Toward Task-Independent Optimal Adaptive Control of a Hip Exoskeleton for Locomotion Assistance in Neurorehabilitation

Qiang Zhang, *Member, IEEE*, Jennie Si, *Fellow, IEEE*, Xikai Tu, *Member, IEEE*, Minhan Li, *Student Member, IEEE*, Michael D. Lewek, and He Huang, *Fellow, IEEE*

***Abstract*—Personalized robotic exoskeleton control is essential in assisting individuals with motor deficits. However, current research still lacks a solution from the end of a practical need of the problem to the end of its successful demonstration in physical environments, namely an end-to-end solution, that enables stable and continuous walking across different tasks. This study addresses this challenge by introducing a hierarchical control framework for the purpose. At the low level, impedance control ensures joint compliance without causing injury to users. At the high level, a reinforcement learning (RL)-based optimal adaptive controller automatically personalizes assistance to both hip extension and flexion (namely, bi-directional) to reach a target range of motion (ROM) under multiple walking conditions. As the first potentially feasible approach to this challenging problem and to meet practical use requirements, we developed a least-square policy iteration-based solution to configure the intrinsic parameters within the well-established finite state machine impedance control (FSM-IC). We successfully tested the control solution on eight young unimpaired participants and one participant post-stroke wearing a hip exoskeleton while walking on an instrumented treadmill. The proposed method can be applied to solving for optimal impedance parameters for individual users and different task scenarios to increase joint ROM. Our next step is to further evaluate this solution framework on additional people with hemiparesis who may benefit from hip joint assistance in therapy or daily activities to restore normative or improve gait patterns.**

***Index Terms*—Assistance personalization, impedance control, optimal adaptive control, rehabilitation exoskeletons, rehabilitation robotics, reinforcement learning (RL).**

Manuscript received 13 February 2024; revised 2 June 2024; accepted 20 August 2024. Date of publication 18 September 2024; date of current version 20 November 2024. This work was supported in part by the National Institute on Disability, Independent Living; in part by the Rehabilitation Research under Grant 90ARHF0004; and in part by the National Science Foundation under Grant 2211739, Grant 2211740, Grant 1808752, Grant 1808898, and Grant 2306660. This article was recommended by Associate Editor X. Jing. *(Corresponding author: He Huang.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institutional Review Board (IRB) at North Carolina State University under Application No. 24671.

Qiang Zhang is with the Department of Mechanical Engineering, The University of Alabama, Tuscaloosa, AL 35401 USA.

Jennie Si is with the Department of Electrical, Computer, and Energy Engineering, Arizona State University, Tempe, AZ 85281 USA.

Xikai Tu is with the Department of Mechanical Engineering, Hubei University of Technology, Wuhan 430068, Hubei, China.

Minhan Li is with the Planning and Control Department, DiDi Labs, Mountain View, CA 94043 USA.

Michael D. Lewek is with the Department of Health Sciences, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599 USA.

He Huang is with the Joint Department of Biomedical Engineering, North Carolina State University and The University of North Carolina at Chapel Hill, Raleigh, NC 27695 USA (e-mail: hhuang11@ncsu.edu).

This article has supplementary material provided by the authors and color versions of one or more figures available at https://doi.org/10.1109/TSMC.2024.3454556.

Digital Object Identifier 10.1109/TSMC.2024.3454556

I. Introduction

ADVANCES in wearable lower-limb exoskeletons have shown great promise in augmenting the movement capability of human wearers [1], [2]. These systems have been developed for different applications, such as military [3], industry [4], and medical rehabilitation [5], [6]. Focusing on rehabilitation implementations, lower-limb exoskeletons have been designed to deliver active assistance to patients with neurological deficits for improving locomotor performance, including people with paraplegia (no motor functions), such as complete spinal cord injury (SCI) [7], [8] and individuals with reduced force-generating capacity caused by incomplete SCI, stroke, cerebral palsy, and multiple sclerosis, etc. [9], [10], [11]. The appeal of modern wearable robotic exoskeletons in physical rehabilitation is their intelligent, active components that can be programmed toward the needs of different patient populations [5]. For example, early designs of powered exoskeletons were for patients with paraplegia, where exoskeletons were programmed to take over the entire lower-limb movement control. Typical control strategies often focused on tracking the kinematics of individual joints during walking or other locomotion tasks [12], [13]. This control technique is quite mature and has been used in the majority of commercial exoskeletons for rehabilitation. Nevertheless, for individuals who still have voluntary motor ability, joint position control is inappropriate and can potentially cause injuries to patients. Instead, compliance is essential to ensure safe human-exoskeleton interactions. Currently, there remains an open question as to how to provide the desired mechanical assistance, tailored to the individual patient, task context, and the environment, despite the emergence of many engineering efforts by the research community [2], [9], [14], [15], [16], [17].

A fundamental challenge for walking assistance control is that individuals with neurological disorders, such as stroke, present with large variability in their motor impairments

and gait patterns [18], [19]. Or to put it intuitively, *the presentation of each person with stroke is different*. The significant variability in an individual's physical and physiological functioning is further confounded by the difficulty in properly and accurately modeling the human-robot system, which has made the associated control problem more challenging. Given the inherent variability and numerous unknowns, a data-driven learning control method makes a natural alternative to address this challenge. As a backdrop, current clinical practice relies on experienced clinicians or experts to tune a robotic device's control parameters while heuristically observing the user's gait performance [14]. The tuning criteria of the control parameters are usually not clarified or difficult to quantify due to the inherent dependence on individual tuning experts' knowledge and the user's condition [20], which may deteriorate the tuning effectiveness. With the increased dimensionality of control parameters, manual tuning significantly reduces the time efficiency and becomes more impractical due to the unexpected long-time tuning procedure.

To provide personalized and optimal assistance (OA), some recent studies, known as human-in-the-loop (HIL) optimization approaches [21], [22], [23], [24], [25], [26], have been proposed. Almost all the studies of HIL optimization approaches used open or closed-loop force/torque control to operate each individual joint of exoskeletons, where a prescribed time-depended reference joint torque with a certain shape is usually parameterized by some timing and magnitude values with the control goal of minimizing the person's metabolic cost during locomotion tasks [22], [23]. More recently, human preference-based learning algorithms for exoskeleton assistance personalization have also attracted more attention [27], [28], [29]. For example, Tucker et al. [28] proposed to implement a coactive learning paradigm in which wearers chose between paired selections and provided improvements for assistance in each trial. The controlled gait parameters were from a *gait library* that was precomputed based on a set of nominal walking gaits over a grid of various parameters. Ingraham et al. [29] measured users' preferences in the applied torque control parameters of bilateral ankle exoskeletons and characterized how preference changed with walking speed, device exposure, and prior device familiarization. However, these HIL optimization approaches usually do not have rigorous formulation from the perspective of control theory/application as they do not focus on modeling and dynamics of the human-robot-interaction system.

Another significant challenge associated with controlling exoskeletons for walking assistance is that the personalized control of robotic exoskeleton assistance needs to adapt to different people, different environments, different walking tasks, and the physical changes of users that occur over time (both positive [e.g., therapy] and negative [e.g., disuse atrophy]). This stringent requirement for adaptation over time means that the control tuning procedure is time-efficient while maintaining meaningful robustness in practical use scenarios. Currently, available results are scarce, while some studies have demonstrated potential. For example, Zhang et al. [22] showed the generality of a HIL optimization approach in single-subject studies employing an ankle exoskeleton device for walking under different locomotion conditions. However, results show that learning to walk under new conditions requires almost the same amount of time as the baseline learning condition. This is caused by the nature of the method, which learns control parameters, but not control policies. Consequently, when walking conditions change, control parameters have to be learned from anew.

To address these above challenges for powered exoskeletons, reinforcement learning (RL) approaches [30], [31], [32], [33], [34] lend themselves as alternative solutions to personalize assistance from exoskeletons. However, in the state-of-the-art, most existing studies either only focused on simulations without any physical systems [35], [36], [37] or only focused on unimpaired human participants under a single locomotion task [32], [33], [34], [38], [39]. The proposed design was inspired by RL-based solutions for powered prostheses [40] used in various locomotion tasks. However, due to the distinct differences between prosthesis (in series, no actuation redundancy) and exoskeleton (in parallel, with actuation redundancy) interactions, the formulation of an OA control problem and its implementation in real-world human testing remained uncertain. This study primarily aimed to increase the hip joint range of motion (ROM) in human wearers, emulating individuals with hip joint motor deficits. Therefore, compared to existing RL-based strategies for human-exoskeleton interactions, the novelty of this study is as follows. It provides an end-to-end hierarchical solution with RL as an essential element to personalize powered exoskeletons to offer customized assistance. This approach is uniquely adaptive and robust across various walking conditions. It has been successfully implemented on both unimpaired participants and persons with chronic hemiparesis post-stroke. To the authors' knowledge, this is the first exoskeleton control strategy capable of simultaneously addressing safety, effectiveness, time efficiency, and adaptation for neurorehabilitation. The study focuses on bilateral hip exoskeletons, chosen for their lower moment of inertia during walking, making them potentially suitable for clinical use in individuals with motor deficits. Additionally, manipulating the hip joints can directly affect step length, step width, and walking cadence [20], [41]. Our comprehensive and hierarchical solution framework was constructed to directly address all the above challenges. At the low level, the finite state machine impedance control (FSM-IC) was used to ensure exoskeleton compliance for safe human-exoskeleton interaction. At the high level, a data-driven RL-based controller was designed to tune impedance control parameters in order to provide OA adaptively for the individual wearer.

The main contributions and innovations of this study are given below. 1) We formulated an end-to-end control solution for the rehabilitation lower-limb exoskeletons that can provide personalized, task-independent, and compliant assistance; 2) Our control design enabled the human-exoskeleton system to achieve increased ROM by tuning the hip joint impedance control parameters while maintaining walking stability and user safety. 3) We successfully demonstrated the effectiveness, efficiency, and adaptation of the proposed control solution in experiments involving eight unimpaired participants and one recovering stroke patient.

