Payload State Estimation for Cooperative Manipulation Using Multiple UAVs

Yuki Ohira, Abner Asignacion Jr., and Satoshi Suzuki*
Chiba University, Japan

http://www.imavs.org/

ABSTRACT

This paper proposes a method for estimating the state of a payload in a cooperative aerial manipulation system using multiple UAVs connected via tethers. The key feature of the proposed approach is that it estimates the tether angles solely from inertial and position data obtained on the UAV side, eliminating the need for any sensors on the payload. These estimated angles are then utilized, in conjunction with a physical model, to infer the payload's position and orientation. The estimation framework is implemented using an Unscented Kalman Filter (UKF) and applied to a system comprising two UAVs and four tethers. The performance of the method is evaluated through both simulations and real-world experiments. The results demonstrate that the proposed approach achieves accurate tracking of both the tether angles and the payload state, even in the presence of sensor noise and modeling inaccuracies, thus validating its effectiveness.

1 INTRODUCTION

In recent years, there has been a growing global interest in small unmanned aerial vehicles (UAVs). Their applications extend beyond conventional fields such as aerial photography, disaster surveys, and infrastructure inspection, and are now being explored for more advanced and complex tasks [1]. Specifically, these include aerial manipulation using UAVs equipped with manipulators [2], and payload transportation in the field of logistics [3]. UAVs are expected to bring numerous benefits to society, including reduced costs and labor requirements, as well as safe access to environments that are difficult or dangerous for humans to enter.

However, most current UAVs are designed for single-unit operation and are therefore limited in the size and weight of payloads they can carry. To address these limitations, cooperative manipulation systems, in which multiple UAVs work together to transport a single payload, have garnered significant attention [4]. Such systems enable the handling of large and heavy payloads that are unmanageable by a single UAV, thereby greatly expanding the potential applications of UAVs.

In cooperative manipulation systems, a widely adopted configuration is to connect each UAV to the payload using cables [5]. While this approach allows for operational flexibility, it also introduces challenges due to the underactuated and deformable nature of the system. These include the risk of inter-UAV collisions, payload oscillations, and instability during flight. Therefore, ensuring the safe and stable operation of such systems requires accurate estimation of the full system state, including the positions and orientations of the UAVs and the payload, as well as cable tensions and angles [6, 7, 8].

Traditionally, estimating the system or payload state has relied on attaching various sensors to the payload. However, these methods face several practical challenges, including the need for dedicated power sources, communication infrastructure, increased cost and weight, and physical constraints on sensor installation. For instance, GPS and IMUs have been used to track the payload's motion [8], while force/torque sensors [9] and vision-based systems [10, 11] have enabled more precise measurements. Even when lightweight sensors are used, the added mass can be non-negligible in small-scale aerial systems with limited payload capacity. In addition, securing power and ensuring reliable data transmission for such sensors impose further constraints that reduce the practicality of these approaches. Consequently, developing estimation techniques that obviate the need for payload-side sensors is crucial for enhancing the practicality and versatility of cooperative aerial systems. In response to these issues, recent research has proposed sensor-minimizing or vision-free methods that estimate the payload state using only onboard UAV measurements [12].

To address this issue, this study proposes a novel framework that estimates the payload state using only onboard UAV sensor data. This sensorless approach on the payload side reduces hardware complexity and eliminates operational constraints. A key feature of the method is that it simultaneously estimates the relative positions between UAVs and the tether attachment points on the payload, which enables the reconstruction of tether angles without requiring direct measurements. This formulation is not restricted to two UAVs and can be extended to systems involving more UAVs due to its generalized modeling of tether geometry.

The remainder of this paper is organized as follows. Section 2 describes the cooperative system and its modeling. The proposed estimation method is explained in Section 3. Then, the effectiveness of the method is validated through simulations in Section 4 and experiments in Section 5. Finally, Section 6 concludes the paper and discusses future work.

*Email address(es): suzuki-s@chiba-u.jp

http://www.imavs.org/

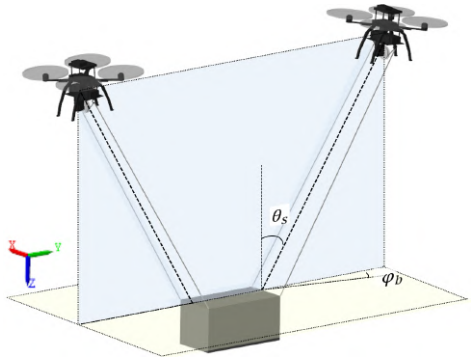


Figure 1: Cooperative manipulation system

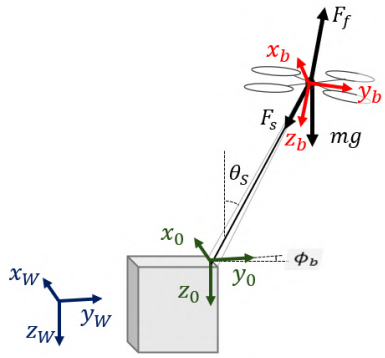


Figure 2: Definition of variables and coordinate system

2 COOPERATIVE MANIPULATION SYSTEM

2.1 System Configuration

This study applies the cooperative manipulation system proposed by Takemura et al. [13]. The system, shown in Figure 1, consists of two UAVs and a payload connected by a total of four tethers at its four vertices. The UAVs are quadrotors, an underactuated system in which rotation and translation are coupled. Tilting the payload in the pitch direction requires the UAV itself to tilt, causing unintended translational motion. To address this , this system is equipped with a winch mechanism on the UAV, allowing for variable tether lengths. The winch mechanism is driven by a single servo motor, where one tether extends as the other retracts. By independently adjusting tether lengths, control of the payload's pitch is achieved.

