Significance of Gradient Information in Bayesian Optimization

Shubhanshu Shekhar
ECE Department, UCSD
shshekha@eng.ucsd.edu

Tara Javidi
ECE Department, UCSD
tjavidi@eng.ucsd.edu

Abstract

We consider the problem of Bayesian Optimization (BO) where the goal is to design an adaptive querying strategy to optimize a function $f\colon [0,1]^d \mapsto \mathbb{R}$. The function is assumed to be drawn from a Gaussian Process $\mathcal{GP}(0, K)$, and can only be accessed through noisy oracle queries. The most commonly used oracle in BO literature is the noisy Zeroth-Order-Oracle (ZOO) which returns noise-corrupted function value $y = f(x)+\eta$ at any point $x \in \mathcal{X}$ queried by the agent. A less studied oracle in BO is the First-Order-Oracle (FOO) which also returns noisy gradient value at the queried point. In this paper we consider the fundamental question of quantifying the possible improvement in regret that can be achieved under FOO access as compared to the case in which only ZOO access is available. Under some regularity assumptions on K, we first show that the expected cumulative regret $\mathcal{R}_n$ with ZOO of any algorithm must satisfy a lower bound of $\Omega(\sqrt{2^d n})$, where n is the query budget. This lower bound captures the appropriate scaling of the regret on both dimension d and budget n, and relies on a novel reduction from BO to a multi-armed bandit (MAB) problem. We then propose a two-phase algorithm which, with some additional prior knowledge, achieves a vastly improved $\mathcal{O}\left(d(\log n)^2\right)$ regret when given access to a FOO. Together, these two results highlight the significant value of incorporating gradient information in BO algorithms.

Proceedings of the 24$^{\text{th}}$ International Conference on Artificial Intelligence and Statistics (AISTATS) 2021, San Diego, California, USA. PMLR: Volume 130.

1 Introduction

We consider the problem of optimizing a function $f : \mathcal{X} = [0,1]^d \mapsto \mathbb{R}$ with the assumption that f is a sample from a zero-mean Gaussian Process (GP)[1] with a covariance function K. The agent can access the unknown objective function f through an oracle, which takes as input any point $x \in \mathcal{X}$ and returns a value u in some space $\mathcal{U}$, which provides some *local* information (Nesterov, 2013, § 1.1.2) about f around x. The most commonly used oracle in Bayesian Optimization (BO) is the Zeroth-Order-Oracle (ZOO), which returns a noisy function evaluation at the queried point, i.e., $u(x) = f(x) + \eta \in \mathbb{R}$. A relatively less considered oracle is the First-Order-Oracle (FOO) which returns both noisy function evaluation and noisy gradient values, i.e, $u(x) = (f(x) + \eta, \nabla f(x) + \xi) \in \mathbb{R}^{d+1}$.

Given a query budget of n samples and the oracle, the goal of an agent in BO is to design an adaptive sampling strategy $\mathcal{A}$ to efficiently learn about the global maximizer x^* of f. An adaptive (non-randomized) strategy $\mathcal{A}$ consists of a sequence of mappings $(A_t)_{t=1}^n$ where $A_t : (\mathcal{X} \times \mathcal{U})^{t-1} \mapsto \mathcal{X}$, which sequentially select a query point x_t at time t based on the history of actions and observations up to time $t-1$. The performance of the sampling strategy $\mathcal{A}$ is usually measured by the cumulative regret $\mathcal{R}_n$, defined as

$$\mathcal{R}_n(\mathcal{A},\, f) = \sum_{t=1}^{n} f(x^*) - f\left(A_t(x_{[1:t-1]}, u_{[1:t-1]})\right). \quad (1)$$

Since FOO provides additional information to the agent, it is natural to expect improvement in the achievable cumulative regret under FOO in comparison to ZOO. Some existing works in literature, such as Wu et al. (2017b,a) and Prabuchandran et al. (2020), have empirically demonstrated benefits of BO algorithms which incorporate gradient information in their execution. However, to the best of our knowledge, no prior work

[1] A related problem, not considered here, called *agnostic* GP bandit considers this problem in a non-Bayesian setting where f is a fixed but unknown element of an RKHS.

in literature quantifies the possible reduction in cumulative regret with FOO access. The main contribution of our paper is to provide a lower bound on the reduction in regret that is possible with ZOO access. We do this in two steps: *first*, we establish the fundamental limit of achievable regret with ZOO for all values of d, and *second* we design an algorithm with FOO access which, under some additional prior information, incurs much smaller regret.

We provide an overview of the results in Sec. 1.1 and discuss some relevant background in Sec. 1.2.

1.1 Overview of Results

Our main contributions are:

- In the first part of the paper (Sec. 3), we focus on establishing the limits of the achievable regret of any algorithm with ZOO. We begin by noting in Prop. 1 that the reduction to binary testing approach of Scarlett (2018) (for $d = 1$) can be directly used to obtain a lower bound on $\mathcal{R}_n$ of $\Omega\left(\sigma\sqrt{n}\right)$ for all $d \geq 1$. This result however does not capture the conjectured exponential scaling of $\mathcal{R}_n$ with d by Scarlett (2018). We rectify this in Theorem 1 by deriving an algorithm-independent lower bound on $\mathcal{R}_n$ with ZOO access matching the conjectured order $\Omega\left(\sqrt{2^d n}\right)$. This result relies on a novel analysis and approach which relates the regret of the BO problem to an appropriately defined multi-armed bandit (MAB) problem, and adapts the lower bounding techniques for MABs to obtain the final result.
- In Sec. 4 we focus on quantifying the possible gain in performance, beyond the limits established in Sec. 3, when the noisy gradient information is also available to the agent. To do so, we propose an algorithm using FOO access, referred to as AlgFOO. We show that under certain technical assumptions (formally stated in Sec. 2.3), AlgFOO can achieve an upper bound on the cumulative regret of $\mathcal{O}\left(d(\log n)^2\right)$ (Theorem 2 and Corollary 1). To the best of our knowledge, this is the first result which formally characterizes the significant benefits of using gradient information in Bayesian Optimization.

1.2 Background

In this section we discuss existing results in literature which provide the background context for our results.

Bayesian Optimization. As mentioned earlier, Bayesian Optimization (BO) refers to the *model based sequential optimization* of a black-box function, in which usually a Gaussian Process (GP) is used to model the function. The prior information about the unknown function is encoded by imposing appropriate restrictions on the kernel (or covariance function) K. Most BO algorithms usually alternate between these two steps: **(i)** update the model of the function (i.e., GP posterior) based on the data observed, and **(ii)** use the updated model to guide the design of the next query point, which is in some sense most informative about the maximizer x^*. The informativeness of a candidate point $x \in \mathcal{X}$ in step (ii) above is usually quantified via an *acquisition function*. The most commonly used acquisition function is the UCB acquisition function which was proposed and analyzed by Srinivas et al. (2012). Other acquisition functions include the *Expected Improvement (EI)*, *Probability of Improvement (PI)* and *Entropy Search* (Hennig and Schuler, 2012; Wang and Jegelka, 2017). An alternative approach is taken by the Thompson Sampling algorithm of Russo and Van Roy (2014) where the query points are drawn randomly with the probability that they are optimal. For a detailed discussion of various aspects of BO, see the surveys by Brochu et al. (2010) and Shahriari et al. (2015).

Lower Bounds. The algorithms mentioned above (and some others in literature) have guarantees on their cumulative regret of the form $\mathcal{O}(\sqrt{n \log n \gamma_n})$. However, in the absence of corresponding algorithm-agnostic lower bounds, it is not clear whether the existing regret bounds are optimal or they can be improved further. To the best of our knowledge, only Grünewälder et al. (2010) and Scarlett (2018) present lower bounds on the regret for BO, under some restrictions. Grünewälder et al. (2010) derived a worst case lower bound on the *simple regret* by constructing specific *hard* Gaussian Process. More relevant to our work, for $d = 1$ case, Scarlett (2018) derived $\Omega(\sigma\sqrt{n})$ lower bound on the average cumulative regret (over two randomly shifted GPs) and also proposed an algorithm with $\mathcal{O}(\sqrt{n})$ upper bound. The lower bound technique of Scarlett (2018) for $d > 1$ doesn't capture the conjectured exponential dependence on d; our first contribution fills this gap in literature by deriving algorithm-independent lower bounds of $\Omega(2^{d/2}\sqrt{n})$ for all $d \geq 1$.

For the non-Bayesian variant of this problem (where f is assumed to lie in the RKHS of kernel K), there exist algorithm-independent lower bounds on the regret with ZOO access in the noise-less setting (Bull, 2011) as well as in the noisy case (Scarlett et al., 2017; Cai and Scarlett, 2020). However, the techniques used in obtaining those results are not directly applicable in the fully Bayesian framework considered in this paper.

BO algorithms using derivatives. The above algorithms only utilize the zeroth order information about the objective function. Surprisingly, unlike other continuous optimization problems (such as convex optimization), there are very few works in BO literature which incorporate gradient information about f to guide the search for the optimizer. Wu et al. (2017b) proposed a derivative based knowledge gradient algorithm and proved its asymptotic consistency and one-step Bayes optimality. Wu et al. (2017a) exploited the first and second order derivative information for improved posterior inference, and applied it to BO and Bayesian Quadrature problems. Both these works also empirically demonstrated the benefits of incorporating gradient information on several benchmark functions. Some other works that use derivative information in BO are Osborne et al. (2009), who exploited the derivatives for better conditioning of covariance matrix and Lizotte (2008), who empirically studied the variants of EI and PI algorithms with derivative information. More recently, Prabuchandran et al. (2020) proposed a new `FOO` algorithm which exploits the fact that the gradient vanishes at the optimum, and empirically demonstrated its improved performance. However, to the best of our knowledge, no attempts have been made to provably quantify the improvement in regret that is achievable when the agent is given access to `FOO`, even in the simplest cases. Our second contribution addresses this issue and shows that there exists an algorithm with `FOO`, that can achieve a regret of $\mathcal{O}\left(d(\log n)^2\right)$ under some technical assumptions stated in Section 2.3, greatly improving upon the the lower bound with `ZOO` of $\Omega(\sqrt{n})$.

