Ensemble Distributionally Robust Bayesian Optimisation

Tigran Ramazyan
Faculty of Computer Science
HSE University
Moscow, Russia
tramazyan@hse.ru

Denis Derkach
Faculty of Computer Science
HSE University
Moscow, Russia
dderkach@hse.ru

Abstract

We study zeroth-order optimisation under context distributional uncertainty, a setting commonly tackled using Bayesian optimisation (BO). A prevailing strategy to make BO more robust to the complex and noisy nature of data is to employ an ensemble as the surrogate model, thereby mitigating the weaknesses of any single model. In this study, we propose a novel algorithm for Ensemble Distributionally Robust Bayesian Optimisation that remains computationally tractable while managing continuous context. We obtain theoretical sublinear regret bounds, improving current state-of-the-art results. We show that our method's empirical behaviour aligns with its theoretical guarantees.

1 Introduction

Design optimisation is fundamental to engineering and scientific discovery, aiming to identify system configurations that maximise operational efficiency or minimise costs [Dorigo, 2023, Baranov et al., 2017]. The underlying objective function is usually unknown, expensive to evaluate, and analytically intractable. These systems are treated as black boxes - complex simulators or real-world experiments that return only point-wise values for queried inputs, offering no direct gradient information to guide the optimisation Boldyrev et al. [2024].

While black-box optimisation spans both gradient-based and gradient-free paradigms, the former is largely restricted to differentiable objectives [Chang et al., 2016, Degrave et al., 2016]. Many critical problems and applications, however, involve non-differentiable landscapes or stochastic simulators with intractable likelihoods. Examples include Monte-Carlo simulations in particle scattering [Sjöstrand et al., 2008] molecular dynamics [Abraham and Gready, 2011] and many more. In these scenarios, Bayesian Optimisation (BO) [Alarie et al., 2021] has emerged as a powerful framework for sample-efficient global optimisation.

A particularly challenging extension is *contextual* BO, where the objective function depends on a controllable design variable $x \in \mathcal{X}$ and an uncontrollable environmental contextual variable $c \in \mathcal{C}$ drawn from an unknown distribution P. While the problem simplifies to standard stochastic optimisation if P is known, practical settings require P to be estimated from empirical data Boldyrev et al. [2025], Ramazyan et al. [2024]. This estimation leads to a *distributional mismatch* between the true and the nominal context distributions, potentially resulting in sub-optimal designs.

The *Distributionally Robust Optimisation* (DRO) paradigm safeguards against such errors [Staib and Jegelka, 2019, Kuhn et al., 2024], and has recently been integrated with BO as *Distributionally Robust Bayesian Optimisation* (DRBO) [Kirschner et al., 2020, Husain et al., 2023, Micheli et al., 2025]. DRO uses statistical divergences or distance metrics to construct an *ambiguity set* - a family of candidate distributions centred around the nominal one. By optimising for the worst-case scenario

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

within this set, DRO provides a robust performance lower bound, ensuring reliability even under distributional shifts.

Contributions. Our primary contributions are as follows:

- We propose *Ensemble Distributionally Robust Bayesian Optimisation* (EDRBO), a novel framework that integrates an ensemble of Gaussian process surrogates with DRO to hedge against both black-box function and contextual distributional uncertainties. Our approach leverages the properties of Wasserstein DRO to derive a computationally tractable acquisition function, effectively resolving the maximin challenge inherent in DRO while preserving efficiency required for practical application.
- We provide a rigorous theoretical analysis, deriving a sublinear cumulative regret of $\mathcal{O}(\gamma_T\sqrt{T})$ order, where T is the time horizon and γ_T denotes the ensemble maximum information gain.
- We demonstrate the efficacy of EDRBO through extensive experiments, showing that it consistently matches or outperforms state-of-the-art methods while maintaining computational tractability.

The main text of this paper is organised as follows. Section 2 presents related works. Section 3 establishes the necessary preliminaries. In Section 4, we detail the EDRBO framework and present our theoretical results. Finally, Section 5 provides an empirical evaluation and discussion of our findings. Section 6 concludes the study.

Notation. We denote the search space by $\mathcal{X} \subseteq \mathbb{R}^{d_x}$ and the context space by $\mathcal{C} \subseteq \mathbb{R}^{d_c}$. We make a common and mild assumption that $\mathcal{X}$ is compact. Let $\Xi = \mathcal{X} \times \mathcal{C}$ and $\xi = [x^T, c^T]^T$. The black-box function $f : \Xi \to \mathbb{R}$ is assumed to be bounded from above. The ensemble surrogate model $\hat{f}$ is comprised of M distinct constituent surrogates $\hat{f}_m,\ m \in \{1, \ldots, M\}, M \geq 2$. Each surrogate is defined by a positive semi-definite kernel $k_m : \Xi \times \Xi \to \mathbb{R}$ with $\mathcal{H}_m$ being the unique RKHS associated with the kernel k_m, equipped with norm $\|\cdot\|_{\mathcal{H}_m}$; k_m is $L_m(x)$-Lipschitz with respect to c. We assume that $\|f\|_{\mathcal{H}_m} \leq B_m, B_m > 0$ and that $\forall \xi, \xi' \in \Xi\ k_m(\xi, \xi') \leq 1$. We use $\mathbb{P}(\mathbb{R}^{d_c})$ to denote the set of all Borel probability measures on $\mathbb{R}^{d_c}$. We use $f(x, \cdot)\#Q$ to denote the pushforward of the context distribution Q through f. We denote the iteration budget by T. At iteration t we use $\mathbb{D}_t = \{(x_\tau, c_\tau, y_\tau)\}_{\tau=1}^T$ for the set of ground truth values and ϵ_t to denote the zero-mean R-sub-Gaussian output noise. We denote expected regret at iteration t by r_t and the cumulative expected regret is defined as $R_T = \sum_{t=1}^T r_t$.

2 Related Works

The UCB algorithm [Srinivas et al., 2012] is one of the main approaches to BO. It fits a Gaussian process to model the unknown objective function. Its posterior mean approximates for the black-box function and the posterior variance highlights areas of search space with high uncertainty. A linear combination of the two, called an acquisition function, provides a high-probability bound on the objective function. Optimising it provides viable candidates for optima. UCB acts as a base for many DRBO algorithms.

The integration of DRO into the BO framework is pioneered by Kirschner et al. [2020]. DRBO-MMD works by discretising the context space and finding a solution that holds up against the worst-case distribution within a set MMD budget. While they show that the problem could be formulated for any divergence, their reliance on Maximum Mean Discrepancy creates significant scaling issues Kirschner et al. [2020]. The resulting minimax problem becomes computationally prohibitive unless the context space is kept discrete and low-dimensional.

Since DRBO-MMD can be computationally heavy, DRBO-MMD Minmax [Tay et al., 2022] offers a faster alternative. They employ Taylor expansions to conduct minimax sensitivity analysis across various ambiguity sets to achieve finer discretization without the same heavy computational overhead. Both DRBO-MMD and DRBO-MMD Minmax methods have linear regret bounds with respect to the iteration budget, limiting their long-term performance guarantees.

On the other hand, Husain et al. [2023] focus on tractability. They extend the DRBO framework to φ-divergences such as Kullback-Leibler divergence and Total Variation distance. They prove that the minimax problem can be reduced to a simpler single-stage optimisation problem, thus making the algorithm much more practical. Nevertheless, their approach introduces implicit constraints regarding the support of the distributions within the ambiguity set. To handle the challenges of

continuous context without relying on discretisation, SBO-KDE [Huang et al., 2024] estimates the context distribution using kernel density estimation and then maximizes the expected UCB through sampling from this estimate within a φ-divergence DRBO setup. Since these density estimates can sometimes be unreliable, DRBO-KDE [Huang et al., 2024] adds a layer of protection by placing a total variation ambiguity set around the estimated distribution. This hedging makes the method more resilient to potential model misspecification.

StableOpt [Bogunovic et al., 2018] takes a different view by looking at robustness through the lens of stability. It identifies the best design by looking for the one that performs best even under the worst plausible context in a set C_t. As suggested by Huang et al. [2024], this set is defined using intervals based on the empirical mean and variance of the context dimensions. This gives us a minimax perspective that explicitly guards against statistically likely but unfavourable context shifts.

Most recently, Micheli et al. [2025] applied Wasserstein DRO to robustify the BO confidence bound acquisition function. Much like Husain et al. [2023], they bypass the maximin bottleneck by utilising the Lipschitz properties of the learner objective over Wasserstein balls. Their work effectively fills a gap in the literature concerning continuous context distributions. WDRBO Micheli et al. [2025] creates a Wasserstein ambiguity set around the observed contexts. The radius is set as $\varepsilon_t = O(1/\sqrt{t})$. It naturally tightens as more data arrives, making the model less conservative over time. Micheli et al. [2025] also propose ERBO, which is a non-robust version of WDRBO that sets the radius to zero. By simply maximising the average performance across observed contexts, ERBO serves as a context-aware non-robust baseline.

Although Gaussian process ensembles are not widely used, various examples exist. For instance Candelieri et al. [2022] fit M Gaussian processes equipped with Squared Exponential kernels with hyperparameters varying over a fixed grid. The Gaussian processes are then combined into their Wasserstein barycentre, which reduces to the arithmetic mean of the individual means and the arithmetic mean of the individual standard deviations. In contrast, we consider Gaussian processes with kernels from distinct families to guarantee orthogonality of base experts, and hence achieve higher predictive performance of the ensemble. In addition to that, we treat the M Gaussian processes as an ensemble, not just distinct runs or configuration of a surrogate model. We do so by considering a posterior radius, which measures how far each individual base expert deviates from the consensus.

3 Preliminaries

3.1 Bayesian Optimisation

Consider optimising a black-box function f with respect to the search space $\mathcal{X}$. f is point-wise observable, that is, there is no direct access to neither f nor its gradients. The learner may only obtain noisy evaluations of f. At iteration t, the learner selects a design point $x_t \in \mathcal{X}$. Then he observes context $c_t \in \mathcal{C}$ and output $y_t = f(x_t, c_t) + \epsilon_t$. Context follows an unknown distribution $P \in \mathbb{P}_2(\mathbb{R})$ and it given by the environment, that is the learner does not choose it. The goal is to find the global maximiser of f:

$$x^* = \arg\max_{x \in \mathcal{X}} \mathbb{E}_{c \sim P}[f(x, c)].$$

The (x_t, c_t, y_t) triplets comprise the set of all ground truth values obtained up to iteration t, $\mathbb{D}_t$. At iteration t, f is modelled by a surrogate model $\hat{f}$ fit on $\mathbb{D}_t$. Using the surrogate's approximation, BO builds an *acquisition function* to guide the optimisation. By optimising the acquisition function, one identifies promising regions of the search space, while accounting all that is known and unknown regarding the objective function at current iteration.

