arXiv:2603.28410v2 [cs.LG] 27 Aug 2026

Active Preference Learning over Latent Preference Archetypes for Many-Objective Bayesian Optimization

Manisha Dubey *[1,2†]
Sebastiaan De Peuter[3,4‡]
Wanrong Wang[1], **Samuel Kaski**[1,3,5]

[1] Department of Computer Science, University of Manchester, UK
[2] Centre for AI in Assistive Autonomy, University of Edinburgh, Scotland, UK
[3] Department of Computer Science, Aalto University, Espoo, Finland
[4] Informatics Institute, University of Amsterdam, Amsterdam, Netherlands
[5] ELLIS Institute Finland, Helsinki, Finland

Abstract

Preference-based many-objective Bayesian optimization typically assumes that all pairwise comparisons arise from a single latent utility function, despite real decision makers often exhibiting multiple latent preference archetypes across contexts. We propose an active preference learning framework for many-objective Bayesian optimization that infers latent preference archetypes from pairwise comparisons, enabling both the identification of the active trade-off strategy and the refinement of its associated preferences. Our framework represents preferences as a Dirichlet-process mixture of latent archetypes and introduces mixture-aware information-theoretic query strategies that separately target archetype identification and within-archetype refinement through a hybrid acquisition policy. Experiments on synthetic benchmarks and real-world chemical process design case-study consistently outperforms state-of-the-art preference-based Bayesian optimization methods while recovering interpretable latent preference structure beyond conventional single-utility models. The proposed mixture-aware diagnostics further quantify archetype recovery and preference calibration, providing insights that are not captured by optimization performance alone.

Introduction

Real-world applications such as materials and additive manufacturing (Hastings et al. 2025; Myung et al. 2025), power systems (Ah King, Rughooputh, and Deb 2005), robotics (Li et al. 2023; Shen et al. 2025), urban design (Liu et al. 2023), and recommender systems (Jannach and Abdollahpouri 2023) require optimizing many competing objectives (Deb, Sindhya, and Hakanen 2016) under expensive evaluations. Bayesian optimization provides an efficient framework for sequentially identifying high-quality solutions in such settings. As the number of objectives grows ($L \geq 4$), however, many-objective Bayesian optimization (MaOBO) becomes increasingly challenging due to rapidly expanding Pareto frontiers, weakened dominance relations, and the difficulty of selecting solutions that reflect a decision maker's priorities (Ishibuchi, Tsukamoto, and Nojima 2008; Binois et al. 2020; Zhang and Golovin 2020). While optimization algorithms can efficiently approximate large Pareto sets, the fundamental challenge increasingly shifts from generating candidate solutions to identifying which trade-offs are actually preferred.

Preference-based Bayesian optimization addresses this challenge by incorporating pairwise preference feedback to guide the search towards Pareto-optimal solutions that align with the decision maker's objectives, avoiding the need to approximate the entire Pareto frontier. Existing methods have demonstrated that actively eliciting preferences can substantially improve optimization efficiency (Astudillo and Frazier 2020; Ozaki et al. 2024; Xu et al. 2024; AV et al. 2022). However, they almost universally assume that all observed comparisons are generated by a single latent utility function. This assumption becomes restrictive in many-objective settings, where users often exhibit latent preference archetypes, for example, prioritizing cost, safety, or sustainability under different circumstances. Compressing these behaviours into a single utility function can blur preference estimates, reduce query efficiency, and ultimately misdirect optimisation towards less relevant regions of the Pareto front.

We address this limitation by introducing a mixture-aware active preference learning framework for many-objective Bayesian optimization that represents heterogeneous decision-maker preferences as latent preference archetypes rather than a single utility function. The inferred posterior over archetypes naturally decomposes preference elicitation into two complementary objectives: identifying the active archetype and refining its associated trade-offs. This leads to inter-archetype, intra-archetype, and hybrid information-theoretic query strategies that actively select informative pairwise comparisons. Combined with many-objective Bayesian optimization, the learned preference posterior guides expensive evaluations towards decision-maker-preferred regions of the Pareto front, enabling more efficient preference-guided optimization. Our contributions can be summarized as follows:

- We introduce a mixture-aware active preference learn-

*Corresponding author.
†This work was carried out while the author was at University of Manchester, UK
‡This work was carried out while the author was at Aalto University, Finland

ing framework for many-objective Bayesian optimization that models heterogeneous decision-maker preferences as latent preference archetypes using a Dirichlet-process mixture, relaxing the single-utility assumption.

- We develop inter-archetype, intra-archetype, and hybrid information-theoretic query strategies that actively identify latent preference archetypes and refine within-archetype trade-offs from pairwise comparisons.
- We integrate the inferred preference posterior into the Bayesian optimization loop to guide expensive evaluations towards decision-maker-preferred regions of the Pareto front, and introduce mixture-aware diagnostics for analysing archetype recovery and preference calibration.
- We demonstrate consistent improvements over state-of-the-art preference-based Bayesian optimization methods on synthetic benchmarks and a real-world chemical process design case study, while recovering interpretable latent preference archetypes beyond conventional single-utility models.

Problem Formulation

We consider many-objective Bayesian optimization (MaOBO) with $L \geq 4$ objectives. Let $\mathcal{X} \subset \mathbb{R}^d$ be a design space and $\mathbf{f} : \mathcal{X} \to \mathbb{R}^L$ an unknown vector-valued objective. Querying a design $x \in \mathcal{X}$ returns a (possibly noisy) observation $\mathbf{y} = \mathbf{f}(x) + \boldsymbol{\epsilon}$ where $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma^2 I)$ and evaluations are assumed expensive. Many-objective problems typically admit a large set of Pareto-optimal solutions. In human-in-the-loop settings, however, the goal is not to recover the entire Pareto front, but to identify solutions that are most desirable to a decision maker (DM). Direct evaluations of the DM's utility are unavailable. Instead, the learner may query pairwise preferences between previously observed outcomes. Given two outcomes $\mathbf{y}_i$ and $\mathbf{y}'_i$, the decision maker returns a binary comparison indicating whether $\mathbf{y}_i \succ \mathbf{y}'_i$. Further, we assume that the DM evaluates outcomes through an unknown utility function. To model heterogeneous trade-offs, we posit that the DM's utility is governed by one of K latent preference modes. Each mode $k \in [K]$ is parameterized by a weight vector $\mathbf{w}_k \in \Delta^{L-1}$, and modes occur with unknown mixture weights $\boldsymbol{\eta} \in \Delta^{K-1}$. This formulation captures either a single DM with context-dependent trade-offs or a heterogeneous population of DMs. We denote the scalar utility induced by mode k as $U(\mathbf{f}(x); \mathbf{w}_k)$. Our objective is to identify designs that maximize the DM's mixture-expected utility,

$$x^\star \in \arg\max_{x \in \mathcal{X}} \sum_{k=1}^{K} \eta_k \, \mathbb{E}[U(\mathbf{f}(\mathbf{x}); \mathbf{w}_k)] \,, \tag{1}$$

where the expectation is taken over observation noise. The learner must therefore infer the latent mixture parameters $\{\eta_k, \mathbf{w}_k\}_{k=1}^K$ from pairwise feedback while sequentially selecting designs to evaluate.

Proposed Methodology

In this section, we propose a preference-based many-objective Bayesian optimization framework that models decision-maker preferences as a mixture of latent preference archetypes, relaxing the conventional single-utility assumption (see Fig. 1). Modeling preferences as a mixture of latent preference archetypes allows multiple preference archetypes to coexist. This introduces two complementary sources of uncertainty: identifying the active preference archetype and learning its associated trade-offs. Accordingly, our framework consists of four components: (i) Gaussian process surrogates for the expensive objectives, (ii) a latent mixture preference model, (iii) variational posterior inference, and (iv) mixture-aware acquisition strategies for selecting both new evaluations and informative preference queries. We also present the algorithm 1 here.

Objective Surrogate Model

We model each objective with an independent Gaussian process (GP). Let $\mathbf{f} : \mathcal{X} \to \mathbb{R}^L$ denote the L objectives. Each objective function f_ℓ is modeled independently with a Gaussian process: $f_\ell \sim \mathcal{GP}(m_\ell, k_\ell), \quad \ell = 1, \ldots, L$. Conditioning on observed data yields predictive means and variances. Given observations $\mathcal{D}_f = \{(x_i, \mathbf{y}_i)\}$, we obtain posterior predictive distributions $f_\ell(x) \mid \mathcal{D}_f \sim \mathcal{N}(\mu_\ell(x), \sigma^2_\ell(x))$. Independence across objectives keeps inference and acquisition scalable in L (Rasmussen and Nickisch 2010; Williams and Rasmussen 2006). One could also employ multi-output gaussian process (Alvarez and Lawrence 2008).

Latent Preference Archetypes

Existing preference-based Bayesian optimization methods typically assume that all pairwise comparisons arise from a single latent utility function (Ozaki et al. 2023; Astudillo and Frazier 2020). We instead represent decision-maker preferences as a mixture of K latent preference archetypes, allowing multiple trade-off strategies to coexist. Unlike a single utility function, the mixture representation allows different pairwise comparisons to be explained by different latent preference archetypes, naturally capturing heterogeneous or context-dependent trade-off strategies. Each archetype corresponds to a distinct scalarization of the objectives, while the mixture captures uncertainty over which archetype generated each observed comparison. Each archetype k is parameterized by a weight vector $\mathbf{w}_k \in \Delta^{L-1}$ that defines a Chebyshev utility over objective outcomes. For a predicted outcome $\mathbf{y}$, the utility under archetype k is

$$U(\mathbf{y}; \mathbf{w}_k) = -\min_\ell \frac{y_\ell}{w_{k\ell}}, \qquad w_{k\ell} > 0, \tag{2}$$

where utilities are computed using the GP posterior predictive means. Rather than fixing a single preference model, we place a Dirichlet-process mixture prior over the archetypes. Using a truncated stick-breaking construction (Blei and Jordan 2006), $v_k \sim \mathrm{Beta}(1, \alpha), \eta_k = v_k \prod_{j<k}(1 - v_j)$, yielding mixture weights $\boldsymbol{\eta} \in \Delta^{K-1}$. Each archetype is assigned a preference weight vector $\mathbf{w}_k \sim \mathrm{Dir}(\boldsymbol{\beta})$ representing a distinct trade-off over the objectives. For each observed comparison $(\mathbf{y}_i, \mathbf{y}'_i)$, a latent archetype assignment $z_i \sim \mathrm{Categorical}(\boldsymbol{\eta})$ selects the preference mode responsible for that comparison. Conditioned on $z_i = k$, pairwise preferences follow the probit

