Article

Journal of Vibration and Control
2021, Vol. 0(0) 1–12

DOI: 10.1177/1077546320982447
journals.sagepub.com/home/jvc
SAGE

Analytical and experimental nonzero-sum differential game-based control of a 7-DOF robotic manipulator

Mostafa Bagheri[1] and Peiman Naseradinmousavi[2]

Abstract

We formulate a Nash-based feedback control law for an Euler–Lagrange system to yield a solution to noncooperative differential game. The robot manipulators are broadly used 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 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.

Keywords

Nonzero-sum differential game, Nash equilibrium, control, high-DOF robot

1. Introduction

The robot manipulators are broadly used in industrial units on the account of their reliable, fast, and precise motions Bagheri et al. (2015), while consuming a significant lumped amount of energy. Therefore, optimization schemes need to be used 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 "noncooperative" game Basar and Olsder (1999), each player pursues its own individual interests, which may result in conflict with the other ones. Note that a solution to the noncooperative game is an equilibrium because 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 Tijs (2003) 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 their own cost function. The control of a high-DOF robot manipulator, with N-player, is an example of nonzero-sum noncooperative differential game. In this problem, each player has their 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 Johnson et al. (2011) is a mathematical representation of a situation in

[1]Chief Technology Officer (CTO) at NXT Robotics, USA
[2]Dynamic Systems and Control Laboratory, San Diego State University, USA

Received: 5 August 2020; accepted: 28 November 2020

Corresponding author:
Peiman Naseradinmousavi, Department of Mechanical Engineering, San Diego State University, 5500 Campanile Drive, E-323C, San Diego 92182, California, USA.
Email: pnaseradinmousavi@sdsu.edu

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 because 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. Nash (1951) provided a solution for a class of general N-player noncooperative games. In 1960s, Isaacs (1965) investigated the multiplayer extension of the dynamic programming solution for the differential games. Ji et al. (2019) proposed a fuzzy linear quadratic regulator game-based control scheme to simultaneously enhance vehicle stability while compensating driver's inappropriate steering reaction in emergency avoidance. Lee et al. (2010) presented the application of advanced optimization techniques to unmanned aerial system mission path planning system (MPPS) using multi-objective evolutionary algorithms. Quintana and Patino (2018) studied a decentralized scheme for active noise control from a game theoretical perspective. They formalized the Nash equilibrium (i.e., the simultaneous best strategy) in the interaction between the controllers. Bhattacharya and Hutchinson (2009) 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. LaValle and Hutchinson (1994) investigated a method to analyze and select time-optimal coordination strategies for n robots whose configurations are constrained to lie on a C-space road map (which could, for instance, represent a Voronoi diagram). Li et al. (2016a, 2016b) proposed a framework to analyze the interactive behaviors of human and robot in physical interactions. They used game theory in describing the system under study, and policy iteration was adopted to provide a solution of Nash equilibrium. Sharma and Gopal (2014) 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. Eslamipoor et al. (2019) used 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. Bugnon and Mohler (1988) 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. Wang et al. (2018) considered a novel coupled state-dependent Riccati equation approach for systematically designing nonlinear quadratic regulator and H_∞ control of mechatronics systems. Doležǎl (1979) 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.

Shafei and Korayem (2017) formulated an open-loop optimal control based on the Pontryagin minimum principle, which yields a 2-point boundary value problem. They used the indirect method to extract the optimality conditions. They calculated the maximum allowed load that a mechanical manipulator with flexible links can carry while traversing an optimal path. Shafei and Shafei (2018) formulated a systematic procedure for the dynamic modeling of a closed-chain robotic system in both the flight and impact phases. In another effort, Korayem et al. (2013) represented a systematic algorithm capable of deriving the equations of motion of N-flexible link manipulators with revolute–prismatic joints. In this study, the links are modeled based on the Euler–Bernoulli beam theory and the assumed mode method. Also, the effects of gravity as well as the longitudinal, transversal, and torsional vibrations have been considered in the formulations. Other efforts for modeling and control of flexible-link robots can be found in Bae and Haug (1987), Saha and Schiehlen (2001), Naudet et al. (2003), Korayem et al. (2011) used the open-loop optimal control approach for generating the optimal trajectory of the flexible mobile manipulators in point-to-point motion, whereas we, here, have the desired end-effector trajectory from our previous deep learning-based path planning Bertino et al. (2019), and therefore, the Nash-based feedback control law is formulated and then experimentally implemented to track the desirable trajectory.