3.2 Surrogate Model

Consider an ensemble of M Gaussian processes (GP). Following Kirschner et al. [2020], Micheli et al. [2025], each base expert solves a regularised least-squares problem in its RKHS. For a regularisation parameter $\lambda > 0$, at iteration t, given $\mathbb{D}_t$, the posterior of the expert with kernel k_m is given by

$$\mu_{m,t}(\xi) = k_{m,t}(\xi)^T \left(K_{m,t} + \lambda I\right)^{-1} y_t \tag{1}$$

and covariance

$$k_{m,t}(\xi, \xi') = \frac{1}{\lambda}\left(k_m(\xi, \xi') - k_{m,t}(\xi)^T \left(K_{m,t} + \lambda I\right)^{-1} k_{m,t}(\xi')\right), \tag{2}$$

where $k_{m,t}(\xi) = [k_m(\xi, \xi_1), \ldots, k_m(\xi, \xi_t)]$, $K_{m,t}$ is the kernel Gram matrix, and $y_t = [y_1, \ldots, y_t]$. We approximate the black-box using the experts mean:

$$\hat{f}(\xi) = \mu_t(\xi) = \frac{1}{M} \sum_{m=1}^{M} \mu_{m,t}(\xi), \tag{3}$$

for which we provide the following finite-sample confidence guarantees.

Lemma 3.1. *Let $\Xi \in \mathbb{R}^d$ and $f : \Xi \to \mathbb{R}$. Each expert assumes that $f \in \mathcal{H}_m$, with $\|f\|_{\mathcal{H}_m} \leq B_m$ and let ϵ_t be R-sub-Gaussian. Let $\delta \in (0, 1)$. With probability at least $1 - \delta$, $\forall \xi \in \Xi$ and $\forall t \geq 1$:*

$$|\mu_t(\xi) - f(\xi)| \leq \beta_t \sigma_t(\xi), \tag{4}$$

with

$$\sigma_t(\xi) = \frac{1}{M} \sum_{m=1}^{M} \sigma_{m,t}(\xi) \tag{5}$$

and

$$\beta_t = \max_m \lambda^{\frac{1}{2}} B_m + R\sqrt{\gamma_{m,t} + 2\log(M/\delta)}, \tag{6}$$

where $\gamma_{m,t}$ is maximum information gain (defined in the following subsection). A full proof is provided in Appendix A.1.

For an ensemble to be effective, the orthogonality, or distinctness, of its base experts is vital. Gaussian processes are fundamentally either equivalent or orthogonal. By the Feldman-Hájek theorem [Feldman, 1958, Hájek, 1958], distinct base experts reside in distinct RKHSs, and hence are orthogonal. We note that even under equivalence two Gaussian processes would yield varying predictions for finite data Rasmussen and Williams [2005]. We equip our method with the following kernels: Square Exponential, Rational Quadratic, and Matérn kernels with $\nu = \frac{3}{2}$ and $\nu = \frac{5}{2}$. While they cover the standard suite of Lipschitz kernels, the future scalability of such ensembles pose an interesting challenge for both the theoretical and empirical aspects of the model.

3.3 Maximum Information Gain

Design points are chosen to obtain the best possible estimates of the black-box function, while attempting to learn f in as few iterations as possible. The reduction in uncertainty about f from observing y_t is called information gain and is the mutual information between f and observed outputs y_t, i.e., $\mathcal{I}(y_t; f)$. In general, finding its maximiser is NP-hard [Ko et al., 1995]. For a Gaussian predictive posterior the problem

$$\gamma_t \overset{\text{def}}{=} \max_{\xi_1, \ldots, \xi_t} \mathcal{I}(y_t; f) \tag{7}$$

is equivalent to choosing the variance-maximising design [Srinivas et al., 2012]. Thus, *maximum information gain* takes the following form:

$$\gamma_{m,t} = \max_{\xi_1, \ldots, \xi_t} \log\det(I + \lambda^{-1} K_{m,t}), \tag{8}$$

where $K_{m,t}$ is the kernel matrix of a Gaussian process surrogate with kernel k_m over $\mathbb{D}_t$. Introduce $\gamma_t = \max_m \gamma_{m,t}$ as the maximum information gain of the ensemble.

3.4 Distributionally Robust Optimisation

The true context distribution P is fundamentally unknown and must be inferred from finite samples. Without P the decision problem cannot be evaluated. A nominal distribution $\tilde{Q}_t$ can be estimated from the observed contexts, but with finite data this estimate is subject to statistical error. To reduce this error, one must increase the sample size, which is costly or impractical within the black-box optimisation framework. This distributional mismatch between P and $\tilde{Q}_t$ leads to the so-called optimiser's curse [Kuhn et al., 2024], whereby improvements fail to yield better decision due to persistent distributional error.

DRO mitigates this by employing an ambiguity set of plausible context distributions. We focus on a common choice of such ambiguity sets - Wasserstein balls. Meaning, we take the 1-Wasserstein

distance as the DRO metric, which results in the following ambiguity set centred at Q_t and radius $\varepsilon_t \geq 0$:

$$\mathbb{B}_{\varepsilon_t}(Q_t) \stackrel{\text{def}}{=} \{Q \in \mathbb{P} : W(Q, Q_t) \leq \varepsilon_t\}. \tag{9}$$

The radius ε_t bounds the allowable distributional deviation. In the DRO framework, the learner has to make the best decision against the worst-case distribution within the ambiguity set. This results in a maximin problem:

$$\max_{x \in \mathcal{X}} \inf_{Q \in \mathbb{B}_{\varepsilon_t}(Q_t)} \mathbb{E}_{c \sim Q}[f(x, c)]. \tag{10}$$

DRO is based on the assumption that the true contextual distribution lies within the ambiguity set: $P \in \mathcal{B}_{\varepsilon_t}(Q_t)$.

4 Method

The key component of BO is the acquisition function. There are several established techniques, most notably the Upper Confidence Bound (UCB) acquisition. It leverages the trade off between exploitation, μ_t, and exploration, σ_t. Directly optimising μ_t is a good worst-case for f [Bull, 2011]. But it may be too greedy, has no mechanism to escape suboptimal solutions, and does not incorporate uncertainty in any form Srinivas et al. [2012]. For UCB, by Lemma 3.1, we can guarantee coverage of the black-box function f with high probability. The UCB acquisition function in the DRBO paradigm is given by the following learner problem:

$$x_t = \arg\max_{x \in \mathcal{X}} \inf_{Q \in \mathcal{B}_{\varepsilon_t}(Q_t)} \mathbb{E}_{c \sim Q}[\mu_t(x, c) + \beta_t \sigma_t(x, c)]. \tag{11}$$

Similarly to Micheli et al. [2025] and unlike Kirschner et al. [2020] or Husain et al. [2023], the algorithm and the following theoretical analysis works for any continuous distribution.

We take advantage of the chosen surrogate mode. Unlike, for example, Bayesian model averaging, which collapses distinct base experts into a single predictive density, we retain the full set of component distributions to characterise the epistemic landscape. This allows us to rigorously disentangle structural ambiguity - disagreement between kernels, from aleatoric uncertainty - lack of data within a kernel.

Consider a meaningful consensus of base experts with respect to the Wasserstein geometry - the Wasserstein Barycentre, which is the solution to the problem:

$$\bar{f}_t(\xi) = \arg\min_{\bar{f}} \frac{1}{M} \sum_{m=1}^{M} W_2(\hat{f}_{m,t}(\xi), \bar{f}(\xi)). \tag{12}$$

In general, finding $\bar{f}(\xi)$ is NP-hard [Agueh and Carlier, 2011]. However, since the posterior of base experts is 1D Gaussian, the ensemble barycentre is analytically tractable. It remains Gaussian, $\bar{f}_t(\xi) \sim \mathcal{N}(\bar{\mu}_t(\xi), \bar{\sigma}_t^2(\xi))$, with $\bar{\mu}_t(\xi) = M^{-1} \sum_{m=1}^{M} \mu_{m,t}(\xi)$ and $\bar{\sigma}_t(\xi) = M^{-1} \sum_{m=1}^{M} \sigma_{m,t}(\xi)$, which match μ_t, Eq. 3, and σ_t, Eq. 5, respectively. The barycentre effectively averages base experts instead of linearly mixing their densities. In contrast with the common single GP paradigm, the posterior radius naturally promotes exploration in regions where experts conflict not only in their predictions but also in their uncertainty estimates, effectively capturing the model-based epistemic uncertainty.

Thus, we can construct the robustness bound to be valid regardless of which expert defines the true context best and the acquisition function defines the exploration budget based on context gap - the distance between the assumed regularity of the most divergent expert and the consensus average. We estimate the radius $\hat{\varepsilon}_t(\xi)$ as the largest 2-Wasserstein distance between any single constituent expert and the consensus barycentre. Leveraging the closed-form Bures-Wasserstein metric for Gaussians, this radius decomposes explicitly into structural and aleatoric components:

Algorithm 1 General EDRBO

Input: Iteration budget T, M GPs, initial data $\mathbb{D}_0$, trade-off parameter β_t
for $t = 1$ **to** T **do**
  for $m = 1$ **to** M **do**
    $\mu_{m,t}, \sigma_{m,t} \overset{\text{fit}}{\leftarrow} \mathbb{D}_{t-1}$
  end for
  $x_t = \arg\max_{x \in \mathcal{X}} \mathbb{E}_{c \sim Q_t} \left[\mu_t(x,c) + \beta_t \sigma_t(x,c) + \frac{1}{\sqrt{2}} \hat{\varepsilon}_t(x,c) \right]$
  Observe $c_t \sim P$
  Observe $y_t = f(x_t, c_t) + \epsilon_t$
  $\mathbb{D}_t \leftarrow \mathbb{D}_{t-1} \bigcup \{(x_t, c_t, y_t)\}$
end for

$$\mathbb{E}_{c \sim Q_t}[\hat{\varepsilon}_t(x,c)] = \max_m W_2(\hat{f}_{m,t}(x,\cdot)\#Q_t, \bar{f}(x,\cdot)\#Q_t) = \tag{13}$$

$$= \max_m \mathbb{E}_{c \sim Q_t} \left[\sqrt{\underbrace{(\mu_{m,t}(x,c) - \mu_t(x,c))^2}_{\text{Structural Disagreement}} + \underbrace{(\sigma_{m,t}(x,c) - \sigma_t(x,c))^2}_{\text{Confidence Dispersion}}} \right]. \tag{14}$$

To simplify the DRO problem, we apply the following Wasserstein robust regulariser.

Lemma 4.1 (Lemma 2 by Gao et al. [2020])**.** *For an $L_g(x)$-Lipschitz function g, $\forall x \in \mathcal{X}$, $\forall Q \in \mathbb{B}_{\varepsilon_t}(Q_t)$ and $\forall t \geq 1$:*

$$|\mathbb{E}_{c \sim Q}[g(x,c)] - \mathbb{E}_{c \sim Q_t}[g(x,c)]| \leq \varepsilon_t L_g(x).$$

For Wasserstein ambiguity sets, Lemma 4.1 bounds the solution to the inner optimisation problem. This result discards the inherit minimax DRO problem and yields a tractable approximation of the learner objective. To apply Lemma 4.1, we first establish the Lipschitzness of the chosen acquisition function, which follows from Lemma 3 by Micheli et al. [2025]. A full proof is provided in Appendix A.2. Thus, the distributionally regularised learner problem is given by the following theorem.

Theorem 4.2. *The distributionally robust learner optimisation problem is given by:*

$$x_t = \arg\max_{x \in \mathcal{X}} \mathbb{E}_{c \sim Q_t} \left[\mu_t(x,c) + \beta_t \sigma_t(x,c) + \frac{1}{\sqrt{2}} \hat{\varepsilon}_t(x,c) \right]. \tag{15}$$

A full proof is provided in Appendix A.3.

