AUT Journal of Electrical Engineering

AUT J. Elec. Eng., 55(1) (2023) 99-106
DOI: 10.22060/eej.2022.21372.5475

# Assistive Control of a Hip Exoskeleton Robot, using a DQN-Adjusted Delayed Output Feedback Method

H. Kalani [1*], S. M. Tahamipour-Z [2], I. Kardan [2], A. Akbarzadeh [2]

[1] Department of Mechanical Engineering, Sadjad University, Mashhad, Iran

[2] Center of Advanced Rehabilitation and Robotic Research (FUM-CARE), Mechanical Engineering Department, Ferdowsi University of Mashhad, Mashhad, Iran

**ABSTRACT:** A major challenge in the development of an assistive exoskeleton robot is to design appropriate control algorithms. These algorithms should be trajectory-independent and require a minimum number of sensors to work in any intended motion and to be easily implementable. As a simple assistive strategy with all promising features, Delayed Output Feedback Control (DOFC) is shown to be effective in assisting the wearers in different types of motion. In this method, the assistive torques are defined in proportion to delayed feedback from the angle difference between the two legs. The authors have recently suggested an intelligent version of DOFC, in which a Deep Q-Network (DQN) was used to adjust the feedback delay according to the speed of the motion. Simulation studies were used to investigate the idea. By conducting some real-world experiments, the present paper extends the results to practical conditions. The provided results clearly verify that if the time delay is adjusted according to the walking speed, the DOFC method can effectively help the users in their motions of any speed. The results also indicated that a fixed or an inappropriate value of the delay may result in resistance against the user motion.

**Review History:**

Received: May, 21, 2022
Revised: Oct. 31, 2022
Accepted: Nov. 05, 2022
Available Online: Mar. 01, 2023

**Keywords:**

Hip Exoskeleton Robot
Delayed Output Feedback Control
Deep Q-Network
Reinforcement Learning
Human-Robot Interaction

## 1- Introduction

Since impaired lower limb mobility would significantly reduce the quality of life, various studies have considered assistive lower-limb exoskeletons to facilitate the daily tasks for the elderlies and patients with physical disabilities [1-4]. A crucial issue in the development of these robots is to design an appropriate control method. A suitable controller for assistive exoskeletons should have a simple structure, work for different motion types, and require a low number of sensors [5-8]. Numerous control methods have been proposed for exoskeleton robots such as force/torque-based controllers [7]-[9-12], impedance/admittance control [13, 14], adaptive oscillator–based control [15, 16] and Reinforcement Learning [8, 17]. Delayed Output Feedback Control (DOFC) [18], is a simple yet effective strategy developed for hip exoskeletons. In this method, the assistive torques are defined in proportion to delayed feedback from the angle difference between the two legs. The authors have recently suggested an intelligent version of DOFC in which a Deep Q-Network (DQN) was used to adjust the feedback delay according to the speed of the motion [17].

Compared to the classical control methods, Reinforcement Learning has considerable advantages for nonlinear systems. Reinforcement Learning (RL) is a powerful, key method in finding optimal policies in complex sequential decision-making problems [19]. There are two categories of classical RL algorithms (i.e. value-based and policy-based). In value-based techniques (such as Q-learning [20], SARSA [21], and Fitted Q-Iteration [22]), the state or state-action value is learned and we don’t store any explicit policy. However, in policy-based techniques (such as Actor-Critic [23], Proximal Policy optimization [24]), we explicitly build a representation of a policy (mapping $\pi$:s→a) that maps state to action. Among these methodologies, Q-Learning has been widely used as a model-free Reinforcement Learning which was limited to tasks with small state spaces [25]. Hamaya et. al [26] used RL method to suggest a task-parametrized assistive strategies for exoskeleton robots. They employed an assistive strategy which improves human-robot interaction data across different tasks, and obtained a generalized method even for an unseen task. They showed the abilities of their approach by reducing the user’s EMGs. Tu et. al [27] developed a new data-driven framework based on RL without using the human-robot dynamics, to provide adaptive personalized torque assistance for reducing human efforts during walking. They used two control timing parameters (peak and offset timings), and Least Square Policy Iteration (LSPI) for learning the parameter tuning policy. They indicated that their approach was reasonable for reducing assistive torque profile of the hip exoskeleton. Yuan et. al [28] proposed a novel trajectory-learning scheme based on RL combined with Dynamic Movement Primitives (DMPs) for a lower limb exoskeleton

