



Data-Driven Optimal Assistance Control of a Lower Limb Exoskeleton for Hemiplegic Patients

Zhinan Peng¹, Rui Luo¹, Rui Huang^{1*}, Tengbo Yu², Jiangping Hu¹, Kecheng Shi¹ and Hong Cheng¹

¹ School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu, China,

² Department of Sports Medicine, The Affiliated Hospital of Qingdao University, Qingdao, China

More recently, lower limb exoskeletons (LLE) have gained considerable interests in strength augmentation, rehabilitation, and walking assistance scenarios. For walking assistance, the LLE is expected to control the affected leg to track the unaffected leg's motion naturally. A critical issue in this scenario is that the exoskeleton system needs to deal with unpredictable disturbance from the patient, and the controller has the ability to adapt to different wearers. To this end, a novel data-driven optimal control (DDOC) strategy is proposed to adapt different hemiplegic patients with unpredictable disturbances. The interaction relation between two lower limbs of LLE and the leg of patient's unaffected side are modeled in the context of leader-follower framework. Then, the walking assistance control problem is transformed into an optimal control problem. A policy iteration (PI) algorithm is utilized to obtain the optimal controller. To improve the online adaptation to different patients, an actor-critic neural network (AC/NN) structure of the reinforcement learning (RL) is employed to learn the optimal controller on the basis of PI algorithm. Finally, experiments both on a simulation environment and a real LLE system are conducted to verify the effectiveness of the proposed walking assistance control method.

OPEN ACCESS

Edited by:

Bin Fang,
Tsinghua University, China

Reviewed by:

Liangtian He,
Anhui University, China
Lei Deng,
Liaocheng University, China

*Correspondence:

Rui Huang
ruihuang@uestc.edu.cn

Received: 25 March 2020

Accepted: 19 May 2020

Published: 03 July 2020

Citation:

Peng Z, Luo R, Huang R, Yu T, Hu J, Shi K and Cheng H (2020) Data-Driven Optimal Assistance Control of a Lower Limb Exoskeleton for Hemiplegic Patients. *Front. Neurobot.* 14:37. doi: 10.3389/fnbot.2020.00037

Keywords: walking assistance control, reinforcement learning, leader-follower multi-agent system, lower limb exoskeleton, hemiplegic patients, actor-critic neural network

1. INTRODUCTION

With the increasing requirement of accomplishing complex or difficult tasks in the fields of industry and human daily life, wearable devices/robots have attracted more attentions (Fang et al., 2018, 2019). As one of wearable devices, the lower limb exoskeleton (LLE) integrates artificial intelligence technologies, control and robotic theory, and has become a hot topic own to its practical applications. Note that so many injuries caused by neuromuscular diseases, and accidents reduce the walking ability, most assistive exoskeletons are developed to aid paraplegic patients (Dollar and Herr, 2008; Strausser and Kazerooni, 2011). On the other hand, stroke has gradually become a global health-care problem, which inspires many researchers to pay attention to walking assistance or rehabilitation case for hemiplegic individuals (Ho et al., 2011; Bortole et al., 2015; Iqbal and Baizid, 2015; Louie and Eng, 2016).

From the functions point of views, the LLE can be categorized in three classes of applications, namely, strength augmentation (Walsh et al., 2006; Huang et al., 2016, 2019), walking

assistance (Esquenazi et al., 2017; Zhang et al., 2017), and rehabilitation (Sankai, 2010; Huo et al., 2014). For strength augmentation, the wearers usually have walking ability, and the influence of human-robot interaction force should be considered in the controller designs. For walking assistance case, one usually uses the LLE to assist patients' walking/training in which the patients lose their ability to walk. Thus, LLE has served as a device for rehabilitation/walking training with paraplegia and hemiplegia. In recent, some researchers have introduced biological signals of human body into the controller designs, such as Electromyography signal (EMG) (Kiguchi et al., 2004) and Electroencephalogram signal (EEG) (Kilicarslan et al., 2013).

In the early research works of rehabilitation and gait recovery of hemiplegia, researcher proposed Ankle-Foot Orthosis (AFO) to achieve good recovery performance (Tyson and Thornton, 2001; Fatone et al., 2009). In order to provide active power assistance for hemiplegic patients, many powered orthosis with active motors have been developed, such as active AFO developed by Blaya and Herr (2004) and Series Elastic Remote Knee Actuator (SERKA) developed by Sulzer et al. (2009). However, these kinds of orthosis are designed for repairing local motion function of hemiplegic patients in particular scenarios, such as the SERKA is design for stroke patient with stiff-knee gait (SKG).

For the assistance control problem of LLE with hemiplegia, one usually focus on how to derive the LLE to generate a normal motion that aid the patients walking or recovering (Maciejasz et al., 2014; Hassan et al., 2018). Sankai developed a single leg exoskeleton system for hemiplegic patients based on the Hybrid Assistive Limb (HAL) (Kawamoto et al., 2009). For the studies on the HAL system with single leg case, motion information of the unaffected side is generated to synchronize gait of the affected side (Kawamoto et al., 2014). Note that the single leg based HAL system should be re-designed as the wearer has different disabled side. In Fisher et al. (2011), a powered exoskeleton was used to improve patients with hemiparesis walking function via robot assisted gait training. In Murray et al. (2014), the authors proposed a control approach of a LLE to provide walking assistance, without giving desired joint angle trajectory, for facilitating recovery. More recently, the walking assistance control problem for a LLE with hemiplegia was investigated via a learning-based control method (Huang et al., 2018).

In most of the existing relevant works, the case of disturbances caused by system or external environment has not been taken into consideration in the designs of controllers. In fact, disturbances caused by system or external environment will affect the control performance of system, which should be considered in controller designs. On the other hand, the precise system dynamics of exoskeleton is difficult to establish, which decreases the control performance of the *model-based* methods in real systems. To solve this issue, the system identification is needed that would introduce new approximation errors. Therefore, the motivation of this paper aims to address these problems.

Motivated by the above-mentioned discussions and observations. In this paper, a data-driven optimal control (DDOC) strategy is proposed for walking assistance of lower exoskeleton with hemiplegic patients. First, the interaction communications between the both two low limbs of LLE and

hemiplegic patient are modeled as a *leader-follower multi-agent system* (LFMAS) framework. Then, a policy iteration (PI) algorithm is employed to compute the optimal assistance controller. Further, in order to improve adaptive performance for walking assistance with different hemiplegic patients, a RL method, called actor-critic neural network (AC/NN), is proposed to achieve better control performance, where the learning process only relays on measurement data from the LLE system. The main contributions of this paper can be summarized as follows:

1. Different from most of the existing control method which is designed in a model-based fashion, a DDOC strategy based on PI algorithm is proposed to learn the optimal assistance controller for walking. The proposed method is designed in a model-free manner without the requirement of the complete knowledge about the accurate dynamics of the exoskeleton system and system identification.
2. An adaptive online-learning based AC/NN structure is employed for the implementation of the controller design, which aims to perform adaptability performance for different patients and achieve good robust against disturbances.

Moreover, the proposed DDOC method is validated through a two degree-of-freedom (2-DOF) simulation environment, and then it is successfully applied on a real LLE system with healthy subjects who simulate paraplegia. Both simulation and experimental results verify that the proposed control approach has robustness performance against disturbances and has adaptive ability for different wearers or even the same wearer with different gait patterns.

