Anchor-Based Heteroscedastic Noise for Preferential Bayesian Optimization

Marshal Sinaga[1,2,†] **Julien Martinelli**[1,2] **Samuel Kaski**[1,2,3]

[1]ELLIS Institute Finland
[2]Aalto University
[3]University of Manchester
[†]Correspondence to marshal.sinaga@aalto.fi

Abstract

Preferential Bayesian optimization (PBO) learns latent utilities from pairwise comparisons, but most existing methods assume homoscedastic comparison noise. This is inadequate in human-in-the-loop settings, where a user may compare some designs reliably and others only hesitantly. We propose a heteroscedastic noise model for PBO: before optimization, the user provides a small set of reliable examples, called *anchors*, and a kernel density estimator (KDE) turns these anchors into an input-dependent map of user uncertainty. We incorporate this map into preferential GP surrogates and derive risk-averse acquisition functions that trade off utility and ease of comparison. We further show that a risk-adjusted variant of the popular expected utility of the best option (EUBO) preserves the one-step Bayes-optimality guarantee up to an additive constant, and that under an idealized i.i.d. anchor model the KDE estimator enjoys standard consistency and concentration rates. Experiments on synthetic problems and human-preference datasets show improved risk-adjusted performance and clarify how anchor placement affects the method.

1. Introduction

Preferential Bayesian Optimization (PBO) learns a latent utility function from pairwise comparisons rather than direct scalar evaluations (Chu and Ghahramani, 2005; González et al., 2017). It is therefore well suited to human-in-the-loop design problems, where users are typically better at choosing between two options than assigning an absolute score (Kahneman and Tversky, 2013). Applications include visual design optimization (Koyama et al., 2020), recommender systems (Brusilovsky et al., 2007), and expert-guided molecular design (Sundin et al., 2022; Nahal et al., 2024).

A key but under-modeled aspect of human feedback is that comparison difficulty is often input-dependent. For example, in expert-guided molecular design, a chemist may confidently compare familiar small molecules but be less reliable on unfamiliar chemical families or materials classes. Prior PBO methods all assume homoscedastic comparison noise, which treats all duels as equally reliable and leads to risk-neutral query selection (González et al., 2017; Takeno et al., 2023; Astudillo et al., 2023; Xu et al., 2024). When two candidate pairs have similar expected utility but different levels of user uncertainty, ignoring this heterogeneity can waste queries on hard-to-judge comparisons.

A natural reaction is to borrow heteroscedastic noise models from standard BO (Griffiths et al., 2021; Cowen-Rivers et al., 2022; Makarova et al., 2021). In preferential learning, however, binary comparisons provide only ordinal information about latent utility, so flexible utility and noise models

are difficult to disentangle. This makes it natural to look for an additional source of information about *where* user judgments are likely to be reliable.

In many expert-facing applications, this information can be elicited in a lightweight way before optimization starts: the user specifies a small set of examples on which they expect their judgments to be reliable, which we call *anchors*. These anchors mark regions where comparisons are likely to be stable, without requiring confidence scores after each duel. The question is then how to turn them into a practical heteroscedastic PBO method that accounts for both utility and ease of comparison.

Contributions. We study PBO under *user-side heteroscedasticity*, where comparison noise depends on how reliable the user is locally in the design space. Our contributions are:

- **Methodological:** We introduce an anchor-based heteroscedastic preference model, where expert-provided anchors define regions of reliable judgment, inducing an input-dependent noise map *via* Kernel Density Estimation (KDE), and we develop risk-aware acquisition functions for this case.
- **Theoretical:** We show that a risk-adjusted expected utility of the best option (EUBO) preserves one-step Bayes-optimality up to an additive constant, and establish consistency and concentration for the KDE variance estimator under an idealized i.i.d. anchor model.
- **Empirical:** We demonstrate improved risk-adjusted performance on synthetic and real preference benchmarks, and study sensitivity to anchor quality and anchor count.

2. Background

The goal of PBO is to identify the optimum $\mathbf{x}^* = \operatorname{argmax}_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x})$ from duel feedback $\mathbf{x} \succ \mathbf{x}'$, signifying a preference of $\mathbf{x}$ over $\mathbf{x}'$. The latent utility function $f : \mathcal{X} \to \mathbb{R}$ is defined on $\mathcal{X} \subset \mathbb{R}^d$, and the observed comparisons are denoted by $\mathcal{D} \triangleq \{\mathbf{x}_k \succ \mathbf{x}'_k\}_{k=1}^m$.

Prior. We place a zero-mean Gaussian process prior on f,

$$f(\mathbf{x}) \sim \mathcal{GP}(0, l(\mathbf{x}, \mathbf{x}')), \tag{1}$$

with kernel $l : \mathcal{X}\times\mathcal{X} \to \mathbb{R}$. For a finite set of inputs $\mathbf{X}$, the corresponding latent values $\mathbf{f} \triangleq [f(\mathbf{x})]_{\mathbf{x}\in\mathbf{X}}$ follow a multivariate Gaussian distribution with covariance matrix $\mathbf{L} \triangleq [l(\mathbf{x}, \mathbf{x}')]_{\mathbf{x},\mathbf{x}'\in\mathbf{X}}$.

Likelihood. Given duel observations $\mathcal{D}$, we write

$$p(\mathcal{D} \mid \mathbf{f}) = \prod_{k=1}^{m} p(\mathbf{x}_k \succ \mathbf{x}'_k \mid f(\mathbf{x}_k), f(\mathbf{x}'_k)). \tag{2}$$

The duel likelihood takes one of the following forms:

$$p(\mathbf{x} \succ \mathbf{x}' \mid f(\mathbf{x}), f(\mathbf{x}')) = \Phi\left(\frac{f(\mathbf{x}) - f(\mathbf{x}')}{\sqrt{\sigma_\varepsilon^2(\mathbf{x}) + \sigma_\varepsilon^2(\mathbf{x}')}}\right), \qquad \text{(probit)} \tag{3}$$

$$p(\mathbf{x} \succ \mathbf{x}' \mid f(\mathbf{x}), f(\mathbf{x}')) = \frac{\exp(f(\mathbf{x})/\lambda(\mathbf{x}))}{\exp(f(\mathbf{x})/\lambda(\mathbf{x})) + \exp(f(\mathbf{x}')/\lambda(\mathbf{x}'))}, \qquad \text{(logistic)} \tag{4}$$

where $\sigma_\varepsilon^2(\mathbf{x})$ and $\lambda(\mathbf{x})$ denote local noise levels. PBO typically assumes the homoscedastic special cases $\sigma_\varepsilon^2(\mathbf{x}) = \sigma_{\text{noise}}^2$ or $\lambda(\mathbf{x}) = \lambda_{\text{noise}}$; we instead model user uncertainty as input-dependent.

Posterior. Combining the GP prior and the preference likelihood yields the posterior

$$p(\mathbf{f} \mid \mathcal{D}) = \frac{p(\mathcal{D} \mid \mathbf{f})\, p(\mathbf{f})}{\int p(\mathcal{D} \mid \mathbf{f})\, p(\mathbf{f})\, d\mathbf{f}}.$$

Because the preference likelihood is non-Gaussian, this posterior is generally intractable and must be approximated. Standard choices include Laplace, EP, and sampling-based approximations; details for the probit and logistic surrogates used in this work are deferred to Appendices A–D.

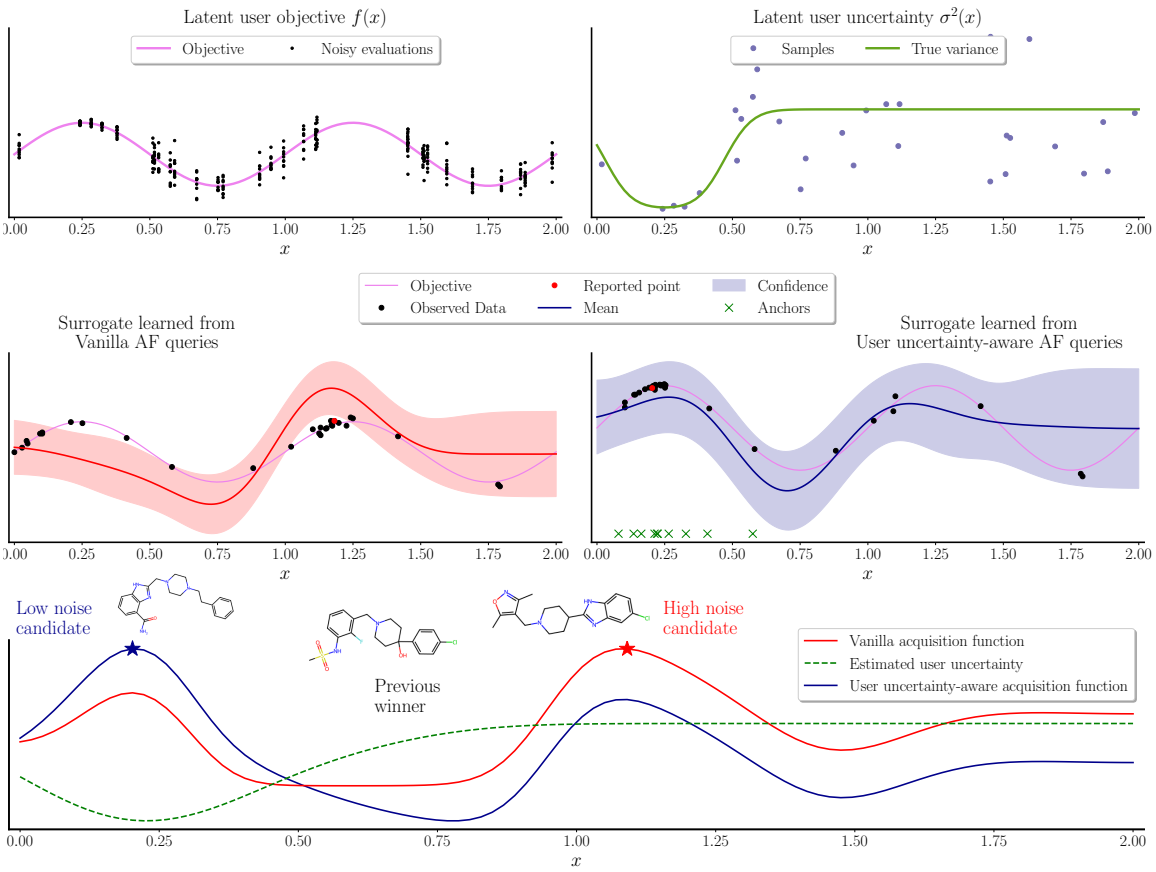

Figure 1. Heteroscedastic preferential Bayesian optimization. Top left: latent user utility with heteroscedastic noisy evaluations (Scalar samples shown only to visualize the latent noise and are not observed by PBO.) Top right: true latent user uncertainty in this toy example. Middle: preferential GP surrogates obtained from queries selected by a vanilla acquisition function (left) and a user-uncertainty-aware acquisition function (right). Our anchor-based model of user uncertainty favors lower-noise yet still promising queries, leading to a better-calibrated surrogate (blue) than the vanilla baseline (red). Importantly, the latent function is inferred from *preferences*, not direct objective evaluations. Bottom: acquisition landscapes. The estimated user uncertainty (green) modifies the acquisition function (blue), shifting its maximizer away from the vanilla one (red) toward a lower-variance candidate, resulting in this illustrative example in two different molecule candidates.

Acquisition function. Given the posterior, PBO selects the next duel via an acquisition function. Some acquisition rules are inherited from single-point BO criteria, of the form

$$\mathbf{x}^{\text{next}} = \operatorname*{argmax}_{\mathbf{x}\in\mathcal{X}} \alpha(\mathbf{x}) := \int u(\mathbf{x}, \mathbf{f})\, p(\mathbf{f} \mid \mathcal{D})\, d\mathbf{f}, \tag{5}$$

where u measures the value of querying $\mathbf{x}$. In PBO, such criteria can be used to select a challenger, which is then paired with the previous winner/current incumbent, a common strategy (Takeno et al., 2023). Alternatively, batch decision-theoretic strategies score a full pair $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2]$ directly,

$$\mathbf{X}^{\text{next}} = \operatorname*{argmax}_{\mathbf{X}\in\mathcal{X}^2} \alpha(\mathbf{X}),$$

as in EUBO (Astudillo et al., 2023). In this case, the two duel elements are optimized jointly, rather than obtained by enumerating all possible pairs. Our heteroscedastic noise model can be used in both single-point and pair-valued acquisition rules.

3. Preferential BO with heteroscedastic noise

We model user uncertainty in the heteroscedastic PBO setting of Section 2. Section 3.1 introduces the input-dependent noise model, Section 3.2 shows how anchors yield a density estimate, and Section 3.3 incorporates the resulting uncertainty map into risk-aware acquisition functions.

3.1. Heteroscedastic noise distribution

For probit surrogates, we replace the constant noise variance by

$$\varepsilon(\mathbf{x}) \sim \mathcal{N}(0, \sigma_\varepsilon^2(\mathbf{x}) = a \exp(-q(\mathbf{x}))), \tag{6}$$

and for logistic surrogates we analogously use

$$\lambda(\mathbf{x}) = a \exp(-q(\mathbf{x})), \tag{7}$$

where $q(\mathbf{x}) \geq 0$ is a reliability score and $a > 0$ sets the overall noise scale. Larger values of $q(\mathbf{x})$ correspond to lower comparison noise. If $q(\mathbf{x})$ is uniform, the model reduces to the usual homoscedastic likelihood. In our anchor-based model below, q controls the spatial variation of user uncertainty, while a calibrates its global magnitude.

