

# Minimal Observations Inverse Reinforcement Learning for Predicting Human Box-Lifting Motions

Maxime Sabbah<sup>1</sup>, Filip Bećanović<sup>3</sup>, Sarmad Mehrdad<sup>2</sup>,  
Ludovic Righetti<sup>2,5</sup>, Bruno Watier<sup>1</sup>, Vincent Bonnet<sup>1,4</sup>

**Abstract**—Heavy-load manual lifting poses a significant risk of injury, motivating the need for personalized robotic assistance. The Minimal Observations Inverse Reinforcement Learning (MO-IRL) algorithm has recently demonstrated strong capabilities in recovering underlying optimality principles from very few demonstrations of simulated robotic motions, and at a very reasonable computational cost. Building on this, the present study integrates ten biomechanically informed cost functions into a direct optimal control formulation to predict human motion during heavy-load manual box-lifting tasks. Contrary to previous literature, thanks to the computational efficiency of MO-IRL, we allow time-varying optimal weights and include a collision-avoidance constraint within the set of cost functions. This constraint represents the subject’s apprehension of hitting the target table. As MO-IRL requires careful tuning of multiple hyperparameters, we employ a grid search to identify the optimal set. With this configuration, the predicted motion achieves an average accuracy of  $11.5 \pm 6.2$ deg across all joint angles, outperforming comparable methods. The inferred cost weights reveal a time-varying control strategy: initially minimizing lower-limb torques, then smoothing the motion through reduced joint accelerations and load velocity, and finally adjusting to avoid table collision. These findings show that biomechanically guided MO-IRL, coupled with direct optimal control, can accurately recover complex, constrained lifting motions while providing interpretable insights into human motor objectives, paving the way for adaptive and user-specific robotic assistance.

## I. INTRODUCTION

Manual handling of heavy objects is a fundamental yet crucial task across various sectors, including manufacturing industries, supermarkets, and online warehouses. It is well known that the repetitive execution of physically demanding tasks can lead to musculoskeletal disorders. In particular, lifting heavy objects and placing them accurately places significant strain on the operator. Such tasks increase ergonomic risks, notably back problems [1]. Therefore, there is a pressing need to develop robotic assistance for physically demanding activities. Assistive robots—whether exoskeletons or collaborative robots—must deliver personalized support to operators. Due to anthropometric differences among individuals, directly applying averaged joint trajectories is not feasible. Furthermore, even for simple tasks like lifting and dropping an object, the human central nervous system—acting as the body’s controller—can select from

an infinite number of joint trajectories due to the high redundancy of the human musculoskeletal system.

Reviews on the simulation and prediction of manual lifting indicate that the lifting task is most commonly modeled with 5 Degrees of Freedom (DoFs) in the sagittal plane [2].

Studies on manual box lifting have employed a variety of cost functions to generate human-like motions. Total muscular effort [3–5], the sum of squared joint torques and a dynamic balance criterion [6, 7], shear and compression forces at the spine [7], ankle torque [8], total torque [9], total energy (work) [10], jerk to produce smoother trajectories [10] were minimized. Motions that maximize the liftable weight were also selected [11]. In all these studies, cost functions were used independently and across the entire lifting motion. Nearly all optimal-control formulations impose joint angle limits, joint torque limits, and collision-avoidance constraints [5–8, 10–14]; joint velocity limits appear less frequently [7, 10, 11, 13]; static postural stability is enforced via center of mass projection [5, 12–14]; and dynamic stability via the Zero-Tilting Moment Point (ZMP) [7, 10]. All biomechanics models treat constraints separately from the cost functions, in contrast to the robotics community, which often imposes soft constraints as costs in model predictive control, for example [15]. Finally, hybrid approaches incorporating tracking error to motion-capture data are considered the most accurate and can achieve Root Mean Squared Error (RMSE) below 15deg [11, 13, 16], but they rely on knowing the motion before synthesizing it, which presents a limitation.

Inferring cost functions from observed human trajectories is a challenging problem, falling under the domains of Inverse Optimal Control (IOC) or Inverse Reinforcement Learning (IRL). Both fields have seen a significant number of contributions, progressively improving robustness to noise and reducing computation times. As recently demonstrated [17, 18], for noisy human motion data, only bilevel optimization approaches are suitable in IOC settings. In a bilevel framework, an outer optimization problem identifies the optimal set of cost function weights based on joint trajectories predicted by an inner optimization process that solves the corresponding Direct Optimal Control (DOC) problem [19]. However, a major limitation is that the gradient of the inner optimization cannot be directly retrieved, making the resolution of the bilevel problem particularly slow. Likely due to this computational burden, to the best of our knowledge, no studies have proposed using time-varying weights to model human motion. Humans are capable of managing time-varying constraints and planning sequences

<sup>1</sup> LAAS-CNRS, Université Paul Sabatier, CNRS, Toulouse, France.

<sup>2</sup> Machines in Motion Laboratory, New York University, USA

<sup>3</sup> School of Electrical Engineering, University of Belgrade, Belgrade

<sup>4</sup> Image and Pervasive Access Laboratory (IPAL), CNRS-UMI, 2955, Singapore.

<sup>5</sup> Artificial and Natural Intelligence Toulouse Institute (ANITI), Toulouse

of tasks. Incorporating such capabilities will likely improve the reproduction of observed human behavior.

Data-driven IRL methods, such as the original maximum entropy formulation [20], rely on the assumption that the sampled set of non-optimal trajectories for the partition function is properly constructed. However, this assumption becomes problematic in large-scale continuous domains. To address this, Finn et al. [21] introduced an iterative reformulation of IRL, called guided cost learning, which shares structural similarities with the bilevel formulation of IOC. This approach jointly optimizes the cost function and the trajectory distribution, enabling it to handle larger-scale problems. It effectively captures complex nonlinear behaviors even when the system dynamics are unknown, and has been widely adopted for imitation learning in robotic manipulation tasks. Nevertheless, it remains fundamentally data-driven, relying on neural networks, and it is still unclear how sampled trajectories should be generated and weighted. Moreover, it still requires many iterations of the inner optimization. Beyond the initial need for demonstration data, many IRL methods must repeatedly process all sampled trajectories throughout training to accurately estimate the distribution over trajectories. This remains computationally expensive. Furthermore, most IRL algorithms rely on probabilistic convergence criteria, such as maximizing entropy, which may not directly reflect how well the learned cost function reproduces expert-like behavior. In contrast, inverse optimal control methods employ gradient-based optimization to estimate cost parameters, thereby avoiding the need for extensive trajectory sampling.

