
Wasserstein Distributionally Robust Bayesian Optimization with Continuous Context

Francesco Micheli
ETH Zürich

Efe C. Balta
inspire AG & ETH Zürich

Anastasios Tsiamis
ETH Zürich

John Lygeros
ETH Zürich

Abstract

We address the challenge of sequential data-driven decision-making under context distributional uncertainty. This problem arises in numerous real-world scenarios where the learner optimizes black-box objective functions in the presence of uncontrollable contextual variables. We consider the setting where the context distribution is uncertain but known to lie within an ambiguity set defined as a ball in the Wasserstein distance. We propose a novel algorithm for Wasserstein Distributionally Robust Bayesian Optimization that can handle continuous context distributions while maintaining computational tractability. Our theoretical analysis combines recent results in self-normalized concentration in Hilbert spaces and finite-sample bounds for distributionally robust optimization to establish sublinear regret bounds that match state-of-the-art results. Through extensive comparisons with existing approaches on both synthetic and real-world problems, we demonstrate the simplicity, effectiveness, and practical applicability of our proposed method.

1 INTRODUCTION

Bayesian Optimization (BO) has emerged as a powerful algorithm for zero-order optimization of expensive-to-evaluate black-box functions, with applications ranging from hyperparameters tuning to scientific discovery and robotics (Ueno et al., 2016; Li et al., 2019; Ru et al., 2020; Shahriari et al., 2015). In the standard BO setting, the learner sequentially selects points to

evaluate the unknown objective function and uses the observed data to update a surrogate model that captures the function’s behavior. In the contextual BO setting, the objective function depends on an additional variable, called the context, which cannot be controlled by the learner (Krause and Ong, 2011; Valko et al., 2013; Kirschner and Krause, 2019). Typically, the context distribution is used to model the uncertainty of the learner related to uncontrollable environmental variables. When the distribution of the context variable is known, the BO algorithm can be used to solve the Stochastic Optimization (SO) problem, where the objective is to maximize the reward of the unknown function in expectation with respect to the context distribution

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

However, in many real-world scenarios, the learner does not have access to the true context distribution, but only to an approximate one. This can happen, e.g., when the context distribution is estimated from historical data, and only a finite number of samples are available. This results in a distributional mismatch between the distribution available to the learner for optimization and the true distribution of the context variable. To formally account for the effect of the distributional mismatch, Distributionally Robust Optimization (DRO) has recently gained considerable attention, especially in the sampled data settings (Rahimian and Mehrotra, 2019; Kuhn et al., 2019; Gao et al., 2024). In DRO, the learner optimizes the reward under the worst-case distribution of the context within a so-called ambiguity set \mathcal{B} that captures the uncertainty of the learner about the true context distribution

$$\max_{x \in \mathcal{X}} \inf_{Q \in \mathcal{B}} \mathbb{E}_{c \sim Q} [f(x, c)] . \quad (1)$$

The advantage of the robust approach is that, by appropriately choosing the ambiguity set \mathcal{B} , we can guarantee that the reward computed for the DRO problem lower-bounds the reward for the true unknown context distribution.

In this work, we introduce Wasserstein Distributionally Robust Bayesian Optimization (WDRBO), a novel algo-

rithm that combines the principles of BO and DRO to address the challenge of sequential data-driven decision-making under context distributional uncertainty. We consider ambiguity sets defined as balls in the Wasserstein distance (Kuhn et al., 2019) which allows for a flexible and intuitive way to model the uncertainty in the context distribution. We design a computationally tractable algorithm and analyze its performance in two settings: the **General WDRBO** setting, where at each time-step the Wasserstein ambiguity set is provided to the learner, and the **Data-Driven WDRBO** setting, in which we assume that the true context distribution is time-invariant and the Wasserstein ambiguity set is built using the past context observations.

Our main contributions are as follows:

- We propose a novel, computationally tractable algorithm for Wasserstein Distributionally Robust Bayesian Optimization that handles continuous context distributions. Our approach exploits an approximate reformulation based on Lipschitz bounds of the acquisition function, circumventing the need for context discretization.
- We establish a cumulative expected regret bound of order $\tilde{O}(\sqrt{T}\gamma_T)$ for the general WDRBO setting, where T is the number of iterations and γ_T is the maximum information gain. For the data-driven setting, we obtain sublinear regret guarantees without requiring assumptions on the rate of decay of the ambiguity set radius.
- We derive novel Lipschitz bounds for the mean and variance estimates, and leverage recent finite-sample bounds for Wasserstein DRO to address the dimensionality challenges in continuous context spaces.
- We provide comprehensive empirical evaluations on synthetic and real-world problems, demonstrating that our method achieves competitive performance with significantly lower computational complexity compared to existing DRBO approaches.

The rest of the paper is organized as follows. In Section 2, we review related work. In Section 3, we introduce the problem formulation. In Section 4, we present the proposed algorithm and provide the theoretical analysis. In Section 5, we present the experimental results. Finally, in Section 6, we conclude the paper and discuss future work.

2 RELATED WORK

The foundation of DRBO was laid by Kirschner et al. (2020), who introduced the concept of distributional

robustness in BO. They propose a BO formulation that is robust to the worst-case context distribution within an ambiguity set defined by the Maximum Mean Discrepancy (MMD) distance. While groundbreaking, the inner worst-case calculation requires at each iteration the solution of a convex optimization problem that renders this approach computationally viable only when the context space is discrete and with low cardinality. A quadrature-based scheme for DRBO is proposed in Nguyen et al. (2020), but their algorithm is limited to the simulator setting where at each iteration the learner is allowed to choose the context. Husain et al. (2024) develops a DRBO formulation for ϕ -divergence-based ambiguity sets, but their formulation has some implicit requirements on the support of the distributions captured by the ambiguity set. Recognizing these computational limitations, Tay et al. (2022) proposed a set of approximate techniques using worst-case sensitivity analysis based on Taylor’s expansions. These methods offer better computational complexity for multiple descriptions of ambiguity sets at the expense of performance and regret bounds that scale linearly with the worst-case sensitivity approximation error. To avoid the challenges of context space discretization, Huang et al. (2024) proposes a kernel density estimation step that uses the available context samples to estimate a continuous context distribution. The estimated context distribution is then sampled and the samples are used in a DRBO formulation where the ϕ -divergence ambiguity sets capture the distributional uncertainty introduced by the density estimation step.

The regret analysis of the existing literature on DRBO builds on the GP-UCB formulation of Srinivas et al. (2009, 2012), we instead exploit self-normalizing concentration bounds in Reproducing Kernel Hilbert Space (RKHS) (Abbasi-Yadkori, 2013; Kirschner et al., 2020; Whitehouse et al., 2023). We address a gap in the DRBO literature and analyze the continuous context distribution setting under the Wasserstein-based ambiguity set. We leverage recent advancements in Wasserstein DRO literature (Gao, 2023; Gao et al., 2024) to provide state-of-the-art regret rates in the data-driven setting.

3 PROBLEM FORMULATION

We consider an unknown objective function $f : \mathcal{X} \times \mathcal{C} \rightarrow \mathbb{R}$, where $\mathcal{X} \subset \mathbb{R}^{d_x}$ is the input space and $\mathcal{C} \subset \mathbb{R}^{d_c}$ is the context space. The learner’s goal is to maximize the expected value of the function under the context distribution by sequentially selecting points to evaluate and receiving noisy observations of the function. More specifically, at each iteration $t = 1, 2, \dots$, the learner selects a point $x_t \in \mathcal{X}$ to query the function, and observes the context $c_t \in \mathcal{C}$ and a noisy output $y_t =$

$f(x_t; c_t) + \epsilon_t$. The context sample c_t is assumed to be an independent sample from some unknown, time-dependent, context distribution P_t , while ϵ_t is a zero-mean R -sub-Gaussian noise, where an upper bound on R is known.

3.1 Wasserstein Distributionally Robust Objective

In this work, we consider the setting of distributionally robust optimization, where the learner does not have access to the true context distribution P_t , but instead optimizes for the expected reward under the worst-case distribution within an ambiguity set

$$\max_{x \in \mathcal{X}} \inf_{Q \in \mathcal{B}_t(\hat{P}_t)} E_{c \sim Q} [f(x; c)] : \quad (2)$$

The time-dependent ambiguity set $\mathcal{B}_t(\hat{P}_t)$ is defined as a ball in the Wasserstein distance centered at the distribution \hat{P}_t and with radius ϵ_t (Kuhn et al., 2019). This is the set of all distributions that are within a Wasserstein distance ϵ_t from the center distribution \hat{P}_t

$$\mathcal{B}_t(\hat{P}_t) = \{Q \in \mathcal{P}(\mathcal{C}) : d_W(Q; \hat{P}_t) \leq \epsilon_t\}$$

The type-1 Wasserstein metric $d_W : \mathcal{M}(\mathcal{Q}) \times \mathcal{M}(\mathcal{Q}) \rightarrow \mathbb{R}_0$ defines the distance between two distributions Q_1 and Q_2 as

$$d_W(Q_1; Q_2) := \inf_{\gamma \in \Pi(Q_1, Q_2)} \int_{\mathcal{Q} \times \mathcal{Q}} \|x_1 - x_2\| d\gamma(x_1, x_2) ;$$

where the transportation map γ takes values in the set of joint distributions of q_1 and q_2 with marginals Q_1 and Q_2 , and $\|\cdot\|$ is the euclidean norm.

3.2 Regularity Assumptions and Surrogate Model

The BO algorithm maintains a surrogate model of the objective function, which is used to guide the selection of the next query point. We use a regularized least squares estimator of the function f in the RKHS (Abbasi-Yadkori, 2013; Kirschner and Krause, 2018) under the assumption that f is an unknown fixed member of the RKHS H_k that is specified by the positive semi-definite kernel $k : Z \times Z \rightarrow \mathbb{R}$, where $Z = \mathcal{X} \times \mathcal{C}$. Here we define $z = [x^\top; c^\top]^\top$ to keep the notation compact. We assume that the spaces \mathcal{X} and \mathcal{C} are compact. We define the norm of a function $g \in H_k$ as $\|g\|_{H_k} = \sqrt{\langle g, g \rangle} = \sqrt{\langle hg, hg \rangle_{H_k}}$. We also assume that the unknown function f has bounded RKHS norm, i.e., $\|f\|_{H_k} \leq B$, for some $B > 0$, and that the kernel k is bounded, i.e., $k(z; z') \leq 1$, for all $z; z' \in Z$. The assumptions made here are common in the BO literature, we point the reader to e.g. Bogunovic and

Krause (2021) for the analysis of bandits optimization with misspecified RKHS. The details of the following derivations are available in the Appendix Section 7.

Given the dataset $D_t = \{(z_i, y_i)\}_{i=1}^t$, and regularization parameter $\lambda > 0$, the regularized least-squares regression problem in RKHS is written as follows:

$$\min_{f \in H_k} \sum_{i=1}^t (y_i - f(z_i))^2 + \lambda \|f\|_{H_k}^2 : \quad (3)$$

The resulting least squares estimator is

$$\hat{f}_t(z) = k_t(z)^\top (K_t + \lambda I)^{-1} y_{1:t} ; \quad (4)$$

where $k_t(z) = [k(z; z_1); \dots; k(z; z_t)]^\top$, $K_t = [k(z_i; z_j)]_{i,j=1}^t$, and $y_{1:t} = [y_1; \dots; y_t]^\top$. We also define

$$\hat{\sigma}_t^2(z) = \frac{1}{t} k_t(z)^\top (K_t + \lambda I)^{-1} k_t(z) : \quad (5)$$

Under suitable assumptions on the prior and the noise distribution, \hat{f}_t and $\hat{\sigma}_t^2$ correspond to the posterior mean and variance of a Gaussian process with kernel k conditioned on the observations $y_{1:t}$ (Schölkopf and Smola, 2002; Williams and Rasmussen, 2006).

