Preferential Multi-Attribute Bayesian Optimization with Application to Exoskeleton Personalization

Raul Astudillo [1] **Kejun Li** [2] **Maegan Tucker** [3] **Chu Xin Cheng** [1] **Aaron D. Ames** [1 3] **Yisong Yue** [1]

Abstract

Preferential Bayesian optimization (PBO) is a framework for optimization of a decision-maker's (DM's) latent preferences. Existing work in PBO assumes these preferences can be encoded by a single latent utility function, which is then estimated from ordinal preference feedback over design variables. In practice, however, it is often challenging for DMs to provide such feedback reliably, leading to poor performance. This is especially true when multiple conflicting latent attributes govern the DM's preferences. For example, in exoskeleton personalization, users' preferences over gait designs are influenced by stability and walking speed, which can conflict with each other. We posit this is a primary reason why inconsistent preferences are often observed in practice. To address this challenge, we propose a framework for preferential multi-attribute Bayesian optimization, where the goal is to help DMs efficiently explore the Pareto front of their preferences over attributes.Within this framework, we propose a Thompson sampling-based strategy to select new queries and show it performs well across three test problems, including a simulated exoskeleton gait personalization task.

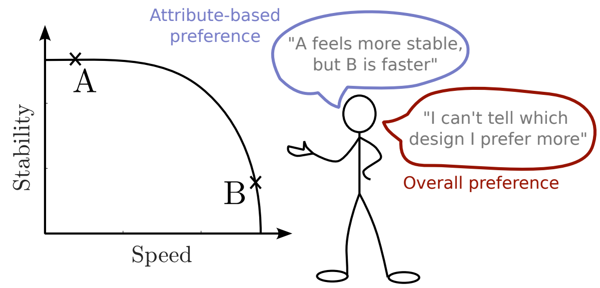


Figure 1: In scenarios where high-quality designs are spread across competing goodness measures (attributes), overall preference feedback can be challenging for DMs to provide. In contrast, it is easier for DMs to directly give feedback on their preferences over attributes and for the algorithm to identify then designs on the Pareto front of preferences over attributes. For example, in the illustration, it is easier for the DM to provide preference feedback regarding the stability and speed of designs A vs. B rather than providing an overall preference between the two designs.

1. Introduction

Bayesian optimization (BO) is a framework for optimizing objective functions with expensive or time-consuming evaluations. It has been successful in real-world applications such as cellular agriculture (Cosenza et al., 2022), chemical design (Griffiths & Hernández-Lobato, 2020), and hyperparameter tuning (Wu et al., 2019). Preferential Bayesian optimization (PBO), a subframework within BO, focuses on settings where the objective function is only measured indirectly through ordinal preference feedback (often in the form of pairwise comparisons) expressed by a decision-maker (DM). This arises, for example, in exoskeleton personalization, where a user assisted by an exoskeleton walks using different gait designs and indicates the one that resulted in more comfortable walking (Tucker et al., 2020a;b).

Prior work on PBO operates under the assumption that the DM's preferences can be encoded by a single latent utility function, which is then estimated from ordinal preference feedback. In practice, however, it is often challenging for DMs to provide such feedback reliably, leading to poor performance. This is particularly true in situations where the DM considers several underlying conflicting attributes when expressing preferences. In such cases, the standard PBO approach expects the DM to be able to aggregate their preferences across attributes to express overall preferences, which can be challenging.

[1]Department of Computing and Mathematical Sciences, Caltech, Pasadena, USA [2]Department of Computation and Neural Systems, Caltech, Pasadena, USA [3]Department of Mechanical and Civil Engineering, Caltech, Pasadena, USA. Correspondence to: Raul Astudillo <raul.astudillo@caltech.edu>.

The Many Facets of Preference Learning Workshop at the International Conference on Machine Learning (ICML), Honolulu, Hawaii, USA, 2023.

For example, in exoskeleton personalization, users' preferences over gait designs are influenced by stability and speed, which can conflict with each other. Users often have difficulty expressing preferences over pairs of gait designs where one produces fast but somewhat unstable walking, and the other produces slow but quite stable walking. However, they can easily express preferences over these two attributes individually (Figure 1). Moreover, users often wish to explore the trade-offs between these attributes, as this allows them to make a more informed decision before committing to a gait design.