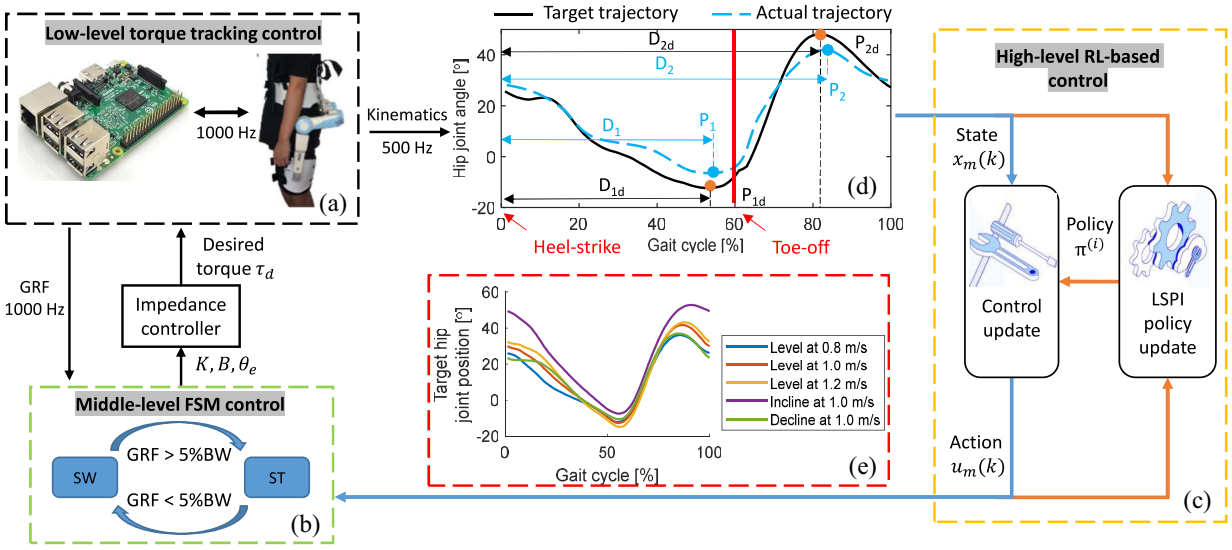


Fig. 1. RL-enabled customization of a hip exoskeleton assistance during walking. (a) Intrinsic low-level control tracks the assistive torque profile. (b) Middle-level control uses the FSM-IC to calculate desired torque profiles. (c) High-level RL control tunes the impedance control parameters and updates the control policy. (d) Target and real hip joint kinematic trajectories during a GC, where the peak flexion/extension locations in the target trajectory and actual trajectory are represented by orange points and blue points, respectively. (e) Hip joint target trajectories for different locomotion tasks.

II. Solution Framework

In this study, we aim to provide a solution that meets the critical requirements of human user safety, the customization of robotic assistance to the human user of different needs, and is potentially transferable to clinical implementations.

A. Addressing User Safety

As previously mentioned, position control is especially useful for people without any voluntary motor function as they have little ability to interact with or influence the control of the exoskeleton. Thus, the control design usually focuses on completely overriding human motor functions. However, for people with residual motor functions, position control restricts the person's residual volitional efforts and necessary gait-to-gait variation for locomotion balance control [14], and the needed flexibility to interact with the exoskeleton [16]. In addition, people with neurological disorders may also have abnormal neuromuscular control, e.g., spasticity. High stiffness in position control or torque control may lead to injuries in these users and lead to a safety concern. Therefore, under these scenarios where people with neurologic deficits have residual motor functions, the compliant behaviors of wearable lower-limb robotic exoskeletons are commonly preferred for locomotion assistance [9], [14]. The basic mechanism of this compliance lies in the notion that the human musculoskeletal system is characterized to have a *spring-like* behavior and the concept of mechanical impedance is presented as a rigorous dynamic generalization of the limb's postural control [42]. Impedance control has been widely used in robotic exoskeletons because the compliant behavior of legs is fundamental for locomotion to adapt to various environments [43].

To guarantee the safety of human users when using wearable lower-limb robotic exoskeletons for locomotion assistance, existing studies have investigated either compliant hardware design [23], [44], [45] or compliant and adaptive control design [9], [14], [46]. Compared to the position control described above, the FSM-based force/torque control for soft robotic devices and impedance control for rigid robotic devices are more compliant, and have applied locomotion assistance to people with motor dysfunction, such as incomplete SCI, cerebral palsy, and stroke [14], [44]. In this study, we restrict our design to two fundamental rules to guarantee the safety of human users, including 1) the intrinsic FSM-IC (see the subsequent section) and 2) the interlimb assistance coordination with moderate assistance level to the hip joints.

B. Intrinsic Impedance Control of the Hip Exoskeleton

Fig. 1 is a comprehensive construct of our solution approach. The behavior of the intrinsic impedance control parameters can be interpreted as a spring-damper system, thus providing necessary compliance to guarantee the safety issue. The personalized gait assistance is realized within a well-established FSM-IC framework [14]. There are two reasons that we chose to apply the FSM-IC framework instead of applying a predefined assistance torque profile, as described in previous studies [22], [23]. First of all, our hip exoskeleton mechanical design is a rigid structure, which may cause a stiff hip motion if the predefined assistance torque magnitude is too high. Second, the shape of the predefined torque profile usually depends on the time duration of the previous gait cycle(s) (GCs) instead of the actual hip joint motion state. Consequently, this may be hazardous if the hip joint suddenly stops moving, causing the wearer to perceive a sensation that the foot is blocked by an obstacle during walking.

The FSM-IC control relies on major gait events, such as heel-strike (HS) and toe-off (TO), to determine transitions between the phases. Transitions are identified according to measurements of hip joint angles θ and vertical ground reaction force (vGRF). Transitions occur according to the

threshold values in the supplementary Table SI. In the current study, the hip joint angle θ was initialized to be 0 at the neutral standing posture, positive during hip flexion, and negative during hip extension. For either phase of stance or swing in a single GC, the FSM determines impedance parameters during either stance or swing phases to be used in the low-level intrinsic torque tracking control [47] [Fig. 1(b)] for generating the desired torque profile τ_d on the targeted hip joint based on

$$\tau_d = K_p(\theta_e - \theta) + K_d\dot{\theta} \tag{1}$$

where K_p, K_d, and θ_e represent the three impedance parameters, corresponding to stiffness gain, damping gain, and equilibrium position. θ and $\dot{\theta}$ represent the angular position and velocity on the hip joint, respectively. Given two distinct phases in each GC, a total of 6 parameters are to be tuned.

C. Personalizing Assistance by Bilateral Hip Exoskeleton

Most of the existing HIL optimization studies of exoskeleton-assisted walking have focused on unimpaired human wearers as device users. As such, metrics like the wearer's energetic or metabolic cost reduction, or muscle electromyography activity reduction have been used as the device control goal [21], [22], [23], [24], [25], [26], [33], [48], [49]. These measures, however, may not serve as an appropriate control goal for people with neurological deficits as they usually have reduced ROM and/or different gait patterns, as shown in [50], [51], and [52]. Therefore, we aim to design an impedance-tuning strategy for wearable hip exoskeletons to provide personalized walking assistance to increase the target ROM. Toward this aim, we formulated the control solution for the human-exoskeleton interactive system as an online impedance parameter tuning problem based on the least-square policy iteration (LSPI) algorithm. There are three levels of the control design, which includes a hierarchy of high-, middle-, and low-level controls. At the high-level control [Fig. 1(c)], the RL adjusts the impedance parameters of the hip exoskeleton for the corresponding FSM. At the middle-level control, an impedance control with two FSMs [Fig. 1(b)] is used to determine the desired assistance torque. At the low-level control [Fig. 1(a)], a closed-loop torque controller [47] is used to accurately track the desired torque.

Refer to Fig. 1(d) for a normative human hip joint movement trajectory (black curve) during one GC, where the stance phase is between the HS event and TO event, while the swing phase is between the TO event and next HS event. The corresponding 0% and 100% represent the time instances between two consecutive HSs. Points P_1 and P_2 (P_{1d} and P_{2d}) represent the actual (target) peak extension and peak flexion positions, respectively, and D_1 and D_2 (D_{1d} and D_{2d}) represent the actual (target) timings that peak extension and peak flexion occur, respectively. As a consequence of lower-limb injury, a realistic gait profile for the affected person may appear as shown in the dashed blue curve. Our goal is to devise a control solution to provide personalized bi-directional assistance (extension during the stance phase and flexion during the swing phase) to increase the affected hip joint ROM. To examine the feasibility of the newly proposed control framework, we include multiple unimpaired human subjects and a person post-stroke in this study. For unimpaired human subjects, the control goal is set as an increased ROM of the targeted hip joint (a 25% increase from the nominal profile in both max flexion and max extension). For the person post-stroke, the control goal is set to generate the normative ROM (matching the unaffected side) for the affected hip joint. Several exampled target hip joint trajectories with increased ROM are shown in Fig. 1(e).

III. Control Solution Approach

A. Problem Formulation

We consider minimizing wearers' kinematic errors as the control objective to be optimized by the RL-based control framework. As demonstrated in Fig. 1(d), we define state variables $x^m \in \mathbb{R}^2$ as the differences between the target and measured peak extension or peak flexion during the stance or swing phase (including both amplitude error and timing error), respectively. The state variables are then defined as

$$x^m = [P_m - P_{md},\ D_m - D_{md}]^T \tag{2}$$

where $m = 1, 2$ represents the stance and swing phases, respectively. The control inputs $u^m \in \mathbb{R}^3$ are the adjustment of the impedance parameters as given subsequently for the stance and swing phases, respectively. Correspondingly, the discrete system dynamical model is given as

$$x^m_{k+1} = f^m\left(x^m_k, u^m_k\right) \tag{3}$$

where $k = 1, 2, 3, \ldots,$ represents the index of the data sample. The state variables and action variables from the current data sample could be used to map the upcoming state variables iteratively through dynamics f^m (unknown). This mapping is denoted by the tuning policy π^m and determines the action variables, defined as $u^m_k = \pi^m(x^m_k)$.

