Proceedings of the ASME 2020 Dynamic Systems and Control Conference
DSCC 2020
October 4-7, 2020, Pittsburgh, Pennsylvania, USA

DSCC2020-23518

EXPERIMENTAL AND ANALYTICAL NONZERO-SUM DIFFERENTIAL GAME-BASED CONTROL OF A 7-DOF ROBOTIC MANIPULATOR

Mostafa Bagheri
Senior Research Engineer
Dynamic Systems and Control Lab.
Dept. of Mechanical and Aero. Eng.
UC San Diego & San Diego State
La Jolla, California 92093
Email: mstfbagheri@ucsd.edu
mbagheri@sdsu.edu

Alexander Bertino
Research Assistant
Dynamic Systems and Control Lab.
Dept. of Mechanical Eng.
San Diego State University
San Diego, California 92182
Email: abertino6245@sdsu.edu

Peiman Naseradinmousavi
Associate Professor
Dynamic Systems and Control Lab.
Dept. of Mechanical Eng.
San Diego State University
San Diego, California 92182
Email: pnaseradinmousavi@sdsu.edu

ABSTRACT

We formulate a Nash-based feedback control law for an Euler-Lagrange system to yield a solution to non-cooperative differential game. The robot manipulators are broadly utilized in industrial units on the account of their reliable, fast, and precise motions, while consuming a significant lumped amount of energy. Therefore, an optimal control strategy needs to be implemented in addressing efficiency issues, while delivering accuracy obligation. As a case study, we here focus on a 7-DOF robot manipulator through formulating a two-player feedback nonzero-sum differential game. First, coupled Euler-Lagrangian dynamic equations of the manipulator are briefly presented. Then, we formulate the feedback Nash equilibrium solution in order to achieve perfect trajectory tracking. Finally, the performance of the Nash-based feedback controller is analytically and experimentally examined. Simulation and experimental results reveal that the control law yields almost perfect tracking and achieves closed-loop stability.

1 Introduction

The robot manipulators are broadly utilized in industrial units on the account of their reliable, fast, and precise motions [1], while consuming a significant lumped amount of energy. Therefore, optimization schemes need to be employed to address efficiency issues, while delivering accuracy obligation. For a dynamically interconnected robotic system, there can be multiple players having different criteria whereas all players intend to execute a task specified. In the so-called "cooperative" game, players are subject to an agreement leading to a best feasible solution for the game, whereas for a "non-cooperative" game [2], each player pursues its own individual interests, which may result in conflict with the other ones. Note that a solution to the non-cooperative game is an equilibrium since it represents a control strategy providing a balance between interests of players. Although the players work to reach the same goal, each player has its own cost function leading to a multi-objective optimization problem to be dealt with.

The optimal control theory was developed as an efficient approach to determine optimal input parameters maximizing performance criteria defined, while satisfying physical constraints. The game theory [3] is an effective approach to address the concerns of having multiple players for complex dynamical systems.

Finding the Nash equilibria has received substantial attention in various disciplines including, but not limited to, mathematics, computer science, economics, and system engineering. Through the Nash strategy, the cost function, for each player, cannot unilaterally be minimized by changing the strategy. The game theory deals with strategic interactions among multiple players making simultaneous decisions, while each player tries to minimize its own cost function. The control of a high-DOF robot manipulator, with N-player, is an example of nonzero-sum non-cooperative differential game. In this problem, each player has its own criterion through minimizing its own control inputs and tracking errors, while all players intend to complete the task specified; tracking desirable joint-space trajectories. Note that a zero-sum game [4] is a mathematical representation of a situation in which gain or loss of each participant is balanced by the other ones. The players' objective functions cannot be easily optimized by ignoring the other ones' choices, since each player affects the actions of the other players. Therefore, it is crucial to achieve the goal of finding the feedback Nash equilibrium solution minimizing all players' cost functions, while completing the task. [5] provided a solution for a class of general N-player non-cooperative games. In 1960s, [6] investigated the multi-player extension of the dynamic programming solution for the differential games. [7] proposed a fuzzy-linear quadratic regulator (LQR) game-based control scheme to simultaneously enhance vehicle stability while compensating driver's inappropriate steering reaction in emergency avoidance. [8] presented the application of advanced optimization techniques to unmanned aerial system mission path planning system (MPPS) using multi-objective evolutionary algorithms (MOEAs). [9] studied a decentralized scheme for active noise control (ANC) from a game-theoretical perspective. They formalized the Nash equilibrium (i.e., the simultaneous best strategy) in the interaction between the controllers. [10] presented a game-theoretic analysis of a visibility-based pursuit—evasion game in a planar environment containing obstacles. In their work, the pursuer and the evader are assumed to be holonomic having bounded speeds. [11] investigated a method to analyze and select time-optimal coordination strategies for n robots whose configurations are constrained to lie on a C-space roadmap (which could, for instance, represent a Voronoi diagram). [12] proposed a framework to analyze the interactive behaviors of human and robot in physical interactions. They employed game theory in describing the system under study, and policy iteration was adopted to provide a solution of Nash equilibrium. [13] tried to achieve a superior performance with fuzzy Markov game based control by hybridizing two game theory based approaches of "fictitious play" and "minimax". They formulated a controller for a two link robot and compared its performance against fuzzy Markov game control and fuzzy Q control. [14] utilized the game theory approach, as a generalized form of nonlinear optimum control, in designing a closed-loop controller for a fixed-base two-link manipulator. [15] provided insight into the solution of difficult problems specific to N versus 1 games. To illustrate further N versus 1 game problems, a nonoptimal scalar case was presented in which the decentralized structure is proven superior. [16] considered a novel coupled state-dependent Riccati equation (SDRE) approach for systematically designing nonlinear quadratic regulator (NLQR) and H_∞ control of mechatronics systems. [17] derived set of necessary optimality conditions, which not only enable the determination of the saddle-point strategies for both participating players, but also the optimal parameters. Based on these conditions an iterative numerical algorithm of gradient type was suggested. More case studies can be found in [18–30].

