Heteroscedastic Preferential Bayesian Optimization with Informative Noise Distributions

Marshal A. Sinaga
Aalto University
marshal.sinaga@aalto.fi

Julien Martinelli
Université de Bordeaux
julien.martinelli@u_bordeaux.fr

Vikas Garg
Aalto University; YaiYai Ltd
vikas.garg@aalto.fi; vikas@yaiyai.fi

Samuel Kaski
Aalto University, University of Manchester
samuel.kaski@aalto.fi

Abstract

Preferential Bayesian optimization (PBO) is a sample-efficient framework for learning human preferences between candidate designs. PBO classically relies on homoscedastic noise models to represent human *aleatoric* uncertainty. Yet, such noise fails to accurately capture the varying levels of human *aleatoric* uncertainty, particularly when the user possesses partial knowledge among different pairs of candidates. For instance, a chemist with solid expertise in glucose-related molecules may easily compare two compounds from that family while struggling to compare alcohol-related molecules. Currently, PBO overlooks this uncertainty during the search for a new candidate through the maximization of the acquisition function, consequently underestimating the risk associated with human uncertainty. To address this issue, we propose a heteroscedastic noise model to capture human *aleatoric* uncertainty. This model adaptively assigns noise levels based on the distance of a specific input to a predefined set of reliable inputs known as *anchors* provided by the human. *Anchors* encapsulate partial knowledge and offer insight into the comparative difficulty of evaluating different candidate pairs. Such a model can be seamlessly integrated into the acquisition function, thus leading to candidate design pairs that elegantly trade informativeness and ease of comparison for the human expert. We perform an extensive empirical evaluation of the proposed approach, demonstrating a consistent improvement over homoscedastic PBO.

1 Introduction

Preferential Bayesian Optimization (PBO, [16]) has become the gold standard method for optimizing black-box functions whose feedback is perceived through the outcome of a pairwise comparison, also referred to as a *duel* [9]. Such a setting naturally appears when interacting with human subjects in design tasks, as humans are better at comparing two options than assessing their value [24]. These tasks include visual design optimization [25], A/B tests, and recommender systems [8]. Importantly, human subject preferences can also be elicited through iterative comparisons and serve as a cheap proxy to evaluate expensive black-box objectives like molecular properties [40].

To learn the latent utility of a human subject, PBO relies on a statistical surrogate, typically a Gaussian process (GP, [7, 32]). The latter features principled uncertainty quantification, thus conveniently describing the *epistemic* uncertainty inherent to the finite sample regime, as well as the *aleatoric* uncertainty stemming from the fact that the observations are noisy. In the case of preferential feedback given by human subjects, a noisy observation is the result of a comparison involving at least one design whose outcome is uncertain, making the comparison cumbersome [37]. Here, we stress that a

key component to model a human's partial knowledge is the input-dependent *aleatoric* uncertainty, as the level of human's uncertainty is not uniform across the design space. Indeed, consider a scenario where a chemist proficient in glucose-related compounds can effortlessly contrast two substances but encounter challenges when dealing with alcohol-related molecules. In such a case, assigning a uniform level of uncertainty to both types of molecules is not reasonable.

This flaw results in the PBO acquisition functions being *risk-neutral*, only seeking to optimize the statistics of latent utility value [16]. Consider the varying *aleatoric* uncertainty results in situations where the latent utility yields two solutions that deliver similar expected function values, yet one may be noisier (Figure 1). Choosing the noisier solution increases the risk of inconsistent preferential feedback from human experts. Over time, this inconsistency can result in suboptimal outcomes for PBO. Consequently, we advocate for adopting a *risk-averse* paradigm, necessitating a balance between maximizing the expected latent utility value and minimizing *aleatoric* uncertainty [28]. For example, a chemist might opt for slightly inferior quality molecules, but this leads to more confident preferential feedback. While there exist risk-averse variants of standard BO [28], this problem has not been studied for PBO.

> **Contributions.** We introduce a novel Heteroscedastic Preferential Bayesian Optimization framework, employing a kernel density estimator-based noise model to accommodate the varying levels of human *aleatoric* uncertainty. This estimator is based on a set of reliable inputs called *anchors*, representing regions of reliable human judgment. Leveraging classical results on Kernel Density Estimation, we conduct a thorough theoretical analysis of the proposed estimator and show that the risk and concentration can be bounded under standard smoothness assumptions over the latent human uncertainty. Finally, we provide empirical evidence on synthetic examples of the superiority of our proposed method over homoscedastic PBO approaches, jointly with several ablation studies.

2 Background

2.1 Preferential Bayesian optimization

The objective of PBO is to determine the optimal solution $\mathbf{x}^* = \underset{\mathbf{x} \in \mathcal{X}}{\operatorname{argmax}} f(\mathbf{x})$ using the duel feedback $\mathbf{x} \succ \mathbf{x}'$, signifying a preference $\mathbf{x}$ over $\mathbf{x}'$. The preference structure is governed by the latent function $f : \mathcal{X} \to \mathbb{R}$, defined over a subset $\mathcal{X} \subset \mathbb{R}^d$. As detailed in [9], the duel outcome is determined by

$$\mathbf{x} \succ \mathbf{x}' \iff f(\mathbf{x}) + \varepsilon(\mathbf{x}) > f(\mathbf{x}') + \varepsilon(\mathbf{x}'), \tag{1}$$

where $\varepsilon(\mathbf{x})$, $\varepsilon(\mathbf{x}')$ denote zero-mean additive noise capturing the human *aleatoric* uncertainty. PBO employs the preferential GP as a surrogate of the latent function in a probabilistic framework [41].

Most works in PBO commonly assume the homoscedastic setting, where the noise $\varepsilon(\mathbf{x}')$ is drawn from a nondegenerate normal distribution $\mathcal{N}(0, \sigma_\varepsilon^2(\mathbf{x}') = \sigma_{\text{noise}}^2)$, for all $\mathbf{x}' \in \mathcal{X}$. While this assumption offers computational efficiency, it is unrealistic since humans may exhibit partial knowledge of the inputs, resulting in varying *aleatoric* uncertainty.

> **Two types of uncertainties.** Our work crucially distinguishes between *epistemic* and *aleatoric* uncertainty. The former is also known as model uncertainty and accounts for the uncertainty around the latent function, as an infinite number of candidate models can accurately fit a finite number of observations. The latter is due to the presence of measurement noise in the data.

We argue the noise process ε should be input-dependent, capturing the user's varying level of uncertainty across the inputs. Next, we recall how GPs work in the context of heteroscedastic noise.

Prior. We begin by placing a zero-mean Gaussian process prior to the latent function, $f(\mathbf{x}) \sim \mathcal{GP}(0, l(\mathbf{x}, \mathbf{x}'))$, with a stationary kernel $l : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$. Therefore, the GP prior density function over a finite set of output values $\mathbf{X}$ is a multivariate Gaussian

$$p(\mathbf{f}) = (2\pi)^{-d/2} \det(\mathbf{L})^{-1/2} \exp\left(-\frac{1}{2}\mathbf{f}^\top \mathbf{L}^{-1}\mathbf{f}\right) \tag{2}$$

where $\mathbf{f} \triangleq \{f(\mathbf{x})\}_{\mathbf{x} \in \mathbf{X}}$ and $\mathbf{L} \triangleq (l(\mathbf{x}, \mathbf{x}'))_{\mathbf{x}, \mathbf{x}' \in \mathbf{X}}$ is the kernel matrix.

Likelihood. Assume we have access to human observations of m duel feedbacks $\mathcal{D} \triangleq \{\mathbf{x}_k \succ \mathbf{x}'_k\}_{k=1}^m$, with $\mathbf{x}_k$ and $\mathbf{x}'_k$ being the winner and the loser of the duel, respectively. Denoting $v \triangleq f(\mathbf{x}') + \varepsilon(\mathbf{x}') - f(\mathbf{x}) - \varepsilon(\mathbf{x})$, we have $\mathbf{v}_m < 0 \triangleq \{v_k < 0\}_{k=1}^m$. Let $\mathbf{X} \triangleq [\mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{x}'_1, \ldots, \mathbf{x}'_m]^\top \in \mathbb{R}^{2m \times d}$ be the concatenation of winners and losers from duels $\mathbf{v}_m$. The likelihood function $p(\mathbf{v}_m < 0 | \mathbf{f})$ adheres to a Gaussian cumulative density function (CDF) [9] defined as follows:

$$p(\mathbf{v}_m < 0 | \mathbf{f}) = \prod_{k=1}^m p(\mathbf{x}_k \succ \mathbf{x}'_k | f(\mathbf{x}_k), f(\mathbf{x}'_k), \varepsilon(\mathbf{x}_k), \varepsilon(\mathbf{x}'_k)) = \prod_{k=1}^m \Phi \left(\frac{f(\mathbf{x}_k) - f(\mathbf{x}'_k)}{\sqrt{\sigma_\varepsilon^2(\mathbf{x}_k) + \sigma_\varepsilon^2(\mathbf{x}'_k)}} \right). \quad (3)$$

It is worth noticing that the likelihood addresses the entanglement of noises by incorporating both noise distribution variances as the normalizing factor for $f(\mathbf{x}_k) - f(\mathbf{x}'_k)$. In the homoscedastic setting, the noise variance $\sigma_\varepsilon^2(\mathbf{x}_k)$ and $\sigma_\varepsilon^2(\mathbf{x}'_k)$ is a constant σ_{noise}^2.

Posterior. Given the GP prior density function $p(\mathbf{f})$ and likelihood $p(\mathbf{v}_m < 0 | \mathbf{f})$, the posterior density function of the latent function $\mathbf{f}$ conditioned on $\mathbf{v}_m < 0$ follows the Bayes's theorem:

$$p(\mathbf{f} | \mathbf{v}_m < 0) = \frac{p(\mathbf{v}_m < 0 | \mathbf{f}) p(\mathbf{f})}{p(\mathbf{v}_m < 0)} = \frac{p(\mathbf{v}_m < 0 | \mathbf{f}) p(\mathbf{f})}{\int p(\mathbf{v}_m < 0 | \mathbf{f}) p(\mathbf{f}) d\mathbf{f}}. \quad (4)$$

It is established that the exact posterior of the preference GP conforms to a skew GP [5].

Acquisition Function. Once the posterior $p(\mathbf{f} | \mathbf{v}_m < 0)$ has been obtained, BO employs an acquisition function (AF) α to determine the next query to evaluate:

$$\mathbf{x}^* = \underset{\mathbf{x} \in \mathcal{X}}{\operatorname{argmax}} \, \alpha(\mathbf{x}) := \int u(\mathbf{x}, \mathbf{f}) \, p(\mathbf{f} | \mathbf{v}_m < 0) d\mathbf{f}, \quad (5)$$

where u is any function measuring the informativeness of $\mathbf{x}$. e.g., $u(\mathbf{x}, \mathbf{f}) = [f(\mathbf{x}) - \max_{f(\mathbf{x}_i) \in \mathbf{f}} f(\mathbf{x}_i)]_+$ leads to the Expected Improvement. Common AFs are typically *risk-neutral*, focusing only on the *epistemic* uncertainty by maximizing u against the posterior of the latent function. Conversely, our approach incorporates *risk-averse* AFs, adding the *aleatoric* uncertainty.

The other design $\mathbf{x}'$ of the pair is often taken as the winner of the previous duel or the current maximizer of the posterior mean. AFs optimizing for whole pairs have also been proposed [2].

2.2 Kernel density estimation

Let $\mathbf{W} \triangleq (\mathbf{w}_1, \ldots, \mathbf{w}_n)$ denote i.i.d. samples with probability density $p(\mathbf{w})$. Using $\mathbf{W}$, Kernel Density Estimation (KDE) estimates $p(\mathbf{w})$ through the following formula:

$$\hat{p}(\mathbf{w} | h, \mathbf{W}) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h^d} k \left(\frac{\|\mathbf{w} - \mathbf{w}_i\|}{h} \right). \quad (6)$$

Here, $k : \mathbb{R} \to \mathbb{R}$ denotes a kernel function satisfying $\int k(u) \, du = 1$. The bandwidth $h > 0$ is critical in estimating the density. A small bandwidth leads to a rough, spiky estimate, whereas a large bandwidth results in over-smoothing. A complete discussion on that matter can be found in [14]. Let us also mention input-specific bandwidth selection procedures such as [26]. If k only takes non-negative values and $\mathbf{W}$ is fixed, then $\hat{p}(\mathbf{w} | h, \mathbf{W})$ is a probability density. We consider a Gaussian kernel for k, which satisfies the latter assumption.