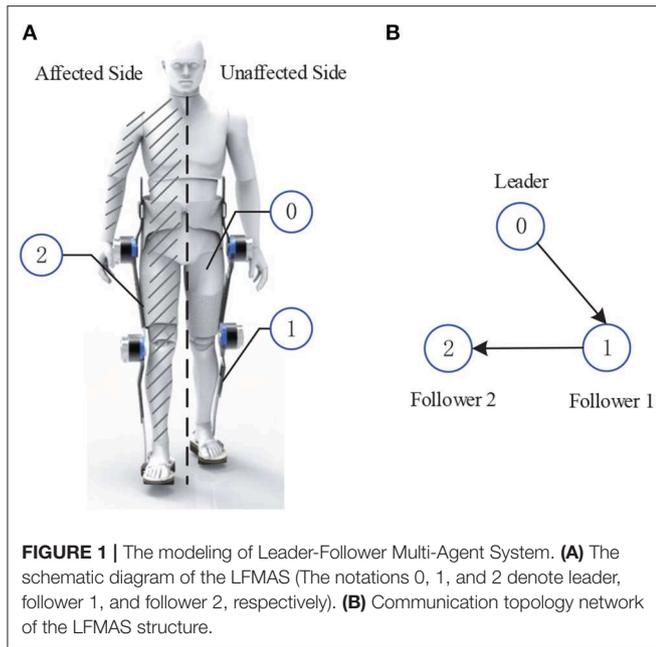
The rest of this paper is organized as follows. In section 2, the modeling process of LFMAS for exoskeleton system with hemiplegic patients is established, the system dynamics of the exoskeleton and problem formulation are given. Then, section 3 proposes the PI based optimal assistance controller designs. Section 4 proposes the data-driven adaptive control strategy by making using of RL framework on the basis of the PI algorithm. In section 5, the proposed control methods are illustrated in simulation scenario and is applied to an actual exoskeleton system with healthy people who simulate hemiplegic patients in section 6. Section 7 gives the conclusions and future work.

2. MODELING AND PROBLEM FORMULATION

In this section, the modeling process for the LLE with hemiplegic patients, namely LFMAS, is given to describe the interaction relations among both lower limbs of LLE and patients' legs. An information exchange rule is introduced for the LFMAS. Then, the system dynamics and control problem are formulated.

2.1. Modeling Exoskeleton System as LFMAS

In this paper, the focus is aim at designing an adaptive assistance controller of a LLE system with both lower extremities to assist hemiplegic individual walking. It should be noted that, for



hemiplegic patients, one of the two legs usually loses walking ability and the other one is normal. Therefore, before introducing the controller designs, it is necessary to tackle how to model the interaction relations among them appropriately such that both low limbs and the LLE can achieve their mutual communication.

In light of the cooperative distributed control, *leader-follower* mechanism has been widely utilized in multi-agent systems control (Hu and Feng, 2010), where the main idea is that information interactions among agents are achieved in a distributed way. In this paper, this mechanism is extended to model the unaffected leg of the hemiplegic patients and the both lower extremities of exoskeleton as a LFMAS, where the structure of the LFMAS for exoskeleton system with hemiplegic individuals is illustrated in **Figure 1A**. That is, the exoskeleton with hemiplegia is divided into three components: one leader agent and two follower agents. In other words, the unaffected leg of patient is regarded as the leader of LFMAS, equipped with an Inertial Measurement Unit (IMU) sensors for measuring its joints' states. Furthermore, both two lower extremities of the LLE system are defined as two follower agents, i.e., follower 1 and follower 2 which can be described as follows:

1. *Follower 1* is the exoskeleton leg of unaffected side, which synchronizes the leader agent's (unaffected side of patient's leg) motion immediately.
2. *Follower 2* is the other side of exoskeleton's limb with the disabled leg of patient, the patient's disabled leg is tightly connected with the exoskeleton.

In the framework of LFMAS, it should be pointed that there is a phase difference between the motion of the affected side and the unaffected side, naturally. In other words, follower 1 first synchronize to the leader's motion trajectory and then follower 2 is expected to track to the leader's trajectory motion after half gait cycle interval.

To guarantee walking assistance control performance, on the basis of LFMAS, the information interaction scheme should be designed for both lower extremities and patient's legs, which means that the information/data (LLE's state and control signal) can be transmitted among them. To this end, the following gives an information exchange rule to describe the evolution of the agents' communication.

- (i) **Information Evolution Rule:** The information update for follower agent i ($i = 1, 2$) includes combining its own information with those received from its neighbors, and Leader can transmit its information to Follower. Assume that each agent has a weight vector $a_i = [a_{ij}]$, in which each element a_{ij} represents that agent i assigns to the information obtained from a neighboring agent j . **Figure 1B** denotes the communication topology network between agents where arrows indicate the direction of information flow.
- (ii) **Weight Rule:** Let $\mathcal{N}(i)$ be the neighbors set of the i th Follower agent. For arbitrary $i \in \{1, 2\}$, if $j \in \mathcal{N}(i)$, $a_{ij} > 0$; if $j \notin \mathcal{N}(i)$, $a_{ij} = 0$. Let $\sum_{j \in \mathcal{N}(i)} a_{ij} = d_i$ be the sum of the neighbors' weights for agent i .

2.2. Dynamics Model of LLE System

In this paper, the dynamics of the LLE system is described as a second-order nonlinear mechanical system (i.e., Euler-lagrange system). Therefore, the dynamics of the both lower extremities, i.e., follower 1 ($i = 1$) and follower 2 ($i = 2$) of the exoskeleton are described as follows:

$$H_i(q_i)\ddot{q}_i + C_i(q_i, \dot{q}_i)\dot{q}_i + G_i(q_i) = \tau_i, \quad i = 1, 2 \quad (1)$$

where $q_i = (q_{ih}, q_{ik})^T \in R^2$ denotes the joints' angle of the LLE, q_{ih} and q_{ik} represent the hip joint and knee joint, respectively. $H_i(q_i)$ denotes inertia matrix, $C_i(q_i, \dot{q}_i)$ represents the centripetal and coriolis matrix. $G_i(q_i)$ denotes the gravitation term, $\tau_i = (\tau_{iu}, \tau_{id})^T$ are the input torques generated by up and down motors for hip and knee joint. Further, we can rewrite Equation (1) as a state-space form:

$$\begin{bmatrix} \dot{q}_i \\ \ddot{q}_i \end{bmatrix} = \begin{bmatrix} \dot{q}_i \\ -H_i^{-1}(C_i\dot{q}_i + G_i) \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ H_i^{-1} \end{bmatrix} \tau_i$$

or equivalently,

$$\dot{\eta}_i(t) = f_i(\eta_i(t)) + g_i(\eta_i(t))u_i, \quad (2)$$

where $\eta_i(t) = [q_i^T, \dot{q}_i^T]^T$, $g_i(\eta_i(t)) = \begin{bmatrix} \mathbf{0} \\ H_i^{-1} \end{bmatrix}$, $f_i(\eta_i(t)) = \begin{bmatrix} \mathbf{0} & I \\ \mathbf{0} & -H_i^{-1}C_i \end{bmatrix} \eta_i(t) + \begin{bmatrix} \mathbf{0} \\ -H_i^{-1}G_i \end{bmatrix}$, $\tau_i = u_i$.

The dynamic of the leader (the motion trajectory of patient's unaffected leg) is given by:

$$\dot{\eta}_r(t) = f(\eta_r(t)), \quad (3)$$

where $\eta_r(t)$ indicates the joint angle collected from human via an IMU sensors matched on the pilot's leg.