Recently, in the Minimal Observations Inverse Reinforcement Learning (MO-IRL) framework [22], it was proposed to use trajectory sampling via a DOC solver to generate trajectories informed by the current estimate of the cost function. This produces more relevant and structured samples, removing the reliance on hand-crafted sampling strategies. Moreover, MO-IRL, which modifies the original maximum entropy IRL formulation, allows for the identification of small variations in the weights, thus naturally favoring moderate and well-distributed weight values. MO-IRL is a new and promising algorithm that has not yet been used to predict human motion or perform complex tasks involving collision constraints. However, it requires the tuning of several parameters, such as regularization coefficients, temporal partitioning, initial guesses, and the number of non-optimal trajectories sampled. Building on these ideas, we propose in this paper to:

- extend the MO-IRL framework by incorporating time-varying, biomechanically meaningful cost functions to accurately predict human motion;
- perform a sensitivity analysis of MO-IRL;
- apply optimally tuned MO-IRL to complex, constrained lifting motions collected from six healthy participants;
- provide insights regarding the time-varying optimality criterion used by humans to lift heavy objects.

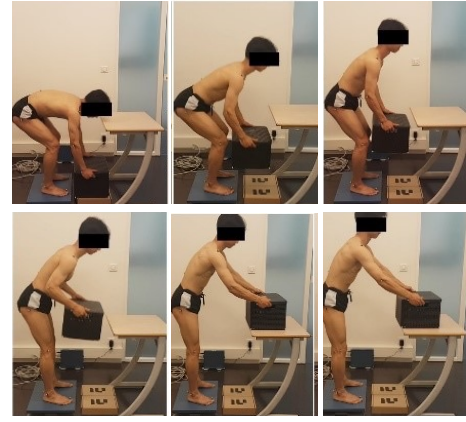


Fig. 1: Investigated human heavy-load lifting task and experimental setup.

## II. METHODS

### A. Experimental setup

Six healthy male adults (age:  $26.9 \pm 4.7$ y.o.; height:  $179.4 \pm 7.6$ cm; weight:  $77.4 \pm 4.9$ kg), with no history of musculoskeletal or neurological disorders, participated after providing informed consent. The experiments were under the national ethics review board of Université Fédérale de Toulouse, IRB#1 N° 00011835-2024-0910-855. Before data collection, participants completed a familiarization session to practice the experimental procedures. Then, they performed repetitive box-lifting tasks involving a 10.0kg box ( $42.7 \times 28.4 \times 21.9$ cm), transferring it between the floor and a table positioned at 0.75m height. Lifting was conducted at each participant's self-selected speed and distance from the box. Four sets of standard lifting, each with five repetitions and 2-minute rests between sets, were evaluated. Each repetition involved lifting the box from the floor to the table and returning it to the floor. Kinematic and kinetic data were collected synchronously using a stereophotogrammetric system (six Flex 13 cameras, 100Hz, Optitrack) and a force plate (1000Hz, AMTI). Retro-reflective markers were placed on nine right-side anatomical landmarks: distal phalanx, calcaneus, lateral malleolus, lateral femoral epicondyle, greater trochanter, T10 vertebra, acromion, lateral humeral epicondyle, and ulnar styloid. Multibody inverse kinematics [23] was performed to retrieve planar joint angles.

Moreover, each participant's inertial parameters were identified using kinematics and kinetics data as in [24]. To excite inertial parameters, participants performed a series of upper and lower limb excitation motions in the sagittal plane, including squat-like and pendulum-like motions, shoulder flexion/extension, and elbow flexion/extension. Each motion was repeated 15 times at increasing speeds.

### B. Biomechanical model

Similarly to most of the literature [2], a 5 DoFs planar sagittal biomechanical model, depicted in Fig. 2, was devised to represent the ankle, knee, hip, shoulder and elbow

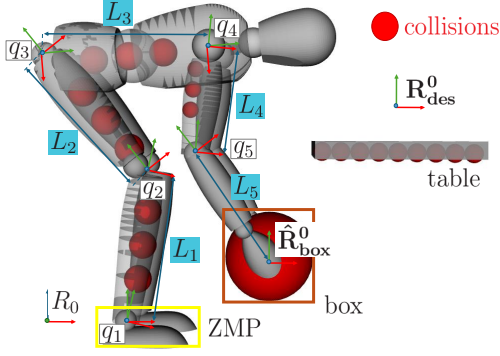


Fig. 2: Definition of the planar biomechanical model. The support polygon is shown in yellow, and the collision spheres in red.

flexion/extension joints  $\mathbf{q}$  connecting the following body segments: foot, shank, thigh, trunk, upper-arm and forearm. The relative positions of each successive joint in its parent frame were determined using the measures of segment lengths  $\mathbf{L} = [L_1, L_2, L_3, L_4, L_5]^T$  given by anatomical motion capture markers. The Forward Kinematics Model (FKM) was used to calculate the pose of the box  $\hat{\mathbf{R}}_{\text{box}}^0 \in SE(3) = (\hat{\mathbf{r}}_{\text{box}}^0, \hat{\mathbf{p}}_{\text{box}}^0)$  with respect to the world  $R_0$  as follows:

$$\hat{\mathbf{R}}_{\text{box}}^0 = \text{FKM}(\mathbf{q}, \mathbf{L}) \quad (1)$$

The dynamical model, Eq. 2, was implemented using the efficient robotics library Pinocchio [25], and inertial parameters composing the mass matrix  $\mathbf{M}$  were obtained from the participant-specific identification pipeline [24].

$$(\boldsymbol{\tau}, \mathbf{f}_{\text{GR}}, \mathbf{m}_{\text{GR}}^0) = \mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}}) + \mathbf{G}(\mathbf{q}) \quad (2)$$

where  $\boldsymbol{\tau}$  is the joint torque vector,  $\mathbf{f}_{\text{GR}}$  and  $\mathbf{m}_{\text{GR}}^0$  are the forces and moment associated with the reaction of the ground expressed with respect to  $R_0$ .  $\mathbf{C}$  is the centripetal and Coriolis matrix and  $\mathbf{G}$  the gravity one.