2 Preliminaries

In this section, we fix the notations used in Sec. 2.1, introduce some definitions in Sec. 2.2 and formally state and discuss all the assumptions in Sec. 2.3.

2.1 Notations

We denote by $f : \mathcal{X} \mapsto \mathbb{R}$ the objective function to be maximized, and set the domain $\mathcal{X} = [0,1]^d$. We endow the domain with the Euclidean norm $\|x\| = \|x\|_2 = \sqrt{\sum_{i=1}^d x_i^2}$ and use $\mathcal{X}^{o}$ and $\partial\mathcal{X}$ to represent the interior and boundary of $\mathcal{X}$ respectively. Furthermore, for any subset A of $\mathcal{X}$ and any $x \in \mathcal{X}$, we use $\|x - A\|$ to denote the distance of x from A, i.e., $\inf_{z\in A} \|x - z\|$. Also, for any $x, z \in \mathcal{X}$ we use $\langle x, z\rangle$ to denote the usual inner product, i.e, $\sum_{i=1}^d x_i z_i$. For any $x \in \mathcal{X}$ and $r > 0$, we use $B(x, r)$ to denote the radius r open-ball around x, i.e., $B(x, r) = \{z \in \mathcal{X} : \|z - x\| < r\}$. For square matrices M, we use $\|M\|$ to represent the spectral norm.

We assume that f is a sample from a zero mean Gaussian Process denoted by $\mathcal{GP}(0, K)$, where $K : \mathcal{X}\times\mathcal{X} \mapsto \mathbb{R}$ is the kernel (or covariance function). In this paper, we restrict ourselves to *stationary kernels*, i.e., kernels which satisfy $K(x, z) = K(x - z)$ for all $x, z \in \mathcal{X}$.

For a positive integer m, we use $[m]$ to denote the set $\{1, 2, \ldots, m\}$. Furthermore, for any finite set S, we use `Unif`(S) to denote the uniform random variable taking values in S. Finally, we will use the notation $\mathcal{M}(L, l, \sigma, \kappa)$ to represent a $(L+1)$ armed multi-armed bandit problem with distributions $(p_1, p_2, \ldots, p_{L+1})$ with $p_i \sim N(1/2 + \kappa, \sigma^2)$ if $i = l$ and $p_i \sim N(1/2, \sigma^2)$ otherwise.

2.2 Definitions

We begin by recalling the definition of Gaussian Processes (GPs).

Definition 1 (`GP`)**.** A zero mean Gaussian Process $\mathcal{GP}(0, K)$ indexed by a set $\mathcal{X}$ is a collection of random variables $\{Z_x : x \in \mathcal{X}\}$ such that for any finite $S \subset \mathcal{X}$, the random vector $\{Z_x : x \in S\}$ is distributed as $N(0, C_S)$ with covariance matrix $C_S = [K(x, z)]_{x,z\in S}$.

If $f \sim \mathcal{GP}(0, K)$ and given noisy `ZOO` observations $\vec{y}_S = (y_1, \ldots, y_t)$ at points in $S = (x_1, \ldots, x_t)$, with the noise $\eta \sim N(0, \sigma^2)$, the posterior distribution over f given $(x_i, y_i)_{i=1}^t$ is again a GP with posterior mean and variance given by

$$\mu_t(x) = K_t(x)^T \left(C_t + \sigma^2 I_t\right)^{-1} \vec{y}_S,$$
$$\sigma_t^2(x) = K(x, x) - K_t(x) \left(C_t + \sigma^2 I_t\right)^{-1} K_t(x),$$

where $K_t(x) = [K(x, x_1), \ldots, K(x, x_t)]^T$, $C_t = [K(x_i, x_j)]_{1\le i,j\le t}$ and I_t is the $t \times t$ identity matrix. In addition, under some smoothness assumptions on K, the derivative of f is also a GP and its posterior (jointly with f) can also be computed in a similar manner. The reader is referred to (Williams and Rasmussen, 2006, § 9.4) for further details.

Next, we formally introduce the `ZOO` and `FOO` oracles used in this paper.

Definition 2 (`Zeroth-Order-Oracle (ZOO)`)**.** The zero order oracle (ZOO) takes as input any point x in the domain $\mathcal{X}$, and returns $y = f(x) + \eta$, where $\eta \sim N(0, \sigma^2)$. The noise η for different calls to the oracle is assumed to be independent.

Definition 3 (`First-Order-Oracle (FOO)`)**.** The first order oracle (FOO) takes as input any point x in the domain $\mathcal{X} = [0,1]^d$, and returns $(y, y') = (f(x) + \eta, g(x) + \xi)$ where $g(x) = \nabla f(x)$, $\eta \sim N(0, \sigma^2)$ and $\xi \sim N(0, \sigma^2 I_d)$ where I_d is the identity matrix.

2.3 Assumptions

We now state the assumptions on K required to derive our results. These assumptions generalize those used

used in the derivation of lower bound in one dimension by Scarlett (2018) to higher dimensions.

Assumption 1. *We assume that $f \sim \mathcal{GP}(0, K)$ where K is stationary, i.e., $K(x_1, x_2) = K(x_1 - x_2)$ for all $x_1, x_2 \in \mathcal{X}$ and also, $K(x, x) = K(0) \leq 1$ for all $x \in \mathcal{X}$. Furthermore, we assume that the observation noise is distributed as $N(0, \sigma^2)$ for* `ZOO` *and $N\left(0, \sigma^2 I_{D+1}\right)$ for* `FOO` *where I_m is the $m \times m$ identity matrix for $m \geq 1$.*

The stationarity and boundedness assumptions on the kernel K are standard in the theoretical Bayesian Optimization literature, while assuming the same noise variance σ^2 for the function values as well as all the partial derivatives is a matter of notational convenience. Our results will easily carry over to the case where the observation noise is $N(0, \Lambda)$ for some diagonal matrix Λ with non-negative entries.

Assumption 2. *For some $\delta_0 > 0$, with probability at least $1 - \delta_0$, f has a unique maximizer x^* and $f(x^*) \geq f(\tilde{x}) + \epsilon_0$ for some $\epsilon_0 > 0$ and $\tilde{x}$ is any other local maximum of f. Furthermore, we assume that f is twice differentiable and $|f(x)| \leq c_0$, $\|\nabla f(x)\| := \|g(x)\| \leq \bar{c}_1$ and $\|\nabla^2 f(x)\| \leq \bar{c}_2$ for all $x \in \mathcal{X}$.*

As mentioned by Scarlett (2018) and De Freitas et al. (2012), the existence of unique x^* occurs w.p. 1 in most cases. Furthermore, as shown in (Ghosal et al., 2006, Theorem 5), the second-order differentiability of f can be almost surely satisfied by imposing smoothness conditions on the covariance function K. Finally, due to the compactness of the domain $\mathcal{X}$, we can obtain high probability bounds on the suprema of the norms of the first and second order derivatives. Thus the term δ_0 in the above assumption can be made arbitrarily small with suitable choices of the constants $c_0, \bar{c}_1, \bar{c}_2$ and ϵ_0.

Assumption 3. *For some $\delta_1 > 0$ and $\rho_0 \in (0, 1/2)$, with probability at least $1 - \delta_1$ the maximizer x^* satisfies $\|x^* - \partial\mathcal{X}\| \geq \rho_0$, i.e., with probability at least $1 - \delta_1$, the maximizer lies ρ_0 distance away from the boundary of the domain, $\partial\mathcal{X}$.*

Remark 1. Assumption 3 requires the maximizer to lie strictly within the interior of the domain $\mathcal{X}$. Unlike Assumption 2, the probability of satisfying this condition (δ_1) cannot, in general, be made arbitrarily small. There exist cases, such as when the length-scale of the GP kernel is large as compared to the diameter of $\mathcal{X}$, when δ_1 is large and the maximizer x^* lies on or close to the boundary $\partial\mathcal{X}$ (Scarlett, 2018). However, in the case of Bayesian Optimization problems, the algorithm designer also wields significant control over the design of the input space $\mathcal{X}$, and can often ensure that the maximizer x^* lies in the interior of $\mathcal{X}$ by suitably selecting the input ranges or by applying appropriate input transforms. An example of input transformations is the *cylindrical-transformation* proposed by Oh et al. (2018), which been used to address the so-called *boundary issue* (Swersky, 2017, § 4.4.1) and prevent algorithms from sampling too many points near $\partial\mathcal{X}$.

To summarize the assumptions, we introduce the following definition.

Definition 4 (Event Ω_0)**.** We will state the results of our paper conditioned on the event Ω_0 under which Assumptions 2 and 3 are assumed to occur. Note that the probability of the event Ω_0 is at least $1 - \delta_0 - \delta_1$.

Finally, as an immediate consequence of the assumptions, we can state the following result which says that under the event Ω_0, the samples of the GP have locally quadratic behavior in the near-optimal region.

Claim 1. *There exist $0 < \underline{c}_2 \leq \bar{c}_2$ such that the following are true under the event Ω_0:*

$$f(x) + \langle g(x),\, z - x\rangle - \frac{\bar{c}_2}{2}\|z - x\|^2 \leq f(z) \text{ and}$$
$$f(x) + \langle g(x),\, z - x\rangle - \frac{\underline{c}_2}{2}\|z - x\|^2 \geq f(z)$$

for all $x, z \in B(x^*, \rho_0)$. In particular, by setting $x = x^*$ we get

$$\frac{\underline{c}_2}{2}\|z - x^*\|^2 \leq f(x^*) - f(z) \leq \frac{\bar{c}_2}{2}\|z - x^*\|^2 \qquad (2)$$

for all $z \in B(x^*, \rho_0)$.

