Letter

A Framework for Human-Robot-Human Physical Interaction Based on N-Player Game Theory

Rui Zou [†] [iD]**, Yubin Liu *, Jie Zhao and Hegao Cai**

State Key Laboratory of Robotics and Systems, Harbin Institute of Technology, Harbin 150001, China; 18b908014@stu.hit.edu.cn (R.Z.); jzhao@hit.edu.cn (J.Z.); hgcai@hit.edu.cn (H.C.)

* Correspondence: liuyubin@hit.edu.cn

† Current address: 92 Xidazhi Street, Nangang District, State Key Laboratory of Robotics and Systems, Harbin Institute of Technology, Harbin 150001, China.

check for updates

Abstract: In order to analyze the complex interactive behaviors between the robot and two humans, this paper presents an adaptive optimal control framework for human-robot-human physical interaction. N-player linear quadratic differential game theory is used to describe the system under study. N-player differential game theory can not be used directly in actual scenerie, since the robot cannot know humans' control objectives in advance. In order to let the robot know humans' control objectives, the paper presents an online estimation method to identify unknown humans' control objectives based on the recursive least squares algorithm. The Nash equilibrium solution of human-robot-human interaction is obtained by solving the coupled Riccati equation. Adaptive optimal control can be achieved during the human-robot-human physical interaction. The effectiveness of the proposed method is demonstrated by rigorous theoretical analysis and simulations. The simulation results show that the proposed controller can achieve adaptive optimal control during the interaction between the robot and two humans. Compared with the LQR controller, the proposed controller has more superior performance.

Keywords: physical human-robot interaction; game theory; adaptive optimal control; robot control

1. Introduction

In the past decade, physical human-robot interaction has attracted the attention of the research community due to the urgent requirement for robot technology in unstructured environment [1–4]. Physical human-robot interaction combines the advantages of humans and robots, which means that humans are good at reasoning and problem solving with high flexibility, while robots perform well in terms of execution as well as guaranteeing the accuracy of task execution [5,6]. The combination of these advantages has led to the wide application of physical human-robot interaction, such as teleoperation [7,8], collaborative assembly [9,10], and collaborative transportation [11–13].

Two types of specific human robot interaction strategies have been widely studied: co-activity type of interaction strategy and master-slave control strategy [14,15]. Co-activity type of interaction strategy is used in typical rehabilitation robots that help limb movement training or intelligent industrial systems that support heavy objects to resist gravity, where robots completely ignore human users' behaviors [16,17]. In contrast, the master–slave control strategy is used in the teleoperated robots or force extender exoskeletons use where robots completely follow the control of human users [18]. However, these strategies can only be used for specific interactive behaviors, the general framework for analyzing various interactive behaviors between robot and humans is still missing [19,20].

It has been pointed out that game theory can be used as a general framework to analyze complex interactive behaviors between multiple agents because different combinations of individual cost

functions and different optimization objectives can be used to describe various interactive behaviors in game theory [21]. In [22], the human and the robot were been regarded as two agents and game theory was used in order to analyze the performance of the two agents. In [23], the optimal control was obtained for a given game with a linear system cost function by solving the coupled Riccati equation. In [24], an optimal control algorithm was developed for human-robot collaboration by solving the Riccati equation in each loop. In [25–28], policy iteration was used to solve the Nash equilibrium solution in order to improve the calculation speed. In [29], cyber-physical human systems was modeled via an interplay between reinforcement learning and game theory. In [30], haptic shared control for human-robot collaboration was modeled by a game-theoretical approach. In [31], human-like motion planning was studied based on game theoretic decision making. In [32], cooperative game was used for human-robot collaborative manufacturing. In [33], a bayesian framework was proposed for nash equilibrium inference in human-robot parallel play. In [19], non-cooperative differential game theory was used to model human-robot interaction system that results in a variety of interaction strategies. However, the above studies only consider two agents, that is, the interaction between one human and one robot. Therefore, the aforementioned methods are not suitable for human-robot-human physical interaction where more than one human interact with one robot physically. It is worth noting that the physical interaction between one robot and two humans will bring greater advantages such as operating larger loads, improving the flexibility and robustness of the system [28,34–37]. These greater advantages are brought by the team collaboration between the robot and two humans. To the authors' acknowledgment, no literature have researched the problem of the physical interaction between one robot and two humans based on game theory.

In the paper, a general adaptive optimal control framework for human-robot-human physical interaction is proposed based on N-player game theory. Accordingly, the robot and two humans can interact with each other optimally by learning each other's control. N-player differential game theory was used to model the human-robot-human interaction system in order to analyze the complex interactive behaviors between the robot and two humans. In N-player differential game theory, humans' control objectives are assumed to be knowledge [38,39]. However, N-player differential game theory can not be used directly in actual scenerie since the robot cannot know humans' control objectives in advance. In order to let the robot know humans' control objectives, the paper presents an online estimation method to identify unknown humans' control objectives based on the recursive least squares algorithm. Subsequently, the Nash equilibrium solution of the multi-human robot physical interaction is obtained by solving the coupled Riccati equation to achieve coupled optimization. Finally, the effectiveness of the proposed method is demonstrated by rigorous theoretical analysis and simulation experiments. This paper makes the following four contributions.

(1) N-player differential game theory is firstly used to model the human-robot-human interaction system.
(2) An online estimation method to identify unknown humans' control objectives based on the recursive least squares algorithm is presented.
(3) A general adaptive optimal control framework for human-robot-human physical interaction is propose based on (1) and (2).
(4) The effectiveness of the proposed method is demonstrated by rigorous theoretical analysis and simulation experiments.

The remainder of this paper is organized, as follows: Section 2 models the human-robot-human physical interaction system based on N-player differential game theory. Section 3 establishes an adaptive optimal control law, and the control performance of the system is analyzed theoretically. Section 4 verifies the effectiveness of the proposed method through simulation experiments. Finally, Section 5 concludes this work.

2. Problem Formulation

2.1. System Description

The system considered contains two humans and one robot. An example scenario is shown in Figure 1, where the robot and the humans collaborate to perform an object transporting task. In this shared control task, when the control objectives of humans' change, the robot should recognize the humans' control objectives and response adaptively and optimally. The forces exerted by the humans on the object are measured by force sensors at the interaction point. It is worth noting that the humans' control objectives are unknown to the robot.
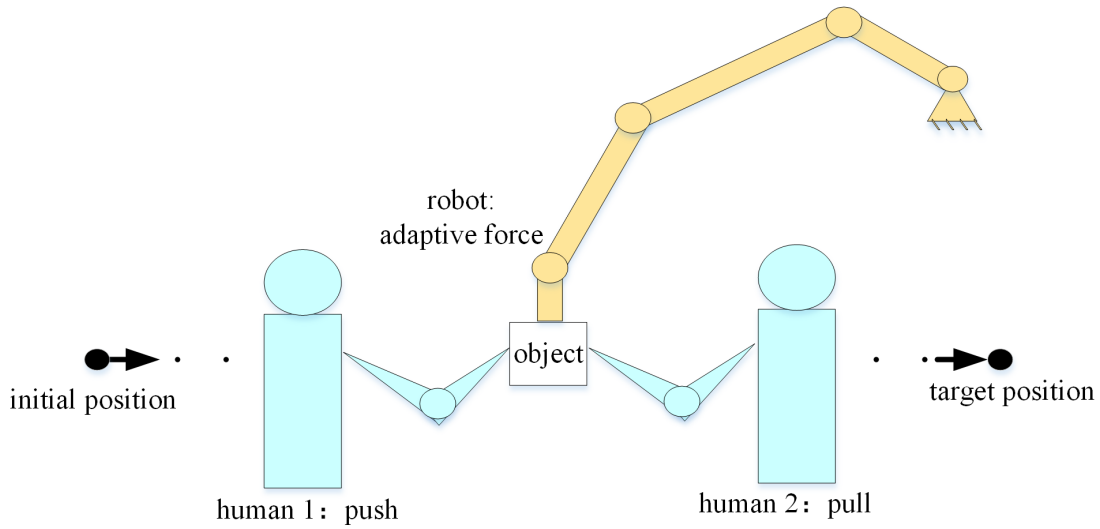

Figure 1. A scenario where the humans and the robot collaborate to perform an object transporting task.

