Adaptive Impedance Controller for Human-Robot Arbitration based on Cooperative Differential Game Theory

Paolo Franceschi[1,2] Nicola Pedrocchi[1] and Manuel Beschi[2]

***Abstract*— The problem addressed in this work is the arbitration of the role between a robot and a human during physical Human-Robot Interaction, sharing a common task. The system is modeled as a Cartesian impedance, with two separate external forces provided by the human and the robot. The problem is then reformulated as a Cooperative Differential Game, which possibly has multiple solutions on the Pareto frontier. Finally, the bargaining problem is addressed by proposing a solution depending on the interaction force, interpreted as the human will to lead or follow. This defines the arbitration law and assigns the role of leader or follower to the robot. Experiments show the feasibility and capabilities of the proposed control in managing the human-robot arbitration during a shared-trajectory following task.**

***Index Terms*— physical Human-Robot Interaction, Role Arbitration, Differential Cooperative Game Theory, Adaptive Control**

I. INTRODUCTION

With the large spread of collaborative robots, collaborative applications involving a human operator and a robot working together to achieve a common goal represents the industry's most recent trend. In contrast to *coexistence* (when the human and the robot are in the same environment but do not interact), *synchronization* (when they work in the same workspace, but at different times) and *cooperation* (when they work in the same workspace at the same time, though each focuses on separate tasks), *collaboration* happens when the human operator and the robot must execute a task together, and the action of the one has immediate consequences on the other [1].

Collaboration requires communication, which typically happens through interaction forces, leading to physical Human-Robot Interaction (pHRI) [2]. As discussed in [3], impedance control is a ubiquitous technique to manage the pHRI. Initially proposed by Hogan [4], impedance control has found an increasing interest due to numerous research works aiming at making it adaptive, in a way such that the robot behavior adapts to the human one. Such adaptation happens mainly in two manners, adaptation of the impedance set-point and adaptation of the mass-spring-damper parameters. To the first category belong [5], and [6], where the damping is made variable since it allows fast/slow motion depending on its value. Examples of the second category can be found in [7] and [8]. Moreover, hybrid controllers exist, which modify both impedance parameters and set-point. This happens in [9], [10], where the updated law is defined by fuzzy logic, in [11], where the parameters are online updated via Model-Based Reinforcement Learning, and in [12], where neural networks (NN) updates the desired position, and the impedance parameters are updated to maintain stability.

In the previous examples, and many other works, the human-robot interaction is described by a static leader-follower paradigm, where the human gives input, and the robot responds consequently, but vice versa never happens. A step ahead is to consider Human-Robot Arbitration, which can be defined as the mechanism that assigns task control to either the human or the robot [13]. Such an issue is addressed, with different approaches, in [14] [15] [16] [17]

As discussed in [18], game theory provides useful tools to analyze complex interactive behaviors involving multiple agents. It provides mathematical models (cooperative, non-cooperative, multi-stages, etc.) of strategic interaction among rational decision-makers and provides the players with "optimal" policies to minimize their objectives, taking into account interaction. Thanks to its capabilities of describing interactions in an "optimal" way, controllers for the pHRI based on game-theoretic formulations are investigated in the literature. In [19], [20], and similarly, in [21] for the shared control with an exoskeleton, the continuous role adaptation is investigated for a Differential Non-Cooperative game. In these works, the concept of Nash Equilibrium is used to update the robot cost function according to the interaction force. The result is a variable impedance control with damping and stiffness updated online. An extension of such approaches, also involving position update, can be found in [22]. The same differential non-cooperative game-theoretic problem is addressed in [23], [24], with a different approach for the solution based on policy iteration. Finally, a general framework for differential game-theoretic modeling of Human-Robot interaction is presented in [25] for the two agents game and extended to multiple agents in [26]. An observer is designed to understand each other's control laws in these works, and different behavior can be addressed.

The previous works show that a game-theoretic description of the human-robot interaction can provide optimal behavior in a non-cooperative framework. Despite this, Nash equilibria (*i.e.* the solutions of the non-cooperative games) are frequently not Pareto optimal. Thus, cooperation can often improve payoffs to all players [27]. Hence, in this paper, Cooperative Game Theory (CGT) is used to describe the

[1]Paolo Franceschi and Nicola Pedrocchi are with the Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing of the National Research Council of Italy (CNR-STIIMA), via Alfonso Corti 12, 20133, Milano, Italy. {*name.surname*}@stiima.cnr.it

[2]Paolo Franceschi and Manuel Beschi are with the Università di Brescia, Dipartimento di Ingegneria Meccanica ed Industriale, via Branze 38, 25123, Brescia, Italy manuel.beschi@unibs.it

interaction model, and Pareto solutions are sought. Cooperation, indeed, can improve the cost of each agent by agreeing. Hence, in a cooperative framework, each agent has an incentive to cooperate since its cost will be lowered compared with the non-cooperative solution. Moreover, a cooperative description of the task allows optimal comanipulation of objects. Finally, as described in this work, Cooperative Game Theory allows for easy role arbitration, keeping the optimal behavior of the robot. To the best of the authors' knowledge, no Cooperative Differential Game-Theoretic frameworks have been used yet to describe HRI. This work aims to show the feasibility of the CGT framework with an application to human-robot arbitration.