2.2.1. Design Objective

The goal is to generate the controller strategy u_i to ensure the trajectory $\eta_i(t)$ generated by Equation (2) can track the trajectory $\eta_r(t)$ in Equation (3). That is, it is desired to make the following tracking error index go to zero:

$$\lim_{t \rightarrow \infty} \|\eta_i(t) - \eta_r(t)\| = 0. \quad (4)$$

In order to achieve control objective, the local neighbor tracking errors of dynamics (2) for follower i are defined as

$$\xi_i(t) = \sum_{j \in \mathcal{N}(i)} a_{ij}(\eta_i(t) - \eta_j(t)) + c_i(\eta_i(t) - \eta_r(t)), \quad (5)$$

where $\mathcal{N}(i)$ and a_{ij} have been defined in section 2.1. $c_i > 0$ denotes the pinning gain, which means agent i can obtain the Leader's information.

Taking the derivation of Equation (5), combining Equation (2) and Equation (3), the dynamics of the tracking errors are written as

$$\dot{\xi}_i(t) = f_{\xi_i} + (d_i + c_i)g_i(t)u_i(t) - \sum_{j \in \mathcal{N}(i)} a_{ij}g_j(t)u_j(t), \quad (6)$$

where $f_{\xi_i}(t) = \sum_{j \in \mathcal{N}(i)} a_{ij}(f_i - f_j) + c_i(f_i - f)$, d_i indicates the sum of the weights of the i^{th} follower agent.

3. POLICY ITERATION BASED CONTROLLER

Based on the system modeling and problem formulation, in this section, the walking assistance control problem will be transformed to an optimal control problem by introducing local cost functions and using optimization theories. Then, the state-of-the-art algorithm called policy iteration (PI) is proposed to obtain the solution to the coupled Hamilton-Jacobi-Bellman (HJB) equation, and thus the optimal controller $u_i^*(t)$ is obtained for solving walking assistance problem.

From the perspective of optimal control (Vamvoudakis and Lewis, 2010) and inspired by RL methods (Mnih et al., 2015, 2016; Sutton and Barto, 2018), we use a local cost function to assess the long-term learning and control performance, which is defined as follows:

$$V_i(\xi_i(t)) = \int_t^{\infty} r_i(\xi_i(s), u_i(s), u_{(j)}(s)) ds, \quad (7)$$

where $u_{(j)}(t)$ denotes the neighbors' control of Follower agent i , and $r_i(\xi_i(t), u_i(t), u_{(j)}(t)) = \xi_i^T(t)Q_{ii}\xi_i(t) + u_i^T(t)P_{ii}u_i(t) + \sum_{j \in \mathcal{N}(i)} u_i^T(t)S_{ij}u_{(j)}(t)$ is the reward function, where the $Q_{ii} > 0$, $P_{ii} > 0$ and $S_{ij} > 0$ are symmetric positive definite weighting matrices, respectively. For the notation simplification, we set $r_i(\xi_i(t), u_i(t), u_{(j)}(t)) = r_i(\xi_i(t), u_i(t))$.

Till now, the walking assistance control problem is transformed into an optimal control problem, which aims to design a distributed controller to guarantee the *Design Objective* as well as minimizing the local cost function (Equation 7).

Further, the Hamilton function is represented as

$$H_i(\xi_i(t), u_i(t), V_i(\xi_i(t))) = r_i(\xi_i(t), u_i(t)) + \nabla V_i^T \xi_i(t), \quad (8)$$

where $V_i(0) = 0$, $\nabla V_i = \partial V_i(\xi_i(t))/\partial \xi_i(t)$ is a partial differential part.

Using the stationary condition for Equation (8), i.e., let $\partial H_i(t)/\partial u_i(t) = 0$, the optimal controller $u_i^*(t)$ is obtained as

$$u_i^*(t) = -\frac{1}{2}(d_i + c_i)P_{ii}^{-1}g_i^T(t)\nabla V_i. \quad (9)$$

The optimal cost function $V_i^*(\xi_i(t))$ satisfies the following coupled Hamilton-Jacobi-Bellman (HJB) equation:

$$H_i(\xi_i(t), u_i^*(t), V_i^*(\xi_i(t))) = r_i(\xi_i(t), u_i^*(t)) + \nabla V_i^{*T} \xi_i(t) = 0. \quad (10)$$

Since the coupled HJB equation Equation (10) exists the nonlinear item and partial differential part, which makes it hard to be solved analytically. Therefore, the PI algorithm (Liu and Wei, 2014; Wang et al., 2014), is introduced to approximate the HJB equation and cope with this issue by a successive iteration way.

Let $u_i^l(t)$ and $V_i^l(\xi_i(t))$ represent iterative control and iterative Q-function, respectively, with l is iteration index. There are two components in PI algorithm, one is *policy evaluation* and the other is *policy improvement*. The detailed iterative performing process can be summarized as follows:

PI Algorithm: Start with admissible initial control u_i^0 .

Step 1. *Policy Evaluation:* Given the control policy u_i^l , solve for value function $V_i^l(\xi_i(t))$ by

$$H_i(\xi_i(t), u_i^l(t), V_i^l(\xi_i(t))) = r_i(\xi_i(t), u_i^l(t)) + \nabla V_i^{lT} \xi_i(t) = 0. \quad (11)$$

Step 2. *Policy Improvement:* Compute the control law u_i^l by

$$u_i^{l+1}(t) = -\frac{1}{2}(d_i + c_i)P_{ii}^{-1}g_i^T(t)\nabla V_i^l. \quad (12)$$

Step 3. If $\|V_i^l(\xi_i) - V_i^{l-1}(\xi_i)\| \leq \epsilon$ (ϵ is a small positive constant), end. Else, let $l = l + 1$, go to step 1.

The PI algorithm is an effective method to solve the various optimal control problems. It has been proved that the iterative cost function and the iterative control strategy in PI will converge to the optimal values $V_i^*(t)$ and $u_i^*(t)$ through iterations (Peng et al., 2019, 2020).

It is worth noting from the above algorithm that the PI algorithm requires the knowledge of system models, i.e., $g_i(t)$ exists in the controller (Equation 12). In this sense, system identification is needed normally (Ghan and Kazerooni, 2006), but it is not suitable for the practical exoskeleton system with different hemiplegic patient. Since for different wearers/patients, the identification process needs to be reconstructed. To overcome this difficulty, the following section will present a data-driven adaptive control strategy with an online-learning fashion. It should be emphasized that this method avoids needing the knowledge of the accurate system dynamics, and no system identification is introduced.

4. IMPLEMENTATION OF CONTROLLER DESIGN

In this section, we will present the DDOC algorithm base on PI algorithm to achieve online-learning-based control and better adaptive performance for different patients via a neural network (NN) framework of RL called AC/NN. In the AC/NN, actor network is used to approximate controller and critic network is introduced to estimate cost function online, respectively. The detailed descriptions are given as follows.

4.1. The Critic NN Modular

First, the critic networks are adopted to approximate the cost function $V_i(t)$ as follows:

$$\hat{V}_i(t) = \hat{w}_{ci}^\top(t) \varphi_{ci}(z_{ci}(t)), \quad (13)$$

where z_{ci} is an input information of the critic modular and information from ξ_i , u_i , and $u_{(j)}$, $\varphi_{ci}(z_{ci})$ denotes the activation function, and \hat{w}_{ci} is the weight vector of the critic network modular.