The forward kinematics of the robot are described as

$$x(t) = \phi(q(t)) \tag{1}$$

where $x(t) \in \mathbb{R}^m$ and $q(t) \in \mathbb{R}^n$ are the positions in Cartesian space and joint space respectively, m and n are degrees of freedom. Derivation of Equation (1) with time can be obtained

$$\dot{x}(t) = J(q(t))\dot{q}(t) \tag{2}$$

where $J(q(t)) \in \mathbb{R}^{m \times n}$ is the Jacobian matrix.

The following impedance model is given in Cartesian space

$$M_d \ddot{x}(t) + C_d \dot{x}(t) = u(t) + f_1(t) + f_2(t) \tag{3}$$

where $M_d \in \mathbb{R}^{m \times m}$ is the desired inertial matrix, $C_d \in \mathbb{R}^{m \times m}$ is the damping matrix, $u(t) \in \mathbb{R}^m$ is the control input in the Cartesian space [40–42], $f_1(t) \in \mathbb{R}^n$ is the contact force between object and human 1, $f_2(t) \in \mathbb{R}^n$ is the contact force between object and human 2.

To track a common and fixed target $x_d \in \mathbb{R}^m$ ($\dot{x}_d \in \mathbb{R}^m$) in cooperative object transporting task, Equation (3) can be transformed, as following

$$M_d(\ddot{x}(t) - \ddot{x}_d) + C_d(\dot{x}(t) - \dot{x}_d) = u(t) + f_1(t) + f_2(t). \tag{4}$$

In order to ease the design of the control, Equation (4) can be rewritten as the following state-space form

$$\dot{z} = Az + B_1 u + B_2 f_1 + B_3 f_2$$

$$z = \begin{bmatrix} x(t) - x_d \\ \dot{x}(t) \end{bmatrix}, A = \begin{bmatrix} 0_m & 1_m \\ 0_m & -M_d^{-1}C_d \end{bmatrix}$$

$$B_1 = B_2 = B_3 = B = \begin{bmatrix} 0_m \\ M_d^{-1} \end{bmatrix}$$

(5)

where 0_m and 1_m denote $m \times m$ zero and unit matrices, respectively.

2.2. Problem Formulation

According to non-cooperative differential game theory, in the paper, the interaction between the robot and the humans is described as a game between N players (in this paper, $N = 3$) [43]. In the game, each player will minimize their respective cost function

$$\Gamma \equiv \int_{t_0}^{\infty} z^T Q z + u^T u d_t$$

$$\Gamma_1 \equiv \int_{t_0}^{\infty} z^T Q_1 z + f_1^T f_1 d_t$$

$$\Gamma_2 \equiv \int_{t_0}^{\infty} z^T Q_2 z + f_2^T f_2 d_t$$

$$Q = \begin{bmatrix} Q_{01} & 0_{n \times n} \\ 0_{n \times n} & Q_{02} \end{bmatrix}$$

(6)

$$Q_1 = \begin{bmatrix} Q_{11} & 0_{n \times n} \\ 0_{n \times n} & Q_{12} \end{bmatrix}$$

$$Q_2 = \begin{bmatrix} Q_{21} & 0_{n \times n} \\ 0_{n \times n} & Q_{22} \end{bmatrix}$$

where $\Gamma, \Gamma_1, \Gamma_2$ are cost functions of the robot, human 1, and human 2, respectively, Q, Q_1, Q_2 are state weights matrices of the robot, human 1 and human 2, respectively. Each player achieves the cooperative object transporting task by minimizing the error to the target while minimizing their own costs. Q, Q_1, Q_2 contain two components corresponding to position regulation and velocity, respectively. Q_{01}, Q_{11}, Q_{21} correspond to position regulation and Q_{02}, Q_{12}, Q_{22} correspond to velocity.

In [27], the N-player game has been studied if the cost functions are known. However, Γ_1, Γ_2 are unknown to the robot because they are determined by the humans. Therefore, a method is proposed in the paper to estimate Γ_1, Γ_2 in order to achieve adaptive optimal control and, thus, the human-robot-human cooperative object transporting task.

2.3. N-Player Differential Game Theory

Based on the differential game theory of linear systems, for *N*-player game the following linear differential equation [43] is considered:

$$\dot{z} = Az + B_1 u_1 + \cdots + B_N u_N, z(0) = z_0.$$

(7)

Each player has a quadratic cost function that they want to minimize:

$$\Gamma_i = \int_0^{\infty} z^T Q_i z + u_i^T u_i d_t, i = 1, \cdots, N$$

(8)

Different types of multi-agent behaviors are defined in game theory, which can be achieved through different concepts of game equilibrium [44,45]. In this paper, Nash equilibrium is considered. In the sense of Nash equilibrium, each player minimizes their cost function:

$$u_i = -\eta_i z, \eta_i = B_i^T P_i$$

$$(A - \sum_{j\neq i}^{N} B_i \eta_i)^T P_i + P_i (A - \sum_{j\neq i}^{N} B_i \eta_i)_i + Q_i - P_i B_i B_i^T P = 0, i = 1, \cdots, N \tag{9}$$

where N is equal to 3 in this paper. In the sense of Nash equilibrium, the humans and the robot minimizes their own cost function:

$$\begin{aligned} u &= -\alpha z \\ \alpha &= B^T P_r \end{aligned} \tag{10a}$$

$$\begin{aligned} f_1 &= -\beta z \\ \beta &= B^T P_1 \end{aligned} \tag{10b}$$

$$\begin{aligned} f_2 &= -\gamma z \\ \gamma &= B^T P_1 \end{aligned} \tag{10c}$$

$$A_r^T P_r + P_r A_r + Q - P_r BB^T P_r = 0_{2n}, A_r = A - B\beta - B\gamma \tag{10d}$$

$$A_1^T P_1 + P_1 A_1 + Q - P_1 BB^T P_1 = 0_{2n}, A_1 = A - B\alpha - B\gamma \tag{10e}$$

$$A_2^T P_2 + P_2 A_2 + Q - P_2 BB^T P_2 = 0_{2n}, A_2 = A - B\alpha - B\beta \tag{10f}$$

where $\alpha \equiv \begin{bmatrix} \alpha_e, \alpha_v \end{bmatrix}$ is the feedback gain of the robot, $\beta \equiv \begin{bmatrix} \beta_e, \beta_v \end{bmatrix}$ is the feedback gain of the human 1, $\gamma \equiv \begin{bmatrix} \gamma_e, \gamma_v \end{bmatrix}$ is the feedback gain of of the human 2. $\alpha_e, \beta_e, \gamma_e$ are the position error gains, $\alpha_v, \beta_v, \gamma_v$ are the velocity gains, P_r, P_1, P_2 are the solutions of the above well-known Riccati equation consisting of Equation (10d–f). The robot and the humans influence each other through A_r, A_1, and A_2 in order to achieve the interactive control and the coupling optimization.

β, γ are unknown to the robot. Therefore, we aim to propose a method to estimate them in the following section.

3. Adaptive Optimal Control

A recursive least squares algorithm with forgetting factors is used in this paper to get the estimate $\hat{\beta}, \hat{\gamma}$ of β, γ in order to estimate the feedback gains of the humans in real time and avoid the data saturation phenomenon caused by the standard least squares algorithm [46]. Subsequently, the estimate $\hat{Q}_1, \hat{Q}_2$ of Q_1, Q_2 can be obtained using Equation (10e,f).

