IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

Cooperative Payload Estimation by a Team of Mocobots

Haoxuan Zhang, C. Lin Liu, *Graduate Student Member, IEEE*, Matthew L. Elwin, Randy A. Freeman, *Member, IEEE*, and Kevin M. Lynch, *Fellow, IEEE*

***Abstract*—For high-performance autonomous manipulation of a payload by a mobile manipulator team, or for collaborative manipulation with the human, robots should be able to discover where other robots are attached to the payload, as well as the payload's mass and inertial properties. In this letter, we describe a method for the robots to autonomously discover this information. The robots cooperatively manipulate the payload, and the twist, twist derivative, and wrench data at their grasp frames are used to estimate the transformation matrices between the grasp frames, the location of the payload's center of mass, and the payload's inertia matrix. The method is validated experimentally with a team of three mobile cobots, or mocobots.**

***Index Terms*—inertial property, payload estimation, mocobot, robot cooperation, human-robot collaboration.**

I. Introduction

CONSIDER a scenario in manufacturing, logistics, or construction where a large, substantially rigid payload must be manipulated in all six degrees of freedom (DoF), perhaps for an assembly or loading task. Multiple distributed contacts with the payload are required to respect the wrench limits of any single manipulator and to minimize stress concentrations for heavy or fragile payloads. If the manipulation task is not one that is easily automated (e.g., it is not a repetitive task performed in a structured environment), then one or more human operators can physically collaborate with a team of mobile cobots, or mocobots [1]. The partnership combines the mocobots' physical strength with human perception and adaptability (Fig. 1).

Received 19 May 2025; accepted 24 July 2025. Date of publication 11 August 2025; date of current version 20 August 2025. This article was recommended for publication by Associate Editor L. Liu and Editor M. Ani Hsieh upon evaluation of the reviewers' comments. This work was supported by Northwestern University and NSF under Grant CMMI-2024774. *(Corresponding author: Kevin M. Lynch.)*

Haoxuan Zhang, C. Lin Liu, and Matthew L. Elwin are with the Department of Mechanical Engineering, Center for Robotics and Biosystems, Northwestern University, Evanston, IL 60208 USA (e-mail: haoxuan.zhang@u.northwestern.edu; lin.liu@u.northwestern.edu; elwin@northwestern.edu).

Randy A. Freeman is with the Center for Robotics and Biosystems, Department of Electrical and Computer Engineering, Northwestern Institute on Complex Systems, Northwestern University, Evanston, IL 60208 USA (e-mail: freeman@northwestern.edu).

Kevin M. Lynch is with the Department of Mechanical Engineering, Center for Robotics and Biosystems, Northwestern Institute on Complex Systems, Northwestern University, Evanston, IL 60208 USA (e-mail: kmlynch@northwestern.edu).

This article has supplementary downloadable material available at https://doi.org/10.1109/LRA.2025.3597893, provided by the authors.

Digital Object Identifier 10.1109/LRA.2025.3597893

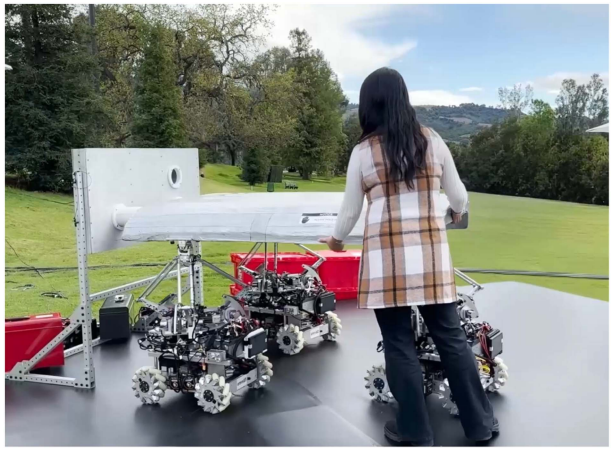

Fig. 1. Three Omnid mocobots collaborate safely and intuitively with a human operator on a simulated plane wing assembly task.

First, a human guides the mobile manipulators to grasp the common payload. Once the grasps are established, to provide optimal model-based assistance to the human, or for high-performance autonomous manipulation, the robots should be able to discover where the other robots are attached to the payload, as well as the payload's mass and inertial properties.

In this letter, we describe a method for the robots to autonomously discover this information. The robots cooperatively manipulate the payload, and the twist, twist derivative, and wrench data at their grasp frames are used to estimate the transformation matrices between the grasp frames, the location of the payload's center of mass, and the payload's inertia matrix. The method is validated experimentally with a team of three mocobots.

A. Related Work

A "mocobot" is a mobile variant of a "cobot," robots designed for physical collaboration with humans, originally introduced in [2]. Mocobots enhance cobots with mobility for diverse tasks. An example is the Omnid mocobot, designed for human-multirobot collaborative manipulation [1].

Cooperative robot manipulation involves multiple robots working together to manipulate objects, with a variety of applications in manufacturing, construction, and hazardous environments [3], [4], [5], [6]. Accurate knowledge of a payload's inertial properties enables planning for dynamic manipulation and high-performance control [7], [8], [9]. Various cooperative

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

control architectures, including centralized and decentralized controllers, have been proposed for manipulation and transportation under the assumption of known inertial properties [10], [11], [12]. This assumption motivates estimation of such properties for unknown payloads.

A recent survey [7] provides a comprehensive review of estimation of payload inertial properties, divided into three main categories: purely visual methods, exploratory methods, and methods where the payload is rigidly attached to or grasped by a robot manipulator. The purely visual method relies only on static images or video of the object. For example, the inertial parameters of an object can be estimated visually from the size and shape of the object under assumptions on its density properties [13]. Most visual methods require prior knowledge or assumptions, however, such as knowledge of the total mass of the object or an assumption of uniform density. Exploratory methods include pushing, poking, or tilting an object (e.g., [14], [15]), but motions are often limited to a plane and only a subset of inertial properties can be estimated. When a payload is firmly attached to a robot manipulator, Newton-Euler equations can be used to estimate payload inertial properties from robot motion and wrist force-torque data [16]. The basic idea can be extended to grasps by multiple robots, as in [17], [18] for a planar payload and [19] for a 3D payload. The approach and goals of [19] are closely related to those of this letter, except [19] does not consider the orientations of the robots' grasp frames and the approach was tested only in simulation.

