ELICITATION-AUGMENTED BAYESIAN OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Human-in-the-loop Bayesian optimization (HITL BO) methods utilize human expertise to improve the cost-efficiency of BO. Most of the HITL BO methods assume that an expert can explicitly quantify their domain knowledge, for instance by pinpointing query locations or specifying their prior beliefs about the location of the maximum as a probability distribution. However, domain knowledge is often tacit and can be expressed only implicitly. We take this into account by eliciting the expert's domain knowledge through pairwise comparisons of designs. We interpret the expert's pairwise judgements as noisy evidence about their mental model, and couple this model with a surrogate model of the objective function via a learned correlation. We develop a principled cost-aware method for combining pairwise comparisons with direct observations. The proposed method can adapt to expert bias by devaluing incorrect information: when pairwise queries are cheap and reliable, it substantially improves sample-efficiency over observation-only BO, and when queries are costly or unreliable, it devalues them.

1 INTRODUCTION

Bayesian optimization (BO) is a sample-efficient framework for optimizing expensive-to-evaluate black-box functions (Shahriari et al., 2016; Frazier, 2018; Garnett, 2023). In many real-world applications—such as material design and molecule finding (e.g., Adachi, 2021; Cissé et al., 2024)—human experts have domain knowledge that is cheaper to elicit compared to directly evaluating the objective function. Hence, there is a growing body of *human-in-the-loop* (HITL) BO methods (e.g., A V et al., 2022; Adachi et al., 2024; Cissé et al., 2024; Hvarfner et al., 2024; Xu et al., 2024a) that utilize this domain knowledge to improve the cost-efficiency of BO.

Most of the HITL BO methods assume that a domain expert can quantify their knowledge, for instance by pinpointing query locations (A V et al., 2022; Gupta et al., 2023; Khoshvishkaie et al., 2023) or by specifying their prior beliefs about the location of the maximum as a probability distribution (Hvarfner et al., 2022; 2024; Cissé et al., 2024). However, human expertise is often tacit and qualitative (Kahneman & Tversky, 1979). Humans rank candidates more reliably than they attach calibrated numbers (Shah et al., 2014). Thus, pairwise comparisons are a natural type of query for knowledge elicitation. Many HITL BO methods also do not consider the possibility of erroneous expert feedback, but assume that the elicited knowledge is accurate. (see Mikkola et al., 2023).

To address these research gaps, we derive a principled probabilistic method for combining data from direct function evaluations and pairwise comparisons. To yield robustness against human error, we model the objective and the expert's mental model by correlated Gaussian processes; learning their correlation allows the method to devalue feedback from a misaligned expert. We then use a cost-aware one-step rule to choose between direct evaluations and pairwise comparisons. For evaluations, this is the knowledge gradient of our mixture posterior; for comparisons, we extend the scalar probit comparison update of Wu & Gardner (2026) to a multi-output objective and a distinct expert model.

We consider both single- and multi-objective BO. In the multi-objective case, we assume that the utility function is elicited or specified prior to the optimization process. We evaluate the proposed method, called Elicitation-Augmented Bayesian Optimization (EA-BO), on standard single- and multi-objective benchmarks. Across cost regimes, EA-BO approaches the convex hull of the single-source trajectories: it allocates the budget to the more cost-effective source at each phase of the optimization, outperforming evaluation-only BO when pairwise queries are cheap, and recovering the performance of standard BO when pairwise queries are costly, noisy or misaligned. EA-BO

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

substantially outperforms the earlier HITL BO method CoExBO (Adachi et al., 2024) that, like our work, assumes that the expert is able to express their domain knowledge via pairwise judgements.

To summarize, our main contributions are: **(i)** We derive a principled BO method for combining direct objective function evaluations with pairwise queries in a cost-aware fashion. Explicitly modelling the correlation between the objective function and the expert's mental model yields robustness against human error. **(ii)** We derive a cost-aware acquisition function with a closed-form posterior mean under the probit comparison likelihood, extending the pairwise knowledge gradient of Wu & Gardner (2026) to a multi-output objective under a known utility. **(iii)** Experiments on standard benchmark functions (Section 6) indicate that the proposed method traces the Pareto-optimal convex hull over the choice of information sources, thus substantially improving the cost-efficiency of BO in the regime where pairwise comparisons are cheaper than evaluations.

2 PRELIMINARIES

We consider the task of optimizing a vector-valued *objective function* $f : \mathcal{X} \to \mathbb{R}^m$ over a bounded *design space* $\mathcal{X} \subset \mathbb{R}^d$. We assume that f is an expensive-to-evaluate black box function, i.e., we do not know its functional form or gradients. Given a known *utility function* $U : \mathbb{R}^m \to \mathbb{R}$, the goal of the optimization process is to find

$$x^* = \arg\max_{x \in \mathcal{X}} U(f(x)),$$

i.e., the design with the highest utility. Standard single-objective BO is a special case with $m = 1$ and $U(y) = y$. When $m > 1$, the utility function U encodes the decision maker's priorities over the outcomes; throughout we take it to be linear, $U(y) = w^\top y$ with known weights $w \in \mathbb{R}^m_+$. We assume that there are two information sources for solving this optimization task: direct evaluations of the objective function and pairwise comparisons of designs.

Direct observations. The experimenter selects $x \in \mathcal{X}$, pays a fixed cost $c_{\text{eval}} \in \mathbb{R}_+$, and receives

$$y = f(x) + \varepsilon_{\text{eval}}, \qquad \varepsilon_{\text{eval}} \sim N(0, \Sigma_{\text{eval}}),$$

where $\Sigma_{\text{eval}} = \text{diag}(\sigma^2_{\text{eval},1}, \ldots, \sigma^2_{\text{eval},m})$ is the covariance matrix defining the observation noise.

Pairwise queries. We assume that there is a data generating process yielding comparisons of two competing designs $x, x' \in \mathcal{X}$; a human expert is the most plausible data source. The expert ranks the designs according to $U \circ g$, where $g : \mathcal{X} \to \mathbb{R}^m$ is the expert's mental model of the objective. When the expert is aligned, $g \equiv f$ and the ranking reduces to $U \circ f$. In a general case, g may differ from f; we model the two as coupled via a learnable correlation and defer the details to Section 3. Formally, the experimenter selects a tuple $(x, x') \in \mathcal{X}^2$, pays a cost $c_{\text{comp}} \in \mathbb{R}_+$, and receives a binary judgement $d \in \{0, 1\}$, where $d = 1$ indicates that x is ranked higher than x'. Using the probit link function, the likelihood can be written as

$$p(d = 1 \mid x, x', g) = \Phi\left(\frac{U(g(x)) - U(g(x'))}{\sqrt{2}\,\sigma_{\text{comp}}}\right),$$

where Φ is the CDF of $N(0, 1)$ and $\sigma_{\text{comp}} > 0$ is the noise parameter of the pairwise comparisons. We treat σ_{comp} as known. A judgement that contradicts the direct observations can be explained either by a noisy expert or by a mismatch between g and f, so σ_{comp} duplicates the role of the learned correlation, and at the data volumes of interest the two are not separable; fixing σ_{comp} makes the correlation the only channel for disagreement.

Notation. We denote direct function evaluations by $\mathcal{D}_{\text{eval}} = \{(x_i, y_i)\}_{i=1}^{n_{\text{eval}}}$, responses to pairwise queries by $\mathcal{D}_{\text{comp}} = \{((x_j, x'_j), d_j)\}_{j=1}^{n_{\text{comp}}}$, and the set of all observations by $\mathcal{D} = \{\mathcal{D}_{\text{eval}}, \mathcal{D}_{\text{comp}}\}$.

3 SURROGATE MODEL

We model the objective function and the expert's mental model jointly, so that pairwise comparisons can inform predictions of the objective according to the inferred alignment between them. We first describe the GP prior and then the variational inference algorithm used to update its posterior.

3.1 GP MODEL

To model the connection between the objective and the expert's mental model, we place a joint GP prior on f and g using a rank-2 linear model of coregionalization (LMC) (Álvarez et al., 2012). For the jth output, one latent GP contributes to both f_j and g_j, while a second independent latent GP representing the expert's bias contributes only to g_j. The latents share a kernel and are mixed as

$$\begin{pmatrix} f_j(x) \\ g_j(x) \end{pmatrix} = \begin{pmatrix} a_j & 0 \\ s_j\,\rho & s_j\sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} u_{j,0}(x) \\ u_{j,1}(x) \end{pmatrix}.$$

Hence, $\mathrm{Var}(f_j) = a_j^2$, $\mathrm{Var}(g_j) = s_j^2$ and $\mathrm{Corr}(f_j, g_j) = \rho$. Correlation ρ is shared across outputs. It is the link that allows information from the expert's mental model to affect the surrogate model of the objective function. When $\rho = 1$, the LMC collapses to $f \equiv g$ up to an output-specific positive rescaling. When $\rho = 0$, g is independent of f once the other model parameters are fixed, making comparisons uninformative about the objective. We denote the remaining GP parameters by θ; their priors and the fixed quantities are given in Appendix D.

We denote the joint posterior moments by $\mu_{\mathcal{D}}(x) := \mathbb{E}\left[(f(x), g(x)) \mid \mathcal{D}\right]$ and $K_{\mathcal{D}}(x, x') := \mathrm{Cov}\left[(f(x), g(x)), (f(x'), g(x')) \mid \mathcal{D}\right]$, dropping the subscript when the dataset is unambiguous, and use $\mu^f, \mu^g, K^{ff}, K^{fg}, K^{gg}$ for the corresponding blocks.

3.2 POSTERIOR INFERENCE

The joint posterior $p(\rho, f, g \mid \mathcal{D}, \theta)$ has no closed form, and we approximate it by sparse variational inference in two stages. We first fit θ by MAP from direct observations alone. Because their Gaussian likelihood depends only on f and the marginal prior on f does not depend on ρ, this fit is independent of ρ. We approximate ρ via R nodes ρ_r and prior masses π_r. The prior is strongly peaked around $\rho = 1$, but retains weight throughout. Details are in Appendix D. For each ρ_r we compute a variational posterior q_r. The posterior is then a mixture over q_r's. Following the standard sparse variational GP construction (Titsias, 2009; Hensman et al., 2013; 2015) extended to coregionalized outputs (Álvarez et al., 2012; Moreno-Muñoz et al., 2018), we introduce M inducing locations $Z \subset \mathcal{X}$ shared by all $2m$ latent processes, with inducing values $u_{j,l} := u_{j,l}(Z) \in \mathbb{R}^M$ and $u := \{u_{j,l}\}$. At each node, q_r factorizes across outputs but retains one joint $2M$-dimensional Gaussian over both latents of each output, so that the shared and the bias latent stay coupled:

$$q_r(u) = \prod_{j=1}^{m} N\left(\begin{pmatrix} u_{j,0} \\ u_{j,1} \end{pmatrix} \middle| \, m_j^{(r)}, S_j^{(r)} \right),$$

with means $m_j^{(r)} \in \mathbb{R}^{2M}$ and covariances $S_j^{(r)} \in \mathbb{R}^{2M \times 2M}$ as the variational parameters.

For each ρ_r, the dual-likelihood structure of our dataset yields a decomposable ELBO with separate observation, comparison, and KL contributions (Wu et al., 2025; Moreno-Muñoz et al., 2018):

$$\begin{aligned} \mathcal{L}_r(q_r) = &\sum_{i=1}^{n_{\mathrm{eval}}} \mathbb{E}_{q_r(f(x_i))}[\log p(y_i \mid x_i, f)] + \sum_{j=1}^{n_{\mathrm{comp}}} \mathbb{E}_{q_r(g(x_j), g(x'_j))}\left[\log p(d_j \mid x_j, x'_j, g)\right] \\ &- \sum_{k=1}^{m} \mathrm{KL}(q_r(u_{k,0}, u_{k,1}) \,\|\, p(u_{k,0}, u_{k,1} \mid \theta)) \,. \end{aligned}$$

Evaluation and KL terms admit standard Gaussian closed forms (Appendix C). For linear U, the comparison term depends on q_r only through the scalar utility difference $\Delta_j = U(g(x_j)) - U(g(x'_j))$, which is Gaussian under q_r. Its expected log-probit likelihood is evaluated by one-dimensional Gauss–Hermite quadrature. We optimize q_r at fixed θ and ρ_r using damped site updates based on conjugate-computation VI (Khan & Lin, 2017); see Appendix B for details.

