User-Adaptive Variable Damping Control Using Bayesian Optimization to Enhance Physical Human-Robot Interaction

Fatemeh Zahedi, Dongjune Chang and Hyunglae Lee*, *Member, IEEE*

***Abstract*—This paper presents a user-adaptive variable damping controller that enhances the overall performance of coupled human-robot systems in terms of stability, agility, user effort, and energy expenditure during physical human-robot interaction. The controller accounts for impedance properties of the human limbs and adaptively changes robotic damping from negative to positive values based on user's intent of motion while minimizing energy of the coupled human-robot system. Bayesian optimization is used to evaluate an unknown objective function and optimize noisy performance, which builds on a Gaussian process to account for the uncertainty of human behaviors and noisy observations. To validate the effectiveness of the presented approach and evaluate its potential applications in real-world scenarios, we performed human experiments using a common robotic arm manipulator. Experimental results from five pilot subjects demonstrated that the controller does not require a long parameter tuning process. Compared to variable damping control without user-adaptive parameter changes, the presented adaptive control strategy could reduce ~45% energy expenditure and achieve average performance improvement of ~20% when several performance metrics of stability, agility, and user effort are considered together.**

***Index Terms*—Physical human-robot interaction, Assistive robotics, impedance control, interaction control**

I. Introduction

THE rise in popularity of physical human-robot interaction field has presented several technical challenges, the most impactful of which is how to maintain user safety during physical interaction. This is most commonly addressed by designing the control mechanism of the robot to prioritize stability of the coupled system but at the expense of performance. One such example is an impedance/admittance controller which adds positive damping to the system to properly dissipate energy and ensure stability [1, 2]. While this approach can guarantee stability [3], it may reduce the user's agility and require additional user effort to overcome the resistive behavior of the robot.

To improve overall performance in pHRI, stability, agility, and user effort all need to be considered when designing robotic controllers [4]. Many efforts have been made to improve the trade-off between stability and agility or to reduce user effort, but not all together. Some task-dependent approaches focused on reducing the user effort by tuning impedance parameters based on demonstration of the desired task by the user [5] or adapting the parameters by detecting deviation from the nominal behavior of the admittance controller [6]. Other less task-dependent approaches change the impedance parameters using some metrics of user intent such as velocity or force at the point of interaction [7] or by minimizing their jerk profiles using reinforcement learning to improve agility [8].

To overcome the limitations of these previous work, the authors have developed a robotic controller to improve the trade-off between stability and agility while reducing user effort [9]. The controller modulated robotic damping from negative to positive damping based on user's intent of motion and knowledge on inherent human limb impedance [10-12]. However, similar to other methods, this work requires a long separate tuning process to determine several important controller parameters. Furthermore, uncertainty of human behavior and noisy observations have not been considered in the tuning process. The controller was also designed based on overall human biomechanics characteristics, and it was not capable of adjusting its parameters in a user-specific manner. However, complexity and redundancy of the human neuromuscular system and individual differences can lead to various responses for the same controller, and thus a control strategy that performed well in one user may perform poorly on another [13, 14]. These differences highlight the necessity of user-specific strategies in pHRI.

Several methods have been developed that investigate each user's performance and find the optimal control parameters through some curve fitting processes [15, 16]. However, computational complexity of these methods increases exponentially with the increase of parameter dimensions, and the methods involve lengthy experiment and evaluation. Human-in-the-loop (HIL) optimization is a promising approach to overcome these challenges by adjusting control parameters based on real-time measurements of biomechanical signals without a lengthy tuning protocol. The HIL optimization method using Bayesian Optimization (BO)

Manuscript received: September 7, 2021; Revised: December 6, 2021; Accepted: January 3, 2022.

This paper was recommended for publication by Editor Tamim Asfour upon evaluation of Associate Editor and Reviewers' comments.

Research supported by National Science Foundation Award #1846885 and #1925110.

Fatemeh Zahedi, Dongjune Chang, and Hyunglae Lee are with the School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ 85287, USA (e-mail: {fzahedi1, Dongjune.Chang, hyunglae.lee}@asu.edu). *: corresponding author

Digital Objective Identifier (DOI): see top of this page.

has been developed to rapidly identify optimal control parameters that minimize the metabolic cost of walking [14]. Another HIL optimization method, benefited from the covariance matrix adaptation evolutionary strategy, identified the best exoskeleton characteristics for each device type and individual user to improve running performance [17]. Although these achievements are impressive, none of the studies have directly addressed an important trade-off problem between stability, agility, and user effort during pHRI.

The goal of this study is to address this trade-off issue and develop a user-adaptive variable damping controller to adjust its parameters based on real-time measurements of human biomechanical signals such as limb kinematics, interaction force, while minimizing energy of the coupled human-robot system. The presented controller not only captures user proficiency in pHRI but also addresses the challenges of the lengthy protocol and tuning process. The controller builds upon BO, which is an efficient global optimization strategy for the evaluation of a noisy, unknown, and expensive objective function [18]. Gaussian process is also used to model noisy observations and uncertainty in human behavior [19, 20].

To evaluate performance of the presented controller and its applicability in real-world scenarios, we performed human experiments in which users interacted with a robotic arm manipulator. Stability, agility, user effort, and energy expendiure were quantified to demonstrate the overall performance improvement of this approach compared to the previous work [9].

II. Methods

A. Problem Statement

Impedance control modulates a set of impedance parameters, stiffness (K_r) and damping (B_r), at the interaction port between the human and robot. The variable damping control varies the damping component of the impedance. The controller is described as (1):

$$F(x, \dot{x}) = B_r \dot{x} + K_r(x - x_e) + g \tag{1}$$

where x is the position, x_e is the desired equilibrium position, $\dot{x}$ is the velocity, g is the gravity compensation force, F is the output controller force, and K_r and B_r are stiffness and damping parameters.