3 Preferential BO with heteroscedastic noise

To solve the heteroscedastic PBO problem introduced in Section 2.1, we explicitly model the *aleatoric* user uncertainty. Section 3.1, introduces a heteroscedastic noise distribution with input-dependent variance. That variance itself involves a distribution obtained by means of Kernel Density Estimation, constructed from a set of user-specified reliable inputs called *anchors* (Section 3.2). Next, Section 3.3 presents the Hallucination Believer, a method adapted from [41] and used to infer the proposed heteroscedastic preferential GP. Finally, the surrogate informs a risk-averse acquisition function described in Section 3.4. The whole process is summarized on a toy case in Figure 1.
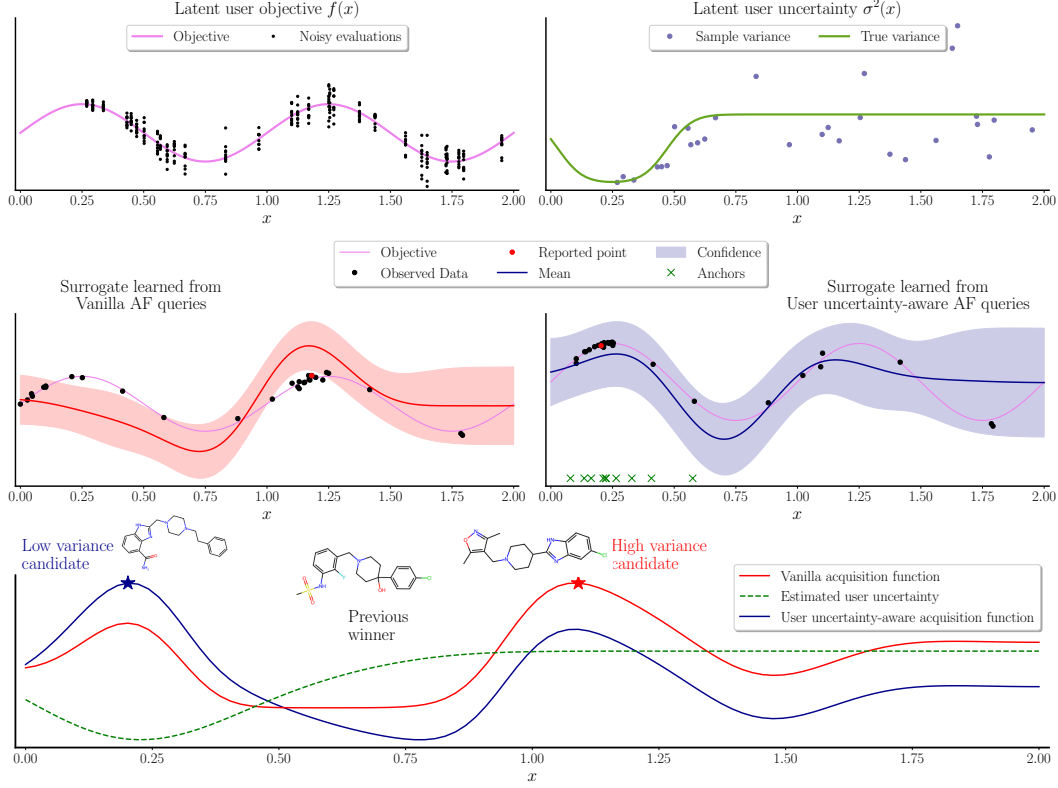
Figure 1: Heteroscedastic Preferential Bayesian Optimization. Top left: latent user utility with heteroscedastic noisy evaluations plotted as an example. Top right: ground truth latent user uncertainty used for this example. Middle: Preferential GP surrogates obtained using queries from vanilla AF (left) and user-uncertainty-aware AF (right). Our anchor-based model of user uncertainty leads to queries associated with lower noise and yet similarly high values, resulting in a better-calibrated surrogate (blue) compared to the vanilla GP (red, left). Bottom: AF landscape. The estimated user uncertainty (green) informs a user-uncertainty-aware acquisition function (blue), leading to a maximizer that differs from the vanilla AF (red) and accurately selects the low-variance design.

3.1 Heteroscedastic noise distribution

Recall that vanilla PBO assumes the noise to be drawn from a normal distribution $\mathcal{N}(0, \sigma^2_{\text{noise}})$, characterized by a constant variance σ^2_{noise}. We chose to model heteroscedastic noise as follows:

$$\varepsilon(\mathbf{x}) \sim \mathcal{N}(0, \sigma^2_{\varepsilon}(\mathbf{x}) \triangleq a \exp(-p(\mathbf{x}))). \tag{7}$$

Here, $a > 0$ is a scaling factor. Substituting the probability density $p(\mathbf{x})$ for the uniform density $p_{\text{uniform}}(\mathbf{x})$ leads to a homoscedastic noise distribution. Conversely, opting for a non-uniform probability density produces heteroscedastic noise distribution. Observe that the variance can be interpreted as a value estimation problem represented by an un-normalized *energy-based model* (EBM), with $p(\mathbf{x})$ specifying the energy function. Notably, the desirable outputs correspond to high energy function values, resulting in low energy scores [27]. In the context of heteroscedastic PBO, it is important to determine the appropriate information conveyed by $p(\mathbf{x})$.

3.2 *Anchors*-based input-dependent noise

In general comparison tasks (including preference), humans exhibit partial knowledge of different inputs [37, 19, 12]. This partial knowledge encompasses easily comparable inputs, which is associated with a low human *aleatoric* uncertainty [37]. For the purposes of this research, we refer to these easily comparable inputs as reliable inputs. We consider $p(\mathbf{x})$ to quantify the reliability of an input $\mathbf{x}$ according to a human expert. However, the precise formulation of $p(\mathbf{x})$ remains unclear.

4

To formalize our approach, we introduce *anchors*, a set of reliable inputs $\mathbf{X}_0 = \{\mathbf{x}_i\}_{i=1}^n$ provided by the human expert. Humans are known to excel in identifying these reliable inputs [29, 4, 34]. We assume the anchors have been collected in an i.i.d. manner, ensuring each $\mathbf{x}_i$ is treated equally. Comparing an input close to the anchors is relatively straightforward for a human expert. Therefore, the anchors encapsulate the region with high probability density $p(\mathbf{x})$. This implies that the noise drawn over these region exhibits low variance.

The anchors $\mathbf{X}_0$ allow us to construct a proxy of the probability density $p(\mathbf{x})$. For this purpose, we perform KDE as formalized in Equation 6 to obtain the density estimator $\hat{p}(\mathbf{x})$ for all $\mathbf{x}$,

$$\hat{p}(\mathbf{x}|h, \mathbf{X}_0) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h^d} \, k\left(\frac{\|\mathbf{x} - \mathbf{x}_i\|_2}{h} \right), \tag{8}$$

where d denotes input dimensionality. The estimator assigns a higher value when the proximity of $\mathbf{x}$ to the *anchors* set $\mathbf{X}_0$ increases. This proximity is represented by the Euclidian distance and a kernel k with a specified bandwidth h. Following Equation 7, we formulate the noise distribution $\varepsilon(\mathbf{x})$:

$$\varepsilon(\mathbf{x}) \sim \mathcal{N}(0, \hat{\sigma}_\varepsilon^2(\mathbf{x}) \triangleq a \exp(-\hat{p}(\mathbf{x}|h, \mathbf{X}_0))). \tag{9}$$

where $a > 0$ is a scaling factor. We assume that noise values across different inputs are independent, given the set of anchors. It is worth noticing that no further assumptions are required about the anchors. For instance, they do not have to be the inputs leading to a better optimum. Further details regarding the preferential likelihood under heteroscedastic noises, hyperparameter optimization, and time complexity of KDE can be found in Appendix A.

3.3 Hallucination believer for inference strategy

We adopt the hallucination believer (HB) [41] to perform inference in the heteroscedastic setting. Recall that the posterior distribution of the preferential GP is characterized as a skew GP, for which inference is notoriously challenging since it involves numerous evaluations of the Cumulative Distribution Function (CDF) of the multivariate normal distribution (MVN) [3]. Takeno *et al.* address this challenge by proposing the HB method, leveraging Gibbs sampling.

Recall the set of duels $\mathbf{v}_m < 0$ specified in Section 2.1. Takeno et al. [41] propose to generate a new sample $\tilde{\mathbf{v}}_m$, called the hallucination drawn from $p(\mathbf{v}_m|\mathbf{v}_m < 0)$ through Gibbs sampling. Note that the sample path from the skew GP denoted as $p(\mathbf{f}_*|\mathbf{v}_m < 0)$ for any output vector of the inputs $\mathbf{f}_* \triangleq [f(\mathbf{x}_1^*), \dots, f(\mathbf{x}_t^*)]^\top$ follows an MVN distribution, implying a regular GP. This characteristic enables a proper posterior computation of $p(\mathbf{f}_*|\mathbf{v}_m < 0, \tilde{\mathbf{v}}_m)$, outperforming other approximate inference techniques like Laplace approximation. Nevertheless, we emphasize that the proposed noise distribution can be applied to any approximate inference method. The algorithm and details of the HB method are outlined in Algorithm 1 and Appendix B, respectively.

Algorithm 1 Hallucination Believer ([41]) for posterior approximation and query acquisitions

---

1: **Input**: Initial dataset $\mathcal{D} = \{\mathbf{x}_k \succ \mathbf{x}_k'\}_{k=1}^m$
2: **for** $t = 1, \dots$ **do**
3:    $\mathbf{x}_t \leftarrow \mathbf{x}_{t-1}$ //set previous winner as first design of the pair
4:    Draw $\tilde{\mathbf{v}}_{t-1}$ from the posterior $p(\mathbf{v}_{t-1}|\mathbf{v}_{t-1} < \mathbf{0})$ via Gibbs sampling of truncated MVN
5:    Sequentially estimate f and $\sigma_\varepsilon^2(\mathbf{x})$ (Section 3.2)
6:    $\hat{\mathbf{x}}_t \leftarrow \arg\max_{\mathbf{x} \in \mathcal{X}} \alpha(\mathbf{x})$ based on GPs $f|\tilde{\mathbf{v}}_{t-1}$ and $\varepsilon(\mathbf{x}) \sim \mathcal{N}(0, \hat{\sigma}_\varepsilon^2(\mathbf{x}))$
7:    Set the winner as $\mathbf{x}_t$ and the loser as $\mathbf{x}_t'$, respectively.
8:    $\mathcal{D}_t \leftarrow \mathcal{D}_{t-1} \cup (\mathbf{x}_t \succ \mathbf{x}_t')$
9: **end for**

3.4 Risk-averse acquisition functions

In the decision-making theory, the aleatoric uncertainty can be interpreted as a form of risk [24]. In our case, we regard the noise variance as the embodiment of this risk as it is associated with the aleatoric uncertainty. It is noteworthy that variance is widely used for risk measurement [31, 22, 15]. To ensure robust querying of new candidates in the presence of heteroscedastic noise, we employ *risk-averse* acquisition functions that penalize high variance [28]. In essence, these acquisition

functions trade off the informativeness of the latent function and the aleatoric uncertainty stemming from the reliability of human judgment.