II. METHOD

This section presents a system modeled as a Cartesian impedance with two separate external forces provided by the human and the robot. The problem is then reformulated as a Differential Cooperative Game, with solutions lying on the Pareto frontier. The bargaining problem is addressed with a formulation depending on the interaction force to define the arbitration law. Finally, it is shown that such a model results in an adaptive impedance controller with variable stiffness and damping.

A. System modeling

Working in the Cartesian space is more intuitive and natural for the human operator; hence the desired robot motion at the end-effector is implemented as an impedance model in the Cartesian space:

$$M_i\ddot{x}(t)+D_i\dot{x}(t)+K_i(x-x_0)(t)=u_h(t)+u_r(t) \qquad (1)$$

where M_i, D_i and $K_i \in \mathbb{R}^{6\times6}$ are the desired inertia, damping and stiffness matrices, respectively, $\ddot{x}(t)$, $\dot{x}(t)$ and $x(t) \in \mathbb{R}^6$ are the Cartesian accelerations, velocities and positions at the end-effector, x_0 the equilibrium position of the virtual spring, and $u_h(t) \in \mathbb{R}^6$ and $u_r(t) \in \mathbb{R}^6$ represent the human and robot effort applied to the system. The Cartesian coordinates in x are defined according to [28], with the vector $x = [p^T\ \phi^T]^T$ where p^T are the position coordinates and ϕ^T the set of Euler angles that defines the rotation matrix describing the end-effector orientation. With a little abuse of notation, to ease the reading, with $\dot{x} = [\dot{p}^T\ \omega^T]^T$ is indicated the vector containing the linear and angular velocities.

After some reformulation, (1) can be rewritten in a linearized state-space formulation around the working point as

$$\dot{z}=Az+B_hu_h+B_ru_r \qquad (2)$$

where $z = [x-x_0\ \dot{x}]^T \in \mathbb{R}^{12}$ is the state space vector, $A = \begin{bmatrix} 0^{6\times6} & J_a \\ -M_i^{-1}K_i & -M_i^{-1}D_i \end{bmatrix}$, $B_h^{12\times6} = B_r^{12\times6} = \begin{bmatrix} 0^{6\times6} \\ M_i^{-1} \end{bmatrix}$, with $0^{6\times6}$ denoting a 6×6 zero matrix and J_a the analytical Jacobian matrix, with the dimensions of the considered Cartesian components.

Since robot controllers typically accept reference positions or velocities in the joint space as a control input, it is worth converting the reference velocities from the Cartesian space to the joint space.

Given the reference velocity in the Cartesian space, the following relation allows obtaining reference velocities in the joint space

$$\dot{q}_{ref}(t) = J(q)^+\dot{x}(t) \qquad (3)$$

where $\dot{q}_{ref}(t) \in \mathbb{R}^n$, where n represents the number of joints, are the reference velocities in the joint space, $J(q)^+$ is the pseudoinverse of the geometric Jacobian matrix. Simple integration allows commanding joint positions instead of velocities to the robot. Considering that today's robots have excellent tracking performance in the frequency range excitable by the operator, in this work, hypothesis $\dot{q} \simeq \dot{q}_{ref}$ is assumed to hold.

B. Differential Cooperative Game Theoretic modeling

Rewriting the model (1) as (2), allows to include it in a Differential Cooperative Game Theory (DCGT) framework, as described in [29].

The system in (2) can be further rewritten as

$$\dot{z}=Az+Bu \qquad (4)$$

with $A^{12\times12}$ defined as before, $B^{12\times12} = [B_h\ B_r]$ and $u = [u_h\ u_r]^T$.

The human and the robot can be seen as two agents, each one with the objective to minimize its quadratic cost function, defined as

$$J_h = \int_0^\infty (zQ_hz+u_hR_hu_h)\,dt \qquad (5)$$

and

$$J_r = \int_0^\infty (zQ_rz+u_rR_ru_r)\,dt \qquad (6)$$

where J_h and J_r are the cost that the human and the robot incur, $Q_h \in \mathbb{R}^{2n\times2n}$ and $Q_r \in \mathbb{R}^{2n\times2n}$ weights on the state and $R_h \in \mathbb{R}^{6\times6}$ and $R_r \in \mathbb{R}^{6\times6}$ weights on the control input. The human weights Q_h and R_h are unknown and cannot be set as the robot parameters Q_r and R_r. A reasonable estimate of the human parameters, $\hat{Q}_h$ and $\hat{R}_h$, can be obtained with inverse model control similarly to [30], [31]. The problem here presented can be summarized as finding the u_r and u_h of (5) and (6) subject to the system dynamics defined in (4), and game theory provides a framework for finding such solutions.