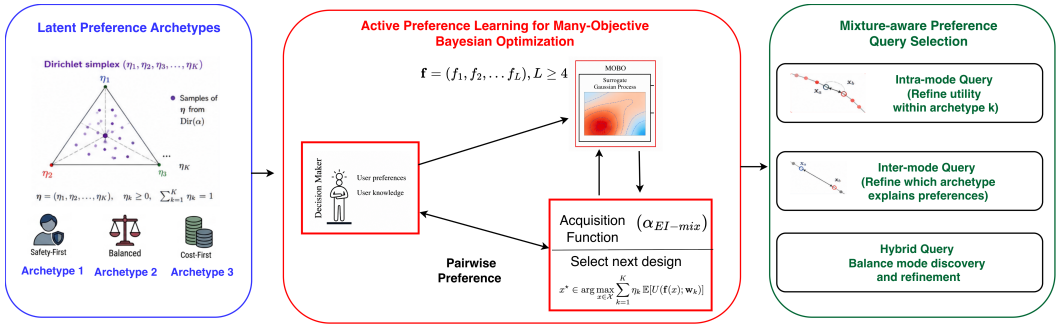


Figure 1: Overview of the proposed active preference learning framework for many-objective Bayesian optimization. Pairwise comparisons are used to infer latent preference archetypes and their mixture weights(left). The inferred posterior guides Gaussian-process-based many-objective Bayesian optimization through a mixture-aware acquisition function and iterative preference feedback from decision maker (centre). Active query selection chooses informative intra-mode, inter-mode, and hybrid comparisons to jointly refine utility estimation and latent preference archetype discovery (right).

random-utility model,

$$P(\mathbf{y}_i \succ \mathbf{y}'_i \mid z_i = k) = \Phi\left(\frac{U(\mathbf{y}_i;\mathbf{w}_k) - U(\mathbf{y}'_i;\mathbf{w}_k)}{\sqrt{2}\sigma_u}\right). \tag{3}$$

Then we marginalize latent assignments as

$$p(\mathcal{D}_{\text{pref}} \mid \boldsymbol{\eta}, \mathbf{w}_{1:K}) = \prod_{i=1}^{N}\sum_{k=1}^{K} \eta_k \Phi\left(\frac{U(\mathbf{y}_i;\mathbf{w}_k) - U(\mathbf{y}'_i;\mathbf{w}_k)}{\sqrt{2}\sigma_u}\right) \tag{4}$$

We adopt the standard probit random-utility model for pairwise comparisons (Astudillo and Frazier 2020; Ozaki et al. 2023). The proposed framework is independent of this choice and can be combined with alternative pairwise likelihoods such as Bradley-Terry-Luce.

Posterior Inference

Let $\mathcal{D}_{\text{pref}} = \{(\mathbf{y}_i, \mathbf{y}'_i)\}_{i=1}^{N}$ denote the observed pairwise comparisons. Given the generative model, our objective is to infer the posterior over the latent preference variables, $p(\boldsymbol{\eta}, \mathbf{w}_{1:K}, \mathbf{z}, \alpha \mid \mathcal{D}_{\text{pref}})$ where $\boldsymbol{\eta}$, $\mathbf{w}_{1:K}$, $\mathbf{z}$, and α denote the mixture weights, preference archetype weights, latent archetype assignments, and concentration parameter, respectively. The corresponding joint distribution is

$$p(\mathcal{D}_{\text{pref}}, \mathbf{z}, \boldsymbol{\eta}, \mathbf{w}_{1:K}, \alpha) \tag{5}$$
$$= p(\alpha)\prod_{k=1}^{K} p(v_k|\alpha)\, p(\mathbf{w}_k) \prod_{i=1}^{N} p(z_i \,|\, \boldsymbol{\eta})\, p(\mathbf{y}_i \succ \mathbf{y}'_i \,|\, z_i, \mathbf{w}_{z_i})$$

Exact posterior inference is intractable because the latent archetype assignments are marginalized and the probit likelihood is non-conjugate. We therefore approximate the posterior using mean-field stochastic variational inference. Under the latent mixture preference model, the unknown variables consist of the mixture weights $\boldsymbol{\eta}$, archetype weights $\mathbf{w}_{1:K}$, latent archetype assignments $\mathbf{z}$, and concentration parameter α. The resulting posterior $p(\boldsymbol{\eta}, \mathbf{w}_{1:K}, \mathbf{z}, \alpha \mid \mathcal{D}_{\text{pref}})$ is analytically intractable because of the latent mixture structure and non-conjugate preference likelihood. We therefore approximate the posterior using a mean-field variational distribution

$$q(\boldsymbol{\eta}, \mathbf{w}_{1:K}, \mathbf{z}, \alpha) = q(\alpha)\prod_{k=1}^{K} q(v_k)\, q(\mathbf{w}_k) \prod_{i=1}^{N} q(z_i). \tag{6}$$

The variational parameters are obtained by maximizing the evidence lower bound (ELBO)

$$\mathcal{L}(q) = \mathbb{E}_q[\log p(\mathcal{D}_{\text{pref}}, \mathbf{z}, \boldsymbol{\eta}, \mathbf{w}_{1:K}, \alpha)] - \mathbb{E}_q[\log q(\boldsymbol{\eta}, \mathbf{w}_{1:K}, \mathbf{z}, \alpha)] \tag{7}$$

using stochastic variational inference (SVI).

Design Acquisition

At iteration t, the learner maintains a posterior over the objective functions and latent preference parameters. Let $\hat{\theta} = (\hat{\boldsymbol{\eta}}, \hat{\mathbf{w}}_{1:K})$ denote the posterior mean estimates of the mixture weights and preference archetype weights obtained from the variational posterior. The GP surrogate induces a posterior $p(\mathbf{f}(x) \mid \mathcal{D}_t)$ over the objective vector at any candidate design x. For a given objective vector $\mathbf{y}$, the mixture utility is computed as $U_{\text{mix}}(\mathbf{y}; \hat{\theta}) = \sum_{k=1}^{K} \hat{\eta}_k\, U(\mathbf{y}; \hat{\mathbf{w}}_k)$, which represents the expected utility under the inferred mixture of preference archetypes. The next design is selected by maximizing the mixture expected improvement (mixture-EI),

$$\alpha_{\text{EI-mix}}(x) = \mathbb{E}_{\mathbf{f}(x)}\left[\max\left(U_{\text{mix}}(\mathbf{f}(x); \hat{\theta}) - U_{\text{best}}, 0\right)\right], \tag{8}$$

where $U_{\text{best}} = \max_{i \le t} U_{\text{mix}}(\mathbf{y}_i; \hat{\theta})$ is the best observed mixture utility under the inferred preference model. The expectation is taken with respect to the GP posterior and is approximated using Monte Carlo samples of $\mathbf{f}(x)$. This acquisition balances exploration and exploitation while incorporating the inferred mixture of latent preference archetypes through the posterior mean preference estimates. Using posterior mean estimates provides a computationally efficient plug-in approximation to the fully Bayesian acquisition, avoiding nested Monte Carlo integration over both objective and preference uncertainty.

Preference Query Acquisition

In addition to evaluating new designs, the learner may select a pairwise comparison between two previously observed

outcomes $(\mathbf{y}_i, \mathbf{y}_j)$. Let $r \in \{0, 1\}$ denote the random response of the decision maker (DM), where $r = 1$ indicates $\mathbf{y}_i \succ \mathbf{y}_j$. Let $\theta = (\boldsymbol{\eta}, \mathbf{w}_{1:K})$ denote the latent preference parameters, and let $q_t(\theta)$ denote the current variational posterior. The predictive probability of the response is $p(r = 1 \mid y_i, y_j, \mathcal{D}_t) = \int p(r = 1 \mid y_i, y_j, \theta) q_t(\theta) d\theta$. The DM's preference r updates the posterior $p(\theta \mid \mathcal{D}_t)$, which in turn influences future design decisions. The value of a preference query can therefore be interpreted as its expected reduction in uncertainty about θ, and hence its expected improvement in downstream decision quality. Computing the exact expected value of a preference query is generally intractable, as it requires updating the posterior for every possible response. Following (Ozaki et al. 2023), we therefore adopt a Bayesian Active Learning by Disagreement (BALD) criterion (Houlsby et al. 2011), which approximates query value by the mutual information between the response r and the latent preference parameters θ.

$$\alpha_{\text{pref}}(i, j) = I(r; \theta \mid \mathbf{y}_i, \mathbf{y}_j, \mathcal{D}_t) \tag{9}$$

Using the entropy decomposition of mutual information, this can be written as

$$I(r; \theta) = H[r] - \mathbb{E}_{\theta \sim p(\theta \mid \mathcal{D}_t)} \left[H[r \mid \theta] \right] \tag{10}$$

where $H[\cdot]$ denotes Bernoulli entropy. The first term measures overall predictive uncertainty, while the second term measures expected uncertainty under a fixed parameter setting. Their difference quantifies how much the response would reduce posterior uncertainty. The conditional preference probability given θ is

$$p(r = 1 \mid \theta) = \sum_{k=1}^{K} \eta_k \, \Phi\left(\frac{U(\mathbf{y}_i; \mathbf{w}_k) - U(\mathbf{y}_j; \mathbf{w}_k)}{\sqrt{2}\sigma_u} \right)$$

The marginal predictive probability is approximated by Monte Carlo samples from the variational posterior $q_t(\theta)$. Under the proposed mixture model, preference queries naturally decompose into two complementary objectives: reducing uncertainty over the active preference archetype (inter-cluster queries) and reducing uncertainty over the preference weights within an archetype (intra-cluster queries). A hybrid acquisition combines these objectives through a convex combination. Please note that other information-theoretic criteria (e.g., predictive entropy) could be substituted within the same framework.

Inter-archetype query In the mixture preference model, uncertainty arises both from which preference archetype governs the DM and from the weights within each mode. The inter-cluster acquisition specifically targets uncertainty about the latent preference archetype. Let K denote the discrete latent archetype variable and $r \in \{0, 1\}$ the response to a comparison $(\mathbf{y}_i, \mathbf{y}_j)$. We define the inter-cluster acquisition as the mutual information between the response and the archetype identity, which can be further decomposed as:

$$I(r; K) = H[r] - \mathbb{E}_K[H[r \mid K]] \,. \tag{11}$$

The entropy form becomes

$$A_{inter} = H[p_{\text{mix}}] - \sum_{k=1}^{K} \eta_k H[p_k]. \tag{12}$$

where mixture predictive is $p_{\text{mix}} = \sum_{k=1}^{K} \eta_k p_k$. This quantity is large when the mixture predictive is uncertain (high $H[p_{\text{mix}}]$) but each individual cluster prediction is confident.