The orientation of each tether in three-dimensional space, which defines the system state, can be uniquely represented by two angles, similar to spherical coordinates: the Inclination Angle (θ_s), tilt from the vertical, and the Body-Relative Azimuth (ϕ_b), horizontal direction relative to the UAV body frame. The objective of this study is to estimate these two angles using only onboard sensors mounted on each UAV.

2.2 Modeling

The system configuration and coordinate frames are illustrated in Figure 2. The world frame ($\mathcal{F}_W$) follows the NED convention, with z_W pointing downward. The standard basis vectors of this frame are denoted as $\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z$, where $\boldsymbol{e}_z = [0, 0, 1]^T$ corresponds to the direction of gravity, and the gravitational vector is given by $\boldsymbol{g} = g\boldsymbol{e}_z$. The system also uses two additional reference frames:

- Payload frame ($\mathcal{F}_0$): located at the midpoint between the two anchor points connecting a single UAV to the payload, serving as the reference for tether orientation.
- Body-fixed frame ($\mathcal{F}_B$): attached to each UAV's center of gravity.

The azimuth is represented by the body-relative angle ϕ_b, measured from y_b toward x_b in the x_b–y_b plane. This angle is used for relative tether dynamics. For use in the world-frame dynamics, this is converted to an absolute azimuth angle ϕ_{abs}. The UAV's yaw angle ψ_{uav} and the payload's yaw angle ψ_{pay} are approximated as equal ($\psi_{pay} \approx \psi_{uav}$) due to the close coupling of the system. The relationship between these angles, illustrated in the top-down view in Figure 3, is computed as shown in Equation (1):

$$\phi_{abs} = (\psi_{uav} + \phi_b) + \frac{\pi}{2} \tag{1}$$

where the $\pi/2$ term adjusts the North-referenced yaw angle ψ_{uav} to align with the East-based azimuth convention. The dynamics follow a spherical pendulum with an accelerating pivot. Depending on whether relative or global angular rates are required, either ϕ_b or ϕ_{abs} is used in the system dynamics.

The tether motion is described using a spherical coordinate system. The elevation angle θ_s, is measured from the local vertical. The azimuth angle ϕ_{abs}, is measured from the North-East direction. The system is characterized by the orthogonal unit vectors $\boldsymbol{u}_r, \boldsymbol{u}_\theta, \boldsymbol{u}_\phi\}$, defined in Equations (2) - (4):

$$\boldsymbol{u}_r = \begin{bmatrix} \sin\theta_s \cos\phi_{abs} \\ \sin\theta_s \sin\phi_{abs} \\ -\cos\theta_s \end{bmatrix} \tag{2}$$

$$\boldsymbol{u}_\theta = \begin{bmatrix} \cos\theta_s \cos\phi_{abs} \\ \cos\theta_s \sin\phi_{abs} \\ -\sin\theta_s \end{bmatrix} \tag{3}$$

$$\boldsymbol{u}_\phi = \begin{bmatrix} -\sin\phi_{abs} \\ \cos\phi_{abs} \\ 0 \end{bmatrix} \tag{4}$$

The UAV's translational dynamics are described in the world frame $\mathcal{F}_W$ by Equation (5) :

$$\ddot{\boldsymbol{p}} = \frac{1}{m}\boldsymbol{R}_f\,\boldsymbol{F}_f - \frac{1}{m}\boldsymbol{F}_s + \boldsymbol{g} \tag{5}$$

Here, $\ddot{\boldsymbol{p}}$ is the acceleration of the UAV in the world frame $\mathcal{F}_W$. The right-hand side accounts for the total force acting on the

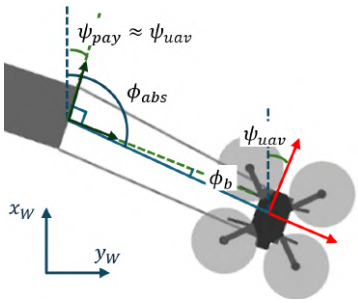


Figure 3: Top-down view of the azimuth angle definitions

UAV, expressed in the world frame. This includes the body-frame thrust $\boldsymbol{F}_f$ (which is transformed into the world frame by the rotation matrix $\boldsymbol{R}_f$), the world-frame tether force $\boldsymbol{F}_s$, and the gravity vector $\boldsymbol{g}$.

The equations of motion for the tether angles are derived from Newton's second law applied to the payload. The model accounts for gravity, tether tension, and the inertial effects of the UAV's acceleration, which is approximated by its thrust. The resulting vector equation is projected onto the spherical coordinate basis vectors ($\boldsymbol{u}_\theta$, $\boldsymbol{u}_\phi$) to yield the component-wise dynamics shown in Equations (6) and (7):

$$\ddot{\theta}_s = -\frac{g}{l}\sin\theta_s + \frac{\boldsymbol{R}_f\boldsymbol{F}_f \cdot \boldsymbol{u}_\theta}{ml} + \dot{\phi}_{abs}^2 \sin\theta_s \cos\theta_s \quad (6)$$

$$\ddot{\phi}_{abs} = \frac{\boldsymbol{R}_f\boldsymbol{F}_f \cdot \boldsymbol{u}_\phi}{ml\sin\theta_s} - 2\dot{\theta}_s\dot{\phi}_{abs}\cot\theta_s \quad (7)$$