The proposed algorithm is presented in Algorithm 1. To measure its performance we introduce the notion of regret. In general it is a difference between a prediction and an oracle. The choice of oracle and the definition of regret is not unique. For a consistent comparison with existing DRBO methods, following Huang et al. [2024] and Micheli et al. [2025], we use instantaneous expected regret. It captures the sub-optimality gap between an algorithm and the solution of the true stochastic optimisation problem with full knowledge of the black-box function f and the true contextual distribution P:

$$r_t = \mathbb{E}_{c \sim P}[f(x^*, c)] - \mathbb{E}_{c \sim P}[f(x_t, c)], \tag{16}$$

$$x^* = \arg\max_{x \in \mathcal{X}} \mathbb{E}_{c \sim P}[f(x,c)]. \tag{17}$$

For our acquisition function, Eq. 15, we achieve the following sublinear cumulative regret bound, which is independent of both the distributional gap, $W(P, Q_t)$, and the ambiguity set radius, ε_t.

Theorem 4.3. *Let $\delta \in (0,1)$ be a probability of failure and $\exists\, \ell \geq 1$. Then with probability at least $1 - \delta$, the cumulative instantaneous expected regret of Algorithm 1 after T iterations is bounded by:*

$$R_T \leq 2(4 + \ell)\beta_T \sqrt{T \left(\gamma_T + 2 \log \frac{6}{\delta} \right)}, \tag{18}$$

which results in the $\mathcal{O}(\gamma_T \sqrt{T})$ general order of cumulative regret, where γ_T is the ensemble maximum information gain. A full proof is provided in Appendix A.4.

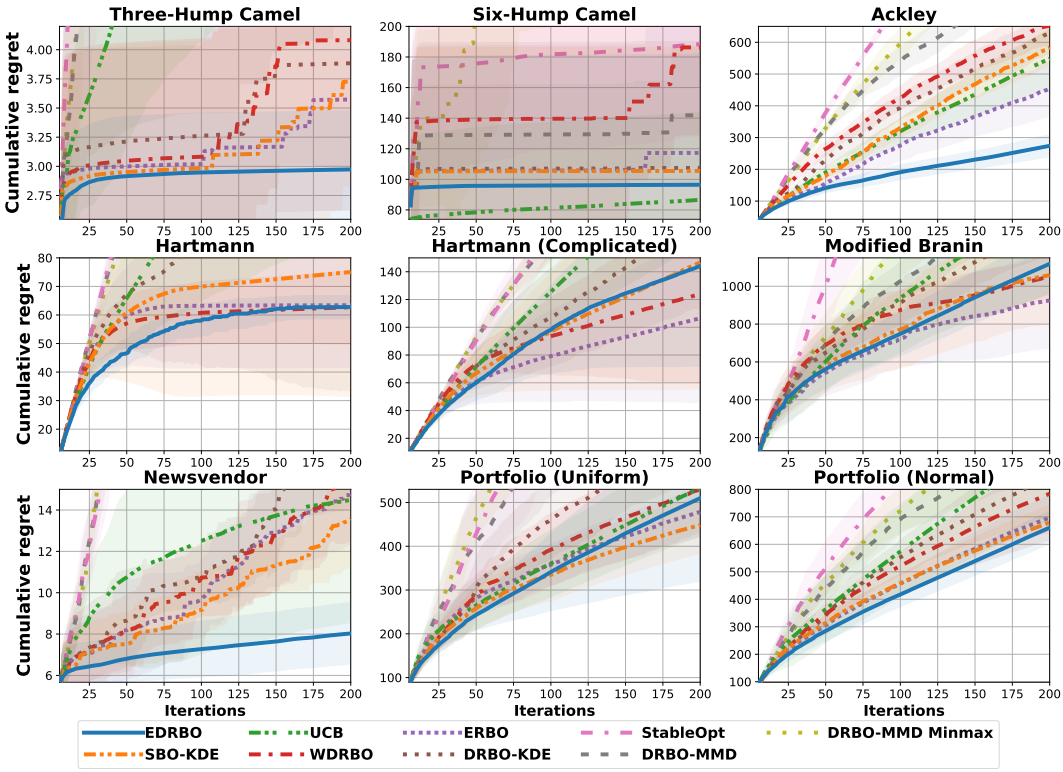


Figure 1: Mean and standard deviation of cumulative expected regret.

5 Experiments

Our algorithm, EDRBO, follows the methodology described in Section 4 and the corresponding Algorithm 1. We compare it against a range of algorithms that take different perspectives on incorporating context and robustness in BO, ranging from purely empirical approaches to principled distributionally robust formulations.

The UCB algorithm [Srinivas et al., 2012] completely sidesteps the context variable. It relies solely on a Gaussian process model of the reward function and optimizes the corresponding UCB, making it an elegant but potentially oversimplified comparator for problems where context plays a central role.

Several of more recent methods explicitly model the context distribution. SBO-KDE [Huang et al., 2024], estimates this distribution via kernel density estimation. The acquisition function then seeks to maximize the expected UCB under the estimated distribution, using sampling to approximate the expectation. Recognizing that KDE-based estimates may be unreliable, DRBO-KDE [Huang et al., 2024] hedges against misspecification by placing a total variation ambiguity set around the KDE-derived empirical distribution. This modification makes the method more resilient to errors in the estimated context model.

A different line of work considers robustness through an MMD-based ambiguity set. DRBO-MMD [Kirschner et al., 2020], discretizes the continuous context space and optimizes the UCB against the worst-case distribution supported on this discretization for a fixed MMD budget. While this formulation offers strong robustness guarantees, it suffers from a cubic computational cost in the number of discretized context points. To address this limitation, DRBO-MMD Minmax [Tay et al., 2022], is a scalable approximation that reduces computational burden through a minmax sensitivity analysis, thereby enabling finer discretization without prohibitive cost.

The StableOpt algorithm [Bogunovic et al., 2018] takes a different view on robustness by focusing directly on stability. At each round it selects $x_t = \arg\max_{x \in X} \min_{c \in C_t} UCB_t(x, c)$, where the uncertainty set C_t contains the most plausible contexts given past observations. Following Huang et al. [2024], this set is constructed by forming an interval for each context dimension centred on the

Table 1: Mean $\pm$ standard deviation of final cumulative regret across all experiments. The best and the second best results per experiment are in bold and underlined, respectively. The Average rank is computed over all experiments.

Method	Three-Hump Camel	Six-Hump Camel	Ackley	Hartmann	Complicated Hartmann
EDRBO (Ours)	**2.97 ± 0.51**	96.45 ± 85.24	**273.96 ± 62.91**	62.78 ± 12.55	144.22 ± 55.01
SBO-KDE	3.73 ± 1.09	105.51 ± 81.19	583.29 ± 160.31	75.04 ± 37.22	146.89 ± 62.65
UCB	7.13 ± 2.03	**86.63 ± 56.36**	547.97 ± 151.63	157.95 ± 58.07	222.12 ± 69.64
WDRBO	4.08 ± 1.11	186.29 ± 65.43	651.87 ± 175.31	**62.71 ± 14.62**	124.04 ± 44.9
ERBO	3.57 ± 1.06	117.26 ± 83.5	453.82 ± 127.19	63.43 ± 22.43	**106.47 ± 43.37**
DRBO-KDE	3.88 ± 1.31	107.48 ± 82.88	628.72 ± 171.38	113.36 ± 56.2	175.69 ± 58.22
StableOpt	56.07 ± 18.31	188.36 ± 65.19	1378.34 ± 388.01	369.69 ± 104.79	325.2 ± 92.84
DRBO-MMD	5.99 ± 1.96	141.89 ± 36.26	809.61 ± 217.32	222.34 ± 62.4	272.38 ± 76.65
DRBO-MMD Minmax	7.66 ± 2.99	287.83 ± 103.97	1085.92 ± 303.7	294.42 ± 79.69	275.39 ± 76.64

Method	Modified Branin	Newsvendor	Portfolio (Uniform)	Portfolio (Normal)	Average rank
EDRBO (Ours)	1118.8 ± 267.61	**8.03 ± 1.46**	509.97 ± 155.87	**660.8 ± 157.71**	**2.0**
SBO-KDE	1059.25 ± 309.47	13.54 ± 2.63	**448.39 ± 105.61**	679.85 ± 160.39	2.9
UCB	1419.76 ± 460.49	14.49 ± 5.5	532.93 ± 134.73	951.73 ± 293.95	4.8
WDRBO	1050.45 ± 233.47	15.37 ± 3.43	529.76 ± 138.79	784.0 ± 202.28	4.0
ERBO	**925.63 ± 266.35**	14.88 ± 3.45	478.87 ± 105.25	697.21 ± 169.05	2.6
DRBO-KDE	1304.87 ± 313.51	17.52 ± 4.05	657.2 ± 179.21	844.75 ± 224.77	5.0
StableOpt	4792.86 ± 1461.35	85.8 ± 24.17	1308.71 ± 366.79	1300.16 ± 354.31	8.9
DRBO-MMD	1502.19 ± 371.17	54.94 ± 15.76	890.06 ± 247.67	981.7 ± 255.29	6.8
DRBO-MMD Minmax	2153.51 ± 588.03	83.82 ± 24.36	920.08 ± 254.66	1110.15 ± 296.39	8.1

empirical mean and scaled by the empirical variance. StableOpt thus embodies a minimax perspective, explicitly guarding against the worst-case realization of context within statistically plausible ranges.

Finally, the WDRBO algorithm [Micheli et al., 2025], which introduces robustness by constructing a Wasserstein ambiguity set around the empirical distribution of the observed contexts. The radius of this set shrinks over time at the rate $\varepsilon_t = O(1/\sqrt{t})$, allowing the method to adapt as more data becomes available. A natural benchmark against this approach is ERBO [Micheli et al., 2025], a non-robust variant that simply sets $\varepsilon_t = 0$. In this setting, the algorithm directly maximizes the empirical risk over the observed contexts, $x_t = \arg\max_{x \in X} \mathbb{E}_{c \sim Q_t}[UCB_t(x, c)]$, thus serving as a context-aware but non-robust baseline.

We test the algorithms on a range of synthetic and real-world problems with various context distributions. Each algorithm runs for 200 iterations, including 5 starting points, at 10 different random seeds. Fig. 1 presents the mean and standard deviation of cumulative regret over these runs. The final values of cumulative regret are presented in Table 1. We adapt existing experimental setups [Huang et al., 2024, Micheli et al., 2025]. A detailed description of all experiments can be found in Appendix B.1. Additional experimental results are presented in Appendix B.3 and an ablation study is presented in Appendix C.

Empirical results provide significant insight into how EDRBO compares with existing approaches. The method achieves the best average ranking of 2.0 across all experiments. While EDRBO notably prevails over baseline methods on Three-Hump Camel, Ackley, Newsvendor, and one of the Portfolio functions, it matches existing state-of-the-art methods on the rest. These findings are also supported by running the Wilcoxon signed-rank test. At significance $\alpha = 10\%$ with Holm-Bonferroni correction, EDRBO results are significantly distinct from all baselines but SBO-KDE, WDRBO, and ERBO. Meaning, that EDRBO is on par with these methods. We report the exact test result in Appendix B.2.

The main advantage of EDRBO is its handling of highly non-convex landscapes with multiple local extrema. In this case, optimisation typically takes place near vanishing gradients regions. Lipschitz-based regularised methods, such as WDRBO, rely on a linear growth assumption, which makes them less stable in these areas. In more severe scenarios, where the cost landscape exhibits localised, sharply varying structures, e.g., Ackley function, the expected cost (μ) and the predictive uncertainty (σ) become strongly intertwined. In contrast to standard DRBO baselines that mostly regularise only first-order deviations, EDRBO's use of the Bures–Wasserstein distance yields a joint second-order stability guarantee.

