

Using Bayesian Optimization to Identify Optimal Exoskeleton Parameters Targeting Propulsion Mechanics: A Simulation Study

GilHwan Kim and Fabrizio Sergi

Abstract—In this study, we determined the feasibility of modeling the relationship between robot control parameters and propulsion mechanics as a Gaussian process. Specifically, we used data obtained in a previous experiment that used pulses of torque applied at the hip and knee joint, at early and late stance, to establish the relationship a 3D control parameter space and the resulting changes in hip extension and propulsive impulse. We estimated Gaussian models both at the group level and for each subject. Moreover, we used the estimated subject-specific models to simulate virtual human-in-the-loop optimization (HIL) experiments based on Bayesian optimization to establish their convergence under multiple combinations of acquisition functions and seed point selection methods.

Results of the group-level model are in agreement with those obtained with linear mixed effect model, thus establishing the feasibility of Gaussian process modeling. The estimated subject-specific optimal conditions have large between-subject variability in the metric of propulsive impulse, with only 31% of subjects featuring a subject-specific optimal point in the surrounding of the group-level optimal point. Virtual HIL experiments indicate that expected improvement is the most effective acquisition method, while no significant effect of seed point selection method was observed. Our study may have practical effects on the adoption of HIL robot-assisted training methods focused on propulsion.

I. INTRODUCTION

Robot assisted gait training is becoming a common method for rehabilitation after neurological injury [1]. With the option of mechanical assistance to multiple joints, and several open parameters for the timing of assistance at each joint, controllers for gait training robots are defined by a large number of open parameters, each corresponding to highly variable outcomes of robotic training. To deal with the large number of open parameters in robot-assisted gait training, real-time optimization methods, or Human-In-the-Loop (HIL) optimization methods, have been introduced [2]. In human-robot interaction, HIL methods are used to identify parameters of a robot controller that are optimal in the sense of a specific cost function. In general, this is done by testing the value of the cost function, collecting data to quantify the subject response at those points, and then iteratively updating the controller parameters in real-time to optimize the cost function.

Several successful implementations of HIL optimization exist, many using different optimization algorithms [3]-[4], with one-dimensional (1D) gradient descent method [5], 4D

Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [4], and Bayesian optimization [3], [6] used in recent studies using exoskeletons supporting walking function. However, most previous HIL optimization experiments focused on reducing metabolic cost to tune control parameters which results limited applicability in robot-assisted gait training, since the major objective of gait training is to induce desired changes in subjects' motor coordination with the ultimate goal of improving their motor function. Moreover, on-going research about optimization algorithm is focused on improving performance of Bayesian optimization based on hyperparameter tuning [7], noise modeling [8], acquisition function definition [9], [10], [11], and definition of seed points [3], but it is currently unclear how these methods apply to HIL optimization methods in biomechanics.

In this work, we seek to apply HIL to establish the subject-specific relationship between exoskeleton control parameters and resulting features of gait, such as those describing propulsion mechanics, a crucial component of walking function [12], [13], [14]. We address the feasibility of modeling previously collected data on the effects of the application of pulses of torque to the hip and knee joint on propulsion mechanics using Gaussian process modeling [15]. Specifically, we establish the relationship between pulse torque conditions and propulsion mechanics at the group and subject-specific level, and evaluate how these relationships differ across subjects. Moreover, we run simulations using the estimated subject-specific Gaussian process models to establish convergence in virtual HIL experiments based on Bayesian optimization. Specifically, in our analysis we establish the effect of two important factors, seed point selection method and acquisition function, in the speed and accuracy of convergence of Bayesian optimization targeting propulsion mechanics.

II. METHODS

A. Data Collection

In previous research, sixteen healthy subjects participated in an experiment based on the application of pulses of torque to the hip and knee joint to modulate propulsion mechanics [15]. Torque conditions were defined as a combination of three parameters: pulse timing, hip pulse amplitude, and knee pulse amplitude. Two levels were used for pulse timing: pulses were applied at 10% of the estimated gait cycle period (early stance), or 45% of the estimated gait cycle period (late stance). Levels for hip and knee pulse amplitude were defined as either zero torque, flexion, or extension (amplitude was set to 15 N·m for the hip joint, and 10 N·m for the

G. Kim is with the Department of Mechanical Engineering, University of Delaware, Newark, DE 19716, USA; secretgh@udel.edu

F. Sergi is with the Department of Biomedical Engineering, University of Delaware, Newark DE, 19713, USA; fabs@udel.edu

*This work is supported in part by NSF-CMMI-1934650 and in part by NSF-CBET-1638007.

knee joint for both flexion and extension). Sixteen conditions were tested, including all combinations of the factors above, with the exclusion of the combination of zero knee and hip torque. When pulses were applied to both joints, they were applied simultaneously. Each condition was repeated ten times per subject, with a random sequence. Pulses were applied to the right leg during single strides, and spaced by at least eight strides of no pulse application. Hip extension (HE) and propulsive impulse (PI) were assessed at multiple strides: prior to pulse application (stride -1), during pulse application (stride 0), and during the three strides following pulse application (stride 1, 2, 3). In previous work, our group used a linear mixed model to establish the relationship between factors including pulse pattern parameters (pulse timing, amount of torque pulses applied to hip and knee joints), subjects, and stride with propulsion mechanics, as defined by HE and PI [15].