Proof. Since x^* is the maximizer, by the second-order-necessary-condition we know that $\nabla^2 f(x^*)$ is a negative semi-definite matrix. This along with the fact that the Hessian is almost sure non-singular at x^* (De Freitas et al., 2012) implies that it is in-fact strictly negative-definite. Also due to the continuously differentiable condition in Assumption 2, we can find a $\rho_0' > 0$ and $\underline{c}_2 > 0$ such that $\lambda_{\min}\left(\nabla^2 f(x)\right) \leq -\underline{c}_2$ for all $x \in B(x^*, \rho_0')$. Finally we can update ρ_0 to $\min\{\rho_0', \rho_0\}$ if needed to get the required statement. □

3 Lower Bound on Regret with ZOO

We first revisit Scarlett (2018) to restate/obtain a $\Omega(\sigma\sqrt{n})$ lower bound on the regret with `ZOO` access in Prop. 1 via a direct extension of the binary hypothesis testing technique of Scarlett (2018) to higher dimensions. This result serves to demonstrate the limitations of the binary testing approach in higher dimensions, since it does not capture the scaling of the regret w.r.t. the dimension d. We then obtain improved $\Omega\left(2^{d/2}\sigma\sqrt{n}\right)$ lower bound in Theorem 1 by using our novel approach of mapping the BO problem to an appropriately constructed multi-armed bandit problem.

Proposition 1. *Consider a problem of optimizing a black-box function $f \sim \mathcal{GP}(0, K)$ with* `ZOO` *access and additive observation noise $\eta \sim N(0, \sigma^2)$. If the kernel K satisfies Assumptions 1, 2, and 3 with $1-\delta_0-\delta_1 > 0$, then for any adaptive optimization scheme $\mathcal{A}$, we have*

$$\mathbb{E}\left[\mathcal{R}_n(\mathcal{A}, f)\right] \geq C_l' \left(1 + \sigma\sqrt{n}\right). \tag{3}$$

where C_l' is a constant which does not depend on the dimension d.

Remark 2. The proof of Prop. 1, as well as the stronger lower bound in Theorem 1, proceed by constructing a collection of randomly shifted GPs, all of which when restricted to the index set $\mathcal{X}$, are distributed as $\mathcal{GP}(0, K)$ (here we use the shift invariance of K). As a result, the expectation in Prop. 1 as well as in Theorem 1 is w.r.t. the noise, the GP as well as the random shifts.

Proof Outline of Prop. 1: Our result follows from a generalization of the above-described proof technique used in (Scarlett, 2018). We present an outline of the steps here for completeness.

- Suppose $\tilde{f} \sim \mathcal{GP}(0, K)$ is a GP indexed by the expanded domain $\tilde{\mathcal{X}} \coloneqq [-\Delta, 1+\Delta]^d$. Next, we introduce $\tilde{f}$ restricted to $\mathcal{X}$ as f_0, i.e., $f_0 = \{\tilde{f}(x) : x \in \mathcal{X}\}$. Assume that the sample f_0 is revealed to the learner. This is the so called *genie argument* (Scarlett, 2018), which informally says that any additional information will only result in a weaker bound, and hence can be used in the lower bound construction.

- Next, we introduce two random variables $W \sim$ `Unif` $([d])$ and $V \sim$ `Unif` $(\{-1, +1\})$, independent of everything else. Let $e_w \in \{0, 1\}^d$ for $w \in [d]$ denote the standard normal unit vector with 1 in the w^{th} coordinate, and define $f_{WV} \coloneqq \tilde{f}(x + V\Delta e_W)$ for $x \in \mathcal{X}$. Due to the translation invariance of the kernel K, f_{WV} is also distributed as $\mathcal{GP}(0, K)$, over the index set $\mathcal{X}$ for all realizations of W, V.

- Now, with f_0 revealed and conditioned on W, following (Scarlett, 2018, Lemma 5) we can relate the regret of any BO algorithm ($\mathcal{A}$) to the probability of error in the binary hypothesis test for the true value of v. This allows us to obtain the bound $\mathbb{E}\left[\mathcal{R}_n(\mathcal{A})|W\right] \geq C_l'\sqrt{n}$ for any adaptive scheme $\mathcal{A}$.

- Finally, taking another expectation with respect to the random variable W completes the proof.

□

As shown in Proposition 1, and as hinted in (Scarlett, 2018, § 5), the reduction to binary testing approach does not capture the d dependence on regret, and new techniques are needed to achieve an exponential dependence in d of the lower bound. In our next result, we present a new approach to obtain a tighter lower bound with exponential d dependence.

Theorem 1. *Consider the problem of optimizing a black-box function $f \sim \mathcal{GP}(0, K)$ with additive observation noise $N(0, \sigma^2)$. Under the assumptions 1 2 and 3, and* `ZOO` *access to f, the cumulative regret of any adaptive scheme $\mathcal{A}$ can bounded as*

$$\mathbb{E}\left[\mathcal{R}_n\left(\mathcal{A}, f\right)\right] \geq \sigma\sqrt{\frac{2^d n}{32}} \tag{4}$$

if n is large enough (See (14) *in Appendix A for exact condition).*

The formal proof of this result is deferred to Appendix A. Here, we present a detailed outline of the proof and describe the key ideas involved.

Proof Outline of Theorem 1: Throughout this discussion we will use $\tilde{\mathcal{X}} = [-\Delta, 1+\Delta]^d$ for some $\Delta > 0$, $\mathcal{Z} = \{-1, +1\}^d$. Suppose the elements of $\mathcal{Z}$ are enumerated as $\{z_1, z_2, \ldots, z_L\}$ for $L = 2^d$ in some fixed order. Introduce the random variable $V \sim$ `Unif` $(\{1, 2, \ldots, L\})$ which is drawn independent of all other quantities.

1. We begin with the GP $\tilde{f} \sim \mathcal{GP}(0, K)$ indexed by the larger set $\tilde{\mathcal{X}}$. Next, for a value of $l \in [L]$, we define $f_l = \{\tilde{f}(x + z_l\Delta) : x \in \mathcal{X}\}$. Note that due to the translation invariance of the kernel K, all of $(f_l)_{l \in [L]}$ are distributed according to $\mathcal{GP}(0, K)$ on the index set $\mathcal{X}$. Finally, for the random variable V introduced above, we define $f = f_V$.

2. Suppose x_l^* is the maximizer of the function f_l for $l \in [L]$. Then due to the local quadratic behavior of f_0 under Assumption 3, we can show that if $f_l(x) > f_l(x_l^*) - \underline{c}_2\Delta^2$, then $f_m(x) < f_m(x_m^*) - \underline{c}_2\left(\|z_l - z_m\|_2 - 1\right)^2\Delta^2$ for $m \in [L] \setminus \{l\}$. We next introduce the sets $\mathcal{X}_l \coloneqq B(x_l^*, \Delta)$ for $l \in [L]$, and $\mathcal{X}_{L+1} = \mathcal{X} \setminus \left(\cup_{l=1}^L \mathcal{X}_l\right)$ and observe that for any algorithm $\mathcal{A}$, the regret given that $V = l$ is lower bounded as follows:
$$\mathcal{R}_n\left(\mathcal{A}, f_l\right) \geq \sum_{l' \in [L]\setminus\{l\}} \underline{c}_2 \sum_{t : x_t \in \mathcal{X}_{l'}} \left(\|z_l - z_{l'}\|_2 - 1\right)^2\Delta^2 + \sum_{t : x_t \in \mathcal{X}_{L+1}} \underline{c}_2\Delta^2.$$

3. To each instance of a GP bandit problem with $V = l$, we can associate a multi-armed bandit problem with $(L+1)$ arms, denoted by $\mathcal{M}\left(L, l, \sigma, \underline{c}_2\Delta^2\right) = (p_1, p_2, \ldots, p_{L+1})$ with $p_i \sim N(0, \sigma^2)$ for $i \neq l$ and $p_l \sim N(\underline{c}_2\Delta^2, \sigma^2)$. Furthermore, for any GP bandits algorithm $\mathcal{A}$, we can associate an algorithm for $\mathcal{M}(L, l, \sigma, \underline{c}_2\Delta^2)$, denoted by $\mathcal{A}^{(L)}$ such that if the

strategy $\mathcal{A}$ results in a point $x_t \in \mathcal{X}_{l'}$, then $\mathcal{A}^{(L)}$ plays the arm l' of $\mathcal{M}(L, l, \sigma, \underline{c}_2\Delta^2)$ (also referred to as $\mathcal{M}_l$). Denoting by $\mathcal{R}_n^{(L)}$, the cumulative regret of the multi-armed bandit problem, we have the following relation:

$$\mathcal{R}_n(\mathcal{A}, f_l) \geq \mathcal{R}_n^{(L)}\left(\mathcal{A}^{(L)}, \mathcal{M}_l\right).$$

4. Finally, we show that for the $L+1$ armed bandit problems $\mathcal{M}\left(L, V, \sigma, \underline{c}_2\Delta^2\right)$, we can lower bound the average regret with the term $\sqrt{\frac{nL\sigma^2}{32}}$ which completes the proof.

□

4 Improved Upper Bound on Regret with FOO

4.1 Algorithm with FOO access

Before describing the steps of our algorithm, we first assume that there exists an algorithm which achieves the optimal regret bound in n with ZOO access, i.e., $\mathcal{O}(\sqrt{n})$. One such algorithm can be constructed by a simple generalization of the algorithm of (Scarlett, 2018) to dimensions larger than one.

Assumption 4. *We assume that there exists an algorithm which uses only ZOO feedback, denoted by `OptAlgZOO` or $\mathcal{A}_0$, that for any $n \geq 1$, satisfies the following properties with probability at least $1 - 1/n$ for all $t \leq n$:*

- *It returns a region of the input space, denoted by $S_t \subset B(x^*, r_t)$ where the radius r_t is non-increasing with t, and*
- *The total regret incurred by the algorithm at any time t is $\mathcal{O}\left(\sqrt{t \log n}\right)$.*

We will refer to the above $1 - \frac{1}{n}$ probability event as Ω_1.

With this assumption, we can now describe the steps of our proposed algorithm, `AlgFOO`, whose pseudo-code is in Algorithm 1.

Outline of Algorithm 1. The algorithm proceeds in two phases consisting of the following steps:

- In the first phase, we implement the `OptAlgZOO` algorithm for t_0 steps, where t_0 is large enough (more precise description of t_0 is in Sec. 4.3) to ensure that the *active region* returned by the algorithm is contained in the ball $B(x^*, \rho_0)$. Note that in this phase, the additional gradient information is utilized only to update the posterior and not directly the query point selection strategy.
- Next, in the second phase we exploit the locally quadratic behavior of the objective function f in the vicinity of the maximizer x^* to perform a version of a gradient ascent algorithm which proceeds in two alternating steps: (i) call the `RepeatQuery` subroutine at the current point x_t to repeatedly query the FOO at x_t in order to construct a *sufficiently accurate* estimate of the true gradient at that point, and (ii) perform a gradient-ascent step using the approximate gradient returned by the `RepeatQuery` subroutine.