To estimate the tether angles for the UKF's observation step, we use the UAV's IMU measurements. An on-board accelerometer measures the specific force, which is the total non-gravitational acceleration. From Equation (5), the specific force expressed in the world frame is $(\ddot{\boldsymbol{p}} - \boldsymbol{g})$. Since the IMU provides this measurement in the UAV's body frame ($\mathcal{F}_B$), we must transform this vector from the world frame to the body frame by multiplying it by the rotation matrix $\boldsymbol{R}_f^T$. Thus, the accelerometer output $\boldsymbol{a}_b$ is modeled as Equation (8):

$$\boldsymbol{a}_b = \frac{1}{m}\left(\boldsymbol{F}_f - \boldsymbol{R}_f^T\,\boldsymbol{F}_s\right) \quad (8)$$

To find the tether orientation, we first estimate the tension force vector $\boldsymbol{F}_s$. Rearranging Equation (8) isolates the tether force $\boldsymbol{F}_s$, which in the world frame can be expressed using the known thrust command $\boldsymbol{F}_f$ and measured acceleration $\boldsymbol{a}_b$, as shown in Equation (9):

$$\boldsymbol{F}_s = \boldsymbol{R}_f(\boldsymbol{F}_f - m\boldsymbol{a}_b) \quad (9)$$

The inclination angle θ_s' is then calculated as the angle between the estimated tension vector $\boldsymbol{F}_s$ and the vertical axis $\boldsymbol{e}_z$. According to the geometric definition of the dot product, the cosine of this angle is the dot product of the two vectors divided by the product of their magnitudes. Since $\boldsymbol{e}_z$ is a unit vector ($\|\boldsymbol{e}_z\| = 1$), we can solve for the angle θ_s' by taking the arccosine:

$$\theta_s' = \arccos\left(\frac{\boldsymbol{F}_s \cdot \boldsymbol{e}_z}{\|\boldsymbol{F}_s\|}\right) \quad (10)$$

Then, the body-relative azimuth angle ϕ_b' is computed as detailed in Equation (11). This is achieved by first calculating the world-frame azimuth from the components of the tension vector $\boldsymbol{F}_s$, and then converting it to the body-relative frame by subtracting the UAV's yaw angle ψ_{uav} and compensating for the coordinate system offset of $\frac{\pi}{2}$, consistent with the convention established in Equation (1):

$$\phi_b' = \operatorname{atan2}\left((\boldsymbol{F}_s)_y, (\boldsymbol{F}_s)_x\right) - \psi_{uav} - \frac{\pi}{2} \quad (11)$$

These estimated angles θ_s' and ϕ_b' are used as observation values in the UKF.

3 STATE ESTIMATION METHOD

3.1 *Tether Angle Estimation*

The following subsections describe in detail the system models used for the UKF, which, unlike the Extended Kalman Filter (EKF), does not require linearization and is therefore well suited for highly nonlinear systems such as the tethered UAV dynamics.

The estimated state of the system, $\hat{\boldsymbol{x}}$, at a discrete time k is defined as follows in Equation (12):

$$\hat{\boldsymbol{x}}^k = \begin{bmatrix}\hat{\theta}_s^k & \hat{\dot{\theta}}_s^k & \hat{\phi}_b^k & \hat{\dot{\phi}}_b^k\end{bmatrix}^T \quad (12)$$

The state transition is modeled by the discrete Wiener process acceleration model shown in Equation (13). This model integrates the state derivatives using the fourth-order Runge-Kutta method (RK_4) [14, 15].

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \mathrm{RK}_4(\boldsymbol{f}(\boldsymbol{x}^k)) + \boldsymbol{\Gamma}\boldsymbol{v}^k \quad (13)$$

where $\boldsymbol{f}(\boldsymbol{x}^k)$ is the state derivative vector, and $\boldsymbol{\Gamma}\boldsymbol{v}^k$ is the process noise term.

The vector $\boldsymbol{f}(\boldsymbol{x}^k)$ is derived by adapting the full physical dynamics from Equations (6) and (7) for implementation in the UKF. To improve computational efficiency and because the effects of these rotational forces are considered negligible for this application, the model is first simplified by neglecting the centrifugal and Coriolis force terms.

The model then incorporates the inertial force $(-m\boldsymbol{a}_0^k)$ resulting from the payload's acceleration, $\boldsymbol{a}_0$. This acceleration is not measured directly and is instead estimated by taking the UAV's absolute acceleration, $\boldsymbol{a}_d$, and subtracting the tether's relative acceleration from the previous step, $\boldsymbol{a}_{rel}(k-1)$, as shown in Equation (14):

$$\boldsymbol{a}_0(k) \approx \boldsymbol{a}_d(k) - \boldsymbol{a}_{rel}(k-1) \quad (14)$$

http://www.imavs.org/

http://www.imavs.org/

Here, the UAV's absolute acceleration is $\boldsymbol{a}_d = \boldsymbol{R}_f \boldsymbol{a}_b + \boldsymbol{g}$. The relative acceleration, $\boldsymbol{a}_{rel}$, is reconstructed from the estimated state using its spherical coordinate components, as defined in Equations (15) and (16):

$$\boldsymbol{a}_{rel} = a_r \boldsymbol{u}_r + a_\theta \boldsymbol{u}_\theta + a_\phi \boldsymbol{u}_\phi \tag{15}$$

$$a_r = -l\dot{\theta}_s^2, \quad a_\theta = l\ddot{\theta}_s, \quad a_\phi = l\ddot{\phi}_{abs} \sin\theta_s \tag{16}$$

Finally, the azimuth dynamics are converted to the body-relative frame used in the state vector. The state vector uses the body-relative azimuth ϕ_b instead of the absolute azimuth ϕ_{abs} because the UKF is formulated in the UAV body frame, and the relative angle directly relates to the tether's motion with respect to the UAV. Therefore, the body-relative angular acceleration is computed by subtracting the UAV's yaw acceleration from the absolute angular acceleration ($\ddot{\phi}_b = \ddot{\phi}_{abs} - \ddot{\psi}_{uav}$).