B. Gaussian Process Modeling

Our first goal is to model the relationship between pulse torque conditions and propulsion mechanics as a Gaussian process. Each subject has 10 measurements of each outcome measure (HE and PI), resulting in a total of 160 data points per stride condition, and thus 800 data points per subject (12800 total measurements), referred to as measurements y_{HE} and y_{PI} . These measurements can be indexed as a function of factors: pulse timing (T), amplitude of hip and knee torque pulses (K and H), stride (S), subject (Sbj), trial index (Rep), as

$$y_{HE} = y_{HE}(T, K, H, S, Sbj, Rep), \quad (1)$$

$$y_{PI} = y_{PI}(T, K, H, S, Sbj, Rep), \quad (2)$$

where $T \in (10, 45)$, $K \in (-10, 0, 10)$, $H \in (-15, 0, 15)$, $S \in (-1, \dots, 3)$, $Sbj \in (1, \dots, 16)$, and $Rep \in (1, \dots, 10)$.

We assumed that the Gaussian process model linking control parameters with the outcome measures of propulsion should follow the characteristics listed below:

- noise from outcome measures is normally distributed with zero mean ($\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$);
- noise is independent of the human response;
- the distribution of the human response under repeated exposure to the same pulse condition is normal;
- the variance of the human response will be constant under different pulse conditions.

1) *Group-level Model Formulation:* Data in variables (1) and (2) are concatenated along the subject dimension obtaining variables $\tilde{y}_{HE}(T, K, H, S, Rep_G)$ for HE, and $\tilde{y}_{PI}(T, K, H, S, Rep_G)$ for PI, with 160 repetitions available for factor Rep_G .

The modified outcome measures (\tilde{y}_{HE} and \tilde{y}_{PI}) can be expressed as the sum of an unknown Gaussian process G , and noise (ϵ) as follows:

$$\tilde{y}_{HE}(T, K, H, S, Rep_G) = G_{HE}(T, K, H, S) + \epsilon_{HE}, \quad (3)$$

$$\tilde{y}_{PI}(T, K, H, S, Rep_G) = G_{PI}(T, K, H, S) + \epsilon_{PI}. \quad (4)$$

The average in the measurements \tilde{y} can be derived by average value among 160 repetitions per each condition. Since noise is assumed to have zero mean, the average value of the unknown Gaussian process G is equal to the average in the measurements \tilde{y} . The variance in the measurements \tilde{y} are expressed as a sum of model variance (variations in the true output arising from repeated exposure to the same conditions σ_G^2), and measurement error (variations in the measurements that are not associated to changes in the true value of the output σ_ϵ^2). Since the noise is constant for all values of pulse torque factors, the following relationships hold true:

$$\sigma_{\tilde{y}_{HE}}^2 = \sigma_{G_{HE}}^2 + \sigma_{\epsilon_{HE}}^2, \quad (5)$$

$$\sigma_{\tilde{y}_{PI}}^2 = \sigma_{G_{PI}}^2 + \sigma_{\epsilon_{PI}}^2. \quad (6)$$

These relationships are used to specify values for the model variance σ_G^2 used for the Gaussian process covariance function. Values of $\sigma_{\tilde{y}_{HE}}^2$ and $\sigma_{\tilde{y}_{PI}}^2$ are calculated as the maximum variance of $\tilde{y}_{HE}(T, K, H, S, Rep_G)$ along the Rep_G dimension, i.e. the one resulting from the combination of pulse parameters associated with the largest variance across subjects and repetitions.

