dot{\boldsymbol{\Omega}}_{k,des} \right), \tag{13}$$

where $\mathbf{K}_R$ and $\mathbf{K}_{\boldsymbol{\Omega}}$ are diagonal positive constant matrices, $\mathbf{e}_{R_k}$ and $\mathbf{e}_{\boldsymbol{\Omega}_k}$ are the orientation and angular velocity errors similarly

defined using eq. (8). The inputs $\mathbf{u}_k^{\perp}$ and $\mathbf{u}_k^{\parallel}$ are designed as

$$\mathbf{u}_k^{\perp} = m_k l_k \boldsymbol{\xi}_k \times \Big[-\mathbf{K}_{\boldsymbol{\xi}_k} \mathbf{e}_{\boldsymbol{\xi}_k} - \mathbf{K}_{\boldsymbol{\omega}_k} \mathbf{e}_{\dot{\boldsymbol{\omega}}_k} - \hat{\boldsymbol{\xi}}_k^2 \dot{\boldsymbol{\omega}}_{k,des} - (\boldsymbol{\xi}_k \cdot \boldsymbol{\omega}_{k,des}) \dot{\boldsymbol{\xi}}_k \Big] - m_k \hat{\boldsymbol{\xi}}_k^2 \boldsymbol{a}_{k,c}, \tag{14}$$

$$\mathbf{u}_k^{\parallel} = \boldsymbol{\mu}_k + m_k l_k \left\| \boldsymbol{\omega}_k \right\|_2^2 \boldsymbol{\xi}_k + m_k \boldsymbol{\xi}_k \boldsymbol{\xi}_k^{\top} \boldsymbol{a}_{k,c}, \tag{15}$$

$$\boldsymbol{a}_{k,c} = \boldsymbol{a}_{L,c} - \boldsymbol{R}_L \hat{\boldsymbol{\rho}}_k \dot{\boldsymbol{\Omega}}_L + \boldsymbol{R}_L \hat{\boldsymbol{\Omega}}_L^2 \boldsymbol{\rho}_k, \tag{16}$$

where $\mathbf{K}_{\boldsymbol{\xi}_k}$ and $\mathbf{K}_{\boldsymbol{\omega}_k}$ are diagonal positive constant matrices, $\mathbf{e}_{\boldsymbol{\xi}_k}$ and $\mathbf{e}_{\boldsymbol{\omega}_k}$ are the cable direction and cable angular velocity errors respectively defined as following

$$\mathbf{e}_{\boldsymbol{\xi}_k} = \boldsymbol{\xi}_{k,des} \times \boldsymbol{\xi}_k, \ \mathbf{e}_{\boldsymbol{\omega}_k} = \boldsymbol{\omega}_k + \boldsymbol{\xi}_k \times \boldsymbol{\xi}_k \times \boldsymbol{\omega}_{k,des}.$$

The reader can refer to [18] for a stability analysis. The aforementioned derivation addresses the challenges related to transitioning this class of controllers from theory to practice by showing a clear hierarchical formulation that can run on real robots. Furthermore, we will show how to integrate it with on board perception based systems and also discuss the scalability of the proposed method in Section IV-D.

III. State Estimation

As previously stated, one of the main goals in this work is to achieve closed-loop control of payload's 6-DoF pose with on board sensors. We consider using a minimal sensor suite, which consists of a single monocular camera and an IMU [22], for autonomous navigation. However, the sensing and perception problems of this system is complex. By inspecting eqs. (6)–(16), we observe that we need to concurrently estimate the position, orientation, and velocity of the carried payload as well as the direction $\boldsymbol{\xi}_k$ and velocity $\dot{\boldsymbol{\xi}}_k$ of each cable connecting the vehicles to the load.

In the following, we propose to solve both problems leveraging the monocular camera and IMU information available on each vehicle. We propose to estimate $\boldsymbol{\xi}_k$ and $\dot{\boldsymbol{\xi}}_k$ in distributed fashion on each vehicle, utilizing system dynamics, motor speeds, and camera information. The payload pose estimation is subsequently solved by sharing the relative position of each robot with respect to the payload and combining it with the on board Visual Inertial Odometry (VIO) information. This allows to infer the payload's pose in case it is not directly measurable by each vehicle. In fact, while it is definitely possible to employ other vision-based object pose algorithms leveraging more sophisticated tags [23] or deep learning methods [24], these would be computationally expensive and slow on small robots with Size, Weight, and Power (SWaP) constraints.