Given the system in 4, the problem reduces to find the control action u as a composition of feedback and feedforward terms, defined as

$$u = -K^{fb}z+K^{ff}z_{ref} \qquad (7)$$

where the apices ^{fb} and ^{ff} denote feedback and feedforward gain matrices, respectively.

While in Non-cooperative Game Theory, the solutions are represented by the Nash equilibria, in Linear Quadratic Cooperative Game Theory, each feasible solution lies on the Pareto frontier (see [29] for definitions).

The cost to be minimized from the two players together is defined as

$$J_\alpha = \alpha J_h + (1-\alpha) J_r = \int_0^\infty (z Q_\alpha z + u R_\alpha u)\, dt \quad (8)$$

where $Q_\alpha = \alpha \hat{Q}_h + (1-\alpha) Q_r$ and $R_\alpha = \begin{bmatrix} \alpha \hat{R}_h & 0^{6\times 6} \\ 0^{6\times 6} & (1-\alpha) R_r \end{bmatrix}$, and $\alpha \in (0,1)$ represents the weight each player's cost has in the overall cost.

To find all cooperative solutions for the feedback linear-quadratic game, one has to solve a regular Linear Quadratic optimal control problem that depends on a parameter α. The solution is given by the matrix P, solution of the infinite horizon Continuous Algebraic Riccati Equation (CARE), given by

$$A^T P + PA - PBR_\alpha^{-1} B^T P + Q_\alpha = 0 \quad (9)$$

The feedback gain matrix is defined as

$$K^{fb} = R_\alpha^{-1} B^T P \quad (10)$$

and the feedback control actions are given by

$$u^{fb} = -K^{fb} z(t) \quad (11)$$

From (11) the feedback control terms demanded from the human and the robot can be computed as

$$u^{fb} = \begin{bmatrix} \bar{u}_h^{fb} \\ u_r^{fb} \end{bmatrix} \quad (12)$$

where $\bar{u}^{fb}$ denotes the nominal effort demanded to the human. A feedforward term is added to allow trajectory tracking.

The feedforward gain matrix is defined as

$$K^{ff} = [K^{fb}\ I] \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ I \end{bmatrix} \quad (13)$$

with A and B as in (4), and C and D output and feedthrough matrices of the state-space system description, respectively. The feedforward term results in

$$u^{ff} = K^{ff} z_{ref}. \quad (14)$$

Since the system matrix is typically not square in this formulation, the pseudoinverse can be used instead of the inverse. This results in a vector of dimension 12, and its components are

$$u^{ff} = \begin{bmatrix} \bar{u}_h^{ff} \\ u_r^{ff} \end{bmatrix} \quad (15)$$

with the superscripts h and r denoting the human and robot contributions.

Finally, the total control input results in $u = u^{ff} + u^{fb}$, and the control action of the robot can be computed as

$$u_r = u_r^{ff} + u_r^{fb} \quad (16)$$

C. Bargaining

In a cooperative environment, it is rational to consider the set of Pareto solutions. Since there are infinity Pareto-optimal solutions, we enter the bargaining theory arena to decide which is the most effective one. The bargaining problem is how to define the appropriate α. In classical bargaining solutions (Nash, Kalay-Smorodinsky, egalitarian), α is defined such that the cost of all players decreases compared to their non-cooperative cost.

In this work, the two players bargain on who leads and follows the task; in this sense, one can accept a higher cost if he is the follower. Hence, α is used as a weighting factor to move the control authority from the robot to the human and vice-versa. Indeed, for high values of α, the robot cost tends to disappear from the overall cost computation J_α. Hence it will be less costly for the robot to put much effort into the system, becoming the leader, resulting in a higher cost for the human. On the contrary, for low values of α, the robot cost increases, and its control input will be dramatically reduced, leading the control authority to the human. Each situation in between represents a cooperative solution where the control authority is shared appropriately.

The selection of the weight parameter α depends on the force applied by the human and is processed by the sigmoid function:

$$\alpha = d - \frac{a}{1 + e^{-b(\|u_h\| - c)}} \quad (17)$$

where the constant parameters a, b, c, d are used to shape the function properly. In particular a defines the height of the function, b the width of the transition phase, c represents an offset in the x direction and d is an offset in the y direction, moreover the negative sign after d means that α is decreasing as $\|u_h\|$ increases.