Then, at each time step, the Hamilton function (8) can be approximated as follows:

$$\begin{aligned} \mathbf{e}_{ci}(t) &= \int_t^{t+\Delta T} r_i(\xi_i, u_i) ds + \hat{V}_i(t + \Delta T) - \hat{V}_i(t) \\ &= \int_t^{t+\Delta T} r_i(\xi_i, u_i) ds + \hat{w}_{ci}^\top(\varphi_{ci}(z_{ci}(t + \Delta T)) - \varphi_{ci}(z_{ci}(t))), \end{aligned} \quad (14)$$

where $\Delta T > 0$ denotes the time interval.

Then, the Equation (14) is utilized to define the approximation error for the critic NNs. Thus, the squared residual error function to be minimized is defined as

$$E_{ci}(t) = \frac{1}{2} \|\mathbf{e}_{ci}(t)\|^2 = \frac{1}{2} \mathbf{e}_{ci}^2(t). \quad (15)$$

Then, by making use of gradient descent based weight update rule (Si and Wang, 2001), the tuning weight law can be adopted as follows

$$\begin{aligned} \dot{\hat{w}}_{ci}(t) &= -\varrho_{ci} \frac{\partial E_{ci}(t)}{\partial \mathbf{e}_{ci}(t)} \frac{\partial \mathbf{e}_{ci}(t)}{\partial \hat{V}_i(t)} \frac{\partial \hat{V}_i(t)}{\partial \hat{w}_{ci}(t)} \\ &= -\varrho_{ci} \varphi_{ci}(z_{ci}) (\hat{w}_{ci}^\top \Delta \varphi_{ci}(z_{ci}) + \int_t^{t+\Delta T} r_i(\xi_i, u_i) ds), \end{aligned} \quad (16)$$

where $\Delta \varphi_{ci}(z_{ci}) = \varphi_{ci}(z_{ci}(t + \Delta T)) - \varphi_{ci}(z_{ci}(t))$, ϱ_{ci} is the learning rate of the critic network modular for agent i .

4.2. The Actor NN Modular

Next, define the actor neural networks, which is employed to estimate the control strategy, as follows:

$$\hat{u}_i(t) = \hat{w}_{ai}^\top(t) \varphi_{ai}(z_{ai}(t)), \quad (17)$$

where z_{ai} is an input vector of the actor network including ξ_i of agent i , $\varphi_{ai}(z_{ai})$ denotes the activation function, and \hat{w}_{ai} is the weight matrix.

Algorithm 1: Optimal Walking Assistance Control Algorithm.

- 1: **Initialization**
- 2: Initialize the values of critic weight $\hat{w}_{ci}(0)$ and actor weight $\hat{w}_{ai}(0)$;
- 3: Set the learning rates of the critic network and actor network to be ρ_{ai} and ρ_{ci} ;
- 4: Choose a sufficiently small computation precision ϵ ;
- 5: Let Q_{ii} , P_{ii} and S_{ij} be positive definite weighting matrices;
- 6: **repeat**
- 7: Calculate the actor network to estimate the control strategy $\hat{u}_i \leftarrow (17)$;
- 8: Calculate the critic network to estimate the cost function $\hat{V}_i \leftarrow (13)$;
- 9: According to the available system data q_i and q_r , compute the error $\xi_i \leftarrow (5)$;
- 10: Calculate the objective function E_{ci} ;
- 11: Update the weights in the critic NNs using $\hat{w}_{ci}(t) \leftarrow (16)$;
- 12: Calculate the objective function E_{ai} ;
- 13: Update the weights in the actor NNs using $\hat{w}_{ai}(t) \leftarrow (20)$;
- 14: **until** $\|\hat{w}'_{ci} - \hat{w}_{ci}\| \leq \epsilon$ (\hat{w}_{ci} and \hat{w}'_{ci} denote the weight of the current time and previous time);

Then, in order to obtain the desired approximation optimal controller to minimize the cost function \hat{V}_i , the error function of the actor network is defined as

$$\mathbf{e}_{ai}(t) = \hat{V}_i(\xi_i(t)) - \mathcal{U}_{obj}, \quad (18)$$

where \mathcal{U}_{obj} is the ultimate objective function. From perspective of the RL, the value of the \mathcal{U}_{obj} is selected according to different purposes of applications.

The squared residual error function to be minimized in the actor network is given by

$$E_{ai}(t) = \frac{1}{2} \|\mathbf{e}_{ai}(t)\|^2 = \frac{1}{2} \mathbf{e}_{ai}^2(t). \quad (19)$$

Similarly, with the aid of the gradient descent rule, the following updating rule for the actor network is obtained

$$\begin{aligned} \dot{\hat{w}}_{ai}(t) &= -\varrho_{ai} \frac{\partial E_{ai}(t)}{\partial \mathbf{e}_{ai}(t)} \frac{\partial \mathbf{e}_{ai}(t)}{\partial \hat{V}_i(t)} \frac{\partial \hat{V}_i(t)}{\partial z_{ci}(t)} \frac{\partial z_{ci}(t)}{\partial \hat{u}_i(t)} \frac{\partial \hat{u}_i(t)}{\partial \hat{w}_{ai}(t)} \\ &= -\varrho_{ai} \varphi_{ai}(z_{ai}) \hat{w}_{ci}^\top \nabla \varphi_{ci}(z_{ci}) \xi_i \varphi_{ci}^\top(z_{ci}) \hat{w}_{ci}, \end{aligned} \quad (20)$$

where $\xi_i = \partial z_{ci} / \partial \hat{u}_i$, $\nabla \varphi_{ci}(z_{ci}) = \partial \varphi_{ci}(z_{ci}) / \partial z_{ci}$ and ϱ_{ai} is a learning rate of the actor NN for agent i .

The procedure of the data-driven adaptive control strategy is presented in Algorithm 1. It should be noted that only the measured system data, i.e., ξ_i and u_i are required in the design of the DDOC algorithm. Thus, this method is a data-driven/model-free approach, which improves the potential application of the proposed control method in real systems.

It is noted that in the neural network based approximated structure, some common forms of the activation functions are polynomial functions, tanh functions, sigmoid functions, and so

on. Further, we found that the appropriate selection of activation function is very important, which leads to the different size of the NN weight parameters. In this paper, the selection of activation function has the same dimension as the input data. These settings can reduce a huge computation burden for implementation.

5. NUMERICAL SIMULATION

In this section, a 2-DOF manipulator system in simulation scenario is first carried out to validate the effectiveness of the proposed data-driven control strategy.

5.1. Dynamic Model of 2-DOF System

For simulation, the simulation environment is set up in Simulink-Matlab. The dynamics of the two followers are the same as Equation (1), where the system matrices are given as follows: $H_i = \begin{bmatrix} m_{i1} + m_{i2} + 2m_{i3} \cos(q_{i2}) & m_{i2} + m_{i3} \cos(q_{i2}) \\ m_{i2} + m_{i3} \cos(q_{i2}) & m_{i2} \end{bmatrix}$, $C_i = \begin{bmatrix} -m_{i3} \dot{q}_{i2} \sin(q_{i2}) & -m_{i3} (\dot{q}_{i1} + \dot{q}_{i2}) \sin(q_{i2}) \\ -m_{i3} \dot{q}_{i1} & 0 \end{bmatrix}$, and the $G_i = \begin{bmatrix} m_{i4} g \cos(q_{i1}) + m_{i5} g \cos(q_{i1} + q_{i2}) \\ m_{i5} g \cos(q_{i1} + q_{i2}) \end{bmatrix}$, $\tau_i = [\tau_{i1}, \tau_{i2}]^T$, m_{ip} ($p = 1, 2, 3, 4, 5$) are the masses. Note that, in simulation case, the given dynamic system can be used to product system data needed in DDOC algorithm.

