



**HAL**  
open science

# A Finite-Horizon Inverse Differential Game Approach for Optimal Trajectory-Tracking Assistance with a Wrist Exoskeleton

Abdelwaheb Hafs, Dorian Verdel, Etienne Burdet, Olivier Bruneau, Bastien Berret

## ► To cite this version:

Abdelwaheb Hafs, Dorian Verdel, Etienne Burdet, Olivier Bruneau, Bastien Berret. A Finite-Horizon Inverse Differential Game Approach for Optimal Trajectory-Tracking Assistance with a Wrist Exoskeleton. IEEE RAS EMBS 10th International Conference on Biomedical Robotics and Biomechanics (BioRob 2024), Sep 2024, Heidelberg, Germany. hal-04443499v2

**HAL Id: hal-04443499**

**<https://hal.science/hal-04443499v2>**

Submitted on 25 Aug 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A Finite-Horizon Inverse Differential Game Approach for Optimal Trajectory-Tracking Assistance with a Wrist Exoskeleton

Abdelwaheb Hafsi , Dorian Verdel , Etienne Burdet , Olivier Bruneau , Bastien Berret 

**Abstract**—Exoskeletons are appealing robotic devices to physically assist humans in various motor tasks. To provide a part of the effort required for performing a task, they should be intuitive to use and adapt to the user’s goal. Differential game theory offers an interesting framework to formalize the shared control problem underlying physical human-robot interaction. In the present paper, we introduce an approach based on finite-horizon inverse differential games, which allows to iteratively infer the user’s internal goals and design a Nash-equilibrium control policy. Here, we focus on a case study in which a user has to move a load along a target trajectory while assisted by a wrist exoskeleton. The method is first validated in simulations and then applied to the control of the HRX-1 wrist interface. The results show that the controller has a positive impact on tracking performance and reduces the joint torque provided by the users while enabling them to remain active. Interestingly, it also yields to a more balanced sharing of task efforts and a better coordination between the robot and the user compared to a user-agnostic linear-quadratic control guidance.

## I. INTRODUCTION

Physical human-robot interaction is at the heart of many applications designed to help humans perform motor tasks, e.g. in industrial or medical settings [1], [2]. Besides hardware [3] and ergonomic considerations [4], the topic of interaction control [5], that is, deciding how to share task control with the human, is of crucial importance for the efficiency and acceptability of the device. Shared control should not only maximize task performance but also enable intuitive interaction. Arguably, a robot controller that integrates the internal motor goals of the human user would lead to improved interaction.

Differential Games (DG) can be used to mathematically formalize shared control problems where all players are dynamically coupled and each player aims at minimizing their own cost function, which results in a so-called non-cooperative DG [6]. Interestingly, Jarrassé et al. provided a DG framework for human-robot interaction and classified interaction paradigms in two-player interactive tasks according to the cost functions that agents aim to minimise [7]. Furthermore, human-human interaction has been found to exhibit the characteristics of optimal strategies in a game-theoretic sense, i.e. Nash equilibria, especially when enough

information can be gathered about the partner’s strategy [8], [9]. In physical human-human and human-robot interaction, it has been shown that haptic feedback is a fundamental modality that allows players called connected agents, to exchange information about common movement goals [10], [11].

Li et al. successfully used non-cooperative DG to optimally assist humans with an endpoint robot interface during reaching movements [12]. They considered an infinite-horizon linear-quadratic DG framework with fixed targets and demonstrated that, when combined with an observer to identify the human user goal, an efficient and stable interaction between the robot and the human user was obtained. Musić et al. generalized this approach to the tracking of arbitrary trajectories [13]. While they considered a finite-horizon linear-quadratic DG problem and also estimated the human user strategy with an adaptive observer, their solution involved taking the horizon limit to infinity. Recently, Pezeshki et al. developed an adaptive assist-as-needed control for rehabilitation with a neural network approximation to identify the human strategy considering infinite-horizon linear-quadratic DG [14]. In summary, previous works used an infinite planning horizon in their practical implementations, which may appear unlikely in real scenarios. Necessarily, the human user will have knowledge of the upcoming trajectory only for a relatively short time horizon. Consequently, if the robot were to infer the intended trajectory of a human, this prediction should also be limited to a short time horizon [15].

In this paper, we wanted to circumvent this infinite-horizon limitation to implement an online game-theoretic controller. To do so, we relied on recent development about inverse DG [16]. Inverse optimal control and inverse DG provide a set of methods to uncover cost functions from the observation of experimental trajectories. Inverse optimal control has been used successfully in the human motor control literature [17] while inverse DG methods are better suited for human-robot interaction. Several methods have been developed for inverse DG [18], [19] but we will rely on a bi-level approach with the aim to make it usable in real-time on a receding finite horizon. This paper focuses on a one degree-of-freedom task that consists of moving a load along a desired trajectory in the sagittal plane with the help of a wrist exoskeleton. The methodology is generalizable to more degrees of freedom but this simple setting allows to test it systematically and in particular verify that (1) the method works in simulation and (2) the game-theoretic controller allows stable and efficient interaction with a human user.

The paper is organized as follows. Section II describes the

AH and BB are with the CIAMS, Université Paris-Saclay, 91405 Orsay, France. CIAMS, Université d’Orléans, Orléans, France. Corresponding e-mail: abdelwaheb.hafs@universite-paris-saclay.fr.

DV and EB are with the Imperial College of Science, Technology and Medicine, W12 0BZ London, United-Kingdom

OB is with the LURPA, ENS Paris-Saclay, 91190 Gif-sur-Yvette, France.

This work is supported in part by the French National Agency for Research (grant ANR-19-CE33-0009) and by the EC H2020 (grant ICT-871803-CONBOTS).

inverse DG approach for the task under consideration. Section III presents simulation results to validate the algorithm with a ground truth for the human cost and experimental results to test the resulting interactive behavior with a human. Concluding remarks and future work are given in Section IV.