Since $\mathcal{L}_r(q_r)$ is a lower bound on the log evidence at node ρ_r, we approximate its posterior weight by $\omega_r \propto \pi_r \exp(\mathcal{L}_r(q_r))$. New observations update the weights through the node ELBOs. As Appendix D shows, comparisons alone carry no information about ρ, and thus direct observations are needed to calibrate it.

4 ACQUISITION FUNCTION

At each step, we choose between a direct evaluation and a pairwise query by their expected gain in the final recommendation per unit cost. We first define this decision problem and then derive the one-step gains under the surrogate model of Section 3.

4.1 DECISION PROBLEM

The optimization proceeds under a fixed total budget $\Lambda \in \mathbb{R}_+$. We select the recommended design after the budget is exhausted as

$$\hat{x} = \arg\max_{x \in \mathcal{X}} \mathbb{E}_f \left[U(f(x)) \mid \mathcal{D} \right].$$

Hence, the utility of the dataset $\mathcal{D}$ is $u(\mathcal{D}) = \max_{x \in \mathcal{X}} \mathbb{E}_f \left[U(f(x)) \,\middle|\, \mathcal{D} \right]$, where the expectation is over the uncertainty of the function approximation.

At each iteration, the experimenter either directly evaluates the objective function at $x \in \mathcal{X}$ or performs a pairwise query for the pair $(x, x') \in \mathcal{X}^2$, i.e., selects an action from the action space $\mathcal{A} = \mathcal{X} \cup \mathcal{X}^2$, and pays the cost c_{eval} or c_{comp} depending on the type of action chosen. This decision problem is an instance of cost-aware multi-information-source optimization (Swersky et al., 2013; Kandasamy et al., 2017; Poloczek et al., 2017). Exact optimization of the full action sequence is intractable; we adopt the standard myopic one-step lookahead (Frazier, 2018).

After selecting action $a \in \mathcal{A}$ and receiving result r ($r \in \mathbb{R}^m$ or $r \in \{0, 1\}$ depending on whether the action was a direct observation or a pairwise comparison, respectively) the updated dataset is $\mathcal{D}' = \mathcal{D} \cup \{(a, r)\}$. The *value of information* (VoI)

$$\alpha(a; \mathcal{D}) = \mathbb{E}_r \left[u(\mathcal{D}') \,\middle|\, a, \mathcal{D} \right] - u(\mathcal{D})$$

of action a is the expected one-step gain obtained by performing it.

We select a that maximizes cost-normalized VoI (Poloczek et al., 2017; Ungredda & Branke, 2023):

$$a^* \in \arg\max_{a \in \mathcal{A}} \frac{\alpha(a; \mathcal{D})}{c(a)}.$$

For an evaluation $a = x \in \mathcal{X}$, the value and cost are $\alpha_{\text{eval}}(x; \mathcal{D})$ and c_{eval}; for a query $a = (x, x') \in \mathcal{X}^2$, they are $\alpha_{\text{comp}}((x, x'); \mathcal{D})$ and c_{comp}. We derive α_{eval} and α_{comp} in Sections 4.2 and 4.3.

Algorithm 1 describes the optimization loop of EA-BO[1]. At each step we fit the posterior, optimize both VoIs to obtain their respective maxima $\alpha^\star_{\text{eval}}$ and $\alpha^\star_{\text{comp}}$, and then select the action with the highest cost-normalized VoI[2]. The optimization loop terminates when the remaining budget falls below $\min(c_{\text{eval}}, c_{\text{comp}})$, and the final recommendation is $\hat{x} = \arg\max_{x \in \mathcal{X}} \mathbb{E}_f \left[U(f(x)) \mid \mathcal{D}_{\text{final}} \right]$.

4.2 VALUE-OF-INFORMATION FOR A DIRECT OBSERVATION

The VoI for a function evaluation, denoted by α_{eval}, is the expected one-step gain in dataset utility after observing y at $x \in \mathcal{X}$. Because the observation likelihood is Gaussian, the posterior of each component is available in closed form after observing y. For linear U, α_{eval} reduces to knowledge gradient (Frazier et al., 2009; Wu & Frazier, 2016) of a mixture posterior.

We adopt the one-shot reformulation (Balandat et al., 2020; Wu et al., 2020). We approximate the expectation over possible observations using fixed quadrature nodes or base draws; for each candidate evaluation point x, these give outcomes y_s and weights v_s. For each outcome, x'_s is a

[1] With the number of inducing points M, and R nodes, the per-iteration cost of EA-BO scales linearly in the number of observations (n = observations + comparisons) and in the number of outputs m: each of the R fits cost $O(m(nM^2 + M^3))$ per fixed-point iteration. The acquisition is independent of n, linear in m and d

[2] Since at least two evaluations are needed to learn ρ effectively, there are two exceptions: (1) the first action is an observation placed in the geometric centre of the domain; (2) if after n_{cal} comparisons no second evaluations has occurred, we force an observation that is chosen to maximize the posterior mutual information about ρ.

Algorithm 1 Elicitation-Augmented Bayesian Optimization (EA-BO).

Input: design space $\mathcal{X} \subset \mathbb{R}^d$, utility U, budget Λ, costs $c_{\text{eval}}, c_{\text{comp}}$, initial design $\mathcal{D}_0 = \emptyset$, hyperparameter priors, calibration trigger n_{cal}.
Output: recommended design $\hat{x} \in \mathcal{X}$.
Fit $\hat{\theta}_0$ and the mixture $\{(\omega_r, q_r)\}_0$ on $\mathcal{D}_0$ (Section 3.2). Set $t \leftarrow 0$, $\Lambda_t \leftarrow \Lambda$.
while $\Lambda_t \geq \min(c_{\text{eval}}, c_{\text{comp}})$ **do**
  if $t = 0$ **then**
    $a_t^* \leftarrow x_{\text{c}}$, the centre of $\mathcal{X}$. ▷ anchor evaluation
  else if $n_{\text{eval}}(\mathcal{D}_t) = 1$ and $n_{\text{comp}}(\mathcal{D}_t) \geq n_{\text{cal}}$ and $\Lambda_t \geq c_{\text{eval}}$ **then**
    $a_t^* \leftarrow \arg\max_{x \in \mathcal{X}} I(\rho;\, y(x) \mid \mathcal{D}_t)$. ▷ calibration evaluation
  else
    $x_t^* \leftarrow \arg\max_x \widehat{\alpha}_{\text{eval}}(x; \mathcal{D}_t)$.
    $(x_a, x_b)_t^* \leftarrow \arg\max_{x_a, x_b} \widehat{\alpha}_{\text{comp}}((x_a, x_b); \mathcal{D}_t)$.
    Restrict to actions affordable under Λ_t.
    $a_t^* \leftarrow \arg\max_i \widehat{\alpha}_i^\star / c_i$ over the affordable actions.
  if $a_t^* \in \mathcal{X}$, i.e. an evaluation **then**
    Observe $y_t \sim N(f(a_t^*), \Sigma_{\text{eval}})$. Set $\mathcal{D}_{t+1} \leftarrow \mathcal{D}_t \cup \{(a_t^*, y_t)\}$, $\Lambda_{t+1} \leftarrow \Lambda_t - c_{\text{eval}}$.
  else
    Query expert on $a_t^* = (x_a, x_b)_t^*$. Set $\mathcal{D}_{t+1} \leftarrow \mathcal{D}_t \cup \{(a_t^*, d_t)\}$, $\Lambda_{t+1} \leftarrow \Lambda_t - c_{\text{comp}}$.
  Refit $\hat{\theta}_{t+1}$ and $\{(\omega_r, q_r)\}_{t+1}$ on $\mathcal{D}_{t+1}$, each q_r warm-started from its own previous solution.
  $t \leftarrow t + 1$.
return $\hat{x} = \arg\max_{x \in \mathcal{X}} \mathbb{E}[U(f(x)) \mid \mathcal{D}_t]$.

candidate for the best design after observing y_s. We jointly optimize the evaluation point x and these recommendation points:

$$\widehat{\alpha}_{\text{eval}}(x, x'_1, \ldots, x'_S; \mathcal{D}) = \sum_{s=1}^{S} v_s \, \mathbb{E}_{\mathcal{D} \cup \{(x, y_s)\}}[U(f(x'_s))] - u(\mathcal{D}).$$

We evaluate the expectation over possible observations numerically; see Appendix C.3 for details.

4.3 Value-of-information for pairwise comparisons

The VoI of a pairwise comparison measures the expected one-step gain from querying the pair (x_a, x_b). The outcome $d \in \{0, 1\}$ has posterior predictive distribution

$$p(d = 1 \mid x_a, x_b, \mathcal{D}) = \sum_{r=1}^{R} \omega_r \, \Phi\left(\frac{\mu_{\Delta_g}^{(r)}}{\sqrt{\sigma_{\Delta_g}^{2\,(r)} + 2\sigma_{\text{comp}}^2}} \right),$$

where $\mu_{\Delta_g}^r$ and $\sigma_{\Delta_g}^{2\,(r)}$ are the mean and variance of $\Delta_g = U(g(x_a)) - U(g(x_b))$ under q_r. The comparison VoI is a discrete two-term expectation over the binary outcome:

$$\alpha_{\text{comp}}((x_a, x_b); \mathcal{D}) = \sum_{d \in \{0,1\}} p(d \mid x_a, x_b, \mathcal{D}) \max_{x \in \mathcal{X}} \mathbb{E}_{\mathcal{D}'}[U(f(x))] - u(\mathcal{D}).$$

Thus, no Monte Carlo approximation is required for the expectation over two possible outcomes.

The key quantity is the posterior mean of f after observing d. Within q_r, for linear U, Δ_g is Gaussian since it is a linear functional of the jointly Gaussian $(g(x_a), g(x_b))$ under q_r. Wu & Gardner (2026) derive a closed-form conditional mean for probit comparisons over a scalar f. We extend their result to a multi-output objective under a known utility U and to a misaligned expert whose mental model g couples to f through the LMC. When $d = 1$, the fantasy posterior component mean is

$$\mu^{f,(r)}_{\mathcal{D} \cup \{((x_a, x_b), 1)\}}(x) = \mu^{f,(r)}(x) + \frac{\phi(\tau_r)}{\Phi(\tau_r)} \cdot \frac{\text{Cov}_r\left[f(x), \Delta_g\right]}{\nu_r},$$

where $\nu_r := \sqrt{\sigma_{\Delta_g}^{2,(r)} + 2\sigma_{\text{comp}}^2}$ and $\tau_r := \mu_{\Delta_g}^r / \nu_r$. Component weights update to $\omega'_r \propto \omega_r \, \Phi(\tau_r)$. For $d = 0$, the final term $-\phi(\tau_r)\text{Cov}_r\left[f(x), \Delta_g\right] / (\Phi(-\tau_r)\nu_r)$ has the opposite sign and inverse

Mills ratio. The proof is deferred to Appendix A. The cross-block covariance $\mathrm{Cov}_r\left[f(x), \Delta_g\right]$ couples the comparison evidence on g to the f-posterior via the LMC cross-block K^{fg}, and reduces to $K^{fg}(x, x_a)w - K^{fg}(x, x_b)w$ for linear $U(y) = w^\top y$. Setting m=1, R=1, $\rho = 1$, $a_1 = s_1$ (so that $g \equiv f$), U to the identity, and $2\sigma^2_{\mathrm{comp}} = 1$ recovers the scalar result of Wu & Gardner (2026).