Using the above as motivation, we propose a framework for preferential multi-attribute BO, where DMs express preferences over multiple attributes of interest instead of overall preferences. Our framework aims to help DMs efficiently explore the Pareto front of their preferences over attributes. Our approach has the following advantages over the standard PBO approach:

1. Preferences over attributes pose a lower cognitive load on DMs and are more reliable than overall preferences, thus resulting in more reliable probabilistic models to guide the search for new queries.
2. Exploring the Pareto front of preferences over attributes provides better support for decision-making, as it allows DMs to understand the trade-offs between competing attributes before committing to a solution.

To our knowledge, our work is the first one pursuing optimization of multiple *latent* attributes using preference feedback, both in and outside the BO framework.

We illustrate our framework in three test problems, including a simulated exoskeleton personalization task. Our results demonstrate the ability of our approach to explore the Pareto front of the DM's preferences over attributes.

2. Related Work

Our work is closely related to three lines of research: preferential Bayesian optimization, multi-objective optimization, and preference aggregation. We discuss connections of our work to these lines of research and also mention other works relevant to our own.

Preferential Bayesian Optimization Our work can be seen as an extension of the PBO framework to the multi-attribute (a.k.a. multi-objective) setting. PBO was first considered by Brochu et al. (2010) in the context of animation design. Since then, most work in this area has focused on applications (Nielsen et al., 2015; Tucker et al., 2020b) and the development of more sophisticated sampling policies (González et al., 2017; Nguyen et al., 2021; Astudillo et al., 2023). In this line of research, our work is most closely related to works considering richer forms of preference feedback. These include the best item in a menu (Siivola et al., 2021; Astudillo et al., 2023), full ranking of items in a menu (Siivola et al., 2021), ranking of a subset of items in a menu (Nguyen et al., 2021), and *projective feedback*, in which the DM indicates a direction in the design space to explore next (Mikkola et al., 2020). To our knowledge, our work is the first to consider preference feedback over multiple attributes within PBO.

Multi-Objective Optimization Multi-objective optimization has been widely studied both in theory and its application to different engineering problems (Miettinen, 1999; Marler & Arora, 2004; Deb, 2013). Among this broad literature, work on multi-objective BO is most closely related to ours (Khan et al., 2002; Knowles, 2006; Belakaria et al., 2019; Daulton et al., 2020). Our work draws inspiration from the work of Knowles (2006), which leverages augmented Chebyshev scalarizations to transform a multi-objective optimization problem into multiple single-objective optimization problems. The incorporation of user preferences to improve efficiency in multi-objective optimization has been actively studied both in and outside the BO framework (Branke & Deb, 2005; Wang et al., 2017; Hakanen & Knowles, 2017; Astudillo & Frazier, 2020; Lin et al., 2022). We note that, unlike in our work, this literature assumes that objectives (i.e., attributes) are observable.

Preference Aggregation Preference aggregation deals with the problem of combining multiple (potentially conflicting) notions of preference either across multiple criteria (i.e., attributes) or multiple users (Young, 1974; Dyer & Sarin, 1979; Baskin & Krishnamurthi, 2009; Baumeister & Rothe, 2016). While, in our framework, preferences across attributes are not aggregated to retain the ability to explore the Pareto front, our work is related to this line of research in that preferences over multiple attributes are considered. We also emphasize that methods in this area have been studied outside the BO framework, which aims for sample efficiency, so they are not readily applicable to our setting.

Other Relevant Works While not crucial to our framework, our numerical experiments use Gaussian process priors to model preferences over attributes. Using Gaussian processes to model preferences was first proposed by Chu & Ghahramani (2005). Approximate posterior inference is performed via the variational inducing point approach of Hensman et al. (2015). Our sampling policy draws inspiration from the self-sparring algorithm for multi-dueling bandits (Sui et al., 2017), which has also been extended to preference-based reinforcement learning (where it is termed dueling posterior sampling) (Novoseller et al., 2020) and PBO (where it is termed batch Thompson sampling).

