2022 IEEE International Conference on Robotics and Automation (ICRA)
May 23-27, 2022. Philadelphia, PA, USA

Admittance Control based Human-in-the-loop Optimization for Hip Exoskeleton: A New Approach

Varun Nalam, Xikai Tu, Minhan Li, Ming Liu, Jennie Si, *Fellow, IEEE* and He (Helen) Huang, *Senior Member, IEEE*

Abstract— **Human-in-the-loop (HIL) optimization usually optimizes assistive torque of exoskeletons to minimize the human's energetic expenditure in walking, quantified by metabolic cost. This formulation can, however, result in altered gait pattern of the human joint from the natural pattern, which is undesired. In this paper, we proposed a novel concept of HIL optimization of a hip exoskeleton. The optimization goal was to maintain the hip kinematics while providing optimal mechanical energy from the exoskeleton by modulating the admittance control. Policy iteration was used to optimize the switching time within the gait phase, at which a single parameter of the admittance controller was altered to provide assistance. The stiffness and equilibrium angle were considered as the two parameters for altering at the switching time, resulting in three possible modes of operation for the algorithm: (i) switching the equilibrium point, (ii) switching stiffness while equilibrium point is set at maximum extension and, (iii) maximum flexion. The optimization algorithm was found to converge for all three modes, with the equilibrium mode resulting in multiple solutions. Further analysis of power injected by the exoskeleton in the three modes showed that the first and third mode reduced human energetic exertion while the second mode increased human exertion. Implications of the results as well as the observed muscle activation patterns in response to assistance are discussed.**

I. Introduction

Over the last two decades, there has been a flurry of research efforts in developing lower limb exoskeletons to augment movement for healthy individuals or provide physical rehabilitation/daily assistance for individuals with motor deficits[1,2]. Specifically, research aimed at gait augmentation focused on able-bodied, health persons wearing single joint exoskeletons, aimed at reducing their energetic expenditure by modulating added torque at the joint in walking[3-5]. Since the hip joint is the most energetically intensive joint during gait[6], several hip exoskeletons have been developed for gait augmentation purposes[7-12].

Achieving gait augmentation requires coordination between human intent and exoskeleton actions. Therefore, the exoskeleton control has been generally structured in a hierarchical way with the higher-level controller detecting the user's intent and the lower-level controller providing the necessary assistance[13]. The user intent was recognized through EMG [14-16] or gait events[17,18] and predetermined corresponding torque was provided to the limb. However, seamlessly coordinating the actions of human and exoskeleton is challenging. The provided torque affects the dynamics of the human-exoskeleton system, which would set the human in a different state compared to the recognized intent. Since the human's response to the exoskeleton torque is not predictable, careful manual tuning of the device control parameters is needed[19], and the approaches found limited adoption for able-bodied users.

To account for changes in human behavior, recent studies have implemented human-in-the-loop (HIL) optimization methods [20-24]. In this approach, HIL optimization, in which a user-robot system is treated as a "black-box", to personalize the robotic control parameters automatically and heuristically. The HIL optimization algorithm iteratively searches for a fixed set of parameters that optimizes a specified objective function. For human gait augmentation, metabolic cost is the most used optimization objective[20,21,24]. The set of parameters are often the parameters that characterize an assistive torque function, such as timings, shape, and magnitude, in a specific gait phase. While HIL optimization methods account for variance in human response to assistance, the current formulation of HIL optimization for gait augmentation has several notable limitations. First, formulating the objective function using metabolic cost can lead to the challenge in the convergence of the algorithm because the metabolic cost measure is noisy and the influence of exoskeleton control on this human state in terms of energetic cost is unclear and indirect. In addition, the metabolic cost measure is impractical for daily use and responds slowly, further causing the optimization procedure to be inefficient[25]. More importantly, the HIL optimization attempts at reducing metabolic cost measures are found to affect human gait kinematics, which can be undesired for healthy users. The second limitation is the structure of action in the exoskeleton control space that focuses the definition of assistive torque profile [20,26]. A fixed torque profile lacks the adaptability to natural variations in human gait, compromising the assistance and restricting movement.

*Research partly supported by National Science Foundation #1563454, #1563921, #1808752 and #1808898. (Corresponding author: He (Helen) Huang; email: hhuang11@ncsu.edu).

V. Nalam, M. Li, M. Liu and H. Huang are with the NCSU/UNC Department of Biomedical Engineering, NC State University, Raleigh, NC, 27695-7115;University of North Carolina at Chapel Hill, Chapel Hill, NC 27599 USA

J. Si is with the Department of Electrical, Computer, and Energy Engineering, Arizona State University, Tempe, AZ, 85281 USA. (email: si@asu.edu).

X. Tu is with the Department of Mechanical Engineering, Hubei University of Technology, Wuhan, Hubei, 430068 China (e-mail: tuxikai@gmail.com).