A. Preliminaries

Since the payload needs to be controlled in the inertial frame $\mathcal{I}$, we propose a method that the team of MAVs can estimate the full 6–DoF of payload's state with respect to the inertial frame $\mathcal{I}$ cooperatively. As mentioned in Section II and shown in Fig. 4, the inertial frame $\mathcal{I}$ is coincident with the initial $\mathcal{L}$ when the payload is still on the ground. The definition of the axes, $\mathbf{e}_1^L, \mathbf{e}_2^L, \mathbf{e}_3^L$, of $\mathcal{L}$ is based on the attach points. As mentioned in Section II-B, there are at least 3 non-collinear attach points on the payload. The 3 non-collinear attach points are numbered clockwisely as point 1, 2, 3 respectively, and used to define the x-y plane of $\mathcal{L}$. The positions of the 3 selected points with respect to $\mathcal{I}$ are denoted as $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3 \in \mathbb{R}^3$ respectively. The vector $\mathbf{e}_3^L$ represents the normal to the x-y plane defined by the 3 selected attach points. $\mathbf{e}_2^L$ is the unit vector of $\mathbf{p}_1 - \mathbf{p}_3$ and $\mathbf{e}_1^L$ is the cross product of $\mathbf{e}_2^L$ and $\mathbf{e}_3^L$. The orientation of payload can then be represented by the rotation matrix

$$\boldsymbol{R}_L = \begin{bmatrix} \mathbf{e}_1^L & \mathbf{e}_2^L & \mathbf{e}_3^L \end{bmatrix}.$$

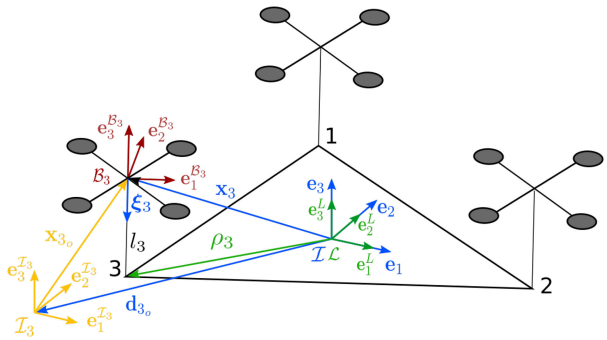

Fig. 4. Representation of the inertial frame $\mathcal{I}$ and the VIO local frame. The vector $\mathbf{d}_{k_0}$ is the offset vector from the origin of $\mathcal{I}$ to the origin of $\mathcal{I}_k$.

As previously mentioned, the local robot position estimate, which can be obtained from the VIO estimation running on each robot using a downward camera and an IMU on board [22], is leveraged to infer the payload state. However, as depicted in Fig. 4, the VIO algorithm on the k^{th} robot estimates the vehicle's states with respect to a fixed local reference frame $\mathcal{I}_k$. The origin of this frame is based on the robot's initial position and we assume that it is initialized to be parallel to the initial $\mathcal{I}$ frame. The initial alignment offset can eventually be estimated at take-off by concurrently using specific objects' visual features (e.g., concurrently using all features used to estimate the $\boldsymbol{\xi}_k$). The offset vector $\mathbf{d}_{k_o}$ from the origin of $\mathcal{I}$ to the origin of $\mathcal{I}_k$ can be measured when the robots are commanded to take–off to a certain height to make the cable taut and are able to detect the attach points. The offset vector $\mathbf{d}_{k_o}$ can then be obtained as follow

$$\mathbf{d}_{k_o} = \boldsymbol{\rho}_k - \mathbf{x}_{k_o} - l_k \boldsymbol{\xi}_k,$$

where $\mathbf{x}_{k_o}$ is the estimated position of the k^{th} robot with respect to $\mathcal{I}_k$ by the on board VIO algorithm, $\boldsymbol{\xi}_k$ is the cable direction estimated by the algorithm in Section III-C, $\boldsymbol{\rho}_k$ and l_k are known geometric constants.

B. Payload Visual Odometry Estimation

Once $\mathbf{d}_{k_o}$ is measured, the position of the k^{th} attach point with respect to $\mathcal{I}$ during the transportation can be obtained

$$\mathbf{p}_k = \mathbf{x}_k + l_k \boldsymbol{\xi}_k = \mathbf{x}_{k_o} + \mathbf{d}_{k_o} + l_k \boldsymbol{\xi}_k, \tag{17}$$

and the $\mathbf{e}_3^L$ is obtained by solving the following equation

$$\mathbf{A} \cdot \mathbf{e}_3^L = \begin{bmatrix} (\mathbf{p}_1 - \mathbf{p}_2)^{\top} \\ \dots \\ (\mathbf{p}_i - \mathbf{p}_j)^{\top} \\ \dots \\ (\mathbf{p}_{m-1} - \mathbf{p}_m)^{\top} \end{bmatrix} \cdot \mathbf{e}_3^L = \mathbf{0}, \ i < j \leq m \leq n,$$

where $m \geq 3$ is the number of co-planar attach points we select to estimate $\mathbf{e}_3^L$. The additional $m - 3$ attach points co-planar to the 3 selected attach points, can be used to obtain a better estimation of $\mathbf{e}_3^L$. Each row of $\mathbf{A}$ is the vector between

two different attach points and $\mathbf{A} \in \mathbb{R}^{\frac{m(m-1)}{2}\times 3}$. Singular value decomposition (SVD) is used to decompose $\mathbf{A}$ and solve the previous equation to derive $\mathbf{e}_3^L$

$$\mathbf{A} = U\Sigma V^\top,\ \mathbf{e}_3^L = V_3/\left\|V_3\right\|_2, \tag{18}$$