Algorithm 1: `AlgFOO` $(\mathcal{A}_1)$.

Input: n, K, $\bar{c}_1, \underline{c}_1, \bar{c}_2, \underline{c}_2, \rho_0$.
Initialize: $t = 1$, $n_e = 0$, $S_t = \mathcal{X}$, $\alpha = \bar{c}_2/(4\underline{c}_2)$
  /* Phase 1: Zoom into the near-optimal region using `OptAlgZOO` */
while $S_t \cap \partial\mathcal{X} \neq \emptyset$ *OR* $diam(S_t) > \rho_0$ **do**
  Run `OptAlgZOO` to update S_t
  $t \leftarrow t + 1$
end
/* Phase 2: Perform gradient ascent with uncertain gradients */
$x_t \sim \texttt{Unif}(S(\rho_0, \delta))$
while $t \leq n$ **do**
  t, g_t, μ_t, σ_t $\leftarrow$ `RepeatQuery`$(x_t, t, \mu_t, \sigma_t, \alpha)$
  $x_t \leftarrow x_t + s_t g_t$
end
Output: x_n

Algorithm 2: `RepeatQuery` Subroutine

Input: x, t, $\mu_t(\cdot)$, $\sigma_t(\cdot)$, α
Initialize: $\tau = 0$, flag = `True`, $b = 2\sqrt{d}/\alpha$.
while *flag* **do**
  Query FOO at x
  $t \leftarrow t + 1$, $\tau \leftarrow \tau + 1$
  Update $\mu_t(\cdot), \sigma_t(\cdot)$.
  /* Check Stopping condition */
  Set $u_{t,i} = \mu_{t,i}(x) + \beta_n\sigma/\sqrt{\tau}$, and
   $l_{t,i} = \mu_{t,i}(x) - \beta_n\sigma/\sqrt{\tau}$
  if $\max_{1 \leq i \leq d} \max\left(l_{t,i}, -u_{t,i}\right) > {}^{2b\beta_n\sigma}/\sqrt{\tau}$ **then**
   flag = `False`
  end
end
$\tilde{g}_t = [\mu_{t,1}(x), \ldots, \mu_{t,d}(x)]^T$
Output: t, $\tilde{g}_t$, $\mu_t(\cdot)$, $\sigma_t(\cdot)$

4.2 Regret Bound for `AlgFOO`

We begin by stating a concentration result for the function and derivative values at the points queried by the algorithm.

Proposition 2. *Suppose the first phase of the algorithm* AlgFOO *ends after $t_0 \leq n$ rounds, and the algorithm queries* FOO *at the points $\{x_t \,:\, t_0+1 \leq t \leq n\}$. Define the event Ω_2 under which we have $|\mu_{t,0}(x_t) - f(x_t)| \leq \beta_t \sigma_{t,0}(x_t)$ and $|\mu_{t,i}(x_t) - g_i(x_t)| \leq \beta_t \sigma_{t,i}(x_t)$ for all $1 \leq i \leq d$ and $t_0 + 1 \leq t \leq n$. Then we have $\mathbb{P}(\Omega_2^c) \leq 1/n$, for $\beta_t = \sqrt{2\log\left(\frac{6nt^2}{(d+1)\pi^2}\right)}$.*

Proof Outline of Prop. 2: The result follows from the following two facts: (1) x_t is a measurable function of the history of observations and actions up to and including time $t-1$ for all $t \geq N_0(\rho_0, \delta) + 1$, and (2) the conditional distribution of $[f(x_t), g_1(x_t), \ldots, g_d(x_t)]^T$ is a multivariate Gaussian (here $g(x) = [g_1(x), \ldots, g_d(x)]$ denotes the gradient of f at x). Combining the two in a manner similar to (Srinivas et al., 2012, Lemma 5.6) gives us the result. □

We can now state the following bound on the regret incurred by AlgFOO.

Theorem 2. *Suppose Assumptions 1, 2, 3 and 4 hold. Then, under the event[2] $\Omega = \Omega_0 \cap \Omega_1 \cap \Omega_2$, which occurs with probability at least $1 - \delta_0 - \delta_1 - 2/n$, there exists an $N_{\Omega_0} < \infty$ (depending on the event Ω_0 introduced in Def. 4) such that for all $n \geq N_{\Omega_0}$, the algorithm* AlgFOO *achieves the following bound:*

$$\mathcal{R}_n(\mathcal{A}_1) = \mathcal{O}\left(d(\log n)^2\right) \tag{5}$$

Remark 3. The term N_{Ω_0} in the above result is the minimum sampling budget which ensures that the regret incurred by phase 2 of AlgFOO, which as we show in the proof of Theorem 2 is $\mathcal{O}\left((\sqrt{d}\log n)^2\right)$, dominates the corresponding regret of the first phase, which is $\mathcal{O}\left(\sqrt{t_0 \log n}\right)$ where t_0 is the (random) time at which phase 1 stops. A sufficient condition for this is that $n \geq N_{\Omega_0} = \exp(t_0)$. Since the distance of x^* from the boundary $\partial\mathcal{X}$ is at least ρ_0 by Assumption 3, the term t_0 can be set to $\min\{t \geq 1 \,:\, r_t < \rho_0/2\}$ where r_t was introduced in Assumption 4.

As mentioned in Remark 1, the term δ_0, unlike δ_1, can be made arbitrarily close to 0 by appropriate choice of the terms ϵ_0, c_0, $\bar{c}_1$ and $\bar{c}_2$. In light of this, we can reformulate the result of Theorem 2 as follows.

Corollary 1. *Conditioned on the $1-\delta_1$ probability event of Assumption 3 (i.e., $\|x^* - \partial\mathcal{X}\| \geq \rho_0$), the regret incurred by* AlgFOO *satisfies $\mathcal{R}_n(\mathcal{A}_1) = \mathcal{O}\left(d(\log n)^2\right)$ with high probability (i.e., at least $1 - {(2/n + \delta_0)}/{(1-\delta_1)}$).*

Remark 4. The above corollary says that if we have the prior knowledge that the optimizer of the unknown function f lies away from the boundary of the domain, $\partial\mathcal{X}$, then AlgFOO achieves a regret of the order $\mathcal{O}\left(d(\log n)^2\right)$. This prior knowledge is often available in the canonical BO application of Hyperparameter Optimization (HPO) of machine learning models. More specifically, in HPO problems, the experimenter has significant control over the choice of the search space $\mathcal{X}$, and often designs $\mathcal{X}$ with the goal of ensuring that x^* lies in the interior, as described in (Swersky, 2017, § 4.4.1). For instance, in the hyperparameter tuning of a convolutional neural network (CNN), it is known that very small and very large choices of kernel size are sub-optimal, and hence the range of kernel size can be chosen to ensure that the optimal value lies in the interior. This suggests that in practical applications, the prior knowledge required for achieving the benefits of incorporating gradient information in BO is often available.

We end this section by stating the upper bound on the expected regret of AlgFOO.

Corollary 2. *As an immediate corollary of Theorem 2 and Assumption 4, the expected regret of* AlgFOO *satisfies*

$$\mathbb{E}\left[\mathcal{R}_n(\mathcal{A})\right] = \mathcal{O}\left(d(\log n)^2 + \delta_1\sqrt{n\log n}\right) \tag{6}$$

Thus for the GPs for which the optimizer can be ensured to lie within the interior of the domain with probability at least $1 - {(\log n)^{3/2}}/{\sqrt{n}}$ the overall regret of AlgFOO is $\mathcal{O}\left(d(\log n)^2\right)$.

4.3 Proof Outline for Theorem 2

Suppose $\mathcal{Z} = \{x_1, x_2, \ldots, x_n\}$ denotes the multiset of points queried by the algorithm $\mathcal{A}_1$. This can be partitioned into $\mathcal{Z}_1$ and $\mathcal{Z}_2$, where $\mathcal{Z}_1$ is the multiset of points queried by OptAlgZOO (or $\mathcal{A}_0$) in the first phase and $\mathcal{Z}_2$ denotes the multiset of points queried by $\mathcal{A}_1$ in the second phase, and the total regret incurred by AlgFOO can also be written accordingly as follows:

$$\mathcal{R}_n(\mathcal{A}_1) = \overbrace{\sum_{z \in \mathcal{Z}_1} f(x^*) - f(z)}^{:=\mathcal{R}_{(1)}} + \overbrace{\sum_{z \in \mathcal{Z}_2} f(x^*) - f(z)}^{:=\mathcal{R}_{(2)}}.$$

Throughout this proof, we assume that the event $\Omega = \cap_{i=0}^2 \Omega_i$ occurs. Recall that we have $\mathbb{P}(\Omega) \geq 1 - \delta_0 - \delta_1 - 2/n$.

We first present a bound on the term $\mathcal{R}_{(1)}$ since it is easier to handle. For this, we recall the stopping time t_0 and term N_{Ω_0} defined as:

$$t_0 = \min\left\{t \geq 1 \,:\, r_t < \rho_0/2\right\}, \qquad N_{\Omega_0} \coloneqq \exp(t_0).$$

In the above display, r_t denotes the radius of the ball centered at x^*, i.e, $B(x^*, r_t)$, within which the *active*

[2] Ω_0 introduced in Def. 4, Ω_1 in Assumption 4 and Ω_2 in Prop. 2

set S_t returned by `OptAlgZOO` in Assumption 4 lies. With these definitions we can state the bound on $\mathcal{R}_{(1)}$.

Lemma 1. *Under the event* Ω*, if* $n \geq N_{\Omega_0}$*, then the regret incurred in the first phase, denoted by* $\mathcal{R}_{(1)}$*, is* $\mathcal{O}(\log n)$.

Proof. First note that due to Assumption 3, the optimizer x^* of any realization of the $\mathcal{GP}(0, K)$ must lie in the interior and hence its distance from the boundary $\|x^* - \partial\mathcal{X}\|_2$ is greater ρ_0.