For ease of discussion, we omit the subscript m from hereon as the solution approach is the same for either the stance or swing phase albeit to say that the respective controllers are independently designed. We consider the human-exoskeleton interaction as an integrated system as depicted in Fig. 1(a) with unknown dynamics. The state variables x_k at the sample index k correspond to the definitions in (2) and (3), while the action variables u_k are defined as the adjustment of impedance parameters in either FSM and noted as $u_k = [\Delta K_p,\ \Delta K_d,\ \Delta\theta_e]^T_k$. The domain of both state and action variables are denoted as $\mathcal{D} \triangleq (x,\ u)|x \in \mathcal{X} \subset \mathbb{R}^2,\ u \in \mathcal{U} \subset \mathbb{R}^3$. The same definitions apply to both stance and swing phases as described in the above.

A stage cost function with a quadratic form in either the stance or swing phase is given below

$$R(x_k,\ u_k) = x^T_k M_x x_k + u^T_k M_u u_k \tag{4}$$

where $M_x \in \mathbb{R}^{2\times2}$ and $M_u \in \mathbb{R}^{3\times3}$ are semi-positive definite and positive definite matrices, respectively. Observations from a dynamic system are usually organized in tuples, noted

as $(x_k,\ u_k,\ R(x_k,\ u_k),\ x_{k+1})$. We define the total cost-to-go function or the state–action Q-function as

$$Q(x_k,\ u_k) = R(x_k,\ u_k) + \sum_{j=1}^{\infty} \gamma^j R\big(x_{k+j},\ u_{k+j}\big) \tag{5}$$

where $0 < \gamma^j < 1$ is the discount factor. Given a deterministic policy π, the value function $Q^{\pi}(x_k,\ u_k)$ is defined as the summation over all possible combinations of state and action variables. It implies that $Q^{\pi}(x_k,\ u_k)$ contains the expected and discounted total cost when selecting policy π and taking action u_k in state x_k [53], thus the Bellman equation is

$$\begin{aligned} Q^{\pi}(x_k,\ u_k) &= R(x_k,\ u_k) + \sum_{t=k+1}^{\infty} \gamma^{t-k} R(x_t,\ u_t) \\ &= R(x_k,\ u_k) + \gamma Q^{\pi}(x_{k+1},\ u_{k+1}). \end{aligned} \tag{6}$$

To extract Q^{π} values for all state–action pairs, a set of linear equations based on the Bellman equation is represented as

$$Q^{\pi} = R + \gamma Q^{\pi} \tag{7}$$

where Q^{π} and R are vectors of appropriate dimensions. The approximate policy iteration algorithm is based on approximated state–action value functions, which are implemented by a linear combination of quadratic basis functions and a set of weighting parameters [53], learned from samples.

The goal of RL control is to attain an optimal policy, which can then be employed to minimize the state–action value function $Q^{\pi}(x_k,\ u_k)$ as below

$$\begin{aligned} Q^*(x_k,\ u_k) &= \min_{\pi} Q^{\pi}(x_k,\ u_k) \\ &= R(x_k,\ u_k) + \gamma \min_{u_{k+1} \in \mathcal{U}} Q^*(x_{k+1},\ u_{k+1}) \\ &= R(x_k,\ u_k) + \gamma Q^*\big(x_{k+1},\ \pi^*(x_{k+1})\big) \end{aligned} \tag{8}$$

where Q^* represents the optimal state–action value function and π^* represents the current optimal policy.

B. Least-Square Policy Iteration Solution

Let i denote the ith iteration in policy iteration. In each iteration, the step of *policy evaluation* calculates the approximation of the value function, denoted as $\hat{Q}^{\pi^{(i)}}$, by solving the Bellman equation. The subsequent step of *policy improvement* defines $\pi^{(i+1)}$, the improved greedy policy, over $\hat{Q}^{\pi^{(i)}}$ as

$$\pi^{(i+1)}(x_k) = \arg\min_{u_k \in \mathcal{U}} \hat{Q}^{\pi^{(i)}}(x_k,\ u_k). \tag{9}$$

Thereafter, the policy $\pi^{(i+1)}$ is not worse than $\pi^{(i)}$, if not better than $\pi^{(i)}$. The two interleaving steps proceed until no change is observed in the policy [40], [53], [54].

A linear architecture of the following form is utilized to approximate the state–action value function

$$\hat{Q}^{\pi^{(i)}}(x,\ u) = \sum_{j=1}^{n} \psi_j(x,\ u) w_j^{(i)} \tag{10}$$

where $w_j^{(i)}$ and $\psi_j(x,\ u)$ represent the parameters and quadratic basis functions, respectively. Define $\psi(x,\ u)$ to be a column vector of size n where each entry j is the corresponding basis function ψ_j computed at $(x,\ u)$

$$\psi(x,\ u) = (\psi_1(x,\ u),\ \psi_2(x,\ u), \ldots,\ \psi_n(x,\ u))^T. \tag{11}$$

Then $\hat{Q}^{\pi^{(i)}}$ can be expressed as

$$\hat{Q}^{\pi^{(i)}} = \psi^T(x,\ u)\ w^{\pi^{(i)}} \tag{12}$$

where $w^{\pi^{(i)}}$ is a column vector with a length of n.

Recall that the value function $Q^{\pi^{(i)}}$ is the solution of the Bellman equation in (7). By substituting the value function approximation $\hat{Q}^{\pi^{(i)}}$ in place of $Q^{\pi^{(i)}}$, we have

$$\psi^T w^{\pi^{(i)}} \approx R + \gamma \psi^T w^{\pi^{(i)}}. \tag{13}$$

In the least-squares sense, the solution of the above over-constrained system is derived as

$$w^{\pi^{(i)}} = \frac{\big(\psi \psi^T\big)^{-1} \psi R}{1 - \gamma}. \tag{14}$$

Note that the above solution minimizes the L_2 norm of the Bellman residual [53].

C. Stability Annotations

Remark 1: To ensure the stability and safety of the human-exoskeleton system during walking, several considerations have been given and implemented in real-time control.

First, note that the convergence of LSPI hinges on the convergence of the approximate policy iteration. The details of LSPI convergence proof can be found in [53, Th. 7.1]. Our RL control policy implementation does not deviate from the LSPI, and therefore, the learning convergence of our implementation is also guaranteed by [53, Th. 7.1]. As such, our RL-tuned control parameters will also converge without going out-of-bounds.

Second, the middle-level control within our hierarchical framework (Fig. 1) was realized as impedance control, which is the most adopted approach to control the interaction force between the human user and the robot [55]. This results in compliant behavior of the wearable exoskeleton to the human wearer, and thus leads to a biologically natural and safe physical interaction. According to the basic principles of impedance control [42], the stability of the human-exoskeleton system during walking is thus inherently ensured in such design realization, in the same sense as how typical human stability is guaranteed by our neuromuscular control.

Third, the high-level RL approach to tuning provided a set-point for each tunable parameter during training, where each set-point was constrained to be within a prescribed human locomotion safety bound (see details in subsequent Section IV-C). Without any further (analytical) assumptions, these safety bounds imposed stability conditions on the human-robot system during the RL-based control parameters tuning.

IV. Evaluation by Human Walking Experiments

The feasibility and effectiveness of the proposed LSPI-based control personalization to provide bi-directional assistance

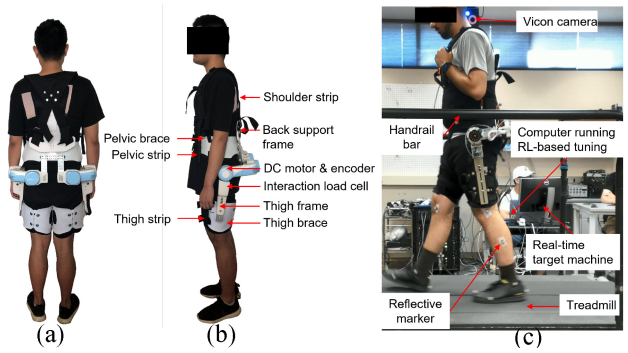


Fig. 2. Overview of a wearable hip exoskeleton designed to personalize robotic assistance during walking. (a) and (b) Lateral and back views of the wearable hip exoskeleton prototype. (c) Walking experimental environment with an instrumented treadmill with split belts and a motion capture system.

were assessed through human walking experiments under multiple scenarios. Below, we provide details of the experimental design, implementation of the LSPI-based control solution, as well as data collection, processing, and analysis. The evaluations also aim to show the *task-independence* nature of the solution approach, by which it refers to the hierarchical control structure does not require any change, but fine-tuning of the impedance parameters only requires a meaningfully short duration.

A. Human Subjects and Experimental Apparatus

The experimental protocol of treadmill walking was approved by the Institutional Review Board (IRB) at North Carolina State University (IRB approval number: 24671). Eight young unimpaired participants (6M/2F, mass: 72.3±6.5 kg, height: 173.2± 2.5 cm, age: 28.8±3.1 years old, identified as A01, A02, . . . ,A08) and a person with chronic hemiparesis post-stroke (F, mass: 54.4 kg, height: 160.1 cm, age: 62 years old, paretic side: right, identified as S01) were included in the study. All participants were instructed on the walking experimental details, and they provided an informed consent form before participating in any experiments.

The treadmill walking experimental setup is shown in Fig. 2(c), where the participant wore the bilateral light-weighted hip exoskeleton affixed to the body. The details of the hip exoskeleton's mechatronics design can be found in the supplementary Fig. S1 and Table SII. Even though the inevitable loss of mechanical power caused by the physical attachment of the device to the 'nonrigid' thigh segment, the robotic assistance torque, when transferred to the human hip joint, could still follow the reference torque profile with a relatively low-tracking error by using the low-level controller. The GRF signals were obtained from two force plates (Bertec, Columbus, OH, USA) positioned beneath the belts. These signals were then utilized to distinguish between the stance and swing phases of each leg.