*Corresponding author’s email: kalani@sadjad.ac.ir

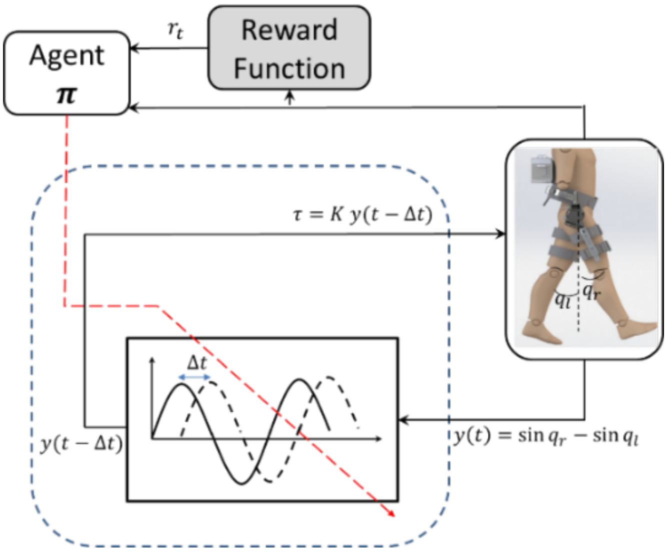


**Fig. 1. Overall block diagram of the adaptive DOFC method.**

robot, aiming to give assistance to human walking. They employed RL to learn the trajectories which learned by DMPs for eliminating the effects of uncertainties in joint space. Although, real-life applications such as robotics, usually have continuous state and action spaces, making it impossible to store exact value functions or policies for each separate state or state-action pair [29]. Therefore, Deep Q-Networks (DQN) suggested covering the full range of states and actions continuously. DQN is one of the most important methodologies of deep Reinforcement Learning [30, 31]. Rastogi [29] applied DQN to learn a 2D bipedal robot for walking. He compared the abilities of DQN with traditional RL methods, like SARSA. He showed that the DQN does seem like a promising way to learn control policies for real systems compared to other traditional RL methods. Xue et. al [32] used Double Q-learning Network (DDQN) to enable mobile robots to learn collision avoidance and navigation capabilities autonomously. In their purposed method, target position, obstacle size, and position are used as input, and the direction of movement of the robot was used as an output. In fact, they developed an obstacle avoidance model based on DQN and applied it successfully on a physical robot platform. Liu et. al [33] used DQN to optimize the average stability margin of the hexapod robot in complex and rugged terrain. They showed that in comparison with the traditional method, the DQN algorithm can plan a free gait with fast convergence, high stability margin and strong adaptability. Rose et. al [34, 35] suggested a model-free deep Reinforcement Learning method for an exoskeleton that can learn to follow a desired gait pattern. In their method, the predefined dynamics models of the user and exoskeleton is not needed.

In the previous paper, we verified by analysis and simulation that the time delay in the DOFC method [17] should be changed according to the walking speed of the wearer. In this paper, we verifiy the simulation results by a custom-made hip exoskeleton robot.

**Fig. 2. View of FUM-Hip Exoskeleton Robot**

This study is organized as follows: Section 2 presents the problem definition and proposes the DQN-adjusted DOFC method. Section 3 provides the simulation and experimental results for walking conditions, and Section 4 concludes this paper.

## 2- Methods and Materials

One of the main challenges in the development of assistive exoskeletons is to design a control algorithm to assist the user in his/her intended motions. Based on a newly developed assistive strategy (Delayed Output Feedback Control (DOFC) by Lim et al. [18]), the authors have recently proposed an adaptive assistive controller [17]. Fig. 1 shows the overall block diagram of the proposed controller that combines Reinforcement Learning and the DOFC method to develop an adaptive assistive strategy. As detailed in the following subsections, the intelligent agent, which is a neural network, learns a policy for the optimal adjustment of the time delay in DOFC method to minimize the energy consumption. This way, unlike the original DOFC method, the adaptive DOFC method will maintain its performance in different motion speeds. While the main idea and initial simulations of the adaptive DOFC method are presented in our recent paper [17], in the present paper a prototype of FUM- Hip Exoskeleton Robot is used to evaluate the performance of the adaptive DOFC in real-world experiments.

### 2- 1- FUM-Hip Exoskeleton Robot

The purpose of developing the robot is to provide appropriate walking assistance in human locomotion. We designed the mechanical structure of the hip exoskeleton robot (Fig. 2), which meets the requirements of convenient wearing and enables natural walking. This robot has an active DOF (flexion/extension) at each hip joint. The custom made FUM-Hip Exoskeleton is actuated by two 70W Maxon motors with Harmonic Drive gearboxes of the ratio 1:100. The joint angles are obtained through measuring the motor angles by 16-bit incremental encoders. Two EPOS2 70/15 controllers are used to drive the motors in the torque mode. The desired torque is defined by the DOFC algorithm, implemented in TwinCAT software, and sent to the drivers through a Beckhoff PLC. Mechanical components are made of Aluminum material and the total weight of the exoskeleton is about 5.5 kg.