3.2. *Anchors*-based input-dependent noise

We introduce *anchors*, a set of reliability markers $\mathbf{X}_0 = \{\mathbf{x}_i\}_{i=1}^n$ provided by the expert, around which judgments are expected to be reliable. Given $\mathbf{X}_0$, we estimate $q(\mathbf{x})$ with a KDE,

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

where k is a kernel and h a bandwidth. A high anchor density implies low comparison noise through Equations (6) and (7). The estimator depends only on anchor locations and can therefore be constructed before the PBO loop.

In standard heteroscedastic BO, a separate GP over the observation noise is natural because repeated scalar observations help disentangle signal and noise. In PBO, binary comparisons are less informative, so utility variation and user uncertainty are more easily confounded. Anchors instead provide external information about *where* comparisons are reliable, and KDE turns this side information into an uncertainty map. This also matches a case-based view of expert judgment, where unfamiliar inputs are assessed by similarity to known cases. We treat anchors as exchangeable user-supplied prototypes; the stronger i.i.d. assumption is only needed for the estimator analysis in Section 4.

3.3. Risk-averse acquisition functions

We interpret heteroscedastic user uncertainty as a source of *risk*: when the local noise level is high, a queried duel is less reliable and thus less useful for optimization. Following prior work on risk-sensitive decision making and risk-averse BO (Kahneman and Tversky, 2013; Hull, 2012; Golub and Tilman, 2000; Makarova et al., 2021; Kuindersma et al., 2013), we favor acquisition functions that trade off high latent utility against low aleatoric uncertainty. In our setting, this uncertainty is quantified by the input-dependent noise variance estimated by the heteroscedastic preference model.

Heuristic acquisition functions. For heuristic PBO rules, we penalize candidates with large estimated noise variance $\sigma_\varepsilon^2(\mathbf{x})$. The first rule is the *aleatoric noise-penalized expected improvement* (ANPEI), which adapts EI to the preferential setting by replacing the incumbent value with the best posterior mean over previously queried points:

$$\alpha_{\text{ANPEI}}(\mathbf{x}) = \mathbb{E}\Big[\big(\mu_{*|\mathcal{D}}(\mathbf{x}) - \mu_{*|\mathcal{D}}(\mathbf{x}^*)\big)_+\Big] - \gamma\sigma_\varepsilon(\mathbf{x}), \tag{9}$$

where $\gamma > 0$ controls the strength of the penalty. Here $\mu_{*|\mathcal{D}}$ and $\sigma^2_{*|\mathcal{D}}$ denote the posterior mean and variance of the preferential GP surrogate, and $\mathbf{x}^*$ is the previously queried point with largest posterior mean. The second rule is the *risk-averse upper confidence bound* (RAHBO). For $\eta, \gamma > 0$,

$$\alpha_{\text{RAHBO}}(\mathbf{x}) = \mu_{*|\mathcal{D}}(\mathbf{x}) + \eta\sigma_{*|\mathcal{D}}(\mathbf{x}) - \gamma\sigma^2_{\varepsilon}(\mathbf{x}), \tag{10}$$

Decision-theoretic acquisition. For decision-theoretic PBO, we build on EUBO (Astudillo et al., 2023), which scores a duel through the expected utility of its best element. To make this criterion risk-averse, we replace utility by the mean-variance style quantity $\text{MV}(\mathbf{x}) \triangleq f(\mathbf{x}) - \alpha\lambda(\mathbf{x})$, where $\lambda(\mathbf{x})$ is the local noise level under the logistic preference model. This yields

$$\alpha_{\text{RAEUBO}}(\mathbf{X}) = \mathbb{E}_m[\max\{\text{MV}(\mathbf{x}_1), \text{MV}(\mathbf{x}_2)\}], \tag{11}$$

with $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2]$ and $\mathbb{E}_m$ the conditional expectation given the current dataset. RAEUBO thus scores the *whole duel* jointly, favoring pairs with high latent utility and low user uncertainty.

Because the penalty terms in (9)–(11) are additive, the resulting policies naturally prefer regions where the user is expected to be more consistent, which in practice often occurs near the anchors. This is related to πBO (Hvarfner et al., 2022), where prior information is injected directly into the acquisition function using a multiplicative prior term, $\alpha_{\pi\text{BO}}(\mathbf{x}) \triangleq \alpha(\mathbf{x})\pi(\mathbf{x})$, whereas our approach uses an additive penalty induced by the estimated heteroscedastic noise.

4. Theoretical Analysis

Our theoretical results support two components of the method. First, we show that RAEUBO retains a one-step decision-theoretic interpretation after replacing utility by its risk-adjusted counterpart. Second, for the anchor-based uncertainty estimator, we derive consistency and concentration results as the number of anchors grows, under an idealized i.i.d. anchor model.

4.1. Risk-averse one-step Bayes optimal policy of RAEUBO

Given $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2]$, define the one-step risk-adjusted value

$$\hat{V}_n^{\lambda}(\mathbf{X}) = \mathbb{E}_n\left[\max_{\mathbf{x}\in\mathbf{X}} \mathbb{E}_n[\text{MV}(\mathbf{x})|\mathbf{X}, r(\mathbf{X})]\right]. \tag{12}$$

Here $\mathbb{E}_n$ denotes expectation conditional on the current dataset $\mathcal{D}_n$, and $r(\mathbf{X}) \in \{1, 2\}$ denotes the index of the winner of the queried pair $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2)$. This is the heteroscedastic counterpart of the qEUBO value analyzed by Astudillo et al. (2023).

Proposition 1. *Suppose human feedback follows the logistic likelihood with a heteroscedastic noise model $\lambda(\mathbf{x}), \forall \mathbf{x} \in \mathcal{X}$. If $\mathbf{X}_* \in \text{argmax}_{\mathbf{X}\in\mathcal{X}^2} \alpha_{\text{RAEUBO}}(\mathbf{X})$, then there exists a constant $\hat{\lambda} > 0$ s.t.*

$$\hat{V}_n^{\lambda}(\mathbf{X}_*) \geq \max_{\mathbf{X}\in\mathcal{X}^2} \hat{V}_n^{0}(\mathbf{X}) - \hat{\lambda}C \tag{13}$$

for $C = L_W((q-1)/e)$ and L_W the Lambert W function.

Thus, maximizing RAEUBO recovers a one-step policy that is optimal up to an additive constant. The proof is deferred to Appendix E.1.

4.2. Consistency analysis of the KDE-based model of user epistemic uncertainty

For the estimator analysis, we study the discrepancy between the estimated local noise variance and the true variance through its mean squared error. To invoke standard KDE consistency results, we strengthen the practical exchangeability view of Section 3.2 and assume that the anchors are i.i.d. We also impose the usual regularity conditions on the underlying density q and variance functions:

Assumption 2. *q belongs to a class of densities $\mathcal{P}(\beta, L)$ defined by*

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

with $\Sigma(\beta, L)$ the Hölder class.

Under this smoothness assumption, the anchor-based estimator converges at the usual KDE rate:

Proposition 3. *Fix $\alpha > 0$ and take $h = \alpha n^{-1/(2\beta+d)}$. Then, for any input $\mathbf{x}$ and any number of anchors $n \geq 1$, the estimated variance satisfies*

$$\sup_{q \in \mathcal{P}(\beta, L)} \mathbb{E}_{\mathbf{X_0}} \left[(\hat{\sigma}_\varepsilon^2(\mathbf{x}) - \sigma_\varepsilon^2(\mathbf{x}))^2 \right] \leq a^2 c_3 n^{-\frac{2\beta}{2\beta+d}}, \tag{14}$$

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

The rate degrades with dimension, which already suggests that anchor-based KDE is primarily useful in low or moderate dimensions. Because the i.i.d. assumption is stronger than our practical anchor-elicitation model, we complement the theory with robustness experiments under few-anchor and non-i.i.d. settings (Section 6 and Figure 3). Of note, this convergence is with respect to the number of anchors, not the number of PBO iterations. Since the KDE map is fixed during a given optimization run, the result quantifies how many anchors are needed to accurately recover the underlying user-uncertainty.

Finally, we complement the MSE analysis with a concentration result that controls the deviation of $\hat{\sigma}_\varepsilon^2(\mathbf{x})$ from the true variance $\sigma_\varepsilon^2(\mathbf{x})$. Under the same assumptions as above, we obtain:

Proposition 4. *For any $\delta > 0$ and $a \geq 1$, there exist constants $c_1, c_2 > 0$ such that, for any fixed $\mathbf{x}$,*

$$\sup_{q \in \mathcal{P}(\beta, L)} \mathbb{P}\left(\left| \sigma_\varepsilon^2(\mathbf{x}) - \hat{\sigma}_\varepsilon^2(\mathbf{x}) \right| \leq a \left(\sqrt{\frac{4c_1}{nh^d} \log \frac{2}{\delta}} + c_2 h^\beta \right) \right) \leq 1 - \delta. \tag{15}$$

Furthermore, under an additional VC-type assumption on the kernel class, there exist constants $c_3, c_4 > 0$ and a bandwidth sequence h_n such that

$$\sup_{q \in \mathcal{P}(\beta, L)} \mathbb{P}\left(\sup_{\mathbf{x} \in \mathcal{X}} \left| \sigma_\varepsilon^2(\mathbf{x}) - \hat{\sigma}_\varepsilon^2(\mathbf{x}) \right| \leq a \left(\sqrt{\frac{1}{c_4 n h_n^d} \log \frac{c_3}{\delta}} + c_2 h_n^\beta \right) \right) \leq 1 - \delta. \tag{16}$$

The first bound gives pointwise high-probability control, the second strengthens this to uniform control over $\mathbf{x}$ under additional VC-type and bandwidth assumptions. Proofs are given in Appendix E.3.

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 (Makarova et al., 2021), surrogate level (Cowen-Rivers et al., 2022; Griffiths et al., 2021) or by adapting distribution-free uncertainty quantification techniques like conformal inference to BO (Stanton et al., 2023). Lastly, recent work studied the heteroscedastic tradeoff between evaluating fewer conditions with more replicates versus more conditions with fewer replicates (Dai et al., 2024).

Preferential Bayesian optimization. As in standard BO, PBO is largely determined by the surrogate and acquisition function. Since its posterior is generally intractable, prior work has relied

on Laplace approximation, skew-GP formulations, EP, and MCMC (Brochu et al., 2010; Benavoli et al., 2021; Takeno et al., 2023). Neural networks have also been used instead of GPs (Huang et al., 2023). On the acquisition side, some work adapted classical strategies to PBO 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 (Siivola et al., 2021). Next, acquisition functions leveraging the preferential nature of the queries were also proposed (González et al., 2017; Astudillo et al., 2023). Recent work proposed theoretically grounded PBO algorithms, with regret bounds (Xu et al., 2024; Pásztor et al., 2024; Kayal et al., 2025). More recently, amortized PBO has been proposed to meta-learn both the surrogate and acquisition function, offering a computationally cheaper alternative to GP-based PBO (Zhang et al., 2025, 2026). Lastly, preferential relations have been leveraged to enhance vanilla BO: Hvarfner et al. (2024) investigate a user-defined prior integrated to the GP surrogate, and (Adachi et al., 2024) allow the prior to be iteratively updated.

6. Experiments

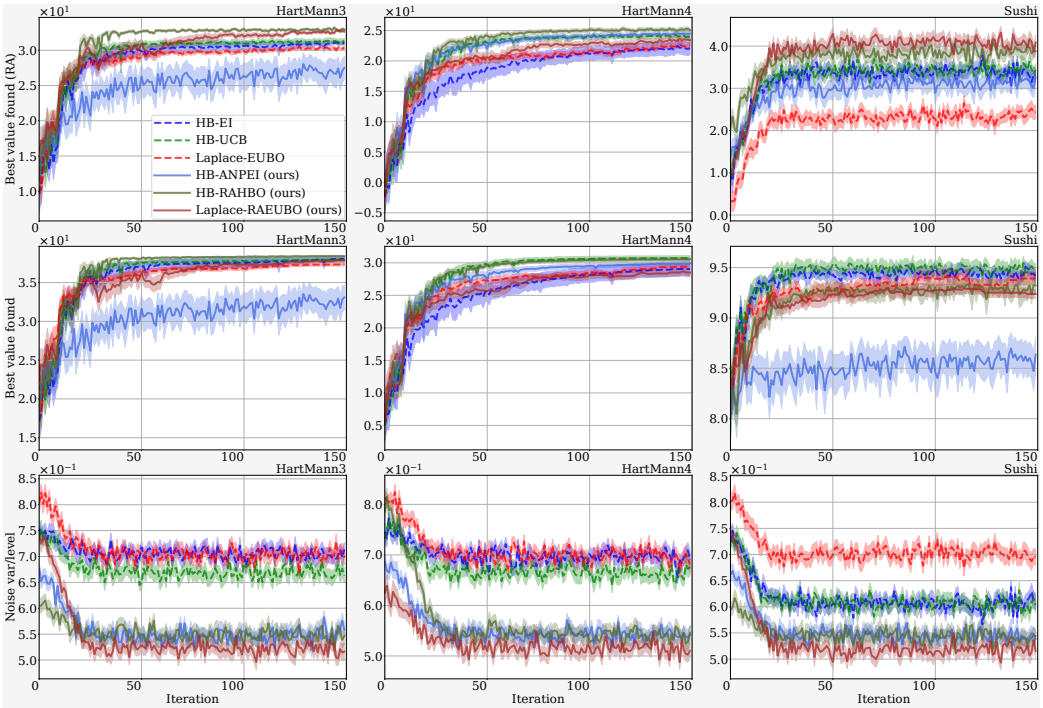

Figure 2. Main results. Top: risk-adjusted best value. Middle: conventional best value. Bottom: noise level of queried pairs. Mean ±0.2 std over 30 runs. **Risk-aware strategies improve the primary risk-adjusted objective and usually query less noisy pairs.**

We evaluate whether modeling user uncertainty changes which duels should be queried across synthetic and real preference benchmarks. Our study covers classical BO test functions (Hartmann3D, Hartmann4D, and Ackley6D) as well as real-world examples (Sushi and Candy, Siivola et al. 2021), together with ablations on anchor quality, anchor count and non-i.i.d. anchor sampling.