2) *Group-level Model Estimation:* A Gaussian process model is estimated from eq. (3) and (4) to approximate the mean and variance of data \tilde{y}_α . In a Gaussian process, the variance is defined based on a covariance (or kernel) function $k_\alpha(x_i, x_j)$, which defines how variance propagates within a dimension, or in different dimensions of the model. In this work, we defined the covariance function (automatic relevance determination squared exponential kernel) as:

$$k_\alpha(x_i, x_j) = (\sigma_{\tilde{y}_\alpha}^2 - \sigma_{\epsilon_\alpha}^2) \exp \left[- \sum_{m=1}^4 \frac{\tilde{x}_m^\top \tilde{x}_m}{2 * l_{m,\alpha}^2} \right], \quad (7)$$

where x is set of pulse torque input parameters, $x = (T, K, H, S)$, $\tilde{x}_m = x_{im} - x_{jm}$, $l_{m,\alpha}$ is the m -th length scale hyperparameter, $\alpha \in (HE, PI)$. (i, j) are indices corresponding to two arbitrary points in the 4D domain of the parameters. Using the covariance function (7), the estimated mean value and variance of a point x_{n+1} , based on n measurements, are

$$\mu_\alpha(x_{n+1}) = p_\alpha^\top (K_\alpha + \sigma_{\epsilon_\alpha}^2 I)^{-1} Y_{\alpha,1:n}, \quad (8)$$

$$\sigma_\alpha^2(x_{n+1}) = k_\alpha(x_{n+1}, x_{n+1}) - p_\alpha^\top (K_\alpha + \sigma_{\epsilon_\alpha}^2 I)^{-1} p, \quad (9)$$

$$p_\alpha = [k_\alpha(x_{n+1}, x_1) \quad k_\alpha(x_{n+1}, x_2) \quad \dots \quad k_\alpha(x_{n+1}, x_n)], \quad (10)$$

$$Y_{\alpha,1:n} = [y_{\alpha,1} \quad \dots \quad y_{\alpha,n}]^\top, \quad (11)$$

and

$$K_\alpha = \begin{bmatrix} k_\alpha(x_1, x_1) & \dots & k_\alpha(x_1, x_n) \\ \vdots & \ddots & \vdots \\ k_\alpha(x_n, x_1) & \dots & k_\alpha(x_n, x_n) \end{bmatrix}. \quad (12)$$

In our case, $n = 80$, referring to the fact that all observations ($Y_{\alpha,1:n}$) are used to estimate a Gaussian process model for G_{HE} and G_{PI} . After a model is estimated given the available n measurements, the estimated average value $\mu(x_{n+1})$ of all

points in the domain is defined as the mean function $m(x)$ of input x .

As discussed above, (3) and (4) define the relationship between control parameters and output using Gaussian processes, i.e., $G_{HE}(x) \sim \mathcal{GP}(m(x), k(x, x'))$, where $k(x, x')$ is the kernel function. As such, the Gaussian processes can be estimated by solving the following least-squares problems in terms of the hyperparameters $l_{m,\alpha}$ and $\sigma_{\epsilon_\alpha}^2$:

$$\text{Find } l_{m,\alpha}, \sigma_{\epsilon_\alpha}^2 \text{ s.t. } \min \sum_{i=1}^n (\tilde{y}_{\alpha,i} - G_\alpha(x_i))^2. \quad (13)$$

To align our fitted model with observations emerging from our previous linear mixed model analysis, we set constraints on the hyperparameters optimized in (13) (Table I).

3) *Subject-specific Model Formulation*: Similarly as the group data, the outcome measures for each subject ($y_{Sbj,HE}$ and $y_{Sbj,PI}$) can be expressed as the sum of an unknown subject-specific Gaussian process G_{Sbj} , and noise (ϵ_{Sbj}):

$$y_{Sbj,HE} = G_{Sbj,HE}(T, K, H, S) + \epsilon_{Sbj,HE}, \quad (14)$$

$$y_{Sbj,PI} = G_{Sbj,PI}(T, K, H, S) + \epsilon_{Sbj,PI}. \quad (15)$$

Measurements collected from each subject correspond to the ten repeated measurements, collected at all combinations of the factors, resulting in 800 measurements. Similarly as done in the group-level model, noise is assumed to be constant for all values of pulse torque factors, hence variances of data and processes are linked by:

$$\sigma_{y_{Sbj,HE}}^2 = \sigma_{G_{Sbj,HE}}^2 + \sigma_{\epsilon_{Sbj,HE}}^2, \quad (16)$$

$$\sigma_{y_{Sbj,PI}}^2 = \sigma_{G_{Sbj,PI}}^2 + \sigma_{\epsilon_{Sbj,PI}}^2. \quad (17)$$

4) *Subject-specific Model Estimation*: Similarly as done for group-level model estimation, a Gaussian process is estimated from (14) and (15). The same kernel function defined in (7) was used for the definition of the subject-specific Gaussian models, with the difference that $\sigma_{y_\alpha}^2$ was defined as the variance resulting from the combination of pulse parameters with the largest within-subject variance across the ten repetitions. Because it is usually impractical to collect a sufficient number of observations from an individual subject to properly estimate subject-specific length-scale hyperparameters, we proceeded to specify for individuals the same values of parameters $l_{m,\alpha}$ estimated for the group. In these conditions, the subject-specific Gaussian processes