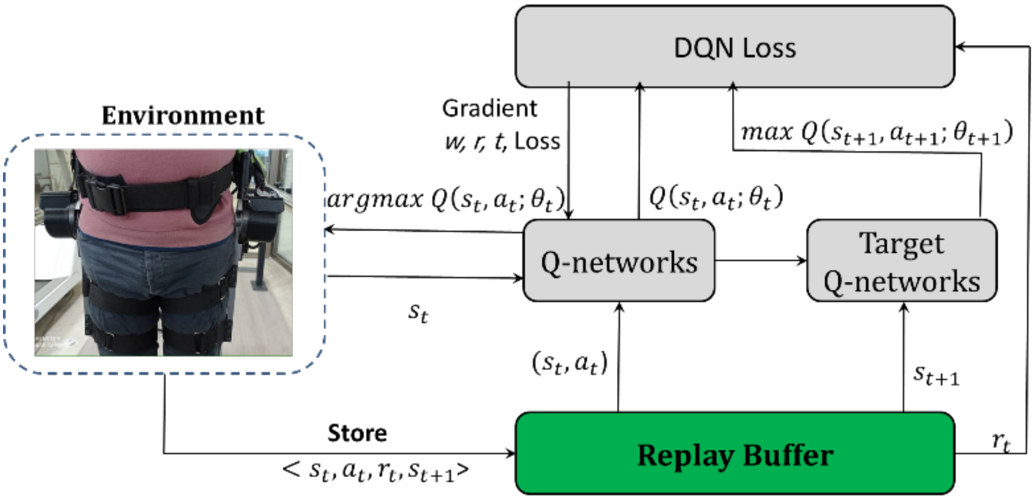


**Fig. 3. The overall structure of the DQN algorithm**

### 2- 2- Assistive Strategy based on Delayed Output Feedback Control

As shown in Fig. 1, the DOFC method uses the low-pass filtered hip angles, $\overline{q}_r$ (t) and $\overline{q}_l$ (t), to calculate the output feedback variable x(t) as:

$$x(t) = \sin \overline{q}_r(t) - \sin \overline{q}_l(t) \quad (1)$$

It is clear that the output feedback variable is just the difference in sinusoidal projections of the hip angles, used to give a normalized variable.

In the implementation of DOFC method, a discrete low-pass filter is applied to the raw values of the joint angles, $q_r$ and $q_l$, as:

$$\begin{aligned} \overline{q}_{r,k} &= (1-\alpha)\overline{q}_{r,k-1} + \alpha q_{r,k} \\ \overline{q}_{l,k} &= (1-\alpha)\overline{q}_{l,k-1} + \alpha q_{l,k} \end{aligned} \quad (2)$$

After calculating the output feedback variable in Eq. (1), a time delay of Δt seconds is applied and the assistive torque, τ(t), is defined as:

$$\tau(t) = \kappa x(t - \Delta t) \quad (3)$$

where κ is the controller gain.

According to Eq. (3), the DOFC method includes two main parameters (i.e. the time delay Δt and the gain κ), that can affect the assistive torque. The gain κ adjusts the magnitude of the assistive torque and Δt affects the timing of the applied torque. As shown in [17], the proper setting of the parameters is crucial for the effective performance of the method. Maladjustment of this parameter may lead to reduce the effectiveness of the assistive strategy and may even lead to resistance against user motions. Due to the complicated relationship between the value of the time delay parameter and the effectiveness of the DOFC method, Deep Q-Network learning is utilized in this study. This network learns the optimal values of the time delay parameter and automatically adjusts it according to the available feedbacks.

### 2- 3- Deep Q-Networks (DQN) Learning

In DQN method, a neural network with weights $\theta_t$ is utilized as the Q-function. This network is trained by minimizing a loss function at every iteration as follows:

$$L_t(\theta_t) \triangleq \mathrm{E}_{s\sim\rho\pi(\cdot)\cdot a\sim(\cdot)}\left[(y_t - Q(s.a;\theta_t))^2\right] \quad (4)$$

where $y_t \triangleq \mathrm{E}_{s'}\,[r_{s,s',a} + \gamma \max_{a'} Q(s',a';\theta_{t-1}) | s,a]$ is a target at iteration *i*, *π(s,a)* is the policy of state s and $\rho_{\pi(\cdot)}$ is the distribution of states under *π(s,a)*. To approximate the optimal action-value function, we have:

$$Q^*(s.a) = \mathrm{E}_{s'}\left[rs.s' + \gamma \max_{a'} Q^*(s'.a') \,|\, s.a\right] \quad (5)$$

where the $Q^*(s,a)$ is the optimal Q value, $r_{s,s',a}$ is the obtained reward, and the next state is s′ and the next action is a′. The γ ∈ [0,1] is the discount factor. If we obtain the weight of neural networks at time *t* as $\theta_t$, Eq. (5) can be employed to approximate the function $Q(s, a; \theta_t)$ such that