3. Problem Setting

We denote the space of designs or inputs by $\mathbb{X}$. We assume there are m attributes and let $f_j : \mathbb{X} \to \mathbb{R}$ denote the DM's latent utility function over attribute j for $j = 1, \ldots, m$. The concatenation of these utility functions is denoted by $f = [f_1, \ldots, f_m] : \mathbb{X} \to \mathbb{R}^m$. The goal of the DM is to find designs such that the corresponding value of each of the m utility functions is as large as possible. As such, our goal is to help the DM explore the *Pareto front* of f. This set is defined using the notion of Pareto-dominance. For a pair of actions $x, x' \in \mathbb{X}$, x Pareto dominates x' if $f_j(x) \geq f_j(x')$ for $j = 1, \ldots, m$ and the inequality is strict for at least one index j. We write $x \succ_f x'$ to denote that x Pareto-dominates x' with respect to f. The Pareto-optimal set of f is $\mathbb{X}^* := \{x : \nexists\, x' \text{ s.t. } x' \succ_f x\}$. The set $\mathbb{Y}^* := \{f(x) : x \in \mathbb{X}^*\}$ is called the Pareto front of f.

To support the DM's goal, an algorithm collects preference feedback interactively (Algorithm 1). Concretely, at every iteration, an algorithm selects a query, $X_n = (x_{n,1}, x_{n,2}) \in \mathbb{X}^2$, where $n = 1, \ldots, N$ denotes the iteration number. The DM then expresses their most preferred design between $x_{n,1}$ and $x_{n,2}$ with respect to each attribute. This response is encoded as a vector $r(X_n) \in \{1, 2\}^m$, where $r_j(X_n)$, the j-th entry of $r(X_n)$, is 1 if the DM prefers $x_{n,1}$ over $x_{n,2}$ with respect to attribute j and 2 otherwise. We shall sometimes denote $r(X_n)$ more compactly by r_n.

For each attribute j, we model noise in the DM's response via a Logistic likelihood of the form

$$\mathbf{P}\left(r_j(X_n) = 1\right) = \frac{\exp\left(f_j(x_{n,1})/\lambda_j\right)}{\exp\left(f_j(x_{n,1})/\lambda_j\right) + \exp\left(f_j(x_{n,2})/\lambda_j\right)},$$

where $\lambda_j > 0$ is the noise-level parameter we can estimate along with other parameters in our model. We assume noise is independent across attributes and interactions. As is standard in BO, we place a prior distribution over f. We denote this prior distribution by p_0. Let $\mathcal{D}_0$ denote the initial data set and $\mathcal{D}_{n-1} = \mathcal{D}_0 \cup \{(X_k, r_k)\}_{k=1}^{n-1}$ denote the preference information collected before the n-th interaction with the DM. The posterior distribution on f given $\mathcal{D}_{n-1}$ is denoted by p_n. In our experiments, we model each attribute independently using a Gaussian process prior. However, more sophisticated priors can be used.

Since our goal is to help DMs explore the Pareto front of their preferences over attributes, we quantify the performance of an algorithm using the hypervolume indicator. Previous work has shown that maximizing hypervolume results in Pareto fronts with good coverage (Zitzler et al., 2003). If $\widehat{\mathbb{Y}}^* = \{y_\ell\}_{\ell=1}^L$ is a finite approximation of the Pareto front of f, its hypervolume is given by $\mathrm{HV}(\widehat{\mathbb{Y}}^*, r) = \lambda_m\left(\bigcup_{\ell=1}^L [r, y_\ell]\right)$, where $r \in \mathbb{R}^m$ is a reference point, λ_m denotes the m-dimensional Lebesgue measure, and $[r, y_\ell]$ denotes the hyper-rectangle bounded by the vertices r and y_ℓ. In our experiments, we report performance by taking $\widehat{\mathbb{Y}}^*$ as the set of Pareto optimal attribute vectors corresponding to designs shown to the DM so far.