Next, we observe that the total number of rounds spent by `AlgFOO` in the first phase is upper bounded by t_0. This is due to the fact that at time t_0, since $S_{t_0} \subset B(x^*, r_{t_0})$, we have $S_{t_0} \cap \partial\mathcal{X} = \emptyset$. Furthermore, since $r_{t_0} < \rho_0/2$, we also have $\mathrm{diam}(S_{t_0}) < \rho_0$. Thus neither of the conditions on Line 2 of Algorithm 1 is satisfied, which implies that this while loop ends at some time $t \leq t_0$.

Finally, since by Assumption 4 the total regret incurred in first phase is $\mathcal{O}\left(\sqrt{t_0 \log n}\right)$, and that $n \geq N_{\Omega_0} \geq \exp(t_0)$ we get that $\mathcal{R}_{(1)} = \mathcal{O}\left(\sqrt{\log n \log N_{\Omega_0}}\right) = \mathcal{O}(\log n)$. ☐

It remains to show that $\mathcal{R}_{(2)} = \mathcal{O}\left(d(\log n)^2\right)$. To obtain this result, we proceed in the following steps:

- First we show, in Lemma 2, that the `RepeatQuery` subroutine returns a sufficiently accurate estimate of the true gradient $g(x)$ at some point x. More formally, that the approximate gradient $\widetilde{g}_t(x)$ satisfies the property that $\|g(x) - \widetilde{g}_t(x)\| \leq \alpha\|g(x)\|$ for $\alpha = \underline{c}_2/4\bar{c}_2$.
- Next, in Lemma 3, we show that if the `RepeatQuery` subroutine at some point $x \in \mathcal{Z}_2$ halts at time t and with τ queries , then the total regret incurred (i.e., $\tau\left(f(x^*) - f(x)\right)$) can be upper bounded by $\mathcal{O}(\log n)$.
- Next, we partition $\mathcal{Z}_2$ into $\mathcal{Z}_{2,1}$ and $\mathcal{Z}_{2,2}$ where $\mathcal{Z}_{2,2} = \{z \in \mathcal{Z}_2 \ : \ \|x^* - z\| \leq 1/\sqrt{\underline{c}_2 n}\}$ and $\mathcal{Z}_{2,1} = \mathcal{Z}_2 \setminus \mathcal{Z}_{2,2}$. Similarly, we can write $\mathcal{R}_{(2)} = \mathcal{R}_{2,1} + \mathcal{R}_{2,2}$, where $\mathcal{R}_{2,i}$ is the contribution to $\mathcal{R}_{(2)}$ by points in $\mathcal{Z}_{2,i}$ for $i = 1, 2$. The term $\mathcal{R}_{2,2}$ is easy to bound, since by Assumption 3, we know that for any $z \in \mathcal{Z}_{2,2}$ we have $f(x^*) - f(z) \leq \bar{c}_2\|x^* - z\|^2 \leq \bar{c}_2/n\underline{c}_2$ which implies that $\mathcal{R}_{2,2} \leq |\mathcal{Z}_{2,2}|/n \leq 1$.

  It remains to show that $\mathcal{R}_{2,1} = \mathcal{O}\left(d(\log n)^2\right)$. Suppose the unique points in $\mathcal{Z}_{2,1}$ are denoted by $z_1, z_2, \ldots, z_{N_1}$. Then we proceed in three steps:

  – First, from Lemma 3, we know that the regret at any z_i is upper bounded by $\mathcal{O}\left(d(\log n)^2\right)$.

  – Then, in Lemma 4, we show that after every gradient-ascent step (i.e, moving from z_i to z_{i+1}), the distance of the new point from the optimizer x^* shrinks at a geometric rate, i.e., $v_{t+1}^2 \leq \left(1 - \frac{\underline{c}_2^2}{64\bar{c}_2^2}\right) v_t^2$.

  – Finally, in Lemma 5 we show that $|\mathcal{Z}_{2,1}| = N_1 = \mathcal{O}(\log n)$ because of the geometric shrinkage of the distance to the optimal proved in Lemma 4. This result, along with Lemma 3, implies that $\mathcal{R}_{2,1} = \mathcal{O}\left(d(\log n)^2\right)$.

- To summarize, under the $1 - \delta_0 - \delta_1 - 2/n$ probability event Ω, we can decompose the regret into $\mathcal{R}_n = \mathcal{R}_{(1)} + \mathcal{R}_{2,1} + \mathcal{R}_{2,2}$. In Lemma 1 we showed that $\mathcal{R}_{(1)} = \mathcal{O}(\log n)$ under the requirement on $n \geq N_{\Omega_0}$. In Lemma 5 we show that $\mathcal{R}_{2,1} = \mathcal{O}\left(d(\log n)^2\right)$, while the a simple computation outlined above implies that $\mathcal{R}_{2,2} = \mathcal{O}(1)$. Together, these statements complete the proof of Theorem 2.

We now present the formal versions of the remaining steps of the proof outlined above.

Lemma 2. *Suppose the* `RepeatQuery` *subroutine halts at time t with* τ *queries at a point* x*. Then the returned gradient estimate satisfies* $\|g(x) - \widetilde{g}_t\| \leq \alpha\|g\|$ *under the event* Ω *(see Theorem 2), where* $\alpha = \underline{c}_2/4\bar{c}_2$.

The proof of this result is in Appendix B.1. The previous proposition shows that the `RepeatQuery` subroutine indeed returns a ‘sufficiently accurate’ gradient estimate. In the next result, we show that the regret incurred by the `RepeatQuery` subroutine in the process of constructing this gradient estimate is not too large.

Lemma 3. *Suppose a point* x *is evaluated* τ *times by the* `RepeatQuery` *subroutine before halting. Under the event* Ω*, the total regret accumulated, i.e.,* $\tau\left(f(x^*) - f(x)\right)$ *is* $\mathcal{O}(\log n)$.

The proof of this result is in Appendix B.2. In our next result, we show that every time performs the noisy gradient-ascent step (i.e., Line 11 of Algorithm 1), the distance of the new point from the optimizer x^* shrinks by at least a constant factor.

Lemma 4. *Suppose the event* Ω *occurs. Then if at some time* t*,* `AlgFOO` *performs the approximate gradient-ascent step to go from* z_t *to* $z_{t+1} = z_t + s_t\widetilde{g}_t$*, we have*

$$\|z_{t+1} - x^*\|^2 \leq \|z_t - x^*\|^2 \left(1 - \frac{\underline{c}_2^2}{64\bar{c}_2}\right). \qquad (7)$$

The proof of this result is in Appendix B.3. Finally, we combine the previous two results to bound $\mathcal{R}_{2,1}$.

Lemma 5. *Under the event* Ω*, we have* $\mathcal{R}_{2,1} = \mathcal{O}\left(d(\log n)^2\right)$.

The proof of this result is in Appendix B.4.

4.4 Numerical Illustration

We now empirically compare the performance of a heuristic variant of `AlgFOO` (denoted by `AlgFOO-h`) against the GP-UCB baseline. Implementing the exact version of `AlgFOO` requires the knowledge of the constants $\underline{c}_1, \overline{c}_1, \underline{c}_2, \overline{c}_2$ and ρ_0, which may not be easy to obtain in practical problems. To address this, we consider the heuristic `AlgFOO-h`, which:

- implements the first phase of Algorithm 1 with a fraction $r \in (0, 1]$ of the budget,
- calls `RepeatQuery` a fixed number (denoted by `reps`) of times in the second phase, and
- uses a fixed step size s in Line 9 of Algorithm 1.

With these three changes, `AlgFOO-h` no longer depends on the above-mentioned parameters. However, this comes at the cost of losing the theoretical performance guarantees.

We compared the performance of `AlgFOO-h` algorithm with the GP-UCB algorithm of Srinivas et al. (2012) on two commonly used optimization benchmark functions: *Himmelblau* function and *Booth* function. In the experiments, we used three instances of `AlgFOO-h`:

- `AlgFOO-h`-1 with $(r, \texttt{reps}, s) = (0.6, 8, 0.0005)$,
- `AlgFOO-h`-2 with $(r, \texttt{reps}, s) = (0.6, 6, 0.0002)$,
- `AlgFOO-h`-3 with $(r, \texttt{reps}, s) = (0.6, 4, 0.0001)$.

For every objective function and algorithm pair, we ran 20 trials with a budget of $n = 100$ and report the performance in Table 1 (for *Himmelblau* function) and Table 2 (for *Booth* function). The results in the tables provide evidence for the fact that the `FOO` algorithms could lead to improved optimization performance on an average. However, we note that the heuristic `FOO` algorithms also demonstrated higher variability in performance over different trials as well as sensitivity to the choice of the hyperparameters $(r, \texttt{reps}, s)$.

Algo.	mean $\mathcal{R}_n$	median $\mathcal{R}_n$	std. $\mathcal{R}_n$
GP-UCB	1235.23	1206.18	104.81
`AlgFOO-h`-1	1108.86	993.57	429.58
`AlgFOO-h`-2	1057.29	890.32	422.32
`AlgFOO-h`-3	1011.61	877.09	407.20

Table 1: Performance of the algorithms on *Himmelblau* objective function.

These preliminary empirical results suggest that the algorithmic ideas analyzed in this paper can be used to design principled heuristics which can lead to improved performance in practical tasks. We leave a systematic empirical investigation of these ideas for future work.

Algo.	mean $\mathcal{R}_n$	median $\mathcal{R}_n$	std. $\mathcal{R}_n$
GP-UCB	3304.62	2954.70	503.28
`AlgFOO-h`-1	2576.29	2287.72	1013.96
`AlgFOO-h`-2	2763.45	2074.64	1431.62
`AlgFOO-h`-3	1971.46	1616.14	588.01

Table 2: Performance of the algorithms on *Booth* objective function.