$Q(s, a;\theta_t) \approx Q^*(s, a)$ when $t \to \infty$. In order to minimize the loss function, the gradient of the function is utilized with respect to the weight as follows:

$$\nabla_{\theta_t} L(\theta_t) = \mathrm{E}_{s\sim\rho\pi(\cdot),a\sim(\cdot)}\left[\left(r_{s,s'} + \gamma \max_{a'} Q^*(s',a') - Q(s,a;\theta_t)\right)\nabla_{\theta_t} Q(s,a;\theta_t)\right] \quad (6)$$

Stochastic Gradient Descent is used to minimize the loss function. We utilized a ε-greedy method which chooses between a random action and the action corresponding to the highest Q-value produced by DQN. During training, the value of *ε* is decreased due to optimal action and less iteration. The key aspect which makes DQN work is the use of experience replay. In this technique, to break harmful correlations and make better use of our experience, at each time step $(s_t, a_t, r_t, s_{t+1})$ is stored in a replay buffer. Next, a

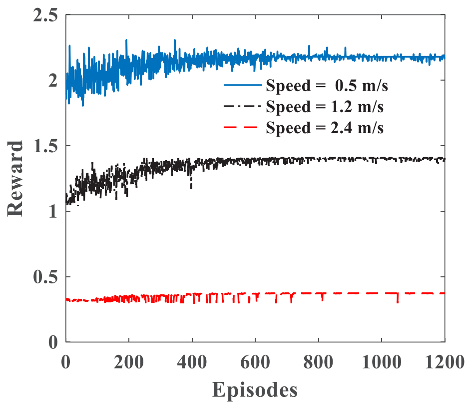


**Fig. 4. Changes in the reward values over the progressing episodes of training for different walking speed.**

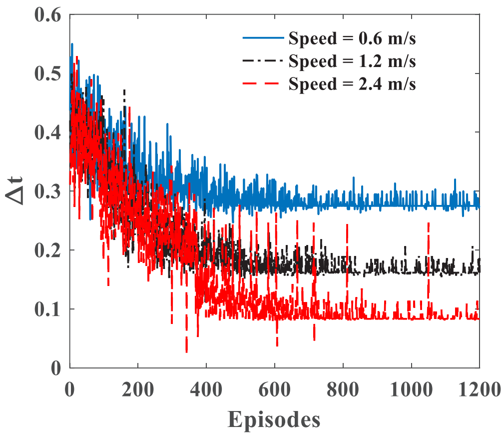


**Fig. 5. Changes in the time delay values over the progressing episodes of training for different walking speed.**

fixed number of samples are extracted from the replay buffer randomly and utilized to train the network. The structure of DQN is shown in Fig. 3, where the energy consumed in each walking cycle is considered as the states, the time-delay Δt is the action, and the inverse of energy consumption in each walking cycle is defined as the reward function. The structure of the fully connected Multi-Layer Perceptron Neural Network (MLPANN) has one hidden layer of 25 nodes which was chosen by performing a small trial and error procedure. The activation function we used in hidden layer is sigmoid function. The last layer has linear activation function. The adaptation law has been developed by the Gradient Descent (GD) method

## 3- Results and Discussion

### 3- 1- Simulations

As detailed in [17], a seven-link biped model is used as a human subject in the sagittal plane to conduct some simulations and evaluate the effectiveness of the proposed method in reducing the energy cost of walking. The model is set to walk at different speeds of 0.6 m/s, 1.2 m/s and 2.4 m/s and energy consumption is calculated for each cycle. This energy is used as the state and its inverse is defined as the reward function. The DQN is set to optimize the value of

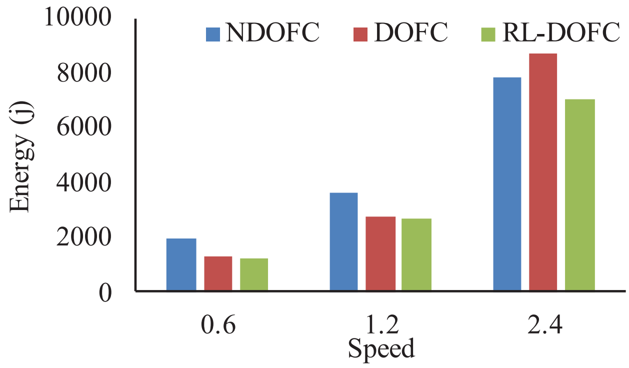


**Fig. 6. Mean energy consumption versus walking speed. NDOFC: Unassisted walking. DOFC: Assisted walking with Delayed Output Feedback Control. RL-DOFC: Assisted walking with RL-tuned DOFC.**

the time delay for each walking speed, such that the reward function is maximized and, consequently, the required energy is minimized.