D. The control law as a variable impedance

Looking at (12), can be divided into two components and the control inputs demanded from the human and the robot are computed as

$$\begin{bmatrix} \bar{u}_h^{fb} \\ u_r^{fb} \end{bmatrix} = - \begin{bmatrix} K_h^{fb} \\ K_r^{fb} \end{bmatrix} z(t) \quad (18)$$

where $K_h \in \mathbb{R}^{6\times m}$ and $K_r \in \mathbb{R}^{6\times m}$ are the components of the matrix K defined in (10). Looking at the robot control input, the feedback part can be rewritten as

$$u_r^{fb} = -K_r^{fb} \begin{bmatrix} x - x_0 \\ \dot{x} \end{bmatrix} \quad (19)$$

Given $K_r^{fb} = [K_{r,k}^{fb}\ K_{r,d}^{fb}]$, the two components associated with the stiffness and damping of the variable impedance cooperative system are defined as

$$K_{imp}^{fb} = K_{r,k}^{fb} \quad (20)$$

and

$$D_{imp}^{fb} = K_{r,d}^{fb} \quad (21)$$

A similar computation can be done for the feedforward terms, resulting in

$$u_r^{ff} = K_r^{ff} x_{ref} \tag{22}$$

where K_r^{ff} represents the components relative to the robot of the matrix (13) and

$$K_{imp}^{ff} = K_r^{ff} \tag{23}$$

The robot control input results in

$$u_r = -D_{imp}^{fb}\dot{x} - K_{imp}^{fb}(x - x_0) + K_{imp}^{ff} x_{ref} \tag{24}$$

Substituting (24) into (1) results in

$$\begin{aligned} M_i\ddot{x}(t) + (D_i + D_{imp}^{fb})\dot{x}(t) + \\ +(K_i + K_{imp}^{fb})\,(x(t) - x_0(t)) - K_{imp}^{ff} x_{ref}(t) = u_h(t) \end{aligned} \tag{25}$$

Because varying α varies matrix P and matrix K accordingly, (25) represents a variable impedance system subject to the human force, with the values of D_{imp} and K_{imp} updated according to the human will to lead or follow, detected by force applied.

III. Experiments

An experiment is designed to test the proposed control method for sharing control authority in human-robot collaboration. The robot has to follow a planar circular trajectory, while the human has a different path to follow, which partially overlaps with the robot one. The nominal robot trajectory is defined as

$$x_{ref,r}(t) = \begin{bmatrix} -\rho\, sin(\omega t) \\ -\rho\, cos(\omega t) \end{bmatrix} \tag{26}$$

where $\rho = 0.2$ is the radius of the circumference and ω is the angular velocity.

The human desired trajectory is piecewise-defined as

$$x_{ref,h}(t) = \begin{cases} x_{ref,r}(t), & t_0 < t < t_1 \\ \begin{bmatrix} -\rho\, sin(\omega t) \\ x_1 + \frac{x_2 - x_1}{t_2 - t_1}(t - t_0) \end{bmatrix}, & t_1 < t < t_2 \\ \begin{bmatrix} -\rho\, sin(\omega t) \\ x_2 + \frac{x_3 - x_2}{t_3 - t_2}(t - t_1) \end{bmatrix}, & t_2 < t < t_3 \\ x_{ref,r}(t), & t_3 < t \end{cases} \tag{27}$$

with $t_0 = 0s, t_1 = 3s, t_2 = 5s, t_3 = 7s$.

The nominal robot trajectory, the desired human trajectory, and the current position are shown in real-time on a screen, and the human has the goal to make the current trajectory as close as possible to the human desired trajectory. Figure 1 shows the experimental setup.

In Figure 2 the nominal trajectory of the robot (dashed red line), the desired trajectory of the human (dashed yellow line), and the actual robot end-effector positions (solid blue line) are shown.

To test and compare the proposed approach, the following indices are computed:

- Trajectory following error, measured as

$$\mathscr{E}_{track} = \int_{T_{start}}^{T_{end}} \left\| x_{ref,h}(t) - x(t) \right\| dt \tag{28}$$

  As low the $\mathscr{E}_{track}$ is, as close the actual trajectory is to the nominal one
- Geometrical following error, measured as

$$\mathscr{E}_{Xcorr} = \int_{T_{start}}^{T_{end}} \left\| x_{ref,h}(t + \delta t) - x(t) \right\| dt \tag{29}$$

  where δt represents the time delay of the actual trajectory with respect to the nominal one, computed by applying cross-correlation between the nominal trajectory of the human and the measured one. In this way, it is possible to compare the capability of the human to track the geometrical reference. As low the $\mathscr{E}_{Xcorr}$ is, as close the actual trajectory is to the geometrical one
- Measured interaction force, measured as

$$\mathscr{F} = \int_{T_{start}}^{T_{end}} \| f(t) \| \, dt \tag{30}$$

  As low the $\mathscr{F}$ is, as less effort the human has to put in the cooperative task
- Mechanical work, measured as

$$\mathscr{W} = \int_{T_{start}}^{T_{end}} \vec{f}(t) \cdot d\vec{S}\, dt \tag{31}$$

  The lower the $\mathscr{W}$ is, the less energy the human has to put into the cooperative task

The goal of the proposed controller is to allow smooth interaction and leader-follower transition with the human, reducing the effort required and allowing good trajectory tracking. Hence the goal is to minimize the indexes defined above together.