The goal of variable damping control is to modulate robotic damping to help the user perform with high stability and agility, and with low effort and energy expenditure. If we know the user's intent of motion, we may modulate robotic damping to maximize efficiency. One effective way to identify user's intent of motion is by collecting kinematic information during pHRI. We define "intent of motion" as the product of velocity and acceleration, $\dot{x}\ddot{x}$, which is a scaled version of the change in kinetic energy of the system. Based on the sign of this term, we can determine when to apply positive or negative damping to the system. As long as the magnitude of the negative robotic damping is less than the magnitude of inherent damping of human user in the extremity, the coupled system can remain passive and stable. Therefore, if user's intent of motion is positive, applying negative robotic damping (injecting energy) can help the user to move faster and when the user's intent of motion is negative, positive robotic damping (energy dissipating) can help the user to slow down and stabilize the limbs and joints.

For the smooth transition between positive and negative damping, a piecewise logistic function was defined as (2):

$$B_r(\dot{x}\ddot{x}) = \begin{cases} \frac{2b_{LB}}{1+e^{-k_p \dot{x}\ddot{x}}} - b_{LB}, & \dot{x}\ddot{x} \geq 0 \\ -\frac{2b_{UB}}{1+e^{-k_n \dot{x}\ddot{x}}} + b_{UB}, & \dot{x}\ddot{x} < 0 \end{cases} \tag{2}$$

where B_r is the robotic damping applied to the system and b_{LB} and b_{UB} are the lower and upper bound of the damping range. Tuning constants k_p and k_n are used as they specify the logistic growth rate of the function, which determine how quickly the transition between b_{LB} and b_{UB} happens (3):

$$k_p = \frac{-ln\left(\frac{1-s}{1+s}\right)}{\dot{x}\ddot{x}_{max}}, \qquad k_n = \frac{-ln\left(\frac{1+s}{1-s}\right)}{\dot{x}\ddot{x}_{min}} \tag{3}$$

where $\dot{x}\ddot{x}_{max}$ and $\dot{x}\ddot{x}_{min}$ are the maximum and minimum user intent, respectively, and s is the sensitivity of the robotic damping function. In this study, $s = 0.95$, which means that the robotic damping becomes $0.95b_{LB}$ at $\dot{x}\ddot{x}_{max}$ and $0.95b_{UB}$ at $\dot{x}\ddot{x}_{min}$. With this piecewise logistic function, positive and negative damping regions can be defined independently.

Careful selection of the controller parameters b_{LB}, b_{UB}, k_p and k_n is critical to maximize the effectiveness of the variable damping controller. In the previous study [9], fixed values were selected for b_{LB} and b_{UB} based on previous studies of characterizing the inherent damping of the human arm [10-12], which did not allow parameter adaptation for different users. In that study, other user-specific parameters, specifically k_p and k_n, were determined with a separate lengthy tuning process before main experiments, and one set of the selected parameters was used throughout the whole experiments without any adjustment.

This method of selecting parameters was successful and resulted in enhancing the trade-off between stability/agility and reducing user effort, but consistent application of the fixed upper and lower bounds of damping across different users limited the efficiency of the controller. Additionally, the tuning process to determine k_p and k_n, was time intensive and did not consider the uncertainty of human behavior and noise of observation such as limb kinematics and interaction force, and user proficiency and adaptability.

If we select each parameter based on the biomechanical characteristics of each user and adjust them according to their proficiency during pHRI, we expect more performance improvement. Further, by leveraging real-time measurements of human biomechanical signals such as limb kinematics, interaction force, and energy of the coupled human-robot system, we can eliminate the lengthy tuning process.

We formulate an optimization problem of finding the controller parameters with a goal to enhance the overall performance by improving the trade-off between stability, agility and reducing user effort (4):

$$\theta^{**} \in \arg\max_{\theta \in \mathbb{R}^d} f(\theta) \tag{4}$$

where $f(\cdot)$ is an objective function in respect to controller parameters $\theta \in \mathbb{R}^d$. As previously mentioned, the basic idea

of variable damping control is to have the robot inject or remove/dissipate energy in a way that the total energy of the system will be passive. Thus, our goal is to decrease the total energy (interaction energy) of the system to be slightly above zero. With this concept, we define an objective function f to minimize the squared energy (5):

$$f = -0.5(E_{interaction})^2 \tag{5}$$

where $E_{interaction}$ is $force * velocity * time$.

The interaction energy ($E_{interaction}$) used to define the objective function depends on user responses that are inherently stochastic due to the uncertainty of human behavior and noisy observations which makes the objective function unknown. Optimizing this objective function analytically is not feasible because there is no direct relationship between the controller parameters and the objective function. Thus, common optimization methods that work based on the derivation of objective function do not work in this problem. Furthermore, the stochasticity of our problem necessitates the need for more robust methods to the effect of errors. The number of local optima or the convexity of the objective function is not determined, while we are seeking the global maximum of the objective function. To overcome these challenges and difficulties, we approach the problem using Bayesian optimization (BO), which is a numerical global optimization method that is particularly well suited for optimizing unknown, black-box objective functions that are expensive to evaluate [18].