Algorithm 1 Preferential Multi-Attribute BO Loop

Input Initial dataset: $\mathcal{D}_0$, and prior distribution over f: p_0.
for $n = 1, \cdots, N$ **do**
  Compute p_n, the posterior on f given $\mathcal{D}_{n-1}$.
  Sample $\tilde{\theta}_n$ uniformly at random over Θ.
  Draw samples $\tilde{f}_{n,1}, \tilde{f}_{n,2} \overset{\text{iid}}{\sim} p_n$.
  Find $x_{n,i} \in \mathrm{argmax}_{x \in \mathbb{X}}\ c(x \mid \tilde{\theta}_n, \tilde{f}_{n,i})\ i = 1, 2$.
  Set $X_n = (x_{n,1}, x_{n,2})$ and observe feedback, r_n.
  Update data set $\mathcal{D}_n = \mathcal{D}_{n-1} \cup \{(X_n, r_n)\}$.
end for

4. Sampling Policy

Our sampling policy combines two key ideas. First, we leverage the ability of augmented Chebyshev scalarizations to *transform* a multi-objective optimization problem into multiple single-objective optimization problems, as we describe below. At every iteration, we then randomly select one of the single-objective problems obtained by fixing an augmented Chebyshev scalarization and sample a pair of designs independently according to the posterior probability of being a solution to this problem. The second step draws inspiration from the self-sparring sampling policy for multi-dueling bandits (Sui et al., 2017) and its extension to PBO (Siivola et al., 2021). We call our sampling policy *scalarized dueling Thompson sampling*.

Augmented Chebyshev Scalarizations For a given objective function, f, and a set of parameters, $\theta \in \Theta = \{\theta \in \mathbb{R}^m : \sum_{j=1}^m \theta_j = 1 \text{ and } \theta_j \geq 0,\ j = 1, \ldots, m\}$, a Chebyshev scalarization function is defined by

$$c(x \mid \theta, f) = \min_{j=1,\ldots,m}\{\theta_j f_j(x)\} + \rho \sum_{j=1}^m \theta_j f_j(x),$$

where ρ is a small positive constant. It can be shown that any solution of the problem $\max_{x \in \mathbb{X}} c(x \mid \theta, f)$ lies in the Pareto-optimal set of f. Conversely, if ρ is small enough, every point in the Pareto-optimal set of f is a solution of $\max_{x \in \mathbb{X}} c(x \mid \theta, f)$ for some $\theta \in \Theta$ (Miettinen, 1999). This is often used as a mechanism to transform a multi-objective optimization problem into multiple single-objective optimization problems obtained by drawing many scalarization parameters uniformly at random over Θ. In Bayesian optimization, in particular, this is the approach pursued by the seminal work of Knowles (2006). We also leverage this to derive a sound sampling policy in our setting.

Scalarized Dueling Thompson Sampling Formally, our sampling policy is defined as follows. At each iteration n, we draw a sample from the scalarization parameters uniformly at random over Θ, denoted by $\tilde{\theta}_n$. We also draw two independent samples, denoted by $\tilde{f}_{n,1}$ and $\tilde{f}_{n,2}$, from the posterior distribution on f given $\mathcal{D}_n$. The next query is then given by $X_n = (x_{n,1}, x_{n,2})$, where

$$x_{n,i} \in \underset{x \in \mathbb{X}}{\operatorname{argmax}}\, c(x \mid \tilde{\theta}_n, \tilde{f}_{n,i}),\ i = 1, 2.$$

Intuitively, our sampling policy works as follows. First, $\tilde{\theta}_n$ determines a subset of the Pareto-optimal set of f, namely, $\mathbb{X}^*_{\tilde{\theta}_n} = \operatorname{argmax}_{x \in \mathbb{X}} c(x \mid \tilde{\theta}_n, \tilde{f}_{n,i})$. Then, each $x_{n,i}$ is sampled according to the probability induced by the posterior distribution on f that $x_{n,i} \in \mathbb{X}^*_{\tilde{\theta}_n}$, in the same vein as standard dueling posterior sampling. The DM's responses allow us to learn about the relative order of $x_{n,1}$ vs. $x_{n,2}$ for each of the attributes, which in turn allow us to learn about $\mathbb{X}^*_{\tilde{\theta}_n}$. Finally, since $\tilde{\theta}_n$ is being drawn independently at each iteration, we are able to learn for a diverse collection of subsets $\mathbb{X}^*_{\tilde{\theta}_n}$ within $\mathbb{X}^*$.