Setup. We compare the risk-neutral baselines EI, UCB, and EUBO with their noise-aware counterparts ANPEI, RAHBO, and RAEUBO. In all experiments, we fix $a = 1$ in the heteroscedastic noise model (Equations 6–7) and select the KDE bandwidth h by leave-one-out cross-validation (Equation 8; details in Appendix A.2). The acquisition penalty parameters are fixed across experiments, using $\gamma = \alpha = 10$ and $\eta = 2$ (Equations 9–11). Probit-based methods use Hallucination Believer inference (Takeno et al., 2023), while logistic ones use Laplace approximation (Astudillo et al., 2023);

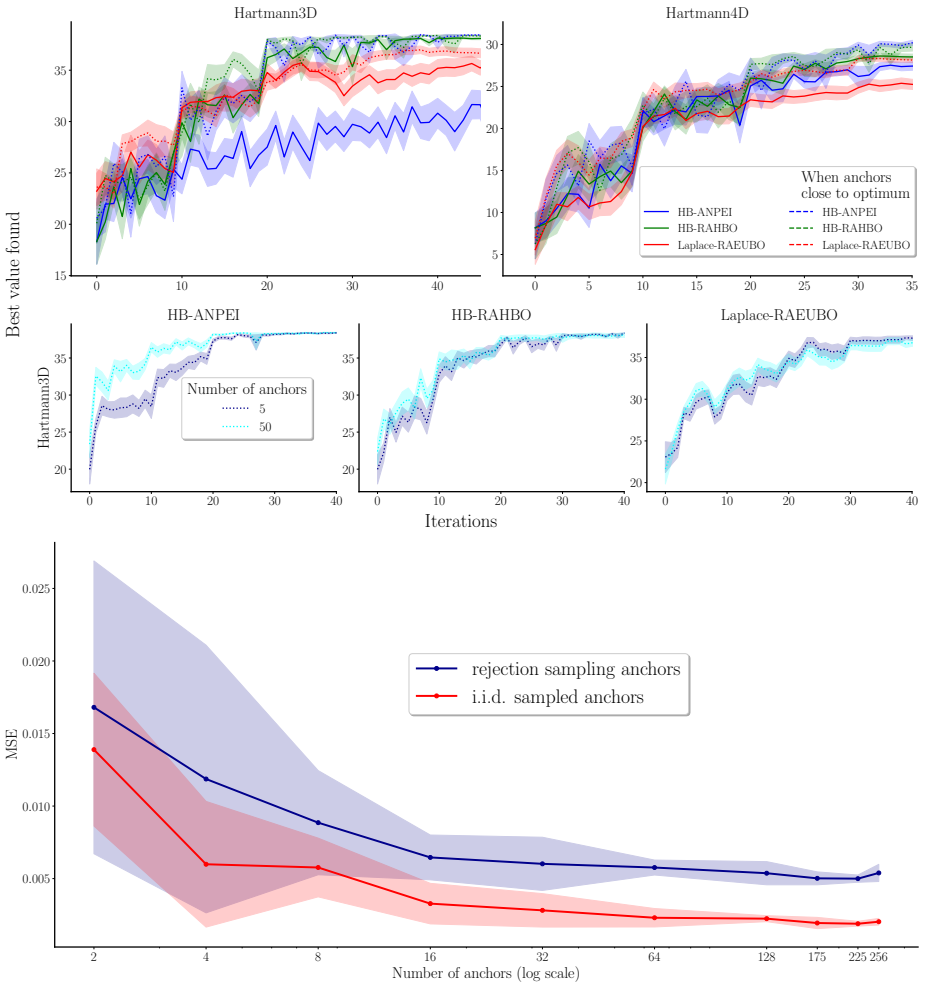

Figure 3. Anchor sensitivity. **Top:** effect of anchor quality on optimization for Hartmann3D and Hartmann4D. **Placing anchors closer to the optimum accelerates convergence**, as the low-uncertainty region better overlaps with the high-utility region. **Middle:** effect of anchor count on Hartmann3D for HB-ANPEI, HB-RAHBO, and Laplace-RAEUBO; **using more anchors improves performance**. **Bottom:** consistency of the KDE-based noise estimator on a 1D problem as a function of the number of anchors, comparing i.i.d. and rejection-sampling anchor processes. Although our theory assumes i.i.d. anchors, **the estimator remains effective under rejection sampling**. Mean ± 0.2 std over 30 random seeds.

the heteroscedastic likelihood can be combined with different posterior approximations, with HB, Laplace, and EP formulations given in the appendix. Surrogate hyperparameters are fit by marginal likelihood maximization.

Metrics. Our primary metric is the risk-adjusted best value $f(\mathbf{x}_t^*) - \rho\, n(\mathbf{x}_t^*)$ with $\rho = 10$, where $\mathbf{x}_t^* \triangleq \operatorname{argmax}_{\mathbf{x}\in\mathcal{X}} \mu_{*|\mathcal{D}}(\mathbf{x}) - \rho n(\mathbf{x})$. Here, $n(\mathbf{x}) = \hat{\sigma}_\varepsilon^2(\mathbf{x})$ for probit surrogates and $n(\mathbf{x}) = \lambda(\mathbf{x})$ for logistic ones. We also report the conventional best-value metric $f(\mathbf{x}_t^*)$, together with the noise level of queried pairs. The risk-adjusted metric is a complementary diagnostic, not a replacement for the conventional best-value objective. It measures whether a method finds candidates that are both promising and expected to be reliably compared by the user, which matters when similar-utility candidates differ in comparison difficulty and cognitive load.

Synthetic functions. To stress-test the method, synthetic anchors are placed away from the optimum, so the low-uncertainty region does not coincide with the most valuable part of the objective landscape. Even in this case, the risk-aware AFs improve the risk-adjusted objective on both Hart-

mann problems (Figure 2, top row). On the conventional best-value metric they remain competitive, while querying systematically lower-noise pairs (bottom row). RAHBO and RAEUBO provide the most stable trade-off, whereas ANPEI is more sensitive to mismatch between low-noise and high-utility regions. On the higher-dimensional Ackley 6D task (Figure S3), only Laplace-RAEUBO remains clearly competitive, consistent with the known deterioration of KDE in higher dimensions.

Real-world examples: learning human preferences on Sushi and Candy. Following Sivola et al. (2021), we turn the 100-item Sushi dataset into a continuous 4D benchmark with a 3-nearest-neighbors regressor. This benchmark construction is inherited from prior PBO work; our contribution is the uncertainty model on top of it. Using 13 representative sushi types as anchors, RAEUBO achieves the best risk-adjusted performance and RAHBO is second (Figure 2, right column). Both remain competitive on the conventional best-value metric while selecting less noisy pairs than their risk-neutral counterparts. We additionally evaluate on a real-world Candy dataset with 85 items, each described by sugar percentage and price percentage, using 12 chocolate-based candies as anchors. On this task, our risk-aware methods again improve over the risk-neutral baselines on both the risk-adjusted and conventional best-value metrics, with the exception of RAEUBO (Figure S4).

Ablation studies on anchors. Figure 3 highlights three practical aspects of the anchor model. First, anchors closer to the optimum accelerate convergence, since the low-uncertainty region better overlaps with the high-utility region (top row). Second, even when anchors are suboptimal, increasing their number improves performance, especially in early iterations (middle row). Third, the bottom panel shows that although our KDE estimator analysis assumes i.i.d. anchors, the estimator remains effective under a non-i.i.d. rejection-sampling anchor process, with error decreasing in a similar way as the number of anchors grows.

7. Discussion and limitations

Anchor-based heteroscedastic noise provides a simple way to inject human-side uncertainty into preferential Bayesian optimization. Empirically, the main effect is not to increase raw utility at any cost, but to redirect the search toward informative and reliable comparisons. This improves the risk-adjusted objective while keeping the conventional best-value metric competitive.

Limitations. Our analysis provides principled support for the proposed approach under a simplified theoretical setting. Extending these guarantees to the full heteroscedastic PBO pipeline under the weaker practical assumption of exchangeable anchors would require additional assumptions and is left for future work. Other natural directions include automatic anchor proposal or validation, and more flexible local-bandwidth noise models.

8. Acknowledgments

This research was supported by the Research Council of Finland (flagship programme: Finnish Center for Artificial Intelligence, FCAI grants 358958, 345604, and 341763), and the UKRI Turing AI World-Leading Researcher Fellowship, EP/W002973/1. We also acknowledge the computational resources provided by the Aalto Science-IT Project.

References

Masaki Adachi, Brady Planden, David Howey, Michael A. Osborne, Sebastian Orbell, Natalia Ares, Krikamol Muandet, and Siu Lun Chau. Looping in the human: Collaborative and explainable Bayesian optimization. In *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*, 2024.

Raul Astudillo, Zhiyuan Jerry Lin, Eytan Bakshy, and Peter Frazier. qeubo: A decision-theoretic acquisition function for preferential bayesian optimization. In *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, 2023.

Alessio Benavoli, Dario Azzimonti, and Dario Piga. Preferential bayesian optimisation with skew gaussian processes. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, 2021.

Sergei Bernstein. On a modification of chebyshev's inequality and of the error formula of laplace. *Ann. Sci. Inst. Sav. Ukraine, Sect. Math*, 1924.

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.

Peter Brusilovsky, Alfred Kobsa, and Wolfgang Nejdl. *The Adaptive Web - Methods and Strategies of Web Personalization*. 01 2007.

Wei Chu and Zoubin Ghahramani. Preference learning with gaussian processes. In *Proceedings of the 22nd international conference on Machine learning*, 2005.

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*, 2022.

Zhongxiang Dai, Quoc Phong Nguyen, Sebastian Tay, Daisuke Urano, Richalynn Leong, Bryan Kian Hsiang Low, and Patrick Jaillet. Batch bayesian optimization for replicable experimental design. *Advances in Neural Information Processing Systems*, 2024.

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*, 2002.

Bennett W Golub and Leo M Tilman. *Risk management: approaches for fixed income markets*, volume 73. 2000.

Javier González, Zhenwen Dai, Andreas Damianou, and Neil D Lawrence. Preferential bayesian optimization. In *International Conference on Machine Learning*, 2017.

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*, 2021.

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

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.

John Hull. *Risk management and financial institutions,+ Web Site*, volume 733. 2012.

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

Carl Hvarfner, Frank Hutter, and Luigi Nardi. A general framework for user-guided bayesian optimization. In *The Twelfth International Conference on Learning Representations*, 2024.

Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I*. 2013.

Aya Kayal, Sattar Vakili, Laura Toni, Da-Shan Shiu, and Alberto Bernacchia. Bayesian optimization from human feedback: Near-optimal regret bounds. In *Proceedings of the 42nd International Conference on Machine Learning*, 2025.

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

Scott R Kuindersma, Roderic A Grupen, and Andrew G Barto. Variable risk control via stochastic optimization. *The International Journal of Robotics Research*, 2013.

Anastasia Makarova, Ilnura Usmanova, Ilija Bogunovic, and Andreas Krause. Risk-averse heteroscedastic bayesian optimization. *Advances in Neural Information Processing Systems*, 2021.

Yasmine Nahal, Janosch Menke, Julien Martinelli, Markus Heinonen, Mikhail Kabeshov, Jon Paul Janet, Eva Nittinger, Ola Engkvist, and Samuel Kaski. Human-in-the-loop active learning for goal-oriented molecule generation. *Journal of Cheminformatics*, 2024.

Jorge Nocedal. Updating quasi-newton matrices with limited storage. *Mathematics of computation*, 1980.

Barna Pásztor, Parnian Kassraie, and Andreas Krause. Bandits with preference feedback: A stackelberg game perspective. *Advances in Neural Information Processing Systems*, 2024.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian processes for machine learning.* MIT Press, 2006.

Vikas C Raykar, Ramani Duraiswami, and Linda H Zhao. Fast computation of kernel estimators. *Journal of Computational and Graphical Statistics*, 2010.