Li et al. (2020a, 2020b) proposed a novel fault-tolerant method with simultaneous fault diagnosis function for motion planning and control of industrial redundant manipulators. The proposed approach is able to adaptively localize which joints run away from the normal state to be fault, and it can guarantee to finish the desired path tracking control even when these fault joints lose their velocity to actuate. Xie et al. (2012) formulated a new nested RRT algorithm to fulfill fault tolerance and some other secondary tasks at the same time. Compared with other existing fault tolerant algorithms, they showed that this new algorithm is more efficient. Jin et al. (2017) reformulated a quadratic program with equality and bound constraints, which is then solved by a discrete-time recurrent neural network. Cheng et al. (2020) proposed a motion planning method based on beetle antennae search algorithm for motion planning of redundant manipulators with the variable joint velocity limit. Li et al. (2020a, 2020b) formulated a novel motion planning strategy with minimal potential energy modulation. Such motion resolution scheme is formulated as an optimization problem and solved by the zeroing dynamics to achieve elegant global convergence. More case studies can be found in Kamalapurkar et al. (2014b,2014a), Dixon (2014), Kamalapurkar et al. (2018), Smyrnakis and Veres (2014), Meng (2008), Korayem et al. (2016), Li et al. (2016a,

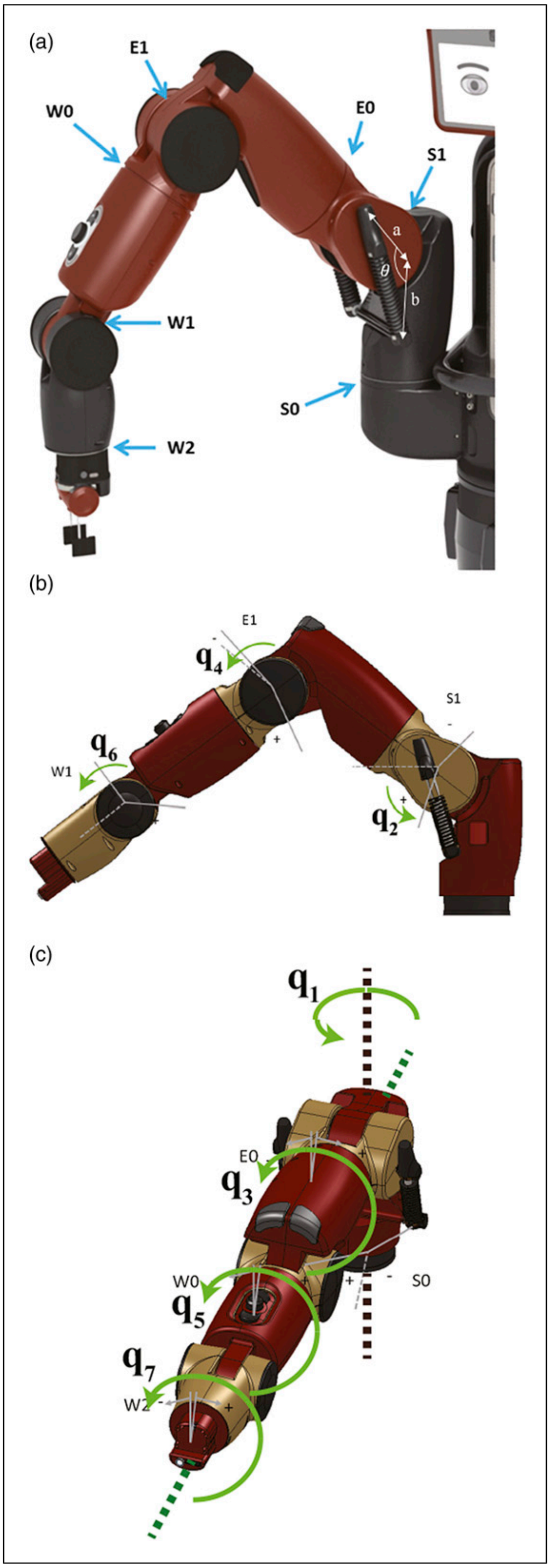

Figure 1. 7-DOF Baxter's arm: (a) Joints' configurations; (b) sagittal view; (c) top view.

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

2016b), Sharma (2016), Xu et al. (2012), Arslan et al. (2007), Jumarie (1977), Ma and Peng (1999).

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, to follow the desired joint-space trajectories. Therefore, we use an approach to solve the nonzero-sum noncooperative 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 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 Figure 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 Bagheri et al. (2019a, 2019b), Bagheri et al. (2018b, 2017), Bagheri (2019), Bagheri et al. (2018c), Bertino et al. (2019), Bagheri et al. (2019a, 2019b), Bagheri et al. (2018a), Bagheri and Naseradinmousavi (2017), Bagheri et al. (2021). 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 noncooperative 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, 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:

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 $\alpha \in \mathbb{R}^{7\times7}$ is a constant positive definite matrix. A state-space model can be developed based on the tracking errors of equations (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 equation (3)

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

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

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

Because the dynamics of system (1) is known, the controller, based on equation (5), is designed as

$$\tau_{7\times1} = M(h - (u_1 + u_2 + \ldots + u_N)) \tag{8}$$

where N is the number of players and $u_1(t), \ldots, u_N \in \mathbb{R}^7$ 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 equation (8) into (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 $A \in \mathbb{R}^{14\times14}$ and $B_i \in \mathbb{R}^{14\times7} (i = 1, 2)$ are defined as

$$A = \begin{bmatrix} -\alpha & I_{7\times7} \\ -\alpha^2 & \alpha \end{bmatrix} \tag{11}$$

$$B_i = [0_{7\times7} \quad I_{7\times7}]^T \quad i = 1, 2 \tag{12}$$

where $I_{7\times7}$ and $0_{7\times7}$ are the 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

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

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

where $t_0 \in \mathbb{R}$ is the initial time and $Q, L \in \mathbb{R}^{14\times14}$ 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 Johnson (2011).