We state here a fundamental result adapted from Abbasi-Yadkori (2013) that provides probabilistic finite-sample confidence guarantees for the least squares estimator (4).

Lemma 1. [(Abbasi-Yadkori, 2013, Th. 3.11)] Let $Z \subseteq \mathbb{R}^d$, where $d = d_x + d_c$, and $f : Z \rightarrow \mathbb{R}$ be a member of H_k , with $\|f\|_{H_k} \leq B$, and let ϵ_t be F_t measurable and R -sub-Gaussian conditionally on F_t . Then, for any $\delta > 0$, with probability $1 - \delta$, we have that for all $z \in Z$ and all $t \geq 1$:

$$|\hat{f}_t(z) - f(z)| \leq \epsilon_t(z) ; \quad (6)$$

with

$$\epsilon_t := R \sqrt{2 \log \frac{\det(I + \lambda^{-1} K_{t-1})^{\frac{1}{2}}}{\lambda}} + \frac{1}{2} B ; \quad (7)$$

where ϵ_{t-1} and ϵ_{t-1} are defined as in equations (4), (5).

We also introduce here the maximum information gain (Srinivas et al., 2009; Chowdhury and Gopalan, 2017; Vakili et al., 2021), a fundamental kernel-dependent quantity that quantifies the complexity of learning in RKHS

$$\gamma_t := \sup_{z_1, z_2, \dots, z_t} \log \det(I + \lambda^{-1} K_{t-1}) :$$

In order to derive the main results in the following sections, we require the kernel to satisfy the following Lipschitz property

Assumption 1 (Lipschitz property). There exists a $L > 0$ such that for any $z, z^0 \in \mathcal{X} \times \mathcal{C}$, $d(z; z^0) := k(z; z^0) - k(z; z^0)_{H_k} \leq L \|z - z^0\|_k$.

As we prove in Lemma 11 in the Appendix, Assumption 1 is verified for popular kernels, e.g. the squared exponential kernel, and some kernels in the Matérn family satisfy.

4 WASSERSTEIN DISTRIBUTIONALLY ROBUST BAYESIAN OPTIMIZATION

In classical BO, following the rich literature of optimism in the face of uncertainty, the learner selects the query point x_t by maximizing the Upper Confidence Bound (UCB) function (Auer, 2002). This provides a trade-off between exploration and exploitation, and results in provable regret guarantees (Srinivas et al., 2009).

Departing from the classical approach, and inspired by Kirschner et al. (2020), we adopt a robust approach, where we consider the optimization of a robustified version of the UCB function

$$x_t = \arg \max_{x \in \mathcal{X}} \inf_{Q \in \mathcal{B}(\hat{\mathcal{P}}_t)} E_{c \sim Q} [\text{UCB}_t(x; c)] ; \quad (8)$$

where

$$\text{UCB}_t(x; c) = \hat{f}_t(x; c) + \hat{\sigma}_t(x; c); \quad (9)$$

with \hat{f}_t , $\hat{\sigma}_t$, and $\hat{\mathcal{P}}_t$ are given in (4), (5), (7) respectively.

Similar to Kirschner et al. (2020), we will analyze two settings, which differ in the way the ambiguity set is obtained. We first consider the General WDRBO setting, where at each time-step the Wasserstein ambiguity set is provided to the learner, and then turn to the Data-Driven WDRBO setting, in which the Wasserstein ambiguity set is built using the past context observations under the assumption that the true context distribution is time-invariant.

All proofs along with supporting derivations and lemmas are provided in Section 8 of the Appendix.

To evaluate the performance of the proposed algorithm we look at the notion of regret. Regret is used to capture the difference in performance between some algorithm and a benchmark algorithm that has access to privileged information. The definition of regret and the choice of benchmark is not unique, and the one chosen here differs from the ones used in the DRBO literature Kirschner and Krause (2019), Husain et al. (2024), Tay et al. (2022).

We will consider the following definitions of instantaneous expected regret :

$$r_t = E_{c \sim \mathcal{P}_t} [f(x_t; c)] - E_{c \sim \mathcal{P}_t} [f(x_t^*; c)]; \quad (10)$$

and cumulative expected regret :

$$R_T = \sum_{t=1}^T r_t ; \quad (11)$$

The benchmark solution x_t^* is the optimal solution to the true stochastic optimization problem at time-step t , given access to the true function f and context distribution \mathcal{P}_t , i.e.,

$$x_t^* = \arg \max_{x \in \mathcal{X}} E_{c \sim \mathcal{P}_t} [f(x; c)] ;$$

Hence, this definition of regret captures the (cumulative) sub-optimality gap, between some proposed algorithm and the optimal solution to the true stochastic optimization problem.

4.1 General WDRBO

In the General WDRBO setting, at each time-step, the center $\hat{\mathcal{P}}_t$ and the radius σ_t of the Wasserstein ambiguity set are provided to the learner. This represents the setting where there is some understanding of what the context distribution is, e.g. with weather or prices forecast, but there is still some uncertainty about its distribution.

To make the robust problem 8 tractable, we introduce a well-known result from the Wasserstein DR optimization literature (Kuhn et al., 2019; Gao et al., 2024) that has been adapted to our problem.

Lemma 2. Let $f : \mathcal{X} \times \mathcal{C} \rightarrow \mathbb{R}$ be a function that is $L_c^f(x)$ -Lipschitz in the context space, i.e. $|f(x; c) - f(x; c^0)| \leq L_c^f(x) \|c - c^0\|_k$, for all $c; c^0 \in \mathcal{C}$. Let $\mathcal{B}(\hat{\mathcal{P}})$ be a Wasserstein ambiguity set defined as a ball of radius σ in the Wasserstein distance centered at the distribution $\hat{\mathcal{P}}$. Then, for any $x \in \mathcal{X}$ and for any distribution $\mathcal{P} \in \mathcal{B}(\hat{\mathcal{P}})$, we have that

$$|E_{c \sim \mathcal{P}} [f(x; c)] - E_{c \sim \hat{\mathcal{P}}} [f(x; c)]| \leq L_c^f(x) \sigma ; \quad (12)$$

Lemma 2 provides a simple Lipschitz-based bound on the worst-case expectation for any distribution in the ambiguity set. By combining Lemma 2 with Assumption 1 we obtain a tractable approximation of the robust maximization problem 8. Thus, at each time-step the query point x_t is selected by the following acquisition function

$$x_t = \arg \max_{x \in \mathcal{X}} E_{c \sim \hat{\mathcal{P}}_t} [\text{UCB}_t(x; c)] - \sigma_t L^{\text{UCB}_t}(x); \quad (13)$$

Algorithm 1 General Algorithm for WDRBO

```

for  $t = 1$  to  $T$  do
     $\hat{c}_{t-1}, \hat{c}_{t-1} \leftarrow \hat{c}_{t-1}(D_{t-1})$ 
     $x_t = \arg \max_{x \in \mathcal{X}} E_{c \sim p_t} [UCB_t(x; c)] - \eta_t L^{UCB_t}(x)$ 
    The environment returns  $c_t$  and  $y_t = f(x_t; c_t) + \epsilon_t$ ,
    where  $c_t \sim P_t \stackrel{D}{=} B^{\eta_t}(\hat{c}_{t-1})$ 
     $D_t \leftarrow D_{t-1} \cup \{(x_t; c_t; y_t)g\}$ 
end for
    
```

where $L^{UCB_t}(x)$ is the Lipschitz constant of the function UCB_t with respect to the context variable c , evaluated at x . The resulting algorithm for WDRBO is provided in Algorithm 1.

We remark that, unlike the algorithm proposed in Kirschner and Krause (2019), Husain et al. (2024), Tay et al. (2022), that rely on discrete context distributions, neither Algorithm 1 nor the following theoretical analysis require a discrete context space or that the distributions in the ambiguity have finite support. The only practical limitation imposed by our algorithm when the center is a continuous context distribution, is the ability to perform the numerical integration required to compute the expectation.

While in Algorithm 1 the Lipschitz constant $L^{UCB_t}(x)$ can be computed at each timestep t from the fitted UCB function UCB_t , in the following lemma we derive a novel upper bound on $L^{UCB_t}(x)$ that will be useful for the theoretical analysis of the regret.

Lemma 3. Let $0 \leq \gamma < 1$ be a failure probability

and let $B_t := \frac{1}{2} R \sqrt{2 \log \frac{\det(I + \gamma K_{t-1})^{\frac{1}{2}}}{\gamma}} + B$
 $\frac{1}{2} R \sqrt{\frac{\gamma}{2 \log 1 + R^2}} + B$.

Then, with probability $1 - \gamma$ for all $t \geq 1$ we have $\|k_{t-1} - k_{H_k}\| \leq B_t$. Further, if Assumption 1 holds we have:

- (i) With probability $1 - \gamma$, for any $z; z^0 \in \mathcal{X} \times \mathcal{C}$: $\|j_{t-1}(z) - j_{t-1}(z^0)\| \leq B_t L_k z - z^0 k$.
- (ii) For any $z; z^0 \in \mathcal{X} \times \mathcal{C}$: $\|j_{t-1}(z) - j_{t-1}(z^0)\| \leq \frac{1}{2} L_k z - z^0 k = B_t L_k z - z^0 k$.

Therefore, with probability $1 - \gamma$, the UCB function is Lipschitz continuous with constant:

$$L^{UCB_t} \leq 2B_t L_k$$

We can now turn to the derivation of the bound on the instantaneous expected regret for the General WDRBO setting.

Theorem 4 (Instantaneous expected regret) Let Assumption 1 hold. Fix a failure probability $0 < \gamma < 1$. With probability at least $1 - \gamma$, for all $t \geq 1$ the instan-

taneous expected regret can be bounded by

$$r_t = E_{c \sim P_t} [2 \eta_{t-1}(x_t; c)] + 2 \eta_t L^{UCB_t}(x_t) \quad (14)$$

We can observe that the first term has the same expression as the instantaneous regret of the GP-UCB Srinivas et al. (2009), while the second term captures the effect of the distributional uncertainty which depends on the maximum distribution shift as specified by η_t , and on a sensitivity term that is bounded by the Lipschitz constants $L^{UCB_t}(x_t)$ computed at the selected input x_t .

Theorem 5 (Cumulative expected regret). Let Assumption 1 hold and let $L^{UCB_t}(x)$ be a Lipschitz constant with respect to the context c for $UCB_t(x; c)$. Fix a failure probability $0 < \gamma < 1$. With probability at least $1 - \gamma$, the cumulative expected regret after T steps can be bounded as:

$$R_T \leq 4 \sqrt{T} + 4 \log\left(\frac{6}{\gamma}\right) + \sum_{t=1}^T \eta_t 2 L^{UCB_t}(x_t); \quad (15)$$

where \sqrt{T} is the maximum information gain at time T .

