

Multifingered force-aware control for humanoid robots

Pasquale Marra^{1,2}, Gabriele M. Caddeo¹, Ugo Pattacini³ and Lorenzo Natale¹

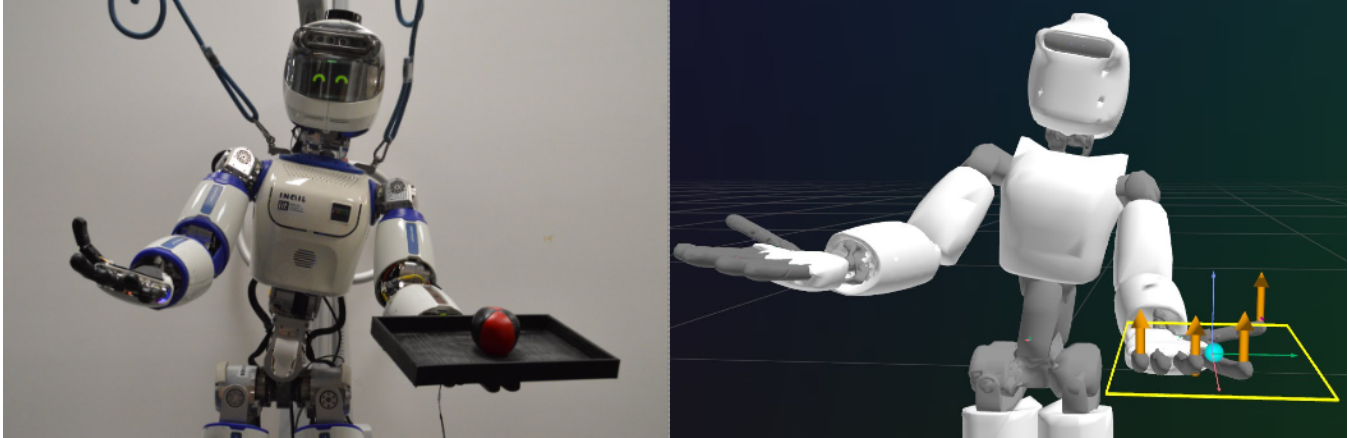


Fig. 1: Real robot (left) and its virtual counterpart [1] (right). ergoCub, equipped with tactile sensors, balances a tray with an unknown object by minimizing the distance between the supporting polygon center and the estimated Center of Pressure (cyan sphere), computed from normal forces (orange arrows).

Abstract—In this paper, we address force-aware control and force distribution in robotic platforms with multi-fingered hands. Given a target goal and force estimates from tactile sensors, we design a controller that adapts the motion of the torso, arm, wrist, and fingers, redistributing forces to maintain stable contact with objects of varying mass distribution or unstable contacts. To estimate forces, we collect a dataset of tactile signals and ground-truth force measurements using five Xela magnetic sensors interacting with indenters, and train force estimators. We then introduce a model-based control scheme that minimizes the distance between the Center of Pressure (CoP) and the centroid of the fingertips contact polygon. Since our method relies on estimated forces rather than raw tactile signals, it has the potential to be applied to any sensor capable of force estimation. We validate our framework on a balancing task with five objects, achieving a 82.7% success rate, and further evaluate it in multi-object scenarios, achieving 80% accuracy. Code and data can be found here <https://github.com/hsp-iit/multifingered-force-aware-control>.

I. INTRODUCTION

Robots meant to live alongside and assist humans must continuously refine their actions through sensory feedback [2]. While computer vision has advanced dramatically—reaching or surpassing human-level performance in some domains—tactile sensing still lags behind. Vision can substitute for some perceptual aspects of touch, but it poorly

captures contact dynamics: the interaction forces are difficult to measure reliably with external cameras. Humans routinely exploit touch for prehensile and non-prehensile manipulation: when holding a box with unknown contents or serving a tray, shifts in mass distribution are compensated by posture and hand adjustments driven by subtle force variations. Accordingly, many tactile sensors have been proposed, including finger-like, flat, and all-around designs based on magnetic [3], [4], capacitive [5], or vision-based principles [6], [7]. Their multimodal capabilities have enabled pose [8], [9], shape/curvature [10], [11], dynamic control [12], and texture estimation [13]. However, outputs differ substantially across technologies, and even within the same sensor family; therefore, manipulation methods that consume raw tactile signals often do not transfer across sensor types. A promising alternative is to map tactile outputs into the force domain: estimating 3D contact forces yields a sensor-agnostic representation that generalizes more robustly across devices [14], [15]. Yet, although force feedback is widely used in manipulation, multifingered strategies are typically limited to prehensile tasks, whereas non-prehensile approaches often rely on single-finger feedback [16]. In this work, we introduce a model-based indirect control strategy to balance in-hand objects with varying mass distribution by distributing contact forces across a multi-fingered hand. We maximize stability by minimizing the distance between the Center of Pressure (CoP), estimated from fingertip force measurements, and the geometric center of the fingertip contact polygon. To enable force feedback, we train a lightweight network to estimate 3D forces from custom magnetic Xela sensors [17] using a calibrated testbed with 3D-printed indenters. The model

¹Humanoid Sensing and Perception, Istituto Italiano di Tecnologia, Genoa, Italy. Acknowledge financial support from the PNRR MUR project PE0000013-FAIR.

²DIBRIS, Università di Genova, Via All’Opera Pia, 13, Genoa, Italy., Genoa, Italy.

³ Mesh Facility, Istituto Italiano di Tecnologia, Genoa, Italy

is deployed on the ergoCub [18] hand equipped with these fingertip sensors. A 3D-printed tray is placed on the hand and an object of unknown mass is positioned arbitrarily on it. Estimated fingertip forces are processed by a controller that computes the CoP and the desired plane rotation, then moves the fingertips to minimize the error between the CoP and the geometric center of the contact polygon, while keeping the fingertip positions on the plane. In parallel, a secondary controller supports finger control to maintain continuous contact with the tray and compensate for sources of error. This layered structure enables stable balancing while remaining agnostic to the tactile sensor type. The approach can serve as a robust low-level primitive within hierarchical frameworks or as a standalone solution for force-based in-hand balancing. We validate it by balancing an unknown object on the tray and preventing it from falling; an example is shown in Fig. 1. In summary, the contributions of this work are threefold:

- We present a dataset containing calibrated 3D forces, relative poses, and tactile outputs for our custom finger-like Xela sensors, collected with a dedicated testbed.
- We provide a comprehensive characterization of the sensor performance and limitations, offering insights into its applicability for force estimation in manipulation.
- We propose a multifingered control algorithm for balancing in-hand objects by acting on force distribution. The controller operates directly in the force domain, making it conceptually applicable across different tactile sensors.

