

Safe Bayesian Optimization for the Control of High-Dimensional Embodied Systems

Yunyue Wei, Zeji Yi, Hongda Li, Saraswati Soedarmadji, Yanan Sui

Tsinghua University

weiy20@mails.tsinghua.edu.cn

zejy@andrew.cmu.edu

lhd21@mails.tsinghua.edu.cn

chenxuying24@mails.tsinghua.edu.cn

ysui@tsinghua.edu.cn

Abstract: Learning to move is a primary goal for animals and robots, where ensuring safety is often important when optimizing control policies on the embodied systems. For complex tasks such as the control of human or humanoid control, the high-dimensional parameter space adds complexity to the safe optimization effort. Current safe exploration algorithms exhibit inefficiency and may even become infeasible with large high-dimensional input spaces. Furthermore, existing high-dimensional constrained optimization methods neglect safety in the search process. In this paper, we propose High-dimensional Safe Bayesian Optimization with local optimistic exploration (HDSAFEBO), a novel approach designed to handle high-dimensional sampling problems under probabilistic safety constraints. We introduce a local optimistic strategy to efficiently and safely optimize the objective function, providing a probabilistic safety guarantee and a cumulative safety violation bound. Through the use of isometric embedding, HDSAFEBO addresses problems ranging from a few hundred to several thousand dimensions while maintaining safety guarantees. To our knowledge, HDSAFEBO is the first algorithm capable of optimizing the control of high-dimensional musculoskeletal systems with high safety probability. We also demonstrate the real-world applicability of HDSAFEBO through its use in the safe online optimization of neural stimulation induced human motion control.

Keywords: Safe Bayesian Optimization, High-dimensional Embodied System

1 Introduction

Some robotics applications require online optimization of control policies for performance while avoiding unsafe parameter tuning that could potentially damage the systems. These scenarios correspond to the problem of *safe exploration*, which involves the sequential optimization of an unknown objective function under the constraint of satisfying unknown safety conditions. Bayesian optimization (BO) is an effective paradigm in optimizing black-box functions, and safe BO methods have been successfully used to tune control parameters for various robotic systems [1, 2, 3]. Most existing safe BO methods use the Gaussian process (GP) to model the underlying safety function, and discriminate safe regions with estimated function’s lower confidence bound to ensure safety with high probability [4]. Such conservative strategies are inefficient for objective optimization, and even infeasible in high-dimensional input settings, such as human or humanoid system control.

A motivating application of our work is the control of musculoskeletal (tendon-driven) systems, where complex motions are driven by dozens to hundreds of muscle-tendon units rather than joints. Such overactuated embodied systems introduce additional control challenges within a large-scale parameter space. Under this high-dimensional input setting, efficiently optimizing the task function while

maintaining safety remains a formidable challenge for existing safe optimization algorithms. Despite considerable efforts to leverage BO for solving high-dimensional constrained optimization problems, these methods often fail to incorporate safety considerations into the optimization procedure [5, 6, 7]. To our best knowledge, there are currently no methods that guarantee safety, or probabilistic safety during the high-dimensional optimization.

In this paper, we introduce High-dimensional Safe Bayesian Optimization with local optimistic exploration (HDSAFEBO) to address probabilistic safety in high-dimensional sequential optimization problems, as shown in Figure 1. To achieve efficient and safe optimization, we propose a *local optimistic strategy* with probabilistic safety guarantee and cumulative safety violation bound. By using isometric embedding for dimension reduction, we enable HDSAFEBO to handle even higher dimensional problems, ranging from a few hundred to several thousand dimensions, while maintaining the probabilistic safety guarantee. Our experimental results show that HDSAFEBO efficiently learns to control a musculoskeletal system with high safety probability - a task where all baseline methods fail. We also demonstrate the success of HDSAFEBO in real-world experiments to safely optimize the control of neural stimulation induced human motion. Our project page is at <https://linsgroup.cc/research/hdsafebo>

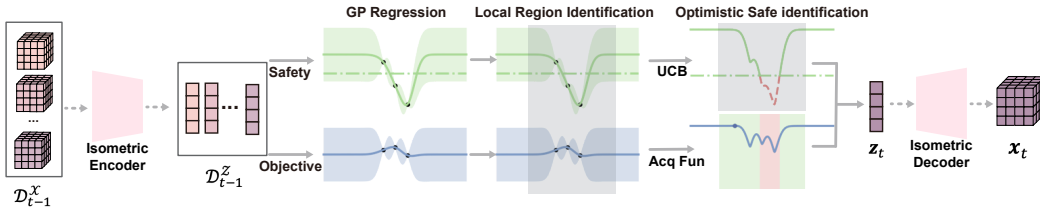


Figure 1: **Workflow of HDSAFEBO.** Local optimistic safe optimization is employed to efficiently optimize the objective function while guaranteeing probabilistic safety. We utilize isometric embedding to reduce the problem dimension to deal with high-dimensional inputs.

2 Related Work

2.1 Safe Bayesian Optimization

The sequential decision-making problem with safety constraints has been extensively studied, varied by the definition of safety. To achieve full safety during exploratory sampling, algorithms have been proposed with theoretical guarantee in near-optimality and safety with high probability [4, 8, 1, 9, 10]. These methods have been applied in safe parameter tuning of quadrotor [11], robot arm [2] and quadruped robots [3]. These safe optimization algorithms conservatively estimate and expand the safe region, leading to inefficient optimization performance.

In contrast to a zero-tolerance approach to unsafe actions, an alternative approach allows for limited constraint violations within a predefined budget, trading safety for more efficient optimization [12, 13, 14]. A recent work CONFIG uses upper confidence bound to optimistically estimate the safe region, enjoying global optimal guarantee as unconstrained methods [15].

Another extreme case, called constrained Bayesian optimization, has also been used for finding safe controller parameters [16, 17, 18, 19, 20]. However, this line of methods aim only to find the best feasible solution, neglecting the safety during the optimization process. Constrained expected improvement (cEI) is a popular constrained BO algorithm that introduces feasibility constraints to acquisition function formulation [21, 16].

All of the aforementioned methods fall in the framework of Bayesian optimization, which is typically limited to low-dimensional problems [22, 23]. LineBO demonstrates success in optimizing problems with dimension up to 40 via searching over 1-dimensional subspace at each iteration [24, 25]. When parameter size exceeds 10, safe optimization algorithms might even struggle to expand the safe region due to sparse discretization of the input space.

2.2 High-dimensional Bayesian Optimization

Recent research efforts have delved into the utilization of Bayesian optimization in high-dimensional problems [26, 27]. In the main paper, our focus lies on dimension reduction-based and local BO methods. Further discussion about other relevant works can be found in Appendix D.

A large body of literature leverages dimension reduction to apply BO over a low-dimensional subspace. Several works use variable selection to identify and optimize important dimensions during optimization [28, 29, 30]. Another popular approach for reducing the search space is random linear embedding, which has been proven to contain the optimal solution with a certain probability relative to the objective function’s effective dimension[31, 32, 33, 34, 35]. Many works also use autoencoder to learn a non-linear mapping between the original space and the latent space [36, 37, 38, 39, 40, 41, 6].

Another line of works utilizes local search to address the over-exploration issue in high-dimensional optimization [42], achieving better empirical performance than global BO methods [43, 44, 45, 46]. This local optimization strategy is also able to be combined with dimension reduction [47] and extended to constrained setting (SCBO [7]). Local BO methods have also been deployed for optimizing locomotion control in simulation [48, 30].

Besides BO methods, evolutionary algorithms such as CMA-ES are competitive to solve high-dimensional problems [49] and can be extended to constrained setting[50, 51]. Although many works attempt to solve high-dimensional constrained optimization problems, to the best of our knowledge, there lacks work that addresses safety in high-dimensional sequential optimization with at least hundreds of variables.

3 Problem Formulation

We aim to optimize an unknown objective function $f : \mathcal{X} \rightarrow \mathbb{R}$ by sequentially sampling points $\mathbf{x}_1, \dots, \mathbf{x}_t \in \mathcal{X}$. We can also get observations of safety measurement from another unknown function $g : \mathcal{X} \rightarrow \mathbb{R}$. We define a point \mathbf{x} is safe when $g(\mathbf{x}) > 0$. We can formally write our optimization problem as follows:

$$\max_{\mathbf{x}_t \in \mathcal{X}} f(\mathbf{x}_t) \quad \text{subject to } g(\mathbf{x}_t) \geq 0, \forall t \geq 1, \quad (1)$$

We consider the above problem formulation widely exists in the application of robotic control, where \mathbf{x} represents a parameterized controller, and $f(\mathbf{x}), g(\mathbf{x})$ are the utility and safety measurement of the controller in a single experimental trial. In our target applications such as the control of embodied systems, the input space $\mathcal{X} \in \mathbb{R}^D$ is high-dimensional, encompassing dozens to thousands of variables.