Specifically, we consider two known risk-averse acquisition functions α, both leveraging the variance of noise distribution $\sigma_\varepsilon^2(\mathbf{x})$ as the penalty criterion. The first one is the *aleatoric* noise-penalized expected improvement (ANPEI) [17] defined as

$$\alpha_{\text{ANPEI}}(\mathbf{x}) = \mathbb{E}[(f(\mathbf{x}^*) - f(\mathbf{x}))_+] - \gamma\sqrt{\sigma_\varepsilon^2(\mathbf{x})}, \tag{10}$$

with $f(\mathbf{x}^*)$ and $\gamma > 0$ denote the incumbent best value and a scaling constant, respectively. As the actual value $f(\mathbf{x}^*)$ is not observed in PBO, the incumbent is usually replaced by the maximum posterior mean value over points in the previous queries $\mu_f(\mathbf{x}^*)$ [36]. Subsequently, we consider the risk-averse upper confidence bound (RAHBO) [28] as

$$\alpha_{\text{RAHBO}}(\mathbf{x}) = \mu_f(\mathbf{x}) + \eta\sigma_f(\mathbf{x}) - \gamma\sigma_\varepsilon^2(\mathbf{x}). \tag{11}$$

The μ_f, σ_f denote the posterior mean and variance of the statistical surrogate accounting for the latent function f, respectively. Additionally, $\eta, \gamma > 0$ are the scaling factors that control the influence of the variance of the latent function and the noise distribution, respectively.

4 Theoretical analysis

Risk analysis. Suppose we can access the true density $p(\mathbf{x})$ for all $\mathbf{x}$. We are interested in assessing the consistency of the variance estimator in relation to the true variance. We approach this by analyzing the mean squared error (MSE) of the variance, denoted as $\mathbb{E}_{\mathbf{X_0}}[(\hat{\sigma}_\varepsilon^2(\mathbf{x}) - \sigma_\varepsilon^2(\mathbf{x}))^2]$. This analysis requires assumptions regarding the smoothness of the probability density function and the characteristic of the kernel under the integral operation, as described below.

Assumption 4.1 *p belongs to a class of densities $\mathcal{P} = \mathcal{P}(\beta, L)$ defined as follows:*

$$\mathcal{P}(\beta, L) \triangleq \left\{ p : p \geq 0, \int p(\mathbf{x})d\mathbf{x} = 1, \ p \in \Sigma(\beta, L) \text{ on } \mathbb{R}^d \right\}$$

with $\Sigma(\beta, L)$ denoting Hölder class. The kernel used in the KDE estimator is a kernel of order ℓ.

Theorem 4.2 *Fix $\alpha > 0$ and take $h = \alpha n^{-1/(2\beta+d)}$. Then, for any input $\mathbf{x}$ and the number of anchors $n \geq 1$, the estimated variance $\hat{\sigma}^2(\mathbf{x})$ satisfies*

$$\sup_{p \in \mathcal{P}(\beta, L)} \mathbb{E}_{\mathbf{X_0}}[(\hat{\sigma}_\varepsilon^2(\mathbf{x}) - \sigma_\varepsilon^2(\mathbf{x}))^2] \leq a^2 c_3 n^{-\frac{2\beta}{2\beta+d}} \tag{12}$$

where $c_3 > 0$ is a constant depending on Hölder class parameter β, constant α, scaling factor $a > 0$, and the kernel bandwidth h.

Of note, the Gaussian kernel satisfies the above assumption. The outline of the proof is as follows. We first bound the MSE of variance as a scaled MSE of density $\mathbb{E}_{\mathbf{X_0}}[(\hat{p}(\mathbf{x}|h, \mathbf{X_0}) - p(\mathbf{x}))^2]$, leveraging *Lipschitz-continuous* property. We then derive the bound for MSE of the density through variance-bias decomposition. The theorem tells that the convergence rate of MSE shrinks polynomially w.r.t. the number of anchors n. It implies that the estimator will be as accurate as the true variance with a sufficient number of anchors. However, the convergence rate is slow as we deal with higher data dimensionality d. Consequently, the accuracy of the variance estimator diminishes with higher dimensions. The detailed proof of the theorem can be found in Appendix C.1.

Concentration analysis. Another way to assess the consistency is understanding the probability of $\hat{\sigma}_\varepsilon^2(\mathbf{x})$ deviating from the true variance $\sigma_\varepsilon^2(\mathbf{x})$.

Theorem 4.3 *For any $\delta > 0, a \geq 1$, there exist constants c_1 and c_2, such that*

$$\sup_{p \in \mathcal{P}(\beta, L)} \mathbb{P}\left(|\sigma_\varepsilon^2(\mathbf{x}) - \hat{\sigma}_\varepsilon^2(\mathbf{x})| > a\left(\sqrt{4c_1 \log(2/\delta)/nh^d} + c_2 h^\beta \right) \right) < \delta \tag{13}$$

for any $\mathbf{x}$. Furthermore, under an additional assumption, the following bound also holds.

$$\sup_{p \in \mathcal{P}(\beta, L)} \mathbb{P}\left(\sup_{\mathbf{x} \in \mathcal{X}} |\sigma_\varepsilon^2(\mathbf{x}) - \hat{\sigma}_\varepsilon^2(\mathbf{x})| > a\left(\sqrt{1/(c_4 n h_n^d) \log(c_3/\delta)} + c_2 h^\beta \right) \right) < \delta \tag{14}$$

for constants c_3, c_4, and h_n depends on n.

The proof employs *Lipschitz continuous* property to bound $|\sigma_\varepsilon^2(\mathbf{x}) - \hat{\sigma}_\varepsilon^2(\mathbf{x})|$. Subsequently, we apply two types of concentration inequality by Bernstein [6] and Giné and Guillou [13]. The first inequality implies that $\hat{\sigma}_\varepsilon^2(\mathbf{x})$ close to $\sigma_\varepsilon^2(\mathbf{x})$ with a high probability as n increases. However, the probability shrinks as d grows. The second inequality provides a stronger bound since this result holds uniformly over all $\mathbf{x}$. Obtaining this inequality requires the kernel to belong to a bounded measurable VC class. In addition, it restricts the value of h as its behavior should depend on n. The proof of the theorem can be found in Appendix C.2.

5 Related work

Bayesian optimization with heteroscedastic noise. In most real-world settings, defining the likelihood based on a fixed variance noise leads to misspecified posteriors, which can harm the optimization by suggesting designs incorrectly believed to yield high function values. Several attempts have been made to tackle this issue at acquisition level [28], surrogate level [10, 17] or by adapting distribution-free uncertainty quantification techniques like conformal inference to BO [38].

Preferential Bayesian optimization. Most of the works in BO revolve around two modeling choices: statistical surrogate and acquisition module. In PBO, the posterior is usually untractable, and hence, a seminal work proposed to carry inference using Laplace approximation [7]. Other works have leveraged the fact that the surrogate can be shown to be a skew GP, for which a posterior sampling scheme can be derived [5]. Approximate inference techniques such as Expectation-Propagation and MCMC have also been shown to work [41]. Let us also mention the existence of a PBO method that employs neural networks as a surrogate rather than GPs [20]. On the acquisition side, some work adapted classical acquisition functions to the PBO case by selecting one of the duel elements as the winner of the previous duel and optimizing for the second element using Expected Improvement or Thompson Sampling [36]. Next, acquisition functions leveraging the preferential nature of the queries were also proposed [16, 2]. Recent work has also investigated preferential relations and queries to encode and elicit user prior beliefs [23, 1]. To the best of our knowledge, employing heteroscedastic noise models to account for human uncertainty has not yet been explored in PBO.

Kernel density estimation in Bayesian optimization. *Anchors*-based modeling was previously introduced by Eduardo and Gutmann [11]. However, in this case, *anchors* referred to promising candidate solutions in the case of vanilla BO and were unrelated to human uncertainty in PBO like us. KDE was also employed to learn the distribution of contextual variables in BO [21].

6 Experiments

We thoroughly assess the performance of our approach on several synthetic black-box functions (Section 6.1). Next, we assess the performance of our method under various experiment and method settings (Section 6.2). These include for instance different distribution for the expert uncertainty, as well as different posterior inference techniques and hyperparameter optimization criteria.

Baselines. Our proposed heteroscedastic PBO approach yields a surrogate explicitly modeling human aleatoric uncertainty. As such, this surrogate enhances vanilla acquisition functions, making them *risk-averse*. Thus, our method leads to two baselines: RAHBO and ANPEI, described in Equations (10) and (11), respectively. Concurrent baselines are risk-neutral versions of these AFs: UCB and EI. In all cases, the second design of the queried duel is obtained as the winner of the previous duel.

Implementation details. Unless stated otherwise, posterior inference for the statistical surrogate is carried out using the HB method presented in Section 3.3 and appendix B. The hyperparameters of the surrogate are obtained using marginal likelihood maximization. Regarding our KDE-based model of human aleatoric uncertainty, the estimation procedure for the bandwidth h is detailed in Appendix A.2. All baselines use the same initial duels and anchors when applicable for the latter. The PBO trials are performed 30 times with different random seeds. All experiments were run on a private cluster consisting of a mixture of Intel®Xeon ®Gold 6248, Xeon ®Gold 6148, Xeon ®E5-2690 v3, and Xeon ®E5-2680 v3 processors. The longest experiments (Hartmann4) took a total of 24 hours.

Performance metrics. Several metrics are considered to assess each baseline's performance. We follow Makarova et al. [28] and define the mean-variance objective $\mathrm{MV}(\mathbf{x}_t) = f(\mathbf{x}_t) - \rho\sigma_\varepsilon^2(\mathbf{x}_t)$, with $\mathbf{x}_t$ the input acquired at round t. Maximizing MV leads to queries that realize a tradeoff between high
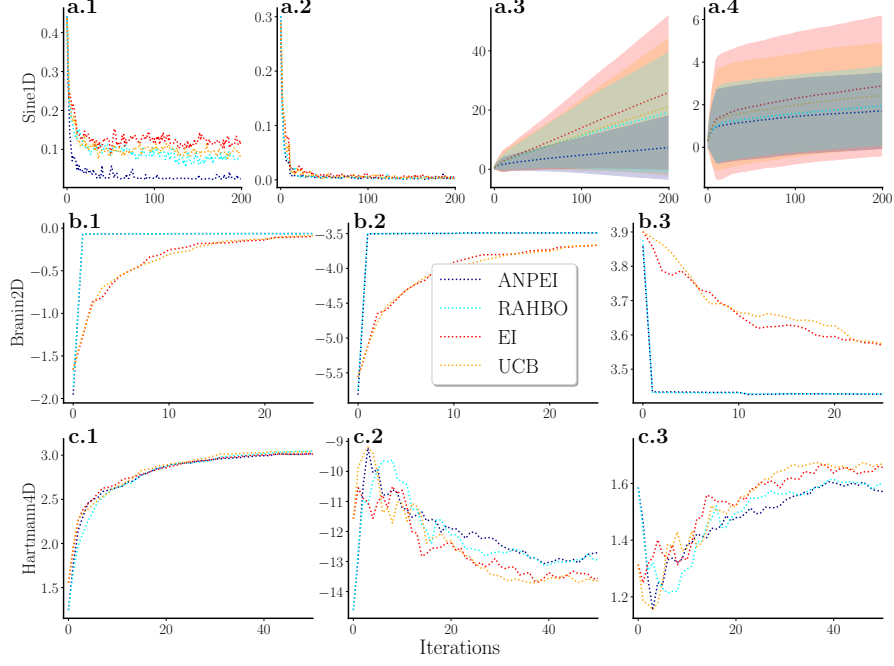
Figure 2: Results for three synthetic problems: Sine1D **(a)**, Branin2D **(b)** and Hartmann4D **(c)**. **(a.1)** Simple regret. **(a.2)** Risk-averse simple regret. **(a.3)** Cumulative regret. **(a.4)** Risk-averse cumulative regret. **(b.1)** Best value found. **(b.2)** Risk-averse best value found **(b.3)** Inferred noise variance $\hat{\sigma}_\varepsilon(\mathbf{x})$. **(c.1)** Best value found. **(c.2)** Risk-averse best value found **(c.3)** Inferred noise variance $\hat{\sigma}_\varepsilon(\mathbf{x})$. For all examples, $\rho = 3f(\mathbf{x}_{\max})$. Mean and standard deviations were computed across 30 random seeds. Risk-averse AFs outperform vanilla AFs, specifically for cumulative regret metrics.

values and low variance, depending on how large ρ is. One should notice that this objective involves the true variance σ_ε^2, which is only accessible in synthetic examples. Our objective being defined, we report the risk-averse *simple* regret $\mathrm{MV}(\mathbf{x}_{\max}) - \mathrm{MV}(\mathbf{x}_t)$ and *cumulative* regret $\sum_{t=1}^{T}(\mathrm{MV}(\mathbf{x}_{\max}) - \mathrm{MV}(\mathbf{x}_t))$, with T the total number of BO rounds, and $\mathbf{x}_{\max} = \mathrm{argmax}_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x})$. Unless stated otherwise in dedicated experiments, we set $\rho = 3f(\mathbf{x}_{\max})$. Lastly, we also report the standard simple and cumulative regrets and the best value found, all of which corresponds to $\rho = 0$.