TABLE I

LOWER AND UPPER BOUNDS FOR ESTIMATED NOISE STANDARD DEVIATION AND FOUR LENGTH SCALE HYPERPARAMETERS

	T [% gait cycle]	S [°]	K [N·m]	H [N·m]	σ_{ϵ_α}
LB	12	0	6	7	0
UB	40	1	15	20	$\max(\sigma_{\tilde{y}_\alpha})$

$G_{Sbj,\alpha}$ are estimated solving the following least-squares problem:

$$\text{Find } \sigma_{\epsilon_{Sbj,\alpha}}^2 \text{ s.t. } \min \sum_{i=1}^n (y_{Sbj,\alpha,i} - G_{Sbj,\alpha}(x_i))^2. \quad (18)$$

5) *Quantifying Variability of Maxima in Subject-specific Models*: To quantify the variability of the optimal points in the estimated subject-specific Gaussian process models $G_{Sbj,HE}$ and $G_{Sbj,PI}$, we calculated the vectors x_{Sbj}^* as the values of T , K , and H that maximized the estimated process value at stride 0 for different subjects, and δx_{Sbj}^* as the values of T , K , and H that maximized the estimated change in outcome measure between stride 0 and stride -1 for different subjects. We thus established for how many subjects (n_w) the subject-specific optimal points fell within a sphere of radius r centered around the optimal points estimated using the group-level model (x^* and δx^*). The optimal point analysis is conducted on the non-dimensional domain where all coordinates are comprised between 0 and 1 using a min-max normalization. The functions $n_w(r)$ are compared to establish variability of maxima in the two outcome measures.

C. Bayesian Optimization Simulation

Virtual HIL experiments are conducted to determine the convergence of Bayesian optimization for each subject-specific Gaussian process model, under different settings of the optimization algorithm. Each subject-specific Gaussian process model is assumed to be the human response (i.e., data generating process) when pulse torque assistance is applied. Three robotic control parameters (T , K , and H) are used as input parameters. In all conditions, the optimization algorithm does not have any knowledge about the subject-specific model at the beginning of each optimization. Two algorithm components are tested using a factorial design: *i*) acquisition function (three levels), and *ii*) and seed point selection method (three levels) are evaluated in terms of the number of iterations required for convergence. In total, the 9 combinations of the two components are tested in simulation, with twenty simulations repeated for each combination for each subject.

1) *Acquisition Function*: Expected Improvement (EI), Probability of Improvement (PoI), and Lower Confidence Bound (LCB) are used in this study. EI finds the next point which maximizes mean value improvement. This expected improvement is defined as

$$EI(x) = \mathbb{E}[\max(0, m(x^+) - f(x))], \quad (19)$$

where x^+ is the best point from data points explored so far and $m(x^+)$ is mean value of Gaussian process model value at point x^+ . PoI finds the next point which has the maximum probability of improvement as

$$PoI(x) = \mathbb{P}[\max(f(x) < m(x^+) - \gamma)], \quad (20)$$

where γ is margin. The LCB acquisition function selects the next point which maximizes the lower bound as

$$LCB(x) = \max(m(x) - \beta \cdot \sigma_x), \quad (21)$$

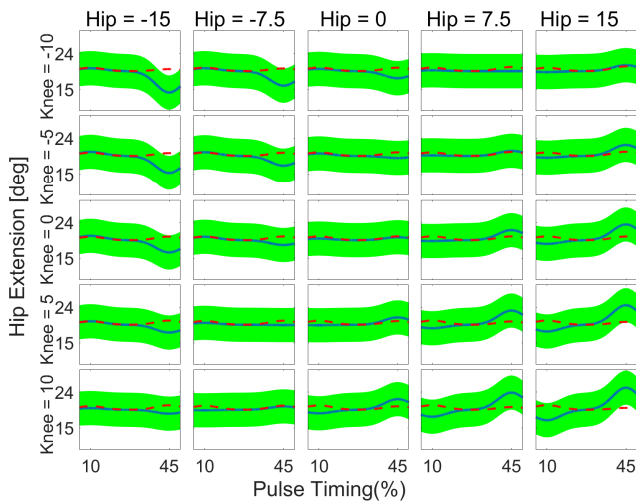


Fig. 1. Gaussian process model of group-level Hip Extension (HE) data. Pulse factors are indexed by the horizontal panel coordinate (hip torque), vertical panel coordinate (knee torque), x axis within a panel (pulse timing). The two lines in each panel indicate different levels of factor stride. Red dashed line: model at the stride before intervention (stride -1); blue line: model at the stride of intervention. Hip and Knee indicate the amplitude of torque pulses applied to hip joint and knee joint respectively. Extension torques are positive in hip and knee. At each point, shaded areas extend one standard deviation beyond the mean.

where σ_x is standard deviation at point x , and $\beta > 0$ is a heuristic trade-off ($\beta = 2$ in this work).