Bodhisattva Sen. *Introduction to Nonparametric Statistics*. 2020.

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)*, 2021.

Samuel Stanton, Wesley Maddox, and Andrew Gordon Wilson. Bayesian optimization with conformal prediction sets. In *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, 2023.

Mervyn Stone. Cross-validatory choice and assessment of statistical predictions. *Journal of the royal statistical society: Series B (Methodological)*, 1974.

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*, 2022.

Shion Takeno, Masahiro Nomura, and Masayuki Karasuyama. Towards practical preferential Bayesian optimization with skew gaussian processes. In *International Conference on Machine Learning*, 2023.

Wenjie Xu, Wenbin Wang, Yuning Jiang, Bratislav Svetozarevic, and Colin Jones. Principled preferential Bayesian optimization. In *Proceedings of the 41st International Conference on Machine Learning*, 2024.

Xinyu Zhang, Daolang Huang, Samuel Kaski, and Julien Martinelli. PABBO: Preferential amortized black-box optimization. In *The Thirteenth International Conference on Learning Representations*, 2025.

Xinyu Zhang, Conor Hassan, Julien Martinelli, Daolang Huang, and Samuel Kaski. In-context multi-objective optimization. In *The Fourteenth International Conference on Learning Representations*, 2026.

Outline.

The appendix is organized as follows:

- Appendix A details the surrogate model, including the probit and logistic likelihoods, hyperparameter optimization, and the complexity of the KDE-based noise estimator.
- Appendices B to D describe the approximate inference schemes used in this work: Hallucination Believer, Laplace approximation, and expectation propagation.
- Appendix E contains the proofs of the main theoretical results, including the one-step guarantee for the risk-adjusted acquisition rule and the consistency and concentration analysis of the KDE-based estimator.
- Appendix F reports additional experimental results, including estimator-consistency studies, ranking plots, computation times, the five-anchor regime, the Candy benchmark, and the Ackley6D experiment.
- Appendix G provides implementation details and a summary of the experimental setup.

A. Details of surrogate model

In this section, we provide the details of the GP surrogate of PBO. Specifically, we discuss the likelihood function, the hyperparameter optimization procedure, and the time complexity of our proposed noise model.

A.1. Likelihood

We consider two likelihoods: probit and logistic likelihoods.

A.1.1. Probit likelihood

As detailed in Chu and Ghahramani (2005), probit likelihood implies that the duel outcome is determined by:

$$\mathbf{x} \succ \mathbf{x}' \iff f(\mathbf{x}) + \varepsilon(\mathbf{x}) > f(\mathbf{x}') + \varepsilon(\mathbf{x}'), \tag{S1}$$

where $\varepsilon(\mathbf{x})$ denote zero-mean additive noise drawn from a normal distribution $\mathcal{N}(0, \sigma_\varepsilon^2(\mathbf{x}))$, for all input $\mathbf{x}$. Given that $\mathbf{x}_k \succ \mathbf{x}'_k$, we model the probit likelihood $\Phi(\mathbf{z}_k)$ as follows:

$$\begin{aligned}
p(\mathbf{x}_k \succ \mathbf{x}'_k | f(\mathbf{x}_k), f(\mathbf{x}'_k)) &= p(\mathbf{x}_k \succ \mathbf{x}'_k | f(\mathbf{x}_k), f(\mathbf{x}'_k), \varepsilon(\mathbf{x}_k), \varepsilon(\mathbf{x}'_k)) \\
&= p(f(\mathbf{x}_k) + \varepsilon(\mathbf{x}_k) - f(\mathbf{x}'_k) - \varepsilon(\mathbf{x}'_k) \geq 0) \\
&= p(\varepsilon(\mathbf{x}'_k) - \varepsilon(\mathbf{x}_k) \leq f(\mathbf{x}_k) - f(\mathbf{x}'_k)) \\
&= p\left(\varepsilon - \varepsilon \leq \frac{f(\mathbf{x}_k) - f(\mathbf{x}'_k)}{\sqrt{a\exp(-\hat{q}(\mathbf{x}_k|h, \mathbf{X}_0)) + a\exp(-\hat{q}(\mathbf{x}'_k|h, \mathbf{X}_0))}}\right) \\
&= \Phi\left(\mathbf{z}_k \triangleq \frac{f(\mathbf{x}_k) - f(\mathbf{x}'_k)}{\sqrt{a\exp(-\hat{q}(\mathbf{x}_k|h, \mathbf{X}_0)) + a\exp(-\hat{q}(\mathbf{x}'_k|h, \mathbf{X}_0))}}\right).
\end{aligned} \tag{S2}$$

Here, $a\exp(-\hat{q}(\mathbf{x}_k|h, \mathbf{X}_0))$ and $a\exp(-\hat{q}(\mathbf{x}'_k|h, \mathbf{X}_0))$ denote the estimate variance of noise $\varepsilon(\mathbf{x}_k)$ and $\varepsilon(\mathbf{x}'_k)$, given a kernel bandwidth h. Φ is the standard Normal distribution c.d.f.

Denoting $v_k \triangleq f(\mathbf{x}'_k) + \varepsilon(\mathbf{x}'_k) - f(\mathbf{x}_k) - \varepsilon(\mathbf{x}_k)$, we have $\mathbf{v}_m < 0 \triangleq \{v_k < 0\}_{k=1}^m$. Recall the notation from the main text $\mathbf{X} \triangleq [\mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{x}'_1, \ldots, \mathbf{x}'_m]^\top \in \mathbb{R}^{2m \times d}$ for the concatenation of winners and

losers from duels $\mathbf{v}_m$, and $\mathbf{f} \triangleq [f(\mathbf{x})]_{\mathbf{x}\in\mathbf{X}}$. The likelihood function $p(\mathcal{D}|\mathbf{f})$ is defined as follows:

$$\begin{aligned} p(\mathcal{D}|\mathbf{f}) &= p(\mathbf{v}_m < 0|\mathbf{f}) \\ &= \prod_{k=1}^{m} \Phi\left(\mathbf{z}_k\right). \end{aligned} \tag{S3}$$

A.1.2. Logistic likelihood

Similarly to the probit likelihood, we make the assumption that the duel outcome may not always be consistent with the latent function f. This inconsistency is now characterized by the input-dependent noise level $\lambda(\mathbf{x})$, giving rise to a logistic likelihood:

$$p(\mathbf{x}_k \succ \mathbf{x}'_k|f(\mathbf{x}_k), f(\mathbf{x}'_k)) = \frac{\exp(f(\mathbf{x}_k)/\lambda(\mathbf{x}_k))}{\exp(f(\mathbf{x}_k)/\lambda(\mathbf{x}_k)) + \exp(f(\mathbf{x}'_k)/\lambda(\mathbf{x}'_k))}. \tag{S4}$$

Using the same notations as in the probit likelihood setting, we define the likelihood function $p(\mathcal{D}|\mathbf{f})$ as follows:

$$p(\mathcal{D}|\mathbf{f}) = \prod_{k=1}^{m} \frac{\exp(f(\mathbf{x}_k)/\lambda(\mathbf{x}_k))}{\exp(f(\mathbf{x}_k)/\lambda(\mathbf{x}_k)) + \exp(f(\mathbf{x}'_k)/\lambda(\mathbf{x}'_k))}. \tag{S5}$$

A.2. Hyperparameter optimization

We minimize the negative log marginal likelihood $p(\mathcal{D}|\theta)$ approximated with Laplace approximation for GP hyperparameter optimization. Following Chu and Ghahramani (2005), the approximation requires obtaining $\mathbf{f}_{\text{MAP}} = \arg\min_{\mathbf{f}} -\log p(\mathbf{f}|\mathbf{v}_m) \approx \arg\min_{\mathbf{f}} S(\mathbf{f})$ where we define $S(\mathbf{f})$ as

$$S(\mathbf{f}) = -\sum_{k=1}^{m} \log p(\mathbf{x}_k \succ \mathbf{x}'_k|f(\mathbf{x}_k), f(\mathbf{x}'_k)) + \frac{1}{2}\mathbf{f}^\top \mathbf{L}^{-1}\mathbf{f} \tag{S6}$$

For the **probit likelihood** function, we find the first and second derivatives are given by

$$\frac{\partial}{\partial \mathbf{f}} S(\mathbf{f}) = \begin{bmatrix} -\frac{1}{\sqrt{a\exp(-\hat{q}(\mathbf{x}_{1:m}|h,\mathbf{X}_0)) + a\exp(-\hat{q}(\mathbf{x}'_{1:m}|h,\mathbf{X}_0))}} \cdot \frac{\phi(\mathbf{z}_{1:m})}{\Phi(\mathbf{z}_{1:m})} \\ \frac{1}{\sqrt{a\exp(-\hat{q}(\mathbf{x}_{1:m}|h,\mathbf{X}_0)) + a\exp(-\hat{q}(\mathbf{x}'_{1:m}|h,\mathbf{X}_0))}} \cdot \frac{\phi(\mathbf{z}_{1:m})}{\Phi(\mathbf{z}_{1:m})} \end{bmatrix} + \mathbf{L}^{-1}\mathbf{f} \quad \in \mathbb{R}^{2m} \tag{S7}$$

$$\frac{\partial^2}{\partial \mathbf{f}\, \partial \mathbf{f}^\top} S(\mathbf{f}) = \Lambda + \mathbf{L}^{-1} \quad \in \mathbb{R}^{2m\times 2m} \tag{S8}$$

$$\Lambda = \begin{bmatrix} \mathbf{c}\cdot \operatorname{diag}\left(\frac{\phi(\mathbf{z}_{1:m})^2}{\Phi(\mathbf{z}_{1:m})^2} + \frac{\phi(\mathbf{z}_{1:m})}{\Phi(\mathbf{z}_{1:m})}\mathbf{z}_{1:m}\right) & \mathbf{c}\cdot \operatorname{diag}\left(-\frac{\phi(\mathbf{z}_{1:m})^2}{\Phi(\mathbf{z}_{1:m})^2} - \frac{\phi(\mathbf{z}_{1:m})}{\Phi(\mathbf{z}_{1:m})}\mathbf{z}_{1:m}\right) \\ \mathbf{c}\cdot \operatorname{diag}\left(-\frac{\phi(\mathbf{z}_{1:m})^2}{\Phi(\mathbf{z}_{1:m})^2} - \frac{\phi(\mathbf{z}_{1:m})}{\Phi(\mathbf{z}_{1:m})}\mathbf{z}_{1:m}\right) & \mathbf{c}\cdot \operatorname{diag}\left(\frac{\phi(\mathbf{z}_{1:m})^2}{\Phi(\mathbf{z}_{1:m})^2} + \frac{\phi(\mathbf{z}_{1:m})}{\Phi(\mathbf{z}_{1:m})}\mathbf{z}_{1:m}\right) \end{bmatrix} \quad \in \mathbb{R}^{2m\times 2m} \tag{S9}$$

$$\mathbf{c} = \left[\frac{1}{a\exp(-\hat{q}_X(\mathbf{x}_{1:m}|h,\mathbf{X}_0)) + a\exp(-\hat{q}_X(\mathbf{x}'_{1:m}|h,\mathbf{X}_0))}\right] \quad \in \mathbb{R}^{m} \tag{S10}$$

where ϕ denotes the probability density function of standard normal distribution. Subsequently, the first and second derivatives of the **logistic likelihood** function are given by

$$\frac{\partial}{\partial \mathbf{f}} S(\mathbf{f}) = \begin{bmatrix} -\frac{1/\lambda(\mathbf{x}_{1:m})\exp(f(\mathbf{x}'_{1:m})/\lambda(\mathbf{x}'_{1:m}))}{\exp(f(\mathbf{x}_{1:m})/\lambda(\mathbf{x}_{1:m})) + \exp(f(\mathbf{x}'_{1:m})/\lambda(\mathbf{x}'_{1:m}))} \\ \frac{1/\lambda(\mathbf{x}'_{1:m})\exp(f(\mathbf{x}'_{1:m})/\lambda(\mathbf{x}'_{1:m}))}{\exp(f(\mathbf{x}_{1:m})/\lambda(\mathbf{x}_{1:m})) + \exp(f(\mathbf{x}'_{1:m})/\lambda(\mathbf{x}'_{1:m}))} \end{bmatrix} + \mathbf{L}^{-1}\mathbf{f} \quad \in \mathbb{R}^{2m} \tag{S11}$$

$$\frac{\partial^2}{\partial \mathbf{f}\, \partial \mathbf{f}^\top} S(\mathbf{f}) = \Lambda + \mathbf{L}^{-1} \quad \in \mathbb{R}^{2m\times 2m} \tag{S12}$$

$$\Lambda = \begin{bmatrix} \frac{\mathbf{c}}{\lambda^2(\mathbf{x}_{1:m})} & \frac{-\mathbf{c}}{\lambda(\mathbf{x}_{1:m})\lambda(\mathbf{x}'_{1:m})} \\ \frac{-\mathbf{c}}{\lambda(\mathbf{x}_{1:m})\lambda(\mathbf{x}'_{1:m})} & \frac{\mathbf{c}}{\lambda^2(\mathbf{x}'_{1:m})} \end{bmatrix} \quad \in \mathbb{R}^{2m\times 2m} \tag{S13}$$

$$\mathbf{c} = \left[\frac{\exp\left(f(\mathbf{x}_{1:m})/\lambda(\mathbf{x}_{1:m}) + f(\mathbf{x}'_{1:m})/\lambda(\mathbf{x}'_{1:m})\right)}{(\exp(f(\mathbf{x}_{1:m})/\lambda(\mathbf{x}_{1:m})) + \exp(f(\mathbf{x}'_{1:m})/\lambda(\mathbf{x}'_{1:m})))^2}\right] \quad \in \mathbb{R}^{m} \tag{S14}$$

The solution of $\frac{\partial}{\partial \mathbf{f}} S(\mathbf{f}) = 0$ provides $\mathbf{f}_{\text{MAP}}$. We use the Newton-Raphson method to obtain $\mathbf{f}_{\text{MAP}}$, following Chu and Ghahramani (2005). The Laplace method approximates $S(\mathbf{f})$ as a Gaussian distribution with mean vector $\mathbf{f}_{\text{MAP}}$ and the covariance matrix $(\mathbf{L}^{-1} + \Lambda_{\text{MAP}})^{-1}$, where $\Lambda_{\text{MAP}} \triangleq \Lambda|_{\mathbf{f}_{\text{MAP}}}$. Given $\mathbf{f}_{\text{MAP}}$ and Λ_{MAP}, we approximate the marginal likelihood as

$$p(\mathcal{D}|\theta) \approx \exp(-S(\mathbf{f}_{\text{MAP}})) \det(\mathbf{I} + \mathbf{L}\Lambda_{\text{MAP}})^{-1/2}. \tag{S15}$$

The hyperparameter optimization problem can be formulated as

$$\theta^* = \arg\min_{\theta} -\log p(\mathcal{D}|\theta), \tag{S16}$$

with θ denoting the lengthscale of GP f. We apply L-BFGS-B (Nocedal, 1980) to obtain θ^*.