Equation (10e) is used as the model for identification. For convenience, we let $\theta_1 = -\beta^T, y_1 = f_1^T, W = z^T$. Subsequently, Equation (10b) can be rewritten as

$$y_1 = W\theta_1. \tag{11}$$

The feedback gain of the human 1 are estimated by minimizing the total prediction error

$$J_1 = \int_0^t exp(-\lambda_1 t) \|y_1(s) - W(s)\hat{\theta}_1\|^2 ds \tag{12}$$

where λ_1 is the constant forgetting factor. The update rule of the parameter θ_1 can be obtained as

$$\begin{aligned} \dot{\hat{\theta}}_1 &= -PW^T e_1 \\ \dot{P} &= \lambda_1 P - PW^T WP \\ e_1 &= \hat{y}_1 - y_1. \end{aligned} \tag{13}$$

The estimated error of $\hat{\theta}_1$ is

$$e_{\theta_1}(t) = exp(-\lambda_1 t)P(t)P^{-1}(0)e_{\theta_1}(0). \tag{14}$$

Thus, the estimate $\hat{\beta}$ can be obtained as

$$\hat{\beta} = -\hat{\theta}_1^T. \tag{15}$$

Similarly, we let $\theta_2 = -\gamma^T$, $y_2 = f_2^T$, $W = z^T$. Afterwards, Equation (10c) can be rewritten as

$$y_2 = W\theta_2. \tag{16}$$

The feedback gain of the human 2 are estimated by minimizing the total prediction error

$$J_2 = \int_0^t exp(-\lambda_2 t)\|y_2(s) - W(s)\hat{\theta}_2\|^2 ds \tag{17}$$

where λ_2 is the constant forgetting factor. The update rule of the parameter θ_2 can be obtained as

$$\begin{aligned} \dot{\hat{\theta}}_2 &= -PW^T e_2 \\ \dot{P} &= \lambda_2 P - PW^T WP \\ e_1 &= \hat{y}_2 - y_2. \end{aligned} \tag{18}$$

The estimated error of $\hat{\theta}_2$ is

$$e_{\theta_2}(t) = exp(-\lambda_2 t)P(t)P^{-1}(0)e_{\theta_2}(0). \tag{19}$$

Thus, the estimate $\hat{\gamma}$ can be obtained as

$$\hat{\gamma} = -\hat{\theta}_2^T. \tag{20}$$

Equations (13), (15), (18) and (20) are critical, because they enable each agent to recognize their partners' control objectives and use Equation (10a–f) to adjust their own control.

In order to ensure the performance of cooperative object transporting task, we let

$$Q + Q_1 + Q_2 \equiv C \tag{21}$$

where C is the total weight. The cooperative object transporting task fixes the task performance through the total weight C and uses Equation (21) to share the the effort between 2 humans and the robot. Equation (21) makes the proposed controller be able to adjust the contributions between the humans and the robot and makes the humans and the robot take complementary roles as well.

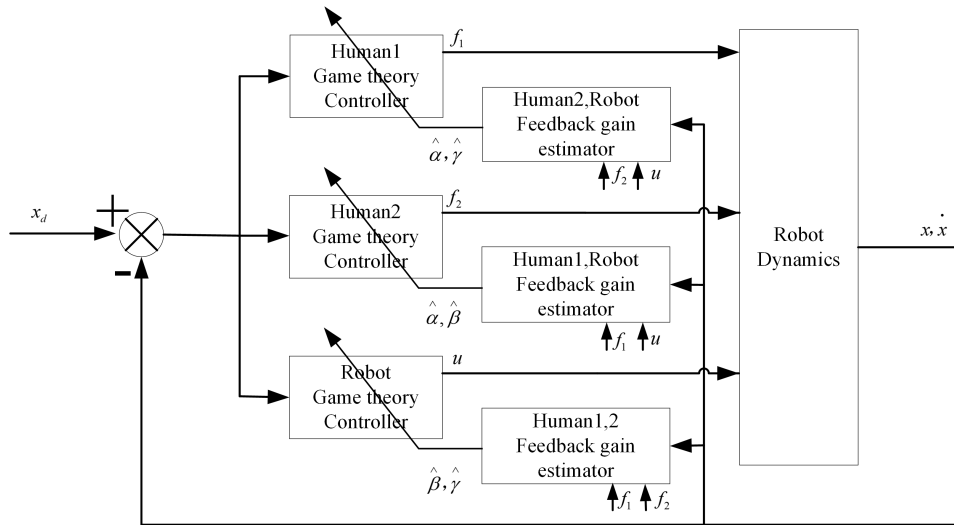The control architecture is shown in Figure 2.

Figure 2. Control Architecture.

A pseudo-code summarizes the implementation procedures of the proposed method as Algorithm 1.

Algorithm 1 Adaptive optimal control algorithm based on N-player game

---

Input: Current state z, target x_d.
Output: Robot's control input u, estimated the humans' cost function state weight $\hat{Q}_1, \hat{Q}_2$ in Equation (10e,f).
Begin
    Define x_d, initialize $Q, \hat{Q}_1, \hat{Q}_2, u, f_1, f_2, \hat{z}, \alpha, \hat{\beta}, \hat{\gamma}, P_r, \hat{P}_1, \hat{P}_2$, set λ_1 in Equation (13), λ_2 in Equation (18), C in Equation (21), the terminal time t_f of one trial.
    While $t < t_f$ do
        Measure the position $x(t)$, velocity $\dot{x}(t)$, and form z.
        Update $\hat{\beta}$ using Equations (13) and (15), Update $\hat{\gamma}$ using Equations (18) and (20).
        Solve the Riccati equation in Equation (10d) to obtain P, and calculate the robot's control input u.
        Calculate estimated the humans' cost function state weights $\hat{Q}_1, \hat{Q}_2$ in Equation (10e,f) using the Riccati equation.
        Compute robot's cost function state weight Q according to Equation (21).

Theorem 1. *Consider the robot dynamics shown in Equation (5). If the robot and the humans estimate the parameters of their partners' controller and adjust their own control according to Equations (10a–f), (13), (15), (18), (20) and (21), then the following conclusions will be drawn:*

- *The closed-loop system is stable, and $z, \alpha, \hat{\beta}, \hat{\gamma}, u$ are bounded.*
- $\lim_{x \to \infty} \hat{Q}_1 = Q_1, \lim_{x \to \infty} \hat{Q}_2 = Q_2,$ *which indicate that $\hat{Q}_1, \hat{Q}_2$ converge to the correct values Q_1, Q_2, if z is persistently exciting.*
- *The Nash equilibrium is achieved for th human-robot-human interaction system.*

Proof of Theorem 1. $\hat{\beta}, \hat{\gamma}$ influence u, f_1, f_2, z as following:

$$\dot{\hat{z}} = A\hat{z} + B\hat{u} + Bf_1 + B_2. \tag{22}$$

By subtracting Equation (5) from Equation (22), we have

$$\dot{e}_z = Ae_z + B(\hat{u} - u) + Be_{f_1} + Be_{f_2} \tag{23}$$

where $e_z = \hat{z} - z$. By considering Equation (10a–c), we have

$$\dot{e}_z = (A - B\alpha)e_z + Be_{\theta_1}^T z + Be_{\theta_2}^T z. \tag{24}$$

Consider the Lyapunov function candidate as following

$$W = \frac{1}{2}z^T z + \frac{1}{2}e_{\theta_1}^T e_{\theta_1} + \frac{1}{2}e_{\theta_2}^T e_{\theta_2} + \frac{\chi}{2}e_z^T e_z \tag{25}$$

where $\chi = min(\frac{2(\lambda_1 - \rho)\pi}{\varphi^2 \|B\|^2}, \frac{2(\lambda_2 - \rho)\pi}{\varphi^2 \|B\|^2})$, with ρ being the upper bound of the maximum eigenvalue of $\dot{P}P^{-1}$, π being the lower bound of the minimum eigenvalue of $B\alpha - A$, φ being the upper bound of $\|z\|$.