B. Contribution

This letter presents a methodology enabling a group of mobile manipulators to estimate properties of an unknown rigid-body payload. These properties include the payload's mass, center of mass, and inertia matrix, as well as the transformation matrices between the robots' grasp frames, which are initially unknown. We have performed extensive experimental validation with mocobots and their force-controlled manipulators, which are well suited to cooperative manipulation and do not have expensive force-torque sensors at their wrists. Results from these experiments highlight the method's robustness and practicality.

II. Cooperative Rigid Payload Estimation Problem Formulation

A team of N robots grasps the rigid body, defining the coordinate frames $\{1\} \dots \{N\}$ at the grasp locations. Let $T_{i_0 i} \in SE(3)$ define the transformation matrix describing frame $\{i\}$, the current location of robot i's grasp, relative to a frame $\{i_0\}$, defined as robot i's "home" configuration. The configuration $T_{i_0 i}$ can be measured by encoders or other sensors on the robot, but the robot has no exteroceptive sensors to directly sense its location relative to other robots or a common world frame.

To determine the properties of the payload and their relative grasp locations, the robots cooperatively manipulate the payload. For example, each robot may attempt to drive its gripper along a periodic reference trajectory using a soft impedance controller. The combination of the reference trajectories should cause the rigid payload to move in all six degrees of freedom, while the impedance control adapts the actual robot trajectories to ensure safe manipulation forces given the unknown connections of the robots to the rigid payload.

Each robot takes measurements at its grasp interface synchronously with the other robots during manipulation. For example, robot i measures the configuration $T_{i_0 i}$, the twist $\mathcal{V}_i = (\omega_i, v_i) \in \mathbb{R}^6$ measured in $\{i\}$, its time derivative $\dot{\mathcal{V}}_i$, and the wrench $\mathcal{F}_i = (m_i, f_i) \in \mathbb{R}^6$ measured in $\{i\}$. A complete data point for robot i is defined as the tuple $\mathcal{D}_i = \{T_{i_0 i}, \mathcal{V}_i, \dot{\mathcal{V}}_i, \mathcal{F}_i\}$. The set of data points collected by robot i at all Q timesteps is denoted $\mathcal{D}_{i*} = \{\mathcal{D}_{iq} \mid q = 1 \dots Q\}$, and the set of data points collected by all robots at timestep q is denoted $\mathcal{D}_{*q} = \{\mathcal{D}_{iq} \mid i = 1 \dots N\}$. The complete data set is denoted $\mathcal{D}_{**} = \{\mathcal{D}_{iq} \mid i = 1 \dots N, q = 1 \dots Q\}$.

The cooperative rigid payload estimation problem can be formulated as follows: given $\mathcal{D}_{**}$, determine the payload's mass $\mathfrak{m}$; the configuration of a frame $\{\mathrm{c}\}$ at the payload's center of mass relative to each robot's grasp frame $(T_{1c}, \dots, T_{Nc})$ such that the frame $\{\mathrm{c}\}$ is aligned with principal axes of inertia of the payload; and the 3×3 positive-definite inertia matrix $\mathcal{I}$ of the payload in the frame $\{\mathrm{c}\}$. This information also implies the configuration of each robot's grasp relative to the others,

$$T_{ij} = \begin{bmatrix} R_{ij} & p_{ij} \\ 0 & 1 \end{bmatrix} \in SE(3),\ i, j \in \{1, \dots, N\}.$$

Payload estimation requires that the data set $\mathcal{D}_{**}$ be sufficiently rich, e.g., the manipulation must cause payload rotations that make estimation of the grasp kinematics and inertial properties well posed. Grasp twist and acceleration data $\mathcal{V}_i$ and $\dot{\mathcal{V}}_i$ may be obtained by using filtered encoder data, IMUs, accelerometers, or a combination, and the wrench $\mathcal{F}_i$ may be obtained by end-effector force-torque sensors, force/torque sensors at individual robot joints, or other means.

III. Estimation of Payload Properties

A. Sequential Estimation

In principle, all data $\mathcal{D}_{**}$ could be used in a single optimization to simultaneously calculate the configurations of the robot grasps relative to each other, the mass and location of the center of mass of the payload, and the inertia of the payload. In this letter, we adopt a sequenced approach, where we first estimate the grasp kinematics using only twist measurements; then estimate the mass and center of mass using the results of the kinematics estimates and wrench measurements when the payload is held stationary; and finally estimate the inertial properties using the results of the previous estimates as well as twist, acceleration, and wrench data during manipulation. This approach (a) allows us to use simple least-squares estimation and (b) requires only the data needed for the particular parameters being estimated; for example, wrench and acceleration data are not needed to estimate the robots' relative grasp frame configurations.

B. Grasp Kinematics

When robots i and j grasp a common rigid body, the twists $\mathcal{V}_i$ and $\mathcal{V}_j$ are related by $\mathcal{V}_i = [\mathrm{Ad}_{T_{ij}}]\mathcal{V}_j$, where $[\mathrm{Ad}_{T_{ij}}] \in \mathbb{R}^{6\times 6}$

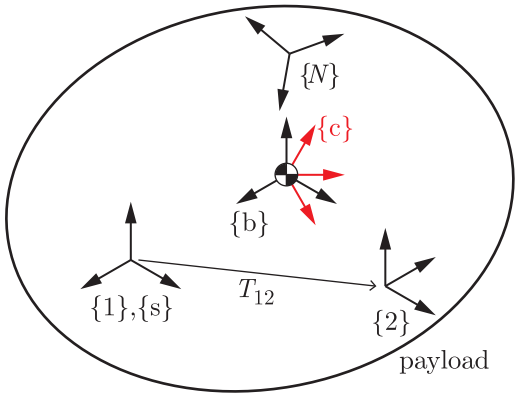

Fig. 2. The N robots grasp the payload at frames $\{1\},\ldots,\{N\}$, and one frame (e.g., $\{1\}$) is chosen as a standard reference frame $\{\mathrm{s}\}$. The frame $\{\mathrm{b}\}$ is located at the center of mass and aligned with $\{\mathrm{s}\}$, and $\{\mathrm{c}\}$ is located at the center of mass and aligned with the principal axes of inertia. The transformation matrix T_{ij} locates frame $\{j\}$ relative to $\{i\}$.

is the adjoint representation of the transformation matrix T_{ij} describing the configuration of the frame $\{j\}$ relative to $\{i\}$ [20]. Frames are illustrated in Fig. 2. We write the complete set of twist measurements as $\mathcal{V}_{i*}, \mathcal{V}_{j*} \in \mathbb{R}^{6\times Q}$, i.e., each individual twist measurement $\mathcal{V}_i$ forms a column of the matrix $\mathcal{V}_{i*}$. Then

$$\mathcal{V}_{i*} = [\mathrm{Ad}_{T_{ij}}]\mathcal{V}_{j*}, \tag{1}$$

or, in expanded form,

$$\begin{bmatrix}\omega_{i*}\\ v_{i*}\end{bmatrix} = \begin{bmatrix}R_{ij} & 0\\ [p_{ij}]R_{ij} & R_{ij}\end{bmatrix}\begin{bmatrix}\omega_{j*}\\ v_{j*}\end{bmatrix}, \tag{2}$$

where $[p_{ij}] \in so(3)$ is the skew-symmetric representation of $p_{ij} \in \mathbb{R}^3$. This equation is the basis for estimating the rotation matrix R_{ij} and the position vector p_{ij} between the two robots.