Figure 1. Modular Hip Exoskeleton attached to an able bodied user

The limitation of current HIL optimization motivated us to reconsider the problem formulation for human gait augmentation via exoskeleton. Fundamentally different from the current concept, we proposed to optimize a hip exoskeleton assistance without changing the gait kinematics in local hip joint. Therefore, the objective function was formulated by the local hip motion, which is directly influenced by the exoskeleton control and is practical and reliable to use. We designed admittance control as the low-level exoskeleton control, which is switched between two different sets of admittance values. The optimal time for switching the admittance value is obtained through HIL optimization. Hence, the approach reduces dimensionality by optimizing only one parameter and ensures that the kinematics of the user is not affected, addressing the limitations of the forementioned approaches. The admittance control ensures compliance in the system to account for natural variations in gait[27]. Further, the admittance parameters are adjusted based on user preference, ensuring user input, and user comfort during assistance.

This paper describes our study to test the feasibility of the proposed approach. We first evaluated the performance of HIL optimization implemented using reinforcement learning (RL) based policy iteration[28,29] in finding an optimal switching time for the objective function. The power input from the exoskeleton as well as the muscle activation of the user were then analyzed to understand (i) if there is a reduction in human exertion during walking and (ii) which admittance parameters have to be switched for best energetic savings. Therefore, our paper made the following contributions. 1) We presented a new formulation of HIL optimization of hip exoskeleton control for gait augmentation.2) We implement the new HIL optimization based on RL, applied to 3 modes using admittance control.3) New knowledge on physical human-robot interaction (e.g., the influence of control timing on switching different impedance values on human energetic exertion) will be obtained. The new knowledge will inform the future design of optimized control for minimizing the user's walking effort.

II. METHODS

A. Hardware

For the purposes of the study, a novel modular hip exoskeleton (Fig. 1) had been utilized. The exoskeleton had been validated to safely interact with human subjects and provide the required assistance torque[26,30]. While the hip exoskeleton was designed to provide assistance along 4 Degrees of Freedom, this study focused only on the Flexion-Extension motion of the hip joint. For each of the two actuators controlling the flexion-extension of both limbs, an admittance controller was implemented with stiffness (K), damping (B) and inertia (I) about an equilibrium angle (θ_e). The input from the load cell(T) placed on the arm connecting the actuator to the limb was taken by the controller to generate a velocity signal satisfying the specified impedance parameters in accordance with (1). The controller was implemented on a real-time PC using EtherCat (TwinCAT 3.1) protocol executing at 1000 Hz.

$$T = K(\theta - \theta_e) + B\dot{\theta} + I\ddot{\theta} \quad (1)$$

B. Assistance protocol

The popular approach of assistance using hip exoskeletons is to set the impedance parameters (K, B, I and θ_e) close to zero for most of gait as a "zero-torque mode" and provide assistance at optimal phases of gait. The optimal assistance is generally provided by switching to a torque controller from the admittance (or impedance) controller and providing a predetermined torque profile. In this approach, we provided non-zero admittance parameters across the entire gait cycle and switched one parameter at an optimal time during gait. The optimal time was determined using an RL algorithm such that the gait kinematics of the hip joint remains similar to that of no assistance (zero torque mode). If the kinematics of the hip joint are the same during both zero torque and assistance modes, the total torque acting on the hip joint should be the same (2,3), with the exoskeleton providing no torque in zero torque condition. Additionally, the work done across the gait

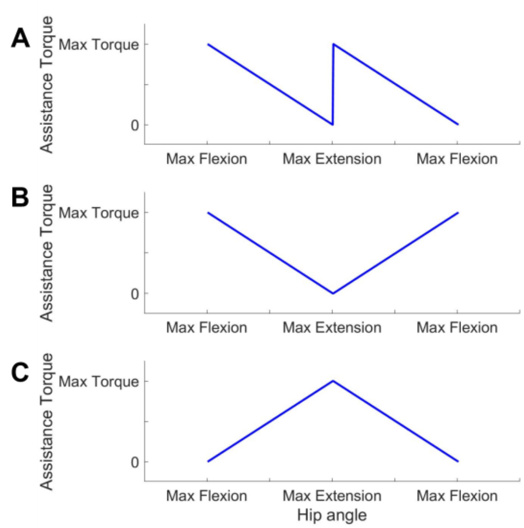

Figure 2. The assistance torque (torque in direction of motion) for **A** Equilibrium mode, **B** Stiffness Extension Mode and **C** Stiffness Flexion Mode