B. Bayesian Optimization (BO) of variable damping control using Gaussian Process (GP)

The BO method works using a response surface (i.e., surrogate model). Response surface-based optimization methods iteratively create a data set of parameters and the related evaluation function as $D = \{\theta, f(\theta)\}$. This data set is used to map the parameters to corresponding evaluation function through building a model (response surface), $\hat{f}(\cdot): \theta \rightarrow f(\theta)$. Using this response surface, the optimization (4) is replaced with a "virtual" optimization process as in (6), where "virtual" is the indication that the optimization problem only requires the evaluation of the learned model ($\hat{f}$), not the true objective function (f) (6):

$$\theta^{**} \in \arg\max_{\theta \in \mathbb{R}^d} \hat{f}(\theta) \tag{6}$$

A Gaussian process (GP) is used as response surface [19]. This probabilistic model allows us to model uncertainty of human behavior and noisy observations. Using probabilistic model as response surface $\hat{f}(\cdot)$ in (6) leads to multi-objective optimization problem [19, 21]. Therefore, an acquisition function $\alpha(\cdot)$ is used for the virtual optimization of this probabilistic model. This acquisition function allows us to scalarize the response surface onto a single function that can be optimized as in (7):

$$\theta^{*} \in \arg\max_{\theta \in \mathbb{R}^d} \alpha(\theta) \tag{7}$$

The parameters to optimize using (4) are b_{LB}, b_{UB}, k_p and k_n. Among these parameters, b_{LB} and b_{UB} are the most important since they define the boundaries of transition between positive and negative damping which has a direct impact on stability, agility, and user effort. Furthermore, depending on the inherent impedance properties of each human user, these parameters can vary significantly across different users [11]. k_p and k_n determine the growth rate of the transition between b_{LB} and b_{UB}. While these are important parameters, they have a much smaller effect on performance and previous studies showed that these two parameters do not vary significantly across different human users [9].

This difference in parameter significance informed our method to include a combination of BO and a real-time tuning method so that b_{LB} and b_{UB} are found directly using the BO and k_p and k_n are found indirectly from the BO method. We defined two data sets as $\theta = [b_{LB}, b_{UB}]$, $D = \{\theta, f(\theta,\beta)\}$ and $\beta = [k_p, k_n], D' = \{\beta, f(\theta,\beta)\}$.

Based on the description of these data sets, the response surface model of GP regression maps becomes $\hat{f}(\theta): \theta \rightarrow f(\theta,\beta)$. A GP is a distribution over functions $f \sim GP(m_f, k_f)$ defined by a prior mean m_f and covariance function k_f. To consider noisy observations, noisy function values are assumed $y = f(\theta,\beta) + \epsilon$, where $\epsilon \sim N(0, \sigma_\epsilon^2)$ is Gaussian noise. As it is conventionally chosen, $m_f \equiv 0$ as prior mean, while the chosen covariance function k_f is the squared exponential described in (8):

$$k_f(\theta_p, \theta_q) = \sigma_f^2 \exp\left(-\frac{1}{2}(\theta_p - \theta_q)^T \Lambda^{-1}(\theta_p - \theta_q)\right) + \sigma_\epsilon^2 \delta_{pq} \tag{8}$$

where $\Lambda = \mathrm{diag}([l_1^2, \dots, l_D^2])$ and $\sigma_\epsilon^2 \delta_{pq}$ is the representation of the white noise kernel in which σ_ϵ^2 is applied only when $p = q$. l_i are the characteristic length-scales, σ_f^2 is the variance of the latent function $f(\cdot)$ and σ_ϵ^2 is the noise variance.

Given training input (prior samples) $X = [\theta_1, \dots, \theta_n]$ and the corresponding training outputs $y = [y_1, \dots, y_n]$, we can define the GP predictive distribution with data set $D = \{X, y\}$ as in (9):

$$p(f(\theta|D,\theta) = N(\mu(\theta), \sigma^2(\theta)) \tag{9}$$

where the mean $\mu(\theta)$ and the variance $\sigma^2(\theta)$ are (10):

$$\mu(\theta) = k_*^T K^{-1} y, \qquad \sigma^2(\theta) = k_{**} - k_*^T K^{-1} k_* \tag{10}$$

respectively, and K is the matrix with $K_{ij} = k(\theta_i, \theta_j), k_{**} = k(\theta,\theta)$ and $k_* = k(X,\theta)$. The selection of hyperparameters of GP model including l_i, σ_f^2 and σ_ϵ^2 is important and they are selected by optimizing the marginal likelihood [19].

Expected improvement (EI) is used as acquisition function because experimental results showed that this function performed better on average than other functions [22, 23]. The EI acquisition function is (11-12):

$$\alpha(\theta) = \begin{cases} (\mu(\theta) - \mu^+ - \zeta)\phi(Z) + \sigma(\theta)\varphi(Z) & if\ \sigma(\theta) > 0 \\ 0 & if\ \sigma(\theta) = 0 \end{cases} \tag{11}$$

Algorithm 1 Bayesian Optimization of Variable Damping Control

1: $D \leftarrow$ if available: $\{\theta, f(\theta,\beta)\}$
2: $D' \leftarrow$ if available: $\{\beta, f(\theta,\beta)\}$
3: Prior $\leftarrow$ if available: Prior of the response surface
4: **While** optimize **do**
5: Train a response surface from D
6: Find θ^* that maximizes the acquisition surface $\alpha(\theta)$
7: Find β^* corresponding to the specific portion of maximum D
8: Evaluate $f(\theta^*,\beta^*)$ on the real system
9: Add $\{\theta^*, f(\theta^*,\beta^*)\}$ to D
10: Add $\{\beta^*, f(\theta^*,\beta^*)\}$ to D'
11: **end While**

$$Z = \begin{cases} (\mu(\theta) - \mu^+ - \zeta)/\sigma(\theta) & if\ \sigma(\theta) > 0 \\ 0 & if\ \sigma(\theta) = 0 \end{cases} \quad (12)$$

$\phi(\cdot)$ is the normal cumulative distribution function, $\varphi(\cdot)$ is the standard normal probability density function, and μ^+ is the best observed value we evaluated so far. ζ determines the amount of exploration during optimization and higher ζ values lead to more exploration than exploitation, which was set to $\zeta = 0.25$ in this study. With this acquisition function, our optimization problem fits the form of (7). In order to numerically find the parameters corresponding to the global maximum of our acquisition function (θ^*), L-BFGS algorithm is used [24]. It is initialized with 50 different random parameter sets and the maximum of the acquisition function is calculated for each set. The maximum is then determined from these 50 maxima.

