

Multi-Objective Human-in-the-Loop Bayesian Optimization of a Lower-Limb Exoskeleton

Neil Janwani, Matthew T. Lerner, Aaron J. Young, Maegan Tucker

Abstract—Human-in-the-loop optimization (HILO) is a common approach for optimizing the control of assistive devices to account for the wearer’s unique biomechanics and subjective preferences. However, despite research suggesting that a person may have a different prioritization of objectives depending on time-varying factors such as the environment, their mood, or energy levels, existing HILO approaches only consider a single objective or enforce a fixed weighting on a set of objectives. Neither approach is capable of representing an individual’s preferences *over objectives*. In this work, we propose Multi-Objective Human-in-the-loop Bayesian Optimization (MO-HILBO), which builds on explicit multi-objective Bayesian optimization to efficiently infer a personalized set of Pareto-optimal controllers. We compare our approach with an existing multi-objective HILO method and experimentally demonstrate MO-HILBO on a lower-limb exoskeleton across two objectives: metabolic cost (efficiency) and ordinal human feedback (comfort). We find that MO-HILBO (1) discovers Pareto-optimal controllers, and (2) that the pairwise ordering of points on the Pareto front itself is consistent with validation trials. Lastly, we open-source `pyPolar`, a Python package for running both HILO and MO-HILBO on wearable devices: <https://anonymous.4open.science/r/pyPolar-52F6/>.

I. INTRODUCTION

Assistive wearable robots like exoskeletons carry an exciting potential for human augmentation and rehabilitation in everyday tasks including walking [1]–[4], running [5], stair-climbing [6], and more [7]–[9]. While the benefits of such devices—typically captured by metrics such as metabolic cost, walking speed, gait symmetry, and subjective preference—largely depend on the underlying controller, it is widely recognized that personalization is required to maximize the benefits for each user [4], [10]–[15]. Specifically, personalization accounts for factors such as unique gait biomechanics and subjective preferences that vary across individuals.

Human-in-the-loop Optimization (HILO) is the standard approach for accomplishing personalization [16]. Broadly, HILO iteratively cycles between 1) selecting new parameters to execute on the device, and 2) collecting corresponding human outcome metrics (e.g. metabolic cost, preference). An often overlooked aspect of HILO is selecting an appropriate outcome to optimize. While metabolic cost is the most commonly selected objective [3], [4], [6], [7], [17]–[30], it can conflict with other measures such as comfort [31]. As such, solely optimizing for one outcome may be detrimental to other important objectives that matter to humans.

In this work, we argue that the wearable device community should move towards optimizing for a *combination of objectives*, rather than only for a single one. One example

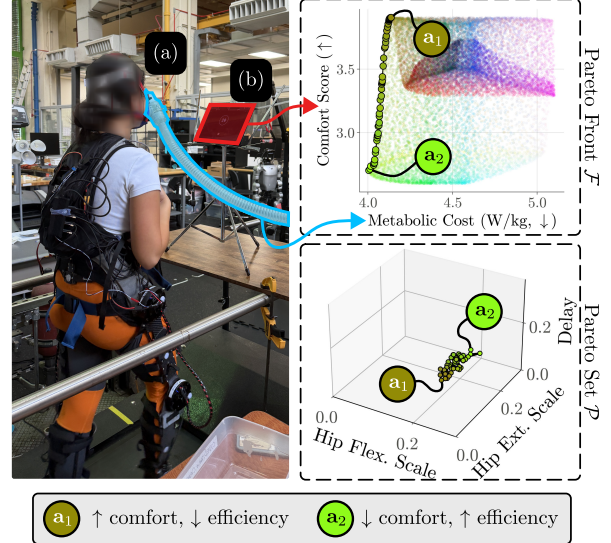


Fig. 1: Our work demonstrates MO-HILBO on a lower-limb exoskeleton (left) over two objectives: (a) metabolic cost, measured by indirect calorimetry, and (b) subjective comfort, recorded using a tablet. The protocol identifies the Pareto set of control parameters (bottom right) defined as the actions associated with the Pareto front, approximated by a learned probabilistic surrogate (top right).

would be speed and efficiency, explored in [32]. Here, a scalar-valued multi-objective cost function was constructed by assigning fixed weights to each objective, scaling each objective’s impact relative to the other. By using two feedback mechanisms, one for each objective, the authors used a traditional HILO setup to complete the optimization. However, this requires that the wearer’s *trade-off* of objectives (i.e., the scalarizing weighting terms assigned to each objective) be known *prior* to the fitting procedure starting. Moreover, a wearer’s preferred trade-off may not be consistent across individuals (or even over time as the user adapts to the device), as implied by preference-based HILO [10]–[12], [14], [15], or be consistent in a single individual across different task settings. For instance, if you are late to a meeting, you might sacrifice efficiency for speed in order to make it on time, despite normally opting for efficiency.

While canonical optimization minimizes a scalar-valued objective, *multi-objective* optimization considers a vector-valued signal with each entry corresponding to a distinct objective. Since no controller can optimize all objectives simultaneously, the goal is to instead infer a *Pareto-optimal set* of controllers which optimally trade off among objectives. A popular approach that, to our knowledge, has not yet been applied to human assistive robotics, is multi-objective Bayesian optimization (MOBO). MOBO takes a sample-