Subsequently, we run leave-one-out (LOO) (Hastie et al., 2009) to optimize the KDE kernel bandwidth h. Note that LOO is a special case of cross-validation, where each fold is of size one. The usually computationally expensive technique remains amenable in the context of our study, where human experts typically provide a small number of *anchors*. Given the *anchors* $\mathbf{X}_0$ and bandwidth h, we aim to minimize the negative LOO:

$$h^* \triangleq \arg\min_{h} -\text{LOO}(h) = -\frac{1}{n} \sum_{\mathbf{x}_0 \in \mathbf{X}_0} \log \hat{q}(\mathbf{x}_0|h, \mathbf{X}_0 \setminus \{\mathbf{x}_0\}). \tag{S17}$$

For each anchor $\mathbf{x}_0 \in \mathbf{X}_0$, we employ the remaining *anchors* $\mathbf{X}_0 \setminus \{\mathbf{x}_0\}$ to estimate $q(\mathbf{x}_0)$. 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 KDE bandwidth $\hat{h}$ derived through cross-validation converges to the optimal h value (Stone, 1974). This optimal bandwidth minimizes the Mean Squared Error (MSE) between the true density $q(\mathbf{x})$ and the density estimator $q(\mathbf{x}|h, \mathbf{X}_0)$ for all $\mathbf{x}$. This theorem leads to the proposition below.

Proposition 5. *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,*

$$\frac{\int (\sigma_\varepsilon^2(\mathbf{x}) - \hat{\sigma}_\varepsilon^2(\mathbf{x}|\hat{h}))^2 \mathrm{d}\mathbf{x}}{\inf_h \int (\sigma_\varepsilon^2(\mathbf{x}) - \hat{\sigma}_\varepsilon^2(\mathbf{x}|h))^2 \mathrm{d}\mathbf{x}} \xrightarrow{a.s.} 1. \tag{S18}$$

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.

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 $\hat{q}$ is $\mathcal{O}(NM)$ (Raykar et al., 2010). 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 as GP exhibits time complexity of $\mathcal{O}(N^3)$ (Rasmussen and Williams, 2006).

B. Details of heteroscedastic Hallucination Believer inference

Given the set of duels $\boldsymbol{v}_m < 0$ specified in Appendix A.1, Takeno et al. (2023) propose to generate a new sample $\tilde{\boldsymbol{v}}_m$, called the hallucination drawn from $p(\boldsymbol{v}_m|\boldsymbol{v}_m < 0)$ 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 $\mathbf{f}_* \triangleq [f(\mathbf{x}_1^*), \ldots, f(\mathbf{x}_t^*)]^\top$ follows an MVN distribution, implying a regular GP. This characteristic enables a proper posterior computation of $p(\mathbf{f}_*|\boldsymbol{v}_m < 0, \tilde{\boldsymbol{v}}_m)$, outperforming other approximate inference techniques like Laplace approximation. Nevertheless, we emphasize that the proposed noise distribution can be applied to any approximate inference method.

Algorithm 1 Hallucination Believer ((Takeno et al., 2023)) for posterior approximation and query acquisitions

Input: Initial dataset $\mathcal{D} = \{\mathbf{x}_k \succ \mathbf{x}'_k\}_{k=1}^m$
for $t = 1, \dots$ **do**
  $\mathbf{x}_t \leftarrow \mathbf{x}_{t-1}$ `//set previous winner as first design of the pair`
  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
  Sequentially estimate f and $\sigma_\varepsilon^2(\mathbf{x})$ (Section 3.2)
  $\hat{\mathbf{x}}_t \leftarrow \mathrm{argmax}_{\mathbf{x}\in\mathcal{X}}\, \alpha(\mathbf{x})$ based on GPs $f|\tilde{\mathbf{v}}_{t-1}$ and $\varepsilon(\mathbf{x}) \sim \mathcal{N}(0, \hat{\sigma}_\varepsilon^2(\mathbf{x}))$
  Set the winner as $\mathbf{x}_t$ and the loser as $\mathbf{x}'_t$, respectively.
  $\mathcal{D}_t \leftarrow \mathcal{D}_{t-1} \cup (\mathbf{x}_t \succ \mathbf{x}'_t)$
end for

B.1. Posterior predictive distribution

Following Takeno et al. (2023), we formalize the predictive posterior as follows. We first model the joint distribution between the test function $\mathbf{f}_*$ and the hallucination $\boldsymbol{v}_m$. We obtain $\boldsymbol{v}_m$ by running the Gibbs sampling algorithm for the truncated MVN.

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

$$\boldsymbol{\Sigma} \triangleq \mathbf{A}(\mathbf{L} + \mathbf{B})\mathbf{A}^\top \in \mathbb{R}^{(t+m)\times(t+m)} \tag{S19}$$

$$\mathbf{A} \triangleq \begin{bmatrix} \mathbf{I}_n & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_m & \mathbf{I}_m \end{bmatrix} \in \mathbb{R}^{(t+m)\times(t+2m)} \tag{S20}$$

$$\mathbf{B} \triangleq \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{\text{noise}} \end{bmatrix} \in \mathbb{R}^{(t+2m)\times(t+2m)}, \tag{S21}$$

where $\mathbf{L} = \{l(\mathbf{x}, \mathbf{x}')\}_{\mathbf{x},\mathbf{x}'\in\mathbf{X}\cup\mathbf{X}^*}$ and $\mathbf{V}_{\text{noise}} = \mathrm{diag}([a\exp(-\hat{q}(\mathbf{x}|h, \mathbf{X}_0))]_{\mathbf{x}\in\mathbf{X}})$. We then rewrite the kernel matrix $\boldsymbol{\Sigma}$ into a block of matrices as follows:

$$\boldsymbol{\Sigma} \triangleq \begin{bmatrix} \boldsymbol{\Sigma}_{*,*} & \boldsymbol{\Sigma}_{*,\boldsymbol{v}_m} \\ \boldsymbol{\Sigma}_{\boldsymbol{v}_m,*} & \boldsymbol{\Sigma}_{\boldsymbol{v}_m,\boldsymbol{v}_m} \end{bmatrix} \tag{S22}$$

where $\boldsymbol{\Sigma}_{*,*} \in \mathbb{R}^{t\times t}$, $\boldsymbol{\Sigma}_{*,\boldsymbol{v}_m} \in \mathbb{R}^{t\times m}$, $\boldsymbol{\Sigma}_{\boldsymbol{v}_m,\boldsymbol{v}_m} \in \mathbb{R}^{m\times m}$. For an output vector $\mathbf{f}_*$ of a given predictive inputs $\mathbf{X}^* = \{\mathbf{x}_1^*, \dots, \mathbf{x}_t^*\}$, the Hallucination Believer predictive posterior $p(\mathbf{f}_*|\mathbf{v}_m < 0)$ is defined as:

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

$$\boldsymbol{\mu}_{*|\mathbf{v}_m} \triangleq \boldsymbol{\Sigma}_{*,\mathbf{v}_m}\boldsymbol{\Sigma}_{\mathbf{v}_m,\mathbf{v}_m}^{-1}\mathbf{v}_{m-1} \tag{S23}$$

$$\boldsymbol{\Sigma}_{*|\mathbf{v}_m} \triangleq \boldsymbol{\Sigma}_{*,*} - \boldsymbol{\Sigma}_{*,\mathbf{v}_m}\boldsymbol{\Sigma}_{\mathbf{v}_m,\mathbf{v}_m}^{-1}\boldsymbol{\Sigma}_{*,\mathbf{v}_m}^\top \tag{S24}$$

B.2. Gibbs sampling for truncated MVN

Here, we provide the Gibbs sampling algorithm for truncated MVN distribution to draw the hallucination $\boldsymbol{v}_m$ in the HB inference algorithm. The algorithm utilizes $\boldsymbol{\Sigma}$ from equation S22 to update each component of $\boldsymbol{v}_m$.

C. Details of heteroscedastic Laplace inference

We describe the formulation of Laplace predictive distribution with the heteroscedastic noise as follows. For an output vector $\mathbf{f}_*$ of a given predictive inputs $\mathbf{X}^* = \{\mathbf{x}_1^*, \dots, \mathbf{x}_t^*\}$, the Laplace predictive

Algorithm 2 Gibbs sampling for truncated MVN (Takeno et al. (2023))

Input: $\boldsymbol{v}_0 = \mathbf{0}, \boldsymbol{\Sigma}$ (Equation S22)
Compute $\boldsymbol{\Sigma}^{-1}$
for $i = 1, \dots$ **do**
  $\boldsymbol{v}_i \leftarrow \boldsymbol{v}_{i-1}$
  for $j = 1, \dots$ **do**
    $\mu_{i,j} \to [\boldsymbol{\Sigma}^{-1}\boldsymbol{v}_i]_j / [\boldsymbol{\Sigma}^{-1}]_{j,j}$
    Set $\boldsymbol{v}_{i,j}$ by sampling from $\mathcal{N}(v_{i,j}|\mu_{i,j}, 1/[\sigma^{-1}]_{j,j})$ with truncation above at 0
  end for
end for

posterior $p(\mathbf{f}_*|\mathbf{v}_m < 0)$ is defined as:

$$p(\mathbf{f}_*|\mathbf{v}_m < 0) \triangleq \mathcal{N}(\mathbf{f}_* |\, \boldsymbol{\mu}_{*|\boldsymbol{v}_m}, \boldsymbol{\Sigma}_{*|\boldsymbol{v}_m})$$
$$\boldsymbol{\mu}_{*|\boldsymbol{v}_m} = \boldsymbol{\Sigma}_{*,\boldsymbol{v}_m} \boldsymbol{\Sigma}_{\boldsymbol{v}_m,\boldsymbol{v}_m}^{-1} \mathbf{f}_{\mathrm{MAP}} \tag{S25}$$
$$\boldsymbol{\Sigma}_{*|\boldsymbol{v}_m} = \boldsymbol{\Sigma}_{*,*} - \boldsymbol{\Sigma}_{*,\boldsymbol{v}_m} (\boldsymbol{\Sigma}_{\boldsymbol{v}_m,\boldsymbol{v}_m} + \Lambda_{\mathrm{MAP}})^{-1} \boldsymbol{\Sigma}_{*,\boldsymbol{v}_m}^\top \tag{S26}$$

where $\boldsymbol{\Sigma}_{*,*}, \boldsymbol{\Sigma}_{*,\boldsymbol{v}_m}, \boldsymbol{\Sigma}_{\boldsymbol{v}_m,\boldsymbol{v}_m}$ follow S22 and $\mathbf{f}_{\mathrm{MAP}}, \Lambda_{\mathrm{MAP}}$ follows Appendix A.2. Note that the forms of $\mathbf{f}_{\mathrm{MAP}}$ and Λ_{MAP} depend on the likelihood function.