The parameters k_p and k_n are calculated in real-time according to (3) after each iteration of BO. To account for high variability in human behavior, the update considers best values of the evaluation function in the previous iterations. Since the BO algorithm gradually converges, the percentage of previous iterations to be considered is also determined to gradually decrease with the number of iterations following $100 * (\text{number of iteration})^{-0.3}$. k_p and k_n for the next iteration are updated by averaging the best values within the selected iterations. With this method, the next β is calculated as β^*, then β^* and θ^* are used to find corresponding f, and they are added to D and D'.

The process of the whole algorithm is summarized in Algorithm 1. As described, D and D' are composed as prior. Then the response surface using the GP and the data set D is built, then according to the built model, $\alpha(\theta)$ is determined. Maximization of this function gives θ^*, and then β^* is found and then $f(\theta^*,\beta^*)$ is evaluated in the real system through pHRI. Then, D and D' will be updated according to new data along with the response surface. This process is repeated until the global maximum is determined numerically. A stopping criterion for our study is described in the next section.

The method is described for a single dimension of movement, but it is worth to note that it can be easily extended to multi-dimensional movement. This is important since nearly all real-world tasks require multi-dimensional movement. For multi-dimensional movement, we can decouple each direction of movement and implement the same process for each direction while modulating the damping value as described in (13):

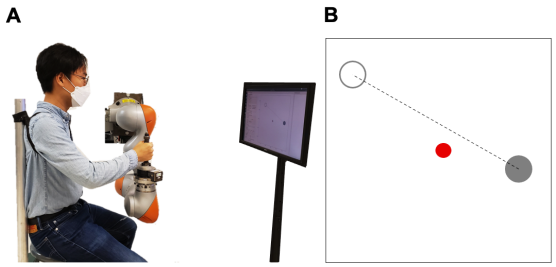

Fig. 1. Experimental setup. **A:** Side view of the human user interacting with the end-effector of a 7-DOF robotic arm and visual feedback display was provided in ~1m distance. **B:** Visual feedback display. Gray solid circle shows the current target, gray hollow circle shows the previous target. The dashed line represented the straightest path between the previous target and the current. The red solid circle presented the current hand position.

$$\boldsymbol{B_r} = B_i\hat{\boldsymbol{i}} + B_j\hat{\boldsymbol{j}} + B_k\hat{\boldsymbol{k}} + \cdots \quad (13)$$

where B_i, B_j, B_k are the robotic damping value modulated in each direction. The number of parameters will increase proportionally to the dimensions considered. For example, implementation of this method in 2 dimensions of movement increases the total number of parameters to 8.

C. Experiments

We performed a human experiment to validate the effectiveness of the presented control approach and investigate its potential applications in real-world scenarios. A 7 degree-of-freedom (DOF) robotic arm (LBR iiwa R820, KUKA, Germany) with a 6-axis load cell (Delta IP60, ATI Industrial Automation, NC) was used as the robotic interface. Both kinematic and force data were recorded at 1 kHz and low-pass filtered using a 4th order Butterworth filter with a cutoff frequency of 20 Hz.

Human users were instructed to perform a target reaching task in the transverse plane while they interacted with the end-effector of robotic arm in a seated position with their trunk securely strapped to a rigid chair to eliminate any effects of the confounding factors due to trunk movement during reaching movement (Fig. 1A) [25, 26]. A visual feedback display was provided at ~1 m to help users in completing target reaching movement tasks (Fig. 1B). The trial started when the new target was shown to the subject and lasted until 2 seconds after the subject first came within 0.5 cm of the target. Once a trial concluded, a new trial started at a randomized interval between 0.5-1.5 seconds. These experiments required movement in both the anterior-posterior (AP) and medial-lateral (ML) directions. The stiffness of the end effector was set to 0 N/m in these directions, and 10^6 N/m in the perpendicular direction to limit the movement of the robot to the transverse plane and prevent the movement in the direction of gravity. The simulated inertia was set to 10 kg. A virtual wall of $36 \times 36\ cm^2$ was implemented around the workspace to ensure the safety of the subjects.

There were multiple blocks of 12 trials. The experiment was divided into blocks to provide subjects a rest period

This article has been accepted for publication in IEEE Robotics and Automation Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LRA.2022.3144511