Note that the cumulative expected regret is a random quantity, as the expectation is taken only with respect to the contexts.

We can combine Lemma 3 with Theorem 5 to derive the regret rate for the cumulative expected regret.

Corollary 6 (General WDRBO Regret Order). Let $0 < \gamma < 1$ be a failure probability and let Assumption 1 hold. Then, with probability $1 - \gamma$, the cumulative expected regret is of the order of

$$R_T = \mathcal{O}\left(\sqrt{T} + \sum_{t=1}^T \eta_t\right);$$

For the Squared Exponential kernel, this reduces to

$$R_T = \mathcal{O}\left(\sqrt{T} + \sum_{t=1}^T \eta_t\right);$$

where \mathcal{O} omits logarithmic terms.

The second term depends on the sum of all radii $\sum_{t=1}^T \eta_t$. Hence, a sufficient condition in order to get sublinear regret guarantees, is that the radii converge to 0 sufficiently fast. If, e.g., $\eta_t = \mathcal{O}(t^{-\frac{1}{2}})$, we obtain $R_T = \mathcal{O}(\sqrt{T})$. This can also occur in certain situations like the data-driven setting that we analyze next.

4.2 Data-Driven WDRBO

In the Data-Driven WDRBO we still rely on Algorithm 1, but differently from the general setting we need to build the Wasserstein ambiguity set using the past context observations. With the assumption that the unknown context distribution is time-invariant, i.e. $P_t := P$, $t = 1, \dots, T$, we build the ambiguity set centered at \hat{P}_t as the empirical distribution of the past observed contexts, i.e.

$$\hat{P}_t = \frac{1}{t} \sum_{i=1}^t \mathbb{I}_{f_{c=c_i^j} g};$$

where $\mathbb{I}_{f_{c=c_i^j} g}$ is the indicator function centered on the context sample c_i^j , and we derive a bound on the sequence of radii r_t using finite-sample concentration results.

Using finite-sample results for the convergence of empirical measures in Wasserstein distance Fournier and Guillin (2015); Fournier (2022) we can bound the size of r_t such that with high probability the true context distribution P is contained in the ambiguity set $B^{(t)}(\hat{P}_t)$. Unfortunately, this approach suffers from the so-called curse of dimensionality with respect to the dimension d_c of the context. To circumvent this issue we propose a novel result that leverages recent finite-sample concentration results from Gao (2023). Instead of focusing on the rate of convergence of the empirical distribution \hat{P}_t to the true unknown P , we focus on the rate at which the worst-case expected cost concentrates around the expected cost under the true context distribution. Since there are multiple possible UCB functions, we need an additional covering argument to apply the result from Gao (2023).

Lemma 7. Define the class of UCB functions as

$$U(A) = \{h : h(z) = \langle z, \bar{c} \rangle + \langle z, c \rangle; k \in H_k, A \in \mathcal{A}\};$$

and let $N_1(\cdot; H_k; A)$ be its covering number under the infinity norm, up to precision ϵ . Let $\text{diam}(X)$, $\text{diam}(C)$ denote the diameters of the sets $X; C$ respectively. Let $0 < \delta < 1$ be a failure probability. Let

$$r_t = \frac{\log \frac{1}{1-\delta} + d_x \log(1+2t \text{diam}(X)) + \log N_1(t^{-1}; H_k; A)}{t},$$

where $\bar{c} = 2 \text{diam}(C)$, and $c_t = 3(1 + LA)t$. Then, with probability at least $1 - \delta$, for all $h \in U(A)$ and all $x \in X$

$$E_{c \sim P} [h(x; c)] - E_{c \sim \hat{P}_t} [h(x; c)] \leq r_t L_c^h(x) \quad t \geq 1$$

We can now use the bound of Lemma 7 to derive the data-driven analogous of Theorem 4 and Theorem 5.

Theorem 8 (Data-driven instantaneous expected regret). Let Assumption 1 hold. Fix a failure probability $0 < \delta < 1$. With probability at least $1 - \delta$, for all $t \geq 1$ the instantaneous expected regret for the data-driven setting can be bounded by

$$r_t = E_{c \sim P_t} [2 \langle c, c_t \rangle (x_t; c)] + 2 r_t L^{UCB_t}(x_t) + 2 r_t \quad (16)$$

Theorem 9 (Data-driven cumulative expected regret). Let Assumption 1 hold and let $L^{UCB_t}(x)$ be a Lipschitz constant with respect to the context c for $UCB_t(x; c)$. Fix a failure probability $0 < \delta < 1$. With probability at least $1 - \delta$, the cumulative expected regret for the data-driven setting can be bounded as:

$$R_T \leq 4 \frac{r}{T} + \frac{1}{T} \sum_{t=1}^T \left(2 L^{UCB_t}(x_t) + 2 r_t \right) \quad (17)$$

We can use Lemma 19 with $A = B_t$ from Lemma 3 to derive a bound on the rate of the cumulative expected regret in the data-driven setting. To specialize the result to the Squared Exponential kernel case, we use a result from Yang et al. (2020) to obtain a bound on $N_1(\cdot; H_k; B)$.

Corollary 10. Let $0 < \delta < 1$ be a failure probability and let Assumption 1 hold. Then, With probability $1 - \delta$, the cumulative expected regret in the data-driven setting is of the order of

$$R_T = O\left(\frac{P}{T} + \sum_{t=1}^T O\left((t^{-1} \log N_1(t^{-1}; H_k; B_t))^{\frac{1}{2}} P\right) + \sum_{t=1}^T O(t^{-1/2} P)\right) \quad (18)$$

For the Squared Exponential kernel, this reduces to a sublinear regret with order

$$R_T^{SE} = O\left(\frac{P}{T}\right)$$

Following the same procedure, using the bounds on the covering number N_1 derived in Yang et al. (2020), it is possible to derive similar bounds for other commonly used kernels. The rate derived in Corollary 10 shows that a sublinear regret is achievable when the dependency on the covering number and the maximum information gain are well behaved. These are linked to the smoothness of the kernel. It is not yet clear whether this is a fundamental limitation or if it is an artifact of our proving technique. We will leave this for future work.

Note that with the proposed data-driven WDRBO, we have a principled way to choose the sequence of radii ϵ_t that provides a probabilistic guarantee on the maximum distance between the expectation under the true context generating distribution P and the expectation under the empirical distribution \hat{P}_t . This makes Algorithm 1 a practical tool to handle continuous context distributions, in contrast with Kirschner et al. (2020), Husain et al. (2024), Tay et al. (2022), where it is assumed that the true distribution is supported on a finite number of contexts. The proposed approach differs also from Huang et al. (2024) where their DRO-KDE algorithm robustifies against the gap between the approximate context distribution obtained by KDE from the observed contexts and the empirical distribution obtained by sampling it.

5 EXPERIMENTS

In this section, we analyze the performance of the proposed algorithm and compare it with the algorithms in the literature. We will start with a simple example that showcases the effect of the robust acquisition function (13) in the general setting. We will then provide an extensive comparison of the algorithms in the data-driven setting, as we consider it the most relevant and more challenging in practice.

To highlight the need for robustness against context distribution shifts, we consider the general DRBO setting with fixed context distributions $\hat{P}_t = N(0.5; 0.1)$ for all $t = 1; \dots; 100$ and $P_t = N(0.6; 0.2)$ for all $t = 1; \dots; 100$, and the unknown function

$$f(x; c) = 1 - \frac{|c - 0.5|}{|x| + 0.2} \cdot \frac{P}{|x| + 0.05} :$$

A plot of the function and its optima under \hat{P}_t and P_t is shown in Fig. 1a. We compare the performance of the proposed WDRBO as in Algorithm 1 with $\epsilon_t = 0.1$ for all $t = 1; \dots; 100$, and ERBO (Empirical Risk BO), the non-robust variant of WDRBO that assumes $\epsilon_t = 0$ for all $t = 1; \dots; 100$.

In Fig. 1b we show that the robust WDRBO results in a lower cumulative regret than ERBO. This is because ERBO solves the stochastic optimization problem assuming that the context is distributed according to the ambiguity set center \hat{P}_t , while WDRBO optimizes for the worst-case distribution in the ambiguity set of radius 0.1. In this simple setting, since the radius ϵ_t remains constant over time, following the result of Theorem 5, the cumulative expected regret shows a linear trend.

For the data-driven DRBO setting we adopt the setup of Huang et al. (2024) and provide a comparison

Figure 1: (a) Function $f(x; c)$ and its optima under optima under \hat{P}_t and P_t . (b) Mean and standard error of the cumulative expected regret.

of the different methods on synthetic function and the realistic problems.¹ We will compare the algorithms' performance based on the cumulative expected regret as in (11).

We will compare the following algorithms:
WDRBO : Data-Driven WDRBO algorithm with robustified acquisition function 13, where the center of the Wasserstein ambiguity set is given by the empirical distribution of the observed contexts and the radius is chosen as $\epsilon_t = O(1/\sqrt{t})$.

ERBO : This is equivalent to WDRBO but we set $\epsilon_t = 0$ in the acquisition function 13, i.e. we maximize the empirical risk with respect to the observed contexts $x_t = \arg \max_{x \in \mathcal{X}} E_{c \sim \hat{P}_t} [UCB_t(x; c)]$.

GP-UCB : Implements the UCB maximization algorithm proposed by Srinivas et al. (2012) ignoring the context variable in both the definition of the Gaussian process model and in the acquisition function maximization.

SBO-KDE : Stochastic BO formulation of Huang et al. (2024). An approximate context distribution is estimated from the observed samples by kernel density estimation. The acquisition function maximizes the expectation of the UCB with respect to the empirical distribution of the context obtained by sampling the approximate context distribution (sample average approximation).

DRBO-KDE : DR formulation of SBO-KDE proposed by Huang et al. (2024). Robustifies the SBO-KDE algorithm by considering DR formulation with a total variation ambiguity set. The ambiguity set is centered on the empirical distribution of the context obtained by sampling from the density estimate.

DRBO-MMD : DRBO formulation with MMD ambiguity set of Kirschner et al. (2020). The continuous context space is discretized and the UCB is maximized

¹The code is available at the following link <https://github.com/frmicheli/WDRBO>.

Figure 2: Mean and standard error of the cumulative expected regret.

for the worst-case distribution supported on the discrete context space for a given MMD budget. The complexity of the robustified acquisition function scales with the cube of the cardinality of the context support.

DRBO-MMD Minmax : Minmax approximate formulation of DRBO-MMD proposed in Tay et al. (2022). The discretization can be seen as the worst-case sensitivity approximation reduces the computational burden of the method.

StableOpt : Implementation of StableOpt Bogunovic et al. (2018). Implements a robust acquisition function $x_t = \arg \max_{x \in \mathcal{X}} \min_{c \in \mathcal{C}_t} \text{UCB}_t(x; c)$, where, following Huang et al. (2024), the set \mathcal{C}_t is chosen at each time t as the set where for each dimension of the context we consider the interval $[\hat{\mu}_t^{c^j} - \hat{\sigma}_t^{c^j}; \hat{\mu}_t^{c^j} + \hat{\sigma}_t^{c^j}]$, with $\hat{\mu}_t^{c^j}$ and $\hat{\sigma}_t^{c^j}$ the empirical mean and variance of the observed contexts.