where $U \in \mathbb{R}^{\frac{m(m-1)}{2}\times\frac{m(m-1)}{2}}$, Σ is the diagonal singular value matrix and $V = [V_1 \quad V_2 \quad V_3] \in \mathbb{R}^{3\times 3}$, where V_3 corresponds to the 0 singular value of $\mathbf{A}$. If $\mathbf{e}_3^L$ is the opposite to the direction of $(\mathbf{p}_3 - \mathbf{p}_1)\times(\mathbf{p}_2 - \mathbf{p}_1)$, we will flip the sign of $\mathbf{e}_3^L$. Then $\mathbf{e}_1^L$ and $\mathbf{e}_2^L$ can be obtained as

$$\mathbf{e}_2^L = \frac{\mathbf{p}_1 - \mathbf{p}_3}{\left\|\mathbf{p}_1 - \mathbf{p}_3\right\|_2},\ \mathbf{e}_1^L = \frac{\mathbf{e}_2^L \times \mathbf{e}_3^L}{\left\|\mathbf{e}_2^L \times \mathbf{e}_3^L\right\|_2}. \tag{19}$$

Finally, the payload position $\mathbf{x}_L$ is derived as

$$\mathbf{x}_L = \frac{\sum_{k=1}^n \mathbf{p}_k - \boldsymbol{R}_L\boldsymbol{\rho}_k}{n}. \tag{20}$$

Further, we employ an Error-State Kalman Filter (ESKF) designed for estimating position, velocity of the center of mass, and orientation, angular velocity of the rigid body load. Compared to the standard EKF, the main advantage of ESKF is the ability to express the error-state orientation in a minimal parametrization. This avoids the cross-correlation among the orientation components that potentially induce singularities of the state covariance matrix. Furthermore, it separates the nominal states and the small error states, which are more amenable to linear approximation of the states in the filtering, particularly for the orientation (in general non-Euclidean elements) [25]. The payload state $\mathbf{X}$ is defined as $\mathbf{X} = [\mathbf{x}_L^\top \quad \mathbf{q}_L^\top \quad \dot{\mathbf{x}}_L^\top \quad \boldsymbol{\Omega}_L^\top]^\top$, where $\mathbf{q}_L$ is the quaternion representation of $\boldsymbol{R}_L$. And the error payload state $\tilde{\mathbf{X}}$ is

$$\tilde{\mathbf{X}} = \begin{bmatrix}\tilde{\mathbf{x}}_L^\top & \delta\boldsymbol{\Theta}_L^\top & \tilde{\dot{\mathbf{x}}}_L^\top & \tilde{\boldsymbol{\Omega}}_L^\top\end{bmatrix}^\top, \tag{21}$$

where standard error (i.e. the error in the estimate $\bar{\mathbf{x}}_L$ of the vector $\mathbf{x}_L$ is defined as $\tilde{\mathbf{x}}_L = \mathbf{x}_L - \bar{\mathbf{x}}_L$) is used for the position, velocity and angular velocity, and the error $\delta\boldsymbol{\Theta}_L^\top$ represents a small angle rotation related to the error quaternion $\delta\mathbf{q}_L^\top$ such that $\delta\mathbf{q}_L^\top \approx [\delta\boldsymbol{\Theta}_L^\top \quad 1]^\top$ as shown in [25].

We assume a constant velocity and angular velocity model and the corresponding system dynamics is

$$\frac{d\mathbf{x}_L}{dt} = \dot{\mathbf{x}}_L,\ \frac{d\mathbf{q}_L}{dt} = \frac{1}{2}\mathbf{q}_L \otimes \boldsymbol{\Omega}_L,\ \frac{d\dot{\mathbf{x}}_L}{dt} = \frac{d\boldsymbol{\Omega}_L}{dt} = \mathbf{0}, \tag{22}$$

where $\otimes$ is the quaternion product operator. By rewriting the eq. (22) into matrix form, the linearized continuous dynamics for the error payload state is defined as

$$\dot{\tilde{\mathbf{X}}} = \mathbf{C}\tilde{\mathbf{X}} + \mathbf{N}, \tag{23}$$

where $\mathbf{C} \in \mathbb{R}^{12\times 12}$, and the system noise is modeled as additive Gaussian white noise $\mathbf{N} \sim \mathcal{N}(\mathbf{0}, Q_\mathbf{N})$ with zero mean and covariance $Q_\mathbf{N} \in \mathbb{R}^{12\times 12}$. The measurement model is

$$\mathbf{Z} = \begin{bmatrix}\mathbf{x}_L \\ \mathbf{q}_L\end{bmatrix} + \mathbf{V}, \tag{24}$$

where the measurement noise is modeled as additive Gaussian white noise $\mathbf{V} \sim \mathcal{N}(\mathbf{0}, Q_\mathbf{V})$ with zero mean and covariance $Q_\mathbf{V} \in \mathbb{R}^{6\times 6}$.

C. Cable State Estimation

To estimate the cable direction, we need a detector that can measure the bearing or cable direction. In this paper, an algorithm of detecting black and white circular tag [26] is used to detect each cable attachment point on the payload for inferring the cable direction. It provides fast, close to camera rate, and precise detection of a black and white roundel as shown in Fig. 1. From the algorithm, the pixel position (u_k, v_k) of the circular tag at the k^{th} attach point in the image is obtained. It should be pointed out that any detector providing just the cable direction estimate can be used.

