Looping in the Human: Collaborative and Explainable Bayesian Optimization

Masaki Adachi[1,2,3] **Brady Planden**[2] **David A. Howey**[2,4] **Michael A. Osborne**[1,2]
Sebastian Orbell[2] **Natalia Ares**[2] **Krikamol Muandet**[5] **Siu Lun Chau**[5]
[1]Machine Learning Research Group, University of Oxford, OX2 6ED, United Kingdom
[2]Department of Engineering Science, University of Oxford, OX1 3PJ, United Kingdom
[3]Toyota Motor Corporation, Shizuoka 410-1107, Japan
[4]The Faraday Institution, Harwell Campus, Didcot OX11 0RA, United Kingdom
[5]CISPA Helmholtz Center for Information Security, 66123 Saarbrücken, Germany

Abstract

Like many optimizers, Bayesian optimization often falls short of gaining user trust due to opacity. While attempts have been made to develop human-centric optimizers, they typically assume user knowledge is well-specified and error-free, employing users mainly as supervisors of the optimization process. We relax these assumptions and propose a more balanced human-AI partnership with our Collaborative and Explainable Bayesian Optimization (CoExBO) framework. Instead of explicitly requiring a user to provide a knowledge model, CoExBO employs preference learning to seamlessly integrate human insights into the optimization, resulting in algorithmic suggestions that resonate with user preference. CoExBO explains its candidate selection every iteration to foster trust, empowering users with a clearer grasp of the optimization. Furthermore, CoExBO offers a no-harm guarantee, allowing users to make mistakes; even with extreme adversarial interventions, the algorithm converges asymptotically to a vanilla Bayesian optimization. We validate CoExBO's efficacy through human-AI teaming experiments in lithium-ion battery design, highlighting substantial improvements over conventional methods. Code is available `https://github.com/ma921/CoExBO`.

Proceedings of the 27th International Conference on Artificial Intelligence and Statistics (AISTATS) 2024, Valencia, Spain. PMLR: Volume 238.

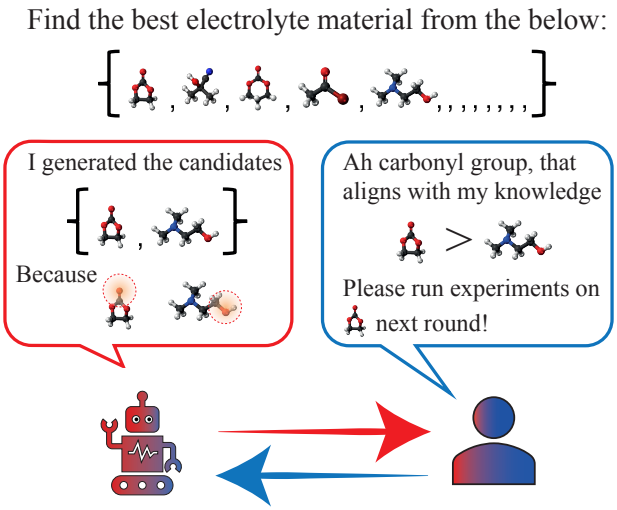

Figure 1: In Collaborative and Explainable Bayesian Optimization (CoExBO), a human expert collaborates with BO to refine electrolyte materials. While experts excel in discerning material differences rather than identifying the best one, pairwise comparisons and explanations boost their feedback accuracy and trust. This guides the BO to produce better candidates, ensuring quicker convergence.

1 Introduction

Bayesian optimization (BO) is a popular blackbox optimizer for expensive-to-evaluate tasks. While it is widely applied in diverse domains (Feurer et al., 2015; Wu et al., 2020; Adachi, 2021), it has yet to fully gain human users' trust. Surveys from NeurIPS2019/ICLR2020 (Bouthillier & Varoquaux, 2020) found that most AI researchers prefer manually tuning hyperparameters. This is surprising given that Bergstra & Bengio (2012) has shown manual search performs worse than simple random search and lacks

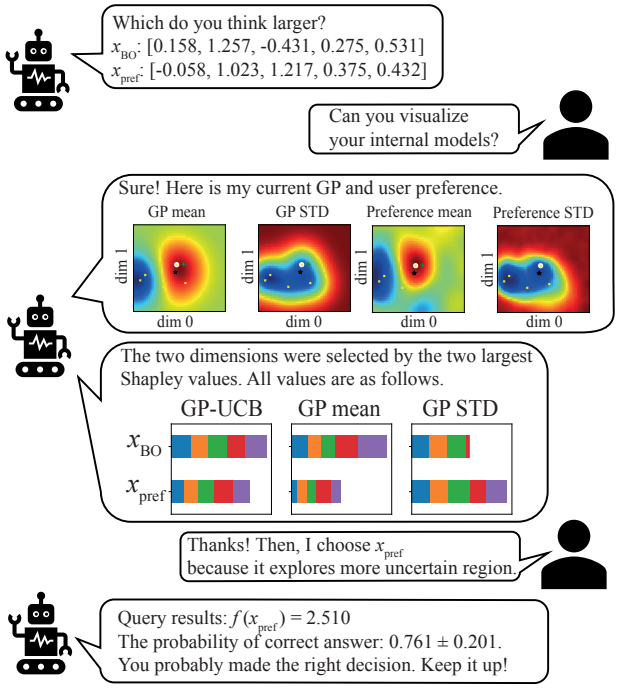

Figure 2: Explanation flow: **Spatial relation:** BO visualizes the surrogate model's predictive distribution and estimated human preference models for the two primary dimensions determined by Shapley values. **Feature importance:** Users' values are provided for both candidates' predictive mean, standard deviation, and acquisition function. **Selection accuracy feedback:** After observing the function value, a post-hoc evaluation of the correct selection probability is given.

global convergence guarantees (Gupta et al., 2023).

To make BO trustworthy, recent research has moved towards human-AI collaborative paradigms (Kanarik et al., 2023). These methods often make contrasting assumptions about human exploration abilities. Suppose humans are superior to BO (Colella et al., 2020; AV et al., 2022), their intervention can enhance convergence—but this potent assumption also implies manual search would be superior to BO, undermining the core justification for using BO. Conversely, if humans are viewed as imperfect agents, existing works such as Gupta et al. (2023); Khoshvishkaie et al. (2023) treat humans as a central optimizer, and BO supports human manual search via exploratory adjustment, that can assure the global convergence even with erroneous human selection. When one side is better, the inferior side's selection is wasteful, leading to a worse convergence rate than vanilla BO (Khoshvishkaie et al., 2023).

To build a balanced human-BO partnership, we believe an ideal method should satisfy the following criteria: (a) **Explainability:** The method should enhance user understanding of the optimization process, promoting transparency. While Li & Adams (2020) introduced explainability by limiting the search space, it may overly restrict it, lacking a global convergence guarantee. (b) **User-centric knowledge integration:** An ideal approach should seamlessly incorporate human insights from user interactions without requiring users to define an exact knowledge model. For instance, materials often exist in a high-dimensional feature space, whereas BO typically uses a reduced low-dimensional feature set due to limited chemistry data (Jordan, 2019). Chemists have discernment to assess materials using information inaccesible to the model but struggle to articulate it quantitatively (Cisse et al., 2023). Hence, users face challenges with existing methods; e.g., Hvarfner et al. (2022) mandates a prior function capturing the users' optimal location belief, while AV et al. (2022) requires users to select the next query point quantitatively. In short, knowledge elicitation in high-dimensional domains is notoriously intricate (Rousseau, 2001; Garthwaite et al., 2005; Mikkola et al., 2023a). (c) **Robustness:** The method should be robust against human errors, offering a no-harm guarantee to ensure that even adversarial interventions do not adversely impact the vanilla BO convergence rate. To the best of our knowledge, only Hvarfner et al. (2022) can theoretically assure the no-harm guarantee. Crucially, none encompass all three comprehensively.

This paper introduces the *Collaborative and Explainable Bayesian Optimization* (CoExBO) framework to tackle the above challenges. For criterion (a), CoExBO employs Shapley values (Shapley, 1953), a cornerstone of explainable AI, to ensure users can effectively understand and interpret the candidate acquisition mechanism. Addressing criterion (b), CoExBO deviates from conventional methods that require users to input an exact knowledge model. Instead, it presents users with candidate pairs, empowering them to select the perceived optimal one. This approach allows CoExBO to implicitly assimilate human insights via preference learning (Bradley & Terry, 1952). This is grounded that humans excel at relative comparisons rather than quantifying an absolute preference for a singular choice (Kahneman & Tversky, 1979). For criterion (c), inspired by Hvarfner et al. (2022), our candidate generation strategy prioritizes expert knowledge in the early optimization stages. As more experimental data accumulates and refines the surrogate model, the influence of human input gradually diminishes. Theorem 2 proves this offers a no-harm guarantee.

Our contributions are summarized as:

1. We introduce CoExBO, a novel framework promoting a balanced human-BO partnership. CoExBO is characterized by its transparency, capacity to assimilate human insights seamlessly, and resilience against human errors.

2. We establish the efficacy of CoExBO via real-world optimization tasks on lithium-ion battery design problems, and the expert-BO team can gain significant speedup over eight popular baselines.

2 Bayesian optimization and existing human-in-the-loop extensions

Bayesian optimization. We aim to optimize the function f when f can only be queried pointwise.

$$x^*_{\text{true}} = \underset{x\in\mathcal{X}}{\text{argmax}}\, f(x), \tag{1}$$

where $\mathcal{X} \subseteq \mathbb{R}^d$ is the d dimensional continuous input domain, and x^*_{true} is the global optimum. We assume that $f(x)$ is costly to query and can only observe a noisy estimate $y = f(x) + \epsilon$ with i.i.d. zero-mean Gaussian noise ϵ. The goal is to find the optimal $f(x)$ under a given number of queries.

BO (Mockus, 1998; Garnett, 2023) is a model-based black box optimizer that employs a Gaussian process (GP) (Williams & Rasmussen, 2006) as a surrogate model. At optimization step t, the GP approximates the true function f using the current observations $\mathbf{D}_{\text{obj}_t} = (\mathbf{X}_{\text{obj}_t}, \mathbf{y}_{\text{obj}_t})$ as $f_t \sim \mathcal{GP}(\mu_t, \kappa_t)$, with

$$\mu_t(x) = k(x, \mathbf{X}_{\text{obj}_t})\mathbf{K}'^{-1}_{\mathbf{XX}t}\mathbf{y}_{\text{obj}_t},$$
$$\kappa_t(x, x') = k(x, x') - k(x, \mathbf{X}_{\text{obj}_t})\mathbf{K}'^{-1}_{\mathbf{XX}t}k(\mathbf{X}_{\text{obj}_t}, x),$$

where μ_t and κ_t are the GP's posterior predictive mean and covariance functions at round t. k is the kernel, $\lambda > 0$ is the Gaussian likelihood variance, $\mathbf{K}_{\mathbf{XX}t} := k(\mathbf{X}_{\text{obj}_t}, \mathbf{X}_{\text{obj}_t})$, $\mathbf{K}'^{-1}_{\mathbf{XX}t} := (\mathbf{K}_{\mathbf{XX}t} + \lambda\mathbf{I})^{-1}$, and $\mathbf{I} \in \mathbb{R}^{d\times d}$ is the identity matrix. Using GP predictive uncertainty, BO solves the blackbox optimization problem as active learning and selects the next query point by maximising an acquisition function (AF). One popular class of AF is the upper confidence bound (UCB) (Srinivas et al., 2010): $\alpha_{f_t}(x) := \mu_t(x) + \beta_t^{1/2}\sigma_t(x)$, where $\sigma_t(x) := \sqrt{\kappa_t(x, x)}$ is the standard deviation of the GP predictive posterior, β_t is the user-specified parameter indicating the trade-off between exploitation (using current knowledge of optimum from μ_t) and exploration (considering the uncertainty from σ_t).

Human-in-the-loop extensions. There are four prevailing approaches to integrate human knowledge into BO: (1) By treating human knowledge as a prior over the input space (Souza et al., 2021; Ramachandran et al., 2020; Hvarfner et al., 2022; Cisse et al., 2023). (2) By adopting a hyperprior over the function space (Hutter et al., 2011; Snoek et al., 2014; Wang et al., 2021). (3) By considering it as a multi-fidelity information source (Song et al., 2019; Huang et al., 2022). (4) Implementing human knowledge as hard constraints (Gelbart et al., 2014; Hernández-Lobato et al., 2015; Adachi et al., 2024). Among these categories, only the work of Hvarfner et al. (2022), belonging to the first category, provides a no-harm guarantee against potential human errors. Notably, all these methods operate under the assumption that human knowledge can be well specified to the algorithm. A more detailed related work section is delineated in Supplementary B.

πBO. CoExBO is inspired by the πBO algorithm (Hvarfner et al., 2022), which characterizes human knowledge as a prior distribution representing their belief in the global optimum location, i.e., $\pi(x) := \mathbb{P}\big(f(x) = \max_{x'\in\mathcal{X}} f(x')\big)$. This prior can then be incorporated into an AF α_t to act as a soft constraint for a warmer start on the optimization. Specifically, at round t, we search for $x_{\text{next}} = \text{argmax}_{x\in\mathcal{X}}\ \alpha_t(x)\pi(x)^{\gamma/t}$ where $\gamma > 1$ controls the decay rate of this constraint. This decay of human contribution is justified as follows: at the start of the BO, expert knowledge can help substantially, whereas, at later stages, the BO will likely have enough data to reach the optimum confidently. Furthermore, this decaying property is the key reason behind the no-harm guarantee in Corollary 1 in Hvarfner et al. (2022).

While πBO's formulation is straightforward, requiring the user to specify a prior over high-dimensional input space could be very challenging in practice (Garthwaite et al., 2005). The following section demonstrates how CoExBO can relax this assumption by interacting with users through preference elicitation.