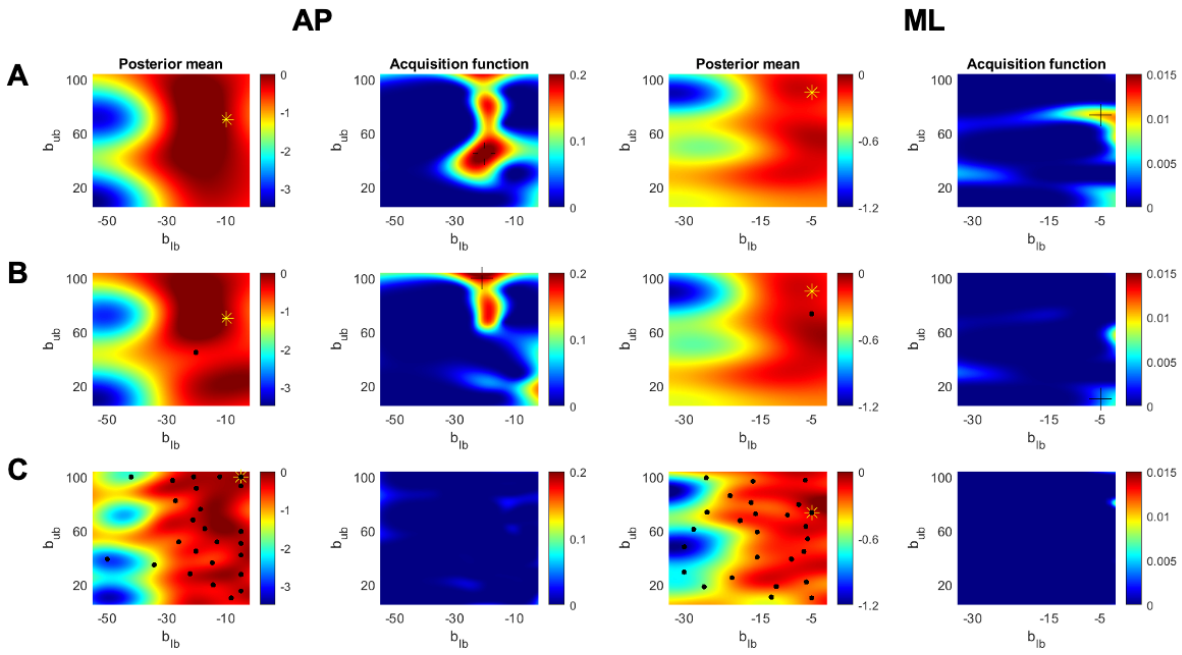

Fig. 2. Bayesian optimization process during the maximization of objective function *f* for a representative subject. There are 2D view of three sample iterations for both the AP and ML directions **A:** Initial iteration (first row), **B:** Second iteration (second row), **C:** Last iteration (third row) of the Bayesian optimization process including the mean posterior of the model prediction, sample posterior points (black dots), expected improvement function (acquisition function). The location of the next parameter to be evaluated (+) and the best selected parameters up to that iteration (yellow *) are shown. Samples are distributed more on the promising areas that are more probable to have the maximum points (red areas). There are some samples out of those promising areas as well (blue or yellow areas), because the acquisition function handles the trade-off between exploration and exploitation in the BO process to search for the global optimum point.

between them to prevent potential fatigue. Each iteration for BO included 4 target reaching trials and the parameters in each iteration remained constant. Therefore, the evaluation of the objective function was done in each iteration based on the average of the 4 trials in that iteration. For these 4 trials, 4 targets were randomly generated in the traverse plane ($20 \times 20\ cm^2$) with a constraint that the total path length in each of the ML and AP ($+\hat{\boldsymbol{\imath}}$, $-\hat{\boldsymbol{\imath}}$, $+\hat{\boldsymbol{\jmath}}$, and $-\hat{\boldsymbol{\jmath}}$) was 20 cm and the minimum length of each path in each direction was 5 cm.

The ranges of parameters b_{LB}^{AP}, b_{UB}^{AP}, b_{LB}^{ML} and b_{UB}^{ML} considered in BO were $[-50, -5]$, $[10, 100]$, $[-30, -5]$, $[10, 100]$ Ns/m, respectively, which were selected based on our previous studies [10, 11]. Thirteen prior samples were evaluated at the beginning of the optimization process. These 13 samples were chosen randomly from a grid made from the ranges of parameters. While k_p^{AP}, k_n^{AP}, k_p^{ML} and k_n^{ML} can be initialized with any value, they were all initialized with 25 according to approximate average values across different subjects from the previous study [9].

The stopping criterion for the optimization process was when the best parameters did not change for 10 iterations of BO consecutively. Since b_{LB}^{AP}, b_{UB}^{AP}, b_{LB}^{ML}, and b_{UB}^{ML} were the parameters found directly based on BO, they were used for the stopping criterion. We might see changes in k_p^{AP}, k_n^{AP}, k_p^{ML}, and k_n^{ML} parameters even in the last iterations of the algorithm, because they were updated at every iteration. Therefore, convergence was only expected for b_{LB}^{AP}, b_{UB}^{AP}, b_{LB}^{ML} and b_{UB}^{ML} parameters that were directly determined from BO.

Five young, healthy subjects (age: 21–34, height: 163–183 cm, weight: 50–78 kg, sex: 4 males and 1 female) participated in this study, which was approved by the Institutional Review Board of Arizona State University (STUDY 00010123). Subjects provided informed, written consent prior to participation. All experimental procedures were performed in accordance with the relevant guidelines and regulations. No subject was informed of the hypotheses of this study.

D. Data Analysis

Several performance metrics representing stability, agility, and user effort were selected to compare the presented controller with the previous variable damping controller without user-adaptive parameter changes [9].

1) Stability: Stability was evaluated in both spatial and time domains. In the spatial domain, overshoot was evaluated by calculating the maximum distance past the target position. In the time domain, stability time was defined as the time between the first time the subject hit the target and when the subject was able to hold the position within the target ($\pm 5\ mm$) for 0.5 *s* continuously.

2) Agility: Agility was evaluated using the maximum and mean speed of the subject's movement. The mean speed was calculated as the average speed from the initiation time, the first time when the subject started moving (move $> 5\ mm$), to the first time to hit the target (within $\pm 5\ mm$ of the center of the target). The maximum speed was the largest magnitude of speed that the subject had during the trial.