To be able to prevent auto-collisions and collisions with the table, collision spheres  $\mathbb{S} = (r, \mathbf{O}(\mathbf{q}))$  characterized by their radius  $r$  and their center position  $\mathbf{O}$ , calculated with the FKM, were introduced. At a given configuration  $\mathbf{q}$ , the distance  $d(\mathbb{S}_1, \mathbb{S}_2)$  between two spheres  $\mathbb{S}_1$  and  $\mathbb{S}_2$  was computed as :

$$d(\mathbb{S}_1, \mathbb{S}_2) = \|\mathbf{O}_1(q) - \mathbf{O}_2(q)\|_2 - (r_1 + r_2) \quad (3)$$

Each segment of the body (except the forearm, since the box was attached to it) was assigned spheres  $\mathbb{S}_{\text{seg}}$  to cover the whole segment length. The radii of the spheres were scaled according to the segment lengths. The box was assigned a large sphere  $\mathbb{S}_{\text{box}}$  corresponding to the box size. The table was assigned multiple spheres  $\mathbb{S}_{\text{tab}}$  whose centers and radii were chosen such that the union of the spheres approximated the surface of the table. The positions of the table spheres were independent of the joint angles  $\mathbf{q}$ .

TABLE I: Objective cost functions. The nine first functions are extracted from [26].

| Label       | Name                     | Equation   |
|-------------|--------------------------|--|
| $\Phi_1$    | Joint torque lower limbs | $\int_0^T \tau_{ll}^2 dt$  |
| $\Phi_2$    | Joint torque shoulder    | $\int_0^T \tau_{sho}^2 dt$   |
| $\Phi_3$    | Joint torque elbow       | $\int_0^T \tau_{elb}^2 dt$   |
| $\Phi_4$    | Joint velocity           | $\int_0^T \dot{q}^2 dt$  |
| $\Phi_5$    | Joint acceleration       | $\int_0^T \ddot{q}^2 dt$   |
| $\Phi_6$    | Cartesian box velocity   | $\int_0^T \dot{p}_x^2 + \dot{p}_z^2 dt$  |
| $\Phi_7$    | Joint torque change      | $\int_0^T \dot{\tau}^2 dt$   |
| $\Phi_8$    | Energy                   | $\int_0^T  \dot{q}\tau  dt$  |
| $\Phi_9$    | Geodesic                 | $\int_0^T (\dot{\mathbf{q}}^T \mathbf{M}(\mathbf{q}) \dot{\mathbf{q}})^{\frac{1}{2}} dt$ |
| $\Phi_{10}$ | Box horizontal position  | $\int_0^T \hat{p}_{box,z}^0 dt$  |

Finally, dynamic balance was taken into account by calculating the ZMP position in the sagittal plane as follows:

$$\text{ZMP} = \frac{m_{GR}^{0,Y}}{f_{GR}^Z} \quad (4)$$

### C. Human motion prediction

Human motion was predicted by solving a DOC problem, formulated as the optimization of a set of  $N_\Phi = 10$  objective cost functions  $(\Phi_i)_{i \in [0, N_\Phi]}$  parameterized by weights  $\boldsymbol{\omega}$  under physical constraints. The list of the objective functions, retained from the literature in human motor control [19, 26], is provided in Table I. Among them, it is proposed to dissociate the lower limb, shoulder, and elbow joint torques to allow independent weighting and tuning of leg drive and arm lift during the box-lifting motion. Additionally, to reflect the participant's cautious behavior near the target table, the horizontal position of the box along the x-axis,  $\hat{p}_{box,x}^0$ , was added as  $\Phi_{10}$ . This is usually treated as a constraint in the literature, which will be discussed in more depth in Section IV.

1) *Direct Optimal Control*: It was hypothesized that participants might not minimize the same objectives throughout the entire motion. For instance, participants are likely to decelerate near the end of the movement to ensure a precise and safe placement of the box on the table. To capture this behavior, each recorded trajectory of duration  $T$  was uniformly discretized into  $N_w$  time windows, each containing  $N_s$  samples. Accordingly, a weight matrix  $\boldsymbol{\omega} \in \mathbb{R}^{N_\Phi \times N_w}$  was introduced to predict the different motion phases during the task. The complete human trajectory  $\mathbf{x} \in \mathbb{R}^{N_s \times N_w}$ , defined as  $\mathbf{x} = (\mathbf{x}_i)_{i \in [0, N_w]} = (\mathbf{q}_i, \dot{\mathbf{q}}_i, \ddot{\mathbf{q}}_i)_{i \in [0, N_w]}$ , was then predicted by solving the following DOC problem:

$$\begin{aligned}
\mathbf{x}^* &= \arg \min_{\mathbf{x}} \sum_{i=1}^{N_w} \sum_{j=1}^{N_\Phi} \omega_{i,j} \Phi_j(\mathbf{x}_i) \\
\text{s.t. } \mathbf{x}(t+1) &= f(\mathbf{x}(t)), \forall t \in \llbracket 0, \lfloor \frac{T}{dt} \rfloor - 1 \rrbracket \\
\hat{\mathbf{R}}_{\text{box}}^0(\lfloor \frac{T}{dt} \rfloor) &= \mathbf{R}_{\text{des}}^0, \\
\mathbf{q}^- \leq \mathbf{q} \leq \mathbf{q}^+, \|\dot{\mathbf{q}}\| &\leq \dot{\mathbf{q}}^+, \\
\boldsymbol{\tau}^- \leq \boldsymbol{\tau} \leq \boldsymbol{\tau}^+, \\
\text{ZMP}^- \leq \text{ZMP} \leq \text{ZMP}^+, \\
\mathbf{q}(0) &= \mathbf{q}_0, \dot{\mathbf{q}}(0) = 0, \\
\hat{p}_{\text{box},z}^0 &\geq 0, \\
d(\mathbb{S}_{\text{seg}}, \mathbb{S}_{\text{box}}) &\geq \epsilon_{\text{col}}, d(\mathbb{S}_{\text{tab}}, \mathbb{S}_{\text{box}}) \geq \epsilon_{\text{col}}
\end{aligned} \tag{5}$$

where the first constraint correspond to the Euler time discretized dynamics  $f$  parametrized by the timestep  $dt$ ;  $\mathbf{R}_{\text{des}}^0 \in SE(3)$  is the pose of the box as demonstrated by the human at the final time;  $\mathbf{q}^-$ ,  $\mathbf{q}^+$  are the lower and upper joint boundaries respectively;  $\dot{\mathbf{q}}^+$  is the maximal joint velocity;  $\boldsymbol{\tau}^-$ ,  $\boldsymbol{\tau}^+$  are the lower and upper joint torque boundaries respectively;  $\text{ZMP}^-$ ,  $\text{ZMP}^+$  are the physical stability boundaries associated to the ZMP;  $\mathbf{q}_0$  is the initial human joint configuration;  $\epsilon_{\text{col}}$  is the threshold at which the collisions must be avoided. If the time dependency is not mentioned, the constraint is assumed to be applied on the whole time horizon. Boundaries values came from [27].