The amount of cooperation between players resulted in different branches of game theory problems. The Nash optimal control scheme is progressed when players have additional information about the other ones. Therefore, it is assumed that the players can observe the actions of the other ones. Motivated by finding the feedback Nash equilibrium solution, we explore the use of differential game theory in formulating a controller to yield a stable Euler-Lagrangian dynamic system, here a robot manipulator, in order to follow the desired joint-space trajectories. Therefore, we employ an approach to solve the nonzero-sum non-cooperative differential game controlling a nonlinear system. Like other classical game theory algorithms, we assume that the players can observe the actions of each other and also know the model information of the game. Then, we formulate the feedback Nash equilibrium solution in order to achieve the perfect tracking. A stability analysis is then carried out to prove that all solutions asymptotically converge to desired trajectories using the Nash-based strategy. Finally, the simulation and experimental results are presented and discussed.

2 Mathematical Modeling

The redundant manipulator, which is being studied here, has 7-DOF as shown in Fig. 1. The manipulator's Denavit-Hartenberg parameters are shown in Table 1 provided by the manufacturer. The robot manipulator is modeled as follows,

$$M(q)\ddot{q}+C(q,\dot{q})\dot{q}+\phi(q)=\tau \tag{1}$$

where, q, $\dot{q}$, and $\ddot{q}\in\mathbb{R}^7$ are angles, angular velocities, and accelerations of joints, respectively, and $\tau\in\mathbb{R}^7$ indicates the vector of joints' driving torques. Also, $M(q)\in R^{7\times7}$, $C(q,\dot{q})\in R^{7\times7}$, and $\phi(q)\in\mathbb{R}^7$ are the mass, Coriolis, and gravitational matrices, respectively, which are symbolically derived using the Euler-Lagrange equation [31–39]. The inertia matrix $M(q)$ is symmetric, positive definite, and consequently invertible. This property is exploited in the subsequent development based on the following assumptions: 1) $M(q), C(q,\dot{q})$, and $\phi(q)$ matrices are known and 2) $q(t)$ and $\dot{q}(t)$ are measurable.

3 Designing the Nash Optimal Controller

Through the game theory, the Nash equilibrium is a solution of a non-cooperative game involving two or more players, in which, each player is assumed to know the equilibrium strategies of the other ones, and no player has anything to gain by changing only its own strategy. The players are committed to follow a predetermined strategy based on the knowledge of initial state,

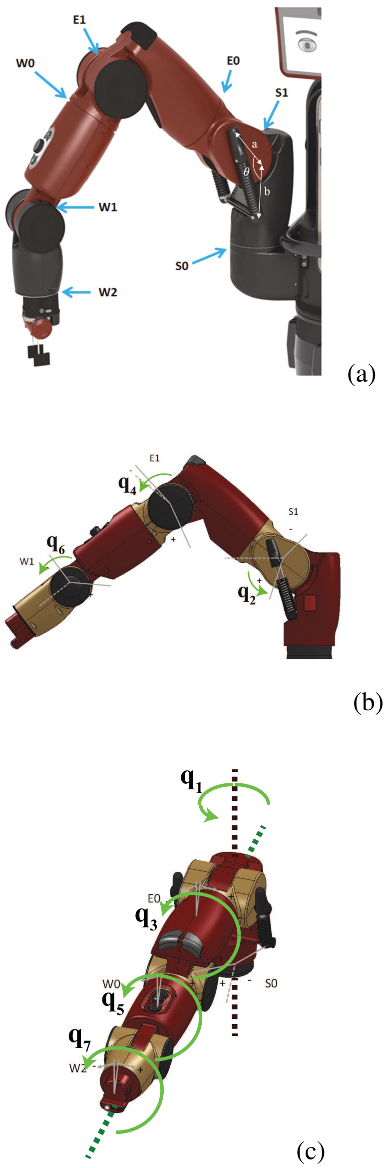

FIGURE 1. The 7-DOF Baxter's arm: (a) The joints' configurations; (b) sagittal view; (c) top view

system model, and cost function to be minimized. Note that solution techniques for the Nash equilibrium can be classified in various ways depending on the amount of information available to the players.

3.1 Error System Development

The control objective includes converging tracking errors to zero such that the generalized coordinates track the desired time-varying joints' trajectories ($q_{\text{des}}(t) \in \mathbb{R}^7$) as well as performance index. Consider the following assumption for the desired joint-space trajectories:

TABLE 1. Baxter's Denavit-Hartenberg Parameters

Link	a_i	d_i	α_i	θ_i
1	0.069	0.27035	$-\pi/2$	θ_1
2	0	0	$\pi/2$	$\theta_2+\pi/2$
3	0.069	0.36435	$-\pi/2$	θ_3
4	0	0	$\pi/2$	θ_4
5	0.010	0.37429	$-\pi/2$	θ_5
6	0	0	$\pi/2$	θ_6
7	0	0.3945	0	θ_7

Assumption 1. *The desired joint-space trajectories, $q_{\text{des}}(t), \dot{q}_{\text{des}}(t)$, and $\ddot{q}_{\text{des}}(t) \in \mathbb{R}^7$, exist and are bounded for all $t \geq 0$.*

To quantify the tracking performance, the angular (e_1) and combined (e_2) tracking errors are defined as

$$e_1 = q_{\text{des}} - q \tag{2}$$

$$e_2 = \dot{e}_1 + \alpha e_1 \tag{3}$$

where $e_1,\ e_2 \in \mathbb{R}^7$ and $\alpha \in \mathbb{R}^{7\times 7}$ is a constant positive definite matrix. A state-space model can be developed based on the tracking errors of Eqs. (2) and (3). According to this model, a controller is derived to improve tracking performance indices subject to the assumption of knowing dynamics of the system, as mentioned earlier. The control term is then established as the solution to the nonzero-sum Nash differential game.