II. RELATED WORKS

Our work is closely related to manipulation control frameworks using tactile sensing. While there remains interest in solving manipulation tasks without tactile feedback [19], there is a growing focus on integrating tactile sensors into robotic hands control [20], [21]. Early works primarily relied on single-finger feedback. For instance, in [16], the structural similarity index measure (SSIM) computed from TacTip images was used to control the closure of a Pisa/IIT SoftHand [22]. Similarly, in [23], Xela uSkin sensors were calibrated for force estimation using an on-hand method, and validated by controlling a single Allegro hand finger [24]. More recent works explore multi-finger control for prehensile tasks. In [25], a joint-level adaptive torque controller tracked desired finger velocities, while BioTac sensors [26] were used to estimate the contact frame orientation. In [27], grasp forces of the Allegro hand were modulated based on shear measurements from five MicroTac sensors. Other approaches estimate pseudo-forces from tactile signals: e.g., [28] employed pseudo-forces reconstructed from Xela uSkin outputs for force direction control, while [29] used pseudo-forces derived from BioTac sensors for grasp force regulation. In [30], fingertip forces were obtained by averaging raw outputs of magnetic tactile sensors to roll a cylinder with two fingers. Whole-body controllers have been proposed for humanoid balancing and locomotion tasks [31], [32]. These approaches similarly exploit a large number of joints and degrees of

freedom to achieve stability. However, they do not control the fingers explicitly, and therefore cannot address multifingered in-hand manipulation. In [33], a tray is fixed at the end of a robotic manipulator to perform non-prehensile transportation. However, it is assumed that the object’s shape and dynamical properties are known. In contrast, we develop an upper-body controller that extends the multi-joint reasoning of whole-body approaches to also include multicontact control of the fingers of a robotic hand, while making no assumption about the object. To the best of our knowledge, this is the first work to combine an upper-body controller with multifingered force distribution for non-prehensile manipulation.

III. METHOD

In our scenario, a humanoid robot equipped with a multifingered hand and fingertip force sensors must balance an object with varying mass distribution, using fingertip force feedback to adjust its hand pose.

Let $\mathcal{F}_{\Pi} = (O_{\Pi}, R_{\Pi})$ be the orthonormal frame with origin in $O_{\Pi} \in \mathbb{R}^3$ and orientation $R_{\Pi} = [b, t, n] \in SO(3)$, with $\dot{O}_{\Pi}, \ddot{O}_{\Pi}$ denoting its linear velocity and acceleration, and $\omega_{\Pi}, \dot{\omega}_{\Pi}$ its angular velocity and angular acceleration, respectively. Let Π be the plane passing through the origin O_{Π} with unit normal vector $n \in \mathbb{R}^3$. Our objective is to design a control architecture that:

- controls the finger positions so as to establish contact with a planar base coincident with the plane Π ;
- employs finger force measurements to adapt the pose of the frame \mathcal{F}_{Π} , and consequently the pose of the plane Π .

In this configuration, a planar object can be placed on the fingertips, establishing a correspondence with the virtual plane Π . This correspondence need not be exact, since the controller ensures fingertip contact with the object to compensate for misalignment. Moreover, we do not attempt to directly control the forces measured at the fingertips, as no assumptions are made about the object’s inertial properties. Fig. 2 shows our pipeline. This section proceeds as follow: in III-A we detail the characterization of the sensors and the dataset collection for force estimation; in III-B we define the finger positions control; then in III-C we show how the fingertip force measurements are used to affect the plane Π pose, leveraging the concept of Center of Pressure (CoP).

A. Sensor characterization and Dataset collection

1) *Setup*: To characterize the sensor and collect the dataset, we use the setup shown in Fig. 3a. A 3D-printed base secures the fingertip, while 3D-printed indenters (similar to those in [34], Fig. 3b) are mounted on the end effector of the Omega3 robot [35]. An ATI Nano17 force/torque sensor is placed between the robot and the indenter to measure applied forces, which are then projected onto the fingertip frame using the known transformation between the F/T sensor and the fingertip. For all experiments, we subtract an offset for each tactile sensor, estimated by averaging a thousand readings when the sensor is not in contact.

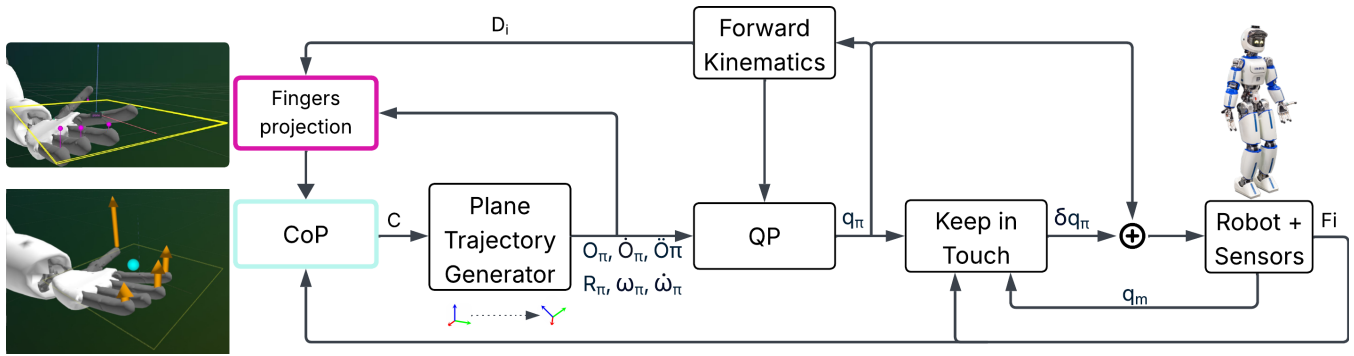


Fig. 2: Overview of the control pipeline. Finger positions (magenta spheres) are projected onto the plane (yellow rectangle), and normal force components (orange arrows) are used to compute the CoP (cyan sphere) in the plane frame. From this, a new target pose for the plane is generated, and finger and joint positions are updated with a corrective term for inaccuracies.