One may observe this behaviour in the cumulative regret plots. Most noticeably for SBO-KDE and WDRBO baselines on both Camel and Newsvendor functions. While the proposed method produces

a stable and smooth cumulative regret curve, existing DRBO methods tend to produce a jagged, staircase-like result.

We also note experiments where the interaction between decision variables and context is strongly nonlinear and non-separable. A key example is the modified Branin function. Similarly, in the Portfolio optimization problems, the objective is not given analytically but instead arises as the posterior mean of a Gaussian process fitted to sampled data. This leads to nonstationary smoothness, local irregularities, and artifacts from finite data.

In conclusion, the proposed method is optimally suited for iterative decision-making under contextual uncertainty in environments where the objective function is: (i) highly non-convex, leading to frequent occurrences of stationary points, and (ii) sensitive to joint first- and second-order moment variations.

6 Conclusions

In this work, we present Ensemble Distributionally Robust Bayesian Optimisation, an ensemble surrogate algorithm for sequential decision making under context distributional uncertainty. The ensemble provides an interpretable and systematic mechanism to encode uncertainty regarding the context distribution. The method relies on robustifying the acquisition function using Lipschitz regularisation within the Wasserstein ambiguity set. In contrast with many existing distributionally robust approaches, ours yields a tractable optimisation problem in practice, while being capable of handling continuous context and having cumulative expected regret guarantees that improve on state-of-the-art methods. On a wide range of problems EDRBO attains superior to existing approaches performance, thus attesting to the theoretical results.

7 Impact Statement

Ensemble Distributionally Robust Bayesian Optimization (EDRBO) contributes to foundational optimization techniques with applications across scientific and engineering design problems. Our method addresses the challenge of optimizing objectives when the operating context differs from training conditions — a common problem in expensive black-box optimization. EDRBO could be applied to molecular design, materials discovery, particle physics simulations, and aerodynamic optimization, where distributional mismatch between estimated and true context distributions introduces uncertainty. The key property of our approach — formal worst-case performance guarantees via the DRO framework — makes BO more reliable in settings where robustness to distributional shift is critical.

Improved optimization efficiency in scientific and engineering design could accelerate innovation and reduce computational costs. In drug discovery and climate science applications, this could support progress on high-impact problems. However, we note that benefits will likely accrue primarily to well-resourced institutions with access to expensive simulators, potentially reinforcing existing technological disparities. A more subtle risk involves false confidence: the worst-case guarantees provided by DRO may instill unwarranted trust if assumptions about the context distribution (e.g., choice of divergence metric) are not validated carefully for the specific application. Additionally, as with any optimization tool, EDRBO could be misused in harmful applications, though this reflects broader dual-use concerns rather than specific vulnerabilities.

To maximize benefits and mitigate risks, we encourage: (i) research on methods to help practitioners select appropriate ambiguity sets for their domains; (ii) collaboration with domain scientists to validate whether DRO assumptions hold in real-world design problems; (iii) development of diagnostic tools to assess when distributional mismatch is significant; and (iv) documentation of assumptions and limitations when EDRBO is applied to new domains. Publishing detailed case studies from scientific applications would help the community better understand real-world impact and appropriate use.

Acknowledgements

The work was supported by the grant for research centers in the field of AI provided by the Ministry of Economic Development of the Russian Federation in accordance with the agreement 000000C313925P4E0002 and the agreement with HSE University № 139-15-2025-009. The compu-

tation for this research was performed using the computational resources of HPC facilities at HSE University [Kostenetskiy et al., 2021].

References

Mark J Abraham and Jill E Gready. Optimization of parameters for molecular dynamics simulation using smooth particle-mesh Ewald in GROMACS 4.5. *Journal of computational chemistry*, 32(9): 2031–2040, 2011.

Martial Agueh and Guillaume Carlier. Barycenters in the Wasserstein space. *SIAM Journal on Mathematical Analysis*, 43(2):904–924, 2011.

Stéphane Alarie, Charles Audet, Aïmen E Gheribi, Michael Kokkolaras, and Sébastien Le Digabel. Two decades of blackbox optimization applications. *EURO Journal on Computational Optimization*, 9:100011, 2021.

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 33*, 2020. URL http://arxiv.org/abs/1910.06403.

A Baranov, E Burnaev, D Derkach, A Filatov, N Klyuchnikov, O Lantwin, F Ratnikov, A Ustyuzhanin, and A Zaitsev. Optimising the Active Muon Shield for the SHiP Experiment at CERN. *Journal of Physics: Conference Series*, 934(1):012050, dec 2017. doi: 10.1088/1742-6596/934/1/012050. URL https://dx.doi.org/10.1088/1742-6596/934/1/012050.

Matteo Barbetti. Lamarr: Lhcb ultra-fast simulation based on machine learning models deployed within gauss, 2024. URL https://arxiv.org/abs/2303.11428.

Ilija Bogunovic, Jonathan Scarlett, Stefanie Jegelka, and Volkan Cevher. Adversarially robust optimization with gaussian processes. *Conference on Neural Information Processing Systems (NeurIPS)*, 2018.

Alexey Boldyrev, Denis Derkach, Fedor Ratnikov, and Andrey Shevelev. Deep learning approaches for lhcb ecal reconstruction. *EPJ Web of Conf.*, 295:09008, 2024. doi: 10.1051/epjconf/202429509008. URL https://doi.org/10.1051/epjconf/202429509008.

Alexey Boldyrev, Fedor Ratnikov, and A. Shevelev. An approach to finding a robust deep learning model. *IEEE Access*, 13:102390–102406, 2025. URL https://api.semanticscholar.org/CorpusID:278886078.

Stephen Boyd, Enzo Busseti, Steven Diamond, Ronald N. Kahn, Kwangmoo Koh, Peter Nystrup, and Jan Speth. Multi-period trading via convex optimization, 2017. URL https://arxiv.org/abs/1705.00109.

Adam D. Bull. Convergence Rates of Efficient Global Optimization Algorithms. *Journal of Machine Learning Research*, 12(88):2879–2904, 2011. URL http://jmlr.org/papers/v12/bull11a.html.

Sait Cakmak, Raul Astudillo Marban, Peter Frazier, and Enlu Zhou. Bayesian optimization of risk measures. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 20130–20141. Curran Associates, Inc., 2020.

Antonio Candelieri, Andrea Ponti, and Francesco Archetti. Gaussian process regression over discrete probability measures: on the non-stationarity relation between euclidean and wasserstein squared exponential kernels, 2022. URL https://arxiv.org/abs/2212.01310.

Michael B. Chang, Tomer D. Ullman, Antonio Torralba, and Joshua B. Tenenbaum. A Compositional Object-Based Approach to Learning Physical Dynamics. *CoRR*, abs/1612.00341, 2016. URL http://arxiv.org/abs/1612.00341.

Sayak Ray Chowdhury and Aditya Gopalan. On kernelized multi-armed bandits. *CoRR*, abs/1704.00445, 2017. URL http://arxiv.org/abs/1704.00445.

Jonas Degrave, Michiel Hermans, Joni Dambre, and Francis Wyffels. A Differentiable Physics Engine for Deep Learning in Robotics. *CoRR*, abs/1611.01652, 2016. URL http://arxiv.org/abs/1611.01652.

T. Dorigo. Toward the end-to-end optimization of particle physics instruments with differentiable programming. *Reviews in Physics*, 10:100085, 2023. ISSN 2405-4283. doi: https://doi.org/10.1016/j.revip.2023.100085. URL https://www.sciencedirect.com/science/article/pii/S2405428323000047.

Jacob Feldman. Equivalence and perpendicularity of gaussian processes. *Pacific Journal of Mathematics*, 8(4):699–708, 1958.

Rui Gao, Xi Chen, and Anton J. Kleywegt. Wasserstein distributionally robust optimization and variation regularization, 2020. URL https://arxiv.org/abs/1712.06050.

Jaroslav Hájek. On a property of normal distributions of any stochastic process. *Czechoslovak Mathematical Journal*, 8(3):610–618, 1958.

Xiaobin Huang, Lei Song, Ke Xue, and Chao Qian. Stochastic bayesian optimization with unknown continuous context distribution via kernel density estimation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(11):12635–12643, Mar. 2024. doi: 10.1609/aaai.v38i11.29158. URL https://ojs.aaai.org/index.php/AAAI/article/view/29158.

Hisham Husain, Vu Nguyen, and Anton van den Hengel. Distributionally robust bayesian optimization with φ-divergences, 2023. URL https://arxiv.org/abs/2203.02128.

Johannes Kirschner, Ilija Bogunovic, Stefanie Jegelka, and Andreas Krause. Distributionally robust bayesian optimization. *International Conference on Artificial Intelligence and Statistics*, pages 2174–2184, 2020.

Chun-Wa Ko, Jon Lee, and Maurice Queyranne. An exact algorithm for maximum entropy sampling. *Operations Research*, 43(4):684–691, 1995.

P. S. Kostenetskiy, R. A. Chulkevich, and V. I. Kozyrev. Hpc resources of the higher school of economics. *Journal of Physics: Conference Series*, 1740(1):012050, jan 2021. doi: 10.1088/1742-6596/1740/1/012050. URL https://dx.doi.org/10.1088/1742-6596/1740/1/012050.

Daniel Kuhn, Peyman Mohajerin Esfahani, Viet Anh Nguyen, and Soroosh Shafieezadeh-Abadeh. Wasserstein distributionally robust optimization: Theory and applications in machine learning, 2024. URL https://arxiv.org/abs/1908.08729.

Francesco Micheli, Efe C. Balta, Anastasios Tsiamis, and John Lygeros. Wasserstein distributionally robust bayesian optimization with continuous context, 2025. URL https://arxiv.org/abs/2503.20341.

Tigran Ramazyan, Mikhail Hushchyn, and Denis Derkach. *Global Optimisation of Black-Box Functions with Generative Models in the Wasserstein Space*. IOS Press, October 2024. ISBN 9781643685489. doi: 10.3233/faia240765. URL http://dx.doi.org/10.3233/FAIA240765.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 11 2005. ISBN 9780262256834. doi: 10.7551/mitpress/3206.001.0001. URL https://doi.org/10.7551/mitpress/3206.001.0001.

Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, 178(11):852–867, 2008.

Niranjan Srinivas, Andreas Krause, Sham M. Kakade, and Matthias W. Seeger. Information-theoretic regret bounds for gaussian process optimization in the bandit setting. *IEEE Transactions on Information Theory*, 58(5):3250–3265, May 2012. ISSN 1557-9654. doi: 10.1109/tit.2011.2182033. URL http://dx.doi.org/10.1109/TIT.2011.2182033.

Matthew Staib and Stefanie Jegelka. Distributionally robust optimization and generalization in kernel methods. *Advances in Neural Information Processing Systems*, pages 9131–9141, 2019.

Sebastian Shenghong Tay, Chuan Sheng Foo, Urano Daisuke, Richalynn Leong, and Bryan Kian Hsiang Low. Efficient distributionally robust bayesian optimization with worst-case sensitivity. *International Conference on Machine Learning*, pages 21180–21204, 2022.