Fig. 4 shows the changes in the reward value for the progressing episodes of training. It can be seen that the DQN has successfully increased the value of the reward function and reduced the required energy.

Different optimal values of Δt in Fig. 5 verify the ideas that the time delay should be adjusted according to the walking speed and be reduced as the walking speed increases. This observation can be justified by noting that the cycle time is smaller for a faster walk. Therefore, the time delay should also be smaller.

Fig. 6 compares the energy consumption for normal walking without the exoskeleton (NDOFC), assisted walking with DOFC method having a fixed Δt (DOFC), and assisted walking with DOFC method having an optimized Δt, as defined by DQN (RL-DOFC). It is seen that assisted walking with the RL-DOFC method has the minimum energy requirement for all walking speeds. The DOFC method, with a fixed value of Δt = 0.3s has also reduced the energy for walking at speeds of 0.6 m/s and 1.2 m/s. However, for the speed of 1.2 m/s, the DOFC method has lost its assistive property and imposed some resistance against the user motions. Therefore, the energy consumption is increased, even over that of unassisted walking.

The results reaffirm the idea that the time delay in the DOFC method should be reduced as the walking speed increases. Improper adjustment of Δt may reduce the assistive effect of the method and may even result in some resistance against the user motions. The results also verify that the DQN has properly adjusted the time delay value for each walking speed.

### 3- 2- Experiments

For further evaluation of the concept in practical conditions, a custom-made hip exoskeleton robot is used to conduct a preliminary experiment. As shown in Fig. 7, the user is asked to wear the robot and walk on a treadmill, running at the speed of about 0.6 m/s. During the walk, the extension/flexion angle of the right and left hips and their

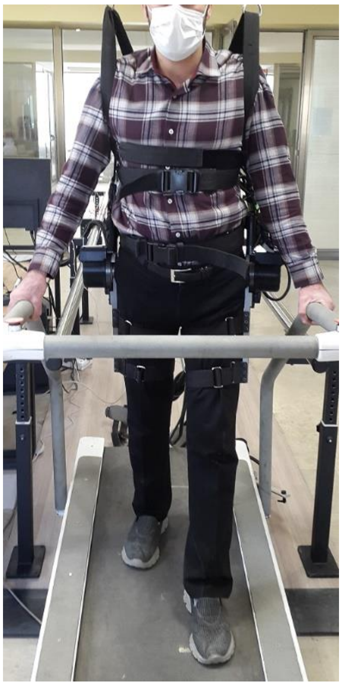

**Fig. 7. A healthy subject, wearing the hip exoskeleton robot.**

angular velocity are recorded through the motor encoders. An overhead weight support system is used for partial support of the robot's weight.

After some familiarization runs, the main experiment is conducted in two steps. In the first step, the user wears the robot in zero-impedance mode and moves freely on the treadmill. In the second step, the DOFC method is implemented with three different values for the feedback delay, namely

$\Delta t = 0.2s$, $\Delta t = 0.3s$, and $\Delta t = 0.4s$.

In Fig. 8, the phase portraits ($\dot{q}_r$ vs $q_r$) of the assisted motions, aided by the DOFC method with different delay values, are compared to the phase portrait of the free motion.

As shown in Fig. 8(a), the delay of $\Delta t = 0.2s$ has led to a smaller phase portrait, compared to that of the free motion. This is in accordance with our expectations that improper setting of the feedback delay will result in resistance against the user motions. The simulation results verify that a delay of $\Delta t = 0.2s$ is too small for the given treadmill speed (i.e. 0.6 m/s).

The enlarged phase portraits in Fig. 8(b) and Fig. 8(c) clearly indicate that when the delay is set to $\Delta t = 0.3s$ or

$\Delta t = 0.4s$, the DOFC method has successfully assisted the user in his intended motion.

For a better comparison, the phase portraits of the motions are depicted in a single plot in Fig. 9. As shown in Fig. 9, an improper setting of the delay ($\Delta t = 0.2s$) leads to resistance against user motions while with a proper setting ($\Delta t = 0.3s$ and $\Delta t = 0.4s$), the DOFC method can provide effective assistance for the user. Fig. 9 also indicates that there is only a slight difference between the results obtained for

$\Delta t = 0.3s$ and $\Delta t = 0.4s$. This observation can be justified by considering the simulation results, indicating that the