5. Numerical Experiments

We compare the performance of our sampling policy (SDTS) against random sampling (Random), which samples each query uniformly at random over the design space. We also include the performance of a standard PBO approach fed with inconsistent overall preference feedback (PBO-DTS-IF), which we describe in detail below. The posterior distributions for both SDTS and PBO-DTS-IF are approximated via the variational inducing point approach of Hensman et al. (2015) using the implementation provided by Astudillo et al. (2023). Approximate samples from the posterior distribution used by both SDTS and I-PBO-DTS are obtained via 1000 random Fourier features (Rahimi & Recht, 2007).

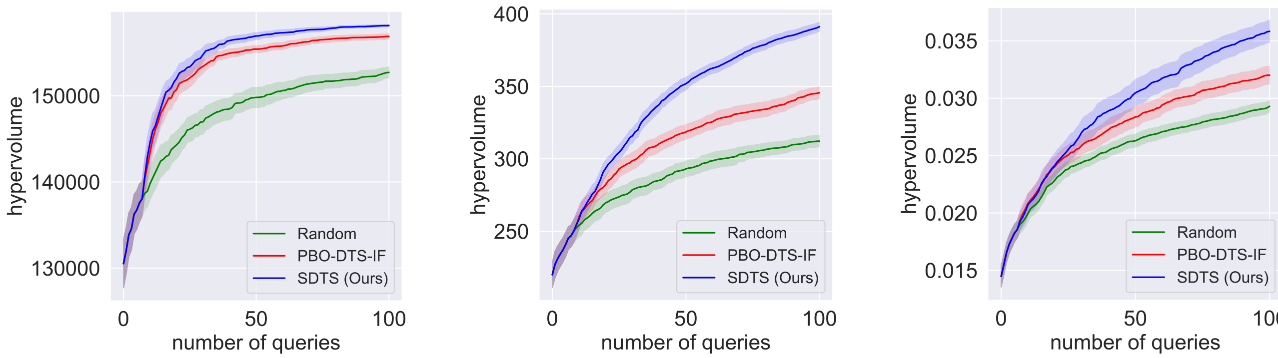
We report performance across two synthetic test problems (DTLZ1 and Car Side Impact) and a simulated exoskeleton gait design task (Exoskeleton). Details for these test problems are provided below. In all problems, an initial data set is obtained using $2(d + 1)$ queries chosen uniformly at random over $\mathbb{X}^2$, where d is the input dimension of the problem. After this initial stage, each algorithm was used to select 100 additional queries sequentially. Figure 2 shows the mean of the hypervolume of the designs included in queries thus far, plus and minus 1.96 times the standard deviation divided by the square root of the number of replications. Each experiment was replicated 30 times using different initial data sets. In all problems, the DM's responses are corrupted by low levels of Gumbel noise (which is consistent with the use of a Logistic likelihood).

PBO with Inconsistent Overall Preference Feedback As a baseline, we also include a standard PBO approach using inconsistent overall preference feedback. Such feedback is produced as follows. At each iteration n, a set of scalarization parameters $\tilde{\theta}_n$, is drawn uniformly at random over Θ. We assume the DM then provides a noisy response to the query "Is $c(x_1; \tilde{\theta}_n, f) > c(x_2; \tilde{\theta}_n, f)$?". These responses are used to fit a (single-output) Gaussian process with a Logistic likelihood. New queries are generated following the standard dueling Thompson sampling (a.k.a. dueling posterior sampling) strategy under this probabilistic model. We argue this baseline mimics a practical scenario where standard PBO is used under inconsistent preference feedback. Inconsistency arises from sampling different scalarization parameters at every iteration, imitating the DM's need to explore the Pareto front before committing to a solution. The performance of this method is expected to be poor when the Pareto-optimal set is large, i.e., when the trade-offs between attributes are significant.