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

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

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

Using the minimum principle Kirk (2012), 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

$$H_1 = \frac{1}{2}e^T Qe + u_1^T R_{11}u_1 + u_2^T R_{12}u_2 + \lambda_1^T(Ae + B_1u_1 + B_2u_2) \tag{18}$$

$$H_2 = \frac{1}{2}e^T Le + u_2^T R_{22}u_2 + u_1^T R_{21}u_1 + \lambda_2^T(Ae + B_1u_1 + B_2u_2) \tag{19}$$

Based on the results of Basar and Olsder (1999), Weeren et al. (1999) 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 equations (18) and (19), respectively, satisfying

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

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 (DREs), given by

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

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

with the following boundary conditions

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

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

Then, the following pair of strategy

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

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 DREs defined in (26) and (27) using boundary conditions (28) and (29). Substituting equations (31) and (32) into (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 equations (13) and (14) are minimized by the control inputs' equations (33) and (34), respectively.

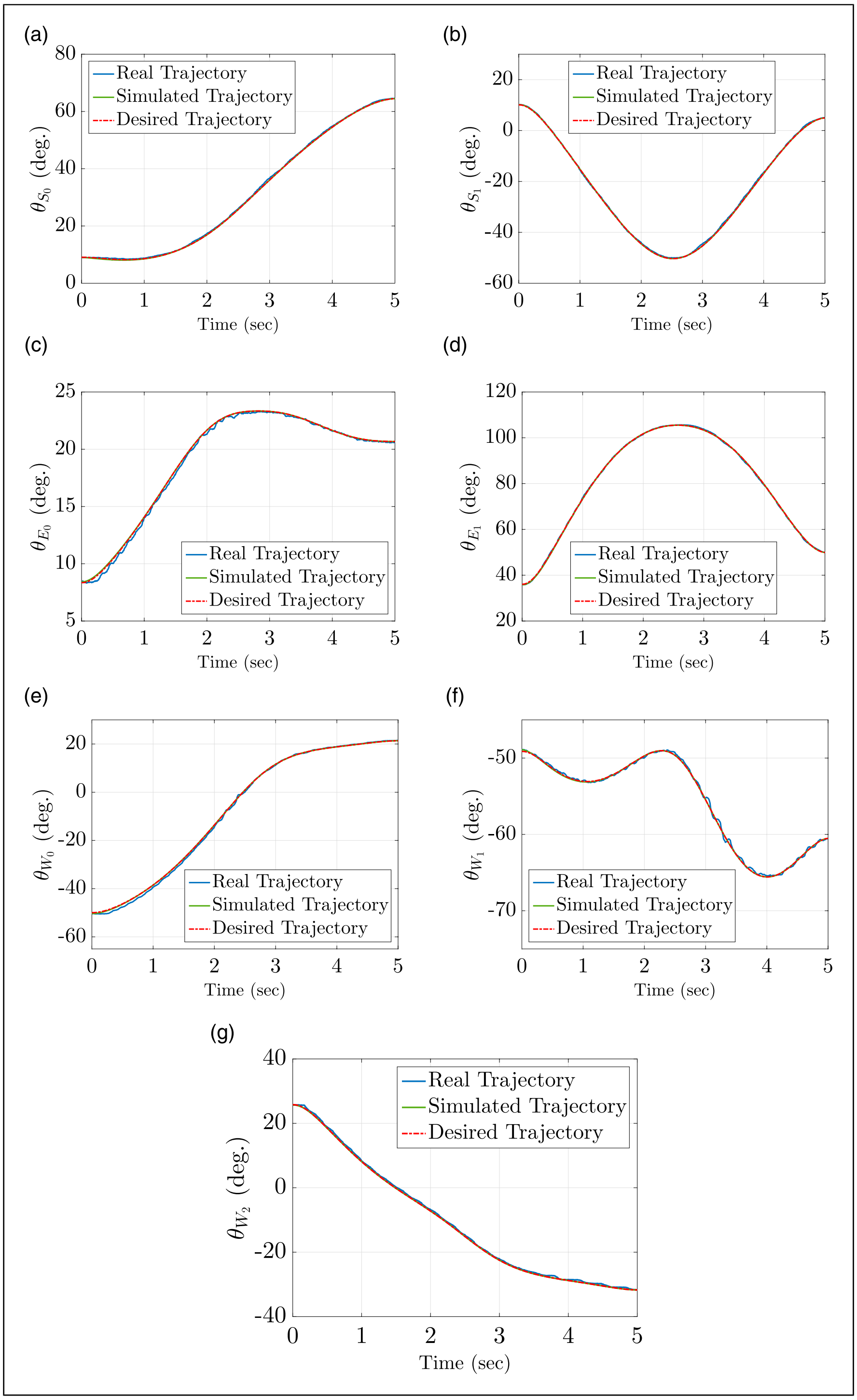

Figure 2. Experimental (blue line), simulated (green line), and desired joint-space trajectories (dash line) for (a) S_0, (b) S_1, (c) E_0, (d) E_1, (e) W_0, (f) W_1, and (g) W_2 joints.