As in the case of VoI for the direct observations, the inner maximum is handled by one-shot reformulation, here with one recommendation point per label (Balandat et al., 2020; Wu et al., 2020):

$$\widehat{\alpha}_{\mathrm{comp}}(x_a, x_b, x'_0, x'_1; \mathcal{D}) = \sum_{d \in \{0,1\}} p(d \mid x_a, x_b, \mathcal{D})\, \mathbb{E}_{\mathcal{D} \cup \{((x_a, x_b), d)\}}[U(f(x'_d))] - u(\mathcal{D}).$$

Both $\widehat{\alpha}_{\mathrm{eval}}$ and $\widehat{\alpha}_{\mathrm{comp}}$ are optimized by screening on a set of scrambled-Sobol candidates and recommendation points, and refining the best jointly with their recommendation points via L-BFGS-B. The action with the largest cost-normalized value is then taken.

5 Related Work

Human-in-the-loop BO. HITL BO methods typically assume that the expert can quantify their domain knowledge explicitly. An approach inspired by classic Bayesian inference requires that the expert specifies their prior beliefs about the location of the maximum of the objective function as a probability distribution (Ramachandran et al., 2020; Souza et al., 2021; Hvarfner et al., 2022; 2024; Cissé et al., 2024; Guay-Hottin et al., 2025). In the human-AI collaborative approach the expert pinpoints potential locations for the next objective function evaluation (A V et al., 2022; Gupta et al., 2023; Khoshvishkaie et al., 2023). In contrast, we assume that the domain knowledge of the expert is tacit, and thus perform the knowledge elicitation implicitly via pairwise queries. An earlier method that uses pairwise queries and is robust against human error is CoExBO (Adachi et al., 2024). The main differences to EA-BO are that CoExBO uses pairwise queries to directly select the location of the next function evaluation, selects the locations of the pairwise queries at random, and does not account for the query costs. Xu et al. (2024a) also utilize tacit domain knowledge of a human expert, but, in contrast to this work, assume that the expert accepts or rejects designs.

Interactive multi-objective BO. Interactive multi-objective BO methods (Astudillo & Frazier, 2020; Ungredda et al., 2021; Lin et al., 2022; Ozaki et al., 2024) utilize pairwise queries. However, they compare two points in the objective space $\mathbb{R}^m$ and use pairwise queries to learn the latent utility function U. In contrast, EA-BO compares two points in the design space $\mathbb{R}^d$, treats expert's judgements as noisy evidence about the behaviour of the objective function f (via expert's mental model g that is assumed to be correlated with f), and assumes that U is either specified or elicited prior to the optimization process. Composite-function BO (Astudillo & Frazier, 2019) assumes a known utility structure, but uses only direct function evaluations as a data source.

Preferential BO. Preferential BO methods (Chu & Ghahramani, 2005; González et al., 2017; Siivola et al., 2021; Mikkola et al., 2020; Benavoli et al., 2021; Takeno et al., 2023; Astudillo et al., 2023; Xu et al., 2024b) utilize pairwise queries in the design space $\mathbb{R}^d$ to learn the latent preference function $g : \mathbb{R}^d \to \mathbb{R}$. However, in preferential BO the task is to optimize g, whereas we consider the task of optimizing the observable objective function f (or $U \circ f$ when f is vector-valued), and use pairwise queries as an additional information source.

Multi-fidelity BO. Multi-fidelity BO methods (Huang et al., 2006; Kandasamy et al., 2016; Sen et al., 2018; Song et al., 2019) use additional cheaper but potentially more unreliable information sources to improve the cost-efficiency of BO. However, these information sources are real-valued, whereas EA-BO combines binary comparisons with continuous observations. In addition, multi-fidelity methods typically, with the exception of the work by Mikkola et al. (2023), assume that the fidelity levels of the information sources are known before the optimization process.

Mixed likelihood GPs. Wu et al. (2025) update a GP-based surrogate model using two different information sources: pairwise queries and Likert-scale data. In contrast, EA-BO combines pairwise queries with direct function evaluations and has two distinct correlated surrogate models for the information sources instead of using both information sources to directly update the same GP.

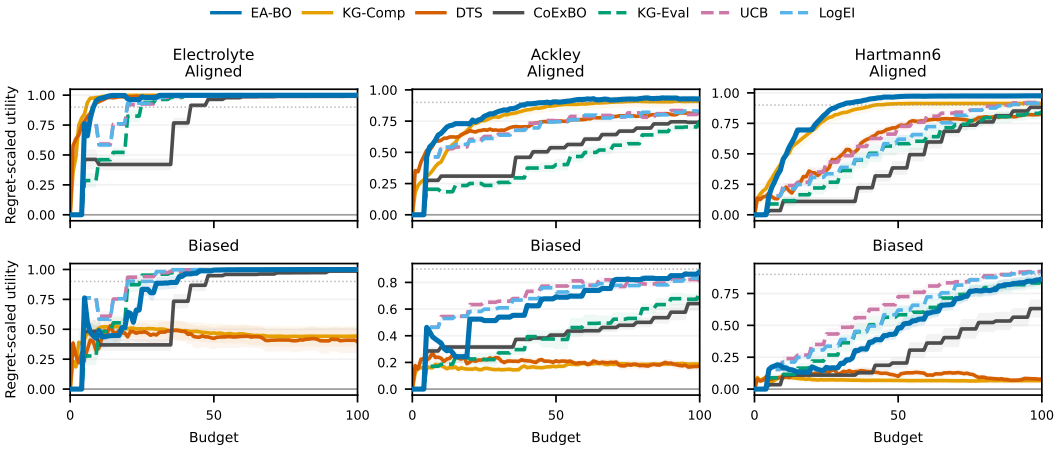


Figure 1: Stop-now recommendation utility on single-objective benchmarks under aligned (top) and biased (bottom) experts. Curves are means over 20 paired repetitions and bands are ± 1 standard error. Solid curves use comparisons and dashed curves are evaluation-only. When the expert is aligned, EA-BO finds good designs faster than the evaluation-only methods; on Ackley and Hartmann6 it also outperforms the comparison-only methods. When the expert is unreliable, the comparison-only methods fail, while EA-BO is able to identify biased advice and recover from it. EA-BO substantially outperforms the HITL BO method CoExBO in both cases.

6 Experimental results

Test problems. We measure the performance of the proposed method using standard benchmark functions Ackley (d=5), Hartmann6 (d=6), VLMOP2 (d=2, m=2), Branin–Currin (d=2, m=2), DTLZ2 (d=8, m=3) and WFG4 (d=10, m=5). In addition, we use the benchmark Electrolyte (d=3) by Adachi et al. (2024) that is based on real data from an electrolyte-design problem. To account for higher dimension, on DTLZ2 we use a calibration trigger threshold n_{cal} of 20 comparisons, on WFG4 30, and 10 on the rest. The budget is $\Lambda = 100$. Observation costs are $c_{\text{eval}} = 5$ and $c_{\text{comp}} = 1$. Noise levels are $\sigma_{\text{eval}} = 0.05$ and $\sigma_{\text{comp}} = 0.1$.

Utility function and expert model. For problems with multiple outputs, we draw positive linear weights from $\text{Dirichlet}(1, \ldots, 1)$, one per seed but matched across different methods. The weights are rescaled so that the utility has unit standard deviation. All outputs are standardized individually. All figures show the utility of the current recommendation $\hat{x}_t$ normalized to $[0, 1]$ between the worst and the best design in the domain, $\big(U(f(\hat{x}_t)) - U_{\min}\big)/\big(U(f(x^*)) - U_{\min}\big)$ with $U_{\min} = \min_x U(f(x))$; before the first observation it is set to 0. We ran each benchmark under two scenarios: an *aligned* (ρ=1) and a *biased* (ρ=0) expert. The mental model g of the biased expert is a random-feature field with a RBF length scale 0.5, independent across outputs and scaled so that its utility has unit marginal standard deviation.

Baselines. We compare EA-BO to single-source methods that use the same surrogate model: evaluation-only LogEI (Ament et al., 2023), UCB (Srinivas et al., 2010) and KG-Eval; and comparison-only KG-Comp and dueling Thompson sampling (DTS). KG-Eval and KG-Comp are the single-source acquisition functions of EA-BO described in Sections 4.2 and 4.3, respectively. On the scalar-output benchmarks, we also compare to the HITL BO method CoExBO (Adachi et al., 2024).[3] Its current implementation does not support vector-valued objectives.

6.1 Results

Aligned expert. EA-BO exploits cheap pairwise comparisons to quickly identify the promising regions of the design space (Figures 1 and 2, top rows). Consequently, it is significantly more cost-efficient than the evaluation-only methods. Even though EA-BO has to use a part of its budget to

[3] We allocate for CoExBO 20 comparisons and 2 evaluations for pre-training according to the ratio used in the original article and deduct the pretraining cost from its budget.

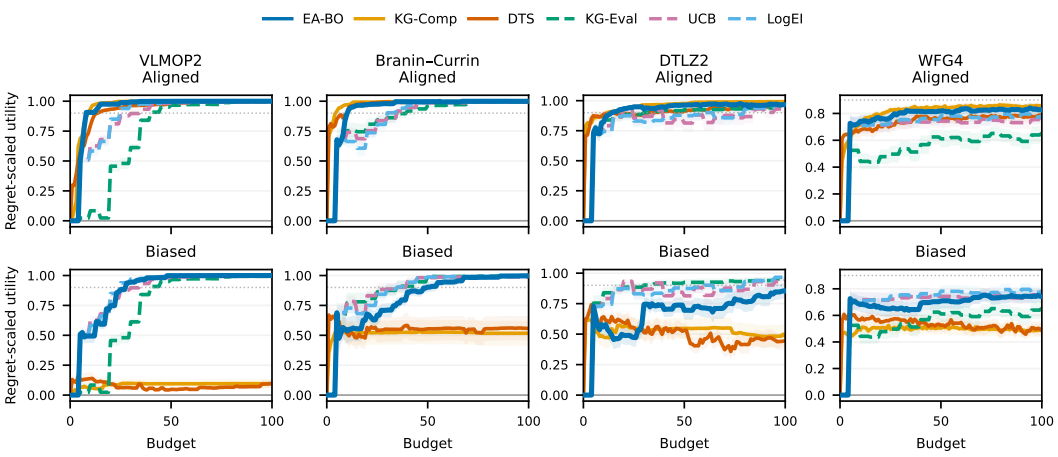


Figure 2: Stop-now recommendation utility on four multi-objective benchmarks (columns) under aligned (top) and biased (bottom) experts. Curves are means over 20 paired repetitions and bands are ± 1 standard error. Solid curves use comparisons and dashed curves are evaluation-only. When the expert is aligned, EA-BO is significantly more cost-efficient than the evaluation-only methods and matches the performance of the strongest comparison-only method KG-Comp. When the expert is unreliable, the comparison-only methods fail, but EA-BO is able to recover from biased advice.

assess the reliability of the expert, it retains the early performance advantage of comparison-only methods: EA-BO outperforms the strongest comparison-only method KG-Comp on Ackley and Hartmann6 test problems, and matches its performance on the other data sets. EA-BO substantially outperforms HITL BO method CoExBO on all the test problems.