5 Conclusion and Future Work

In this paper we took the first step towards quantifying the improvement in regret ($\mathcal{R}_n$) that can be achieved in Bayesian Optimization when the agent has access to gradient information, in addition to the usual noisy function evaluations. To do this, we first derived algorithm-independent lower bound on $\mathcal{R}_n$ with `ZOO` access for all $d \geq 1$. This result captures the exponential scaling of $\mathcal{R}_n$ with dimension d, and relies on a novel approach which proceeds by connecting the regret of the BO problem to that of a certain multi-armed bandit problem. Next, we consider the case in which the agent has `FOO` access to f, and construct an algorithm which achieves a regret bound of $\mathcal{O}\left(d(\log n)^2\right)$ under the condition that the optimizer x^* lies in the interior of the domain. Together these two results imply that exploiting gradient information can be very beneficial in Bayesian optimization.

Our results in this paper open several questions for future work. *First*, it is interesting to investigate whether the lower bounding technique for `ZOO` used in Theorem 1 of this paper can be used to obtain tight lower bounds for related problems such as contextual GP bandits, additive GP bandits and GP level set estimation. *Second*, it is also crucial to obtain the algorithm-independent lower bounds for BO with `FOO` access. For the case of $d = 1$, a combination of the reduction to binary hypothesis testing of Scarlett (2018) along with appropriate KL-divergence bounds for `FOO` due to Raginsky and Rakhlin (2011) might work, but obtaining the lower bound for $d > 1$ may require some new ideas. *Finally*, similar to Algorithm 1 of Scarlett (2018), our proposed algorithm `AlgFOO` is primarily a theoretical device to show that faster convergence can be achieved with gradient information, and is not suitable for practical applications. Thus an important question for future work is to design practically viable algorithms using gradient information, which can also provably achieve tighter than $\Omega(\sqrt{2^d n})$ regret bound.

References

Brochu, E., Cora, V. M., and De Freitas, N. (2010). A Rutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning. *arXiv preprint arXiv:1012.2599*.

Bull, A. D. (2011). Convergence rates of efficient global optimization algorithms. *Journal of Machine Learning Research*, 12(10).

Cai, X. and Scarlett, J. (2020). On lower bounds for standard and robust gaussian process bandit optimization. *arXiv preprint arXiv:2008.08757*.

De Freitas, N., Smola, A. J., and Zoghi, M. (2012). Exponential regret bounds for Gaussian process bandits with deterministic observations. In *Proceedings of the 29th International Coference on International Conference on Machine Learning*, pages 955–962.

Ghosal, S., Roy, A., et al. (2006). Posterior consistency of gaussian process prior for nonparametric binary regression. *The Annals of Statistics*, 34(5):2413–2429.

Grünewälder, S., Audibert, J.-Y., Opper, M., and Shawe-Taylor, J. (2010). Regret Bounds for Gaussian Process Bandit Problems. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, pages 273–280.

Hennig, P. and Schuler, C. J. (2012). Entropy Search for Information-Efficient Global Optimization. *Journal of Machine Learning Research*, 13(Jun):1809–1837.

Lattimore, T. and Szepesvári, C. (2020). *Bandit algorithms*. Cambridge University Press.

Lizotte, D. J. (2008). *Practical bayesian optimization*. University of Alberta.

Nesterov, Y. (2013). *Introductory Lectures on Convex Optimization: A Basic Course*, volume 87. Springer Science & Business Media.

Oh, C., Gavves, E., and Welling, M. (2018). Bock: Bayesian optimization with cylindrical kernels. In *International Conference on Machine Learning*, pages 3868–3877.

Osborne, M. A., Garnett, R., and Roberts, S. J. (2009). Gaussian processes for global optimization. In *3rd international conference on learning and intelligent optimization (LION3)*, pages 1–15.

Prabuchandran, K., Penubothula, S., Kamanchi, C., and Bhatnagar, S. (2020). Novel first order bayesian optimization with an application to reinforcement learning. *Applied Intelligence*, pages 1–15.

Raginsky, M. and Rakhlin, A. (2011). Information-based Complexity, Feedback and Dynamics in Convex Programming. *IEEE Transactions on Information Theory*, 57(10):7036–7056.

Russo, D. and Van Roy, B. (2014). Learning to optimize via posterior sampling. *Mathematics of Operations Research*, 39(4):1221–1243.

Scarlett, J. (2018). Tight Regret Bounds for Bayesian Optimization in One Dimension. In *International Conference on Machine Learning*, pages 4500–4508.

Scarlett, J., Bogunovic, I., and Cevher, V. (2017). Lower bounds on regret for noisy gaussian process bandit optimization. In *Conference on Learning Theory*, pages 1723–1742. PMLR.

Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., and De Freitas, N. (2015). Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proceedings of the IEEE*, 104(1):148–175.

Shekhar, S. and Javidi, T. (2018). Gaussian Process Bandits with Adaptive Discretization. *Electronic Journal of Statistics*, 12(2):3829–3874.

Srinivas, N., Krause, A., Kakade, S. M., and Seeger, M. W. (2012). Information-Theoretic Regret Bounds for Gaussian Process Optimization in the Bandit Setting. *IEEE Transactions on Information Theory*, 58(5):3250–3265.

Swersky, K. J. (2017). *Improving Bayesian Optimization for Machine Learning using Expert Priors*. PhD thesis, University of Toronto.

Wang, Z. and Jegelka, S. (2017). Max-Value Entropy Search for Efficient Bayesian Optimization. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 3627–3635. JMLR. org.

Williams, C. K. and Rasmussen, C. E. (2006). *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA.

Wu, A., Aoi, M. C., and Pillow, J. W. (2017a). Exploiting Gradients and Hessians in Bayesian Optimization and Bayesian Quadrature. *arXiv preprint arXiv:1704.00060*.

Wu, J., Poloczek, M., Wilson, A. G., and Frazier, P. (2017b). Bayesian Optimization with Gradients. In *Advances in Neural Information Processing Systems*, pages 5267–5278.

A Proof of Theorem 1.

Notations. We first recall some notation required for presenting the proof. Given a kernel K, we define a GP $\widetilde{f}$ distributed according to $\mathcal{GP}(0, K)$ on the extended index set $\widetilde{X} = [-\Delta, \Delta]^d$ for some $\Delta > 0$ to be decided later. For $L = 2^d$, let $\{z_1, z_2, \ldots, z_L\}$ denote the set $\{-1, 1\}^d$ where the $(z_l)_{l=1}^L$ are ordered in some arbitrary but fixed way, and let $V \sim \texttt{Unif}([L])$ be a random variable independent of everything else.

With $z_0 = (0, \ldots, 0) \in \mathbb{R}^d$, we introduce the Gaussian Processes $f_l = \{\widetilde{f}(x+\Delta z_l) \ : \ x \in \mathcal{X}\}$ for $l \in \{0\} \cup [L]$. Note that due to the translation invariance of the kernel K, all of f_l have the same $\mathcal{GP}(0, K)$ distribution (although they are not independent). We will use f_l^* and x_l^* for all $l \in [L]$ to denote the maximum value and the maximizer of f_l on the domain $\mathcal{X}$.

We first state a result which says that there exists no point x at which both f_l and f_m for $l \neq m$ are arbitrarily close to their respective maximum values.

Lemma 6. *Suppose the event Ω_0 (introduced in Definition 4) holds, and Δ is small enough to ensure that $(2\sqrt{d}+1)\Delta < \rho_0$. Then, for any $m \neq l$ and $x \in \mathcal{X}$, if $f_l^* - f_l(x) < \underline{c}_2 \Delta^2$ then $f_m^* - f_m > \underline{c}_2(\|z_m - z_l\|_2 - 1)\Delta^2$.*

Proof. The statement above follows as a consequence of the local quadratic behavior of GP samples under the event Ω_0 as formalized in (2) in Claim 1. The proof is a generalization of the corresponding result in Eq. (28) in Lemma 3 of (Scarlett, 2018) to the case of higher dimensions, and we present the details here for completion.

Suppose x is such that $f_l^* - f_l(x) \leq \underline{c}_2 \Delta^2$, then by the result of Claim 1, we have that $\underline{c}_2 \|x - x_l^*\|^2 \leq f_l^* - f(x) \leq \underline{c}_2 \Delta^2$ which implies that $\|x_l^* - x\| < \Delta$.

Next, we know by definition of x_l^* and x_m^* that $\|x_l^* - x_m^*\|$ is equal to $\|z_m - z_l\|\Delta$. By triangle inequality and the fact that $\|x_l^* - x\| < \Delta$, this implies that $\|x_m^* - x\| > (\|z_m - z_l\| - 1)\Delta$, which by another application of the result of Claim 1 implies that $f_m^* - f_m(x) \geq \underline{c}_2 \|x - x_m^*\|^2 > \underline{c}_2 (\|z_m - z_l\| - 1)^2 \Delta^2$. ☐

Suppose a GP bandit algorithm $\mathcal{A}$ selects the following query points $x_1, x_2, \ldots, x_n$ on the objective f_l. Let $T_i(n, \mathcal{A}, l)$ denote the (random) number of times the algorithm queries points in the subset $\mathcal{X}_i$. Then the regret incurred by algorithm $\mathcal{A}$ on f_l, denoted by $\mathcal{R}_n(\mathcal{A}, f_l)$, satisfies the following inequality

$$\mathcal{R}_n(\mathcal{A}, f_l) = \sum_{t=1}^n f_l^* - f_l(x_t) \stackrel{(a)}{\geq} \sum_{t=1}^n \underline{c}_2 \Delta^2 \left(\sum_{l' \neq l} (\|z_l - z_{l'}\|^2 - 1) \mathbb{1}_{\{x_t \in \mathcal{X}_{l'}\}} + \mathbb{1}_{\{x_t \in \mathcal{X}_{L+1}\}} \right)$$

$$\stackrel{(b)}{=} \underline{c}_2 \Delta^2 \left(T_{L+1}(n, \mathcal{A}, f_l) + \sum_{l' \neq l} (\|z_{l'} - z_l\|_2 - 1) T_{l'}(n, \mathcal{A}, f_l) \right) \tag{8}$$

$$\geq \underline{c}_2 \Delta^2 \left(n - T_l(n, \mathcal{A}, f_l) \right). \tag{9}$$

The inequality **(a)** in the above display follows from an application of Lemma 6 by noting that points which are not in $\mathcal{X}_l$ are at least Δ away from the optimizer x_l^*, and hence at least $\underline{c}_2 \Delta^2 (\|z_l - z_{l'}\| - 1)$ sub-optimal, while **(b)** uses the fact that $\|z_l - z_{l'}\| \geq 2$ for all l, l'.