3 Collaborative and Explainable BO

In this section, we present our Collaborative and Explainable Bayesian Optimization (CoExBO) algorithm. While its objective aligns with the conventional BO objective (Eq. 1), CoExBO differentiates itself by explaining the acquisition process and incorporating human knowledge through preference learning. Specifically, the query procedure consists of the following steps: At round $t > 1$ with surrogate GP f_t and preference model $\hat{\pi}_t$, we have

$$\Gamma(f_t, \hat{\pi}_t) \to (x_1, x_2), \qquad \text{(Acquire candidates)}$$
$$\text{E}(f_t, x_1, x_2) \to (\boldsymbol{\phi}_1, \boldsymbol{\phi}_2) \qquad \text{(Explain acquisition)}$$
$$\text{H}\left(\{(x_i, \boldsymbol{\phi}_i)\}_{i=1}^2\right) \to \tilde{x} \in (x_1, x_2), \quad \text{(Elicit preference)}$$
$$\mathbf{\Pi}\left(\hat{\pi}_t, \tilde{x}, x_1, x_2\right) \to \hat{\pi}_{t+1}, \qquad \text{(Update } \hat{\pi}_t\text{)}$$

and finally we run the experiment with $\tilde{x}$ to obtain $y_{\text{next}} = f(\tilde{x})$. We denote Γ as the candidate generation function (see §3.2) that takes in the surrogate and current preference models and generates a pair of candidates. E is an explanation function that explains the acquisition process (see §3.3) and returns

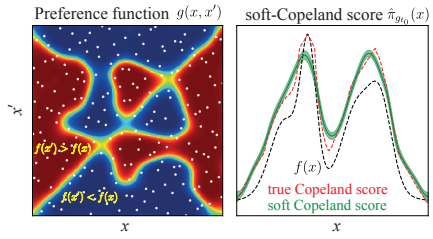

Figure 3: Preference learning concepts: we aim to model the ordinal relationship $f(x) < f(x')$ and its inverse with GP, utilizing the dataset $D^{t_0}_{\text{pref}}$, represented by white dots. Soft-Copeland score is used for the proxy of true function estimate.

explanation $\boldsymbol{\phi}_i$ for the i^{th} candidate. H denotes the human users' choice when pairs of candidates and the subsequent explanations are given. Preference function $\hat{\pi}_t$ is then updated to $\hat{\pi}_{t+1}$ by taking into account the human preference through an update function $\mathbf{\Pi}$ (see §3.1), and at last we end the iteration by running an experiment on the chosen candidate $\tilde{\boldsymbol{x}}$.

3.1 Model human knowledge through preference learning

While πBO requires the user to provide the preference function π explicitly—which might be challenging to elicit in practice—we relax this assumption and estimate π by $\hat{\pi} : \mathcal{X} \to \mathbb{R}_+$ using preference learning. At its core, preference learning aims to model the order relationships among a candidate set, $\mathcal{X}$. For this paper, our emphasis is on binary preference learning, as detailed in (Bradley & Terry, 1952; Chau et al., 2022b). However, this concept can be straightforwardly extended to more complex preference models such as choice functions (Benavoli et al., 2023b).

To initiate the optimization process, we randomly select candidate pairs from $\mathcal{X}$, then solicit the users' opinion on which one is more likely the optimal location. Formally, at $t = t_0$, we sample J_{t_0} binary comparisons, denoted as $D^{t_0}_{\text{pref}} := \{x_1^{(j)}, x_2^{(j)}, y_{\text{pref}}^{(j)}\}_{j=1}^{J_{t_0}}$, where y_{pref} is 1 if $x_1^{(j)}$ is preferred over $x_2^{(j)}$, and 0 otherwise. Using $D^{t_0}_{\text{pref}}$, we can construct a binary preference function $g : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ based on the following likelihood model for any candidate pair x_1, x_2:

$$\mathbb{P}(y_{\text{pref}} \mid x_1, x_2) = S(y_{\text{pref}}; g(x_1, x_2)),$$

where $S(y_{\text{pref}}; z) := z^{y_{\text{pref}}}(1-z)^{1-y_{\text{pref}}}$ is the Bernoulli likelihood and $g(x_1, x_2)$ denotes the users degree of preference of x_1 over x_2. There are various ways to learn such a g (Bradley & Terry, 1952; Chau et al., 2022a). We choose to model g with a GP to encapsulate the inherent estimation uncertainty, pivotal for designing an acquisition in §3.2 that considers both surrogate and preference model's uncertainties. As the exact method we choose to learn g is not the primary focus of this work, we defer this discussion in Supplementary C.

Figure 3 visualizes the GP preference function g_{t_0} on the left, and $\hat{\pi}_{t_0}$ on the right, estimated by:

$$\hat{\pi}_{g_{t_0}}(x) := \int_{\mathcal{X}} g_{t_0}(x, x_2) dx_2. \tag{2}$$

This approach mirrors the soft-Copeland score[1] in González et al. (2017) and models the (unnormalized) likelihood of x being the Condorcet winner[2]. Hence, we can integrate out g_{t_0} and obtain the following representation of our estimated user preference

$$\hat{\pi}_{g_{t_0}}(x) \sim \mathcal{N}(\mathbb{E}_{g_{t_0}}[\hat{\pi}_{g_{t_0}}(x)], \mathbb{V}_{g_{t_0}}[\hat{\pi}_{g_{t_0}}(x)]).$$

Importantly, while the soft-Copeland score does not replicate the original function $f(x)$, the location of its maximum still corresponds to the maximum of $f(x)$. Furthermore, the maximum of the soft-Copeland score converges to the true maximum as the dataset size increases, namely $\lim_{t\to\infty} \operatorname{argmax}_{x\in\mathcal{X}} \pi_{g_t}(x) = \operatorname{argmax}_{x\in\mathcal{X}} f(x)$ if $|\mathcal{X}| < \infty$. As the optimization progresses and the acquisition of more binary comparison data, we can iteratively update the posterior of g_{t-1} via Bayes' theorem and recalibrate the preference model $\hat{\pi}_t$ at each iteration t.

3.2 Candidate generation with no-harm guarantee

New acquisition function. Following Hvarfner et al. (2022), we can use $\mathbb{E}_{g_t}[\hat{\pi}_{g_t}]$ as a discounting factor for α_{f_t} and redefine it as $\alpha_{f_t}(\cdot)\mathbb{E}_{g_t}[\hat{\pi}_{g_t}(\cdot)]^{\frac{\gamma}{t}}$. However, this approach does not consider the predictive uncertainty of $\hat{\pi}_{g_t}$ represented by the GP model g, i.e., $\mathbb{V}_{g_t}[\hat{\pi}_{g_t}(x)] = \mathbb{E}_{g_t}[\hat{\pi}_{g_t}(x)(1-\hat{\pi}_{g_t}(x))]$. This could result in overly optimistic acquisitions. πBO's performance depends on the peak centre of π, which must be close to the true global optimum for speedup. However, the peak centre of $\mathbb{E}_{g_t}[\hat{\pi}_{g_t}]$ can be unreliable, especially when the user inputs are insufficient or inconsistent to confidently construct $\hat{\pi}_{g_t}$. Hence, we aim to utilize $\mathbb{E}_{g_t}[\hat{\pi}_{g_t}]$ information only when it is confident, i.e., when $\mathbb{V}_{g_t}[\hat{\pi}_{g_t}(x)]$ is sufficiently small.

To account for uncertainties in both the surrogate and preference model, we multiply the two Gaussians. For any $x \in \mathcal{X}$, $f_t(x)$ is a Gaussian random variable in the target space of f, and $\hat{\pi}_{g_t}$ is a Gaussian random variable indicating the likelihood of x being the Condorcet winner. We therefore scale the preference function to

[1] True Copeland score is calculated by Eq.(2) but using true π instead of estimated $\hat{\pi}$, thus there is no uncertainty.

[2] typically defined as the most favoured player within $\mathcal{X}$.

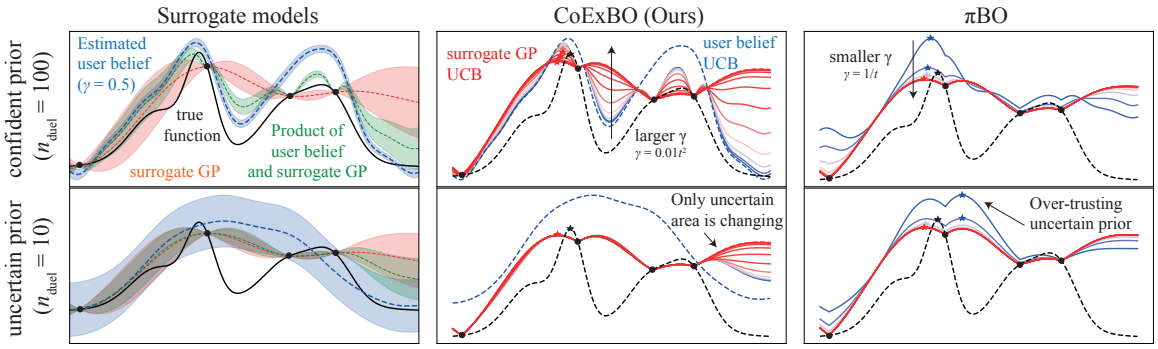

Figure 4: The CoExBO AF synthesizes GPs, utilizing one GP to represent the true function (red) and another to reflect user belief (blue), leading to the product GP (green). This product GP effectively assimilates the uncertainty inherent in user belief, adjusting the level of user belief integration during the acquisition process. Whereas the CoExBO AF is designed to adaptively manage the integration of uncertain user beliefs, the πBO approach tends to excessively depend on user belief, overlooking the uncertainty.

align with the surrogate model's scale. Using the property that the product of Gaussians is Gaussian, we derive a UCB-style AF.

Proposition 1. *Given* $f_t(x) \sim \mathcal{N}(\mu_{f_t}(x), \kappa_{f_t}(x,x))$, $\hat{\pi}_{g_t}(x) \sim \mathcal{N}(\mu_{g_t}(x), \sigma^2_{g_t}(x))$ *and a scaling function* ρ *that maps* $\hat{\pi}_{g_t}$ *to the scale of* f_t *and* $\gamma > 0$*, our new acquisition function* $\alpha_{f,\pi}$ *takes the following form:*

$$\alpha_{f_t,\hat{\pi}_t}(x) := \mu_{f_t,\hat{\pi}_t}(x) + \beta^{\frac{1}{2}} \sigma_{f_t,\hat{\pi}_t}(x) \tag{3}$$

where

$$\mu_{f_t,\hat{\pi}_t}(x) = \frac{\sigma^2_{f_t,\hat{\pi}_t}(x)}{\sigma^2_{\hat{\pi}_t}(x)} \mu_{\hat{\pi}_t}(x) + \frac{\sigma^2_{f_t,\hat{\pi}_t}(x)}{\sigma^2_{f_t}(x)} \mu_{f_t}(x), \tag{4}$$

$$\sigma^2_{f_t,\hat{\pi}_t}(x) = \frac{\sigma^2_{\hat{\pi}_t}(x)\sigma^2_{f_t}(x)}{\sigma^2_{\hat{\pi}_t}(x) + \sigma^2_{f_t}(x)}, \tag{5}$$

$$\mu_{\hat{\pi}_t}(x) := \rho(\mathbb{E}_{g_t}[\hat{\pi}_{g_t}(x)]), \tag{6}$$

$$\sigma^2_{\hat{\pi}_t}(x) := \rho^2(\mathbb{V}_{g_t}[\hat{\pi}_{g_t}(x)]) + \gamma t^2 \sigma^2_{f_t}(x), \tag{7}$$

$$\rho(x) := \mathbb{E}[\boldsymbol{y}_{obj_t}]x + \sqrt{\mathbb{V}[\boldsymbol{y}_{obj_t}]} \tag{8}$$

The new AF adheres to the following principle:

1. (Uncertainty in Preference Estimation) The information provided by $\hat{\pi}_t(x)$ becomes valuable only when its variance, denoted as $\mathbb{V}_{g_t}[\hat{\pi}_{g_t}(x)]$, is smaller than the variance of the surrogate model, represented as $\sigma^2_{f_t}(x)$.
2. (Uncertainty in Efficacy of User Information) $\hat{\pi}_{g_t}(x)$ becomes less significant as iterations proceed. This mirrors πBO's principle that human knowledge is most valuable in the early stages.

The simple product of two Gaussian distributions offers a heuristic solution to the above assumptions. Firstly, the resulting Gaussian mean is a weighted sum of two means weighted by their variances, aligning with our first assumption, where the mean $\mu_{f_t,\hat{\pi}_t}$ stays unchanged from μ_{f_t} when $\mathbb{V}_{g_t}[\hat{\pi}_{g_t}(x)] \gg \sigma^2_{f_t}(x)$. To address the second, we introduce a decay hyperparameter γ term in Eq. 7. When $\gamma t^2 \geq 1$, $\sigma^2_{\hat{\pi}_t}(x) > \sigma^2_{f_t}(x)$ holds, causing information from $\hat{\pi}_{g_t}$ to decay. While other methods could meet these principles, we choose the computationally simplest one.

Figure 4 illustrates typical behaviors of πBO and CoExBO. With a confident and accurate $\hat{\pi}_{g_t}$, both methods perform well. However, when dealing with uncertain $\hat{\pi}_{g_t}$, the peaks do not align with the true global maximum location. πBO tends to rely on user belief regardless of uncertainty, while CoExBO mitigates over-reliance on user belief in uncertain situations.

Candidate generation. Given f_t and $\hat{\pi}_t$, we generate a pair of candidates as follows:

$$x_1 = \arg\max_{x \in \mathcal{X}} \alpha_{f_t}(x) \qquad \text{(standard UCB)}$$

$$x_2 = \arg\max_{x \in \mathcal{X}} \alpha_{f_t,\hat{\pi}_t}(x) \qquad (\hat{\pi} \text{ incorporated UCB})$$

This approach is similar to other human-AI collaborative BO methods (Gupta et al., 2023; Khoshvishkaie et al., 2023), involving a direct comparison of BO with human recommendations. Opting for x_2 speeds up convergence if human input is superior while choosing x_1 is optimal if BO performs better. This represents a greedy approach to optimizing choices for both sides. A greedy approach is optimal in both scenarios since we assume that either human or BO is superior. The key difference lies in the decision-making process: in previous approaches, each agent independently selects their preferred option, necessitating separate queries. In contrast, our method employs BO to generate both candidates and then makes a selection. As our acquisition function $\alpha_{f,\hat{\pi}}$ gradually converges to the standard UCB, the selection process becomes equivalent over time. This prevents budget wastage resulting from suboptimal choices made by either side.