Biased expert. Comparison-only methods KG-Comp and DTS fail since they rely purely on biased comparisons (Figures 1 and 2, bottom rows). In contrast, EA-BO compares expert's judgments with evidence from direct evaluations of the objective function. When these data sources disagree, the posterior mean of ρ falls, comparisons lose value, and the policy returns to direct evaluations. Thus, after a small initial plateau required to identify biased advice, EA-BO is able to match the progress of the evaluation-only methods: EA-BO reaches a comparable final utility at budget exhaustion to the evaluation-only methods on all the test problems except DTLZ2. HITL BO method CoExBO is also able to recover from biased advice, but EA-BO again outperforms it on all the test problems.

6.2 Ablations

Source allocation. We study the optimization performance (we summarize the optimization performance by the average area under curve [AUC] of the optimization trajectories) of EA-BO and its source allocation between evaluations and comparisons as a function of the cost and the reliability of the information sources. In summary, EA-BO is able to adapt to the reliability of the information sources in a cost-aware fashion by favoring cheap and reliable sources, and its performance is relatively robust w.r.t. noise and human error. When evaluations and comparisons are equally costly, EA-BO uses more accurate evaluations, and shifts its allocation towards comparisons as the evaluation cost increases (Figure 3, bottomleft). By exploiting cheaper comparisons, EA-BO is able to maintain a solid optimization performance even when the evaluation cost increases tenfold (Figure 3, topleft; Table 7). When the reliability of the expert (ρ) increases, a higher proportion of the budget is allocated to comparisons (Figure 3, bottommiddle). On the effect of ρ to the optimization performance (Figure 3, bottommiddle; Table 6), there is a slight non-monotonicity: the results indicate that identifying a moderate ρ is harder than a complete or zero correlation, reflected in a higher standard deviation. When the comparison noise increases (Figure 3, bottomright), EA-BO relies more on evaluations instead of comparisons. A modest drop in the optimization performance (Figure 3, topright; Table 8) reflects the lack of a reliable and cheap information source.

Sensitivity analysis. We study the EA-BO's sensitivity to changes in priors and methodology. The figures and tables can be found in Appendix F. We assess the cost of the robustness (in the case when

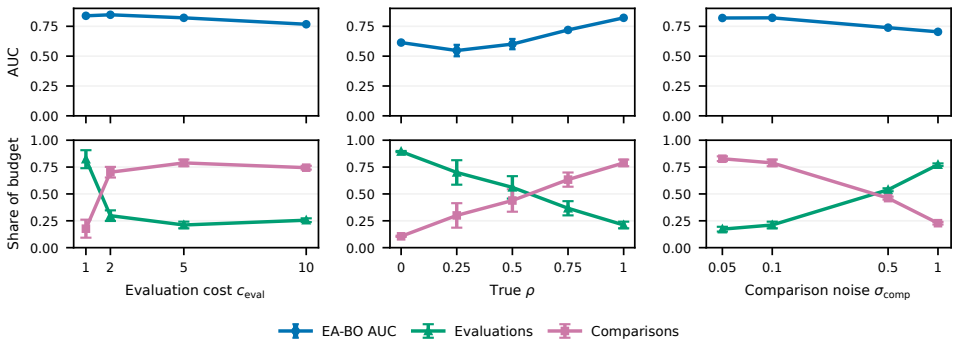


Figure 3: Ablation study on Ackley test problem. Optimization performance measured on the average AUC of the optimization trajectories (top), and source allocation measured by the shares of budget spent on evaluations and on comparisons (bottom) as a function of the evaluation cost c_{eval} (left), the expert reliability ρ (middle), and the comparison noise σ_{comp} (right). The defaults for the non-varied parameters are $c_{\text{eval}} = 5$, $\rho = 1$, and $\sigma_{\text{comp}} = 0.1$. Points are averages and error bars are ± 1 standard errors over nine repetitions. As the evaluation cost rises (left), EA-BO utilizes more comparisons, while retaining most of its optimization performance. When the expert reliability improves (middle), EA-BO utilizes more comparisons; the effect of ρ to the optimization performance is slightly non-monotonic. When the comparison noise increases (right), EA-BO favours evaluations, and its AUC dips modestly.

the expert is reliable) by comparing EA-BO with the default prior for the correlation parameter ρ to the version where ρ is fixed to 1 (i.e., the expert is assumed to be 100% reliable); see Figure 5. The optimization performance of the default version where ρ is learned is almost equal to the version with the fixed ρ. This indicates that the cost of robustness is minimal when the expert is reliable. We study the effect of computing the posterior predictive distribution by integrating over the posterior of the parameter ρ instead of using a point estimate with different true values of ρ (Table 10). The fully Bayesian approach outperforms the MAP estimate in most of the cases. We study the impact of the choice of the prior for ρ (Table 11). Specifically, we compare the default prior against a diffuse prior and a weaker informative prior towards alignment. The default prior performs best on the aligned problem, while giving up some AUC but still achieving the best final utility on the biased problem.

7 Discussion

We considered the task of eliciting tacit domain knowledge of the expert to improve the cost-efficiency of BO, and derived a principled probabilistic method for solving this task in the Bayesian decision-theoretic framework. First, we defined a joint model for the objective function (informed by direct observations) and the mental model of the expert (informed by pairwise comparisons). We then derived the cost-aware one-step optimal policy under this model. Experiments on standard benchmark functions indicate that this policy enables combining the strengths of the single-source strategies by exploiting cheap pairwise queries in the early stages of the optimization to quickly identify promising regions of the design space, and then using more accurate direct objective function evaluations to find the maximum from these regions. Further, our experiments verify that explicitly modelling the correlation between the mental model of the expert and the surrogate model of the objective function enables identifying and devaluing incorrect information, making the proposed method robust against human error.

Limitations and future work. Our model represents the relationship between the objective function and the expert's mental model through one global ρ shared across outputs. Output-specific or spatially varying coregionalization could represent experts whose reliability differs between the components of the objective function or regions of the design space, while allowing negative correlation would enable modelling adversarial feedback. We also assume that the utility U is elicited before optimization, as in *a priori* (see, e.g., Branke et al., 2008, p. 18-21) multi-objective optimization. We leave extending the method by learning U in an interactive fashion during the optimization process as future work.

AI USE STATEMENT

Beyond the purposes stated below, we have not used AI in this work. We have used AI for debugging and to generate code which processes results for visualization. The authors have read the generated code to make sure it matches what is described in the paper and that the generated figures match the raw data. We have used AI to identify relevant literature and information. All sources found by AI have been read by the authors to verify the contents. We take responsibility for the final content of this work, including text, claims or artefacts produced with the aid of generative AI.

ETHICS STATEMENT

The work has no immediate ethical implications.

REPRODUCIBILITY STATEMENT

The appendix contains all relevant parameter and hyper-parameter values to get the displayed results using the algorithm presented. Source code for exact implementation is available in the supplementary material. A proof is provided in the appendix for the closed form posterior.

REFERENCES

Arun Kumar A V, Santu Rana, Alistair Shilton, and Svetha Venkatesh. Human-ai collaborative bayesian optimisation. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (eds.), *Advances in Neural Information Processing Systems*, volume 35, pp. 16233–16245. Curran Associates, Inc., 2022.

Masaki Adachi. High-dimensional discrete bayesian optimization with self-supervised representation learning for data-efficient materials exploration. In *NeurIPS 2021 AI for Science Workshop*, 2021.

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 *International Conference on Artificial Intelligence and Statistics*, pp. 505–513. PMLR, 2024.

Mauricio A. Álvarez, Lorenzo Rosasco, and Neil D. Lawrence. Kernels for vector-valued functions: A review. *Foundations and Trends in Machine Learning*, 4(3):195–266, 2012.

Sebastian Ament, Samuel Daulton, David Eriksson, Maximilian Balandat, and Eytan Bakshy. Unexpected improvements to expected improvement for Bayesian optimization. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

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

Raul Astudillo and Peter I. Frazier. Bayesian optimization of composite functions. In *Proceedings of the 36th International Conference on Machine Learning (ICML)*, volume 97, pp. 354–363. PMLR, 2019.

Raul Astudillo, Zhiyuan Jerry Lin, Eytan Bakshy, and Peter Frazier. qeubo: A decision-theoretic acquisition function for preferential bayesian optimization. In *International conference on artificial intelligence and statistics*, pp. 1093–1114. PMLR, 2023.

Adelchi Azzalini. *The Skew-Normal and Related Families*. Institute of Mathematical Statistics Monographs. Cambridge University Press, 2013.

Maximilian Balandat, Brian Karrer, Daniel R. Jiang, Samuel Daulton, Benjamin Letham, Andrew Gordon Wilson, and Eytan Bakshy. BoTorch: A framework for efficient Monte-Carlo Bayesian optimization. In *Advances in Neural Information Processing Systems*, volume 33, 2020.

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)*, pp. 1842–1850, 2021.

J. Branke, K. Deb, Kaisa Miettinen, and R. Slowinski. *Multiobjective optimization: Interactive and evolutionary approaches*. Springer, Germany, 2008.

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

Abdoulatif Cissé, Xenophon Evangelopoulos, Sam Carruthers, Vladimir V. Gusev, and Andrew I. Cooper. Hypbo: Accelerating black-box scientific experiments using experts' hypotheses. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, pp. 3881–3889, 2024. doi: 10.24963/ijcai.2024/429.

Peter I. Frazier. A tutorial on Bayesian optimization. *arXiv preprint arXiv:1807.02811*, 2018.

Peter I. Frazier, Warren B. Powell, and Savas Dayanik. The knowledge-gradient policy for correlated normal beliefs. *INFORMS Journal on Computing*, 21(4):599–613, 2009.

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

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

Rose Guay-Hottin, Lison Kardassevitch, Hugo Pham, Guillaume Lajoie, and Marco Bonizzato. Robust prior-biased acquisition function for human-in-the-loop Bayesian optimization. *Knowledge-Based Systems*, pp. 113039, 2025.

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

James Hensman, Nicolò Fusi, and Neil D. Lawrence. Gaussian processes for big data. In *Proceedings of the Twenty-Ninth Conference on Uncertainty in Artificial Intelligence*, 2013.

James Hensman, Alexander G. de G. Matthews, and Zoubin Ghahramani. Scalable variational Gaussian process classification. In *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics*, volume 38 of *Proceedings of Machine Learning Research*. PMLR, 2015.

Deng Huang, Theodore T Allen, William I Notz, and R Allen Miller. Sequential kriging optimization using multiple-fidelity evaluations. *Structural and Multidisciplinary Optimization*, 32(5):369–382, 2006.

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 *International Conference on Learning Representations*, 2024.

Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291, 1979.

Kirthevasan Kandasamy, Gautam Dasarathy, Junier B Oliva, Jeff Schneider, and Barnabás Póczos. Gaussian process bandit optimisation with multi-fidelity evaluations. *Advances in neural information processing systems*, 29, 2016.

Kirthevasan Kandasamy, Gautam Dasarathy, Jeff Schneider, and Barnabás Póczos. Multi-fidelity Bayesian optimisation with continuous approximations. In *Proceedings of the 34th International Conference on Machine Learning (ICML)*, pp. 1799–1808, 2017.

Mohammad Emtiyaz Khan and Wu Lin. Conjugate-computation variational inference: Converting variational inference in non-conjugate models to inferences in conjugate models. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 54 of *Proceedings of Machine Learning Research*, pp. 878–887, 2017.

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

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

Petrus Mikkola, Milica Todorović, Jari Järvi, Patrick Rinke, and Samuel Kaski. Projective preferential Bayesian optimization. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*, pp. 6884–6892, 2020.

Petrus Mikkola, Julien Martinelli, Louis Filstroff, and Samuel Kaski. Multi-fidelity Bayesian optimization with unreliable information sources. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 206 of *Proceedings of Machine Learning Research*, pp. 7425–7454. PMLR, 2023.

Pablo Moreno-Muñoz, Antonio Artés-Rodríguez, and Mauricio A. Álvarez. Heterogeneous multi-output Gaussian process prediction. In *Advances in Neural Information Processing Systems*, volume 31, 2018.