As a comparison, three different types of impedance controllers are used. The CGT controller is compared with an LQR controller, which can be seen as the CGT with $\alpha = 1$ always, and two classical impedance controllers with different values of stiffness K. One has high stiffness (HIC), and the other has a low value of stiffness (LIC). For the

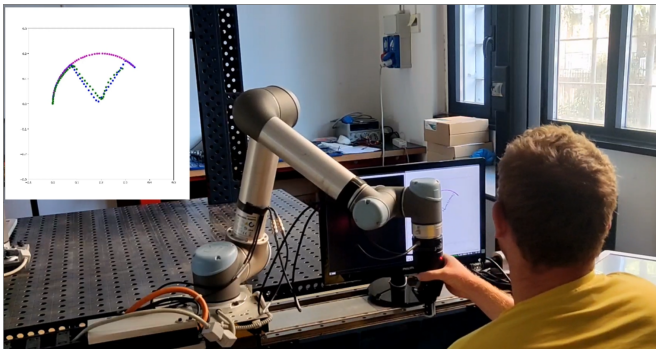


Fig. 1: The experimental setup, showing the application and the monitor displaying the three trajectories: magenta - robot desired trajectory, blue - human desired trajectory, green - actual trajectory.

two classical impedance controllers (HIC, LIC), the value of the damping to achieve the critical damping $D = 2D_{cr}\sqrt{KM}$. For the LQR and CGT controllers, the values of damping and stiffness depend on the matrices $\hat{Q}_h$, $\hat{R}_h$, Q_r and R_r, defined as $\hat{Q}_h = diag(50,50,1,1)$, $\hat{R}_h = diag(10,10,0,0)$, $Q_r = diag(20,20,1,1)$ and $R_r = diag(1,1,0.01,0.01)$ for the CGT case, $R_r = diag(1,1,0.1,0.1)$ for the LQR case, otherwise the stiffness was too high and too much force was required to barely move the robot. The $\hat{Q}_h$ and $\hat{R}_h$ used are average values. The mass matrix for all the cases is defined as $M = diag(10,10)$, the base damping and stiffness for the CGT and LQR cases are $K = diag(0,0)$ and $D = diag(40,40)$. The mass-spring-damper parameters used or computed for the experiments are presented in Table I, while Figure 3 shows the changing parameters during the task. For the CGT case, the value of α is bounded such that $0.01 \leq \alpha \leq 0.99$, and the sigmoid parameters are chosen to be $a = 0.98$, $b = 0.7$, $c = 7$, $d = 0.99$.

TABLE I: The mass, spring and damping parameters used for the experiments.

	CGT	HIC	LIC	LQR
K	55 ÷ 550	100	20	222
D	53 ÷ 135	57	25.5	84
D/D_{cr}	1.13 ÷ 0.92	0.9	0.9	0.89
M	10	10	10	10

IV. Results

In the test campaign, five subjects (age 30 ± 1) are asked to perform the task five times for each controller. Before starting the test, the subjects are allowed to practice as long as needed to feel confident with the current controller, then five trials in a row with the same controller are recorded. This procedure is used for each of the four controllers, selected randomly for each subject.

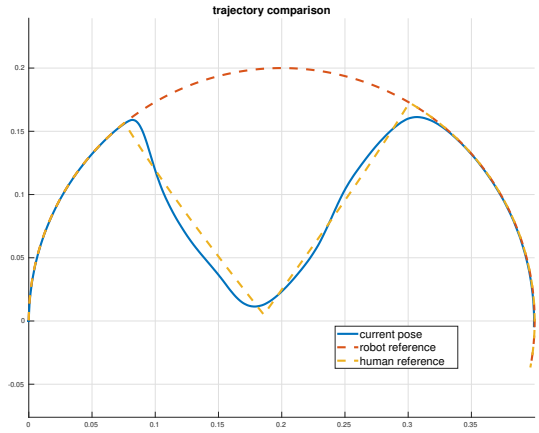


Fig. 2: Comparison between the nominal trajectory of the robot, the nominal trajectory of the human and the measured

A UR5 robot is used, controlled by joint velocities. (2) is used to compute the Cartesian reference positions, and (3) provides the robot joint velocity tracking controller with the reference values. A Robotiq FT 300, mounted on the robot tip, is used to measure the force applied by the human.