2) *Sensor characterization*: We employ a custom version of the Xela sensors specifically designed for the ErgoCub hand (called uSCu). Each fingertip consists of five magnetic taxels that can displace in three directions (x, y, z) under applied force. Before collecting the dataset, we first characterize the sensor behavior with respect to three aspects most relevant to our scenario: **repeatability** of the measurements, **consistency** across different sensors, and the effect of **gravity** on tactile readings.

To assess repeatability, we repeatedly press an indenter on the fingertip by controlling the robot’s position with 0.01 mm precision. The robot maintains contact for five seconds, rests for another five seconds, and repeats the motion. This experiment is conducted three times at reference positions corresponding to contact forces of approximately 2 N, 4 N, and 6 N. Higher forces are not considered, since the fingertips are mechanically tested only up to 5 N. While the sensor produces consistent outputs at 2 N, performance deteriorates at higher forces due to hysteresis. Fig. 3c illustrates this effect: the top plot shows the sequence of applied forces, while the bottom plot reports the output of one taxel, highlighting the growing influence of hysteresis with repeated touches.

To validate consistency, we perform a similar test across different fingertips. The robot is controlled in position along the z-axis, with the reference incremented by 0.2 mm at each step, producing gradually increasing contact forces. Fig. 4 shows the results: despite identical inputs, the taxels of different sensors produce different outputs, confirming variability across sensors.

Finally, we mount the fingers on the ErgoCub hand and perform random 6D movements. We observe that tactile readings vary significantly with orientation. More critically, the displacement of the taxels follows gravity, producing opposite responses w.r.t. to contact cases. For instance, when the fingertip is inverted, multiple taxels move outward due to gravity, a behavior that never occurs under actual contact.

3) *Data collection*: To estimate forces, we train a neural network to predict the 3D fingertip force vector from raw sensor readings. To collect data, we define contact points with Poisson sampling on the sensor surface, and for each point we generate random indentation trajectories. These are

constrained within a cylinder centered at the contact point (radius 1 mm to prevent tangential slip, height 5 mm to avoid excessive forces). Data are collected by controlling the robot probe in position, ensuring contact across a wide variety of fingertip poses. During collection, the fingertip is fixed facing upward. To mitigate the gravity effect described in Sec. III-A.2, we augment the dataset with additional no-contact samples acquired by mounting the sensors on the robot and applying random wrist and finger motions, to avoid out-of-distribution inputs during experiments. In total, we collect about 200k force–tactile pairs using all indenters, which are then balanced and reduced to 20k per sensor. Results of the calibration are reported in Section IV.

B. Finger positions control

We now describe how we control the fingers. Consider d fingers and let $D_i \in \mathbb{R}^3$, $\dot{D}_i \in \mathbb{R}^3$, $\ddot{D}_i \in \mathbb{R}^3$, $i = 1, \dots, d$ be the position, linear velocity, and linear acceleration of the i -th finger, respectively. Notice that the kinematic chain up to the hand palm is shared by all the fingers. Therefore, the joint positions vector for the i -th finger can be expressed as: $q_i = [q_c^T \ q_{fi}^T]^T$, where $q_c \in \mathbb{R}^{n_C}$ are the joints common to the finger kinematic chains ($n_C = 10$, that is torso-shoulder-elbow-wrist) and $q_{fi} \in \mathbb{R}^{n_{Fi}}$ are those specific of the i -th finger ($n_F = 2$ for thumb and index, $n_F = 1$ for the others). We can construct a joint position vector:

$$q = \begin{bmatrix} q_c \\ q_{f1} \\ \vdots \\ q_{fd} \end{bmatrix} \in \mathbb{R}^{n_C + n_{F1} + \dots + n_{Fd}} \quad (1)$$

The joint velocities \dot{q} and accelerations \ddot{q} are related to the finger linear acceleration as follows:

$$\ddot{D}_i = \dot{J}_{P,i} \cdot \dot{q} + J_{P,i} \cdot \ddot{q} \quad (2)$$

where $J_{P,i} \in \mathbb{R}^{3 \times (n_C + n_{F1} + \dots + n_{Fd})}$ is the positional part of the geometric Jacobian and $\dot{J}_{P,i}$ its time derivative. For each finger position, we are interested in calculating its

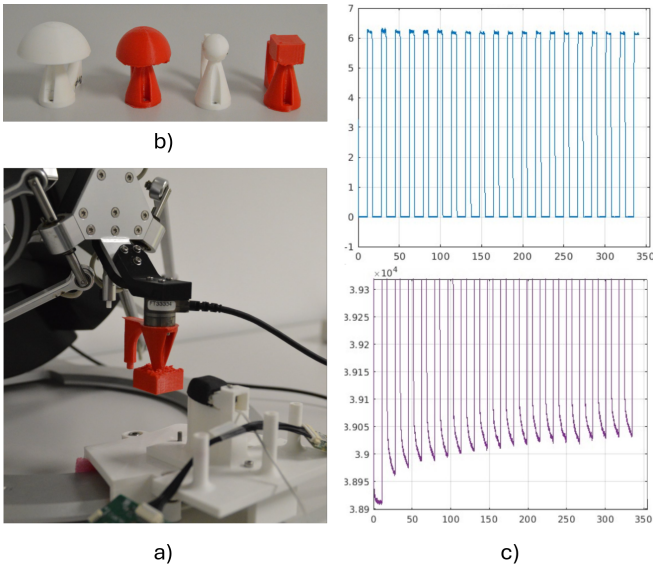


Fig. 3: a) Data collection setup. b) Indenters. c) Effect of hysteresis on the tactile readings. The graphic above shows the indented force (N) over time (s), while the bottom graphic shows the output of the sensor's taxel over time.

orthonormal projection $D_i^\Pi \in \mathbb{R}^3$ on the plane Π :

$$D_i^\Pi = O_\Pi + D_{t,i}^\Pi = D_i + D_{n,i}^\Pi \quad (3a)$$

$$D_{t,i}^\Pi = (I_{3 \times 3} - P_n^\Pi)(D_i - O_\Pi) \quad (3b)$$

$$D_{n,i}^\Pi = P_n^\Pi \cdot (O_\Pi - D_i) \quad (3c)$$

where $P_n^\Pi \triangleq n \cdot n^T$ is the orthogonal projector onto the line spanned by n ; $D_{t,i}^\Pi$ is the displacement, that lies on Π , between the projection D_i^Π and the origin O_Π ; $D_{n,i}^\Pi$ is the displacement, parallel to the unit normal vector n , between the projection D_i^Π and the finger position D_i . Using the previously defined quantities, in the following we define the two modules composing the finger positions control: the QP problem, whose unknowns are the joint accelerations \ddot{q} , and the *Keep in Touch* module.