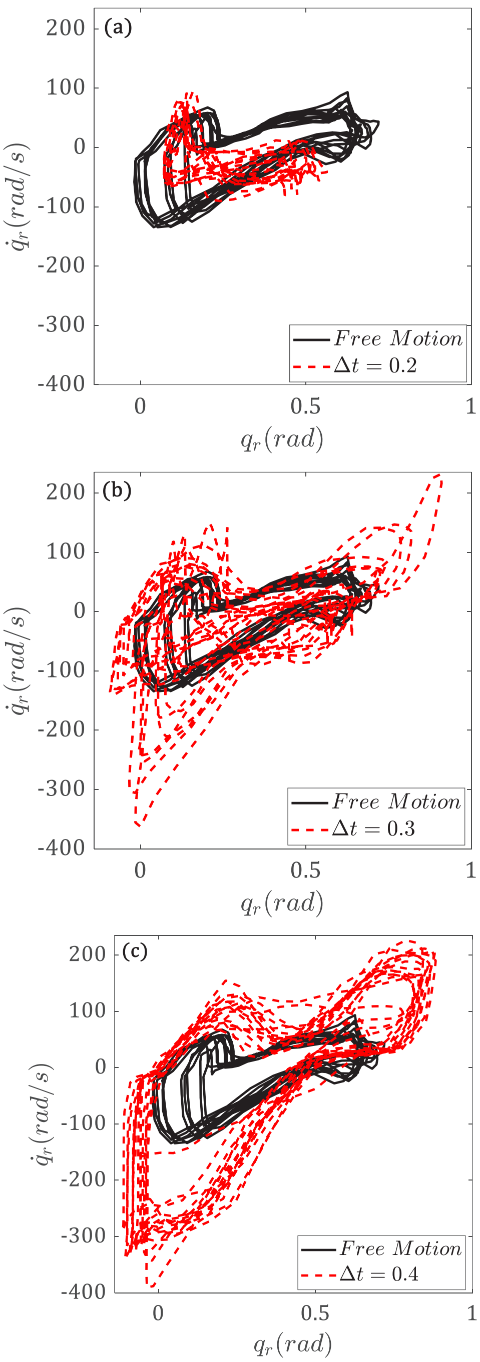


**Fig. 8. Phase portraits of the motions assisted by the DOFC method with different feedback delays, compared to the phase portrait of the free motion.**

optimal delay value for the given treadmill speed (i.e. 0.6 m/s), should be a value between 0.3s and 0.4s. A larger or smaller value of the delay may lead to suboptimal results, as in the cases of $\Delta t = 0.3s$ and $\Delta t = 0.4s$, or may result in resistance against user motions, as in the case of $\Delta t = 0.2s$. The good agreement between simulation and experiment results reaffirms the fact that the performance of the DOFC method is mostly related to the kinematics of the motion, rather than its dynamics. This also verifies the good calibration of the model for the specific subject.

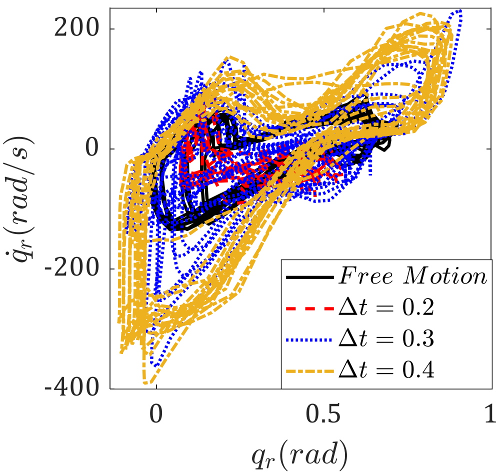


**Fig. 9. Overall comparison of the phase portraits of the free and assisted**

## 4- Conclusion

In this paper, an adaptive assistive strategy is developed to assist the motions of the users wearing hip exoskeletons. The proposed strategy is an intelligent or adaptive version of the Delayed Output Feedback Controller (DOFC), developed by Lim et. al [18]. A Deep Q-Network (DQN) is used to adjust the time delay value in DOFC according to the walking speed. Therefore, unlike the original DOFC method, the DQN-adjusted DOFC method will maintain its performance in different motion speeds. The results obviously indicate that DQN can effectively improve the performance of the DOFC method, by proper adjustment of the time delay. Some simulation tests are conducted to prove the idea that different time delays are required for different walking speeds. The simulation results also verify the successful performance of DQN in proper adjustment of the time delay. Finally, some experiments are conducted to verify the performance of the adaptive DOFC method in real-world applications. The experimental results clearly show that the phase portraits of the user motions are enlarged if DOFC is implemented by appropriate time delay values. However, if the time delay is not properly adjusted, the implementation of the DOFC method may make the phase portraits smaller, indicating that there is some resistance against user motions. A good agreement is observed between simulation and experimental results which may be justified by the fact that the performance of the DOFC method is mostly related to the kinematics of the motion, rather than its dynamics.