2) *Seed Point Selection Method*: Three seed point selection methods are used in this work. In the first method (*divide*), eight points are selected by dividing the parameter search space in eight regions (the domain for each parameter is divided into two equal parts), and then each of the eight points are selected randomly within each of the eight regions. In the second method (*random*), eight points are randomly defined in the search space. In the third method (*optimal*), the set of eight seed points is composed of seven random points in the search space, plus the optimal point of the group-level Gaussian process model.

3) *Virtual Optimization*: Virtual Bayesian optimization is conducted to maximize HE and PI at the stride of pulse application, and the outcome measure is defined as the number of iterations n_x and n_y required to achieve one of two convergence criteria.

If $\hat{x}_{opt,i}$ is the normalized coordinates optimal point estimated via virtual optimization after i iterations, and x_{opt} is the known optimal point for that subject, n_x is defined as the minimum value of i where the normalized difference x_{diff} between the two quantities is smaller than 10%, where

$$x_{diff,i} = \sqrt{\frac{1}{3} \sum_{q=1}^3 \left(\frac{\hat{x}_{opt,i} - x_{opt}}{p_{q,max} - p_{q,min}} \right)^2}, \quad (22)$$

where p is the range of each component of x .

The second criterion is based on the value of the estimated outcome after a certain number of iterations. Specifically, n_y is defined as the minimum number of iterations where the normalized estimation error $f_{diff}(x_i)$ is smaller than 5%

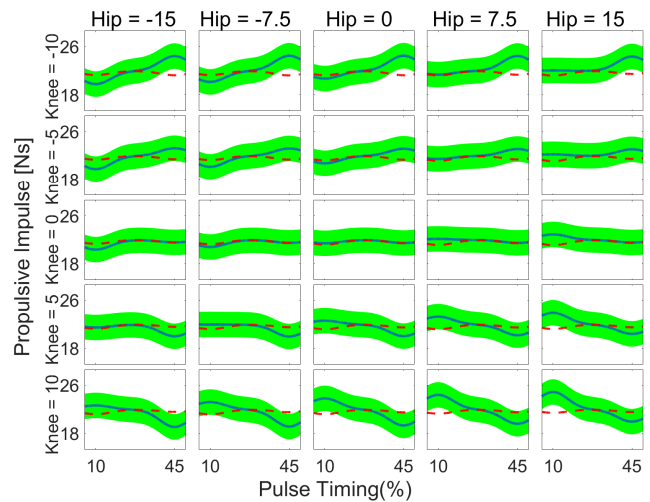


Fig. 2. Gaussian process model of group-level Propulsive Impulse (PI) data.

of the known subject-specific model maximum $G_{sbj,\alpha}(x^*)$, where

$$f_{diff}(x_i) = \sqrt{\left(\frac{\hat{f}(x_i) - G_{sbj,\alpha}(x^*)}{G_{sbj,\alpha}(x^*)} \right)^2}, \quad (23)$$

All simulations were run for 80 iterations. For each criterion, if no convergence was achieved within 80 simulations, n was set to 80.

Four two-way ANOVA, one for each outcome measure (combination of convergence criterion – $n(x)$ and $n(y)$) – and propulsive metric – HE and PI – were conducted to quantify the effects of the two factors (i.e., acquisition function, three levels, and seed point selection methods, three levels) on convergence speed.

III. RESULTS

A. Gaussian Process Modeling

1) *Group-level Model*: Results for group-level Gaussian process modeling are shown in Fig. 1 and Fig. 2. The process variances σ_G^2 calculated from group-level propulsion are equal to 4.31 deg² for HE and 2.12 N²s². Estimated noise standard deviation and hyperparameter values are listed in Table II.