Manfred Opper and Cédric Archambeau. The variational Gaussian approximation revisited. *Neural Computation*, 21(3):786–792, 2009. doi: 10.1162/neco.2008.08-07-592.

Ryota Ozaki, Kazuki Ishikawa, Youhei Kanzaki, Shion Takeno, Ichiro Takeuchi, and Masayuki Karasuyama. Multi-objective Bayesian optimization with active preference learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 14490–14498, 2024.

Matthias Poloczek, Jialei Wang, and Peter I. Frazier. Multi-information source optimization. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

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

Alexandra M. Schmidt and Alan E. Gelfand. A Bayesian coregionalization approach for multivariate pollutant data. *Journal of Geophysical Research: Atmospheres*, 108(D24):8783, 2003.

Rajat Sen, Kirthevasan Kandasamy, and Sanjay Shakkottai. Multi-fidelity black-box optimization with hierarchical partitions. In *International conference on machine learning*, pp. 4538–4547. PMLR, 2018.

Nihar B Shah, Sivaraman Balakrishnan, Joseph Bradley, Abhay Parekh, Kannan Ramchandran, and Martin Wainwright. When is it better to compare than to score? *arXiv preprint arXiv:1406.6618*, 2014.

Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P. Adams, and Nando de Freitas. Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE*, 104(1): 148–175, 2016.

Eero Siivola, Akash Kumar Dhaka, Michael Riis Andersen, Javier González, Pablo García Moreno, and Aki Vehtari. Preferential batch Bayesian optimization. In *IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, 2021.

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

Artur Souza, Luigi Nardi, Leonardo B Oliveira, Kunle Olukotun, Marius Lindauer, and Frank Hutter. Bayesian optimization with a prior for the optimum. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 265–296. Springer, 2021.

Niranjan Srinivas, Andreas Krause, Sham M. Kakade, and Matthias W. Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. In *Proceedings of the 27th International Conference on Machine Learning*, pp. 1015–1022, 2010.

Kevin Swersky, Jasper Snoek, and Ryan P. Adams. Multi-task Bayesian optimization. In *Advances in Neural Information Processing Systems*, volume 26, 2013.

Shion Takeno, Masahiro Nomura, and Masayuki Karasuyama. Towards practical preferential Bayesian optimization with skew Gaussian processes. In *Proceedings of the 40th International Conference on Machine Learning (ICML)*, pp. 33516–33533, 2023.

Michalis K. Titsias. Variational learning of inducing variables in sparse Gaussian processes. In *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics*, volume 5 of *Proceedings of Machine Learning Research*. PMLR, 2009.

Juan Ungredda and Jürgen Branke. When to elicit preferences in multi-objective Bayesian optimization. In *Proceedings of the Companion Conference on Genetic and Evolutionary Computation (GECCO '23 Companion)*. ACM, 2023.

Juan Ungredda, Juergen Branke, Mariapia Marchi, and Teresa Montrone. One step preference elicitation in multi-objective bayesian optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pp. 195–196, 2021.

Homer F. Walker and Peng Ni. Anderson acceleration for fixed-point iterations. *SIAM Journal on Numerical Analysis*, 49(4):1715–1735, 2011. doi: 10.1137/10078356X.

Jian Wu and Peter I. Frazier. The parallel knowledge gradient method for batch Bayesian optimization. In *Advances in Neural Information Processing Systems*, volume 29, 2016.

Jian Wu, Saul Toscano-Palmerin, Peter I. Frazier, and Andrew Gordon Wilson. Practical multi-fidelity Bayesian optimization for hyperparameter tuning. In *Uncertainty in Artificial Intelligence (UAI)*, 2020.

Kaiwen Wu and Jacob R. Gardner. Knowledge gradient for preference learning. arXiv preprint arXiv:2601.22335, 2026.

Kaiwen Wu, Craig Sanders, Benjamin Letham, and Phillip Guan. Mixed likelihood variational Gaussian processes. arXiv preprint arXiv:2503.04138, 2025.

Wenjie Xu, Masaki Adachi, Colin Jones, and Michael A Osborne. Principled bayesian optimization in collaboration with human experts. In *Advances in Neural Information Processing Systems*, volume 37, pp. 104091–104137. Curran Associates, Inc., 2024a.

Wenjie Xu, Wenbin Wang, Yuning Jiang, Bratislav Svetozarevic, and Colin N. Jones. Principled preferential Bayesian optimization. In *Proceedings of the 41st International Conference on Machine Learning (ICML)*, pp. 55305–55336, 2024b.

A PROOF OF THE QUERY POSTERIOR MEAN UPDATE

We prove the identity stated in Section 4.3 for the fantasy posterior mean after a pairwise comparison, adapting Lemma 3.1 of Wu & Gardner (2026) to expert noise σ_{comp}, linear utility U, and an LMC-coupled mental model g. The identity is applied to each mixture component q_r separately, so we drop the index r throughout this appendix. The component weights update by Bayes' rule via the predictive probability, $\omega_r' \propto \omega_r \Pr_r(d)$.

A.1 SETUP

Under one variational component q_r, the triple $(f(x), g(x_a), g(x_b))$ is jointly Gaussian with means $(\mu^f(x), \mu^g(x_a), \mu^g(x_b))$ and covariances inherited from the component's posterior kernel blocks K^{ff}, K^{gg}, K^{fg} defined in Section 3. Let $U : \mathbb{R}^m \to \mathbb{R}$ be linear, $U(y) = w^\top y$ for some $w \in \mathbb{R}^m$. Define the utility difference

$$\Delta_g := U(g(x_a)) - U(g(x_b)) = w^\top (g(x_a) - g(x_b)).$$

Under the variational posterior Δ_g is Gaussian with moments

$$\mu_{\Delta_g} = w^\top (\mu^g(x_a) - \mu^g(x_b)),$$
$$\sigma^2_{\Delta_g} = w^\top \left[K^{gg}(x_a, x_a) + K^{gg}(x_b, x_b) - 2K^{gg}(x_a, x_b)\right] w,$$

and covariance with $f(x)$

$$\gamma := \text{Cov}\left[f(x), \Delta_g\right] = K^{fg}(x, x_a)w - K^{fg}(x, x_b)w \in \mathbb{R}^m.$$

The comparison likelihood of Section 3 encodes the event $\{d = 1\}$ as $\{\Delta_g + \varepsilon_{\text{comp}} > 0\}$ for $\varepsilon_{\text{comp}} \sim N(0, 2\sigma^2_{\text{comp}})$ independent of (f, g). Writing

$$\nu^2 := \sigma^2_{\Delta_g} + 2\sigma^2_{\text{comp}}, \qquad \tau := \frac{\mu_{\Delta_g}}{\nu},$$

the marginal probability of the event is $\Pr(d = 1) = \Phi(\tau)$ by the standard Gaussian tail identity. We claim

$$\mathbb{E}\left[f(x) \mid d = 1\right] = \mu^f(x) + \frac{\phi(\tau)}{\Phi(\tau)} \cdot \frac{\gamma}{\nu}, \tag{1}$$

which is the formula used in Section 4.3. The proof follows by the logic presented in Wu & Gardner (2026), but we present a version of it below for clarity.

A.2 PROOF

Introduce the auxiliary variable $T := \Delta_g + \varepsilon_{\text{comp}}$. Since $\varepsilon_{\text{comp}}$ is independent of (f, g) and Gaussian, the pair $(f(x), T)$ is jointly Gaussian with

$$\mathbb{E}\left[T\right] = \mu_{\Delta_g}, \qquad \text{Var}\left[T\right] = \sigma^2_{\Delta_g} + 2\sigma^2_{\text{comp}} = \nu^2, \qquad \text{Cov}\left[f(x), T\right] = \gamma.$$

The event of interest is $\{T > 0\}$, a half-space constraint on the auxiliary Gaussian.

We use the classical identity for the mean of a jointly Gaussian random variable under a half-space constraint (Azzalini, 2013). If (X, T) is jointly Gaussian with $\mathbb{E}\left[X\right] = \mu_X$, $\mathbb{E}\left[T\right] = \mu_T$, $\text{Var}\left[T\right] = \sigma^2_T$, and $\text{Cov}\left[X, T\right] = \gamma_X$, then

$$\mathbb{E}\left[X \mid T > 0\right] = \mu_X + \frac{\gamma_X}{\sigma_T} \cdot \frac{\phi(\mu_T/\sigma_T)}{\Phi(\mu_T/\sigma_T)}.$$

The identity applies componentwise when X is vector-valued, with γ_X a vector of covariances. Applying it with $X = f(x)$, $\mu_X = \mu^f(x)$, $\mu_T = \mu_{\Delta_g}$, $\sigma_T = \nu$, and $\gamma_X = \gamma$ yields

$$\mathbb{E}\left[f(x) \mid d = 1\right] = \mu^f(x) + \frac{\gamma}{\nu} \cdot \frac{\phi(\tau)}{\Phi(\tau)},$$

which is equation 1. The case $d = 0$ follows by replacing the event $\{T > 0\}$ with $\{T < 0\}$, equivalently by replacing Δ_g with $-\Delta_g$. Both the standardized mean and the covariance with the auxiliary variable change sign, giving

$$\mathbb{E}\left[f(x) \mid d = 0\right] = \mu^f(x) - \frac{\gamma}{\nu} \cdot \frac{\phi(\tau)}{\Phi(-\tau)}.$$

B SITE SOLVER

At a fixed node ρ_r and fixed $\hat{\theta}$, we optimize q_r by conjugate-computation VI (Khan & Lin, 2017). Each evaluation and comparison term depends on q_r only via the observed output value or the utility difference Δ_g, which are Gaussian. Each observation contributes one Gaussian site, a precision and a linear term, as in Opper & Archambeau (2009). An observation site enters the block of its output. A comparison site enters every block. The sites are updated with damping, accepting the first step that keeps every block precision positive definite and does not increase the objective, and the fixed-point iteration is accelerated by Anderson mixing (Walker & Ni, 2011).

C COMPUTATIONAL DETAILS

This appendix collects the closed-form expressions and approximations used in the per-node ELBO (Section 3.2) and the VoI acquisition (Sections 4.2 and 4.3). Throughout, U is linear, $U(y) = w^\top y$, so U commutes with Gaussian expectations, and θ is fixed at the target-only MAP estimate $\hat{\theta}$ of Section 3.2.

We begin with the ELBO derivation at a fixed node ρ_r. The marginal log-likelihood $\log p(\mathcal{D} \mid \rho_r, \hat{\theta})$ is intractable due to the probit factors. Let $u = \{u_{j,l}\}$ denote the full collection of inducing values across latents, as in Section 3.2. Writing $\log p(\mathcal{D} \mid \rho_r, \hat{\theta})$ as an expectation under the joint $p(f, g, u \mid \rho_r, \hat{\theta})$ and applying Jensen's inequality gives

$$\log p(\mathcal{D} \mid \rho_r, \hat{\theta}) \geq \mathbb{E}_{q_r(f,g,u)} \left[\log \frac{p(f, g, u \mid \rho_r, \hat{\theta})\, p(\mathcal{D} \mid f, g)}{q_r(f, g, u)} \right] =: \mathcal{L}_r(q_r).$$

Substituting the SVGP factorization $q_r(f, g, u) = p(f, g \mid u, \rho_r, \hat{\theta})\, q_r(u)$ cancels the conditional terms in numerator and denominator, and the dataset factorization $p(\mathcal{D} \mid f, g) = \prod_i p(y_i \mid f) \prod_j p(d_j \mid g)$ yields the decomposed form of the main text. Since $\mathcal{L}_r \leq \log p(\mathcal{D} \mid \rho_r, \hat{\theta})$ at every node, the weights $\omega_r \propto \pi_r \exp(\mathcal{L}_r)$ of Section 3.2 are the posterior component masses with each node's evidence replaced by its lower bound.