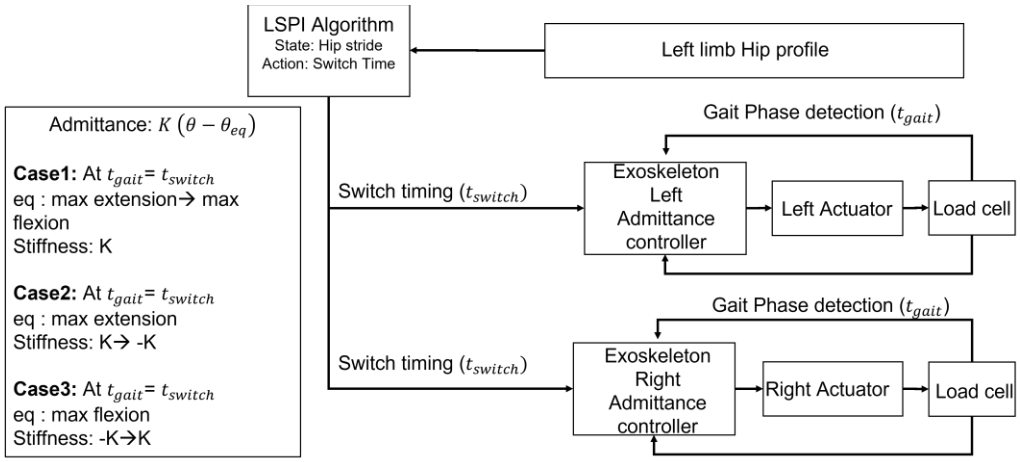

Figure 3. Realization of the Reinforcement Learning Implementation.

cycle would be the same in both cases if the kinematics were not affected. Since the work done by the exoskeleton in the zero-torque mode would be 0, any positive work done by the exoskeleton would result in lower amount of work done by the human and thereby reducing human exertion (5).

$$T = I_{tot}\ddot{\theta} \tag{2}$$

$$T_{assist} = T_{exo} + T_{human} = T_{zero\ torque} \tag{3}$$

$$Work = \int T\, d\theta \tag{4}$$

$$Work = \int (T_{exo}) d\theta + \int (T_{Human})\, d\theta \tag{5}$$

To test the hypothesis, the stiffness (K) and equilibrium angle (θ_e) were considered for switching during the algorithm, leading to three possible cases each having unique torque profile characteristics. In the first case, θ_e was switched from maximum extension to maximum flexion as observed during zero torque mode. In the second case, the stiffness was switched from positive (K) to negative (-K) and θ_e was fixed at maximum extension, while in the third the stiffness was switched from negative to positive with θ_e fixed at maximum flexion. The stiffness (K) was predetermined based on user comfort and the optimal time is determined using reinforcement learning. The hip joint kinematics and the power injected by the exoskeleton for each of the three cases was analyzed to validate the hypothesis. The three cases are referred to as the equilibrium mode, stiffness extension mode and stiffness flexion mode for the rest of the paper (Fig. 2).

C. Reinforcement Learning Control

The goal of reinforcement learning control was to obtain the optimal switch times of assistance to be validated for each of the three modes. The heel strike of each foot was taken as the start of a gait cycle and the switch time was normalized to the gait cycle. The stride angle of the hip, which is the angle between maximum flexion and extension across a gait cycle was compared to the observed stride angle in zero torque mode to assess the optimality of the switch time. For the reinforcement learning (RL) problem, the state (x) was considered to be the error in stride angle of the hip joint and the action (u) was the change in switch time. An initial switch time was chosen randomly at the start of the algorithm.

The control u_k, the action at the k[th] iteration, was obtained from the control policy (π) as shown in (6) in which δ_k was discounted noise added to facilitate exploration. Additionally, a cost (C) for each iteration is estimated using (7), where Q and R are positive definite matrices.

$$u_k = \pi(x_k) + \delta_k \tag{6}$$

$$C(x_k, u_k) = x_k^t Q x_k + u_k^t R u_k \tag{7}$$

We define the state-action Q-value function as

$$\mathrm{Q}(x_k, u_k) = \mathrm{C}(x_k, u_k) + \textstyle\sum_{j=1}^{\infty} \gamma \mathrm{C}\left(x_{k+j}, \pi\left(x_{k+j}\right)\right) \tag{8}$$

where $\mathrm{C}(x_k, u_k)$ is the stage cost or instantaneous cost function and γ is a discount factor. Note that the $\mathrm{Q}(x_k, u_k)$ value is a performance measure when action u_k is applied at state x_k and the control policy π is followed thereafter. We used least-squares policy iteration to solve for the optimal policy π^* from the optimal $\mathrm{Q}*$ value.

A sample consisting of the state, the action and the reward was obtained every iteration. The sample was collected across 8 gait cycles obtained with the updated switch time (Fig. 3) and the state was obtained by averaging the hip stride angle across the last 5 gait cycles, allowing 3 cycles for transition. The policy (π) was updated every 8 iterations using the least squares policy iteration [28] on collected samples to further accelerate the convergence. The switch time was found to have converged when the cost was within a predetermined margin for three consecutive samples.