Note that our AF $\alpha_{f,\hat{\pi}}$ combines GP information and user beliefs, meaning it does not always align with user preferences. The GP component corrects any uncertainties or inaccuracies in user beliefs, preventing a persistent bias toward selecting x_2.

Regret analysis. By following Srinivas et al. (2010), we analyse the regret of preference-based AF $r_{\hat{\pi}_t} := f(x^*_{\text{true}}) - f(x_2)$ and the standrd UCB regret $r_t := f(x^*_{\text{true}}) - f(x_1)$. We assume good and bad user beliefs. A good user belief assumes to contain the true function within the standard deviation, whereas a bad user belief does with extra error with mean estimation.

Theorem 2. *Fix $t \geq 1$ and $\gamma > 0$. If $|f(\boldsymbol{x}) - \mu_{f_{t-1}}(\boldsymbol{x})| \leq \beta_t^{1/2}\sigma_{f_{t-1}}(\boldsymbol{x})$ for all $\boldsymbol{x} \in D$, $|D| < \infty$ hold, the ratio of regrets for with and without π augmentation $r_t, r_{\hat{\pi}_t}$ is bounded by:*
(Good user belief) If $|f(\boldsymbol{x}) - \mu_{f_{t-1},\hat{\pi}_{t-1}}(\boldsymbol{x})| \leq \beta_t^{1/2}\sigma_{f_{t-1},\hat{\pi}_{t-1}}(\boldsymbol{x})$ holds,

$$r_{\hat{\pi}_t} {r_t}^{-1} \leq R_{\hat{\pi}_t} < 1, \tag{9}$$

where
$R_{\hat{\pi}_t} = \sqrt{\frac{\rho^2(\mathbb{V}_{g_{t-1}}[\hat{\pi}_{g_{t-1}}(x_2)]) + \gamma(t-1)^2\sigma^2_{t-1}(x_2)}{\rho^2(\mathbb{V}_{g_{t-1}}[\hat{\pi}_{g_{t-1}}(x_2)]) + \gamma(t-1)^2\sigma^2_{t-1}(x_2) + \sigma^2_{t-1}(x_1)}}$.
(Bad user belief) If $|f(\boldsymbol{x}) - \mu_{f_{t-1},\hat{\pi}_{t-1}}(\boldsymbol{x})| \leq |\mu_{f_{t-1}}(\boldsymbol{x}) - \mu_{f_{t-1},\hat{\pi}_{t-1}}(\boldsymbol{x})| + \beta_t^{1/2}\sigma_{f_{t-1},\hat{\pi}_{t-1}}(\boldsymbol{x})$ holds,

$$r_{\hat{\pi}_t} {r_t}^{-1} \leq \Delta\mu_t + R_{\hat{\pi}_t t}, \tag{10}$$

where $\Delta\mu_t = \frac{|\mu_{t-1}(x_1) - \mu_{f_{t-1},\hat{\pi}_{t-1}}(x_2)|}{2\beta_t^{1/2}\sigma_{f_{t-1}}(x_1)}$.

The proof is given in Supplementary A. Using Theorem 2, we obtain the convergence rate of $\alpha_{f_{t-1},\hat{\pi}_{t-1}}$. This trivially follows the original convergence rate on UCB as in Srinivas et al. (2010):

Lemma 3. *(No harm guarantee) Given the regret in Theorem 2, The regret of a preference-based acquisition function, α_{pref_t}, asymptotically equals to the regret of an upper confidence bound strategy, UCB:*

$$\lim_{t\to\infty} r_t^{\pi} {r_t}^{-1} = 1, \tag{11}$$

so we obtain a convergence rate for $\alpha_{f_T,\hat{\pi}_T}$ of $\mathcal{O}(\sqrt{T\gamma_T \log T})$, an original UCB convergence rate.

Hence, we can ensure that the worst-case convergence rate remains unaffected, even with inaccurate user beliefs, for large t, where $\gamma t^2 \gg 1$. While short-term performance may not match the standard UCB, it often yields better empirical results. Particularly, a good user belief has a provably better regret bound. In our scenario, humans choose candidates from the UCB or $\alpha_{f_t,\hat{\pi}_t}$, reinforcing the no-harm guarantee. The human selection process does not impact the convergence rate because $\alpha_{f_t,\hat{\pi}_t}$ determines tighter or looser bounds than the UCB, depending on user beliefs. In practice, human knowledge evolves over iterations, positively influencing convergence, as demonstrated in our experiments section. The parameter γt^2 balances the integration of evolving human knowledge with the no-harm guarantee. It is worth noting that our approach differs from that of multitask GPs, as multitask GPs are vulnerable to unreliable low-fidelity GPs (Mikkola et al., 2023b).

3.3 Explaining candidate generation through Shapley values

To foster trust in the black-box optimizer among users, we employ Shapley values (Shapley, 1953), a popular solution concept from game theory adopted by the machine learning community (Lundberg & Lee, 2017; Chau et al., 2022c; Hu et al., 2022) to provide feature attributions for the acquisitions and the surrogate model. This provides users with a clearer understanding of the factors influencing the selection of candidates.

Shapley values follow a set of favourable rationality axioms, setting them apart from heuristic methods like extracting the length scale from a GP kernel. For a given function $h : \mathcal{X} \to \mathbb{R}$, a query location x, the Shapley value for feature j is expressed as

$$\phi_{j,x}(h) = \sum_{S \subseteq [d]\setminus\{j\}} c_{|S|} \left(\nu_{x,h}(S \cup i) - \nu_{x,h}(S)\right)$$

where $[d] := \{1, \ldots, d\}$, $c_{|S|} = \frac{1}{d}\binom{d-1}{|S|}^{-1}$, X_S is the subfeature vector of X for features in S and $\nu_{x,h}(S)$ measures a notion of contribution features S has to the prediction $h(x)$. We utilize the recently introduced GPSHAP (Chau et al., 2023) to explain the surrogate GP f. We illustrate how to estimate the Shapley values for α_f, but extending to $\alpha_{f,\pi}$ is straightforward. While in the Shapley explanation literature, it is suggested one should take the conditional expectation of the to-be-explained function, i.e. $\mathbb{E}_X[\alpha_f(X) \mid X_S = x_s]$, to structure the cooperative game. However, for computational reasons (see appendix), we opted to establish the game using its upper bound instead:

$$\nu_{x,f}(S) := \mathbb{E}[\mu_f(X) \mid X_S = x_S] + \beta_t^{1/2}\sqrt{\mathbb{E}[\kappa_f(X,X) \mid X_S = x_S]}.$$

$\nu_{x,f}(S)$ can be interpreted as the significance of the feature subset S measured by how much the upper bound has changed if we remove the contribution from other features in S^c by integration. This formulation allows us to estimate the quantity in analytical form:

Proposition 4. *Given $f \sim \mathcal{GP}(\mu, \kappa)$, for a given feature subset $S \subseteq \{1, \ldots, d\}$, and location x, $\nu_{x,f}(S)$ can*

Table 1: Comparisons between our proposed CoExBO with baselines used in the ablation study.

				Novel contributions		
Baselines	Human selection	BO	π augmentation	Interative π update	Uncertainty in π estimation	Explanation
random	✓	✗	✗	✗	✗	✗
manual search	✓	✗	✗	✗	✗	✗
UCB/TS	✗	✓	✗	✗	✗	✗
prior sampling	✗	✗	✓	✗	✗	✗
batch UCB/TS	✓	✓	✗	✗	✗	✗
πBO (Hvarfner et al., 2022)	✗	✓	✓	✗	✗	✗
CoExBO (πBO)	✓	✓	✓	✓	✗	✗
CoExBO	✓	✓	✓	✓	✓	✓

be estimated from observations as

$$\mathbf{B}_S(\mathbf{x})^\top \tilde{\mathbf{f}} + \beta_t^{1/2} \sqrt{\mathbf{B}_S(\mathbf{x})^\top \tilde{\mathbf{K}}_{\mathbf{XX}} \mathbf{B}_S(\mathbf{x})} \tag{12}$$

where $\mathbf{B}_S(\mathbf{x}) = (\mathbf{K}_{\mathbf{X}_S\mathbf{X}_S} + \lambda_S I)^{-1} k_S(\mathbf{X}_S, \mathbf{x}_S)$, $\lambda_S > 0$ *is a regularisation parameter,* $\tilde{\mathbf{f}}$ *is the posterior mean, and* $\tilde{\mathbf{K}}_{\mathbf{X},\mathbf{X}}$ *is the posterior covariance matrix of the GP.*

To obtain this quantity, we utilized the fact that the conditional expectation of GPs also admits an analytical form; see Chau et al. (2021a,b) for further details. Other explanation features based on Shapley values are detailed in Supplementary F.

4 Experiments

CoExBO has been tested for synthetic and real-world tasks and is implemented using PyTorch (Paszke et al., 2019), GPyTorch (Gardner et al., 2018), BoTorch (Balandat et al., 2020), and SOBER (Adachi et al., 2023a). All experiments were averaged over 10 repeats, computed with a laptop PC[3]. We set the initial random samples for objective queries as $n_{\text{obj}} = 10$ and for preferential learning $n_{\text{pref}} = 100$, respectively.

Table 1 provides a summary of the baseline algorithms we evaluated. These include **Random**: This method generates a pair of i.i.d. samples uniformly, after which a human selects the preferred one. **Manual Search**: In this approach, a human selects the next query without any algorithmic assistance. **UCB** (Srinivas et al., 2010) and **Thompson Sampling (TS)** (Thompson, 1933): Both methods autonomously select the next query without human intervention. **Prior Sampling**: This technique involves choosing the next query point as an i.i.d. sample from the estimated prior $\hat{\pi}$, derived from the initial preference samples n_{pref}, without BO assistance. **BatchUCB** (Azimi et al., 2010) and **BatchTS** (Kandasamy et al., 2018): These algorithms generate pairs of candidates, from which a human selects one. They do not integrate human knowledge $\hat{\pi}$ in the candidate generation process. **πBO** (Hvarfner et al., 2022): This algorithm selects the next query point based on a $\hat{\pi}$-augmented AF, incorporating human knowledge through $\hat{\pi}$. However, it is not interactive (as it is fixed before running the BO) and does not account for uncertainty in $\hat{\pi}$ estimation or human interactive selection. Our proposed algorithm, **CoExBO**, incorporates all these elements. Its variant, **CoExBO (πBO)**, specifically analyzes the efficacy of our new AF by replacing it with the πBO's AF, $\alpha_{f_t}(\cdot)\mathbb{E}_{g_t}[\hat{\pi}_{g_t}(\cdot)]^{\frac{\gamma}{t}}$. Following the methodology in the original πBO paper, we set the decaying hyperparameter to 10 for πBO and γ to 0.01 for our algorithm.

[3]MacBook Pro 2019, 2.4 GHz 8-Core Intel Core i9, 64 GB 2667 MHz DDR4

4.1 Synthetic Functions with Synthetic Human Selection

Synthetic functions. First, CoExBO was tested with synthetic functions and a synthetic human selection as $\mathrm{H}(x_1, x_2)$ such that $f_{\text{human}}(x_1) > f_{\text{human}}(x_2)$, where $f_{\text{human}}(x) := f(x) + \epsilon_{\text{pref}}$, and $\epsilon_{\text{pref}} \sim \mathcal{N}(x; 0, \sigma^2_{\text{pref}})$. The correct human selection rate can be modified by changing σ^2_{pref}. $\sigma^2_{\text{pref}} = 0.1$ throughout this experiment. We have chosen the five commonly used test functions (Surjanovic & Bingham, 2023) (see details in Supplementary G.1).

Figure 5 shows the results on simple regret. CoExBO consistently outperforms four baselines except for the Rosenbrock function. This suggests human feedback is particularly effective for multimodal functions with many local minima. This is evident as there are no big differences between the algorithms with and without human intervention in the Rosenbrock function. Further details and computational complexity analysis can be found in Supplementary G.1.

Robustness evaluation. We tested CoExBO's robustness with the Ackley function regarding (a) uncertain prior and (b) incorrect human selections. We varied the prior confidence by changing the number of initial random samples ($n_{\text{pref}} = 10, 100, 500$). To vary human selection correctness, we adjusted

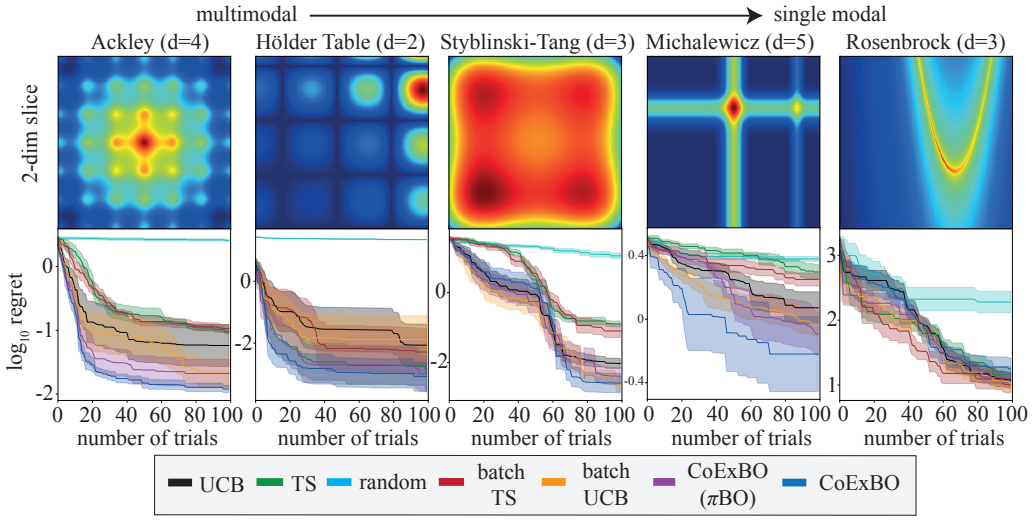

Figure 5: Convergence plot of simple regret for 5 synthetic functions with the synthetic selection accuracy ($\epsilon_{\text{pref}} := \mathcal{N}(0, 0.1^2)$). Lines and shaded area denote mean ± 1 standard error. CoExBO consistently outperforms all six baselines except for the Rosenbrock function. The dark red region is the global maximum.