By incorporating these three modifications, the final state derivative vector $\boldsymbol{f}(\boldsymbol{x}^k)$ is constructed as shown in Equation (17):

$$\boldsymbol{f}(\boldsymbol{x}^k) = \begin{bmatrix} \hat{\dot{\theta}}_s^k \\ -\frac{g}{l}\sin\hat{\theta}_s^k + \frac{(\boldsymbol{R}_f \boldsymbol{F}_f - m\boldsymbol{a}_0^k)\cdot \boldsymbol{u}_\theta^k}{ml} \\ \hat{\dot{\phi}}_b^k \\ \frac{(\boldsymbol{R}_f \boldsymbol{F}_f - m\boldsymbol{a}_0^k)\cdot \boldsymbol{u}_\phi^k}{ml\sin\hat{\theta}_s^k} - \ddot{\psi}_{uav}^k \end{bmatrix} \tag{17}$$

Finally, the process noise model is defined. We assume two independent noise sources affecting the inclination and azimuth accelerations, respectively. These are represented by a zero-mean Gaussian noise vector $\boldsymbol{v}^k$, defined in Equation (18):

$$\boldsymbol{v}^k = \begin{bmatrix} v_\theta^k \\ v_\phi^k \end{bmatrix} \tag{18}$$

The covariance of this input noise is given by the diagonal matrix $\boldsymbol{\Sigma}$ in Equation (19). Its diagonal elements, σ_θ^2 and σ_ϕ^2, are the variances of the respective noise components v_θ^k and v_ϕ^k.

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_\theta^2 & 0 \\ 0 & \sigma_\phi^2 \end{bmatrix} \tag{19}$$

The noise distribution matrix $\boldsymbol{\Gamma}$, which maps the input noise vector $\boldsymbol{v}^k$ to the state space. Its terms, T_s and $T_s^2/2$, result from integrating acceleration noise into velocity and angle. It is defined in Equation (20):

$$\boldsymbol{\Gamma} = \begin{bmatrix} T_s^2/2 & 0 \\ T_s & 0 \\ 0 & T_s^2/2 \\ 0 & T_s \end{bmatrix} \tag{20}$$

The final process noise covariance matrix $\boldsymbol{Q}$ is then constructed by combining the distribution matrix $\boldsymbol{\Gamma}$ from Equation (20) and the input noise covariance $\boldsymbol{\Sigma}$ from Equation (19) using the sandwich product, as shown in Equation (21):

$$\boldsymbol{Q} = \boldsymbol{\Gamma}\boldsymbol{\Sigma}\boldsymbol{\Gamma}^T \tag{21}$$

In the UKF, the process noise covariance $\boldsymbol{Q}$ is added during prediction. The Kalman gain is then computed as $\boldsymbol{K} = \boldsymbol{P}xy\boldsymbol{P}yy^{-1}$, where $\boldsymbol{P}xy$ denotes the cross-covariance between state and measurement, and $\boldsymbol{P}yy$ is the predicted measurement covariance. This gain blends the predicted state with measurements, and the updated covariance $\boldsymbol{P}$ reflects the reduced estimation uncertainty.

3.2 Geometric Payload Pose Estimation

This section describes the method for estimating the 6-DOF (degrees of freedom) state of the payload using geometric relationships. The estimation relies primarily on the tether angle states for each UAV, as obtained from the UKF.

Payload Position ($\boldsymbol{p}_p$) First, the 3D position of each payload anchor point, $\boldsymbol{p}_{pa,i}$, is calculated. The tether's unit direction vector, $\hat{\boldsymbol{r}}_i$, is reconstructed from the estimated angles ($\hat{\theta}_{s,i}$, $\hat{\phi}_{b,i}$) and the UAV's yaw ($\psi_{uav,i}$), consistent with the definitions in Equations (1) and (2). The anchor point position is then found via forward kinematics as shown in Equation (22):

$$\boldsymbol{p}_{pa,i} = \boldsymbol{p}_{ua,i} + l_i \boldsymbol{u}_{r,i} \tag{22}$$

where $\boldsymbol{p}_{ua,i}$ is the UAV's anchor position. The payload's center, $\boldsymbol{p}_p$, is then calculated as the midpoint of these two points in Equation (23):

$$\boldsymbol{p}_p = \frac{\boldsymbol{p}_{pa,1} + \boldsymbol{p}_{pa,2}}{2} \tag{23}$$

Roll Angle (φ_p) The payload's roll angle, φ_p, is determined by the angle of the vector connecting the two anchor points ($\boldsymbol{p}_{pa,2} - \boldsymbol{p}_{pa,1}$) when projected onto the world frame's y_W–z_W plane, as calculated in Equation (24):

$$\varphi_p = \text{atan2}\left((\boldsymbol{p}_{pa,2} - \boldsymbol{p}_{pa,1})_z, (\boldsymbol{p}_{pa,2} - \boldsymbol{p}_{pa,1})_y\right) \tag{24}$$