A state-space model, based on the tracking error, is formulated through premultiplying the inertia matrix by the time derivative of Eq. (3),

$$\begin{aligned} M\dot{e}_2 &= M\ddot{q}_{\text{des}} + C\dot{q}_{\text{des}} + (M\alpha - C)e_2 \\ &\quad + (-M\alpha^2 + C\alpha)e_1 + G - \tau \\ \rightarrow \dot{e}_2 &= \ddot{q}_{\text{des}} + M^{-1}C\dot{q}_{\text{des}} + (\alpha - M^{-1}C)e_2 \\ &\quad + (-\alpha^2 + M^{-1}C\alpha)e_1 + M^{-1}G - M^{-1}\tau \end{aligned} \tag{4}$$

which yields,

$$\dot{e}_2 = \alpha e_2 - \alpha^2 e_1 + h - M^{-1}\tau \tag{5}$$

where $h \in \mathbb{R}^7$ is a nonlinear function defined as

$$h = \ddot{q}_{\text{des}} + M^{-1}(C\dot{q}_{\text{des}} + G + C\alpha e_1 - Ce_2) \tag{6}$$

and the state-space model of error dynamics becomes,

$$\dot{e} = f(e,\tau) = \begin{bmatrix} e_2 - \alpha e_1 \\ \alpha e_2 - \alpha^2 e_1 + h - M^{-1}\tau \end{bmatrix} \tag{7}$$

Since the dynamics of system (1) is known, the controller, based on Eq. (5), is designed as

$$\tau_{7\times 1} = M(h - (u_1 + u_2 + \cdots + u_N)) \tag{8}$$

where N is the number of players and $u_1(t), \cdots, u_N \in \mathbb{R}^7$ are auxiliary players' control inputs, which are formulated to minimize their own cost functions including the tracking errors. For two players having a feasible computation cost in real-time operation, substituting Eq. (8) into Eq. (5) results in the closed-loop error signal for $e_2(t)$ as

$$\dot{e}_2 = -\alpha^2 e_1 + \alpha e_2 + u_1 + u_2 \tag{9}$$

Finally, the state-space model for error dynamics is derived as follows,

$$\dot{e} = Ae + B_1 u_1 + B_2 u_2 \tag{10}$$

where $e = [e_1^T, e_2^T]^T \in \mathbb{R}^{14}$, and $A \in \mathbb{R}^{14\times 14}$ and $B_i \in \mathbb{R}^{14\times 7}$ $(i = 1,2)$ are defined as

$$A = \begin{bmatrix} -\alpha & I_{7\times 7} \\ -\alpha^2 & \alpha \end{bmatrix} \tag{11}$$

$$B_i = [0_{7\times 7} \;\; I_{7\times 7}]^T \quad i = 1,2 \tag{12}$$

where $I_{7\times 7}$ and $0_{7\times 7}$ are identity and zero matrices, respectively.

Note that through the Nash equilibrium solution, the performance of each player cannot be improved by a unilateral change of strategy. To determine the two-player feedback Nash nonzero-sum differential game solution, we define the following cost functions $J_1(e,u_1,u_2)$ and $J_2(e,u_1,u_2) \in \mathbb{R}$ as

$$J_1 = \frac{1}{2}\int_{t_0}^{\infty} \left(e^T Qe + u_1^T R_{11} u_1 + u_2^T R_{12} u_2\right) dt \tag{13}$$

$$J_2 = \frac{1}{2}\int_{t_0}^{\infty} \left(e^T Le + u_2^T R_{22} u_2 + u_1^T R_{21} u_1\right) dt \tag{14}$$

where $t_0 \in \mathbb{R}$ is the initial time and Q, $L \in \mathbb{R}^{14\times 14}$ are symmetric semi-definite constant matrices defined as

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix}, \quad L = \begin{bmatrix} L_{11} & L_{12} \\ L_{12}^T & L_{22} \end{bmatrix} \tag{15}$$

where Q and L impose penalties on the tracking errors. Also, $R_{ij} \in \mathbb{R}^{7\times 7}$ is a constant positive definite matrix. Note that we here focus on a game with memoryless perfect state information. Therefore, the controller's information set contains the initial conditions e_0 as well as the current state estimates $e(t)$ at time t. The actions of the players are completely determined by the relations $(u_1,u_2) = (\gamma_1(e_0,e), \gamma_2(e_0,e))$, where $(\gamma_1(e_0,e), \gamma_2(e_0,e))$ is the pair of strategies [40].