DTLZ1 and Car Side Impact These test functions are standard benchmarks from the multi-objective optimization literature. DTLZ1 has $m = 2$ attributes and $d = 6$ design variables. We refer the reader to Deb et al. (2005) for more details. The car side impact test function is designed to emulate various metrics of interest in the context of crashworthiness vehicle design. This problem has $m = 4$ attributes and $d = 7$ design variables. We refer the reader to Tanabe & Ishibuchi (2020) for further details. Results for these two experiments can be found in Figures 2b and 2c. Our approach outperforms PBO with inconsistent preference feedback and random sampling consistently.

Atalante Exoskeleton Simulation We evaluate our algorithm on a surrogate preference landscape model of the lower-body exoskeleton Atalante over a gait design space. The surrogate model was built by fitting an independent (regular) Gaussian process to each attribute using the results of 500 simulations over gait designs drawn uniformly at random over the design space.

We parameterize the gait design space through the following five constraints used in the gait generation process: step length, minimum center of mass position with respect to stance foot in sagittal and coronal plane, minimum foot clearance, and percentage of the gait at which the minimum foot clearance is enforced. For each set of constraints, a nonlinear optimization problem is solved using FROST toolbox (Hereid & Ames, 2017) to generate the gait. This gait is then simulated in Mujoco to obtain the corresponding attributes. In our experiment, we assume that user preference can be described by the following four attributes: average speed (faster speed is preferred), maximum pelvis acceleration (as an approximation of trajectory smoothness), the center of



(a) DTLZ1 (b) Car Side Impact (c) Exoskeleton

Figure 2: Our framework was demonstrated on three test problems: DTLZ1 (left); Car Side Impact (center); and Exoskeleton (right). As illustrated, our proposed method (SDTS) achieves larger hypervolume values in fewer samples compared to random sampling (Random) and standard PBO with inconsistent preference feedback (PBO-DTS-IF).

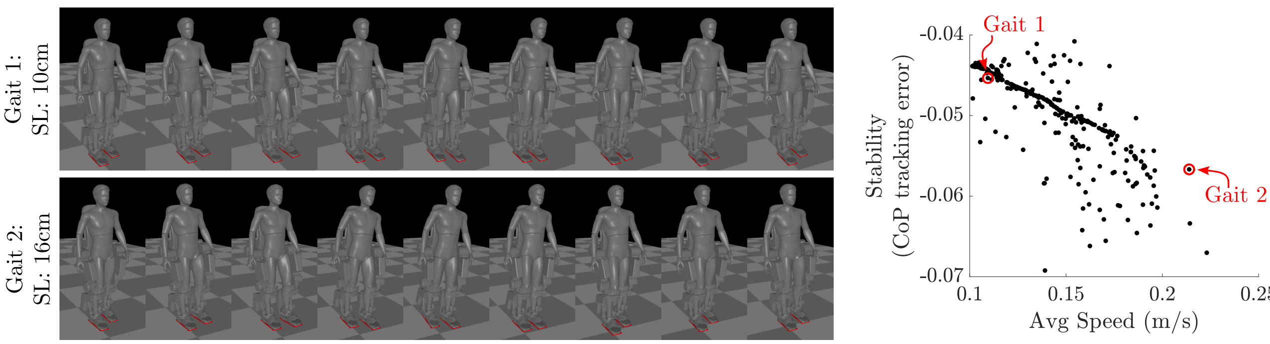


Figure 3: Example of two exoskeleton gaits with different attributes. The top gait illustrates a gait with a high stability and low speed, while the bottom gait illustrates a gait with low stability but high speed.

mass tracking error (as an approximation of stability), and center of pressure tracking error (as another approximation of stability). Example gait tiles can be found in Figure 3.

The results of this experiment are illustrated in Figure 2c. As in the previous examples, our proposed approach accelerated the rate of Pareto front exploration (evaluated by hypervolume) compared to benchmark methods.