The expressions in (**??**) and (9) resemble the regret decomposition in a multi-armed bandit (MAB) problem with $L+1$ arms. Motivated by this, for any GP bandit problem with objective function f_l, we associate a corresponding $(L+1)-$armed MAB problem, denoted by $\mathcal{M}_l = \mathcal{M}\left(L, l, \sigma, \underline{c}_2 \Delta^2\right) = \left(P_1^{(l)}, P_2^{(l)}, \ldots, P_{L+1}^{(l)}\right)$ where $P_i^{(l)} \sim N\left(\underline{c}_2 \Delta^2 \mathbb{1}_{\{i=l\}}, \sigma^2\right)$. Now, corresponding to any GP bandit algorithm $\mathcal{A}$, we can define an algorithm for the corresponding $(L+1)$ armed MAB problem, denoted by $\mathcal{A}^{(L)}$, such that if the point x_t played by $\mathcal{A}$ lies in $\mathcal{A}$ lies in $\mathcal{X}_i$, the scheme $\mathcal{A}^{(L)}$ plays arm $a_t = i$. With this setup, and from (9), we can now relate the regret incurred by $\mathcal{A}^{(L)}$ on $\mathcal{M}_l$, denoted by $\mathcal{R}_n^{(L)}(\mathcal{A}^{(L)}, \mathcal{M}_l)$, to the regret of the original GP bandits problem.

$$\mathcal{R}_n(\mathcal{A},\, f_l) \geq \mathcal{R}_n^{(L)}\left(\mathcal{A}^{(L)},\, \mathcal{M}_l\right). \tag{10}$$

This suggests that in order to lower bound $\mathcal{R}_n$, it suffices to obtain a lower bound on $\mathcal{R}_n^{(L)}$. We follow this approach, and in particular, show that

$$\inf_{\mathcal{A}\in\Pi^{GP}} \mathbb{E}\left[\mathcal{R}_n\left(\mathcal{A}, f_V\right)\right] \geq \inf_{\mathcal{A}^{(L)}:\mathcal{A}\in\Pi^{GP}} \mathbb{E}\left[\mathcal{R}_n^{(L)}\left(\mathcal{A}^{(L)}, \mathcal{M}_V\right)\right] \geq \inf_{A\in\Pi^{MAB}} \mathbb{E}\left[\mathcal{R}_n^{(L)}\left(A, \mathcal{M}_V\right)\right]. \tag{11}$$

In the above display, $\Pi^{(GP)}$ and $\Pi^{(MAB)}$ denote the set of all feasible GP bandit and MAB algorithms respectively. Thus to complete the proof, we need to obtain a lower bound on the expected regret of any MAB algorithm for problem instances chosen randomly from the set $\{M_l; l \in [L]\}$. For any bandit algorithm A, and a MAB instance $\mathcal{M}_l$, we use the notation $T_i(A, \mathcal{M}_l, n)$ to denote the (random) number of times the algorithm A pulled arm $i \in [L+1]$. Denoting by $\mathbb{E}_i[\cdot]$, the expectation conditioned on $V = i$ for $i \in [L]$, we then have

$$\begin{aligned}
\mathbb{E}\left[\mathcal{R}_n^{(L)}\left(A,\, \mathcal{M}_V\right)\right] &= \frac{1}{L}\sum_{l=1}^{L} \mathbb{E}_l\left[\mathcal{R}_n^{(L)}\left(A,\, \mathcal{M}_V\right)\right] \geq \frac{1}{L}\underline{c}_2\Delta^2\mathbb{E}_i\left[\sum_{i\neq l} T_i\left(A, \mathcal{M}_l, n\right)\right] \\
&= \frac{1}{L}\sum_{i=1}^{L} \underline{c}_2\Delta^2\left(n - \mathbb{E}_i\left[T_i\left(A, \mathcal{M}_l, n\right)\right]\right) \\
&= \underline{c}_2\Delta^2\left(n - \frac{1}{L}\sum_{l=1}^{L}\mathbb{E}_l\left[T_l\left(A, \mathcal{M}_l, n\right)\right]\right).
\end{aligned} \tag{12}$$

Next, we need to obtain an upper bound on the term $\frac{1}{L}\sum_{l=1}^{L}\mathbb{E}_l[T_l\left(A, \mathcal{M}_l, n\right)]$, which is the average number of times the algorithm A pulls the optimal arm on an MAB instance drawn uniformly at random from the $[L]$ options. The proof of this statement follows by adapting the existing lower bounding techniques in MAB and statistics literature. Introduce the notation $\mathcal{M}_0$ to denote an MAB with all arms with $N(0, \sigma^2)$ and let $\mathbb{E}_0$ denote the corresponding expectation. Furthermore, let ρ_i^A for $i \in \{0, \ldots, L\}$ denote the joint distribution on the action-observation sequences $\mathcal{J}_n := (\mathbb{R} \times [L+1])^n$ induced by the strategy A and MAB $\mathcal{M}_i$. For $i \in [L] \cup \{0\}$ and an MAB algorithm A, we introduce $\tau_0(i, A, n) \subset [n] \cup \{0\}$ defined as $\tau_0(i, A, n) = \{t \in [n] \cup \{0\} : \rho_i^A\left(\{T_i(A, \mathcal{M}_i, n) = t\}\right) \geq \rho_0^A\left(\{T_i(A, \mathcal{M}_0, n) = t\}\right)\}$. Then we have the following

$$\begin{aligned}
\mathbb{E}_l\left[T_l\left(A, \mathcal{M}_l, n\right)\right] - \mathbb{E}_0\left[T_l\left(A, \mathcal{M}_l, n\right)\right] &= \sum_{t=0}^{n} t\left(\rho_i^A\left(\{T_i = t\}\right) - \rho_0^A\left(\{T_i = t\}\right)\right) \\
&\overset{(i)}{\leq} \sum_{t\in\tau_0(i,A,n)} t\left(\rho_i^A\left(\{T_i = t\}\right) - \rho_0^A\left(\{T_i = t\}\right)\right) \\
&\overset{(ii)}{\leq} n\sup_{E\subset\mathcal{J}_n}\left|\rho_0^A(E) - \rho_i(E)\right| = n\, d_{TV}\left(\rho_i^A, \rho_0^A\right) \\
&\overset{(iii)}{\leq} n\sqrt{\frac{1}{2}d_{KL}\left(\rho_0^A, \rho_i^A\right)} \\
\Rightarrow \mathbb{E}_l\left[T_l\left(A, \mathcal{M}_l, n\right)\right] \leq \mathbb{E}_0\left[T_l\left(A, \mathcal{M}_l, n\right)\right] &+ n\sqrt{\frac{1}{2}d_{KL}\left(\rho_0^A, \rho_i^A\right)}.
\end{aligned} \tag{13}$$

In the previous display,
(i) follows from the fact that by definition the set $E_0(i, A, n)$ denotes the values of t for which the terms in the expectation are non-negative,
(ii) follows by upper-bounding t with n, and then taking the supremum over all possible subsets of the action-observations sequences $\mathcal{J}_n$, and
(iii) follows by an application of Pinsker's inequality.

Next, we proceed as follows:

$$\begin{aligned}
\frac{1}{L}\sum_{l=1}^{L}\mathbb{E}_l\left[T_l\left(A,\mathcal{M}_l,n\right)\right] &\overset{(i)}{\leq} \frac{1}{L}\sum_{l=1}^{L}\mathbb{E}_0\left[T_l\left(A,\mathcal{M}_l,n\right)\right] + n\frac{1}{L}\sum_{l=1}^{L}\sqrt{\frac{1}{2}d_{KL}\left(\rho_0^A,\rho_l^A\right)} \\
&\overset{(ii)}{=} \frac{n}{L} + \frac{n}{L}\sum_{l=1}^{L}\sqrt{\frac{1}{2}d_{KL}\left(\rho_0^A,\rho_l^A\right)} \\
&\overset{(iii)}{=} \frac{n}{L} + \frac{n\underline{c}_2\Delta^2}{2L\sigma}\sum_{l=1}^{L}\sqrt{\mathbb{E}_0\left[T_l(n)\right]} \\
&\overset{(iv)}{\leq} \frac{n}{L} + \frac{n\underline{c}_2\Delta^2}{2L\sigma}\sqrt{Ln} = \frac{n}{L}\left(1+\frac{\underline{c}_2\Delta^2\sqrt{nL}}{2\sigma}\right)
\end{aligned}$$

In the previous display,
(i) follows from Eq. (13),
(ii) uses the fact that $\sum_{l=1}\mathbb{E}_0\left[T_l\right] \leq n$,
(iii) uses the standard kl-divergence decomposition rule for multi-armed bandits (Lattimore and Szepesvári, 2020, Lemma 15.1),
(iv) follows from an application of Cauchy-Schwarz inequality to get $\sum_l \sqrt{\mathbb{E}_0\left[T_l\right]} \leq \sqrt{L}\sqrt{\mathbb{E}_0\left[\sum_l T_l\right]} = \sqrt{Ln}$.

Finally, combining this result with (12), we get the required inequality as follows:

$$\begin{aligned}
\mathbb{E}\left[\mathcal{R}_n^{(L)}\left(A,\,\mathcal{M}_V\right)\right] &\geq \underline{c}_2\Delta^2 n\left(1-\frac{1}{L}\left(1+\frac{\underline{c}_2\Delta^2\sqrt{nL}}{2\sigma}\right)\right) \\
&\overset{(i)}{=} \sigma\sqrt{\frac{nL}{2}}\left(1-\frac{1}{L}\left(1+\frac{L}{4}\right)\right) \overset{(ii)}{\geq} \sigma\sqrt{\frac{nL}{2}}\left(1-3/4\right) = \sigma\sqrt{\frac{nL}{32}}.
\end{aligned}$$

In the previous display,
(i) follows from our choice of $\Delta = \left(L\sigma^2/2n\underline{c}_2^2\right)^{1/4}$, and
(ii) follows from the assumption that $d \geq 1$, which implies that $L \geq 2$ and hence $1 + L/4 \leq L/2 + L/4 = 3L/4$.