The main contribution of this paper is the implementation of a Deep Q-Network (DQN) technique, which have the ability to reasonably learn control schemes without the need for dynamics model of hip exoskeleton robot. Moreover, we focused on assistive strategy, where for every walking speed, the user doesn't need to interact with a robot to learn an appropriate assistive strategy. Finally, we developed an experimental platform with a hip exoskeleton robot to verify the effectiveness of our method.

In future works, we plan to validate our results by analyzing the electromyography signals of hip muscles and measuring the metabolic energy of volunteers in treadmill walking.

## Acknowledgment

This research is supported by grant #101120 from the Ferdowsi University of Mashhad and grant #962297 from the National Institute for Medical Research Development of Iran. We want to thank all members of the FUM Center Advanced of Rehabilitation and Robotics (FUMCARE) at the Ferdowsi University of Mashhad for their kind participation and cooperation.

## References

[1] M. Grimmer, R. Riener, C.J. Walsh, A. Seyfarth, Mobility related physical and functional losses due to aging and disease-a motivation for lower limb exoskeletons, Journal of neuroengineering and rehabilitation, 16(1) (2019) 1-21.

[2] A. Rodríguez-Fernández, J. Lobo-Prat, J.M. Font-Llagunes, Systematic review on wearable lower-limb exoskeletons for gait training in neuromuscular impairments, Journal of neuroengineering and rehabilitation, 18(1) (2021) 1-21.

[3] F. Hussain, R. Goecke, M. Mohammadian, Exoskeleton robots for lower limb assistance: A review of materials, actuation, and manufacturing methods, Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, 235(12) (2021) 1375-1385.

[4] T. Wang, B. Zhang, C. Liu, T. Liu, Y. Han, S. Wang, J.P. Ferreira, W. Dong, X. Zhang, A Review on the Rehabilitation Exoskeletons for the Lower Limbs of the Elderly and the Disabled, Electronics, 11(3) (2022) 388.

[5] J. Zhou, S. Yang, Q. Xue, Lower limb rehabilitation exoskeleton robot: A review, Advances in Mechanical Engineering, 13(4) (2021) 16878140211011862.

[6] D. Shi, W. Zhang, W. Zhang, X. Ding, A review on lower limb rehabilitation exoskeleton robots, Chinese Journal of Mechanical Engineering, 32(1) (2019) 1-11.

[7] R. Baud, A.R. Manzoori, A. Ijspeert, M. Bouri, Review of control strategies for lower-limb exoskeletons to assist gait, Journal of NeuroEngineering and Rehabilitation, 18(1) (2021) 1-34.

[8] S. Luo, G. Androwis, S. Adamovich, H. Su, E. Nunez, X. Zhou, Reinforcement Learning and Control of a Lower Extremity Exoskeleton for Squat Assistance, Frontiers in Robotics and AI, 8 (2021).

[9] K. Kiguchi, K. Tamura, Y. Hayashi, Estimation of joint force/torque based on EMG signals, in: 2013 IEEE Workshop on Robotic Intelligence in Informationally Structured Space (RiiSS), IEEE, 2013, pp. 20-24.

[10] U. Nagarajan, G. Aguirre-Ollinger, A. Goswami, Integral admittance shaping: A unified framework for active exoskeleton control, Robotics and Autonomous Systems, 75 (2016) 310-324.

[11] G. Aguirre-Ollinger, J.E. Colgate, M.A. Peshkin, A. Goswami, Inertia compensation control of a one-degree-of-freedom exoskeleton for lower-limb assistance: Initial experiments, IEEE transactions on neural systems and rehabilitation engineering, 20(1) (2012) 68-77.

[12] G. Lv, J. Lin, R.D. Gregg, Trajectory-free control of lower-limb exoskeletons through underactuated total energy shaping, IEEE Access, 9 (2021) 95427-95443.

[13] T. Zhang, M. Tran, H. Huang, Admittance shaping-based assistive control of SEA-driven robotic hip exoskeleton, IEEE/ASME Transactions on Mechatronics, 24(4) (2019) 1508-1519.

[14] L.D. da Silva, T.F. Pereira, V.R. Leithardt, L.O. Seman, C.A. Zeferino, Hybrid Impedance-Admittance Control for Upper Limb Exoskeleton Using Electromyography, Applied Sciences, 10(20) (2020) 7146.

[15] K. Seo, K. Kim, Y.J. Park, J.-K. Cho, J. Lee, B. Choi, B. Lim, Y. Lee, Y. Shim, Adaptive oscillator-based control for active lower-limb exoskeleton and its metabolic impact, in: 2018 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2018, pp. 6752-6758.

[16] T. Xue, Z. Wang, T. Zhang, M. Zhang, Adaptive oscillator-based robust control for flexible hip assistive exoskeleton, IEEE Robotics and Automation Letters, 4(4) (2019) 3318-3323.