Definition 1. A pair of strategies (γ_1^*, γ_2^*) is a Nash equilibrium for the differential game, if for all strategies (γ_1, γ_2) the following inequalities hold,

$$J_1(\gamma_1^*, \gamma_2^*) \leq J_1(\gamma_1, \gamma_2^*) \tag{16}$$

$$J_2(\gamma_1^*, \gamma_2^*) \leq J_2(\gamma_1^*, \gamma_2) \tag{17}$$

Using the minimum principle [41], we define the Hamiltonians $H_1(e,u_1,u_2)$ and $H_2(e,u_1,u_2)$ of the control inputs u_1 and u_2, respectively, as

$$\begin{aligned} H_1 = \frac{1}{2}e^T Qe &+ u_1^T R_{11} u_1 + u_2^T R_{12} u_2 \\ &+ \lambda_1^T (Ae + B_1 u_1 + B_2 u_2) \end{aligned} \tag{18}$$

$$\begin{aligned} H_2 = \frac{1}{2}e^T Le &+ u_2^T R_{22} u_2 + u_1^T R_{21} u_1 \\ &+ \lambda_2^T (Ae + B_1 u_1 + B_2 u_2) \end{aligned} \tag{19}$$

Based on the results of [2,42] provided for this information structure and using the following Theorems, the feedback Nash solution for nonzero-sum differential game is obtained.

Theorem 1. *Let the strategies (γ_1^*, γ_2^*) be such that there exist* *solutions* (λ_1, λ_2) *to the following differential equations,*

$$\dot{\lambda}_1 = -\frac{\partial H_1}{\partial e}(e, \gamma_1^*, \gamma_2^*, \lambda_1) - \frac{\partial H_1}{\partial u_2}(e, \gamma_1^*, \gamma_2^*, \lambda_1) \times \frac{\partial \gamma_2^*}{\partial e}(e_0, e) \tag{20}$$

$$\dot{\lambda}_2 = -\frac{\partial H_2}{\partial e}(e, \gamma_1^*, \gamma_2^*, \lambda_2) - \frac{\partial H_2}{\partial u_1}(e, \gamma_1^*, \gamma_2^*, \lambda_2) \times \frac{\partial \gamma_1^*}{\partial e}(e_0, e) \tag{21}$$

where H_1 *and* H_2 *are defined in Eq.* (18) *and Eq.* (19)*, respectively, satisfying*

$$\frac{\partial H_i}{\partial u_i}(e, \gamma_1^*, \gamma_2^*, \lambda_i) = 0 \qquad i = 1, 2 \tag{22}$$

and e satisfies

$$\dot{e} = Ae + B_1\gamma_1^* + B_2\gamma_2^* \tag{23}$$

Then, (γ_1^*, γ_2^*) *is a Nash equilibrium with respect to the memoryless perfect state information structure, and the following equalities hold:*

$$u_1^* = \gamma_1^* = -R_{11}^{-1}B_1^T\lambda_1 \tag{24}$$

$$u_2^* = \gamma_2^* = -R_{22}^{-1}B_2^T\lambda_2 \tag{25}$$

Theorem 2. *Suppose* (P, S) *satisfy the coupled differential Riccati equations (DRE), given by*

$$\dot{P} = -A^TP - PA - Q + PB_1R_{11}^{-1}B_1^TP + PB_2R_{22}^{-1}B_2^TS + SB_2R_{22}^{-1}B_2^TP - SB_2R_{22}^{-1}R_{12}R_{22}^{-1}B_2^TS \tag{26}$$

$$\dot{S} = -A^TS - SA - L + SB_2R_{22}^{-1}B_2^TS + SB_1R_{11}^{-1}B_1^TP + PB_1R_{11}^{-1}B_1^TS - PB_1R_{11}^{-1}R_{21}R_{11}^{-1}B_1^TP \tag{27}$$

with the following boundary conditions,

$$P(T) = 0 \tag{28}$$

$$S(T) = 0 \tag{29}$$

Then, the following pair of strategy

$$(\gamma^{*1}, \gamma^{*2}) = \left(-R_{11}^{-1}B_1^TP(t)e, -R_{22}^{-1}B_2^TS(t)e\right) \tag{30}$$

is a feedback Nash equilibrium law, and the solutions to the equations of (20) *and* (21) *are as follows,*

$$\lambda_1 = Pe \tag{31}$$

$$\lambda_2 = Se \tag{32}$$

We can simultaneously solve differential Riccati equations (DRE) defined in (26) and (27) using boundary conditions (28) and (29). Substituting Eqs. (31) and (32) into Eqs. (24) and (25), respectively, yields the following Nash-based controllers for two players,

$$u_1^* = -R_{11}^{-1}B_1^TPe \tag{33}$$

$$u_2^* = -R_{22}^{-1}B_2^TSe \tag{34}$$

Based on the feedback Nash strategy, the cost functions defined in Eqs. (13) and (14) are minimized by the control inputs Eqs. (33) and (34), respectively.

4 Experimental Results

We implement the two-player Nash-based feedback controller for the 7-DOF Baxter manipulator through a pick-and-place task, while each player tries to minimize its own cost function. We take the advantage of this controller, using the Theorems 1 and 2, in order to globally asymptotically stabilize the manipulator due to the fact that all the assumptions are valid for the robot's dynamics. Then, we thoroughly investigate the performance of this controller through simulations and experiments.