B. Experimental Protocol

Three sessions of walking experiments were conducted on each participant: 1) baseline walking trials; 2) LSPI-enabled walking as control tuning trials; and 3) evaluation walking trials. Due to the different locomotion abilities of unimpaired participants and those with stroke, some adjustments were made to the experimental protocol. Prior to any trial in training and evaluation, all unimpaired participants performed baseline walking trials under different walking task scenarios using a similar setup as in Fig. 2. In the baseline walking trials, each unimpaired subject walked on the treadmill for 100 GCs under 12 scenarios (four speeds of 0.6, 0.8, 1.0, and 1.2 m/s and three slopes of 5° incline, level, and 5° decline), while the person post-stroke walked on the treadmill for 100 GCs with a level terrain at the preferred speed of 0.8 m/s. Based on baseline walking from unimpaired participants, the results of reference peak flexion and extension are determined. Details can be found in Fig. S2 in the supplementary file.

Control tuning and evaluation were conducted for all unimpaired participants with five out of the 12 walking task scenarios, including level terrain with speeds of 0.8, 1.0, and 1.2 m/s, as well as 5° incline and 5° decline terrains with the speed of 1.0 m/s. For each of those five walking task scenarios, a target hip joint trajectory was determined by scaling the respective normative joint trajectory by a factor of 1.25 to simulate an increased hip joint ROM for the participants. Thereafter, the walking task scenario-dependent target trajectory, as shown in Fig. 1(e), was used in each trial of the LSPI-based training and evaluation sessions. For the participant post-stroke, training and evaluation were conducted only for the preferred walking condition. The desired/reference hip joint trajectory of the affected side was determined to be the trajectory of the unaffected side.

During each LSPI-enabled walking training trial for each participant, one post-learning trial (of up to 25-min duration or 375 impedance parameter updates) and one evaluation trial (of 100 GCs) were conducted. If a training trial was not completed within 25 min or 375 impedance parameter updates, it was considered a failure, and the training trial was terminated. Detailed implementation procedures of the LSPI-based tuning in the learning trials are displayed in the next section. The evaluation trials included walking with the hip exoskeleton under two conditions: 1) walking in individualized OA mode for 50 GCs and 2) walking in the zero impedance (ZI) or walking without assistance mode for 50 GCs.

During all training trials, in order to mitigate the potential influence of human fatigue, participants were instructed to engage in multiple 5-min walking segments, interspersed with 3-min rest intervals (could be longer if needed). To improve the training efficiency, all impedance parameters during both stance and swing phases were auto-tuned according to the kinematics measurements from the right hip joint. However, for unimpaired participants, to maintain symmetry between the left and right hip joints as well as to guarantee the user's safety, the fine-tuned impedance parameters set originally assigned to the right hip joint was also applied to the left hip joint, so as to provide bilateral assistance on hip joints during walking. For the participant post-stroke, only impedance parameters of the affected (right) hip joint were tuned while the unaffected hip joint remained unassisted at all times.

TABLE I
TUNING RANGE OF EACH IMPEDANCE PARAMETER DURING THE WALKING STANCE AND SWING PHASES

Phase in FSM	Stance phase	Swing phase
Stiffness [Nm/rad]	[0 31]	[0 31]
Damping [Nms/rad]	[-0.1 0.1]	[-0.1 0.1]
Equilibrium position [rad]	[-0.4 0]	[0.6 1.0]

C. Impedance Tuning and Policy Update Using LSPI

1) Hyperparameters: In this study, to mitigate gait-to-gait variations during walking, a measurement sample that contains the current state x_k, action u_k, stage cost $R(x_k, u_k)$, and next state x_{k+1}, is defined as the respective average values across three consecutive GCs. The number of samples N_p is usually at least twice the number of weights in (14). Therefore, N_p was 30 in each training trial. For either the swing phase or stance phase, the action update occurred at intervals of every three GCs. In order to gather samples for the solution of the Q-value by utilizing the LSPI approach, the policy update, as shown in (9), was executed every 30 samples ($i = \lfloor k/30 \rfloor \in \mathbb{N}$).

The weighting matrices M_x and M_u in (4) were set as $M_x = \text{diag}([10, 0.01])$ and $M_u = \text{diag}([0.01, 0.1, 0.1])$, respectively. The discount factor γ was set as 0.9. For the regeneration of the target hip joint trajectory, we designed appropriate tolerance levels of state variables (specifically, $\pm 5^\circ / \pm 3^\circ$ for peak flexion/extension amplitude errors and 2% for peak flexion/extension timing errors) and action variables 10% range of variation of the corresponding impedance parameters to account for the inherent walking variability, which was used to derive the stage cost tolerance levels of 0.06, noted as ε_e, and 0.15, noted as ε_f, for hip extension and flexion, respectively. It was considered that the impedance parameters tuning procedures were successful if the stage cost values stayed within the tolerance levels for 8 out of the very recent 12 consecutive samples. If both stance and swing phases become successful, the impedance tuning trial was considered to reach the stopping criterion. To prevent any possible harm to the human participants stemming from unsafe impedance parameters, a set of ranges for each impedance parameter was predefined [56] in Table I, where positive and negative equilibrium positions represent the flexion and extension directions, respectively. We also designed safety bounds for both ROM on the targeted hip joint and the ROM-associated gait symmetry (GS), e.g., 0.5 rad $\leq$ ROM $\leq$ 1.6 rad and $-0.1 \leq$ GS ≤ 0.1. If any impedance parameter was beyond the safety range, any ROM or GS values were beyond the safety bounds, all control parameters were reset to the safe and feasible initial values below.

2) Policy and Impedance Parameters Initialization and Personalization: We applied one set of initial impedance parameters ($K_p = 2$ Nm/rad, $K_d = -0.01$ Nms/rad, and $\theta_e = 0.6$ rad for flexion, as well as $K_p = 2$ Nm/rad, $K_d = -0.01$ Nms/rad, and $\theta_e = -0.2$ rad for extension) to initialize a policy. In the impedance parameters tuning trials, each participant conducted treadmill walking with the ZI mode for the first 10 GCs. Subsequently, the hip exoskeleton started to provide both hip extension and flexion assistance with the initial impedance parameters above. The impedance parameters were tuned according to a pretrained policy from a user who has much experience in using the hip exoskeleton as the initial policy $\pi^{(0)}$ for online training. To acquire the pretrained initial policy, the walking task scenario was set as the level terrain at 1.0 m/s with a pseudo-random initial policy and the same initial impedance parameters as described above. Seven pretraining trials on the same experienced user were conducted using different pseudo-random initial policies. Eventually, the policy derived from the most recently successful trial was subsequently employed with the same participant in the remaining four walking task scenarios and on all other unimpaired participants under all five walking task scenarios and on the stroke patient under her most preferred walking condition. The implementation of the proposed LSPI-based control is described as pseudocodes in Algorithm 1 of the supplementary file.

D. Data Collection and Processing

Throughout the impedance parameters tuning trials according to LSPI, we recorded temporally discrete samples of all state and action variables, impedance parameters, stage costs, and policies for both stance and swing phases every three GCs. Onboard data measurements of the exoskeleton were online sampled and collected simultaneously at 500 Hz in MATLAB (R2022a, MathWorks, MA, USA), including angle, velocity, assistive torque, and interaction torque on both hip joints, in addition, GRF signals from both left and right force plates on the instrumented treadmill were recorded at 500 Hz in MATLAB. The low-level torque control of the hip exoskeleton was implemented in C++ and operated at 1000 Hz. All onboard signals and GRF signals were filtered by using a 4th-order, zero-lag, low-pass Butterworth filters, where the cut-off frequencies were set as 10 Hz and 20 Hz, respectively.

V. RESULTS

A. Online Training Results

Given the limited space, results from representative participants are shown in this section. Additional subject-specific and group-level demonstrations can be found in the supplementary file. The mean and standard deviation (SD) values of iteration numbers when reaching the termination criteria was 150.0±59.5 across these seven pretraining trials from the experienced participant during the determination of the pretrained initial policy. Among all online training trials other than the pretrained initial policy determination, take participants A01 and S01 as examples with the top performance, during the online LSPI-based impedance parameters post-tuning trials, the iterative changes of cost function values in each walking task scenario are demonstrated in Fig. 3. For participant A01, the iteration numbers when reaching the stopping criterion for these five walking task scenarios were 23, 63, 62, 11, and 93, respectively. Correspondingly, the impedance parameters tuning duration for these five walking task scenarios were 1.55, 4.02, 3.63, 0.67, and 6.02 min, respectively. For the participant post-stroke (S01), the iteration number and tuning duration

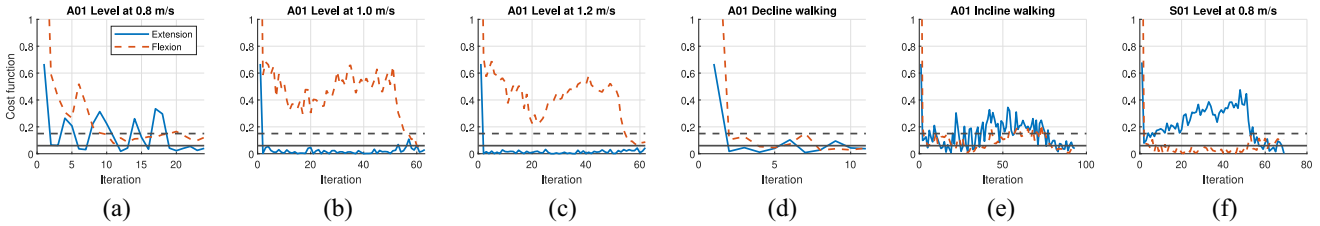


Fig. 3. Iterative cost function values of both hip extension and flexion phases during the LSPI-based control parameters tuning procedures. The presented results in (a)–(e) are from participant A01 walking under five scenarios (level walking at 0.8, 1.0, and 1.2 m/s, and 5° decline and 5° incline walking at 1.0 m/s), and results in (f) are from participant S01 under level walking at 0.8 m/s. The horizontal black solid and dashed lines represent the ultimate stage cost tolerance values of both hip extension and hip flexion.