## II. METHODS

### A. Dyadic system model

Let us consider a physical interaction between a wrist exoskeleton and a human user during a trajectory-tracking task in a common dynamic setting including gravity and friction. The interaction dynamics can thus be described by the following equation:

$$u_r + u_h = I\ddot{q} + D\dot{q} + G(q) \quad (1)$$

where  $u_r$  and  $u_h$  are respectively the robot and the human control variables that can change the exoskeleton's state  $\mathbf{x} \triangleq [q, \dot{q}]^\top$ . In the rest of the paper, the subscript  $r$  will refer to the robotic exoskeleton and  $h$  to the human user. The parameters  $I$ ,  $D$ , and  $G$ , represent the exoskeleton's inertia, damping and gravitational terms (including the potential load attached to it) respectively. Note that  $G$  is a nonlinear function proportional to  $\cos(q)$ .

Here, the task is to track a desired trajectory in the vertical plane which is *a priori* known, as shown in Figure 1a. This desired trajectory defined in the exoskeleton's joint coordinates is denoted by  $q_d(t)$  for  $t \in [0, T]$ , where  $T$  is the total duration of the tracking task. The desired state trajectory will be denoted by  $\mathbf{x}_d(t) \triangleq [q_d(t), \dot{q}_d(t)]^\top$ . There is an associated control  $u_d(t)$  satisfying (1), that is  $u_d = I\ddot{q}_d + D\dot{q}_d + G(q_d)$ .

By linearizing the dynamics in (1) around this reference trajectory-control pair, we get the following affine system to describe the interaction dynamics:

$$\begin{aligned} \dot{\boldsymbol{\xi}} &= \mathbf{A}\boldsymbol{\xi} + \mathbf{B}(u_r + u_h) + \mathbf{c}, & \boldsymbol{\xi} &\triangleq \mathbf{x} - \mathbf{x}_d \\ \mathbf{A} &\triangleq \begin{bmatrix} 0 & 1 \\ -G'(q_d)/I & -D/I \end{bmatrix}, & \mathbf{B} &\triangleq \begin{bmatrix} 0 \\ 1/I \end{bmatrix} \\ \mathbf{c} &\triangleq -\mathbf{B}u_d \end{aligned} \quad (2)$$

where  $G' = dG/dq$  represents the derivative of  $G$ . Note that  $\mathbf{A}$  and  $\mathbf{c}$  are time-varying functions after linearization of the state-space dynamics around  $\mathbf{x}_d(t)$ .

### B. Human-Exoskeleton differential game

Here, we focus on the problem of effort sharing between the exoskeleton and the human user. To this aim, we assume that the two agents (human user and exoskeleton) continuously play a DG over a finite time horizon (referred to as the *planning horizon*). Where each agent minimises a quadratic cost function as follows:

$$J_r = \int_{t_c}^{t_c + \Delta_p} \boldsymbol{\xi}(t)^\top \mathbf{Q}_r \boldsymbol{\xi}(t) + u_r(t)^2 + u_h(t)^2 dt \quad (3)$$

$$J_h = \int_{t_c}^{t_c + \Delta_p} \boldsymbol{\xi}(t)^\top \mathbf{Q}_h \boldsymbol{\xi}(t) + u_h(t)^2 dt \quad (4)$$

where  $t_c$  is the current (initial) time and  $\Delta_p$  is the planning horizon used for the fixed-time DG problem. The matrices  $\mathbf{Q}_r$  and  $\mathbf{Q}_h \in \mathbb{R}^{2 \times 2}$  are positive semi-definite (assumed to be diagonal here for simplicity) that define the weights associated with tracking error  $\boldsymbol{\xi}$ .

The cost functions defined in (3) and (4) state that the human user aims at minimising tracking error and their own effort expenditure  $u_h$ . Similarly, the exoskeleton aims at minimising tracking error and its effort  $u_r$ . However, for the assistance to be beneficial to the human user, we further assume that the exoskeleton also aims at minimising the effort of the human user  $u_h$  leading to a teacher-student relationship as in [7], [13].

For the DG problem, a Nash equilibrium exists, and the optimal control of each agent can be obtained by solving the following set of ordinary differential equations backwards in time [6, p. 323]:

$$-\dot{\mathbf{P}}_r = \mathbf{F}^\top \mathbf{P}_r + \mathbf{P}_r \mathbf{F} + \mathbf{Q}_r + \mathbf{P}_r \mathbf{B} \mathbf{B}^\top \mathbf{P}_r + \mathbf{P}_h \mathbf{B} \mathbf{B}^\top \mathbf{P}_h \quad (5a)$$

$$-\dot{\mathbf{P}}_h = \mathbf{F}^\top \mathbf{P}_h + \mathbf{P}_h \mathbf{F} + \mathbf{Q}_h + \mathbf{P}_h \mathbf{B} \mathbf{B}^\top \mathbf{P}_h \quad (5b)$$

$$-\dot{\boldsymbol{\alpha}}_r = \mathbf{F}^\top \boldsymbol{\alpha}_r + \mathbf{P}_r \boldsymbol{\beta} + \mathbf{P}_r \mathbf{B} \mathbf{B}^\top \boldsymbol{\alpha}_r + \mathbf{P}_h \mathbf{B} \mathbf{B}^\top \boldsymbol{\alpha}_h \quad (5c)$$

$$-\dot{\boldsymbol{\alpha}}_h = \mathbf{F}^\top \boldsymbol{\alpha}_h + \mathbf{P}_h \boldsymbol{\beta} + \mathbf{P}_h \mathbf{B} \mathbf{B}^\top \boldsymbol{\alpha}_h \quad (5d)$$

where

$$\mathbf{F} \triangleq \mathbf{A} - \mathbf{B} \mathbf{B}^\top (\mathbf{P}_r + \mathbf{P}_h) \quad (6a)$$

$$\boldsymbol{\beta} \triangleq \mathbf{c} - \mathbf{B} \mathbf{B}^\top (\boldsymbol{\alpha}_r + \boldsymbol{\alpha}_h). \quad (6b)$$

With no terminal cost, the values at time  $t_f = t_c + \Delta_p$  are  $\boldsymbol{\alpha}_h(t_f) = \boldsymbol{\alpha}_r(t_f) = \mathbf{0}$  and  $\mathbf{P}_h(t_f) = \mathbf{P}_r(t_f) = \mathbf{0}$ .