D. Experimental Protocol and analysis

The purpose of this study was to verify if an optimal switching time could be obtained for the three modes using RL and, further investigate the power and muscle activation characteristics of the three modes during optimal switching times. One subject (Age:29, Sex: Male, Height: 171 cm and Weight: 78 Kg) had been recruited to perform the trial. The study was approved by the institutional review board at the University of North Carolina at Chapel Hill and the participant provided informed consent. Surface electromyograpy (EMG) sensors (Motion-labs MA-300, LA, USA), were placed on the Gluteus Maximus (GLU), Bicep Femoris (BF), Semitendinosus (Semi), Vastus Lateralis (Vast) and Rectus Femoris (RF) muscles of both limbs of the subject. The entire study was performed while the subject was instructed to walk on a treadmill set to 1.2 m/s.

Prior to the tuning trial, the subject was instructed to walk without the exoskeleton and with the exoskeleton in zero torque mode for 2 minutes to obtain the baseline EMG and the reference stride angle for the hip joint. Three tuning trials were then performed for each of the three modes resulting in a total of 9 tuning trials. The admittance parameters K, B and I were set to 10 Nm/rad, 0.5 Nms/rad and 0.1 Kgm^2 and the stiffness was lowered in case of user discomfort. The tuning trials were performed till convergence unless the algorithm failed to converge within 8 min. If the tuning trial converged, a validation trial was performed with the optimal switching time obtained from the tuning trial. The subject was instructed to walk on the treadmill for 2 min for the validation trial, while the exoskeleton provides assistance. A total of 20 trials, including 9 tuning and 9 validation trials were performed for the study. Sufficient rest was provided between trials to ensure subject comfort and prevent fatigue.

To analyze the validity of each of the three approaches, the number of iterations taken to converge as well as the optimal switching time behavior of the algorithms were analyzed. In addition, the kinematics of the hip joint was compared between the zero-torque mode and the optimal assistance mode for each of the three modes. To compare the kinematic behavior, the variance accounted for (%VAF or R^2) and Pearson's coefficient (correlation coefficient) for each trial is estimated. Finally, the power injected by the exoskeleton into the hip-exoskeleton complex as well as the muscle activation of predominant flexor and extensor muscles is evaluated across the gait cycle. The muscle activations are normalized using the observed muscle activations of corresponding muscles during normal walking without the exoskeleton. The observed results and inferences are described below.

III. Results

We performed tuning trials for all three modes. While the subject found the stiffness during equilibrium mode to be comfortable, the stiffness was reduced to 5 Nm/rad during stiffness extension and stiffness flexion mode to ensure user comfort. The convergence behavior, the kinematics along with the power and EMG behavior are described below.

A. Convergence

As shown in Fig. 4, the algorithm converged to an optimal solution in all the three modes. For the equilibrium point mode, the algorithm converged between 16-24 iterations for all three trials. Since the policy updated every 8 iterations, the third policy could be inferred to be the ideal policy for each trial. In comparison, two of the trials in the stiffness extension mode converged around 24-28 iterations, while the third one took over 40 iterations to converge. Based on the oscillations in the state value, it can be inferred that the hip joint-exoskeleton complex was unstable leading to varying human behavior, thereby affecting convergence. Finally, the stiffness flexion trials converged the fastest. Two of the trials converged even without an updated policy, which implies that the random policy was sufficient to aid in convergence of the switch time.