1) Rotation Matrix: Let $\omega_{i*}, \omega_{j*} \in \mathbb{R}^{3\times Q}$ be the angular velocities of the twists measured at each grasp. The relationship between the two sets is

$$\omega_{i*} = R_{ij}\omega_{j*}. \tag{3}$$

The problem is to estimate a rotation matrix $\hat{R}_{ij}$ that best fits the data in Equation (3) by solving the following least-squares problem based on the Frobenius norm:

$$\hat{R}_{ij} = \underset{R_{ij}\in SO(3)}{\operatorname{argmin}} \|R_{ij}\omega_{j*} - \omega_{i*}\|_F^2. \tag{4}$$

This problem is known as Wahba's problem, which has been solved by the Kabsch-Umeyama algorithm [21], [22], [23]. Defining the matrix $X = \omega_{j*}\omega_{i*}^{\intercal}$, utilizing SVD decomposition to get $X = U\Sigma V^{\intercal}$, and defining $S = \mathrm{diag}(1, 1, \det(VU^{\intercal}))$, the solution to the optimization problem is

$$\hat{R}_{ij} = VSU^{\intercal}. \tag{5}$$

At least three pairs of angular velocities are required to uniquely determine the rotation matrix R_{ij}. We assume the measurement data are sufficiently rich to ensure that X is full rank.

2) Position Vector: Let $v_{i*}, v_{j*} \in \mathbb{R}^{3\times Q}$ be the linear velocities of the twists measured at each grasp. By (2), the linear velocities satisfy

$$v_{i*} = [p_{ij}]R_{ij}\omega_{j*} + R_{ij}v_{j*}. \tag{6}$$

Due to the skew-symmetric property $[\omega]p = -[p]\omega$, (6) can be rearranged to

$$v_{i*} = -[R_{ij}\omega_{j*}]p_{ij} + R_{ij}v_{j*}. \tag{7}$$

Plugging in the estimate $\hat{R}_{ij}$, the estimate $\hat{p}_{ij}$ is found by solving the least-squares problem

$$\hat{p}_{ij} = \underset{p_{ij}}{\operatorname{argmin}} \|[\hat{R}_{ij}\omega_{j*}]p_{ij} - (\hat{R}_{ij}v_{j*} - v_{i*})\|_2, \tag{8}$$

provided $[\hat{R}_{ij}\omega_{j*}]$ is full rank.

The solution $\hat{T}_{ij} = (\hat{R}_{ij}, \hat{p}_{ij})$, via (4) and (8), depends only on pairwise robot data. To obtain a complete and consistent representation of the robots' relative grasp frames, $N-1$ such solutions are needed, e.g., $\hat{T}_{12}, \hat{T}_{13}, \ldots, \hat{T}_{1N}$. Then the configuration of any frame $\{i\}$ relative to another frame $\{j\}$ may be calculated as $\hat{T}_{ij} = \hat{T}_{1i}^{-1}\hat{T}_{1j}$. While this method permits efficient linear least-squares computation, it does not simultaneously take into account all combinations of robots' twist data nor loop-closure constraints among three or more robots, e.g., $T_{ij}T_{jk}T_{ki} = I$. Such constraints allow squeezing more information out of the collected data at the cost of greater computational complexity and nonlinear optimization.

We employed an iterative gradient-based nonlinear optimization to incorporate all the data and loop constraints to further refine the $\hat{T}_{ij}$ estimates from the original linear least-squares solutions. The optimization minimizes the weighted sum of squared errors from data from all pairwise combinations of grasping frames and all closed loops. Our implementation uses the scipy.optimize library and the Broyden-Fletcher-Goldfarb-Shanno method for estimating gradients [24], but other nonlinear optimization methods incorporating combinatorial loop-closure constraints could also be employed [25], [26].

C. Mass and Center of Mass (CoM)

The mass and center of mass are estimated using the grasp kinematics solution and a set of static wrench measurements. A common reference frame $\{\mathrm{s}\}$ is chosen from among the grasp frames $\{\{1\},\ldots,\{N\}\}$; for example, $\{\mathrm{s}\}$ may be chosen to be coincident with $\{1\}$ (Fig. 2). We also define a frame $\{\mathrm{s}_0\}$ coincident with $\{\mathrm{s}\}$ but oriented such that its $\hat{z}$-axis is opposite the gravity vector $\mathfrak{g} \in \mathbb{R}^3$. The center of mass of the payload is located at the origin of a to-be-estimated frame $\{\mathrm{c}\}$.

With local wrenches $\mathcal{F}_i = (m_i, f_i)$ from each robot, the force and moment static equilibrium conditions are

$$\mathfrak{m}\mathfrak{g} = -\sum_{i=1}^{N} R_{s_0 i} f_i \tag{9}$$

$$[p_{sc}](R_{ss_0}\mathfrak{m}\mathfrak{g}) = -\sum_{i=1}^{N}\left([p_{si}](R_{si}f_i) + R_{si}m_i\right), \tag{10}$$

where $\mathfrak{m}$ is the mass of the payload, f_i is the measured force component of $\mathcal{F}_i$, and p_{sc} and p_{si} represent the origin of $\{\mathrm{c}\}$ and $\{i\}$ in $\{\mathrm{s}\}$ coordinates, respectively.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

The mass $\mathfrak{m}$ can be estimated using (9) and one or more static measurements by the robots, while p_{sc} can be estimated using (10) and two or more static measurements, provided the orientations of the payload during measurement differ by rotation about an axis not aligned with the gravity vector $\mathfrak{g}$.