In Bayesian optimization, Gaussian process is usually used as the surrogate model to learn the unknown functions. Taking g as an example, for samples at points $X_t = [\mathbf{x}_1 \dots \mathbf{x}_t]^T$, we have noise-perturbed observations $\mathbf{y}_t = [y_1^g \dots y_t^g]^T$. The GP posterior over g is also Gaussian with mean $\mu_t(\mathbf{x})$, covariance $k_t(\mathbf{x}, \mathbf{x}')$ and variance $\sigma_t^2(\mathbf{x}, \mathbf{x}')$ under kernel function k :

$$\begin{aligned} \mu_t(\mathbf{x}) &= \mathbf{k}_t(\mathbf{x})^T (\mathbf{K}_t + \sigma^2 \mathbf{I})^{-1} \mathbf{y}_t \\ k_t(\mathbf{x}, \mathbf{x}') &= k(\mathbf{x}, \mathbf{x}') - \mathbf{k}_t(\mathbf{x})^T (\mathbf{K}_t + \sigma^2 \mathbf{I})^{-1} \mathbf{k}_t(\mathbf{x}') \\ \sigma_t^2(\mathbf{x}) &= k_t(\mathbf{x}, \mathbf{x}), \end{aligned} \quad (2)$$

where $\mathbf{k}_t(\mathbf{x}) = [k(\mathbf{x}_1, \mathbf{x}), \dots, k(\mathbf{x}_t, \mathbf{x})]$ is the covariance between \mathbf{x} and sampled points, \mathbf{K}_t is the covariance of sampled positions: $[k(\mathbf{x}, \mathbf{x}')]_{\mathbf{x}, \mathbf{x}' \in X_t}$. Similarly we can use GP to derive posterior of f . Using the posterior of GP, we can define $u_t(\mathbf{x}) := \mu_{t-1}(\mathbf{x}) + \beta_t \sigma_{t-1}(\mathbf{x})$ as the upper confidence bound (UCB) of the function estimation, where β_t is a scalar which can be properly set to contain $g(\mathbf{x})$ with desired probability. we make the following regularity assumptions that are commonly used in the field of Bayesian optimization:

Assumption 3.1. g and f are samples of two Gaussian processes defined by the kernels $k(\cdot, \cdot), k'(\cdot, \cdot)$ respectively. The observations are perturbed by i.i.d. Gaussian noise: $y^g(\mathbf{x}_t) = g(\mathbf{x}_t) + n_t$, $y^f(\mathbf{x}_t) = f(\mathbf{x}_t) + n'_t$ where $n_t \sim \mathcal{N}(0, \sigma^2), n'_t \sim \mathcal{N}(0, \sigma'^2)$

Algorithm 1 HDSAFEBO

Input Sample set \mathcal{X} , GP priors GP^f, GP^g , safety threshold h , acquisition function A , initial dataset $\mathcal{D}_0^{\mathcal{X}}$, trust region length l_0 , isometric encoder Π and decoder Π^{-1}

- 1: $c_s, c_f \leftarrow 0$
- 2: $\mathbf{x}^*, y^* \leftarrow \operatorname{argmax}_{\{\mathbf{x}, y^f\} \in \mathcal{D}_0, y^g(\mathbf{x}) > 0} y^f(\mathbf{x})$
- 3: **for** $t = 1$ to \dots **do**
- 4: $D_{t-1}^{\mathcal{Z}} \leftarrow \Pi(D_{t-1}^{\mathcal{X}})$
- 5: Update GP^f, GP^g using $\mathcal{D}_{t-1}^{\mathcal{Z}}$
- 6: $L_t \leftarrow \{\mathbf{z} \in \mathcal{Z} \mid \Pi(\mathbf{x}^*) - l_t \leq \mathbf{z} \leq \Pi(\mathbf{x}^*) + l_t\}$
- 7: $S_t \leftarrow \{\mathbf{z}' \in L_t \mid \mu_{t-1}(\mathbf{z}') + \beta_t \sigma_{t-1}(\mathbf{z}') \geq 0\}$
- 8: $\mathbf{z}_t \leftarrow \operatorname{argmax}_{\mathbf{z} \in S_t} (A(\mathbf{z}))$
- 9: $\mathbf{x}_t \leftarrow \Pi^{-1}(\mathbf{z}_t)$
- 10: $y_t^f \leftarrow f(\mathbf{x}_t) + n_t$
- 11: $y_t^g \leftarrow g(\mathbf{x}_t) + n_t$
- 12: $\mathcal{D}_t \leftarrow \mathcal{D}_{t-1} \cup \{\mathbf{x}_t, y_t^f, y_t^g\}$
- 13: $\mathbf{x}^*, y^*, c_s, c_f, l_t \leftarrow$
 $\text{SAFEUPDATE}(\mathbf{x}_t, y_t^f, \mathbf{x}^*, y^*, c_s, c_f, l_t)$
- 14: **end for**

Algorithm 2 SAFEUPDATE

Input Current sample \mathbf{x}_t, y_t^f , current best sample \mathbf{x}^*, y^* , current counters c_s, c_t current trust region length l_t , success tolerance τ_s , failure tolerance τ_f , initial trust region length l_0 , trust region limits l_{\max}, l_{\min}

- 1: **if** $y_t^g(\mathbf{x}_t) > 0$ and $y_t^f > y^*$ **then**
- 2: $\mathbf{x}^*, y^* \leftarrow \mathbf{x}_t, y_t^f; c_s \leftarrow c_s + 1; c_f \leftarrow 0$
- 3: **else**
- 4: $c_s \leftarrow 0; c_f \leftarrow c_f + 1$
- 5: **end if**
- 6: **if** $c_s = \tau_s$ **then**
- 7: $l_t \leftarrow \min(2l_{t-1}, l_{\max}); c_s, c_f \leftarrow 0$
- 8: **else if** $c_f = \tau_f$ **then**
- 9: $l_t \leftarrow \max(\frac{1}{2}l_{t-1}, l_{\min}); c_s, c_f \leftarrow 0$
- 10: **if** $l_t = l_{\min}$ **then**
- 11: $l_t \leftarrow l_0$
- 12: **end if**
- 13: **else**
- 14: $l_t \leftarrow l_{t-1}$
- 15: **end if**

4 High-dimensional Safe Bayesian Optimization

In high-dimensional space, existing safe optimization methods are too conservative to efficiently optimize the objective function, and would be infeasible due to sparse discretization. Therefore we aim to improve sample efficiency by slightly relaxing the safety constraint to a probabilistic version. The probabilistic safety means each sample point is safe with a probability above predefined threshold α , that is $\Pr(g(\mathbf{x}_t) \geq 0) \geq \alpha, \forall t \leq T$. In this sense, we introduce HDSAFEBO, an innovative algorithm designed for ensuring probabilistic safety while optimizing in a high-dimensional space. The workflow of this algorithm is illustrate in Figure 1 and Algorithm 1.

HDSAFEBO first uses the isometric encoder Π to reduce the problem dimension, converting the original dataset from \mathcal{X} to the low-dimensional latent space $\mathcal{Z} \in \mathbb{R}^d$ (Line 4). Isometry means the embedded subspace is able to preserve the distance of the original space according to the corresponding metric $d_{\mathcal{X}}, d_{\mathcal{Z}}$, i.e. $d_{\mathcal{X}}(\mathbf{x}, \mathbf{x}') = d_{\mathcal{Z}}(\Pi(\mathbf{x}), \Pi(\mathbf{x}'))$. Leveraging the historical record of function observations, we compute the posterior and confidence interval of both the objective and safety functions through distinct Gaussian processes. (Line 5). Then we define a local region to search over (Line 6), and identify the safe space within the local region using GP upper confidence bound (Line 7). We optimize the acquisition function over the safe space, and project the recommended latent point back to original input space using the decoder Π^{-1} (line 8-9). Thompson sampling (TS) [52] was selected as the acquisition function A due to its compatibility with the discrete nature of our safety estimation and search space, and its innate ability for batch optimization by sampling the GP posterior—an appropriate choice for high-dimensional tasks that support parallel evaluations. Finally the history data is updated via evaluating new inputs. (Line 10-12), and the local region parameters are updated based on the sample results (Line 13).

In conjunction with isometric embedding, we highlight two important components to improve the efficiency and safety: **Optimistic Safety Identification** and **Local Search via Trust Region**. For clear illustration, we first introduce algorithm details under identity mapping $\mathcal{I}(\mathbf{x}) = \mathbf{x}$, which is a special case of isometric embedding. In this way, the embedded subspace is equivalent to the original space \mathcal{X} , and inherits all the assumptions defined in Section 3.

4.1 Optimistic safety identification

To mitigate inefficiency concerns, HDSAFEBO distinguishes the safe region through the Gaussian process upper confidence bound of the safety function. This optimistic strategy allows for optimization over distinct safe regions and is viable in high-dimensional problems. In many real-world applications, the search space is inherently constrained by domain priors, where a certain proportion of decisions are safe even under random search. Therefore, we contend that optimistic safety identification is a more practical strategy. Here we derive the appropriate choice of the scalar β_t to ensure step-wise probabilistic safety.

Proposition 4.1. *Let Assumptions 3.1 holds for the latent safety function g , and set β_t satisfying $\Phi(\beta_t) \leq 1 - \alpha$. Then at every time step t :*

$$\Pr(g(\mathbf{x}) \geq \mu_{t-1}(\mathbf{x}) + \beta_t \sigma_{t-1}(\mathbf{x})) \geq \alpha, \forall \mathbf{x} \in \mathcal{X}, \quad (3)$$

where $\Phi(\cdot)$ is the cumulative distribution function (CDF) of the standard normal distribution $\mathcal{N}(0, 1)$.

Our use of β_t here is different from the common-used setting [53], which would make $u_t(\mathbf{x}_t) > g(\mathbf{x})$ with a high probability and violate the probabilistic requirement. Our results indicate that the upper confidence bound is permissible when the probability threshold is less than 0.5, representing the maximum acceptable level of safety violation. Otherwise, the algorithm reverts to a more conservative strategy, utilizing the lower confidence bound to identify safety. We also derive the upper bound of cumulative safety violation $V_T = \sum_{t=1}^T \max(0, -g(\mathbf{x}_t))$ of our optimistic strategy, where the proper choice of β_t contributes to the reduction of safety violations:

Theorem 4.2. *Let assumptions 3.1 holds for safety function g . Define $\beta_T^c := 2\log(|\mathcal{X}|T^2\pi^2/6\delta)$ and $C_1 := 8/\log(1 + \sigma^{-2})$. With probability at least $1 - \delta$, the sample points of Algorithm 1 at time step T satisfy*

$$\mathbb{E}[V_T] \leq (1 - \alpha)\sqrt{C_1 T \beta_T^c \gamma_T}, \quad (4)$$

where γ_T is the maximum information gain for g over \mathcal{X} .