For all the algorithms considered we fixed the value of the UCB trade-off parameter $\gamma_t = 1:5$. This has been done to be consistent with the engineering practice and earlier works such as Huang et al. (2024). We consider a set of artificial and real-world problems and different types of context distributions. For each problem and algorithm, we ran 100 iterations and repeated over 15 random seeds. Fig. 2 shows the resulting cumulative expected regret for WDRBO, ERBO, GP-UCB, SBO-KDE, and DRBO-KDE. More details about the specifics of the test problems and the implementations are left in the Appendix. We also leave the results for

DRBO-MMD, DRBO-MMD Minmax, and StableOpt to the Appendix as their performance was not competitive with the other methods. The performance of DRBO-MMD is limited by the coarseness of the context discretization which is required to have a computationally tractable inner convex optimization step. The performance of DRBO-MMD Minmax is mainly limited by the worst-case sensitivity that introduces a linear term in the resulting regret bound. StableOpt suffers from the fact that it is solving a robust optimization problem.

We tracked the time required by each algorithm for a 100-iteration-long experiment. We report in Table 1 the computational times in seconds for the Ackley and Branin functions. The computational times are affected both by algorithm specific characteristics, e.g. an inner convex optimization problem is solved at each iteration, and by specific parameters choices, e.g. the discretization grid-size. The reported times have been obtained by running the algorithm on CPU only, as some of the algorithms have not been implemented to exploit the potential speed-ups resulting from running on GPU. For this test we used an Intel(R) Core(TM) i9-9900K@3.60GHz. GP-UCB has the smallest computational time as it ignores the context, thus also reducing the regression step complexity. ERBO and SBO-KDE are not robust approaches, the extra time required by SBO-KDE is due to the KDE step. One of the advantage of the proposed WDRBO is that it is able to add robustness against context distribution uncertainty without the large overheads of the other robust

methods. The extra computational time required by WDRBO compared to ERBO is related to the calculation of the Lipschitz constant from the UCB expression. The main computational bottleneck for DRBO-MMD and DRBO-KDE is related to the solution of the inner minimization problems.

	Ackley		Branin	
WDRBO (ours)	44:3	2:2	54:5	2:6
ERBO (ours)	43:8	1:6	45:2	2:5
GP-UCB	15:7	1:4	15:1	1:0
SBO-KDE	46:7	0:7	49:7	1:5
DRBO-KDE	599:7	33:0	525:0	71:7

Table 1: Mean and standard error of computational times in seconds for the Ackley ($d_x = 1$, $d_c = 1$) and Branin ($d_x = 2$, $d_c = 2$) functions.

We can see in Fig. 2 that the performance of WDRBO, ERBO, and SBO-KDE is extremely compelling, particularly when considering the computational complexity of the other algorithms. We argue that the performance of DRBO-KDE does not justify the extra computation required to solve the inner two-dimensional optimization problem. While we observe very strong performances for ERBO and SBO-KDE in the data-driven setting, with the smallest computational complexities, we want to highlight that, contrarily to WDRBO and DRBO-KDE, they do not compute a robust solution. This might lead to disappointing performance as shown in the first example where the ambiguity set does not collapse to the true distribution as the number of iterations grows, and for which a robust solution might be preferable.

6 CONCLUSIONS AND FUTURE WORK

In this paper, we introduced Wasserstein Distributionally Robust Bayesian Optimization (WDRBO), a novel algorithm that addresses the challenge of sequential data-driven decision-making under context distributional uncertainty. We developed a computationally tractable algorithm for WDRBO that can handle continuous context distributions, leveraging an approximate reformulation based on Lipschitz bounds of the acquisition function. This approach extends the existing literature on Distributionally Robust Bayesian Optimization by providing a principled method to handle continuous context distributions within a Wasserstein ambiguity set, allowing for a flexible and intuitive way to model uncertainty in the context distribution while maintaining computational feasibility.

Our theoretical analysis provides an cumulative ex-

pected regret bounds that match state-of-the-art results. Notably, for the data-driven setting, the bound does not require assumptions on the rate of decay of the ambiguity set radius but relies on finite-sample concentration results, making our approach more broadly applicable to real-world situations. Lastly, we conducted a comprehensive empirical evaluation demonstrating the effectiveness and practical applicability of WDRBO on both synthetic and real-world benchmarks. Our results show that the proposed WDRBO algorithm exhibits promising performance in terms of regret while avoiding computationally expensive inner optimization steps.

The promising results of WDRBO open up exciting opportunities for further research and development in the field of Distributionally Robust Bayesian optimization. Extending the WDRBO framework to risk measures, such as Conditional Value at Risk (CVaR), could broaden its applicability to risk-sensitive domains, such as robotics and finance.

References

- Tsuyoshi Ueno, Trevor David Rhone, Zhufeng Hou, Teruyasu Mizoguchi, and Koji Tsuda. Combo: An efficient bayesian optimization library for materials science. *Materials discovery*, 4:18 21, 2016.
- Cheng Li, Santu Rana, Sunil Gupta, Vu Nguyen, Svetha Venkatesh, Alessandra Sutti, David Rubin, Teo Slezak, Murray Height, Mazher Mohammed, et al. Accelerating experimental design by incorporating experimenter hunches. *arXiv preprint arXiv:1907.09065*, 2019.
- Binxin Ru, Ahsan Alvi, Vu Nguyen, Michael A Osborne, and Stephen Roberts. Bayesian optimisation over multiple continuous and categorical inputs. In *International Conference on Machine Learning* pages 8276 8285. PMLR, 2020.
- Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P Adams, and Nando De Freitas. Taking the human out of the loop: A review of bayesian optimization. *Proceedings of the IEEE*, 104(1):148 175, 2015.
- Andreas Krause and Cheng Ong. Contextual gaussian process bandit optimization. *Advances in neural information processing systems* 24, 2011.
- Michal Valko, Nathaniel Korda, Rémi Munos, Ilias Flaounas, and Nelo Cristianini. Finite-time analysis of kernelised contextual bandits. *arXiv preprint arXiv:1309.6869*, 2013.
- Johannes Kirschner and Andreas Krause. Stochastic bandits with context distributions. *Advances in Neural Information Processing Systems* 32, 2019.
- Hamed Rahimian and Sanjay Mehrotra. Distribution-

- ally robust optimization: A review. arXiv preprint arXiv:1908.05659, 2019.
- Daniel Kuhn, Peyman Mohajerin Esfahani, Viet Anh Nguyen, and Soroosh Shafadeh-Abadeh. Wasserstein distributionally robust optimization: Theory and applications in machine learning. In *Operations research & management science in the age of analytics*, pages 130–166. INFORMS, 2019.
- Rui Gao, Xi Chen, and Anton J Kleywegt. Wasserstein distributionally robust optimization and variation regularization. *Operations Research* 72(3): 1177–1191, 2024.
- Johannes Kirschner, Ilija Bogunovic, Stefanie Jegelka, and Andreas Krause. Distributionally robust bayesian optimization. In *International Conference on Artificial Intelligence and Statistics*, pages 2174–2184. PMLR, 2020.
- Thanh Nguyen, Sunil Gupta, Huong Ha, Santu Rana, and Svetha Venkatesh. Distributionally robust bayesian quadrature optimization. In *International Conference on Artificial Intelligence and Statistics*, pages 1921–1931. PMLR, 2020.
- Hisham Husain, Vu Nguyen, and Anton van den Hengel. Distributionally robust bayesian optimization with α -divergences. *Advances in Neural Information Processing Systems* 36, 2024.
- Sebastian Shenghong Tay, Chuan Sheng Foo, Urano Daisuke, Richalynn Leong, and Bryan Kian Hsiang Low. Efficient distributionally robust bayesian optimization with worst-case sensitivity. In *International Conference on Machine Learning* pages 21180–21204. PMLR, 2022.
- Xiaobin Huang, Lei Song, Ke Xue, and Chao Qian. Stochastic bayesian optimization with unknown continuous context distribution via kernel density estimation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 12635–12643, 2024.
- Niranjan Srinivas, Andreas Krause, Sham M Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. arXiv preprint arXiv:0912.3995, 2009.
- 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, 2012.
- Yasin Abbasi-Yadkori. Online learning for linearly parametrized control problems. 2013.
- Justin Whitehouse, Zhiwei Steven Wu, and Aaditya Ramdas. Improved self-normalized concentration in hilbert spaces: Sublinear regret for gp-ucb. arXiv preprint arXiv:2307.07539, 2023.
- Rui Gao. Finite-sample guarantees for wasserstein distributionally robust optimization: Breaking the curse of dimensionality. *Operations Research* 71(6): 2291–2306, 2023.
- Johannes Kirschner and Andreas Krause. Information directed sampling and bandits with heteroscedastic noise. In *Conference On Learning Theory*, pages 358–384. PMLR, 2018.
- Ilija Bogunovic and Andreas Krause. Misspecified gaussian process bandit optimization. *Advances in neural information processing systems* 34:3004–3015, 2021.
- Bernhard Schölkopf and Alexander J Smola. *Learning with kernels: support vector machines, regularization, optimization, and beyond* MIT press, 2002.
- Christopher KI Williams and Carl Edward Rasmussen. *Gaussian processes for machine learning* volume 2. MIT press Cambridge, MA, 2006.
- Sayak Ray Chowdhury and Aditya Gopalan. On kernelized multi-armed bandits. In *International Conference on Machine Learning* pages 844–853. PMLR, 2017.
- Sattar Vakili, Kia Khezeli, and Victor Picheny. On information gain and regret bounds in gaussian process bandits. In *International Conference on Artificial Intelligence and Statistics*, pages 82–90. PMLR, 2021.
- P Auer. Finite-time analysis of the multiarmed bandit problem, 2002.
- Nicolas Fournier and Arnaud Guillin. On the rate of convergence in wasserstein distance of the empirical measure. *Probability theory and related fields*, 162(3):707–738, 2015.
- Nicolas Fournier. Convergence of the empirical measure in expected wasserstein distance: non asymptotic explicit bounds in \mathbb{R}^d . arXiv preprint arXiv:2209.00923, 2022.
- Zhuoran Yang, Chi Jin, Zhaoran Wang, Mengdi Wang, and Michael I Jordan. On function approximation in reinforcement learning: Optimism in the face of large state spaces. arXiv preprint arXiv:2011.04622, 2020.
- Ilija Bogunovic, Jonathan Scarlett, Stefanie Jegelka, and Volkan Cevher. Adversarially robust optimization with gaussian processes. *Advances in neural information processing systems* 31, 2018.
- Motonobu Kanagawa, Philipp Hennig, Dino Sejdinovic, and Bharath K Sriperumbudur. Gaussian processes and kernel methods: A review on connections and equivalences. arXiv preprint arXiv:1807.02582, 2018.
- Henk J Van Waarde and Rodolphe Sepulchre. Kernel-based models for system analysis. *IEEE Transactions on Automatic Control*, 68(9):5317–5332, 2022.

Sattar Vakili and Julia Olkhovskaya. Kernelized reinforcement learning with order optimal regret bounds. *Advances in Neural Information Processing Systems* 36:4225–4247, 2023.

Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge university press, 2018.

Maximilian Balandat, Brian Karrer, Daniel Jiang, Samuel Daulton, Ben Letham, Andrew G Wilson, and Eytan Bakshy. Botorch: A framework for efficient monte-carlo bayesian optimization. *Advances in neural information processing systems* 33:21524–21538, 2020.