Algorithm 1: MIXTURE-BASED MAOBO

Require: $\mathbf{f} : \mathcal{X} \to \mathbb{R}^L$, truncation K, query mode
1: Initialize dataset $\mathcal{D}_f$ and preference set $\mathcal{D}_{pref}$
2: **for** $t = 1, 2, \ldots$ **do**
3: Fit L independent GPs on $\mathcal{D}_f$
4: Update DP-mixture posterior via SVI to obtain $(\hat{\eta}, \{\hat{w}_k\}_{k=1}^{K})$
5: $\mathbf{x}_t \leftarrow \arg\max_{\mathbf{x} \in \mathcal{X}} \mathsf{EI}_{\mathsf{mix}}(\mathbf{x}; \hat{\eta}, \{\hat{\mathbf{w}}_k\})$ (using Eq. 8)
6: Evaluate $\mathbf{y}_t = \mathbf{f}(\mathbf{x}_t)$ and update $\mathcal{D}_f$
7: $(\mathbf{y}_a, \mathbf{y}_b) \leftarrow$ SELECTPAIR$(\{\mathbf{y}_i\}, \hat{\eta}, \{\hat{\mathbf{w}}_k\}, \mathsf{mode})$
8: Query for preference label $\mathbf{y}_{\text{pref}} \succ \mathbf{y}_{\text{other}}$
9: Set $\mathcal{D}_{pref} \leftarrow \mathcal{D}_{pref} \cup \{(\mathbf{y}_{\text{pref}}, \mathbf{y}_{\text{rej}})\}$
10: **end for**

Intra-archetype query While the inter-mode criterion targets uncertainty over the active archetype, the intra-mode criterion targets uncertainty over the archetype weights within a particular preference archetype. For a candidate comparison between two previously observed outcomes $(\mathbf{y}_i, \mathbf{y}_j)$ with response $r \in \{0, 1\}$ and target archetype $c \in [K]$,we define the intra-mode acquisition as the mutual information between r and the archetype-c weight vector:

$$A^{(c)}_{\text{intra}}(i, j) = I(r; \mathbf{w}_c \mid K = c, \, \mathbf{y}_i, \mathbf{y}_j, \, \mathcal{D}_t) \tag{13}$$

Using the BALD entropy decomposition, we obtain

$$A^{(c)}_{\text{intra}}(i, j) = H(\bar{p}_c) - \mathbb{E}_{\mathbf{w} \sim q_t(\mathbf{w}_c)} \left[H(p_{\mathbf{w}}) \right]$$

where $\bar{p}_c = \mathbb{E}_{\mathbf{w} \sim q_t(\mathbf{w}_c)}[p_{\mathbf{w}}]$ and $q_t(\mathbf{w}_c)$ denotes the current approximate posterior (variational factor) over $\mathbf{w}_c$ after observing $\mathcal{D}_t$.

Hybrid query The hybrid acquisition balances two complementary objectives: identifying the active preference archetype (inter-cluster exploration) and refining the preference weights within that archetype (intra-cluster refinement).

$$A_{\text{hybrid}} = \lambda A_{\text{inter}} + (1 - \lambda) A_{\text{intra}}, \quad \lambda \in [0, 1]. \tag{14}$$

The hyperparameter λ controls the exploration trade-off between mode disambiguation and weight refinement.

Under exact GP inference, the dominant per-iteration computational cost arises from fitting L independent Gaussian process (GP) surrogates over n observations, resulting in a complexity of $\mathcal{O}(Ln^3)$. The additional cost of variational mixture inference scales approximately linearly with the number of mixture components K and the number of preference observations q, giving $\mathcal{O}(Kq)$ (or $\mathcal{O}(Kn)$ when q scales with n). Consequently, a practical upper bound on the per-iteration computational complexity is $\mathcal{O}(Ln^3 + Kq)$ where GP updates remain the dominant computational bottleneck, while the overhead introduced by the mixture model is approximately linear in the number of active components.

Theoretical Proofs

Our analysis builds on standard high-probability concentration bounds for Gaussian process (GP) regression (Srinivas et al. 2010; Kandasamy et al. 2016; AV et al. 2022), which hold uniformly over arbitrary evaluation sequences and are

therefore independent of the acquisition strategy. We combine these bounds with Lipschitz properties of the Chebyshev utility to translate objective-level approximation errors into errors in the inferred mixture utility. Since our acquisition jointly depends on GP surrogate accuracy and latent preference estimation, we derive a high-probability simple regret decomposition that separates the contributions of GP approximation, preference archetype estimation, mixture-weight estimation, and acquisition optimization.

Theorem 1 (Simple regret decomposition)**.** *Let $\mathcal{X}$ be compact and let $\mathbf{f} = (f_1, \dots, f_L) : \mathcal{X} \to \mathbb{R}^L$ be an L-objective function. Assume:*

(A1) Each objective f_ℓ lies in an RKHS $\mathcal{H}_\ell$ with kernel k_ℓ and $\|f_\ell\|_{\mathcal{H}_\ell} \le B_f$.

(A2) For all $x \in \mathcal{X}$, $\|\mathbf{f}(x)\|_\infty \le B_y$.

(A3) Preference weights satisfy $w_{k\ell} \ge c_w > 0$ for all modes k and objectives ℓ.

(A4) Observations are corrupted by independent σ^2–sub-Gaussian noise.

Define the Chebyshev utility for minimization

$$U(\mathbf{y}; \mathbf{w}) = -\min_{1 \le \ell \le L} \frac{y_\ell}{w_\ell},$$

and the mixture utility

$$U_{\mathrm{mix}}(\mathbf{y}; \eta, \{w_k\}) = \sum_{k=1}^{K} \eta_k U(\mathbf{y}; w_k).$$

Let $(\eta^\star, w_k^\star)$ denote the true mixture parameters and define

$$U^\star = \sup_{x \in \mathcal{X}} U_{\mathrm{mix}}(\mathbf{f}(x); \eta^\star, \{w_k^\star\}).$$

After T evaluations $x_1, \dots, x_T$, define the simple regret

$$R_T = U^\star - \max_{t \le T} U_{\mathrm{mix}}(\mathbf{f}(x_t); \eta^\star, \{w_k^\star\}).$$

Then, with probability at least $1 - \delta$,

$$R_T \;\le\; C_1 \sqrt{\frac{\beta_T \gamma_T}{T}} + C_2\, \Delta_w(T) + C_3\, \Delta_\eta(T) + \varepsilon_T,$$

where γ_T is the maximal GP information gain, β_T is the standard GP confidence parameter, $\Delta_w(T) = \max_k \|\hat{w}_k - w_k^\star\|_1$, $\Delta_\eta(T) = \|\hat{\eta} - \eta^\star\|_1$, ε_T is the optimization error in maximizing the acquisition, C_1, C_2, C_3 depend only on B_y and c_w.

Remark 1. Theorem 1 provides an error decomposition for the final simple regret rather than a classical sublinear cumulative regret guarantee. Its primary purpose is to identify the distinct sources of error introduced by objective approximation, latent preference estimation, and numerical optimization.

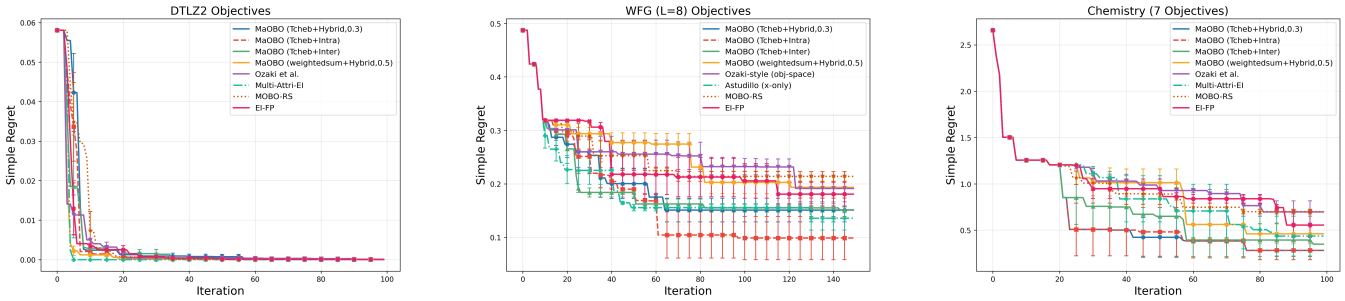
Under consistent preference estimation, i.e., when $\Delta_w(T) \to 0$, $\Delta_\eta(T) \to 0$, and the acquisition optimization error $\varepsilon_T \to 0$, the bound is dominated by the GP approximation term, which decreases at the standard GP convergence rate $\mathcal{O}\Big(\sqrt{\beta_T \gamma_T / T}\Big)$. The detailed proof is provided in the supplementary.