Since the robots rigidly grasp the payload, p_{sc}, p_{si}, and R_{si} are constant, and only p_{sc} remains unknown. Given $q = 1 \dots Q$ measurements by the robots, we define the scalars

$$A_q = [0,0,1]\mathfrak{g}, \quad b_q = -[0,0,1]\left(\sum_{i=1}^{N} R^q_{s_0 i} f^q_i\right) \tag{11}$$

and the vectors

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_Q \end{bmatrix} \in \mathbb{R}^{Q\times 1}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_Q \end{bmatrix} \in \mathbb{R}^{Q\times 1}. \tag{12}$$

Plugging this data into the $\hat{z}$-component of Equation (9), we get

$$A\mathfrak{m} = b, \tag{13}$$

which can be solved for the mass estimate $\hat{\mathfrak{m}}$ using least squares.

Similarly, the data can be plugged into (10) to estimate the center-of-mass location $p_{sc} \in \mathbb{R}^3$:

$$A_q = [R^q_{ss_0}\hat{\mathfrak{m}}\mathfrak{g}] \in \mathbb{R}^{3\times 3}, \tag{14}$$

$$b_q = \sum_{i=1}^{N} ([p_{si}](R^q_{si} f^q_i) + R^q_{si} m^q_i) \in \mathbb{R}^{3\times 1}, \tag{15}$$

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_Q \end{bmatrix} \in \mathbb{R}^{3Q\times 3}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_Q \end{bmatrix} \in \mathbb{R}^{3Q\times 1}, \tag{16}$$

$$Ap_{sc} = b, \tag{17}$$

which can be solved for the estimate $\hat{p}_{sc} \in \mathbb{R}^3$ using least squares.

D. Inertia Matrix

We define a frame $\{\mathrm{b}\}$, shown in Fig. 2, at the center of mass $\hat{p}_{sc}$ and aligned with $\{\mathrm{s}\}$. After estimating the inertia matrix in $\{\mathrm{b}\}$, we identify the principal axes of inertia in $\{\mathrm{b}\}$, and then define a final center-of-mass frame $\{\mathrm{c}\}$ with axes aligned with the principal axes of inertia.

The Newton-Euler dynamics of a rotating rigid body are

$$m_b = \mathcal{I}_b\alpha_b + [\omega_b]\mathcal{I}_b\omega_b, \tag{18}$$

where m_b is the moment acting on the body, $\mathcal{I}_b \in \mathbb{R}^{3\times 3}$ is the positive-definite inertia matrix of the payload, ω_b is the angular velocity, and α_b is the angular acceleration, all expressed in the frame $\{\mathrm{b}\}$. For N robots supporting the payload with local wrenches $\mathcal{F}_i = (m_i, f_i)$, the moment can be expressed as

$$m_b = \sum_{i=1}^{N} ([p_{bi}](R_{bi} f_i) + R_{bi} m_i). \tag{19}$$

Equations (18) and (19) can be reorganized to solve for $\mathcal{I}_b$ using linear least squares. Since $\mathcal{I}_b$ is symmetric, it contains six unique elements: $\mathcal{I}_{xx}, \mathcal{I}_{xy}, \mathcal{I}_{xz}, \mathcal{I}_{yy}, \mathcal{I}_{yz}$, and $\mathcal{I}_{zz}$. We define

$$\mathcal{I}_{\mathrm{reg}} = [\mathcal{I}_{xx}, \mathcal{I}_{xy}, \mathcal{I}_{xz}, \mathcal{I}_{yy}, \mathcal{I}_{yz}, \mathcal{I}_{zz}]^\intercal. \tag{20}$$

At timestep q, we construct the matrices $A_q, B_q \in \mathbb{R}^{3\times 6}$ (based on the angular acceleration $\alpha_b = (\alpha_x, \alpha_y, \alpha_z)$ and angular velocity ω_b measured in $\{\mathrm{s}\}$) and the vector y_q based on force measurements at each of the N robots:

$$A_q = \begin{bmatrix} \alpha_x & \alpha_y & \alpha_z & 0 & 0 & 0 \\ 0 & \alpha_x & 0 & \alpha_y & \alpha_z & 0 \\ 0 & 0 & \alpha_x & 0 & \alpha_y & \alpha_z \end{bmatrix}, \tag{21}$$

$$B_q = \begin{bmatrix} 0 & -\omega_x\omega_z & \omega_x\omega_y & -\omega_y\omega_z & -\omega_z^2+\omega_y^2 & \omega_y\omega_z \\ \omega_x\omega_z & \omega_y\omega_z & \omega_z^2-\omega_x^2 & 0 & -\omega_x\omega_y & -\omega_x\omega_z \\ -\omega_x\omega_y & -\omega_y^2+\omega_x^2 & -\omega_y\omega_z & \omega_x\omega_y & \omega_x\omega_z & 0 \end{bmatrix}, \tag{22}$$

$$y_q = \sum_{i=1}^{N} \left([\hat{p}_{bi}](\hat{R}_{bi} f_{iq}) + \hat{R}_{bi} m_{iq}\right). \tag{23}$$

Combining (18) and (19), the dynamics can be written

$$(A_q + B_q)\mathcal{I}_{\mathrm{reg}} = y_q. \tag{24}$$

For Q measurements, we define

$$X = \begin{bmatrix} A_1 + B_1 \\ A_2 + B_2 \\ \vdots \\ A_Q + B_Q \end{bmatrix} \in \mathbb{R}^{3Q\times 6}, \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_Q \end{bmatrix} \in \mathbb{R}^{3Q\times 1}, \tag{25}$$

yielding

$$X\mathcal{I}_{\mathrm{reg}} = y, \tag{26}$$

which $\hat{\mathcal{I}}_{\mathrm{reg}}$ solves in a least-squares sense. The elements of $\hat{\mathcal{I}}_{\mathrm{reg}}$ form the entries of the estimated inertia matrix $\hat{\mathcal{I}}_b$.[1]

Let $R_{bc} = [r_1, r_2, r_3]$, where r_i is the ith eigenvector of $\hat{\mathcal{I}}_b$ and $\{\mathrm{c}\}$ is a frame coincident with $\{\mathrm{b}\}$ with axes aligned with the principal axes of inertia. Then the estimated diagonal inertia matrix in $\{\mathrm{c}\}$ is

$$\hat{\mathcal{I}}_c = R^\intercal_{bc}\hat{\mathcal{I}}_b R_{bc}. \tag{27}$$

IV. Experimental Implementation With the Omnid Mocobots

We performed payload estimation using the Omnid mocobots (Fig. 3), which are designed for human-robot collaborative manipulation [1]. Each Omnid mocobot consists of a mecanum-wheel mobile base, a 3-DoF Delta parallel mechanism driven

[1]If noisy data causes the least-squares estimate to produce an $\hat{\mathcal{I}}_b$ with a negative eigenvalue, the inertia estimate can either be discarded or projected to the space of positive semidefinite matrices by performing an eigenvalue decomposition and replacing any negative eigenvalue with a zero eigenvalue.