When considering function $V = \frac{1}{2}z^T z$ and differentiating V with respect to time, we obtain

$$\dot{V} = z^T \dot{z} = -z^T (B\alpha + B\beta + B\gamma - A)z. \tag{26}$$

According to Equation (10d), $B\alpha + B\beta + B\gamma - A$ is positive definite if Q is positive definite, it follows $\lim_{t\to\infty}\|z\| = 0$. Therefore, z is bounded and we define φ as the upper bound of $\|z\|$. By differentiating Equation (25), with respect to time, and considering Equations (14), (19) and (24), we obtain

$$
\begin{aligned}
\dot{W} =& z^T \dot{z} + e_{\theta_1}^T \dot{e}_{\theta_1} + e_{\theta_2}^T \dot{e}_{\theta_2} + \chi e_z^T \dot{e}_z \\
=& -z^T (B\alpha + B\beta + B\gamma - A)z \\
& - \lambda_1 e_{\theta_1}^T e_{\theta_1} + e_{\theta_1}^T \dot{P}P^{-1}e_{\theta_1} - \lambda_2 e_{\theta_2}^T e_{\theta_2} + e_{\theta_2}^T \dot{P}P^{-1}e_{\theta_2} \\
& - \chi e_z^T (B\alpha - A)e_z + \chi e_z^T Be_{\theta_1}^T z + \chi e_z^T Be_{\theta_2}^T z \\
\leq& -z^T (B\alpha + B\beta + B\gamma - A)z \\
& - \lambda_1 \|e_{\theta_1}\|^2 + \rho\|e_{\theta_1}\|^2 - \lambda_2\|e_{\theta_2}\|^2 + \rho\|e_{\theta_2}\|^2 \\
& - \chi\pi\|e_z\|^2 + \chi\varphi\|B\|\|e_z\|\|e_{\theta_1}\| + \chi\varphi\|B\|\|e_z\|\|e_{\theta_2}\| \\
=& -z^T (B\alpha + B\beta + B\gamma - A)z \\
& - (\sqrt{\lambda_1 - \rho}\|e_{\theta_1}\| - \sqrt{\frac{\chi\pi}{2}}\|e_z\|)^2 \\
& - 2\sqrt{\lambda_1 - \rho}\sqrt{\frac{\chi\pi}{2}}\|e_{\theta_1}\|\|e_z\| + \chi\varphi B\|e_z\|\|e_{\theta_1}\| \\
& - (\sqrt{\lambda_2 - \rho}\|e_{\theta_2}\| - \sqrt{\frac{\chi\pi}{2}}\|e_z\|)^2 \\
& - 2\sqrt{\lambda_2 - \rho}\sqrt{\frac{\chi\pi}{2}}\|e_{\theta_2}\|\|e_z\| + \chi\varphi\|B\|\|e_z\|\|e_{\theta_2}\| \\
\leq& -z^T (B\alpha + B\beta + B\gamma - A)z \\
& + (-2\sqrt{\lambda_1 - \rho}\sqrt{\frac{\chi\pi}{2}} + \chi\varphi\|B\|)\|e_z\|\|e_{\theta_1}\| \\
& + (-2\sqrt{\lambda_2 - \rho}\sqrt{\frac{\chi\pi}{2}} + \chi\varphi\|B\|)\|e_z\|\|e_{\theta_2}\| \\
\leq& 0
\end{aligned} \tag{27}
$$

According to Equations (26) and (27), we have $\lim_{t\to\infty}\|z\| = 0$, $\lim_{t\to\infty}\|e_z\| = 0$. Therefore, $z(t)$ is bounded and $\lim_{t\to\infty}\|\dot{e}_z\| = 0$. According to Equation (27), we have $\lim_{t\to\infty}\|e_{\theta_1}\| = 0$, $\lim_{t\to\infty}\|e_{\theta_2}\| = 0$. Because of $e_{\theta_1} = \hat{\theta}_1 - \theta_1 = (-\hat{\beta})^T - (-\beta)^T = \beta^T - \hat{\beta}^T$, $e_{\theta_2} = \hat{\theta}_2 - \theta_2 = (-\hat{\gamma})^T - (-\gamma)^T = \gamma^T - \hat{\gamma}^T$, we can obtain $\lim_{t\to\infty}\|\beta^T - \hat{\beta}^T\| = 0$, $\lim_{t\to\infty}\|\gamma^T - \hat{\gamma}^T\| = 0$. β, γ are assumed to be bounded, since they are the feedback gains of the

humans. Therefore, $\hat{\beta}, \hat{\gamma}$ are also bounded. According to Equation (10a–c), P_1, P_2 are also bounded. According to Equation (10d), A_r is bounded. Therefore, P, α and u are bounded.

According to Equation (10e), we can calculate the estimated errors $e_{Q_1} = \hat{Q}_1 - Q_1, e_{Q_2} = \hat{Q}_2 - Q_2$. e_{Q_1}, e_{Q_2} are due to the errors e_P, e_{P_1}, e_{P_2}. Because e_P, e_{P_1}, e_{P_2} converge to zero, we have $\lim_{t \to \infty} \|e_{Q_1}\| = 0$, $\lim_{t \to \infty} \|e_{Q_2}\| = 0$, that is $\lim_{t \to \infty} \hat{Q}_1 = Q_1, \lim_{t \to \infty} \hat{Q}_2 = Q_2$.

Multiplying Equation (10d) by $\hat{z}^T$ on the left side and by $\hat{z}$ on the right side, and considering Equation (13), we have

$$
\begin{aligned}
0 =& \hat{z}^T Q \hat{z} + \hat{z}^T P_r B B^T P_r \hat{z} + \hat{z}^T P_r \dot{\hat{z}} \\
&+ \hat{z} P_r \dot{\hat{z}}^T + \hat{z}^T P_r H e_z + \hat{z} P_r H e_z^T \\
\equiv& \hat{\sigma}.
\end{aligned}
\tag{28}
$$

Considering $\lim_{t \to \infty} e_z = 0, \lim_{t \to \infty} \dot{e}_z = 0$, we can obtain

$$
\begin{aligned}
\lim_{t \to \infty} \sigma \equiv \lim_{t \to \infty} (&z^T Q Z + z^T P_r B B^T P_r z \\
&+ z^T P_r \dot{z} + z P_r \dot{z}^T) \\
=& 0.
\end{aligned}
\tag{29}
$$

Similarly, we can obtain

$$
\begin{aligned}
\lim_{t \to \infty} \sigma_1 \equiv \lim_{t \to \infty} (&z^T Q_1 Z + z^T P_1 B B^T P_1 z \\
&+ z^T P_1 \dot{z} + z P_1 \dot{z}^T) \\
=& 0 \\
\lim_{t \to \infty} \sigma_2 \equiv \lim_{t \to \infty} (&z^T Q_2 Z + z^T P_2 B B^T P_2 z \\
&+ z^T P_2 \dot{z} + z P_2 \dot{z}^T) \\
=& 0.
\end{aligned}
\tag{30}
$$

$\lim_{t \to \infty} \sigma = 0, \lim_{t \to \infty} \sigma_1 = 0$ and $\lim_{t \to \infty} \sigma_2 = 0$ indicate that the Nash equilibrium is achieved for the human-robot-human interaction system. $\square$

4. Simulations and Results

4.1. Experimental Design and Ssimulation Settings

With the development of the robot technology, in the future, robots will enter our homes and become a member of family in our daily lives. In our daily lives, we often need to carry various objects. Some objects (e.g., objects with smaller size and lower weight) can be successfully carried by one human; some objects (e.g., objects with medium size and medium weight) need to be carried successfully by two humans; some objects (e.g., objects with larger size and higher weight) can be carried successfully by three or more humans. Consider one scenario: In our home, we have an object (such as a table with a relatively larger size and higher weight) that need to be carried by three humans. However, there are only two humans in the home. In this case, we can let the robot help us carry the object together with the two humans. The robot can play the same role as one human. A simulation is conducted with CoppeliaSim in order to verify the control performance of the controller proposed in this paper. The version of CoppeliaSim that we used is CoppeliaSim 4.0.0 (CoppeliaSim Edu, Windows). Figure 3 demonstrates the CoppeliaSim simulation scenario of cooperative object transporting task. The humans cooperate with the robot to transport the object between -10 cm and $+10$ cm back and forth along the horizontal direction.
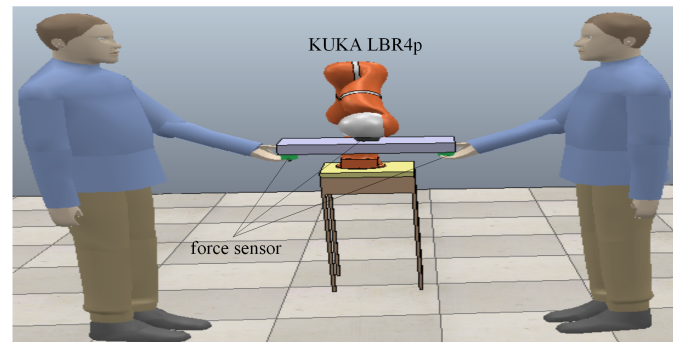
Figure 3. Simulation of cooperative object transporting task. The humans cooperate with the robot to transport the object back and forth between -10 cm and $+10$ cm along the horizontal direction. The forces that are exerted by the humans on the object are measured by force sensors at the interaction point.