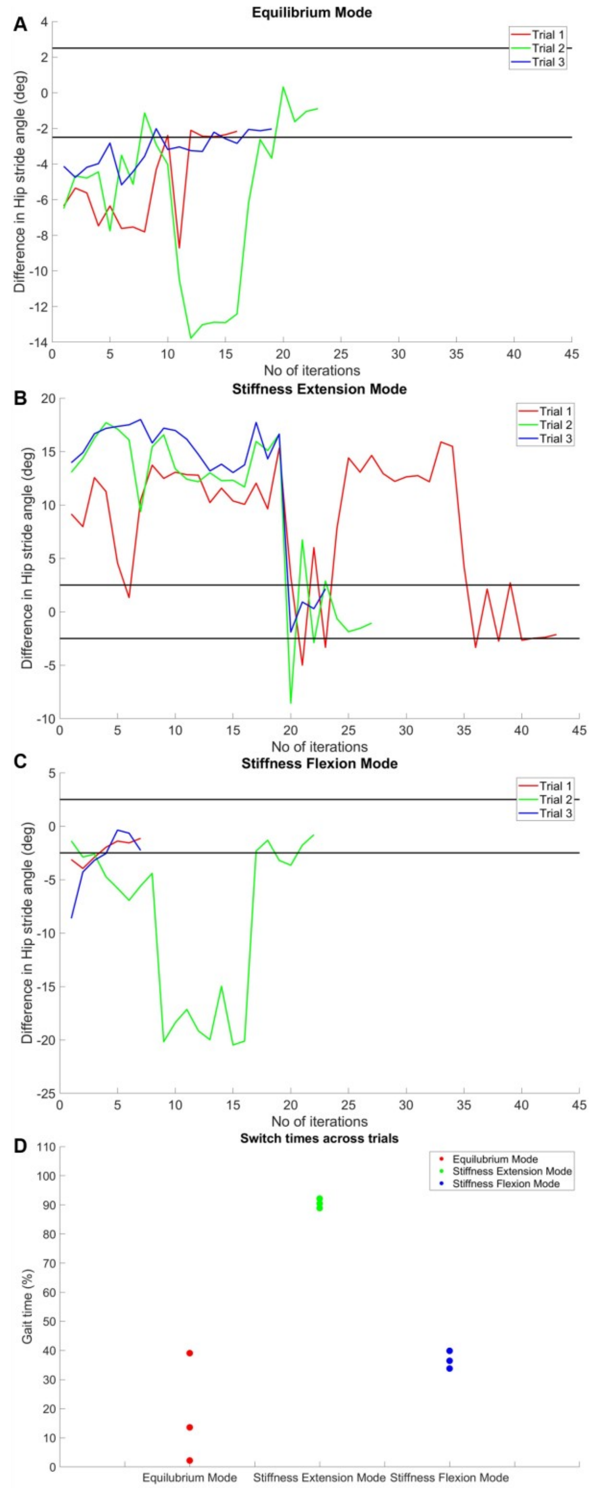

Figure 4. Tuning performance for **A** Equilibrium mode, **B** Stiffness extension mode and, **C** Stiffness Flexion Mode. The algorithm only converged once the angles were within ± 2.5 degrees of hip stride angle observed in zero torque mode **D** Optimal switch timing across trials for all three modes.

The third trial, converged in 23 iterations and seemed to perform similar to the other two trials as soon as the policy was updated. Overall, the algorithm seemed to converge faster for stiffness flexion mode and the slowest for the stiffness extension mode.

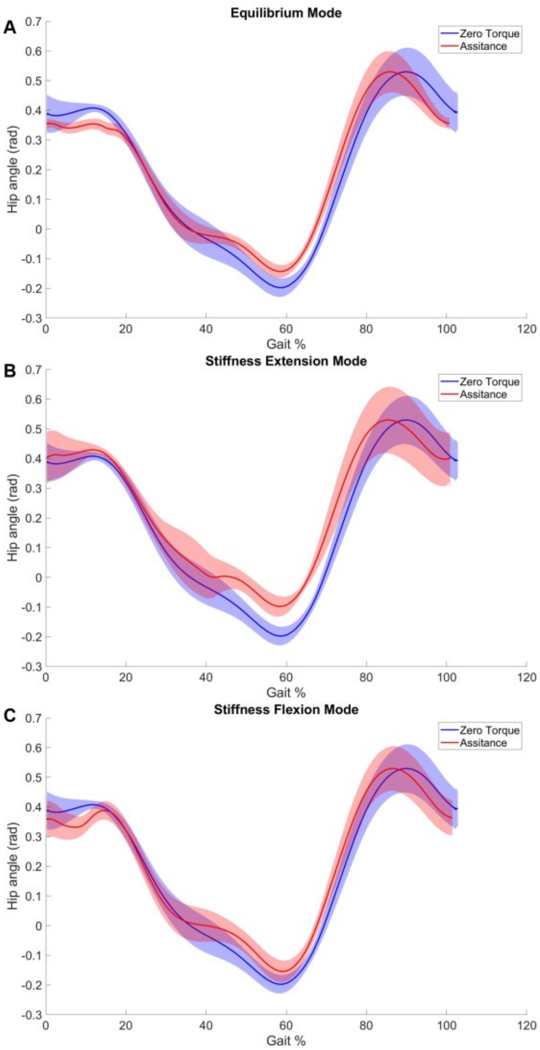

Figure 5. Comparison of hip gait kinematic profiles for the three modes

TABLE I. MEAN AND STD DEVIATION OF % VAF AND CORRELATION COEFFICIENT ON COMPARING HIP KINEMATICS FOR EACH OF THE THREE MODES WITH ZERO TORQUE CONDITION

Assistance Mode	***VAF (%)***	***Correlation Coefficient***
Equilibrium Mode	97.46 (0.28)	0.99 (<0.001)
Stiffness Extension Mode	93.17 (3.28)	0.97 (0.01)
Stiffness Flexion Mode	93.01 (1.53)	0.97 (0.01)

Looking at the final switching times, the equilibrium point mode converged to different optimal times while both extension and flexion stiffness mode converged to similar switching time for all three trials. The results imply that there is only one optimal solution for the trials in which stiffness is switched but multiple optimal solutions when equilibrium point is switched. Understanding the differences in power input during the various solutions could help in developing optimal targets for the algorithm.