The initial conditions are selected based on the accuracy of the joints' sensors,

$$q_0 = q_{d0} + 0.05[\text{rand}(0,1), \cdots, \text{rand}(0,1)]^T$$
$$\dot{q}_0 = 0_{7\times 1}$$

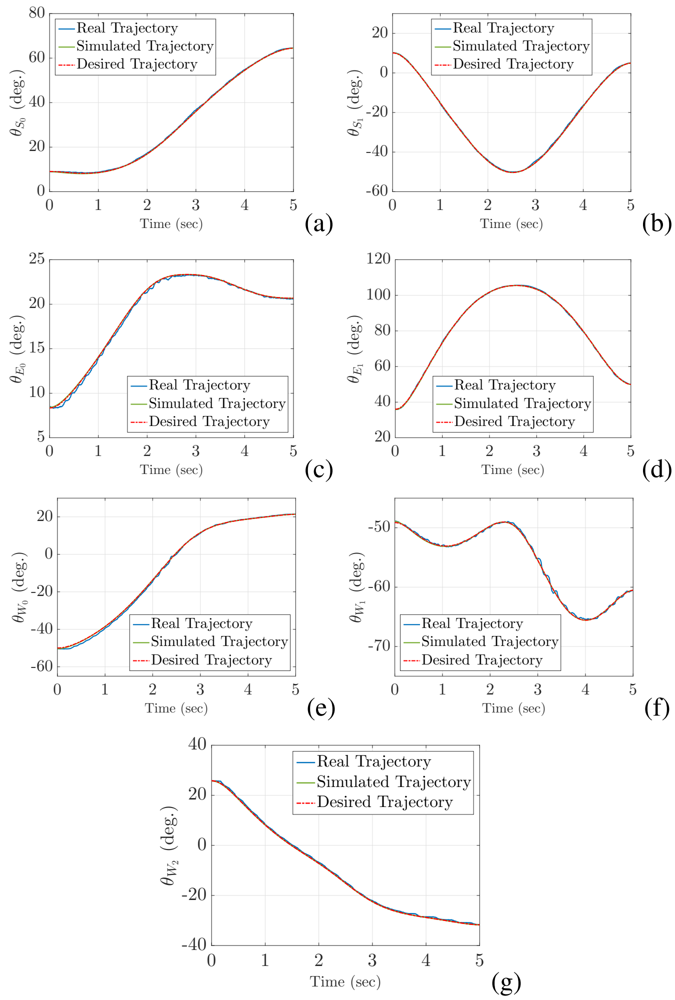

FIGURE 2. The experimental (blue line), simulated (green line), and desired joint-space trajectories (red dash line) for (a) S_0, (b) S_1, (c) E_0, (d) E_1, (e) W_0, (f) W_1, and (g) W_2 joints

The weighting matrices L and Q, and the Nash gains R_{11}, R_{12}, R_{21}, and R_{22} are selected as follows:

$$Q_{11} = \text{diag}\{7.0, 9.0, 7.0, 10.0, 5.0, 8.0, 5.0\}$$
$$Q_{12} = 6 \times I_{7\times7}, \qquad Q_{22} = 0.5 \times Q_{11}$$
$$L_{11} = \text{diag}\{7.0, 9.0, 70, 15, 25, 80, 25\}$$
$$L_{12} = -6 \times I_{7\times7}, \qquad L_{21} = 10 \times I_{7\times7}$$
$$R_{11} = 10 \times I_{7\times7}, \quad R_{12} = R_{21} = I_{7\times7}, \quad R_{22} = 7.5 \times I_{7\times7}$$

Figs. 2 and 3 present the experimental and simulated joint-space trajectories and tracking errors, respectively. As can be observed, the simulation results reveal that the manipulator perfectly tracks the desired trajectories, while the experimental ones present a highly acceptable tracking process. The negligible experimental tracking errors mainly root on the inaccuracy of sensors and actuators.

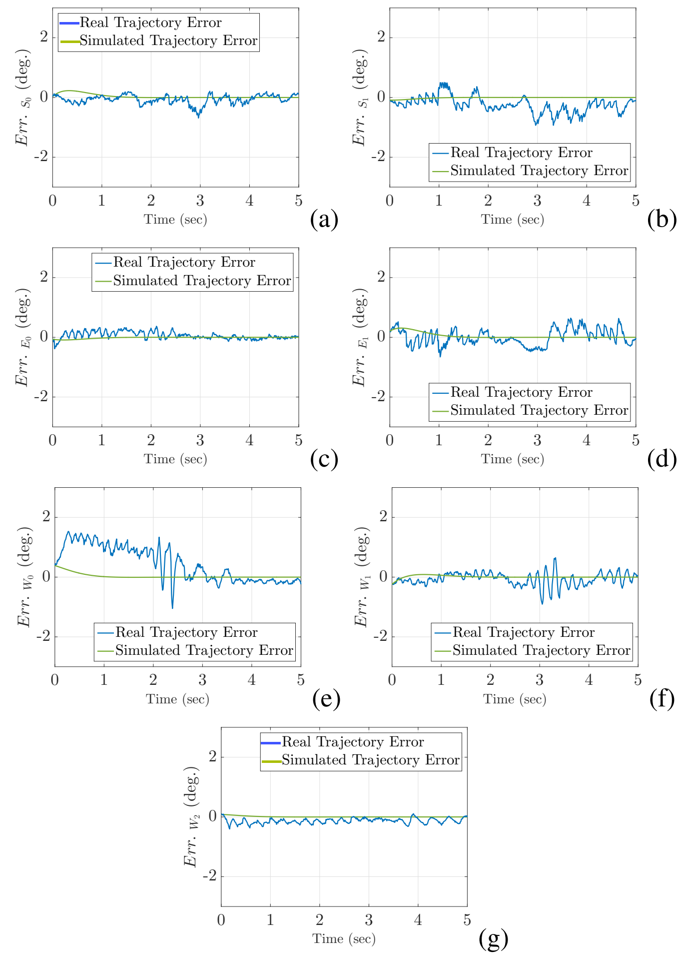

FIGURE 3. The experimental (blue line) and simulated (green line) Nash-based tracking errors for (a) S_0, (b) S_1, (c) E_0, (d) E_1, (e) W_0, (f) W_1, and (g) W_2 joints

The simulation results, shown in Fig. 3, indicate that the tracking errors asymptotically converge to zero, as expected. However, the tracking errors of the experimental work do not necessarily converge to zero due to the lack of sufficient accuracies of sensors and actuators as well as the joints' backlash.