2) *Minimal Observations Inverse Reinforcement Learning*: In maximum entropy inverse reinforcement learning [20], the expert trajectory  $\mathbf{x}^*$  is assumed to be exponentially more likely if it has a lower cost  $C(\mathbf{x}^*, \boldsymbol{\omega}) = \boldsymbol{\omega}^\top \Phi(\mathbf{x}^*)$  than any other trajectory found in  $\bar{\mathbf{x}}$  shaping  $Z$ , the partition function. Mathematically, this is described by the following probability:

$$P(\mathbf{x}^* | \bar{\mathbf{x}}, \boldsymbol{\omega}) = \frac{1}{Z} \exp(-C(\mathbf{x}^*, \boldsymbol{\omega})) \tag{6}$$

Thus, IRL methods based on (6) seek to maximize the log likelihood of the optimal trajectory given any set of observed trajectories  $\bar{\mathbf{x}}$ :

$$\begin{aligned}
\boldsymbol{\omega}^* &= \arg \min_{\boldsymbol{\omega}} -\log P(\mathbf{x}^* | \bar{\mathbf{x}}, \boldsymbol{\omega}) \\
\text{s.t. } 0 &\leq \boldsymbol{\omega}
\end{aligned} \tag{7}$$

The challenge in applying this method to real-world scenarios lies in the proper computation of  $Z$ , that is to say, producing a representative set of suboptimal trajectories to better discriminate between cost functions when solving (7). Ideally, the recovered weights should reproduce the demonstrated optimal trajectory when applied in the DOC. However, maximizing the probability of the observed trajectory does not guarantee this outcome. In practice, the partition function is approximated using a finite, often limited, set of non-optimal trajectories. As a result, there may exist trajectories outside this set that have a higher probability than the demonstration. To overcome this, the newly proposed

MO-IRL uses a line-search procedure relying on a minimal set of non-optimal trajectories to solve the problem iteratively by taking small steps  $\Delta\boldsymbol{\omega}$ . Following [22], Eq.7 can be reformulated as:

$$\begin{aligned}
\Delta\boldsymbol{\omega}^* &= \arg \min_{\Delta\boldsymbol{\omega}} \sum_d -\theta_d \log \frac{1}{1 + \sum_{\mathbf{x} \in \bar{\mathbf{x}}_d} \gamma_{\mathbf{x}} e^{-\Delta\boldsymbol{\omega}^\top \bar{\Phi}_{d,\mathbf{x}}}} \\
&\quad + \lambda |\Delta\boldsymbol{\omega}| + \frac{\beta}{2} \|\Delta\boldsymbol{\omega}\|_2^2 \\
\text{s.t. } 0 &\leq \boldsymbol{\omega} + \Delta\boldsymbol{\omega}
\end{aligned} \tag{8}$$

where the subscript  $d$  corresponds to the sub-sampling strategy that will segment the non-optimal trajectories into  $N_w$  sub-trajectories containing only  $d$  samples for given time windows. This allows the addition of another layer in the time-space by weighting trajectories of different lengths differently using  $\theta_d = \frac{\lfloor \frac{T}{dt} \rfloor - d + 1}{\lfloor \frac{T}{dt} \rfloor + 1}$ . Therefore,  $\bar{\mathbf{x}}_d$  is the set of non-optimal trajectories of size  $d$ .  $\bar{\Phi}_{d,\mathbf{x}} = \Phi_d(\mathbf{x}) - \Phi(\mathbf{x}^*)$  is the difference in costs between sub-sampled non-optimal trajectories and the optimal trajectory.  $\gamma_{\mathbf{x}} = e^{-\boldsymbol{\omega}^\top \bar{\Phi}_{d,\mathbf{x}}}$  comes from the reformulation in  $\Delta\boldsymbol{\omega}$  and represents a weighting for the more probable non-optimal trajectories. Finally,  $\lambda$  and  $\beta$  are the parameters inducing the elastic net regularization on  $\Delta\boldsymbol{\omega}$ .

Once  $\Delta\boldsymbol{\omega}$  is determined, the line-search procedure will accept a step  $\alpha$  in the direction of  $\Delta\boldsymbol{\omega}$  based on the following merit function  $m(\boldsymbol{\omega} + \Delta\boldsymbol{\omega}) = \|\tilde{\mathbf{q}} - \mathbf{q}^*\|_2$ , where  $\tilde{\mathbf{q}}$  is the predicted joint angle trajectories after taking an  $\alpha$  step and  $\mathbf{q}^*$  are the optimal joint angle trajectories. This merit function has empirically proved to achieve better results in terms of joint angle accuracy. When an appropriate step is found, the output joint angle trajectory is added to the pool of non-optimal trajectories  $\bar{\mathbf{x}}$ . If no acceptable step is found within 10 iterations, we accept a minimal step, and the process continues. Then, the algorithm reiterates until reaching a maximum number of iterations threshold or sufficient accuracy in joint angle prediction. Finally, the parameter  $N$  is set as the number of non-optimal trajectories considered in  $\bar{\mathbf{x}}$ . For instance, if  $N = 1$ , we only take the last computed non-optimal trajectory as the set of non-optimal trajectories. All the optimization problems were solved in Python using the CasADi library [28] and the Ipopt solver [29].

#### D. Algorithm parameters tuning

As exposed in Section II-C.2, the MO-IRL algorithm relies on several parameters, whose influences are still debatable. To optimally tune these parameters for the lifting task and to analyze MO-IRL sensitivity, a grid search procedure, analogously to what is done in the machine learning community, was devised. Since the MO-IRL algorithm ran very quickly, numerous parameters were investigated:

- 1) the number of time windows  $N_w$  with values equals to (4, 6, 8). Note that this parameter also affected how the non-optimal trajectories' sub-sampling was done;

TABLE II: RMSE on the joint angles predicted either with MO-IRL tuned with participant-specific optimal parameter set and measured with the motion capture during the first trial; or with MO-IRL tuned with  $\Pi$  and measured with the motion capture during the 19 remaining trials.