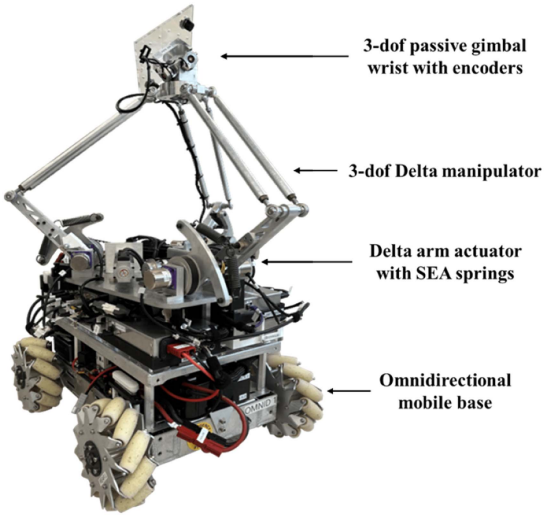

Fig. 3. The structure of the Omnid mocobot.

by series-elastic actuators (SEAs), and a 3-DoF passive gimbal wrist. In this letter, the mobile bases are stationary, and all manipulation is performed by the Delta-plus-gimbal manipulators.

The Delta mechanism and gimbal wrist of each Omnid are equipped with encoders, allowing Omnid i to measure its wrist configuration $T_{i_0 i}$. Since the payload is rigidly attached to the gimbal, the wrist configuration and the grasp configuration are equivalent. The twist $\mathcal{V}_i$ and its derivative $\dot{\mathcal{V}}_i$ are calculated by filtering encoder data. Wrenches at the wrist take the form $\mathcal{F}_i = (m_i, f_i) = (0, f_i)$, where the linear force f_i is calculated based on torques measured at the SEA joints and the moment m_i is zero due to the passive gimbal wrist. Data is collected at a fixed sampling rate of 100 Hz.

Experimental video is at https://youtu.be/fAvCyRhzfC4 or https://tinyurl.com/mocobot-estimation.

A. Experimental Configuration

To experimentally validate the methodology, we used four different payloads: two structures constructed from aluminum extrusion, each with two different sets of grasp locations (Fig. 4).

The payloads are rigidly attached to the Omnid gimbal wrists during experiments. The Omnid grasp frames are denoted $\{1\}$, $\{2\}$, and $\{3\}$, and frame $\{1\}$ is selected as the reference frame $\{\mathrm{s}\}$. Fig. 5 shows the three Omnids carrying payload (d).

The ground truth values of the transformation matrices, mass, center of mass, and inertia matrix for experimental validation were derived from CAD and are summarized in Table I. Rotation matrices are summarized in the axis-angle (exponential coordinates) representation $\omega\beta \in \mathbb{R}^3$, where ω is the unit rotation axis and β is the angle of rotation, expressed in this letter in degrees.

B. Data Processing

As shown in Fig. 6, $\phi = (\phi_1, \phi_2, \phi_3)$ represent the actuated proximal joint angles of the Delta mechanism and $\theta = (\theta_x, \theta_y, \theta_z)$ represent the angles of the gimbal wrist, all measured

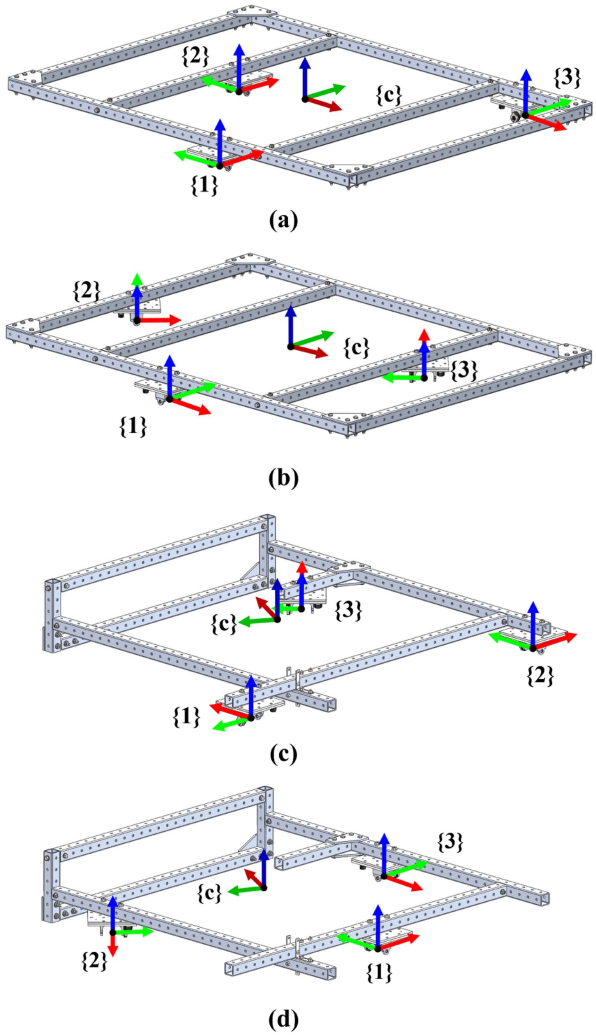

Fig. 4. Four payloads with associated grasp frames for experimental validation. Frame $\{1\}$ is chosen as the reference frame $\{\mathrm{s}\}$. The axes of the frames are visually differentiated by colors, with red corresponding to the x-axis, green to the y-axis, and blue to the z-axis.

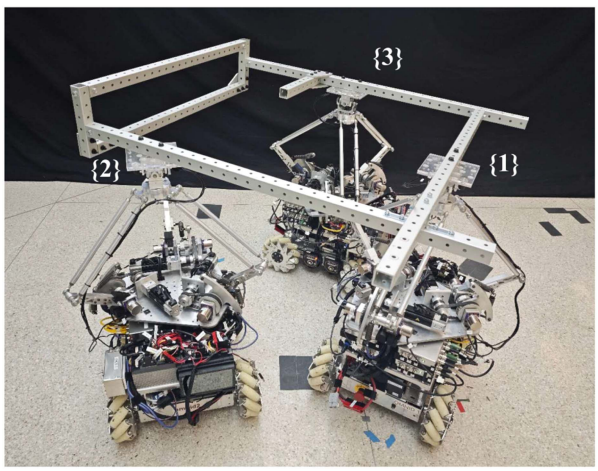

Fig. 5. Experimental setup with payload configuration (d).