6.1 Synthetic experiments

Sine function. We consider the sine function example presented in Figure 1. The input domain $\mathcal{X} = [0, 2]$ contains two similar optima $\mathbf{x}_{\max} \in \{0.25, 1.25\}$. To simulate human-like aleatoric uncertainty, we construct an oracle by placing a normal distribution $\mathcal{N}(\mathbf{x}_{\max} = 0.25, 0.125)$. We then utilize the oracle to specify the noise distribution and draw 30 anchors from it. The scaling factor for the noise variance estimator is set to $a = 0.1$. Figure 2a presents the results. For $\rho = 0$, panel **(a.2)** shows that all baselines quickly reach the best value in a similar fashion. However, when requiring the optimized queries to display a low variance, incentivized by setting $\rho = 3f(\mathbf{x}_{\max})$, panel **(a.1)** shows that PBO-RAHBO and PBO-ANPEI exhibit lower risk-averse simple regret. The cumulative regret plots provided in panel **(a.3-4)** further demonstrate the superiority of risk-averse AFs, notably even when $\rho = 0$. Additional visualizations provided in Figures S1 and S2 illustrate how risk-averse AFs indeed lead to low-variance queries, contrarily to their risk-neutral counterparts.

Branin and Hartmann4. Next, we look at two classical black-box functions: Branin and Hartmann4. We simulate human aleatoric uncertainty by placing a normal distribution centered at points near and far from the optimum for the Branin and Hartmann4 functions, respectively. The scaling factor is set to $a = 1.0$ and $a = 2.0$ for Branin and Hartmann4, respectively. Panels **(b)** and **(c)** from Figure 2 report the mean-variance objective MV for $\rho = 3f(\mathbf{x}_{\max})$ (**b.1** and **c.1**) and $\rho = 0$ (**b.2** and **c.2**). Lastly, **(b.3)** and **(c.3)** present the noise variance $\sigma_\varepsilon^2(\mathbf{x})$. Again, heteroscedastic PBO AFs prefer the query with lower variance, thus achieving higher risk-averse best value $\mathrm{MV}(\mathbf{x}_t)$. Despite the lower variances, the queries obtained by heteroscedastic PBOs yield competitive best values found.

6.2 Ablation study

Posterior inference techniques. We evaluate the impact of different approximate inference methods on the outcomes of PBO. Specifically, we perform heteroscedastic PBO on the Hartmann4 function using both the Laplace approximation and expectation propagation. All other experimental settings remain as described above. Our heteroscedastic PBO consistently outperforms vanilla PBO in terms of mean-variance objective across inference techniques, without compromising the best value found.
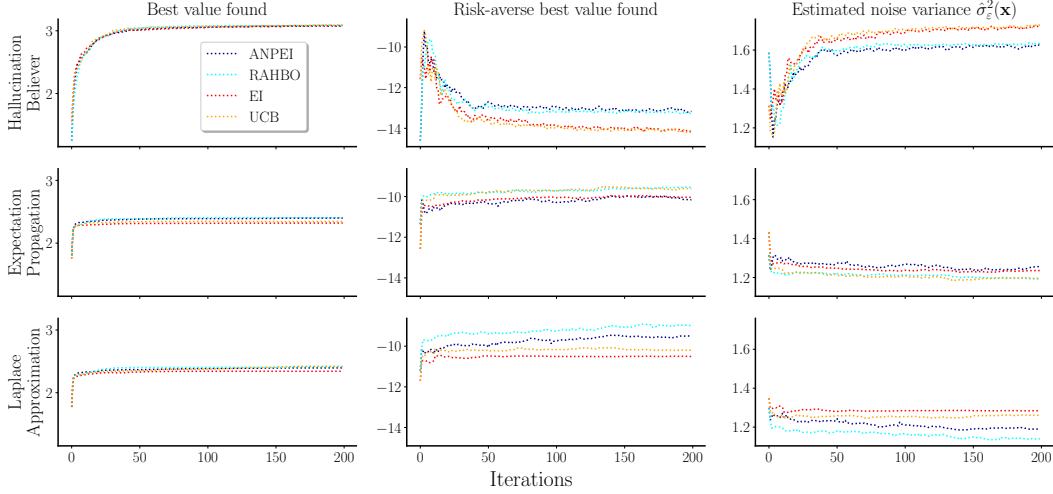

Figure 3: Results for the Hartmann4D test function using different approximated inference techniques.

Bandwidth optimization method. Instead of optimizing the bandwidth h of the estimator kernel k using the marginal likelihood, we maximize the Leave-One-Out (LOO) criterion. Results indicate this approach does not impact simple and cumulative regret (Figure S3). However, the risk-averse simple and cumulative regret of PBO-RAHBO degrade compared to the main results. A similar study was carried for the Hartmann4 function in Appendix D.3 and yielded similar results.

Noise distributions. Lastly, we study the performance of heteroscedastic PBO under a non-Gaussian user uncertainty distribution $p(\mathbf{x})$. Specifically, we set $p(\mathbf{x})$ as a student-t density function with $\nu = 5$ degrees of freedom. Our method remains consistent in terms of best value found (Figure S5). However, the objective $\mathrm{MV}(\mathbf{x}_t)$ PBO-ANPEI degrade compared to the main results. Further details can be found in Appendix D.4.

7 Conclusions

In this paper, we considered the problem of learning human preferences through dueling feedback between candidate pairs, a task that falls into the realm of Preferential Bayesian Optimization. Through several examples, we illustrated the necessity to account for the different levels of human uncertainty, as experts typically exhibit varying levels of knowledge over the input domain. To that end, we enhanced preferential GPs with an input-dependent noise model, built using a set of inputs over which the user has low uncertainty, so-called *anchors*. This led to faster convergence towards low-variance optima in BO trials, as we demonstrated in various synthetic examples.

Limitations and future work. By bounding the MSE between the true noise variance $\sigma_\varepsilon^2(\mathbf{x})$ and its estimator $\hat{\sigma}_\varepsilon^2(\mathbf{x})$ defined in Equation (9), Theorem 4.2 assesses how accurate the estimated user uncertainty is. $\sigma_\varepsilon^2(\mathbf{x})$ and $\hat{\sigma}_\varepsilon^2(\mathbf{x})$ respectively lead to the true likelihood from Equation (3) and a "misspecified" likelihood, as the latter contains an estimate of the noise variance instead of the true one. Therefore, a promising research direction is to leverage Theorem 4.2 to derive the convergence rate of the misspecified likelihood to the true one. Additionally, releasing the i.i.d. assumption on the anchors set would allow their collection in a sequential manner.

References

[1] Masaki Adachi, Brady Planden, David A. Howey, Krikamol Muandet, Michael A. Osborne, and Siu Lun Chau. Looping in the human: Collaborative and explainable bayesian optimization, 2023.

[2] Raul Astudillo, Zhiyuan Jerry Lin, Eytan Bakshy, and Peter Frazier. qeubo: A decision-theoretic acquisition function for preferential bayesian optimization. In Francisco Ruiz, Jennifer Dy, and Jan-Willem van de Meent, editors, *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pages 1093–1114. PMLR, 25–27 Apr 2023. URL https://proceedings.mlr.press/v206/astudillo23a.html.

[3] Adelchi Azzalini. *The skew-normal and related families*, volume 3. Cambridge University Press, 2013.

[4] Lawrence W Barsalou. Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of experimental psychology: learning, memory, and cognition*, 11(4):629, 1985.

[5] Alessio Benavoli, Dario Azzimonti, and Dario Piga. Preferential bayesian optimisation with skew gaussian processes. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, GECCO '21, page 1842–1850, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383516. doi: 10.1145/3449726.3463128. URL https://doi.org/10.1145/3449726.3463128.

[6] Sergei Bernstein. On a modification of chebyshev's inequality and of the error formula of laplace. *Ann. Sci. Inst. Sav. Ukraine, Sect. Math*, 1(4):38–49, 1924.

[7] Eric Brochu, Vlad M Cora, and Nando De Freitas. A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint arXiv:1012.2599*, 2010.

[8] Peter Brusilovsky, Alfred Kobsa, and Wolfgang Nejdl. *The Adaptive Web - Methods and Strategies of Web Personalization*. 01 2007. ISBN 978-3-540-72078-2. doi: 10.1007/978-3-540-72079-9.

[9] Wei Chu and Zoubin Ghahramani. Preference learning with gaussian processes. In *Proceedings of the 22nd international conference on Machine learning*, pages 137–144, 2005.

[10] Alexander Cowen-Rivers, Wenlong Lyu, Rasul Tutunov, Zhi Wang, Antoine Grosnit, Ryan-Rhys Griffiths, Alexandre Maravel, Jianye Hao, Jun Wang, Jan Peters, and Haitham Bou Ammar. Hebo: Pushing the limits of sample-efficient hyperparameter optimisation. *Journal of Artificial Intelligence Research*, 74, 07 2022.

[11] Afonso Eduardo and Michael U. Gutmann. Bayesian optimization with informative covariance, 2023.

[12] Paul J Feltovic, Kenneth M Ford, and Robert R Hoffman. Expertise in context: Human and machine. *AI Magazine*, 22(3):62, 2001.

[13] Evarist Giné and Armelle Guillou. Rates of strong uniform consistency for multivariate kernel density estimators. In *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, volume 38, pages 907–921. Elsevier, 2002.

[14] Alexander Goldenshluger and Oleg Lepski. Bandwidth selection in kernel density estimation: Oracle inequalities and adaptive minimax optimality. *The Annals of Statistics*, 39(3):1608 – 1632, 2011. doi: 10.1214/11-AOS883. URL https://doi.org/10.1214/11-AOS883.

[15] Bennett W Golub and Leo M Tilman. *Risk management: approaches for fixed income markets*, volume 73. John Wiley & Sons, 2000.

[16] Javier González, Zhenwen Dai, Andreas Damianou, and Neil D Lawrence. Preferential bayesian optimization. In *International Conference on Machine Learning*, pages 1282–1291. PMLR, 2017.

[17] Ryan-Rhys Griffiths, Alexander A Aldrick, Miguel Garcia-Ortegon, Vidhi Lalchand, et al. Achieving robustness to aleatoric uncertainty with heteroscedastic bayesian optimisation. *Machine Learning: Science and Technology*, 3(1):015004, 2021.

[18] Trevor Hastie, Robert Tibshirani, Jerome H Friedman, and Jerome H Friedman. *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer, 2009.

[19] Robert R Hoffman. *The psychology of expertise: Cognitive research and empirical AI*. Psychology Press, 2014.

[20] Daolang Huang, Louis Filstroff, Petrus Mikkola, Runkai Zheng, Milica Todorovic, and Samuel Kaski. Augmenting bayesian optimization with preference-based expert feedback. In *ICML 2023 Workshop The Many Facets of Preference-Based Learning*, 2023. URL https://openreview.net/forum?id=OJXfwj2hiP.

[21] Xiaobin Huang, Lei Song, Ke Xue, and Chao Qian. Stochastic bayesian optimization with unknown continuous context distribution via kernel density estimation, 2023.

[22] John Hull. *Risk management and financial institutions,+ Web Site*, volume 733. John Wiley & Sons, 2012.

[23] Carl Hvarfner, Frank Hutter, and Luigi Nardi. A general framework for user-guided bayesian optimization, 2023.

[24] Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I*, pages 99–127. World Scientific, 2013.