4.2 Local search via trust region

In addition to optimistic safety identification, a trust region method is employed to dynamically pinpoint local search regions, which has demonstrated impressive empirical performance over high-dimensional problems [43, 48, 7]. At each optimization round, a local search space is defined as a hypercube trust region around the current best safe point in the sample dataset. We design a safety sensitive strategy to update the trust region state, as illustrated in Algorithm 2. Specifically, a sampling round is considered “successful” if it finds a better reward while maintaining comprehensive safety (line 1-2). Conversely, it is labeled a “failure” if any unsafe points are found or if there is no discernible improvement (line 3-5). The side length is adjusted—increased for successes and decreased for failures—upon reaching a preset threshold (line 6-9). Unlike conventional local BO methods that discard all data and restart when the side length reaches its minimum, our approach resets l_t to its initial length and retains all previous samples, ensuring a different, safer trajectory sampling than the initial instance (line 10-12).

We also provide the theoretical implications of our local search strategy in reducing the safety violations. The maximum information gain γ_T is positively correlated with the size of search space \mathcal{X} . During the optimization, the actual search region of HDSAFEBO is restricted by the adaptive trust region, resulting in a lower maximum information gain compared to entire input space. Therefore the safety violation bound in Theorem 4.2 can be further reduced compared to global search.

4.3 Safe optimization with isometric embedding

In addition to identity mapping, when applying safe optimization over the subspace of other embeddings, a pertinent question arises regarding whether safety guarantees persist in the original space.

Here we demonstrate that, through the use of isometric embedding, the probabilistic safety guarantee in the embedded subspace can still be fulfilled in the original space when using stationary kernels.

Proposition 4.3. *Define $d_{\mathcal{X}}(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|$, $d_{\mathcal{Z}}(\mathbf{z}, \mathbf{z}') = \|\mathbf{z} - \mathbf{z}'\|$, and $u_t^{\mathcal{Z}}(\mathbf{z}) := \mu_{t-1}(\mathbf{z}) + \beta_t \sigma_{t-1}(\mathbf{z})$ is the upper confidence bound estimated from GP over the embedded subspace \mathcal{Z} . Suppose the GP kernel is stationary and $d_{\mathcal{Z}}(\mathbf{z}, \mathbf{z}') = d_{\mathcal{X}}(\Pi^{-1}(\mathbf{z}), \Pi^{-1}(\mathbf{z}'))$. At every time step t , if the point \mathbf{z} satisfies that $\Pr(u_t^{\mathcal{Z}}(\mathbf{z}) \geq 0) \geq \alpha$, then we have $\Pr(g(\Pi^{-1}(\mathbf{z})) \geq 0) \geq \alpha$*

The choice of the isometry embedding can be some linear mappings, such as principal component analysis (PCA) when the number of principal components exceeds the effective dimension of the problem. Recent research efforts have also explored the utilization of deep neural networks to learn a subspace with both good representation and approximated distance-preserving property. We will show in experiment that HDSAFEBO is still able to maintain high safe probability when using approximated isometric embeddings.

5 Experiments

In this section, we first utilize synthetic functions to evaluate the algorithm performance. Then we apply HDSAFEBO to safely optimize the control of high-dimensional musculoskeletal system. Finally we demonstrate the potential of HDSAFEBO in optimizing neural stimulation induced human motion control through both simulation and real experiments. Additional experiments are presented in Appendix C.

Evaluation metrics. We assess the performance of the algorithm according to three metrics: the best feasible objective function value (Objective), the safe decision ratio of all samples (Safety), and the cumulative safety violation (Violation). The presented plots and tables display the means along with one standard error.

Isometric embedding. We utilize PCA in synthetic function and musculoskeletal system control tasks. In the neural stimulation induced human motion control task, we explore the use of isometric regularized variational autoencoder (IRVAE [54]) as the dimension reduction component.

5.1 Synthetic Function Optimization

To evaluate the algorithm performance under full assumption satisfaction, we sample both objective and safety functions from Gaussian processes, with an effective dimension $d_e = 40$, which is much lower than the input dimension $D = 1000$. We contrast the optimization performance with the following competitive algorithms target for high-dimensional constrained optimization: LineBO [24], SCBO [7], CONFIG [15], cEI [21] and CMA-ES [49]. We attempted to run SafeOpt [4], but it failed to expand the safe region from the initial points. We set PCA subspace dimension as $d = 50$. We show the optimization result in Table 1. HDSAFEBO achieves significantly better optimization performance, higher safety decision ratio, and lower cumulative safety violation compared against all baselines.

Table 1: Algorithm performance on GP synthetic functions. The total sample budget is 500 including 200 initial points. We show the averaged performance over 100 independent runs.

Metric	HDSAFEBO	LineBO	SCBO	CONFIG	cEI	CMAES
Objective (\uparrow)	3.96 \pm 0.15	3.07 \pm 0.04	2.97 \pm 0.06	2.91 \pm 0.05	2.9 \pm 0.03	2.7 \pm 0.04
Safety (\uparrow)	0.81 \pm 0.02	0.78 \pm 0.0	0.77 \pm 0.0	0.77 \pm 0.0	0.77 \pm 0.0	0.78 \pm 0.01
Violation (\downarrow)	27.42 \pm 4.01	36.59 \pm 0.82	39.05 \pm 0.92	38.96 \pm 0.96	39.61 \pm 0.58	38.65 \pm 1.89

5.2 Optimization for the Control of a Musculoskeletal System

We establish an upper limb control task utilizing a musculoskeletal system [55]. The objective is to optimize the activities of 55 hand-related muscles to rotate and hold a bottle in vertical position

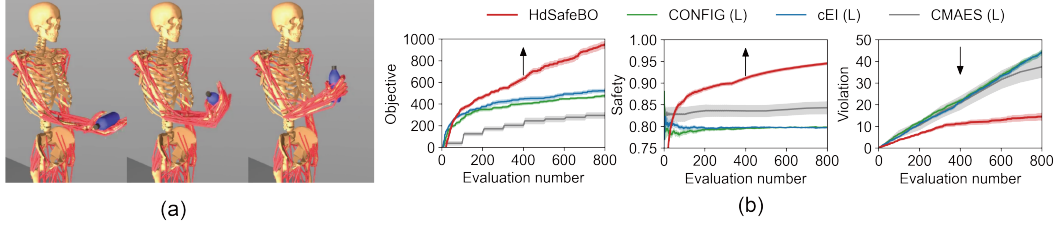


Figure 2: **Optimization for the control of a musculoskeletal system.** (a) Task illustration. (b) Optimization performance averaged over 50 independent runs. Arrows indicate the better direction.

(Figure 2 (a)). As in Mania et al. [56], we formulate the original reinforcement learning (RL) problem as a sampling problem, where the algorithms need to optimize a linear policy: $x \in \mathbb{R}^{|a| \times |o|}$, where $|a| = 55$ and $|o| = 65$ are the dimensions of the action space (number of muscles) and the observation space, respectively. The policy to be optimized has $D = 3575$ parameters, making it a very high-dimensional task. The objective function is set as the accumulated reward from the environment, and the safety function is defined as the landing speed of the bottle (see Appendix B.5). We collected muscle activation and use PCA to build the muscle synergies of performing the task, reducing action dimension from 55 to 5 (3575 to 325 for policy dimension). While the search space is significantly reduced, the remaining optimization problem is still high-dimensional.

We find that *all baselines fail to improve the objective when optimizing over the original parameter space*. Therefore we conducted baseline runs over the subspace derived from PCA (denoted with (L)), where dimension reduction facilitated effective optimization. The optimization results are presented in Figure 2 (b). SCBO (L) and LineBO (L) are omitted from the figure as they fail to attain a positive reward. We observe even within the PCA subspace, all the shown baselines demonstrated lower efficiency and sampled more unsafe parameters compared to HDSAFEBO. Our proposed method stands as the pioneering algorithm to achieve efficient and safe optimization over high-dimensional musculoskeletal system control. We also conducted an ablation study on the components of HDSAFEBO in Appendix C.2, where both optimistic safe identification and local search were shown to contribute to a safer optimization process.

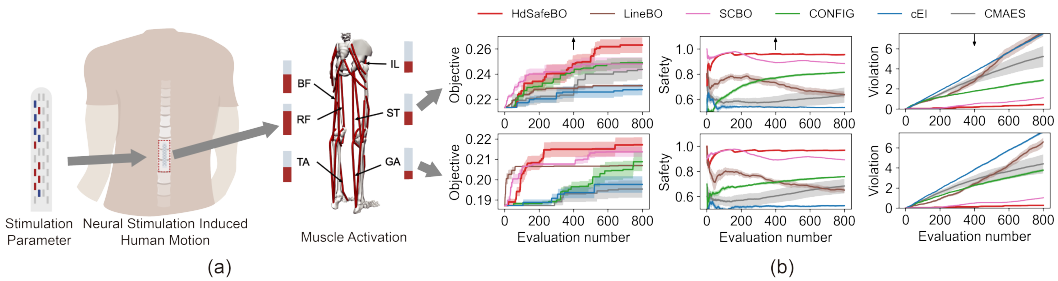


Figure 3: **Optimization for the control of neural stimulation induced human motion.** (a) Task illustration. IL, RF, TA, BF, ST, GA are different group of target muscle on the lower limb. (b) Optimization performance on the control of neuromuscular model for semitendinosus (ST), and gastrocnemius (GA), averaged over 10 independent runs. Arrows indicate the better direction.

5.3 Optimization for the Control of Neural Stimulation Induced Human Motion

Apart from direct activation of muscle-tendon units, locomotion of musculoskeletal systems can also be governed by muscle synergies arising from spinal cord stimulation, as observed in the central nervous system of vertebrates and certain neuromuscular robotic designs [57, 58]. While the neuromuscular system holds the potential to achieve robust control performance with higher energy efficiency compared to direct muscle actuation, the mapping from stimulation inputs to motion becomes less straightforward under complex neural systems. In this section, we showcase the success of HDSAFEBO in optimizing the control of intricate neural stimulation-actuated human motion through simulation and real experiments.