The leader system (desired trajectories) is expressed by

$$q_r = \begin{bmatrix} q_{1r} \\ q_{2r} \end{bmatrix} = \begin{bmatrix} 0.5 \cos(t) + 0.2 \sin(3t) \\ 0.3 \cos(3t) - 0.5 \sin(2t) \end{bmatrix}. \quad (21)$$

We select the structure of the AC/NN as 3-layers back propagation (BP) NN (Goh, 1995). The initial values of critic NN weights and actor NN weights are set to be zero, and setting the value of the computation precision as $\epsilon = 10^{-5}$. The weight learning rates of the actor network and the critic network are chosen as $\rho_{ai} = 0.03$, $\rho_{ci} = 0.06$. The activation functions φ_{ai} and φ_{ci} are selected as the hyperbolic tangent functions, i.e., $\tanh(x) = (e^x - e^{-x}) / (e^x + e^{-x})$.

5.2. Simulation Results and Analysis

As shown in **Figure 2**, we can see that after 2 s learning process, the critic NN weights and the actor NN weights are convergent, and thus the optimal weights parameters are obtained. Therefore, the approximate optimal controller can be obtained in Equation (17). On the basis of the optimal controller, the trajectory of joint angles $q_1 = (q_{11}, q_{12})^T$ of follower 1 achieves a good tracking performance to the leader at $3 \text{ s} < t < 6 \text{ s}$, which is illustrated in **Figure 3**.

In order to further verify the performance of adaption against uncertainty, we add some disturbance signal (white noise) to the system at $t \in [6, 7] \text{ s}$. In **Figure 2**, the AC/NN weights are retrained for learning again adaptively until converge from $t = 6 \text{ s}$ to $t = 8 \text{ s}$, and thus the optimal controller has been modified correspondingly. With the help of the modified optimal controller, from **Figure 3**, it is seen that joint angle trajectories of two links of follower 1 are synchronized with the leader again quickly after $t = 8 \text{ s}$. These simulation results illustrate the

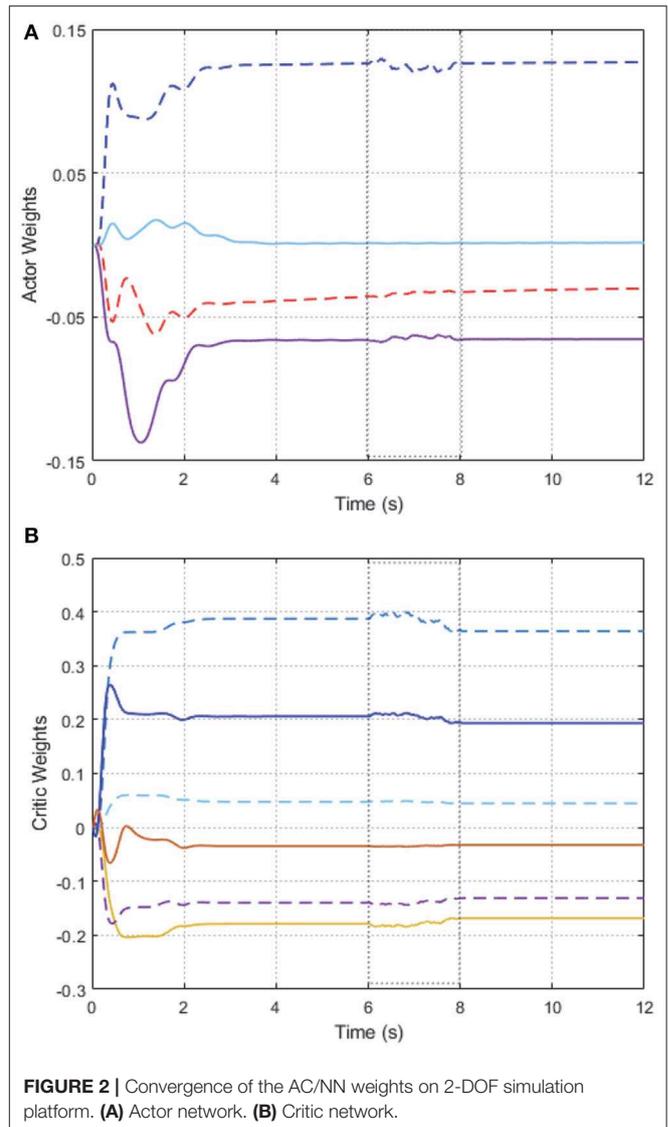


FIGURE 2 | Convergence of the AC/NN weights on 2-DOF simulation platform. (A) Actor network. (B) Critic network.

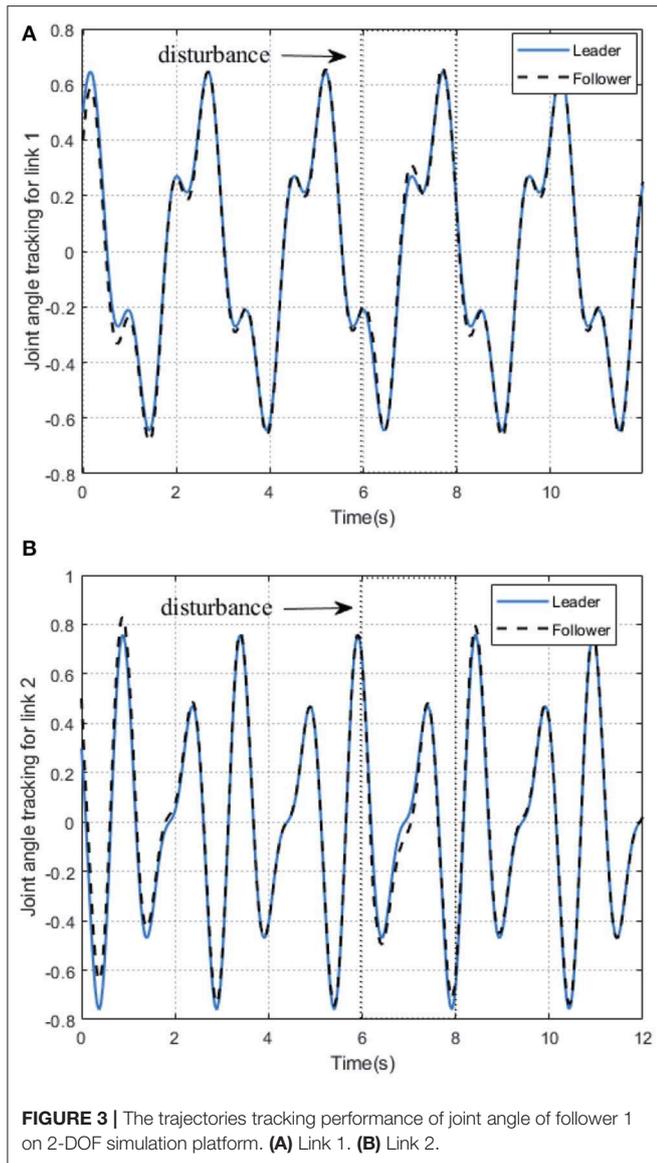
better control performance of the proposed DDOC algorithm, which has ability to respond to disturbances online in the system operation. It is proved that our proposed control method has good robustness against uncertainties.

6. EXPERIMENTS ON A REAL LLE SYSTEM

In this section, to further verify the control performance of the proposed data-driven control strategy, *walking assistance* experiments on an actual LLE system are performed.