| -           | MO-IRL Optimal | MO-IRL $\Pi$   |
|-------------|----------------|----------------|
| $q_1$ [deg] | $3.0 \pm 1.2$  | $3.9 \pm 1.9$  |
| $q_2$ [deg] | $6.4 \pm 1.5$  | $10.6 \pm 4.4$ |
| $q_3$ [deg] | $7.7 \pm 3.2$  | $10.8 \pm 3.8$ |
| $q_4$ [deg] | $8.2 \pm 2.2$  | $13.6 \pm 3.6$ |
| $q_5$ [deg] | $14.0 \pm 4.5$ | $18.5 \pm 5.6$ |
| Mean [deg]  | $7.8 \pm 4.5$  | $11.5 \pm 6.2$ |

- 2) the L1-Norm regularization parameter  $\lambda$  with values equals to  $(1e^{-6}, 1e^{-3}, 1e^{-1})$ ;
- 3) the L2-Norm regularization parameter  $\beta$  with values equals to  $(1e^{-4}, 1e^{-2}, 1)$ ;
- 4) the scale  $\sigma$  that sets the initial values of  $\omega$  before starting the algorithm. This parameter allowed the algorithm to explore different warm starts. Its value was set equal to  $(0, 1e^{-3}, 1e^{-1})$ ;
- 5) the maximum number of non-optimal trajectories  $N$ , whose values in  $(\text{all}, 1)$ , where the parameter all refers to taking the whole non-optimal trajectories set at each iteration.

The grid search procedure was performed on the first box lifting trial for each participant in this study. Once the best parameters were found, the remaining participants' trials were treated with the same parameter values, called  $\Pi$ . Indeed,  $\Pi$  will simplify the evaluation of the algorithm across all the participants, and the analysis of the contributions of costs driving the motion. The grid search analysis, as well as the accuracy of prediction across all participants and trials, are presented in Section III.

### III. RESULTS

The grid-search procedure was run on each participant's first trial to identify the set of MO-IRL parameter values that yielded the most accurate predictions when applied to the remaining 19 trials. Spanning 162 grid points per participant, this procedure produced many candidate parameter sets. To select the optimal values, the prediction accuracy assessment reported in Table III was performed. For each participant, the parameter values associated with the 20 lowest ("top-20") and the 20 highest ("bottom-20") RMSEs for joint-angle predictions were retained. We then tallied, across all participants, how often each parameter value occurred in the top-20 and bottom-20 sets and computed the corresponding averaged RMSE. For example,  $N_w = 8$  appeared 51 times among the top-20 results, with an average RMSE of  $10.29 \pm 1.53\text{deg}$ , and 41 times among the bottom-20 results, with an average RMSE of  $19.17 \pm 1.73\text{deg}$ . Our selection criterion thus balanced high frequency in the top-20, low frequency in

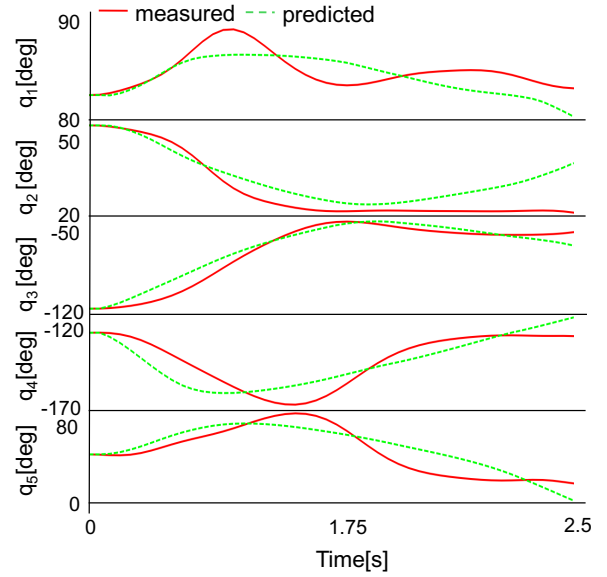


Fig. 3: Representative comparison of measured and estimated joint angles during a randomly selected lifting trial.

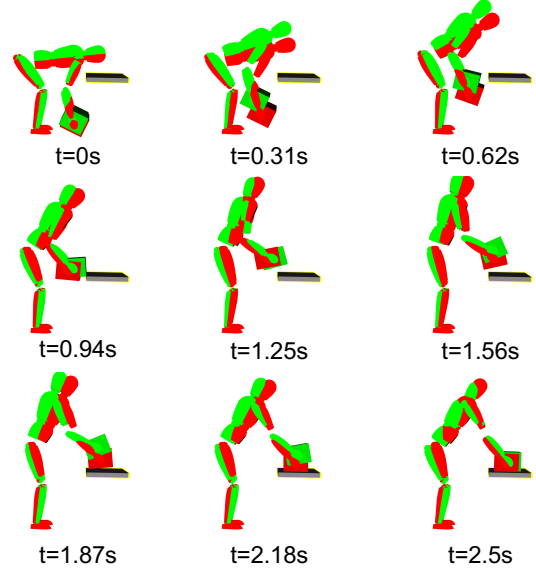


Fig. 4: Representative animation of measured and estimated motion during a randomly selected lifting trial.

the bottom-20, and low overall RMSE. The chosen parameter set  $\Pi$  was:

- $\Pi = (N_w = 8, \lambda = 0.1, \beta = 1, \sigma = .001 \text{ and } N = 1)$

Fig. 3 and 4 show the comparison between measured (red) and predicted (green) joint angle trajectories for one representative trial. The corresponding RMSE was  $7.9 \pm 3.5\text{deg}$  across all joints. Regarding joint angle prediction accuracy, Table II presents two different results: first, the most "optimal" results obtained with MO-IRL tuned with participant-specific best parameters values that were found by the grid search on the first trial; and secondly on the