As shown in Figure 3 (a), we want to improve the control of lower limb muscles via an electrode array implanted in the human spinal cord. By setting different parameters, the electrical stimulation delivered by the electrode array can induce patient’s muscle activities, allowing us to control the lower limb movements of the patient. The stimulation parameter space consists of 16-contact spatial configuration (discrete) and current intensity (continuous). Our optimization goal is the selectivity index of target muscle groups (see Appendix B.6). A higher selectivity index indicates better control over the target muscle group and less influence over non-target muscles. To enhance the physical interpretability of discrete input representation, we transformed 17-dimensional vectors into an electric field map with 52×14 pixels using simplified computation. Subsequently, we generate unlabeled synthetic data with clinical priors and train an IRVAE with a 16-dimensional latent space to embed and reconstruct the electric field map, with distance-preserving checks in Appendix C.3.

Simulation over neuromuscular model. In simulation, we use a human neuromuscular model as the function oracle, which is capable of computing the evoked electric field around the spinal cord given certain stimulation parameter, and inferring the lower limb muscle activation (see Appendix B.6). The maximum induced muscle activation is used as the safety function to avoid causing harm to the patient during optimization. We choose to optimize the selectivity of emitendinosus (ST) and gastrocnemius (GA) for their importance during human walking. The simulation results are depicted in Figure 3 (b). Despite the subspace space dimensions being comparable to the original space, HDSAFEBO achieves safe exploration by optimizing on the continuous manifold, which obtains the best control performance while maintaining the highest safety selection ratio and lowest cumulative safety violation compared to other algorithms.

Real-world experiments on paraplegic patient. We further applied HDSAFEBO to improve the motor function of a paraplegic patient with the same electrode array implanted. Starting with 132 initial configurations from clinical prior, a total of 504 additional trials were conducted with the patient over a 1-month period. We observed selectivity improvement of 7 out of 8 target muscles compared to the baseline (left IL: 0.112, left RF: 0.143, left TA: 0.097, left BF: 0.380, right IL: 0.00, right RF: 0.266, right TA: 0.216, right BF: 0.141). During the whole experimental procedure, only three configurations recommended by HDSAFEBO were rated as unsafe, which evoked large lower limb movements but no physical damage or pain. Our real-world experiment underscores the practicality of HDSAFEBO in safely optimize the control of complex neuromuscular systems.

6 Conclusion

We develop HDSAFEBO for optimizing of the control over high-dimensional embodied systems under safety constraints. Our proposed method employs a local optimistic safe strategy to optimize the objective function and expand the safe region, with probabilistic safety guarantee and cumulative safety violation bound. HDSAFEBO is able to optimize high-dimensional input ranging from a few dozen to several thousand variables with safety guarantee. The algorithm can efficiently optimize the control of high-dimensional human musculoskeletal systems with high safety probability, and successfully optimize human motion control via neural stimulation in real clinical experiments. HDSAFEBO has great potential to safely optimize the control of real-world high-dimensional embodied systems online.

Limitations. While we provide the probabilistic safety guarantee for HDSAFEBO, real-world applications may fail to fully satisfy the theoretical assumptions regarding function regularity. The imperfection of the trained embedding could lead to unsafe behavior when optimizing over a reduced subspace. It is important to pre-check the distance-preserving quality of this subspace before conducting online optimization, and improvements can be made by synthesizing additional unsupervised data based on domain knowledge. The parameter space in our real human experiment has been restricted by domain prior. Directly applying HDSAFEBO to a completely unexplored problem may cause more unsafe decisions due to its optimism to safety. The total evaluation number of HDSAFEBO is constrained by cubic complexity of Gaussian process. The computational complexity prevents the use of HDSAFEBO in very long time horizon optimization problems.

Acknowledgments

This project is funded by STI 2030-Major Projects 2022ZD0209400 and Tsinghua DuShi fund.

References

- [1] F. Berkenkamp, A. P. Schoellig, and A. Krause. Safe controller optimization for quadrotors with gaussian processes. In *2016 IEEE international conference on robotics and automation (ICRA)*, pages 491–496. IEEE, 2016.
- [2] B. Sukhija, M. Turchetta, D. Lindner, A. Krause, S. Trimpe, and D. Baumann. Scalable safe exploration for global optimization of dynamical systems. *arXiv preprint arXiv:2201.09562*, 2022.
- [3] D. Widmer, D. Kang, B. Sukhija, J. Hübotter, A. Krause, and S. Coros. Tuning legged locomotion controllers via safe bayesian optimization. In *Conference on Robot Learning*, pages 2444–2464. PMLR, 2023.
- [4] Y. Sui, A. Gotovos, J. Burdick, and A. Krause. Safe exploration for optimization with gaussian processes. In *International conference on machine learning*, pages 997–1005. PMLR, 2015.
- [5] R.-R. Griffiths and J. M. Hernández-Lobato. Constrained bayesian optimization for automatic chemical design using variational autoencoders. *Chemical science*, 11(2):577–586, 2020.
- [6] P. Notin, J. M. Hernández-Lobato, and Y. Gal. Improving black-box optimization in vae latent space using decoder uncertainty. *Advances in Neural Information Processing Systems*, 34: 802–814, 2021.
- [7] D. Eriksson and M. Poloczek. Scalable constrained bayesian optimization. In *International Conference on Artificial Intelligence and Statistics*, pages 730–738. PMLR, 2021.
- [8] Y. Sui, V. Zhuang, J. Burdick, and Y. Yue. Stagewise safe bayesian optimization with gaussian processes. In *International conference on machine learning*, pages 4781–4789. PMLR, 2018.
- [9] M. Turchetta, F. Berkenkamp, and A. Krause. Safe exploration for interactive machine learning. *Advances in Neural Information Processing Systems*, 32, 2019.
- [10] D. Baumann, A. Marco, M. Turchetta, and S. Trimpe. Gosafe: Globally optimal safe robot learning. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4452–4458. IEEE, 2021.
- [11] F. Berkenkamp, A. Krause, and A. P. Schoellig. Bayesian optimization with safety constraints: safe and automatic parameter tuning in robotics. *Machine Learning*, pages 1–35, 2021.
- [12] X. Zhou and B. Ji. On kernelized multi-armed bandits with constraints. *Advances in Neural Information Processing Systems*, 35:14–26, 2022.
- [13] C. Lu and J. A. Paulson. No-regret bayesian optimization with unknown equality and inequality constraints using exact penalty functions. *IFAC-PapersOnLine*, 55(7):895–902, 2022.
- [14] H. Guo, Z. Qi, and X. Liu. Rectified pessimistic-optimistic learning for stochastic continuum-armed bandit with constraints. In *Learning for Dynamics and Control Conference*, pages 1333–1344. PMLR, 2023.
- [15] W. Xu, Y. Jiang, B. Svetozarevic, and C. Jones. Constrained efficient global optimization of expensive black-box functions. In *International Conference on Machine Learning*, pages 38485–38498. PMLR, 2023.
- [16] M. A. Gelbart, J. Snoek, and R. P. Adams. Bayesian optimization with unknown constraints. *arXiv preprint arXiv:1403.5607*, 2014.

- [17] J. M. Hernández-Lobato, M. A. Gelbart, R. P. Adams, M. W. Hoffman, and Z. Ghahramani. A general framework for constrained bayesian optimization using information-based search. 2016.
- [18] A. Marco, D. Baumann, M. Khadiv, P. Hennig, L. Righetti, and S. Trimpe. Robot learning with crash constraints. *IEEE Robotics and Automation Letters*, 6(2):1439–1446, 2021.
- [19] J. R. Gardner, M. J. Kusner, Z. E. Xu, K. Q. Weinberger, and J. P. Cunningham. Bayesian optimization with inequality constraints. In *ICML*, volume 2014, pages 937–945, 2014.
- [20] A. Marco, A. von Rohr, D. Baumann, J. M. Hernández-Lobato, and S. Trimpe. Excursion search for constrained bayesian optimization under a limited budget of failures. *arXiv preprint arXiv:2005.07443*, 2020.
- [21] M. Schonlau, W. J. Welch, and D. R. Jones. Global versus local search in constrained optimization of computer models. *Lecture notes-monograph series*, pages 11–25, 1998.
- [22] P. I. Frazier. A tutorial on bayesian optimization. *arXiv preprint arXiv:1807.02811*, 2018.
- [23] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. De Freitas. Taking the human out of the loop: A review of bayesian optimization. *Proceedings of the IEEE*, 104(1):148–175, 2015.
- [24] J. Kirschner, M. Mutny, N. Hiller, R. Ischebeck, and A. Krause. Adaptive and safe bayesian optimization in high dimensions via one-dimensional subspaces. In *International Conference on Machine Learning*, pages 3429–3438. PMLR, 2019.
- [25] S. De Blasi, A. Neifer, and A. Gepperth. Multi-pronged safe bayesian optimization for high dimensions. In *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1966–1973. IEEE, 2021.
- [26] R. Turner, D. Eriksson, M. McCourt, J. Kiili, E. Laaksonen, Z. Xu, and I. Guyon. Bayesian optimization is superior to random search for machine learning hyperparameter tuning: Analysis of the black-box optimization challenge 2020. In *NeurIPS 2020 Competition and Demonstration Track*, pages 3–26. PMLR, 2021.
- [27] M. Binois and N. Wycoff. A survey on high-dimensional gaussian process modeling with application to bayesian optimization. *ACM Transactions on Evolutionary Learning and Optimization*, 2(2):1–26, 2022.
- [28] B. Chen, R. Castro, and A. Krause. Joint optimization and variable selection of high-dimensional gaussian processes. *International Conference on Machine Learning*, 2012.
- [29] M. Zhang, H. Li, and S. Su. High dimensional bayesian optimization via supervised dimension reduction. *arXiv preprint arXiv:1907.08953*, 2019.
- [30] L. Song, K. Xue, X. Huang, and C. Qian. Monte carlo tree search based variable selection for high dimensional bayesian optimization. *Advances in Neural Information Processing Systems*, 35:28488–28501, 2022.
- [31] Z. Wang, F. Hutter, M. Zoghi, D. Matheson, and N. de Freitas. Bayesian optimization in a billion dimensions via random embeddings. *Journal of Artificial Intelligence Research*, 55:361–387, 2016.
- [32] A. Nayebi, A. Munteanu, and M. Poloczek. A framework for bayesian optimization in embedded subspaces. In *International Conference on Machine Learning*, pages 4752–4761. PMLR, 2019.
- [33] C. Li, S. Gupta, S. Rana, V. Nguyen, S. Venkatesh, and A. Shilton. High dimensional bayesian optimization using dropout. *arXiv preprint arXiv:1802.05400*, 2018.