Steven Diamond and Stephen Boyd. Cvxpy: A python-embedded modeling language for convex optimization. *Journal of Machine Learning Research* 17(83): 1–5, 2016.

Checklist

1. For all models and algorithms presented, check if you include:
 - (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model. [Yes]
 - (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm. [Yes]
 - (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries. [Yes]
2. For any theoretical claim, check if you include:
 - (a) Statements of the full set of assumptions of all theoretical results. [Yes]
 - (b) Complete proofs of all theoretical results. [Yes]
 - (c) Clear explanations of any assumptions. [Yes]
3. For all figures and tables that present empirical results, check if you include:
 - (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL). [Yes]
 - (b) All the training details (e.g., data splits, hyperparameters, how they were chosen). [Yes]
 - (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times). [Yes]
- (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider). [Yes]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:
 - (a) Citations of the creator If your work uses existing assets. [Yes]
 - (b) The license information of the assets, if applicable. [Yes]
 - (c) New assets either in the supplemental material or as a URL, if applicable. [Yes]
 - (d) Information about consent from data providers/curators. [Not Applicable]
 - (e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content. [Not Applicable]
5. If you used crowdsourcing or conducted research with human subjects, check if you include:
 - (a) The full text of instructions given to participants and screenshots. [Not Applicable]
 - (b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable. [Not Applicable]
 - (c) The estimated hourly wage paid to participants and the total amount spent on participant compensation. [Not Applicable]

Wasserstein Distributionally Robust Bayesian Optimization with Continuous Context: Supplementary Materials

7 BACKGROUND ON RKHS AND KERNEL RIDGE REGRESSION

In this paper, we consider the frequentist perspective and formulate the surrogate model as the solution of a regularized least-squares regression problem in the Reproducing Kernel Hilbert Space (RKHS). A similar formulation can be derived following the Bayesian perspective of Gaussian Process Regression under suitable assumptions on the Gaussian Process prior and observation noise Kanagawa et al. (2018).

Consider an RKHS $(H_k; h; \langle \cdot, \cdot \rangle_{H_k})$ with reproducing kernel $k : Z \times Z \rightarrow \mathbb{R}$. Define the inner product of the RKHS as $\langle f, g \rangle = \langle hf, hg \rangle_{H_k}$ and the outer product as $fg^\top = \langle f, hg \rangle_{H_k}$. Let $\phi_t := (k(\cdot; z_1), \dots, k(\cdot; z_t))^\top$ be the feature map of the RKHS for a sequence of points $z_1, \dots, z_t \in Z$. Define the kernel matrix $K_t = \phi_t \phi_t^\top$ and the covariance operator $V_t = \int \phi_t \phi_t^\top$. The RKHS norm of a function $f \in H_k$ is defined as $\|f\|_{H_k}^2 = \langle hf, hf \rangle_{H_k}$. By the reproducing property of the kernel, we have that $f(z) = \langle hf, k(\cdot; z) \rangle_{H_k}$ for all $z \in Z$.

With a slight abuse of notation we write the following equality which will be useful in the upcoming derivations

$$(\phi_t \phi_t^\top + I)^{-1} = I - \phi_t (\phi_t \phi_t^\top + I)^{-1} \phi_t^\top; \quad (19)$$

where it should be clear from the context that I is either the identity matrix or the identity operator in the RKHS. We also define the short-hand notation $V_t = V_t + I = \int \phi_t \phi_t^\top + I$.

Given the observed data $D_t = \{(z_i, y_i)\}_{i=1}^t$, the regularized least-squares regression problem in RKHS is defined as follows:

$$\min_{f \in H_k} \sum_{i=1}^t (y_i - f(z_i))^2 + \|f\|_{H_k}^2; \quad (20)$$

The solution to this problem is given by:

$$\begin{aligned} \hat{f}_t &= V_t^{-1} \int \phi_t y_{1:t} \\ &= (V_t + I)^{-1} \int \phi_t y_{1:t} \\ &= \int (\phi_t \phi_t^\top + I)^{-1} \phi_t y_{1:t} \end{aligned} \quad (21)$$

where $y_{1:t} = (y_1, \dots, y_t)^\top$ is the vector of observed responses. By the representation theorem, we can compute at some new point $z \in Z$ as follows:

$$\begin{aligned} \hat{f}_t(z) &= \langle \hat{f}_t, k(\cdot; z) \rangle_{H_k} \\ &= \int (\phi_t \phi_t^\top + I)^{-1} \phi_t y_{1:t}^\top k(\cdot; z) \\ &= k_t(z)^\top (K_t + I)^{-1} y_{1:t}; \end{aligned} \quad (22)$$

where $k_t(z) = (k(z; z_1), \dots, k(z; z_t))^\top$ is the vector of kernel evaluations at z .

We can also compute

$$\begin{aligned} \hat{\sigma}_t^2(z) &:= k k(\cdot; z) k_{V_t}^2 \\ &= \frac{1}{\phi_t^\top (\phi_t \phi_t^\top + I)^{-1} \phi_t} k_t(z)^\top (K_t + I)^{-1} k_t(z); \end{aligned} \quad (23)$$

where $k k(\cdot; z) k_{V_t}^2 = k_t(z)^\top (\phi_t \phi_t^\top + I)^{-1} k_t(z)$, and we use (19) to get the final equation.

We are using the notation $\hat{f}_t(z)$ and $\hat{\sigma}_t^2(z)$ to align with the Gaussian Process Regression literature Kanagawa et al. (2018), where $\hat{f}_t(z)$ and $\hat{\sigma}_t^2(z)$ would represent the mean and variance of the Gaussian Process posterior at z respectively.

7.1 Kernels that satisfy the Lipschitz condition in Assumption 1

Assumption 1 is satisfied for commonly used kernels. For example, it is satisfied with $L = 1$ for the squared exponential kernel and the Matérn kernel for $\nu = 3/2$ (Van Waarde and Sepulchre, 2022, Proposition 2). In fact, all smooth, positive definite, stationary kernels that have zero derivatives at zero satisfy Assumption 1. This, in turn, implies that Assumption 1 is satisfied for Matérn kernels with $\nu = p + 1/2$, for $p = 1; 2; \dots$.

Lemma 11. Let k be a positive definite, stationary kernel such that $k(x; x^0) = r(kx - x^0k)$, for some function $r : \mathbb{R} \rightarrow \mathbb{R}$ that is continuously twice differentiable in a neighborhood of the origin with first derivative $r^{(1)}(0) = 0$. Then, the kernel-induced distance

$$d(x; x^0) := \sqrt[p]{\frac{k(x; x) + k(x^0; x^0) - k(x; x^0) - k(x^0; x)}{2}} = M \|kx - x^0k\|;$$

for some constant $M > 0$.

Proof. Replacing k with r in the expression of $d(x; x^0)$ we write

$$d(x; x^0) = \sqrt[p]{\frac{2r(0) - 2r(kx - x^0k)}{2}};$$

Pick any $\epsilon > 0$. Let's consider two cases:

Case 1: For $kx - x^0k \leq \epsilon$, by the positive definite property, we have $|r(kx - x^0k) - r(0)| \leq \epsilon$ for any $x; x^0$. Therefore:

$$d(x; x^0) = \sqrt[p]{\frac{2r(0) - 2r(kx - x^0k)}{2}} \leq \sqrt[p]{\frac{2\epsilon}{2}} = M_1 \|kx - x^0k\|;$$

where $M_1 = \sqrt[p]{2\epsilon/r(0)}$.

Case 2: For $kx - x^0k > \epsilon$, using the Taylor remainder formula and the fact that $r^{(1)}(0) = 0$:

$$r(kx - x^0k) = r(0) + \frac{kx - x^0k^2}{2} r^{(2)}(s);$$

for some $s \in [0; kx - x^0k]$. Since $r(kx - x^0k) \leq r(0)$, we have $kx - x^0k^2 r^{(2)}(s) \leq 0$, which, in turn, implies that $r^{(2)}(s) \leq 0$. As a result,

$$d(x; x^0) = \sqrt[p]{\frac{kx - x^0k^2 r^{(2)}(s)}{2}} \leq \sqrt[p]{\frac{kx - x^0k^2}{2} \max_{s \in [0; \epsilon]} r^{(2)}(s)} = M_2 \|kx - x^0k\|;$$

Since the function has a continuous second derivative in the interval $[0; \epsilon]$ and $r^{(2)}(s) \leq 0$ for all $s \in [0; \epsilon]$, the maximum $M_2 = \sqrt[p]{\max_{s \in [0; \epsilon]} r^{(2)}(s)}$ is well-defined and finite.

The result follows by taking

$$M = \max\{M_1; M_2\} = \max\left(\sqrt[p]{\frac{2\epsilon}{2r(0)}}, \sqrt[p]{\max_{s \in [0; \epsilon]} r^{(2)}(s)}\right)$$

□

8 MAIN PROOFS

We can state here a well-known result from the Wasserstein DR optimization literature Kuhn et al. (2019); Gao et al. (2024).

Lemma 12. Consider a function $g : \mathbb{R}^g \rightarrow \mathbb{R}$ that is L^g -Lipschitz, i.e. $|g(\theta) - g(\theta')| \leq L^g \|\theta - \theta'\|$, for all $\theta; \theta' \in \mathbb{R}^g$. Let $B^g(\tilde{\theta})$ be a Wasserstein ambiguity set defined as a ball of radius ϵ in the Wasserstein distance centered at the distribution $\tilde{\theta}$. Then,

$$\sup_{Q \in \mathcal{B}^g(\tilde{\theta})} \mathbb{E}_Q[g(\theta)] \leq \mathbb{E}_{\tilde{\theta}}[g(\theta)] + L^g \epsilon; \quad (24)$$

Similarly,

$$\inf_{Q \in \mathcal{B}^g(\tilde{\theta})} \mathbb{E}_Q[g(\theta)] \geq \mathbb{E}_{\tilde{\theta}}[g(\theta)] - L^g \epsilon; \quad (25)$$

As a consequence of Lemma 12, we can state the following result.

Lemma 13. Let $f : X \times C \rightarrow \mathbb{R}$ be a function that is $L_c^f(x)$ -Lipschitz in the context space, i.e. $|f(x; c) - f(x; c^0)| \leq L_c^f(x) \|c - c^0\|$, for all $c; c^0 \in C$. Let $B^\pi(\tilde{\mathbf{P}})$ be a Wasserstein ambiguity set defined as a ball of radius π in the Wasserstein distance centered at the distribution $\tilde{\mathbf{P}}$. Then, for any $x \in X$ and for any distribution $\mathbf{P} \in B^\pi(\tilde{\mathbf{P}})$, we have that

$$|E_{c \sim \mathbf{P}}[f(x; c)] - E_{c \sim \tilde{\mathbf{P}}}[f(x; c)]| \leq L_c^f(x) \pi \quad (26)$$

Lemma 14 (Lemma 3 in the main text). Let $0 < \epsilon < 1$ be a failure probability and let

$$B_t := \frac{1}{2} R \sqrt{\frac{1}{2 \log \frac{\det(I + \frac{1}{2} K_t)}{\det(I + \frac{1}{2} K_{t-1})}} + B} + \frac{1}{2} R \sqrt{\frac{1}{2 \log \frac{1}{2} + R^2 \frac{1}{2} + B}} : \quad (27)$$

Then, with probability $1 - \epsilon$ for all $t \geq 1$ we have $k_{t-1} k_{H_k} \leq B_t$. Further, if Assumption 1 holds (i.e., $k(\cdot; z) = k(\cdot; z^0) k_{H_k} - L_k z - z^0$), we have:

(i) With probability $1 - \epsilon$, for any $z; z^0 \in X \times C : |j_{t-1}(z) - j_{t-1}(z^0)| \leq B_t L_k z - z^0$:

(ii) For any $z; z^0 \in X \times C : |j_{t-1}(z) - j_{t-1}(z^0)| \leq \frac{1}{2} L_k z - z^0 = B_t L_k z - z^0$:

Therefore, with probability $1 - \epsilon$, the UCB function is Lipschitz continuous with constant: $L^{UCB_t} \leq 2B_t L$.