2) *Subject-specific Results*: The mean of the distribution of subject-specific process variances $\sigma_{G_{sbj}}^2$ were equal to 69.04 deg² and 33.86 N²s² for HE and PI, respectively. The number of optimal points for subject-specific models within

TABLE II
LENGTH SCALE HYPERPARAMETER VALUES AND ESTIMATED NOISE STANDARD DEVIATION

	T [% gait cycle]	S [°]	K [N·m]	H [N·m]	$\sigma_{\epsilon,\alpha}$
HE	26.71	0.73	6.17	7.43	0.21 deg
PI	12.18	0.50	8.66	11.73	0.78 Ns

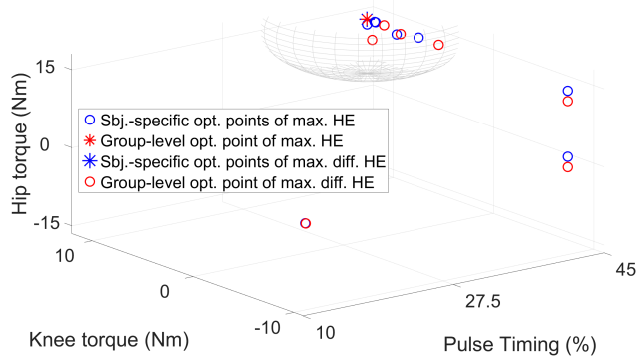


Fig. 3. Distribution of the optimal points of subject-specific models (circles) and of the group-level model (asterisks) for HE. The 20% range is indicated by a sphere centered around the optimal point of the group-level model.

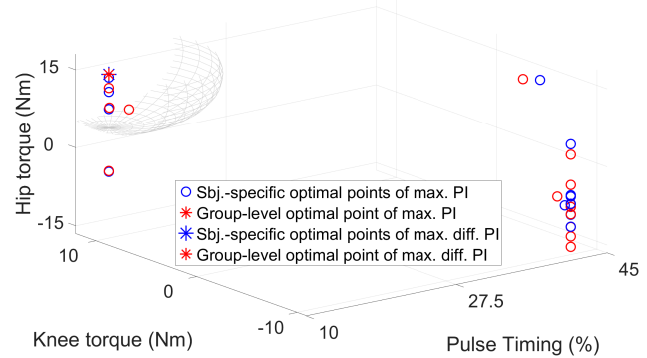


Fig. 4. Distribution of the points of sixteen subject-specific models (circles) and of the group-level model (asterisk dots) for PI. 20 % difference are indicated by grid sphere around optimal point of group-level model.

percentage difference of the group-level model is shown in Fig. 5. The optimal points are for maximal outcomes during stride 0 and maximum change in outcome (between stride 0 and stride -1) for each subject-specific model. The optimal points for the subject-specific models are shown in Fig. 3 and Fig. 4 for HE and PI, respectively. 12 subject-specific Gaussian process models have optimal points that have a percentage difference of less than 20 % compared to the optimal point of group-level Gaussian process model for HE (Fig. 3). For propulsive impulse case (Fig. 4), only 5 optimal points of subject-specific Gaussian process model are located within 20% difference from the optimal point of group-level Gaussian process model. For PI, the minority of subjects exhibited subject-specific optimal points within a reasonable neighborhood of the group-level optimal point. In fact, more optimal points of subject-specific models in PI case are located in late stance (45 %) than early stance (10 %) where optimal point of group-level model is located.

B. Bayesian Optimization Simulation

Virtual Bayesian optimization experiments achieved convergence within 80 iterations in 80.58 % of runs (94.40 ± 6.06 % when using EI, 85.57 ± 13.94 % when using PoI, 61.77 ± 26.66 % when using LCB). The results of the ANOVA analysis are reported in Table III and Fig. 6. Based on the ANOVA analysis, the seed point selection method had no significant effect on the outcome measure in any condition. Instead, the choice of an acquisition function showed to be associated with the number of iterations required for convergence ($p < 0.001$ in all conditions tested). Specifically, EI was the only acquisition function type that was estimated to have negative coefficient in the linear model when using *random* method and PoI as neutral level of the factors, indicating that its effect is to decrease the number of iterations to obtain convergence. Post-hoc tests are conducted to distinguish the significant difference among groups under different combinations of acquisition functions and seed point selection methods (Fig. 6). Post-hoc tests indicate that within the same group of acquisition function types, there exist no significant differences ($p > 0.001$ for all comparisons). Moreover, post-hoc tests comparing different

acquisition functions showed that EI was significantly better than LCB in all conditions tested, and afforded a significantly greater accuracy in identifying optimal control parameters compared to PoI.

IV. DISCUSSION AND CONCLUSIONS

In this study, we determined the feasibility of using Gaussian process modeling to establish modeling the relationship between three exoskeleton control parameters (hip torque pulse amplitude, knee torque pulse amplitude, and timing of the pulses), and outcome measures relevant for propulsion. To estimate parameters of a Gaussian process, we estimated the noise variance and tuned the hyperparameters of a Gaussian process describing the response at the group level. Based on the group-level Gaussian process model, we estimated 16 subject-specific models and used those as data-generating process for virtual HIL Bayesian optimization simulations. In the virtual Bayesian optimizations, we established how many iterations would be required for convergence of a hypothetical experiment, and quantified the effects of different acquisition functions and seed point selection methods on convergence speed.