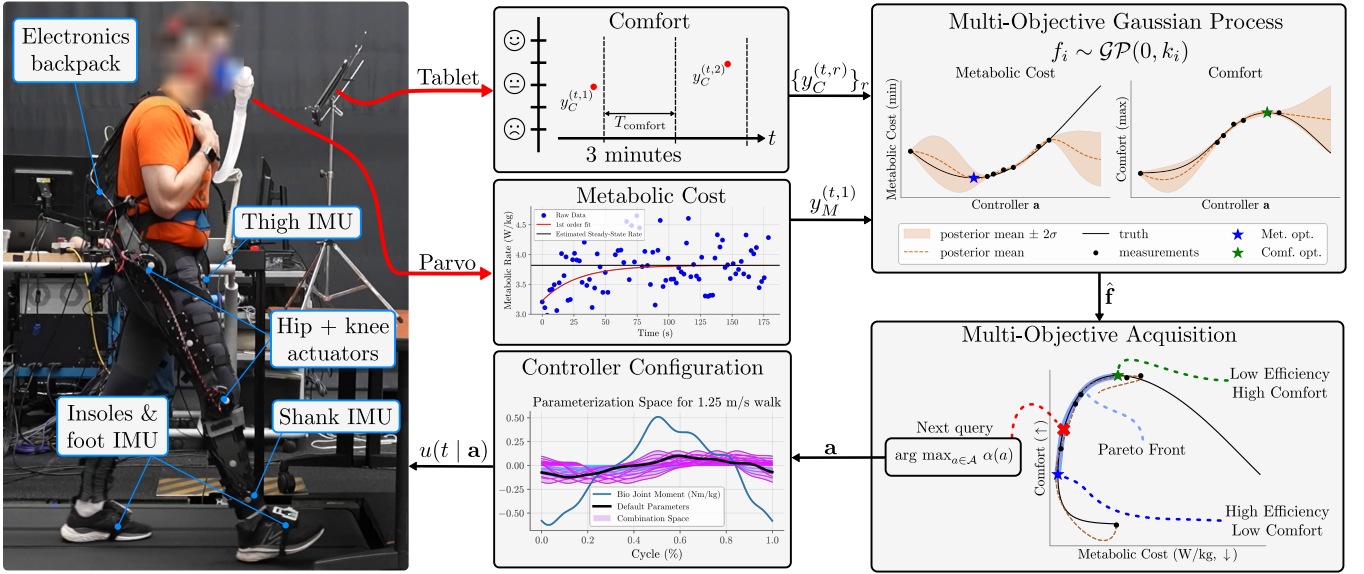


Fig. 2: MO-HILBO aims to efficiently identify the Pareto set of mutually-optimal actions across multiple objectives. Similar to traditional human-in-the-loop Bayesian optimization, MO-HILBO cycles over three main steps: 1) collecting feedback from human users, 2) updating the belief of the underlying objective function(s) based on the provided feedback, and 3) selecting new queries for the subsequent iteration using an acquisition function. In our work, we explore the application of MO-HILBO on a lower-limb exoskeleton across two objectives: subjective evaluations of comfort using ordinal labels, and metabolic cost measured by indirect calorimetry.

efficient, and noise-aware approach by intelligently querying different controllers to balance exploration and exploitation.

Of particular relevance is the work of [33], which introduces non-Bayesian multi-objective HILO by using an evolutionary algorithm: Non-dominated Sorting Genetic Algorithm (NSGA-II) [34]. However, with a limited budget, this approach often finds sparse Pareto-optimal points that may not describe the full Pareto set, and must roll out entire generations of parameters to the user before making an update. In comparison, MOBO is a more sample-efficient optimizer that infers dense Pareto sets via a continuous approximation of the objective space.

In this work, we demonstrate the first, to our knowledge, multi-objective human-in-the-loop Bayesian optimization (MO-HILBO) protocol for wearable devices, as illustrated in Fig. 1. Our explicit contributions are:

- 1) We introduce MO-HILBO and showcase its improved performance on limited sample, high noise regimes compared to previous genetic-based multi-objective HILO approaches.
- 2) We develop a MO-HILBO protocol and demonstrate it towards optimizing 3 parameters of a lower-limb exoskeleton controller over two objectives: metabolic cost and subjective comfort.
- 3) We introduce and demonstrate a validation procedure for multi-objective HILO, showing that MO-HILBO identifies Pareto-optimal points and that the ordering of the front with respect to a single objective is largely consistent for each user.
- 4) We provide `pypolar`¹: a Python framework for using both MO-HILBO and single-objective Bayesian Optimization for HILO on assistive robots.

¹Code available at: <https://anonymous.4open.science/r/pyPolar-52F6/>

II. MO-HILBO

The MO-HILBO framework, illustrated in Fig. 2, has two main components: the multi-objective Gaussian Process (MO-GP) model, which infers the Pareto-optimal controller set, and the multi-objective acquisition function, which picks the next controller to query based on the MO-GP's current estimate of the Pareto set. While this acquisition can be modified for different end goals, we select an acquisition function to efficiently explore the trade-off space, with the goal of identifying a Pareto front with the largest hypervolume as quickly as possible. Our algorithm is open-sourced through the Python package `pypolar`¹, making it available for practitioners who wish to perform sample-efficient multi-objective HILO.

A. Problem Setup

The goal of MO-HILBO is to identify a set of actions, with each action $\mathbf{a} \in \mathcal{A} \subset \mathbb{R}^n$ being a tuple of n tuneable parameters, which optimally trade off across a set of m objectives $\mathbf{f} : \mathcal{A} \rightarrow \mathbb{R}^m$ (i.e., $\mathbf{f}(\mathbf{a}) = [f_1(\mathbf{a}), \dots, f_m(\mathbf{a})]^\top$, with $f_i : \mathcal{A} \rightarrow \mathbb{R}$). In other words, we wish to optimize these m objectives simultaneously. However, since these objectives may conflict, no single controller is necessarily optimal for all objectives at once. Instead, MO-HILBO solves for the Pareto set of controllers $\mathcal{P}_{\mathbf{f}}$, which contains only controllers that are *non-dominated*:

$$\mathcal{P}_{\mathbf{f}} := \{\mathbf{a} \in \mathcal{A} : \nexists \mathbf{a}' \in \mathcal{A} \text{ with } \mathbf{a}' \succ \mathbf{a}\}, \quad (1)$$

where $\succ$ denotes Pareto dominance over the m objectives $f_1, \dots, f_m$:

$$\mathbf{a}' \succ \mathbf{a} \iff \begin{cases} f_j(\mathbf{a}') \geq f_j(\mathbf{a}) & \forall j \in \{1, \dots, m\}, \\ f_j(\mathbf{a}') > f_j(\mathbf{a}) & \text{for some } j. \end{cases} \quad (2)$$

Plainly stated, controller $\mathbf{a}$ is excluded from $\mathcal{P}_f$ whenever another controller $\mathbf{a}'$ can perform as well as $\mathbf{a}$ on every objective and strictly improve on at least one. With this, we define the Pareto front $\mathcal{F}$ as

$$\mathcal{F}_f := \mathbf{f}(\mathcal{P}_f) = \{\mathbf{f}(\mathbf{a}) : \mathbf{a} \in \mathcal{P}_f\} \subset \mathbb{R}^m. \quad (3)$$

For notational consistency in this section, we seek to *maximize* all objectives. While this is not always the case (e.g., in practice we seek to *minimize* metabolic cost), we use a standardization procedure in MO-HILBO's initialization that transforms all objectives into maximization space.