D. Details of heteroscedastic expectation propagation inference

Following (Takeno et al., 2023), we apply expectation propagation (EP) directly on $\boldsymbol{v}_m|\boldsymbol{v}_m < 0$. Given $\boldsymbol{v}_m \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{v}_m,\boldsymbol{v}_m})$, we define the EP approximation as follows:

$$p(\boldsymbol{v}_m|\boldsymbol{v}_m < 0) \propto \mathcal{N}(\boldsymbol{v}_m|\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{v}_m,\boldsymbol{v}_m}) \prod_{i=1}^m \mathbb{I}(v_i < 0) \tag{S27}$$
$$\approx \mathcal{N}(\boldsymbol{v}_m|\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{v}_m,\boldsymbol{v}_m}) \prod_{i=1}^m \mathcal{N}(v_i|\tilde{\mu}_i, \tilde{\sigma}_i^2) \tag{S28}$$
$$= \mathcal{N}(\boldsymbol{v}_m|\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{v}_m,\boldsymbol{v}_m}) \mathcal{N}(\boldsymbol{v}_m|\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}) \tag{S29}$$
$$\propto \mathcal{N}(\boldsymbol{v}_m|\hat{\boldsymbol{\mu}} = \hat{\boldsymbol{\Sigma}}(\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{\mu}} + \boldsymbol{\Sigma}_{\boldsymbol{v}_m,\boldsymbol{v}_m}^{-1}\mathbf{0}), \hat{\boldsymbol{\Sigma}} = (\boldsymbol{\Sigma}_{\boldsymbol{v}_m,\boldsymbol{v}_m}^{-1} + \tilde{\boldsymbol{\Sigma}}^{-1})^{-1}), \tag{S30}$$

where $\tilde{\boldsymbol{\mu}} \in \mathbb{R}^m \triangleq [\tilde{\mu}_i, \dots, \tilde{\mu}_m]^\top$ and $\tilde{\boldsymbol{\Sigma}} \in \mathbb{R}^{m \times m} \triangleq \mathrm{diag}([\tilde{\sigma}_i^2, \dots, \tilde{\sigma}_m^2])$. We first consider the *cavity distribution*

$$\tilde{p}_{\backslash j}(\boldsymbol{v}_m) = \mathcal{N}(\boldsymbol{v}_m|\bar{\boldsymbol{\mu}}_{\backslash j}, \bar{\boldsymbol{\Sigma}}_{\backslash j})$$
$$\propto \mathcal{N}(\boldsymbol{v}_m|\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{v}_m,\boldsymbol{v}_m}) \prod_{i=1, i\neq j}^m \mathcal{N}(v_i|\tilde{\mu}_i, \tilde{\sigma}_i^2), \tag{S31}$$

where

$$\bar{\mu}_{\backslash jj} = \mathbb{E}_{p_{\backslash j}}[v_j] = \bar{\sigma}_{\backslash jj}^2 (\hat{\mu}_j/\hat{\sigma}_j^2 - \tilde{\mu}_j/\tilde{\sigma}_j^2) \tag{S32}$$
$$\bar{\sigma}_{\backslash jj}^2 = \mathbb{V}_{\tilde{p}_{\backslash j}}[v_j] = (1/\hat{\sigma}_j^2 - 1/\tilde{\sigma}_j^2)^{-1}. \tag{S33}$$

Given $\tilde{p}_{\backslash j}(\boldsymbol{v}_m)$, we define *tilted* distribution as

$$\tilde{p}_{\text{-}j}(v_j) = \underbrace{\tilde{p}_{\backslash j}(v_j)}_{\text{tilted dist.}} \underbrace{\frac{\mathbb{I}(v_j < 0)}{\phi(\zeta)}}_{\text{truncated MVN}}, \tag{S34}$$

where $\zeta = -\bar{\mu}_{\backslash jj}/\bar{\sigma}_{\backslash jj}$. The expectation propagation iteratively optimizes $\tilde{\boldsymbol{\mu}}$ and $\tilde{\boldsymbol{\Sigma}}$ as a part of the global approximation. This is done by minimizing the following Kullback-Leibler (KL) -divergence:

$$\mathrm{KL}[\tilde{p}_{\backslash j}(v_j)\|\tilde{p}_{\text{-}j}(v_j)\mathcal{N}(v_j|\tilde{\mu}_j,\tilde{\sigma}_j^2)] = \int \tilde{p}_{\backslash j}(v_j)\log\frac{\tilde{p}_{\backslash j}(v_j)}{\tilde{p}_{\text{-}j}(v_j)\mathcal{N}(v_j|\tilde{\mu}_j,\tilde{\sigma}_j^2)}\mathrm{d}v_j \tag{S35}$$

In practice, the minimization requires updating $\tilde{\mu}_j$ and $\tilde{\sigma}_j^2$ using the following rules:

$$\tilde{\mu}_j \leftarrow \mathbb{E}_{\tilde{p}_{\backslash j}}[v_j] + \left(\zeta + \frac{\phi(\zeta)}{\Phi(\zeta)}\right)\bar{\sigma}_{\backslash jj} \tag{S36}$$

$$\tilde{\sigma}_j^2 \leftarrow \left(\left(\frac{\zeta\phi(\zeta)}{\Phi(\zeta)} + \left(\frac{\phi(\zeta)}{\Phi(\zeta)}\right)^2\right)^{-1} - 1\right)\bar{\sigma}_{\backslash jj}^2 \tag{S37}$$

For an output vector $\mathbf{f}_*$ of a given predictive inputs $\mathbf{X}^* = \{\mathbf{x}_1^*,\ldots,\mathbf{x}_t^*\}$, the EP predictive posterior $p(\mathbf{f}_*|\mathbf{v}_m < 0)$ is defined as:

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

$$\boldsymbol{\mu}_{*|\boldsymbol{v}_m} = \boldsymbol{\Sigma}_{*,\boldsymbol{v}_m}\left(\boldsymbol{\Sigma}_{\boldsymbol{v}_m,\boldsymbol{v}_m} + \tilde{\boldsymbol{\Sigma}}\right)^{-1}\tilde{\boldsymbol{\mu}} \tag{S38}$$

$$\boldsymbol{\Sigma}_{*|\boldsymbol{v}_m} = \boldsymbol{\Sigma}_{*,*} - \boldsymbol{\Sigma}_{*,\boldsymbol{v}_m}\left(\boldsymbol{\Sigma}_{\boldsymbol{v}_m,\boldsymbol{v}_m} + \tilde{\boldsymbol{\Sigma}}\right)^{-1}\boldsymbol{\Sigma}_{*,\boldsymbol{v}_m}^\top \tag{S39}$$

where $\tilde{\boldsymbol{\mu}}$ and $\tilde{\boldsymbol{\Sigma}}$ are obtained via EP.

E. Proofs

E.1. Risk-averse one-step Bayes optimal policy of RAEUBO

Except for the number of data m, we use the same notations as Astudillo et al. (2023). Recall that we consider the noise level parameter $\lambda(x) = a\exp(-\hat{q}(x|h,\mathbf{X_0}))$. Note that $\lambda(x)$ is bounded for all x, i.e., $\lambda_{\min} \le \lambda(x) \le \lambda_{\max}$, where $\lambda_{\min} = \exp(-1)a$ and $\lambda_{\max} = \exp(0)a$. Our objective is to maximize the risk-averse objective $\mathrm{MV}(\mathbf{x}) \triangleq f(\mathbf{x}) - \alpha\lambda(\mathbf{x})$ for $\alpha > 0$. By definition of MV, we have that

$$\mathbb{E}_m[\max_{\mathbf{x}}\mathbb{E}_m[\mathrm{MV}(\mathbf{x})|\mathbf{X}, r(\mathbf{X})]] = \mathbb{E}_m[\max_{\mathbf{x}}\mathbb{E}_m[f(\mathbf{x}) - \alpha\lambda(\mathbf{x})|\mathbf{X}, r(\mathbf{X})]], \tag{S40}$$

where $r(\mathbf{X}) = r(\mathbf{x}_1,\mathbf{x}_2) = 1$ if $\mathbf{x}_1$ is preferred over $\mathbf{x}_2$ and 2 otherwise. $\hat{V}_m^\lambda$ is the risk-averse counterpart of V_n^λ defined in [Section A.2] of Astudillo et al. (2023). Following Astudillo et al. (2023), we first define:

$$\hat{U}_m^\lambda(X) = \mathbb{E}_m[\mathrm{MV}(\mathbf{x}_{r(\mathbf{X})})] \tag{S41}$$

Then, a parallel with [Lemma A.2] from Astudillo et al. (2023) can be formalized as follows:

Lemma 6. *$\hat{U}_n^\lambda(X) \ge \mathrm{RAEUBO}(\mathbf{X}) - \hat{\lambda}C$, where*

$$\hat{\lambda} \ge \frac{(\mathrm{MV}(\mathbf{x}_2) - \mathrm{MV}(\mathbf{x}_1))\lambda_{\min}\lambda_{\max}}{\mathrm{MV}(\mathbf{x}_2)\lambda_{\max} - \mathrm{MV}(\mathbf{x}_1)\lambda_{\min}}, \tag{S42}$$

and $\mathrm{MV}(\mathbf{x}_2) \ge \mathrm{MV}(\mathbf{x}_1)$ without loss of generality.

Proof:

We note that

$$\mathbb{E}[\mathrm{MV}(\mathbf{x}_{r(\mathbf{X})})|f(\mathbf{X})] = \sum_{i=1}^{2}\frac{\exp(f(\mathbf{x}_i)/\lambda(\mathbf{x}_i))}{\sum_{j=1}^{2}\exp(f(\mathbf{x}_j)/\lambda(\mathbf{x}_j))}\mathrm{MV}(\mathbf{x}_i) \tag{S43}$$

Note that for any $\lambda_i := \lambda(\mathbf{x}_i)$, we can bound [Lemma A.1] in Astudillo et al. (2023), i.e, for any $s_i \in \mathbb{R}$ and $i \in \{1, 2\}$ by choosing $\hat{\lambda} \geq (s_2 - s_1)\lambda_{\min}\lambda_{\max}/(s_2\lambda_{\max} - s_1\lambda_{\min})$. We then find that

$$\sum_{i=1}^{2} \frac{\exp(s_i/\lambda_i)}{\sum_{j=1}^{2} \exp(s_j/\lambda_j)} \geq \sum_{i=1}^{2} \frac{\exp(s_i/\hat{\lambda})}{\sum_{j=1}^{2} \exp(s_j/\hat{\lambda})} \geq \max\{s_1, s_2\} - \hat{\lambda}C \tag{S44}$$

where $s_2 > s_1$ without loss of generality. Next, adapting the bound of [Lemma A.1] of Astudillo et al. (2023), it holds that

$$\mathbb{E}[\mathrm{MV}(\mathbf{x}_{r(\mathbf{X})})|f(\mathbf{X})] \geq \max\{\mathrm{MV}(\mathbf{x}_1), \mathrm{MV}(\mathbf{x}_2)\} - \hat{\lambda}C \tag{S45}$$

Taking expectations $\mathbb{E}_n$ over both sides of the inequality yields the desired result. In our case, this also stems from the fact that our KDE-based noise model $\lambda(\cdot)$ does not depend on the dataset $\mathcal{D}_n$, only on the anchors $\mathbf{X_0}$.

Subsequently, adapting [Lemma A.3] in Astudillo et al. (2023) provides the following:

Lemma 7. $\hat{V}_n^{\lambda}(\mathbf{X}) \geq \hat{U}_n^{\lambda}(\mathbf{X})$ *for all* $\mathbf{X}$.

Proof:

Observe that

$$\begin{aligned}
\hat{V}_n^{\lambda}(\mathbf{X}) &= \mathbb{E}_n[\max_{\mathbf{x}\in\mathcal{X}} \mathbb{E}[\mathrm{MV}(\mathbf{x})|\mathbf{X}, r(\mathbf{X})]] && \text{(S46)}\\
&\geq \mathbb{E}_n[\mathbb{E}[\mathrm{MV}(\mathbf{x}_{r(\mathbf{X})})|\mathbf{X}, r(\mathbf{X})]] && \text{(S47)}\\
&= \mathbb{E}_n[f(\mathbf{x}_{r(\mathbf{X})}) - \alpha\lambda(\mathbf{x}_{r(\mathbf{X})})] && \text{(S48)}\\
&= \hat{U}_n^{\lambda}(\mathbf{X}) && \text{(S49)}
\end{aligned}$$

We can derive a risk-averse version of the Lemmas derived in Astudillo et al. (2023) with these risk-averse versions of [Theorem 2]. Based on the notations from Astudillo et al. (2023), let $\mathbf{X}^* \in \mathrm{argmax}_{\mathbf{X}\in\mathbb{X}^2} \mathrm{RAEUBO}(\mathbf{X})$. Subsequently, let $\mathbf{X}^{**} \in \mathrm{argmax}_{\mathbf{X}\in\mathbb{X}^2} \hat{V}_n^0(\mathbf{X})$. Recall that in our case, $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2)$ and let $\mathbf{X}^+(\mathbf{X}^{**}) = (\mathrm{argmax}_{\mathbf{x}\in\mathbb{X}} \mathbb{E}_n[f(\mathbf{x})|\mathbf{X}^{**}, 1], \mathrm{argmax}_{\mathbf{x}\in\mathbb{X}} \mathbb{E}_n[f(x)|X^{**}, 2])$, we have that

$$\begin{aligned}
\hat{V}_n^{\lambda}(\mathbf{X}^*) &\geq \hat{U}_n^{\lambda}(\mathbf{X}^*) && \text{(S50)}\\
&\geq \hat{U}_n^{0}(\mathbf{X}^*) - \hat{\lambda}C && \text{(S51)}\\
&= \mathrm{RAEUBO}(\mathbf{X}^*) - \hat{\lambda}C && \text{(S52)}\\
&\geq \mathrm{RAEUBO}(\mathbf{X}^+(\mathbf{X}^{**})) - \hat{\lambda}C && \text{(S53)}\\
&\geq \hat{V}_n^{0}(\mathbf{X}^{**}) - \hat{\lambda}C && \text{(S54)}\\
&= \max_{\mathbf{X}\in\mathbf{X}^2} \hat{V}_n^{0}(\mathbf{X}) - \hat{\lambda}C && \text{(S55)}
\end{aligned}$$

The first line follows from Lemma 7. The second line follows from adapted Lemma 6. The third line follows from the definition of $\hat{U}_n^0$. The fourth line follows from the definition of $\mathbf{X}^*$. The fifth line can be obtained as in the proof of [Theorem 1] from Astudillo et al. (2023) (noise-free setting). Finally, the last line follows from the definition of $\mathbf{X}^{**}$.