- [34] B. Letham, R. Calandra, A. Rai, and E. Bakshy. Re-examining linear embeddings for high-dimensional bayesian optimization. *Advances in neural information processing systems*, 33: 1546–1558, 2020.
- [35] L. Papenmeier, L. Nardi, and M. Poloczek. Increasing the scope as you learn: Adaptive bayesian optimization in nested subspaces. *Advances in Neural Information Processing Systems*, 35: 11586–11601, 2022.
- [36] R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, and A. Aspuru-Guzik. Automatic chemical design using a data-driven continuous representation of molecules. *ACS central science*, 4(2):268–276, 2018.
- [37] R. Moriconi, M. P. Deisenroth, and K. Sesh Kumar. High-dimensional bayesian optimization using low-dimensional feature spaces. *Machine Learning*, 109(9):1925–1943, 2020.
- [38] A. Tripp, E. Daxberger, and J. M. Hernández-Lobato. Sample-efficient optimization in the latent space of deep generative models via weighted retraining. *Advances in Neural Information Processing Systems*, 33:11259–11272, 2020.
- [39] A. Deshwal and J. Doppa. Combining latent space and structured kernels for bayesian optimization over combinatorial spaces. *Advances in Neural Information Processing Systems*, 34: 8185–8200, 2021.
- [40] A. Grosnit, R. Tutunov, A. M. Maraval, R.-R. Griffiths, A. I. Cowen-Rivers, L. Yang, L. Zhu, W. Lyu, Z. Chen, J. Wang, et al. High-dimensional bayesian optimisation with variational autoencoders and deep metric learning. *arXiv preprint arXiv:2106.03609*, 2021.
- [41] E. Siivola, A. Paleyes, J. González, and A. Vehtari. Good practices for bayesian optimization of high dimensional structured spaces. *Applied AI Letters*, 2(2):e24, 2021.
- [42] C. Oh, E. Gavves, and M. Welling. Bock: Bayesian optimization with cylindrical kernels. In *International Conference on Machine Learning*, pages 3868–3877. PMLR, 2018.
- [43] D. Eriksson, M. Pearce, J. Gardner, R. D. Turner, and M. Poloczek. Scalable global optimization via local bayesian optimization. *Advances in neural information processing systems*, 32, 2019.
- [44] S. Müller, A. von Rohr, and S. Trimpe. Local policy search with bayesian optimization. *Advances in Neural Information Processing Systems*, 34:20708–20720, 2021.
- [45] Q. Nguyen, K. Wu, J. Gardner, and R. Garnett. Local bayesian optimization via maximizing probability of descent. *Advances in Neural Information Processing Systems*, 35:13190–13202, 2022.
- [46] K. Wu, K. Kim, R. Garnett, and J. R. Gardner. The behavior and convergence of local bayesian optimization. *arXiv preprint arXiv:2305.15572*, 2023.
- [47] N. Maus, H. Jones, J. Moore, M. J. Kusner, J. Bradshaw, and J. Gardner. Local latent space bayesian optimization over structured inputs. *Advances in Neural Information Processing Systems*, 35:34505–34518, 2022.
- [48] L. Wang, R. Fonseca, and Y. Tian. Learning search space partition for black-box optimization using monte carlo tree search. *Advances in Neural Information Processing Systems*, 33:19511–19522, 2020.
- [49] N. Hansen. The cma evolution strategy: a comparing review. *Towards a new evolutionary computation: Advances in the estimation of distribution algorithms*, pages 75–102, 2006.
- [50] O. Kramer. A review of constraint-handling techniques for evolution strategies. *Applied Computational Intelligence and Soft Computing*, 2010:1–19, 2010.

- [51] D. V. Arnold and N. Hansen. A (1+ 1)-cma-es for constrained optimisation. In *Proceedings of the 14th annual conference on Genetic and evolutionary computation*, pages 297–304, 2012.
- [52] K. Kandasamy, A. Krishnamurthy, J. Schneider, and B. Póczos. Parallelised bayesian optimisation via thompson sampling. In *International Conference on Artificial Intelligence and Statistics*, pages 133–142. PMLR, 2018.
- [53] N. Srinivas, A. Krause, S. M. Kakade, and M. Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. *arXiv preprint arXiv:0912.3995*, 2009.
- [54] L. Yonghyeon, S. Yoon, M. Son, and F. C. Park. Regularized autoencoders for isometric representation learning. In *International Conference on Learning Representations*, 2021.
- [55] C. Zuo, K. He, J. Shao, and Y. Sui. Self model for embodied intelligence: Modeling full-body human musculoskeletal system and locomotion control with hierarchical low-dimensional representation. *arXiv preprint arXiv:2312.05473*, 2023.
- [56] H. Mania, A. Guy, and B. Recht. Simple random search provides a competitive approach to reinforcement learning. *arXiv preprint arXiv:1803.07055*, 2018.
- [57] O. Aydin, X. Zhang, S. Nuethong, G. J. Pagan-Diaz, R. Bashir, M. Gazzola, and M. T. A. Saif. Neuromuscular actuation of biohybrid motile bots. *Proceedings of the National Academy of Sciences*, 116(40):19841–19847, 2019.
- [58] S. Kurumaya, K. Suzumori, H. Nabae, and S. Wakimoto. Musculoskeletal lower-limb robot driven by multifilament muscles. *Robomech Journal*, 3:1–15, 2016.
- [59] S. R. Chowdhury and A. Gopalan. On kernelized multi-armed bandits. In *International Conference on Machine Learning*, pages 844–853. PMLR, 2017.
- [60] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [61] 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*, 33:21524–21538, 2020.
- [62] E. Todorov, T. Erez, and Y. Tassa. Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ international conference on intelligent robots and systems*, pages 5026–5033. IEEE, 2012.
- [63] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *International conference on machine learning*, pages 1861–1870. PMLR, 2018.
- [64] A. Thomson. Fifth annual report of the committee of collective investigation of the anatomical society of great britain and ireland for the year 1893-94. *Journal of Anatomy*, 29(Pt 1):35–60, 1894.
- [65] R. E. Mccotter. Regarding the length and extent of the human medulla spinalis. *The Anatomical Record*, 10(9):559–564, 1916.
- [66] M. W. Zhou, W. T. Wang, H. S. Huang, G. Y. Zhu, Y. P. Chen, and C. M. Zhou. Microsurgical anatomy of lumbosacral nerve rootlets for highly selective rhizotomy in chronic spinal cord injury. *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology*, 293, 2010.
- [67] A. Rowald, S. Komi, R. Demesmaeker, E. Baaklini, S. D. Hernandez-Charpak, E. Paoles, H. Montanaro, A. Cassara, F. Becce, B. Lloyd, et al. Activity-dependent spinal cord neuromodulation rapidly restores trunk and leg motor functions after complete paralysis. *Nature medicine*, 28(2):260–271, 2022.

- [68] T. Kameyama, Y. Hashizume, and G. Sobue. Morphologic features of the normal human cadaveric spinal cord. *Spine*, 21(11):1285, 1996.
- [69] J. Ladenbauer, K. Minassian, U. S. Hofstoetter, M. R. Dimitrijevic, and F. Rattay. Stimulation of the human lumbar spinal cord with implanted and surface electrodes: A computer simulation study. *IEEE Transactions on Neural Systems Rehabilitation Engineering A Publication of the IEEE Engineering in Medicine Biology Society*, 18(6):637, 2010.
- [70] W. Sharrard. The segmental innervation of the lower limb muscles in man. *Annals of the Royal College of Surgeons of England*, 35(2):106, 1964.
- [71] C. M. Schirmer, J. L. Shils, J. E. Arle, G. R. Cosgrove, and S. Magge. Heuristic map of myotomal innervation in humans using direct intraoperative nerve root stimulation. *J Neurosurg Spine*, 15(1):64–70, 2011.
- [72] Rattay and Frank. Analysis of models for external stimulation of axons. *IEEE transactions on bio-medical engineering*, 33(10):974, 1986.
- [73] C. R. Butson and C. C. McIntyre. Role of electrode design on the volume of tissue activated during deep brain stimulation. *Journal of Neural Engineering*, 3:1–8, 3 2006. ISSN 1741-2560. doi:10.1088/1741-2560/3/1/001.
- [74] M. W. Seeger, C. K. Williams, and N. D. Lawrence. Fast forward selection to speed up sparse gaussian process regression. In *International Workshop on Artificial Intelligence and Statistics*, pages 254–261. PMLR, 2003.
- [75] E. Snelson and Z. Ghahramani. Sparse gaussian processes using pseudo-inputs. *Advances in neural information processing systems*, 18, 2005.
- [76] J. Hensman, N. Fusi, and N. D. Lawrence. Gaussian processes for big data. *arXiv preprint arXiv:1309.6835*, 2013.
- [77] K. Kandasamy, J. Schneider, and B. Póczos. High dimensional bayesian optimisation and bandits via additive models. In *International conference on machine learning*, pages 295–304. PMLR, 2015.
- [78] J. Gardner, C. Guo, K. Weinberger, R. Garnett, and R. Grosse. Discovering and exploiting additive structure for bayesian optimization. In *Artificial Intelligence and Statistics*, pages 1311–1319. PMLR, 2017.
- [79] Z. Wang, C. Gehring, P. Kohli, and S. Jegelka. Batched large-scale bayesian optimization in high-dimensional spaces. In *International Conference on Artificial Intelligence and Statistics*, pages 745–754. PMLR, 2018.
- [80] M. Mutny and A. Krause. Efficient high dimensional bayesian optimization with additivity and quadrature fourier features. *Advances in Neural Information Processing Systems*, 31, 2018.
- [81] D. Eriksson and M. Jankowiak. High-dimensional bayesian optimization with sparse axis-aligned subspaces. In *Uncertainty in Artificial Intelligence*, pages 493–503. PMLR, 2021.