The controller that is proposed in this paper implements interactive control because every agent considers the control of other partners. In order to present the advantages of the proposed controller, we compare the proposed controller with the linear quadratic regulators (LQR) optimal controller. The LQR controller can be obtained by setting $A_r = A$, $A_1 = A$, $A_2 = A$ in Equation (10d–f). The LQR controller allows each agent to form its own control input optimally, but it ignores the controls of other partners. Let $Q = Q_1 = Q_2 = diag(100, 0)$.

The cost functions of the humans usually change during the physical human-robot-human interaction. The robot needs to identify the change and adaptively adjust its own cost function in order to complete the cooperative object transporting task. In order to verify the ability of the robot to adaptively interact with two humans when humans' cost functions change, we simulated a scenario where the robot cooperated with the humans to perform an object transporting task. The task performance is achieved by setting the value of C in Equation (21). Let $C = diag(300, 0)$. The cost functions of the human 1 and the human 2 change randomly according to $Q_1 = diag(50, 0) + \rho \cdot diag(50, 0)$, $Q_2 = diag(50, 0) + \rho \cdot diag(50, 0)$ (ρ is a uniformly distributed random number between $[0, 1]$).

The human-robot-human cooperative object transporting task can be fulfilled with less effort with the proposed controller. In order to make this affirmation, we made a comparison with a human-robot cooperative object transporting task. In simulation of the human-robot-human cooperative object transporting task, we let $Q = Q_1 = Q_2 = diag(100, 0)$. In simulation of the human-robot cooperative object transporting task, we let $Q = diag(100, 0)$, $Q_1 = diag(100, 0)$, $Q_2 = diag(0, 0)$.

We assume that the humans and the robot do not have prior knowledge of each other (thus, initially $\hat{\alpha} \equiv 0, \hat{\beta} \equiv 0, \hat{\gamma} \equiv 0$). The control input of the robot are generated by Equations (5), (10a–f), (13), (15), (18) and (20). The simulated interaction forces f_1, f_2 of the human 1 and the human 2 are generated by a similar set of equations. The simulation time is 40 s. Let the inertia of the robot $M_d = 6$ kg, the damping of the robot $C_d = -0.2 \ N \cdot m^{-1}$ [19], the real-time least squares algorithm forgetting factor $\lambda_1 = \lambda_2 = 0.95$. Simulation time step is 0.005 s.

4.2. Results

Figure 4 depicts the change in position of the end effector with respect to time. The results plotted in Figure 4 is a smooth curve that looks like a sinusoidal signal. This smooth curve is determined by Equation (3). In Equation (3), $u(t), f_1(t), f_2(t)$ are iteratively calculated by our proposed controller based on game theory. Due to the fact that the humans and the robot do not transport the object at a constant speed using our method, the end effector follows a curve signal rather than a straight line signal. As can be seen from Figure 4, the end effector can reach the target position with the proposed controller which means that the cooperative object transporting task is successfully fulfilled. In contrast, the end effector can not reach the target position with the LQR controller, which means that

the cooperative object transporting task is not successfully fulfilled. The reason why the cooperative object transporting task can be successfully fulfilled with the proposed controller rather than with the LQR controller is that the proposed controller considers the interaction with other partners. When one partner decreases effort, the other partners will gradually increase their efforts to ensure the successful fulfillment of the cooperative object transporting task. In contrast, the LQR controller does not consider the interaction with other partners, so the cooperative object transporting task cannot be guaranteed to be successfully fulfilled.
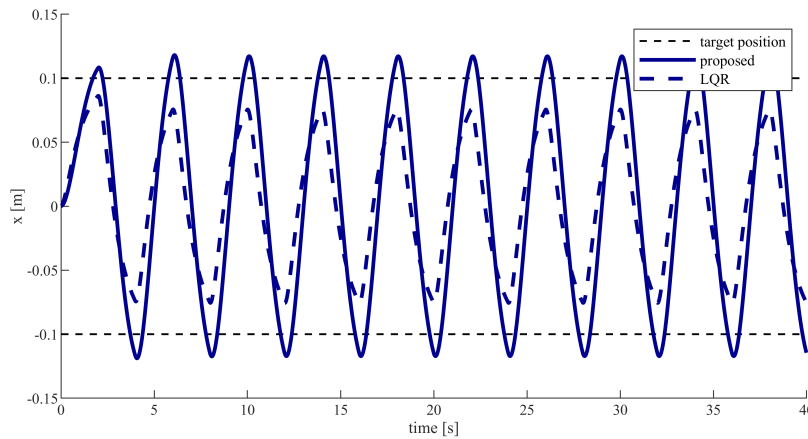

Figure 4. The end effector position value.

In Figure 5, we can see that the estimated humans' feedback gains converge to the real values in a few seconds. This means that the humans' feedback gains can be successfully estimated by the proposed method.
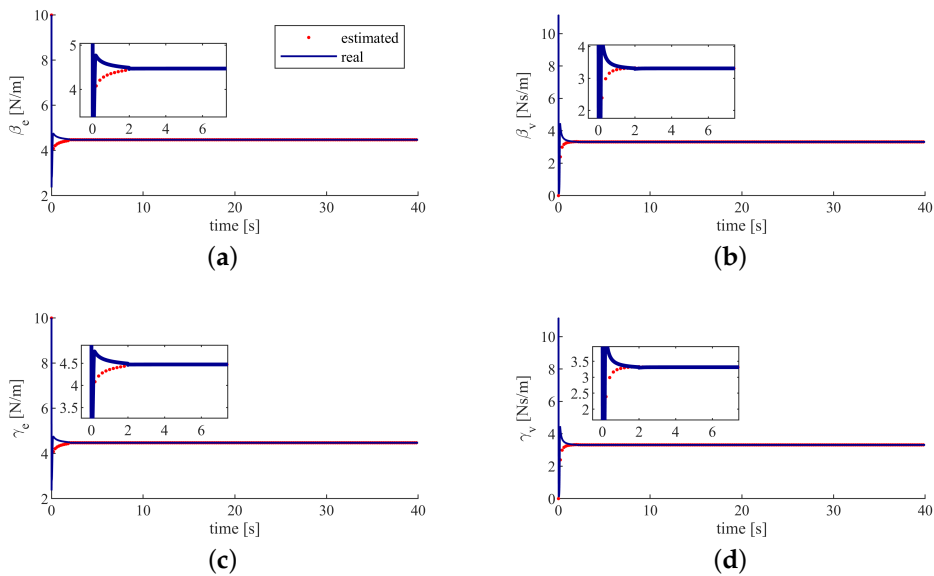

Figure 5. Control gains of humans. (**a**) the position error feedback gain of the human 1. (**b**) the velocity feedback gain of the human 1. (**c**) the position error feedback gain of the human 2. (**d**) the velocity feedback gain of the human 2.