A t-test compares the CGT controller in pairs with all the others for each of the computed indexes. The t-test allows defining the statistical relevance of the computed values. The results are in Figure 4, with the mean and standard deviations, as well as the relative p-values.

The tracking error, computed as in (28), is shown in Figure 4a. The CGT controller shows the best performance, and the t-test assesses the statistical relevance of the measured values. This result can be explained by the robot's high stiffness/low damping ratio while the robot is leading and the low stiffness/high damping ratio while the human is leading. On the one hand, the high stiffness/low damping ratio allows fast and reliable trajectory tracking, as well as fast recovery after a trajectory modification (*i.e.* the robot is faster in getting back to its nominal trajectory after the human leaves control). On the other hand, a low stiffness/high damping ratio allows the human to move the robot smoothly and precisely while leading. Compared to the LIC, constant high stiffness values allow better trajectory tracking in HIC and LQR.

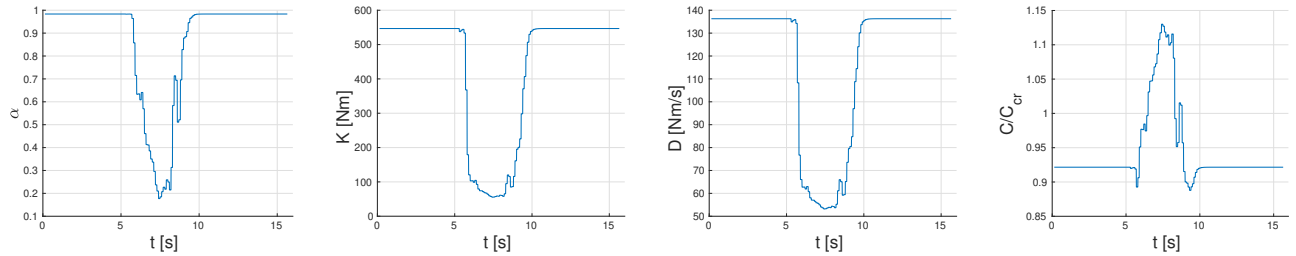
The geometrical error, computed as in (29) is shown in 4b. The t-test shows that no significant differences between the four controllers can be appreciated. This shows that the tasks are executed correctly, and the comparison of the controllers for the computed indexes is fair. Indeed, this shows that similar results can be obtained in terms of geometrical path following with all the controllers.

The force evaluation is presented in 4c. Since the exchanged force mainly depends on the robot stiffness (measured as $f = K(x_{ref,r} - x)$), the lower the stiffness is, the lower the exchanged force. This effect results in the LIC with the best performance concerning the required force. Despite this, the CGT controller shows better performance with respect to LQR and HIC because when interaction happens, its stiffness lowers. Notably, performances similar to the LIC can be obtained by a different tuning of the values of matrices Q_r, R_r.

Finally, similar considerations can be derived for the mechanical work shown in Figure 4d. Indeed, depending on the interaction force, similar behavior is observed with respect to the force results.

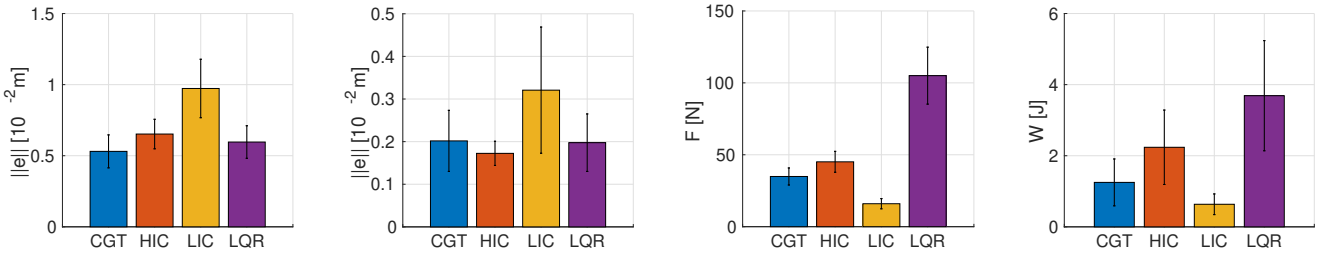
V. Discussion

The results show that being equal to the capability of the four controllers in following the nominal geometrical path under human-robot cooperation, the CGT controller outperforms the others in trajectory following. Moreover, even if the LIC requires less interaction force to accomplish the task, it shows the lowest capability in trajectory tracking, and the CGT also requires low interaction forces. As a remark, a similar interaction force can be obtained for the CGT controller by tuning the parameters differently.