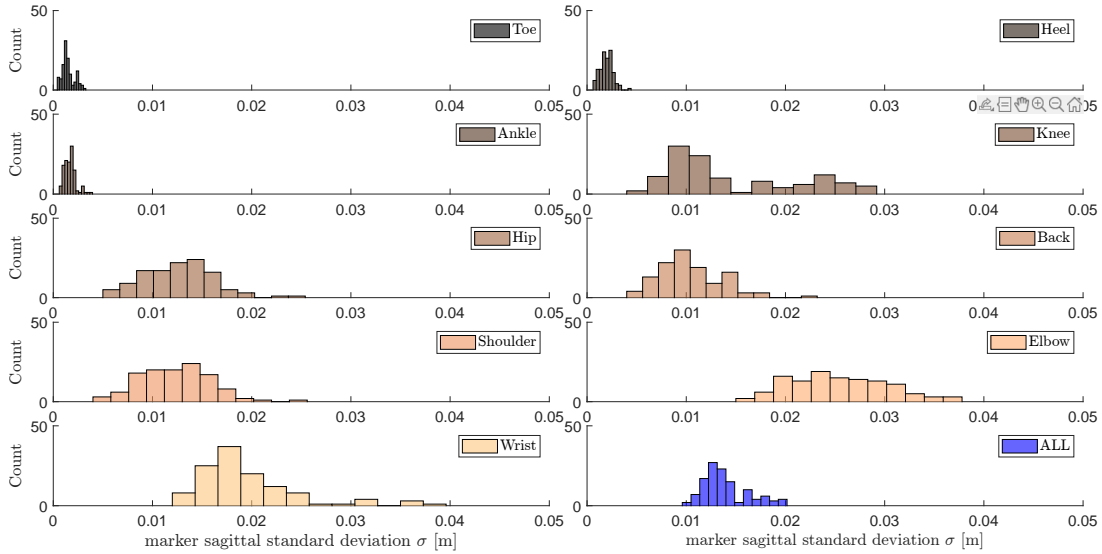


Fig. 5: Histograms of the standard deviations of markers from the sagittal plane across all collected trials.

TABLE III: Results of the hyper-parameter grid search. Columns are grouped by parameter ( $N_w$ ,  $\lambda$ ,  $\beta$ ,  $\sigma$ ,  $N$ ). For every candidate value we report: (i) the mean  $\pm$  standard deviation of the joint-angle RMSE computed over each participant’s 20 best runs (“top-20”, first block of rows) and 20 worst runs (“bottom-20”, second block); and (ii) how many participants’ top-20 (or bottom-20) sets contain that value summed over the 6 participants (“Occurrences”). Within each parameter group **boldface** marks (1) the lowest mean RMSE in the top-20 block, (2) the highest occurrence count in the top-20 block, and (3) the lowest occurrence count in the bottom-20 block. The values bolded in the header row are the hyperparameter settings ultimately chosen for the final model (they satisfy at least one of the above optimality criteria, typically all four).

| Parameter               | $N_w$        |            |              | $\lambda$ [ $\log_{10}$ ] |            |              | $\beta$    |            |              | $\sigma$   |            |              | $N$          |            |
|-------------------------|--------------|------------|--------------|---------------------------|------------|--------------|------------|------------|--------------|------------|------------|--------------|--------------|------------|
| Values                  | 4            | 6          | <b>8</b>     | −6                        | −3         | −1           | .0001      | .01        | <b>1</b>     | 0          | .001       | .1           | 1            | all        |
| mean $\pm$ std.         | 10.46        | 10.48      | <b>10.29</b> | <b>10.29</b>              | 10.54      | 10.44        | 10.60      | 10.49      | <b>10.22</b> | 10.45      | 10.42      | <b>10.37</b> | <b>10.37</b> | 10.45      |
| RMSE [deg] (top-20)     | $\pm 1.47$   | $\pm 1.40$ | $\pm 1.53$   | $\pm 1.44$                | $\pm 1.60$ | $\pm 1.42$   | $\pm 1.67$ | $\pm 1.48$ | $\pm 1.35$   | $\pm 1.53$ | $\pm 1.40$ | $\pm 1.63$   | $\pm 1.59$   | $\pm 1.33$ |
| Occurrences (top-20)    | 27           | 42         | <b>51</b>    | 40                        | 30         | <b>50</b>    | 20         | 43         | <b>57</b>    | 48         | <b>50</b>  | 22           | <b>62</b>    | 58         |
| mean $\pm$ std.         | <b>18.55</b> | 19.57      | 19.17        | 19.26                     | 19.08      | <b>18.73</b> | 19.05      | 19.82      | <b>18.23</b> | 18.80      | 19.69      | <b>18.32</b> | <b>18.51</b> | 19.32      |
| RMSE [deg] (bottom-20)  | $\pm 1.35$   | $\pm 1.78$ | $\pm 1.73$   | $\pm 1.70$                | $\pm 1.79$ | $\pm 1.12$   | $\pm 1.47$ | $\pm 1.97$ | $\pm 1.53$   | $\pm 1.86$ | $\pm 1.79$ | $\pm 1.51$   | $\pm 1.95$   | $\pm 1.47$ |
| Occurrences (bottom-20) | 42           | <b>37</b>  | 41           | 50                        | 39         | <b>31</b>    | 57         | 35         | <b>28</b>    | <b>22</b>  | 57         | 41           | <b>43</b>    | 77         |

remaining 19 trials across all the participants when joint angles were predicted using MO-IRL and the parameter set **II**. On average, the RMSE was  $7.8 \pm 4.5\text{deg}$  for the most “optimal” set of parameters and  $11.5 \pm 6.2\text{deg}$  when using **II**.

Finally, Fig. 6 shows the average and standard deviation of the identified weights  $\omega$  across all participants for each different cost function at each movement phase.

#### IV. DISCUSSION

The influence of the different values of each parameter on the MO-IRL outcome is contrasted. Therefore, finding a way to build **II** was not trivial. Some parameters, like  $N_w$ ,  $\lambda$ , and  $\beta$ , seemed to dramatically affect the type of solution obtained. Indeed, higher  $N_w$  led to more accurate predictions, with more frequent occurrences in the top-20 and lower RMSEs. However, higher  $N_w$  also resulted in increased computational time and likely overfitting. For the sake of accuracy,  $N_w = 8$  was chosen for this study, but it appears that  $N_w = 6$  could also represent a good trade-

off between accuracy, computational time, and overfitting. Secondly,  $\lambda$  and  $\beta$  tuned the elastic net regularization, especially regarding whether the recovered weights  $\omega$  were sparse or not, and whether they varied significantly between iterations. For  $\beta$ , it was fortunately straightforward to observe that  $\beta = 1$  outperformed the other values on all criteria. This indicates that MO-IRL favored low-magnitude solutions with non-exact zero weights, implying that all the cost functions considered have a relative impact on accurately predicting a lifting motion. For  $\lambda$ , the conclusion is less clear, with  $\lambda = 1e^{-6}$  and  $\lambda = 1e^{-1}$  both contributing to most of the top-20 and best RMSEs results. This may indicate that, in some trials,  $\omega$  was found to be more sparse, suggesting certain objective functions were more critical to minimize, while in other trials, this distinction was less evident. Since the goal was to strongly discriminate between all objective functions, the choice  $\lambda = 0.1$  and  $\beta = 1$  was made. For the last two parameters,  $\sigma$  and  $N$ , their influence appeared less significant.  $\sigma$  was intended to induce different warm-starts for the algorithm, but it was unclear which led to better