E.2. Consistency analysis of the KDE-based model of user epistemic uncertainty

E.2.1. 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:

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

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

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

Definition 8. *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*

$$|D^{|\mathbf{s}|} f(\mathbf{x}) - D^{|\mathbf{s}|} f(\mathbf{y})| \leq L\|\mathbf{x} - \mathbf{y}\| \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{X} \tag{S56}$$

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

Definition 9. *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 $r \mapsto r^j k(r), j = 0, \dots, \ell$ are integrable and satisfy*

$$\int k(r)dr = 1, \quad \int r^j k(r)dr = 0 \quad j = 1, \dots, \ell \tag{S57}$$

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

Definition 10. *A function $f : \mathcal{X} \subseteq \mathbb{R}^d \to \mathbb{R}$ is Lipschitz-continuous if there exists $K > 0$ such that, for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}$*

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq K\|\mathbf{x} - \mathbf{y}\| \tag{S58}$$

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

Subsequently, we mention all the propositions responsible for developing Lemma 13. Later on, we also employ these propositions to prove Proposition 4. The propositions provide the bound for the variance and the bias of the KDE, respectively.

Proposition 11. *(Sen (2020)) Suppose that the density $q(\mathbf{x})$ satisfies $q(\mathbf{x}) \leq q_{\max} < \infty$ 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

$$\mathbb{V}\left[\hat{q}(\mathbf{x}|h, \mathbf{X}_0)\right] \leq \frac{c_1}{nh^d} \tag{S59}$$

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

Proposition 12. *(Sen (2020)) Assume $q \in \mathcal{P}(\beta, L)$ 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*

$$|\mathbb{E}[\hat{q}(\mathbf{x}|h, \mathbf{X}_0)] - q(\mathbf{x})| \leq c_2 h^\beta \tag{S60}$$

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

The proof of Proposition 3 relies on the following lemma, which provides the bound of the MSE between the density estimator and the true probability density function whenever Assumption 2 holds.

Lemma 13. *(Sen (2020)) Suppose that Assumption 2 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 $\hat{q}(\mathbf{x}|h, \mathbf{X}_0)$ satisfies*

$$\sup_{q \in \mathcal{P}(\beta, L)} \mathbb{E}_{\mathbf{X_0}}[(\hat{q}(\mathbf{x}|h, \mathbf{X}_0) - q(\mathbf{x}))^2] \leq c_3 n^{-\frac{2\beta}{2\beta+d}} \tag{S61}$$

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

Lemma 13 obtains the bound by decomposing MSE as the addition of variance and bias squared. Subsequently, it leverages Proposition 11 and Proposition 12 to obtain the bound of variance and bias, respectively. Equipped with Lemma 13, we can now state and prove Proposition 3.

Proposition 3. *Fix $\alpha > 0$ and take $h = \alpha n^{-1/(2\beta+d)}$. Then, for any input $\mathbf{x}$ and any number of anchors $n \geq 1$, the estimated variance satisfies*

$$\sup_{q \in \mathcal{P}(\beta, L)} \mathbb{E}_{\mathbf{X_0}}\left[(\hat{\sigma}^2_\varepsilon(\mathbf{x}) - \sigma^2_\varepsilon(\mathbf{x}))^2\right] \leq a^2 c_3 n^{-\frac{2\beta}{2\beta+d}}, \tag{14}$$

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

Proof:

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

$$(\hat{\sigma}^2_\varepsilon(\mathbf{x}) - \sigma^2_\varepsilon(\mathbf{x}))^2 = \left(|\hat{\sigma}^2_\varepsilon(\mathbf{x}) - \sigma^2_\varepsilon(\mathbf{x})|\right)^2 = a^2 \left(|\exp(-\hat{q}(\mathbf{x}|h, \mathbf{X}_0)) - \exp(-q(\mathbf{x}))|\right)^2. \tag{S62}$$

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

$$|\exp(-\hat{q}(\mathbf{x}|h, \mathbf{X}_0)) - \exp(-q(\mathbf{x}))| \leq |\hat{q}(\mathbf{x}|h, \mathbf{X}_0) - q(\mathbf{x})| \tag{S63}$$

We then apply Lemma 13 to obtain:

$$\begin{aligned}\sup_{q \in \mathcal{P}(\beta, L)} \mathbb{E}_{\mathbf{X_0}}[a^2 \left(|\exp(-\hat{q}(\mathbf{x}|h, \mathbf{X}_0)) - \exp(-q(\mathbf{x}))|\right)^2] &\leq a^2 \sup_{q \in \mathcal{P}(\beta, L)} \mathbb{E}_{\mathbf{X_0}}[(\hat{q}(\mathbf{x}|h, \mathbf{X}_0) - q(\mathbf{x}))^2] \\ &\leq a^2 c_3 n^{-\frac{2\beta}{2\beta+d}}\end{aligned} \tag{S64}$$

The choice of h is based on the optimum bandwidth of the regular KDE (Sen, 2020).

E.3. Concentration analysis

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

Definition 14. *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 15. *(covering number) $N(\mathcal{S}, \|.\|, \omega) = \min\{n : \exists\, \omega - \text{covering over} \mathcal{S} \text{ of size } n\}$*

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

Assumption 16. *The KDE kernel k belongs to a collection of measurable functions $\mathcal{F}_h = \left\{\mathbf{z} \mapsto k\left(\frac{\|\mathbf{x}-\mathbf{z}\|}{h}\right), \mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^d, h > 0\right\}$, with $\mathcal{F}_h$ satisfies*

$$\sup_{\mathbb{P}} N\left(\mathcal{F}_h, L_2(\mathbb{P}), \omega\|F\|_2\right) \leq \left(\frac{m}{\omega}\right)^v \tag{S65}$$

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 $\mathbb{R}^d$.

Proposition 4 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 (Giné and Guillou, 2002).

Lemma 17. *(Bernstein (1924)) 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.*

$$\mathbb{P}\left(|\bar{W} - \mu| > \epsilon\right) \leq 2\exp\left(-\frac{n\epsilon^2}{2\sigma^2 + 2b\epsilon/3}\right) \tag{S66}$$

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

Lemma 18. *(Giné and Guillou (2002)) Suppose the kernel k satisfies Assumption 16. Given a fixed $h > 0$, there exists constants $c_3, c_4 > 0$, s.t. $\forall \epsilon > 0$ and all large n,*

$$\mathbb{P}\left(\sup_{\mathbf{x}\in\mathcal{X}} |\hat{q}(\mathbf{x}|h, \mathbf{X}_0) - \mathbb{E}[\hat{q}(\mathbf{x}|h, \mathbf{X}_0)]| > \epsilon\right) < c_3 \exp\left(-c_4 n h^d \epsilon^2\right) \tag{S67}$$

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

$$\epsilon_n \geq \sqrt{\frac{|\log h_n|}{nh_n^d}} \tag{S68}$$

Then for all large n, Equation (S67) holds with h and ϵ are replaced by h_n and ϵ_n, respectively.

We are now ready to prove Proposition 4.

Proposition 4. *For any $\delta > 0$ and $a \geq 1$, there exist constants $c_1, c_2 > 0$ such that, for any fixed $\mathbf{x}$,*

$$\sup_{q\in\mathcal{P}(\beta,L)} \mathbb{P}\left(\left|\sigma_\varepsilon^2(\mathbf{x}) - \hat{\sigma}_\varepsilon^2(\mathbf{x})\right| \leq a\left(\sqrt{\frac{4c_1}{nh^d}\log\frac{2}{\delta}} + c_2 h^\beta\right)\right) \leq 1 - \delta. \tag{15}$$

Furthermore, under an additional VC-type assumption on the kernel class, there exist constants $c_3, c_4 > 0$ and a bandwidth sequence h_n such that

$$\sup_{q\in\mathcal{P}(\beta,L)} \mathbb{P}\left(\sup_{\mathbf{x}\in\mathcal{X}} \left|\sigma_\varepsilon^2(\mathbf{x}) - \hat{\sigma}_\varepsilon^2(\mathbf{x})\right| \leq a\left(\sqrt{\frac{1}{c_4 nh_n^d}\log\frac{c_3}{\delta}} + c_2 h_n^\beta\right)\right) \leq 1 - \delta. \tag{16}$$

Proof:

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

$$\begin{aligned}
&|\hat{\sigma}_\varepsilon^2(\mathbf{x}) - \sigma_\varepsilon^2(\mathbf{x})| = |a\exp(-\hat{q}(\mathbf{x}|h, \mathbf{X}_0)) - a\exp(-q(\mathbf{x}))| \\
&\leq |a\exp(-\hat{q}(\mathbf{x}|h, \mathbf{X}_0)) - a\exp(-\mathbb{E}[\hat{q}(\mathbf{x}|h, \mathbf{X}_0)])| + |a\exp(-\mathbb{E}[\hat{q}(\mathbf{x}|h, \mathbf{X}_0)]) - a\exp(-q(\mathbf{x}))| \\
&\leq a\left(|\hat{q}(\mathbf{x}|h, \mathbf{X}_0) - \mathbb{E}[\hat{q}(\mathbf{x}|h, \mathbf{X}_0)]| + |\mathbb{E}[\hat{q}(\mathbf{x}|h, \mathbf{X}_0)] - q(\mathbf{x})|\right)
\end{aligned} \tag{S69}$$

The last inequality follows the *Lipschitz continuous* property as in the proof of Proposition 3. Next, we bound the second term using Proposition 12, yielding $a|\mathbb{E}[\hat{q}(\mathbf{x}|h,\mathbf{X}_0)] - q(\mathbf{x})| \leq a\, c_2 h^\beta$. Subsequently, we address the first term of Equation (S69). Define that $\hat{q}(\mathbf{x}|h,\mathbf{X}_0) \triangleq \frac{1}{n}\sum_{i=1}^{n} G^{(i)}$ 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 11 provides $\mathbb{V}[G^{(i)}] \leq \frac{c_1}{h^d}$. By applying *Bernstein's inequality*, we obtain the following bound:

$$\begin{aligned}\mathbb{P}\left(|\hat{q}(\mathbf{x}|h,\mathbf{X}_0) - \mathbb{E}[\hat{q}(\mathbf{x}|h,\mathbf{X}_0)]| > \epsilon\right) &\leq 2\exp\left(-\frac{n\epsilon^2}{2c_1 h^{-d} + 2b_1 h^{-d}\epsilon/3}\right) \\ &\leq 2\exp\left(-\frac{nh^d\epsilon^2}{4c_1}\right) \end{aligned} \tag{S70}$$

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.

$$\mathbb{P}\left(|\hat{q}(\mathbf{x}|h,\mathbf{X}_0) - \mathbb{E}[\hat{q}(\mathbf{x}|h,\mathbf{X}_0)]| > \sqrt{\frac{4c_1}{nh^d}\log\frac{2}{\delta}}\right) \leq \delta \tag{S71}$$

Applying Equation (S71) to Equation (S69) results in Equation (15). To prove the second result, we can derive

$$\begin{aligned}&\sup_{\mathbf{x}\in\mathcal{X}} \left|\hat{\sigma}_\varepsilon^2(\mathbf{x}) - \sigma_\varepsilon^2(\mathbf{x})\right| = \sup_{\mathbf{x}\in\mathcal{X}} |a\exp(-\hat{q}(\mathbf{x}|h,\mathbf{X}_0)) - a\exp(-q(\mathbf{x}))| \\ &\leq a\sup_{\mathbf{x}\in\mathcal{X}} |\exp(-\hat{q}(\mathbf{x}|h,\mathbf{X}_0)) - \exp(-\mathbb{E}[\hat{q}(\mathbf{x}|h,\mathbf{X}_0)])| \\ &\quad + a\sup_{\mathbf{x}\in\mathcal{X}} |\exp(-\mathbb{E}[\hat{q}(\mathbf{x}|h,\mathbf{X}_0)]) - \exp(-q(\mathbf{x}))| \\ &\leq a\left(\sup_{\mathbf{x}\in\mathcal{X}} |\hat{q}(\mathbf{x}|h,\mathbf{X}_0) - \mathbb{E}[\hat{q}(\mathbf{x}|h,\mathbf{X}_0)]| + \sup_{\mathbf{x}\in\mathcal{X}} |\mathbb{E}[\hat{q}(\mathbf{x}|h,\mathbf{X}_0)] - q(\mathbf{x})|\right) \\ &\leq a\left(\sup_{\mathbf{x}\in\mathcal{X}} |\hat{q}(\mathbf{x}|h,\mathbf{X}_0) - \mathbb{E}[\hat{q}(\mathbf{x}|h,\mathbf{X}_0)]| + c_2 h^\beta\right). \end{aligned} \tag{S72}$$

The third inequality follows the *Lipschitz continuous* property as in the proof of Proposition 3. By applying Lemma 18 into Equation (S72) and choose $\epsilon_n = \sqrt{\frac{1}{c_4 nh_n^d}\log\left(\frac{c_3}{\delta}\right)}$, 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 18, then the following bound holds

$$\mathbb{P}\left(\sup_{\mathbf{x}\in\mathcal{X}} |\hat{q}(\mathbf{x}|h,\mathbf{X}_0) - \mathbb{E}[\hat{q}(\mathbf{x}|h,\mathbf{X}_0)]| > \sqrt{\frac{1}{c_4 nh_n^d}\log\frac{c_3}{\delta}}\right) < \delta \tag{S73}$$

Applying Equation (S73) to Equation (S72) results in Equation (16).