(a) Arbitration parameter α (b) Variable stiffness (c) Variable damping (d) Variable critical damping

Fig. 3: Variable parameters for the CGT



(a) Tracking error
p = 0.0081, <0.0001, 0.1316

(b) X-correlated error
p = 0.2204, 0.0261, 0.8767

(c) Force
p = <0.001, <0.0001, <0.0001

(d) Mechanical work
p = 0.0093, 0.0093, <0.0001

Fig. 4: Indexes comparison

Interestingly, from Table I and Figure 3d, it can be noted that the critical damping value varies from lower than critical to higher than critical, according to the role adaptation. This effect allows the robot to have a more damped behavior when the human is leading, allowing smooth motion, while a low damping ratio allows better and faster trajectory tracking when the robot is leading.

Some limitations of the proposed control can be identified. First, the human control objective (*i.e.* the human cost function parameters) can only be estimated offline with some inverse techniques. Future works will address online estimation techniques for the human control objective estimation to adapt the robot behavior accordingly. Indeed, it is reasonable to suppose that the human may change his cost function parameters due to fatigue by repeating a task all day long. Secondly, many parameters have to be defined arbitrarily. In particular, it is difficult to predict the variable damping and stiffness values. Hence, it may be interesting to define some tuning rules according to the required performances. Moreover, the arbitration law introduced in this work is a simple function of the force, but it can be improved and modified according to other sensing capabilities and different human intention estimations. Future work will address a more complex and improved arbitration law involving visual feedback and human intention estimation, along with force feedback. In such a way, with a more complex sensor-fusion, a more precise estimate of the human will to lead or follow can be obtained. Some future applications of this framework will involve the co-manipulation and the co-transportation of heavy objects, as well as lightweight parts and deformable objects as composite materials plies. Another interesting application can be for rehabilitation purposes, where the robot behavior is adjusted by tuning the parameter α, to make the robot contribution high when the patient is barely able to perform a task and gradually lower the robot contribution as the patient starts to recover some motion capabilities.

VI. Conclusion

This work proposes a framework for physical Human-Robot interaction based on a Cooperative Game theory formulation. The bargaining problem is addressed with the law definition for the leader-follower role arbitration. The results show the capability of the proposed method in managing the leader-follower transition continuously and show that with the proposed controller, high performance in tracking a trajectory, different from the nominal one of the robot, can be achieved through role arbitration.

In conclusion, the proposed framework shows good capabilities in describing a human-robot cooperative task. It allows easy integration of different modules (human cost function identification, human intention identification) and easy implementation of different robot behaviors (leader-follower adaptation, co-manipulation, co-transportation), which will be investigated in future works. Future works will consider different solutions to the bargaining problem to provide the robot with the proper control input, possibly involving constraints on the maximum allowed input.

ACKNOWLEDGMENT

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101006732 (Drapebot).

REFERENCES

[1] E. Matheson, R. Minto, E. G. G. Zampieri, M. Faccio, and G. Rosati, "Human–robot collaboration in manufacturing applications: A review," *Robotics*, vol. 8, no. 4, 2019. [Online]. Available: https://www.mdpi.com/2218-6581/8/4/100

[2] S. Haddadin and E. Croft, *Physical Human–Robot Interaction*. Cham: Springer International Publishing, 2016, pp. 1835–1874.

[3] A. De Santis, B. Siciliano, A. De Luca, and A. Bicchi, "An atlas of physical human–robot interaction," *Mechanism and Machine Theory*, vol. 43, no. 3, pp. 253–270, 2008.

[4] N. Hogan, "Impedance Control: An Approach to Manipulation: Part I—Theory," *Journal of Dynamic Systems, Measurement, and Control*, vol. 107, no. 1, pp. 1–7, 03 1985.

[5] J. Dong, J. Xu, Q. Zhou, and S. Hu, "Physical human–robot interaction force control method based on adaptive variable impedance," *Journal of the Franklin Institute*, vol. 357, no. 12, pp. 7864–7878, 2020.

[6] F. Ficuciello, L. Villani, and B. Siciliano, "Variable impedance control of redundant manipulators for intuitive human–robot physical interaction," *IEEE Transactions on Robotics*, vol. 31, no. 4, pp. 850–863, 2015.

[7] D. P. Losey and M. K. O'Malley, "Trajectory deformations from physical human–robot interaction," *IEEE Transactions on Robotics*, vol. 34, no. 1, pp. 126–138, 2018.

[8] S. S. Ge, Y. Li, and H. He, "Neural-network-based human intention estimation for physical human-robot interaction," in *2011 8th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, 2011, pp. 390–395.