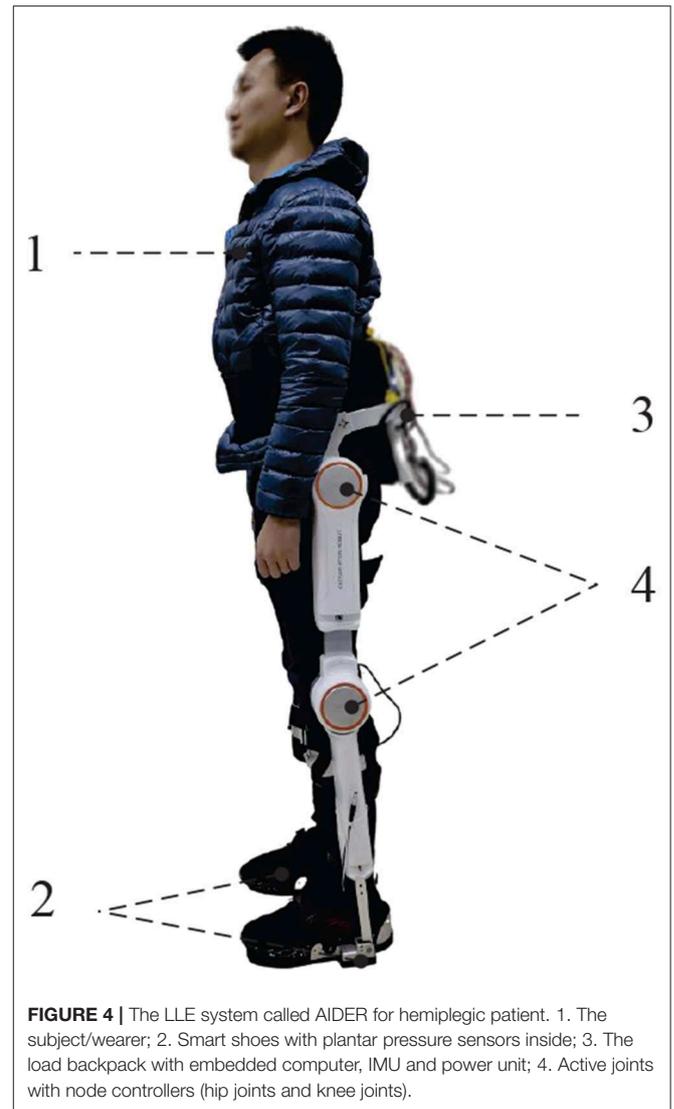
6.1. Experimental Setup

To demonstrate the effectiveness and adaptability of the proposed control strategy, a practical LLE system, called AIDER, which is shown in **Figure 4**, is designed for walking assistance case to help hemiplegia. A distributed control system is embedded in AIDER



which consists of a main controller and four node controllers. The main controller is placed on the backpack to compute the control algorithm. Node controllers are fixed near by the corresponding active joints position of LLE robot which aims to receive sensor data and execute control commands according to the main controller.

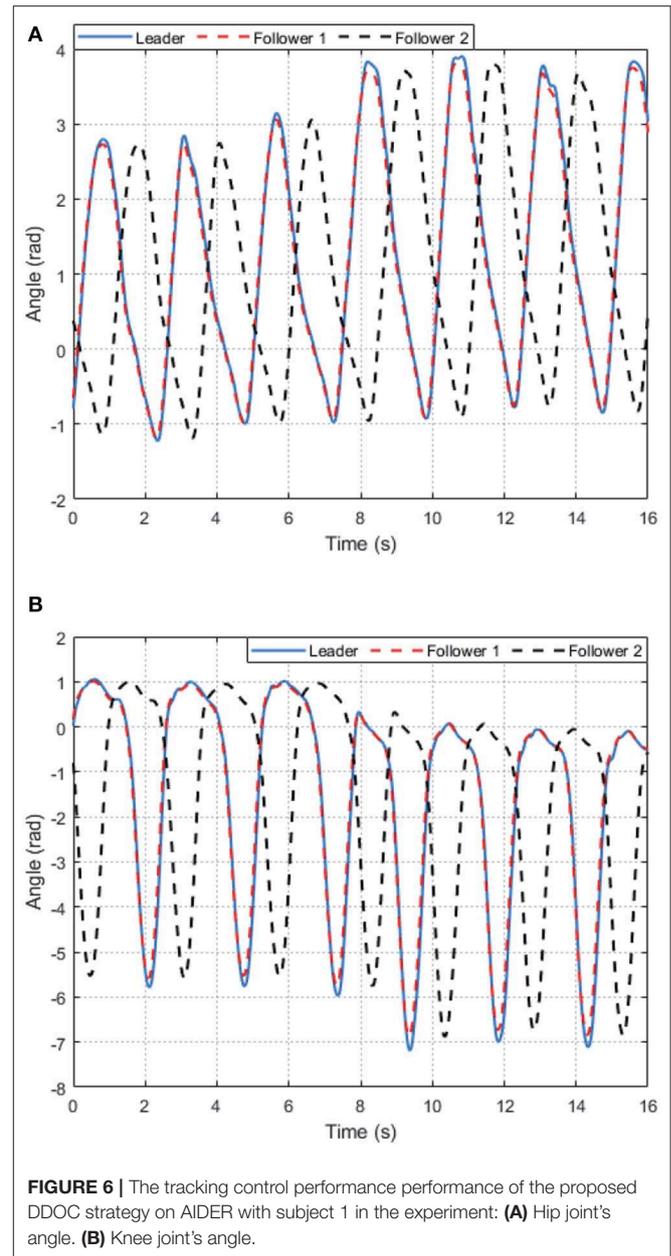
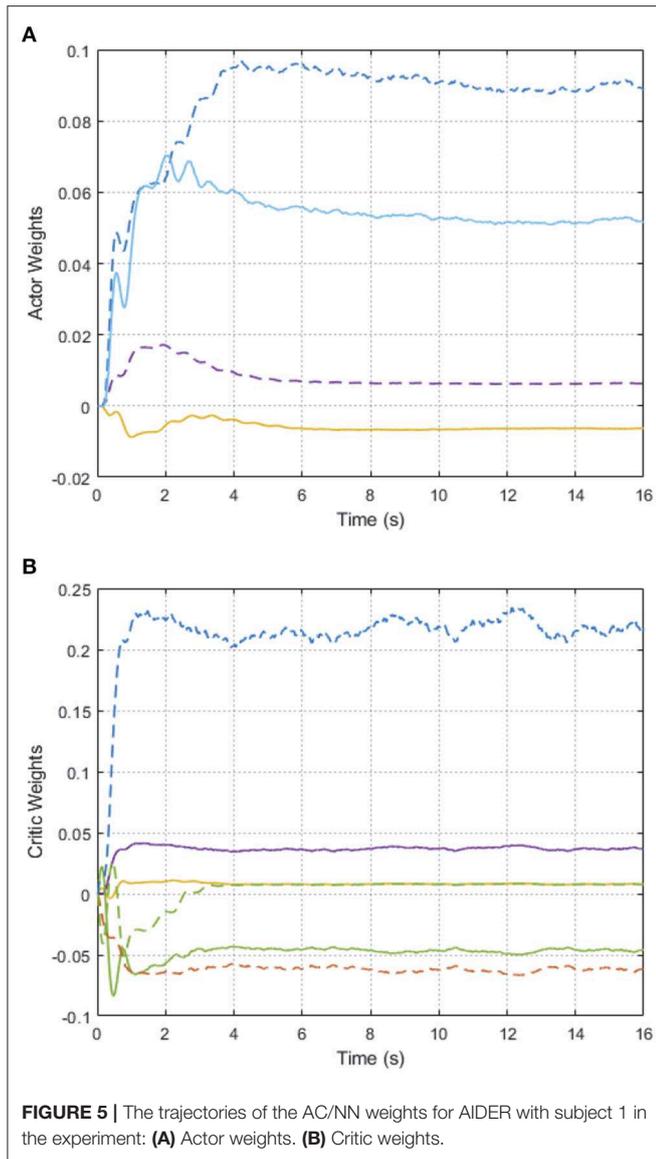
During the experiments on the AIDER system, three healthy subjects (1, 2, 3) with different heights (165, 176, 180 cm) are selected to participate this experiment and operate the wearable LLE robot. All wearers are simulated as hemiplegic patients, and the right legs of the subjects are simulated as the affected leg. In the walking assistance task for all wearers, each wearer is asked to walk for 50 s via the AIDER system. All the pilot's walking speed is varying from 0.1 to 0.4m/s. Further, the AIDER is equipped with accelerometer and the wearable sensory system for measuring system data.