[17] H. Kalani, S.M. Tahamipour-Z, I. Kardan, A. Akbarzadeh, Application of DQN Learning for Delayed Output Feedback Control of a Gait-Assist Hip Exoskeleton, in: 2021 9th RSI International Conference on Robotics and Mechatronics (ICRoM), IEEE, 2021, pp. 341-345.

[18] B. Lim, J. Lee, J. Jang, K. Kim, Y.J. Park, K. Seo, Y. Shim, Delayed Output Feedback Control for gait assistance with a robotic hip exoskeleton, IEEE Transactions on Robotics, 35(4) (2019) 1055-1062.

[19] H. Kalani, A. Akbarzadeh, Application of Reinforcement Learning for Navigation of a Planar Snake Robot in Serpentine Locomotion, Journal Of Applied and Computational Sciences in Mechanics, 26(1) (2015) 97-118.

[20] E.S. Low, P. Ong, K.C. Cheah, Solving the optimal path planning of a mobile robot using improved Q-learning, Robotics and Autonomous Systems, 115 (2019) 143-161.

[21] X. Wenxia, B. Yu, L. Cheng, Y. Li, X. Cao, Multi-fuzzy Sarsa learning-based sit-to-stand motion control for walking-support assistive robot, International Journal of Advanced Robotic Systems, 18(5) (2021) 17298814211050190.

[22] M. Gaeta, V. Loia, S. Miranda, S. Tomasiello, Fitted Q-iteration by Functional Networks for control problems, Applied mathematical modelling, 40(21-22) (2016) 9183-9196.

[23] Y. Ouyang, L. Dong, Y. Wei, C. Sun, Neural network based tracking control for an elastic joint robot with input constraint via actor-critic design, Neurocomputing, 409 (2020) 286-295.

[24] L.C. Melo, M.R.O.A. Máximo, Learning humanoid robot running skills through proximal policy optimization, in: 2019 Latin american robotics symposium (LARS), 2019 Brazilian symposium on robotics (SBR) and 2019 workshop on robotics in education (WRE), IEEE, 2019, pp. 37-42.

[25] A. Konar, I.G. Chakraborty, S.J. Singh, L.C. Jain, A.K. Nagar, A deterministic improved Q-learning for path planning of a mobile robot, IEEE Transactions on Systems, Man, and Cybernetics: Systems, 43(5) (2013) 1141-1153.

[26] M. Hamaya, T. Matsubara, T. Noda, T. Teramae, J. Morimoto, Learning task-parametrized assistive strategies for exoskeleton robots by multi-task Reinforcement Learning, in: 2017 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2017, pp. 5907-5912.

[27] X. Tu, M. Li, M. Liu, J. Si, H.H. Huang, A data-driven Reinforcement Learning solution framework for optimal and adaptive personalization of a hip exoskeleton, in: 2021 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2021, pp. 10610-10616.

[28] Y. Yuan, Z. Li, T. Zhao, D. Gan, DMP-based motion generation for a walking exoskeleton robot using Reinforcement Learning, IEEE Transactions on Industrial Electronics, 67(5) (2019) 3830-3839.

[29] D. Rastogi, Deep Reinforcement Learning for Bipedal Robots, (2017).

[30] [30] I. Osband, C. Blundell, A. Pritzel, B. Van Roy, Deep exploration via bootstrapped DQN, Advances in neural information processing systems, 29 (2016).

[31] X. Chen, C. Fu, J. Huang, A Deep Q-Network for robotic odor/gas source localization: Modeling, measurement and comparative study, Measurement, 183 (2021) 109725.

[32] X. Xue, Z. Li, D. Zhang, Y. Yan, A deep Reinforcement Learning method for mobile robot collision avoidance based on double dqn, in: 2019 IEEE 28th International Symposium on Industrial Electronics (ISIE), IEEE, 2019, pp. 2131-2136.

[33] Y. Liu, Y. Xu, Free Gait Planning of Hexapod Robot Based on Improved DQN Algorithm, in: 2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT, IEEE, 2020, pp. 488-491.

[34] L. Rose, M.C. Bazzocchi, G. Nejat, End-to-end deep Reinforcement Learning for exoskeleton control, in: 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, 2020, pp. 4294-4301.

[35] L. Rose, M.C. Bazzocchi, G. Nejat, A model-free deep Reinforcement Learning approach for control of exoskeleton gait patterns, Robotica, 40(7) (2022) 2189-2214.



**HOW TO CITE THIS ARTICLE**

*H. Kalani, S. M. Tahamipour-Z, I. Kardan, A. Akbarzadeh, Assistive Control of a Hip Exoskeleton Robot, using a DQN-Adjusted Delayed Output Feedback Method , AUT J. Elec. Eng., 55(1) (2023) 99-106.*

**DOI:** 10.22060/eej.2022.21372.5475