Related Work

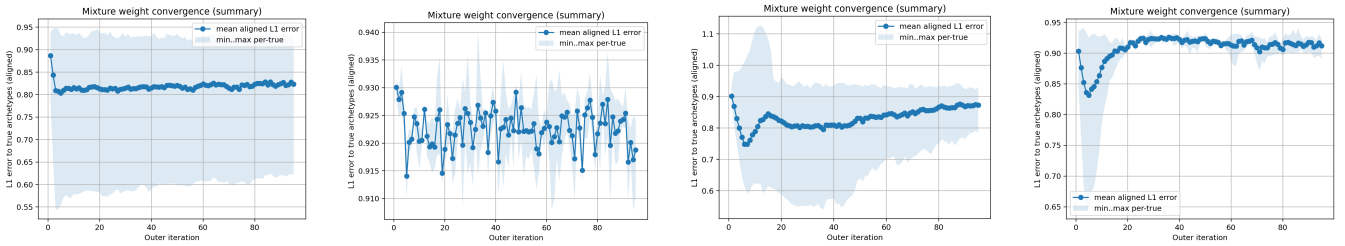
While there are several recent efforts along the direction of multi-objective Bayesian optimization (Li et al. 2025; Haddadnia, Grashoff, and Strieth-Kalthoff 2025; Hung et al. 2025), a line of works integrate human-in-the-loop for Bayesian optimization (Xu et al. 2024; AV et al. 2022). Towards interactive multi-objective optimization, (Ozaki et al. 2024) proposes a framework that optimizes multiple objectives by actively querying the decision-maker in order to steer Bayesian optimization towards user's preferred trade-offs. There are limited efforts in the direction of many-objectove optimization. Classical many-objective Bayesian optimization increasingly avoids recovering large Pareto sets and instead targets a principled single compromise. The Kalai-Smorodinsky (KS) solution equalizes benefit ratios from a disagreement point to utopia; subsequent work further introduces a copula-invariant variant (CKS) to remove sensitivity to monotone rescalings, together with a GP-based stepwise-uncertainty-reduction (SUR) scheme that scales to 4-9 objectives (Binois et al. 2020). These methods elegantly collapse decision making to a point solution, but they assume access to full objective vectors (rather than comparisons) and a single underlying taste of utility. A complementary line seeks a small set of solutions that collaboratively "cover" many objectives. Recent work proposes Tchebycheff-set (TCH-Set) scalarization and a smooth variant (STCH-Set) that mitigates the non-smooth max operator, with theory and empirical studies showing that a handful of solutions can handle tens of objectives effectively (Lin et al. 2024). Such approaches reduce downstream decision complexity, yet still optimize from full objective feedback and typically posit a single, fixed scalarization rather than heterogeneous preferences. Orthogonal to both is objective reduction where detecting redundant objectives via similarity between GP predictive distributions and dropping them to save evaluations. This can preserve solution quality on toy, synthetic, and real setups while cutting cost (Martín and Garrido-Merchán 2021), but it maintains a single user model and does not reason about preference heterogeneity. Relative to single–compromise and few–solution paradigms (Binois et al. 2020; Lin et al. 2024; Jiang, Huang, and Li 2026), our framework operates directly on preferences and embraces heterogeneity as signal, not noise. Relative to objective–reduction (Martín and Garrido-Merchán 2021), we keep all objectives but learn which trade–off modes matter and when, delivering interpretable decisions under many objectives.

Experiments

Datasets We evaluate on two standard many-objective optimization benchmarks, DTLZ2 (Deb et al. 2005) and WFG9 (Huband et al. 2005), representing smooth and highly non-separable optimization landscapes, respectively. To simulate heterogeneous preferences, we generate pairwise comparisons from K latent preference archetypes, each defined by a Chebyshev utility with weight vector $\mathbf{w}_k$. For DTLZ2 ($L = 6$) and WFG9 ($L = 8$), archetypes emphasize different objective groups by assigning 80% of the weight to one group and distributing the remainder across the others. We consider both an i.i.d. setting, where latent modes are sampled independently from $\mathrm{Categorical}(\eta^\star)$, and a persistent setting, where mode identity evolves according to a sticky Markov process with persistence probability ρ. We further evaluate our method on a real-world chemical process design case study



(a) DTLZ2 (b) WFG (c) Chemistry



(d) Clusterless (e) Intra (f) Inter (g) Hybrid

Figure 2: (Top) Mixture-aware policies reduce regret faster and achieve lower final regret than unimodal and random-scalarization baselines, with Hybrid performing best overall. This is visible from the steeper early decline and lower terminal curves of Inter and Hybrid across DTLZ2, WFG, and chemical task, while Clusterless and scalarization methods plateau higher with greater variability. Curves show mean simple regret over five independent runs; vertical bars denote $\pm$ one standard error.(Bottom)Hybrid most accurately and stably recovers true archetypes, showing the largest early drop and lowest final error with the narrowest band; Inter drops quickly but plateaus, Intra refines slowly with higher variance, and Clusterless remains flat. We report mean aligned L1 error per outer iteration (band = per-true min–max; lower is better). Dataset: DTLZ2, persistent context.

for polyethylene terephthalate (PET) production (Wang et al. 2022), comprising 10,000 simulated designs with 12 decision variables and seven objectives (return on investment and six life-cycle assessment metrics). We construct three stakeholder archetypes representing economic, environmental, and health priorities, with true mixture weights $\eta = [0.40, 0.35, 0.25]$ and persistence $\rho = 0.8$. Additional dataset and simulation details are provided in the Supplementary.

Learning and Policies We adopt Chebyshev scalarization for its ability to represent non-convex trade-offs and its standard use in many-objective optimization, though the framework can accommodate alternative scalarizations. Each objective is modeled with an independent GP, and preference modes are learned via a truncated stick-breaking mixture using SVI with priors $v_k \sim \mathrm{Beta}(1, \alpha)$ and $\mathbf{w}_k \sim \mathrm{Dirichlet}(\boldsymbol{\beta})$. Although we fix a truncation level K, the stick-breaking prior shrinks redundant components toward negligible mass, so effective modes correspond to components with non-trivial posterior weight. The hybrid parameter λ is treated as a hyperparameter and selected via validation.

Baselines We compare against following optimization baselines: (i)*Ozaki* (Ozaki et al. 2023), learns preference vector interactively using MCMC with single utility function; (ii) *Multi-Attr-EI* (Astudillo and Frazier 2020), which infers a single preference vector using MAP estimation; (iii) *MOBO-RS* (Paria, Kandasamy, and Póczos 2020), based on random scalarizations; (iv) *EI-FP*, expected improvement with known fixed preferences; acting as upper performance bound and (v) *MaOBO-WS*, our framework using weighted-sum scalarization instead of Chebyshev. For preference query acquisition, we compare *Random*, *Clusterless* (unimodal BALD using the mixture-mean preference vector), *Inter* (mode identification; Eq. 11), *Intra* (within-mode refinement; Eq. 13), and *Hybrid* (convex combination of Inter and Intra; Eq. 14). Existing preferential BO methods assume a single latent utility, while random scalarization explores diverse trade-offs without inferring heterogeneous preferences. In contrast, our method explicitly models multi-modal, dynamically switching preference archetypes.

Metrics We report (i) *Simple Regret*, $r_t = U^* - \max_{i \le t} U(f(x_i), w_{\text{true}})$ (lower is better); (ii) *Archetype Recovery*, measured as Hungarian-aligned mean ℓ_1 error between inferred and true archetypes; (iii) *Mixture-weight trajectories*, tracking convergence of posterior means $\hat{\eta}$ to the true mixture weights; and (iv) *Per-component ℓ_1 error*, which reveals mode collapse, label switching, and convergence of individual archetypes.

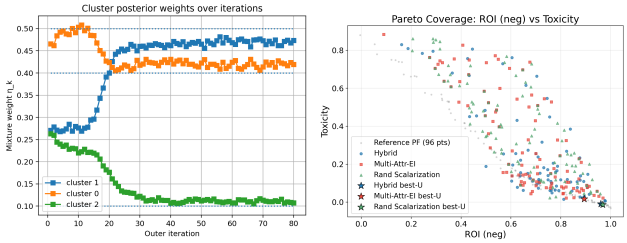
Results and Analysis

Figure 2 (Top) compares simple regret under the true mixture utility across outer iterations, with uncertainty estimate over five seeds. Across benchmarks, the mixture-aware methods achieve faster regret reduction and lower final regret than unimodal and random-scalarization baselines. For instance,

(Ozaki et al. 2023) achieves regret as 0.697 as opposed to 0.282 with 60% reduction, reflecting that prior methods assume a single utility, whereas proposed method models heterogeneous preferences via latent archetypes. On DTLZ2, all methods improve rapidly due to the smooth landscape, but Inter and Hybrid converge more consistently to near-zero regret with lower variance, while Clusterless and fixed-preference baselines lag slightly, indicating consistent gains from explicit mode identification even in well-behaved settings. On the more irregular WFG benchmark, the separation is clearer: Hybrid exhibits the steepest early decline and the lowest final regret, whereas Clusterless and random scalarization plateau at higher levels, suggesting that preference averaging fails to resolve complex trade-offs. In the chemical process design task, the gap is most pronounced: Hybrid converges faster and stabilizes at lower regret, while unimodal and scalarization baselines converge more slowly and with higher variance, highlighting the benefit of modeling heterogeneous archetypes in structured, multimodal trade-off landscapes. Figure 2 (Bottom) shows archetype recovery for a persistent user on DTLZ2. Clusterless exhibits a flat curve with a wide band, indicating mode averaging and poor archetype separation. Intra-queries reduce error within active modes but produce unstable overall alignment, reflected in persistent variability. Inter-queries yield a sharp early drop by rapidly disambiguating archetypes, followed by slower refinement. Hybrid combines both effects: rapid initial mode identification followed by steady within-mode calibration, achieving lowest final aligned error and the narrowest uncertainty band.

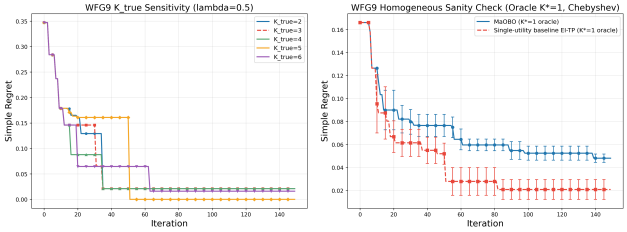
Figure 5 illustrates the evolution of inferred mixture weights under persistent preferences with true weights $(0.40, 0.35, 0.25)$. Hybrid rapidly reallocates probability mass during the early iterations to identify the latent preference modes before converging to stable estimates close to the ground truth. The smallest archetype remains slightly underestimated, reflecting the increased difficulty of learning minority preference modes from limited preference queries. Importantly, no spurious dominant component emerges and the mixture weights stabilize after approximately 20 iterations. Simple regret alone does not reveal how different preference archetypes are explored. We therefore visualize evaluated solutions relative to a reference Pareto front. As shown in Fig. 5, Hybrid explores multiple regions of the Pareto frontier before concentrating on the mixture-preferred region. Compared with competing methods, it achieves better localization of solutions near the preferred trade-off surface, particularly along the ROI–Toxicity projection.

Robustness and Scalability We evaluate the method under varying preference complexity and model structure. Figure 4(a) varies the true number of preference archetypes from $K^\star = 2$ to $K^\star = 6$. Increasing $K^\star$ generally slows convergence because more preference observations are required to distinguish the latent archetypes. Nevertheless, all settings attain low final regret, indicating that the method scales gracefully with increasing preference heterogeneity. Figure 4(b) considers a homogeneous oracle with $K^\star = 1$. As expected, a specialized single-utility baseline performs best in this setting, while the proposed mixture model remains competitive by



(a) Hybrid: correct mode (b) ROI vs Toxicity

Figure 3: Mixture weight trajectory for chemical process design process. (left) Hybrid quickly identifies and stabilizes near the correct mode proportions. (right) Grey: reference Pareto front; coloured: evaluated solutions; stars: final best-utility solution. Hybrid converges to the preferred Pareto region.