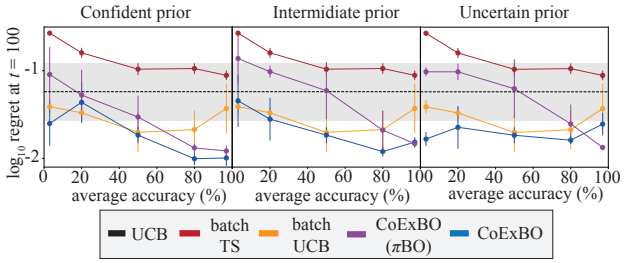

Figure 6: Convergence after 100 iterations on Ackley function ($d = 4$) with three prior confidence levels and five selection accuracy levels. Lines and error bars denote mean ± 1 standard error.

the noise variance of the synthetic human function ($\sigma^2_{\text{pref}} = 0.1, 1, 100$). We simulated adversarial selection cases by flipping the feedback from $\sigma^2_{\text{pref}} = 0.1, 1$.

Figure 6 demonstrates that CoExBO is robust against uncertain prior knowledge and incorrect human selections. While the optimistic πBO AF becomes less effective with reduced selection accuracy, CoExBO maintains its effectiveness better. The key distinction between πBO and CoExBO AFs is that the former only modifies the UCB in uncertain areas, whereas πBO adjusts it regardless of uncertainty (see Figure 4). In other words, once the wrongly believed position is queried, the uncertainty in that area decreases, leading to unbiased UCB. This feature offers greater resilience to incorrect and uncertain human selections compared to πBO AF. Note that both πBO and CoExBO are guaranteed to asymptotically approach the standard UCB, so longer iterations should yield similar results. Both batchUCB and CoExBO exhibited robustness against adversarial selection. However, the difference in adversarial selection falls within the standard error of UCB, indicating that it does not outperform the standard UCB in adversarial cases. Relative differences in tendencies provide more reliable insights. Further details are explained in Supplementary G.1.

4.2 Real-World Tasks with Human Experts

Lithium-ion batteries are the key to realizing the electrification of many sectors as a climate action. However, due to their complicated chemical nature, a perfect simulator that predicts required material properties under all operational and degradation conditions does not exist. Hence, researchers need to repeat costly laboratory experiments to find the best combinations of materials from which to build new lithium-ion batteries.

We assessed expert advice's effectiveness with four battery researchers and compared results with and without explainability features to gauge their impact on selection accuracy. The problem involves finding the best electrolyte material combination to maximize ionic conductivity. This task is challenging due to complex solvation and many-body effects (Gering, 2017), making prediction with a simulator-based approach difficult. We applied CoExBO to two electrolyte design problems: one involving four materials (EC-DMC-EMC-$LiPF_6$) (Dave et al., 2022), a well-known combination, and the other comprising MA-DMC-EMC-$LiPF_6$ (Logan et al., 2018), an unfamiliar composition to all participating experts. While materials science knowledge can deduce the effect on lithium-ion solvation states by changing from carbonate to acetate non-aqueous solvents, their knowledge is qualitative and not quantitative.

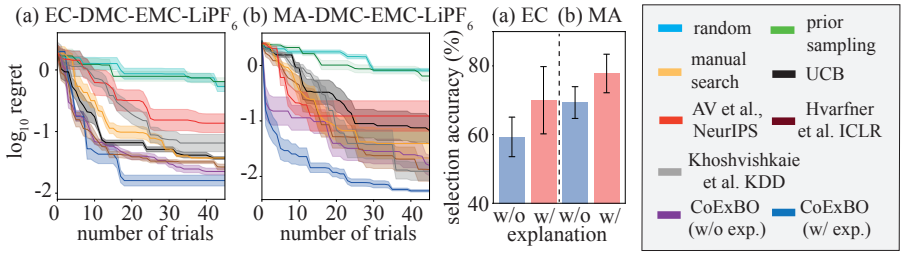

Figure 7: Convergence plot and selection accuracy of two battery material design tasks (a) EC (b) MA-based system. The explainability strengthens the selection accuracy and accelerates convergence.

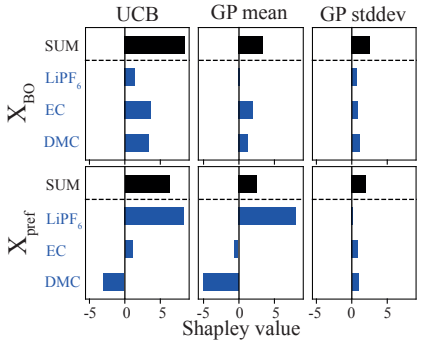

Figure 8: An example of Shapley values for UCB, GP predictive mean and standard deviation during the iteration at two recommended locations.

Figure 7 shows that CoExBO with expert knowledge accelerates convergence, even without explainability features. The addition of explainability results in a significant speedup in time-to-accuracy and enhancement in selection accuracy, which outperformed eight baselines. Figure 8 exemplifies a typical case of Shapley values. The black bar's sum of Shapley values shows that X_{BO} has a higher UCB value, indicating that selecting X_{BO} is natural in standard BO. However, the preference-based X_{pref} attributes more reasonable importance to conductivity, consistent with chemical expertise—highlighting $LiPF_6$ as the key material. We illustrate the CoExBO's effectiveness with two distinct cases: One participant initially relied heavily on BO suggestions, a phenomenon known as automation bias (Cummings, 2004). Goddard et al. (2012); Skitka et al. (2000) have shown that explaining the process and holding users accountable for their decision accuracy can reduce such biases, both of which CoExBO achieved. Another participant consistently trusted their own preferences, even when the GP improved. CoExBO's no-harm guarantee shepherds users to global convergence.

In summary, expert knowledge can be particularly helpful to the GP in two ways: (A) Assessing the reliability of noisy ionic conductivity measurements and disregarding noise to infer the true function shape. (B) Applying their chemical knowledge to roughly optimize the composition, exploiting the knowledge of each material's importance. On the other hand, CoExBO can help experts in three ways: (A) BO is better at fine-tuning more precisely than experts, as expert knowledge only focuses on the main effect. (B) Shapley values provide accurate importance rankings conditioned on x, which updates experts' knowledge. (C) The feedback and explaining feature can correct experts' wrong understanding and guide them in the true optimal direction.

5 Discussions and Limitations

Our approach accelerates convergence when experts possess accurate comparative knowledge that the GP cannot access, which remains a strong assumption, albeit weaker than the conventional ones with a well-defined and error-free belief function or the optimal query. Expert knowledge can be particularly helpful to the GP in three ways: (a) as a good warm starter, allowing it to begin from a promising region; (b) as a good encoder, compensating for the information lost during simplification to a low-dimensional search space; and (c) as a noise reducer, providing a more accurate estimation of experimental noise. These aspects align well with fields such as chemistry and scientific experiments, and experts can convey this complex information through simple selections. Our proofs are specific to the UCB setting, but the decay property can ensure convergence for any AF. In our experiments with experts, we observed that there is a tendency to expect both surrogate and explanation models to provide an 'oracle' understanding of the whole scientific process. We emphasise this is not the purpose of explainable BO as we are by definition, operating under a small data regime. However, our collaboration and explanation framework allows us to demystify the BO process and thus mitigate over-trust.

There is a growing interest in putting humans back into the optimization cycle. A prime example is the RLHF to fine-tune LLMs (Christiano et al., 2017; Rafailov et al., 2024). We also observe concurrent works centered on enhancing human-AI collaboration (AV et al., 2024) and explainable BO (Chakraborty et al., 2023, 2024), showcasing this as a promising direction of research.

Acknowledgements

We thank Philipp Dechent and Katie Lukow for participating in the real-world experiments, and anonymous reviewers who gave useful comments. Masaki Adachi was supported by the Clarendon Fund, the Oxford Kobe Scholarship, the Watanabe Foundation, and Toyota Motor Corporation.

References

Masaki Adachi. High-dimensional discrete Bayesian optimization with self-supervised representation learning for data-efficient materials exploration. In *NeurIPS 2021 AI for Science Workshop*, 2021. doi: https://openreview.net/forum?id=xJhjehqjQeB.

Masaki Adachi, Satoshi Hayakawa, Martin Jørgensen, Harald Oberhauser, and Michael A Osborne. Fast Bayesian inference with batch Bayesian quadrature via kernel recombination. *Advances in Neural Information Processing Systems*, 35, 2022. doi: https://doi.org/10.48550/arXiv.2206.04734.

Masaki Adachi, Satoshi Hayakawa, Saad Hamid, Martin Jørgensen, Harald Oberhauser, and Michael A Osborne. SOBER: Highly parallel Bayesian optimization and Bayesian quadrature over discrete and mixed spaces. In *ICML 2023 Workshop: Sampling and Optimization in Discrete Space*, 2023a. doi: https://doi.org/10.48550/arXiv.2301.11832.

Masaki Adachi, Yannick Kuhn, Birger Horstmann, Arnulf Latz, Michael A Osborne, and David A Howey. Bayesian model selection of lithium-ion battery models via Bayesian quadrature. *IFAC-PapersOnLine*, 56(2):10521–10526, 2023b. doi: https://doi.org/10.1016/j.ifacol.2023.10.1073.

Masaki Adachi, Satoshi Hayakawa, Martin Jørgensen, Xingchen Wan, Vu Nguyen, Harald Oberhauser, and Michael A Osborne. Adaptive batch sizes for active learning: A probabilistic numerics approach. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2024. doi: https://doi.org/10.48550/arXiv.2306.05843.

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

Raul Astudillo, Zhiyuan Jerry Lin, Eytan Bakshy, and Peter Frazier. qEUBO: A decision-theoretic acquisition function for preferential Bayesian optimization. In *International Conference on Artificial Intelligence and Statistics*, pp. 1093–1114. PMLR, 2023.

Arun Kumar AV, Santu Rana, Alistair Shilton, and Svetha Venkatesh. Human-AI collaborative Bayesian Optimisation. *Advances in Neural Information Processing Systems*, 35:16233–16245, 2022.

Arun Kumar AV, Alistair Shilton, Sunil Gupta, Santu Runa, Stewart Greenhill, and Svetha Venkatesh. Enhanced bayesian optimization via preferential modeling of abstract properties. *arXiv preprint arXiv:2402.17343*, 2024.

Javad Azimi, Alan Fern, and Xiaoli Fern. Batch Bayesian optimization via simulation matching. *Advances in Neural Information Processing Systems*, 23, 2010.

Maximilian Balandat, Brian Karrer, Daniel Jiang, Samuel Daulton, Ben Letham, Andrew G Wilson, and Eytan Bakshy. BoTorch: a framework for efficient Monte-Carlo Bayesian optimization. *Advances in neural information processing systems*, 33:21524–21538, 2020.

Alessio Benavoli, Dario Azzimonti, and Dario Piga. Bayesian optimization for choice data. In *Proceedings of the Companion Conference on Genetic and Evolutionary Computation*, pp. 2272–2279, 2023a.

Alessio Benavoli, Dario Azzimonti, and Dario Piga. Learning choice functions with Gaussian processes. In *The 39th Conference on Uncertainty in Artificial Intelligence*, 2023b.

James Bergstra and Yoshua Bengio. Random search for hyper-parameter optimization. *Journal of machine learning research*, 13(2), 2012.

Julian Berk, Sunil Gupta, Santu Rana, and Svetha Venkatesh. Randomised Gaussian process upper confidence bound for Bayesian optimisation. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, pp. 2284–2290, 2020.

Xavier Bouthillier and Gaël Varoquaux. Survey of machine-learning experimental methods at NeurIPS2019 and ICLR2020. *HAL 02447823*, 2020. doi: https://hal.science/hal-02447823.

Ralph Allan Bradley and Milton E. Terry. Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, 39(3/4):324–345, 1952.

Eric Brochu, Nando Freitas, and Abhijeet Ghosh. Active preference learning with discrete choice data. *Advances in neural information processing systems*, 20, 2007.

Jerry F Casteel and Edward S Amis. Specific conductance of concentrated solutions of magnesium salts in water-ethanol system. *Journal of Chemical and Engineering Data*, 17(1):55–59, 1972.

Tanmay Chakraborty, Christian Wirth, and Christin Seifert. Post-hoc rule based explanations for black

box bayesian optimization. In *European Conference on Artificial Intelligence*, pp. 320–337. Springer, 2023.

Tanmay Chakraborty, Christin Seifert, and Christian Wirth. Explainable bayesian optimization. *arXiv preprint arXiv:2401.13334*, 2024.

Siu Lun Chau, Shahine Bouabid, and Dino Sejdinovic. Deconditional downscaling with Gaussian processes. *Advances in Neural Information Processing Systems*, 34:17813–17825, 2021a.

Siu Lun Chau, Jean-Francois Ton, Javier González, Yee Teh, and Dino Sejdinovic. Bayesimp: Uncertainty quantification for causal data fusion. *Advances in Neural Information Processing Systems*, 34:3466–3477, 2021b.

Siu Lun Chau, Mihai Cucuringu, and Dino Sejdinovic. Spectral ranking with covariates. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 70–86. Springer, 2022a.

Siu Lun Chau, Javier González, and Dino Sejdinovic. Learning inconsistent preferences with Gaussian processes. In *International Conference on Artificial Intelligence and Statistics*, pp. 2266–2281. PMLR, 2022b.

Siu Lun Chau, Robert Hu, Javier Gonzalez, and Dino Sejdinovic. RKHS-SHAP: Shapley values for kernel methods. *Advances in Neural Information Processing Systems*, 35:13050–13063, 2022c.

Siu Lun Chau, Krikamol Muandet, and Dino Sejdinovic. Explaining the uncertain: Stochastic Shapley values for Gaussian process models. *arXiv preprint arXiv:2305.15167*, 2023.

Paul F Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, and Dario Amodei. Deep reinforcement learning from human preferences. *Advances in neural information processing systems*, 30, 2017.

Wei Chu and Zoubin Ghahramani. Preference learning with Gaussian processes. In *Proceedings of the 22nd International Conference on Machine Learning*, pp. 137–144, 2005.

Abdoulatif Cisse, Xenophon Evangelopoulos, Sam Carruthers, Vladimir V Gusev, and Andrew I Cooper. HypBO: Expert-guided chemist-in-the-loop Bayesian search for new materials. *arXiv preprint arXiv:2308.11787*, 2023.