Pitch Angle (θ_p) The pitch angle, θ_p, is controlled by the winch mechanism (Section 2). It is therefore estimated directly from a pre-calibrated function of the winch encoder measurements, as shown conceptually in Equation (25):

$$\theta_p = f_{\text{winch}}(l_{\text{winch},1}, l_{\text{winch},2}) \tag{25}$$

Yaw Angle (ψ_p) The payload's yaw, ψ_p, is assumed to align with the average horizontal direction of the tethers. First, the absolute compass heading of each tether, $\phi_{r,i}$, is calculated in Equation (26):

$$\phi_{r,i} = \psi_{uav,i} + \phi_{b,i} \tag{26}$$

The payload's yaw is then estimated as the circular average of these two headings to correctly handle angle wrapping, as shown in Equation (27):

$$\psi_p = \text{atan2}\left(\sin\phi_{r,1} + \sin\phi_{r,2}, \cos\phi_{r,1} + \cos\phi_{r,2}\right) \tag{27}$$

This completes the geometric estimation of the payload's 6-DOF pose.

http://www.imavs.org/

4 VERIFICATION WITH SIMULATION

To validate the proposed state estimator, a simulation environment was constructed using MATLAB/Simulink and Simscape Multibody. The physical and filter design parameters used in the simulation are summarized in Table 1.

Reference trajectories for the payload were predefined, and UAV trajectories were computed by inverse kinematics and tracked using local PID control, making the payload motion purely feedforward-driven. The UKF was run offline using logged simulation data. Sensor noise was excluded; robustness will be evaluated in real experiments.

Attitude and position maneuvers were performed, and the logged data provided inputs for the UKF and ground-truth tether angles for evaluation.

4.1 Estimation During Attitude Maneuvers

To evaluate the estimation performance, the payload was commanded to perform a series of attitude maneuvers. Based on this simulation data, the proposed tether angle estimation was conducted offline. The results are presented in Figure 4, which shows the estimated inclination θ_s and azimuth ϕ_b angles for the tethers of each UAV.

The inclination angle θ_s was estimated with high accuracy throughout the maneuver. For the azimuth angle ϕ_b, although the overall estimation is acceptable, transient errors are observed at the beginning of attitude changes. These discrepancies are attributed to the inertial effects neglected in the simplified dynamics model. However, since the estimation converges shortly afterward, the estimator is considered effective during attitude maneuvers. The payload's attitude, calculated from these tether angle estimates, was compared to the ground truth (Figure 5). The high-precision tracking confirms the effectiveness of the proposed method and validates the underlying tether angle estimator.

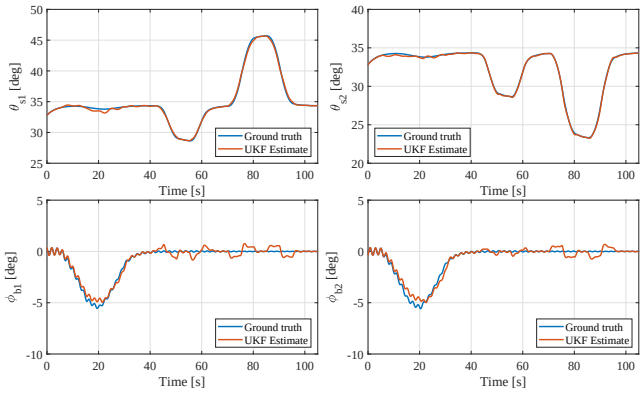
4.2 Estimation During Position Maneuvers

The performance during payload position maneuvers is shown in Figure 6. The estimated azimuth angle ϕ_b captured the payload's pendulum-like swing, seen as x-axis oscillations in Figure 7. The inclination angle θ_s is mostly accurate, but errors appear during vertical (z-axis) motion. This discrepancy is likely due to unmodeled dynamics at the payload-side tether attachment points. Furthermore, the payload's position was calculated from these tether estimates, and the result is compared to the ground truth in Figure 7. Since the estimation errors in the tether angles were minor, the resulting payload position also tracks the true trajectory with high accuracy. This confirms that the proposed estimation method is sufficiently precise for payload state tracking during position maneuvers.

Notably, the estimation of ϕ_b is more accurate here than during attitude maneuvers. This is because the primarily translational motion during position maneuvers aligns better with the model's assumptions than the complex rotational motion seen during attitude maneuvers.

Physical Parameters	
Payload size	$0.2\text{m} \times 0.2\text{m} \times 0.4\text{m}$
Payload mass	0.5kg
UAV mass	1.64kg
Tether length	1.2m
UKF Design Parameters	
Proc. noise std. (σ_θ)	1.0×10^{-3}
Proc. noise std. (σ_ϕ)	5.0×10^{-4}
Obs. noise cov. ($\boldsymbol{R}$)	$\text{diag}\left(1.0 \times 10^{-3}, 1.0 \times 10^{-3}\right)$