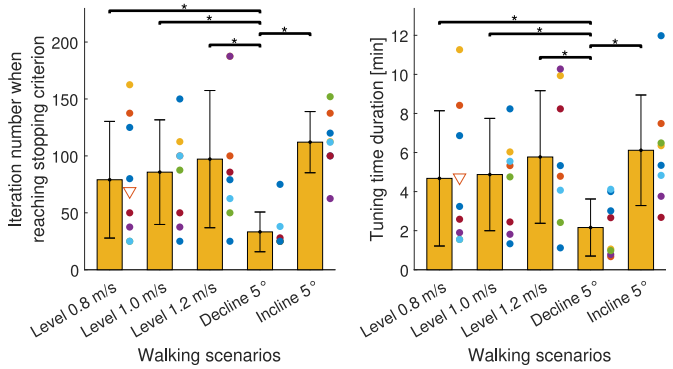


Fig. 4. Individual values of the iteration numbers and tuning duration (the elapsed time for auto-tuning during the treadmill walking experiments counted by MATLAB timer) when reaching the stopping criterion under each walking task scenario as well as the mean and SD across all participants. Different dots represent data from all eight unimpaired participants and the reverse triangle represents data from the participant post-stroke. The asterisks represent the significant difference level of $p < 0.05$.

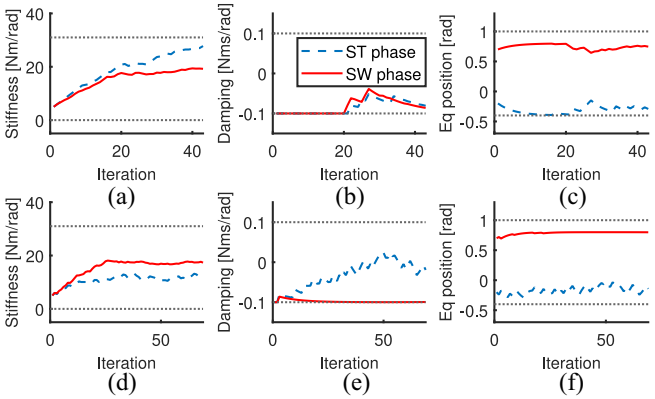


Fig. 5. Iterative impedance parameters (i.e., stiffness, damping, and equilibrium position) of both stance (blue dashed curves) and swing (red solid curves) phases during the training trial. Data in (a) to (c) and in (d) to (f) are from participants A05 and S01 under the level walking at 0.8-m/s scenario. Positive and negative equilibrium positions represent the flexion and extension directions, respectively.

when reaching the stopping criterion under the preferred walking condition were 69 and 4.75 min, respectively.

The bar plots in Fig. 5 show the mean and SD values of the tuning iteration numbers and tuning duration when reaching the stopping criterion under each walking task scenario across all eight young healthy participants. It is noted that all bars are composed of 8 data points, but one data point shown with a triangles represents that from the stroke participant. For the 8 unimpaired participants across walking conditions, the overall number of impedance updates was an average of 85, corresponding to 4.72 min of walking. Results of unimpaired participants from the Shapiro–Wilk parametric hypothesis test validated that each group followed the normal distribution, and results from one-way ANOVA indicated that the iteration number or tuning duration under the walking scenario of 5° decline walking was significantly less than the other four walking task scenarios (all $p < 0.05$), while no statistically significant difference was found among other four scenarios.

To inspect the dynamic behavior of impedance tuning using LSPI from the beginning to the point meeting the termination criteria, we present representative results from participants A05 and S01 under the level walking at 0.8 m/s in Fig. 5, which shows that impedance parameters, e.g., stiffness and equilibrium position, converged to relatively constant values when the LSPI-based learning trial for both stance and swing phases reached the ending, and these relatively constant values were considered as the personalized impedance parameters. Even though the damping parameters were not quite settled, both tuning trials reached the stopping criteria. This indicates that the robotic assistance was not sensitive to the damping parameters as the predefined range was relatively small.

B. Evaluation Results—Hip Joint Kinematics

For both unimpaired participants and the participant post-stroke, we focused on the kinematics of the right hip joint. Fig. 6 illustrates the hip joint trajectories from participant A03 under the five walking task scenarios and from the S01 under the level walking at 0.8-m/s scenario. In each plot, the blue dashed curve represents the desired/reference hip joint trajectory under each scenario, and the green (red) curves and corresponding light-colored areas represent the mean and SD values of the actual hip joint trajectories across all GCs in ZI (OA) mode of the hip exoskeleton. Given the primary control goal of the proposed control framework is not a point-to-point position tracking (rigid) on the targeted hip joint, but an adaptive optimal control solution (compliant) to personalize robotic assistance based on RL, the point-to-point errors between the desired and actual trajectories do not need to be minimized as long as the magnitude error and timing error at the peak extension and peak flexion meet the predefined threshold values. The direct effect is the joint ROM

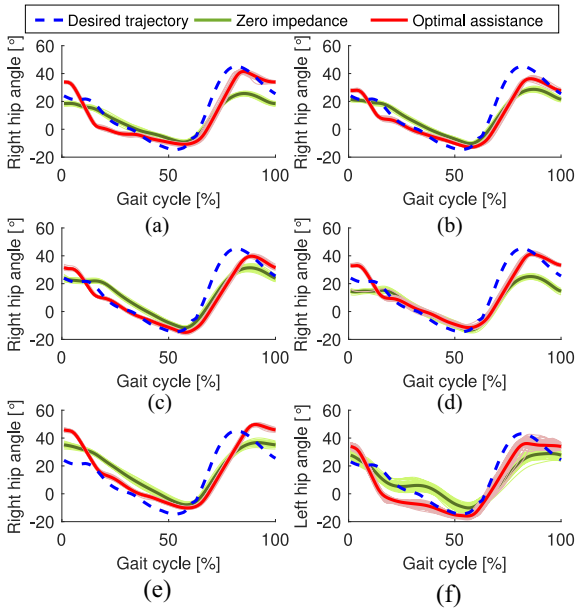


Fig. 6. Desired and real hip joint trajectories from one unimpaired participant and the person post-stroke under multiple walking task scenarios. The blue dashed curves represent the desired hip joint trajectories for different walking task scenarios, the green curves represent the actual hip joint trajectories under ZI mode, and the red curves represent the actual hip joint trajectories under OA mode. Data in (a) to (e) are from participant A03 during level walking at 0.8 m/s, level walking at 1.0 m/s, level walking at 1.2 m/s, 5° decline walking at 1.0 m/s, and 5° incline walking at 1.0 m/s, respectively. Data in (f) are from a person post-stroke during level walking at 0.8 m/s.

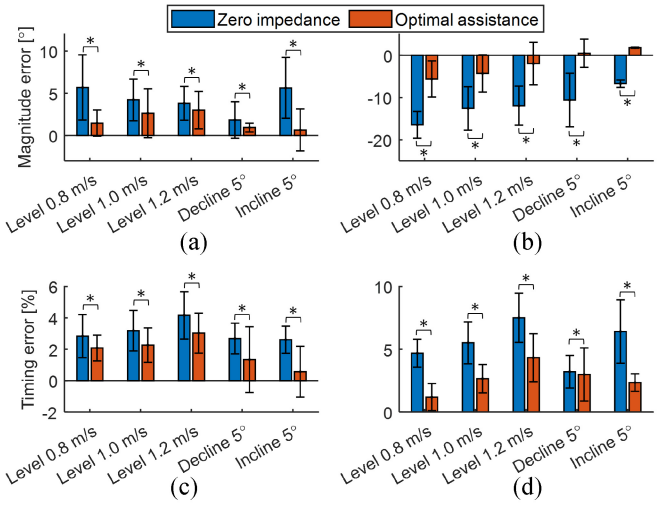


Fig. 7. Magnitude error and timing error result at peak extension and peak flexion points between ZI mode and OA mode. Presented bars and error bars are the mean and SD values of each metric on the right hip joint across all unimpaired participants, respectively. (a) and (c) Stance phase and (b) and (d) Swing phase. The statistically significant difference is denoted by * at a level of $p < 0.05$.

increase after adding the personalized assistance from the hip exoskeleton. From the results of evaluation trials in Fig. 6, it is observed that A03's right hip joint ROM significantly increased from 34.96±1.17° to 42.59±1.08° ($p < 0.05$), from 39.07±0.92° to 49.29±1.02° ($p < 0.05$), from 43.10±1.31° to 54.75±0.97° ($p < 0.05$), from 38.49±1.14° to 52.90±0.91° ($p < 0.05$), and from 44.88±1.28° to 60.08±0.77° ($p < 0.05$) under those five walking conditions, respectively. It is also observed that S01's right hip joint ROM significantly increased from 39.57±2.72° to 54.24±4.14° ($p < 0.05$) during level walking at 0.8 m/s. Results of the ROM values of the right hip joint from both unimpaired participants and our participant recovering from stroke under each walking task scenario are summarized in Table SIII of the supplementary materials.

The outcomes of the proposed control framework regarding hip joint kinematics were also evaluated by the magnitude error and timing error at peak flexion and peak extension points between the desired and actual hip joint trajectories. Fig. 7 is a summary of kinematic errors of hip joint motion (magnitude and timing) for all unimpaired participants in all evaluation scenarios, which indicates these errors were effectively reduced in OA mode regardless of walking task scenarios. Averaged across all unimpaired participants under each walking task scenario, the improvement percentage of the peak extension/flexion magnitude and timing errors by using the OA mode are summarized in Table II.