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

[26] Samory Kpotufe and Vikas Garg. Adaptivity to local smoothness and dimension in kernel regression. In C.J. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 26. Curran Associates, Inc., 2013. URL https://proceedings.neurips.cc/paper_files/paper/2013/file/28fc2782ea7ef51c1104ccf7b9bea13d-Paper.pdf.

[27] Yann LeCun, Sumit Chopra, Raia Hadsell, M Ranzato, and Fujie Huang. A tutorial on energy-based learning. *Predicting structured data*, 1(0), 2006.

[28] Anastasia Makarova, Ilnura Usmanova, Ilija Bogunovic, and Andreas Krause. Risk-averse heteroscedastic bayesian optimization. *Advances in Neural Information Processing Systems*, 34: 17235–17245, 2021.

[29] Douglas L Medin and Andrew Ortony. Psychological essentialism. *Similarity and analogical reasoning*, 179:195, 1989.

[30] Jorge Nocedal. Updating quasi-newton matrices with limited storage. *Mathematics of computation*, 35(151):773–782, 1980.

[31] Svetlozar T Rachev, Stefan Mittnik, Frank J Fabozzi, Sergio M Focardi, et al. *Financial econometrics: from basics to advanced modeling techniques*. John Wiley & Sons, 2007.

[32] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian processes for machine learning*. Adaptive computation and machine learning. MIT Press, 2006. ISBN 026218253X.

[33] Vikas C Raykar, Ramani Duraiswami, and Linda H Zhao. Fast computation of kernel estimators. *Journal of Computational and Graphical Statistics*, 19(1):205–220, 2010.

[34] Brian H Ross. Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, memory, and cognition*, 15(3):456, 1989.

[35] Bodhisattva Sen. *Introduction to Nonparametric Statistics*. 2020.

[36] Eero Siivola, Akash Kumar Dhaka, Michael Riis Andersen, Javier González, Pablo García Moreno, and Aki Vehtari. Preferential batch bayesian optimization. In *2021 IEEE 31st International Workshop on Machine Learning for Signal Processing (MLSP)*, pages 1–6, 2021. doi: 10.1109/MLSP52302.2021.9596494.

[37] Paul Slovic and Amos Tversky. *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press, 1982.

[38] Samuel Stanton, Wesley Maddox, and Andrew Gordon Wilson. Bayesian optimization with conformal prediction sets. In Francisco Ruiz, Jennifer Dy, and Jan-Willem van de Meent, editors, *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pages 959–986. PMLR, 25–27 Apr 2023. URL https://proceedings.mlr.press/v206/stanton23a.html.

[39] Mervyn Stone. Cross-validatory choice and assessment of statistical predictions. *Journal of the royal statistical society: Series B (Methodological)*, 36(2):111–133, 1974.

[40] Iiris Sundin, Alexey Voronov, Haoping Xiao, Kostas Papadopoulos, Esben Jannik Bjerrum, Markus Heinonen, Atanas Patronov, Samuel Kaski, and Ola Engkvist. Human-in-the-loop assisted de novo molecular design. *Journal of Cheminformatics*, 14(1), 2022.

[41] Shion Takeno, Masahiro Nomura, and Masayuki Karasuyama. Towards practical preferential bayesian optimization with skew gaussian processes. In *International Conference on Machine Learning*, pages 33516–33533. PMLR, 2023.

Appendix — Heteroscedastic Preferential Bayesian Optimization with Informative Noise Distributions

Outline.

The appendix is organized as follows. In Appendix A, further details regarding our proposed *anchors*-based noise model are provided. In Appendix B, we describe the Hallucination Believer algorithm used for inference and inspired from [41]. Next, Appendix C contains the proofs of Theorems 4.2 and 4.3. Appendix D displays additional figures supporting the experiments carried in the main text. Finally, Appendix E provides the mathematical expression of the test functions employed throughout the paper.

A Details of *anchors*-based input-dependent noise

We consider the settings in Section 2.1 for this model.

A.1 Likelihood

We derive the likelihood $\Phi(\mathbf{z}_k)$ as follows:

$$p(\mathbf{x}_k \succ \mathbf{x}'_k | f(\mathbf{x}_k), f(\mathbf{x}'_k), \varepsilon(\mathbf{x}_k), \varepsilon(\mathbf{x}'_k)) = \Phi\left(\mathbf{z}_k \triangleq \frac{f(\mathbf{x}_k) - f(\mathbf{x}'_k)}{\sqrt{a \exp(-\hat{p}(\mathbf{x}_k | h, \mathbf{X}_0)) + a \exp(-\hat{p}(\mathbf{x}'_k | h, \mathbf{X}_0))}}\right) \tag{S1}$$

where $a \exp(-p(\mathbf{x}_k | h, \mathbf{X}_0))$ and $a \exp(-p(\mathbf{x}'_k | h, \mathbf{X}_0))$ denotes the variance of noise $\varepsilon(\mathbf{x}_k)$ and $\varepsilon(\mathbf{x}'_k)$, respectively.

A.2 Hyperparameter optimization

We minimize the negative log marginal likelihood approximated with Laplace approximation for preferential GP hyperparameter optimization. This involves obtaining $\mathbf{f}_{\mathrm{MAP}} = \arg\min_\mathbf{f} - \log p(\mathbf{f}|\mathbf{v}_m) \approx \arg\min_\mathbf{f} S(\mathbf{f})$ where we define $S(\mathbf{f})$ as

$$S(\mathbf{f}) = -\sum_{k=1}^{m} \log \Phi(\mathbf{z}_k) + \frac{1}{2} \mathbf{f}^\top \mathbf{L}^{-1} \mathbf{f} \tag{S2}$$

The first and second derivatives, respectively, are given by

$$\nabla_\mathbf{f} S(\mathbf{f}) = \begin{bmatrix} -\frac{1}{\sqrt{a \exp(-a\hat{p}(\mathbf{x}_{1:m}|h,\mathbf{X}_0)) + a \exp(-\hat{p}(\mathbf{x}'_{1:m}|h,\mathbf{X}_0))}} \cdot \frac{\phi(\mathbf{z}_{1:m})}{\Phi(\mathbf{z}_{1:m})} \\ \frac{1}{\sqrt{a \exp(-\hat{p}(\mathbf{x}_{1:m}|h,\mathbf{X}_0)) + a \exp(-\hat{p}(\mathbf{x}'_{1:m}|h,\mathbf{X}_0))}} \cdot \frac{\phi(\mathbf{z}_{1:m})}{\Phi(\mathbf{z}_{1:m})} \end{bmatrix} + \mathbf{L}^{-1}\mathbf{f} \quad \in \mathbb{R}^{2m} \tag{S3}$$

$$\nabla_\mathbf{f}^2 S(\mathbf{f}) = \Lambda + \mathbf{L}^{-1} \quad \in \mathbb{R}^{2m \times 2m} \tag{S4}$$

$$\Lambda = \begin{bmatrix} \mathbf{c}^\top \mathrm{diag}\left(\frac{\phi(\mathbf{z}_{1:m})^2}{\Phi(\mathbf{z}_{1:m})^2} + \frac{\phi(\mathbf{z}_{1:m})}{\Phi(\mathbf{z}_{1:m})}\mathbf{z}_{1:m}\right) & \mathbf{c}^\top \mathrm{diag}\left(-\frac{\phi(\mathbf{z}_{1:m})^2}{\Phi(\mathbf{z}_{1:m})^2} - \frac{\phi(\mathbf{z}_{1:m})}{\Phi(\mathbf{z}_{1:m})}\mathbf{z}_{1:m}\right) \\ \mathbf{c}^\top \mathrm{diag}\left(-\frac{\phi(\mathbf{z}_{1:m})^2}{\Phi(\mathbf{z}_{1:m})^2} - \frac{\phi(\mathbf{z}_{1:m})}{\Phi(\mathbf{z}_{1:m})}\mathbf{z}_{1:m}\right) & \mathbf{c}^\top \mathrm{diag}\left(\frac{\phi(\mathbf{z}_{1:m})^2}{\Phi(\mathbf{z}_{1:m})^2} + \frac{\phi(\mathbf{z}_{1:m})}{\Phi(\mathbf{z}_{1:m})}\mathbf{z}_{1:m}\right) \end{bmatrix} \quad \in \mathbb{R}^{2m \times 2m} \tag{S5}$$

$$\mathbf{c} = \begin{bmatrix} \frac{1}{a \exp(-p_X(\mathbf{x}_{1:m})) + a \exp(-p_X(\mathbf{x}'_{1:m}))} \end{bmatrix} \quad \in \mathbb{R}^m \tag{S6}$$

The solution of the gradient $\nabla_\mathbf{f} S(\mathbf{f}) = 0$ provides $\mathbf{f}_{\mathrm{MAP}}$. Subsequently, we introduce Λ_{MAP} which depends on the inverse Hessian given at $\mathbf{f}_{\mathrm{MAP}}$, i.e., $\Lambda_{\mathrm{MAP}} \triangleq \left(\nabla_\mathbf{f}^2 S(\mathbf{f})\right)^{-1} - \mathbf{L}^{-1}\big|_{\mathbf{f}_{\mathrm{MAP}}}$. We utilize the Newton-Raphson method to obtain $\mathbf{f}_{\mathrm{MAP}}$, following [9]. Given $\mathbf{f}_{\mathrm{MAP}}$ and Λ_{MAP}, we construct the evidence as

$$p(\mathcal{D}|\theta) \approx \exp(-S(\mathbf{f}_{\mathrm{MAP}})) \det(\mathbf{L}\Lambda_{\mathrm{MAP}})^{-1/2} \tag{S7}$$

We then formulate the hyperparameter optimization problem as

$$\theta^* = \arg \min_{\theta} - \log p(\mathcal{D}|\theta) \tag{S8}$$

with $\theta = \{\lambda\}$ denotes the hyperparameters. Specifically, λ denotes the length scale of preferential GP f. We utilize L-BFGS-B [30] to obtain the solution of Equation (S8).

Subsequently, we apply leave-one-out (LOO) [18] to optimize the KDE kernel bandwidth h. Note that LOO is a special case of cross-validation, where the k-fold is fixed to one. The rationale behind choosing LOO lies in the context of our study, where human expert input typically yields a small number of *anchors*. Consequently, utilizing k-fold to one aligns with the constraints imposed by our data. Given the *anchors* $\mathbf{X}_0$ and bandwidth h, we aim to minimize the negative LOO, formulated as follows:

$$h^* \triangleq \arg \min_{h} -\text{LOO}(h) = -\frac{1}{n} \sum_{\mathbf{x}_0 \in \mathbf{X}_0} \log \hat{p}(\mathbf{x}_0|h, \mathbf{X}_0 \setminus \{\mathbf{x}_0\}) \tag{S9}$$

For each anchor $\mathbf{x}_0 \in \mathbf{X}_0$, we employ the remaining *anchors* $\mathbf{X}_0 \setminus \{\mathbf{x}_0\}$ to estimate $p(\mathbf{x}_0)$. We then take the negative average of the logarithmic estimator for all *anchors*. We also employ L-BFGS-B to obtain h^*.

According to Stone's theorem, the Kernel Density Estimation (KDE) bandwidth $\hat{h}$ derived through cross-validation converges to the optimal h value [39]. This optimal bandwidth minimizes the Mean Squared Error (MSE) between the true density $p(\mathbf{x})$ and the density estimator $p(\mathbf{x}|h, \mathbf{X}_0)$ for all $\mathbf{x}$. Following this theorem, we obtain the following proposition.

Proposition A.1 *Suppose that $\sigma_\varepsilon^2(\mathbf{x})$ is bounded. Let $\hat{\sigma}_\varepsilon^2(\mathbf{x}|h)$ denote the kernel variance estimator with bandwidth h and let $\hat{\sigma}^2(\mathbf{x}|\hat{h})$ denote the bandwidth chosen by leave-one-out. Then,*

$$\frac{\int (\sigma_\varepsilon^2(\mathbf{x}) - \hat{\sigma}_\varepsilon^2(\mathbf{x}|\hat{h}))^2 d\mathbf{x}}{\inf_h \int (\sigma_\varepsilon^2(\mathbf{x}) - \hat{\sigma}_\varepsilon^2(\mathbf{x}|h))^2 d\mathbf{x}} \xrightarrow{a.s.} 1 \tag{S10}$$

Note that we slightly abuse the notation of the estimator variance to differentiate the bandwidth being utilized. The result follows the *Lipschitz continuous* property of the noise variance shown as in the proof of Theorem 4.2.