1) **QP formulation**: we cast the finger positions control as:

$$\min_{\ddot{q}} T_1 + T_2 + T_3 + T_4 \quad (4)$$

$$\text{s.t.} \quad c_L \leq \ddot{q} \leq c_U$$

where the joint acceleration constraints are defined as ([36]):

$$c_U = \frac{k_l}{\delta t} \cdot \left(\frac{q^U - q}{\delta t} - \dot{q} \right) \quad (5a)$$

$$c_L = \frac{k_l}{\delta t} \cdot \left(\frac{q^L - q}{\delta t} - \dot{q} \right) \quad (5b)$$

with $k_l \in (0, 1]$, δt is the control period and $[q^L, q^U]$ is the joint range.

2) **Bring the fingers on the plane**: we start by defining the terms of the QP. We want each fingertip position to lie on the plane Π , so as to arrange the fingers on a planar support. We do so by imposing the following dynamics:

$$\ddot{D}_{n,i}^\Pi + K_{n,d} \cdot \dot{D}_{n,i}^\Pi + K_{n,p} \cdot D_{n,i}^\Pi = 0, \quad i = 1, \dots, d \quad (6)$$

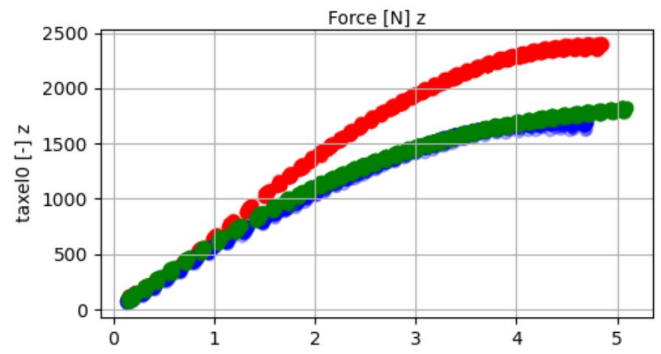


Fig. 4: Consistency validation across sensors shows variability in tactile outputs under identical applied forces.

with $\dot{D}_{n,i}^\Pi$ and $\ddot{D}_{n,i}^\Pi$ first and second order time derivative of (3c), $K_{n,d}$ and $K_{n,p}$ positive scalars. This is an asymptotically stable dynamics and converges to $D_{n,i}^\Pi = 0$, meaning that the i -th finger position converges to its projection on the plane Π . Substituting (2) and the second-order time derivative of (3c) in (6), we obtain:

$$A_{1,i} \cdot \ddot{q} = b_{1,i}, \quad i = 1, \dots, d \quad (7)$$

with:

$$A_{1,i} = P_n^\Pi \cdot J_{P,i} \quad (8a)$$

$$b_{1,i} = K_{n,d} \cdot \dot{D}_{n,i}^\Pi + K_{n,p} \cdot D_{n,i}^\Pi + P_n^\Pi \cdot [\ddot{O}_\Pi - \dot{J}_{P,i} \cdot \dot{q}] \quad (8b)$$

for $i = 1, \dots, d$. We define the relative cost function term as follows:

$$T_1 \triangleq \sum_{i=1}^d \|A_{1,i} \cdot \ddot{q} - b_{1,i}\|_{W_1}^2 \quad (9)$$

where $W_1 > 0$ and $\|x\|_W^2 = x^T W x$.

3) **Align the centroid of the finger projections to the origin of the plane frame**: we want the position of the fingertips centroid projection to correspond to the origin of the plane frame, because this point is also the target for the CoP control (III-C). We define the error:

$$e_t \triangleq O_\Pi - \frac{1}{d} \sum_{i=1}^d D_i^\Pi = -\frac{1}{d} \sum_{i=1}^d D_{t,i}^\Pi \quad (10)$$

and impose the error dynamics:

$$\ddot{e}_t + K_{t,d} \cdot \dot{e}_t + K_{t,p} \cdot e_t = 0 \quad (11)$$

with \dot{e}_t and \ddot{e}_t first and second order time derivative of e_t , $K_{t,d}$ and $K_{t,p}$ positive scalars. The dynamics (11) ensures that the geometric center of the fingertips converges to the origin O_Π . Substituting (2) and the second-order time

derivative of (3b) in (11), we obtain:

$$A_2 \cdot \ddot{q} = b_2 \quad (12a)$$

$$A_2 = (I_{3 \times 3} - P_n^\Pi) \frac{1}{d} \sum_{i=1}^d J_{P,i} \quad (12b)$$

$$b_2 = \frac{1}{d} \sum_{i=1}^d [\ddot{P}_n^\Pi \cdot (D_i - O_\Pi) + 2\dot{P}_n^\Pi \cdot (\dot{D}_i - \dot{O}_\Pi) + (I_{3 \times 3} - P_n^\Pi) \cdot (\ddot{O}_\Pi - \dot{J}_{P,i} \cdot \dot{q}) + K_{t,d} \cdot \dot{e}_t + K_{t,p} \cdot e_t] \quad (12c)$$

and define the relative cost function term:

$$T2 \triangleq \|A_2 \cdot \ddot{q} - b_2\|_{W_2}^2 \quad (13)$$

with $W_2 > 0$.

4) **Joint positions control:** we further add a postural term to keep the robot joint positions close to the initial configuration q^0 :

$$-\ddot{q} - K_{d,q} \cdot \dot{q} + K_{p,q} \cdot (q - q^0) = 0 \quad (14)$$

and define:

$$T3 \triangleq \sum_{i=1}^d \|\ddot{q} - b_3\|_{W_3}^2 \quad (15)$$

with: $b_3 = -K_{d,q} \cdot \dot{q} + K_{p,q} \cdot (q - q^0)$ and $W_3 > 0$.