For the implementation of the proposed data-driven control strategy on the AIDER system. Note that the proposed data-driven control strategy DDOC has a learning process using the online system data at the beginning, which aims to adapt different subjects. After the learning stage, the optimal control policies can be obtained, and then walking assistance can be realized for the LLE system with pilots. We choose the AC/NN as 3-layer Back propagation (BP) NNs structure (Goh, 1995), that is, input layer, hidden layer and output layer. The initial values of weights \hat{w}_{ci} and \hat{w}_{ai} of the critic and actor are all set to be zero, and the activation functions φ_{ai} and φ_{ci} are chosen as hyperbolic tangent functions $\tanh(x) = (e^x - e^{-x}) / (e^x + e^{-x})$. The learning rates are the same as in the simulation, that is $\rho_{ai} = 0.03$, $\rho_{ci} = 0.06$.

6.2. Experimental Results and Discussions

For participant 1, from **Figures 5A,B**, we can see that, after about 5 s training, the weights of AC/NN are bounded convergent, i.e., uniformly ultimately bounded because of the disturbances



and uncertainties always exist in LLE system. The tracking performance of the hip joint and knee joint for the LLE system with wearer 1 is depicted in **Figures 6A,B**, which states that with the help of the learned optimal control policies, the hip joint and knee joint of two limbs of the exoskeleton (follower 1,2) can achieve synchronization with the desired (leader's) motion trajectories. Moreover, it is noted that there are different walking motion patterns in the procedure of walking, which means our proposed method has capability of adapting different gait patterns. It should be pointed out that the affected side of wearer with exoskeleton's side (follower 2) has a half gait cycle delay to the side which has walking ability (leader), which is marked with blue dashed line as shown in **Figure 6**. In summary, the experimental results illustrate the effectiveness of the proposed DDOC approach for walking assistance of the exoskeleton with different pilots.

7. CONCLUSIONS

In this paper, a DDOC control strategy has been proposed for a lower exoskeleton system to assist hemiplegic patient walking. A LFMAS structure has been established to model the interaction relation among LLE system and hemiplegic individual. The walking assistance problem has been transformed to an optimal control problem. The PI algorithm has been introduced to obtain optimal assistance controller. On the basis of the PI algorithm, in order to adapt different patients, the AC/NN framework has been presented for the implementation of the proposed approach in an online-learning manner. It highlights that the controller design only relies on the measured system data, rather than the

accurate system model. Finally, we have successfully validated the proposed method on two situations: *2-DOF manipulator* in simulation environment and *walking assistance* experiment on a real LLE system called AIDER. Experimental results have confirmed the effectiveness of the proposed control method. In the future, we will focus on more practical control issues, and consider the RL-based controller designs for exoskeleton system with actuator faults and input time-delay.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary materials, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

All procedures performed in studies were approved by Research Ethical Committee of University of Electronic Science and Technology of China. Written informed consent

was obtained from the participant for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

ZP designed the control methods, performed the experiments, and drafted the manuscript. RL, RH, and KS participated in the design of the controllers and assisted in the manuscript writing. JH and HC guided writing paper and doing experiments. TY designed procedure of experiments and helped to revise and improve the paper significantly.

FUNDING

This work was made possible by support from National Key Research and Development Program of China (No. 2017YFB1302300), National Natural Science Foundation of China (NSFC) (Nos. 6150020696, 61503060), and Sichuan Science and Technology Program (No. 2020YFSY0012).