Ensemble Distributionally Robust Bayesian Optimisation: Supplementary Materials

A. Missing Proofs 13

A.1 Regression confidence guarantees 13
A.2 Posterior radius Lipschitzness 14
A.3 Posterior radius bounds 15
A.4 Regret bounds 17

B. Experiment details 19

B.1 Problems definition 19
B.2 Wilcoxon signed-rank test 20
B.3 Additional empirical results 21

C. Ablation study 22

A Missing proofs

A.1 Regression confidence guarantees

To prove Lemma 3.1 we require the following lemma:

Lemma A.1 (Theorem 2 by Chowdhury and Gopalan [2017]). *Let $k : \Xi \times \Xi \to \mathbb{R}$ be the positive semidefinite kernel defining the GP surrogate, with RKHS $\mathcal{H}_k$ and norm $\|\cdot\|_{\mathcal{H}_k}$. Let $f : \Xi \to \mathbb{R}$, $f \in \mathcal{H}_k$, $\|f\|_{\mathcal{H}_k} \leq B_m$, and let ϵ_t be R-sub-Gaussian. Let $\delta \in (0, 1)$ be a probability of failure. Then with probability at least $1 - \delta$, $\forall \xi \in \Xi$ and $\forall t \geq 1$*

$$|\mu_{m,t}(\xi) - f(\xi)| \leq \beta_{m,t}\sigma_{m,t}(\xi), \tag{19}$$

with

$$\beta_{m,t} = \lambda^{\frac{1}{2}} B_m + R\sqrt{\gamma_{m,t} + 2\log(1/\delta)} \tag{20}$$

Lemma A.2 (Lemma 3.1 in the main text). *Let $\Xi \in \mathbb{R}^d$ and $f : \Xi \to \mathbb{R}$. Each expert assumes that $f \in \mathcal{H}_m$, with $\|f\|_{\mathcal{H}_m} \leq B_m$ and let ϵ_t be R-sub-Gaussian. Let $\delta \in (0, 1)$ be a probability of failure. With probability at least $1 - \delta$, $\forall \xi \in \Xi$ and $\forall t \geq 1$:*

$$|\mu_t(\xi) - f(\xi)| \leq \beta_t\sigma_t(\xi), \tag{21}$$

with

$$\sigma_t(\xi) = \frac{1}{M}\sum_{m=1}^{M} \sigma_{m,t}(\xi) \tag{22}$$

and

$$\beta_t = \max_m \lambda^{\frac{1}{2}} B_m + R\sqrt{\gamma_{m,t} + 2\log(M/\delta)}. \tag{23}$$

Proof.

$$|\mu_t(\xi) - f(\xi)| = \left| \frac{1}{M} \sum_{m=1}^{M} \mu_{m,t}(\xi) - \frac{1}{M} \sum_{m=1}^{M} f(\xi) \right| = \tag{24}$$

$$= \left| \frac{1}{M} \sum_{m=1}^{M} (\mu_{m,t}(\xi) - f(\xi)) \right| \leq \tag{25}$$

$$\overset{\text{(i)}}{\leq} \frac{1}{M} \sum_{m=1}^{M} |\mu_{m,t}(\xi) - f(\xi)| \leq \tag{26}$$

$$\overset{\text{(ii)}}{\leq} \frac{1}{M} \sum_{m=1}^{M} \beta_{m,t}\sigma_{m,t}(\xi) \leq \tag{27}$$

$$\overset{\text{(iii)}}{\leq} \beta_t \sigma_t(\xi), \tag{28}$$

where (i) follows from sub-additivity of absolute value, (ii) follows from applying Lemma A.1 that holds with probability at least $1 - \frac{\delta}{M}$ for each base expert, and (iii) follows from taking $\beta_t = \max_m \beta_{m,t}$. □

A.2 Posterior radius Lipschitzness

Lemma A.3 (Lemma 3 by Micheli et al. [2025])**.** *Let $\delta \in (0,1)$. Then $\sigma_{m,t}$ is $\lambda^{-\frac{1}{2}} L_m(x)$-Lipschitz; with probability at least $1-\delta$, $\mu_{m,t}$ is $\lambda^{-\frac{1}{2}}\beta_{m,t}L_m(x)$-Lipschitz. Therefore, since $\beta_{m,t} \geq 1$, with probability at least $1-\delta$, $\mu_{m,t} + \beta_{m,t}\sigma_{m,t}$ is $2\lambda^{-\frac{1}{2}}\beta_{m,t}L_m(x)$-Lipschitz.*

Consequently, by applying this result to our surrogate, μ_t is $\max_m \lambda^{-\frac{1}{2}}\beta_{m,t}L_m(x)$ - Lipschitz with probability at least $1-\delta$ for some probability of failure $\delta \in (0,1)$; σ_t is $\max_m \lambda^{-\frac{1}{2}} L_m(x)$ - Lipschitz, and $\mu_t + \beta_t\sigma_t$ is $L_t(x)$-Lipschitz with $L_t(x) = \max_m 2\lambda^{-\frac{1}{2}}\beta_{m,t}L_{m,t}(x) = 2\lambda^{-\frac{1}{2}}\beta_t \max_m L_{m,t}(x)$. This also gives the following property of the ensemble posterior radius:

Corollary A.4. *Let $\delta \in (0,1)$. Then, with probability at least $1-\delta$, $\hat{\varepsilon}_t$ is $L_t(x)\sqrt{2}$-Lipschitz.*

Proof. Without loss of generality fix $t \geq 1$, $x \in \mathcal{X}$, and m, for which $\hat{\varepsilon}_t$ is computed. Put $\psi(\xi) = \mu_{m,t}(\xi) - \mu_t(\xi)$ and $\varphi(\xi) = \sigma_{m,t}(\xi) - \sigma_t(\xi)$ for some $\xi \in \Xi$. Then $\hat{\varepsilon}_t(\xi) = \sqrt{\psi^2(\xi) + \varphi^2(\xi)} = \|(\psi(\xi), \varphi(\xi))\|_2$. Consider each component $\forall \xi, \xi \in \Xi$ and apply Lemma A.3:

$$|\psi(\xi) - \psi(\xi')| = |(\mu_{m,t}(\xi) - \mu_t(\xi)) - (\mu_{m,t}(\xi') - \mu_t(\xi'))| \leq \tag{29}$$

$$\leq |\mu_{m,t}(\xi) - \mu_{m,t}(\xi')| + |\mu_t(\xi) - \mu_t(\xi')| \leq \tag{30}$$

$$\leq 2\lambda^{-\frac{1}{2}}\beta_t \max_m L_m(x)\|\xi - \xi'\| = L_t(x)\|\xi - \xi'\|, \tag{31}$$

and identically for φ

$$|\varphi(\xi) - \varphi(\xi')| = |(\sigma_{m,t}(\xi) - \sigma_t(\xi)) - (\sigma_{m,t}(\xi') - \sigma_t(\xi'))| \leq \tag{32}$$

$$\leq |\sigma_{m,t}(\xi) - \sigma_{m,t}(\xi')| + |\sigma_t(\xi) - \sigma_t(\xi')| \leq \tag{33}$$

$$\leq 2\lambda^{-\frac{1}{2}} \max_m L_m(x)\|\xi - \xi'\| \leq L_t(x)\|\xi - \xi'\|. \tag{34}$$

Since norm is 1-Lipschitz and $\beta_t \geq 1$, combining the two gives

$$|\hat{\varepsilon}_t(\xi) - \hat{\varepsilon}_t(\xi')| \leq \sqrt{(\psi(\xi) - \psi(\xi'))^2 + (\varphi(\xi) - \varphi(\xi'))^2} \leq \tag{35}$$

$$\leq \sqrt{(L_t(x)\|\xi - \xi'\|)^2 + (L_t(x)\|\xi - \xi'\|)^2} \leq \tag{36}$$

$$\leq L_t(x)\sqrt{2}\|\xi - \xi'\|. \tag{37}$$

□

A.3 Posterior radius bounds

Theorem A.5 (Theorem 4.2 in main text)**.** *The distributionally robust learner optimisation problem is given by:*

$$x_t = \arg\max_{x \in \mathcal{X}} \mathbb{E}_{c \sim Q_t} \left[\mu_t(x, c) + \beta_t \sigma_t(x, c) + \frac{1}{\sqrt{2}} \hat{\varepsilon}_t(x, c) \right]. \tag{38}$$

Proof. Consider the non-robust learner objective as the following minimax problem:

$$x_t = \arg\max_{x \in \mathcal{X}} \inf_{Q \in \mathbb{B}_{\varepsilon_t}(Q_t)} \mathbb{E}_{c \sim Q_t} \left[\mu_t(x, c) + \beta_t \sigma_t(x, c) + \frac{1}{\sqrt{2}} \hat{\varepsilon}_t(x, c) \right]. \tag{39}$$

The posterior radius, $\hat{\varepsilon}_t$, is considered ensemble-uncertainty optimistic, weighted to preserve the UCB trade-off against the design variable x. By Lemma A.3 and Corollary A.4 the non-robust Eq. 39 is $2L_t(x)$-Lipschitz. By Lemma 4.1 we bound the inner optimisation problem, and hence robustify against the unknown contextual distribution using:

$$x_t = \arg\max_{x \in \mathcal{X}} \mathbb{E}_{c \sim Q_t} \left[\mu_t(x, c) + \beta_t \sigma_t(x, c) + \frac{1}{\sqrt{2}} \hat{\varepsilon}_t(x, c) \right] - 2\varepsilon_t L_t(x). \tag{40}$$

Since a Gaussian process surrogate is almost always equipped with a stationary kernel, the Lipschitz constant holds for both arguments. Meaning, $L_t(x) = L_t$, which makes the robustification term constant with respect x. The same holds for the base experts chosen for our study. Hence the maximiser of the robustified acquisition function, Eq. 40, is equivalent to the maximiser of the learner's problem at Q_t, Eq. 15:

$$x_t = \arg\max_{x \in \mathcal{X}} \mathbb{E}_{c \sim Q_t} \left[\mu_t(x, c) + \beta_t \sigma_t(x, c) + \frac{1}{\sqrt{2}} \hat{\varepsilon}_t(x, c) \right] - 2\varepsilon_t L_t(x) = \tag{41}$$

$$= \arg\max_{x \in \mathcal{X}} \mathbb{E}_{c \sim Q_t} \left[\mu_t(x, c) + \beta_t \sigma_t(x, c) + \frac{1}{\sqrt{2}} \hat{\varepsilon}_t(x, c) \right] - 2\varepsilon_t L_t = \tag{42}$$

$$= \arg\max_{x \in \mathcal{X}} \mathbb{E}_{c \sim Q_t} \left[\mu_t(x, c) + \beta_t \sigma_t(x, c) + \frac{1}{\sqrt{2}} \hat{\varepsilon}_t(x, c) \right]. \tag{43}$$

□

Lemma A.6. *Let $\delta \in (0, 1)$ be a probability of failure. Then with probability at least $1 - \delta$, $\forall t \geq 1, \forall \xi \in \Xi$ the posterior ensemble radius, $\hat{\varepsilon}_t(\xi)$, is bounded as follows*

$$\hat{\varepsilon}_t(\xi) \leq 2\sqrt{2}\beta_t \max_m \sigma_{m,t}(\xi). \tag{44}$$