The camera is modeled as a pinhole camera and the affine transformation between the pixel coordinates and the unit vector of tag position in $\boldsymbol{\mathcal{C}}_k$ is $\mathbf{n}^{\boldsymbol{\mathcal{C}}_k} = \lambda\mathbf{K}^{-1}[u_k \quad v_k \quad 1]^\top$, where λ is a constant for the scale ambiguity, $\mathbf{K}$ is the camera intrinsic matrix which we assume to be known from camera calibration. Considering the cable taut we have (see Fig. 2)

$$\mathbf{p}_k^{\boldsymbol{\mathcal{B}}_k} = \mathbf{x}_{\boldsymbol{\mathcal{C}}_k}^{\boldsymbol{\mathcal{B}}_k} + d_k\mathbf{R}_{\boldsymbol{\mathcal{C}}_k}^{\boldsymbol{\mathcal{B}}_k}\mathbf{n}^{\boldsymbol{\mathcal{C}}_k}, \tag{25}$$

where $\mathbf{p}_k^{\boldsymbol{\mathcal{B}}_k}$, $\mathbf{x}_{\boldsymbol{\mathcal{C}}_k}^{\boldsymbol{\mathcal{B}}_k}$ are position of the k^{th} attach point and the camera correspondingly with respect to $\boldsymbol{\mathcal{B}}_k$, $\mathbf{R}_{\boldsymbol{\mathcal{C}}_k}^{\boldsymbol{\mathcal{B}}_k}$ is the rotation matrix from $\boldsymbol{\mathcal{B}}_k$ to $\boldsymbol{\mathcal{C}}_k$. After $\mathbf{n}^{\boldsymbol{\mathcal{C}}_k}$ is obtained, the depth d_k of the tag in $\boldsymbol{\mathcal{C}}_k$ can be obtained by solving eq. (26)

$$l_k = \left\|\mathbf{x}_{\boldsymbol{\mathcal{C}}_k}^{\boldsymbol{\mathcal{B}}_k} + d_k\mathbf{R}_{\boldsymbol{\mathcal{C}}_k}^{\boldsymbol{\mathcal{B}}_k}\mathbf{n}^{\boldsymbol{\mathcal{C}}_k}\right\|. \tag{26}$$

The solution that corresponds to the attach point oriented below the robot is selected. An Extended Kalman filter is designed to estimate the k^{th} cable direction $\boldsymbol{\xi}_k$ and its velocity $\dot{\boldsymbol{\xi}}_k$. The filter state and input are defined as

$$\mathbf{X}_k = \begin{bmatrix}\boldsymbol{\xi}_k^\top & \dot{\boldsymbol{\xi}}_k^\top\end{bmatrix}^\top,\ \mathbf{U}_k = \begin{bmatrix}\boldsymbol{a}_k^\top & \mathbf{u}_k^\top\end{bmatrix}^\top, \tag{27}$$

where $\boldsymbol{a}_k$ is defined in the dynamic model and the k^{th} attach point and $\mathbf{u}_k$ is obtained based on motor speed values from the electronic speed controllers as

$$\mathbf{u}_k = \sum_{j=1}^4 k_f\omega_{mj}^2\boldsymbol{R}_k\mathbf{e}_3, \tag{28}$$

where k_f is the known motor constant, ω_{mj} is the motor speed of the j^{th} motor. Based on the dynamics of the cable in eq. (4), we can obtain the resulting process model as

$$\begin{aligned}\dot{\mathbf{X}}_k &= \begin{bmatrix}\dot{\boldsymbol{\xi}}_k \\ \ddot{\boldsymbol{\xi}}_k\end{bmatrix} = f\left(\mathbf{X}_k, \mathbf{U}_k, \mathbf{N}_k\right) \\ &= \begin{bmatrix}\dot{\boldsymbol{\xi}}_k \\ \frac{1}{m_k l_k}\hat{\boldsymbol{\xi}}_k^2\left(\mathbf{u}_k - m_k\boldsymbol{a}_k\right) - \left\|\dot{\boldsymbol{\xi}}_k\right\|_2^2\boldsymbol{\xi}_k\end{bmatrix} + \mathbf{N}_k,\end{aligned} \tag{29}$$

where the process noise $\mathbf{N}_k \in \mathbb{R}^6$ is modeled as additive Gaussian white noise $\mathbf{N}_k \sim \mathcal{N}(\mathbf{0}, Q_{\mathbf{N}_k})$ with zero mean and covariance $Q_{\mathbf{N}_k} \in \mathbb{R}^{6\times 6}$. The measurement model is

$$\begin{aligned}\mathbf{Z}_k &= \begin{bmatrix}\mathbf{p}_k^{\boldsymbol{\mathcal{B}}_k} \\ \dot{\mathbf{p}}_k^{\boldsymbol{\mathcal{B}}_k}\end{bmatrix} = g\left(\mathbf{X}_k, \mathbf{V}_k\right) \\ &= \begin{bmatrix}\boldsymbol{R}_k^\top l_k\boldsymbol{\xi}_k \\ l_k\left(\boldsymbol{R}_k^\top\dot{\boldsymbol{\xi}}_k - \boldsymbol{\Omega}_k \times \boldsymbol{R}_k^\top\boldsymbol{\xi}_k\right)\end{bmatrix} + \mathbf{V}_k,\end{aligned} \tag{30}$$

where the measurement noise is modeled as additive Gaussian white noise $\mathbf{V}_k \sim \mathcal{N}(\mathbf{0}, Q_{\mathbf{V}_k})$ with zero mean and standard deviation $Q_{\mathbf{V}_k} \in \mathbb{R}^{6\times 6}$.