A.3 Time complexity of KDE estimator

For a set of anchors $\mathbf{X}_0$ with N samples and M evaluation inputs $\hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_m \in \mathcal{X}$, the time complexity of the density estimator $\hat{p}$ is $\mathcal{O}(NM)$ [33]. This arises from constructing kernel matrix sized $N \times M$ and the number of operations to obtain the density estimator for m evaluation inputs. Note that this estimator does not hurt the time complexity of our preferential-GP-based latent function, which can lead to $\mathcal{O}(N^3)$ in the classical scenario [32].

B Details of Hallucination Believer algorithm

B.1 Posterior predictive distribution

We formalize the predictive posterior as follows. We first model the joint distribution between the test function $\mathbf{f}_*$ and the sampled duels $\mathbf{v}_m$.

$$\begin{bmatrix} \mathbf{f}_* \\ \mathbf{v}_m \end{bmatrix} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}) \tag{S11}$$

$$\mathbf{\Sigma} \triangleq \mathbf{A}(\mathbf{L} + \mathbf{B})\mathbf{A}^\top \in \mathbb{R}^{(t+m) \times (t+m)} \tag{S12}$$

$$\mathbf{A} \triangleq \begin{bmatrix} \mathbf{I}_n & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_m & \mathbf{I}_m \end{bmatrix} \in \mathbb{R}^{(t+m) \times (t+2m)}$$

$$\mathbf{B} \triangleq \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{\text{noise}} \end{bmatrix} \in \mathbb{R}^{(t+2m) \times (t+2m)},$$

with $\mathbf{V}_{\mathrm{noise}} = \mathrm{Cov}[\varepsilon(\mathbf{x}), \varepsilon(\mathbf{x}')]_{\mathbf{x}, \mathbf{x}' \in \mathbf{X}}$, and $\mathbf{X}$ is defined as in Section 2.1. It is worth noticing that $\mathbf{V}_{\mathrm{noise}}$ is a diagonal matrix. We then rewrite the kernel matrix $\boldsymbol{\Sigma}$ into a block of matrices as follows:

$$\boldsymbol{\Sigma} \triangleq \begin{bmatrix} \boldsymbol{\Sigma}_{*,*} & \boldsymbol{\Sigma}_{*,\mathbf{v}} \\ \boldsymbol{\Sigma}_{\mathbf{v},*} & \boldsymbol{\Sigma}_{\mathbf{v},\mathbf{v}} \end{bmatrix}$$

where $\boldsymbol{\Sigma}_{*,*} \in \mathbb{R}^{t \times t}$, $\boldsymbol{\Sigma}_{*,\mathbf{v}} \in \mathbb{R}^{t \times m}$, $\boldsymbol{\Sigma}_{\mathbf{v},\mathbf{v}} \in \mathbb{R}^{m \times m}$. Thus, for an arbitrary output vector $\mathbf{f}_*$, we obtain

$$p(\mathbf{f}_* | \mathbf{v}_m < 0, \mathbf{v}_m) = p(\mathbf{f}_* | \mathbf{v}_m) \triangleq \mathcal{N}(\boldsymbol{\mu}_{*|\mathbf{v}}, \boldsymbol{\Sigma}_{*|\mathbf{v}}) \tag{S13}$$

$$\boldsymbol{\mu}_{*|\mathbf{v}} \triangleq \boldsymbol{\Sigma}_{*,\mathbf{v}} \boldsymbol{\Sigma}_{\mathbf{v},\mathbf{v}}^{-1} \mathbf{v}_{m-1} \tag{S14}$$

$$\boldsymbol{\Sigma}_{*|\mathbf{v}} \triangleq \boldsymbol{\Sigma}_{*,*} + \mathbf{V}_{\mathrm{noise}*} - \boldsymbol{\Sigma}_{*,\mathbf{v}} \boldsymbol{\Sigma}_{\mathbf{v},\mathbf{v}}^{-1} \boldsymbol{\Sigma}_{*,\mathbf{v}}^{\top} \tag{S15}$$

B.2 Gibbs sampling for truncated MVN

Here, we provide the Gibbs sampling algorithm for truncated MVN distribution utilized to draw the hallucination sample in the HB algorithm.

Algorithm 2 Gibbs sampling for truncated MVN (Takeno et al. [41])

1: **Input**: $\mathbf{v}_0, \boldsymbol{\Sigma}$
2: Compute $\boldsymbol{\Sigma}^{-1}$
3: **for** $i = 1, \ldots$ **do**
4: $\quad \mathbf{v}_i \leftarrow \mathbf{v}_{i-1}$
5: $\quad$ **for** $j = 1, \ldots$ **do**
6: $\qquad \mu_{i,j} \to [\boldsymbol{\Sigma}^{-1} \mathbf{v}_i]_j / [\boldsymbol{\Sigma}^{-1}]_{j,j}$
7: $\qquad$ Set $\mathbf{v}_{i,j}$ by sampling from $\mathcal{N}(v_{i,j} | \mu_{i,j}, 1/[\sigma^{-1}]_{j,j})$ with truncation above at 0
8: $\quad$ **end for**
9: **end for**

C Proofs

C.1 Risk analysis

The proof begins by providing the definitions of *Hölder class* and the *kernel of order ℓ*, which are used to construct the necessary assumptions.

We first introduce *multi-index* notation as follows:

$$|\mathbf{s}| = s_1 + \cdots + s_d, \quad \mathbf{x}^s = x_1^{s_1} \ldots x_d^{s_d}$$

with $\mathbf{s} \in \mathbb{N}_0^d, \mathbf{x} \in \mathbb{R}^d$, and $|.|$ denotes the magnitude of $\mathbf{s}$. Using *multi-index* notation, the derivative of a function $f : \mathbb{R}^d \to \mathbb{R}$ is denoted by

$$D^{|\mathbf{s}|} f = \frac{\partial^{|\mathbf{s}|} f}{\partial x_1^{s_1} \ldots \partial x_d^{s_d}}$$

Definition C.1 *Assume that $\mathcal{X} \subseteq \mathbb{R}^d$ and let $\beta, L > 0$. The Hölder class $\Sigma(\beta, L)$ on $\mathcal{X}$ is defined as the set of $|\mathbf{s}| = \beta - 1$ times differentiable functions $f : \mathcal{X} \to \mathbb{R}$ whose derivative $D^{|\mathbf{s}|} f$ satisfies*

$$|D^{|\mathbf{s}|} f(\mathbf{x}) - D^{|\mathbf{s}|} f(\mathbf{y})| \leq L \|\mathbf{x} - \mathbf{y}\| \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{X} \tag{S16}$$

with $\|.\|$ denotes the metric on $\mathcal{X}$.

Definition C.2 *Let $\ell \geq 1$ be an integer. We say that $k : \mathbb{R} \to \mathbb{R}$ is a kernel of order ℓ if the functions $r \mapsto r^j k(r), j = 0, \ldots, \ell$ are integrable and satisfy*

$$\int k(r) dr = 1, \quad \int r^j k(r) dr = 0 \quad j = 1, \ldots, \ell \tag{S17}$$

It is also important to introduce the notion of *Lipschitz continuity* as the foundation for our proofs.

Definition C.3 *A function $f : \mathcal{X} \subseteq \mathbb{R}^d \to \mathbb{R}$ is Lipschitz-continuous if there exists $K > 0$ such that*

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq K\|\mathbf{x} - \mathbf{y}\| \tag{S18}$$

with K and $\|.\|$ denote the Lipschitz constant and metric on $\mathcal{X}$, respectively.

Subsequently, we mention all the propositions responsible for developing Lemma C.6. Later on, we also employ these propositions to prove Theorem 4.3. The propositions provide the bound for the variance and the bias of the KDE, respectively.

Proposition C.4 *(Sen [35]) Suppose that the density $p(\mathbf{x})$ satisfies $p(\mathbf{x}) \leq p_{\max} < \infty$ for all $\mathbf{x} \in \mathbb{R}^d$. Let $k : \mathbb{R} \to \mathbb{R}$ be the kernel such that*

$$\int k^2(r)dr \leq \infty$$

Then, for any $\mathbf{x} \in \mathbb{R}^d$, $h > 0$, and $n \geq 1$ we have

$$\mathbb{V}[\hat{p}(\mathbf{x}|h, \mathbf{X}_0)] \leq \frac{c_1}{nh^d} \tag{S19}$$

with $c_1 \triangleq p_{\max} \int k^2(r)dr$.

Proposition C.5 *(Sen [35]) Assume $p \in \mathcal{P}(\beta, L)$ and let k be a kernel of order $\ell = \lfloor \beta \rfloor$. Then for any $\mathbf{x} \in \mathbb{R}^d, h > 0$, and $n \geq 1$ we have*

$$|\mathbb{E}[\hat{p}(\mathbf{x}|h, \mathbf{X}_0)] - p(\mathbf{x})| \leq c_2 h^\beta \tag{S20}$$

where $c_2 \triangleq \frac{L}{\ell!} \int |r|^\beta k(r)dr$

The proof of Theorem 4.2 relies on the following lemma, which provides the bound of the MSE between the density estimator and the true probability density function whenever Assumption 4.1 holds.

Lemma C.6 *(Sen [35]) Suppose that Assumption 4.1 holds. Fix $\alpha > 0$ and take $h = \alpha n^{-1/(2\beta+d)}$. Then, for any $\mathbf{x}$ and $n \geq 1$, the estimated variance $\hat{p}_X(\mathbf{x}|h, \mathbf{X}_0)$ satisfies*

$$\sup_{p \in \mathcal{P}(\beta, L)} \mathbb{E}_{\mathbf{X_0}}[(\hat{p}(\mathbf{x}|h, \mathbf{X}_0) - p(\mathbf{x}))^2] \leq c_3 n^{-\frac{2\beta}{2\beta+d}} \tag{S21}$$

where $c_3 > 0$ is a constant depending only on β, α and on the kernel bandwidth h.

Lemma C.6 obtains the bound by decomposing MSE as the addition of variance and bias squared. Subsequently, it leverages Proposition C.4 and Proposition C.5 to obtain the bound of variance and bias, respectively. Equipped with Lemma C.6, we can now state and prove Theorem 4.2.

Theorem 4.2 *Fix $\alpha > 0$ and take $h = \alpha n^{-1/(2\beta+d)}$. Then, for any input $\mathbf{x}$ and the number of anchors $n \geq 1$, the estimated variance $\hat{\sigma}^2(\mathbf{x})$ satisfies*

$$\sup_{p \in \mathcal{P}(\beta, L)} \mathbb{E}_{\mathbf{X_0}}[(\hat{\sigma}_\varepsilon^2(\mathbf{x}) - \sigma_\varepsilon^2(\mathbf{x}))^2] \leq a^2 c_3 n^{-\frac{2\beta}{2\beta+d}} \tag{12}$$

where $c_3 > 0$ is a constant depending on Hölder class parameter β, constant α, scaling factor $a > 0$, and the kernel bandwidth h.