C.1 ELBO TERMS

Observation term. Each observation term is a Gaussian expectation of a Gaussian log density,

$$\begin{aligned}\mathbb{E}_{z \sim N(\mu_q, \Sigma_q)}[\log N(z \mid \mu_p, \Sigma_p)] = -\tfrac{1}{2}\big[& m \log 2\pi + \log|\Sigma_p| + \operatorname{tr}(\Sigma_p^{-1}\Sigma_q) \\ & + (\mu_q - \mu_p)^\top \Sigma_p^{-1}(\mu_q - \mu_p)\big],\end{aligned}$$

evaluated at the f-marginal of the component, $\mu_q = \mu^{(r)}(x_i)$ and $\Sigma_q = K_{ff}^{(r)}(x_i, x_i)$, against the observation $\mu_p = y_i$, $\Sigma_p = \Sigma_{\text{eval}}$.

Query term. For $U(y) = w^\top y$, the utility difference $\Delta_{g,j} = U(g(x_j)) - U(g(x'_j))$ is Gaussian under q_r with mean $\mu_{\Delta_{g,j}} = w^\top(\mu^g(x_j) - \mu^g(x'_j))$ and variance $\sigma^2_{\Delta_{g,j}} = w^\top[K^{gg}(x_j, x_j) + K^{gg}(x'_j, x'_j) - 2K^{gg}(x_j, x'_j)]w$, all moments taken under q_r. The expected log-likelihood reduces to a one-dimensional integral,

$$\mathbb{E}_{q_r}\big[\log p(d_j \mid x_j, x'_j, g)\big] = \mathbb{E}_{\Delta \sim N(\mu_{\Delta_{g,j}}, \sigma^2_{\Delta_{g,j}})} \left[\log \Phi\left(\frac{(2d_j - 1)\,\Delta}{\sqrt{2}\,\sigma_{\text{comp}}} \right) \right],$$

computed by one-dimensional Gauss–Hermite quadrature.

KL term. Under the LMC, the KL divergence is a sum over outputs of the KL between the joint variational block $q_r(u_{j,0}, u_{j,1}) = N(m_j^{(r)}, S_j^{(r)})$ and its prior $p(u_{j,0}, u_{j,1} \mid \hat{\theta}) = N(m_j^0, K_j)$, where $K_j = \operatorname{blockdiag}(K_{ZZ}^{(j)}, K_{ZZ}^{(j)})$ since both latents of output j share the kernel k_j, and $m_j^0 =$

$(c_j/a_j \cdot \mathbf{1}, \mathbf{0})$ carries the constant mean of the shared latent:

$$\sum_j \mathrm{KL}(q_r(u_{j,0}, u_{j,1}) \,\|\, p(u_{j,0}, u_{j,1} \mid \hat{\theta})) = \tfrac{1}{2} \sum_j \big[\mathrm{tr}(K_j^{-1} S_j^{(r)}) + (m_j^{(r)} - m_j^0)^\top K_j^{-1} (m_j^{(r)} - m_j^0) - 2M + \log|K_j| - \log|S_j^{(r)}|\big].$$

In the implementation the inducing values are whitened, $u = K_j^{1/2}\tilde{u} + m_j^0$, so the prior on $\tilde{u}$ is $N(0, I)$ and the KL loses the K_j factors.

C.2 Fantasy posterior updates

Both fantasies act on each component q_r and on the weights ω_r; the variational parameters of the components are held fixed, so the fantasy is the exact one-step Bayesian update of the mixture, not a refit. We write the component moments without the index r.

Sample fantasy. A Gaussian observation y of f at x_{cand} updates both the f- and g-marginals of each component:

$$\mu^f_{\mathcal{D}'}(x) = \mu^f(x) + K^{ff}(x, x_{\mathrm{cand}}) \left[K^{ff}(x_{\mathrm{cand}}, x_{\mathrm{cand}}) + \Sigma_{\mathrm{eval}}\right]^{-1} (y - \mu^f(x_{\mathrm{cand}})),$$
$$\mu^g_{\mathcal{D}'}(x) = \mu^g(x) + K^{gf}(x, x_{\mathrm{cand}}) \left[K^{ff}(x_{\mathrm{cand}}, x_{\mathrm{cand}}) + \Sigma_{\mathrm{eval}}\right]^{-1} (y - \mu^f(x_{\mathrm{cand}})),$$

and the weights by the component predictive density, $\omega_r' \propto \omega_r\, N\big(y \mid \mu^{(r)}(x_{\mathrm{cand}}),\, K^{(r)}_{ff}(x_{\mathrm{cand}}, x_{\mathrm{cand}}) + \Sigma_{\mathrm{eval}}\big)$. The cross-block K^{gf} in the second update is the LMC channel by which a sample observation indirectly sharpens the g-marginal, improving the next iteration's query VoI. Both updates are exact and independent of U.

Query fantasy. Let $s_d := 2d-1$, $\mathcal{D}' := \mathcal{D} \cup \{((x_a, x_b), d)\}$, and $\tau_d := s_d \mu_{\Delta_g}/\nu$. The comparison updates both marginal means:

$$\mu^f_{\mathcal{D}'}(x) = \mu^f(x) + s_d \frac{\phi(\tau_d)}{\Phi(\tau_d)} \cdot \frac{\gamma_f(x)}{\nu},$$
$$\mu^g_{\mathcal{D}'}(x) = \mu^g(x) + s_d \frac{\phi(\tau_d)}{\Phi(\tau_d)} \cdot \frac{\gamma_g(x)}{\nu},$$

where

$$\gamma_f(x) := K^{fg}(x, x_a)w - K^{fg}(x, x_b)w,$$
$$\gamma_g(x) := K^{gg}(x, x_a)w - K^{gg}(x, x_b)w,$$

and $\nu := \sqrt{\sigma^2_{\Delta_g} + 2\sigma^2_{\mathrm{comp}}}$. This is the closed-form update derived in Appendix A; choosing $s_d = -1$ gives the required sign reversal for $d = 0$. The weights update by the component's predictive probability of the observed label, $\omega_r' \propto \omega_r\, \Phi(\tau_{r,d})$. In both fantasies the post-observation expected utility is $\mathbb{E}_{\mathcal{D}'}\left[U(f(x))\right] = \sum_r \omega_r'\, w^\top \mu^{f,(r)}_{\mathcal{D}'}(x)$.

C.3 Dataset utility and VoI evaluation

Dataset utility. Since U is linear, $u(\mathcal{D}) = \max_x \sum_r \omega_r\, w^\top \mu^{(r)}_{\mathcal{D}}(x)$. The maximum is taken over a fixed scrambled-Sobol support of 257 points and polished by L-BFGS-B on $[0,1]^d$ to a projected gradient below 10^{-5}. The same procedure gives $\hat{x}$.

Sample VoI. For $m = 1$ the outcome expectation is Gauss–Legendre quadrature on $[\min_r(\mu_r - 10\sigma_r), \max_r(\mu_r + 10\sigma_r)]$ with 129 nodes in the screen and 65 in the refinement. For $m \geq 2$ it is a sample average over $S = 128$ antithetic scrambled-Sobol base normals, each assigned to a component by systematic sampling of ω and mapped to $y_s(x) = \mu^{(k_s)}(x) + \mathrm{diag}(\sigma^{(k_s)}(x))\, z_s$. The design is drawn once per step and shared by all candidates and stages, and the current recommendation acts as a control variate.

Query VoI. Evaluated exactly by the two-term sum of Section 4.3.

Optimization. Three stages per action: (i) score every support point, or every pair of support points, with the recommendation restricted to the support; (ii) polish the four best by L-BFGS-B, recommendation still on the support; (iii) refine the two best jointly with free recommendation points, one per outcome node or per label, by L-BFGS-B on $[0,1]^d$ (at most 300 iterations, projected gradient 10^{-5}). The largest attained value is kept. Components with $\omega_r < 10^{-8}$ are dropped first.

D LMC PARAMETERIZATION AND PRIORS

This appendix details the parameterization of the mixing matrices A_j of Section 3, the priors on θ and ρ, and the identifiability argument behind the calibration evaluations of Algorithm 1.

D.1 PARAMETERIZATION

For each output j, the independent latent processes have priors $u_{j,0} \sim \mathcal{GP}(c_j/a_j, k_j)$ and $u_{j,1} \sim \mathcal{GP}(0, k_j)$, with the same squared-exponential kernel k_j. The shared latent's mean gives f_j the constant mean c_j. Latent pairs are independent across output dimensions conditional on θ and ρ.

For each output j the mixing matrix is lower-triangular,

$$A_j(\rho,\theta) = \begin{pmatrix} a_j & 0 \\ s_j\,\rho & s_j\sqrt{1-\rho^2} \end{pmatrix},$$

with $a_j > 0$ the amplitude of f_j, $s_j > 0$ the scale of the expert's output, and $\rho \in (0,1)$ shared across outputs. The objective is standardized per output before optimization and the weights of the utility are standardized such that the utility has unit standard deviation. Because we assume no correlation between outputs, we must choose whether the prior on a_j matches the variance of the utility or that of the individual outputs. We choose to match utility, with prior mean $1/\|w\|$ for every a_j, the prior of the utility has unit standard deviation. This is the scale at which recommendations and comparisons act. a_j is part of θ.

The expert's mental model is assumed to be on the same scale as the individual outputs, so $s_j = 1$ is fixed. Writing the g-row in (ρ, s_j) rather than in two free entries puts the prior directly on ρ and makes ρ the only quantity the comparison channel has to resolve. Negative ρ (an anti-aligned expert) is excluded by construction. The lower-triangular form removes rotational invariance (Schmidt & Gelfand, 2003). The kernel output-scale is fixed at one, so that output variance is carried by A_j.

We present the full generative model here:

$$\begin{aligned}
\rho \sim p(\rho), \qquad \theta \sim p(\theta),& \\
u_{j,0} \sim \mathcal{GP}(c_j/a_j,\, k_j(\cdot,\cdot;\theta)), \quad u_{j,1} \sim \mathcal{GP}(0,\, k_j(\cdot,\cdot;\theta)) \perp\!\!\!\perp j = 1,\dots,m,& \\
(f_j, g_j)^\top = A_j(\rho,\theta)\,(u_{j,0}, u_{j,1})^\top,& \\
y_i \mid x_i, f \sim N(f(x_i), \Sigma_{\text{eval}}),& \\
d_j \mid (x_j, x_j'), g \sim \text{Bern}\Big(\Phi\Big(\tfrac{U(g(x_j)) - U(g(x_j'))}{\sqrt{2}\,\sigma_{\text{comp}}}\Big)\Big),&
\end{aligned}$$

D.2 PRIORS AND FIXED QUANTITIES

θ collects, per output, the ARD lengthscales $\ell_j \in \mathbb{R}^d_+$, the amplitude a_j and the constant mean c_j; both latents of output j share k_j. Because the direct observations have a Gaussian likelihood and the marginal prior on f is independent of ρ, the target-only marginal likelihood is available in closed form. We maximize its product with the hyperprior independently for each output using L-BFGS, then hold θ fixed during the ρ-node variational fits. The noise levels are known and fixed. The prior on ρ assigns probability 0.95 to $\rho = 1 - 10^{-6}$ and spreads the remaining 0.05 uniformly over $(0,1)$. We approximate the uniform part using nine equal-cell midpoint nodes and retain the near-one atom as a tenth node, giving prior weights $\pi_r = 0.05/9$ for $r = 1,\dots,9$ and $\pi_{10} = 0.95$ (prior mean approximately 0.975).

Table 1: Priors and fixed quantities. Output dimensions $j = 1, \ldots, m$ each have their own copy of (ℓ_j, a_j, c_j).