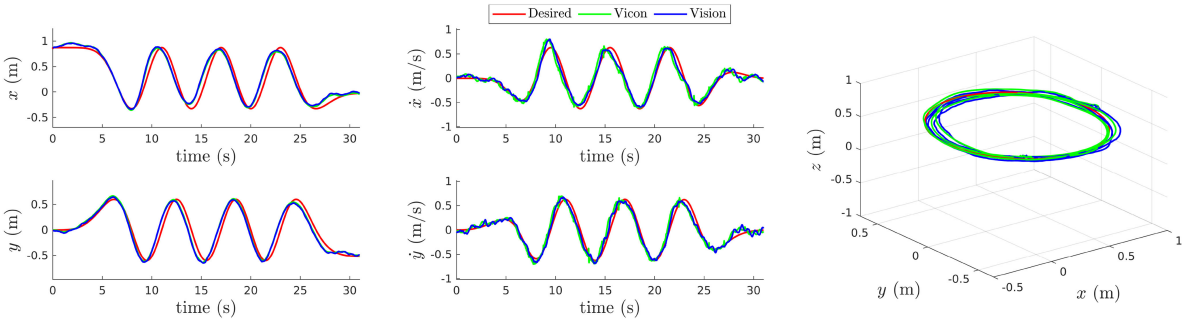

Fig. 5. Tracking results of the payload following a circular trajectory. Cartesian payload position, velocity and 3D path.

IV. Experimental Results

The experiments are conducted in an indoor testbed with a flying space of $7 \times 5 \times 4$ m^3 of the ARPL lab at New York University. The ground truth data is collected using a Vicon[1] motion capture system at 100 Hz. We conduct the experiments with a triangular payload suspended from 3 quadrotors. The mass of the payload is 250 g, which exceeds the payload capacity of a single quadrotor used in the experiments. Each quadrotor platform used in the experiments is equipped with a Qualcomm®Snapdragon™ board for on board computing [22]. The control and estimation frameworks are developed in ROS.[2]

A. Implementation

The controller gains are tuned by evaluating the response of the payload to a step input in simulation and then fine-tuned by using an external motion capture system. The pose of each quadrotor is estimated using the onboard VIO algorithm at the IMU rate, which is approximately 250 Hz. For more details of the VIO algorithm, the reader can refer to our previous work [22]. The EKF developed in Section III-C, which estimates the cable states: direction and velocity, is also running at the same frequency. The EKF will conduct a process update when a new IMU measurement is received, using the closest estimated rotation matrix, control input, and acceleration of the attached point. The measurement update runs asynchronously when a new image is processed by the detector and the $\mathbf{p}_k^{\mathcal{B}_k}$, $\dot{\mathbf{p}}_k^{\mathcal{B}_k}$ are obtained. Together with the position of the robot and measured fixed coordinated frame offset, the position of the corresponding attach point is estimated by applying eq. (17). The three robots, synchronized using Chrony,[3] will share the estimation of the positions of their attach points and the odometry of the payload is then estimated using the algorithm developed in Section III-B.

B. Payload Trajectory Tracking

We test our approach considering a payload's trajectory tracking problem in two different types of trajectory: $\infty-$shape Lissajous trajectory and circular trajectory with radius $r = 0.6$ m. The circular trajectory is designed as

$$\mathbf{x}_{L,des}^{c}(t) = \begin{bmatrix} r\cos\frac{2\pi t}{T_c} & r\sin\frac{2\pi t}{T_c} & h \end{bmatrix}^{\top} \tag{31}$$

with period $T_c = 6$ s, and the Lissajous trajectory as

$$\mathbf{x}_{L,des}^{l}(t) = \begin{bmatrix} A_x\left(1-\cos\frac{2\pi t}{T_l}\right) & A_y\sin\frac{4\pi t}{T_l} & h \end{bmatrix}^{\top} \tag{32}$$

TABLE II
Payload Trajectory Tracking Error and RSME

	Component	Lissajous		Circular	
		MEAN	RMSE	MEAN	RMSE
Position	x (m)	0.05398	0.06521	0.08765	0.1115
	y (m)	0.07835	0.09220	0.08808	0.1074
	z (m)	0.02969	0.03242	0.02059	0.04933
Euler Angle	roll (rad)	0.07979	0.08352	0.1075	0.1162
	pitch (rad)	0.06234	0.07312	0.07797	0.09696
	yaw (rad)	0.09793	0.1136	0.06725	0.09472

TABLE III
Mean and STD of the Cable Estimation Errors

Trajectory	Robot ID	e_{angle} (rad)		$e_{\dot{\xi}}$ (1/s)	
		MEAN	STD	MEAN	STD
Circular	1	0.02097	0.009218	0.1222	0.06856
	2	0.008134	0.003879	0.1255	0.06745
	3	0.03679	0.004616	0.1351	0.06956
Lissajous	1	0.01671	0.006406	0.1242	0.1084
	2	0.008968	0.004869	0.1281	0.07214
	3	0.03540	0.004647	0.1306	0.07574

with period $T_l = 16$ s, amplitude $A_x = 1.2$ m, $A_y = 0.8$ m and a constant height $h = 0.3$ m. The desired orientation of the payload is set to be parallel to the inertial frame during the trajectory tracking experiments. Both the payload controller and the robot controller run in real-time by using the estimated states from our proposed algorithms.