Proof. First consider

$$|\mu_{m,t}(\xi) - \mu_t(\xi)| = |\mu_{m,t}(\xi) - f(\xi) + f(\xi) - \mu_t(\xi)| \leq \tag{45}$$

$$\overset{\text{(i)}}{\leq} |\mu_{m,t}(\xi) - f(\xi)| + |\mu_t(\xi) - f(\xi)| \leq \tag{46}$$

$$\overset{\text{(ii)}}{\leq} \beta_{m,t}\sigma_{m,t}(\xi) + \beta_t \sigma_t(\xi) \leq \tag{47}$$

$$\overset{\text{(iii)}}{\leq} \beta_t \sigma_{m,t}(\xi) + \beta_t \sigma_t(\xi) = \tag{48}$$

$$= \beta_t(\sigma_{m,t}(\xi) + \sigma_t(\xi)), \tag{49}$$

where (i) follows from sub-additivity of absolute value, (ii) follows by applying Lemma A.1 and Lemma 3.1, and (iii) follows from the definition of β_t, Eq. 6. Then

$$\hat{\varepsilon}_t(\xi) = \max_m \sqrt{(\mu_{m,t}(\xi) - \mu_t(\xi))^2 + (\sigma_{m,t}(\xi) - \sigma_t(\xi))^2} \leq \tag{50}$$

$$\overset{\text{(i)}}{\leq} \max_m \sqrt{\left(\beta_t(\sigma_{m,t}(\xi) + \sigma_t(\xi))\right)^2 + (\sigma^2_{m,t}(\xi) + \sigma^2_t(\xi))} \leq \tag{51}$$

$$\overset{\text{(ii)}}{\leq} \max_m \sqrt{2\beta_t^2 \left(\sigma^2_{m,t}(\xi) + \sigma^2_t(\xi)\right) + \beta_t^2(\sigma^2_{m,t}(\xi) + \sigma^2_t(\xi))} \leq \tag{52}$$

$$\overset{\text{(iii)}}{\leq} \max_m \sqrt{6\beta_t^2 \sigma^2_{m,t}(\xi)} \leq \tag{53}$$

$$\leq 2\sqrt{2}\beta_t \max_m \sigma_{m,t}(\xi), \tag{54}$$

where (i) follows from applying the bound from Eq.54 and the full square, (ii) follows from $(a+b)^2 \leq 2(a^2+b^2)\ \forall a, b$ and $\beta_t \geq 1$, and (iii) follows from the fact that the maximum is never less than the average. □

Lemma A.7. *$\forall m = 1, \ldots, M$ and $\forall \xi \in \Xi$ the posterior variance of a GP with regularisation parameter $\lambda > 0$ after $t \geq 1$ iterations is bounded as follows:*

$$\sigma^2_{m,t}(\xi) \geq \frac{1}{n_t + \lambda}. \tag{55}$$

Proof. Without loss of generality fix $t \geq 1$ and $m \in 1, \ldots, M$. For a positive semidefinite kernel $k_m : \Xi \times \Xi \to \mathbb{R}$ with RKHS $(\mathcal{H}_m, \langle \cdot, \cdot \rangle_{\mathcal{H}_m})$, let $\Phi_{m,t} = (k_m(\cdot, \xi_1), \ldots, k_m(\cdot, \xi_t))^T$ be the RKHS feature map for $\xi_1, \ldots, \xi_t \in \Xi$. Then the kernel matrix is given by $K_{m,t} = \Phi_{m,t}\Phi^T_{m,t}$ and the covariance operator $\Sigma_{m,t} = \Phi^T_{m,t}\Phi_{m,t}$.

First, we note that $\|\Phi_{m,t}\|_{\mathcal{H}_m} = \|\Phi^T_{m,t}\|_{\mathcal{H}_m}$ and $\forall a \in \mathbb{R}^t \ \ \|\Phi^T_{m,t}a\| = \langle a, \Phi_{m,t}\Phi^T_{m,t}a\rangle_{\mathcal{H}_m} = a^T K_{m,t} a$. So $\|\Phi_{m,t}\|^2_{\mathcal{H}_m} = \sup_{\|a\|=1} a^T K_{m,t} a = \lambda_{\max}(K_{m,t})$, where $\lambda_{\max}$ denotes the largest eigenvalue. Then $\forall g \in \mathcal{H}_m$

$$\langle g, \Sigma_{m,t} g\rangle_{\mathcal{H}_m} = \|\Phi_{m,t} g\|^2_{\mathcal{H}_m} \leq \lambda_{\max}(K_{m,t})\|g\|^2_{\mathcal{H}_m} \leq \operatorname{tr}(K_{m,t})\|g\|^2_{\mathcal{H}_m}, \tag{56}$$

where the last inequality follows from $K_{m,t} \succeq 0$. By the Cauchy-Schwartz inequality, $\forall \xi \in \Xi$

$$\begin{aligned}
\|k_m(\cdot, \xi)\|^2_{\mathcal{H}_m} &= \langle (\lambda I + \Sigma_{m,t})(\lambda I + \Sigma_{m,t})^{-1} k_m(\cdot, \xi), k_m(\cdot, \xi)\rangle_{\mathcal{H}_m} \leq && (57)\\
&\leq \langle (\lambda I + \Sigma_{m,t})(\lambda I + \Sigma_{m,t})^{-1} k_m(\cdot, \xi), (\lambda I + \Sigma_{m,t})^{-1} k_m(\cdot, \xi)\rangle^{\frac{1}{2}}_{\mathcal{H}_m} \cdot && (58)\\
&\qquad \cdot \langle (\lambda I + \Sigma_{m,t}) k_m(\cdot, \xi), k_m(\cdot, \xi)\rangle^{\frac{1}{2}}_{\mathcal{H}_m} = && (59)\\
&= \langle k_m(\cdot, \xi), (\lambda I + \Sigma_{m,t})^{-1} k_m(\cdot, \xi)\rangle^{\frac{1}{2}}_{\mathcal{H}_m} \cdot && (60)\\
&\qquad \langle k_m(\cdot, \xi), (\lambda I + \Sigma_{m,t}) k_m(\cdot, \xi)\rangle^{\frac{1}{2}}_{\mathcal{H}_m} = && (61)\\
&= \langle k_m(\cdot, \xi), (\lambda I + \Sigma_{m,t})^{-1} k_m(\cdot, \xi)\rangle^{\frac{1}{2}}_{\mathcal{H}_m} \cdot && (62)\\
&\qquad \left(\lambda \|k_m(\cdot, \xi)\|^2_{\mathcal{H}_m} + \|\Phi_{m,t}\, k_m(\cdot, \xi)\|^2_{\mathcal{H}_m}\right)^{\frac{1}{2}} \leq && (63)\\
&\leq \langle k_m(\cdot, \xi), (\lambda I + \Sigma_{m,t})^{-1} k_m(\cdot, \xi)\rangle^{\frac{1}{2}}_{\mathcal{H}_m} \cdot \|k_m(\cdot, \xi)\|_{\mathcal{H}_m} (\lambda + \operatorname{tr}(K_{m,t}))^{\frac{1}{2}}, && (64)
\end{aligned}$$

where the last inequality follows from applying Eq. 56. The posterior variance is given by $\sigma^2_{m,t}(\xi) = \langle k_m(\cdot, \xi), (\lambda I + \Sigma_{m,t})^{-1} k_m(\cdot, \xi)\rangle_{\mathcal{H}_m}$. This, together with the kernel unit-variance bound, $\forall \xi, \xi' \in \Xi,\ k_m(\xi, \xi') \leq 1$, yields the desired bound

$$\begin{aligned}
\sigma_{m,t}(\xi) &\geq \frac{\|k_m(\cdot, \xi)\|^2_{\mathcal{H}_m}}{\|k_m(\cdot, \xi)\|_{\mathcal{H}_m} \sqrt{\lambda + \operatorname{tr}(K_{m,t})}} = && (65)\\
&= \frac{\|k_m(\cdot, \xi)\|^2_{\mathcal{H}_m}}{\sqrt{\lambda + \operatorname{tr}(K_{m,t})}} \geq && (66)\\
&\geq \frac{1}{\sqrt{\lambda + t}}. && (67)
\end{aligned}$$

□

A.4 Regret bounds

Lemma A.8. *Let $\delta \in (0,1)$ be a probability of failure and $\exists\, \ell \geq 1$. Then with probability at least $1-\delta$, $\forall x \in \mathcal{X}$ and $\forall t \geq 1$ the instantaneous expected regret, Eq. 16, can be bounded as follows:*

$$r_t \leq (4+\ell)\beta_t \max_m \mathbb{E}_{c\sim P}[\sigma_{m,t}(x_t, c)]. \tag{68}$$

Proof.

$$\begin{aligned}
r_t &= \mathbb{E}_{c\sim P}[f(x^*, c)] - \mathbb{E}_{c\sim P}[f(x_t, c)] \leq && (69)\\
&\overset{\text{(i)}}{\leq} \mathbb{E}_{c\sim P}[\mu_t(x^*, c) + \beta_t\sigma_t(x^*, c)] - \mathbb{E}_{c\sim P}[\mu_t(x_t, c) - \beta_t\sigma_t(x_t, c)] \leq && (70)\\
&\overset{\text{(ii)}}{\leq} \mathbb{E}_{c\sim Q_t}[\mu_t(x^*, c) + \beta_t\sigma_t(x^*, c)] + \varepsilon_t L_t(x^*) + \frac{1}{\sqrt{2}}\mathbb{E}_{c\sim Q_t}[\hat{\varepsilon}_t(x^*, c)] - && (71)\\
&\qquad - \mathbb{E}_{c\sim P}[\mu_t(x_t, c) - \beta_t\sigma_t(x_t, c)] \\
&\overset{\text{(iii)}}{\leq} \mathbb{E}_{c\sim Q_t}[\mu_t(x_t, c) + \beta_t\sigma_t(x_t, c) + \hat{\varepsilon}_t(x_t, c)/\sqrt{2}] + \varepsilon_t L_t(x^*) - && (72)\\
&\qquad - \mathbb{E}_{c\sim P}[\mu_t(x_t, c) - \beta_t\sigma_t(x_t, c)] = \\
&\overset{\text{(iv)}}{\leq} \mathbb{E}_{c\sim P}[\mu_t(x_t, c) + \beta_t\sigma_t(x_t, c) + \hat{\varepsilon}_t(x_t, c)/\sqrt{2}] + \varepsilon_t(L_t(x^*) + 2L_t(x_t)) - && (73)\\
&\qquad - \mathbb{E}_{c\sim P}[\mu_t(x_t, c) - \beta_t\sigma_t(x_t, c)] = \\
&\overset{\text{(v)}}{=} 2\beta_t\mathbb{E}_{c\sim P}[\sigma_t(x_t, c)] + \frac{1}{\sqrt{2}}\mathbb{E}_{c\sim Q_t}[\hat{\varepsilon}_t(x_t, c)] + 3\varepsilon_t L_t \leq && (74)\\
&\overset{\text{(vi)}}{\leq} 2\beta_t\mathbb{E}_{c\sim P}[\sigma_t(x_t, c)] + 2\beta_t \max_m \mathbb{E}_{c\sim P}[\sigma_{m,t}(x_t, c)] + 3\varepsilon_t(2\lambda^{-\frac{1}{2}}\beta_t \max_m L_m) && (75)\\
&\overset{\text{(vii)}}{\leq} 4\beta_t \max_m \mathbb{E}_{c\sim P}[\sigma_{m,t}(x_t, c)] + \ell\beta_t \max_m \mathbb{E}_{c\sim P}[\sigma_{m,t}(x_t, c)] = && (76)\\
&= (4+\ell)\beta_t \max_m \mathbb{E}_{c\sim P}[\sigma_{m,t}(x_t, c)], && (77)
\end{aligned}$$

where (i) follows from Lemma 3.1, (ii) follows from Lemma 4.1 and adding the weighted posterior radius as it is non-negative, (iii) follows from x_t being the maximiser of 15, (iv) follows from adding and subtracting $2\varepsilon_t L_t(x_t)$ and applying Lemma 4.1, (v) follows from applying the fact that for stationary kernels the Lipschitz constant holds for both arguments (see proof of Theorem A.5), (vi) follows from applying Lemma A.6 and multiplying by $\beta_t \geq 1$, (vii) follows from Lemma A.7 and, following Micheli et al. [2025], taking $\varepsilon_t = \mathcal{O}(t^{-\frac{1}{2}})$ and, following Chowdhury and Gopalan [2017], taking $\lambda = 1 + \frac{2}{T}$: $\exists\, \ell \geq 1, \ell \perp t : \sigma_t(\xi) \gtrsim \varepsilon_t L_t/(\beta_t \ell)$. □