Table 1: Parameters used in the simulation



(a) Tether angles: UAV1 (b) Tether angles: UAV2

Figure 4: Tether angle estimation during attitude maneuvers

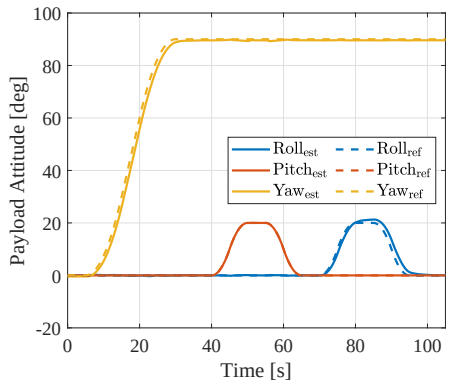


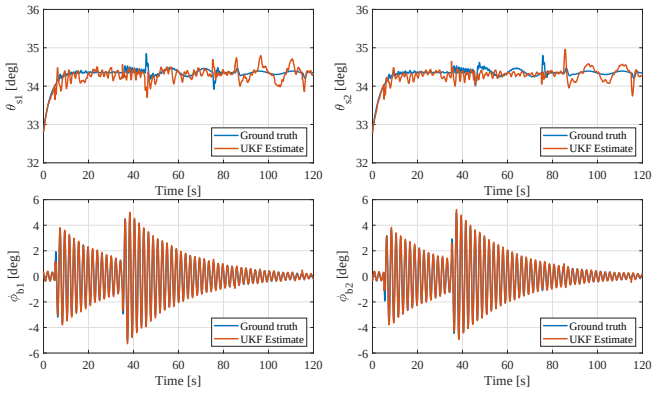
Figure 5: Estimation result of payload attitude

5 VERIFICATION WITH EXPERIMENT

5.1 Experimental Setup

Flight experiments were conducted to verify the proposed estimation method in a real-world environment. The experiments aimed to evaluate its accuracy under sensor noise and environmental disturbances.

The experiments were conducted in a Motion Capture (MoCap) environment with two quadrotors carrying a pay-



(a) Tether angles: UAV1 (b) Tether angles: UAV2

Figure 6: Tether angle estimation during position maneuvers

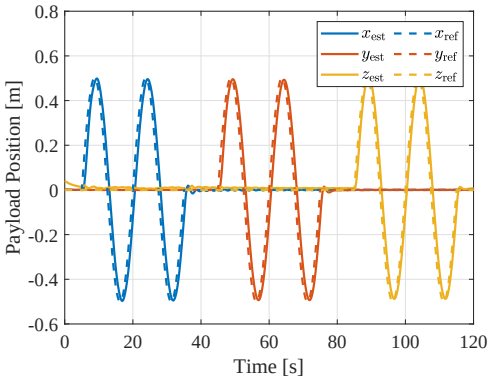


Figure 7: Estimation result of payload position

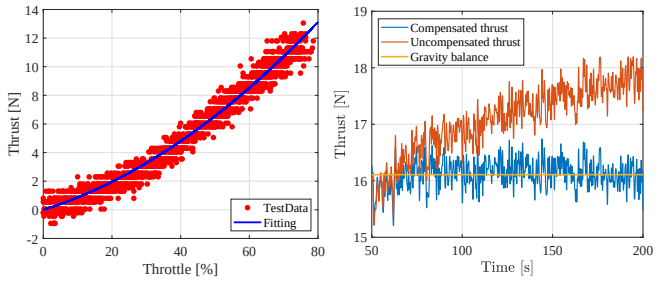
load. The MoCap system provided ground-truth data for evaluation and position feedback for the UAVs, which followed predefined trajectories. UAV synchronization was achieved through a centralized control system using ROS(Robot Operating System) over Wi-Fi. While the physical parameters were identical to the simulation (Table 1), the UKF design parameters were empirically tuned to account for the differing sensor noise and unmodeled disturbances present in the real experiment. The final values are summarized in Table 2.

5.2 *Thrust Estimation Model and Voltage Compensation*

The accuracy of the proposed estimation method depends on an accurate thrust value $\boldsymbol{F}_f$ as an input to the prediction model (Equations (6) and (7)) and the observation model (Equation (9)). However, direct in-flight measurement of

UKF Design Parameters	
Proc. noise std. (σ_θ)	2.5×10^{-2}
Proc. noise std. (σ_ϕ)	1.0×10^{-2}
Obs. noise cov. ($\boldsymbol{R}$)	diag $(0.25, 0.20)$

Table 2: UKF parameters tuned for the experiment



(a) Static thrust test result (b) Voltage-compensated thrust

Figure 8: Experimental analysis of thrust estimation

thrust is infeasible. While prior work has treated thrust as a constant value [14], this assumption is not valid for our dynamic maneuvers and would degrade estimation accuracy. Therefore, we adopted a model that compensates for the battery voltage drop in-flight. This involved a two-step process.

First, a static thrust test was conducted to establish a baseline thrust-throttle relationship, shown in Figure 8a. The result was approximated by Equation (28):

$$\|\boldsymbol{F}_f\| = 0.0011x^2 + 0.76x \tag{28}$$

where $\|\boldsymbol{F}_f\|$ is the generated thrust [N], and x is the throttle percentage [%].

Second, to compensate for battery voltage drop, this baseline model was adjusted in real-time using the measured voltage V, as shown in Equation (29):

$$\|\boldsymbol{F}_f\| = 0.8\left(\frac{V}{16.0}\right)^2 (0.0011x^2 + 0.76x) \tag{29}$$

where V is the measured battery voltage, $V_0 = 16.0$ V is the nominal voltage used in the calibration, and the factor 0.8 reflects the saturation at 80% throttle, introduced as a safety margin to prevent the propulsion system from exceeding its current rating.