The tracking results of the circular trajectory and the Lissajous trajectory are shown in the Fig. 5 and Fig. 6 respectively. As we can observe from the plots, the payload is able to track the desired position and velocity with very small tracking errors by using our visual estimation as feedback to the closed loop control. The tracking errors of the payload pose are reported in Table II. The maximum speed of the payload reached around 1 m/s during the experiments. The experimental results validate that our proposed system is able to control the cable suspended payload to follow a desired trajectory with onboard real-time visual estimation.

C. Evaluation of Visual State Estimation

We evaluate the performance of our proposed visual state estimation algorithms, which estimates the cable direction, velocity and the payload odometry with respect to the inertial frame. Let

[1] www.vicon.com
[2] www.ros.org
[3] chrony.tuxfamily.org

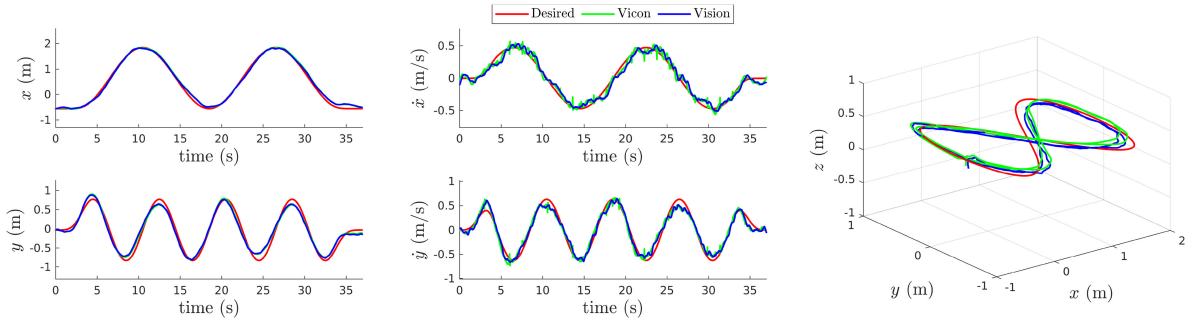

Fig. 6. Tracking results of the payload following a Lissajous trajectory. Cartesian payload position, velocity and 3D path.

TABLE IV
MEAN AND STD OF ABSOLUTE ERROR AND RSME OF VISUAL PAYLOAD POSE ESTIMATION

Trajectory	Statistics	Position (m)			Euler Angle (rad)		
		x	y	z	roll	pitch	yaw
Circular	MEAN	0.01800	0.01544	0.01068	0.06028	0.07577	0.07172
	STD	0.008234	0.008778	0.004570	0.03988	0.05193	0.03749
	RMSE	0.01979	0.01776	0.01161	0.07228	0.09186	0.08093
Lissajous	MEAN	0.02496	0.02080	0.03855	0.06551	0.1224	0.04141
	STD	0.01814	0.01040	0.009478	0.01818	0.07543	0.02930
	RMSE	0.03086	0.02325	0.03970	0.06799	0.1438	0.05072

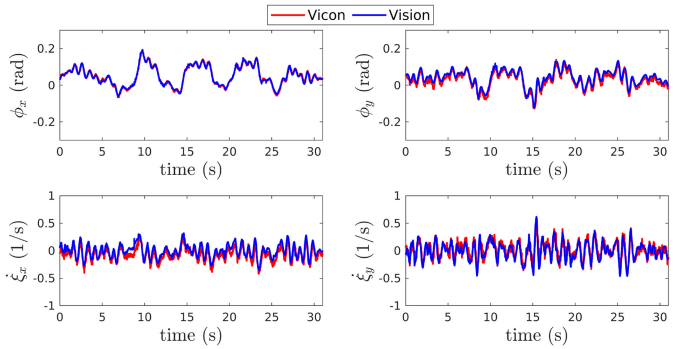

Fig. 7. Estimation results of the cable direction and velocity.

the subscript $\cdot_{gt}$ and $\cdot_v$ denote the ground truth and the visual estimation respectively, and let $\cdot_{1,2,3}$ be the vector components along the axes of the inertial frame $\mathcal{I}$.