3) User Effort: User effort was evaluated using force at the interaction point. The mean root-mean-squared (RMS) and maximum RMS interaction forces were used to quantify the user effort. The mean RMS interaction force was calculated from the initiation time to the stability time, while

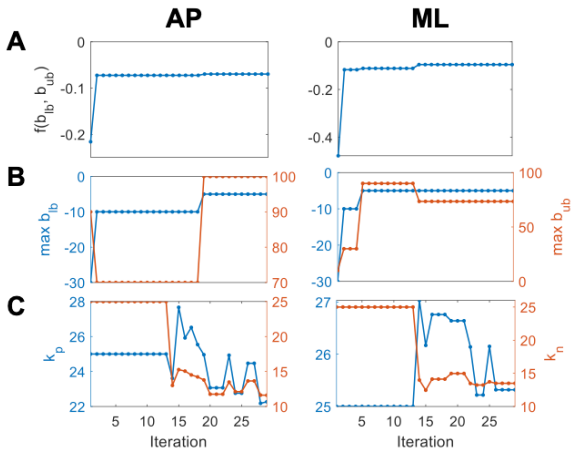

Fig. 3. The Maximum objective function value up to the current iteration and the best selected parameters over iterations in both the AP and ML directions. **A:** Maximum value of the objective function, **B:** Best selected b_{LB} and b_{UB}, **C:** Best selected k_p and k_n with respect to iterations.

the maximum RMS force was defined as the highest value of interaction force during the trial.

All these performance metrics were calculated in trials with the optimized parameters found from our presented controller and with the fixed parameters used in the previous variable damping controller. The percentages of improvement for each metric were calculated as shown in (14), except for agility which was the negative of (14), because we aimed to increase agility.

$$\%\, improvement = \frac{(metrics_i^{fix} - metrics_i^{optimized})}{mean(metrics_i^{fix}, metrics_i^{optimized})} \quad (14)$$

where $mean(metrics_i^{fix}, metrics_i^{optimized})$ denotes the average of $metrics_i$ in the fixed parameter condition and in the parameter condition with BO.

III. Results

Results of human experiments demonstrate how fast and successful the presented BO algorithm was in finding the optimal parameters of the variable damping controller despite unknown objective function, uncertainty of human behavior, and noisy observations. Quantitative results of one representative subject and detailed results of all subjects are described in this section.

The representative subject's response surface model and acquisition function in both the AP and ML directions in the initial, second, and last iterations of the BO process are shown (Fig. 2) to demonstrate how the response surface evolved by the addition of new samples from the acquisition function. In the initial iteration (Fig. 2A), the response surface was built based on the available prior samples. The corresponding acquisition function of this model was calculated, and its maximum value provided the information for the next parameters to be evaluated (as denoted with (+) in Fig. 2). The next parameters selected from the acquisition function in the initial iteration was added to the data set for the next iteration (second iteration) and the response surface was updated accordingly (Fig. 2B). Based on the added new samples, the response surface was updated iteratively until the optimization process converged. The last iteration (Fig. 2C) shows the updated response surface with all posterior samples. The value of acquisition function with respect to all parameters was similar (similar color) in the last iteration of the optimization algorithm.

The evolution of the maximum value of the objective function and the best corresponding parameters are shown in Fig. 3. The best parameters for this representative subject were found in 29 iterations (13 prior iterations and 16 BO iterations), demonstrating the fast convergence of the presented optimization process. It also shows how the best parameter values evolved over the course of the optimization process. The timings for changes of b_{LB} and b_{UB} matched those of the objective function as they were directly found from the objective function. On the contrary, k_p and k_n showed a different pattern of evolution since they were updated every iteration.

The optimal parameters and the number of iterations until convergence for all subjects are summarized in Table I. The optimal value for b_{lb}^{AP} and b_{lb}^{ML} ranged -20.5 to -5.0 Ns/m and -10.0 to -5.0 Ns/m, respectively, and that for b_{ub}^{AP} and b_{ub}^{ML} ranged 39.9 to 100.0 Ns/m and 60 to 90 Ns/m, respectively. In addition, the optimal value for k_p^{AP} and k_p^{ML} varied across different subjects from 22.8 to 55.7 and from 24.0 to 57.9, respectively, and that for k_n^{AP} and k_n^{ML} changed from 12.5 to 42.2 and from 13.7 to 46.1, respectively. A clear difference in the optimal parameters across different subjects emphasizes the importance of designing the variable damping controller in a user-specific manner by considering user proficiency and biomechanical characteristics during pHRI. While the optimal parameters

TABLE I. Optimal Parameters and the Number of Iterations Until Convergence

Optimal Parameters	b_{lb}^{AP}	b_{ub}^{AP}	b_{lb}^{ML}	b_{ub}^{ML}	k_p^{AP}	k_n^{AP}	k_p^{ML}	k_n^{ML}	Number of Iterations until Convergence
Subject1	-10.0	70.0	-5.0	90.0	36.1	19.8	39.0	22.6	15
Subject2	-5.0	100.0	-5.0	73.5	22.8	12.5	24.0	13.7	16
Subject3	-20.5	100.0	-10.0	60.0	33.4	20.6	30.2	16.4	13
Subject4	-9.7	39.9	-10.0	90.0	55.7	42.2	57.9	46.1	16
Subject5	-8.2	72.0	-10.0	90.0	45.3	28.0	47.3	34.8	12
Mean (std.)	-10.7 (5.2)	76.4 (22.4)	-8.0 (2.4)	80.7 (12.2)	38.7 (11.1)	24.6 (10.1)	39.7 (12.0)	26.6 (12.0)	14 (2)

TABLE II. PERCENTAGE IMPROVEMENT OF THE PRESENTED USER-ADAPTIVE CONTROLLER WITH RESPECT TO THE PREVIOUS VARIABLE DAMPING CONTROLLER WITH FIXED PARAMETERS [9]