Theorem A.9 (Theorem 4.3 in main text)**.** *Let $\delta \in (0,1)$ be a probability of failure and $\exists\, \ell \geq 1$. Then with probability at least $1-\delta$, the cumulative instantaneous expected regret of Algorithm 1 after T iterations is bounded by:*

$$R_T \leq 2(4+\ell)\beta_T\sqrt{T\left(\gamma_T + 2\log\frac{6}{\delta}\right)}, \tag{78}$$

which results in the $\mathcal{O}(\gamma_T\sqrt{T})$ general order of cumulative regret, where γ_T is the ensemble maximum information gain.

Proof.

$$R_T = \sum_{t=1}^{T} r_t \leq \tag{79}$$

$$\overset{\text{(i)}}{\leq} \sum_{t=1}^{T} (4+\ell)\beta_t \max_m \mathbb{E}_{c\sim P}[\sigma_{m,t}(x_t, c)] \leq \tag{80}$$

$$\overset{\text{(ii)}}{\leq} (4+\ell)\beta_T \sum_{t=1}^{T} \max_m \mathbb{E}_{c\sim P}[\sigma_t(x_t, c)] = \tag{81}$$

$$\overset{\text{(iii)}}{\leq} (4+\ell)\beta_T \sqrt{T \sum_{t=1}^{T} \left(\max_m \mathbb{E}_{c\sim P}\left[\sigma_{m,t}(x_t, c)\right]\right)^2} \leq \tag{82}$$

$$\overset{\text{(iv)}}{\leq} (4+\ell)\beta_T \sqrt{T \sum_{t=1}^{T} \max_m \mathbb{E}_{c\sim P}\left[\sigma^2_{m,t}(x_t, c)\right]} \leq \tag{83}$$

$$\overset{\text{(v)}}{\leq} (4+\ell)\beta_T \sqrt{T \left(2\sum_{t=1}^{T} \max_m \sigma^2_{m,t}(x_t, c_t) + 8\log\frac{6}{\delta}\right)} \leq \tag{84}$$

$$\overset{\text{(vi)}}{\leq} (4+\ell)\beta_T \sqrt{T \left(2\sum_{t=1}^{T} \max_m 2\log\left(1+\sigma^2_{m,t}(x_t, c_t)\right) + 8\log\frac{6}{\delta}\right)} \leq \tag{85}$$

$$\overset{\text{(vii)}}{\leq} (4+\ell)\beta_T \sqrt{T \left(4\sum_{t=1}^{T} \gamma_t + 8\log\frac{6}{\delta}\right)} \leq \tag{86}$$

$$\overset{\text{(viii)}}{\leq} 2(4+\ell)\beta_T \sqrt{T \left(\gamma_T + 2\log\frac{6}{\delta}\right)} \tag{87}$$

where (i) follows from Lemma A.8, (ii) follows from $\beta_T \geq \beta_t,\ \forall t \geq 1$, (iii) follows from the Cauchy-Schwartz inequality, (iv) follows from the Jensen inequality, (v) follows from Lemma 7 from [Kirschner et al., 2020] and noting the assumption that $\forall m, \forall \xi, \xi' \in \Xi, k_m(\xi, \xi') \leq 1$, (vi) follows from $\forall z : |z| \leq a, z \leq 2a\log(1+z)$, (vii) and (viii) follow from the definition of maximum information gain and ensemble maximum information gain. To establish the general order:

$$R_T \leq 2(4+\ell)\beta_T \sqrt{T \left(\gamma_T + 2\log\frac{6}{\delta}\right)} \leq \tag{88}$$

$$\overset{\text{(ix)}}{\leq} 2(4+\ell) \left(\max_m \left(B_m + \sigma_\epsilon \sqrt{\gamma_{m,T} + 2\log(M/\delta)}\right)\right) \sqrt{T \left(\gamma_T + 2\log\frac{6}{\delta}\right)} \leq \tag{89}$$

$$\overset{\text{(x)}}{\leq} 2(4+\ell) \left(B + \sigma_\epsilon \sqrt{\gamma_T + 2\log(M/\delta)}\right) \sqrt{T \left(\gamma_T + 2\log\frac{6}{\delta}\right)} = \tag{90}$$

$$= \mathcal{O}(\gamma_T \sqrt{T}), \tag{91}$$

where (ix) follows from the definition of β_T - Eq. 6, and (x) follows from taking $B = \max_m B_m$ and $\gamma_T = \max_m \gamma_{m,T}$.

□

B Experiment details

Our code is available at **HERE**.

All experimental results can be produced my running the 'main_multiple_runs.py' script using Python 3.11.13. All required packages are provided in the 'requirements.txt' file. A 'Results' directory will be created where all obtained results will be saved.

All experiments were run using two CPU cores of Intel Xeon Gold 6152 2.1 GHz and one GPU core of NVIDIA Tesla V100 32 Gb NVLink. We provide the mean and the standard deviation (over all experiments) of average computing time per each method in seconds in Table 2.

We adapt the experimental setup of Micheli et al. [2025] (available at https://github.com/frmicheli/WDRBO), who in their turn, adapted the experimental setup of Huang et al. [2024] (available at https://github.com/lamda-bbo/sbokde).

We adapt implementations UCB, SBO-KDE, DRBO-KDE, DRBO-MMD, DRBO-MMD Minmax, StableOpt, WDRBO, and ERBO. All algorithms, including our EDRBO, are based on the BOTorch Python package Balandat et al. [2020]. Baseline methods are equipped with the original, default RBF kernel GP regression surrogate model. All acquisition functions are optimised using the BOTorch optimiser.

For the test problems, we reuse the Three-Hump Camel, Six-Hump Camel, Ackley, Hartmann, Hartmann (Complicated), Modified Branin, Newsvendor, Portfolio (Normal) and Portfolio (Uniform) functions.

Table 2: Mean and standard deviation (over all experiments) of average computing time in seconds.

Method	**Mean $\pm$ Std (sec.)**
EDRBO (Ours)	843.16 ± 330.48
SBO-KDE	381.21 ± 434.97
UCB	216.15 ± 287.53
WDRBO	416.98 ± 452.97
ERBO	401.09 ± 291.74
DRBO-KDE	3369.58 ± 2653.04
StableOpt	411.38 ± 494.65
DRBO-MMD	2026.80 ± 951.64
DRBO-MMD Minmax	828.01 ± 702.28

B.1 Problem definition

- **Three-Hump Camel.** For $x \in [0,1]^2$ and $c \sim \mathcal{N}(0.5, 0.2^2)$ with both dimensions clipped to $[0,1]$

$$f(x,c) = 2x^2 - 1.05x^4 + \frac{x^6}{6} + xc + c^2. \tag{92}$$

- **Six-Hump Camel.** For $x \in [0,1]^2$ and $c \sim \mathcal{N}(0.6, 0.2^2)$ with both dimensions clipped to $[0,1]$

$$f(x,c) = \left(4 - 2.1x_1^2 + \frac{x_1^4}{3}\right) x_1^2 + x_1 x_2 + \left(-4 + 4x_2^2\right) x_2^2. \tag{93}$$

- **Ackley.** For $x \in [0,1]^2$ and $c \sim \mathcal{N}(0.5, 0.2^2)$ clipped to $[0,1]$

$$f(x,c) = a \exp\left(-b\sqrt{\frac{1}{3}\left(\tilde{x}_1^2 + \tilde{x}_2^2 + \tilde{c}^2\right)}\right) - \tag{94}$$

$$- \exp\left(\sqrt{\frac{1}{3}\left(\cosh h\tilde{x}_1 + \cosh h\tilde{x}_2 + \cosh hc\right)}\right) + a + e, \tag{95}$$

where $a = 20, b = 0.2, h = 2\pi, \tilde{x}_i = 65.536x_i - 32.768, \tilde{c} = 65.536c - 32.768$.

- **Hartmann.** For $x \in [0,1]^5$ and $c \sim \mathcal{N}(0.5, 0.2^2)$ clipped to $[0,1]$, $z = [x_1, \ldots, x_5, c]^T$

$$f(x,c) = \sum_{i=1}^{4} \alpha_i \exp\left(\sum_{j=1}^{6} A_{ij}\left(z_j - P_{ij}\right)^2\right), \tag{96}$$

where $\alpha = [1.0, 2.0, 3.0, 3.2]^T$, $A = \begin{bmatrix} 10 & 3 & 17 & 3.50 & 1.7 & 8 \\ 0.05 & 10 & 17 & 0.1 & 8 & 14 \\ 3 & 3.5 & 1.7 & 10 & 17 & 8 \\ 17 & 8 & 0.05 & 10 & 0.1 & 14 \end{bmatrix}$, and $P = 10^{-4} \begin{bmatrix} 1312 & 1696 & 5569 & 124 & 8283 & 5886 \\ 2329 & 4135 & 8307 & 3736 & 1004 & 9991 \\ 2348 & 1451 & 3522 & 2883 & 3047 & 6650 \\ 4047 & 8828 & 8732 & 5743 & 1091 & 381 \end{bmatrix}$.

- **Complicated Hartmann.** For $x \in [0,1]^5$ and $c \in [0,1]$, which is a mixture of six Gaussians $\mathcal{N}(0.1, 0.02^2), \mathcal{N}(0.3, 0.075^2), \mathcal{N}(0.4, 0.1^2), \mathcal{N}(0.5, 0.1^2), \mathcal{N}(0.7, 0.075^2), \mathcal{N}(0.8, 0.03^2)$, and two Cauchy distributions $Cauchy(0.2, 0.02), Cauchy(0.8, 0.02)$, clipped to $[0,1]$.
- **Modified Branin.** For $x \in [0,1]^2$ and $c \sim \mathcal{N}(0.5, 0.2^2)$ with both dimensions clipped to $[0,1]$

$$f(x,c) = -\sqrt{h(15x_1 - 5, 15c_1)h(15c_2 - 5, 15x_2)}, \tag{97}$$

where

$$h(u,v) = a\left(v - bu^2 + cu - r\right)^2 + s(1-t) * \cos(u) + s, \tag{98}$$

and $a = 1$, $b = 5.1/(4\pi^2)$, $c = 5/\pi$, $r = 6$, $s = 10$, $t = 1/(8\pi)$.
- **Newsvendor.** For $x \in [0,1]^2$

$$f(x,c) = 9\min\{x,c\} + \max\{0, x-c\} - 5x, \tag{99}$$

where c $\sim$ Burr Type XII distribution with PDF

$$p(c;\alpha,\beta) = \alpha\beta\frac{c^{\alpha-1}}{(1+c^{\alpha})^{\beta+1}} \tag{100}$$

with $\alpha = 2$, $\beta = 20$ and c being clipped to $[0,1]$ when it is out of the range.
- **Portfolio Optimisation.** For $x \in [0,1]^3$ and $c \sim \mathrm{U}(0,1)$ (Uniform case) or $c \sim \mathcal{N}(0.5, 0.2^2)$ clipped to $[0,1]$ (Normal case), the function is evaluated as the posterior mean of a Gaussian process fitted on 3,000 samples generated from CVXPortfolio Boyd et al. [2017], which were re-scaled to $[0,1]$ after the generation and provided by Cakmak et al. [2020].