[9] L. Roveda, S. Haghshenas, M. Caimmi, N. Pedrocchi, and L. Molinari Tosatti, "Assisting operators in heavy industrial tasks: On the design of an optimized cooperative impedance fuzzy-controller with embedded safety rules," *Frontiers in Robotics and AI*, vol. 6, p. 75, 2019.

[10] L. Roveda, S. Haghshenas, A. Prini, T. Dinon, N. Pedrocchi, F. Braghin, and L. M. Tosatti, "Fuzzy impedance control for enhancing capabilities of humans in onerous tasks execution," in *2018 15th International Conference on Ubiquitous Robots (UR)*, 2018, pp. 406–411.

[11] L. Roveda, J. Maskani, P. Franceschi, A. Abdi, F. Braghin, L. Molinari Tosatti, and N. Pedrocchi, "Model-based reinforcement learning variable impedance control for human-robot collaboration," *Journal of Intelligent & Robotic Systems*, vol. 100, no. 2, pp. 417–433, Nov 2020.

[12] Y. Li and S. S. Ge, "Human–robot collaboration based on motion intention estimation," *IEEE/ASME Transactions on Mechatronics*, vol. 19, no. 3, pp. 1007–1014, 2014.

[13] 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, no. 1, 02 2018, 010804.

[14] L. Peternel, N. Tsagarakis, D. Caldwell, and A. Ajoudani, "Robot adaptation to human physical fatigue in human–robot co-manipulation," *Autonomous Robots*, vol. 42, no. 5, pp. 1011–1021, Jun 2018.

[15] R. J. Ansari and Y. Karayiannidis, "Task-based role adaptation for human-robot cooperative object handling," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3592–3598, 2021.

[16] B. Nemec, N. Likar, A. Gams, and A. Ude, "Human robot cooperation with compliance adaptation along the motion trajectory," *Autonomous Robots*, vol. 42, no. 5, pp. 1023–1035, Jun 2018.

[17] X. Li, Y. Pan, G. Chen, and H. Yu, "Adaptive human–robot interaction control for robots driven by series elastic actuators," *IEEE Transactions on Robotics*, vol. 33, no. 1, pp. 169–182, 2017.

[18] N. Jarrassé, T. Charalambous, and E. Burdet, "A framework to describe, analyze and generate interactive motor behaviors," *PLOS ONE*, vol. 7, no. 11, pp. 1–13, 11 2012.

[19] Y. Li, K. P. Tee, W. L. Chan, R. Yan, Y. Chua, and D. K. Limbu, "Continuous role adaptation for human–robot shared control," *IEEE Transactions on Robotics*, vol. 31, no. 3, pp. 672–681, 2015.

[20] Y. Li, K. P. Tee, W. L. Chan, R. Yan, Y. Chua, and D. Limbu, "Role adaptation of human and robot in collaborative tasks," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 5602–5607.

[21] W. Bi, X. Wu, Y. Liu, and Z. Li, "Role adaptation and force, impedance learning for physical human-robot interaction," in *2019 IEEE 4th International Conference on Advanced Robotics and Mechatronics (ICARM)*, 2019, pp. 111–117.

[22] Y. Li, C. Yang, and W. He, "Towards coordination in human-robot interaction by adaptation of robot's cost function," in *2016 International Conference on Advanced Robotics and Mechatronics (ICARM)*, 2016, pp. 254–259.

[23] Y. Li, K. P. Tee, R. Yan, W. L. Chan, Y. Wu, and D. K. Limbu, "Adaptive optimal control for coordination in physical human-robot interaction," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 20–25.

[24] Y. Li, K. P. Tee, R. Yan, W. L. Chan, and Y. Wu, "A framework of human–robot coordination based on game theory and policy iteration," *IEEE Transactions on Robotics*, vol. 32, no. 6, pp. 1408–1418, 2016.

[25] 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, no. 1, pp. 36–43, Jan 2019.

[26] R. Zou, Y. Liu, J. Zhao, and H. Cai, "A framework for human-robot-human physical interaction based on n-player game theory," *Sensors*, vol. 20, no. 17, 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/17/5005

[27] A. Seierstad, "Pareto improvements of nash equilibria in differential games," *Dynamic Games and Applications*, vol. 4, 01 2011.

[28] B. Siciliano and L. Villani, "Robot force control," 2000.

[29] J. Engwerda, *LQ dynamic optimization and differential games*. John Wiley & Sons, 2005.

[30] S. Rothfuß, J. Inga, F. Köpf, M. Flad, and S. Hohmann, "Inverse optimal control for identification in non-cooperative differential games," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 14 909–14 915, 2017, 20th IFAC World Congress. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405896317334602

[31] F. Köpf, J. Inga, S. Rothfuß, M. Flad, and S. Hohmann, "Inverse reinforcement learning for identification in linear-quadratic dynamic games," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 14 902–14 908, 2017, 20th IFAC World Congress.