Then, the actual optimal controls of each agent, denoted by  $u_r^*$  and  $u_h^*$  respectively, can be expressed as follows:

$$u_r^* = -\mathbf{B}^\top (\mathbf{P}_r \boldsymbol{\xi} + \boldsymbol{\alpha}_r), \quad u_h^* = -\mathbf{B}^\top (\mathbf{P}_h \boldsymbol{\xi} + \boldsymbol{\alpha}_h) \quad (7)$$

We note that optimal controls have both open-loop  $\boldsymbol{\alpha}_{h,r}$  and feedback gain  $\mathbf{P}_{h,r}$  terms. In practice,  $\boldsymbol{\alpha}_h$  and  $\mathbf{P}_h$  are unknown but could be estimated if we knew the parameters of the human cost function, i.e. the weights defined in  $\mathbf{Q}_h$ .

### C. Bi-level estimation method

We present below a method to estimate the human cost function parameters, which corresponds to an inverse DG problem [16]. The inverse DG was resolved using a bi-level approach. The method consists in repeatedly solving the above DG problem (Eqs. 3, 4) on the current planning horizon starting at time  $t_c$  (referred to as the lower-level problem) while minimising the control estimation error over some past time interval (referred to as the upper-level problem) as follows:

$$\min_{\boldsymbol{\theta}} \int_{t_c - \Delta_e}^{t_c} \|u_{h,\boldsymbol{\theta}}^*(t) - u_{h,\hat{\boldsymbol{\theta}}}^*(t)\|^2 dt \quad (8)$$

where  $\boldsymbol{\theta} = [\theta_1 \theta_2]^\top$  is the vector of the human cost parameters (here  $\theta_1$  is the position error weight and  $\theta_2$  is the velocity error weight),  $u_{h,\boldsymbol{\theta}}^*$  is the optimal human control with these true (yet unknown) cost parameters  $\boldsymbol{\theta}$ ,

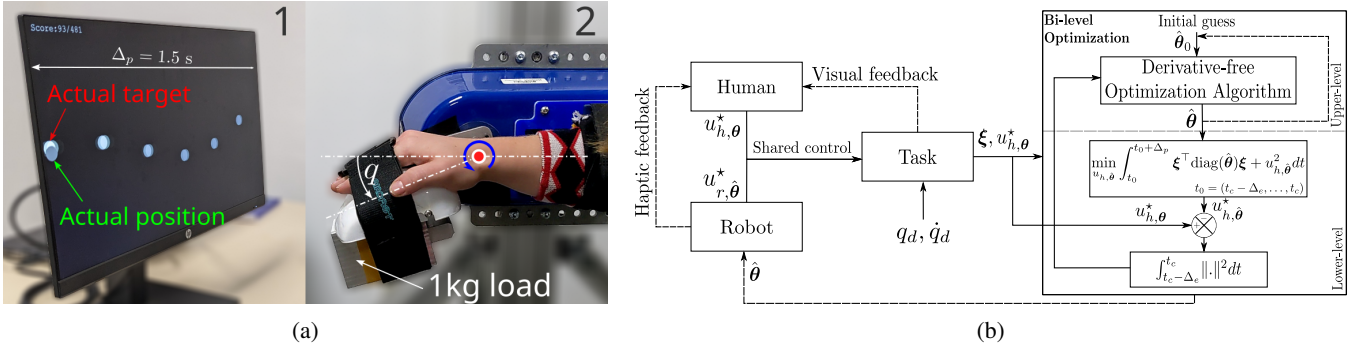


Fig. 1: Task description. (a): Experimental setup. (1) The desired trajectory displayed on a screen as white sliding dots with the actual position depicted as a grey dot. (2) The human and the exoskeleton (HRX-1) that can collaborate to accomplish the task. (b): Control and estimation workflow.  $\xi$  is the state deviation,  $u_{h,\theta}^*$  is the (optimal) human control,  $\hat{\theta}$  and  $u_{h,\hat{\theta}}^*$  are the estimated human cost parameters and corresponding (optimal) human control,  $u_{r,\hat{\theta}}^*$  is the exoskeleton (optimal) control based on the estimated human cost parameters.  $\hat{\theta}_0$  is the initial guess of the parameters of the cost function.  $\Delta_p$  and  $\Delta_e$  are the planning horizon and estimation epoch respectively.

and  $u_{h,\hat{\theta}}^*$  is the optimal control of the DG with the currently estimated cost parameters  $\hat{\theta}$ . The estimation time interval is  $[t_c - \Delta_e, t_c]$ , i.e. we use past observations where  $u_{h,\theta}^*$  can be measured (e.g., via interaction torque) or computed (e.g., via inverse dynamics) to evaluate the estimation error in the upper level. Note that while the tracking error could be incorporated, its computation was not necessary in these simple settings.

The main difficulty with this bi-level approach is that it can be computationally costly, which can limit its application in real-time control. However, it is possible to use only a few iterations of the upper-level optimization to improve  $\hat{\theta}$  and to use the current estimate as the initial guess for the next time step of the exoskeleton control loop [20] (see Fig. 1b). Finally, it is worth noting that the method is applied in a receding horizon fashion, so the problem will be different as new observations come in at each time step of the exoskeleton control loop.