Figure 6 demonstrates that fulfilling the cooperative object transporting task requires larger control gains β, γ with the LQR controller compared with the controller proposed in this paper. It means that accomplishing the same task requires less effort using the proposed controller. This is because that the

proposed controller considers the interaction with other partners and calculates the minimal effort for the humans and the robot to complete the task. In contrast, the LQR controller doesn't consider the interaction with other partners, so the humans and the robot only minimize their own cost function and may, therefore, require larger effort.
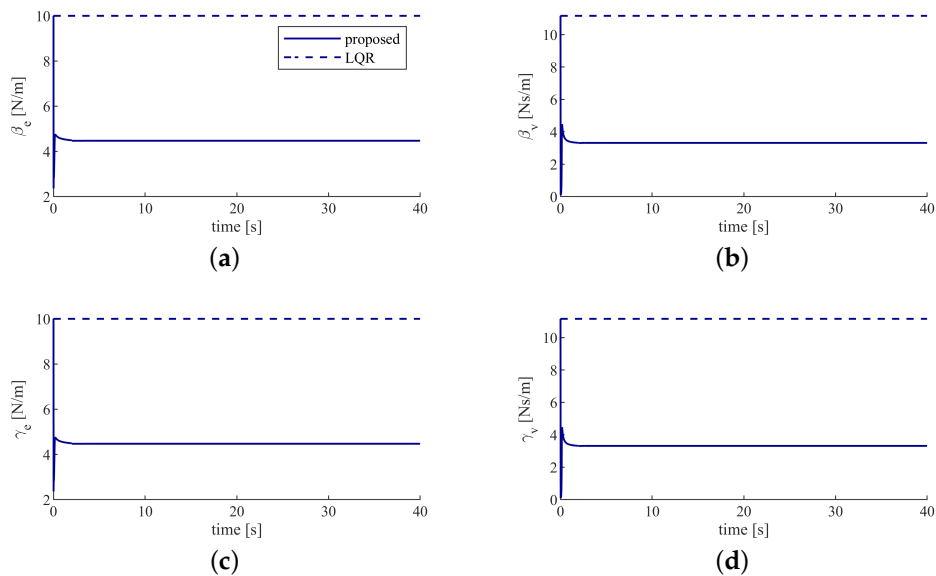

Figure 6. Humans' control gains (**a**) the position error feedback gain of the human 1. (**b**) the velocity feedback gain of the human 1. (**c**) the position error feedback gain of the human 2. (**d**) the velocity feedback gain of the human 2.

The feedback gains are affected by the state weights of the cost functions. In order to verify the advantages of the proposed controller when the state weights vary, we let Q_1 vary from 0 to $10Q$ with $Q_2 = diag(100, 0)$ and let Q_2 vary from 0 to $10Q$ with $Q_1 = diag(100, 0)$ respectively. It can be seen from Figure 7 that accomplishing the same task always requires less effort using the proposed controller. We can also see that the difference between the control gains with our proposed controller and the control gains with LQR controller becoming smaller when Q_1/Q or Q_2/Q increases, this is because that the robot's relative influence decreases.

From Figures 4–7, we can conclude that the human-robot-human cooperative object transporting task can be fulfilled with less effort and the system can be kept stable using the proposed controller.

It can be seen from Figure 8 that, when the cost functions of the human 1 and the human 2 change, the cost function of the robot will also change adaptively. When the sum of the state weights of the human 1 and the human 2 $Q_1 + Q_2$ increases, the state weight of the Robot Q decreases accordingly. Conversely, when the sum of the state weights of the human 1 and the human 2 $Q_1 + Q_2$ decreases, the state weight of the robot Q increases accordingly. The reason why the robot can change adaptively is that we set the constant C value in Equation (21). Equation (21) makes the proposed controller able to adjust the contributions between the humans and the robot and makes the humans and the robot take complementary roles as well.

Figure 9 shows that, using the proposed controller, the adaptive cooperative object transporting task can be fulfilled and the system can be kept stable.
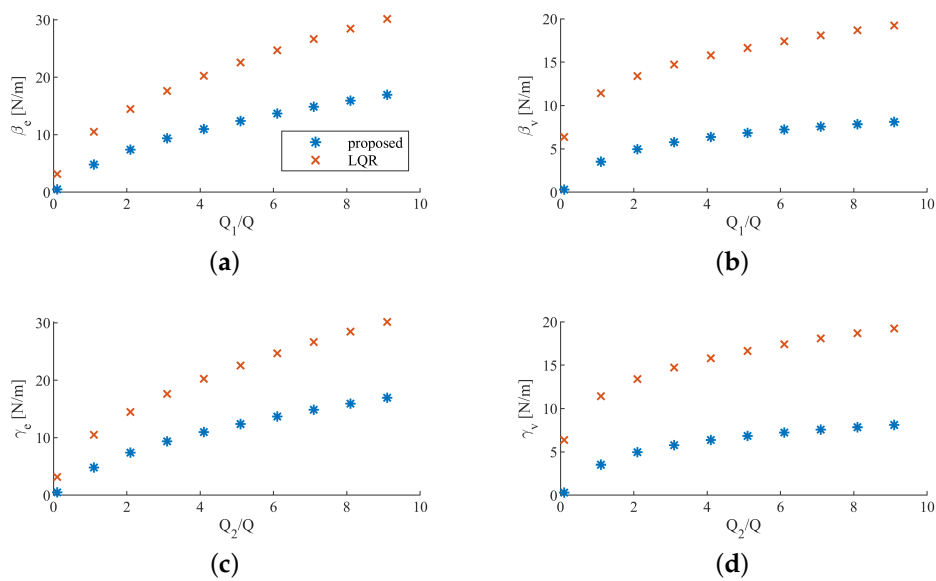
Figure 7. Control gains for different values of humans' state weights. (**a**) and (**b**) the state weight of the human 1 vary. (**c**) and (**d**) the state weight of the human 2 vary.
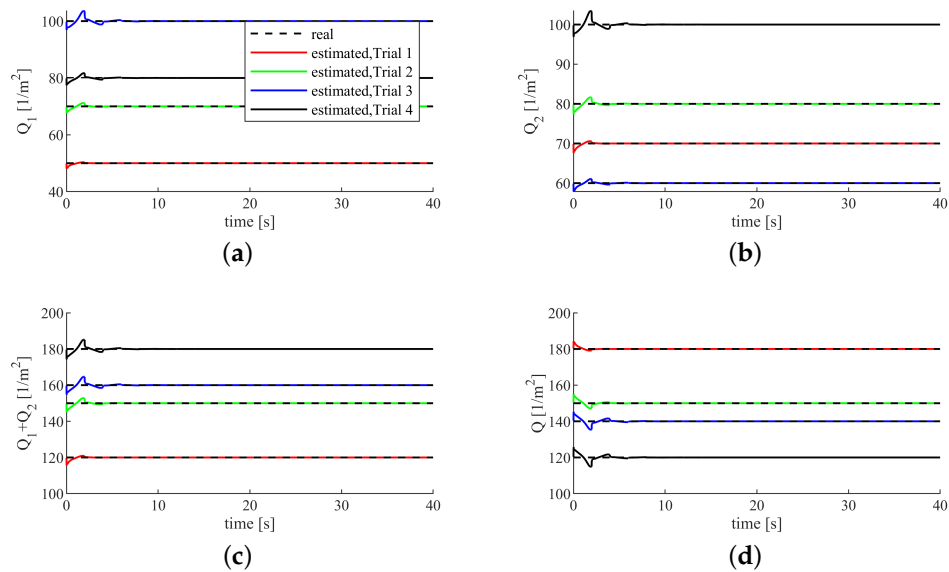

Figure 8. Humans' state weights. (**a**) the state weight of the human 1. (**b**) the state weight of the human 2. (**c**) the sum of the state weights of the human 1 and human 2. (**d**) the state weight of the robot.

From Figures 8 and 9, we can conclude that the adaptive cooperative object transporting task can be fulfilled with the proposed controller. During the physical interaction, the robot can successfully identify the change of each human's cost function, and then adaptively adjust its own cost function to achieve interactive optimal control.