Fabio Colella, Pedram Daee, Jussi Jokinen, Antti Oulasvirta, and Samuel Kaski. Human strategic steering improves performance of interactive optimization. In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization*, pp. 293–297, 2020.

Mihai Cucuringu. Sync-rank: Robust ranking, constrained ranking and rank aggregation via eigenvector and SDP synchronization. *IEEE Transactions on Network Science and Engineering*, 3(1):58–79, 2016.

Mary Cummings. Automation bias in intelligent time critical decision support systems. In *AIAA 1st intelligent systems technical conference*, pp. 6313, 2004.

Adarsh Dave, Jared Mitchell, Sven Burke, Hongyi Lin, Jay Whitacre, and Venkatasubramanian Viswanathan. Autonomous optimization of non-aqueous Li-ion battery electrolytes via robotic experimentation and machine learning coupling. *Nature communications*, 13(1):5454, 2022.

Matthias Feurer, Aaron Klein, Katharina Eggensperger, Jost Springenberg, Manuel Blum, and Frank Hutter. Efficient and robust automated machine learning. *Advances in neural information processing systems*, 28, 2015.

Jacob Gardner, Geoff Pleiss, Kilian Q Weinberger, David Bindel, and Andrew G Wilson. GPyTorch: Blackbox matrix-matrix Gaussian process inference with GPU acceleration. In *Advances in Neural Information Processing Systems*, pp. 7576–7586, 2018.

Roman Garnett. *Bayesian optimization*. Cambridge University Press, 2023.

Paul H Garthwaite, Joseph B Kadane, and Anthony O'Hagan. Statistical methods for eliciting probability distributions. *Journal of the American statistical Association*, 100(470):680–701, 2005.

Michael A Gelbart, Jasper Snoek, and Ryan P Adams. Bayesian optimization with unknown constraints. In *Proceedings of the Thirtieth Conference on Uncertainty in Artificial Intelligence*, pp. 250—-259, 2014. doi: https://doi.org/10.48550/arXiv.1403.5607.

Kevin L Gering. Prediction of electrolyte conductivity: results from a generalized molecular model based on ion solvation and a chemical physics framework. *Electrochimica Acta*, 225:175–189, 2017.

Joseph Giovanelli, Alexander Tornede, Tanja Tornede, and Marius Lindauer. Interactive hyperparameter optimization in multi-objective problems via preference learning. *arXiv preprint arXiv:2309.03581*, 2023.

Kate Goddard, Abdul Roudsari, and Jeremy C Wyatt. Automation bias: a systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, 19(1): 121–127, 2012.

Javier González, Michael Osborne, and Neil Lawrence. Glasses: Relieving the myopia of Bayesian optimisation. In *Artificial Intelligence and Statistics*, pp. 790–799. PMLR, 2016.

Javier González, Zhenwen Dai, Andreas Damianou, and Neil D Lawrence. Preferential Bayesian optimization. In *Proceedings of the 34th International Conference on Machine Learning*, pp. 1282–1291, 2017.

Sunil Gupta, Alistair Shilton, Arun Kumar AV, Shannon Ryan, Majid Abdolshah, Hung Le, Santu Rana, Julian Berk, Mahad Rashid, and Svetha Venkatesh. BO-Muse: A human expert and AI teaming framework for accelerated experimental design. *arXiv preprint arXiv:2303.01684*, 2023.

Satoshi Hayakawa, Harald Oberhauser, and Terry Lyons. Positively weighted kernel quadrature via subsampling. *Advances in Neural Information Processing Systems*, 35:6886–6900, 2022.

José Miguel Hernández-Lobato, Michael Gelbart, Matthew Hoffman, Ryan Adams, and Zoubin Ghahramani. Predictive entropy search for Bayesian optimization with unknown constraints. In *International conference on machine learning*, pp. 1699–1707. PMLR, 2015.

Robert Hu, Siu Lun Chau, Jaime Ferrando Huertas, and Dino Sejdinovic. Explaining preferences with shapley values. *Advances in Neural Information Processing Systems*, 35:27664–27677, 2022.

Daolang Huang, Louis Filstroff, Petrus Mikkola, Runkai Zheng, and Samuel Kaski. Bayesian optimization augmented with actively elicited expert knowledge. *arXiv preprint arXiv:2208.08742*, 2022.

Frank Hutter, Holger H Hoos, and Kevin Leyton-Brown. Sequential model-based optimization for general algorithm configuration. In *Learning and Intelligent Optimization: 5th International Conference, LION 5, Rome, Italy, January 17-21, 2011. Selected Papers 5*, pp. 507–523. Springer, 2011.

Carl Hvarfner, Danny Stoll, Artur Souza, Marius Lindauer, Frank Hutter, and Luigi Nardi. πBO: Augmenting acquisition functions with user beliefs for bayesian optimization. In *International Conference on Learning Representations*, 2022.

Michael I Jordan. Artificial intelligence—the revolution hasn't happened yet. *Harvard Data Science Review*, 1(1):1–9, 2019.

Daniel Kahneman and Amos Tversky. On the interpretation of intuitive probability: A reply to Jonathan Cohen. *Cognition*, 7(4):409–411, 1979.

Motonobu Kanagawa, Bharath K Sriperumbudur, and Kenji Fukumizu. Convergence guarantees for kernel-based quadrature rules in misspecified settings. *Advances in Neural Information Processing Systems*, 29, 2016.

Keren J Kanarik, Wojciech T Osowiecki, Yu Lu, Dipongkar Talukder, Niklas Roschewsky, Sae Na Park, Mattan Kamon, David M Fried, and Richard A Gottscho. Human–machine collaboration for improving semiconductor process development. *Nature*, 616 (7958):707–711, 2023.

Kirthevasan Kandasamy, Akshay Krishnamurthy, Jeff Schneider, and Barnabás Póczos. Parallelised Bayesian optimisation via Thompson sampling. In *International Conference on Artificial Intelligence and Statistics*, pp. 133–142. PMLR, 2018.

Ali Khoshvishkaie, Petrus Mikkola, Pierre-Alexandre Murena, and Samuel Kaski. Cooperative Bayesian optimization for imperfect agents. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 475–490. Springer, 2023.

Yuki Koyama, Issei Sato, and Masataka Goto. Sequential gallery for interactive visual design optimization. *ACM Transactions on Graphics (TOG)*, 39(4):88–1, 2020.

Michael Y Li and Ryan P Adams. Explainability constraints for Bayesian optimization. In *6th ICML Workshop on Automated Machine Learning*, 2020.

Zhiyuan Jerry Lin, Raul Astudillo, Peter Frazier, and Eytan Bakshy. Preference exploration for efficient Bayesian optimization with multiple outcomes. In *International Conference on Artificial Intelligence and Statistics*, pp. 4235–4258. PMLR, 2022.

Dong C Liu and Jorge Nocedal. On the limited memory BFGS method for large scale optimization. *Mathematical programming*, 45(1-3):503–528, 1989.

Sulin Liu, Qing Feng, David Eriksson, Benjamin Letham, and Eytan Bakshy. Sparse Bayesian optimization. In *International Conference on Artificial Intelligence and Statistics*, pp. 3754–3774. PMLR, 2023.

ER Logan, Erin M Tonita, KL Gering, Jing Li, Xiaowei Ma, LY Beaulieu, and JR Dahn. A study of the physical properties of Li-ion battery electrolytes containing esters. *Journal of The Electrochemical Society*, 165(2):A21, 2018.

R Duncan Luce. On the possible psychophysical laws. *Psychological review*, 66(2):81, 1959.

Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. In *Advances in neural information processing systems*, pp. 4765–4774, 2017.

Petrus Mikkola, Milica Todorović, Jari Järvi, Patrick Rinke, and Samuel Kaski. Projective preferential Bayesian optimization. In *International Conference on Machine Learning*, pp. 6884–6892. PMLR, 2020.

Petrus Mikkola, Osvaldo A. Martin, Suyog Chandramouli, Marcelo Hartmann, Oriol Abril Pla, Owen Thomas, Henri Pesonen, Jukka Corander, Aki Vehtari, Samuel Kaski, Paul-Christian Bürkner, and

Arto Klami. Prior Knowledge Elicitation: The Past, Present, and Future. *Bayesian Analysis*, pp. 1 – 33, 2023a. doi: 10.1214/23-BA1381.

Petrus Mikkola, Julien Martinelli, Louis Filstroff, and Samuel Kaski. Multi-fidelity Bayesian optimization with unreliable information sources. In *International Conference on Artificial Intelligence and Statistics*, pp. 7425–7454. PMLR, 2023b.

Dimitrios Milios, Raffaello Camoriano, Pietro Michiardi, Lorenzo Rosasco, and Maurizio Filippone. Dirichlet-based Gaussian processes for large-scale calibrated classification. *Advances in Neural Information Processing Systems*, 31, 2018.

Jonas Mockus. The application of Bayesian methods for seeking the extremum. *Towards global optimization*, 2:117, 1998.

Anthony O'Hagan. Bayes–Hermite quadrature. *Journal of statistical planning and inference*, 29(3):245–260, 1991.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, and Luca Antiga. PyTorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. *Advances in Neural Information Processing Systems*, 36, 2024.

Anil Ramachandran, Sunil Gupta, Santu Rana, Cheng Li, and Svetha Venkatesh. Incorporating expert prior in Bayesian optimisation via space warping. *Knowledge-Based Systems*, 195:105663, 2020.

Denise M Rousseau. Schema, promise and mutuality: The building blocks of the psychological contract. *Journal of occupational and organizational psychology*, 74(4):511–541, 2001.

Lloyd S Shapley. A value for n-person games. *Contributions to the Theory of Games*, 2(28):307–317, 1953.

Linda J Skitka, Kathleen Mosier, and Mark D Burdick. Accountability and automation bias. *International Journal of Human-Computer Studies*, 52(4):701–717, 2000.

Jasper Snoek, Kevin Swersky, Rich Zemel, and Ryan Adams. Input warping for Bayesian optimization of non-stationary functions. In *International Conference on Machine Learning*, pp. 1674–1682. PMLR, 2014.

Il'ya Meerovich Sobol'. On the distribution of points in a cube and the approximate evaluation of integrals. *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, 7(4):784–802, 1967.

Jialin Song, Yuxin Chen, and Yisong Yue. A general framework for multi-fidelity Bayesian optimization with Gaussian processes. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pp. 3158–3167. PMLR, 2019.

Artur Souza, Luigi Nardi, Leonardo B Oliveira, Kunle Olukotun, Marius Lindauer, and Frank Hutter. Bayesian optimization with a prior for the optimum. In *Machine Learning and Knowledge Discovery in Databases. Research Track: European Conference, ECML PKDD 2021, Bilbao, Spain, September 13–17, 2021, Proceedings, Part III 21*, pp. 265–296. Springer, 2021.

Niranjan Srinivas, Andreas Krause, Sham M Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. In *International Conference on International Conference on Machine Learning*, pp. 1015–1022, 2010.

S. Surjanovic and D. Bingham. Virtual library of simulation experiments: Test functions and datasets. Retrieved October 6, 2023, from `http://www.sfu.ca/~ssurjano`, 2023.

Shion Takeno, Yu Inatsu, and Masayuki Karasuyama. Randomized Gaussian process upper confidence bound with tight Bayesian regret bounds. In *International Conference on Machine Learning*, volume 202, pp. 33490–33515, 2023a.

Shion Takeno, Masahiro Nomura, and Masayuki Karasuyama. Towards practical preferential Bayesian optimization with skew Gaussian processes. In *International Conference on Machine Learning*, volume 202, pp. 33516–33533, 2023b.

Kendall Taylor, Huong Ha, Minyi Li, Jeffrey Chan, and Xiaodong Li. Bayesian preference learning for interactive multi-objective optimisation. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 466–475, 2021.

William R Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3-4):285–294, 1933.

Zi Wang, George E Dahl, Kevin Swersky, Chansoo Lee, Zelda Mariet, Zachary Nado, Justin Gilmer, Jasper Snoek, and Zoubin Ghahramani. Pre-trained Gaussian processes for Bayesian optimization. *arXiv preprint arXiv:2109.08215*, 2021.

Christopher KI Williams and Carl Edward Rasmussen. *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006.

Jian Wu, Saul Toscano-Palmerin, Peter I Frazier, and Andrew Gordon Wilson. Practical multi-fidelity Bayesian optimization for hyperparameter tuning. In *Uncertainty in Artificial Intelligence*, pp. 788–798. PMLR, 2020.

1. For all models and algorithms presented, check if you include:
   (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model. [Yes]
   (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm. [Yes in supplementary]
   (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries. [Yes at https://anonymous.4open.science/r/CoExBO-4B06/]
2. For any theoretical claim, check if you include:
   (a) Statements of the full set of assumptions of all theoretical results. [Yes]
   (b) Complete proofs of all theoretical results. [Yes]
   (c) Clear explanations of any assumptions. [Yes]
3. For all figures and tables that present empirical results, check if you include:
   (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL). [Yes/No/Not Applicable]
   (b) All the training details (e.g., data splits, hyperparameters, how they were chosen). [Yes in text and supplementary]
   (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times). [Yes]
   (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider). [Yes]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:
   (a) Citations of the creator If your work uses existing assets. [Yes]
   (b) The license information of the assets, if applicable. [Not Applicable]
   (c) New assets either in the supplemental material or as a URL, if applicable. [Not Applicable]
   (d) Information about consent from data providers/curators. [Yes in supplementary]
   (e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content. [Not Applicable]
5. If you used crowdsourcing or conducted research with human subjects, check if you include:
   (a) The full text of instructions given to participants and screenshots. [Yes]
   (b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable. [Not Applicable]
   (c) The estimated hourly wage paid to participants and the total amount spent on participant compensation. [Not Applicable]

Part I
Appendix

Table of Contents

A Proof of theorem

A.1 Regret analysis of normal UCB policy

We begin with the finite case, $|D| < \infty$. We recall two lemmas from Srinivas et al. (2010).

Lemma 5. *Pick $\delta \in (0,1)$ and set $\beta_t = 2\log(|D|p_t/\delta)$, where $\sum_{t\geq 1} p_t^{-1} = 1$, $p_t > 0$. Then,*

$$|f(\boldsymbol{x}) - \mu_{f_{t-1}}(\boldsymbol{x}))| \leq \beta_t^{1/2}\sigma_{f_{t-1}}(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in D,\, \forall t \geq 1 \tag{13}$$

holds with probability $\geq 1-\delta$.