Percentage Improvement	Overshoot	Stability time	Mean speed	Max speed	Mean force	Max force	Overall improvement	Energy
Subject1	100.8	-15.7	-11.8	-24.1	35.6	43.7	21.4	68.1
Subject2	86.1	43.4	-18.2	-21.7	31.2	32.6	25.6	57.1
Subject3	73.2	61.6	-10.9	-1.7	11.7	4.8	23.1	23.6
Subject4	55.2	12.9	-13.8	-20.9	12.7	4.6	8.4	43.9
Subject5	80.6	25.8	0.8	-9.0	5.1	12.5	19.3	31.6
Mean (std.)	79.2 (15.0)	25.6 (26.4)	-10.8 (6.3)	-15.5 (8.6)	19.2 (11.9)	19.6 (15.8)	19.6 (5.9)	44.9 (16.3)

were highly variable across different subjects, the number of iterations until the algorithm converged showed much smaller variation with a range of 12 to 16. This demonstrates the consistent speed of our algorithm in determining the optimal parameters for different users.

The presented user-adaptive damping controller showed considerable overall performance improvement compared to the previous variable damping controller without user-adaptive parameter changes. Percentage improvement for different performance metrics (stability, agility, and user effort) for all subjects are summarized in Table II. Stability through both metrics (overshoot and stability time) considerably improved in most of the subjects. When averaged across all subjects, the improvement was 79.2% and 25.6% in overshoot and stability time, respectively. User effort also improved consistently in all subjects according to both mean RMS force and max RMS force metrics. There were average of 19.2% and 19.6% improvements across subjects for mean RMS force and max RMS force, respectively. An average reduction in the agility was observed compared to the previous variable damping controller: -10.8% and -15.5% for mean and max speed, respectively. This reduction is mainly because the presented user-adaptive controller determined the optimal parameters by considering all performance metrics simultaneously although indirectly via the energy function. The overall performance, calculated by averaging results of all three performance metrics, showed consistent improvement in all subjects. When averaged across all subjects, the overall performance improvement was 19.6%.

As expected from the use of energy in the objective function (Eq. (4)), the presented controller reduced energy of the coupled system considerably. On average, an energy reduction of 44.9% was observed compared to the previous controller. It is important to note that while we did not directly use performance metrics (stability, agility, and user effort) in the optimization process, targeting the reduction of interaction energy in the optimization process led to substantial overall performance improvement in pHRI.

IV. DISCUSSION

This paper presented a user-adaptive variable damping controller that can be applied in a diverse set of applications to enhance the overall performance of coupled human-robot systems. Based on the framework of BO with GP, the presented controller minimizes the energy of the coupled human-robot system without violating the passivity constraint. The algorithm incorporates the inherent impedance properties of each human user's limbs to initiate the optimization procedure and adaptively changes the controller parameters according to the user's proficiency in physical interaction measured by human biomechanical signals including limb kinematics, interaction force and energy.

The presented approach overcomes many drawbacks and limitations of previous approaches, by effectively avoiding local optima via the acquisition function that addresses the trade-off between exploration of the search space and exploitation of promising areas, explicitly modeling noisy observations and uncertainties of the human user response, and efficiently and effectively selecting parameters without a need of long experiments and tuning sessions. This approach could successfully determine the controller parameters on an individual basis. With an optimal set of controller parameters that minimizes interaction energy, the controller could substantially enhance the overall performance in pHRI by improving the trade-off between stability and agility and reducing user effort.

Experimental results from 5 pilot subjects interacting with a popular robotic arm manipulator confirmed the effectiveness of the presented user-adaptive controller in enhancing the overall performance of coupled human-robot systems beyond that of the previous variable damping controller with fixed controller parameters [9].

Compared to the previous controller, the presented adaptive control strategy reduced about 45% energy expenditure and achieve average performance improvement of about 20% when several performance metrics of stability, agility, and user effort are considered together. All subjects participated in this pilot experiments consistently showed a clear prioritization on stability and user effort over agility to enhance the overall performance during interaction with the robotic arm. Specifically, all subjects showed notable improvements in terms of stability and user effort: when averaged across subjects, stability and user effort metrics showed improvements of 52.4% and 19.4%, respectively. However, arm movements were consistently slowed down and agility metrics showed about 10% and 15% of reduction in mean and max speeds, respectively.

These results are somewhat expected as performance metrics were not directly used in the optimization process, but energy expenditure was used. In fact, since there exist clear trade-offs between these performance metrics (stability, agility, and user effort), using them in the optimization process would not lead to an improvement in every performance metric.

It is worth to note that the presented controller can still improve agility compared to popular positive damping controllers. According to the previous study [9], the variable damping controller without user-adaptive parameter changes improved the mean and max speeds by 19.4% and 56.1%, respectively. Thus, the presented controller in this paper is still expected to achieve a better agility performance than the positively damped controllers.

The previous controller included two separate lengthy tuning sessions before the main experiments (one with 126 trials (~15 sec for each trial) to determine b_{LB} and b_{UB} and the other with 60 trials (~15 sec for each trial) to determine k_p and k_n), while the uncertainty of human behavior and noise of observation such as limb kinematics and interaction force, and user proficiency and adaptability were not considered in the tuning process. The current adaptive control strategy, on the contrary to the previous one, eliminated these extra tuning processes and the controller parameters were adaptively determined throughout the main experiments based on user proficiency. The whole experimental protocol took on average of 14 iterations of BO that includes 56 trials without any extra tuning process.

Although the human experiments with 5 pilot subjects in this study used simple 2D arm reaching tasks to validate the effectiveness of the presented controller, an additional study with a larger set of subjects are warranted in more complicated task conditions (e.g., 3D arm movement, irregular movement, and obstacle avoidance) to fully evaluate its potential applicability in real-world scenarios. In addition, other future work will incorporate variable stiffness into the existing controller framework to fully implement a user-adaptive variable impedance controller, which we expect to further improve the performance in pHRI beyond the variable damping controller.