B. Kinematics

Since the algorithm only considered the maximum extension and flexion of the hip joint, it is possible that the kinematics of the hip joint were radically altered with assistance even when the hip stride angle was the same. On investigating the % VAF and Pearson's coefficient, the equilibrium point mode was found to have a %VAF greater than 97% with the correlation coefficient greater than 0.99 when the observed hip profile is compared to that of the zero-torque condition. For the stiffness modes, the %VAF was found to be greater than 93% with correlation greater than 0.96 for both cases, which shows that the gait kinematics of the hip remained consistent even with assistance. The kinematic profile of the hip for the three modes is shown in Fig. 5 while Table I shows the mean and standard deviations of %VAF and correlation for the three modes.

C. Power and EMG

It is observed that the total power injected into the system during a gait cycle is 25.14 Nm/s for equilibrium point mode, -7.82 Nm/s for stiffness extension mode and 15.73 Nm/s for stiffness flexion mode (Fig. 6). From (5), it can be asserted that both equilibrium mode and stiffness flexion mode result in reduced exertion for the subject during gait, while stiffness extension mode resulted in increased exertion. Further investigating the temporal nature of the power input (Fig. 7), a later switching time during equilibrium point mode resulted in a higher net positive power during extension but a similarly higher net negative power during flexion. Hence, while the switching times for all three trials were different, the total power was similar. Looking at the EMG responses, the equilibrium point mode resulted in reduction in muscle activity of the extensor muscles (BF and Semi) with a slight increase in flexor muscle activation. The increased activation in flexor muscles correlated with the negative power observed in the later part of gait in equilibrium point mode as shown in Fig. (). In comparison, stiffness extension mode showed an increase in extensor muscle activity compared to flexor muscle activity, while stiffness flexion mode shows a reduced activity in flexor muscles as well as one of the extensor muscles (BF).

IV. DISCUSSION

The purpose of the study was to test the feasibility of the HIL approach aimed at reducing human exertion without affecting the hip kinematics during gait. The viability of using reinforcement learning to obtain the optimal timings as well as the impact of admittance parameters modulated to ensure reduction in human exertion were analyzed. The convergence results have shown that the algorithm is capable of finding an optimal switching time of an admittance control parameters without changing the gait pattern, while the results of exoskeleton augmentation power validate the admittance control approach in reducing human exertion for two of the three modes analyzed.

A. Implications of the observed results

The reinforcement learning algorithm was found to converge for all three modes analyzed in the results. Based on ease of convergence, equilibrium and stiffness flexion mode were found to be more stable approaches. One possible explanation for the multiple switching times in equilibrium mode could be that the hip stride angle is robust to changes in switching time. However, that possibility can be eliminated as the state was found to vary with varying switching times as shown in the convergence plot. Interestingly, since the total

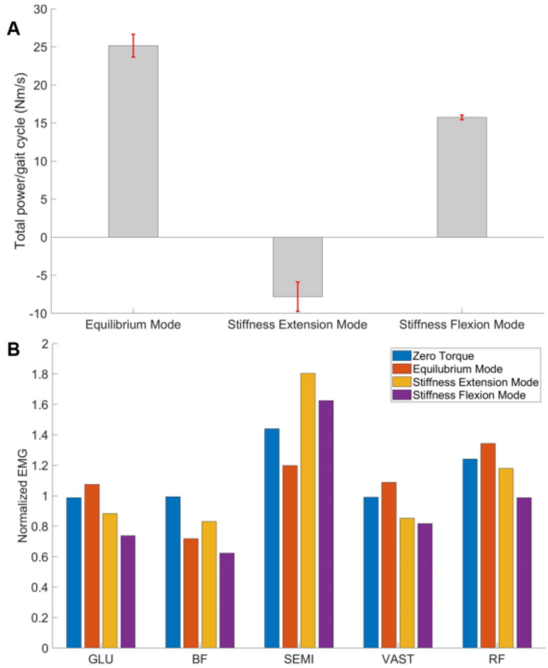

Figure 6. **A** Total Power across the gait cycle and **B** Normalized EMG responses of predominant thigh muscles across gait

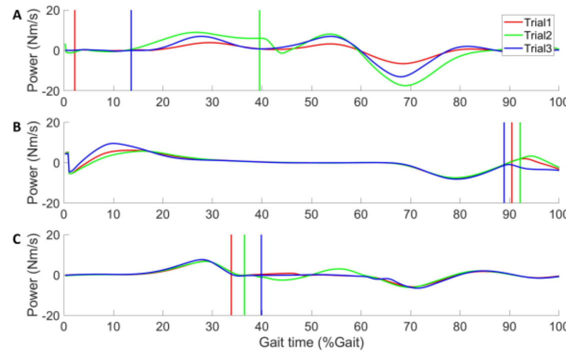

Figure 7. Temporal distibution of power across the gait cycle for **A** Equilubrium mode, **B** Stiffness extension mode and, **C** Stiffness Flexion Mode. The vertical lines denote switching times for each trial

power injected into the system during trials was similar, the optimal timing from the possible solutions can be chosen based on user preference without any energetic penalty.