Quantity	Prior or value
Lengthscale $\ell_{j,k}$	$\mathrm{Gamma}(64,\ 64/\bar{\ell}_{j,k})$ (mean $\bar{\ell}_{j,k}$, CV 12.5%)
Amplitude a_j	$\mathrm{Gamma}(20, 20\|w\|)$ (mean $1/\|w\|$, CV 22%)
Constant mean c_j	$N(0, 0.25^2)$
Expert scale s_j	fixed at 1
σ_{eval}, σ_{comp}	fixed at 0.05, 0.1
ρ	atom–uniform mixture; $R = 10$
Inducing points Z	M scrambled-Sobol points, fixed; $M = 128, 256, 384$ for $d = 2, 3, 5$–6

The per-dimension prior means $\bar{\ell}$ of Appendix D are the medians of exact ARD-RBF maximum a posteriori fits to noise-free standardized evaluations of each benchmark at 64 to 512 scrambled-Sobol points under a weak prior; at $m \geq 2$ one vector per output. A practitioner facing a real black box would not have these; every method shares them.

Table 2: Per-dimension ARD lengthscale-prior means $\bar{\ell}$ on the unit cube. For multi-output benchmarks, vectors are listed in output order; VLMOP2 uses one vector shared by both outputs.

Benchmark	$\bar{\ell}$
Electrolyte	$(0.550, 2.000, 2.000)$
Ackley	$(0.385, 0.414, 0.436, 0.334, 0.487)$
Hartmann6	$(0.230, 0.370, 0.500, 0.280, 0.290, 0.290)$
VLMOP2	$(0.242, 0.242)$ *(shared)*
Branin–Currin	$(0.273, 1.149)$
	$(0.249, 0.500)$
DTLZ2	$(0.791, 0.803, 1.534, 1.629, 1.492, 1.655, 1.605, 1.627)$
	$(0.817, 0.814, 1.632, 1.645, 1.501, 1.793, 1.595, 1.614)$
	$(0.621, 2.000, 1.434, 1.511, 1.452, 1.531, 1.444, 1.467)$
WFG4	$(0.734, 0.768, 0.654, 0.748, 0.620, 0.671, 0.716, 0.783, 0.324, 0.292)$
	$(0.469, 0.482, 0.516, 0.473, 0.405, 0.528, 0.976, 1.133, 0.564, 0.498)$
	$(0.402, 0.335, 0.358, 0.376, 0.695, 0.803, 1.202, 1.204, 0.906, 0.932)$
	$(0.265, 0.262, 0.502, 0.547, 1.494, 1.351, 1.187, 1.264, 0.991, 1.089)$
	$(0.217, 0.196, 1.411, 1.532, 1.546, 1.554, 1.267, 1.458, 1.064, 1.160)$

D.3 Identifiability of ρ

With no direct observations and k_j shared by both latents, the prior covariance of any g-contrast is $\rho^2 K + (1-\rho^2)K = K$, so the comparison likelihood does not depend on ρ: comparisons alone carry exactly no information about it, whatever their number. Direct observations pin the shared latent, and only then do comparisons discriminate between ρ values. This is why Algorithm 1 places one evaluation before any comparison and a second, information-maximizing one if sufficiently many comparisons arrive before ordinary VoI buys it; the second is placed at the support point maximizing $I(\rho; y(x) \mid \mathcal{D}) = H(\omega) - \mathbb{E}_y[H(\omega'(y))]$, with $\omega'(y)$ the weights after the sample fantasy of Appendix C.2 and the expectation computed as in Appendix C.3.

E Extended experimental results

Utility at budget exhaustion. Table 4 complements Table 3. On Electrolyte, VLMOP2 and Branin–Currin, methods using evaluations reach near-optimal utility by budget exhaustion; on Hartmann6, DTLZ2 and WFG4, final utility still varies across methods. The AUC shows how quickly each method improves.

Table 3: Cost-normalized area under the stop-now utility trajectories of Figures 1 and 2 across the single- and multi-output benchmarks, measured over the first 100 budget units. The current recommendation is held between actions and its normalized utility is integrated over this interval, so higher values indicate that good designs are found earlier. Entries are means over 20 paired repetitions (standard error in parentheses); the best available method per row is bold. Evaluation-only entries repeat across expert rows, and dashes mark an unavailable method. Utility at cost 100 is reported in Table 4.

Benchmark	Expert	EA-BO	KG-Comp	DTS	LogEI	UCB	KG-Eval	CoExBO
Scalar-output benchmarks								
Electrolyte	aligned	0.94 (0.00)	**0.97** (0.00)	0.96 (0.00)	0.90 (0.00)	0.89 (0.00)	0.85 (0.01)	0.75 (0.02)
Electrolyte	biased	0.83 (0.01)	0.46 (0.08)	0.42 (0.08)	0.89 (0.00)	**0.90** (0.00)	0.85 (0.01)	0.72 (0.02)
Ackley	aligned	**0.81** (0.01)	0.78 (0.04)	0.71 (0.01)	0.67 (0.01)	0.68 (0.01)	0.41 (0.03)	0.49 (0.01)
Ackley	biased	0.61 (0.01)	0.17 (0.02)	0.20 (0.02)	0.67 (0.01)	**0.68** (0.01)	0.39 (0.04)	0.40 (0.02)
Hartmann6	aligned	**0.83** (0.00)	0.79 (0.04)	0.59 (0.02)	0.56 (0.02)	0.61 (0.02)	0.50 (0.05)	0.42 (0.03)
Hartmann6	biased	0.46 (0.03)	0.07 (0.02)	0.11 (0.01)	0.56 (0.02)	**0.61** (0.02)	0.50 (0.05)	0.28 (0.04)
Vector-output benchmarks								
VLMOP2	aligned	0.93 (0.00)	**0.95** (0.00)	0.92 (0.01)	0.87 (0.01)	0.86 (0.01)	0.71 (0.01)	—
VLMOP2	biased	0.86 (0.01)	0.09 (0.03)	0.07 (0.02)	**0.87** (0.01)	0.86 (0.01)	0.71 (0.01)	—
Branin–Currin	aligned	0.93 (0.00)	**0.97** (0.00)	0.96 (0.00)	0.86 (0.01)	0.87 (0.01)	0.87 (0.02)	—
Branin–Currin	biased	0.80 (0.03)	0.52 (0.10)	0.55 (0.09)	0.87 (0.01)	**0.87** (0.02)	0.87 (0.02)	—
DTLZ2	aligned	0.89 (0.01)	**0.95** (0.01)	0.90 (0.02)	0.82 (0.03)	0.82 (0.04)	0.86 (0.03)	—
DTLZ2	biased	0.66 (0.04)	0.52 (0.07)	0.50 (0.04)	0.82 (0.03)	0.82 (0.04)	**0.86** (0.03)	—
WFG4	aligned	0.76 (0.04)	**0.81** (0.02)	0.74 (0.04)	0.72 (0.04)	0.70 (0.05)	0.54 (0.04)	—
WFG4	biased	0.67 (0.05)	0.49 (0.05)	0.53 (0.05)	**0.72** (0.04)	0.70 (0.05)	0.54 (0.04)	—

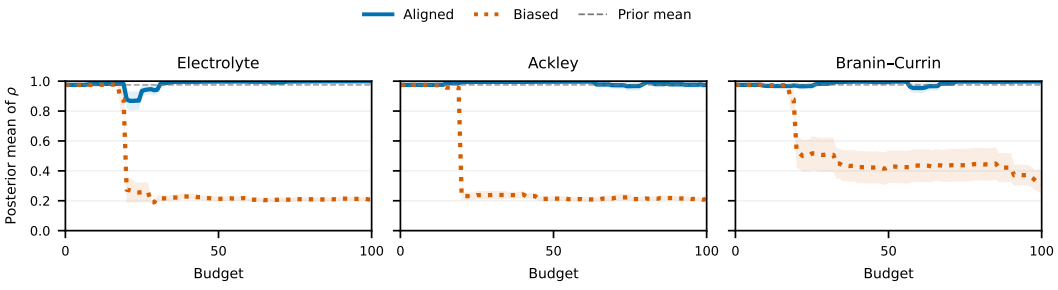


Figure 4: Posterior mean of the expert-alignment parameter ρ against cumulative budget on Electrolyte, Ackley, and Branin–Currin (columns). Solid and dotted curves denote aligned and biased experts. Curves are means over 20 paired repetitions and bands are ± 1 standard error; the horizontal reference is the prior mean, approximately 0.975. Under alignment the posterior remains near one, while under bias it falls as evaluations and comparisons reveal the disagreement.

Table 4: Normalized true utility of the final recommendation at budget exhaustion (cost 100) on the single- and multi-output benchmarks; higher is better. Entries are means over 20 paired repetitions (standard error in parentheses), with the best per row in bold. Evaluation-only entries repeat across expert rows because those methods do not observe comparisons. Dashes mark an unavailable method.

Benchmark	Expert	EA-BO	KG-Comp	DTS	LogEI	UCB	KG-Eval	CoExBO
Electrolyte	aligned	**1.00** (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	0.99 (0.01)
Electrolyte	biased	**1.00** (0.00)	0.44 (0.09)	0.40 (0.10)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	0.99 (0.01)
Ackley	aligned	**0.93** (0.02)	0.91 (0.05)	0.82 (0.01)	0.82 (0.01)	0.83 (0.02)	0.76 (0.02)	0.74 (0.02)
Ackley	biased	**0.88** (0.02)	0.19 (0.02)	0.17 (0.03)	0.83 (0.02)	0.81 (0.02)	0.73 (0.02)	0.64 (0.05)
Hartmann6	aligned	**0.98** (0.00)	0.91 (0.05)	0.83 (0.02)	0.91 (0.01)	0.93 (0.01)	0.84 (0.04)	0.88 (0.02)
Hartmann6	biased	0.86 (0.04)	0.07 (0.02)	0.08 (0.02)	0.91 (0.01)	**0.93** (0.01)	0.84 (0.04)	0.63 (0.07)
Vector-output benchmarks								
VLMOP2	aligned	**1.00** (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	0.99 (0.01)	1.00 (0.00)	—
VLMOP2	biased	**1.00** (0.00)	0.10 (0.04)	0.10 (0.03)	1.00 (0.00)	0.99 (0.01)	1.00 (0.00)	—
Branin–Currin	aligned	1.00 (0.00)	1.00 (0.00)	0.99 (0.00)	**1.00** (0.00)	0.99 (0.01)	1.00 (0.00)	—
Branin–Currin	biased	1.00 (0.00)	0.52 (0.10)	0.56 (0.10)	**1.00** (0.00)	0.99 (0.01)	1.00 (0.00)	—
DTLZ2	aligned	0.97 (0.02)	**0.99** (0.00)	0.97 (0.01)	0.97 (0.01)	0.94 (0.04)	0.96 (0.02)	—
DTLZ2	biased	0.85 (0.06)	0.48 (0.09)	0.45 (0.07)	**0.97** (0.01)	0.94 (0.04)	0.96 (0.02)	—
WFG4	aligned	0.83 (0.04)	**0.86** (0.02)	0.78 (0.04)	0.78 (0.04)	0.74 (0.04)	0.65 (0.06)	—
WFG4	biased	0.74 (0.05)	0.49 (0.05)	0.48 (0.06)	**0.78** (0.04)	0.74 (0.04)	0.65 (0.06)	—

Table 5: EA-BO source allocation across four 25-unit budget quartiles through cost 100. Q1–Q4 cover costs 0–25, 25–50, 50–75 and 75–100. Each entry is the mean number of evaluations/comparisons selected in that span and completed by cost 100; n gives the number of runs.