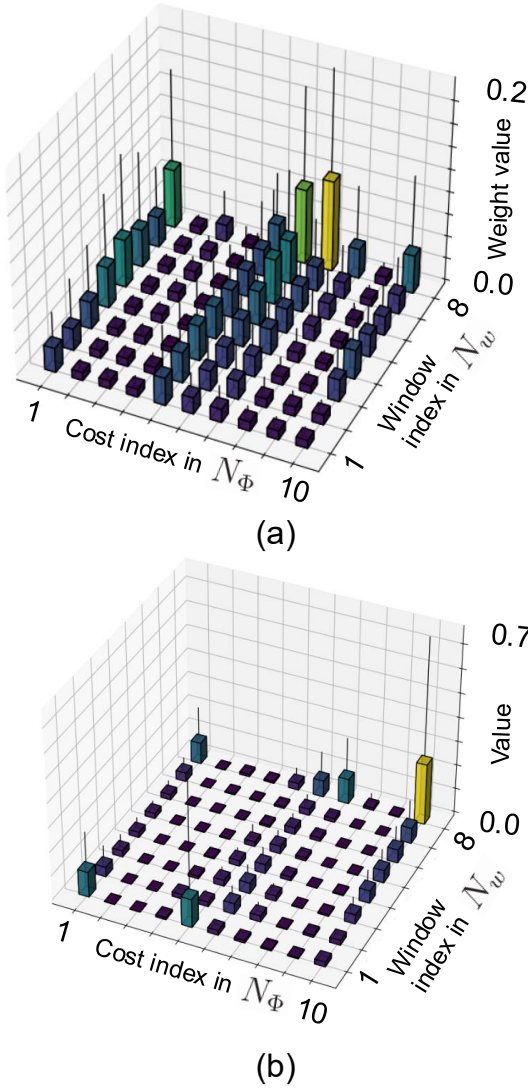


Fig. 6: (a) Average (+1std) distribution of the weights  $\omega$  across all trials and all participants. (b) Average (+1std) contributions of each weighted cost (i.e.,  $\omega^T \times \Phi$ ) across all trials and all participants.

outcomes. Since  $\sigma = 1e^{-3}$  appeared more frequently in the top-20 and produced the best result among all trials, it was the retained value. For  $N$ , it seems that  $N = 1$  yielded slightly better results. This can be explained by the fact that, in terms of accuracy, retaining the most accurate recent solution in the pool of non-optimal trajectories tended to shift the algorithm toward increasingly accurate solutions. However, this may lead to poor generalization of the solution to unseen demonstrations.

Regarding joint angles accuracy, it can be seen that the lower limbs joint angles were predicted with a good  $8.4 \pm 3.4\text{deg}$  accuracy. However, the upper limb joints were less accurately predicted. The average RMSE for the shoulder and elbow was  $16.1 \pm 4.6\text{deg}$ . This can likely be explained by the planar modeling hypothesis. As shown in Fig. 5, showing standard deviations of the markers out of the

sagittal plane, the markers located at the upper limbs show much higher deviation. The deviations were on average of  $0.65 \pm 0.68\text{cm}$  and  $1.69 \pm 0.73\text{cm}$  for the lower and upper limbs, respectively. Using simple trigonometric relationships between this error and the segment lengths led to errors in the joint angles as high as  $6.5\text{deg}$  for the shoulder, and  $12\text{deg}$  for the elbow. These values are of the same order of magnitude as the reported error for the predicted joint angles. Thus, future works will investigate the use of a planar model for the lower limbs and a 3D model for the upper limbs.

The study of the recovered weights  $\omega$  allowed us to gain a deeper understanding of the strategies used when lifting a heavy load. As shown in Fig. 6a, the distribution of  $\omega$  throughout the task duration exhibits varying cost contributions. First, the lower limb joint torque weights remained high across the entire motion. Interestingly, this results in a  $\Phi_1 * \omega_1$  contribution that is relatively low except at the beginning and end of the motion, as illustrated in Fig. 6b. This suggests that participants primarily prioritized a low-strain back strategy, i.e., keeping their trunk as vertical as possible. A higher weight for the elbow joint torque is also noted in the final windows, clearly indicating that participants approached the box landing more smoothly, i.e., compensating for the box's inertia. In addition, three other biomechanical cost functions,  $\Phi_5$ ,  $\Phi_6$ , and  $\Phi_7$ , corresponding respectively to joint accelerations, cartesian box velocity, and joint torque change, were increasingly minimized throughout the motion. This indicates that participants preferred to perform the task at a constant velocity and highly prioritized slowing the box to ensure a safe landing on the table. Finally,  $\Phi_{10}$ , which may relate to apprehension about striking the table with the box, did not play a significant role in the initial movement phases. However, it became more engaged in the middle phase, likely to help avoid the table, and again at the end of the motion. While this makes sense in the middle of the motion, its contribution at the end is harder to interpret. During the motion, minimizing this cost tends to bring the box closer to the body. Since the final pose of the box is constrained, minimizing this cost may instead pull the body toward the table. In this sense, it functions similarly to the minimization of lower limb joint torques. This overlap illustrates a drawback of treating a constraint as a cost function: different cost functions can serve similar roles. Indeed, this may indicate overfitting due to the chosen basis of cost functions. While this is not problematic for motion prediction, it is a limitation when attempting to explain the motion. To address this issue, results from IOC, particularly studies related to the linear dependencies in the DOC's Lagrangian matrix and the Karush-Kuhn-Tucker conditions [17,30], should be explored to automatically reduce the cost function basis. This can potentially avoid certain cost terms to unintentionally bias the predicted motion away from natural or safe movement strategies.

## V. CONCLUSION

The present study demonstrates the application of the recently released Minimal Observations Inverse Reinforcement

Learning algorithm for predicting complex human box-lifting motions. To appropriately tune all the parameters of the algorithm, a sensitivity analysis was conducted to find the best set of parameters. Once determined, the algorithm was able to accurately predict the measured joint angles. Future work will focus on modeling the human upper limbs in 3D, studying the Direct Optimal Control problem in greater depth to determine a minimal basis of objective functions, and applying the method to other types of motion. Furthermore, these biomechanical cost functions can be integrated into a model-predictive control framework to govern a collaborative robot that anticipates human motion. This could also be applied to physical interaction in lifting tasks, enabling more natural assistance for workers.

#### ACKNOWLEDGMENTS

This project was partially funded by the French ANR-23-CE33-0010 HERCULES project.