REFERENCES

- Blaya, J. A., and Herr, H. (2004). Adaptive control of a variable-impedance ankle-foot orthosis to assist drop-foot gait. *IEEE Trans. Neural Syst. Rehabil. Eng.* 12, 24–31. doi: 10.1109/TNSRE.2003.823266
- Bortole, M., Venkatakrishnan, A., Zhu, F., Moreno, J. C., Francisco, G. E., Pons, J. L., et al. (2015). The H2 robotic exoskeleton for gait rehabilitation after stroke: early findings from a clinical study. *J. NeuroEng. Rehabil.* 12:54. doi: 10.1186/s12984-015-0048-y
- Dollar, A. M., and Herr, H. (2008). Lower extremity exoskeletons and active orthoses: challenges and state-of-the-art. *IEEE Trans. Robot.* 24, 144–158. doi: 10.1109/TRO.2008.915453
- Esquenazi, A., Talaty, M., and Jayaraman, A. (2017). Powered exoskeletons for walking assistance in persons with central nervous system injuries: a narrative review. *PM&R* 9, 46–62. doi: 10.1016/j.pmrj.2016.07.534
- Fang, B., Sun, F., Liu, H., and Liu, C. (2018). 3D human gesture capturing and recognition by the immu-based data glove. *Neurocomputing* 277, 198–207. doi: 10.1016/j.neucom.2017.02.101
- Fang, B., Wei, X., Sun, F., Huang, H., Yu, Y., and Liu, H. (2019). Skill learning for human-robot interaction using wearable device. *Tsinghua Sci. Technol.* 24, 654–662. doi: 10.26599/TST.2018.9010096
- Fatone, S., Gard, S. A., and Malas, B. S. (2009). Effect of ankle-foot orthosis alignment and foot-plate length on the gait of adults with poststroke hemiplegia. *Arch. Phys. Med. Rehabil.* 90, 810–818. doi: 10.1016/j.apmr.2008.11.012
- Fisher, S., Lucas, L., and Adam Thrasher, T. (2011). Robot-assisted gait training for patients with hemiparesis due to stroke. *Top. Stroke Rehabil.* 18, 269–276. doi: 10.1310/tsr1803-269
- Ghan, J., and Kazerooni, H. (2006). “System identification for the Berkeley lower extremity exoskeleton (bleex),” in *IEEE International Conference on Robotics and Automation* (Orlando, FL), 3477–3484. doi: 10.1109/ROBOT.2006.1642233
- Goh, A. T. (1995). Back-propagation neural networks for modeling complex systems. *Artif. Intell. Eng.* 9, 143–151. doi: 10.1016/0954-1810(94)00011-S
- Hassan, M., Kadone, H., Ueno, T., Hada, Y., Sankai, Y., and Suzuki, K. (2018). Feasibility of synergy-based exoskeleton robot control in hemiplegia. *IEEE Trans. Neural Syst. Rehabil. Eng.* 26, 1233–1242. doi: 10.1109/TNSRE.2018.2832657
- Ho, N., Tong, K., Hu, X., Fung, K., Wei, X., Rong, W., et al. (2011). “An EMG-driven exoskeleton hand robotic training device on chronic stroke subjects: task training system for stroke rehabilitation,” in *IEEE International Conference on Rehabilitation Robotics* (Zurich), 1–5. doi: 10.1109/ICORR.2011.5975340
- Hu, J., and Feng, G. (2010). Distributed tracking control of leader-follower multi-agent systems under noisy measurement. *Automatica* 46, 1382–1387. doi: 10.1016/j.automatica.2010.05.020
- Huang, R., Cheng, H., Guo, H., Chen, Q., and Lin, X. (2016). “Hierarchical interactive learning for a human-powered augmentation lower exoskeleton,” in *IEEE International Conference on Robotics and Automation* (Stockholm), 257–263. doi: 10.1109/ICRA.2016.7487142
- Huang, R., Cheng, H., Qiu, J., and Zhang, J. (2019). Learning physical human-robot interaction with coupled cooperative primitives for a lower exoskeleton. *IEEE Trans. Autom. Sci. Eng.* 16, 1566–1574. doi: 10.1109/TASE.2018.2886376
- Huang, R., Peng, Z., Cheng, H., Hu, J., Qiu, J., Zou, C., et al. (2018). “Learning-based walking assistance control strategy for a lower limb exoskeleton with hemiplegia patients,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Madrid), 2280–2285. doi: 10.1109/IROS.2018.8594464
- Huo, W., Mohammed, S., Moreno, J. C., and Amirat, Y. (2014). Lower limb wearable robots for assistance and rehabilitation: a state of the art. *IEEE Syst. J.* 10, 1068–1081. doi: 10.1109/JSYST.2014.2351491
- Iqbal, J., and Baizid, K. (2015). Stroke rehabilitation using exoskeleton-based robotic exercisers: mini review. *Biomed. Res.* 26, 197–201.
- Kawamoto, H., Hayashi, T., Sakurai, T., Eguchi, K., and Sankai, Y. (2009). “Development of single leg version of HAL for hemiplegia,” in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Minneapolis, MN), 5038–5043. doi: 10.1109/IEMBS.2009.5333698
- Kawamoto, H., Kandone, H., Sakurai, T., Ariyasu, R., Ueno, Y., Eguchi, K., et al. (2014). “Development of an assist controller with robot suit HAL for hemiplegic patients using motion data on the unaffected side,” in *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Chicago, IL), 3077–3080. doi: 10.1109/EMBC.2014.6944273
- Kiguchi, K., Tanaka, T., and Fukuda, T. (2004). Neuro-fuzzy control of a robotic exoskeleton with EMG signals. *IEEE Trans. Fuzzy Syst.* 12, 481–490. doi: 10.1109/TFUZZ.2004.832525
- Kilicarslan, A., Prasad, S., Grossman, R. G., and Contreras-Vidal, J. L. (2013). “High accuracy decoding of user intentions using EEG to control a lower-body exoskeleton,” in *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Osaka), 5606–5609. doi: 10.1109/EMBC.2013.6610821
- Liu, D., and Wei, Q. (2014). Policy iteration adaptive dynamic programming algorithm for discrete-time nonlinear systems. *IEEE Trans. Neural Netw. Learn. Syst.* 25, 621–634. doi: 10.1109/TNNLS.2013.2281663

- Louie, D. R., and Eng, J. J. (2016). Powered robotic exoskeletons in post-stroke rehabilitation of gait: a scoping review. *J. NeuroEng. Rehabil.* 13:53. doi: 10.1186/s12984-016-0162-5
- Maciejasz, P., Eschweiler, J., Gerlach-Hahn, K., Jansen-Troy, A., and Leonhardt, S. (2014). A survey on robotic devices for upper limb rehabilitation. *J. NeuroEng. Rehabil.* 11:3. doi: 10.1186/1743-0003-11-3
- Mnih, V., Badia, A. P., Mirza, M., Graves, A., Lillicrap, T., Harley, T., et al. (2016). "Asynchronous methods for deep reinforcement learning," in *International Conference on Machine Learning* (New York, NY), 1928–1937.
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., et al. (2015). Human-level control through deep reinforcement learning. *Nature* 518, 529–533. doi: 10.1038/nature14236
- Murray, S. A., Ha, K. H., Hartigan, C., and Goldfarb, M. (2014). An assistive control approach for a lower-limb exoskeleton to facilitate recovery of walking following stroke. *IEEE Trans. Neural Syst. Rehabil. Eng.* 23, 441–449. doi: 10.1109/TNSRE.2014.2346193
- Peng, Z., Hu, J., Shi, K., Luo, R., Huang, R., Ghosh, B. K., et al. (2020). A novel optimal bipartite consensus control scheme for unknown multi-agent systems via model-free reinforcement learning. *Appl. Math. Comput.* 369:124821. doi: 10.1016/j.amc.2019.124821
- Peng, Z., Zhao, Y., Hu, J., and Ghosh, B. K. (2019). Data-driven optimal tracking control of discrete-time multi-agent systems with two-stage policy iteration algorithm. *Inf. Sci.* 481, 189–202. doi: 10.1016/j.ins.2018.12.079
- Sankai, Y. (2010). Hal: Hybrid assistive limb based on cybernics. *Robot. Res.* 66, 25–34. doi: 10.1007/978-3-642-14743-2_3
- Si, J., and Wang, Y.-T. (2001). Online learning control by association and reinforcement. *IEEE Trans. Neural Netw.* 12, 264–276. doi: 10.1109/72.914523
- Strausser, K. A., and Kazerooni, H. (2011). "The development and testing of a human machine interface for a mobile medical exoskeleton," in *IEEE/RSJ International Conference on Intelligent Robots and Systems* (San Francisco, CA), 4911–4916. doi: 10.1109/IROS.2011.6095025
- Sulzer, J. S., Roiz, R. A., Peshkin, M. A., and Patton, J. L. (2009). A highly backdrivable, lightweight knee actuator for investigating gait in stroke. *IEEE Trans. Robot.* 25, 539–548. doi: 10.1109/TRO.2009.2019788
- Sutton, R. S., and Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. MIT Press.
- Tyson, S., and Thornton, H. (2001). The effect of a hinged ankle foot orthosis on hemiplegic gait: objective measures and users' opinions. *Clin. Rehabil.* 15, 53–58. doi: 10.1191/026921501673858908
- Vamvoudakis, K. G., and Lewis, F. L. (2010). Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem. *Automatica* 46, 878–888. doi: 10.1016/j.automatica.2010.02.018
- Walsh, C., Paluska, D., Pasch, K., Grand, W., Valiente, A., and Herr, H. (2006). "Development of a lightweight, under-actuated exoskeleton for load-carrying augmentation," in *IEEE International Conference on Robotics and Automation* (Orlando, FL), 3485–3491.
- Wang, D., Liu, D., and Li, H. (2014). Policy iteration algorithm for online design of robust control for a class of continuous-time nonlinear systems. *IEEE Trans. Autom. Sci. Eng.* 11, 627–632. doi: 10.1109/TASE.2013.2296206
- Zhang, J., Fiers, P., Witte, K. A., Jackson, R. W., Poggensee, K. L., Atkeson, C. G., et al. (2017). Human-in-the-loop optimization of exoskeleton assistance during walking. *Science* 356, 1280–1284. doi: 10.1126/science.aal5054

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Peng, Luo, Huang, Yu, Hu, Shi and Cheng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.