6. Conclusion and Future Work

In this work, we proposed a framework for preferential multi-attribute BO, where the goal is to help decision-makers explore the Pareto front of their preferences over attributes of interest. To our knowledge, our work is the first one to consider optimization of multiple *latent* attributes using preference feedback, both in and outside the BO framework. We argued and provided empirical evidence that this approach has multiple advantages over the standard PBO approach. Within our proposed framework, we developed a Thompson sampling-based strategy to select queries. We showed this strategy provides a better exploration of the Pareto front than random sampling and a standard preferential Bayesian optimization approach fed with inconsistent preference feedback across three test problems, including a simulated exoskeleton gait customization task.

There are many exciting directions for future work. Our framework is currently limited to the use of preferences over individual attributes. However, in some situations, the decision-maker may be able to articulate overall preferences over a pair of designs. A framework able to combine such sporadically available overall preferences with preferences over individual attributes could help focus the search of new queries in regions of the Pareto front that are more relevant to the decision-maker, thus increasing sampling efficiency (Astudillo & Frazier, 2020; Lin et al., 2022). We are also interested in developing more principled sampling policies within our framework. For example, the recent work of Astudillo et al. (2023) provided an efficient scheme to approximately compute the one-step optimal policy in the standard PBO setting and showed it outperforms several popular baseline methods. It would be interesting to study if such an approach could be extended to our setting.

References

Astudillo, R. and Frazier, P. Multi-attribute Bayesian optimization with interactive preference learning. In *International Conference on Artificial Intelligence and Statistics*, pp. 4496–4507. PMLR, 2020.

Astudillo, R., Lin, Z. J., Bakshy, E., and Frazier, P. qEUBO: A decision-theoretic acquisition function for preferential Bayesian optimization. In *International Conference on Artificial Intelligence and Statistics*, pp. 1093–1114. PMLR, 2023.

Baskin, J. P. and Krishnamurthi, S. Preference aggregation in group recommender systems for committee decision-making. In *Proceedings of the Third ACM Conference on Recommender Systems*, pp. 337–340, 2009.

Baumeister, D. and Rothe, J. Preference aggregation by voting. In *Economics and Computation*, pp. 197–325. Springer, 2016.

Belakaria, S., Deshwal, A., and Doppa, J. R. Max-value entropy search for multi-objective Bayesian optimization. *Advances in Neural Information Processing Systems*, 32, 2019.

Branke, J. and Deb, K. Integrating user preferences into evolutionary multi-objective optimization. In *Knowledge Incorporation in Evolutionary Computation*, pp. 461–477. Springer, 2005.

Brochu, E., Brochu, T., and De Freitas, N. A Bayesian interactive optimization approach to procedural animation design. In *Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pp. 103–112, 2010.

Chu, W. and Ghahramani, Z. Preference learning with Gaussian processes. In *Proceedings of the 22nd international conference on Machine learning*, pp. 137–144, 2005.

Cosenza, Z., Astudillo, R., Frazier, P. I., Baar, K., and Block, D. E. Multi-information source Bayesian optimization of culture media for cellular agriculture. *Biotechnology and bioengineering*, 119(9):2447–2458, 2022.

Daulton, S., Balandat, M., and Bakshy, E. Differentiable expected hypervolume improvement for parallel multi-objective Bayesian optimization. *Advances in Neural Information Processing Systems*, 33:9851–9864, 2020.

Deb, K. Multi-objective optimization. In *Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques*, pp. 403–449. Springer, 2013.

Deb, K., Thiele, L., Laumanns, M., and Zitzler, E. Scalable test problems for evolutionary multiobjective optimization. In *Evolutionary Multiobjective Optimization*, pp. 105–145. Springer, 2005. doi: 10.1007/1-84628-137-7_6.

Dyer, J. S. and Sarin, R. K. Group preference aggregation rules based on strength of preference. *Management Science*, 25(9):822–832, 1979.

González, J., Dai, Z., Damianou, A., and Lawrence, N. D. Preferential Bayesian optimization. In *International Conference on Machine Learning*, pp. 1282–1291. PMLR, 2017.

Griffiths, R.-R. and Hernández-Lobato, J. M. Constrained Bayesian optimization for automatic chemical design using variational autoencoders. *Chemical science*, 11(2): 577–586, 2020.

Hakanen, J. and Knowles, J. D. On using decision maker preferences with ParEGO. In *Evolutionary Multi-Criterion Optimization*, pp. 282–297. Springer, 2017.