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

TABLE I
GROUND TRUTH VALUES FOR THE FOUR EXPERIMENTAL CONFIGURATIONS

Rotation Matrix in Axis Angle (deg)			
Config.	R_{12}	R_{23}	R_{31}
(a)	(0, 0, 0)	(0, 0, −90)	(0, 0, 90)
(b)	(0, 0, 45)	(0, 0, 90)	(0, 0, −135)
(c)	(0, 0, −90)	(0, 0, 45)	(0, 0, 45)
(d)	(0, 0, −135)	(0, 0, 45)	(0, 0, 90)
Position Vector (m)			
Config.	p_{12}	p_{23}	p_{31}
(a)	(0.647, 0.533, 0)	(0.457, −0.838, 0)	(−0.305, −1.105, 0)
(b)	(−0.686, 0.572, 0)	(0.916, −0.647, 0)	(−0.242, 0.835, 0)
(c)	(−0.115, −1.143, 0)	(−0.229, 0.800, 0)	(−1.132, 0.162, 0)
(d)	(−0.382, 0.800, −0.038)	(−0.485, 0.862, 0.038)	(0.533, −0.571, 0)
Mass (kg), Center of Mass (m), and Principal Axes (deg)			
Config.	m	p_{1c}	R_{1c}
(a)	11.672	(0.537, 0.156, 0.072)	(0.043, 0.085, -84.787)
(b)	11.672	(0.034, 0.535, 0.072)	(0.115, −0.115, 2.465)
(c)	10.478	(0.547, −0.670, 0.057)	(−1.034, 0.355, −24.643)
(d)	10.478	(0.089, 0.593, 0.055)	(−1.662, −0.702, 68.937)
Inertia Matrix (kg·m^2)			
Config.	$\mathcal{I}_{xx}$	$\mathcal{I}_{yy}$	$\mathcal{I}_{zz}$
(a)	2.318	3.215	5.524
(b)	2.153	3.390	5.535
(c)	1.824	2.438	4.208
(d)	1.711	2.122	3.772

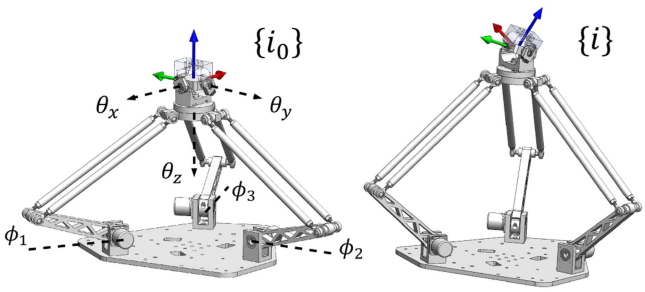

Fig. 6. Frame $\{i\}$ of the 6-DoF manipulator of Omnid i is positioned at the center of the gimbal wrist. The $\hat{z}$-axis encoder is installed at the base of the gimbal wrist, which is not visible in the figure. The left image shows the manipulator in its home position.

by encoders. The transformation matrix $T_{i_0 i}$ can be derived through forward kinematics as

$$T_{i_0 i}(\theta, \phi) = \begin{bmatrix} g(\theta) & h(\phi) \\ 0 & 1 \end{bmatrix}. \tag{28}$$

The twist $\mathcal{V}_i = (\omega_i, v_i)$ is calculated as

$$T_{i_0 i}^{-1} \dot{T}_{i_0 i} = [\mathcal{V}_i] = \begin{bmatrix} [\omega_i] & v_i \\ 0 & 0 \end{bmatrix} \in se(3). \tag{29}$$

The wrench at $\{i\}$ is $\mathcal{F}_i = (0, f_i)$, where $f_i = (\partial h/\partial \phi)^{-\mathsf{T}} \tau_i$ and τ_i is the Delta's SEA joint torque vector.

The derivative $\dot{T}_{i_0 i}$ and the twist derivative $\dot{\mathcal{V}}_i$ are calculated using central differencing post-processing of the 100 Hz estimates of $T_{i_0 i}$ and $\mathcal{V}_i$, respectively. To smooth the data and reduce the effect of amplification of encoder quantization in this numerical differencing, we apply a third-order Butterworth low-pass filter to the encoder data prior to calculating the $T_{i_0 i}$ estimates. The same filter is applied to smooth the SEA torque data. To avoid filter edge effects and other transients, the first and last two seconds of each trial's data are discarded.

TABLE II
EXPERIMENTAL RESULTS FOR RELATIVE GRASP FRAME ROTATION MATRIX ESTIMATION. RESULTS ARE REPRESENTED IN ABSOLUTE DEGREES OF ROTATION AROUND THE AXIS THAT ROTATES THE ESTIMATED ROTATION MATRIX TO THE ACTUAL ROTATION MATRIX

Config.	**Parameters**	**Absolute Error (deg)**	
		Mean	**Standard Deviation**
(a)	R_{12}	0.94	0.21
	R_{23}	0.99	0.16
	R_{31}	1.07	0.53
(b)	R_{12}	1.77	0.19
	R_{23}	0.90	0.42
	R_{31}	1.93	0.42
(c)	R_{12}	0.66	0.36
	R_{23}	2.21	0.59
	R_{31}	1.67	0.70
(d)	R_{12}	2.55	0.27
	R_{23}	3.57	0.32
	R_{31}	1.49	0.38

Results are represented in absolute degrees of rotation around the axis that rotates the estimated rotation matrix to the actual rotation matrix.

C. Experimental Results

1) Grasp Kinematics: To generate sufficient twist information for accurate estimation of relative grasp frame locations, the payload was randomly moved in all six DoF by assigning independent random trajectories to each Omnid manipulator. Each random trajectory consisted of a series of nominally straight-line paths for the gimbal wrist between via points, with the randomized transit times and dwell times at the vias chosen to be between 0.5 and 0.8 s. Because these independent random trajectories are not kinematically compatible with rigid grasps on the payload, we took advantage of mechanical compliance and used low motion control gains at each Omnid's force-controlled manipulator to allow trajectory tracking error to keep internal forces on the payload low.

Encoder data from the Omnid manipulators were collected to calculate $\mathcal{V}_{i*}, i = 1, 2, 3$. The transformations for the individual robot pairs $\hat{T}_{12}, \hat{T}_{23}$, and $\hat{T}_{31}$ were first calculated using the least-squares methods of Section III-B, and these estimates were then refined using nonlinear optimization incorporating loop-closure constraints. In our experiments, this second optimization step did not change the estimates significantly. Six independent trials were conducted for each payload, and the random trajectories for each trial consisted of 80 random via points.