The joints' torques, shown in Fig. 4, reveal that the incremental tracking errors expectedly demand more control torques to be applied. Therefore, it is straightforward to conclude that the experimental torques are higher than those of the simulated ones, since the experimental tracking errors are higher than the simulated ones, as shown in Fig. 3. It is worth mentioning that, in addition to the control torques, the gravity compensation torques need to be applied in order to overcome the effect of gravity; this is a basic mode which is, by default, active for the onset of

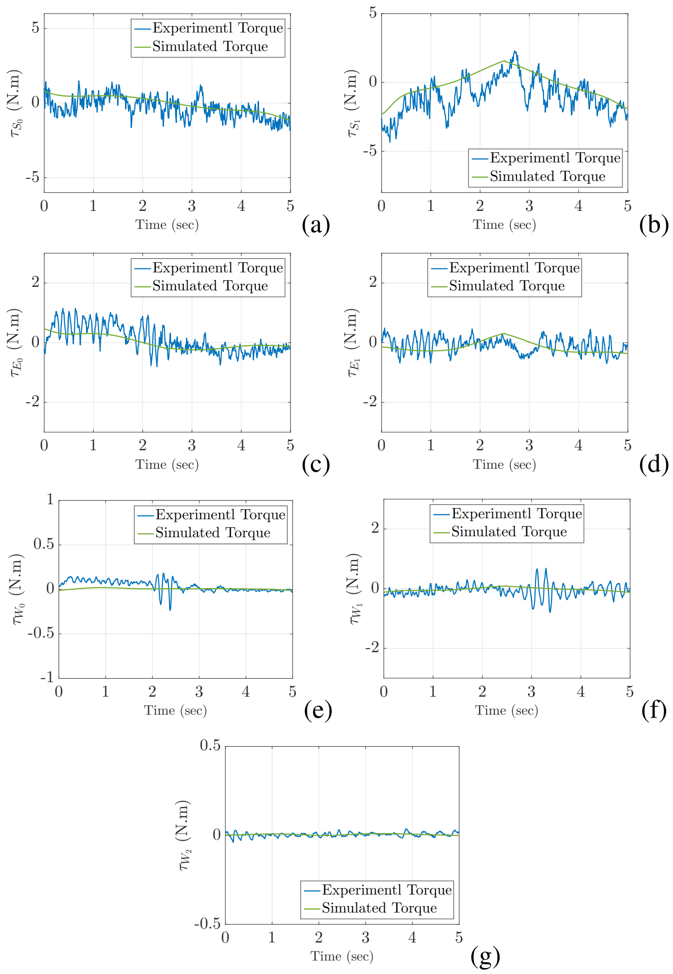

FIGURE 4. The experimental (blue line) and simulated (green line) Nash-based torques for (a) S_0, (b) S_1, (c) E_0, (d) E_1, (e) W_0, (f) W_1, and (g) W_2 joints

manipulator operation.

The simulated and experimental minimization and convergence processes of the cost functions for the two players are shown in Figs. 5 and 6, respectively. Shown in Fig. 5 reveals that both the cost functions asymptotically converge to the optimal values, as expected, although the experimental ones do not converge to optimal values perfectly (Fig. 6). These happen due to the fact that the experimental tracking process cannot be perfectly achieved, while the errors do not converge to absolute zero. However, the robot manipulator is stabilized using the Nash-based feedback control law and has an acceptable tracking process. Fig. 7 presents the experimental work carried out at our Dynamic Systems and Control Laboratory (DSCL) to examine the Nash-based control law for a simple obstacle avoidance pick-and-place task defined.

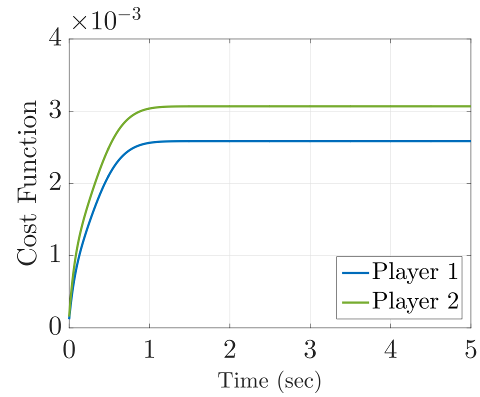

FIGURE 5. The simulated cost functions for u_1 and u_2

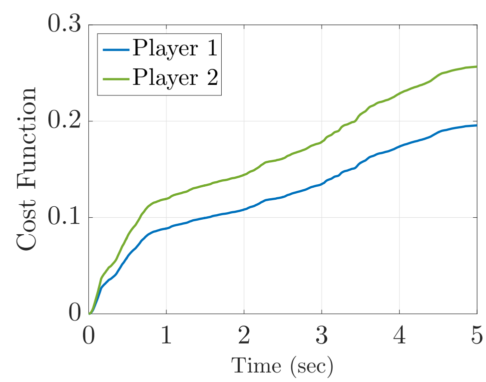

FIGURE 6. The experimental cost functions for u_1 and u_2

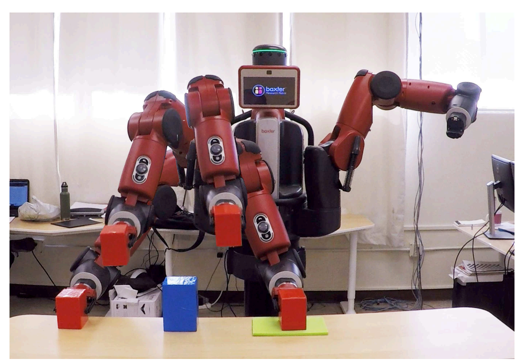

FIGURE 7. A stable obstacle avoidance pick-and-place task using the Nash-based feedback control law