TABLE II
IMPROVEMENT (IN PERCENTAGE) OF MAGNITUDE/TIMING ERROR FOR PEAK FLEXION/EXTENSION UNDER EACH WALKING SCENARIO FROM ZI MODE TO OA MODE

Scenario	0.8 m/s	1.0 m/s	1.2 m/s	Decline	Incline
Stance ΔP	74.16%	37.32%	21.35%	48.38%	88.43%
Stance ΔD	26.69%	28.88%	27.21%	50.18%	78.01%
Swing ΔP	66.07%	65.73%	83.65%	95.46%	73.38%
Swing ΔD	74.65%	51.86%	42.35%	7.03%	63.41%

C. Evaluation Results—Hip Joint Kinetics

Given the customized impedance parameters for each participant, once the hip joint angular position and velocity signals were measured, the assistance torque during hip extension and flexion in OA mode could be calculated based on (1). As stated in the experimental protocol, the human-exoskeleton interaction torque was measured by using the embedded load cell after interacting with the wearer's thigh during walking. Representative interaction torque profiles (mean and SD values across all GCs in the evaluation session) under multiple walking task scenarios from both participant A03 and participant S01 are presented in Fig. 8. The positive and negative values represent hip flexion and extension assistance torques, respectively. The green (red) curves and corresponding light-colored areas present the mean and SD values of the interaction torque profiles across all GCs in the ZI (OA) mode of the hip exoskeleton. Consistent results were obtained for all other participants. Results from all participants showed that the interaction torque between the exoskeleton and human thigh was between −20 N·m and 20 N·m. To quantify the assistance level throughout the GC, the root mean square (RMS) values of the interaction torque and assistive power within each GC are calculated, whose mean and SD values across all recorded GCs and all participants from each walking task scenario in the evaluation session are summarized in Table SIV of the supplementary materials. It is observed that the interaction torque and assistive power were not completely zero in the ZI mode because a necessary interaction torque/assistive power (with a reverse direction from the exoskeleton to wearers) was

required to achieve backdrivability. Compared to the ZI mode, the OA mode significantly increased the interaction torque and assistance power throughout the GC, which assisted the hip joint in flexing and extending during walking.

VI. Discussion and Conclusion

A. Interpretation of LSPI-Based Assistance Personalization

This study presents a safe, efficient, and effective adaptive control solution for providing personalized walking assistance to individuals with motor impairments. Our experimental results, involving eight unimpaired participants and one stroke survivor (as shown in Fig. 3 through Fig. 5), demonstrate that the LSPI-based hierarchical control framework effectively optimizes the high-dimensional control parameters of the exoskeleton. The proposed approach successfully increased the targeted hip joint ROM to the desired levels across five different walking task scenarios. Notably, we observed significant variability in the final optimal impedance parameters for each scenario, highlighting the need for customized assistance tailored to individual differences, which ultimately led to similar improvements in hip joint kinematics across participants.

We observed the highest-learning time efficiency during decline walking, and the lowest-learning time efficiency during incline walking for our unimpaired participants (Fig. 4), indicating that the proposed LSPI-based control framework tended to aid participants in reaching a target hip trajectory faster if the given target is lower [Fig. 1(e)]. Remarkably, the learning time efficiency of the participant post-stroke was very similar to the average performance from the unimpaired participants, which implies the possibility of automatically personalizing robotic assistance for neurologically impaired users with only a short period of training/tuning. As an efficient and generic learning-based control solution, the proposed LSPI-based assistance personalization can also be implemented in other wearable assistive devices, including upper or lower-limb exoskeletons, prostheses, and neuroprosthetics [40], [57], [58], [59].

Our finding in regards to the simultaneous optimization of 6 control parameters in about 5 min is appropriate for clinical rehabilitation applications according to the detailed evidence in the supplementary file. This performance would be unlikely to achieve by using manual updates for high-dimensional control parameters, as detailed in the supplementary file. Based on the discussed evidence, we believe that our finding with optimization of close to 5 min while auto-tuning six control parameters in the current study is the most responsive technique shown in the literature. We also want to point out other practical values of our proposed new method. Our auto-tuning algorithm does not require additional sensors. In addition, many current HIL optimization methods do not converge, rather the investigators use time limits (20 to 25 min) as the termination criterion [22], [23]. Finally, compared to manual tuning, our automatic tuning method is more accurate and time-efficient. Therefore, we believe our proposed method is significant and feasible for potential clinical use in the future.

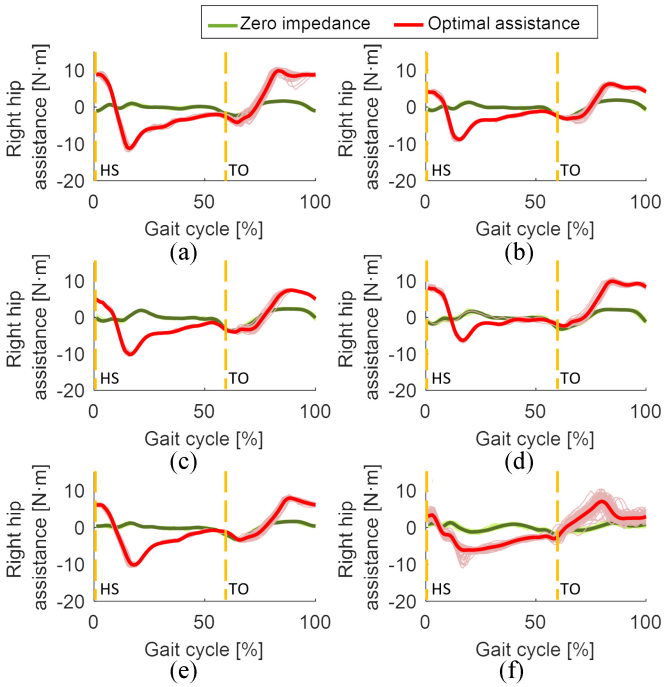


Fig. 8. Interaction torque profiles on the right hip joint from one unimpaired participant and the participant following stroke under multiple walking task scenarios. The green curves represent the interaction torque under ZI mode while the red curves represent the interaction torque under OA mode. Data in (a) to (e) are from participant A03 during level walking at 0.8 m/s, level walking at 1.0 m/s, level walking at 1.2 m/s, 5° decline walking at 1.0 m/s, and 5° incline walking at 1.0 m/s, respectively. Data in (f) are from S01 (post-stroke) during level walking at 0.8 m/s.

B. Insights From Evaluation Outcomes

The primary goal of the proposed control framework was not to achieve a point-to-point good tracking performance like traditional position control approaches. Consequently, we observed a relatively high error between the desired/reference and actual hip joint trajectories other than the moments of peak extension and peak flexion, as shown in Fig. 6. This is a fact that all FSM-IC approaches encounter, especially when the number of FSMs is small, like the two FSMs in the current study. In this sense, the hip exoskeleton provided compliance to some degree without strictly constraining the human hip joint. This could be the main reason that all participants felt comfortable with the hip assistance and could easily get used to walking with external robotic assistance. Based on a post hoc analysis of human-exoskeleton interaction torque in Fig. 8, the profile shape was very similar for both unimpaired participants and the individual post-stroke. The difference in magnitude and any timing shift mainly depend on various personalized impedance parameter sets and hip joint kinematics from person to person. It is also observed that the variation of the interaction torque on the person with stroke was much higher than that of the unimpaired participants. We attribute this finding to larger GC-to-cycle variability in our participant with stroke, compared to the unimpaired participants. Furthermore, the results from [60] reported that the overall hip joint torque remained invariant despite substantial external assistance from the hip exoskeleton, according to

inverse dynamics. We also found similar observations in the current study (see representative results in the supplementary Fig. S17), while the hip joint kinematics were significantly affected as shown in Fig. 6. By subtracting the robotic assistance torque from the overall hip joint torque, we hypothesize that the biological torque on the hip joint could be reduced, thus suggesting a potential energy expenditure reduction when using personalized robotic assistance. Given the focus of this article is to demonstrate the preliminary implementation of a new adaptive optimal control solution under multiple walking conditions, we will examine this hypothesis in the future by calculating inverse dynamics from more participants and systematic evaluations on whether and how energetic alterations are achieved.

As similar results reported in [61], we also found participants tended to have longer stride lengths under scenarios with faster level walking speed and incline ramp with both ZI and OA modes (see the supplementary Fig. S16). In addition, the stride length of unimpaired participants' left and right sides (as well as the right side of the person with stroke) were significantly increased by applying the personalized OA for each walking task scenario. The most likely reason could be the improvement of the hip joint ROM after imposing the OA mode, which indicates potential walking speed increase if the walking cadence is stable during over-ground walking in the real world, a promising outcome for patients with mobility deficits. Although a kinematics change at the knee joint might also contribute to the stride length increase, this effect would be minor compared to a change in hip joint kinematics. Lastly, one note is that even though impedance parameters were learned based on the kinematics from the right hip joint, the current control framework is expected to be extensible for assistance personalization on both sides when considering motor deficits for individuals with other neurological disorders.

C. Limitations and Future Work

Although our proposed LSPI-based control design and experimental study are promising, there are still some limitations. The primary limitation is the small population size and limited walking task scenario of evaluation on human participants with mobility impairments due to the focus here being the feasibility investigation and proof-of-concept study of the proposed novel assistance personalization method instead of a clinical implementation. Second, considering that our applied target hip joint trajectories, as suggested in Fig. 1(e), were walking task scenario-dependent and designed with a scaling factor of 1.25 based on data from the lookup library that was established according to the normative treadmill walking data under 12 different scenarios, a possible limitation is that the target hip joint trajectory is not adaptive online or directly appropriate for daily use overground walking task scenarios with time-variant speed. Also, no cross-task generalization results were investigated in the current study. To address these issues, future work may consider, for example, building a high-dimensional regression model for generating a trajectory library according to the real-time estimation of the walking speed and terrain slope, such as the approaches in [62] and [63].

Third, two FSMs (stance and swing phases) were applied in the current study, and the optimal personalized impedance parameters from the LSPI-based tuning were constant throughout each FSM under each walking task scenario. Switching between the stance and swing phases during locomotion as well as the distinct impedance parameters in these two FSMs is likely to create a sharp change in the desired torque command. In future work, one possible solution for addressing this issue is to increase the number of FSMs with a more reliable detection approach or build a continuous impedance model throughout the GC [64]. Lastly, the multiple 5-min walking sections in the experimental protocol were just an elapsed time approximation due to the reasons that human participants may change GC counts in one training section at different gait speeds. The 3-min resting period during walking experiments was provided in order to avoid the potential fatigue of the participants, and the de-adaptation during this resting period is not within the scope of the current study and will be an interesting point for future study.