(a) Scalability with increasing preference heterogeneity (b) Homogenous Oracle ($K^* = 1$)

Figure 4: Robustness Analysis for WFG

concentrating posterior mass on a single dominant archetype. Additional studies on K-misspecification, utility misspecification, and sensitivity to the hybrid parameter λ are reported in the Supplementary.

Conclusion

We introduced a mixture-aware active preference learning framework for many-objective Bayesian optimization that moves beyond the standard assumption of a single latent utility function. By representing decision-maker preferences as latent preference archetypes, the proposed framework jointly infers heterogeneous preference patterns from pairwise comparisons and actively acquires informative preference queries. This representation naturally decomposes preference elicitation into archetype identification and within-archetype refinement, leading to inter-archetype, intra-archetype, and hybrid information-theoretic query strategies that improve preference-guided optimization. Experiments on synthetic benchmarks and a real-world chemical process design case study demonstrate that the proposed framework more efficiently identifies preferred regions of the Pareto front while consistently improving optimization performance over existing preference-based Bayesian optimization methods. Beyond optimization, the learned preference archetypes recover interpretable latent preference archetypes, and the proposed diagnostics quantify archetype recovery and preference calibration beyond conventional regret-based evaluation.

Together, these results establish latent preference archetypes as a principled foundation for active preference

learning in many-objective Bayesian optimization. The proposed framework is orthogonal to existing many-objective Bayesian optimization methods and can be combined with alternative surrogate models, scalarizations, or acquisition functions. While our evaluation uses simulated preferences and a real-world engineering case study, an important next step is validation with human decision makers, broader real-world applications, and adaptive models of non-stationary preferences.

References

Ah King, R. T.; Rughooputh, H. C.; and Deb, K. 2005. Evolutionary multi-objective environmental/economic dispatch: Stochastic versus deterministic approaches. In *International conference on evolutionary multi-criterion optimization*, 677–691. Springer.

Alvarez, M.; and Lawrence, N. 2008. Sparse convolved Gaussian processes for multi-output regression. *Advances in neural information processing systems*, 21.

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

AV, A. K.; Rana, S.; Shilton, A.; and Venkatesh, S. 2022. Human-AI collaborative Bayesian optimisation. *Advances in neural information processing systems*, 35: 16233–16245.

Binois, M.; Picheny, V.; Taillandier, P.; and Habbal, A. 2020. The Kalai-Smorodinsky solution for many-objective Bayesian optimization. *Journal of Machine Learning Research*, 21(150): 1–42.

Blei, D. M.; and Jordan, M. I. 2006. Variational inference for Dirichlet process mixtures.

Deb, K.; Sindhya, K.; and Hakanen, J. 2016. Multi-objective optimization. In *Decision sciences*, 161–200. CRC Press.

Deb, K.; Thiele, L.; Laumanns, M.; and Zitzler, E. 2005. Scalable test problems for evolutionary multiobjective optimization. In *Evolutionary multiobjective optimization: theoretical advances and applications*, 105–145. Springer.

Haddadnia, M.; Grashoff, L.; and Strieth-Kalthoff, F. 2025. BoTier: multi-objective Bayesian optimization with tiered objective structures. *Digital Discovery*.

Hastings, T.; Mulukutla, M.; Khatamsaz, D.; Salas, D.; Xu, W.; Lewis, D.; Person, N.; Skokan, M.; Miller, B.; Paramore, J.; et al. 2025. Accelerated multi-objective alloy discovery through efficient bayesian methods: Application to the fcc high entropy alloy space. *Acta Materialia*, 121173.

Houlsby, N.; Huszár, F.; Ghahramani, Z.; and Lengyel, M. 2011. Bayesian active learning for classification and preference learning. *arXiv preprint arXiv:1112.5745*.

Huband, S.; Barone, L.; While, L.; and Hingston, P. 2005. A scalable multi-objective test problem toolkit. In *International Conference on Evolutionary Multi-Criterion Optimization*, 280–295. Springer.

Hung, Y.-H.; Lin, K.-J.; Lin, Y.-H.; Wang, C.-Y.; Sun, C.; and Hsieh, P.-C. 2025. BOFormer: Learning to Solve Multi-Objective Bayesian Optimization via Non-Markovian RL. *arXiv preprint arXiv:2505.21974*.

Ishibuchi, H.; Tsukamoto, N.; and Nojima, Y. 2008. Evolutionary many-objective optimization: A short review. In *2008 IEEE congress on evolutionary computation (IEEE world congress on computational intelligence)*, 2419–2426. IEEE.

Jannach, D.; and Abdollahpouri, H. 2023. A survey on multi-objective recommender systems. *Frontiers in big Data*, 6: 1157899.

Jiang, C.; Huang, J.; and Li, M. 2026. Do We Really Need to Approach the Entire Pareto Front in Many-Objective Bayesian Optimisation. *arXiv preprint arXiv:2604.09417*.

Kandasamy, K.; Dasarathy, G.; Oliva, J. B.; Schneider, J.; and Póczos, B. 2016. Gaussian process bandit optimisation with multi-fidelity evaluations. *Advances in neural information processing systems*, 29.

Li, B.; Di, Z.; Lu, Y.; Qian, H.; Wang, F.; Yang, P.; Tang, K.; and Zhou, A. 2025. Expensive multi-objective bayesian optimization based on diffusion models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 27063–27071.

Li, G.; Guo, H.; Wang, Z.; and Wang, M. 2023. Online trajectory optimization for safe autonomous overtaking with active obstacle avoidance. *Robotics and Autonomous Systems*, 169: 104528.

Lin, X.; Liu, Y.; Zhang, X.; Liu, F.; Wang, Z.; and Zhang, Q. 2024. Few for many: Tchebycheff set scalarization for many-objective optimization. *arXiv preprint arXiv:2405.19650*.

Liu, K.; Xu, X.; Huang, W.; Zhang, R.; Kong, L.; and Wang, X. 2023. A multi-objective optimization framework for designing urban block forms considering daylight, energy consumption, and photovoltaic energy potential. *Building and Environment*, 242: 110585.

Martín, L. A.; and Garrido-Merchán, E. C. 2021. Many objective bayesian optimization. *arXiv preprint arXiv:2107.04126*.

Myung, J. I.; Deneault, J. R.; Chang, J.; Kang, I.; Maruyama, B.; and Pitt, M. A. 2025. Multi-objective Bayesian optimization: a case study in material extrusion. *Digital Discovery*, 4(2): 464–476.

Ozaki, R.; Ishikawa, K.; Kanzaki, Y.; Suzuki, S.; Takeno, S.; Takeuchi, I.; and Karasuyama, M. 2023. Multi-objective bayesian optimization with active preference learning. *arXiv preprint arXiv:2311.13460*.

Ozaki, R.; Ishikawa, K.; Kanzaki, Y.; Takeno, S.; Takeuchi, I.; and Karasuyama, M. 2024. Multi-objective Bayesian optimization with active preference learning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 38, 14490–14498.

Paria, B.; Kandasamy, K.; and Póczos, B. 2020. A flexible framework for multi-objective bayesian optimization using random scalarizations. In *Uncertainty in Artificial Intelligence*, 766–776. PMLR.

Rasmussen, C. E.; and Nickisch, H. 2010. Gaussian processes for machine learning (GPML) toolbox. *The Journal of Machine Learning Research*, 11: 3011–3015.

Shen, N.; You, H.; Li, J.; and Song, P. 2025. Real-time trajectory planning for collaborative robots using incremental multi-objective optimization. *Intelligent Service Robotics*, 18(1): 43–59.

Srinivas, N.; Krause, A.; Kakade, S. M.; and Seeger, M. 2010. Gaussian process optimization in the bandit setting: No regret and experimental design. In *ICML*.

Wang, Z.; Li, J.; Rangaiah, G. P.; and Wu, Z. 2022. Machine learning aided multi-objective optimization and multi-criteria decision making: Framework and two applications in chemical engineering. *Computers & Chemical Engineering*, 165: 107945.

Williams, C. K.; and Rasmussen, C. E. 2006. *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA.

Xu, W.; Adachi, M.; Jones, C. N.; and Osborne, M. A. 2024. Principled Bayesian Optimization in Collaboration with Human Experts. *Advances in Neural Information Processing Systems*, 37: 104091–104137.

Zhang, R.; and Golovin, D. 2020. Random hypervolume scalarizations for provable multi-objective black box optimization. In *International conference on machine learning*, 11096–11105. PMLR.

Supplementary Material

Detailed Experimental Details

Datasets

We evaluate on two standard many-objective benchmarks, DTLZ2 (Deb et al. 2005) and WFG9 (Huband et al. 2005), together with a real-world chemical process design benchmark for polyethylene terephthalate (PET) production (Wang et al. 2022).

DTLZ2: DTLZ2 uses $L = 6$ objectives and $d = 7$ decision variables over $[0, 1]^7$. All objectives are minimized.

WFG9: WFG9 uses $L = 8$ objectives and $d = 34$ variables ($k = 14$, $l = 20$). Variables are normalized to $[0, 1]$. The benchmark exhibits bias and mixed separability, making it challenging for preference learning.

PET process design: The PET benchmark contains 10,000 simulated process designs with 12 normalized decision variables and seven objectives consisting of ROI and six Life Cycle Assessment (LCA) indicators (GWP, terrestrial acidification, water consumption, fossil depletion, surplus ore and human toxicity). We treat the dataset as a fixed black-box many-objective optimization problem.

Simulated Datasets

We use well-known MOO benchmarks for our experimental setup, called DTLZ (Deb et al. 2005) and WFG suite (Huband et al. 2005). We evaluate on DTLZ2 with $L = 6$ objectives and $d = 7$ decision variables (minimization), search space $[0, 1]^d$. At each design $\mathbf{x} \in [0, 1]^7$ we observe $\mathbf{y} = \mathbf{f}(\mathbf{x})$. We also use WFG9 with $L = 8$ objectives and $d = 34$ decision variables. Following common practice, we set the WFG position parameter $k = 2(L - 1) = 14$ and distance parameter $l = 20$ so $d = k + l$. Variables are scaled to $[0, 1]$ and all objectives are minimized. A design $\mathbf{x} \in [0, 1]^d$ evaluates to an objective vector $\mathbf{y} = \mathbf{f}(\mathbf{x}) \in \mathbb{R}^8$. WFG9 induces challenging biases and mixed separability, making it a stronger testbed for query policies.