1) Cable State: To evaluate the visual estimation of the cable state, we use the angle errors of the cable direction

$$e_{angle} = \frac{|\phi_{x,gt} - \phi_{x,v}| + |\phi_{y,gt} - \phi_{y,v}|}{2}, \tag{33}$$

where

$$\phi_x = \tan^{-1}\left(-\boldsymbol{\xi}_{k_2}/\boldsymbol{\xi}_{k_3}\right), \phi_y = -\tan^{-1}\left(-\boldsymbol{\xi}_{k_1}/\boldsymbol{\xi}_{k_3}\right),$$

and use $e_{\dot{\boldsymbol{\xi}}}$ to evaluate the estimation of the cable velocity

$$e_{\dot{\boldsymbol{\xi}}} = \left\|\dot{\boldsymbol{\xi}}_{k,gt} - \dot{\boldsymbol{\xi}}_{k,v}\right\|_2. \tag{34}$$

In Fig. 7 we show the cable state direction estimation on one robot during a Lissajous trajectory. The other two robots have a similar performance of the cable state estimates. The mean and Standard Deviation (STD) of the errors of cable direction and velocity estimation on all three robots during different flight trajectories are shown in the Table III.

2) Payload Odometry: In this subsection, we report the performance of our visual payload odometry estimation method. The estimation results are evaluated using absolute errors and Root Mean Square Errors (RMSE). The mean and standard deviation of the absolute errors and the RMSE of the payload pose and velocity comparing with the ground truth are shown in Table IV and Table V. We can conclude from the table that the estimation algorithm can infer the payload odometry with small errors comparing with the ground truth from the Vicon. This validates our proposed estimation algorithms that it can provide accurate payload odometry in real time for closed-loop control.

D. Scalability Discussion

It is worth discussing the scalability of the proposed approach with respect to the number of vehicles n both at the control and perception levels. The complexity of the controller is $O(n^2)$ and the one of the state estimation algorithm is $O(m^2)$, $m \leq n$, where m is defined in Section III-B. This implies that, computationally, the proposed approach can theoretically scale to a good number of vehicles even with small CPUs. If the bandwidth of the MAVs communication network is limited, we can reduce the number of shared co-planar attach points' positions m as long as $m \geq 3$. The overall number of agents is generally limited by physical constraints imposed by the payload size. Deciding the flight geometry is a complex task. There is probably not a unique way to solve this problem. Generally, it is good practice to add the vehicles towards the periphery of object to maximize the moments that can be exerted on the payload given a force. Moreover, we chose to equally space them for energy-based reasons to avoid one of the platforms generating substantially higher thrust than the others.

V. CONCLUSION

In this letter, we addressed the vision-based control and estimation problems for cooperative transportation of cable-suspended rigid body payloads with multiple quadrotors using monocular vision and inertial sensing. We demonstrated that our

TABLE V
MEAN AND STD OF ABSOLUTE ERROR AND RSME OF VISUAL PAYLOAD LINEAR AND ANGULAR VELOCITY ESTIMATION

Trajectory	Statistics	Linear Velocity (m/s)			Angular Velocity (rad/s)		
		x	y	z	roll	pitch	yaw
Circular	MEAN	0.08111	0.06008	0.02277	0.1242	0.09768	0.1608
	STD	0.04747	0.04633	0.01792	0.09755	0.06988	0.1198
	RMSE	0.09398	0.07586	0.02898	0.1580	0.1201	0.2005
Lissajous	MEAN	0.03790	0.03845	0.01925	0.07190	0.08953	0.1552
	STD	0.03103	0.02911	0.01500	0.05535	0.07485	0.1138
	RMSE	0.04898	0.04822	0.024406	0.09074	0.1167	0.1925

proposed state estimation method enables closed-loop control of rigid body cable-suspended payload in full 6–DoF state space with on board sensors. We successfully showed multiple trajectory tracking experiments with MAVs equipped with monocular camera and IMU.

Future works will involve improving the controller by considering aerodynamic effects like rotor drags or disturbances such as wind, which become particularly relevant at high speed and in outdoor settings. A possible direction to tackle this problem is to leverage Bayesian learning techniques to predict the unmodeled residual nonlinear dynamics effects. This can correct the nominal system model and further be exploited to design more appropriate controllers that incorporate the uncertainty information such as adaptive controllers or stochastic MPC control. We would like to study how to optimally exploit the extra degrees of freedom in the control inputs once n increases. This enables additional flexibility in controlling the relative configuration of links to account for additional tasks such as obstacle avoidance and minimum safety distance among vehicles. We will also study the use of different types of connections to diversify the robot/object mobility properties and analyze the benefits/drawbacks in terms of perception and control, while still keeping the ability to assign the object's pose in 3D space. We will relax the assumption of known payload characteristics by distributively estimating the relevant parameters using on board sensors.

ACKNOWLEDGMENT

The authors acknowledge Manling Li, Yang Zhou, and Jueun Kwon for their help and support on the experiments.

REFERENCES

[1] T. Tomic *et al.*, "Toward a fully autonomous UAV: Research platform for indoor and outdoor urban search and rescue," *IEEE Robot. Automat. Mag.*, vol. 19, no. 3, pp. 46–56, Sep. 2012.

[2] T. Ozaslan *et al.*, "Autonomous navigation and mapping for inspection of penstocks and tunnels with MAVs," *IEEE Robot. Automat. Lett.*, vol. 2, no. 3, pp. 1740–1747, Jul. 2017.

[3] G. Loianno, J. Thomas, and V. Kumar, "Cooperative localization and mapping of MAVs using RGB-D sensors," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2015, pp. 4021–4028.