B. Multi-Objective Gaussian Process

With finding $\mathcal{P}_f$ as the express goal of MO-HILBO, it is tempting to simply evaluate $\mathbf{f}$ on many actions and choose the non-dominated ones as $\mathcal{P}_f$. However, simply evaluating each objective f_i requires eliciting biomechanical or psychometric feedback from the human, making it a costly operation. Thus, we instead learn a *probabilistic surrogate* of the true objectives $\mathbf{f}$ using an independent GP for each objective:

$$f_j \sim \mathcal{GP}(0, k_j(\cdot, \cdot)), \quad j = 1, \dots, m, \quad (4)$$

where each GP in the MO-GP represents a non-parametric distribution over the unknown true objective function f_j . We denote k_j as the kernel function which determines the prior covariance of any two actions $\mathbf{a}, \mathbf{a}'$. Several kernels exist in the Gaussian process literature; we use the automatic relevance determination (ARD) squared-exponential kernel:

$$k_j(\mathbf{x}, \mathbf{x}') = \sigma_{f,j}^2 \exp\left(-\frac{1}{2} \sum_{i=1}^n \frac{(x_i - x'_i)^2}{\ell_{j,i}^2}\right), \quad (5)$$

for its commonplace use in regression of black-box functions. Here, $\{\sigma_{f,j}, \ell_{j,i}\} \in \mathbb{R}_+$ are hyperparameters, denoting the signal variance and lengthscale, respectively. Intuitively, signal variance defines the expected scale of the objective and lengthscale defines the correlation between two actions. We optionally set a minimum lengthscale $\ell_{\min}$ that prevents each GP from overfitting by collapsing the lengthscale.

There are two primary benefits of using a Gaussian process as a surrogate model. First, is that we obtain a continuous, easy-to-evaluate, posterior of each objective over the action space. Second, we can obtain the confidence of the surrogate by evaluating the variance of the Gaussian. Specific formulations are provided in the model update section below.

C. Pareto Set Discovery

As summarized in Algorithm 1, MO-HILBO begins with an initialization phase, followed by iteratively cycling through three steps: (1) choosing the next set of control parameters to roll out to the user, (2) eliciting the corresponding feedback from the human user, and (3) updating the surrogate model. We repeat this for a fixed budget T . In practice, we choose the largest T that does not fatigue the user, maximizing the amount of data collected.

Algorithm 1 MO-HILBO

Require: action space $\mathcal{A}$, objectives 1.. m with directions ρ_j , budget T , initialization allotment T_0

- 1: $\mathcal{D}_j \leftarrow \emptyset$ for $j = 1, \dots, m$
- 2: **for** $t = 1, \dots, T$ **do**
- 3: **if** $t \leq T_0$ **then**
- 4: $\mathbf{a}^{(t)} \sim \text{Unif}(\mathcal{A})$ $\triangleright$ no posterior yet
- 5: **else**
- 6: $\mathbf{a}^{(t)} \leftarrow \text{qLogNParEGO}$
- 7: **end if**
- 8: Send $\mathbf{a}^{(t)}$ to the device
- 9: Measure $\{y_j^{(t,r)}\}$ in parallel, including repeats.
- 10: Store feedback in $\mathcal{D}_j^{(t)}$
- 11: Refit μ and v by maximizing (15)
- 12: **end for**
- 13: **return** (μ and v) and the corresponding Pareto set $\mathcal{P}_f^{(T)}$

Initialization. We initialize the algorithm by setting the bounds on the action space $[\underline{a}_i, \bar{a}_i]$ for each dimension $i \in [1, n]$. We instantiate a normalizing transform φ which maps $\mathbf{a} \in \mathcal{A}$ to the unit-box:

$$\mathbf{x} = \varphi(\mathbf{a}) := \left[\frac{a_i - \underline{a}_i}{\bar{a}_i - \underline{a}_i} \right]_{i=1}^n \in [0, 1]^n. \quad (6)$$

This is convenient for dealing with actions that span different ranges or exist in different units. Second, given the number of objectives m and whether each is to be maximized or minimized, we instantiate a standardizing transform:

$$z_j = \rho_j \frac{y_j - \hat{\mu}_j}{\hat{\sigma}_j}, \quad \rho_j = \begin{cases} +1 & \text{if } f_j \text{ is maximized} \\ -1 & \text{otherwise} \end{cases}, \quad (7)$$

where y_j is the feedback measure provided for objective j , and $\hat{\mu}_j$ and $\hat{\sigma}_j$ are per-objective running mean and standard deviation, respectively, computed from the entire feedback dataset. Superscripts (i.e., $y_j^{(\cdot)}$) will be used later to index per-iteration feedback, but are omitted here for clarity. Lastly, we randomly sample T_0 actions to begin the algorithm.

Feedback Collection. After initialization, MO-HILBO enters a feedback collection stage where the selected action(s) (in our case, control parameters) is rolled out to the user and their feedback is elicited. Standard MOBO assumes that a single feedback point is obtained per objective per action. We relax this structure into *decoupled feedback collection* by appending to the following dataset:

$$\mathcal{D}_j^{(t)} = \left\{ (\mathbf{x}^{(t')}, y_j^{(t',r)}) : t' = 1, \dots, t, \quad r = 1, \dots, c_j^{(t')} \right\}, \quad (8)$$

$$N_j^{(t)} = \sum_{t'=1}^t c_j^{(t')}, \quad N_j^{(t)} \neq N_{j'}^{(t)} \text{ in general}, \quad (9)$$

where $t \in \mathbb{N}$ is the current iteration number, and r indexes $c_j^{(t)} \in \mathbb{N}$ repeated measurements for objective j in trial t . For example, for iteration t there could have been three repeated measurements of objective one ($\{y_1^{(t,1)}, y_1^{(t,2)}, y_1^{(t,3)}\}$, $c_1^{(t)} = 3$) and only one for objective two ($y_2^{(t,1)}$, $c_2^{(t)} = 1$).

We found this parallelized, decoupled scheme to significantly reduce the full data collection procedure's duration as it enables multiple feedback points on one objective to be collected while a different, more time-consuming objective is being evaluated. Additionally, when feedback is noisy, evaluating an objective multiple times on a single action helps improve the quality of the surrogate model.

Updating the Multi-Objective Gaussian Process. After collecting feedback, we fit a closed-form GP [35] to each $\mathcal{D}_j^{(t)}$, with mean and variance, respectively computed as:

$$\mu_j^{(t)}(\mathbf{x}_*) = \mathbf{k}_{j,*}^\top (K_j + \sigma_{n,j}^2 I)^{-1} \mathbf{z}_j^{(t)}, \quad (10)$$

$$v_j^{(t)}(\mathbf{x}_*) = k_j(\mathbf{x}_*, \mathbf{x}_*) - \mathbf{k}_{j,*}^\top (K_j + \sigma_{n,j}^2 I)^{-1} \mathbf{k}_{j,*}, \quad (11)$$

where $\mathbf{z}_j^{(t)}$ is the standardized feedback and vector $[\mathbf{k}_{j,*}]_{(t',r)}$ and Gram Matrix K_j are computed as:

$$[K_j]_{(t',r),(t'',r'')} = k_j(\mathbf{x}^{(t')}, \mathbf{x}^{(t'')}), \quad (12)$$

$$[\mathbf{k}_{j,*}]_{(t',r)} = k_j(\mathbf{x}_*, \mathbf{x}^{(t')}), \quad (13)$$

for any algorithm iterations t', t'' . See that, due to repeated feedback, there may exist repeated values in K_j , making it singular. However, because each objective is assumed to be noisy, we assign $\sigma_{n,j}$, a non-zero user-defined prior per-objective variance, that is added to the diagonal of K_j above. We can now define the surrogate Pareto set $\mathcal{P}_{\hat{\mathbf{f}}}^{(t)} \triangleq \mathcal{P}_{\boldsymbol{\mu}^{(t)}}$, constructed by the vector-valued mean of the individual GPs:

$$\boldsymbol{\mu}^{(t)}(\mathbf{x}) = [\mu_1^{(t)}(\mathbf{x}), \dots, \mu_m^{(t)}(\mathbf{x})]^\top. \quad (14)$$