Proof Fix $t \geq 1$ and $\boldsymbol{x} \in D$. Conditioned on $\boldsymbol{y}_{t-1} = (y_1, \ldots, y_{t-1})$, $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{t-1}\}$ are deterministic, and $f(\boldsymbol{x}) \sim \mathcal{N}(\mu_{f_{t-1}}(\boldsymbol{x}), \sigma^2_{f_{t-1}}(\boldsymbol{x}))$. Now, if $r \sim \mathcal{N}(0,1)$, then

$$\mathbb{P}(r > c) = \exp\left(-\frac{c^2}{2}\right)(2p)^{-1/2}\int \exp\left(-\frac{(r-c)^2}{2} - c(r-c)\right)\mathrm{d}r, \tag{14}$$

$$\leq \exp\left(-\frac{c^2}{2}\right)\mathbb{P}(r > 0), \tag{15}$$

$$= \frac{1}{2}\exp\left(-\frac{c^2}{2}\right) \tag{16}$$

for $c > 0$, since $\exp(-c(r-c)) \leq 1$ for $r > c$. Therefore, $\mathbb{P}\left(|f(\boldsymbol{x}) - \mu_{f_{t-1}}(\boldsymbol{x}))| \leq \beta_t^{1/2}\sigma_{f_{t-1}}(\boldsymbol{x})\right) \leq \exp\left(-\frac{\beta_t}{2}\right)$, using $r = \frac{f(\boldsymbol{x})-\mu_{f_{t-1}}(\boldsymbol{x}))}{\sigma_{f_{t-1}}(\boldsymbol{x})}$ and $c = \beta_t^{1/2}$. Applying the union bound,

$$|f(\boldsymbol{x}) - \mu_{f_{t-1}}(\boldsymbol{x}))| \leq \beta_t^{1/2}\sigma_{f_{t-1}}(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in D \tag{17}$$

holds with probability $\geq 1 - |D|\exp\left(-\frac{\beta_t}{2}\right)$. Choosing $|D|\exp\left(-\frac{\beta_t}{2}\right) = \frac{\delta}{p_t}$ and using the union bound for $t \in \mathbb{N}$, the statement holds.

Lemma 6. *Fix $t \geq 1$. If $|f(\boldsymbol{x}) - \mu_{f_{t-1}}(\boldsymbol{x}))| \leq \beta_t^{1/2}\sigma_{f_{t-1}}(\boldsymbol{x})$ for all $\boldsymbol{x} \in D$, then the regret r_t is bounded by $2\beta_t^{1/2}\sigma_{f_{t-1}}(\boldsymbol{x})$.*

Proof By definition of $\boldsymbol{x}_{\text{bo}} := \text{argmax}_{x\in\mathcal{X}}\, \mu_{f_{t-1}}(\boldsymbol{x}) + \beta_t^{1/2}\sigma_{f_{t-1}}(\boldsymbol{x})$, we have $\mu_{f_{t-1}}(\boldsymbol{x}_{\text{bo}}) + \beta_t^{1/2}\sigma_{f_{t-1}}(\boldsymbol{x}_{\text{bo}}) \geq \mu_{f_{t-1}}(\boldsymbol{x}^*_{\text{true}}) + \beta_t^{1/2}\sigma_{f_{t-1}}(\boldsymbol{x}^*_{\text{true}}) \geq f_{\text{true}}(\boldsymbol{x}^*_{\text{true}})$. Therefore

$$r_t = f(\boldsymbol{x}^*_{\text{true}}) - f(\boldsymbol{x}_{\text{bo}}) \leq \beta_t^{1/2}\sigma_{f_{t-1}}(\boldsymbol{x}_{\text{bo}}) + \mu_{f_{t-1}}(\boldsymbol{x}_{\text{bo}}) - f(\boldsymbol{x}_{\text{bo}}) \tag{18}$$

$$\leq 2\beta_t^{1/2}\sigma_{f_{t-1}}(\boldsymbol{x}_{\text{bo}}) \tag{19}$$

A.2 Proof of Regrets

Recall the definition of the good user belief is $|f(\boldsymbol{x}) - \mu_{f_{t-1},\hat{\pi}_{t-1}}(\boldsymbol{x}))| \leq \beta_t^{1/2}\sigma_{f_{t-1},\hat{\pi}_{t-1}}(\boldsymbol{x})$.

Proof of good user belief regrets

$$\frac{r_{\hat{\pi}_t}}{r_t} = \frac{f(x^*_{\text{true}}) - f(x_2)}{f(x^*_{\text{true}}) - f(x_1)} \tag{20}$$

$$\leq \frac{2\beta_t^{1/2}\sigma_{f_{t-1},\hat{\pi}_{t-1}}(x_2)}{2\beta_t^{1/2}\sigma_{f_{t-1}}(x_1)} \qquad \text{(Lemma 5)} \tag{21}$$

$$= \frac{\sigma_{f_{t-1},\hat{\pi}_{t-1}}(x_2)}{\sigma_{f_{t-1}}(x_1)} \tag{22}$$

$$= \frac{\sigma_{\hat{\pi}_{t-1}}(x_2)}{\sqrt{\sigma^2_{\hat{\pi}_{t-1}}(x_2) + \sigma^2_{f_{t-1}}(x_1)}} \qquad \text{(Proposition 1)} \tag{23}$$

$$= \sqrt{\frac{\rho^2(\mathbb{V}_{g_{t-1}}[\hat{\pi}_{g_{t-1}}(x_2)]) + \gamma(t-1)^2\sigma^2_{f_{t-1}}(x_2)}{\rho^2(\mathbb{V}_{g_{t-1}}[\hat{\pi}_{g_{t-1}}(x_2)]) + \gamma(t-1)^2\sigma^2_{f_{t-1}}(x_2) + \sigma^2_{f_{t-1}}(x_1)}} \qquad \text{(Proposition 1)} \tag{24}$$

$$= R_{\hat{\pi}_t} \tag{25}$$

$$< 1 \qquad (\sigma^2_{f_{t-1}}(x_1) > 0) \tag{26}$$

Proof of bad user belief regrets The bad user belief case trivially follows the same steps with the additional term:

$$\frac{r_{\hat{\pi}_t}}{r_t} = \frac{f(x^*_{\text{true}}) - f(x_2)}{f(x^*_{\text{true}}) - f(x_1)} \tag{27}$$

$$\leq \frac{2\beta_t^{1/2}\sigma_{f_{t-1},\hat{\pi}_{t-1}}(x_2)}{2\beta_t^{1/2}\sigma_{f_{t-1}}(x_1)} + \frac{|\mu_{f_{t-1}}(x_1) - \mu_{f_{t-1},\hat{\pi}_{t-1}}(x_2)|}{2\beta_t^{1/2}\sigma_{f_{t-1}}(x_1)} \qquad \text{(Bad user belief definition)} \tag{28}$$

$$= \Delta\mu_t + R_{\hat{\pi}_t} \tag{29}$$

These proofs are for finite decision sets. We can extend this proof to a general decision set by following Srinivas et al. (2010) steps. We omit this procedure, but it essentially boils down to the same procedure and similar results with slight coefficient differences.

Proof of No Harm Guarantee We first review the general large t limit properties of π-augmented GP.

Lemma 7. *At the $t \to \infty$ limit, the posterior GP is asymptotically equal to the original GP.*

$$\lim_{t\to\infty} \sigma^2_{f_t,\hat{\pi}_t}(x) = \sigma^2_{f_t}(x), \tag{30}$$

$$\lim_{t\to\infty} \mu_{f_t,\hat{\pi}_t}(x) = \mu_{f_t}(x), \tag{31}$$

$$\lim_{t\to\infty} \alpha_{f_t,\hat{\pi}_t}(x) = \alpha_{f_t}(x), \tag{32}$$

$$\lim_{t\to\infty} x_2 = x_1, \tag{33}$$

Proof of Lemma 6

$$\lim_{t\to\infty} \sigma^2_{f_t,\hat{\pi}_t}(x) = \lim_{t\to\infty} \frac{\sigma^2_{\hat{\pi}_t}(x)\sigma^2_{f_t}(x)}{\sigma^2_{\hat{\pi}_t}(x) + \sigma^2_{f_t}(x)} \tag{34}$$

$$= \lim_{t\to\infty} \frac{(\rho^2(\mathbb{V}_{g_t}[\hat{\pi}_{g_t}(x)]) + \gamma t^2\sigma^2_{f_t}(x))\sigma^2_{f_t}(x)}{\rho^2(\mathbb{V}_{g_t}[\hat{\pi}_{g_t}(x)]) + \gamma t^2\sigma^2_{f_t}(x) + \sigma^2_{f_t}(x)} \tag{35}$$

$$= \lim_{t\to\infty} \frac{(\rho^2(\mathbb{V}_{g_t}[\hat{\pi}_{g_t}(x)])/t^2 + \gamma\sigma^2_{f_t}(x))\sigma^2_{f_t}(x)}{\rho^2(\mathbb{V}_{g_t}[\hat{\pi}_{g_t}(x)])/t^2 + \gamma\sigma^2_{f_t}(x) + \sigma^2_{f_t}(x)/t^2} \tag{36}$$

$$= \frac{\gamma\sigma^2_{f_t}(x)\sigma^2_{f_t}(x)}{\gamma\sigma^2_{f_t}(x)} \tag{37}$$

$$= \sigma^2_{f_t}(x) \tag{38}$$

$$\lim_{t\to\infty} \mu_{f_t,\hat{\pi}_t}(x) = \lim_{t\to\infty} \frac{\sigma^2_{f_t,\hat{\pi}_t}(x)}{\sigma^2_{\hat{\pi}_t}(x)}\mu_{\hat{\pi}_t}(x) + \lim_{t\to\infty} \frac{\sigma^2_{f_t,\hat{\pi}_t}(x)}{\sigma^2_{f_t}(x)}\mu_{f_t}(x) \tag{39}$$

$$= \lim_{t\to\infty} \frac{\sigma^2_{f_t}(x)}{\sigma^2_{\hat{\pi}_t}(x)}\mu_{\hat{\pi}_t}(x) + \frac{\sigma^2_{f_t}(x)}{\sigma^2_{f_t}(x)}\mu_{f_t}(x) \tag{40}$$

$$= \lim_{t\to\infty} \frac{\sigma^2_{f_t}(x)}{\rho^2(\mathbb{V}_{g_t}[\hat{\pi}_{g_t}(x)]) + \gamma t^2\sigma^2_{f_t}(x)}\mu_{\hat{\pi}_t}(x) + \mu_{f_t}(x) \tag{41}$$

$$= \mu_{f_t}(x) \tag{42}$$

$$\lim_{t\to\infty} \alpha_{f_t,\hat{\pi}_t}(x) = \lim_{t\to\infty} \mu_{f_t,\hat{\pi}_t}(x) + \beta_t^{1/2} \lim_{t\to\infty} \sigma^2_{f_t,\hat{\pi}_t}(x) \tag{43}$$

$$= \mu_{f_t}(x) + \beta_t^{1/2}\sigma^2_{f_t}(x) \tag{44}$$

$$= \alpha_{f_t}(x) \tag{45}$$

$$\lim_{t\to\infty} x_2 = \operatorname*{argmax}_{x\in\mathcal{X}} \lim_{t\to\infty} \alpha_{f_t,\hat{\pi}_t}(x) \tag{46}$$

$$= \operatorname*{argmax}_{x\in\mathcal{X}} \alpha_{f_t}(x) \tag{47}$$

$$= x_1 \tag{48}$$

Proof of Lemma 2 By definition and lemma 6,

$$\lim_{t\to\infty} r_t^{\pi} = \lim_{t\to\infty} |f(x^*_{\text{true}}) - f(x_2)| \tag{49}$$

$$= |f(x^*_{\text{true}}) - f(x_1)| \tag{50}$$

$$= r_t \tag{51}$$

$$\lim_{t\to\infty} \frac{r_{\hat{\pi}_t}}{r_t} = 1. \tag{52}$$

B Related work

Eliciting human knowledge for Bayesian optimzation There are four ways of incorporating human knowledge, (a) prior over input space (Souza et al., 2021; Ramachandran et al., 2020; Hvarfner et al., 2022; Cisse et al., 2023), (b) Hyperprior over function space (Hutter et al., 2011; Snoek et al., 2014; Wang et al., 2021), (c) multi-fidelity information source (Song et al., 2019; Huang et al., 2022), (d) unknown constraints (Gelbart et al., 2014; Hernández-Lobato et al., 2015; Adachi et al., 2024). They assume fixed and given prior knowledge; we assume implicit and dynamic one. Only πBO and ours offer theoretical guarantees for misspecified π. π is a soft constraint, distinct from the hard one in category (d). They overly limit the search space if incorrect.

Human preference as the objective Preferential BO (PBO) is to find the sample that maximize the probability to be Condorcet winner via interaction with users (González et al., 2017; Astudillo et al., 2023; Takeno

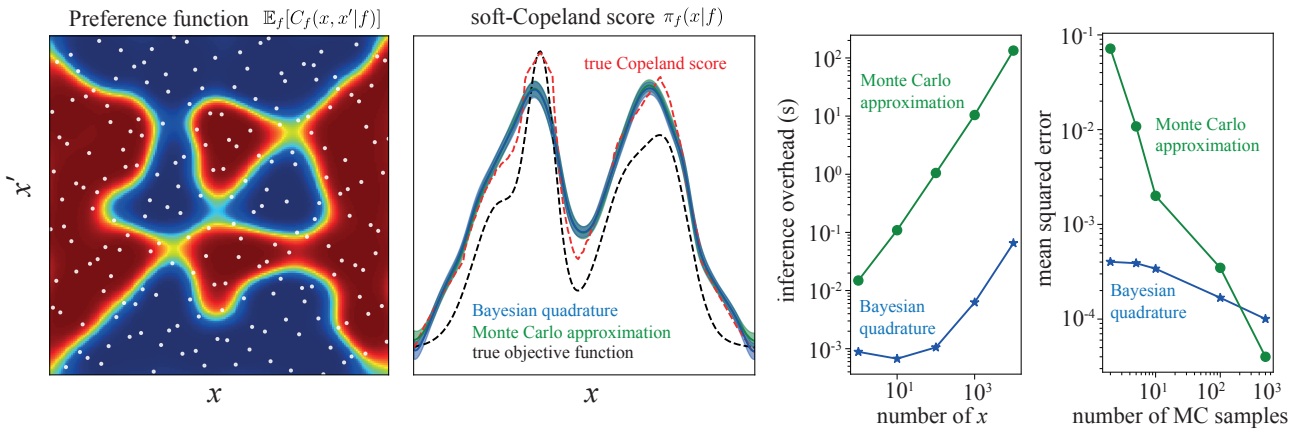

Figure 9: Comparison of Monte Carlo (MC) approximation and Bayesian quadrature approximation on marginalisation for soft-Copeland score. Overhead refers to the wall-clock time to compute the soft-Copeland score at x with the 100 MC samples. The mean squared error is the error between the estimated soft-Copeland score and the ground truth computed with massive MC samples.

et al., 2023b), then extended to preference exploration for multi-objective BO (Lin et al., 2022; Astudillo & Frazier, 2020; Taylor et al., 2021; Giovanelli et al., 2023), non-duel-based PBO (Brochu et al., 2007; Koyama et al., 2020; Mikkola et al., 2020; Benavoli et al., 2023a). While they treat preferences as the objective, ours regards it as an additional information source of the original optimization task.