To construct distinct but overlapping preference archetypes, we partition the objectives into disjoint groups and assign structured weight vectors. For the DTLZ suite ($L = 6$), we define three groups $\mathcal{G}_1 = \{0, 1\}$, $\mathcal{G}_2 = \{2, 3\}$, $\mathcal{G}_3 = \{4, 5\}$ while for the WFG suite ($L = 8$) we define four groups $\mathcal{G}_1 = \{0, 1\}$, $\mathcal{G}_2 = \{2, 3\}$, $\mathcal{G}_3 = \{4, 5\}$, $\mathcal{G}_4 = \{6, 7\}$. Each mode allocates 80% of its total mass uniformly across its dominant group $\mathcal{G}_k$ and distributes the remaining 20% uniformly over the other objectives. The resulting vector is renormalized to ensure $\mathbf{w}_k \in \Delta^{L-1}$. This construction yields clearly separated trade-off profiles while maintaining nonzero sensitivity to all objectives.

Let $Z_t \in \{1, \dots, K\}$ denote the active preference mode at query t. Under an iid context regime, each query is generated independently according to a fixed mixture $Z_t \sim \mathrm{Categorical}(\eta^\star)$, with $\eta^\star = (0.5, 0.3, 0.2)$ for DTLZ and $\eta^\star = \{(0.60, 0.25, 0.15, 0.10),\ 0.4, 0.3, 0.3, 0.1\}$ for WFG. Under a persistent context regime, the mode evolves according to a stay-or-resample process with persistence parameter $\rho \in [0, 1)$. The initial mode is drawn as $Z_1 \sim \mathrm{Categorical}(\eta^\star)$, and at each subsequent query the mode remains unchanged with probability ρ and is resampled from $\mathrm{Categorical}(\eta^\star)$ with probability $1-\rho$. The expected run length within a single mode is $1/(1 - \rho)$, and the expected number of switches over T queries is $(1 - \rho)(T - 1)$.

Real-World: PET Process Design

We evaluate the proposed framework on a real-world multi-objective process design problem for polyethylene terephthalate (PET) production. The dataset comprises 10,000 previously evaluated process designs. Each design is represented by a 12-dimensional continuous decision vector $\mathbf{x} \in \mathbb{R}^{12}$ corresponding to operating and design parameters such as temperatures, pressures, residence times, and feed ratios. Each variable is bounded within physically meaningful limits and normalized to $[0, 1]$ for modeling purposes. For each design, we consider $L = 7$ objectives consisting of Return on Investment (ROI), life-Cycle Assessment (LCA) indicators, including global warming potential (GWP), terrestrial acidification, water consumption, fossil depletion, surplus ore, and human toxicity. For the PET benchmark we construct three stakeholder archetypes representing economic, environmental and health-oriented priorities. Mixture weights are fixed to $\eta = (0.40, 0.35, 0.25)$ with persistence $\rho = 0.8$.

Interaction Contexts

We consider two preference-generation regimes.

I.i.d: The active archetype is sampled independently at each comparison, $Z_t \sim \mathrm{Categorical}(\eta^\star)$.

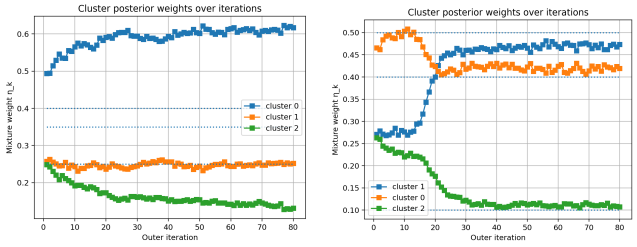
Persistent: Under the persistent regime, the active archetype remains unchanged with probability ρ and is resampled from $\mathrm{Categorical}(\eta^\star)$ with probability $1 - \rho$, yielding temporally consistent preferences with occasional switches. This models temporally consistent preferences with occasional switches.

Implementation Details

Experiments are implemented in Python using JAX, NumPyro, GPy and Optuna. Each objective is modeled by an independent RBF Gaussian process. Preference inference uses a truncated stick-breaking Dirichlet-process mixture optimized with stochastic variational inference (Adam, learning rate 10^{-3}, 500 updates). Preference queries are selected using the proposed clusterless, inter-, intra-, or hybrid criteria, while new evaluations are chosen via Monte Carlo expected improvement (12 MC samples, 10 L-BFGS-B restarts). All methods share identical initial designs, random seeds, and evaluation budgets. Results are reported as mean$\pm$std over three independent runs. All experiments were conducted on a MacBook Pro equipped with an Apple M2 Pro processor (10-core CPU, 16-core GPU), 32 GB unified memory, running macOS. The implementation used Python 3.12 with JAX, NumPyro, GPy, and Optuna. No distributed computing or external GPU cluster was used.

Symbol	Meaning
L	number of objectives
K	truncation level (max modes)
ℓ	objective index
k	preference mode index
$\mathbf{w}_k$	weight vector for mode k
η_k	mixture weight of mode k
r_i	latent mode for comparison i
σ_u	preference noise parameter

Table 1: Notation summary.



(a) Clusterless (b) Hybrid

Figure 5: Mixture weight trajectory for PET production process. Clusterless (left) collapses to a dominant component, while Hybrid (right) quickly identifies and stabilizes near the correct mode proportions.

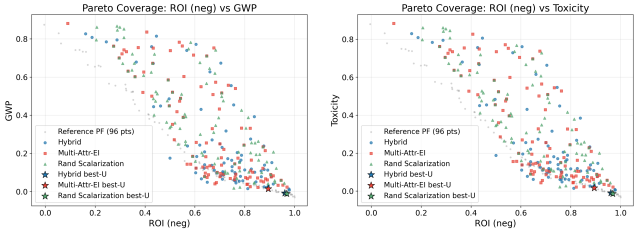
Additional Metrics

Besides simple regret, we report 1) Mixture-weight calibration by comparing inferred $\hat{\eta}$ with ground-truth $\eta^\star$ after Hungarian alignment. 2) Preference estimation error using mixture-mean, MAP-cluster and expected L_1 errors.

Results and Analysis

Additional Preference-Recovery Diagnostics

Simple regret measures optimization performance under the true mixture utility, but does not reveal whether the inferred preference model recovers the underlying archetype structure.



(a) ROI vs GWP (b) ROI vs Toxicity

Figure 6: Pareto coverage in objective space. Grey points-reference Pareto front; colored points - evaluated solutions; stars - final best-utility solution under the true mixture preference. Efficient trade-offs lie along the grey frontier. Hybrid achieves broader frontier coverage and identifies final solutions closest to the mixture-preferred Pareto region.

We therefore report additional diagnostics for mixture calibration, archetype recovery, and localization on the Pareto front.

Mixture-weight recovery: Figure 5 compares the evolution of the posterior mixture weights on the PET process-design case study. The Clusterless policy increasingly concentrates probability on a single component, indicating mode averaging and weak identification of minority archetypes. In contrast, Hybrid maintains three non-negligible components and stabilizes substantially closer to the true mixture proportions $(0.40, 0.35, 0.25)$. The minority component remains slightly underestimated, reflecting the difficulty of recovering less frequent archetypes from a finite number of comparisons.

Localization on the Pareto front: Figure 6 projects the evaluated PET designs onto the ROI–GWP and ROI–Toxicity planes. Different archetypes may favour distinct regions of the Pareto front, which aggregate regret does not expose. Both Hybrid and the unimodal baseline sample broadly along the projected efficient frontier, but Hybrid produces a final best-utility solution slightly closer to the region preferred under the true mixture utility. These projections provide qualitative evidence that mixture-aware preference learning improves localization in decision-relevant regions, rather than merely reducing scalar regret.

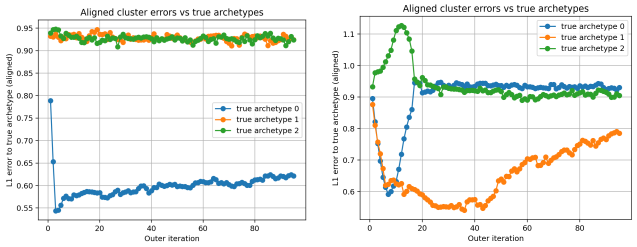
Archetype recovery and mode collapse: Figure 7 reports component-wise L_1 errors after Hungarian alignment with the true archetypes. Under Clusterless querying, only one inferred component improves substantially while the remaining components stay at high error, indicating mode collapse. Inter queries reduce errors across multiple components by selecting comparisons that differentiate competing archetypes, although non-monotone adjustments remain as the inferred component alignment changes during learning.

Mixture calibration across query policies: Figure 8 compares inferred and true mixture weights after alignment. Random querying assigns excessive mass to a single component, while Intra primarily refines the currently dominant archetype and does not reliably calibrate the global mixture. Inter querying produces the closest correspondence between inferred and true component masses, supporting its intended role in archetype identification.

Robustness Analysis

We evaluate the robustness of the proposed framework under four challenging settings: homogeneous preferences, utility misspecification, varying degrees of preference heterogeneity, and misspecified mixture complexity.

Homogeneous preferences: Figure 9 considers the degenerate case where all decision makers share a single preference archetype ($K^* = 1$). As expected, the specialized single-utility baseline performs best because the true preference model is unimodal. Nevertheless, the proposed mixture-based framework remains competitive on both DTLZ2 and WFG, indicating that the additional flexibility introduced by the mixture model does not substantially degrade optimization performance when heterogeneous preferences are absent.



(a) Clusterless (b) Inter

Figure 7: L_1 error trajectories per inferred component for a persistent user on DTLZ. After Hungarian alignment (lower is better), Inter queries prevent mode collapse and rapidly reduce error for at least one archetype, while Clusterless largely collapses, with only one component improving and others remaining high

Utility misspecification: Figure 10 evaluates robustness when the optimization model assumes a different utility representation from the oracle generating preferences (Chebyshev model versus weighted-sum oracle). Performance remains nearly identical to the well-specified setting across both benchmarks, suggesting that the proposed method is not overly dependent on the exact functional form of the utility model. Instead, it primarily exploits the relative preference structure revealed through pairwise comparisons.