Lastly, because hyperparameters $\boldsymbol{\theta}_j = \{\sigma_{f,j}^2, \ell_{j,i}, \sigma_{n,j}^2\}$ are not necessarily known in advance, we estimate them from data, as commonly done in the literature [35]. However, rather than maximizing the log marginal likelihood $\log p(\mathbf{z}_j^{(t)} | X_j^{(t)}, \boldsymbol{\theta}_j)$ alone, we incorporate the practitioner-specified prior on the measurement noise and obtain a maximum a posteriori:

$$\boldsymbol{\theta}_j^* = \underset{\boldsymbol{\theta}_j}{\operatorname{argmax}} \log p(\mathbf{z}_j^{(t)} | X_j^{(t)}, \boldsymbol{\theta}_j) + \log p(\sigma_{n,j}^2). \quad (15)$$

with $X_j^{(t)}$ denoting the normalized training set of actions at iteration t and $\mathbf{z}_j^{(t)}$ the corresponding feedback standardized by Eqn (7).

Choosing the Next Query. To select a query/action to execute on the system during iteration t , denoted $\mathbf{a}^{(t)}$, we utilize a multi-objective acquisition function $\alpha : [0, 1]^n \rightarrow \mathbb{R}$ that trades off between exploitation and exploration given the collected data. We specifically choose qLogNParEGO as implemented in BoTorch [36]. This acquisition function is a log-domain, noise-aware form of ParEGO [37]–[39], which works well in cases with several objectives and noisy feedback.

First, qLogNParEGO draws a trade-off vector uniformly from the simplex:

$$\mathbf{w} \sim \operatorname{Unif}(\Delta^{m-1}), \quad w_j \geq 0, \quad \sum_j w_j = 1, \quad (16)$$

and then collapses the vector-valued problem into a scalar one with an *augmented Chebyshev* scalarization:

$$s_{\mathbf{w}}(\mathbf{u}) = \min_j (w_j u_j) + \alpha_{\text{aug}} \sum_{j=1}^m w_j u_j, \quad \alpha_{\text{aug}} = 0.05. \quad (17)$$

Importantly, $\mathbf{w}$ is chosen randomly each time α is invoked. This directly enables MO-HILBO to explore the Pareto set through different prioritizations of objectives, rather than choosing a fixed one as in [32].

Second, given $\mathbf{w}$, qLogNParEGO scores the scalarization candidate using the expected improvement over the best scalarized value of the measured feedback:

$$\alpha_{\mathbf{w}}(\mathbf{x}) = \mathbb{E} \left[(s_{\mathbf{w}}(\psi^{(t)}(\hat{\mathbf{f}}(\mathbf{x}))) - \max_{\mathbf{x}'} s_{\mathbf{w}}(\psi^{(t)}(\hat{\mathbf{f}}(\mathbf{x}'))))^+ \right], \quad (18)$$

with $(\cdot)^+ := \max(\cdot, 0)$, $\hat{\mathbf{f}}$ is a posterior sample drawn from the MO-GP given data $\mathcal{D}^{(t)}$, and $\psi^{(t)} : \mathbb{R}^m \rightarrow [0, 1]^m$ is a normalizing transform that ensures relative comparison of objectives on different scales. Lastly, we maximize the acquisition function using BoTorch [36] to procure the next controller:

$$\mathbf{x}^{(t)} = \underset{\mathbf{x} \in [0, 1]^n}{\operatorname{argmax}} \alpha_{\mathbf{w}}(\mathbf{x}), \quad \mathbf{a}^{(t)} = \varphi^{-1}(\mathbf{x}^{(t)}) \quad (19)$$

and send $\mathbf{a}^{(t)}$ to the device, completing the loop.

III. SYNTHETIC EXPERIMENTS

We compare MO-HILBO with Zhang et al. [33], which is to our knowledge, the only other HILO approach that attempts to approximate the Pareto set of controllers. Zhang et al. utilize the Non-dominated Sorting Genetic Algorithm (NSGA-II) for their core optimization framework. Below, we demonstrate that MO-HILBO results in improved sample efficiency and yields a more accurate approximation of the Pareto front, as indicated by a lower inverted generational distance plus (IGD⁺) [40]. We evaluate on a test domain that mimics the conditions of a human-subject study: high noise and a small feedback budget.

A. Comparison Setup

We construct an ideal-point synthetic function which represents the ground truth objectives. Each objective is optimized at a randomly drawn single action $\mathbf{a}_i^* \sim \operatorname{Unif}(\mathcal{A})$:

$$f_i(\mathbf{a}) = -r_i \|\mathbf{a} - \mathbf{a}_i^*\|_2^2 + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \sigma_{\text{syn}}), \quad (20)$$

where $r_i \in [0.1, 1]$ is uniformly randomly drawn and σ_{syn} is used to inject noise into the synthetic model, representing unavoidable measurement noise in human feedback. When training the GP, we set the prior noise $\sigma_{n,j}$ to within 30% of the true noise σ_{syn} to simulate the realistic case where one does not precisely know the true measurement noise of the groundtruth objective $\mathbf{f}$.

We also wrap each $f_i(\mathbf{a})$ in a tanh function to bound and re-scale the output to $[-1, 1]$. We ablate along three directions: measurement noise (σ_{syn}), action dimension size (n), and number of objectives (m). Unless being swept, we set $\sigma_{\text{syn}} = 0.1, n = 3, m = 2$, mirroring the action

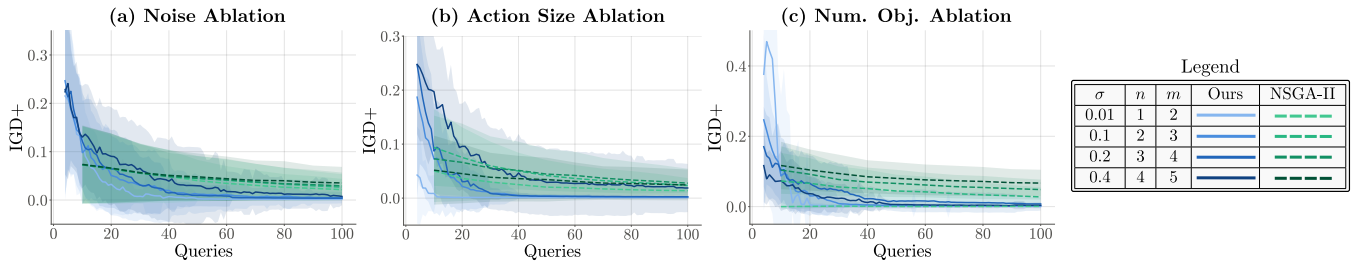


Fig. 3: Synthetic experiments on high noise and limited sample budget. Across the ablation studies (left three plots) MO-HILBO achieves a lower IGD^+ score across different noise levels, action dimensions, and numbers of objectives than NSGA-II. All curves are averaged over 50 runs to better measure performance.

and objective dimensions in our real-world experiment. We perform 50 runs per ablation per algorithm and cap synthetic trials at $T = 100$ to visualize how sample budget affects fit quality and inform our experimental protocol. Lastly, we compute IGD^+ using `pymoo` [41].