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 Theorems 1 and 2, to globally asymptotically stabilize the manipulator because of 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.

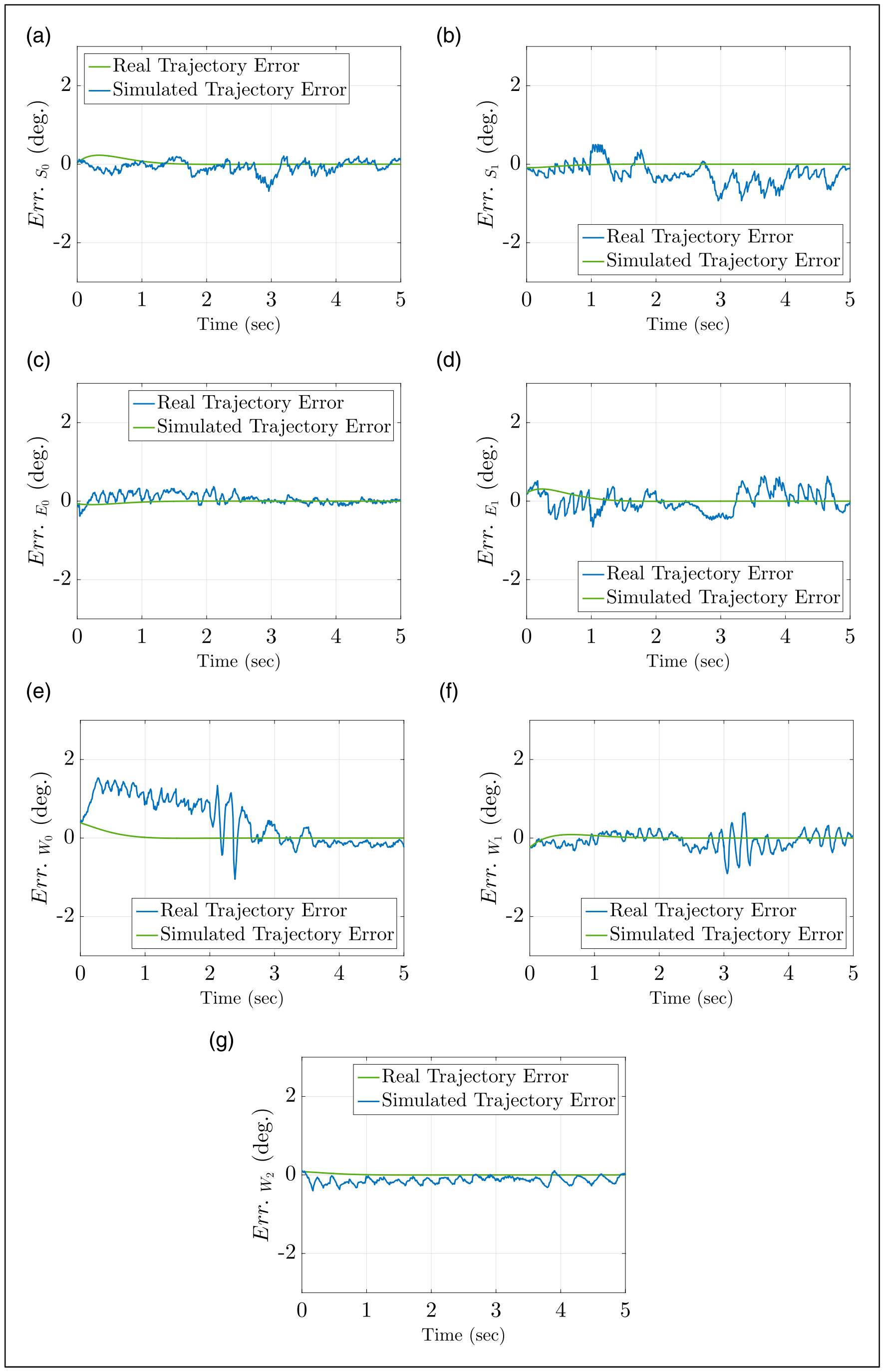

Figure 3. Experimental (blue line) and simulated (green line) Nash-based tracking errors for (a) S_0, (b) S_1, (c) E_0, (d) E_1, (e) W_0, (f) W_1, and (g) W_2 joints.