#### D. Human-Exoskeleton Co-activity

For the sake of comparison, a linear-quadratic controller was implemented. This controller ignores the human, where equation (3) becomes:

$$J_r = \int_{t_c}^{t_c + \Delta_p} \xi(t)^\top \mathbf{Q}_r \xi(t) + u_r(t)^2 dt \quad (9)$$

with the corresponding ordinary differential equations [21]:

$$-\dot{\mathbf{P}}_r = \mathbf{A}^\top \mathbf{P}_r + \mathbf{P}_r \mathbf{A} + \mathbf{Q}_r - \mathbf{P}_r \mathbf{B} \mathbf{B}^\top \mathbf{P}_r \quad (10a)$$

$$-\dot{\alpha}_r = (\mathbf{A} - \mathbf{B} \mathbf{B}^\top \mathbf{P}_r)^\top \alpha_r + \mathbf{P}_r \mathbf{c}. \quad (10b)$$

#### E. Simulation settings

To test our method, we first simulated the motor task shown in Figure 1a. The desired trajectory is a sum of sine waves with different frequencies:

$$q_d(t) = \frac{\pi}{16} [1 - \cos(0.97t) - \cos(2.34t) - \cos(4.11t)] \quad (11)$$

The sampling time was  $t_s = 10$  ms in our simulations. We modeled a human hand for a person with a height

of 1.70 m and weight of 70 kg and extracted the dynamics parameters using anthropometric tables [22]. We also modeled a cylindrical 1-DoF wrist exoskeleton with mass of  $m = 1$  kg, length of 20 cm, and a radius of 3 cm. The dynamic parameters of the exoskeleton were set as  $I = 0.0152$  kg.m<sup>2</sup>, and  $D = 0.5$  kg.m<sup>2</sup>/s. The gravity term was  $G(q) = mgl \cos(q)$  with  $l = 0.1$  m. The simulated human was assumed to learn the exoskeleton control strategy (i.e. its cost function) by the haptic feedback [10], while the simulated exoskeleton had to estimate the human control strategy. The exoskeleton cost matrix was set as  $\mathbf{Q}_r = \text{diag}(30, 0.1)$  while we simulated five different couples of parameters of the human cost matrix to test the robustness of the method. The human cost matrix  $\mathbf{Q}_h$  was assumed to be diagonal  $\mathbf{Q}_h = \text{diag}(\theta)$  where  $\theta = [\theta_1, \theta_2]$ . The total simulation duration was set to  $T = 5$  s. The planning window ( $\Delta_p = 1.5$  s) and estimation window ( $\Delta_e = 0.25$  s) were selected pragmatically to offer good performance with the experiment setup limited by the control rate. To solve the upper-level problem from Equation (8), we employed a derivative free optimization algorithm called BOBYQA. We used the NLOpt toolbox in MATLAB [23], [24], with the following configuration settings: a maximum number of iterations set to 10 and an initial trust-region radius of 0.2. Finally, we ran the simulation with randomly attributed initial guesses  $\theta_0 \in [0, 30] \times [0, 1]$ .

#### F. Experimental setup

1) *Participants*: The method was then tested on ten healthy participants (4 males and 6 females, with age =  $24.6 \pm 2.27$  years, height =  $1.68 \pm 0.10$  m, and weight =  $63.3 \pm 12.18$  kg). The experimental protocol was approved by the ethical committee for research (CER-Paris-Saclay-2022-071) and the written consent of participants was obtained.

2) *HRX-1 exoskeleton*: The experiment was conducted on the 1-DoF wrist robotic interface HRX-1 (HumanRobotiX, UK) designed to study upper limb neuromechanics. We added an extra 1 kg load to simulate a load-carrying task (see Fig. 1a). This interface (load + exoskeleton link) can be

modeled as in Equation (2) with  $I = 0.0109 \text{ kg.m}^2$ ,  $D = 0.006 \text{ kg.m}^2/\text{s}$ ,  $m = 1.2 \text{ kg}$  and  $l = 0.0954 \text{ m}$ . The human user was attached to the exoskeleton using straps. In addition to the haptic feedback inherent to human-exoskeleton interaction, visual feedback of the current position and upcoming desired trajectory was provided on a screen (see Fig. 1a). The control rate of HRX-1 was 10 ms and uses the cost function defined in Equation (3), with  $\mathbf{Q}_r = \text{diag}(30, 0.1)$ . The BOBYQA parameters were set as in simulation with null initial guess. An adaptive initial trust-region radius reduction was added based on the estimation error, where the initial radius was fixed to (0.1, 0.0001), which dynamically adapted in proportion to the estimation error defined in Equation (8). This adaptive mechanism triggers when the error reaches 0.2 Nm, making the algorithm less sensitive to the noise in the human controller when the fitting error is small enough. We also set an upper boundary to the second estimated parameter of the human cost  $\hat{\theta}_2$  (velocity error weight) to 0.1, to ensure the system stability. Finally, the estimation time interval was set to  $\Delta_e = 0.25 \text{ s}$  and the planning horizon was  $\Delta_p = 1.5 \text{ s}$ .

3) *Evaluation task*: The desired trajectory was the same sum of sine waves as in Equation (11), which was displayed as a sliding sequence of dots on a finite horizon. This horizon corresponded to the planning horizon. To compare our game-theoretic controller to other controllers, we conducted an experiment with three conditions in separate blocks. *No assistance* (NA) condition was without assistance of the exoskeleton. In this case, the exoskeleton torque  $u_r$  was simply set to zero so that the exoskeleton would not assist the movement. A second *Co-activity* (LQ) condition was with assistance from a standard linear-quadratic regulator [7]. A third *Collaboration* (DG) condition was with the proposed DG-based controller [7].

Each participant was instructed to follow the displayed trajectory regardless of the exoskeleton contribution to the task. The experiment started with two familiarization trials of 1 minute each, which were both performed in the NA condition. Then the three conditions were tested during separate 2-minute blocks of continuous tracking, which were performed in a random order across participants. A 1-minute break was taken between blocks to avoid effects of fatigue.