5 Conclusions

Throughout this effort, we presented the formulation of the two-player Nash-based feedback control law for an Euler-Lagrangian system, and then the controller was experimentally implemented for the 7-DOF Baxter manipulator. Toward formulating the controller, we assumed and then validated some properties for the robot's operation. We formulated the Nash-based feedback controller using the Theorems 1 and 2, and then investigated its performance. The experimental results revealed the stable operation of the manipulator, and the robot could expectedly track the desired joint-space trajectories.

We also presented that the simulated tracking errors, using the Nash-based feedback controller, asymptotically converge to zero guaranteed through the Theorems 1 and 2. Although, the experimental results revealed an acceptable tracking process, which is due to the inaccuracy involved with sensors and actuators in addition to the joints' backlash. The simulated and experimental torques were shown for all joints along with both the players' cost functions. The simulation results indicate the asymptotic convergence of the cost functions of players. However, since the errors of the experimental work did not converge to absolute zero, the experimental torques and cost functions did not converge to $M(q)h$ and optimal values, respectively. Note that the experiment revealed the stable operation of the manipulator, while the robot tracked the desired joint-space trajectory in an acceptable fashion.

ACKNOWLEDGMENT

This article is based upon work supported by the National Science Foundation under Award #1823951. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

The DSCL website for the experimental work: http://peimannm.sdsu.edu

References

[1] Bagheri, M., Ajoudani, A., Lee, J., Caldwell, D. G., and Tsagarakis, N. G., 2015. "Kinematic analysis and design considerations for optimal base frame arrangement of humanoid shoulders". In 2015 IEEE Int. Conf. on Robotics and Automation (ICRA), pp. 2710–2715.

[2] Basar, T., and Olsder, G. J., 1999. *Dynamic noncooperative game theory*, Vol. 23. Siam.

[3] Tijs, S., 2003. *Introduction to game theory*. Springer.

[4] Johnson, M., Bhasin, S., and Dixon, W. E., 2011. "Nonlinear two-player zero-sum game approximate solution using a policy iteration algorithm". In Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on, IEEE, pp. 142–147.

[5] Nash, J., 1951. "Non-cooperative games". *Annals of mathematics*, pp. 286–295.

[6] Isaacs, R., 1965. *Differential games: a mathematical theory with applications to warfare and pursuit, control and optimization*. John Wiley.

[7] Ji, X., Yang, K., Na, X., Lv, C., Liu, Y., and Liu, Y., 2019. "Feedback game-based shared control scheme design for emergency collision avoidance: A fuzzy-linear quadratic regulator approach". *Journal of Dynamic Systems, Measurement, and Control,* **141**(8), mar.

[8] Lee, D., Periaux, J., and Gonzalez, L. F., 2010. "UAS mission path planning system (MPPS) using hybrid-game coupled to multi-objective optimizer". *Journal of Dynamic Systems, Measurement, and Control,* **132**(4), jun.

[9] Quintana, R., and Patino, D., 2018. "A game-theoretical perspective for decentralized active noise control". *Journal of Vibration and Control,* **24**(21), pp. 5062–5071.

[10] Bhattacharya, S., and Hutchinson, S., 2009. "On the existence of nash equilibrium for a two player pursuit-evasion game with visibility constraints". In *Algorithmic Foundation of Robotics VIII*. Springer, pp. 251–265.

[11] LaValle, S. M., and Hutchinson, S. A., 1994. "Path selection and coordination for multiple robots via nash equilibria". In Proceedings of the 1994 IEEE International Conference on Robotics and Automation, IEEE, pp. 1847–1852.

[12] Li, Y., Tee, K. P., Yan, R., Chan, W. L., and Wu, Y., 2016. "A framework of human–robot coordination based on game theory and policy iteration". *IEEE Transactions on Robotics,* **32**(6), pp. 1408–1418.

[13] Sharma, R., and Gopal, M., 2014. "Fictitious play based markov game control for robotic arm manipulator". In Proceedings of 3rd International Conference on Reliability, Infocom Technologies and Optimization, IEEE, pp. 1–6.

[14] Eslamipoor, R., Nobari, A., and Nobari, O., 2019. "Optimal point-to-point path planning of manipulator by using vibration damping optimization algorithm and game theory method". *Journal of Testing and Evaluation,* **47**(4).

[15] Bugnon, F. J., and Mohler, R. R., 1988. "Structural choices of 2 versus 1 pursuit-evasion differential games". *Journal of Dynamic Systems, Measurement, and Control,* **110**(2), jun, pp. 160–167.

[16] Wang, X., Yaz, E. E., and Schneider, S. C., 2018. "Coupled state-dependent riccati equation control for continuous time nonlinear mechatronics systems". *Journal of Dynamic Systems, Measurement, and Control,* **140**(11), jun.

[17] Dolezal, J., 1979. "Parameter optimization for two-player zero-sum differential games". *Journal of Dynamic Systems, Measurement, and Control,* **101**(4), dec, pp. 345–349.