[4] F. Forte, R. Naldi, and L. Marconi, "Impedance control of an aerial manipulator," in *Proc. Amer. Control Conf.*, 2012, pp. 3839–3844.

[5] N. Michael *et al.*, "Collaborative mapping of an earthquake-damaged building via ground and aerial robots," *J. Field Robot.*, vol. 29, no. 5, pp. 832–841, 2012.

[6] G. Loianno *et al.*, "Localization, grasping, and transportation of magnetic objects by a team of MAVs in challenging desert-like environments," *IEEE Robot. Automat. Lett.*, vol. 3, no. 3, pp. 1576–1583, Jul. 2018.

[7] K. Sreenath, T. Lee, and V. Kumar, "Geometric control and differential flatness of a quadrotor UAV with a cable-suspended load," in *Proc. 52nd IEEE Conf. Decis. Control*., 2013, pp. 2269–2274.

[8] J. Stolaroff, C. Samaras, E. O'Neill, A. Lubers, A. Mitchell, and D. Ceperley, "Energy use and life cycle greenhouse gas emissions of drones for commercial package delivery," *Nature Commun.*, vol. 9, no. 1, pp. 1–13, Dec. 2018.

[9] D. Mellinger, M. Shomin, N. Michael, and V. Kumar, "Cooperative grasping and transport using multiple quadrotors," *Springer Tracts Adv. Robot.*, Berlin, Heidelberg: Springer, vol. 83, pp. 545–558, 2012.

[10] G. Loianno and V. Kumar, "Cooperative transportation using small quadrotors using monocular vision and inertial sensing," *IEEE Robot. Automat. Lett.*, vol. 3, no. 2, pp. 680–687, Apr. 2018.

[11] N. Michael, J. Fink, and V. Kumar, "Cooperative manipulation and transportation with aerial robots," *Auton. Robots*, vol. 30, no. 1, pp. 73–86, Jan. 2011.

[12] K. Sreenath and V. Kumar, "Dynamics, control and planning for cooperative manipulation of payloads suspended by cables from multiple quadrotor robots," in *Proc. Robot.: Sci. Syst.*, Berlin, Germany, Jun. 2013, doi: 10.15607/RSS.2013.IX.011.

[13] M. Bernard and K. Kondak, "Generic slung load transportation system using small size helicopters," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2009, pp. 3258–3264.

[14] M. Bernard, K. Kondak, I. Maza, and A. Ollero, "Autonomous transportation and deployment with aerial robots for search and rescue missions," *J. Field Robot.*, vol. 28, no. 6, pp. 914–931, 2011.

[15] V. P. Tran, F. Santoso, M. Garratt, and S. Anavatti, "Distributed artificial neural networks-based adaptive strictly negative imaginary formation controller for unmanned aerial vehicles in time-varying environments," *IEEE Trans. Ind. Informat.*, vol. 17, no. 6, pp. 3910–3919, Jun. 2021.

[16] M. Gassner, T. Cieslewski, and D. Scaramuzza, "Dynamic collaboration without communication: Vision-based cable-suspended load transport with two quadrotors," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2017, pp. 5196–5202.

[17] A. Tagliabue, M. Kamel, S. Verling, R. Siegwart, and J. Nieto, "Collaborative transportation using MAVs via passive force control," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2017, pp. 5766–5773.

[18] T. Lee, "Geometric control of quadrotor UAVs transporting a cable-suspended rigid body," in *Proc. 53rd IEEE Conf. Decis. Control*, 2014, pp. 6155–6160.

[19] G. Wu and K. Sreenath, "Geometric control of multiple quadrotors transporting a rigid-body load," in *Proc. 53rd IEEE Conf. Decis. Control.*, 2014, pp. 6141–6148.

[20] G. Li and G. Loianno, "Design and experimental evaluation of distributed cooperative transportation of cable suspended payloads with micro aerial vehicles," in *Proc. Exp. Robot. (ISER 2020), Springer Proc. Adv. Robot.*, 2021, vol. 19, pp. 28–36.

[21] J. Geng and J. Langelaan, "Cooperative transport of a slung load using load-leading control," *J. Guid., Control, Dyn.*, vol. 43, no. 7, pp. 1313–1331, 2020.

[22] G. Loianno, C. Brunner, G. McGrath, and V. Kumar, "Estimation, control, and planning for aggressive flight with a small quadrotor with a single camera and imu," *IEEE Robot. Automat. Lett.*, vol. 2, no. 2, pp. 404–411, Apr. 2017.

[23] E. Olson, "AprilTag: A robust and flexible visual fiducial system," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2011, pp. 3400–3407.

[24] S. Peng, Y. Liu, Q. Huang, X. Zhou, and H. Bao, "Pvnet: Pixel-wise voting network for 6dof pose estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 4556–4565.

[25] J. Solà, "Quaternion Kinematics for the error-state kalman filter," *CoRR*, vol. abs/1711.02508, 2017. [Online]. Available: http://arxiv.org/abs/1711.02508

[26] T. Krajník *et al.*, "A practical multirobot localization system," *J. Intell. Robotic Syst.*, vol. 76, no. 3, pp. 539–562, 2014.