Sensitivity to the hybrid query parameter: Figure 11 studies the influence of the hybrid query parameter λ, which balances inter-mode exploration against intra-mode refinement. Performance is relatively stable across a broad range of values, while intermediate settings ($\lambda \approx 0.3$–0.5) consistently achieve the lowest regret. Very small λ values focus excessively on local utility refinement before identifying the correct archetype, whereas very large values spend too many queries distinguishing modes at the expense of optimization.

Scalability with increasing preference heterogeneity: Figure 12 investigates increasing numbers of latent preference archetypes. As K_{true} increases, convergence naturally becomes slower because more preference information is required to distinguish among competing archetypes. Despite this increased identification complexity, the final regret remains nearly unchanged, demonstrating that the framework scales gracefully with growing preference diversity.

Misspecified number of preference archetypes: Figure 13 evaluates robustness when the assumed number of latent preference modes differs from the true oracle ($K^* = 4$). Moderate under-specification produces only a small degradation in performance because similar archetypes are effectively merged. Conversely, over-specification primarily slows convergence, as additional components receive little posterior probability before being discarded during inference. In both cases, the final optimization performance remains close to that of the correctly specified model, indicating robustness to moderate errors in selecting the mixture size.

Theoretical Proofs

In this section, we detail the proof of the theorem stated in the main paper. We state model setup and assumptions. Then we prove Lipschitz properties of the utility. Further, we discuss surrogate mismatch bound. Further, we combine all the pieces for the regret decomposition theorem.

Model Assumptions

We consider m objectives to be *minimized*. For any outcome vector $y \in \mathbb{R}^m$ and weights $w \in \Delta^{m-1}$, define the Chebyshev utility

$$U(y, w) := -\min_{1 \le j \le m} \frac{y_j}{w_j}. \tag{15}$$

Thus larger utility is preferred. We assume all weights are bounded away from the simplex boundary: there exists $c_w > 0$ such that $w_j \ge c_w$ for all j for every true archetype $w_k^\star$ and every estimated archetype $\hat{w}_{k,t}$.

Let the true latent preferences be a mixture of $K_\star$ archetypes with weights $\eta^\star \in \Delta^{K_\star - 1}$ and archetypes $\{w_k^\star\}_{k=1}^{K_\star} \subset \Delta^{m-1}$. Define the true mixture utility at design $x \in \mathcal{X}$ by

$$\mathcal{U}^\star(x) := \sum_{k=1}^{K_\star} \eta_k^\star \, U\big(f(x), w_k^\star\big), \qquad \mathcal{U}^\star := \sup_{x \in \mathcal{X}} \mathcal{U}^\star(x). \tag{16}$$

The algorithm fits one GP per objective f_j and maintains posterior mean $\mu_{t-1,j}(x)$ and standard deviation $\sigma_{t-1,j}(x)$ after $t-1$ evaluations. Let $\mu_{t-1}(x) \in \mathbb{R}^m$ be the vector of means. At round t, the preference model outputs estimates $\hat{\eta}_t \in \Delta^{K-1}$ and $\{\hat{w}_{k,t}\}_{k=1}^{K}$. Define the surrogate mixture utility used for acquisition

$$\widehat{\mathcal{U}}_t(x) := \sum_{k=1}^{K} \hat{\eta}_{k,t} \, U\big(\mu_{t-1}(x), \hat{w}_{k,t}\big). \tag{17}$$

Assume the selected point x_t is ε_t-optimal for the surrogate:

$$\widehat{\mathcal{U}}_t(x_t) \ge \sup_{x \in \mathcal{X}} \widehat{\mathcal{U}}_t(x) - \varepsilon_t. \tag{18}$$

Assumption 1 (Bounded Objectives)**.** $\|f(x)\|_\infty \le B_y \quad \forall x \in \mathcal{X}$

Alignment and estimation errors Let π_t be a permutation (e.g., Hungarian matching) aligning estimated archetypes to true archetypes. Pad $\eta^\star$ with zeros if $K > K_\star$. Define

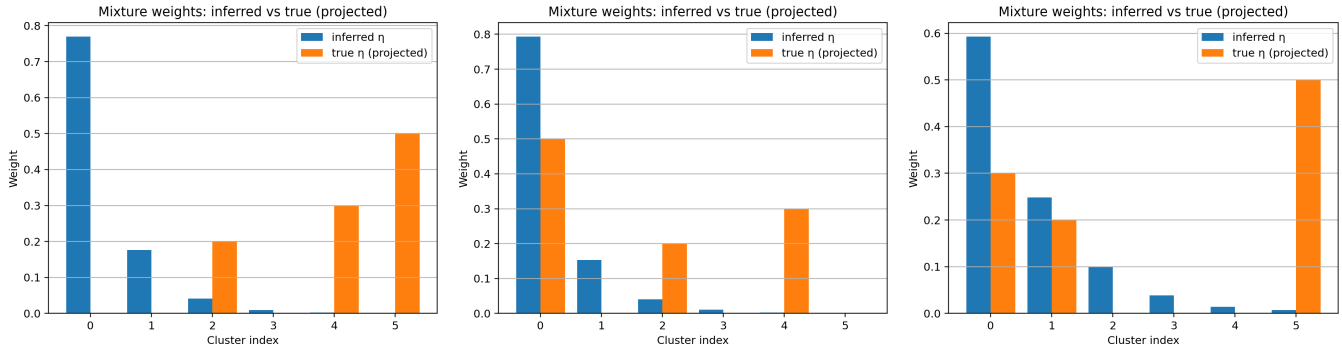
$$\Delta_t^w := \sum_{k=1}^{K_\star} \eta_k^\star \, \|\hat{w}_{\pi_t(k),t} - w_k^\star\|_1, \qquad \Delta_t^\eta := \|\hat{\eta}_t - \eta^\star\|_1. \tag{19}$$

Lemma 2 (Lipschitzness of Chebyshev utility)**.** *Let U be as in* (15) *and assume $w_j, w_j' \ge c_w > 0$ for all j. Then for any $y, y' \in \mathbb{R}^m$ and any such $w, w' \in \Delta^{m-1}$,*

$$\big|U(y, w) - U(y', w)\big| \le \frac{1}{c_w} \|y - y'\|_\infty, \tag{20}$$

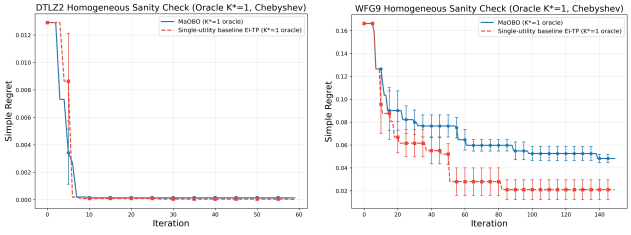
$$\big|U(y, w) - U(y, w')\big| \le \frac{\|y\|_\infty}{c_w^2} \|w - w'\|_1. \tag{21}$$

Proof. For (20), fix w and define $g_j(y) := -y_j / w_j$. Each g_j is $(1/c_w)$-Lipschitz under $\|\cdot\|_\infty$ because $|g_j(y) - g_j(y')| = |y_j - y_j'|/w_j \le \|y - y'\|_\infty / c_w$. Since $U(y, w) = \min_j g_j(y)$



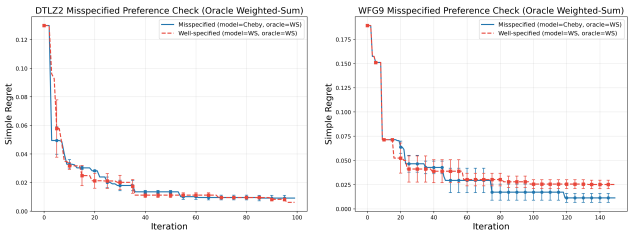
(a) Random (b) Intra (c) Inter

Figure 8: Mixture-weight calibration for persistent user for DTLZ dataset. Inter queries best match the true mixture, while random collapses to one mode and intra largely ignores η. Blue bars are inferred $\hat{\eta}$, orange bars are projected true η^* via Hungarian alignment. (a) Random: blue mass on one index and orange-only bars $\rightarrow$ mode collapse. (b) Intra: blue focuses on one component $\rightarrow$ poor allocation across modes. (c) Inter: clear blue–orange alignment $\rightarrow$ best identification and calibration across modes



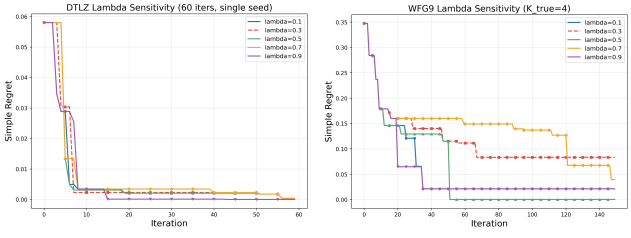
(a) DTLZ (b) WFG

Figure 9: Homogenous Oracle ($K^* = 1$)



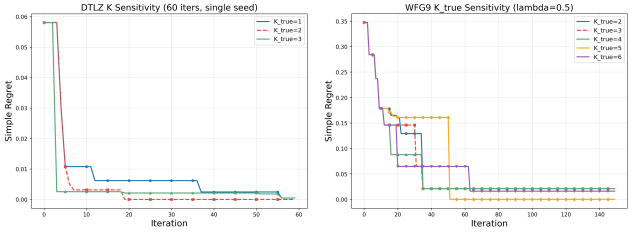
(a) DTLZ (b) WFG

Figure 10: Utility Misspecification



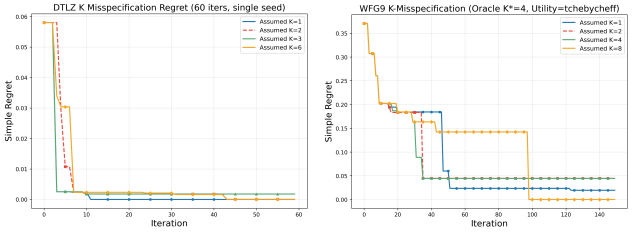
(a) DTLZ Sensitivity (b) WFG

Figure 11: Lambda Sensitivity



(a) DTLZ (b) WFG

Figure 12: Scalability with increasing preference heterogeneity



(a) DTLZ (b) WFG

Figure 13: Sensitivity to misspecified preference modes

is the pointwise minimum of $(1/c_w)$-Lipschitz functions, it is itself $(1/c_w)$-Lipschitz, proving (20).