A Theoretical Analysis

A.1 Proof of Proposition 4.1

Proof. Fix $t \geq 1$ and $\mathbf{z} \in \mathcal{Z}$. Conditioned on $\mathbf{y}_{t-1} = (y_1^g, \dots, y_{t-1}^g)$, $\{\mathbf{z}_1, \dots, \mathbf{z}_{t-1}\}$ are deterministic, and $g(\mathbf{z}) \sim \mathcal{N}(\mu_{t-1}(\mathbf{z}), \sigma_{t-1}(\mathbf{z}))$. For $r \sim \mathcal{N}(0, 1)$, $\Pr(r > c) = 1 - \Phi(c)$.

Therefore, by applying $r = (g(\mathbf{z}) - \mu_{t-1}(\mathbf{z}))/\sigma_{t-1}(\mathbf{z})$ and $c = \beta_t$, the statement holds. \square

A.2 Proof of Theorem 4.2

We first introduce the results in Chowdhury and Gopalan [59] which prove the selection of confidence scalar to contain the function with high probability, and the upper bound of cumulative standard deviations.

Lemma A.1. (*Lemma 5.1 in Srinivas et al. [53]*) *Let assumptions 3.1 and ?? hold for g . For any $\delta \in (0, 1)$, with probability at least $1 - \delta$, the following holds for all $\mathbf{x} \in \mathcal{X}$ and $1 \leq t \leq T, T \in \mathbb{N}$,*

$$|g(\mathbf{x}) - \mu_{t-1}(\mathbf{x})| \leq (\beta_t^c)^{\frac{1}{2}} \sigma_{t-1}(\mathbf{x}), \quad (5)$$

where $\beta_t^c = 2 \log(|\mathcal{X}| t^2 \pi^2 / 6\delta)$

Then we can bound the instantaneous safety violation of HDSAFEBO.

Lemma A.2. *Let assumptions 3.1 hold for safety function g . With probability at least $1 - \delta$, the sample points of Algorithm 1 for all time steps $1 \leq t \leq T$ satisfy*

$$\mathbb{E}[v_t] = \mathbb{E}[\max(0, -g(\mathbf{x}_t))] \leq 2(1 - \alpha)(\beta_t^c)^{\frac{1}{2}} \sigma_{t-1}(\mathbf{x}_t). \quad (6)$$

Proof. We denote $[\cdot]^+ := \max(\cdot, 0)$. Then we have

$$\mathbb{E}[v_t] = \mathbb{E}[[-g(\mathbf{x}_t)]^+] \quad (7)$$

$$= \mathbb{E}[[-g(\mathbf{x}_t) - u_t(\mathbf{x}_t) + u_t(\mathbf{x}_t)]^+] \quad (8)$$

$$\leq \mathbb{E}[[-g(\mathbf{x}_t) + u_t(\mathbf{x}_t)]^+ + [-u_t(\mathbf{x}_t)]^+] \quad (9)$$

$$= \mathbb{E}[[-g(\mathbf{x}_t) + u_t(\mathbf{x}_t)]^+] \quad (10)$$

$$= \mathbb{E}[[-g(\mathbf{x}_t) + u_t(\mathbf{x}_t)]^+ \mathbb{1}\{g(\mathbf{x}_t) \geq u_t(\mathbf{x}_t)\} + [-g(\mathbf{x}_t) + u_t(\mathbf{x}_t)]^+ \mathbb{1}\{g(\mathbf{x}_t) < u_t(\mathbf{x}_t)\}] \quad (11)$$

$$= (1 - \alpha)[-g(\mathbf{x}_t) + u_t(\mathbf{x}_t)]^+ \quad (12)$$

$$\leq (1 - \alpha)[-(\mu_{t-1}(\mathbf{x}_t) - (\beta_t^c)^{\frac{1}{2}} \sigma_{t-1}(\mathbf{x}_t)) + \mu_{t-1}(\mathbf{x}_t) + (\beta_t^c)^{\frac{1}{2}} \sigma_{t-1}(\mathbf{x}_t)] \quad (13)$$

$$= 2(1 - \alpha)(\beta_t^c)^{\frac{1}{2}} \sigma_{t-1}(\mathbf{x}_t), \quad (14)$$

where the inequality (9) follows by the fact that $[a + b]^+ \leq [a]^+ + [b]^+, \forall a, b \in \mathbb{R}$, the equality (10) is derived from the fact that $u_t(\mathbf{x}_t) \geq 0$, the equality (12) is derived from Proposition 4.1, and the inequality (13) is derived from Lemma A.1. \square

Lemma A.3. (*Theorem 1 in Srinivas et al. [53].*) *Let $\mathbf{x}_1, \dots, \mathbf{x}_T$ be the points selected by the algorithms. With $C_1 := 8/\log(1 + \sigma^{-2})$,*

$$\sum_{t=1}^T 2(\beta_t^c)^{\frac{1}{2}} \sigma_{t-1}(\mathbf{x}_t) \leq \sqrt{C_1 T \beta_T^c \gamma_T} \quad (15)$$

Finally we bound the summation of instantaneous safety violations.

Proof.

$$\mathbb{E}[V_T] = \mathbb{E}\left[\sum_{t=1}^T v_t\right] \quad (16)$$

$$= \sum_{t=1}^T \mathbb{E}[v_t] \quad (17)$$

$$\leq \sum_{t=1}^T 2(1-\alpha)(\beta_t^c)^{\frac{1}{2}}\sigma_{t-1}(\mathbf{x}_t) \quad (18)$$

$$\leq \sum_{t=1}^T 2(1-\alpha)(\beta_T^c)^{\frac{1}{2}}\sigma_{t-1}(\mathbf{x}_t) \quad (19)$$

$$\leq (1-\alpha)\sqrt{C_1 T \beta_T^c \gamma_T}, \quad (20)$$

where the inequality (18) follows by Lemma A.2, the inequality (19) follows by the monotonicity of β_t^c , and the inequality (20) follows by Lemma A.3. \square

A.3 Proof of Proposition 4.3

Proof. Define $k^{\mathcal{X}}, k^{\mathcal{Z}}$ are the kernel function of GP with same hyperparameter over \mathcal{X} and \mathcal{Z} respectively. For stationary kernel, the kernel function value between \mathbf{x} and \mathbf{x}' is only depend on $|\mathbf{x} - \mathbf{x}'|$, therefore we can write the $k^{\mathcal{X}}, k^{\mathcal{Z}}$ as a function of metric $d_{\mathcal{X}}, d_{\mathcal{Z}}$:

$$k^{\mathcal{X}}(\mathbf{x}, \mathbf{x}') = k^{\mathcal{X}}(d_{\mathcal{X}}(\mathbf{x}, \mathbf{x}')), k^{\mathcal{Z}}(\mathbf{z}, \mathbf{z}') = k^{\mathcal{Z}}(d_{\mathcal{Z}}(\mathbf{z}, \mathbf{z}')). \quad (21)$$

We define $k^{\mathcal{X}}$ as the kernel function of GP over \mathcal{X} , and $u_t^{\mathcal{X}}$ as the estimated upper confidence bound of GP over \mathcal{X} . Utilizing the property of isometric embedding, we can obtain equivalent GP estimation between original space and embeded subspace:

$$k^{\mathcal{Z}}(d_{\mathcal{Z}}(\mathbf{z}, \mathbf{z}')) = k^{\mathcal{X}}(d_{\mathcal{X}}(\Pi^{-1}(\mathbf{z}), \Pi^{-1}(\mathbf{z}'))). \quad (22)$$

Therefore the estimated upper confidence bound is also equivalent. For every sample \mathbf{z} from HDSAFEBO, we have

$$\Pr(g(\Pi^{-1}(\mathbf{z})) \geq 0) \geq \Pr(u_t^{\mathcal{X}}(\Pi^{-1}(\mathbf{z})) \geq 0) \quad (23)$$

$$= \Pr(u_t^{\mathcal{Z}}(\mathbf{z}) \geq 0) \quad (24)$$

$$\geq \alpha, \quad (25)$$

the statement holds. \square

B Experimental Details

The full implementation of our experiments can be found on our project page: <https://lmsgroup.cc/research/hdsafebo>. Our musculoskeletal model will be released soon. In the meantime, the model can be accessed for research purposes upon request (ysui@tsinghua.edu.cn).

B.1 PCA training

We employ PCA using the scikit-learn library ¹ with default parameter settings.

¹<https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html>

B.2 Autoencoder training

We employ IRVAE on neural stimulation task and digital generation task by directly using the paper’s original repository². We use MLP as the VAE module for all tasks, and list the model detail in Table 2. We train all models for 300 epochs using Adam[60] optimizer with a learning rate of 0.0001.

Task	Layer number	Hiddien number	Latent dimension
Neural stimulation	4	256	16
Digital generation	4	256	6

Table 2: Autoencoder model detail.

B.3 Algorithm Implementation

For implementation of HDSAFEBO, We use BoTorch as the GP inference component [61]. We also use BoTorch to replicate LineBO, SCBO, CONFIG, cEI and cEI-Prob. For LineBO, we choose to implement the random line embedding version shown in the main paper. We use the package pycma³ to run CMA-ES on benchmarks.