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

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

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

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

The immediate challenge of experimental work emerged as the computation time of the optimal control scheme in

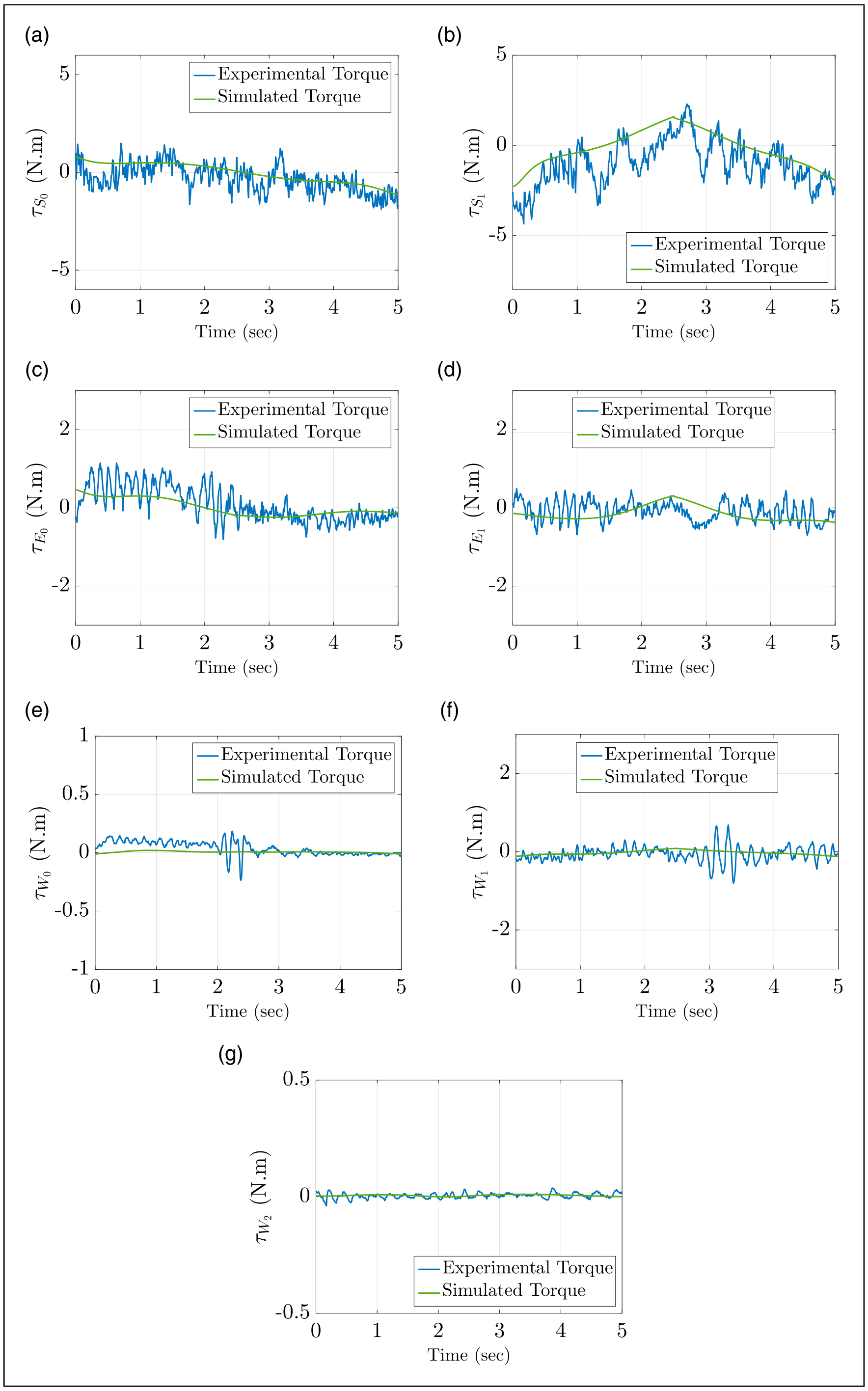

Figure 4. Experimental (blue line) and simulated (green line) Nash-based torques for (a) S_0, (b) S_1, (c) E_0, (d) E_1, (e) W_0, (f) W_1, and (g) W_2 joints.

each loop, which was incompatible with the minimum time step ($\Delta t = 0.001$ s or $f = 1$ kHZ) of Baxter. The computation time of the optimal control scheme was in the range of $0.003\ s \leq t_c \leq 0.005\ s$ leading to a destabilizing time delay in each control loop. Therefore, we addressed a critical trade-off between the high accuracy and computational cost. To resolve this problem, we increased the Baxter's time step to $\Delta t = 0.01$ s or $f = 0.1$ kHZ to avoid such a harmful time delay.