Figure 10 demonstrates that fulfilling the human-robot-human cooperative object transporting task requires smaller control gains β_e, β_v as compared with the human-robot cooperative object transporting task. It means that accomplishing the same task requires less effort by means of the human-robot-human physical interaction. This is because the human-robot-human cooperative object transporting task considers the interaction with more partners (two partners) and calculates minimal effort for the humans and the robot to complete the task. In contrast, the human-robot cooperative

object transporting task consider the interaction with less partners (only one partner), so the human and the robot may therefore require larger effort.
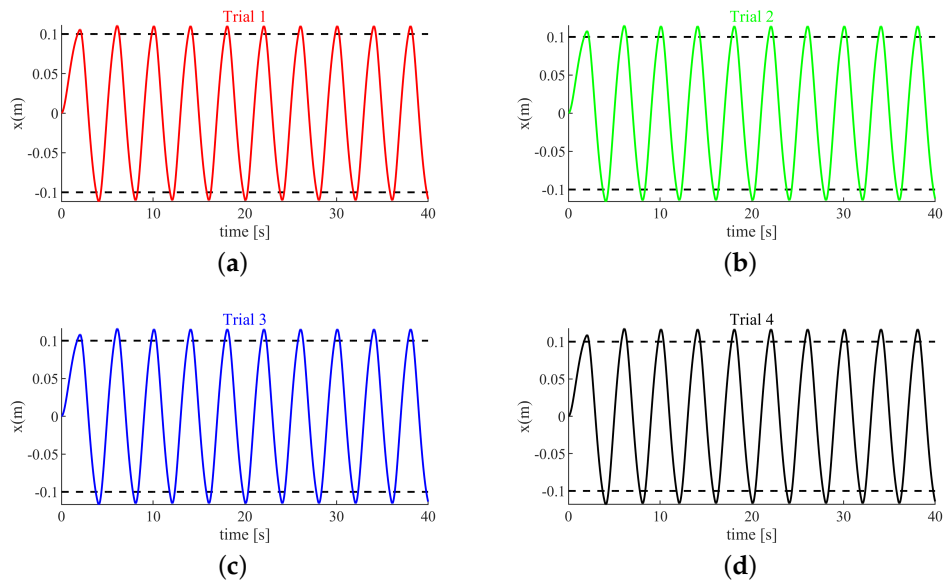

Figure 9. The end effector position value. (**a**) The end effector position value in Trial 1. (**b**) The end effector position value in Trial 2. (**c**) The end effector position value in Trial 3. (**d**) The end effector position value in Trial 4.
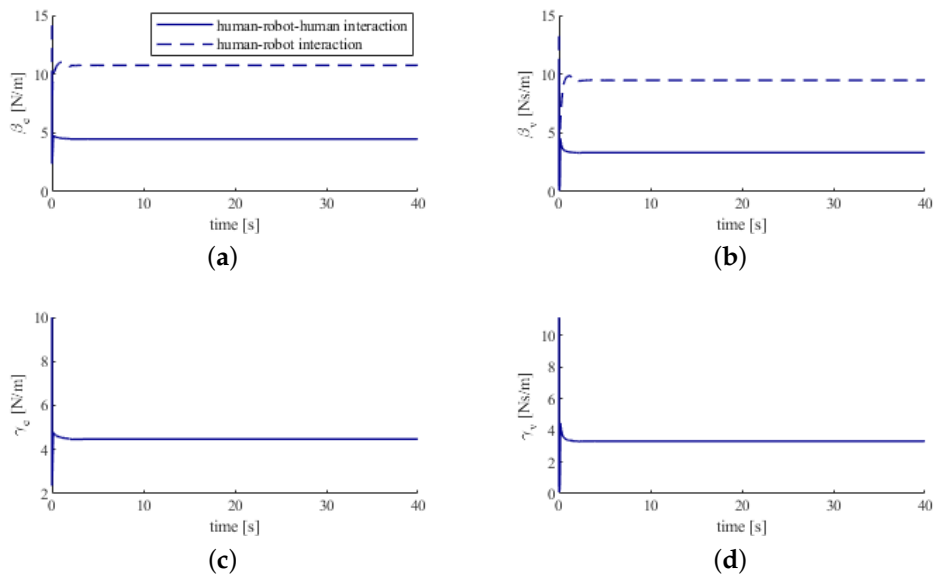

Figure 10. Humans' control gains. The dashed lines correspond to the human-robot cooperative object transporting task. The solid lines correspond to the human-robot-human cooperative object transporting task. (**a**) the position error feedback gain of the human 1. (**b**) the velocity feedback gain of the human 1. (**c**) the position error feedback gain of the human 2. (**d**) the velocity feedback gain of the human 2.

5. Conclusions

In this paper, the human-robot-human physical interaction problem has been studied. An adaptive optimal control framework for the human-robot-human physical interaction has been proposed based on N-player game theory. The recursive least squares algorithm based on forgetting factor has

been used to identify unknown control parameters of the humans online. The performance of the controller proposed in this paper has been verified by simulations of cooperative object transporting task. The simulation results show that the proposed controller can achieve adaptive optimal control during the interaction between the robot and two humans and keep the system stable. Compared with the LQR controller, the proposed controller has more superior performance. Compared with the human-robot physical interaction, accomplishing the same cooperative object transporting task requires less effort by means of the human-robot-human physical interaction based on the approach proposed in the paper. Although this paper only conducts simulations on the physical interaction between one robot and two humans, it is worth mentioning that the framework that is proposed in this paper has the potential to be generalized to the situation where multiple robots physically interact with multiple humans. As future work, we will extend the framework to the interaction between multiple robots and multiple humans.

Author Contributions: R.Z. conceived the original ideas, designed all the experiments, and subsequently drafted the manuscript. Y.L. provided supervision and funding support for the project. J.Z. provided supervision and funding support for the project. H.C. provided supervision. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. De Santis, A.; Siciliano, B.; De Luca, A.; Bicchi, A. An atlas of physical human-robot interaction. *Mech. Mach. Theory* **2008**, *43*, 253–270. [CrossRef]
2. Carolina, P.; Angelika, P.; Martin, B. A survey of environment-, operator-, and task-adapted controllers for teleoperation systems. *Mechatronics* **2010**, *20*, 787–801.
3. Losey, D.P.; McDonald, C.G.; Battaglia, E.; O'Malley, M.K. A review of intent detection, arbitration, and communication aspects of shared control for physical human-robot interaction. *Appl. Mech. Rev.* **2018**, *70*, 010804. [CrossRef]
4. Aslam, P.; Jeha, R. Safe physical human robot interaction-past, present and future. *J. Mech. Sci. Technol.* **2008**, *22*, 469.
5. Li, Y.; Ge, S.S. Human–robot collaboration based on motion intention estimation. *IEEE-ASME Trans. Mechatron.* **2013**, *19*, 1007–1014. [CrossRef]
6. Li, Y.; Ge, S.S. Force tracking control for motion synchronization in human-robot collaboration. *Robotica* **2016**, *34*, 1260–128. [CrossRef]
7. Sandra, H.; Martin, B. Human-oriented control for haptic teleoperation. *Proc. IEEE* **2012**, *100*, 623–647.
8. Chen, Z.; Huang, F.; Yang, C.; Yao, B. Adaptive fuzzy backstepping control for stable nonlinear bilateral teleoperation manipulators with enhanced transparency performance. *IEEE Trans. Ind. Electron.* **2019**, *67*, 746–756. [CrossRef]
9. Liu, C.; Masayoshi, T. Modeling and controller design of cooperative robots in workspace sharing human-robot assembly teams. In Proceedings of the IROS 2014, Chicago, IL, USA, 14–18 September 2014; pp. 1386–1391.
10. Zanchettin, A.M.; Casalino, A.; Piroddi, L.; Rocco, P. Prediction of human activity patterns for human-robot collaborative assembly tasks. *IEEE Trans. Ind. Inform.* **2018**, *15*, 3934–3942. [CrossRef]
11. Alexander, M.; Martin, L.; Ayse, K.; Metin, S.; Cagatay, B.; Sandra, H. The role of roles: Physical cooperation between humans and robots. *Int. J. Robot. Res.* **2012**, *31*, 1656–1674.
12. Costa, M.J.; Dieter, C.; Veronique, L.; Johannes, C.; El-Houssaine, A. A structured methodology for the design of a human-robot collaborative assembly workplace. *Int. J. Adv. Manuf. Technol.* **2019**, *102*, 2663–2681.
13. Daniel, N.; Jan, K. A problem design and constraint modelling approach for collaborative assembly line planning. *Robot. Comput. Integr. Manuf.* **2019**, *55*, 199–207.
14. Selma, M.; Sandra, H. Control sharing in human-robot team interaction. *Annu. Rev. Control* **2017**, *44*, 342–354.