For (21), fix y and write $U(y,w) = \min_j h_j(w)$ with $h_j(w) \coloneqq -y_j/w_j$. For any j,

$$\begin{aligned} |h_j(w) - h_j(w')| &= |y_j| \left| \frac{1}{w_j} - \frac{1}{w'_j} \right| \\ &= \frac{|y_j|\,|w_j - w'_j|}{w_j w'_j} \\ &\le \frac{\|y\|_\infty}{c_w^2}\,|w_j - w'_j|. \end{aligned} \tag{22}$$

Therefore

$$\begin{aligned} |U(y,w) - U(y,w')| &\le \max_j |h_j(w) - h_j(w')| \\ &\le \frac{\|y\|_\infty}{c_w^2} \max_j |w_j - w'_j| \le \frac{\|y\|_\infty}{c_w^2} \|w - w'\|_1. \end{aligned} \tag{23}$$

which proves (21). □

GP confidence

We use a standard high-probability confidence bound for each objective GP. Assume sub-Gaussian observation noise and conditions under which the standard GP-UCB confidence event holds for each objective (e.g., (Srinivas et al. 2010)).

Assumption 2 (Per-objective GP confidence event)**.** *There exists a nondecreasing sequence $\{\beta_t\}$ such that with probability at least $1-\delta$, simultaneously for all $t \ge 1$, all $x \in \mathcal{X}$, and all $j \in \{1,\dots,m\}$,*

$$|f_j(x) - \mu_{t-1,j}(x)| \le \sqrt{\beta_t}\,\sigma_{t-1,j}(x). \tag{24}$$

Under (24), we have for all x,

$$\|f(x) - \mu_{t-1}(x)\|_\infty \le \sqrt{\beta_t}\,\|\sigma_{t-1}(x)\|_\infty. \tag{25}$$

Main result: simple regret decomposition

Define the simple regret under the *true* mixture utility:

$$R_T \coloneqq \mathcal{U}^\star - \max_{t \le T} \mathcal{U}^\star(x_t). \tag{26}$$

Theorem 3 (Simple regret decomposition)**.** *On the GP confidence event in Assumption 2, for any $T \ge 1$,*

$$\begin{aligned} R_T \le\ & \underbrace{\frac{2}{c_w} \min_{t \le T} \|f(x_t) - \mu_{t-1}(x_t)\|_\infty}_{\text{Objective GP error}} + \underbrace{\frac{2B_y}{c_w^2} \min_{t \le T} \Delta_t^w}_{\text{Archetype error}} \\ &+ \underbrace{\frac{2B_y}{c_w} \min_{t \le T} \Delta_t^\eta}_{\text{Mixture-weight error}} + \underbrace{\max_{t \le T} \varepsilon_t}_{\text{Acquisition optimization error}}. \end{aligned} \tag{27}$$

Moreover, still on the same event, the objective GP term admits an existence-type rate:

$$\begin{aligned} \min_{t \le T} \|f(x_t) - \mu_{t-1}(x_t)\|_\infty &\le \frac{1}{T} \sum_{t=1}^{T} \|f(x_t) - \mu_{t-1}(x_t)\|_\infty \\ &\lesssim \sqrt{\frac{\beta_T\,\gamma_T}{T}}. \end{aligned} \tag{28}$$

where γ_T is a (max) information-gain term for the objective GPs (taking the maximum over objectives if kernels differ), and $\lesssim$ hides constants depending only on the kernels and noise.

Proof. Let $x^\star \in \arg\max_{x \in \mathcal{X}} \mathcal{U}^\star(x)$. Fix any round $t \ge 1$. Add and subtract the surrogate utility (17) at $x^\star$ and x_t:

$$\begin{aligned} \min_{t \le T} \|f(x_t) - \mu_{t-1}(x_t)\|_\infty &\le \frac{1}{T} \sum_{t=1}^{T} \|f(x_t) - \mu_{t-1}(x_t)\|_\infty \\ &\lesssim \sqrt{\frac{\beta_T\,\gamma_T}{T}}. \end{aligned} \tag{29}$$

where γ_T denotes the maximum information gain of the objective GPs (maximized over objectives if different kernels are used), and $\lesssim$ hides constants depending only on the kernels and observation noise. By ε_t-optimality (18), the middle term is at most ε_t. Therefore,

$$\mathcal{U}^\star(x^\star) - \mathcal{U}^\star(x_t) \le \left|\mathcal{U}^\star(x^\star) - \widehat{\mathcal{U}}_t(x^\star)\right| + \left|\mathcal{U}^\star(x_t) - \widehat{\mathcal{U}}_t(x_t)\right| + \varepsilon_t. \tag{30}$$

It remains to bound $|\mathcal{U}^\star(x) - \widehat{\mathcal{U}}_t(x)|$ for a generic x.

Decomposition Using (16) and (17),

$$\mathcal{U}^\star(x) - \widehat{\mathcal{U}}_t(x) = \sum_{k=1}^{K_\star} \eta_k^\star U(f(x), w_k^\star) - \sum_{k=1}^{K} \hat{\eta}_{k,t} U(\mu_{t-1}(x), \hat{w}_{k,t}).$$

Insert and subtract $\sum_{k=1}^{K_\star} \eta_k^\star U(\mu_{t-1}(x), w_k^\star)$ and $\sum_{k=1}^{K_\star} \eta_k^\star U(\mu_{t-1}(x), \hat{w}_{\pi_t(k),t})$, then apply triangle inequality:

$$\begin{aligned} \left|\mathcal{U}^\star(x) - \widehat{\mathcal{U}}_t(x)\right| \le\ & \sum_{k=1}^{K_\star} \eta_k^\star \left|U(f(x), w_k^\star) - U(\mu_{t-1}(x), w_k^\star)\right| \\ &+ \sum_{k=1}^{K_\star} \eta_k^\star \left|U(\mu_{t-1}(x), w_k^\star) - U(\mu_{t-1}(x), \hat{w}_{\pi_t(k),t})\right| \\ &+ \left| \sum_{k=1}^{K} (\eta_k^\star - \hat{\eta}_{k,t}) \times U(\mu_{t-1}(x), \hat{w}_{k,t}) \right|. \end{aligned} \tag{31}$$

Bounding terms using Lemma 2 For the first line, apply Lipschitzness in y (20):

$$\left|U(f(x), w_k^\star) - U(\mu_{t-1}(x), w_k^\star)\right| \le \frac{1}{c_w} \|f(x) - \mu_{t-1}(x)\|_\infty.$$

Summing over k with weights $\eta_k^\star$ gives

$$\sum_{k=1}^{K_\star} \eta_k^\star \left|U(f(x), w_k^\star) - U(\mu_{t-1}(x), w_k^\star)\right| \le \frac{1}{c_w} \|f(x) - \mu_{t-1}(x)\|_\infty. \tag{32}$$

For the second line, apply Lipschitzness in w (21) with $\|\mu_{t-1}(x)\|_\infty \le B_y$:

$$\left|U(\mu_{t-1}(x), w_k^\star) - U(\mu_{t-1}(x), \hat{w}_{\pi_t(k),t})\right| \le \frac{B_y}{c_w^2} \|\hat{w}_{\pi_t(k),t} - w_k^\star\|_1.$$

Thus the second line is bounded by $(B_y/c_w^2)\Delta_t^w$.

For the third line, use $|U(\mu_{t-1}(x), \hat{w}_{k,t})| \le$

$\|\mu_{t-1}(x)\|_\infty / c_w \le B_y / c_w$ to obtain

$$\left| \sum_{k=1}^{K} (\eta_k^\star - \hat{\eta}_{k,t})\, U(\mu_{t-1}(x), \hat{w}_{k,t}) \right| \le \frac{B_y}{c_w} \|\hat{\eta}_t - \eta^\star\|_1 = \frac{B_y}{c_w} \Delta_t^\eta. \tag{33}$$

Combining (31), (32) and (33), we conclude that for all x,

$$\left|\mathcal{U}^\star(x) - \widehat{\mathcal{U}}_t(x)\right| \le \frac{1}{c_w} \|f(x) - \mu_{t-1}(x)\|_\infty + \frac{B_y}{c_w^2} \Delta_t^w + \frac{B_y}{c_w} \Delta_t^\eta. \tag{34}$$

Convert to simple regret Apply (34) at $x = x^\star$ and $x = x_t$ in (30):

$$\mathcal{U}^\star(x^\star) - \mathcal{U}^\star(x_t) \le \frac{1}{c_w} \sum_{z \in \{x^\star, x_t\}} \|f(z) - \mu_{t-1}(z)\|_\infty + \frac{2B_y}{c_w^2} \Delta_t^w + \frac{2B_y}{c_w} \Delta_t^\eta + \varepsilon_t. \tag{35}$$

Now take the minimum over $t \le T$ on the right-hand side. Since $R_T = \min_{t \le T} \left(\mathcal{U}^\star(x^\star) - \mathcal{U}^\star(x_t)\right)$ and $\max_{t \le T} \varepsilon_t$ upper bounds ε_t, we obtain (27). On the GP confidence event (24), the bound

$$\|f(x) - \mu_{t-1}(x)\|_\infty \le \sqrt{\beta_t}\, \|\sigma_{t-1}(x)\|_\infty$$

holds for all $x \in \mathcal{X}$, including $x^\star$.

Existence-type GP rate On the confidence event (24), $\|f(x_t) - \mu_{t-1}(x_t)\|_\infty \le \sqrt{\beta_t}\, \|\sigma_{t-1}(x_t)\|_\infty \le \sqrt{\beta_T}\, \|\sigma_{t-1}(x_t)\|_\infty$. Under standard variance-sum bounds for GP regression (e.g., Lemma X),

$$\sum_{t=1}^{T} \|\sigma_{t-1}(x_t)\|_\infty^2 \le C\, \gamma_T,$$

which implies

$$\min_{t \le T} \|\sigma_{t-1}(x_t)\|_\infty \le \frac{1}{T} \sum_{t=1}^{T} \|\sigma_{t-1}(x_t)\|_\infty \le \sqrt{\frac{C\, \gamma_T}{T}}.$$

□

Limitations and Future Work

The proposed method evaluates the proposed framework in simulated settings and on a real-world process design dataset, but does not include human-subject experiments. Due to time and resource constraints, validation with real decision makers was beyond the scope of the current study and remains an important direction for future work. We also plan to test the method on a broader range of real-world multi-objective datasets to assess generalization across domains. Finally, extending the framework to more adaptive mixture structures instead of stationary latent archetypes could further improve flexibility and practical applicability.