Human-AI teaming BO An interactive human-BO collaboration setting has recently been proposed to accommodate time-varying human knowledge. There are two categories: (a) Human rectifies BO suggestions; Colella et al. (2020); AV et al. (2022), (b) BO supports human manual search via exploratory adjustment Gupta et al. (2023); Khoshvishkaie et al. (2023). While (a) has no guarantee on worst-case, (b) has a worse convergence rate than the standard BO (Khoshvishkaie et al., 2023). We proposed the first approach with a no-harm guarantee in an interactive setting without losing human initiative.

Explainability in BO There are two prior works on explainable BO. Li & Adams (2020) integrate explainability as hard constraints to the optimization that restrict the search space to where humans can interpret. Liu et al. (2023) introduced sparsity to find simple and interpretable configurations via L0 regularization. Both methods do not consider human interaction, but these are orthogonal to ours and may be beneficial to combine.

C Modelling Human Preference via Gaussian Process

Preferential Gaussian process modelling While many algorithms, rooted in either probabilistic methods (Bradley & Terry, 1952; Luce, 1959) or spectral approaches (Cucuringu, 2016; Chau et al., 2022a), can model human preferences, we opted for Gaussian Processes (GP). This choice enables us to consider epistemic uncertainty more effectively during modeling. However, the classical preferential GP (PGP)(Chu & Ghahramani, 2005) has several limitations; it is computationally expensive and unable to learn preferences that might be inconsistent and with heteroscedastic noise. We combined two existing simple approaches instead of PGP; Dirichlet-based GP (DGP) (Milios et al., 2018), and skew-symmetric data-augmentation (Chau et al., 2022b). DGP translates classification problem as regression one via transforming classification labels to the coefficients of a degenerate Dirichlet distribution. This offers a fast and heteroscedastic GP classifier that has essentially the same accuracy and uncertainty quantification as the original GP classifier. Skew-symmetric data augmentation is to add the symmetric data of original data $Y_{j,i} = 1 - Y_{i,j}$ for the duel (x_i, x_j). Additionally, we used Bayesian quadrature for fast approximation of integral in Eq. (2). However, our setting is not limited to this GP and Bayesian quadrature approximation.

Bayesian quadrature modelling As the integration Eq. (2) is intractable, we have to approximate it. The original preferential BO (González et al., 2017) adopts Monte Carlo (MC) approximation, but it is slow as the authors admit there should be a better way. We adopted Bayesian quadrature (BQ) O'Hagan (1991), particularly

BASQ (Adachi et al., 2022, 2023b) for fast approximation. Figure 9 compares the overhead and accuracy of MC integration and BQ. BQ yields faster computation than MC samples at inference (BQ: $\mathcal{O}((n_{\mathrm{MC}} + n_{\mathrm{duel}})^2)$ vs. MC: $\mathcal{O}(n_{\mathrm{MC}} n_{\mathrm{func}} n_{\mathrm{duel}}^2)$), which repeats the computation in the acquisition function optimization loop. Of course, training the BQ model takes additional cost ($\mathcal{O}(n_{\mathrm{mBCG}} n_{\mathrm{duel}}^2)$ for mBCG algorithm (Gardner et al., 2018)), but this can be understood as "pretraining" to avoid the quartic cost of MC integration at each inference point. As RBF kernel assumption of BQ is misspecified against true Bernoulli distribution, the convergence over sample size is limiting; still, BQ works well as it is robust against misspecification (Kanagawa et al., 2016). In larger MC sample sizes, MC integration outperforms BQ, but it produces non-negligible overhead. It should be noted that BQ can derive the closed form of the denominator of Eq. (2), whereas MC approximation requires another higher-level MC approximation of quintic cost, which is prohibitively slow. So, we adopt BQ in this paper thanks to its good balance of computational accuracy and cost; however, our setting is not limited to BQ.

D Bayesian Quadrature for Fast Soft-Copeland Score Approximation

We have GP classifier $g \sim \mathcal{GP}(\mu_g, \kappa_g)$, and Monte Carlo integration via transforming sampled function with the link function can estimate the expectation of binary probability as Bernoulli distribution:

$$\mathbb{E}_g[g(x, x'|g, \mathbf{D}_{\mathrm{pref}})] = \int \frac{\exp(g_i)}{\sum \exp(g_i)} \mathbb{P}(g_i|x, x', \mathbf{D}_{\mathrm{pref}}) \mathrm{d}g \tag{53}$$

As this is intractable, the forthcoming Condorset winner $\hat{\pi}_g(x)$ marginalisation is also intractable.

Therefore, we apply Bayesian quadrature (BQ). Bayesian quadrature is the model-based approximation technique for intractable integration; typically, GP is applied for the surrogate model for the integrand. We apply the simple GP with RBF kernel to the integrand. Namely, we place the surrogate model GP with the pair of dataset:

$$\mathbf{x}_{\mathrm{bq}} := \mathbf{x}_{\mathrm{pref}} = (\mathbf{x}_1, \mathbf{x}_2), \tag{54}$$
$$\mathbf{y}_{\mathrm{bq}} := \mathbb{E}_g[g(\mathbf{x}_{\mathrm{pref}} \mid \mathbf{D}_{\mathrm{pref}})], \tag{55}$$
$$\mathbf{D}_{\mathrm{bq}} := \left(\mathbf{x}_{\mathrm{bq}}, \mathbf{y}_{\mathrm{bq}}\right), \tag{56}$$

Then, all integration in equation 2 becomes analytical:

$$f_{\mathrm{bq}} \sim \mathcal{GP}(\mu_{\mathrm{bq}}, \kappa_{\mathrm{bq}} \mid \mathbf{D}_{\mathrm{bq}}), \tag{57}$$
$$\mu_{\mathrm{bq}}(X) = k(X, \mathbf{x}_{\mathrm{bq}}) \left[k(\mathbf{x}_{\mathrm{bq}}, \mathbf{x}_{\mathrm{bq}}) + \lambda \mathbf{I}\right]^{-1} \mathbf{y}_{\mathrm{bq}}, \tag{58}$$
$$= k(X, \mathbf{x}_{\mathrm{bq}})^\top \boldsymbol{\omega}, \tag{59}$$
$$= v' \mathcal{N}(X; \mathbf{x}_{\mathrm{bq}}, \mathbf{W})^\top \boldsymbol{\omega}, \tag{60}$$
$$\kappa_{\mathrm{bq}}(X, X') = k(X, X') - k(X, \mathbf{x}_{\mathrm{bq}}) \left[k(\mathbf{x}_{\mathrm{bq}}, \mathbf{x}_{\mathrm{bq}}) + \lambda \mathbf{I}\right]^{-1} k(\mathbf{x}_{\mathrm{bq}}, X), \tag{61}$$
$$= v' \mathcal{N}(X; x', \mathbf{W}) - v'^2 \mathcal{N}(X; \mathbf{x}_{\mathrm{bq}}, \mathbf{W}) \boldsymbol{\Omega} \mathcal{N}(X; \mathbf{x}_{\mathrm{bq}}, \mathbf{W})^T, \tag{62}$$
$$= v' \mathcal{N}(X; x', \mathbf{W}) - v'^2 \sum_{i,j}^{N} \Omega_{i,j} \mathcal{N}(X; x_{bq,i}, \mathbf{W}) \mathcal{N}(X; x_{bq,j}, \mathbf{W}), \tag{63}$$
$$= v' \mathcal{N}(X; X', \mathbf{W}) - v'^2 \mathcal{N}(x_{bq,i}; x_{bq,j}, 2\mathbf{W}) \sum_{i,j}^{N} \Omega_{i,j} \mathcal{N}\left(X; \frac{x_{bq,i} + x_{bq,j}}{2}, \frac{\mathbf{W}}{2}\right), \tag{64}$$

where

$$X := (x, x'), \tag{65}$$
$$k(X, X') := v' \mathcal{N}(X; X', \mathbf{W}), \tag{66}$$
$$v' = v\sqrt{|2\pi \mathbf{W}|}, \tag{67}$$
$$\mathbf{W} := \ell^2 \mathbf{I}, \tag{68}$$
$$\boldsymbol{\omega} := \left[k(\mathbf{x}_{\mathrm{bq}}, \mathbf{x}_{\mathrm{bq}}) + \lambda \mathbf{I}\right]^{-1} \mathbf{y}_{\mathrm{bq}}, \tag{69}$$
$$\boldsymbol{\Omega} := \left[k(\mathbf{x}_{\mathrm{bq}}, \mathbf{x}_{\mathrm{bq}}) + \lambda \mathbf{I}\right]^{-1}, \tag{70}$$

v and ℓ are the output scale and length scale of the RBF kernel, and λ is the Gaussian likelihood variance. Posterior predictive mean and variance are just a mixture of Gaussians; thus integration is tractable.

Then, we consider the soft-Copeland score, which is the marginalisation of x' from $x = (x, x')$. Marginalisation of Gaussian is just extracting the corresponding elements. We have

$$X = \begin{bmatrix} x \\ x' \end{bmatrix} \qquad \mathbf{x}_{\text{bq}} = \begin{bmatrix} x_{\text{bq}} \\ x'_{\text{bq}} \end{bmatrix} \qquad \mathbf{W} = \begin{bmatrix} \mathbf{W}' & \mathbf{0} \\ \mathbf{0} & \mathbf{W}'' \end{bmatrix} \tag{71}$$

Consequently, the soft-Copeland score is reduced to be:

$$\hat{\pi}_g(x) := \int_{\mathcal{X}} g(x, x') \mathrm{d}x', \tag{72}$$

$$= \mathrm{V}_{\mathcal{X}}^{-1} \int g(x, x') \mathrm{d}x', \tag{73}$$

where $\mathrm{V}_{\mathcal{X}} = \int \mathbb{E}_g[\hat{\pi}_g(x)]\mathrm{d}x$ is the normalizing constant to make $\hat{\pi}_g(x)$ become a probability density function.

$$\mathbb{E}_g\left[\hat{\pi}_g(x)\right] = \mathrm{V}_{\mathcal{X}}^{-1} \int \mathbb{E}_g\left[g(x, x')\right] \mathrm{d}x', \tag{74}$$

$$= \mathrm{V}_{\mathcal{X}}^{-1} \int \mu_{\text{bq}}(x, x') \mathrm{d}x', \tag{75}$$

$$= v' \mathrm{V}_{\mathcal{X}}^{-1} \int \mathcal{N}\left(X; \mathbf{x}_{\text{bq}}, \mathbf{W}\right)^\top \mathrm{d}x' \boldsymbol{\omega}, \tag{76}$$

$$= v' \mathrm{V}_{\mathcal{X}}^{-1} \int \mathcal{N}\left(\begin{bmatrix} x \\ x' \end{bmatrix}; \begin{bmatrix} x_{\text{bq}} \\ x'_{\text{bq}} \end{bmatrix}, \begin{bmatrix} \mathbf{W}' & \mathbf{0} \\ \mathbf{0} & \mathbf{W}'' \end{bmatrix} \right)^\top \mathrm{d}x' \boldsymbol{\omega}, \tag{77}$$

$$= v' \mathrm{V}_{\mathcal{X}}^{-1} \mathcal{N}(x; x_{\text{bq}}, \mathbf{W}')^\top \boldsymbol{\omega}, \tag{78}$$

Similarly, we place another GP on the variance, $\mathbb{V}_g\left[g(x, x')\right]$, then the procedure is the same:

$$\mathbb{V}_g\left[\hat{\pi}_g(x)\right] = \mathrm{V}_{\mathcal{X}}^{-1} \int \mathbb{E}_g\left[g(x, x')(1 - g(x, x'))\right] \mathrm{d}x', \tag{79}$$

$$\approx \mathrm{V}_{\mathcal{X}}^{-1} \int \mu'_{\text{bq}}(x, x') \mathrm{d}x', \tag{80}$$

$$= v' \mathrm{V}_{\mathcal{X}}^{-1} \mathcal{N}(x; x''_{\text{bq}}, \mathbf{W}''')^\top \boldsymbol{\omega}, \tag{81}$$

Then, the soft-Copeland score can be approximated by moment-matching the original Bernoulli distribution with the Gaussian distribution.

$$\pi_g(x) \approx \mathcal{N}(\hat{\pi}_g(x); \mathbb{E}_g\left[\hat{\pi}_g(x)\right], \mathbb{V}_g\left[\hat{\pi}_g(x)\right]). \tag{82}$$

This is a coarse approximation of Condorcet winner probability $\mathbb{P}(y = 1|x)$ and is probabilistically wrong (Gaussian is not bounded in [0, 1]). Still, recall our original motivation is to model the probability of global optimal location, which is not bounded in [0, 1]. In this sense, precisely computing Bernoulli distribution is not required. We further use this soft-Copeland score function to model the prior distribution on the continuous value y, which is not the Bernoulli distribution and rather assumes Gaussian distribution. Thus we adopt Gaussian moment-matching approximation.