B. Baseline

We compare the underlying optimization framework of MO-HILBO (MOBO) against NSGA-II [33], [34]. At a high level, NSGA-II sorts a population of sampled points into successive non-dominated fronts and assigns each point a crowding distance that favors points isolated from their neighbors in objective space. Together these define the crowded-comparison operator, which orders points first by front and only then by crowding distance. Offspring are generated from the better-ranked points by crossover and mutation, after which parents and offspring are pooled and truncated back to the population size under the same operator, so non-dominated solutions are never lost (elitism). Over successive generations the population converges to an approximation of the Pareto front. While NSGA-II is very computationally efficient, it does not fit a surrogate model like MOBO does. Thus, approximation only comes through a large number of objective evaluations. For our experiments, we set our population size $P = 10$, matching the size used in the baseline approach [33].

C. Results

The results of the synthetic experiments and comparison to NSGA-II are illustrated in Fig. 3. Across our ablation study, MOBO reaches better IGD^+ scores than NSGA-II with fewer samples, regardless of noise level, action dimension size, or number of objectives. Across σ_{syn} , MOBO reaches a lower IGD^+ score within 40 queries and converges by 100 queries. In contrast, NSGA-II convergence slows as noise increases. Indeed, NSGA-II has no way of distinguishing a falsely high-performing action from a truly non-dominated one once observed. MOBO avoids this problem through its sampling strategy and confidence-aware surrogate model, trialing the same region many times to learn an accurate estimate of the true performance.

On ablating n , MOBO runs with $n \in \{1, 2, 3\}$ converge quickly. While MOBO $n = 4$ outperforms NSGA-II $n = 4$, its higher IGD^+ score suggests that higher action dimensions are a challenge for MOBO. Lastly, MOBO shows promising

performance for all $m \in \{2, 3, 4, 5\}$ objective cases. In contrast, NSGA-II begins to plateau at higher IGD^+ values as the number of objectives increases. Overall, these results underscore the regime in which MOBO outperforms evolutionary methods, like NSGA-II: high noise, small action spaces, and limited sample budgets.

IV. HUMAN-STUDY EXPERIMENT

We construct a real-world human subject study protocol for MO-HILBO and demonstrate optimization of three exoskeleton control parameters across two objectives: metabolic cost (efficiency) and ordinal human feedback (comfort). Moreover, we formulate and carry out a new validation procedure specifically tailored towards multi-objective HILO protocols.

Using the results from the synthetic study (Fig. 3), we determined that for $n = 3$ and $m = 2$, $T = 24$ queries is sufficient to yield informative results while remaining feasible for subjects to complete without fatiguing. To combat high noise observed in the pilot study and a smaller sample budget than ideal, we set the prior noise (in standardized units) to $\sigma_{n,j} = 0.5 \forall j$. We also set $T_0 = 3, \ell_{\min} = 0.1$.

A. Hardware

We utilize a modified autonomous hip/knee exoskeleton originally developed by X, the Moonshot Factory described in Molinaro and Scherpereel et al. The sensing layer of the device consists of IMUs (3DM-GX5-25, HKB MicroStrain, Williston, VT, USA) embedded in the skin-tight force-transmission leggings, pressure-sensing insoles and foot IMUs placed inside and on the user's shoes (XSENSOR X4, XSENSOR, Calgary, AB, Canada). As described in prior work, the wearable sensor data is mapped to the user's biological joint moment using a temporal convolutional neural network (TCN) for use as the basis of device control [9], [42]. A single board Linux computer (NVIDIA Jetson Orin Nano, NVIDIA, Santa Clara, CA, USA) manages data acquisition from the sensors, TCN model inference, and communication with the device's four actuators (AK80-9, CubeMars, Jiangxi Province, China). Interfacing with the exoskeleton is handled through an offboard laptop, and participant comfort feedback is collected on an offboard tablet. Both offboard devices are connected to a WiFi network hosted by the onboard microprocessor. The device and mechatronics are pictured in Fig. 2.

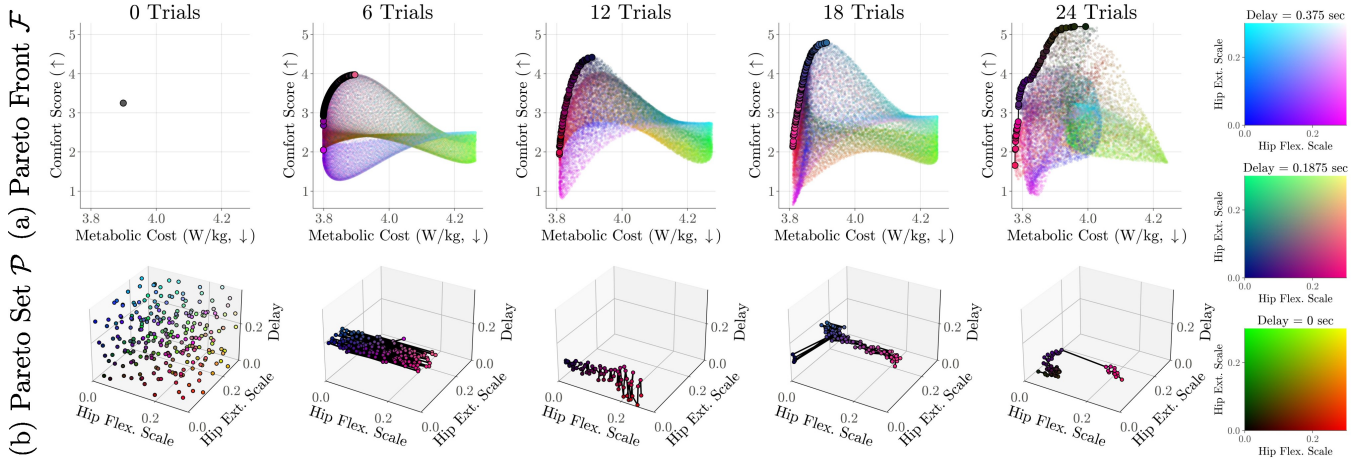


Fig. 4: Subject MB03’s Pareto set progression throughout the optimization procedure. The Pareto set begins as the entire action space, and eventually clusters around low delay, low hip extension scale. For MB03, hip flexion scale was the primary decision variable for comfort versus efficiency. In fact, MB03’s most comfortable gait was when the exoskeleton was turned off, while their most efficient controller was when the hip flexion torque scale was nearly maximized. In short, by performing MO-HILBO, we are able to reduce the search space of controllers from the entire action box, visualized on the bottom left, to a simpler, single-dimension Pareto set, on the bottom right.