Figures 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

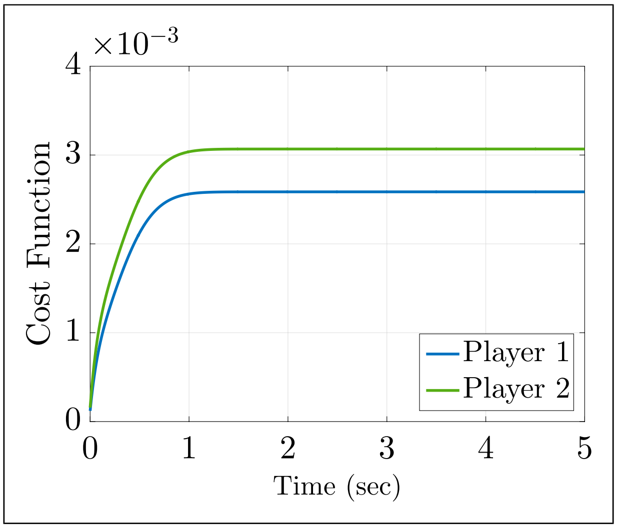

Figure 5. Simulated cost functions for u_1 and u_2.

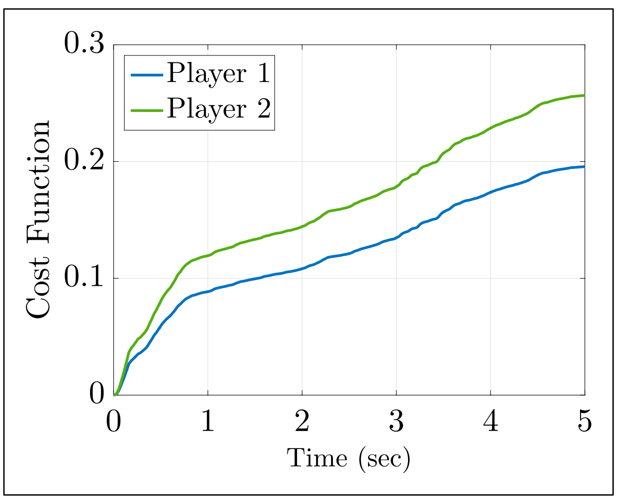

Figure 6. Experimental cost functions for u_1 and u_2.

manipulator perfectly tracks the desired trajectories, whereas the experimental ones present a highly acceptable tracking process. The negligible experimental tracking errors mainly root on the inaccuracy of sensors and actuators.

The simulation results, shown in Figure 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 because of the lack of sufficient accuracies of sensors and actuators, measurement noise, as well as the joints' backlash. From another aspect, the measurement of discrete time control, for Baxter, is subject to noise. The noise affects the joint position/velocity sensors. The joint angular positions are measured directly through the use of encoders, whereas the joint angular velocities are calculated by taking the derivative of the angular positions (finite-difference approximation). This amplifies the noise in the angular position

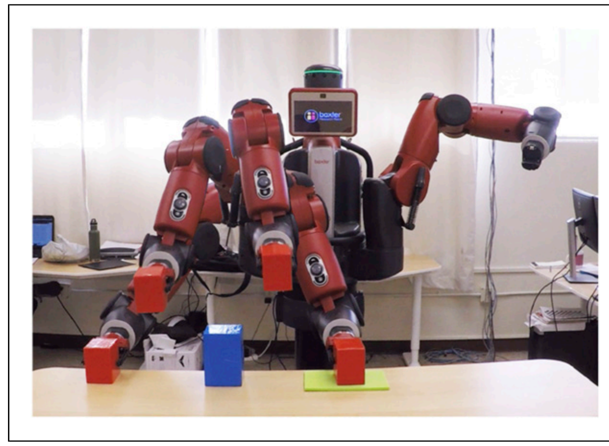

Figure 7. Stable obstacle avoidance pick-and-place task using the Nash-based feedback control law.

measurement, as well as introducing new noise due to variations in sampling time as well as small uncertainties in angular position based on the resolution of the encoder. This noise in the angular velocity obviously affects the computation of the joint torques. The joints' torques, shown in Figure 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 because the experimental tracking errors are higher than the simulated ones, as shown in Figure 3. It is worth mentioning that in addition to the control torques, the gravity compensation torques need to be applied to overcome the effect of gravity; this is a basic mode which is, by default, active for the onset of manipulator operation.

The simulated and experimental minimization and convergence processes of the cost functions for the two players are shown in Figures 5 and 6, respectively. Shown in Figure 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 (Figure 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. Figure 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.

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 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 Theorems 1 and 2. Although, the experimental results revealed an acceptable tracking process, which is because of the inaccuracy involved with sensors and actuators, measurement noise, and 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, because 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.

Based on efforts reported in Zhou et al. (2015), Liu et al. (2016), Nammoto and Kosuge (2012), we will investigate the joints' limit avoidance as our future study.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: 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. Visit the DSCL website for the experimental work: http://peimannm.sdsu.edu.

ORCID iD

Peiman Naseradinmousavi https://orcid.org/0000-0002-0724-4070

References

Arslan G, Marden JR and Shamma JS (2007) Autonomous vehicle-target assignment: a game-theoretical formulation. *Journal of Dynamic Systems, Measurement, and Control* 129(5): 584–596.