All GP-based methods uses matérn kernel and fits kernel parameters after each iteration. During the experiment, we set $\tau_s, \tau_f, l_0, l_{\max}, l_{\min}$ for HDSAFEBO and SCBO (the default setting of SCBO). We use Thompson sampling for LineBO and CONFIG, and use inherent acquisition function for other BO baselines. Confidence scalar β is set as 2 for HDSAFEBO, LineBO and CONFIG across all experiments. We set the latent optimization bound as the upper bound and lower bound of training points in the latent space. Other baseline parameters are set to default values as in the original implementation.

During the experiment, we set the sample size to 10 for all tasks in the main paper, and sample one point each iteration in constrained digital generation task.

B.4 Synthetic Function

We sample objective and safety function from separate Gaussian process with Matérn kernel and length scale as 0.05, which is implemented using GPyTorch⁴. We set the safety threshold to -0.75 . For each independent run, we use a random linear projection $\Pi_{\text{rand}} \in \mathbb{R}^{D \times d}$ to create d -dimensional latent space, and randomly select d_e variables as the function effective dimension.

B.5 Musculoskeletal Model Control

We use a full-body musculoskeletal model which actuates locomotion by controlling muscle activation. Here we only control the right hand part (below elbow), and fix other joints, leading to 55 muscles and 28 joints. The primary task is to control hand muscles to maintain a steady vertical grip on a bottle. At the beginning of the episode, the bottle is initially positioned horizontally in the hand. The initial task involves first rotating the bottle to achieve and maintain a vertical orientation. At each time step, the reward from the environment is computed as follows:

$$r = r_{\text{pose}} + r_{\text{bonus}} - 10 * r_{\text{penalty}} + r_{\text{grasp}} + 2 * r_{\text{survive}} - r_{\text{activation}} - 100 * r_{\text{drop}} \quad (26)$$

where r_{pose} is the difference between the bottle and vertical orientation, computed by Euler angle. r_{bonus} is the reward when the difference falls below a predefined threshold. r_{penalty} is positive when the bottle position is out of from the predefined range. r_{grasp} is the distance between the centroid of

²<https://github.com/Gabe-YHLee/IRVAE-public>

³<https://github.com/CMA-ES/pycma>

⁴<https://gpytorch.ai/>

the bottle and palm joints. r_{survive} is the reward for not dropping the bottle during the current time step. $r_{\text{activation}}$ is the penalty for large muscle activations. r_{drop} is the penalty when the bottle drop from hand. The overall simulation is based on Mujoco [62].

When the height of the bottle is below 0.4m, we consider it to be dropped from the hand and the episode concludes. We record the speed of episode ending as the landing speed of the bottle. We use a safety threshold of 4.4, which is the average landing speed when randomly sampling the environment. We train a Soft Actor-Critic agent [63] for 150k time steps under different observation setting, and rollout for 1000 episode to build a muscle activation dataset with 107, 439 datapoints.

The performance video of HDSAFEBO and other baselines (in the subspace from PCA) is shown in the supplementary file. HDSAFEBO is capable of quickly rotating the bottle and hold vertically and steadily, while other algorithms either hold the bottle non-vertically, or learn drop the bottle safely to avoid penalty of wrong orientation.

B.6 Modeling of the human neuromuscular system

We developed an average model of the human spinal cord based on anatomical statistics(model paper under review, [64, 65, 66, 67, 68]). The model contains gray matter, white matter, nerve roots, cerebrospinal fluid (CSF), and dura mater of T12-S2 segments of the spinal cord which are related with the motor control of lower limbs. The specific conductivity values of the modeled tissues were set refer to [69]. Electric fields induced by different stimulation parameters were derived using finite element method (FEM). To calculate the stimulation effects for different muscles, we redistricted the cord model according to reported results of the segmental innervation for lower limb muscles([70, 71]). Six groups of muscles of bilateral lower limbs were studied: iliopsoas (IL), vastus lateralis and rectus femoris (VL&RF), tibialis anterior (TA), biceps femoris muscle and gluteus maximus (BF&GM), semitendinosus (ST), and gastrocnemius (GA).

To evaluate the selectivity of stimulation for certain muscle, we used a selectivity index (SI) to characterize the distribution of the electric field. The selectivity index for the i th muscle was defined as follows:

$$SI_i = \mu_i - \frac{1}{m_{\text{neighbor}} - 1} \sum_{j \neq i}^{m_{\text{neighbor}}} \mu_j \quad (27)$$

where m_{neighbor} represents the number of muscles whose motor neuron pools are adjacent to the i th muscle's. The selectivity index ranges from -1 to 1, where -1 represents the maximum of activation of all undesired muscles with a complete absence of activation of the targeted muscle, 0 indicates that all muscles are activated at the same level, and 1 means the targeted muscle is activated at the greatest extent while no undesired muscles are activated. And μ_i is the normalized activation of the i th muscle and is defined as follows in the simulation:

$$\mu_i = \frac{\iiint_{\Omega_i} f(x, y, z) dx dy dz}{\iiint_{\Omega_i} 1 dx dy dz} \quad (28)$$

$$f(x, y, z) = \begin{cases} 1, & \text{if } AF(x, y, z) > AF_{\text{threshold}} \\ 0, & \text{if } AF(x, y, z) \leq AF_{\text{threshold}} \end{cases} \quad (29)$$

Ω_i is the segmental volume of the i th muscle in the cord. AF is the activating function, defined as the second spatial derivative of extracellular voltage along an axon([72, 73]).

We use the spinal model to traverse all stimulation parameters with 1 cathode with anodes no more than 3, and 2 cathodes with anodes no more than 2, leading to a spinal cord stimulation (SCS) dataset with 218,000 stimulation parameters and predicted muscle activation. We compute the objective function using 27, and compute the safety function as $g(x) = 1 - \max_i(\mu_i)$. The selectivity index distribution is shown in Figure 4. We set the safe threshold as 0.05, with nearly half of the traversed parameters are safe.

We convert electrode parameters from 17d vector to a 2d electric field image using simplified computation. In concrete, we map contact combinations to the spatial position in the electrode, linearly compute the diffusion of electrical field from each cathode and anode, and multiply the map by current intensity. One example of generated electric field image is shown in Figure 5.

We use the constructed SCS dataset as the function oracle. For a given 2d map, we set its function value as the function value of the nearest unevaluated point in the dataset measured by the electrical field map.

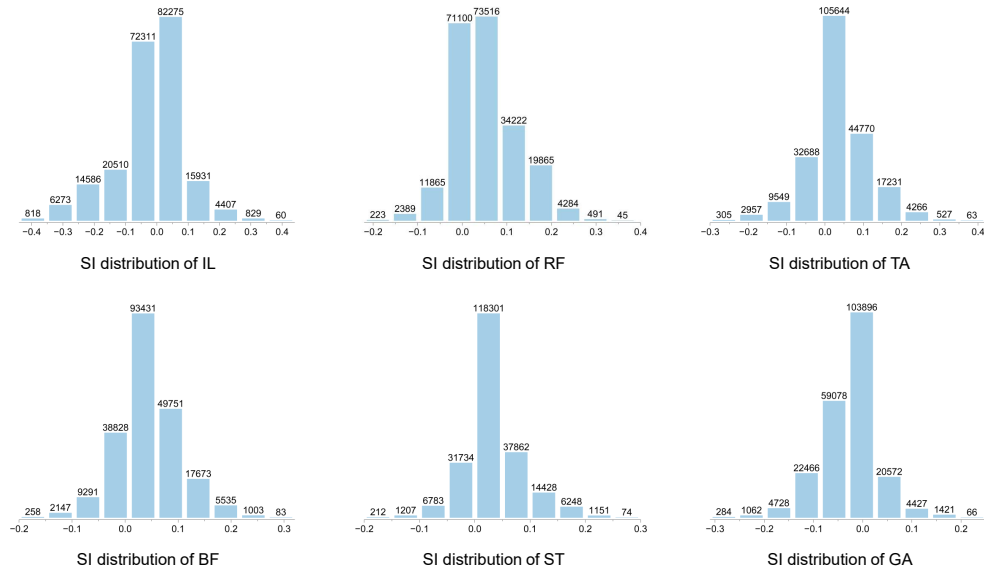


Figure 4: Distribution of SI for six muscle groups of different configurations used in SCS simulation experiment

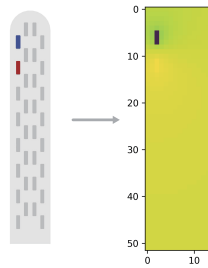


Figure 5: 2d electrical map computation.

B.7 Clinical experiment of neuromuscular system control

We employ HDSAFEBO in the treatment of spinal cord stimulation to find more selective stimulation parameters for different muscles. The clinical experiments received approval from the Institutional Review Board (IRB) of the hospital. All the trials were conducted under the supervision of therapists. The patient was seated in the wheelchair in a comfortable way and was told to relax. During the first period, typical parameters which were usually used in the therapy (e.g. bipolar stimulation) were delivered to the patient while the evoked muscle activities were recorded using EMG. These data (132 trials) were used to initialize HDSAFEBO. Except for the first 132 trials as the initial data, 441 out of 504 trials are recommended by HDSAFEBO. The other trials were conducted by the therapist.

We focus on 8 groups of muscles: iliopsoas (IL), rectus femoris (RF), tibialis anterior (TA), and biceps femoris (BF) for both sides. The clinical selectivity index was defined as following:

$$SI_i = \frac{\mu_i}{1 + \sum_{j \neq i}^m \mu_j} \quad (30)$$

where μ_i represents the normalized peak-to-peak value of the evoked EMG for the i th muscle and m is the total number of the target muscles.