#### REFERENCES

- [1] L. Lei, P. G. Dempsey, J.-g. Xu, L.-n. Ge, and Y.-x. Liang, "Risk factors for the prevalence of musculoskeletal disorders among chinese foundry workers," *Int. J. Ind. Ergon.*, vol. 35, pp. 197–204, 2005.
- [2] E. A. Chumacero-Polanco and J. Yang, "A review on human motion prediction in sit to stand and lifting tasks," in *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, vol. 50077. American Society of Mechanical Engineers, 2016, p. V01AT02A066.
- [3] S. M. Hsiang and M. Ayoub, "Development of methodology in biomechanical simulation of manual lifting," *Int. J. Ind. Ergon.*, vol. 13, pp. 271–288, 1994.
- [4] S. M. Hsiang and R. W. McGorry, "Three different lifting strategies for controlling the motion patterns of the external load," *Ergonomics*, vol. 40, pp. 928–939, 1997.
- [5] T. Bernard, M. Ayoub, and C. Lin, "Evaluation of a biomechanical simulation model for sagittal plane lifting," *Int. J. Indus. Ergon.*, vol. 24, pp. 157–171, 1999.
- [6] Y. Xiang, J. S. Arora, S. Rahmatalla, T. Marler, R. Bhatt, and K. Abdel-Malek, "Human lifting simulation using a multi-objective optimization approach," *Multibody Syst. Dyn.*, vol. 23, pp. 431–451, 2010.
- [7] Y. Xiang, S. Rahmatalla, H.-J. Chung, J. Kim, R. Bhatt, A. Mathai, S. Beck, T. Marler, J. Yang, J. S. Arora *et al.*, "Optimization-based dynamic human lifting prediction," SAE Technical Paper, Tech. Rep., 2008.
- [8] A. L. Shoushtari, "What strategy central nervous system uses to perform a movement balanced? biomechatronical simulation of human lifting," *Appl Bionics Biomech*, vol. 10, pp. 113–124, 2013.
- [9] M. J. Dysart and J. C. Woldstad, "Posture prediction for static sagittal-plane lifting," *J. Biomech.*, vol. 29, pp. 1393–1397, 1996.
- [10] H. Arisumi, S. Miossec, J.-R. Chardonnet, and K. Yokoi, "Dynamic lifting by whole body motion of humanoid robots," in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2008, pp. 668–675.
- [11] R. Rakshit, Y. Xiang, and J. Yang, "Dynamic-joint-strength-based two-dimensional symmetric maximum weight-lifting simulation: Model development and validation," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 234, pp. 660–673, 2020.
- [12] C. Lin, M. Ayoub, and T. Bernard, "Computer motion simulation for sagittal plane lifting activities," *Int. J. Indus. Ergon.*, vol. 24, pp. 141–155, 1999.
- [13] J. Song, X. Qu, and C.-H. Chen, "Lifting motion simulation using a hybrid approach," *Ergonomics*, vol. 58, pp. 1557–1570, 2015.
- [14] —, "Simulation of lifting motions using a novel multi-objective optimization approach," *Int. J. Indus. Ergon.*, vol. 53, pp. 37–47, 2016.
- [15] A. Meduri, H. Zhu, A. Jordana, and L. Righetti, "Mpc with sensor-based online cost adaptation," in *Proceedings of the 2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 996–1002.
- [16] Y. Xiang, J. S. Arora, and K. Abdel-Malek, "Hybrid predictive dynamics: a new approach to simulate human motion," *Multibody System Dynamics*, vol. 28, pp. 199–224, 2012.
- [17] J. Colombel, D. Daney, and F. Chappillet, "On the reliability of inverse optimal control," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 8504–8510.
- [18] F. Bečanović, J. Miller, V. Bonnet, K. Jovanović, and S. Mohammed, "Assessing the quality of a set of basis functions for inverse optimal control via projection onto global minimizers," in *2022 IEEE 61st Conference on Decision and Control (CDC)*. IEEE, 2022, pp. 7598–7605.
- [19] N. Sylla, V. Bonnet, G. Venture, N. Armande, and P. Fraisse, "Human arm optimal motion analysis in industrial screwing task," in *5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics*, 2014, pp. 964–969.
- [20] B. D. Ziebart, A. L. Maas, J. A. Bagnell, A. K. Dey *et al.*, "Maximum entropy inverse reinforcement learning," in *Aaai*, vol. 8. Chicago, IL, USA, 2008, pp. 1433–1438.
- [21] C. Finn, S. Levine, and P. Abbeel, "Guided cost learning: Deep inverse optimal control via policy optimization," in *International conference on machine learning*. PMLR, 2016, pp. 49–58.
- [22] S. Mehrdad, A. Meduri, and L. Righetti, "Cost function estimation using inverse reinforcement learning with minimal observations," 2025. [Online]. Available: <https://arxiv.org/abs/2505.08619>
- [23] T.-W. Lu and J. O'connor, "Bone position estimation from skin marker co-ordinates using global optimisation with joint constraints," *J. Biomech.*, vol. 32, pp. 129–134, 1999.
- [24] V. Bonnet and G. Venture, "Fast determination of the planar body segment inertial parameters using affordable sensors," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, pp. 628–635, 2015.
- [25] J. Carpentier, G. Saurel, G. Buondonno, J. Mirabel, F. Lamiroux, O. Stasse, and N. Mansard, "The pinocchio c++ library – a fast and flexible implementation of rigid body dynamics algorithms and their analytical derivatives," *IEEE/SICE*, 2019.
- [26] B. Berret, E. Chiovetto, F. Nori, and T. Pozzo, "Evidence for composite cost functions in arm movement planning: an inverse optimal control approach," *PLoS computational biology*, vol. 7, no. 10, p. e1002183, 2011.
- [27] T. Robert, J. Causse, L. Denninger, and X. Wang, "A 3d analysis of the joint torques developed during driver's ingress–egress motion," *Ergonomics*, vol. 57, pp. 1008–1020, 2014.
- [28] J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "CasADi – A software framework for nonlinear optimization and optimal control," *Math. Program. Comput.*, vol. 11, pp. 1–36, 2019.
- [29] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Math. Program.*, vol. 106, pp. 25–57, 2006.
- [30] F. Becanovic, K. Jovanović, and V. Bonnet, "Reliability of single-level equality-constrained inverse optimal control," in *2024 IEEE-RAS 23rd International Conference on Humanoid Robots (Humanoids)*. IEEE, 2024, pp. 623–630.