Our results of group-level Gaussian process model (Fig. 1 and Fig. 2) demonstrate that it is feasible to construct a model between the 3D space of control parameters and outcomes using a Gaussian process with estimated noise variance. Based on previous predictions using linear mixed model [15],

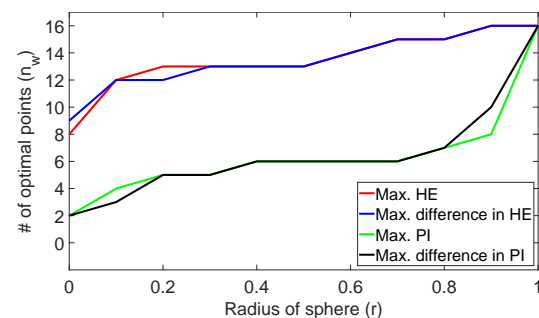


Fig. 5. Number of subjects (n_w) whose subject-specific model maxima are within a specified distance (r) from the group-level model maximum.

group-level Gaussian process models for both HE and PI are expected to have a similar response when pulses are applied to the knee and hip joints. In agreement with the results of the previous study [15], hip extension torque increases HE in late stance but decreases HE in early stance; knee extension and hip extension pulse torque increase PI in early stance; knee extension pulse torque decreases PI in late stance. The variability of optimal points for subject-specific models is moderate in HE and high in PI. Specifically for the PI case, the maxima of subject-specific models fell within 20% normalized distance from the optimal point for group-level model only in the minority of cases (5/16). This indicates that the optimal point for group-level model is likely to be distant from the optimal point for subject-specific models. In this case, in fact, optimal points for subject-specific models are separated to make two clusters of optimal points in early and late stance. Our analysis in HE case is aligned with previous work targeting reduction in metabolic cost, indicating that one subject's optimal point can be optimal for another subject [3].

In contrast with previous research selecting seed points by dividing the search space in N regions and randomly selecting seed points from all the N regions [3], seed point selection method is not estimated to make a significant difference in a single pulse torque application experiment ($p > 0.12$ for the different conditions tested, Table III). This results are consistent in both HE and PI cases. Since we included *optimal* method in seed point selection method, this results also indicate that optimal point location relationship between group-level optimal point and subject-specific optimal points in HE case do not effect results of simulations. However, the type of acquisition function used for optimization did significantly contribute to the number of iterations required to achieve convergence in all cases of outcome measures (HE and PI) and point locations ($p < 0.001$). Since the interaction term between types of acquisition function and seed point selection methods is not significant ($p > 0.1$ for all conditions tested), the optimal acquisition function (in terms of convergence speed) of HIL Bayesian optimization is estimated to be EI. Therefore, our expectation toward HIL Bayesian optimization setting is that LCB will require the largest number of iterations to establish convergence, while EI will require the smallest number for iterations for

TABLE III

TWO WAY ANOVA TEST RESULTS FOR NUMBER OF ITERATIONS TO ESTABLISH CONVERGENCE

Source	HE		Point location (HE)	
	F ratio	Prob> F	F ratio	Prob> F
Acq. func.	27.9950	< 0.001	150.1179	< 0.001
Seed pt.	0.0702	0.9322	0.2270	0.7969
Interaction	0.2350	0.9187	1.2166	0.3017

Source	PI		Point location (PI)	
	F ratio	Prob> F	F ratio	Prob> F
Acq. func.	100.9153	< 0.001	137.0745	< 0.001
Seed pt.	0.2063	0.8136	0.1297	0.8784
Interaction	0.1021	0.9818	0.5217	0.7198

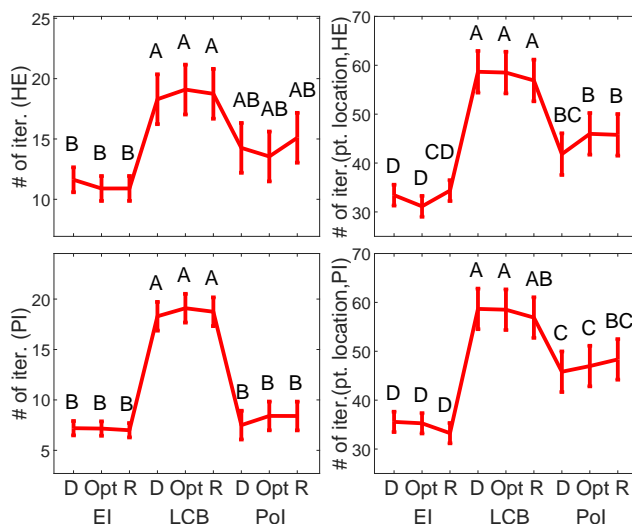


Fig. 6. 2-way ANOVA interaction plot with post-hoc test results. D, Opt, and R in x-axis indicate the *divide* method, *optimal* method, and *random* method for seed point selection method respectively. Pairs of groups that share the same letter label are not significantly different.

convergence regardless of seed point selection methods.