Proof:

By invoking the definition of the variance of the noise distribution, we derive:

$$(\hat{\sigma}_\varepsilon^2(\mathbf{x}) - \sigma_\varepsilon^2(\mathbf{x}))^2 = \left(|\hat{\sigma}_\varepsilon^2(\mathbf{x}) - \sigma_\varepsilon^2(\mathbf{x})|\right)^2 = a^2 \left(|\exp(-\hat{p}(\mathbf{x}|h, \mathbf{X}_0)) - \exp(-p(\mathbf{x}))|\right)^2. \tag{S22}$$

Define the map $g : s \mapsto \exp(-s)$ for $s \in [0, 1]$. g is *Lipschitz-continuous* with Lipschitz constant $K = 1$. Furthermore, $p(\mathbf{x})$ and $\hat{p}(\mathbf{x}|h, \mathbf{X}_0)$ are quantities bounded in $[0, 1]$ for all $\mathbf{x}$. Then, the following bound holds:

$$|\exp(-\hat{p}(\mathbf{x}|h, \mathbf{X}_0)) - \exp(-p(\mathbf{x}))| \leq |\hat{p}(\mathbf{x}|h, \mathbf{X}_0) - p(\mathbf{x})| \tag{S23}$$

We then apply Lemma C.6 to obtain:

$$\sup_{p\in\mathcal{P}(\beta,L)} \mathbb{E}_{\mathbf{X_0}}[a^2\left(|\exp(-\hat{p}(\mathbf{x}|h,\mathbf{X}_0))-\exp(-p(\mathbf{x}))|\right)^2] \leq a^2 \sup_{p\in\mathcal{P}(\beta,L)} \mathbb{E}_{\mathbf{X_0}}[|\hat{p}(\mathbf{x}|h,\mathbf{X}_0)-p(\mathbf{x})|]$$

$$\leq a^2 c_3 n^{-\frac{2\beta}{2\beta+d}} \tag{S24}$$

The choice of h is based on the optimum bandwidth of the regular KDE [35].

C.2   Concentration analysis

This analysis requires an assumption that depends on $\omega-$covering number. As a starting point, we define $\omega-$covering as follows.

Definition C.7 *Let $(\mathcal{X},\|.\|)$ be a metric space and $\mathcal{S}\subset\mathcal{X}$. $\{\mathbf{x}_1,\ldots,\mathbf{x}_n\}\in\mathcal{X}^n$ is an $\omega-$covering of $\mathcal{S}$ if $\forall\mathbf{y}\in\mathcal{S}$, $\exists i$ such that $\|\mathbf{y}-\mathbf{x}_i\|\leq\omega$*

Based on the definition above, $\omega-$covering number tells the number of ω-balls to cover a given space $\mathcal{S}$ by allowing the overlaps between the balls. The formal definition of $\omega-$covering number is provided below.

Definition C.8 *(covering number)* $N(\mathcal{S},\|.\|,\omega)=\min\{n:\exists\,\omega-covering\;over\,\mathcal{S}\;of\;size\,n\}$

Here, we additionally assume the kernel k belongs to a bounded measurable VC class, as described below.

Assumption C.9 *The KDE kernel k belongs to a collection of measurable functions $\mathcal{F}_h = \left\{k\left(\frac{\|\mathbf{x}-.\|}{h}\right),\mathbf{x}\in\mathcal{X}\subseteq\mathbb{R}^d,h>0\right\}$, with $\mathcal{F}_h$ satisfies*

$$\sup_{\mathbb{P}} N\left(\mathcal{F}_h,L_2(\mathbb{P}),\omega\|F\|_2\right)\leq\left(\frac{m}{\omega}\right)^v \tag{S25}$$

where $N\left(\mathcal{F}_h,L_2(\mathbb{P}),\omega\|F\|_2\right)$ denotes the $\omega-$covering number of metric space $(\mathcal{F}_h,L_2(\mathbb{P}))$, F is the envelope function of $\mathcal{F}$, the constants $m,v>0$ are the VC characteristics of $\mathcal{F}_h$, and the supremum is taken over the set of all probability measures on $\mathbb{R}^d$.

Theorem 4.3 is built upon the following Lemmas. The first lemma is the renowned *Bernstein's inequality*. The second lemma tells the appropriate ϵ value to guarantee the probability of KDE exponentially deviating from its expectation depending on the number of *anchors* as well as the dimensionality of the data [13].

Lemma C.10 *(Bernstein [6]) Suppose that $W_1,\ldots,W_n$ are i.i.d. random variables with mean μ, $\mathbb{V}[W_i]\leq\sigma^2$, and $|W_i|\leq b$. Then for any $\epsilon>0$, the following inequality holds.*

$$\mathbb{P}\left(|\bar{W}-\mu|>\epsilon\right)\leq 2\exp\left(\frac{n\epsilon^2}{2\sigma^2+2b\epsilon/3}\right) \tag{S26}$$

with $\bar{W}$ denotes the sample mean.

Lemma C.11 *(Giné and Guillou [13]) Suppose the kernel k satisfies Assumption C.9. Given a fixed $h>0$, there exists constants $c_3,c_4>0$, s.t. $\forall\epsilon>0$ and all large n,*

$$\mathbb{P}\left(\sup_{\mathbf{x}\in\mathcal{X}}|\hat{p}(\mathbf{x}|h,\mathbf{X}_0)-\mathbb{E}[\hat{p}(\mathbf{x}|h,\mathbf{X}_0)]|>\epsilon\right)<c_3\exp\left(-c_4 nh^d\epsilon^2\right) \tag{S27}$$

Let $h_n\to 0$ as $n\to\infty$ in such a way that $\frac{nh_n^d}{|\log h_n^d|}\to\infty$. Let

$$\epsilon_n\geq\sqrt{\frac{|\log h_n|}{nh_n^d}} \tag{S28}$$

Then for all large n, Equation (S27) holds with h and ϵ are replaced by h_n and ϵ_n, respectively.

We are now ready to prove Theorem 4.3.

17

Theorem 4.3 *For any $\delta > 0, a \geq 1$, there exist constants c_1 and c_2, such that*

$$\sup_{p \in \mathcal{P}(\beta,L)} \mathbb{P}\left(|\sigma_\varepsilon^2(\mathbf{x}) - \hat{\sigma}_\varepsilon^2(\mathbf{x})| > a\left(\sqrt{4c_1 \log(2/\delta)/nh^d} + c_2 h^\beta\right)\right) < \delta \tag{13}$$

for any $\mathbf{x}$. Furthermore, under an additional assumption, the following bound also holds.

$$\sup_{p \in \mathcal{P}(\beta,L)} \mathbb{P}\left(\sup_{\mathbf{x} \in \mathcal{X}}|\sigma_\varepsilon^2(\mathbf{x}) - \hat{\sigma}_\varepsilon^2(\mathbf{x})| > a\left(\sqrt{1/\left(c_4\,n\,h_n^d\right)\log\left(c_3/\delta\right)} + c_2 h^\beta\right)\right) < \delta \tag{14}$$

for constants c_3, c_4, and h_n depends on n.

Proof:

By invoking the definition of the noise variance and applying the triangle inequality, we derive the following bounds:

$$|\sigma_\varepsilon^2(\mathbf{x}) - \hat{\sigma}_\varepsilon^2(\mathbf{x})| = |a \exp(-\hat{p}(\mathbf{x}|h, \mathbf{X}_0)) - a \exp(-p(\mathbf{x}))|$$
$$\leq |a \exp(-\hat{p}(\mathbf{x}|h, \mathbf{X}_0)) - a \exp(-\mathbb{E}[\hat{p}(\mathbf{x}|h, \mathbf{X}_0)])| + |a \exp(-\mathbb{E}[\hat{p}(\mathbf{x}|h, \mathbf{X}_0)]) - a \exp(-p(\mathbf{x}))|$$
$$\leq a\left(|\hat{p}(\mathbf{x}|h, \mathbf{X}_0) - \mathbb{E}[\hat{p}(\mathbf{x}|h, \mathbf{X}_0)]| + |\mathbb{E}[\hat{p}(\mathbf{x}|h, \mathbf{X}_0)] - p(\mathbf{x})|\right) \tag{S29}$$

The last inequality follows the *Lipschitz continuous* property as in the proof of Theorem 4.2. Next, we bound the second term using Proposition C.5, yielding $a|\mathbb{E}[\hat{p}(\mathbf{x}|h, \mathbf{X}_0)] - p(\mathbf{x})| \leq a\,c_2 h^\beta$. Subsequently, we address the first term of Equation (S29). Define that $\hat{p}(\mathbf{x}|h, \mathbf{X}_0) \triangleq \frac{1}{n}\sum_{i=1}^n G^{(i)}$ with the random variable $G^{(i)} \triangleq \frac{1}{h^d}k\left(\frac{\|\mathbf{X}_0^{(i)} - \mathbf{x}\|_2}{h}\right)$. Notably, $|G^{(i)}| \leq \frac{b_1}{h^d}$ where $b_1 = k(\mathbf{0})$. Proposition C.4 provides $\mathbb{V}[G^{(i)}] \leq \frac{c_1}{h^d}$. By applying *Bernstein's inequality*, we obtain the following bound:

$$\mathbb{P}\left(|\hat{p}(\mathbf{x}|h, \mathbf{X}_0) - \mathbb{E}[\hat{p}(\mathbf{x}|h, \mathbf{X}_0)]| > \epsilon\right) \leq 2\exp\left(-\frac{n\epsilon^2}{2c_1 h^{-d} + 2b_1 h^{-d}\epsilon/3}\right)$$
$$\leq 2\exp\left(-\frac{nh^d \epsilon^2}{4c_1}\right) \tag{S30}$$

whenever $\epsilon \leq 3c_1$ and $b_1 = 1$. If we choose $\epsilon = \sqrt{4c_1 \log(2/\delta)/nh^d}$, then the following bound holds.

$$\mathbb{P}\left(|\hat{p}(\mathbf{x}|h, \mathbf{X}_0) - \mathbb{E}[\hat{p}(\mathbf{x}|h, \mathbf{X}_0)]| > \sqrt{\frac{4c_1 \log(2/\delta)}{nh^d}}\right) \leq \delta \tag{S31}$$

Applying Equation (S31) to Equation (S29) results in Equation (13). To prove the second result, we can derive

$$\sup_{\mathbf{x} \in \mathcal{X}}|\hat{\sigma}^2(\mathbf{x}|h, \mathbf{X}_0) - \sigma_\varepsilon^2(\mathbf{x})| = \sup_{\mathbf{x} \in \mathcal{X}}|a\exp(-\hat{p}(\mathbf{x}|h, \mathbf{X}_0)) - a\exp(-p(\mathbf{x}))|$$
$$\leq \sup_{\mathbf{x} \in \mathcal{X}}|a\exp(-\hat{p}(\mathbf{x}|h, \mathbf{X}_0)) - a\exp(-\mathbb{E}[\hat{p}(\mathbf{x}|h, \mathbf{X}_0)])| + \sup_{\mathbf{x} \in \mathcal{X}}|a\exp(-\mathbb{E}[\hat{p}(\mathbf{x}|h, \mathbf{X}_0)]) - a\exp(-p(\mathbf{x}))|$$
$$\leq a\left(\sup_{\mathbf{x} \in \mathcal{X}}|\exp(-\hat{p}(\mathbf{x}|h, \mathbf{X}_0)) - \exp(-\mathbb{E}[\hat{p}(\mathbf{x}|h, \mathbf{X}_0)])| + \sup_{\mathbf{x} \in \mathcal{X}}|\exp(-\mathbb{E}[\hat{p}(\mathbf{x}|h, \mathbf{X}_0)]) - a\exp(-p(\mathbf{x}))|\right)$$
$$\leq a\left(\sup_{\mathbf{x} \in \mathcal{X}}|\hat{p}(\mathbf{x}|h, \mathbf{X}_0) - \mathbb{E}[\hat{p}(\mathbf{x}|h, \mathbf{X}_0)]| + \sup_{\mathbf{x} \in \mathcal{X}}|\mathbb{E}[\hat{p}(\mathbf{x}|h, \mathbf{X}_0)] - p(\mathbf{x})|\right)$$
$$\leq a\left(\sup_{\mathbf{x} \in \mathcal{X}}|\hat{p}(\mathbf{x}|h, \mathbf{X}_0) - \mathbb{E}[\hat{p}(\mathbf{x}|h, \mathbf{X}_0)]| + c_2 h^\beta\right) \tag{S32}$$

The third inequality follows the *Lipschitz continuous* property as in the proof of Theorem 4.2. By applying Lemma C.11 into Equation (S32) and choose $\epsilon_n = \sqrt{\frac{1}{c_4 n h_n^d}\log\left(\frac{c_3}{\delta}\right)}$, such that $\frac{1}{c_4}\log\left(\frac{c_3}{\delta}\right) \geq |\log h_n|$ where c_3, c_4, ϵ_n and h_n satisfy Lemma C.11, then the following bound holds

$$\mathbb{P}\left(\sup_{\mathbf{x} \in \mathcal{X}}|\hat{p}(\mathbf{x}|h, \mathbf{X}_0) - \mathbb{E}[\hat{p}(\mathbf{x}|h, \mathbf{X}_0)]| > \sqrt{\frac{1}{c_4 n h_n^d}\log\left(\frac{c_3}{\delta}\right)}\right) < \delta \tag{S33}$$

Applying Equation (S33) to Equation (S32) results in Equation (14).