REFERENCES

[1] E. Colgate, N. Hogan, "The interaction of robots with passive environments: application to force feedback control," in *Waldron K.J. (eds) Advanced Robotics: 1989*. Springer, Berlin, Heidelberg, 1989.

[2] N. Hogan, "Impedance control: an approach to manipulation," in *1984 American Control Conf.*, San Diego, CA, USA, pp. 304-313, 1984.

[3] C. Ott, A. Albu-Schaffer, A. Kugi and G. Hirzinger, "On the passivity-based impedance control of flexible joint robots," in *IEEE Trans. Robotics*, vol. 24, no. 2, pp. 416-429, Apr. 2008.

[4] S.P. Buerger, and N. Hogan, "Impedance and interaction control," in T.R. Kurfess (Ed.), *Robotics and Automation Handbook*, Boca Raton, FL: CRC Press, 2004, pp. 368-391.

[5] E. Gribovskaya, A. Kheddar, and A. Billard, "Motion learning and adaptive impedance for robot control during physical interaction with humans," *IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2011, pp. 4326–4332.

[6] C. T. Landi, F. Ferraguti, L. Sabattini, C. Secchi and C. Fantuzzi, "Admittance control parameter adaptation for physical human-robot interaction," *IEEE Int. Conf. on Robotics and Automation (ICRA)*, Singapore, 2017, pp. 2911-2916.

[7] V. Duchaine and C. M. Gosselin, "General model of human-robot cooperation using a novel velocity based variable impedance control," *2nd Joint Euro Haptics Conf. and Symp. on Haptic Interfaces for Virtual Environ. and Teleoperator Systems (WHC'07)*, Tsukaba, 2007, pp. 446-451.

[8] F. Dimeas and N. Aspragathos, "Reinforcement learning of variable admittance control for human robot co-manipulation," *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, Hamburg, 2015, pp. 1011-1016.

[9] F. Zahedi, J. Arnold, C. Phillips and H. Lee, "Variable damping control for pHRI: considering stability, agility, and human effort in controlling human interactive robots," in *IEEE Trans. on Human-Machine Systems*, vol. 51, no. 5, pp. 504-513, 2021.

[10] F. Zahedi, T. Bitz, C. Phillips, and H. Lee, "Regulation of 2D arm stability against unstable, damping-defined environments in physical human-robot interaction," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2020.

[11] F. Zahedi and H. Lee, "Human arm stability in relation to damping-defined mechanical environments in physical interaction with a robotic arm," in *Int. Conf. on Robotics and Automation* (ICRA), 2021.

[12] H. Lee, H. I. Krebs and N. Hogan, "Multivariable dynamic ankle mechanical impedance with active muscles," *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 5, pp. 971-981, Sept. 2014.

[13] W. Jackson and S. H. Collins, "Heuristic-based ankle exoskeleton control for co-adaptive assistance of human locomotion," *IEEE Trans-actions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 10, pp. 2059–2069, 2019.

[14] 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.

[15] P. Malcolm, D. M. Rossi, C. Siviy, S. Lee, B. T. Quinlivan, M. Grimmer, and C. J. Walsh, "Continuous sweep versus discrete step protocols for studying effects of wearable robot assistance magnitude," *Journal of Neuroengineering and Rehabilitation*, vol. 14, no. 1, pp. 1–13, 2017.

[16] W. Felt, J. C. Selinger, J. M. Donelan, and C. D. Remy, ""body-in-the-loop": Optimizing device parameters using measures of instantaneous energetic cost," *PloS one*, vol. 10, no. 8, p. e0135342, 2015.

[17] K. A. Witte, P. Fiers, A. L. Sheets-Singer, and S. H. Collins, "Improving the energy economy of human running with powered and unpowered ankle exoskeleton assistance," *Science Robotics*, vol. 5, no. 40, 2020.

[18] D. D. Cox and S. John, "Sdo: A statistical method for global optimization," in *Multidisciplinary Design Optimization: State-of-the-Art*, 1997, pp. 315–329.

[19] C. K. Williams and C. E. Rasmussen, Gaussian processes for machine learning. *MIT press Cambridge*, MA, 2006, vol. 2, no. 3.

[20] M. A. Osborne, R. Garnett, and S. J. Roberts, "Gaussian processes for global optimization," *in 3rd Int. Conf. on Learning and Intelligent Optimization (LION3)*, 2009, pp. 1–15.

[21] D. R. Jones, "A taxonomy of global optimization methods based on response surfaces," *Journal of global optimization*, vol. 21, no. 4, pp.345–383, 2001.

[22] Mockus, V. Tiesis, and A. Zilinskas, "The application of Bayesian methods for seeking the extremum," *Towards Global Optimization*, vol. 2, no. 117-129, p. 2, 1978.

[23] P. Hennig and C. J. Schuler, "Entropy search for information-efficient global optimization." *Journal of Machine Learning Research*, vol. 13, no. 6, 2012.

[24] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization," *SIAM Journal on scientific computing*, vol. 16, no. 5, pp. 1190–1208, 1995.

[25] E. J. Perreault, R. F. Kirsch, and P. E. Crago, "Effects of voluntary force generation on the elastic components of endpoint stiffness," *Experimental Brain Research*, vol. 141, no. 3, pp. 312–323, 2001.

[26] E. Burdet, R. Osu, D. W. Franklin, T. E. Milner, and M. Kawato, "The central nervous system stabilizes unstable dynamics by learning optimal impedance," *Nature*, vol. 414, no. 6862, pp. 446–449, 2001.