The errors between the estimated transformation matrices and the ground truth transformation matrices (Table I) are reported in Tables II and III for the rotation matrices and position vectors, respectively. Rotation matrix errors are expressed as the rotation angle β (in degrees) of $[\omega]\beta = \log(R_{ij}^{\mathsf{T}} \hat{R}_{ij})$.

TABLE III
EXPERIMENTAL RESULTS FOR RELATIVE GRASP FRAME POSITION ESTIMATION

Config.	Parameters	Absolute Error (m)		Percentage Error	
		Mean	Standard Deviation	Mean	Standard Deviation
(a)	p_{12}	0.007	0.003	0.9%	0.3%
	p_{23}	0.042	0.005	4.4%	0.5%
	p_{31}	0.051	0.008	4.5%	0.7%
(b)	p_{12}	0.016	0.002	1.8%	0.2%
	p_{23}	0.039	0.006	3.5%	0.6%
	p_{31}	0.039	0.002	4.5%	0.3%
(c)	p_{12}	0.038	0.003	3.3%	0.2%
	p_{23}	0.046	0.004	5.6%	0.4%
	p_{31}	0.063	0.007	5.5%	0.6%
(d)	p_{12}	0.019	0.002	2.1%	0.3%
	p_{23}	0.057	0.004	5.8%	0.4%
	p_{31}	0.029	0.004	3.8%	0.5%

Each 3-vector position error is converted to a scalar using the Euclidean norm, and percentage errors are relative to the Euclidean norm of the ground truth p_{ij} .

TABLE IV
EXPERIMENTAL RESULTS FOR MASS AND CENTER OF MASS ESTIMATION

Config.	Parameters	Absolute Error	Percentage Error
(a)	m (kg)	0.11	0.9%
	p_{1c} (m)	0.021	3.7%
(b)	m (kg)	0.07	0.6%
	p_{1c} (m)	0.018	3.3%
(c)	m (kg)	0.11	1.1%
	p_{1c} (m)	0.039	4.5%
(d)	m (kg)	0.18	1.8%
	p_{1c} (m)	0.023	3.8%

The 3-vector position error is converted to a scalar using the Euclidean norm, and the percentage error is relative to the Euclidean norm of the ground truth p_{1c}.

2) Mass and Center of Mass: To estimate the mass and center of mass of the payload, the Omnids were commanded to hold the payload stationary at six distinct orientations. Data collection was performed during periods when the measured forces along all axes remained constant (with a tolerance of 0.01 N) for a continuous duration of six seconds. The results of the least-squares estimates from Section III-C are summarized in Table IV.

3) Inertia Matrix: To estimate the inertia matrix, the Omnids were tasked with generating motion to sufficiently excite the payload's rotational dynamics. Two types of trajectories were employed: the random trajectories used previously in the grasp kinematics experiments, and periodic trajectories constructed based on the estimated grasp transformations. For the periodic trajectories, all three Omnid wrists moved along periodic trajectories on the surface of the sphere for which the origins of the three grasp frames lie on a common meridian. In total, six experimental trials for each configuration were conducted: three trials using random trajectories, each consisting of 80 configurations per Omnid, and three trials using controlled periodic trajectories, each lasting one minute. No significant difference in the data quality was observed between the two trajectory types.

Table V summarizes the results from the six experimental trials. Similar to the labeling of axes in the ground truth inertia matrix, the labels $\hat{\mathcal{I}}_{xx}$, $\hat{\mathcal{I}}_{yy}$, and $\hat{\mathcal{I}}_{zz}$ are given to the axes with principal inertia estimates in ascending order.

TABLE V
EXPERIMENTAL RESULTS FOR THE INERTIAS OF THE PRINCIPAL AXES OF INERTIA FOR DIFFERENT PAYLOADS

Config.	Parameters	Absolute Error (kg·m^2)		Percentage Error	
		Mean	Standard Deviation	Mean	Standard Deviation
(a)	$\mathcal{I}_{xx}$	0.083	0.036	3.6%	1.6%
	$\mathcal{I}_{yy}$	0.130	0.022	4.1%	0.7%
	$\mathcal{I}_{zz}$	0.102	0.166	1.8%	3.0%
(b)	$\mathcal{I}_{xx}$	0.067	0.065	3.1%	3.0%
	$\mathcal{I}_{yy}$	0.086	0.102	2.5%	3.0%
	$\mathcal{I}_{zz}$	0.246	0.056	4.4%	1.0%
(c)	$\mathcal{I}_{xx}$	0.005	0.073	0.3%	4.0%
	$\mathcal{I}_{yy}$	0.039	0.120	1.6%	4.9%
	$\mathcal{I}_{zz}$	0.158	0.091	3.8%	2.2%
(d)	$\mathcal{I}_{xx}$	0.048	0.058	2.8%	3.4%
	$\mathcal{I}_{yy}$	0.016	0.060	0.8%	2.8%
	$\mathcal{I}_{zz}$	0.005	0.165	0.1%	4.4%

V. DISCUSSION

The experimental results demonstrate the validity of the approach, producing relative grasp rotation matrix mean errors of less than 4° and position vector magnitude mean errors of less than 6%, a mass estimate error of less than 2%, and a center of mass position error of less than 5%. Inertia matrix estimation, which relies on acceleration estimates from twice-differenced encoder data, exhibits a bit more error, but it is still less than 7% for the principal moments of inertia when accounting for one standard deviation.

The experimentally-derived kinematic and mass parameters may be directly inserted into the mocobot gravity compensation mode reported in [1], and the inertial parameters can be used to improve the performance of dynamic model-based cooperative control, eliminating the need for an *a priori* model of the payload and the grasp locations. There are a number of limitations to the method, however.

First, grasp frame velocities and accelerations are obtained by differencing encoder data, including encoders at the proximal joints of the Delta mechanism, distant from the wrist. The intervening mechanism and the differencing process introduce noise to the estimation. To address this, the Omnid sensor suite could be supplemented by IMUs at the gimbals. Independent of the sensing modality, however, the inherent compliance of the manipulators, due to their SEAs, can introduce vibrations at the wrists, which may be exacerbated by a poor choice of wrist trajectories to estimate the payload inertial properties. In practice, our excitation trajectories resulted in relatively smooth motions of the payload, as shown in the video.

Second, manipulation forces at the wrist are estimated based on the Delta's proximal joint SEAs. This measurement assumes zero friction at joints and bearings of the Delta mechanisms, but it eliminates the need for costly and potentially fragile end-effector force-torque sensors. Friction in the Omnid Delta joints is quite low, resulting in good wrist force sensing.