5) **Bound Joint Accelerations:** this term bounds the required joint accelerations for the task and helps the solver convergence.

$$T4 \triangleq \|\ddot{q}\|_{W_4}^2 \quad (16)$$

6) **Keep in touch:** in addition to the QP generated references q_Π , we introduce a contribution δq_Π to compensate for small inaccuracies in the model, and mechanical imprecision due to finger tendon actuations. We call this module the *Keep in Touch* controller. This contribution is added only to the finger joints q_{f_i} and it is evaluated from the finger force F_i . Given a force threshold F_{th} , if $F_{th} > F_i$, then we update the offset so as to close the finger:

$$\delta q_\Pi = \min(\max(K_{p,f}(F_{th} - F_i) + q_m^{f_i} - q_\Pi^{f_i}, 0), q_u^{f_i} - q_\Pi^{f_i}) \quad (17)$$

where $q_u^{f_i}$ is the upper bound for the range of q_{f_i} and $q_m^{f_i}$ is the measured joint position, and $K_{p,f}$ positive scalar. On the other hand, when $F_{th} < F_i$ and the finger is closing, if necessary, we only update this offset in order to guarantee that $\delta q_\Pi < q_u^{f_i} - q_\Pi^{f_i}$. The threshold is set to account for hysteresis effects.

C. Plane pose evolution under the effect of the finger force measurements

Based on the knowledge of the measured forces at the fingertips, we want to implement an **indirect** control of the CoP by adjusting the pose of the plane Π . The idea is to move the CoP towards the plane frame origin, which is also the target position of the fingers centroid control (see III-B.3). We define F_i as the force measured by the i -th fingertip.

1) **CoP calculation:** for each F_i , we calculate the magnitude of the normal component $R_{n,i} \in \mathbb{R}$ and the resultant $R_n \in \mathbb{R}$ as:

$$R_{n,i} \triangleq n^T \cdot F_i \quad (18a)$$

$$R_n \triangleq \sum_{i=1}^d R_{n,i} \quad (18b)$$

Notice that $R_{n,i} \geq 0$ because the fingertips can only push on the plane. Then, assuming that the force F_i is applied at the finger projection D_i^Π , we calculate the CoP position relative to the origin O_Π as [37]:

$$O_\Pi C = \sum_{i=1}^d \frac{R_{n,i}}{R_n} \cdot D_{t,i}^\Pi \quad (19)$$

2) **Plane trajectory:** We introduce the rotation matrix R_e as:

$$R_e = \text{Rotation}(e) \quad (20a)$$

$$e = K_{p,e} s + K_{i,e} \int s \quad (20b)$$

$$s = O_\Pi C \times n \quad (20c)$$

where $\text{Rotation}(\cdot)$ calculates the rotation matrix from an axis-angle representation, $K_{p,e}$ and $K_{i,e}$ are positive scalars, and s is the cross product between (19) and the plane normal. The rotation matrix R_e is interpreted as the rotation that, applied to the initial plane orientation R_Π^i , could bring the CoP towards the origin O_Π . Note that the integral action in (20b) is introduced because the rotation due to the proportional action alone may be insufficient to cause the object to slide on the plane. We generate the plane trajectory as in the following:

$$R_\Pi = R_e \cdot R_\Pi^i, \quad \omega_\Pi = 0, \quad \dot{\omega}_\Pi = 0 \quad (21a)$$

$$O_\Pi = O_\Pi^i, \quad \dot{O}_\Pi = 0 \quad \ddot{O}_\Pi = 0 \quad (21b)$$

with O_Π^i is the (desired) initial plane position.

IV. EXPERIMENTAL RESULTS

We evaluate our method in two stages: first, by validating the precision of the force estimation network, and second, by testing the controller on balancing performance on a set of objects. The task consists in balancing an object on top of a 3D-printed tray held by the ergoCub hand. Because the hand is equipped with four Xela sensors, only four fingers are used. The tray weighs 250 g, measures 27 cm x 17 cm, and includes 2 cm-high barriers (0.5 cm thick) along the edges. These barriers prevent small objects from falling but still allow objects of higher mass to fall, ensuring the task remains non-trivial. The tray dimensions are constrained by the robot's hand span, with the thumb-little finger distance being 25 cm. We test five objects spanning different shapes, materials, and internal dynamics:

- Object 1–3: a paper tea box, filled with varying amounts of small plastic balls containing modeling clay of different weights (122.7 g, 152.6 g, 216.2 g), producing different mass distributions.

TABLE I: Results of the experiments. Force estimation uses different networks for each finger (first five rows) and one network for all the fingers (last row). Each metric is evaluated at different positions (1–5) and on average.

Position	Time (s)						Distance CoP (cm)						Success rate (%)					
	1	2	3	4	5	Mean	1	2	3	4	5	Mean	1	2	3	4	5	Mean
Object 1	9.2	6.5	6.4	42.1	40.5	20.9	35.0	13.2	13.6	33.4	31.9	25.4	66.7	100.0	100.0	100.0	100.0	93.3
Object 2	8.8	9.5	5.2	24.6	28.9	15.4	14.1	6.4	3.1	12.6	15.0	10.2	100.0	100.0	66.7	100.0	66.7	86.7
Object 3	9.0	7.4	5.1	29.3	/	12.7	9.3	2.3	6.7	21.7	/	10.0	100.0	100.0	100.0	66.7	0.0	73.3
Object 4	12.3	28.4	11.9	22.2	22.8	19.52	44.0	168.0	15.2	93.2	25.0	69.1	66.7	66.7	33.3	100.0	66.7	66.7
Object 5	42.3	20.94	39.7	5.7	43.0	30.3	34.6	21.7	10.3	21.0	33.2	24.2	66.7	100.0	100.0	100.0	100.0	93.3
Object 3 (SN)	/	/	26.8	5.1	11.7	11.5	/	/	5.8	1.2	2.7	2.9	0	0	67	100	67	47

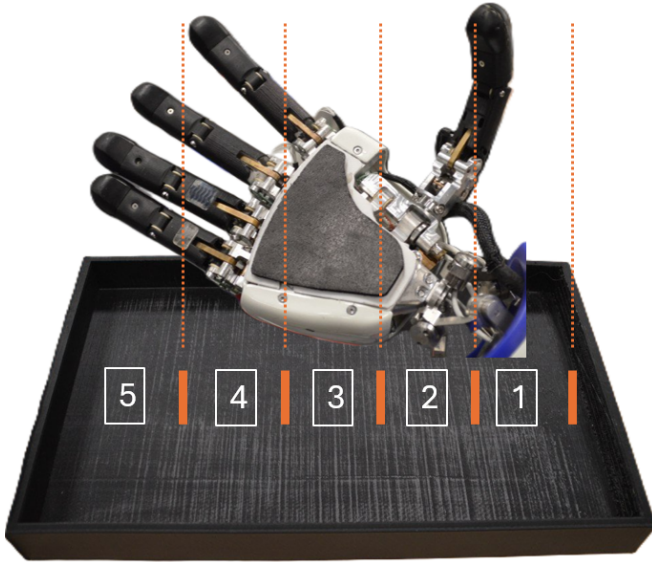


Fig. 5: Starting position on the tray. The superimposed hand shows the position of the hand under the tray.