Benchmark	Expert	n	Q1 E/C	Q2 E/C	Q3 E/C	Q4 E/C
Electrolyte	aligned	15	2.7/12.6	2.4/12.4	3.3/9.2	3.3/6.8
Electrolyte	biased	9	3.1/10.1	5.0/0.0	5.0/0.0	4.8/0.4
Ackley	aligned	15	2.0/15.0	0.7/21.9	1.0/20.7	1.1/18.3
Ackley	biased	9	3.0/10.0	5.0/0.1	5.0/0.0	4.9/0.4
Hartmann6	aligned	20	2.0/15.0	0.6/22.2	0.1/24.1	0.0/25.0
Hartmann6	biased	20	2.6/11.8	4.4/3.6	4.8/1.2	4.5/1.2
VLMOP2	aligned	20	4.2/4.8	5.0/1.1	4.7/2.0	3.9/3.5
VLMOP2	biased	20	4.8/3.5	5.0/0.1	5.0/0.1	4.2/1.8
Branin–Currin	aligned	15	2.9/12.3	3.0/9.9	3.0/9.5	2.4/12.0
Branin–Currin	biased	9	3.1/10.3	4.6/2.4	4.8/1.1	3.7/5.6
DTLZ2	aligned	14	1.0/20.0	2.1/15.3	2.0/16.1	1.1/17.6
DTLZ2	biased	14	1.5/18.7	4.1/3.9	4.2/4.3	4.0/4.2
WFG4	aligned	14	1.0/20.0	2.6/12.0	4.0/5.1	2.8/10.8
WFG4	biased	14	1.0/20.0	3.0/10.1	5.0/0.2	4.9/0.4

F ABLATIONS

Table 6: EA-BO on Ackley as the expert's true ρ varies. The table reports mean numbers of completed evaluations (E) and comparisons (C), and AUC with standard errors in parentheses. All runs use $c_{\text{eval}} = 5$, $c_{\text{comp}} = 1$, $\sigma_{\text{eval}} = 0.05$ and $\sigma_{\text{comp}} = 0.1$.

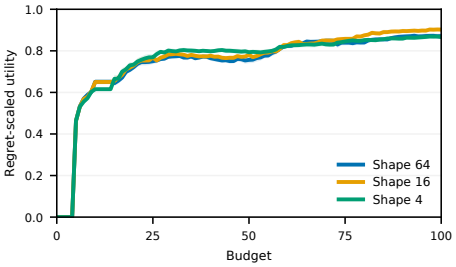
True ρ	E	C	EA-BO AUC
0	17.9	10.6	0.614 (0.008)
0.25	14.0	30.0	0.547 (0.048)
0.5	11.2	43.9	0.601 (0.043)
0.75	7.3	63.3	0.719 (0.012)
1	4.2	78.9	0.821 (0.009)

Table 7: Evaluation-cost sensitivity on aligned Ackley. The table panel reports mean numbers of completed evaluations (E) and comparisons (C) for EA-BO, its AUC, and its paired AUC gain over KG-Eval (Δ AUC); standard errors are in parentheses. Comparison cost is 1, $\sigma_{\text{eval}} = 0.05$, and $\sigma_{\text{comp}} = 0.1$.

c_{eval}	E	C	EA-BO AUC	Δ AUC
1	82.3	17.7	0.838 (0.007)	+0.069 (0.017)
2	14.9	70.2	0.846 (0.009)	+0.168 (0.015)
5	4.2	78.9	0.821 (0.009)	+0.430 (0.042)
10	2.6	74.4	0.767 (0.007)	+0.574 (0.006)

Table 8: Comparison-noise sensitivity on aligned Ackley. The table reports mean completed evaluations (E) and comparisons (C), EA-BO AUC, and its paired AUC gain over the common evaluation-only control, KG-Eval (Δ AUC); standard errors are in parentheses. All runs use $c_{\text{eval}} = 5$, $c_{\text{comp}} = 1$, and $\sigma_{\text{eval}} = 0.05$.

σ_{comp}	E	C	EA-BO AUC	Δ AUC
0.05	3.4	82.8	0.819 (0.005)	+0.428 (0.038)
0.1	4.2	78.9	0.821 (0.009)	+0.430 (0.042)
0.5	10.8	46.1	0.739 (0.008)	+0.347 (0.041)
1	15.4	22.8	0.704 (0.011)	+0.313 (0.045)

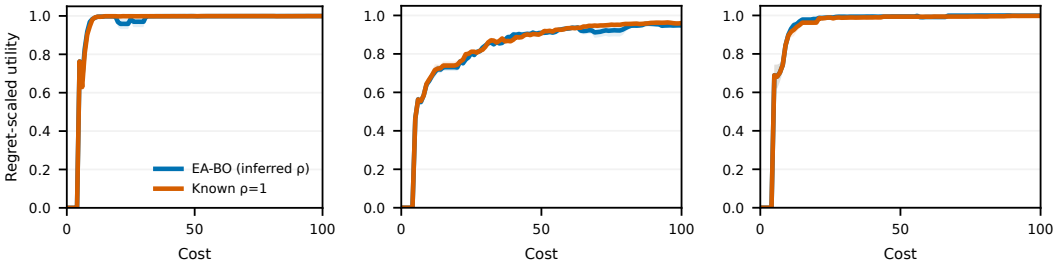


(a) EA-BO trajectories.

Shape	E	C	EA-BO AUC	Δ AUC
64	14.8	24.5	0.740 (0.006)	—
16	12.4	37.3	0.752 (0.005)	+0.012 (0.004)
4	11.1	44.7	0.749 (0.008)	+0.009 (0.008)

(b) AUC and action counts through cost 100.

Table 9: Target-lengthscale prior breadth on aligned Ackley under the uninformative ρ prior with 33 sample-quadrature nodes. The Gamma priors share each coordinate's calibrated mean; at the representative mean $\bar{\ell} = 0.4$, shapes 64, 16, and 4 have standard deviations 0.05, 0.10, and 0.20, and central 95% prior intervals $[0.31, 0.50]$, $[0.23, 0.62]$, and $[0.11, 0.88]$, respectively. The left panel plots mean stop-now normalized true utility through cost 100 for EA-BO with ± 1 standard error bands over 15 paired repetitions. The right panel reports mean numbers of completed evaluations (E) and comparisons (C), cost-normalized AUC, and paired AUC differences from shape 64 (Δ AUC), all through cost 100; standard errors are in parentheses. Evaluation cost is 5, comparison cost is 1, and observation and comparison noise are 0.05 and 0.1.

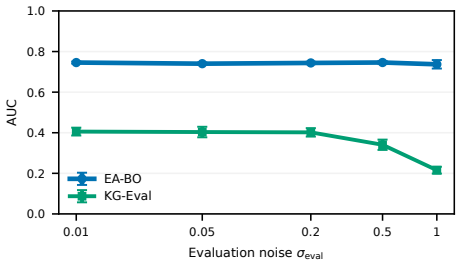


(a) Electrolyte. (b) Ackley. (c) Branin–Currin.

Problem	E	C	Inferred AUC	Known AUC	Δ AUC
Electrolyte	8.6	57.2	0.937 (0.003)	0.939 (0.001)	+0.002 (0.003)
Ackley	2.1	89.4	0.821 (0.009)	0.829 (0.003)	+0.008 (0.008)
Branin–Currin	10.4	47.8	0.929 (0.002)	0.925 (0.003)	−0.004 (0.003)

(d) Cost-100 outcomes and action counts.

Figure 5: Knowing full alignment on Electrolyte, Ackley, and Branin–Currin. The three upper panels show normalized true utility through budget 100 for EA-BO with inferred ρ (blue) and with $\rho = 1$ fixed (orange). Curves and bands are means and ± 1 standard error over nine paired repetitions. The table below reports mean completed evaluations (E) and comparisons (C) under known ρ, each method's cost-normalized AUC, and the paired AUC change (Δ AUC), with standard errors in parentheses.



(a) AUC by evaluation noise.

σ_{eval}	E	C	EA-BO AUC	KG-Eval AUC
0.01	14.6	25.6	0.746 (0.005)	0.406 (0.019)
0.05	14.8	24.5	0.740 (0.006)	0.403 (0.026)
0.2	14.3	27.9	0.744 (0.006)	0.402 (0.021)
0.5	13.0	33.1	0.746 (0.008)	0.341 (0.025)
1	11.9	38.1	0.737 (0.021)	0.216 (0.017)

(b) Action counts and both AUCs.

Figure 6: Evaluation-noise sensitivity on aligned Ackley under a $\mathrm{Beta}(1.2, 0.8)$ prior on ρ. The left panel plots cost-normalized AUC of stop-now normalized true utility through cost 100 for EA-BO and its noise-matched evaluation-only control (KG-Eval). Points and error bars are means and ± 1 standard error over 15 paired repetitions. The right panel reports mean EA-BO evaluations (E) and comparisons (C), EA-BO AUC, and KG-Eval AUC; standard errors are in parentheses. Source runs had budget 200 and are truncated to the first 100 cost units. Evaluation cost is 5, comparison cost is 1, and comparison-noise standard deviation is 0.1.

Table 10: Posterior-mixture versus data-fitted point-MAP inference on Ackley and Hartmann6 with a shared flat $\mathrm{Beta}(1, 1)$ prior on ρ. All runs use budget 200, evaluation cost 5, comparison cost 1, and 15 paired repetitions. The Δ columns give paired point MAP minus mixture differences in cost-normalized stop-now AUC and normalized utility at budget exhaustion, as mean (standard error); positive values favor point MAP. Final mean and bottom-four entries give mixture/MAP utility in that order. Bottom four is each method's mean over its four lowest final-utility runs, with the four selected separately per method.

Benchmark	True ρ	Δ AUC	Δ final	Final mean mix/MAP	Bottom four mix/MAP
Ackley	0	+0.010 (0.011)	+0.001 (0.010)	0.895/0.895	0.855/0.841
Ackley	0.5	−0.065 (0.035)	−0.091 (0.050)	0.895/0.804	0.866/0.595
Ackley	1	+0.049 (0.012)	+0.031 (0.018)	0.896/0.926	0.850/0.873
Hartmann6	0	−0.117 (0.072)	−0.201 (0.093)	0.955/0.754	0.889/0.164
Hartmann6	0.5	−0.047 (0.013)	−0.032 (0.022)	0.938/0.906	0.886/0.806
Hartmann6	1	−0.019 (0.011)	−0.007 (0.009)	0.973/0.966	0.942/0.933

Table 11: Effect of the ρ prior on EA-BO for Ackley under aligned and biased experts. Diffuse uses $\mathrm{Beta}(1.2, 0.8)$. 80/20 and the default prior (Appendix D) put mass on a near-one atom and the remainder on a uniform slab. All runs use budget 100, $c_{\text{eval}} = 5$, $c_{\text{comp}} = 1$ and nine paired repetitions. Δ AUC is the paired difference from diffuse within each expert condition. Final is utility at budget exhaustion, and E/C is the mean number of completed evaluations/comparisons. AUC, Δ AUC and final are means with standard errors in parentheses. Stronger near-one mass accelerates aligned search; under bias it gives up some early AUC but recovers final utility.

Expert	Prior	AUC	Δ AUC	Final	E/C
aligned	Diffuse	0.745 (0.008)	–	0.868 (0.014)	14.7/26.7
aligned	80/20	0.789 (0.007)	+0.043 (0.010)	0.901 (0.019)	7.0/65.0
aligned	Default	0.821 (0.009)	+0.076 (0.011)	0.948 (0.012)	4.2/78.9
biased	Diffuse	0.639 (0.010)	–	0.835 (0.018)	18.9/5.6
biased	80/20	0.613 (0.006)	−0.026 (0.008)	0.863 (0.017)	17.6/12.2
biased	Default	0.614 (0.008)	−0.026 (0.015)	0.880 (0.016)	17.9/10.6