15. Mahdi, K.; Aude, B. A dynamical system approach to task-adaptation in physical human-robot interaction. *Auton. Robot.* **2019**, *43*, 927–946.

16. Roberto, C.; Vittorio, S. Rehabilitation Robotics: Technology and Applications. In *Rehabilitation Robotics*; Colombo, R., Sanguineti, V., Eds.; Academic Press: London, UK, 2018; pp. xix–xxvi.

17. Colgate, J.E.; Decker, P.F.; Klostermeyer, S.H.; Makhlin, A.; Meer, D.; Santos-Munne, J.; Peshkin, M.A.; Robie, M. Methods and Apparatus for Manipulation of Heavy Payloads with Intelligent Assist Devices. U.S. Patent 7,185,774, 6 March 2007.

18. Zoss, A.B.; Kazerooni, H.; Chu, A. Biomechanical design of the Berkeley lower extremity exoskeleton (BLEEX). *IEEE-ASME Trans. Mechatron.* **2006**, *11*, 128–138. [CrossRef]

19. Li, Y.; Carboni, G.; Gonzalez, F.; Campolo, D.; Burdet, E. Differential game theory for versatile physical human-robot interaction. *Nat. Mach. Intell.* **2019**, *1*, 36–43. [CrossRef]

20. Li, Y.; Tee, K.P.; Yan, R.; Chan, W.L.; Wu, Y. A framework of human-robot coordination based on game theory and policy iteration. *IEEE Trans. Robot.* **2016**, *32*, 1408–1418. [CrossRef]

21. Nathanaël, J.; Themistoklis, C.; Etienne, B. A framework to describe, analyze and generate interactive motor behaviors. *PLoS ONE* **2012**, *7*, e49945.

22. Li, Y.; Tee, K.P.; Yan, R.; Chan, W.L.; Wu, Y.; Limbu, D.K. Adaptive optimal control for coordination in physical human-robot interaction. In Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, 28 September–3 October 2015; pp. 20–25.

23. Kirk, D.E. Optimal control theory: An introduction. In *Optimal Control Theory*; Dover Publications: Mineola, NY, USA, 2004.

24. Li, Y.; Tee, K.P.; Chan, W.L.; Yan, R.; Chua, Y.; Limbu, D.K. Continuous role adaptation for human-robot shared control. *IEEE Trans. Robot.* **2015**, *31*, 672–681. [CrossRef]

25. Lewis, F.L.; Vrabie, D. Reinforcement learning and adaptive dynamic programming for feedback control. *IEEE Circuits Syst. Mag.* **2009**, *9*, 32–50. [CrossRef]

26. Vamvoudakis, K.G.; Lewis, F.L. Multi-player non-zero-sum games: Online adaptive learning solution of coupled Hamilton–Jacobi equations. *Automatica* **2011**, *47*, 1556–1569. [CrossRef]

27. Zhang, H.; Wei, Q.; Liu, D. An iterative adaptive dynamic programming method for solving a class of nonlinear zero-sum differential games. *Automatica* **2011**, *47*, 207–214. [CrossRef]

28. Liu, D.; Li, H.; Wang, D. Online synchronous approximate optimal learning algorithm for multi-player non-zero-sum games with unknown dynamics. *IEEE Trans. Syst. Man Cybern. Syst.* **2014**, *44*, 1015–1027. [CrossRef]

29. Albaba, B.M.; Yildiz, Y. Modeling cyber-physical human systems via an interplay between reinforcement learning and game theory. *Annu. Rev. Control* **2019**, *48*, 1–21. [CrossRef]

30. Music, S.; Hirche, S. Haptic Shared Control for Human-Robot Collaboration: A Game-Theoretical Approach. In Proceedings of the 21st IFAC World Congress, Berlin, Germany, 12–17 July 2020.

31. Turnwald, A.; Wollherr, D. Human-like motion planning based on game theoretic decision making. *Int. J. Soc. Robot.* **2019**, *11*, 151–170. [CrossRef]

32. Liu, Z.; Liu, Q.; Xu, W.; Zhou, Z.; Pham, D.T. Human-robot collaborative manufacturing using cooperative game: Framework and implementation. *Procedia CIRP* **2018**, *72*, 87–92. [CrossRef]

33. Bansal, S.; Xu, J.; Howard, A.; Isbell, C. A Bayesian Framework for Nash Equilibrium Inference in Human-Robot Parallel Play. *arXiv* **2020**, arXiv:2006.05729.

34. Antonelli, G.; Chiaverini, S.; Marino, A. A coordination strategy for multi-robot sampling of dynamic fields. In Proceedings of the 2012 IEEE International Conference on Robotics and Automation, Saint Paul, MN, USA, 14–18 May 2012; pp. 1113–1118.

35. Yan, Z.; Jouandeau, N.; Cherif, A.A. A survey and analysis of multi-robot coordination. *Int. J. Adv. Robot. Syst.* **2013**, *10*, 399. [CrossRef]

36. Martina, L.; Alessandro, M.; Stefano, C. A distributed approach to human multi-robot physical interaction. In Proceedings of the 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), Bari, Italy, 6–9 October 2019.

37. Kim, W.; Marta, L.; Balatti, P.; Wu, Y.; Arash, A. Towards ergonomic control of collaborative effort in multi-human mobile-robot teams. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Macau, China, 3–8 November 2019.

38. Starr, A.W.; Ho, Y.-C. Nonzero-sum differential games. *J. Optim. Theory Appl.* **1969**, *3*, 184–206. [CrossRef]

39. Fudenberg, D.; Tirole, J. Noncooperative game theory for industrial organization: An introduction and overview. *Handb. Ind. Organ.* **1989**, *1*, 259–327.

40. Hegan, N. Impedance Control: An Approach To Manipulation: Part I-Theory Part II-Implementation Part III-Applications. *J. Dyn. Syst. Meas. Control* **1985**, *107*, 1–24. [CrossRef]

41. Blank, A.A.; Okamura, A.M.; Whitcomb, L.L. Task-dependent impedance and implications for upper-limb prosthesis control. *Int. J. Robot. Res.* **2014**, *33*, 827–846. [CrossRef]

42. Vogel, J.; Haddadin, S.; Jarosiewicz, B.; Simeral, J.D.; Bacher, D.; Hochberg, L.R.; Donoghue, J.P.; van der Smagt, P. An assistive decision-and-control architecture for force-sensitive hand–arm systems driven by human–machine interfaces. *Int. J. Robot. Res.* **2015**, *34*, 763–780. [CrossRef]

43. Basar, T.; Olsder, G.J. *Dynamic Noncooperative Game Theory*, 2nd ed.; Society for Industrial and Applied Mathematics; The Math Works Inc.: Natick, MA, USA, 1999.

44. Shima, T.; Rasmussen, S. UAV cooperative decision and control: Challenges and practical approaches. In *UAV Cooperative Decision and Control*; SIAM: Philadelphia, PA, USA, 2009.

45. Hudas, G.; Vamvoudakis, K.G.; Mikulski, D.; Lewis, F.L. Online adaptive learning for team strategies in multi-agent systems. *J. Def. Model. Simul.* **2012**, *9*, 59–69. [CrossRef]

46. Tan, H.J.; Chan, S.C.; Lin, J.Q.; Sun, X. A New Variable Forgetting Factor-Based Bias-Compensated RLS Algorithm for Identification of FIR Systems With Input Noise and Its Hardware Implementation. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2019**, *67*, 198–211. [CrossRef]