Third, the limited workspace of the manipulators places constraints on the payload trajectories that can be used to determine inertial properties. In particular, the payload trajectories in our experiments were limited to less than $\pm 10^{\circ}$ rotation about any axis. Payload trajectories with larger rotations could provide a higher signal-to-noise ratio.

Finally, computational aspects of the method (e.g., data filtering during post-processing) could be further optimized.

VI. Conclusion

We present a methodology for cooperative payload estimation using a group of mocobots, focusing on the estimation of the mass, center of mass, and inertia matrix of a rigid body payload, as well as the transformations between the grasp frames of each robot. Despite challenges of sensor noise and unmodeled dynamics, the proposed approach demonstrates promising results for payload estimation using Omnid mocobots. Future work will focus on extending the methodology to identify the kinematic and inertial properties of articulated payloads, and eventually continuously deformable payloads.

References

[1] M. L. Elwin, B. Strong, R. A. Freeman, and K. M. Lynch, "Human-multirobot collaborative mobile manipulation: The Omnid mocobots," *IEEE Robot. Autom. Lett.*, vol. 8, no. 1, pp. 376–383, 2023.

[2] J. E. Colgate, W. Wannasuphoprasit, and M. A. Peshkin, "Cobots: Robots for collaboration with human operators," in *Proc. ASME Int. Mech. Eng. Congr. Expo.*, 1996, vol. 15281, pp. 433–439.

[3] F. Caccavale and M. Uchiyama, "Cooperative manipulation," in *Springer Handbook of Robotics*. Berlin, Germany: Springer, 2016, pp. 989–1006.

[4] A. Cherubini, R. Passama, A. Crosnier, A. Lasnier, and P. Fraisse, "Collaborative manufacturing with physical human–Robot interaction," *Robot. Comput.- Integr. Manuf.*, vol. 40, pp. 1–13, 2016.

[5] J. Werfel, K. Petersen, and R. Nagpal, "Designing collective behavior in a termite-inspired robot construction team," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2014, pp. 3271–3278.

[6] J. Trevelyan, W. R. Hamel, and S.-C. Kang, "Robotics in hazardous applications," in *Springer Handbook of Robotics*. Berlin, Germany: Springer, 2016, pp. 1521–1548.

[7] N. Mavrakis and R. Stolkin, "Estimation and exploitation of objects' inertial parameters in robotic grasping and manipulation: A survey," *Robot. Auton. Syst.*, vol. 124, 2020, Art. no. 103374.

[8] N. Mavrakis et al., "Analysis of the inertia and dynamics of grasped objects, for choosing optimal grasps to enable torque-efficient post-grasp manipulations," in *Proc. 2016 IEEE-RAS 16th Int. Conf. Humanoid Robots (Humanoids)*, 2016, pp. 171–178.

[9] N. Mavrakis, E. A. M. Ghalamzan, and R. Stolkin, "Safe robotic grasping: Minimum impact-force grasp selection," in *Proc. 2017 IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2017, pp. 4034–4041.

[10] H. Farivarnejad and S. Berman, "Multirobot control strategies for collective transport," *Annu. Rev. Control, Robot., Auton. Syst.*, vol. 5, no. 1, pp. 205–219, 2022.

[11] C. K. Verginis and D. V. Dimarogonas, "Energy-optimal cooperative manipulation via provable internal-force regulation," in *Proc. 2020 IEEE Int. Conf. Robot. Automat.*, 2020, pp. 9859–9865.

[12] A. Petitti, A. Franchi, D. Di Paola, and A. Rizzo, "Decentralized motion control for cooperative manipulation with a team of networked mobile manipulators," in *Proc. 2016 IEEE Int. Conf. Robot. Automat.*, 2016, pp. 441–446.

[13] B. Mirtich, "Fast and accurate computation of polyhedral mass properties," *J. Graph. Tools*, vol. 1, no. 2, pp. 31–50, 1996.

[14] K. Lynch, "Estimating the friction parameters of pushed objects," in *Proc. 1993 IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 1993, vol. 1, pp. 186–193.

[15] G. Habibi, Z. Kingston, W. Xie, M. Jellins, and J. McLurkin, "Distributed centroid estimation and motion controllers for collective transport by multi-robot systems," in *Proc. 2015 IEEE Int. Conf. Robot. Automat.*, 2015, pp. 1282–1288.

[16] C. G. Atkeson, C. H. An, and J. M. Hollerbach, "Estimation of inertial parameters of manipulator loads and links," *Int. J. Robot. Res.*, vol. 5, no. 3, pp. 101–119, 1986.

[17] A. Franchi, A. Petitti, and A. Rizzo, "Distributed estimation of the inertial parameters of an unknown load via multi-robot manipulation," in *Proc. 52nd IEEE Conf. Decis. Control*, 2014, pp. 6111–6116.

[18] A. Franchi, A. Petitti, and A. Rizzo, "Distributed estimation of state and parameters in multiagent cooperative load manipulation," *IEEE Trans. Control Netw. Syst.*, vol. 6, no. 2, pp. 690–701, Jun. 2019.

[19] A. Marino and F. Pierri, "A two stage approach for distributed cooperative manipulation of an unknown object without explicit communication and unknown number of robots," *Robot. Auton. Syst.*, vol. 103, pp. 122–133, 2018.

[20] K. M. Lynch and F. C. Park, *Modern Robotics*. Cambridge, MA, USA: Cambridge Univ. Press, 2017.

[21] W. Kabsch, "A solution for the best rotation to relate two sets of vectors," *Acta Cryst. Lographica Sect. A: Cryst. Phys., Diffraction, Theor. Gen. Crystallogr.*, vol. 32, no. 5, pp. 922–923, 1976.

[22] W. Kabsch, "A discussion of the solution for the best rotation to relate two sets of vectors," *Acta Crystallographica Section A: Cryst. Phys., Diffraction, Theor. Gen. Crystallogr.*, vol. 34, no. 5, pp. 827–828, 1978.

[23] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 04, pp. 376–380, Apr. 1991.

[24] J. Nocedal and S. J. Wright, "Numerical Optimization," in *Large-Scale Unconstrained Optimization*. Springer, 2006, pp. 164–192.

[25] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. Cambridge, MA, USA: MIT Press, 2005.

[26] F. Lu and E. Milios, "Globally consistent range scan alignment for environment mapping," *Auton. Robots*, vol. 4, pp. 333–349, 1997.