- Object 4: a fabric ball filled with sand (98.6 g).
- Object 5: an aluminum box filled with modeling clay (125.6 g).

Objects are chosen to weigh between 100 g and 300 g, so that they could be reliably sensed by the Xela sensors, and they remain below the tested fingertip limit of 5 N of sustainable load.

Each fingertip uses a dedicated force estimation network trained on its own calibration dataset to maximize accuracy. To assess the necessity of this choice, we conduct an ablation in which a single shared network is used across all fingers (evaluated on Object 3). We perform all the experiments with the following parameters: $W_1 = W_2 = 10.0$, $W_3 = 0.01$, $W_4 = 2.5$, $K_{p,q} = 15.0$, $K_{d,q} = 5.0$, $K_{t,p} = K_{n,p} = 140.0$, $K_{t,d} = K_{n,d} = 10.0$, $k_l = 0.9$, $\delta_t = 0.01$, $K_{p,\omega} = 100$, $K_{p,e} = 8000 * (\pi/180)$, $K_{i,e} = 25 * (\pi/180)$, $K_{p,f} = 50.0 * (\pi/180.0)$, $F_{th} = 0.15$ N. We limit the maximum plane orientation to 30° along all three axes to constraint the speed of the object while moving on the tray. In the following we report results of the force estimation, of the balancing task employing single and multiple objects.

1) *Force estimation*: We train a regressor using the dataset described in Sec. III-A.3. The regressor is a multilayer perceptron (MLP) with four linear layers, interleaved with LayerNorm and GELU activation functions. Following the

characterization study in Sec. III-A.2, we address sensor-to-sensor variability by training a separate regressor for each sensor. We validate these regressors on test datasets collected by touching the sensors in poses different from those used for training. For comparison, we also train the same architecture on the combined data from all sensors and evaluate it on all test sets. On average, the single-sensor regressors achieve a normalized MAE of 9%, while the multi-sensor regressor reaches 13%.

2) *Single object balancing experiments*: For each object, we perform 15 trials by placing it in five different positions on the tray, with three repetitions per position. The positions are chosen by uniformly dividing the tray along its major axis. Fig. 5 shows the positions on the tray. We evaluate performance using three metrics. First, the time to convergence. Second, the distance traveled by the CoP, indicating how many corrections are required to balance the object (movements under 5 mm are ignored to filter sensor noise). Third, the success rate, when convergence is reached within 120 s, also accounting for those cases in which the object falls. Convergence is obtained by maintaining the distance between the CoP and the plate centroid under 3 cm threshold for at least 5 s. This threshold ensures the CoP remains inside the fingertip contact polygon, while accounting for sensor noise and static friction effects, which prevent motions from small rotations of the plane, thus not allowing small corrections. Only successful runs are included in the Time and CoP Distance metrics. The first five rows of Table I summarize the results. The best performance is obtained with the lightest objects (Object 1 and 5), since their lower friction and small momentum when contacting the barriers are insufficient to lead to failure cases. In contrast, Objects 3 and 4 show the lowest performance. Balancing Object 3 is difficult because higher friction requires larger plane inclinations to induce motion, often leading to unrecoverable situations. Object 4, instead, can roll easily; even small rotations generate high momentum, causing strong impacts against the barriers and eventual falls. Regarding object position, *Position 4* is the easiest to balance, being closest to the center of the support polygon (around the index/middle finger), whereas *Positions 1, 2, and 3* are farther away and therefore harder to control. Small plane inclinations, needed to correct minor unbalances, strongly affect Object 4, while Object 3 demands larger inclinations due to static friction. When the CoP is near the support polygon center, the controller also reacts more

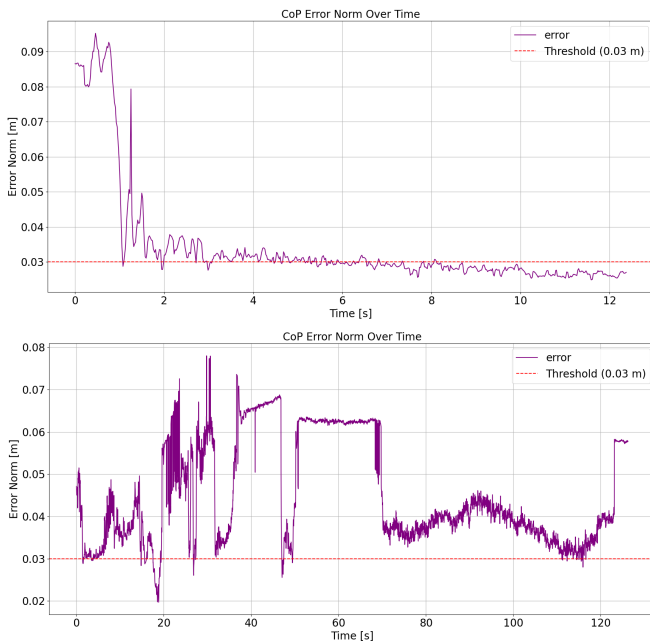


Fig. 6: CoP error for a success (top) and a failure (bottom) episode. In the failure case, the error drops below the threshold but is highly unstable.

slowly. This behavior is reflected in the *CoP Distance* metric: Object 4 exhibits much larger CoP movements, while Object 3 remains comparable to the others. Finally, to validate our choice of training a separate force estimator for each finger, we conduct an ablation using a single shared network (SN) on Object 3, shown in the last row of Table I. Results show a 26% performance drop, confirming the benefit of finger-specific estimators: when force estimation is unreliable, the system struggles to converge. This outcome is consistent with the findings in Section IV-1.