D   Additional experiments and visualizations

D.1   Further sine experiment visualizations

We provide additional visualizations for query acquisitions for sine experiments in Figure S1 and Figure S2. The visualizations demonstrate that PBO-UCB and PBO-EI prefer to exploit the region with higher noise variance. On the contrary, PBO-RAHBO and PBO-ANPEI acquire points with lower aleatoric uncertainty inherited in noise.
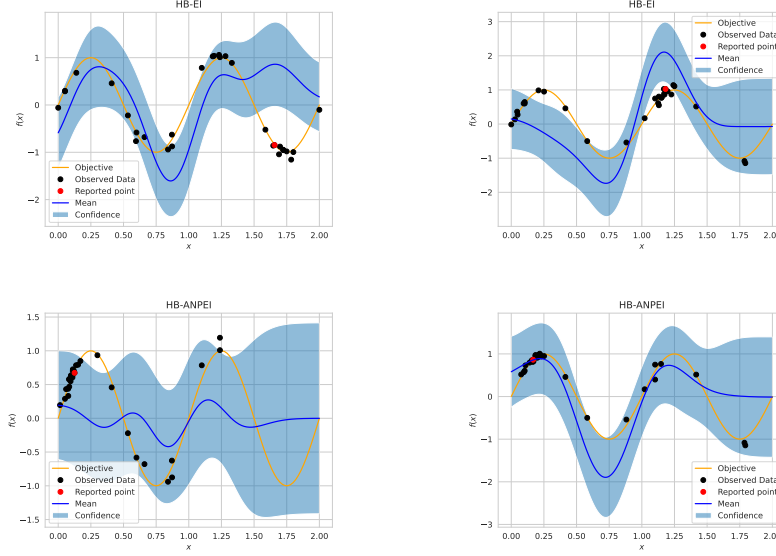

Figure S1: Additional visualizations of query acquisition in the sine experiments. Each column corresponds to a certain initialization. For the same initialization, PBO-ANPEI tends to exploit regions with low noise, whereas PBO-EI's queries are concentrated in high-variance regions.
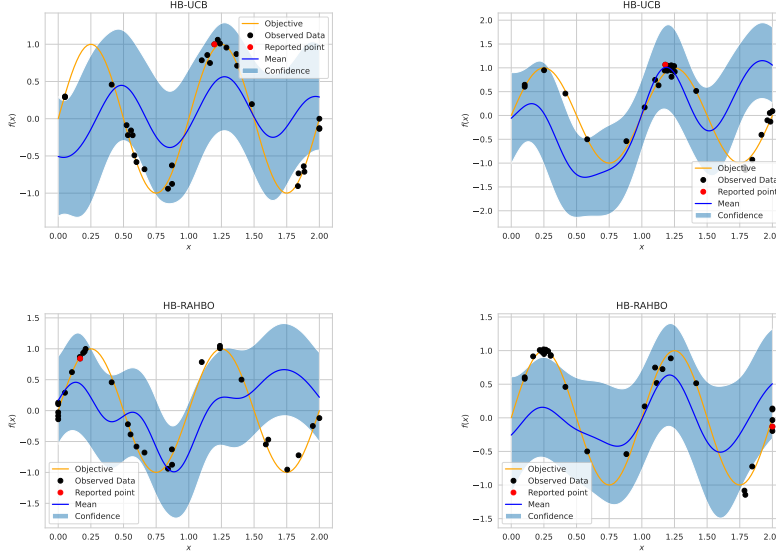

Figure S2: Additional visualizations of query acquisition in the sine experiments. Each column corresponds to a certain initialization. For the same initialization, PBO-RAHBO tends to exploit regions with low noise, whereas PBO-UCB's queries are concentrated in high-variance regions.

D.2 Sine experiments with bandwidth selection via leave-one-out

In this experiment, we maximize the leave-one-out (LOO) criterion to optimize the bandwidth h of the estimator kernel k while keeping all other settings consistent with the main experiments. The results indicate that this approach does not impact either simple or cumulative regret. However, the risk-averse simple and cumulative regret of PBO-ANPEI degrade compared to the main results.
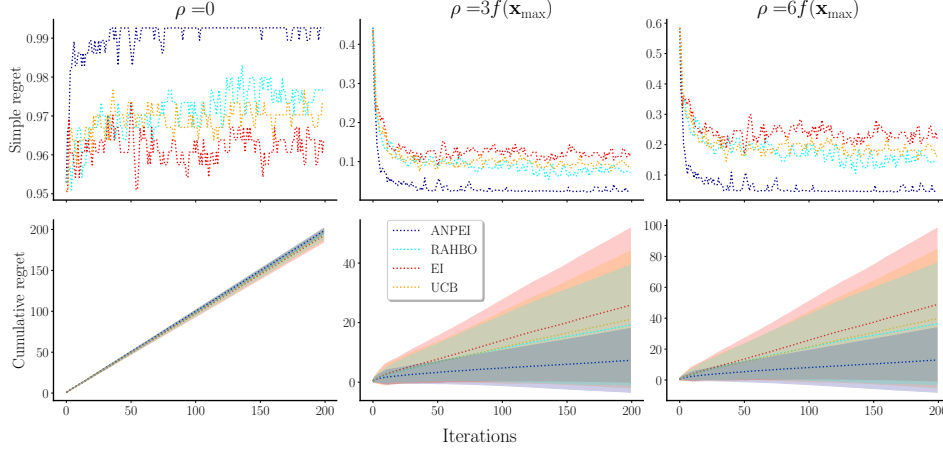

Figure S3: Results for the Sine1D problem, using Leave-One-Out cross-validation to optimize the bandwidth h, for various level of ρ. Mean and standard deviations were computed across 30 random seeds.

D.3 Synthetic function experiments with bandwidth selection via marginal likelihood

We also perform an ablation study on Hartmann4. In this study, we simultaneously select the lengthscale λ of the preferential GP and the bandwidth h by maximizing the marginal log-likelihood. All other settings are kept consistent with the main experiments. The results indicate that this approach does not impact the best value found. However, the objective $\mathrm{MV}(\mathbf{x}_t)$ PBO-RAHBO degrade compared to the main results.
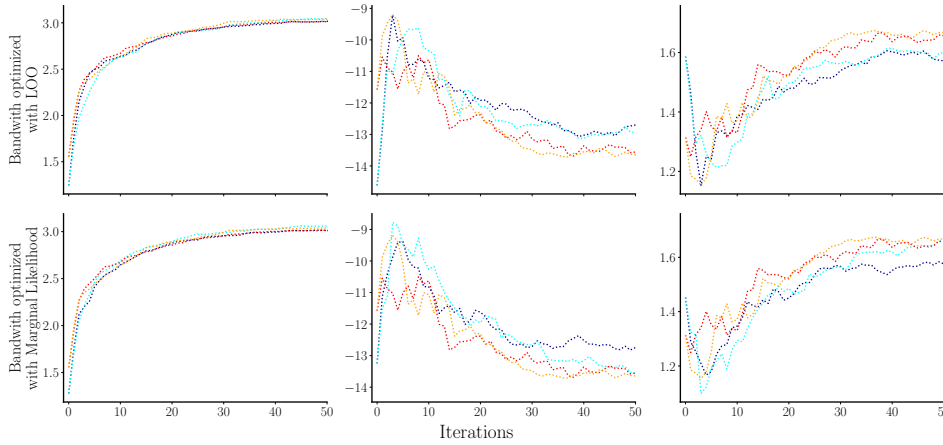

Figure S4: Results for the Hartmann4 problem, using either leave-out-out criterion to optimize the bandwidth h (top row), or marginal log-likelihood (bottom row). $\rho = 0$ (left column), $\rho = 3f(\mathbf{x}_{\max})$ (middle column) and noise variance $\sigma_\varepsilon^2(\mathbf{x})$. Mean BO trajectories computed across 30 random seeds.

D.4 Synthetic functions experiments with student-t noise distribution

In this experiment, we aim to study the performance of heteroscedastic PBO when $p(\mathbf{x})$ is chosen to be less smooth. For the purpose of this experiment, we set $p(\mathbf{x})$ to be a student-t density function with the degree of freedom $\nu > 0$ set to 5. It is worth noting that as ν increases, the density becomes smoother, eventually converging to a normal distribution as ν approaches infinity. All other settings are kept consistent with the main experiments. The results indicate that this approach does not impact the best value found. However, the objective $\mathrm{MV}(\mathbf{x}_t)$ PBO-ANPEI degrade compared to the main results.
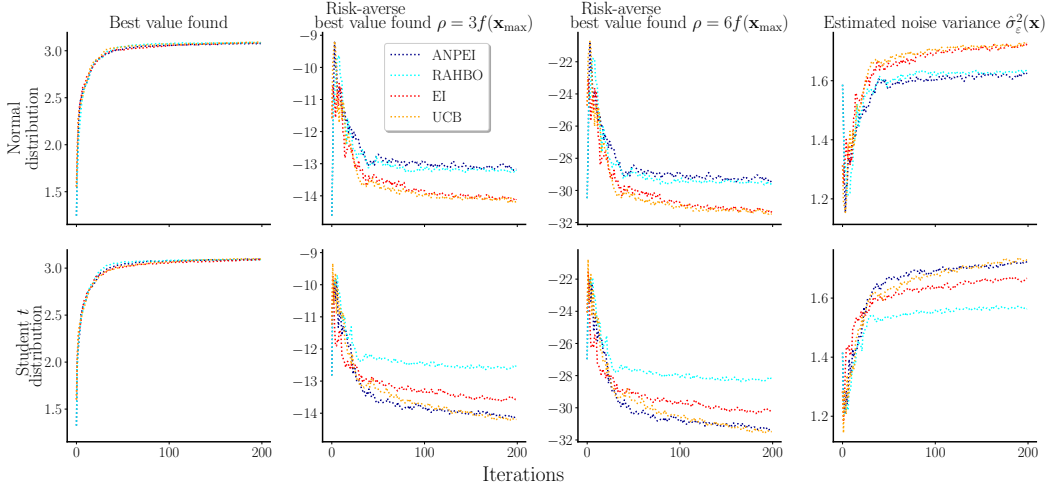

Figure S5: Results for the Hartmann4D function under different noise distributions for the user uncertainty. Mean BO trajectories were computed across 30 random seeds.

E   Experiment details

Branin-2D function:

$$f(\mathbf{x}) = a(x_2 - bx_1^2 + cx_1 - r)^2 + s(1 - t)\cos(x_1) + s$$
$$a = 1,\ b = \frac{5.1}{4\pi^2},\ c = \frac{5}{\pi},\ r = 6\ s = 10,\ t = \frac{1}{8\pi}$$

defined over $\mathcal{X} = [-5, 10] \times [0, 15]$.

Hartmann-4D function:

$$f(\mathbf{x}) = \frac{1}{0.839}\left(1.1 - \sum_{i=1}^{4}\alpha_i\exp\left(-\sum_{j=1}^{4}A_{ij}(x^{(j)} - P_{ij})\right)\right)$$
$$\alpha = (1.0, 1.2, 3.0, 3.2)^T$$
$$\mathbf{A} = \begin{pmatrix} 10 & 3 & 17 & 3.5 \\ 0.05 & 10 & 17 & 0.1 \\ 3 & 3.5 & 1.7 & 10 \\ 17 & 8 & 0.05 & 10 \end{pmatrix}$$
$$\mathbf{P} = 10^{-4}\begin{pmatrix} 1312 & 1696 & 5569 & 124 \\ 2329 & 4135 & 8307 & 3736 \\ 2348 & 1451 & 3522 & 2883 \\ 4047 & 8828 & 8732 & 5743 \end{pmatrix}$$

defined over $\mathcal{X} = [0, 1]^4$.