This study has limitations. Since virtual HIL Bayesian optimization rely on our assumption that subject-specific Gaussian process models can fully represent actual human response, the validity of these models directly relates to the validity of the results of Bayesian optimization obtained in this study. Another limitation is that noise is not fully eliminated to construct a model that relate control input parameters and propulsive mechanics. As shown in Fig. 1 and Fig. 2, the model at the stride before intervention (red dashed line) have changes with control input parameters, which is impossible as a change in output cannot anticipate a change in input. This result indicates that our optimization procedure for estimating hyperparameters and noise variance may need to be improved, or that the relationship between input parameters and improvement of outcome measures (HE and PI) should be considered to cancel out noise from stride at intervention (stride 0) and stride before intervention (stride -1) Gaussian process models.

Overall, our study identified a new method based on Gaussian process modeling to estimates the relationship between exoskeleton control parameters and specific gait features (HE and PI), and established the effects of seed point selection method and acquisition function types in number of iterations for convergence. These results make foundation of promising approach for planning further HIL robotic gait training experiment by focusing on specific gait features to target with training.

REFERENCES

- [1] L. J. Holanda, P. M. Silva, T. C. Amorim, M. O. Lacerda, C. R. Simão, and E. Morya, "Robotic assisted gait as a tool for rehabilitation of individuals with spinal cord injury: a systematic review," *Journal of neuroengineering and rehabilitation*, vol. 14, no. 1, pp. 1–7, 2017.
- [2] 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.

- [3] 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, p. eaar5438, 2018.
- [4] J. Zhang, P. Fiers, K. A. Witte, R. W. Jackson, K. L. Poggensee, C. G. Atkeson, and S. H. Collins, "Human-in-the-loop optimization of exoskeleton assistance during walking," *Science*, vol. 356, no. 6344, pp. 1280–1284, 2017.
- [5] 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 *Robotics: Science and Systems*, 2016, pp. 1–10.
- [6] M. Kim, Y. Ding, P. Malcolm, J. Speeckaert, C. J. Siviyy, C. J. Walsh, and S. Kuindersma, "Human-in-the-loop bayesian optimization of wearable device parameters," *PloS one*, vol. 12, no. 9, p. e0184054, 2017.
- [7] J. Snoek, H. Larochelle, and R. P. Adams, "Practical bayesian optimization of machine learning algorithms," in *Advances in neural information processing systems*, 2012, pp. 2951–2959.
- [8] B. Letham, B. Karrer, G. Ottoni, E. Bakshy *et al.*, "Constrained bayesian optimization with noisy experiments," *Bayesian Analysis*, vol. 14, no. 2, pp. 495–519, 2019.
- [9] J. Kim and S. Choi, "On local optimizers of acquisition functions in bayesian optimization," *arXiv preprint arXiv:1901.08350*, 2019.
- [10] M. Balandat, B. Karrer, D. R. Jiang, S. Daulton, B. Letham, A. G. Wilson, and E. Bakshy, "Botorch: Programmable bayesian optimization in pytorch," *arXiv preprint arXiv:1910.06403*, 2019.
- [11] H. Wang, B. van Stein, M. Emmerich, and T. Back, "A new acquisition function for bayesian optimization based on the moment-generating function," in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2017, pp. 507–512.
- [12] H. Hsiao, B. A. Knarr, J. S. Higginson, and S. A. Binder-Macleod, "The relative contribution of ankle moment and trailing limb angle to propulsive force during gait," *Human movement science*, vol. 39, pp. 212–221, 2015.
- [13] C. L. Peterson, J. Cheng, S. A. Kautz, and R. R. Neptune, "Leg extension is an important predictor of paretic leg propulsion in hemiparetic walking," *Gait & posture*, vol. 32, no. 4, pp. 451–456, 2010.
- [14] M. G. Bowden, C. K. Balasubramanian, R. R. Neptune, and S. A. Kautz, "Anterior-posterior ground reaction forces as a measure of paretic leg contribution in hemiparetic walking," *Stroke*, vol. 37, no. 3, pp. 872–876, 2006.
- [15] R. L. McGrath and F. Sergi, "Single-stride exposure to pulse torque assistance provided by a robotic exoskeleton at the hip and knee joints," in *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*. IEEE, 2019, pp. 874–879.