B. Exoskeleton Controller Parametrization

The exoskeleton controller is parametrized by:

$$\mathbf{a} = [\beta_1, \beta_2, \delta]^T \in \mathbb{R}^3, \quad (21)$$

with $\delta \in [0, 0.375]$ being the temporal delay (in seconds) from the estimated biological joint moment $\tau_{bio}(t - \delta)$ obtained from the TCN, and $\{\beta_1, \beta_2\} \in [0, 0.3]^2$ denoting the scaling coefficients applied during positive and negative biological moment phases, respectively. Given these parameters, the commanded exoskeleton torque is computed as:

$$u(t | \mathbf{a}) = \begin{cases} \beta_1 \tau_{bio}(t - \delta), & \tau_{bio}(t - \delta) \geq 0, \\ \beta_2 \tau_{bio}(t - \delta), & \tau_{bio}(t - \delta) < 0. \end{cases} \quad (22)$$

For our study, assistance is applied at the hip while the knee actuators were set to provide zero torque. This reduces the parameter space and therefore the human-subject experiment duration for this initial study of MO-HILBO.

C. Experimental Design

Three participants (two males, one female; age: 24.7 ± 2.3 ; body mass 75.3 ± 14 kg; one experienced with the device, two who had not worn the device before) participated in a single-session protocol consisting of MO-HILBO optimization and validation (3-4 hours per subject). Informed consent was obtained from each participant under Georgia Institute of Technology Review Board protocol H21184.

Metabolic cost and user comfort were collected as participants walked on a treadmill at 1.2 m/s wearing the hip-knee exoskeleton device. Each optimization iteration lasted 3 minutes. Metabolic cost was estimated by fitting a first-order dynamical model to transient metabolic data calculated from oxygen intake and carbon dioxide exhaust measured from each breath (TrueOne 2400, ParvoMedics) using the modified Brockway equation [3], [43]. User comfort was measured using a tablet (Apple iPad Air) which queried the user of their comfort level on a 1-5 scale (not comfortable to very

Subject	Hypervolume		Pairwise Ordering	
	Pareto	Anti-Pareto	Metabolic Cost	Comfort
MB01	4.02	1.300	100% (3/3)	100% (3/3)
MB02	1.55	0.366	100% (3/3)	100% (3/3)
MB03	1.02	0.128	67% (2/3)	100% (3/3)

TABLE I: Validation data across subjects. In all cases, the identified Pareto sets were validated to Pareto-dominate the *anti-Pareto* set (built by flipping the sign of each objective). MO-HILBO also correctly predicted 94% (17/18) of the pairwise orderings of Pareto-optimal points. A valid reference point was chosen for each subject’s hypervolume computation, resulting in different scales.

comfortable) once per minute. We utilized MO-HILBO’s decoupled feedback capability to enable the user to answer just one query, or use the remaining two to change their mind, as pilot studies indicated that subjects often changed their minds regarding comfort over the duration of a trial.

D. Validation Protocol

To evaluate the predictive capability of the learned surrogate model, we validate both the Pareto dominance of the inferred Pareto set and the relative pairwise ranking of its actions. Specifically, we sample three actions from the inferred Pareto set $\mathcal{P}_{\hat{f}}$, with two on the endpoints of the front and one in between. We similarly sample three actions from the *anti-Pareto set* $\mathcal{P}_{-\hat{f}}$ (found by negating the objectives). These six controllers are evaluated for 5 minutes each by the user in a randomized order during treadmill walking at 1.2 m/s while collecting metabolic cost and user comfort.

Using these validation measurements, we compute two tests: (1) a 3-point estimate of the true hypervolume for both sets to confirm Pareto dominance, and (2) the surrogate’s accuracy in predicting the pairwise ordering among the three sampled Pareto actions. For each objective, the surrogate model predicts relative rankings across the three pairwise combinations of these actions. We then compare these predictions against the true measured rankings, yielding an accuracy score based on three binary outcomes per objective.

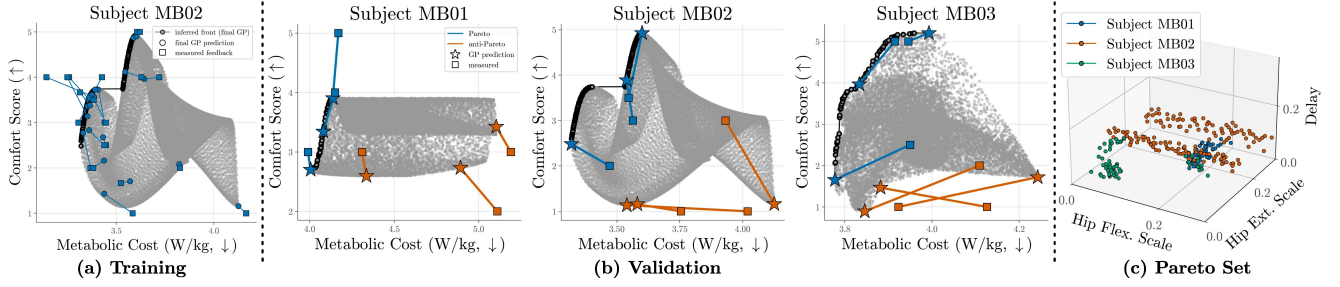


Fig. 5: Predicted versus measured points (a), (b) and discovered Pareto sets (c). (a) shows a fitted MO-GP on MB02, where the training data lie near the Pareto front—a result of MO-HILBO’s acquisition function exploring and refining the Pareto set. (b) shows the validation points (stars)—which were not seen during training—spaced out over the front. (c) shows the overlaid Pareto sets of all subjects. For the vast majority of points, each subject’s Pareto set does not overlap in action space, except for part of MB03’s and MB01’s. While the subjects’ Pareto-optimal controllers featured a variety of flexion or extension torques, none had a high delay.