REFERENCES

[1] H. Kazerooni, "Exoskeletons for human power augmentation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2005, pp. 3459–3464.
[2] B. S. Rupal, S. Rafique, A. Singla, E. Singla, M. Isaksson, and G. S. Virk, "Lower-limb exoskeletons: Research trends and regulatory guidelines in medical and non-medical applications," *Int. J. Adv. Robot. Syst.*, vol. 14, no. 6, 2017, Art. no. 1729881417743554.
[3] A. B. Zoss, H. Kazerooni, and A. Chu, "Biomechanical design of the Berkeley lower extremity exoskeleton (BLEEX)," *IEEE/ASME Trans. Mechatron.*, vol. 11, no. 2, pp. 128–138, Apr. 2006.
[4] M. P. De Looze, T. Bosch, F. Krause, K. S. Stadler, and L. W. O'Sullivan, "Exoskeletons for industrial application and their potential effects on physical work load," *Ergonomics*, vol. 59, no. 5, pp. 671–681, 2016.
[5] J. L. Pons, "Rehabilitation exoskeletal robotics," *IEEE Eng. Med. Biol. Mag.*, vol. 29, no. 3, pp. 57–63, May 2010.
[6] D. Shi, W. Zhang, W. Zhang, and X. Ding, "A review on lower limb rehabilitation exoskeleton robots," *Chin. J. Mech. Eng.*, vol. 32, no. 1, pp. 1–11, 2019.
[7] R. J. Farris, H. A. Quintero, S. A. Murray, K. H. Ha, C. Hartigan, and M. Goldfarb, "A preliminary assessment of legged mobility provided by a lower limb exoskeleton for persons with paraplegia," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 3, pp. 482–490, May 2014.
[8] M. Mekki, A. D. Delgado, A. Fry, D. Putrino, and V. Huang, "Robotic rehabilitation and spinal cord injury: A narrative review," *Neurotherapeutics*, vol. 15, no. 3, pp. 604–617, 2018.
[9] T. Yan, M. Cempini, C. M. Oddo, and N. Vitiello, "Review of assistive strategies in powered lower-limb orthoses and exoskeletons," *Robot. Auton. Syst.*, vol. 64, pp. 120–136, Feb. 2015.
[10] B. Chen et al., "Recent developments and challenges of lower extremity exoskeletons," *J. Orthop. Translat.*, vol. 5, pp. 26–37, Apr. 2016.
[11] M. Grimmer, R. Riener, C. J. Walsh, and A. Seyfarth, "Mobility related physical and functional losses due to aging and disease-a motivation for lower limb exoskeletons," *J. Neuroeng. Rehabil.*, vol. 16, no. 1, pp. 1–21, 2019.
[12] M. Bortole et al., "The H2 robotic exoskeleton for gait rehabilitation after stroke: Early findings from a clinical study," *J. Neuroeng. Rehabil.*, vol. 12, no. 1, p. 54, 2015.
[13] Y. Long, Z. Du, L. Cong, W. Wang, Z. Zhang, and W. Dong, "Active disturbance rejection control based human gait tracking for lower extremity rehabilitation exoskeleton," *ISA Trans.*, vol. 67, pp. 389–397, Mar. 2017.
[14] S. A. Murray, K. H. Ha, C. Hartigan, and M. Goldfarb, "An assistive control approach for a lower-limb exoskeleton to facilitate recovery of walking following stroke," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 3, pp. 441–449, May 2015.

[15] W. Huo, S. Mohammed, J. C. Moreno, and Y. Amirat, "Lower limb wearable robots for assistance and rehabilitation: A state of the art," *IEEE Syst. J.*, vol. 10, no. 3, pp. 1068–1081, Sep. 2016.

[16] A. J. Young and D. P. Ferris, "State of the art and future directions for lower limb robotic exoskeletons," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 25, no. 2, pp. 171–182, Feb. 2017.

[17] A. Rodríguez-Fernández, J. Lobo-Prat, and J. M. Font-Llagunes, "Systematic review on wearable lower-limb exoskeletons for gait training in neuromuscular impairments," *J. Neuroeng. Rehabil.*, vol. 18, no. 1, pp. 1–21, 2021.

[18] G. Kwakkel, B. Kollen, and E. Lindeman, "Understanding the pattern of functional recovery after stroke: Facts and theories," *Restor. Neurol. Neurosci.*, vol. 22, nos. 3–5, pp. 281–299, 2004.

[19] K. K. Patterson et al., "Gait asymmetry in community-ambulating stroke survivors," *Arch. Phys. Med. Rehabil.*, vol. 89, no. 2, pp. 304–310, 2008.

[20] K. Seo, J. Lee, Y. Lee, T. Ha, and Y. Shim, "Fully autonomous hip exoskeleton saves metabolic cost of walking," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2016, pp. 4628–4635.

[21] J. R. Koller, D. H. Gates, D. P. Ferris, and C. D. Remy, "'Body-in-the-loop'optimization of assistive robotic devices: A validation study," in *Proc. Robot., Sci. Syst.*, 2016, pp. 1–10.

[22] J. Zhang et al., "Human-in-the-loop optimization of exoskeleton assistance during walking," *Science*, vol. 356, no. 6344, pp. 1280–1284, 2017.

[23] Y. Ding, M. Kim, S. Kuindersma, and C. J. Walsh, "Human-in-the-loop optimization of hip assistance with a soft exosuit during walking," *Sci. Robot.*, vol. 3, no. 15, 2018, Art. no. eaar5438.

[24] J. Kim et al., "Reducing the metabolic rate of walking and running with a versatile, portable exosuit," *Science*, vol. 365, no. 6454, pp. 668–672, 2019.

[25] S. Song and S. H. Collins, "Optimizing exoskeleton assistance for faster self-selected walking," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 786–795, Apr. 2021.

[26] G. M. Bryan et al., "Optimized hip-knee-ankle exoskeleton assistance reduces the metabolic cost of walking with worn loads," *J. Neuroeng. Rehabil.*, vol. 18, no. 1, pp. 1–13, 2021.

[27] M. Tucker et al., "Human preference-based learning for high-dimensional optimization of exoskeleton walking gaits," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2020, pp. 3423–3430.

[28] M. Tucker et al., "Preference-based learning for exoskeleton gait optimization," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2020, pp. 2351–2357.

[29] K. Ingraham, C. Remy, and E. Rouse, "The role of user preference in the customized control of robotic exoskeletons," *Sci. Robot.*, vol. 7, no. 64, 2022, Art. no. eabj3487.

[30] Y. Zhang, S. Li, K. J. Nolan, and D. Zanotto, "Adaptive assist-as-needed control based on actor-critic reinforcement learning," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2019, pp. 4066–4071.

[31] R. Huang, H. Cheng, J. Qiu, and J. Zhang, "Learning physical human–robot interaction with coupled cooperative primitives for a lower exoskeleton," *IEEE Trans. Autom. Sci. Eng.*, vol. 16, no. 4, pp. 1566–1574, Oct. 2019.

[32] Y. Yuan, Z. Li, T. Zhao, and D. Gan, "DMP-based motion generation for a walking exoskeleton robot using reinforcement learning," *IEEE Trans. Ind. Electron.*, vol. 67, no. 5, pp. 3830–3839, May 2020.

[33] X. Tu, M. Li, M. Liu, J. Si, and H. H. Huang, "A data-driven reinforcement learning solution framework for optimal and adaptive Personalization of a hip exoskeleton," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2021, pp. 10610–10616.

[34] Q. Zhang et al., "Imposing healthy hip movement pattern and range by exoskeleton control for Individualized assistance," *IEEE Robot. Autom. Lett.*, vol. 7, no. 4, pp. 11126–11133, Oct. 2022.

[35] S. G. Khan, M. Tufail, S. H. Shah, and I. Ullah, "Reinforcement learning based compliance control of a robotic walk assist device," *Adv. Robot.*, vol. 33, no. 24, pp. 1281–1292, 2019.

[36] L. Rose, M. C. Bazzocchi, and G. Nejat, "End-to-end deep reinforcement learning for exoskeleton control," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, 2020, pp. 4294–4301.

[37] S. Luo, G. Androwis, S. Adamovich, H. Su, E. Nunez, and X. Zhou, "Reinforcement learning and control of a lower extremity exoskeleton for squat assistance," *Front. Robot. AI*, vol. 8, Jul. 2021, Art. no. 702845.

[38] R. Huang, H. Cheng, H. Guo, X. Lin, and J. Zhang, "Hierarchical learning control with physical human-exoskeleton interaction," *Inf. Sci.*, vol. 432, pp. 584–595, Mar. 2018.

[39] Z. Peng et al., "Data-driven reinforcement learning for walking assistance control of a lower limb exoskeleton with hemiplegic patients," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2020, pp. 9065–9071.

[40] M. Li, Y. Wen, X. Gao, J. Si, and H. Huang, "Toward expedited impedance tuning of a robotic prosthesis for personalized gait assistance by reinforcement learning control," *IEEE Trans. Robot.*, vol. 38, no. 1, pp. 407–420, Feb. 2022.

[41] B. Lim et al., "Delayed output feedback control for gait assistance with a robotic hip exoskeleton," *IEEE Trans. Robot.*, vol. 35, no. 4, pp. 1055–1062, Aug. 2019.

[42] N. Hogan, "The mechanics of multi-joint posture and movement control," *Biol. Cybern.*, vol. 52, no. 5, pp. 315–331, 1985.

[43] A. F. Azocar, L. M. Mooney, J.-F. Duval, A. M. Simon, L. J. Hargrove, and E. J. Rouse, "Design and clinical implementation of an open-source bionic leg," *Nat. Biomed. Eng.*, vol. 4, no. 10, pp. 941–953, 2020.

[44] L. N. Awad et al., "A soft robotic exosuit improves walking in patients after stroke," *Sci. Transl. Med.*, vol. 9, no. 400, 2017, Art. no. eaai9084.