Proof. The UCB is defined as:

$$UCB_t(z) = j_{t-1}(z) + j_{t-1}(z);$$

where $z = (x; c) \in X \times C$.

We have

$$j_t(z) = h(I + V_t)^{-1} \sum_{i=1:t} y_{i:t} k(\cdot; z); \quad (28)$$

and

$$k_{t-1} k_{H_k} = k(I + V_t)^{-1} \sum_{i=1:t} y_{i:t} k_{H_k} \quad (29)$$

$$= k(I + V_t)^{-1} \sum_{i=1:t} (f(z_{i:t}) + \frac{1}{2} k_{H_k}) \quad (30)$$

$$= k(I + V_t)^{-1} \sum_{i=1:t} (f + \frac{1}{2} k_{H_k}) \quad (31)$$

$$= k(I + V_t)^{-1} \sum_{i=1:t} f k_{H_k} + k(I + V_t)^{-1} \sum_{i=1:t} \frac{1}{2} k_{H_k} \quad (32)$$

$$= k(I + V_t)^{-1} V_t f k_{H_k} + k(I + V_t)^{-1} \sum_{i=1:t} \frac{1}{2} k_{H_k} \quad (33)$$

$$= k(I + V_t)^{-1} V_t k_{H_k} k f k_{H_k} + k(I + V_t)^{-1} \sum_{i=1:t} \frac{1}{2} k_{H_k} k(I + V_t)^{-1} \sum_{i=1:t} \frac{1}{2} k_{H_k} \quad (34)$$

$$= B + \frac{1}{2} k(I + V_t)^{-1} \sum_{i=1:t} k_{H_k}; \quad (35)$$

where we used the assumption that $k f k_{H_k} \leq B$ and the fact that $k(I + V_t)^{-1} V_t k_{H_k} \leq 1$ for $\epsilon > 0$.

Applying Corollary 3.6 of Abbasi-Yadkori (2013), with probability at least $1 - \epsilon$ and for all $t \geq 1$, we have

$$k(I + V_t)^{-1} \sum_{i=1:t} k_{H_k} \leq \frac{1}{2} R \sqrt{\frac{1}{2 \log \frac{\det(I + \frac{1}{2} V_t)}{\det(I + \frac{1}{2} V_{t-1})}} + B} = \frac{1}{2} R \sqrt{\frac{1}{2 \log \frac{1}{2} + R^2 \frac{1}{2} + B}} : \quad (36)$$

Thus, we obtain

$$k_{t-1} k_{H_k} \leq \frac{1}{2} R \sqrt{\frac{1}{2 \log \frac{\det(I + \frac{1}{2} K_t)}{\det(I + \frac{1}{2} K_{t-1})}} + B} = B_t : \quad (37)$$

If Assumption 1 holds, for any $z; z^0 \in \mathcal{C}$, we have:

$$j_{t-1}(z) - j_{t-1}(z^0) = j_{h_{t-1}; k(\cdot; z) - k(\cdot; z^0)} i_{H_k} j \quad (38)$$

$$k_{t-1} k_{H_k} k k(\cdot; z) - k(\cdot; z^0) k_{H_k} \quad (39)$$

$$B_t L k z - z^0 k \quad (40)$$

where we used the Cauchy-Schwarz inequality and Assumption 1.

For the term $j_{t-1}(z)$, we need to analyze the Lipschitz property of $j_{t-1}(z)$. We know by the Woodbury identity:

$$j_{t-1}(z) = h k(\cdot; z); (I + V_{t-1})^{-1} k(\cdot; z) i_{H_k} = \frac{1}{K_t + I} k(z; z) - k_{t-1}(z)^T (K_t + I)^{-1} k_{t-1}(z) \quad (41)$$

This gives us:

$$j_{t-1}(z) = k(I + V_{t-1})^{-1/2} k(\cdot; z) k_{H_k} \quad (42)$$

Now, for the Lipschitz property, by the triangle inequality:

$$j_{t-1}(z) - j_{t-1}(z^0) \leq k(I + V_{t-1})^{-1/2} (k(\cdot; z) - k(\cdot; z^0)) k_{H_k} \quad (43)$$

$$k(I + V_{t-1})^{-1/2} k_{op} k k(\cdot; z) - k(\cdot; z^0) k_{H_k} \quad (44)$$

$$\frac{1}{2} L k z - z^0 k \quad (45)$$

where we used the fact that $k(I + V_{t-1})^{-1/2} k_{op} \leq \frac{1}{2}$ and Assumption 1.

Therefore:

$$j_{t-1}(z) - j_{t-1}(z^0) \leq \frac{1}{2} L k z - z^0 k \quad (46)$$

Note that, recalling the definition of t :

$$t := R^{\frac{1}{2}} \frac{\log \frac{1}{\delta}}{\det(I + K_{t-1})^{\frac{1}{2}}} + \frac{1}{2} B; \quad (47)$$

we can observe that $t^{-\frac{1}{2}} = B_t$.

Combining the results, with probability $1 - \delta$, we have:

$$j_{UCB}(z) - UCB(z^0) = j_{t-1}(z) + j_{t-1}(z) - j_{t-1}(z^0) - j_{t-1}(z^0) \quad (48)$$

$$j_{t-1}(z) - j_{t-1}(z^0) + j_{t-1}(z) - j_{t-1}(z^0) \quad (49)$$

$$B_t L k z - z^0 k + t^{-\frac{1}{2}} L k z - z^0 k \quad (50)$$

$$= (B_t L + B_t L) k z - z^0 k \quad (51)$$

$$= 2 B_t L k z - z^0 k \quad (52)$$

where we used the fact that $t^{-\frac{1}{2}} = B_t$.

Thus, with probability $1 - \delta$, the Lipschitz constant of the UCB function is:

$$L^{UCB_t} = 2 B_t L \quad (53)$$

which concludes the proof. \square

Theorem 15 (Instantaneous expected regret -Thm. 4 in the main text). Let Assumption 1 hold. Fix a failure probability $0 < \delta < 1$. With probability at least $1 - \delta$, for all $t \geq 1$ the instantaneous expected regret can be bounded by

$$r_t = E_{\mathcal{C}P_t} [2 j_{t-1}(x_t; c)] + 2 t^{-\frac{1}{2}} L^{UCB_t}(x_t) \quad (54)$$

Proof. Recall that the benchmark solution x_t is the optimal solution to the stochastic optimization problem at time-step t , given access to the true function f and context distribution P_t

$$x_t = \arg \max_{x \in \mathcal{X}} E_{c \sim P_t} [f(x; c)] :$$

Whereas x_t is the solution to the robustified UCB acquisition function as given in (13)

$$x_t = \arg \max_{x \in \mathcal{X}} E_{c \sim \hat{P}_t} [\text{UCB}_t(x; c)] - \beta_t L^{\text{UCB}_t}(x) ;$$

From the definition of instantaneous regret, we can write:

$$\begin{aligned} r_t &= E_{c \sim P_t} [f(x_t; c)] - E_{c \sim \hat{P}_t} [f(x_t; c)] \\ &\stackrel{(i)}{=} E_{c \sim P_t} [\text{UCB}_t(x_t; c)] - E_{c \sim \hat{P}_t} [\text{LCB}_t(x_t; c)] \\ &\stackrel{(ii)}{=} E_{c \sim \hat{P}_t} [\text{UCB}_t(x_t; c)] + \beta_t L^{\text{UCB}_t}(x_t) - E_{c \sim P_t} [\text{LCB}_t(x_t; c)] \\ &\stackrel{(iii)}{=} E_{c \sim \hat{P}_t} [\text{UCB}_t(x_t; c)] - \beta_t L^{\text{UCB}_t}(x_t) + 2\beta_t L^{\text{UCB}_t}(x_t) - E_{c \sim P_t} [\text{LCB}_t(x_t; c)] \\ &\stackrel{(iv)}{=} E_{c \sim P_t} [\text{UCB}_t(x_t; c)] + 2\beta_t L^{\text{UCB}_t}(x_t) - E_{c \sim P_t} [\text{LCB}_t(x_t; c)] \\ &\stackrel{(v)}{=} E_{c \sim P_t} [\beta_t f(x_t; c) + \beta_t f(x_t; c) - \beta_t f(x_t; c) + \beta_t f(x_t; c)] + 2\beta_t L^{\text{UCB}_t}(x_t) \\ &= E_{c \sim P_t} [2\beta_t f(x_t; c)] + 2\beta_t L^{\text{UCB}_t}(x_t) : \end{aligned} \tag{55}$$

Where (i) holds with probability 1 by definition of UCB and LCB, with x_t chosen as in Lemma 1. (ii) holds by applying Lemma 2 under the assumption that $P_t \in \mathcal{B}^{\beta_t}(\hat{P}_t)$. In (iii) we add and subtract $\beta_t L^{\text{UCB}_t}(x_t)$ and use the fact that x_t is the maximizer of the acquisition function (13). The inequality (iv) follows from another application of Lemma 2, and finally (v) follows again from the definitions of the UCB and the LCB. \square

Theorem 16 (cumulative expected regret - Thm. 5 in the main text). Let Assumption 1 hold. Fix a failure probability $0 < \delta < 1$. With probability at least $1 - \delta$, the cumulative expected regret after T steps can be bounded as:

$$R_T \leq 4\gamma_T T + 4 \log \frac{6}{\delta} + \sum_{t=1}^T \beta_t 2L^{\text{UCB}_t}(x_t) ; \tag{56}$$

where γ_T is the maximum information gain (Srinivas et al., 2009; Chowdhury and Gopalan, 2017; Vakili et al., 2021) at time T , which is defined as

$$\gamma_t := \sup_{z_1, z_2, \dots, z_t} \log \det I + \beta_t K_{t-1} :$$

Proof. Starting from the definition of cumulative expected regret:

$$\begin{aligned} R_T &= \sum_{t=1}^T r_t = \sum_{t=1}^T E_{c \sim P_t} [f(x_t; c)] - E_{c \sim \hat{P}_t} [f(x_t; c)] \\ &\stackrel{(i)}{=} \sum_{t=1}^T E_{c \sim P_t} [2\beta_t f(x_t; c)] + \sum_{t=1}^T \beta_t 2L^{\text{UCB}_t}(x_t) \\ &\quad - 2\gamma_T \sum_{t=1}^T E_{c \sim P_t} [\beta_t f(x_t; c)] + \sum_{t=1}^T \beta_t 2L^{\text{UCB}_t}(x_t) \\ &\stackrel{(ii)}{=} 2\gamma_T \sum_{t=1}^T \frac{\sum_{t=1}^T E_{c \sim P_t} [\beta_t f(x_t; c)]^2}{T} + \sum_{t=1}^T \beta_t 2L^{\text{UCB}_t}(x_t) \\ &\stackrel{(iii)}{=} 2\gamma_T \sum_{t=1}^T \frac{\sum_{t=1}^T E_{c \sim P_t} [\beta_t f(x_t; c)^2]}{T} + \sum_{t=1}^T \beta_t 2L^{\text{UCB}_t}(x_t) \end{aligned} \tag{57}$$

Where the inequality (i) follows from Theorem 15, (ii) follows from the Cauchy-Schwarz inequality, and (iii) follows from Jensen's inequality.