Next, we analyze the primary reasons for failure for each object. Object 3 and 4 consistently fail because of strong momentum that impacts the barrier. Object 2 fails once because of the momentum and once because of exceeding convergence time, Object 5 for strong impact with barriers; in the *SN* case, Object 3 fails seven times because of exceeding convergence time, while one time it fails because its speed generates a large impact with the border of the table. In general, potential inaccuracies in the force estimation may affect the outcome of the experiment. The results are expected, as the heavier object (Object 3) and the object with less static friction (Object 5) tends to accumulate more momentum. Fig. 6 shows the CoP error over time for a failure and for a success episode. The accompanying video provides a visual demonstration of the objects and experiments.

3) *Multiple objects balancing experiments*: To further check the robustness of our approach, we conduct experiments with multiple objects simultaneously. In particular, we use two aluminum boxes filled with modeling clay (as in Object 5), for a total weight of 299 g. We randomly place the boxes on the tray, varying their position. The boxes are placed in three predefined configurations (opposite sides, same side, and center), with exact positions randomized

across five trials each, obtaining 80%, 60%, and 100% success rate respectively. All failures result from excessive plane tilting caused by strong impacts with the objects. In successful trials, the average convergence time is 11.3 s, and the CoP travels 13 cm on average. These outcomes are consistent with the single-object experiments reported in Section IV-2. When the objects are placed on the same side, the CoP error increases, requiring larger plane inclinations; this amplifies object momentum and leads to strong barrier impacts. Such failures occur primarily when the objects move simultaneously, behaving as a single mass. Failures with objects set apart arise for a similar reason: the plane inclines to reduce the error, one object slides rapidly toward the other, and the impact with the plane border causes failure. These outcomes are expected, as they mirror the dynamics observed in the single-object case.

4) *Ablation on control frequency*: To assess the importance of the controller reactivity when dealing with object with varying mass, we repeat the same tests for single objects reported in Section IV-2, lowering the control frequency to 50 Hz. The lower frequency reduces the system reactivity and hinders proper recovery, especially for fast-moving or rolling objects (e.g., Objects 3 and 4): on average, we experience a 9% drop in success rate.

V. LIMITATIONS

In this section, we discuss the limitations of our work. The first concerns discrepancies between the 3D hand model and tendon friction, which cause misalignment between desired and actual finger motion. Nevertheless, the approach is still able to drive the CoP in most cases, showing robustness to modeling errors. Furthermore, the *Keep in Touch* control is currently implemented as a separate module: a more systematic approach would integrate this compensation directly into the QP. The second limitation is the plane trajectory module: the trajectory generation in (21) only accounts for the angular position of the plane Π , while its linear position remains constant. Exploiting the full pose of the plane could enhance the balancing performance and also enable the execution of more complex tasks, such as trajectory tracking with the arm while balancing an object. Finally, contact with the palm represent an additional source of error: in some trials, the plane tilts and touches the unsensorized palm, breaking the assumption of fully sensorized contact surfaces. A sensorized palm would mitigate this issue.

VI. CONCLUSION

In this paper, we presented a model-based force-aware control strategy for multifingered robotic hands. By operating in the force domain, our method adapts the motion of the torso, arm, wrist, and fingers to maintain stable contact with objects of varying mass, under the assumption of planar contact; we validated the approach on single-object and multi-object balancing tasks involving diverse shapes and weights, demonstrating robust performance across all scenarios. As future works, we plan to integrate the proposed

algorithm as a low-level policy within hierarchical manipulation frameworks, particularly for tasks involving bimanual manipulation.