B.2 Wilcoxon signed-rank test

Table 3: Pairwise comparison of EDRBO against each baseline on final cumulative regret over the nine experiments ($n = 9$ paired observations per comparison). Differences are computed as baseline minus EDRBO, so positive values favour EDRBO. Wilcoxon signed-rank p-values are Holm-Bonferroni-corrected over the eight comparisons; paired t-test p-values are uncorrected. A Friedman omnibus test across all nine methods rejects the hypothesis of equal performance ($\chi^2_F = 58.13$, $p = 1.08 \times 10^{-9}$).

			Wilcoxon signed-rank					
Baseline	**Mean diff.**	**Median diff.**	W statistic	p value	p_{Holm}	**Significant at** $\alpha = 10\%$	**Paired** t**-test** p **value**	**Significant at** $\alpha = 10\%$
DRBO-KDE	108.39	50.58	0	0.0039	0.0313	Yes	0.0261	Yes
DRBO-MMD	222.57	159.56	0	0.0039	0.0313	Yes	0.0074	Yes
DRBO-MMD Minmax	371.20	231.64	0	0.0039	0.0313	Yes	0.0127	Yes
StableOpt	769.69	306.91	0	0.0039	0.0313	Yes	0.0795	Yes
UCB	118.08	77.90	3	0.0195	0.0781	Yes	0.0283	Yes
WDRBO	58.95	7.34	12	0.2500	0.7500	No	0.2185	No
SBO-KDE	26.39	5.51	15	0.4258	0.8516	No	0.4935	No
ERBO	−1.87	0.65	21	0.9102	0.9102	No	0.9548	No

B.3 Additional empirical results

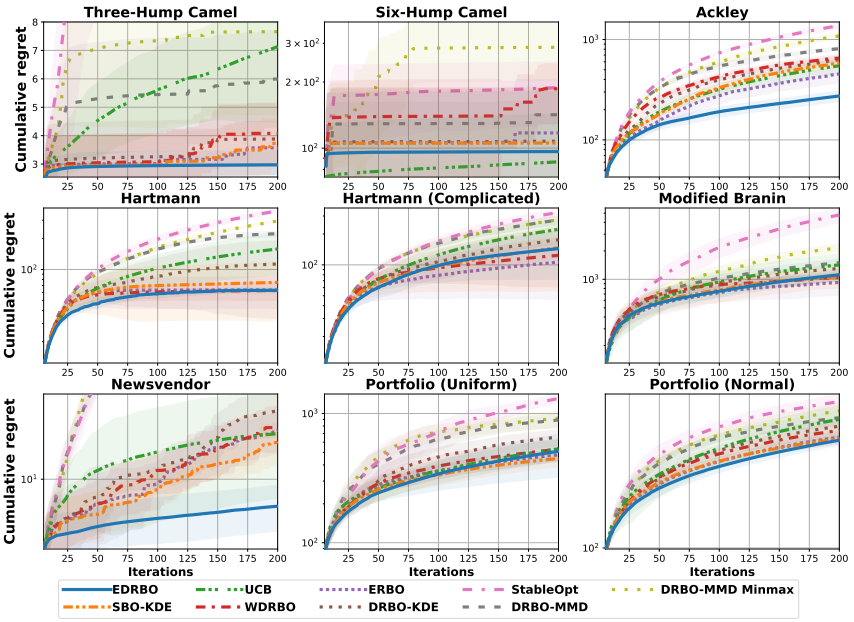


Figure 2: Mean and standard deviation of cumulative expected regret in logarithmic scale. The lower, the better.

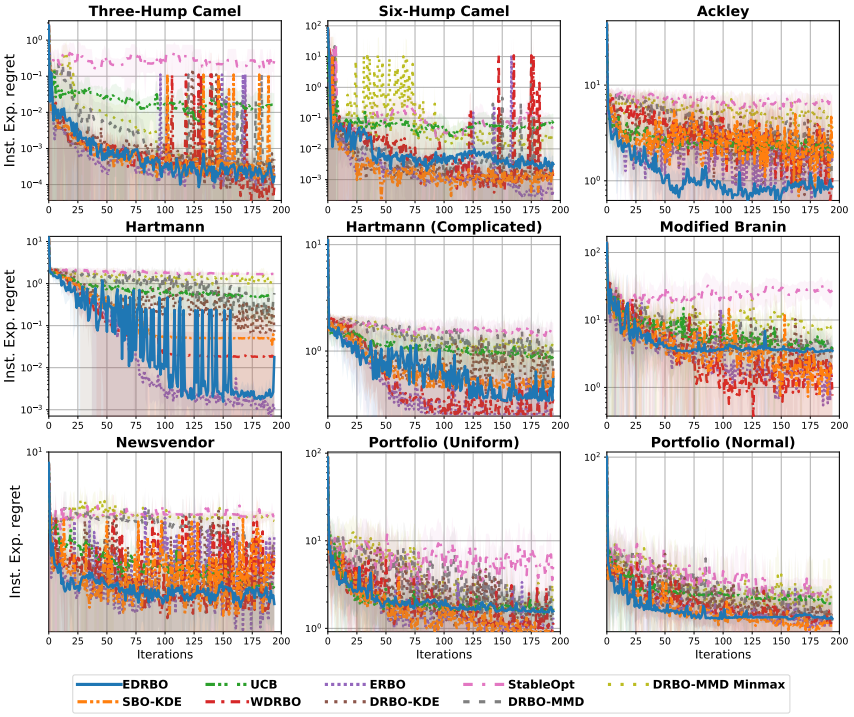


Figure 3: Mean and standard deviation of instantaneous expected regret. The lower, the better.

C Ablation study

Table 4: Paired Wilcoxon signed-rank and paired t-tests ($n = 9$ pairs per baseline).

Method	Mean diff.	Median diff.	W	Wilcoxon p_{Holm}	t-test p_{Holm}
SBO-KDE	26.389	5.513	15	1.000	1.000
WDRBO	58.953	7.341	12	1.000	1.000
ERBO	-1.871	0.654	21	1.000	1.000
Ensemble WDRBO	40.266	21.237	3	0.273	0.723
UCB	118.082	77.896	3	0.273	0.396
EDRBO (RBF UCB)	54.412	19.461	4	0.328	0.532
EDRBO (RQ - Matérn $\frac{3}{2}$)	-30.745	-10.665	6	0.602	1.000
EDRBO (Matérn $\frac{5}{2}$ - Matérn $\frac{3}{2}$)	-15.838	-8.853	8	0.977	0.723
EDRBO (SE - RQ)	31.340	12.891	9	1.000	0.823
EDRBO (SE - RQ - Matérn $\frac{5}{2}$)	-0.450	-0.044	22	1.000	1.000
EDRBO (SE - Matérn $\frac{5}{2}$)	16.127	1.528	15	1.000	1.000
EDRBO (RQ - Matérn $\frac{5}{2}$)	-3.710	-8.797	16	1.000	1.000
EDRBO (SE - RQ - Matérn $\frac{3}{2}$)	-9.062	0.009	19	1.000	1.000
EDRBO (SE - Matérn $\frac{3}{2}$)	-16.079	-0.190	16	1.000	1.000

We run an ablation on EDRBO ensemble kernels. We consider different submodels of the full EDRBO ensemble. The experimental setup is identical to the one described in the main text, see Section 5 and Appendix B for more detail. For better clarity we keep only the best performing baseline methods. In addition to that, we examine EDRBO with the UCB portion of the acquisition function being computed using RBF GP only and WDRBO with our full ensemble.

The Friedman omnibus statistic is $F = 55.658$ with p-value of 6.67×10^{-7}, meaning that there is difference in model performance. For a particular problem a specific combination of kernels may yield a better result than the full ensemble. However, both the Wilcoxon signed-rank and t-tests, presented in Table 4, show that no comparison with full-model EDRBO is statistically significant after Holm-Bonferroni correction.

We note that in general BO methods are considered over, e.g., evolutionary algorithms, when the black-box function is computationally heavy Barbetti [2024]. Meaning, that fitting an ensemble of GP surrogates and employing EDRBO, should not add significant compute to the overall runtime of an experiment - we refer to the average computing time across our experiments presented in Table 2. Without prior knowledge on the shape of the objective function, one should consider the largest feasible ensemble to maximise his safeguard and the predictive capability of the surrogate model.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction accurately reflect the proposed methodology, its theoretical and experimental results, which align with the content provided in Section 4, Section 5 together with Appendix A and Appendix B.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Section 5 discusses the obtained empirical results, which demonstrate the limitations of the proposed method.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: Notation together with the main set of assumptions is given in the corresponding paragraph in the beginning 1. All theoretical claims, including supporting lemmas, are stated and provided with a full proof in Appendix A, or are cited.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: All experiments are described in Section 5 and Appendix B. Our method implementation follows the proposed methodology, Section 4, and baseline methods are described and cited in Section 5 and implementations are adapted from published research. Instructions are accompanied with the code.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in

some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: Code to reproduce the results in full is provided in supplementary materials. Instructions are accompanied with the code.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: All experimental details are provided in Section 5 and Appendix B.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: Standard deviations (over 10 random seeds per experiment) for all results are reported. Namely, in the cumulative regret plots, Fig. 1, instantaneous expected regret plots, Fig. 3, and final cumulative regret table, Table **??**, and average computing time, Table 2.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.

- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: Computer resources used in our experiments are reported in Appendix B, including mean and standard deviation of each method, Table 2, allowing for estimation of computational cost.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The research involves algorithmic development and evaluation on standard test functions for optimisation. It does not involve human subjects or obvious ethically sensitive applications, thus leading us to believe that it conforms to the NeurIPS Code of Ethics.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: We refer to Section 7.

Guidelines:

- The answer [N/A] means that there is no societal impact of the work performed.
- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: The paper proposes a new optimisation method. Neither the algorithm nor the standard test functions pose a high risk for misuse so as to require any safeguards.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [No]

    Justification: We adapt publicly available on GitHub existing implementations of both benchmarks and baselines from published research. This is stated in Appendix B, provided with links to original implementations and corresponding publications. However, specific licences or terms of use for these assets are not explicitly mentioned in either the main text or appendix.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.

- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

    Justification: The primary new asset is the source code for the proposed method, which is provided in the supplemental materials. Instructions are accompanied with the code.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

Guidelines:

- The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: The core development does not involve LLMs as any important, original, or non-standard component.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.