4) *Data acquisition and analysis*: The angular positions of the human user and the exoskeleton were filtered (moving average filter with a window size of 5 samples) before numerical differentiation, which allowed for the computation of the real-time velocity and acceleration. This procedure allowed us to compute an estimation of  $u_h$  in real time using inverse dynamics. The resulting  $u_h$  was low-pass filtered (Butterworth, second order, 5 Hz cut-off frequency) after the experiment for analyses. Main possible effects of the tested conditions on the interaction were first assessed by a one-way ANOVA with  $p < 0.05$  as significance level. Whenever a significant difference was found, the  $\eta^2$  was provided to report the effect size. Pairwise t-tests were employed to assess the main effect of the conditions, with a significance level set at  $p < 0.05$ . Whenever a significant difference was

detected, Cohen's D was provided to report the effect size. All the statistical analyses were conducted with MATLAB.

### III. RESULTS

Below we present the simulation and experimental results for the proposed game-theoretic approach for the one-dimensional tracking task described above.

#### A. Simulation results

Figure 2 illustrates the simulated and estimated human cost parameters with the corresponding optimal control inputs. The simulation results exhibit a rapid estimation of parameters  $\theta$  across different values of  $\mathbf{Q}_h$ , with a final error smaller than 0.01%, even when starting from randomly assigned initial guesses. They also show the convergence of both the estimated human torque and the exoskeleton torque to the correct Nash-equilibrium solution. Therefore, these simulations validated the method, yielding correct results across a wide range of human cost parameters.

#### B. Experimental results

In this section, we present the results of the game-theoretic controller for all participants and the three exoskeleton control conditions.

1) *Estimation*: Figure 3 illustrates the estimated cost parameters for a representative participant over a 10-second window. Note that cost parameters were estimated in the three conditions using the same algorithm. Notably, in the DG condition, the parameters quickly reached a plateau in all cases, which is compatible with a Nash equilibrium. The steadiness of the estimated cost parameters further corroborates that the DG condition fosters a more consistent and reliable interaction between the human and the exoskeleton. In contrast, the other control conditions displayed fluctuations in the estimated cost parameters, thus failing to achieve any clear equilibrium state throughout the block.

We performed further analyses on all the participants to quantify these observations. Figure 4 illustrates the mean values of human cost parameters for each participant during the second half of each block. First, we observe that there are inter-individual differences on those parameters, suggesting that the method can adapt to each individual. Second, the position error weights  $\theta_1$  take higher values in the NA condition. This could be expected due to the absence of the exoskeleton's involvement in the task, leading to error corrections handled by the human alone. In contrast, the LQ condition led to smaller estimated weights. In this condition the human essentially became a follower and the exoskeleton a leader. Furthermore, the value of the estimated weights in the DG condition appeared to be quite stable compared to the other conditions (lower standard deviation). This reinforces the idea that this cost parameter could be proper to an individual in the task. This stability also confirms that the parameter reached a clear plateau in the DG condition. Regarding the velocity error weight  $\hat{\theta}_2$ , we observed no repeatable difference in the mean values among the three control conditions but the inter-individual variability was

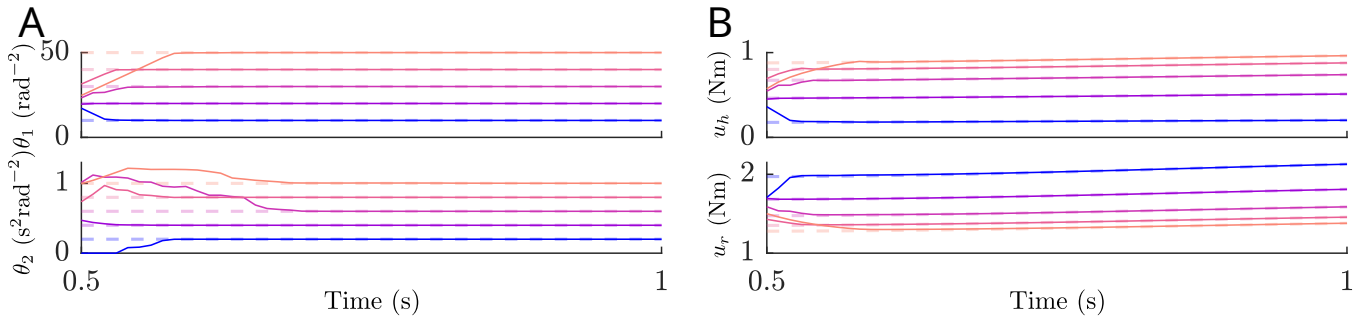


Fig. 2: Estimation of the human cost parameters in simulation. A: The true human cost parameters (position error weight  $\theta_1$  and velocity error weight  $\theta_2$ ) are dashed lines and the estimated parameters are solid lines. B: The true Nash equilibrium torques are dashed lines and the torques recovered the inverse DG method are solid lines.

again quite low in the DG condition. The lack of a clear pattern across conditions might be due to the relatively low values assigned to this weight and its limited importance for a system with a small inertia.

2) *Control performance*: To analyse the effects of the differences in the estimated cost functions, we examined the torque applied by the human user and the exoskeleton in all conditions. In Figure 5, we depict the human and exoskeleton torques alongside the system trajectory over the same time window. In the NA condition, both the tracking error and the human torque were higher. This was expected since the exoskeleton torque was null and the human performed the task alone in this condition. However, in the DG condition the torques were more evenly distributed between the human and the exoskeleton, reflecting a shared control of task-related efforts. The tracking errors were also smaller. In the LQ condition, the human torque approached zero and even attained negative values indicating that the human could transiently resist the applied exoskeleton torque, or just be passive [12]. The tracking error was low in this condition as well because the exoskeleton was active and aimed at performing the tracking task. These results are consistent with what could be expected from the different control conditions and support the relevance of our estimation method. They also suggest that shared task control may be more effective in the DG condition. Figure 6 provides a detailed analysis of the mean absolute torques of the human and the exoskeleton throughout the entire block duration in the three control conditions. The mean values of all absolute tracking errors are presented in the same figure to appreciate task performance concurrently. The results revealed a significant difference in the human torque for all conditions ( $p < 10^{-23}$ ,  $\eta^2 > 0.98$ ) and a larger level of human torque in DG compared to LQ ( $p < 10^{-10}$ ,  $d > 0.77$ ). There was no significant difference in tracking error between the two conditions ( $p > 0.44$  for the position and  $p > 0.29$  for the velocity, see Fig. 7) despite the important effect size returned by ANOVA ( $\eta^2 > 0.75$  for the position and  $\eta^2 > 0.91$  for the velocity). This observation reveals that in the DG condition, the human contribution is augmented but it still achieves a good tracking performance through the shared control with the exoskeleton, in contrast to the NA condition. Moreover,

in the DG condition, there was a decrease of the absolute exoskeleton torque compared to LQ ( $p < 10^{-15}$ ,  $d > 0.8$ ), indicating a shift toward a more balanced sharing of task control with the human user.