Even though the algorithm uses only hip stride angle as the state for the algorithm, the resulting kinematics of the hip joint correlated with the zero-torque condition ($r>0.97$ and $r^2>93\%$). So, it is not necessary to add additional state information for the algorithm to match the kinematics of the hip joint to that of no assistance case, reducing the complexity of the required algorithm. More importantly, since the kinematics were not altered, any positive energy injected by the exoskeleton would result in reduction in human exertion which was further validated through the observed EMG measures. A positive power input during earlier phases of gait as seen in the temporal power distribution graph of equilibrium mode (Fig. 7) correlated with reduction in extensor muscle activation, while positive power input in the later parts of gait observed in both stiffness modes resulted in lower activation of flexor muscles. Overall, positive power across gait was observed to correlate with reduced extensor muscle activity, specifically to that of the Bicep Femoris. BF activation can therefore be used as a state to further optimize for reduced human exertion in future iterations.

B. *Significance of the current approach*

HIL approaches aimed at reducing metabolic costs were found to alter gait kinematics. However, changes in gait kinematics might not be preferred by the user even with the resulting energetic cost savings. The current approach reduces human exertion without any changes to the gait kinematics by compensating for some of the torque generated by the user and could be preferred by users compared to other approaches. The choice of using an admittance controller and altering parameters at an optimal time is inherently safer. Admittance controllers have compliance incorporated into the system enabling the controller to adapt to any changes in the natural gait pattern of the user[27]. In comparison, the optimized torque profiles cannot adapt to any changes in the gait pattern after optimization. So, the torque profile could have adverse effects in response to natural variation in human gait. In addition, even the simplest realizations of torque profile require multiple parameters that must be optimized, resulting in longer duration of optimization which is eliminated in this approach. The admittance parameters are predetermined based on user preference and the only parameter optimized is the switching time during the gait cycle. This ensures that the optimization algorithm is simplified while also taking into account user preference for assistance.

Finally, the optimal torque profiles may not translate across varying conditions such as incline and decline walking, varying environments or walking speeds, requiring separate optimization process for each condition. The current approach uses least squares policy iteration[28], which finds the optimal policy in addition to the switching time. While the switching time might not translate across the various walking conditions, the policy function relates to the behavior of the hip-exoskeleton complex and should be independent of the walking conditions and environment. Hence the proposed approach based on policy iteration is better equipped to adapt to different walking conditions and environments by utilizing the optimal policy deduced by the optimization.

C. *Future Scope*

While the current approach is shown to have reduced human exertion, it is possible that human exertion could be further reduced during gait. Since the RL algorithm can easily be scaled, additional switching times could be introduced in a gait cycle in either the equilibrium mode, stiffness flexion mode or a combination of both. Understanding the relationship between the admittance parameters and switching time could aid in development of better tuning protocols that would emphasize user preference and ensure user safety during optimization. The optimization could be performed at low admittance parameter values to mitigate effects of non-optimal switch times and on convergence, admittance parameters preferred by the user could be tuned using the relationship. Successful realization of this approach would simultaneously facilitate user preferred assistance and reduction in energetic cost of locomotion.

REFERENCES

[1] N. Li, L. Yan, H. Qian, H. Wu, J. Wu, and S. Men, "Review on lower extremity exoskeleton robot," *The Open Automation and Control Systems Journal,* vol. 7, no. 1, 2015.

[2] A. J. Young and D. P. Ferris, "State of the art and future directions for lower limb robotic exoskeletons," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 25, no. 2, pp. 171-182, 2016.

[3] K. Seo, J. Lee, and Y. J. Park, "Autonomous hip exoskeleton saves metabolic cost of walking uphill," in *2017 International Conference on Rehabilitation Robotics (ICORR)*, 2017: IEEE, pp. 246-251.

[4] M. Shepertycky, S. Burton, A. Dickson, Y.-F. Liu, and Q. Li, "Removing energy with an exoskeleton reduces the metabolic cost of walking," *Science,* vol. 372, no. 6545, pp. 957-960, 2021.

[5] S. Galle, P. Malcolm, S. H. Collins, and D. De Clercq, "Reducing the metabolic cost of walking with an ankle exoskeleton: interaction between actuation timing and power," *Journal of neuroengineering and rehabilitation,* vol. 14, no. 1, pp. 1-16, 2017.

[6] D. J. Farris and G. S. Sawicki, "The mechanics and energetics of human walking and running: a joint level perspective," *Journal of The Royal Society Interface,* vol. 9, no. 66, pp. 110-118, 2012.