V. RESULTS AND DISCUSSION

Our results are shown in Fig. 4 and Fig. 5. For all experiments, we compute $\mathcal{P}_{\hat{f}}$ by evaluating the MO-GP mean on a dense sample of the action space. Fig. 4 showcases a single optimization procedure for subject MB03. MO-HILBO began by considering all controllers $\mathbf{a} \in \mathcal{A}$ as Pareto-optimal (bottom left), and iteratively inferred more selective Pareto sets over decoupled feedback updates (bottom right). In addition to identifying the surrogate Pareto set $\mathcal{P}_{\hat{f}}$, MO-HILBO also learns a posterior over all $\mathbf{a} \in \mathcal{A}$. This is insightful for understanding what the subject *didn’t* like. For instance, according to Fig. 4, MB03 found controllers with high torque and delay uncomfortable, and ones with high delay and flexion torque as inefficient.

A secondary goal of this study was to discover if the Pareto sets shared structure across subjects. Qualitatively, the answer is mixed. All subjects found high delay to be neither comfortable nor efficient. However, MB02 and MB03 found low torque as most comfortable, while MB01 found high flexion and extension torque as both comfortable and efficient. Of interest is that MB02 and MB03 had no prior experience in the exoskeleton device, while MB01 did have prior experience. This suggests that there could be non-stationary factors which influence the location of $\mathcal{P}_{\hat{f}}$ in addition to other subject characteristics, like physical embodiment and gait style. This is consistent with findings from prior work that demonstrate substantial improvements in metabolic outcomes with training [24]. Thus, an interesting area of future work would be to investigate how these Pareto fronts evolve over extended user exposure and adaptation.

Fig. 5 and Table I summarize the experimental validation results. Across all subjects, the controllers in the surrogate Pareto set $\mathcal{P}_{\hat{f}}$ identified by MO-HILBO empirically dominated those from the anti-Pareto set $\mathcal{P}_{-\hat{f}}$. The surrogate also accurately predicted the relative pairwise ordering across nearly all validation Pareto actions for both objectives. These were novel controllers not seen during training, indicating that MO-HILBO successfully learns a continuous, probabilistic surrogate of the true Pareto set $\mathcal{P}_f$. Effectively, MO-HILBO mapped the 3-dimensional action space into a 1-dimensional manifold of controllers that optimally trade off among the objectives. This reduced representation offers a

significantly simplified search space for real-time human-in-the-loop tuning, an exciting future research direction.

VI. CONCLUSION

This paper introduces the first application of Multi-Objective Human-in-the-Loop Bayesian Optimization (MO-HILBO) for wearable assistive devices. By leveraging probabilistic surrogate models, MO-HILBO improves sample-efficiency compared to prior multi-objective HILO techniques. Through both synthetic benchmarking and lower-limb exoskeleton trials across metabolic cost and user comfort, we demonstrate that MO-HILBO efficiently maps a continuous Pareto set within a practical trial budget ($T = 24$). Further, we construct a validation procedure for MO-HILBO and use it to confirm that MO-HILBO reliably identifies Pareto-dominant controllers and accurately predicts relative pairwise rankings along the learned Pareto front. To support broader adoption, we open-source `pypolar`¹, a Python package for deploying sample-efficient HILO and MO-HILBO algorithms on wearable systems.

There are several exciting future directions. First, standard Bayesian optimization assumes a stationary objective, which may not capture human adaptation, fatigue, or shifting preferences; future work could model these dynamics as distinct regions of the Pareto set. Second, in cases with many objectives, interactive search algorithms may be required to assist users in navigating large Pareto sets, ensuring tractable, user-centered control.

VII. ACKNOWLEDGEMENTS

Models from Anthropic and OpenAI assisted in writing code for this project.

REFERENCES

- [1] E. Tricomi, F. Missiroli, M. Xiloyannis, N. Lotti, X. Zhang, M. Stefanakis, M. Theisen, J. Bauer, C. Becker, and L. Masia, “Soft robotic shorts improve outdoor walking efficiency in older adults,” *Nature Machine Intelligence*, vol. 6, no. 10, pp. 1145–1155, 2024.
- [2] S. Seungmoon, R. Ricardo, M. P O’Donovan *et al.*, “Optimized hip-knee-ankle exoskeleton assistance at a range of walking speeds,” *Journal of NeuroEngineering and Rehabilitation (Web)*, vol. 18, no. 1, pp. 1–12, 2021.
- [3] 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.