Finally, to ensure the above choice of $\Delta = L\sigma^2/2n\underline{c}_2^2$ satisfies the condition on Δ required in Lemma 6, i.e., $\Delta < \rho_0/(2\sqrt{d}+1)$, we need that n is large enough, i.e.,

$$n > \frac{2^d\sigma^2}{2\underline{c}_2^2}\left(\frac{2\sqrt{d}+1}{\rho_0}\right)^4. \tag{14}$$

This concludes the proof of Theorem 1.

B Details of Proof of Theorem 2

B.1 Proof of Lemma 2

Introduce the notation $v = \max_{1\leq i\leq d} \max(l_{t,i},\, -u_{t,i})$, which is achieved at the coordinate j. Furthermore, assume that $l_{t,j} > 0$ (the other case can be handled in exactly the same way). Then we have $2b\beta_n\sigma/\sqrt{\tau} \leq l_{t,j} \leq \|g(x)\|_\infty \leq u_{t,j} = l_{t,j} + 4b\beta_n\sigma/\sqrt{\tau}$.

$$\|\tilde{g}_t - g(x)\|_2 = \left(\sum_{i=1}^{d} |\mu_{t,i}(x) - g_j(x)|^2\right)^{1/2} \overset{(i)}{\leq} \frac{2\sqrt{d}\beta_n\sigma}{\sqrt{\tau}}$$
$$\overset{(ii)}{\leq} \frac{2\sqrt{d}}{b}\frac{b\beta_n\sigma}{\sqrt{\tau}} \overset{(iii)}{\leq} \frac{2}{b}\|g(x)\|_\infty \overset{(iv)}{\leq} \frac{2\sqrt{d}}{b}\|g(x)\|_2.$$

In the above display:
(i) follows from the fact that under the event Ω, we have $|\mu_{t,i}(x) - g_i(x)| \leq \beta_n\sigma_{t,j}(x) \leq \beta_n\sigma/\sqrt{\tau}$ for all t, τ, and $1 \leq i \leq d$. The last inequality follows from an application of the first part of (Shekhar and Javidi, 2018, Proposition 3).
(ii) simply involves multiplication and division by the non-zero term b
(iii) uses the fact (discussed at the beginning of this proof) that due to the stopping rule, we must have $b\beta_n\sigma/\sqrt{\tau} \leq \|g(x)\|_\infty$, while
(iv) uses the fact that $\|z\|_\infty \leq \|z\|_2$ for any $z \in \mathbb{R}^d$.

B.2 Proof of Lemma 3

As mentioned earlier we assume that the event Ω (which occurs with probability at least $1 - \delta_0 - \delta_1 - 2/n$) holds throughout this proof.

For $\tau = 1$, the regret incurred is just $f(x^*) - f(x)$ which can be bounded trivially by the constant $2c_0$, where the term c_0 was introduced in Assumption 1. Thus for the remaining part of the proof we will assume $\tau > 1$.

Suppose the distance of the point x from the optimizer x^* is $v := d(x^*, x)$. Then from Assumption 2, we have $\underline{c}_2 v \leq \|g(x)\|_2 \leq \bar{c}_2 v$ and $\underline{c}_2 v^2 \leq f(x^*) - f(x) \leq \bar{c}_2 v^2$. We have the following chain of inequalities:

$$\underline{c}_2\, v \leq \|g(x)\|_2 \leq \sqrt{d}\|g(x)\|_\infty \leq \frac{\sqrt{d}(b+2)\beta_n\sigma}{\sqrt{\tau - 1}}.$$

This result implies the upper bound on the number of queries made by `RepeatQuery` subroutine, τ,

$$\frac{\tau}{2} \leq \tau - 1 \leq \frac{(b+2)^2 d\beta_n^2\sigma^2}{\underline{c}_2^2 v^2} \quad \Rightarrow \quad \tau \leq \frac{2(b+2)^2 d\beta_n^2\sigma^2}{\underline{c}_2^2 v^2} \tag{15}$$

Now the total regret incurred is $\tau\left(f(x^*) - f(x)\right) \leq \tau\bar{c}_2 v^2$, which gives us the bound

$$\tau\left(f(x^*) - f(x)\right) \leq \bar{c}_2 v^2 \left(\frac{(b+2)^2 d\beta_n^2\sigma^2}{\underline{c}_2^2 v^2}\right) = \frac{\bar{c}_2(b+2)^2 d\sigma^2\beta_n^2}{\underline{c}_2^2} = \mathcal{O}\left(\beta_n^2\right)$$

This completes the proof, since $\beta_n^2 = 2\log\left(\frac{6n^3}{(d+1)\pi^2}\right) = \mathcal{O}\left(\log n\right)$.

B.3 Proof of Lemma 4

Suppose $\{z_1, z_2, \ldots, z_{N_1}\}$ denote the set of unique points resulting from the gradient steps taken by the algorithm and introduce the notation $v_t \coloneqq \|z_t - x^*\|_2$. Then we have the following:

$$\begin{aligned}
v_{t+1}^2 &= \|z_t - x^* + s_t\tilde{g}_t\|_2^2 = \|z_t - x^* + s_t g_t + s_t(\tilde{g}_t - g_t)\|_2^2 \\
&= v_t^2 + s_t^2\left(\|g_t\|_2^2 + \|g_t - \tilde{g}_t\|_2^2 + 2\langle g_t, \tilde{g}_t - g_t\rangle\right) + 2s_t\langle z_t - x^*, g_t\rangle + 2s_t\langle z_t - x^*, \tilde{g}_t - g_t\rangle \\
&\stackrel{(i)}{\leq} v_t^2 + s_t^2\bar{c}_2^2\left(v_t^2 + \alpha^2 v_t^2 + 2\alpha v_t^2\right) + 2s_t\langle z_t - x^*, g_t\rangle + 2s_t\langle z_t - x^*, \tilde{g}_t - g_t\rangle \\
&\stackrel{(ii)}{\leq} v_t^2 + s_t^2\bar{c}_2^2(1+\alpha)^2 v_t^2 - s_t\underline{c}_2 v_t^2 + 2s_t\langle z_t - x^*, \tilde{g}_t - g_t\rangle \\
&\stackrel{(iii)}{\leq} v_t^2 + s_t^2\bar{c}_2^2(1+\alpha)^2 v_t^2 - s_t\underline{c}_2 v_t^2 + 2s_t\bar{c}_2\alpha v_t^2 \\
&\stackrel{(iv)}{\leq} v_t^2\left(1 + 4s_t^2\bar{c}_2^2 + 2s_t\bar{c}_2\alpha - s_t\underline{c}_2\right) \\
&\stackrel{(v)}{=} v_t^2\left(1 + 4s_t^2\bar{c}_2^2 - s_t\underline{c}_2/2\right) \\
&\stackrel{(v)}{=} v_t^2\left(1 - s_t\left(\frac{\underline{c}_2}{2} - 4s_t\bar{c}_2\right)\right) = v_t^2\left(1 - \frac{\underline{c}_2^2}{64\bar{c}_2}\right).
\end{aligned}$$

In the above display,

- **(i)** uses the fact that $\|g_t\|_2^2 \leq \bar{c}_2 v_t^2$ and the fact that $\|g_t - \tilde{g}_t\|_2 \leq \alpha\|g_t\|_2$.
- **(ii)** uses the fact that due to the local quadratic behavior (under the event Ω_0) of the function f in the near optimal region as described in Claim 1, we have

  $$\begin{aligned}
  f(x^*) \leq&\ f(z_t) + \langle g_t, x^* - z_t\rangle - \frac{\underline{c}_2}{2}v_t^2 \\
  \Rightarrow \langle g_t, z_t - x^*\rangle \leq&\ f(z_t) - f(x^*) - \frac{\underline{c}_2}{2}v_t^2 \ \leq\ 0 - \frac{\underline{c}_2}{2}v_t^2,
  \end{aligned}$$

  which implies that $2s_t\langle z_t - x^*, g_t\rangle \leq -s_t\underline{c}_2 v_t^2$.
- **(iii)** uses the Cauchy-Schwarz inequality, the fact that $\|g_t - \tilde{g}_t\|_2 \leq \alpha\|g_t\|_2$ and the bound $\|g_t\|_2 \leq \bar{c}_2 v_t$.
- **(iv)** simply bounds $(1+\alpha)^2$ with 4, since $\alpha \leq 1$.
- **(v)** uses the fact that the choice of $\alpha = \underline{c}_2/4\bar{c}_2$ implies that $2\alpha\bar{c}_2 = \underline{c}_2/2$, which means that $2s_t\bar{c}_2\alpha - s_t\underline{c}_2 = -s_t\underline{c}_2/2$.
- **(v)** follows by plugging in the value of $s_t = \underline{c}_2/16\bar{c}_2$.

B.4 Proof of Lemma 5

This result follows from a combination of Lemma 3 and Lemma 4. More specifically, recall that $\mathcal{Z}_{2,1} = \{z_1, z_2, \ldots, z_{N_1}\}$ for some $N_1 \geq 1$. Then to show the required result, it suffices to prove that $N_1 = \mathcal{O}(\log n)$, because this along with the fact that the total contribution of each z_i to $\mathcal{R}_{2,1}$ is $\mathcal{O}(d\log n)$ (Lemma 3) implies the required $\mathcal{O}\left(d(\log n)^2\right)$ bound.

Now, Lemma 4 tells us that $\|z_{i+1} - x^*\|_2^2 \leq C\|z_i - x^*\|_2^2$ where $C = \left(1 - \frac{\underline{c}_2^2}{64\bar{c}_2}\right) < 1$. Repeated application of this implies that $\|z_{N_1} - x^*\|_2^2 \leq C^{N_1-1}\|z_1 - x^*\|_2^2 \leq C^{N_1-1}\rho_0^2$. Furthermore, by the definition of the set $\mathcal{Z}_{2,1}$ we must have $\|z_{N_1} - x^*\|_2^2 \geq 1/(\bar{c}_2 n)$. Together, these two conditions imply that $N_1 \leq 1 + \frac{\log\left(\rho_0^2\bar{c}_2 n\right)}{\log(1/C)} + 1 = \mathcal{O}(\log n)$. This completes the proof.