[45] D. G. Schmitz et al., "Modulation of Achilles tendon force with load carriage and exosuit assistance," *Sci. Robot.*, vol. 7, no. 71, 2022, Art. no. eabq1514.

[46] A. Martinez, B. Lawson, and M. Goldfarb, "A controller for guiding leg movement during overground walking with a lower limb exoskeleton," *IEEE Trans. Robot.*, vol. 34, no. 1, pp. 183–193, Feb. 2018.

[47] U. Nagarajan, G. Aguirre-Ollinger, and A. Goswami, "Integral admittance shaping: A unified framework for active exoskeleton control," *Robot. Auton. Syst.*, vol. 75, pp. 310–324, Jan. 2016.

[48] M. Kim et al., "Human-in-the-loop Bayesian optimization of wearable device parameters," *PloS one*, vol. 12, no. 9, 2017, Art. no. e0184054.

[49] M. Kim et al., "Bayesian optimization of soft exosuits using a metabolic estimator stopping process," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, 2019, pp. 9173–9179.

[50] D. C. Kerrigan, L. W. Lee, J. J. Collins, P. O. Riley, and L. A. Lipsitz, "Reduced hip extension during walking: Healthy elderly and fallers versus young adults," *Arch. Phys. Med. Rehabil.*, vol. 82, no. 1, pp. 26–30, 2001.

[51] M. P. Steultjens, J. Dekker, M. E. van Baar, R. A. Oostendorp, and J. W. Bijlsma, "Range of joint motion and disability in patients with osteoarthritis of the knee or hip," *Rheumatology*, vol. 39, no. 9, pp. 955–961, 2000.

[52] L. E. Diamond, F. L. Dobson, K. L. Bennell, T. V. Wrigley, P. W. Hodges, and R. S. Hinman, "Physical impairments and activity limitations in people with femoroacetabular impingement: A systematic review," *Br. J. Sports Med.*, vol. 49, no. 4, pp. 230–242, 2015.

[53] M. G. Lagoudakis and R. Parr, "Least-squares policy iteration," *J. Mach. Learn. Res.*, vol. 4, pp. 1107–1149, Dec. 2003.

[54] D. P. Bertsekas, "Approximate policy iteration: A survey and some new methods," *Int. J. Control. Theory Appl.*, vol. 9, no. 3, pp. 310–335, 2011.

[55] M. Vukobratović, "How to control robots interacting with dynamic environment," *J. Intell. Robot. Syst.*, vol. 19, pp. 119–152, Jun. 1997.

[56] B. Koopman, E. H. Van Asseldonk, and H. Van Der Kooij, "Estimation of human hip and knee multi-joint dynamics using the LOPES gait trainer," *IEEE Trans. Robot.*, vol. 32, no. 4, pp. 920–932, Aug. 2016.

[57] Q. Zhang, D. Sun, W. Qian, X. Xiao, and Z. Guo, "Modeling and control of a cable-driven rotary series elastic actuator for upper limb rehabilitation robot," *Front. Neurorobot.*, vol. 14, p. 13, Feb. 2020.

[58] Y. Wen, J. Si, A. Brandt, X. Gao, and H. H. Huang, "Online reinforcement learning control for the personalization of a robotic knee prosthesis," *IEEE Trans. Cybern.*, vol. 50, no. 6, pp. 2346–2356, Jun. 2020.

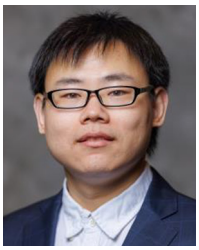
[59] V. Molazadeh, Q. Zhang, X. Bao, and N. Sharma, "An iterative learning controller for a switched cooperative allocation strategy during sit-to-stand tasks with a hybrid exoskeleton," *IEEE Trans. Control Syst. Technol.*, vol. 30, no. 3, pp. 1021–1036, May 2022.

[60] C. L. Lewis and D. P. Ferris, "Invariant hip moment pattern while walking with a robotic hip exoskeleton," *J. Biomech.*, vol. 44, no. 5, pp. 789–793, 2011.

[61] S. Agiovlasitis, J. A. McCubbin, J. Yun, G. Mpitsos, and M. J. Pavol, "Effects of down syndrome on three-dimensional motion during walking at different speeds," *Gait Posture*, vol. 30, no. 3, pp. 345–350, 2009.

[62] K. R. Embry, D. J. Villarreal, R. L. Macaluso, and R. D. Gregg, "Modeling the kinematics of human locomotion over continuously varying speeds and inclines," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 12, pp. 2342–2350, Dec. 2018.

[63] I. Kang, D. D. Molinaro, S. Duggal, Y. Chen, P. Kunapuli, and A. J. Young, "Real-time gait phase estimation for robotic hip exoskeleton control during multimodal locomotion," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 3491–3497, Apr. 2021.
[64] T. K. Best, C. G. Welker, E. J. Rouse, and R. D. Gregg, "Data-driven variable impedance control of a powered knee–ankle prosthesis for adaptive speed and incline walking," *IEEE Trans. Robot.*, vol. 29, no. 3, pp. 2151–2169, Jun. 2023.



Qiang Zhang (Member, IEEE) received the Ph.D. degree in biomedical engineering from The University of North Carolina at Chapel Hill (UNC), Chapel Hill, NC, USA, and The North Carolina State University (NCSU), Raleigh, NC, in 2021.

From 2021 to 2023, he was an Advanced Rehabilitation Research and Training Postdoctoral Research Fellow of the Closed Loop Engineering for Advanced Rehabilitation with the Joint Department of Biomedical Engineering, UNC/NCSU. He is currently an Assistant Professor with the Department of Mechanical Engineering and the Department of Chemical and Biological Engineering, University of Alabama, Tuscaloosa, AL, USA. His current research interests include advanced controls, wearable robotics and control, neuromuscular modeling and control, biomedical engineering, and human biomechanics.

Jennie Si (Fellow, IEEE) received the B.S. and M.S. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1985 and 1988, respectively, and the Ph.D. degree in electrical engineering from the University of Notre Dame, Notre Dame, IN, USA, in 1992.

She has been a Faculty Member with the School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ, USA, since 1991. Her research focuses on reinforcement learning control utilizing tools from optimal control theory, reinforcement learning, and neural networks. Her recent work also involves optimal adaptive control of wearable robots.

Dr. Si was a recipient of the NSF/White House Presidential Faculty Fellow Award in 1995 and the Motorola Engineering Excellence Award in the same year. She is a Distinguished Lecturer of the IEEE Computational Intelligence Society and an IEEE Fellow. She consulted for Intel, Arizona Public Service, and Medtronic, all in Phoenix, Arizona. She has served on several professional organizations' executive boards and international conference committees. She was an Advisor to the NSF Social Behavioral and Economical Directory. She served on several proposal review panels. She is a Past Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, the IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING, the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, and a Past Action Editor of *Neural Networks*. She is a Current Senior Editor of the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS.

Xikai Tu (Member, IEEE) received the M.E. and Ph.D. degrees in control science and engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2011 and 2016, respectively.

From 2019 to 2021, he was a Research Scientist with the Joint Department of Biomedical Engineering, North Carolina (NC) State University and the University of North Carolina at Chapel Hill, Raleigh, NC, USA. He is currently an Associate Professor with the School of Mechanical Engineering, Hubei University of Technology, Wuhan. His research interests include the application of advanced artificial intelligence control algorithms to rehabilitation robots, assistive robots, humanoid robots, and smart equipment powered by fuel cells.

Minhan Li (Student Member, IEEE) received the B.S. and M.S. degrees in mechanical engineering from Tianjin University, Tianjin, China, in 2015 and 2018, respectively, and the Ph.D. degree in biomedical engineering from North Carolina State University, Raleigh, NC, USA, and the University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, in 2022.

He is currently a Software Engineer with DiDi Labs, Mountain View, CA, USA. His current research interests include wearable robots, intelligent control, machine learning, computer vision, and rehabilitation engineering.

Michael D. Lewek received the B.S. degree in exercise science from Ithaca College, Ithaca, NY, USA, in 1997, and the MPT and Ph.D. (biomechanics and movement science) degrees from the University of Delaware, Newark, DE, USA, in 1999 and 2004, respectively.

He worked as a Postdoctoral Fellow with the Rehabilitation Institute of Chicago, Chicago, IL, USA, from 2004 to 2006. He is an Associate Professor with the Division of Physical Therapy, Department of Health Sciences, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. He is a Core Faculty Member and an Associate Director of the Human Movement Sciences Curriculum. Dr. Lewek is an Affiliate Faculty with the Joint Department of Biomedical Engineering of UNC/NC State where he serves as an Associate Director of the Closed Loop Engineering for Advanced Rehabilitation Core. His research focuses on the restoration and maintenance of walking in individuals who have had a stroke, or those who live with Parkinson disease.

He (Helen) Huang (Fellow, IEEE) received the B.S. degree from the School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an, China, in 2000, and the M.S. and Ph.D. degrees in bioengineering from Arizona State University, Tempe, AZ, USA, in 2002 and 2006, respectively.

She worked as a Postdoctoral Research Associate with the Rehabilitation Institute of Chicago, Chicago, IL, USA, from 2006 to 2008. She is the Jackson Family Distinguished Professor with the Joint Department of Biomedical Engineering, North Carolina State University and The University of North Carolina at Chapel Hill, Raleigh, NC, USA, and the Director for the Closed-Loop Engineering for Advanced Rehabilitation Core. Her research interest lies in neural-machine interfaces for prostheses and exoskeletons, human–robot interaction, adaptive and optimal control of wearable robots, and human movement control.

Dr. Huang was a recipient of the Delsys Prize for Innovation in Electromyography, the NIDILRR Switzer Fellowship, the NSF CAREER Award, the ASA Statistics in Physical Engineering Sciences Award, and the NC State ALCOA Foundation Distinguished Engineering Research Award. She is currently the Editor-in-Chief for the IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING and an Editorial Board Member for IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING. She is a Fellow of AIMBE and a NC State Faculty Scholar.