We can now apply the concentration of conditional mean result from Lemma 7 of Kirschner et al. (2020) (see also Lemma 3 of Kirschner and Krause (2018)), with probability at least $1 - \delta$ we obtain for all T :

$$\begin{aligned}
 & \sum_{t=1}^T \mathbb{E}_{c|P_{t-1}} \left(\sum_{i=1}^I (x_t; c_i)^2 \right) \\
 (i) \quad & \sum_{t=1}^T \left(\sum_{i=1}^I (x_t; c_i)^2 + 8 \log \frac{6}{\delta} \right) \\
 (ii) \quad & \sum_{t=1}^T \left(2 \log(1 + \sum_{i=1}^I (x_t; c_i)^2) + 8 \log \frac{6}{\delta} \right) \\
 (iii) \quad & 4T + 16 \log \frac{6}{\delta}
 \end{aligned} \tag{58}$$

where (i) follows from Lemma 7 of Kirschner et al. (2020) noting that $k(z; z^0) \leq 1$ by assumption, (ii) follows from the fact that $x \leq 2 \log(1 + x)$ for all $x \geq 0$, and (iii) follows from the definition of maximum information gain. By substituting this result into the cumulative regret expression we get with probability $1 - \delta$:

$$R_T \leq 4T + 16 \log \frac{6}{\delta} + \sum_{t=1}^T \sum_{i=1}^I 2L^{UCB_i}(x_t); \tag{59}$$

which concludes the proof. \square

In order to provide rates for the cumulative expected regret we want to provide high probability bounds for the Lipschitz constants $L^{UCB_i}(x)$.

Corollary 17 (Corollary 6 in the main text: General WDRBO Regret Order). Let $0 < \delta < 1$ be a failure probability and let Assumption 1 hold. Then, with probability $1 - \delta$, the cumulative expected regret is of the order of

$$R_T = \mathcal{O} \left(\sum_{t=1}^T \sum_{i=1}^I \frac{1}{\sqrt{t}} \right);$$

For the Squared Exponential kernel, this reduces to

$$R_T = \mathcal{O} \left(\sum_{t=1}^T \sum_{i=1}^I \frac{1}{\sqrt{t}} \right);$$

where \mathcal{O} omits logarithmic terms.

Proof. We can combine the results of Theorem 5 and Lemma 6 to write

$$\begin{aligned}
 R_T & \leq 4T + 16 \log \frac{6}{\delta} + \sum_{t=1}^T \sum_{i=1}^I 2L^{UCB_i}(x_t) \\
 & \leq 4T + 16 \log \frac{6}{\delta} + \sum_{t=1}^T \sum_{i=1}^I \left(\frac{1}{\sqrt{t}} \left(R \frac{1}{2 \log \frac{1}{\delta}} + R^p \frac{1}{2t} + B \right) L \right) \\
 & \leq 4 \left(R^p \frac{1}{2 \log \frac{1}{\delta}} + R^p \frac{1}{2T} + \frac{1}{2} B \right) T + 16 \log \frac{6}{\delta} + \sum_{t=1}^T \sum_{i=1}^I \left(\frac{1}{\sqrt{t}} \left(R \frac{1}{2 \log \frac{1}{\delta}} + R^p \frac{1}{2t} + B \right) L \right) \tag{60} \\
 & = \mathcal{O} \left(\sum_{t=1}^T \sum_{i=1}^I \frac{1}{\sqrt{t}} \right) + \sum_{t=1}^T \sum_{i=1}^I \frac{1}{\sqrt{t}} \mathcal{O} \left(\frac{1}{\sqrt{t}} \sum_{i=1}^I \frac{1}{\sqrt{t}} \right) \\
 & = \mathcal{O} \left(\sum_{t=1}^T \sum_{i=1}^I \frac{1}{\sqrt{t}} \right) + \sum_{t=1}^T \sum_{i=1}^I \frac{1}{\sqrt{t}} \mathcal{O} \left(\frac{1}{\sqrt{t}} \right);
 \end{aligned}$$

which proves the first statement.

The maximum information gain for the Squared Exponential kernel can be bounded as Vakili et al. (2021):

$$\gamma_t \leq \mathcal{O}(\log^{d+1}(t)) .$$

Thus, the rate for the cumulative expected regret is

$$R_T = \tilde{\mathcal{O}} \left(\sqrt{T} + \sum_{t=1}^T \varepsilon_t \right)$$

□

8.1 Proofs of the data-driven scenario

We can now show how this result translates to the data-driven formulation, and provide a rate for the cumulative expected regret.

Define the class of UCB functions as

$$\mathcal{U}(B) = \{h : h(z) = \mu(z) + \beta_t \sigma_{t-1}(z), \|\mu\|_{\mathcal{H}_k} \leq B\} ,$$

and let $\mathcal{N}_\infty(\rho, \mathcal{H}_k, B)$ be its ρ -covering number. We use a result from Yang et al. (2020) to obtain a bound on $\mathcal{N}_\infty(\rho, \mathcal{H}_k, B)$, specialized to the squared exponential kernel. The proposed analysis works in more general settings with kernels experiencing either exponential or polynomial eigendecay, see e.g. Yang et al. (2020); Vakili and Olkhovskaya (2023).

Lemma 18 (Adapted from Lemma D.1 of Yang et al. (2020)). *Let $k(z, z')$ be the squared exponential kernel, $\|f\|_{\mathcal{H}_k} \leq B$, and $k(z, z) \leq 1$. Then, there exist global constant C_N such that*

$$\log \mathcal{N}_\infty(\rho, \mathcal{H}_k, B) \leq \frac{C_N}{d} \left(1 + \log \frac{B}{\rho} \right)^{1+d} + \frac{C_N}{d} \left(1 + \log \frac{1}{\rho} \right)^{1+2d}$$

The following result is an adaptation of Corollary 2 of Gao (2023). Let $\text{diam}(\mathcal{X})$, $\text{diam}(\mathcal{C})$ denote the diameters of the sets \mathcal{X}, \mathcal{C} respectively.

Lemma 19. *Let $0 < \delta < 1$ be a failure probability. Let*

$$\varepsilon_t(\rho) = \frac{\tau}{t} \frac{\log 1/\delta + d_x \log(1 + 2\rho^{-1} \text{diam}(\mathcal{X})) + \log \mathcal{N}_\infty(\rho, \mathcal{H}_k, B)}{t},$$

where $\tau = 2\text{diam}(\mathcal{C})$. With probability at least $1 - \delta$

$$\forall h \in \mathbb{S}_B, \forall x \in \mathcal{X} : \mathbb{E}_{c \sim \mathcal{P}} [h(x, c)] \leq \mathbb{E}_{c \sim \hat{\mathcal{P}}_t} [h(x, c)] + \varepsilon_t(\rho) L_c^h(x) + 3(1 + LB)\rho. \quad (61)$$

Proof. Let $h \in \mathbb{S}_B$, $x \in \mathcal{X}$. Following the notation of Gao (2023), let for any $\varepsilon > 0$

$$\mathcal{R}_{\mathcal{P}}(\varepsilon, h(x, \cdot)) = \sup_{Q \in \mathcal{B}^+(\mathcal{P})} \mathbb{E}_{c \sim Q}(h(x, c)) - \mathbb{E}_{c \sim \mathcal{P}}(h(x, c)).$$

Using the bound on the Lipschitz constant derived in 3, it holds that

$$\mathcal{R}_{\mathcal{P}}(\varepsilon, h(x, \cdot)) \leq \varepsilon 2L\bar{B}_T \quad (62)$$

Let $\tilde{h} \in \mathbb{S}_B$, $\tilde{x} \in \mathcal{X}$. Following the proof of Corollary 2 in Gao (2023), we have

$$|\mathcal{R}_{\mathcal{P}}(\varepsilon, -h(x, \cdot)) - \mathcal{R}_{\mathcal{P}}(\varepsilon, -\tilde{h}(\tilde{x}, \cdot))| \leq \sup_{Q \in \mathcal{B}^+(\mathcal{P})} |\mathbb{E}_{c \sim Q}(\tilde{h}(\tilde{x}, c)) - \mathbb{E}_{c \sim Q}(h(x, c))|.$$

Using

$$\begin{aligned} |h(x, c) - \tilde{h}(\tilde{x}, c)| &= |h(x, c) - \tilde{h}(x, c) + \tilde{h}(x, c) - \tilde{h}(\tilde{x}, c)| \\ &\leq \|h - \tilde{h}\|_\infty + BL\|x - \tilde{x}\|, \end{aligned}$$

we obtain

$$\mathcal{R}_{\mathcal{P}}(\varepsilon, -h(x, \cdot)) - \mathcal{R}_{\mathcal{P}}(\varepsilon, -\tilde{h}(\tilde{x}, \cdot)) \leq \|h - \tilde{h}\|_\infty + BL\|x - \tilde{x}\| \quad (63)$$

Similarly

$$\begin{aligned} |\mathbb{E}_{c \sim \mathcal{P}}(\tilde{h}(\tilde{x}, c)) - \mathbb{E}_{c \sim \mathcal{P}}(h(x, c))| &\leq \|h - \tilde{h}\|_\infty + BL\|x - \tilde{x}\| \\ |\mathbb{E}_{c \sim \mathcal{P}_t}(\tilde{h}(\tilde{x}, c)) - \mathbb{E}_{c \sim \mathcal{P}_t}(h(x, c))| &\leq \|h - \tilde{h}\|_\infty + BL\|x - \tilde{x}\| \end{aligned} \quad (64)$$

Let \mathcal{X}_ρ be a covering of \mathcal{X} of precision ρ with respect to the Euclidean norm and $S_{\rho, B}$ be a covering of \mathbb{S}_B of precision ρ with respect to the infinity norm. We have that $|\mathcal{X}_\rho| \leq (1 + 2\text{diam}(\mathcal{X})\rho^{-1})^{d_x}$ (Vershynin, 2018, Ch. 4), while by definition $|S_{\rho, B}| = \mathcal{N}_\infty(\rho, \mathcal{H}_k, B)$.