During the optimization, We restrict the number of cathodes and anodes to only evaluate configurations consistent with clinical priors. We used the Electromyography (EMG) to compute the selectivity index and queried safety scores from the patient and the therapists. For each trial, our algorithm recommended the parameter based on the history data and configured it onto the stimulator, which would deliver electrical stimulation to the patient for 800 ms at a frequency of 10 Hz. Peak-to-peak values were averaged and normalized to obtain selectivity indices of different muscles after stimulation. The calculated feedback and queried safety score were used to update the optimizer and it would recommend a new parameter. The tasks were refined sequentially and all the history data were reused when optimizing a new task.

C Additional Experiment

C.1 Baselines performance over embedded subspace

We run baselines on the same embedded space as HdSafeBO in optimizing GP synthetic functions (Table 3). We observe that, overall, the optimization and safety performances slightly improved. However, HdSafeBO still outperforms these baselines, as the reduced space remains too high-dimensional for them.

Table 3: Algorithm performance on GP synthetic functions. We show the averaged performance of 500 evaluations over 100 independent runs. "L" indicates baselines optimize over the embedded latent space, "O" indicates baselines optimize over the original input space.

Method	HdSafeBO	LineBO		SCBO		CONFIG		cEI		CMAES	
		L	O	L	O	L	O	L	O	L	O
Objective (†)	3.96 ± 0.15	3.03 ± 0.05	3.07 ± 0.04	3.15 ± 0.06	2.95 ± 0.04	2.99 ± 0.04	2.91 ± 0.05	2.92 ± 0.03	2.9 ± 0.03	2.69 ± 0.04	2.7 ± 0.04
Safety (†)	0.81 ± 0.02	0.79 ± 0.0	0.78 ± 0.0	0.77 ± 0.0	0.77 ± 0.0	0.77 ± 0.0	0.77 ± 0.0	0.77 ± 0.0	0.77 ± 0.0	0.77 ± 0.01	0.78 ± 0.01
Violation (‡)	27.42 ± 4.01	36.26 ± 1.28	36.59 ± 0.82	38.51 ± 0.64	38.47 ± 0.69	39.08 ± 0.6	38.96 ± 0.96	39.15 ± 0.64	39.61 ± 0.58	38.01 ± 1.94	38.65 ± 1.89

C.2 Ablation on algorithm components

In the musculoskeletal system control task, we conducted an ablation study of two components in HdSafeBO: local search and optimistic safe identification (Figure 6). We observed that without local search, the algorithm tends to over-explore, leading to degraded optimization and safety performance in this high-dimensional problem. Without optimistic safe identification, the algorithm makes more unsafe selections during the early stages of optimization. Combining these two components enables a safe and efficient optimization procedure.

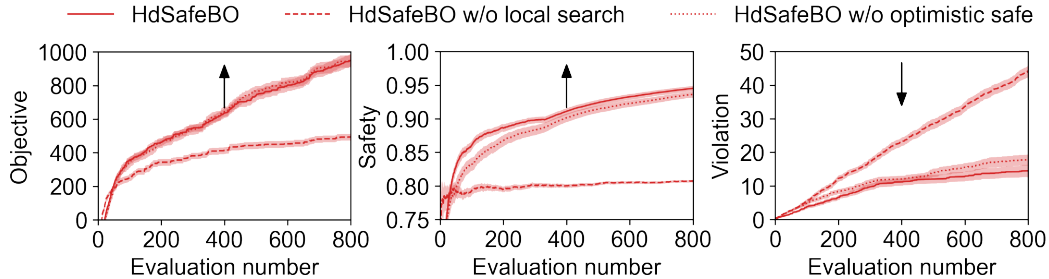


Figure 6: Ablation on components of HdSafeBO. Optimization performance averaged over 50 independent runs.

C.3 Distance Preserving of IRVAE

In the implementation of HDSAFEBO, we explore the use of IRVAE to learn a mapping with distance-preserving property. We randomly sample 10,000 datapoints from SCS dataset. Using conventional VAE and IRVAE trained from same dataset, we compute the point-wise distance as well as GP estimation difference between the original and latent space, as shown in Table 4. The results

Using the 218,000 synthetic training data in human neural stimulation control, we plotted the pair-wise distance of points in both original input space and latent space learns by IRVAE (Figure 7). We can observe that the learned embedding space exhibits approximately scaled isometry.

Model	Linear correlation of distance (\uparrow)	GP mean estimation difference (\downarrow)	GP variance estimation difference (\downarrow)
IRVAE	0.9729	1.2936	2.9953
conventional VAE	0.8923	2.1839	4.5608

Table 4: Distance preserving comparison between IRVAE and conventional VAE

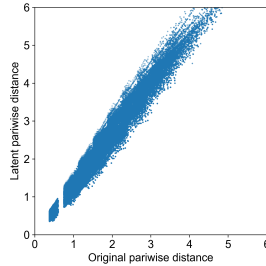


Figure 7: Pair-wise Distance of training data points in neural stimulation induced human motion control task.

C.4 Ablation on Confidence Bound Scalar

Here we run HDSAFEBO with $\beta = 0, 2, 4, 8, 16$ on musculoskeletal model control task, and shown the results in Figure 8. The safety and violation metric may become slightly worse as β increases. We observe the algorithm performance is similar under a wide range choice of β on this high-dimensional task.

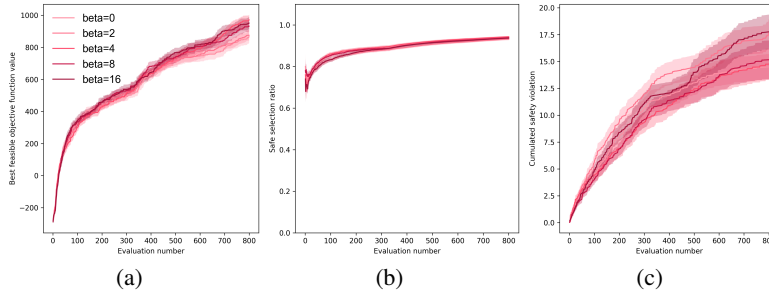


Figure 8: Ablation study on confidence bound scalar β .

C.5 Comparison with random-embedding BO

We additionally run HesBO and BAXUS on the musculoskeletal model control task and the neural stimulation task with no safety constraints in simulation. Due to the algorithmic mechanism of HesBO and BAXUS, we cannot directly use the same initial point as HDSAFEBO. Therefore we randomly sample initial points from their corresponding latent space. In HesBO, we set the same latent dimension number as in HDSAFEBO. Table 5 shows the best objective function values found by algorithms (shown as mean \pm 1 std).

Table 5: Best objective function values found by algorithms.

Algorithm	SCS-ST	SCS-GA
HDSAFEBO	0.26 ± 0.02	0.22 ± 0.01
HesBO	0.23 ± 0.0	0.19 ± 0.02
BAXUS	0.24 ± 0.03	0.19 ± 0.05

We observe HDSAFEBO still outperforms HesBO and BAXUS across all tasks, even when optimizing under safety constraint. We think using IRVAE enables utilizing the pre-collected unlabeled data to learn a better representation than random projection.

C.6 Hand-writing Digital Generation

In hand-writing digital generation task, the goal is to generate images of target digit as thick as possible, while keeping the image valid for the required number. Using this task we can test the algorithm performances when the latent dimension is low. We trained a fully-connected IRVAE with a latent dimension of 6 and use a two-layer CNN model as the predictor. We set the objective function as the sum of image pixel intensities and the safety function as the prediction probability of target number. Since the CNN prediction is very sharp, we wrap the CNN output via a softmax layer with temperature as 200. We set the sample budget as 200 including with 20 images of target digit as the initial data.

We summarize the averaging performance in Table 6. We observed HDSAFEBO outperforms all baselines in terms of optimization performance and safety violation. Note that while HDSAFEBO efficiently finds highest objective, its safety violations is 63% less than the second best method original SCBO.

Table 6: Experiment results of constrained hand-writing digital generation. We evaluate algorithm performance of generating digital from 0 to 9 in terms of best found feasible objective value (higher is better) and cumulative safety violation (lower is better). Objective values are normalized by best feasible point in the MNIST dataset. The results are shown as mean performance ± one standard deviation across ten tasks.

Metric	HDSAFEBO	SCBO	CONFIG	cEI	CMAES
Objective	1.14 ± 0.12	1.08 ± 0.18	0.77 ± 0.26	0.77 ± 0.26	0.68 ± 0.12
Violation	18.99 ± 10.82	52.01 ± 20.28	72.09 ± 17.85	72.04 ± 17.82	72.01 ± 7.06

C.7 Run time performance

We ran HdSafeBO and baseline algorithms on the human neural stimulation control problem and recorded the average run-time of each iteration in Table 7. Most baseline algorithms have a processing time of around 10 seconds.

In our problem setting, we consider the run-time difference to be marginal compared to the actual experiment time for each trial. For instance, in our real experiments, applying recommended stimulation parameters typically takes 1-2 minutes. Therefore, we compared optimization performances based on the number of evaluations.

Method	HdSafeBO	LineBO	SCBO	CONFIG	cEI	CMAES
Iteration time	11.75 ± 1.35s	9.88 ± 1.01s	11.99 ± 1.21s	9.49 ± 1.05s	20.38 ± 1.21s	9.59 ± 0.72s

Table 7: Run-time per iteration in neural induced human motion control task. Results show mean ± 1 std averaged over 80 iterations.

D Additional Related Work

Here we additionally discuss more works about high-dimensional Bayesian optimization besides dimension-reduction based BO and local BO.

Due to the inversion of kernel function matrix, the complexity of GP inference scales exponentially with the sample number, limiting the search budget of high-dimensional problems. Sparse GP or variational GP is used to achieve scalable sampling over the high-dimensional space [74, 75, 76].

Another line of work assume the additive structure of the objective functions, and decomposes the function to solve the low-dimensional sub-problem decentrally [77, 78, 79, 80].

To overcome over-exploration issue over the high-dimensional space, several works also propose to shape kernel prior to sample points near the search center [42, 81].