Bae D-S and Haug EJ (1987) A recursive formulation for constrained mechanical system dynamics: part I. Open loop systems. *Mechanics of Structures and Machines* 15(3): 359–382.

Bagheri M (2019) *Adaptive and delay-compensating robot controllers*. PhD Thesis, UC San Diego and SDSU.

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): 4153–4167.

Bagheri M, Ajoudani A, Lee J, et al. (2015) Kinematic analysis and design considerations for optimal base frame arrangement of humanoid shoulders. In: 2015 IEEE international conference on robotics and automation (ICRA), WA, USA, 26–30 May 2015, pp. 2710–2715. IEEE.

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, VA, USA, 13–11 October 2017. The American Society of Mechanical Engineers.

Bagheri M, Krstić M and Naseradinmousavi P (2018a) Analytical and experimental predictor-based time delay control of baxter robot. In: ASME 2018 dynamic systems and control conference, GA, USA, 30 September–3 October 2018. The American Society of Mechanical Engineers.

Bagheri M, Krstić M and Naseradinmousavi P (2018b) Joint-space trajectory optimization of a 7-dof baxter using multivariable extremum seeking. In: 2018 IEEE american control conference (ACC). WI, USA, 27–29 June 2018, pp. 2176–2181. IEEE

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

Bagheri M, Naseradinmousavi P and Krstić M (2019a) Feedback linearization based predictor for time delay control of a high-dof robot manipulator. *Automatica* 108: 108485.

Bagheri M, Naseradinmousavi P and Krstić M (2019b) Time delay control of a high-dof robot manipulator through feedback linearization based predictor. In: ASME 2019 dynamic systems and control conference, UT, USA, 8–11 October 2019. The American Society of Mechanical Engineers.

Bagheri M, Karafyllis I, Naseradinmousavi P, et al. (2021) Adaptive control of a two-link robot using batch least-square identifier. *IEEE/CAA Journal Automatica Sinica* 8(1): 86–93.

Basar T and Olsder GJ (1999) *Dynamic Noncooperative Game Theory*. Siam.

Bertino A, Bagheri M, Naseradinmousavi P, et al. (2019) Experimental autonomous deep learning-based 3d path planning for a 7-dof robot manipulator. In: ASME 2019 dynamic systems and control conference, UT, USA, 8–11 October 2019. The American Society of Mechanical Engineers.

Bhattacharya S and Hutchinson S (2009) On the existence of nash equilibrium for a two player pursuit-evasion game with visibility constraints. In: Chirikjian GS, Choset H, Morales M, et al. (eds). *Algorithmic Foundation of Robotics VIII. Springer Tracts in Advanced Robotics*. Springer, 251–265.

Bugnon FJ and Mohler RR (1988) Structural choices of 2 versus 1 pursuit-evasion differential games. *Journal of Dynamic Systems, Measurement, and Control* 110(2): 160–167.

Cheng Y, Li C, Li S, et al. (2020) Motion planning of redundant manipulator with variable joint velocity limit based on beetle antennae search algorithm. *IEEE Access* 8: 138788–138799.

Dixon W (2014) *Optimal Adaptive Control and Differential Games by Reinforcement Learning Principles*. IEEE.

Doležǎl J (1979) Parameter optimization for two-player zero-sum differential games. *Journal of Dynamic Systems, Measurement, and Control* 101(4): 345–349.

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): 1–7.

Isaacs R (1965) *Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization*. John Wiley.

Ji X, Yang K, Na X, et al. (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): 081005.

Jin L, Liao B, Liu M, et al. (2017) Different-level simultaneous minimization scheme for fault tolerance of redundant manipulator aided with discrete-time recurrent neural network. *Frontiers in Neurorobotics* 11: 50.

Johnson MA (2011) *Differential Game-Based Control Methods for Uncertain Continuous-Time Nonlinear Systems*. Citeseer.

Johnson M, Bhasin S and Dixon WE (2011) Nonlinear two-player zero-sum game approximate solution using a policy iteration algorithm. In: 2011 50th Decision and control and European control conference (CDC-ECC), FL, USA, 12–15 December 2011, pp. 142–147. IEEE.

Jumarie G (1977) Coordination concepts in multilevel hierarchical differential games. *Journal of Dynamic Systems, Measurement, and Control* 99(3): 201–208.

Kamalapurkar R, Klotz J and Dixon WE (2014a) Model-based reinforcement learning for on-line feedback-nash equilibrium solution of n-player nonzero-sum differential games. In: American control conference (ACC), OR, USA, 4–6 June 2014, pp. 3000–3005. IEEE.

Kamalapurkar R, Klotz JR and Dixon WE (2014b) Concurrent learning-based approximate feedback-nash equilibrium solution of n-player nonzero-sum differential games. *IEEE/CAA Journal of Automatica Sinica* 1(3): 239–247.