F. Additional experiments

F.1. Noise model's consistency experiments

This experiment aims to validate the consistency analysis provided in Section 4.2, particularly on Proposition 3. The results are shown in Figure S1. First, we assess the impact of the number of anchors on the estimator's MSE. This experiment considers a univariate normal distribution with mean 0.5 and variance 20 as the ground truth density q. The MSE is computed from 500 data samples. The results confirm that, as predicted by Proposition 3, the MSE decreases as the number of anchors increases. Subsequently, we examine the effect of data dimensionality d on the estimator's MSE. In this setting, we consider uncorrelated multivariate normal distribution with zero means, fixing the number of anchors at 50 for all scenarios. The results show that the MSE increases with higher data dimensionality, again consistent with Proposition 3. Moreover, the optimal kernel bandwidth increases as the dimensionality d grows.

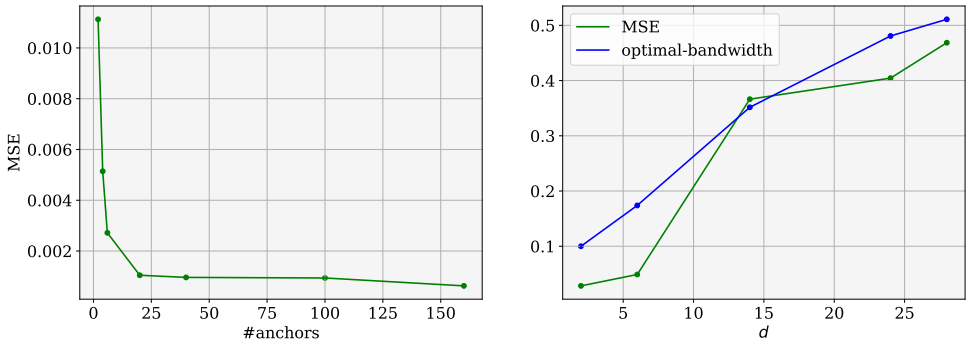

Figure S1. Noise model estimator's consistency results. Left: the estimator's MSE w.r.t. the number of anchors. Right: the estimator's MSE w.r.t. the data dimensionality. Both results are consistent with Proposition 3.

F.2. Ranking Plot

Another way to assess the superiority of our proposed AFs over risk-neutral AFs is to display the average rank as a function of the PBO trial iterations, instead of the achieved reward value as done in Figure 2. For each PBO trial, at every iteration t, we compute the best value found for each given AF, and rank these values. We then compute the average ranking across all trials. The ranking index starts from 0. A lower average ranking indicates a better best value found. The results are presented in Figure S2. Our proposed AFs consistently outperform the baseline in risk-averse settings, except for the ANPEI acquisition function. Conversely, in the risk-neutral scenario, the baseline generally performs better than our methods, except for the Hartmann3 problem, where the RAHBO acquisition function achieves the best ranking.

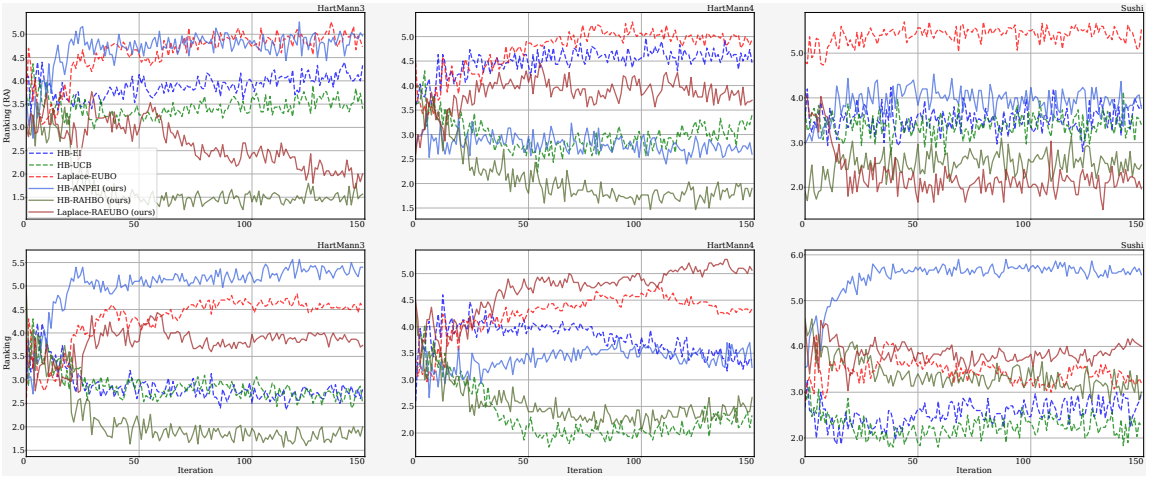

Figure S2. The mean ranking plots of the main experiment presented in Figure 2. The top row depicts the rankings under the risk-averse setting, while the bottom row corresponds to the risk-neutral scenario.

F.3. Computation time

We compare the computational cost of our method and the baselines across acquisition functions, reporting the mean and standard deviation of the per-iteration runtime. As expected, our method is somewhat slower because it additionally estimates local noise levels, but the overhead remains manageable. Higher-dimensional tasks such as Hartmann4D and Sushi increase runtime for all methods. EUBO and RAEUBO are the most expensive because they jointly optimize duel pairs. Results are reported in Table S1.

Table S1. Average computation time (in seconds) and standard deviation per iteration for different acquisition functions.

Problem / AF	UCB	EI	EUBO	RAHBO	ANPEI	RAEUBO
HartMann3	25.3 ± 2.9	26.4 ± 3.5	136.8 ± 20.2	52.4 ± 6.4	22.8 ± 2.8	315.1 ± 29.1
HartMann4	37.9 ± 2.5	35.6 ± 13.5	213.3 ± 40.6	57.6 ± 7.05	33.0 ± 7.3	562.4 ± 64.3
Sushi	38.2 ± 2.5	33.8 ± 8.2	231.3 ± 69.9	45.2 ± 6.7	34.9 ± 4.4	539.3 ± 79.1

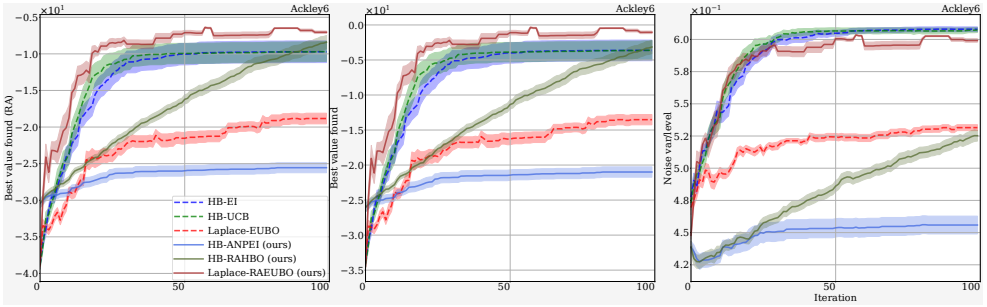

Figure S3. Results on the Ackley 6D benchmark. We report performance over 100 PBO iterations, averaged over 20 repetitions. (**Left**) Risk-adjusted best value found. (**Middle**) Conventional best value found. (**Right**) Queried noise variance/level over time. Laplace-RAEUBO remains competitive in this higher-dimensional setting, while the KDE-based risk-aware methods degrade more substantially, in line with the known difficulty of KDE in higher dimensions.

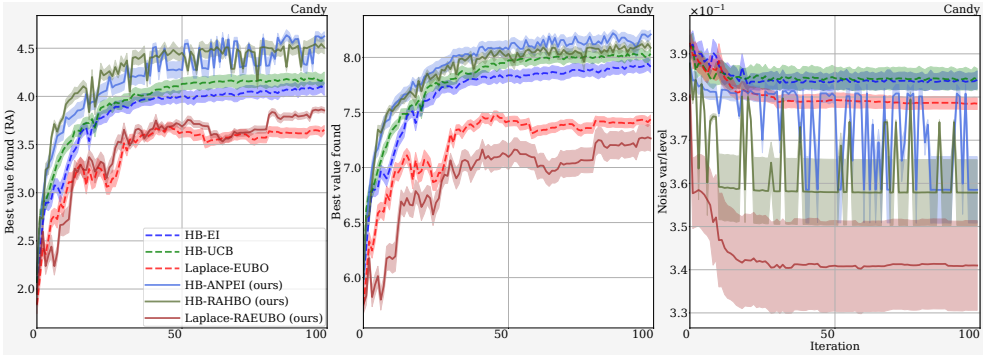

Figure S4. Results on the Candy benchmark. We report performance after 100 PBO iterations on the real-world Candy task, averaged over 20 repetitions. (**Left**) Risk-adjusted best value. (**Middle**) Conventional best value found. (**Right**) Estimated queried noise variance/level. **Risk-aware methods generally achieve stronger risk-adjusted performance while remaining competitive on the standard best-value metric, selecting less noisy comparisons than their risk-neutral counterparts.**

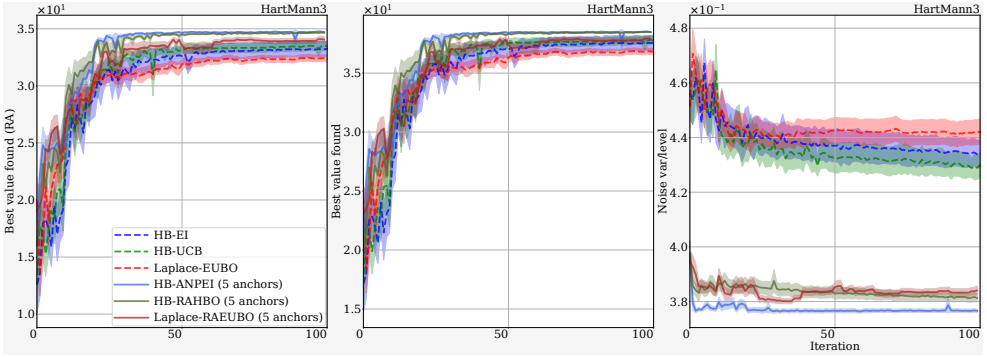

Figure S5. (**Left**) Risk-averse best value found, (**Middle**) overall best value found, and (**Right**) noise variance/level plot for PBO on the Hartmann 3D task after 100 iterations. The results are averaged over 20 repetitions.

G. Further implementation and experiment details

We provide the training details of our experiments to ensure reproducibility and transparency. We set the hyperparameter a defined in Equations 6 and 7 to 1.0. We run 30 PBO trials for 150 rounds, each time with a different initial seed. Regarding our method's hyperparameters, the lengthscale θ is being optimized within $[0.1, 1]$ and the KDE bandwidth h within $[1.0, 2.0]$. The optimal hyperparameters are selected every 10 rounds to adjust to the evolving model complexity as more data becomes available. In terms of acquisition function, we set $\gamma, \alpha = 10.0$ and $\eta = 2.0$, following the standard choice in Makarova et al. (2021). In the case of the ANPEI acquisition strategy, its closed form is computed:

$$\alpha_{\text{ANPEI}} = (\mu_{*|\boldsymbol{v}_m}(\mathbf{x}) - \mu_{*|\boldsymbol{v}_m}(\mathbf{x}^*))\Phi\left(\frac{\mu_{*|\boldsymbol{v}_m}(\mathbf{x}) - \mu_{*|\boldsymbol{v}_m}(\mathbf{x}^*)}{\sigma^2_{*|\boldsymbol{v}_m}(\mathbf{x})}\right) + \sigma_{*|\boldsymbol{v}_m}(\mathbf{x})\phi\left(\frac{\mu_{*|\boldsymbol{v}_m}(\mathbf{x}) - \mu_{*|\boldsymbol{v}_m}(\mathbf{x}^*)}{\sigma^2_{*|\boldsymbol{v}_m}(\mathbf{x})}\right) - \gamma\sigma^2_{\varepsilon}(\mathbf{x}), \tag{S74}$$

with Φ and ϕ denote the cumulative and probability density function, respectively. Finally, to kickstart the GP surrogate, each PBO trial is initialized with 8 duels (16 data points) through Sobol sampling.