Furthermore, Figure 8 introduces a Coordination Index (CI) (i.e. correlation coefficient between the human and exoskeleton torques) to quantify how the exoskeleton and the human dynamically synchronize to track the trajectory. It was computed for each participant separately during the first 10 seconds and the last 10 seconds in both LQ and DG conditions. Interestingly, the results reveal a significant difference on CI in DG compared to LQ in the last 10 seconds ( $p < 10^{-3}$ ,  $d > 2.1$ ). This means that after 2 minutes of adaptation in a block, the human and the exoskeleton had a shared task control with better synchronization in the sense that their torque inputs were better correlated ( $r = 0.77$  on average). Moreover, CI tended to significantly improve within a block in the *Collaboration* condition ( $p = 0.0471$ ,  $d > 0.95$ ).

#### IV. CONCLUSION

We have presented a game-theoretic method for interaction control between a human and a robotic exoskeleton. The rationale of our approach was to estimate the human cost parameters to predict their upcoming contribution to the task and exploit this knowledge in the exoskeleton control policy [20]. For a trajectory-tracking task, we formulated a finite-horizon affine-quadratic DG and used a bi-level optimization to estimate the human cost parameters in real time. On the methodological side, this extends previous approaches in which DG controllers assumed an infinite horizon in practice [12], [13], [14]. Our premise was that humans may only consider the desired trajectory on a relatively short time horizon in practice. After validating the proposed method in simulation for a 1-DOF wrist system with known ground truth, we successfully implemented the method during a real task with a robotic wrist exoskeleton. Efforts in this task mainly originated from gravitational torques due to the carried load. It is worth noting that the nonlinear effects of gravity can be easily integrated in our settings although it leads to time-varying affine dynamics. As a consequence, we could not compare our method to previous DG methods assuming infinite horizon settings since they cannot handle time-varying affine systems.

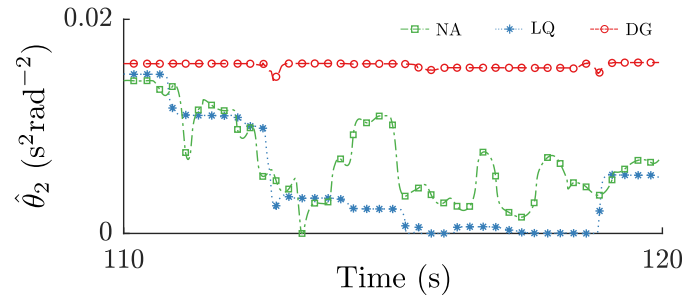
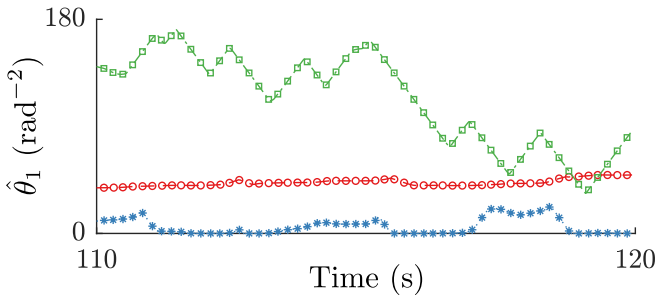


Fig. 3: Estimated parameters of the human cost parameters in experiment for one representative subject during 10 seconds of tracking in the different exoskeleton control conditions.

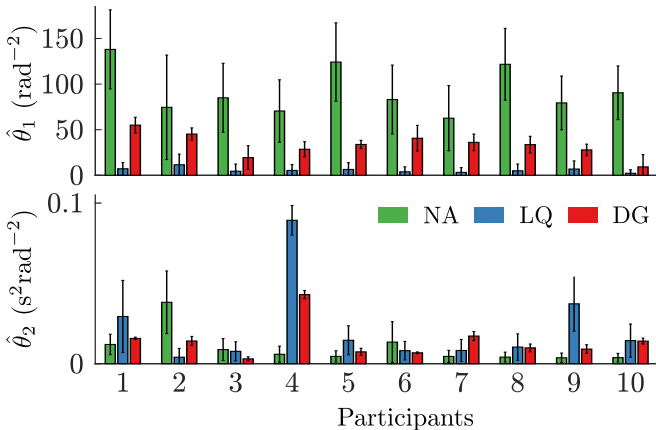


Fig. 4: Estimated human cost parameters for all participants during the last minute of the block.

Our findings indicate that the *Collaboration* condition promotes a more effective and collaborative control strategy that goes beyond the classical leader/follower dichotomy. Here, the exoskeleton was mostly a leader in the *Co-activity* condition whereas it was mostly a follower in the *No assistance* condition. Our results confirm that the *Collaboration* condition may lead to more intuitive collaborations with the exoskeleton as it naturally adapted to the human torque contribution without compromising the tracking performance. Another interesting finding is the higher coordination (or synchronization) between the human and the exoskeleton obtained with the game-theoretic controller, which tends to improve through practice as if the human user was able to learn the robot behavior (e.g., cost function). A limitation was the relatively small velocity weights that could be tested and identified. Simulation findings indicate that this weight may become more important for systems with a larger inertia.