Kamalapurkar R, Klotz JR, Walters P, et al. (2018) Model-based reinforcement learning in differential graphical games. *IEEE Transactions on Control of Network Systems* 5(1): 423–433.

Kirk DE (2012) *Optimal Control Theory: An Introduction*. Courier Corporation.

Korayem MH, Nohooji HR and Nikoobin A (2011) Path planning of mobile elastic robotic arms by indirect approach of optimal control. *International Journal of Advanced Robotic Systems* 8(1): 10.

Korayem MH, Shafei AM and Dehkordi SF (2013) Systematic modeling of a chain of n-flexible link manipulators connected by revolute–prismatic joints using recursive gibbs-appell formulation. *Archive of Applied Mechanics* 84(2): 187–206.

Korayem MH, 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): 5197–5205.

LaValle SM and Hutchinson SA (1994) Path selection and coordination for multiple robots via nash equilibria. In: Proceedings of the 1994 IEEE international conference on robotics and automation. San Diego, CA, May 8–13 1994, pp. 1847–1852. IEEE.

Lee D, Periaux J and Gonzalez LF (2010) UAS mission path planning system (MPPS) using hybrid-game coupled to multiobjective optimizer. *Journal of Dynamic Systems, Measurement, and Control* 132(4): 041005.

Li S, He J, Li Y, et al. (2016a) Distributed recurrent neural networks for cooperative control of manipulators: a game-theoretic perspective. *IEEE Transactions on Neural Networks and Learning Systems* 28(2): 415–426.

Li Y, Tee KP, Yan R, et al. (2016b) A framework of human–robot coordination based on game theory and policy iteration. *IEEE Transactions on Robotics* 32(6): 1408–1418.

Li Z, Li C, Li S, et al. (2020a) A fault-tolerant method for motion planning of industrial redundant manipulator. *IEEE Transactions on Industrial Informatics* 16(12): 7469–7478.

Li Z, Zuo W and Li S (2020b) Zeroing dynamics method for motion control of industrial upper-limb exoskeleton system with minimal potential energy modulation. *Measurement* 163: 107964.

Liu W, Chen D and Steil J (2016) Analytical inverse kinematics solver for anthropomorphic 7-DOF redundant manipulators with human-like configuration constraints. *Journal of Intelligent & Robotic Systems* 86(1): 63–79.

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): 326–331.

Meng Y (2008) Multi-robot searching using game-theory based approach. *International Journal of Advanced Robotic Systems* 5(4): 44.

Nammoto T and Kosuge K (2012) An analytical solution for a redundant manipulator with seven degrees of freedom. *International Journal of Automation and Smart Technology* 2(4): 339–346.

Nash J (1951) Non-cooperative games. *Annals of Mathematics* 54: 286–295.

Naudet J, Lefeber D, Daerden F, et al. (2003) Forward dynamics of open-loop multibody mechanism using a efficient recursive algorithm based on canonical momenta. *Multibody System Dynamics* 10(1): 45–59.

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

Saha SK and Schiehlen WO (2001) Recursive kinematics and dynamics for parallel structured closed-loop multibody system. *Mechanics of Structures and Machines* 29(2): 143–175.

Shafei AM and Korayem MH (2017) Theoretical and experimental study of dynamic load-carrying capacity for flexible robotic arms in point-to-point motion. *Optimal Control Applications and Methods* 38(6): 963–972.

Shafei A and Shafei H (2018) Dynamic modeling of planar closed-chain robotic manipulators in flight and impact phases. *Mechanism and Machine Theory* 126: 141–154.

Sharma R (2016) Lyapunov theory based stable Markov game fuzzy control for non-linear systems. *Engineering Applications of Artificial Intelligence* 55: 119–127.

Sharma R and Gopal M (2014) Fictitious play based markov game control for robotic arm manipulator. In: Proceedings of 3rd international conference on reliability, infocm technologies and optimization, Noida, India, 8−10 October 2014, pp. 1−6. IEEE.

Smyrnakis M and Veres SM (2014) Coordination of control in robot teams using game-theoretic learning. *IFAC Proceedings Volumes* 47(3): 1194–1202.

Tijs S (2003) *Introduction to Game Theory*. Springer.

Wang X, Yaz EE and Schneider SC (2018) Coupled state-dependent riccati equation control for continuous time nonlinear mechatronics systems. *Journal of Dynamic Systems, Measurement, and Control* 140(11): 111013.

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

Xie B, Zhao J and Liu Y (2012) Fault tolerant motion planning of robotic manipulators based on a nested RRT algorithm. *Industrial Robot: An International Journal* 39(1): 40–46.

Xu P, Groff RE and Burg T (2012). Dynamics and control of the shoot-the-moon tabletop game. *Journal of Dynamic Systems, Measurement, and Control* 134(5): 051006.

Zhou D, Ji L, Zhang Q, et al. (2015) Practical analytical inverse kinematic approach for 7-DOF space manipulators with joint and attitude limits. *Intelligent Service Robotics* 8(4): 215–224.