The validation was performed in hover, where the thrust can be directly inferred from the equilibrium of forces, providing a reliable baseline for evaluating the model. The estimated thrust from both the original and compensated equations was compared with the gravitational force. The comparison in Figure 8b shows the compensated model more accurately captures the true thrust required for hovering, validating the voltage compensation approach.

5.3 *Experimental Results and Discussion*

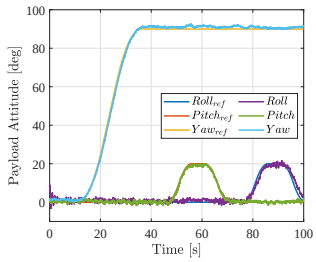
An attitude maneuver experiment was conducted, as shown in Figure 9a. The resulting payload attitude is presented in Figure 9b. Figures 9c and 9d show the accelerations of the UAVs, while Figure 10 shows the thrust from the voltage-compensated model (Equation (29)).

The tether angles and 6-DOF payload state were estimated from the flight data, as shown in Figure 11 and Figure 12. A quantitative evaluation is summarized in Table 3.

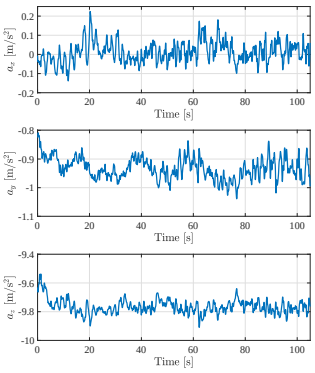
http://www.imavs.org/

http://www.imavs.org/

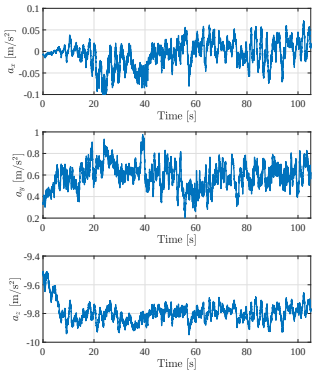
(a) Overview of the experiment



(b) Result of payload attitude



(c) UAV1 acceleration



(d) UAV2 acceleration

Figure 9: Flight test for payload attitude maneuver

The results confirm the proposed framework is effective, achieving attitude RMSE within a few degrees and position RMSE within several centimeters.

The primary performance limitation observed was high-frequency oscillations in the estimates, a phenomenon not present in the simulation. These oscillations can be traced directly to high-frequency noise in the input signals, such as the measured UAV accelerations (Figures 9c and 9d) and the thrust estimate (Figure 10). This noise propagates through the model, causing oscillations in the final estimates. This effect was most pronounced in the lateral and rotational axes, where the Y-axis exhibited the largest position RMSE (4.30 cm), alongside significant errors in Roll and Yaw. In contrast, the Pitch estimate remained highly accurate with an RMSE of only 0.76 deg. This accuracy stems from its winch-based model, independent of the noisy IMU and thrust inputs, reinforcing the conclusion that input signal quality is the dominant error source for the other states. However, this simplified model has its own limitations, as it did not fully capture the minor pitching oscillations caused by the payload's natural swing from aerodynamic effects. Additionally, the omission of centrifugal and Coriolis terms, though less significant than sensor noise, remains a modeling limitation for more dynamic maneuvers.

State	RMSE	Max Error
Attitude Errors [deg]		
Roll	2.392	8.363
Pitch	0.760	2.108
Yaw	2.034	6.811
Position Errors [m]		
X	0.0271	0.1230
Y	0.0430	0.1328
Z	0.0258	0.0717

Table 3: RMSE and maximum error of payload estimation

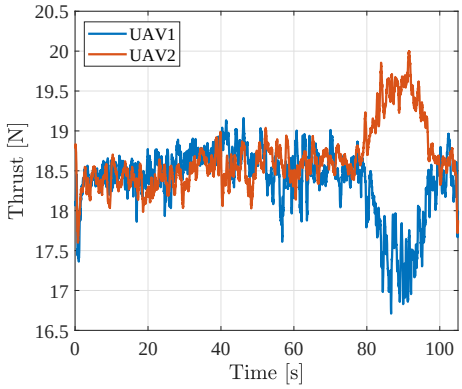
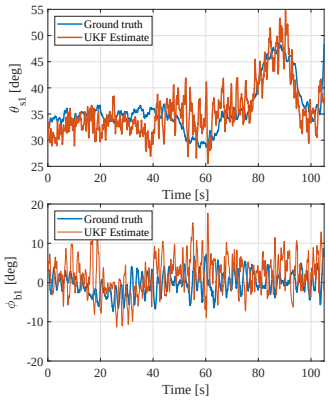
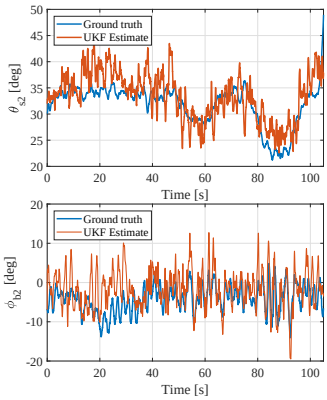