Overall, our results demonstrate the effectiveness of our method to address real-world problems for 1-DOF systems when the upcoming trajectory can be predicted over a short time horizon. Future work should investigate the minimal planning horizon that is needed to maintain an efficient interaction control with this approach. When the desired trajectory is not known *a priori*, complementary methods should also be used to make a prediction on some time horizon [10], but this is beyond the scope of this paper.

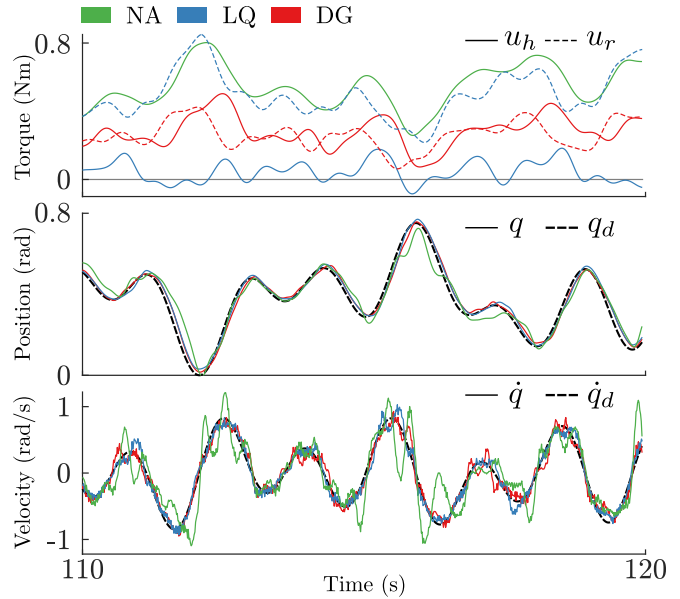


Fig. 5: Applied torques and tracking performance of one representative subject during 10 seconds with different exoskeleton control conditions.

Considering 1-DOF system with known trajectory is already interesting because effective wearable robotic exoskeletons often consider only one joint and rhythmical patterns (e.g., during walking), and rehabilitation applications often target a specific joint and can impose trajectories during exercises (e.g. wrist rehabilitation) [25], [26], [27], [28]. Extensions to multi-DOF systems are possible with a similar approach but the bi-level optimization would be more challenging.

## REFERENCES

- [1] A. Ajoudani, A. M. Zanchettin, S. Ivaldi, A. Albu-Schäffer, K. Kogure, and O. Khatib, "Progress and prospects of the human-robot collaboration," *Autonomous Robots*, vol. 42, p. 957–975, Oct. 2017.
- [2] D. P. Losey, C. G. McDonald, E. Battaglia, and M. K. O'Malley, "A review of intent detection, arbitration, and communication aspects of shared control for physical human-robot interaction," *Applied Mechanics Reviews*, vol. 70, Jan. 2018.
- [3] B. Vanderborght, A. Albu-Schaeffer, A. Bicchi, E. Burdet, D. Caldwell, R. Carloni, M. Catalano, O. Eiberger, W. Friedl, G. Ganesh, M. Garabini, M. Grebenstein, G. Grioli, S. Haddadin, H. Hoppner, A. Jafari, M. Laffranchi, D. Lefeber, F. Petit, S. Stramigioli, N. Tsagarakis, M. Van Damme, R. Van Ham, L. Visser, and S. Wolf, "Variable impedance actuators: A review," *Robotics and Autonomous Systems*, vol. 61, p. 1601–1614, Dec. 2013.

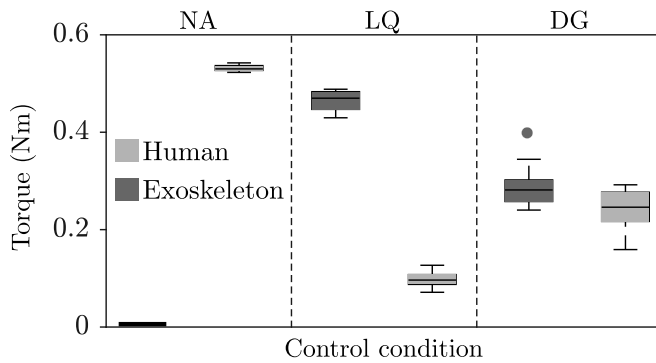


Fig. 6: Boxplot illustrating the mean of absolute values of human and exoskeleton torques, across all participants during the last minute of task duration in three control conditions

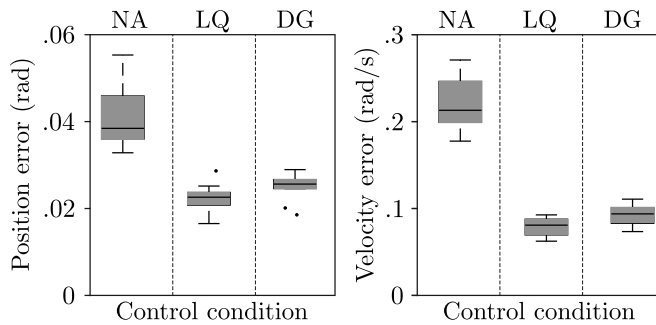


Fig. 7: Boxplots illustrating the mean values of absolute tracking position errors and velocity errors, across all participants during the entire task duration in three control conditions

[4] M. Lorenzini, M. Lagomarsino, L. Fortini, S. Gholami, and A. Ajoudani, "Ergonomic human-robot collaboration in industry: A review," *Frontiers in Robotics and AI*, vol. 9, 2023.

[5] Y. Li, A. Sena, Z. Wang, X. Xing, J. Babič, E. v. Asseldonk, and E. Burdet, "A review on interaction control for contact robots through intent detection," *Progress in Biomedical Engineering*, vol. 4, p. 032004, Aug. 2022.

[6] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, 2nd Edition. Classics in Applied Mathematics, Society for Industrial and Applied Mathematics, Jan. 1998.

[7] N. Jarrassé, T. Charalambous, and E. Burdet, "A framework to describe, analyze and generate interactive motor behaviors," *PLOS ONE*, vol. 7, p. e49945, Nov. 2012.