REFERENCES

- [1] Rerun Development Team, “Rerun: A visualization sdk for multimodal data,” Online, 2024, available from <https://www.rerun.io/> and <https://github.com/rerun-io/rerun>. [Online]. Available: <https://www.rerun.io>
- [2] A. G. Billard, “Robots that share our spaces require proprioception and a sense of touch,” *Science*, vol. 389, no. 6761, p. eaea2492, 2025. [Online]. Available: <https://www.science.org/doi/abs/10.1126/science.aea2492>
- [3] L. Jamone, L. Natale, G. Metta, and G. Sandini, “Highly sensitive soft tactile sensors for an anthropomorphic robotic hand,” *IEEE Sensors Journal*, vol. 15, no. 8, pp. 4226–4233, 2015.
- [4] J. Li, H. Qin, Z. Song, L. Hou, and H. Li, “A tactile sensor based on magnetic sensing: Design and mechanism,” *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–9, 2024.
- [5] N. Jamali, M. Maggiali, F. Giovannini, G. Metta, and L. Natale, “A new design of a fingertip for the icub hand,” in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 2705–2710.
- [6] A. Sipos, W. van den Bogert, and N. Fazeli, “Gelslim 4.0: Focusing on touch and reproducibility,” 2024. [Online]. Available: <https://arxiv.org/abs/2409.19770>
- [7] M. Lambeta, P.-W. Chou, S. Tian, B. Yang, B. Maloon, V. R. Most, D. Stroud, R. Santos, A. Byagowi, G. Kammerer, D. Jayaraman, and R. Calandra, “DIGIT: A novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation,” *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 3838–3845, jul 2020.
- [8] G. M. Caddeo, N. A. Piga, F. Bottarel, and L. Natale, “Collision-aware in-hand 6d object pose estimation using multiple vision-based tactile sensors,” 2023. [Online]. Available: <https://arxiv.org/abs/2301.13667>
- [9] M. Bauza, A. Bronars, and A. Rodriguez, “Tac2pose: Tactile object pose estimation from the first touch,” *The International Journal of Robotics Research*, vol. 42, no. 13, p. 1185–1209, Sep. 2023. [Online]. Available: <http://dx.doi.org/10.1177/02783649231196925>
- [10] G. M. Caddeo, A. Maracani, P. D. Alfano, N. A. Piga, L. Rosasco, and L. Natale, “Sim2real bilevel adaptation for object surface classification using vision-based tactile sensors,” in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 15 128–15 134.
- [11] A.-H. Shahidzadeh, S. J. Yoo, P. Mantripragada, C. D. Singh, C. Fermüller, and Y. Aloimonos, “Actexplore: Active tactile exploration of unknown objects,” 2024. [Online]. Available: <https://arxiv.org/abs/2310.08745>
- [12] M. Oller, D. Berenson, and N. Fazeli, “Tactile-driven non-prehensile object manipulation via extrinsic contact mode control,” 2024. [Online]. Available: <https://arxiv.org/abs/2405.18214>
- [13] A. Böhm, T. Schneider, B. Belousov, A. Kshirsagar, L. Lin, K. Doerschner, K. Drewing, C. A. Rothkopf, and J. Peters, “What matters for active texture recognition with vision-based tactile sensors,” in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 15 099–15 105.
- [14] A. H. Shahidzadeh, G. Caddeo, K. Alapati, L. Natale, C. Fermüller, and Y. Aloimonos, “Feelanyforce: Estimating contact force feedback from tactile sensation for vision-based tactile sensors,” *ArXiv*, vol. abs/2410.02048, 2024. [Online]. Available: <https://api.semanticscholar.org/CorpusID:273097993>
- [15] Z. Chen, N. Ou, X. Zhang, Z. Wu, Y. Zhao, Y. Wang, N. Lepora, L. Jamone, J. Deng, and S. Luo, “General force sensation for tactile robot,” 2025. [Online]. Available: <https://arxiv.org/abs/2503.01058>
- [16] N. F. Lepora, C. Ford, A. Stinchcombe, A. Brown, J. Lloyd, M. G. Catalano, M. Bianchi, and B. Ward-Cherrier, “Towards integrated tactile sensorimotor control in anthropomorphic soft robotic hands,” in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 1622–1628.
- [17] Xela Robotics, “uscu & alha - tactile sensors,” <https://www.xelarobotics.com/sensor-collection/uscu-alha>, 2025, accessed: 2025-04-23.
- [18] ErgoCub team, “Project ergocub,” <https://ergocub.eu/>, 2023.
- [19] G. Khandate, C. P. Mehlman, X. Wei, and M. Ciocarlie, “Dexterous in-hand manipulation by guiding exploration with simple sub-skill controllers,” in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 6551–6557.
- [20] J. Egli, B. Forrai, T. Buchner, J. Su, X. Chen, and R. K. Katzschmann, “Sensorized soft skin for dexterous robotic hands,” in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 18 127–18 133.
- [21] Y. Li, P. Wang, R. Li, M. Tao, Z. Liu, and H. Qiao, “A survey of multifingered robotic manipulation: Biological results, structural evolvments, and learning methods,” *Frontiers in Neurorobotics*, vol. Volume 16 - 2022, 2022. [Online]. Available: <https://www.frontiersin.org/journals/neurorobotics/articles/10.3389/fnbot.2022.843267>
- [22] M. Catalano, G. Grioli, E. Farnioli, A. Serio, C. Piazza, and A. Bicchi, “Adaptive synergies for the design and control of the pisa/itit soft hand,” *The International Journal of Robotics Research*, vol. 33, no. 5, pp. 768–782, 2014. [Online]. Available: <https://doi.org/10.1177/0278364913518998>
- [23] E. Chelly, A. Cherubini, P. Fraise, F. B. Amar, and M. Khoramshahi, “Tactile-based force estimation for interaction control with robot fingers,” 2025. [Online]. Available: <https://arxiv.org/abs/2411.13335>
- [24] W. Robotics, “Allegro hand,” 2024, accessed: 2025-04-27. [Online]. Available: <https://www.allegrohand.com/>
- [25] F. Khadivar and A. Billard, “Adaptive fingers coordination for robust grasp and in-hand manipulation under disturbances and unknown dynamics,” *IEEE Transactions on Robotics*, vol. 39, no. 5, pp. 3350–3367, 2023.
- [26] N. Wettels, J. A. Fishel, and G. E. Loeb, “Multimodal tactile sensor,” in *The Human Hand as an Inspiration for Robot Hand Development*, ser. Springer Tracts in Advanced Robotics, R. Balasubramanian and V. J. Santos, Eds. Springer International Publishing, 2014, vol. 95, pp. 405–429. [Online]. Available: https://doi.org/10.1007/978-3-319-03017-3_19
- [27] C. Ford, H. Li, M. Catalano, M. Bianchi, E. Psomopoulou, and N. Lepora, “Shear-based grasp control for multi-fingered underactuated tactile robotic hands,” 03 2025.
- [28] D. Kitouni, E. Chelly, M. Khoramshahi, and V. Perdereau, “Fingertip contact force direction control using tactile feedback,” 2024. [Online]. Available: <https://arxiv.org/abs/2406.11545>
- [29] Z. Deng, Y. Jonetzko, L. Zhang, and J. Zhang, “Grasping force control of multi-fingered robotic hands through tactile sensing for object stabilization,” *Sensors*, vol. 20, no. 4, p. 1050, 2020.
- [30] H. Lee, Y. Kim, V. M. Staven, and C. Sloth, “Trajectory optimization for in-hand manipulation with tactile force control,” 2025. [Online]. Available: <https://arxiv.org/abs/2503.08222>
- [31] R. Calandra, S. Ivaldi, M. P. Deisenroth, and Peters, “Learning torque control in presence of contacts using tactile sensing from robot skin,” in *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, 2015, pp. 690–695.
- [32] J. Rozlivek, A. Roncone, U. Pattacini, and M. Hoffmann, “Harmonious—human-like reactive motion control and multimodal perception for humanoid robots,” *IEEE Transactions on Robotics*, vol. 41, pp. 378–393, 2025.
- [33] M. Selvaggio, A. Garg, F. Ruggiero, G. Oriolo, and B. Siciliano, “Non-prehensile object transportation via model predictive non-sliding manipulation control,” *IEEE Transactions on Control Systems Technology*, vol. 31, no. 5, pp. 2231–2244, 2023.
- [34] D. F. Gomes, P. Paoletti, and S. Luo, “Generation of gelsight tactile images for sim2real learning,” *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 4177–4184, 2021.
- [35] Force Dimension, “Force dimension omega3,” <https://www.forcedimension.com/products/omega3>,
- [36] H. Durrant-Whyte, N. Roy, and P. Abbeel, *Real-Time Prioritized Kinematic Control under Inequality Constraints for Redundant Manipulators*. MIT Press, 2012, pp. 145–152. [Online]. Available: <https://ieeexplore.ieee.org/document/6301071>
- [37] P. Sardain and G. Bessonnet, “Forces acting on a biped robot. center of pressure-zero moment point,” *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 34, no. 5, pp. 630–637, 2004.