[7] T. Zhang, M. Tran, and H. Huang, "Design and Experimental Verification of Hip Exoskeleton with Balance Capacities for Walking Assistance," *IEEE/ASME Transactions on Mechatronics,* vol. PP, no. 99, pp. 1-1, 2018, doi: 10.1109/TMECH.2018.2790358.

[8] T. Zhang, M. Tran, and H. Huang, "Admittance shaping-based assistive control of SEA-driven robotic hip exoskeleton," *IEEE/ASME Transactions on Mechatronics,* vol. 24, no. 4, pp. 1508-1519, 2019.

[9] A. T. Asbeck, K. Schmidt, and C. J. Walsh, "Soft exosuit for hip assistance," *Robotics and Autonomous Systems,* vol. 73, pp. 102-110, 2015/11/01/ 2015, doi: https://doi.org/10.1016/j.robot.2014.09.025.

[10] K. Junius, M. Degelaen, N. Lefeber, E. Swinnen, B. Vanderborght, and D. Lefeber, "Bilateral, misalignment-compensating, full-DOF hip exoskeleton: design and kinematic validation," *Applied bionics and biomechanics,* vol. 2017, 2017.

[11] A. Parri *et al.*, "A portable active pelvis orthosis for ambulatory movement assistance," in *Wearable Robotics: Challenges and Trends*: Springer, 2017, pp. 75-80.

[12] I. Kang, H. Hsu, and A. J. Young, "Design and validation of a torque controllable hip exoskeleton for walking assistance," in *Dynamic Systems and Control Conference*, 2018, vol. 51890: American Society of Mechanical Engineers, p. V001T12A002.

[13] V. Kumar, Y. V. Hote, and S. Jain, "Review of exoskeleton: history, design and control," in *2019 3rd International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE)*, 2019: IEEE, pp. 677-682.

[14] H. Kawamoto, S. Lee, S. Kanbe, and Y. Sankai, "Power assist method for HAL-3 using EMG-based feedback controller," in *SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme-System Security and Assurance (Cat. No. 03CH37483)*, 2003, vol. 2: IEEE, pp. 1648-1653.

[15] K. Gui, H. Liu, and D. Zhang, "A practical and adaptive method to achieve EMG-based torque estimation for a robotic exoskeleton," *IEEE/ASME Transactions on Mechatronics,* vol. 24, no. 2, pp. 483-494, 2019.

[16] C. Fleischer, C. Reinicke, and G. Hommel, "Predicting the intended motion with EMG signals for an exoskeleton orthosis controller," in *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2005: IEEE, pp. 2029-2034.

[17] C. L. Lewis and D. P. Ferris, "Invariant hip moment pattern while walking with a robotic hip exoskeleton," *Journal of biomechanics,* vol. 44, no. 5, pp. 789-793, 2011.

[18] K. Seo, J. Lee, Y. Lee, T. Ha, and Y. Shim, "Fully autonomous hip exoskeleton saves metabolic cost of walking," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016: IEEE, pp. 4628-4635.

[19] J. Zhang, C. C. Cheah, and S. H. Collins, "Experimental comparison of torque control methods on an ankle exoskeleton during human walking," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015: IEEE, pp. 5584-5589.

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

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

[22] S. Song and S. H. Collins, "Optimizing exoskeleton assistance for faster self-selected walking," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 29, pp. 786-795, 2021.

[23] D. Wei *et al.*, "Human-in-the-loop control strategy of unilateral exoskeleton robots for gait rehabilitation," *IEEE Transactions on Cognitive and Developmental Systems,* 2019.

[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] K. A. Ingraham, E. J. Rouse, and C. D. Remy, "Accelerating the estimation of metabolic cost using signal derivatives: implications for optimization and evaluation of wearable robots," *IEEE Robotics & Automation Magazine,* vol. 27, no. 1, pp. 32-42, 2019.

[26] X. Tu, M. Li, M. Liu, and J. Si, "A Data-Driven Reinforcement Learning Solution Framework for Optimal and Adaptive Personalization of a Hip Exoskeleton," *arXiv preprint arXiv:2011.06116,* 2020.

[27] C. Liang and T. Hsiao, "Walking strategies and performance evaluation for human-exoskeleton systems under admittance control," *Sensors,* vol. 20, no. 15, p. 4346, 2020.

[28] M. G. Lagoudakis and R. Parr, "Least-squares policy iteration," *The Journal of Machine Learning Research,* vol. 4, pp. 1107-1149, 2003.

[29] 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 Transactions on Robotics,* 2021.

[30] X. Tu, M. Li, M. Liu, J. Si, and H. Huang, "Reinforcement Learning Control of a Hip Exoskeleton for Personalized Walking Assistance," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, Xi'an, China 2021, vol. in press.