[8] D. A. Braun, P. A. Ortega, and D. M. Wolpert, "Nash equilibria in multi-agent motor interactions," *PLOS Computational Biology*, vol. 5, p. e1000468, Aug. 2009.

[9] V. T. Chackochan and V. Sanguineti, "Incomplete information about the partner affects the development of collaborative strategies in joint action," *PLOS Computational Biology*, vol. 15, p. e1006385, Dec. 2019.

[10] A. Takagi, G. Ganesh, T. Yoshioka, M. Kawato, and E. Burdet, "Physically interacting individuals estimate the partner's goal to enhance their movements," *Nature Human Behaviour*, vol. 1, Mar. 2017.

[11] A. Takagi, F. Usai, G. Ganesh, V. Sanguineti, and E. Burdet, "Haptic communication between humans is tuned by the hard or soft mechanics of interaction," *PLOS Computational Biology*, vol. 14, p. e1005971, Mar. 2018.

[12] Y. Li, G. Carboni, F. Gonzalez, D. Campolo, and E. Burdet, "Differential game theory for versatile physical human-robot interaction," *Nature Machine Intelligence*, vol. 1, p. 36–43, Jan. 2019.

[13] S. Musić and S. Hirche, "Haptic shared control for human-robot collaboration: A game-theoretical approach," *IFAC-PapersOnLine*, vol. 53, p. 10216–10222, Jan. 2020.

[14] L. Pezeshki, H. Sadeghian, M. Keshmiri, X. Chen, and S. Haddadin, "Cooperative assist-as-needed control for robotic rehabilitation: A two-

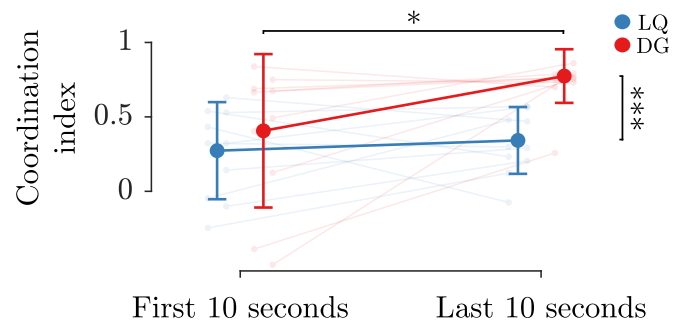


Fig. 8: Human-exoskeleton coordination index during the first and last 10 seconds of the task. Significant effects on the tested conditions are represented by \* if  $p < .05$  and \*\*\* if  $p < .001$

player game approach," *IEEE Robotics and Automation Letters*, vol. 8, p. 2852–2859, May 2023.

[15] L. Bashford, D. Kobak, J. Diedrichsen, and C. Mehring, "Motor skill learning decreases movement variability and increases planning horizon," *Journal of Neurophysiology*, vol. 127, p. 995–1006, Apr. 2022.

[16] T. L. Molloy, J. Inga Charaja, S. Hohmann, and T. Perez, *Inverse Optimal Control and Inverse Noncooperative Dynamic Game Theory: A Minimum-Principle Approach*. Communications and Control Engineering, Springer International Publishing, 2022.

[17] 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, p. e1002183, Oct. 2011.

[18] J. Inga, E. Bischoff, T. L. Molloy, M. Flad, and S. Hohmann, "Solution sets for inverse non-cooperative linear-quadratic differential games," *IEEE Control Systems Letters*, vol. 3, p. 871–876, Oct. 2019.

[19] T. L. Molloy, J. Inga, M. Flad, J. J. Ford, T. Perez, and S. Hohmann, "Inverse open-loop noncooperative differential games and inverse optimal control," *IEEE Transactions on Automatic Control*, vol. 65, p. 897–904, Feb. 2020.

[20] A. Hafs, D. Verdel, W. Gomes, O. Bruneau, and B. Berret, "Optimizing human-robot interactions through differential gamecontrol," in *Workshop on Multilimb Coordination in Human Neuroscience and Robotics: Classical and Learning Perspectives, IROS*, (Detroit (MI), United States), Oct. 2023.

[21] H. J. Kappen, *Optimal control theory and the linear Bellman equation*, p. 363–387. Cambridge: Cambridge University Press, 2011.

[22] D. A. Winter, *The biomechanics and motor control of human gait: normal, elderly and pathological*. University of Waterloo Press, 2nd ed ed., 1991.

[23] S. G. Johnson, "The NLOpt nonlinear-optimization package." <https://github.com/stevengj/nlopt>, 2007.

[24] M. J. D. Powell, "The BOBYQA algorithm for bound constrained optimization without derivatives," Tech. Rep. NA2009/06, Department of Applied Mathematics and Theoretical Physics, Cambridge University, Cambridge, UK, 2009.

[25] R. Colombo, F. Pisano, S. Micera, A. Mazzone, C. Delconte, M. Carrozza, P. Dario, and G. Minuco, "Robotic techniques for upper limb evaluation and rehabilitation of stroke patients," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 13, p. 311–324, Sept. 2005.

[26] N. Jarrassé, T. Proietti, V. Crocher, J. Robertson, A. Sahbani, G. Morel, and A. Roby-Brami, "Robotic exoskeletons: A perspective for the rehabilitation of arm coordination in stroke patients," *Frontiers in Human Neuroscience*, vol. 8, Dec. 2014.

[27] C. Siviý, L. M. Baker, B. T. Quinlivan, F. Porciuncula, K. Swaminathan, L. N. Awad, and C. J. Walsh, "Opportunities and challenges in the development of exoskeletons for locomotor assistance," *Nature Biomedical Engineering*, vol. 7, p. 456–472, Dec. 2022.

[28] P. Slade, M. J. Kochenderfer, S. L. Delp, and S. H. Collins, "Personalizing exoskeleton assistance while walking in the real world," *Nature*, vol. 610, p. 277–282, Oct. 2022.