Hensman, J., Matthews, A., and Ghahramani, Z. Scalable variational Gaussian process classification. In *Artificial Intelligence and Statistics*, pp. 351–360. PMLR, 2015.

Hereid, A. and Ames, A. D. FROST: Fast robot optimization and simulation toolkit. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 719–726. IEEE, 2017.

Khan, N., Goldberg, D. E., and Pelikan, M. Multi-objective Bayesian optimization algorithm. In *Proceedings of the 4th Annual Conference on Genetic and Evolutionary Computation*, pp. 684–684, 2002.

Knowles, J. ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Transactions on Evolutionary Computation*, 10(1):50–66, 2006.

Lin, Z. J., Astudillo, R., Frazier, P., and Bakshy, E. Preference exploration for efficient Bayesian optimization with multiple outcomes. In *International Conference on Artificial Intelligence and Statistics*, pp. 4235–4258. PMLR, 2022.

Marler, R. T. and Arora, J. S. Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization*, 26:369–395, 2004.

Miettinen, K. Nonlinear multiobjective optimization. International Series in Operations Research & Management Science. Springer, 1999.

Mikkola, P., Todorović, M., Järvi, J., Rinke, P., and Kaski, S. Projective preferential Bayesian optimization. In *International Conference on Machine Learning*, pp. 6884–6892. PMLR, 2020.

Nguyen, Q. P., Tay, S., Low, B. K. H., and Jaillet, P. Top-k ranking Bayesian optimization. In *Proceedings of the*

AAAI Conference on Artificial Intelligence, volume 35, pp. 9135–9143, 2021.

Nielsen, J. B. B., Nielsen, J., and Larsen, J. Perception-based personalization of hearing aids using Gaussian processes and active learning. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(1):162–173, 2015.

Novoseller, E., Wei, Y., Sui, Y., Yue, Y., and Burdick, J. Dueling posterior sampling for preference-based reinforcement learning. In *Conference on Uncertainty in Artificial Intelligence*, pp. 1029–1038. PMLR, 2020.

Rahimi, A. and Recht, B. Random features for large-scale kernel machines. *Advances in Neural Information Processing Systems*, 20, 2007.

Siivola, E., Dhaka, A. K., Andersen, M. R., González, J., Moreno, P. G., and Vehtari, A. Preferential batch Bayesian optimization. In *2021 IEEE 31st International Workshop on Machine Learning for Signal Processing*, pp. 1–6. IEEE, 2021.

Sui, Y., Zhuang, V., Burdick, J. W., and Yue, Y. Multi-dueling bandits with dependent arms. *arXiv preprint arXiv:1705.00253*, 2017.

Tanabe, R. and Ishibuchi, H. An easy-to-use real-world multi-objective optimization problem suite. *Applied Soft Computing*, 89:106078, 2020.

Tucker, M., Cheng, M., Novoseller, E., Cheng, R., Yue, Y., Burdick, J. W., and Ames, A. D. Human preference-based learning for high-dimensional optimization of exoskeleton walking gaits. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3423–3430. IEEE, 2020a.

Tucker, M., Novoseller, E., Kann, C., Sui, Y., Yue, Y., Burdick, J. W., and Ames, A. D. Preference-based learning for exoskeleton gait optimization. In *2020 IEEE International Conference on Robotics and Automation*, pp. 2351–2357. IEEE, 2020b.

Wang, H., Olhofer, M., and Jin, Y. A mini-review on preference modeling and articulation in multi-objective optimization: Current status and challenges. *Complex & Intelligent Systems*, 3:233–245, 2017.

Wu, J., Chen, X.-Y., Zhang, H., Xiong, L.-D., Lei, H., and Deng, S.-H. Hyperparameter optimization for machine learning models based on Bayesian optimization. *Journal of Electronic Science and Technology*, 17(1):26–40, 2019.

Young, H. P. A note on preference aggregation. *Econometrica: Journal of the Econometric Society*, pp. 1129–1131, 1974.

Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C. M., and Da Fonseca, V. G. Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Transactions on Evolutionary Computation*, 7(2):117–132, 2003.