By the definition of the coverings and (63), (64), we have that

$$\begin{aligned} \forall \tilde{h} \in S_{\rho, B}, \forall \tilde{x} \in \mathcal{X}_\rho \quad \mathbb{E}_{c \sim \mathcal{P}}[\tilde{h}(\tilde{x}, c)] &\leq \mathbb{E}_{c \sim \mathcal{P}_t}[\tilde{h}(\tilde{x}, c)] + \mathcal{R}_{\mathcal{P}}(\varepsilon, \tilde{h}(\tilde{x}, \cdot)) \Rightarrow \\ \forall h \in \mathbb{S}_B, \forall x \in \mathcal{X} \quad \mathbb{E}_{c \sim \mathcal{P}}[h(x, c)] &\leq \mathbb{E}_{c \sim \mathcal{P}_t}[h(x, c)] + \mathcal{R}_{\mathcal{P}}(\varepsilon, h(x, \cdot)) + 3 \min_{\tilde{h} \in S_{\rho, B}} \|h - \tilde{h}\|_\infty + 3BL \min_{\tilde{x} \in \mathcal{X}_\rho} \|x - \tilde{x}\| \\ &\leq \mathbb{E}_{c \sim \mathcal{P}_t}[h(x, c)] + \mathcal{R}_{\mathcal{P}}(\varepsilon, h(x, \cdot)) + 3(1 + BL)\rho. \end{aligned}$$

Hence,

$$\begin{aligned} \mathbb{P}(\forall h \in \mathbb{S}_B, \forall x \in \mathcal{X} : \mathbb{E}_{c \sim \mathcal{P}}[h(x, c)] &\leq \mathbb{E}_{c \sim \mathcal{P}_t}[h(x, c)] + \mathcal{R}_{\mathcal{P}}(\varepsilon, h(x, \cdot)) + 3(1 + BL)\rho) \\ &\geq 1 - \mathbb{P}(\mathbb{E}_{c \sim \mathcal{P}}[\tilde{h}(\tilde{x}, c)] \geq \mathbb{E}_{c \sim \mathcal{P}_t}[\tilde{h}(\tilde{x}, c)] + \mathcal{R}_{\mathcal{P}}(\varepsilon, \tilde{h}(\tilde{x}, \cdot))) \\ &\quad \text{for } \tilde{h} \in S_{\rho, B}, \tilde{x} \in \mathcal{X}_\rho \\ &\geq 1 - |\mathcal{X}_\rho| |S_{\rho, B}| e^{-\varepsilon^2 t / \tau}, \end{aligned}$$

where the last inequality follows from Theorem 1 in Gao (2023). The result follows from picking $\varepsilon = \varepsilon_t(\rho)$. \square

By integrating the result from Lemma 19 in place of Lemma 2, and following the same steps for the derivation of the instantaneous expected regret, we obtain the following cumulative expected regret for the data-driven scenario:

$$R_T \leq 4\beta_T \frac{1}{T\gamma_T + 4\log \frac{1}{\delta}} + \sum_{t=1}^T \varepsilon_t 2L^{\text{UCB}_t}(x_t^*) + \rho_t,$$

Lemma 20. Fix a failure probability $0 \leq \delta \leq 1$, and let

$$\varepsilon_t = \frac{1}{2\text{diam}(\mathcal{C})^2} \frac{\log 1/\delta + d_x \log(1 + 2t\text{diam}(\mathcal{X})) + \log \mathcal{N}_\infty((t)^{-1}, \mathcal{H}_k, \bar{B}_t)}{t},$$

and

$$\bar{B}_t = 2\lambda^{-\frac{1}{2}} \left(R \frac{1}{2\log \frac{1}{\delta}} + R \frac{1}{2\gamma_t} + B \right),$$

as in Lemma 3. Then, for the Squared Exponential kernel, the cumulative expected regret for the data-driven scenario is bounded by

$$R_T \leq \tilde{\mathcal{O}}(\beta_T \frac{1}{T\gamma_T} + \frac{1}{(d\gamma_T + \log 1/\delta)T}).$$

Proof. The proof of the first part follows by applying Lemma 19 to $-h(x, c)$, with $\rho = 1/t$, $B = \bar{B}_t$, and using the fact that $\|\text{UCB}_t\|_{\mathcal{H}_k} \leq \bar{B}_t$ from Lemma 3. We can combine these results to obtain bounds on the Lipschitz

$L^{\text{UCB}_t}(x_t^*)$ and on ρ_t . Thus,

$$\begin{aligned} R_T &\leq 4\beta_T \sqrt{\frac{6}{T\gamma_T + 4\log \frac{6}{\delta}}} + \sum_{t=1}^T \varepsilon_t 2L^{\text{UCB}_t}(x_t^*) + \rho_t \\ &\leq 4\beta_T \sqrt{\frac{6}{T\gamma_T + 4\log \frac{6}{\delta}}} + \sum_{t=1}^T \varepsilon_t 2\bar{B}_t L + 3(1 + \bar{B}_t L)/t \end{aligned}$$

For the Squared Exponential kernel we have that $\gamma_t \leq \mathcal{O}(\log^{d+1}(t))$ and the covering number bound for Squared Exponential kernels from Lemma 18.

Under the selections for ρ_t and ε_t , $\bar{B}_t = \mathcal{O}(\sqrt{\log 1/\delta + \log^{d/2} t})$, and

$$\log \mathcal{N}_\infty(1/t, \mathcal{H}_k, \bar{B}_t) = \mathcal{O}(d \log(\log t \{\sqrt{\log 1/\delta + \log^{d/2} t}\})) = \tilde{\mathcal{O}}(d),$$

where $\tilde{\mathcal{O}}$ hides logarithmic terms of t and $\log 1/\delta$. Hence,

$$\varepsilon_t = \tilde{\mathcal{O}} \left(\frac{\sqrt{\log 1/\delta + d}}{t} \right)$$

We can derive the rate for the cumulative expected regret

$$R_T \leq \tilde{\mathcal{O}}(\beta_T \sqrt{\frac{1}{T\gamma_T}} + \sqrt{\frac{1}{(d\gamma_T + \log 1/\delta)T}}).$$

$$\gamma_t \leq \mathcal{O}(\log^{d+1}(t)).$$

We have

$$R_T \leq \tilde{\mathcal{O}}(\sqrt{\frac{1}{T\gamma_T}}) = \tilde{\mathcal{O}}(\sqrt{T})$$

□

9 EXPERIMENTS DETAILS AND ADDITIONAL EXPERIMENTS

The experimental setup is based on an adaptation of the work of Huang et al. (2024). We exploit their implementation of the methods GP-UCB, SBO-KDE, DRBO-KDE, DRBO-MMD, DRBO-MMD Minmax, and StableOpt, available at <https://github.com/lamda-bbo/sbokde>, with only minor changes to make the code compatible with our hardware. We refer to Appendix B.1 of Huang et al. (2024) for more details. The algorithms, including ERBO and WDRBO, are implemented in BoTorch Balandat et al. (2020), with the inner convex optimization problems of DRBO-KDE and DRBO-MMD solved using CVXPY Diamond and Boyd (2016). Our code is available at the following link <https://github.com/frmicheli/WDRBO>.

We also exploited the implementations of the functions Ackley, Hartmann, Modified Branin, Newsvendor, Portfolio (Normal), and Portfolio (Uniform) of Huang et al. (2024). We refer to Appendix B.2 of Huang et al. (2024) for more details. The only variation has been in the choice of context distribution for Ackley, Hartmann, Modified Branin, Portfolio (Normal) where we used $c \sim \mathcal{N}(0.5, 0.2^2)$ with c clipped to $[0, 1]$. We implemented the Three Humps Camel function that is a standard benchmark function for global optimization algorithms. The input space is two-dimensional, we restricted it to the domain $x \in [-1, 1]$ and $c \in [-1, 1]$, and chose a uniform distribution for the context c . We function is defined as follows:

$$f(x) = 2x^2 - 1.05x^4 + \frac{x^6}{6} + xc + c^2. \quad (65)$$

In Fig. 3 we compare the performance of all the algorithms including DRBO-MMD, DRBO-MMD Minmax, and StableOpt, which we did not show in the main text. In Fig. 4 we show the instantaneous regret for all the algorithms which can help better compare the asymptotic performance of the different algorithms.

Table 2 reports the mean and standard error of computational times in seconds for the Ackley and Branin functions for all the considered algorithms.

	Ackley	Branin
WDRBO (ours)	44.3 ± 2.2	54.5 ± 2.6
ERBO (ours)	43.8 ± 1.6	45.2 ± 2.5
GP-UCB	15.7 ± 1.4	15.1 ± 1.0
SBO-KDE	46.7 ± 0.7	49.7 ± 1.5
DRBO-KDE	599.7 ± 33.0	525.0 ± 71.7
DRBO-MMD	644.9 ± 45.6	131.6 ± 9.6
DRBO-MMD Minimax	104.8 ± 3.9	21.1 ± 1.8
StableOpt	77.6 ± 2.4	64.7 ± 4.2

Table 2: Mean and standard error of computational times in seconds for the Ackley ($d_x = 1$, $d_c = 1$) and Branin ($d_x = 2$, $d_c = 2$) functions.

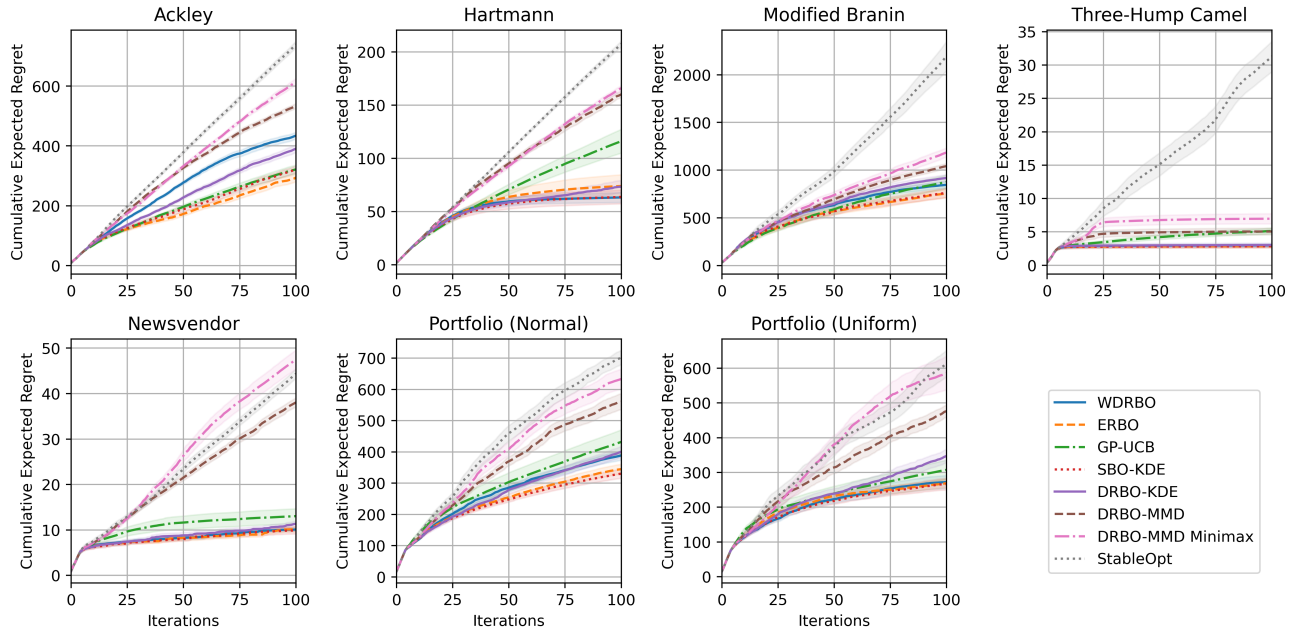


Figure 3: Mean and standard error of the cumulative expected regret.