- [4] P. Slade, M. J. Kochenderfer, S. L. Delp, and S. H. Collins, "Personalizing exoskeleton assistance while walking in the real world," *Nature*, vol. 610, no. 7931, pp. 277–282, 2022.
- [5] P. R. Shetty, J. A. Menezes, S. Song, A. J. Young, and M. K. Shepherd, "Ankle exoskeleton control via data-driven gait estimation for walking, running, and inclines," *IEEE Robotics and Automation Letters*, vol. 10, no. 6, pp. 5855–5862, 2025.
- [6] D. Park, J. An, D. Lee, I. Kang, and A. J. Young, "Human-in-the-loop optimization of hip exoskeleton assistance during stair climbing," *IEEE Transactions on Biomedical Engineering*, vol. 72, no. 7, pp. 2147–2156, 2025.
- [7] P. Kantharaju, H. Jeong, S. Ramadurai, M. Jacobson, H. Jeong, and M. Kim, "Reducing squat physical effort using personalized assistance from an ankle exoskeleton," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 1786–1795, 2022.
- [8] W. Huo, H. Moon, M. A. Alouane, V. Bonnet, J. Huang, Y. Amirat, R. Vaidyanathan, and S. Mohammed, "Impedance modulation control of a lower-limb exoskeleton to assist sit-to-stand movements," *IEEE Transactions on Robotics*, vol. 38, no. 2, pp. 1230–1249, 2021.
- [9] D. D. Molinaro, K. L. Scherperceel, E. B. Schonhaut, G. Evangelopoulos, M. K. Shepherd, and A. J. Young, "Task-agnostic exoskeleton control via biological joint moment estimation," *Nature*, vol. 635, no. 8038, pp. 337–344, 2024.
- [10] M. Tucker, M. Cheng, E. Novoseller, R. Cheng, Y. Yue, J. W. Burdick, and A. D. Ames, "Human preference-based learning for high-dimensional optimization of exoskeleton walking gaits," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 3423–3430.
- [11] M. Tucker, E. Novoseller, C. Kann, Y. Sui, Y. Yue, J. W. Burdick, and A. D. Ames, "Preference-based learning for exoskeleton gait optimization," in *2020 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2020, pp. 2351–2357.
- [12] K. A. Ingraham, C. D. Remy, and E. J. Rouse, "The role of user preference in the customized control of robotic exoskeletons," *Science robotics*, vol. 7, no. 64, p. eabj3487, 2022.
- [13] K. A. Ingraham, M. Tucker, A. D. Ames, E. J. Rouse, and M. K. Shepherd, "Leveraging user preference in the design and evaluation of lower-limb exoskeletons and prostheses," *Current Opinion in Biomedical Engineering*, vol. 28, p. 100487, 2023.
- [14] U. H. Lee, V. S. Shetty, P. W. Franks, J. Tan, G. Evangelopoulos, S. Ha, and E. J. Rouse, "User preference optimization for control of ankle exoskeletons using sample efficient active learning," *Science Robotics*, vol. 8, no. 83, p. eadg3705, 2023.
- [15] P. Arens, D. A. Quirk, W. Pan, Y. Yacoby, F. Doshi-Velez, and C. J. Walsh, "Preference-based assistance optimization for lifting and lowering with a soft back exosuit," *Science Advances*, vol. 11, no. 15, p. eadu2099, 2025.
- [16] P. Slade, C. Atkeson, J. M. Donelan, H. Houdijk, K. A. Ingraham, M. Kim, K. Kong, K. L. Poggensee, R. Riener, M. Steinert *et al.*, "On human-in-the-loop optimization of human-robot interaction," *Nature*, vol. 633, no. 8031, pp. 779–788, 2024.
- [17] 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.
- [18] 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 *Robot ics: Science and Systems*, vol. 2016, 2016, pp. 1–10.
- [19] M. Kim, Y. Ding, P. Malcolm, J. Speeckaert, C. J. Sivi, 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.
- [20] 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.
- [21] M. Kim, C. Liu, J. Kim, S. Lee, A. Meguid, C. J. Walsh, and S. Kuindersma, "Bayesian optimization of soft exosuits using a metabolic estimator stopping process," in *2019 international conference on robotics and automation (icra)*. IEEE, 2019, pp. 9173–9179.
- [22] F. Haufe, P. Wolf, and R. Riener, "Human-in-the-loop optimization of a multi-joint wearable robot for movement assistance," *Proceedings on Automation in Medical Engineering*, vol. 1, no. 1, p. 023, 2020.
- [23] 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, p. eaay9108, 2020.
- [24] K. L. Poggensee and S. H. Collins, "How adaptation, training, and customization contribute to benefits from exoskeleton assistance," *Science Robotics*, vol. 6, no. 58, p. eabf1078, 2021.
- [25] G. M. Bryan, P. W. Franks, S. C. Klein, R. J. Peuchen, and S. H. Collins, "A hip-knee-ankle exoskeleton emulator for studying gait assistance," *The International Journal of Robotics Research*, vol. 40, no. 4-5, pp. 722–746, 2021.
- [26] P. W. Franks, G. M. Bryan, R. M. Martin, R. Reyes, A. C. Lakmazaheri, and S. H. Collins, "Comparing optimized exoskeleton assistance of the hip, knee, and ankle in single and multi-joint configurations," *Wearable Technologies*, vol. 2, p. e16, 2021.
- [27] D. F. Gordon, C. McGreavy, A. Christou, and S. Vijayakumar, "Human-in-the-loop optimization of exoskeleton assistance via online simulation of metabolic cost," *IEEE Transactions on Robotics*, vol. 38, no. 3, pp. 1410–1429, 2022.
- [28] J. Kim, B. T. Quinlivan, L.-A. Deprey, D. Arumukhom Revi, A. Eckert-Erdheim, P. Murphy, D. Orzel, and C. J. Walsh, "Reducing the energy cost of walking with low assistance levels through optimized hip flexion assistance from a soft exosuit," *Scientific reports*, vol. 12, no. 1, p. 11004, 2022.
- [29] P. Kantharaju, S. S. Vakacherla, M. Jacobson, H. Jeong, M. N. Mevada, X. Zhou, M. J. Major, and M. Kim, "Framework for personalizing wearable devices using real-time physiological measures," *IEEE Access*, vol. 11, pp. 81 389–81 400, 2023.
- [30] J. C. Powell, E. B. Schonhaut, D. D. Molinaro, and A. J. Young, "Optimized mappings from biological hip moment estimates to exoskeleton torque can personalize assistance across users and generalize across tasks," *bioRxiv*, pp. 2025–08, 2025.
- [31] A. J. Young, J. Foss, H. Gannon, and D. P. Ferris, "Influence of power delivery timing on the energetics and biomechanics of humans wearing a hip exoskeleton," *Frontiers in bioengineering and biotechnology*, vol. 5, p. 4, 2017.
- [32] A. Lakmazaheri, S. Song, B. Vuong, B. Biskner, D. Kado, and S. Collins, "Optimizing exoskeleton assistance to improve walking speed and energy economy for older adults," *Journal of NeuroEngineering and Rehabilitation*, vol. 21, 01 2024.
- [33] X. Zhang, Y.-X. Liu, R. Wang, and E. M. Gutierrez-Farewik, "Multi-objective human-in-the-loop optimization of exoskeleton assistance for dropfoot gait," *IEEE Robotics and Automation Letters*, vol. 10, no. 8, pp. 8586–8593, 2025.
- [34] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.
- [35] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006, vol. 2, no. 3.
- [36] M. Balandat, B. Karrer, D. Jiang, S. Daulton, B. Letham, A. G. Wilson, and E. Bakshy, "Botorch: A framework for efficient monte-carlo bayesian optimization," *Advances in neural information processing systems*, vol. 33, pp. 21 524–21 538, 2020.
- [37] J. Knowles, "Parego: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 1, pp. 50–66, January 2006.
- [38] S. Daulton, M. Balandat, and E. Bakshy, "Differentiable expected hypervolume improvement for parallel multi-objective bayesian optimization," *Advances in neural information processing systems*, vol. 33, pp. 9851–9864, 2020.
- [39] S. Ament, S. Daulton, D. Eriksson, M. Balandat, and E. Bakshy, "Unexpected improvements to expected improvement for bayesian optimization," *Advances in neural information processing systems*, vol. 36, pp. 20 577–20 612, 2023.
- [40] H. Ishibuchi, H. Masuda, Y. Tanigaki, and Y. Nojima, "Modified distance calculation in generational distance and inverted generational distance," in *Evolutionary Multi-Criterion Optimization*, A. Gaspar-Cunha, C. Henggeler Antunes, and C. C. Coello, Eds. Cham: Springer International Publishing, 2015, pp. 110–125.
- [41] J. Blank and K. Deb, "pymoo: Multi-objective optimization in python," *IEEE Access*, vol. 8, pp. 89 497–89 509, 2020.
- [42] D. D. Molinaro, I. Kang, and A. J. Young, "Estimating human joint moments unifies exoskeleton control, reducing user effort," *Science robotics*, vol. 9, no. 88, p. eadi8852, 2024.
- [43] J. M. Brockway, "Derivation of formulae used to calculate energy expenditure in man," *Human Nutrition. Clinical Nutrition*, vol. 41, no. 6, pp. 463–471, 1987.