[18] Kamalapurkar, R., Klotz, J. R., and Dixon, W. E., 2014. "Concurrent learning-based approximate feedback-nash equilibrium solution of n-player nonzero-sum differential games". *IEEE/CAA journal of Automatica Sinica,* **1**(3), pp. 239–247.

[19] Kamalapurkar, R., Klotz, J., and Dixon, W. E., 2014. "Model-based reinforcement learning for on-line feedback-nash equilibrium solution of n-player nonzero-sum differential games". In American Control Conference (ACC), 2014, IEEE, pp. 3000–3005.

[20] Dixon, W., 2014. Optimal adaptive control and differential games by reinforcement learning principles.

[21] Kamalapurkar, R., Klotz, J. R., Walters, P., and Dixon, W. E., 2018. "Model-based reinforcement learning in differential graphical games". *IEEE Transactions on Control of Network Systems,* **5**(1), pp. 423–433.

[22] Smyrnakis, M., and Veres, S. M., 2014. "Coordination of control in robot teams using game-theoretic learning". *IFAC Proceedings Volumes,* **47**(3), pp. 1194–1202.

[23] Meng, Y., 2008. "Multi-robot searching using game-theory based approach". *International Journal of Advanced Robotic Systems,* **5**(4), p. 44.

[24] Korayem, M., Esfandiar, H., and Dargahi, R., 2016. "Determining load carrying capacity of a manipulator by game theory: Closed-loop nonzero-sum differential game approach". *Journal of Mechanical Science and Technology,* **30**(11), pp. 5197–5205.

[25] Li, S., He, J., Li, Y., and Rafique, M. U., 2016. "Distributed recurrent neural networks for cooperative control of manipulators: A game-theoretic perspective". *IEEE transactions on neural networks and learning systems,* **28**(2), pp. 415–426.

[26] Sharma, R., 2016. "Lyapunov theory based stable markov game fuzzy control for non-linear systems". *Engineering Applications of Artificial Intelligence,* **55**, pp. 119–127.

[27] Xu, P., Groff, R. E., and Burg, T., 2012. "Dynamics and control of the shoot-the-moon tabletop game". *Journal of Dynamic Systems, Measurement, and Control,* **134**(5), jun.

[28] Arslan, G., Marden, J. R., and Shamma, J. S., 2007. "Autonomous vehicle-target assignment: A game-theoretical formulation". *Journal of Dynamic Systems, Measurement, and Control,* **129**(5), apr, pp. 584–596.

[29] Jumarie, G., 1977. "Coordination concepts in multilevel hierarchical differential games". *Journal of Dynamic Systems, Measurement, and Control,* **99**(3), sep, pp. 201–208.

[30] Ma, W.-H., and Peng, H., 1999. "Solutions of continuous-time preview two-player differential games". *Journal of Dynamic Systems, Measurement, and Control,* **121**(2), jun, pp. 326–331.

[31] Bagheri, M., Naseradinmousavi, P., and Krstić, M., 2019. "Feedback linearization based predictor for time delay control of a high-dof robot manipulator". *Automatica,* **108**, p. 108485.

[32] Bagheri, M., Krstić, M., and Naseradinmousavi, P., 2018. "Joint-space trajectory optimization of a 7-dof baxter using multivariable extremum seeking". In IEEE American Control Conference (ACC 2018), pp. 2176–2181.

[33] Bagheri, M., Naseradinmousavi, P., and Morsi, R., 2017. "Experimental and novel analytical trajectory optimization of a 7-dof baxter robot: Global design sensitivity and step size analyses". In ASME 2017 Dynamic Systems and Control Conference, American Society of Mechanical Engineers, American Society of Mechanical Engineers, p. V001T30A001.

[34] Bagheri, M., 2019. "Adaptive and delay-compensating robot controllers". PhD thesis, UC San Diego and SDSU.

[35] Bagheri, M., Krstić, M., and Naseradinmousavi, P., 2018. "Multivariable extremum seeking for joint-space trajectory optimization of a high-degrees-of-freedom robot". *Journal of Dynamic Systems, Measurement, and Control,* **140**(11), p. 111017.

[36] Bertino, A., Bagheri, M., Naseradinmousavi, P., and Krstić, M., 2019. "Experimental autonomous deep learning-based 3d path planning for a 7-dof robot manipulator". In ASME 2019 Dynamic Systems and Control Conference, American Society of Mechanical Engineers.

[37] Bagheri, M., Naseradinmousavi, P., and Krstić, M., 2019. "Time delay control of a high-dof robot manipulator through feedback linearization based predictor". In ASME 2019 Dynamic Systems and Control Conference, American Society of Mechanical Engineers.

[38] Bagheri, M., Krstić, M., and Naseradinmousavi, P., 2018. "Analytical and experimental predictor-based time delay control of baxter robot". In ASME 2018 Dynamic Systems and Control Conference, American Society of Mechanical Engineers, pp. V001T04A011–V001T04A011.

[39] Bagheri, M., and Naseradinmousavi, P., 2017. "Novel analytical and experimental trajectory optimization of a 7-dof baxter robot: global design sensitivity and step size analyses". *The International Journal of Advanced Manufacturing Technology,* **93**(9-12), December, pp. 4153–4167.

[40] Johnson, M. A., 2011. *Differential game-based control methods for uncertain continuous-time nonlinear systems*. Citeseer.

[41] Kirk, D. E., 2012. *Optimal control theory: an introduction*. Courier Corporation.

[42] Weeren, A., Schumacher, J., and Engwerda, J. C., 1999. "Asymptotic analysis of linear feedback nash equilibria in nonzero-sum linear-quadratic differential games". *Journal of Optimization Theory and Applications,* **101**(3), pp. 693–722.