Figure 10: Thrust calculations during experiment



(a) Tether angles: UAV1



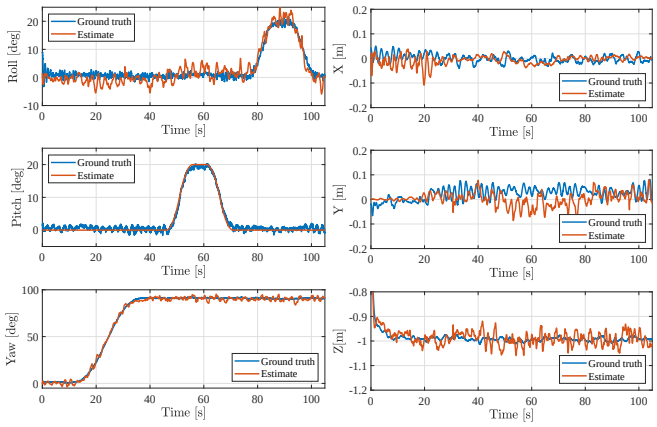
(b) Tether angles: UAV2

Figure 11: Tether angle estimation in experiment

In summary, this experiment validates the proposed strategy under real-world conditions and identifies input signal quality as the main performance limitation. Future work should therefore focus on improved input filtering and more robust dynamic modeling for practical deployment.

6 CONCLUSION

This study developed and validated a UKF-based estimation framework for a cooperative manipulation system. We successfully demonstrated that the payload's 6-DOF state can be determined with practical accuracy using only each UAV's onboard IMU and position data. The method's effectiveness



(a) Estimated payload attitude (b) Estimated payload position

Figure 12: Payload state estimation in experiment

was confirmed through both dynamic simulations and real-world flight experiments, establishing a baseline for sensor-based payload state awareness in such systems.

The experimental validation highlighted the primary challenges for real-world deployment: performance degradation due to unmodeled dynamics, such as aerodynamic effects, and the propagation of sensor noise. A key challenge for real-time implementation is filter divergence, which requires robust modeling and recovery strategies. Future work will refine the dynamic model and implement the estimator in a real-time control loop. Successfully tackling these challenges is a critical step toward the practical deployment of autonomous and dynamic cooperative manipulation systems.

REFERENCES

[1] Anibal Ollero, Marco Tognon, Alejandro Suarez, Dongjun Lee, and Antonio Franchi. Past, present, and future of aerial robotic manipulators. *IEEE Transactions on Robotics*, 38(1):626–645, 2022.

[2] Miguel Ángel Trujillo, José Ramiro Martínez-de Dios, Carlos Martín, Antidio Viguria, and Aníbal Ollero. Novel aerial manipulator for accurate and robust industrial ndt contact inspection: A new tool for the oil and gas inspection industry. *Sensors*, 19(6), 2019.

[3] Ivana Palunko, Rafael Fierro, and Patricio Cruz. Trajectory generation for swing-free maneuvers of a quadrotor with suspended payload: A dynamic programming approach. In *2012 IEEE International Conference on Robotics and Automation*, pages 2691–2697, 2012.

[4] Stamatina C. Barakou, Costas S. Tzafestas, and Kimon P. Valavanis. A survey of modeling and control approaches for cooperative aerial manipulation. In *2024 International Conference on Unmanned Aircraft Systems (ICUAS)*, pages 610–617, 2024.

[5] Koushil Sreenath and Vijay Kumar. Dynamics, control and planning for cooperative manipulation of payloads suspended by cables from multiple quadrotor robots. In *Robotics: Science and Systems*, 2013.

[6] Kristian Klausen, Chris Meissen, Thor I. Fossen, Murat Arcak, and Tor Arne Johansen. Cooperative control for multirotors transporting an unknown suspended load under environmental disturbances. *IEEE Transactions on Control Systems Technology*, 28(2):653–660, 2020.

[7] Dario Sanalitro, Heitor J. Savino, Marco Tognon, Juan Cortés, and Antonio Franchi. Full-pose manipulation control of a cable-suspended load with multiple uavs under uncertainties. *IEEE Robotics and Automation Letters*, 5(2):2185–2191, 2020.

[8] Junyi Geng and Jack W. Langelaan. Cooperative transport of a slung load using load-leading control. *Journal of Guidance, Control, and Dynamics*, 43(7), 2020.

[9] Shubham Barawkar and K. Kumar. Force-torque (f-t) based multi-drone cooperative transport using fuzzy logic and low-cost and imprecise f-t sensor. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 237(11):2517–2544, 2023.

[10] Yasheng Li, Sifan Ge, and Giuseppe Loianno. Cooperative transportation of cable suspended payloads with mavs using monocular vision and inertial sensing. *IEEE Robotics and Automation Letters*, 6(3):516–523, 2021.

[11] Sara Tang, Thomas Wüest, and Vijay Kumar. Aggressive flight with suspended payloads using vision-based control. *IEEE Robotics and Automation Letters*, 3(4):3410–3417, 2018.

[12] Heng Xie, Kaixu Dong, and Pakpong Chirarattananon. Cooperative transport of a suspended payload via two aerial robots with inertial sensing. *IEEE Access*, 10:81764–81776, 2022.

[13] Yuichi Takemura, Yuki Ohira, and Satoshi Suzuki. Reserch on cooperative manipulation by multiple drones with winch mechanisms. In *Proceedings of the 42nd Annual Conference of the Robotics Society of Japan (RSJ2024)*, pages 3J4–01, 2024. (In Japanese).

[14] Sergei Lupashin and Raffaello D'Andrea. Stabilization of a flying vehicle on a taut tether using inertial sensing. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2432–2438, 2013.

[15] Yaakov Bar-Shalom, X. Rong Li, and Thia Kirubarajan. *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, Inc., New York, NY, 2002.

http://www.imavs.org/