However, this soft-Copeland score is not normalised over the domain, so we need to take further integral over x domain. Bayesian quadrature makes this integral analytical:

$$\mathrm{V}_{\mathcal{X}} = \int_{\mathcal{X}} \mathbb{E}_g\left[\hat{\pi}_g(x)\right] \mathrm{d}x \tag{83}$$

$$= v' \mathrm{V}_{\mathcal{X}}^{-1} \int \mathcal{N}(x; x_{\text{bq}}, \mathbf{W}')^\top \mathrm{d}x \boldsymbol{\omega}, \tag{84}$$

$$= v' \mathrm{V}_{\mathcal{X}}^{-1} \mathbf{1}^\top \boldsymbol{\omega}, \tag{85}$$

$$= \sqrt{v' \mathbf{1}^\top \boldsymbol{\omega}}, \tag{86}$$

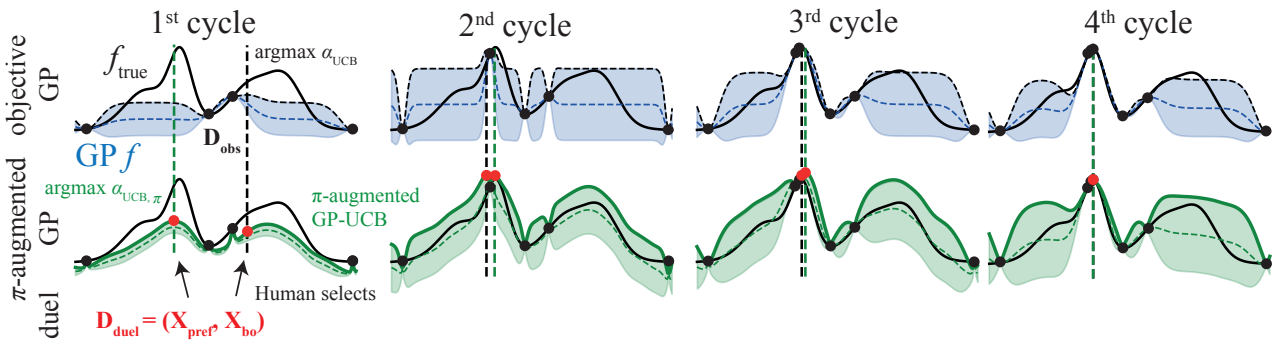

Figure 10: Dueling candidate generation algorithm. While the optimistic sample is selected by maximizing the π-augmented GP, the pessimistic sample is selected by the original objective GP. The distance between pairwise samples asymptotically decreases over iterations. (We set $\gamma = 0.1$).

$$\pi(x) := \frac{\mathbb{E}_g[\pi_g(x]}{\int_{\mathcal{X}} \mathbb{E}_g[\hat{\pi}_g(x)]\mathrm{d}x}, \tag{87}$$

$$= \frac{1}{\mathbf{1}^\top \boldsymbol{\omega}} \mathcal{N}(x; x_{\mathrm{bq}}, \mathbf{W}')^\top \boldsymbol{\omega} \tag{88}$$

Now we have the closed-form soft-Copeland score approximation. It should be noted that the original $g(x, x')$ distribution is the Bernoulli distribution, whereas this BQ-GP is the Gaussian distribution, which is a crude approximation. So only predictive mean is reliable. To estimate the variance of $\mathbb{V}_g[\hat{\pi}_g(x)]$ at the same time, we need to have another GP for variance estimation. Even the predictive mean of BQ-GP can be a crude approximation. One simple solution to boost the accuracy is to increase the number of data $\mathbf{D}_{\mathrm{bq}}$, as this can be augmented cheaply by f_{pref}. However, increasing the number can lead to large computational overhead as training GP costs cubic complexity $\mathcal{O}(n^3)$. We wish to minimise the number of augmented data. So we adopt BASQ Adachi et al. (2022) for selecting the next query point. This allows provably small predictive uncertainty (see Theorem 1 in Adachi et al. (2022) and also Hayakawa et al. (2022)).

E Dueling Acquisition Function

Figure 10 visualizes the dueling acquisition function. In the first cycle, the distance between preference-based and standard UCB-based recommendations is large. But it gradually decreases over iterations, and it is almost the same in the last (fourth) iteration. Furthermore, this figure also shows that the human preference successfully avoids climbing up the wrong left peak by placing the belief over the left peak, which accelerates early-stage exploration.

F Explaining Bayesian Optimization

We have three explanation features explained in Figure 2; Spatial relation, feature importance, and selection accuracy feedback. For the spatial relation, we select the two primary dimensions using Shapley values. As Shapley values are conditioned on x, we need to select x. For simplicity, we take average of Shapley values at the pairwise candidates, x_1 and x_2, then take two dimensions with top 2 mean Shapley values. For the drawing range, we first compute the rectangle which bounds the following three points; x_1, x_2, and the current best observed point $x_{\mathrm{current}} := \arg\max \mathbf{D}_{obj_t}$. Then, we expand this bounded rectangle 2 times as large as the original for visibility. We set this expanded rectangle as the visualisation range. This procedure is shared with the preference model visualisation. For feature importance, we simply visualise the Shapley values at x_1 and x_2 as a bar plot.

For selection accuracy feedback, we first update the GP surrogate function with the queried point $x_t \in (x_1, x_2)$ and the queried value $y_t = f(x_t)$. For the sake of argument, we assume $x_t = x_1$. Then, we compute the probability

of correct selection, given by:

$$\mathbb{P}(f(x_1) \geq f(x_2)) \sim \mathcal{N}(\mathbb{E}_f[\ell(x_1, x_2)], \mathbb{V}_f[\ell(x_1, x_2)]), \tag{89}$$

$$\ell(x_1, x_2) := \Phi\left(\frac{f(x_1) - f(x_2)}{\sqrt{\lambda}}\right), \tag{90}$$

where Φ is the cumulative density function of standard normal distribution $\mathcal{N}(0, 1)$. We compute the expectation and variance over f space by Monte Carlo integration.

G Experimental details

We have tested CoExBO for 5 synthetic functions against 6 baselines. We use a constant-mean GP with an RBF kernel. In each iteration of the active learning loop, the outputs are standardized to have zero mean and unit variance. We optimize the hyperparameter by maximizing the marginal likelihood (type-II maximum likelihood estimation) using L-BFGS-B optimizer (Liu & Nocedal, 1989) implemented with BoTorch (Balandat et al., 2020). We also maximized the acquisition function using the same optimizer. The initial data sets consist of ten data points drawn by Sobol sequence (Sobol', 1967) and corresponding noisy observations. We generate 10 samples for objective dataset and 100 samples for preferential dataset construction. We adopt the log regret as the evaluation metric using the test dataset. The models are implemented in GPyTorch (Gardner et al., 2018). All experiments are repeated ten times with different initial data sets via different random seeds (the seeds are shared with baseline methods).

G.1 Synthetic functions with Synthetic Human Selection

G.1.1 Synthetic Functions

Ackely Ackley funciton is defined as:

$$f(x) := -a \exp\left[-b\sqrt{\frac{1}{d}\sum_{i=1}^{d} x_i^2}\right] - \exp\left[\frac{1}{d}\sum_{i=1}^{d}\cos(cx_i)\right] + a + \exp(1) \tag{91}$$

where $a = 20, c = 2\pi, d = 4$. We take the negative Ackley function as the objective of BO to make this optimisation problem maximisation. This is a 4-dimensional function bounded by $x \in [-1, 1]^d$. The global optimum is $x^*_{\text{true}} = [0, 0, 0, 0]$ and $f(x^*_{\text{true}}) = 0$.

Hölder Table Hölder Table funciton is defined as:

$$f(x) := \left|\sin(x_1)\cos(x_2)\exp\left(\left|1 - \frac{\sqrt{x_1^2 + x_2^2}}{\pi}\right|\right)\right| \tag{92}$$

where x_i is the i-th dimensional input. This is a 2-dimensional function bounded by $x \in [0, 10]^d$. The global optimum is $x^*_{\text{true}} = [8.05502, 9.66459]$ and $f(x^*_{\text{true}}) = 19.2085$.

Styblinski-Tang Styblinski-Tang funciton is defined as:

$$f(x) := \frac{1}{2}\sum_{i=1}^{d}(x_i^4 - 16x_i^2 + 5x_i) \tag{93}$$

where x_i is the i-th dimensional input. This is a 3-dimensional function bounded by $x \in [-5, 5]^d$. The global optimum is $x^*_{\text{true}} = [-2.903534]^d$ and $f(x^*_{\text{true}}) = 39.166166d$.

Michalewicz Michalewicz funciton is defined as:

$$f(x) := \sum_{i=1}^{d}\sin(x_i)\sin^{2m}\left(\frac{ix_i^2}{\pi}\right) \tag{94}$$

where x_i is the i-th dimensional input and $m = 10$. This is a 5-dimensional function bounded by $x \in [0, \pi]^d$. The global optimum is $f(x^*_{\text{true}}) = 4.687658$.

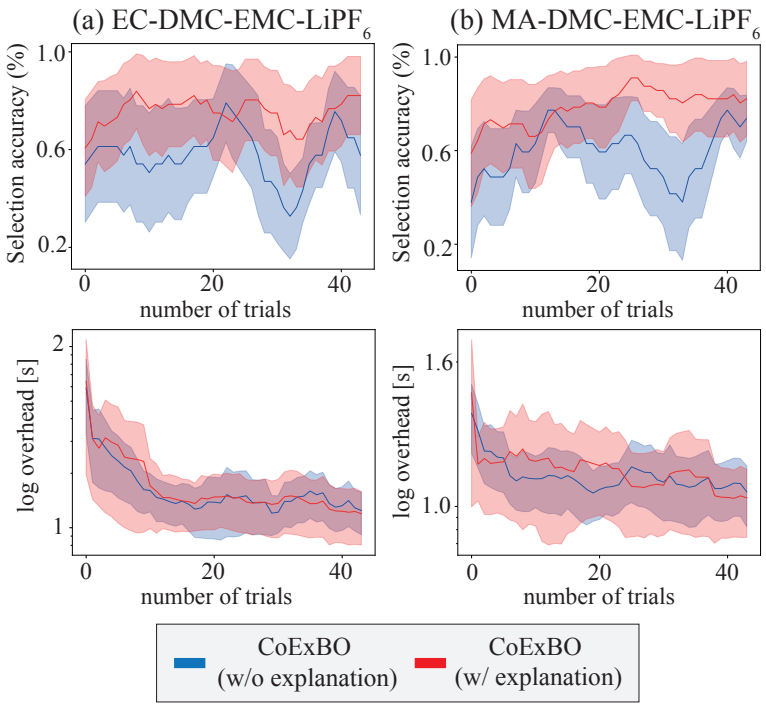

Figure 11: Selection accuracy and overhead analysis over iterations for the tasks (a) EC-DMC-EMC-LiPF$_6$ and (b) MA-DMC-EMC-LiPF$_6$. The solid lines and shaded areas refer to the mean and 1 standard error of the results of four participants. To smooth out the noisy results, we take the moving average with the window size of 3 trials for visualizing the trend.

Rosenbrock Rosenbrock funciton is defined as:

$$f(x) := \sum_{i=1}^{d-1} \left[100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2\right] \tag{95}$$

where x_i is the i-th dimensional input. This is a 3-dimensional function bounded by $x \in [-5, 10]^d$. The global optimum is $x^*_{\text{true}} = [1]^d$ and $f(x^*_{\text{true}}) = 0$.

G.1.2 Robustness evaluation

In the adversarial selection, both CoExBO and batchUCB superficially surpass the original UCB. We first point out that this is within the standard error, but there may be possible reasons why these two are better than others even in adversarial settings. For CoExBO, this may come from randomised effect. Recent work shows randomizing β parameter of UCB yields faster convergence than original (Berk et al., 2020; Takeno et al., 2023a). Our CoExBO can be understood as randomising β, which may effect positively. Still, we can confirm the trend that confident and correct human feedback can accelerate convergence. For batchUCB, this may come from a nonmyopic effect. González et al. (2016) pointed out the similarity between hallucination and one-step lookahead BO, which empirically yields better convergence than the original UCB.

G.2 Real-world tasks

G.2.1 Designing battery

The problem involves finding the best electrolyte material combination to maximize ionic conductivity. Ionic conductivity is dependent on both lithium salt molarity and the cosolvent composition. They show the complex non-linear relationship due to the solvation effect and cannot predict even with the state-of-the-art quantum chemistry simulator. We create the true functions by fitting the experimental data of EC-DMC-EMC-LiPF$_6$ (Dave et al., 2022) and MA-DMC-EMC-LiPF$_6$ (Logan et al., 2018) systems using the Casteel-Amis equation

(Casteel & Amis, 1972). Note that the Casteel-Amis equation is just for the interpolation of experimental data to be continuous, and is not capable of predicting different cosolvent. Both tasks are the three-dimensional continuous input function. The input features are (1) the lithium salt ($LiPF_6$) molarity, (2) DMC/EMC cosolvent ratio, and (3) (EC or MA)/carbonates cosolvent ratio, respectively. The inputs are bounded with $x_1 \in [0, 2]$, $x_2 \in [0, 1]$, and $x_3 \in [0, 1]$. The noisy output is generated by adding i.i.d. zero-mean Gaussian noise with the 3^2 variance to the noiseless $f(x)$.

G.2.2 Selection Accuracy and Complexity Analysis

We further analyzed the real-world task results based on selection accuracy and overhead over iterations. Figure 11 illustrates the results. For selection accuracy, while results with explanation remain stable over iterations, the ones without explanation fluctuate largely, particularly in the later rounds. Over iterations, the pairwise candidates become closer due to the no-harm guarantee. Hence, the later iterations are more difficult to select the correct one. The explanation feature can help users distinguish the slight differences by the quantitative Shapley values, leading to accurate selection even for the later iterations. The bottom row of Figure 11 shows the overhead of the candidate selection process, including pairwise candidate generation, explanation feature, and human selection time. We can observe the general decrease trend over iterations regardless of the explanation feature, and the difference in overhead between with and without explanation features is negligible. This is because the most time-consuming part is the human selection process. In the early stage, human users are also uncertain and need more time to decide which to select. Over time, it becomes more confident and smoother to select, resulting in quicker selection. This suggests the algorithmic overhead is negligible when compared to the human selection process.